
Development of an Artificial Peroxidase Based on a Human Carbonic Anhydrase Protein

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Iori Morita

2024

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel

<https://edoc.unibas.ch>

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von:

Erstbetreuer: Prof. Dr. Thomas R. Ward

Zweitbetreuer: Prof. Dr. Olivier Baudoin

Externer Experte: Prof. Dr. Gerard Roelfes

Basel, 30. April 2024

Prof. Dr. Marcel Mayor

The Dean of Faculty

「理想を失わない現実主義者にならないといけないんです。理想のない現実主義者ならいくらでもいるんですよ。」

宮崎駿

Acknowledgement

First of all, I would like to thank **the Nakajima Foundation** for the financial support during my PhD years.

Throughout my doctoral journey spanning over four years, I have been fortunate to receive guidance and support from several esteemed mentors. First and foremost, I would like to thank **Prof. Dr. Thomas Ward**, for being always supportive and for giving me invaluable pieces of advice in both my research and career. I am equally grateful to **Prof. Dr. Olivier Baudoin**, my secondary advisor, who provided precious insights and advice from the perspective of organic chemistry. I would like to thank **Prof. Dr. Gerard Roelfes** for graciously accepting the role of external examiner.

Special thanks are also due to **Dr. Valentin Köhler** for his invaluable support, especially in the deep insight into various analytical instruments. I would like to thank **Prof. Dr. Anastassia Vorobieva** and **Prof. Dr. Bruno Correia** for generously providing the library of chimeric hCAII proteins to our group. I deeply appreciate the contribution from **Dr. Adriana Faraone** for the cloning, expression, and purification of the key protein essential to my thesis. I would like to thank **Elias**, without whom the latter part of this doctoral thesis would not have been possible. I am immensely grateful to **Kailin**, **Dr. Roman P. Jakob**, and **Eli** for their assistance in the X-ray crystallography of my enzymes. I appreciate the contribution from **Xiaochun** for her assistance and guidance in binding assays and quantitative determination of the thermal stability of my proteins. I thank **all the coworkers** in the Ward group since I joined the group for the comfortable and supportive atmosphere.

I am deeply thankful to **the Japanese community in Basel** for their support throughout my stay, from organizing social gatherings to exchanging valuable information. My heartfelt appreciation also goes to **my family in Japan** for their constant support during my stay.

Finally, I express my deepest gratitude to my wife, **Natsuka**, for her decision to relocate to Europe and her unwavering support, which have greatly enriched our lives in Switzerland.

いつもありがとう。

Table of Contents

ACKNOWLEDGEMENT	III
TABLE OF CONTENTS	IV
ABSTRACT	1
1 INTRODUCTION	2
1.1 C–H Hydroxylation by Fe-TAML Complexes	2
1.2 Artificial Metalloenzymes	6
2 GOAL OF THE THESIS	24
3 DIRECTED EVOLUTION OF AN ARTIFICIAL PEROXIDASE BASED ON A HUMAN CARBONIC ANHYDRASE PROTEIN	25
3.1 Abstract.....	25
3.2 Introduction	25
3.3 Results and Discussion	27
3.4 Conclusion	37
3.5 Supporting Information	39
3.6 References	95
4 INVESTIGATION OF THE IMPACT OF A GRAFTED MOTIF NEAR THE ACTIVE SITE	98
4.1 Introduction	98
4.2 Results and Discussion	98
4.3 Conclusion	104
4.4 Supporting Information	105
4.5 References	142
5 CONCLUDING REMARKS	143
5.1 Summary and Outlook	143
5.2 References	144

Abstract

Artificial metalloenzymes are constructed through the incorporation of a metal catalyst into a protein scaffold. The protein scaffold provides second coordination sphere around the catalytic moiety, endowing it with various attributes to the reaction that the free catalyst does not possess. In this PhD thesis, we will present the development of an artificial peroxidase (ArPase) by combining an Fe-TAML (TAML = Tetra Amide Macrocyclic Ligand) catalyst with a human carbonic anhydrase scaffold. We will show that the protein scaffold enhances the peroxidase activity of the catalyst. Through directed evolution facilitated by screening on cell lysates, the activity of ArPase was improved both in terms of total turnover number and enantiomeric excess. Additionally, a library of chimeric hCAII scaffolds was explored for further improving the performance of the ArPase.

Overall, this thesis highlights the advantage of “genetic” optimization of a Fe-TAML catalyst through the assembly of an ArPase.

1 Introduction

1.1 C–H Hydroxylation by Fe-TAML Complexes

1.1.1 C–H Hydroxylation with Metal-oxo Complexes

The hydroxylation of C–H bonds is a key element among synthetic strategies, considering that many natural compounds bear a hydroxyl group. Late-stage conversion of a C–H bond to a C–OH moiety can have a significant impact given that such modification can lead to considerable changes in the molecular properties^[1]. One of the multiple mechanisms for C–H hydroxylation is the C–H bond abstraction by a metal-centered oxidant followed by C–O bond formation without the generation of an organometallic intermediate. This strategy is of particular interest considering the well-defined, tunable catalyst properties and the use of earth-abundant metals, typically iron and manganese^[2]. This mechanism is widely found among natural enzymes such as i) heme-dependent enzymes, including cytochrome P450 monooxygenases^[3], ii) non-heme mononuclear iron-dependent enzymes including α -ketoglutarate-dependent dioxygenases^[4], iii) dinuclear iron-dependent enzymes including soluble methane monooxygenases^[5], and iv) copper-dependent enzymes including lytic polysaccharide monooxygenases^[6]. Inspired by natural metal active sites, various synthetic ligands for C–H hydroxylation have been developed including porphyrins, aminopyridines, and salens^[7]. Nevertheless, regio- and stereoselective hydroxylation of unactivated C–H bonds with such metal complexes remains challenging due to the abundance of such C–H bonds present in organic molecules. Therefore, as discussed in Section 1.2, employing protein scaffolds as the second coordination sphere around a catalyst with a well-defined first coordination sphere holds promise for enhancing the selectivity and activity of such metal complexes.

1.1.2 Tetra Amide Macrocyclic Ligand (TAML)

Fe-TAML catalysts were initially introduced by Collins and coworkers as biomimetic complexes for the oxidation of organic molecules^[8]. The ligand has been studied over the last two decades, leading to multiple variants of its structure. Notably, the Fe-TAML catalysts can operate under

aqueous conditions, enabling them to degrade environmentally threatening water contaminants.

The catalytic mechanism has been well established (Figure I-1). The Fe(III)-TAML complex reacts with an oxidant (peracids, sodium hypochlorite or peroxides) and forms Fe(V)-oxo species via heterolytic cleavage of the O–X bond (X = Cl, O). Homolytic cleavage is only proposed for tBuOOH^[9]. The Fe(V)-oxo species undergoes fast comproportionation with a Fe(III)-TAML present in the reaction mixture to afford (Fe(IV))₂-μ-oxo species^[10]. This dimer species can be oxidized again to afford the Fe(V)-oxo species. The C–H bond abstraction mediated by this Fe(V)-oxo species affords a carbon-centered radical, which then forms the C–O bond with an oxygen source in the medium. It is known that the C–H bond abstraction by the (Fe(IV))₂-μ-oxo species is orders of magnitude slower than the Fe(V)-oxo species.

One of the variants, Fe-bTAML (b for biuret, Figure I-2) has been reported to catalyze selective oxidation of C–H bonds in various hydrocarbons, including complex natural products in an aqueous medium^[11]. Therefore, we selected this complex as the starting point of our investigation for the development of an artificial peroxidase.

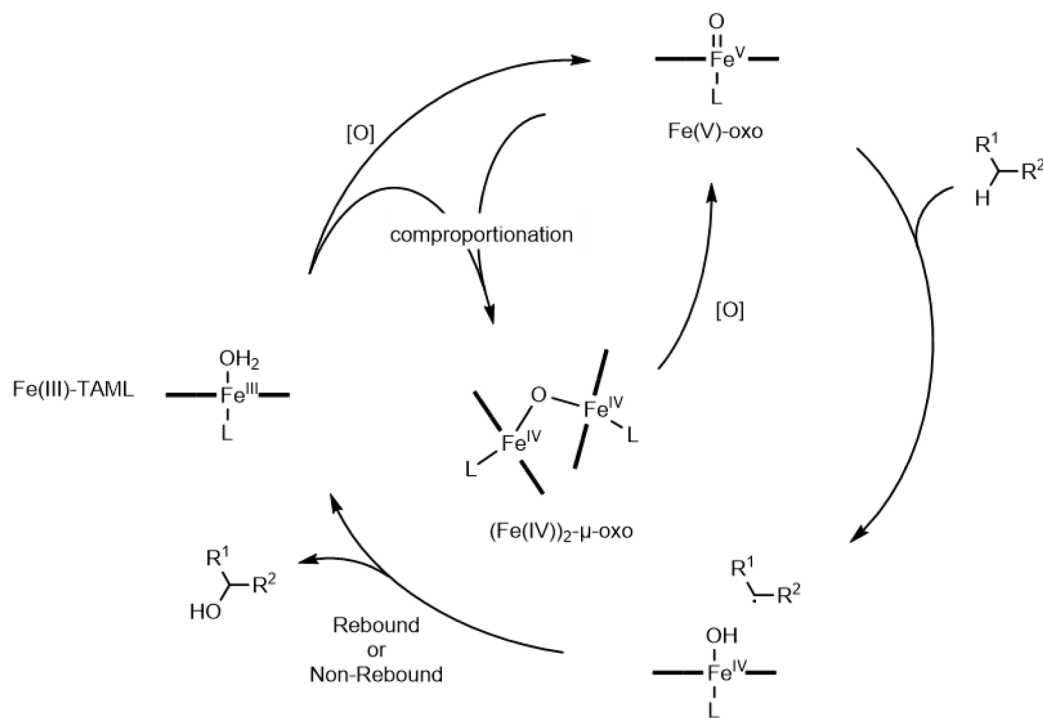


Figure I-1. Mechanism of C–H hydroxylation by a Fe-TAML complex.

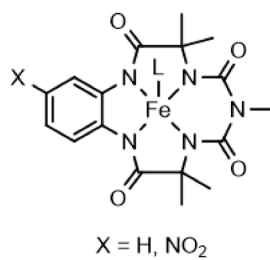


Figure I-2. The structure of the Fe-bTAML complexes.

1.1.3 References

- [1] a) K. J. Rossiter, *Chem. Rev.* **1996**, *96*, 3201-3240; b) J. M. Weber, J. O. Leung, S. J. Swanson, K. B. Idler, J. B. McAlpine, *Science* **1991**, *252*, 114-117; c) N. M. Davies, *Clin. Pharmacokinet.* **1998**, *34*, 101-154.
- [2] a) M. C. White, J. Zhao, *J. Am. Chem. Soc.* **2018**, *140*, 13988-14009; b) M. Costas, *Coord. Chem. Rev.* **2011**, *255*, 2912-2932; c) X. Engelmann, I. Monte-Pérez, K. Ray, *Angew. Chem. Int. Ed.* **2016**, *55*, 7632-7649.
- [3] B. Meunier, S. P. de Visser, S. Shaik, *Chem. Rev.* **2004**, *104*, 3947-3980.
- [4] V. Purpero, G. R. Moran, *J. Biol. Inorg. Chem.* **2007**, *12*, 587-601.
- [5] R. Banerjee, J. C. Jones, J. D. Lipscomb, *Annu. Rev. Biochem.* **2019**, *88*, 409-431.
- [6] P. H. Walton, G. J. Davies, *Curr. Opin. Chem. Biol.* **2016**, *31*, 195-207.
- [7] a) J. Chen, Z. Jiang, S. Fukuzumi, W. Nam, B. Wang, *Coord. Chem. Rev.* **2020**, *421*; b) M. Milan, M. Bietti, M. Costas, *Chem. Commun.* **2018**, *54*, 9559-9570.
- [8] T. J. Collins, A. D. Ryabov, *Chem. Rev.* **2017**, *117*, 9140-9162.
- [9] M. R. Mills, A. E. Burton, D. I. Mori, A. D. Ryabov, T. J. Collins, *J. Coord. Chem.* **2015**, *68*, 3046-3057.
- [10] S. Kundu, J. V. Thompson, A. D. Ryabov, T. J. Collins, *J. Am. Chem. Soc.* **2011**, *133*, 18546-18549.
- [11] S. Jana, M. Ghosh, M. Ambule, S. Sen Gupta, *Org. Lett.* **2017**, *19*, 746-749.

1.2 Artificial Metalloenzymes¹

1.2.1 Introduction

Artificial metalloenzymes (ArMs) are constructed by anchoring a small metal catalyst within a protein scaffold as a cofactor^[12]. This strategy localizes the catalyst within the protein, providing numerous benefits to the reaction due to the well-defined micro-environment surrounding the cofactor. These advantages encompass not only the well-defined chiral environment, but also substrate orientation and localization, catalyst protection, the presence of acid/base or redox-active residues, metal coordination by amino acids, and the influence of a local electric field^[13]. Moreover, advancements in molecular biology have facilitated protein engineering, including directed evolution^[14], thereby expediting the development of ArMs. The increasing repertoire of synthetic metal catalysts and the diversity of available protein scaffolds have enabled researchers to pursue the development of a variety of ArMs relying on different anchoring strategies. To date, ArMs have been shown to catalyze various challenging natural and non-natural reactions, including transfer hydrogenation, C–C bond formation, C–X bond formation, etc. Furthermore, the advancement in this field has facilitated the gradual transition towards *in vivo* abiotic metal catalysis^[15]. In this review, we aim to underscore recent advances in the field of ArMs, categorizing them based on the reactions they catalyze. Noteworthy *in vivo* studies will be highlighted in the relevant sections. The ArM studies covered herein are summarized in Table I-1.

¹ This section was written as a manuscript for a review article in Current Opinion in Chemical Biology. The manuscript is currently under revision.

Reaction	Metal cofactor	Protein	Anchoring strategy	Optimization ^a	In vivo	Ref
<p>Cyclopropanation</p> <p>Cyclopropanation</p> <p>R = Br, Cl, t-Bu, NO₂, CN, I, AcNH, AcO, MeO</p>	Ir(Me) ₂ -MPIX	CYP119 mutants	Supramolecular	Genetic	-	[16]
<p>Cyclopropanation</p> <p>Cyclopropanation</p> <p>Carvone (I-2)</p>	Ir(Me) ₂ -MPIX	CYP119-C317G/L155W/T213GN/254L	Supramolecular	-	-	[17]
<p>C-H carbene insertion</p> <p>C-H carbene insertion</p>	Biotinylated Cu(I) tris(pyrazolyl)borate complex	Streptavidin mutants	Supramolecular	Genetic	-	[18]
<p>Cyclopropanation</p> <p>Cyclopropanation</p> <p>styrene (I-5)</p>	Heme	Closed α helical solenoid repeat protein	Dative	Computational, genetic	-	[19]
<p>Ring-closing metathesis</p> <p>Ring-closing metathesis</p> <p>Umbelliprenin</p> <p>Prodrug I-8</p>	HG catalyst, linked to aminocoumarin via PEG	HSA	Supramolecular	Chemical	+	[20]

Table I-1. Summary of all the ArMs covered in this review. ^aChemical optimization includes the optimization of the ligand, spacer or anchor structure as well as reaction condition. Genetic optimization includes mutation of the protein scaffold. Computational optimization includes a large-scale *in silico* screening of the protein scaffold. Ir(Me)₂-MPIX = Ir(Me)₂-mesoporphyrin IX, CYP119 = Cytochrome P450 119, HG = Hoveyda-Grubbs, PEG = polyethylene glycol, HSA = human serum albumin.

Reaction	Metal cofactor	Protein	Anchoring strategy	Optimization ^a	In vivo	Ref
<p>28 examples with tosyl, benzoyl, ether, benzylic ketone, and halogenated moieties 11 among them were followed by 1,4-elimination</p> <p>Homo coupling: </p> <p>Cross coupling: </p> <p>Prodrug I-9 </p> <p>I-10 </p> <p>I-11 </p>	<p>Iodine-substituted HG catalyst, linked to aminocoumarin via PEG</p>	HSA	Supramolecular	Chemical	+	[22]
<p>Ring-closing/opening metathesis</p>	HG catalyst linked to 6-chlorohexanoic acid	HaloTag-SNAP1 tag fusion protein	Covalent	-	+	[23]
<p>Michael addition</p>	p-AF and Cu(II)-phen	LmrR mutants	Genetic incorporation (pAF) and supramolecular (Cu(II)-phen)	Genetic	-	[24]

Table I-1. Summary of all the ArMs covered in this review (continued). pAF = *p*-aminophenylalanine, Cu(II)-phen = Cu(II)-phenanthroline, LmrR = lactococcal multidrug resistance regulator.

Reaction	Metal cofactor	Protein	Anchoring strategy	Optimization*	In vivo	Ref																		
Friedel-Crafts alkylation, Diels-Alder reaction	Cu(II)-phen	LmrR mutants	Supramolecular	Genetic	+	[25]																		
							<table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>R¹</td> <td>R²</td> </tr> <tr> <td>Me</td> <td>H</td> </tr> <tr> <td>H</td> <td>OMe</td> </tr> <tr> <td>H</td> <td>H</td> </tr> <tr> <td>H</td> <td>Cl</td> </tr> </table>	R ¹	R ²	Me	H	H	OMe	H	H	H	Cl							
							R ¹	R ²																
Me	H																							
H	OMe																							
H	H																							
H	Cl																							
Cycloaddition	Cp*Rh(III) catalyst linked to maleimide	Chimeric nitrobindin	Covalent	Genetic	-	[26]																		
Diazo cross-coupling	Dirhodium catalyst linked to an exo-bicyclo[6.1.0]nonyne	POP mutants	Covalent	Genetic	-	[27]																		
							<table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>R¹</td> <td>R²</td> <td>R³</td> </tr> <tr> <td>OMe</td> <td>OMe</td> <td>OEt</td> </tr> <tr> <td>H</td> <td>OMe</td> <td>OEt</td> </tr> <tr> <td>Cl</td> <td>OMe</td> <td>OEt</td> </tr> <tr> <td>Br</td> <td>OMe</td> <td>OEt</td> </tr> <tr> <td>OMe</td> <td>OMe</td> <td>NEt₂</td> </tr> <tr> <td>OMe</td> <td>OMe</td> <td>OBn</td> </tr> <tr> <td>Cl</td> <td>Me</td> <td>OEt</td> </tr> </table>	R ¹	R ²	R ³	OMe	OMe	OEt	H	OMe	OEt	Cl	OMe	OEt	Br	OMe	OEt	OMe	OMe
R ¹	R ²	R ³																						
OMe	OMe	OEt																						
H	OMe	OEt																						
Cl	OMe	OEt																						
Br	OMe	OEt																						
OMe	OMe	NEt ₂																						
OMe	OMe	OBn																						
Cl	Me	OEt																						
N-H carbene insertion	Ir(Me)-MPIX	CYP119 mutants	Supramolecular	Genetic	+	[29]																		
C-H nitrene insertion	Biotinylated [Cp*Ir(III)(Boc-amidoquinoline)Cl]	Streptavidin mutants	Supramolecular	Chemical, genetic	-	[30]																		
							<p>R₁ = Me, F, Br, MeO, F₃CO, I, MeO₂C, Ph R₂ = Me, F, CF₃</p>																	

Table I-1. Summary of all the ArMs covered in this review (continued). POP = prolyl oligopeptidase.

Reaction	Metal cofactor	Protein	Anchoring strategy	Optimization ^a	In vivo	Ref
C-H hydroxylation		<i>Pseudomonas aeruginosa</i> azurin mutant	Dative	Genetic	-	[31]
	<p>Cu(II) bound to a histidine-brace</p>					
C-H hydroxylation		Streptavidin mutants	Supramolecular	Chemical, genetic	-	[32]
	<p>Biotinylated [Fe(TAML)]⁻ catalyst</p> <p>R¹ = Me, Et, nPr, iPr R² = OMe, Cl, Br</p>					
Nucleophilic substitution		<i>p</i> -boronophenylalanine	Genetic incorporation	Genetic	-	[33]
	<p>Benzoin:</p>					

Table I-1. Summary of all the ArMs covered in this review (continued). TAML = tetra amido macrocyclic ligand, bpy = 2,2'-bipyridine, dppz = dipyrido[3,2-*a*:2',3'-*c*]phenazine.

Reaction	Metal cofactor	Protein	Anchoring strategy	Optimization ^a	In vivo	Ref
<p>Photo-SET/ET</p> <p>R = H, , </p>	[Ru(bpy) ₂ dppz] ²⁺	Apo riboflavin-binding protein	Supramolecular	-	-	[34]
<p>Photoreductive cyclization, photocycloaddition</p> <p>I-30 + I-31 + I-32 + I-33</p>	[Ru(bpy) ₃] ²⁺ derivatives	POP mutants	Covalent or supramolecular	Chemical, genetic	-	[35]
<p>Photocycloaddition</p> <p>I-34 + I-35 + I-36 + I-37</p>	[Ir(ppy) ₂ bpy] ³⁺	POP mutant	Covalent	Chemical	-	[36]
<p>Sulfite reduction</p> <p>SO₃²⁻ + 6 e⁻ + 7 H⁺ → HS⁻ + 3 H₂O</p>	Heme b and [4Fe-4S] cluster	Cytochrome c peroxidase	Dative	Genetic	-	[37]
<p>CO₂ reduction</p> <p>CO₂ + C₁₁H₁₆ → </p>	Doubly biotinylated [4Fe-4S] cluster	Streptavidin mutants	Supramolecular	Chemical, genetic	-	[38]
<p>Transfer hydrogenation</p> <p>I-33 + HCO₂Na → </p>	[Cp*Ir(III)(aminosulfonamide)Cl] catalyst linked to an optimized anchor	<i>Thermoanaerobacter brookii</i> alcohol dehydrogenase wild type and a mutant	Supramolecular	Computational, chemical	-	[39]

Table I-1. Summary of all the ArMs covered in this review (continued). ppy = 2-phenylpyridine.

1.2.2 C–C Bond Formation

1.2.2.1 Carbene C–H Insertion and cyclopropanation

An iridium-porphyrin complex serves as an efficient and selective carbene transfer catalyst when integrated into a protein scaffold. Hartwig and coworkers investigated the impact of CYP119 (a cytochrome P450 monooxygenase) mutant scaffolds on the site selectivity of C–H carbene insertion reactions^[16]. They observed that the ArMs composed of a CYP119-T213G-C317G double mutant and Ir(Me)-MPIX (MPIX = mesoporphyrin IX) catalyzed the C–H insertion of 4-substituted phthalans **I-1** with negligible site selectivity. Through directed evolution, they identified two series of variants favoring carbene insertion into either para- (3.2:1 d.r.) or meta- (1:3.4 d.r.) benzylic C–H bond relative to the substituent on the aromatic ring (resulting in products **I-1a** and **I-1b**, respectively). Additionally, they found that peripheral mutations, distant from the active site, enhanced the turnover number (TON) of the reaction. Characterization with circular dichroism (CD) spectrometry, differential scanning calorimetry, and control enzymatic assays revealed that these peripheral mutations did not affect the thermal stability of the protein scaffold but improved its stability against the reaction products. This illustrates how subtle variations in the protein scaffold, even at distant positions, can enhance the catalytic performance by modulating the protein's stability in the reaction conditions.

Interestingly, Ir(Me)-MPIX-CYP119 exhibits unique kinetics. It was observed that there is an approximately 30-second induction period in the formation of the iridium-carbene intermediate during the cyclopropanation of carvone (**I-2**)^[17]. Detailed kinetic studies and X-ray crystallography revealed that the Ir(Me)-MPIX cofactor exists in equilibrium between an active and an inactive binding mode, necessitating a transition to the latter before the reaction begins. Upon addition of the carbene source, ethyl diazoacetate (EDA), the equilibrium shifts towards the active form after a certain induction period, resulting in the sigmoidal progression of the reaction over time. This represents a unique kinetic behavior of ArMs resulting from the combination of an artificial cofactor and a protein scaffold.

The widely used streptavidin (Sav)-biotin technology was employed for a carbene insertion reaction. Ward and coworkers described intramolecular C–H carbene insertion using a biotinylated Cu(I) tris(pyrazolyl)borate complex^[18]. They genetically engineered the ArM (biotC4–TazCu·Sav mutants) by mutating the close-lying residues near the active site, namely S112, K121, and L124. The resulting ArMs, double or triple mutants, exhibited significant total TON (TTON, >1000) toward cyclization of diazoacetamides **I-3** and **I-4** in general. Notably, they obtained ArM variants that predominantly formed β -lactam via non-benzylic C–H activation (**I-3b**, **I-4b**), highlighting the ArM's capacity to catalyze challenging carbene insertions. The best-performing variant demonstrated catalytic activity on a tertiary C–H bond with a TTON of 4627 and a yield of 92.5% (**I-4b**).

Furthermore, a carbene transfer reaction was accomplished on a *de novo*-designed protein scaffold. Green, Baker and coworkers initiated the design process based on a previously reported hyperstable structure, α -helical solenoid repeat architecture, and designed a heme binding site in the center of the protein^[19], where a histidine side chain was introduced for heme coordination. Experimental characterization of the designed ArMs with ultraviolet-visible (UV/Vis) spectra indicated successful coordination of histidine to the heme in one of the designs. The resulting hyperstable ArM enabled substantial engineering of the residues near the active site without compromising overall protein stability. Through either directed evolution or computational design alone, they successfully engineered the ArM towards peroxidase or stereoselective carbene transferase, respectively. The best carbene transferases catalyzed the formation of cyclopropane **I-5a** from styrene (**5**) and EDA with 99.5: 0.5 d.r. and 3:97 e.r. for the (*S,S*)-**I-5a**, and 99: 1 d.r. and 86:14 e.r. for the (*R,R*)-**I-5a**.

1.2.2.2 Olefin Metathesis

A deep hydrophobic pocket within a protein can serve as an ideal environment for protecting the catalyst from other potentially inhibitory molecules in the reaction medium, particularly within biological contexts. Tanaka and coworkers reported an enhanced biocompatibility of the Hoveyda-Grubbs ruthenium (HG) catalyst by incorporating it into human serum albumin (HSA), showcasing its therapeutic potential^[20]. Upon confirming the ability of the HG catalyst to bind

to the hydrophobic drug site of HSA via an aminocoumarin anchor, they demonstrated that the ArM can catalyze olefin (**I-6**) and ene-yne (**I-7**) metathesis even in the presence of up to 20 mM glutathione (GSH), a significant poison for metal catalysts in cytoplasm. By comparing the performance of the ArM with that in the presence of dodecanethiol (DdSH), they suggested that this protective effect against GSH largely arises from charge repulsion near the entrance of the binding site. Additionally, they glycosylated the surface of the ArM and showed that the glycosylated ArM (GARm) accumulated on the surface of SW620, A549, and HeLa cancer cells. They further demonstrated that GARm-catalyzed activation of prodrug **I-8** to form umbelliprenin was effective against all three cancer cell lines, whereas the free cofactor exhibited comparably lower efficiency. Subsequently, inspired by the study by Skowerski and coworkers,^[21] they discovered that substituting the chlorine ligands in the HG catalyst with iodine resulted in a more active ArM^[22] with enhanced stability in blood. Upon surface modification of the HSA with a cyclic-Arg-Gly-Asp (cRGD) pentapeptide to target cancer cells, the ArM suppressed the growth of implanted tumors in mice by activating prodrug **I-9** more strongly than the drug itself. Recently, Bai and coworkers introduced a novel strategy for protecting ArMs within cellular environments. They investigated *in vivo* ArM-catalyzed reactions leveraging liquid-liquid phase separation (LLPS) within *E. coli* cells^[23]. By employing HaloTag-SNAPTag fusion proteins along with protein-specific chemical cross-linkers, they achieved compartmentalization of a metal active site within a liquid condensate in the cytoplasm. Upon incorporating a HG catalyst containing an anchor for HaloTag, they demonstrated ring-closing metathesis of olefin **I-10** and ring-opening metathesis of cyclic alkene **I-11** within the liquid condensate. Furthermore, they illustrated a cascade of Cu-mediated alkyne-azide cycloaddition and ring-closing metathesis occurring within the liquid condensate, underscoring the robustness of LLPS-assisted catalysis.

1.2.2.3 Michael Addition

Synergistic catalysis, wherein two substrates are activated by distinct catalysts, greatly benefit from confinement within the protein's active site. While such mechanisms are prevalent in natural enzymes, designing an ArM to achieve this synergy poses a general challenge. Nevertheless, Roelfes and coworkers reported an ArM capable of performing a synergistic

Michael addition facilitated by a Lewis-acidic copper and a *p*-aminophenylalanine (pAF)^[24]. They introduced a pAF residue into the lactococcal multidrug resistance regulator (LmrR) by expressing *p*-azidophenylalanine (pAzF) using an orthogonal translation system followed by Staudinger reduction. Cu(1,10-phenanthroline)(NO₃)₂ (Cu(II)-phen) was incorporated by supramolecular assembly into the binding pocket of LmrR. The resulting ArM catalyzed a synergistic Michael addition using 2-acyl-1-methylimidazole **I-12** and an α,β -unsaturated aldehyde **I-13**. Notably, they observed that the reaction proceeds with good enantioselectivity (>80%) only when all reaction components were present. Method optimization using crotonaldehyde (**I-13** where R² = Me) as the acceptor resulted in decent yield and high e.e. (54% yield, d.r. 3.5:1, 98/82% e.e.). A mutation study further enhanced the yield up to 82% and e.e. up to 99/93% for each diastereomer. This study exemplified the advantage of placing multiple catalytic components close together within a protein scaffold to perform synergistic catalysis.

1.2.2.4 Friedel-Crafts Alkylation/Diels-Alder Reaction

More recently, the supramolecular assembly of LmrR and Cu(II)-(phen) complex has been demonstrated to occur *in vivo*. Roelfes and coworkers unambiguously established assembly formation within *E. coli* cells through enzymatic activity assays, inductively coupled plasma-optical emission spectrometry, cell fractionation, inhibitor assays, and in-cell nuclear magnetic resonance (NMR) studies^[25]. The strong correlation between *in vivo* and *in vitro* activity enabled them to conduct directed evolution of the ArM using whole cells. Screening of the whole cells for Friedel-Crafts alkylation resulted in improved yields and the e.e. for substrates **I-14** and **I-15**. The catalytic efficiency of the ArM increased nearly threefold (from 73.3 M⁻¹min⁻¹ to 185.9 M⁻¹min⁻¹). Directed evolution for Diels-Alder reaction with azachalcone **I-16** and cyclopentadiene **I-17** resulted in a double mutant LmrR_A92E_V15A, exhibiting a moderate yield (52%) and good e.e. (84%).

1.2.2.5 Cycloaddition

One of the challenges encountered when using non-enzyme protein scaffolds for ArMs is that their active site architecture often tends to be more open compared to natural enzymes. One potential approach to address this problem is to graft a motif from other proteins onto the scaffold to create a chimeric protein capable of more efficiently shielding the active site. Onoda, Hayashi and coworkers introduced an ArM based on a chimeric nitrobindin (NB) protein scaffold harboring a Cp*Rh(III) complex^[26]. They grafted a loop domain from a fatty acid binding protein around the entrance of the binding pocket of NB. An initial small-scale screening of the chimeric structure followed by directed evolution resulted in a significant enhancement in catalytic efficiency in the Rh-catalyzed cycloaddition of the oxime **I-18** and the alkyne **I-19**. The best-performing variant, NBHLH1 (Y119A / G149P), demonstrated a more than 35-fold increase in k_{cat}/K_m compared to wild-type NB, thus clearly illustrating the advantage of shielding the active site with a grafted motif.

1.2.2.6 Diazo Cross-Coupling

Prolyl oligopeptidase (POP), as a natural enzyme with a large cavity, has proven to be versatile as a scaffold for ArM. Lewis and coworkers demonstrated that an ArM based on the POP scaffold and a dirhodium complex could be integrated in a biocatalytic cascade, tolerating numerous reaction components^[27]. These ArMs were constructed via strain-promoted azide-alkyne cycloaddition (SPAAC) between an *exo*-bicyclo[6.1.0]nonyne anchor and a pAzF residue incorporated by an orthogonal translation system. Following initial screening for ideal diazo coupling reactivity, they improved the chemo- and stereoselectivity of the reaction through directed evolution employing NNK codons and combinatorial codon mutagenesis. The resulting ArM achieved a remarkable TON of 44612 for the desired *E*-alkene and *E/Z* ratio of 14.9. Cascade reactions performed in combination with an ene reductase and a glucose dehydrogenase yielded enantioenriched succinate derivatives, with most tested substrates **I-20** and **I-21** exhibiting moderate yield and excellent e.e. This indicates that the ArM does not have a negative effect on the catalytic activity of the other two enzymes. Furthermore, the minimal occurrence of water O–H insertion of the Rh-carbene intermediate, along with the

absence of either glucose O–H insertion or protein modification, highlights the clear advantage of the deep cavity in efficiently isolating the catalyst from other reaction components, including the bulk water. The authors hypothesized that the closed state of the protein scaffold is favored due to a plausible Rh-histidine cross linking, which was supported by molecular dynamics simulations.

1.2.3 C–X Bond Formation

1.2.3.1 N–H Carbene Insertion

Several biocatalytic carbene insertion reactions have been reported; however, the inherent reactivity of the metal-carbene intermediate hampers its application in cellular contexts. Nevertheless, Tomás-Gamasa and Mascareñas reported a copper-catalyzed N–H insertion reaction of a carbene to form benzoquinoxalines in mammalian cells for the first time^[28]. These studies would accelerate *in vivo* studies of ArMs in terms of therapeutic applications.

Hartwig and coworkers capitalized on the *in vivo* N–H insertion of a carbene for efficient directed evolution campaign of their CYP119 harboring Ir(Me)MPIX^[29]. By coexpressing a heme transport system with the protein scaffold, they established a whole-cell screening method enabling the screening of over a thousand mutants within less than two weeks. With this protocol, they screened more than 4000 colonies to obtain a reactive and enantioselective ArM toward insertion of a carbene into N–H bonds of *N*-methyl anilines **I-22**. Together with condition optimization, they achieved 82% e.e. and 8569 TON.

1.2.3.2 Nitrene Insertion

Nitrene insertion is also within the scope of nonnative biocatalytic reactions. Recently, Ward and coworkers reported an artificial nitrene insertase (ANlase) based on mutants of Sav and a biotinylated Cp*Ir-pianostool cofactor^[30]. They selected a dioxazolone **I-23a** as the model substrate for the chemogenetic optimization of the ANlase. The best performing variants, either single or double mutants, catalyzed intramolecular nitrene insertion into both benzylic and aliphatic C–H bonds with γ -lactam preference, with up to 92% e.e. On the basis of quantum mechanics-molecular mechanics (QM-MM) studies, they hypothesized that the

enantioselectivity stems from hydrogen bonding between the carbonyl group in the substrate and the cationic residue at position 121 (either lysine or arginine).

1.2.3.3 C–H Hydroxylation

C–H hydroxylation is a prevalent natural reaction primarily catalyzed by heme-, non-heme iron, or copper-dependent enzymes. Lu and coworkers reported an artificial lytic polysaccharide monooxygenase (LPMO) based on *Pseudomonas aeruginosa* azurin that harbors a non-native Type-2 copper site^[31]. Following the characterization of the active site with UV/Vis and electron paramagnetic resonance spectroscopy, they confirmed its ability to activate both O₂ and H₂O₂. The ArM was demonstrated to catalyze the oxidative cleavage/depolymerization of model substrates α - and β -4-nitrophenyl-D-glucopyranoside (α -/ β -4-NPGP) (**I-24**), phosphoric acid swollen cellulose, and starch (**I-25**). In all the cases, multiple turnover was detected. Especially, the ArM showed 52 TON in an optimized condition with β -4-NPGP.

Ward and coworkers reported artificial hydroxylases based on a synthetic iron complex^[32]. They synthesized a cofactor by attaching a biotin linker to [Fe(TAML)]⁻ complex (TAML = Tetra Amido Macrocyclic Ligand) to ensure its binding into the cavity of Sav. Screening of the Sav mutants indicated the effects of surrounding amino acid residues on both TTON and e.e. for the hydroxylation of benzylic C–H bonds (compounds **I-26**, **I-27**, and **I-28**). Further investigation revealed that the ArM performs kinetic resolution of ethylbenzene and tetralin. Notably, the reaction with tetralin in the presence of excess H₂O₂ resulted in >99% e.e. after kinetic resolution. Additionally, they demonstrated that the reaction can occur in an enzymatic cascade with a glucose oxidase to provide H₂O₂ in the reaction mixture, enabling the reaction to proceed with only molecular oxygen as the terminal oxidant.

1.2.3.4 Enantioselective Nucleophilic Substitution

Boron, classified as a metalloid, had not been previously applied in ArMs until recently. Roelfes and coworkers reported on an ArM relying on *p*-boronophenylalanine (pBoF) incorporation into LmrR by an orthogonal translation system^[33]. After identifying the optimal position for pBoF, they conducted directed evolution toward kinetic resolution in nucleophilic oxime formation

from a benzoin substrate. The final variant, BOS-EHL, exhibited high E values (up to 147) and 33-fold improvement in the rate constant. High resolution mass spectrometry (HRMS), ^{11}B NMR and X-ray crystallography suggested that the reaction proceeds through an sp^3 hybridized cyclic boronate ester intermediate.

1.2.4 Photocatalysis

The properties of photocatalysts can be fine-tuned through incorporation into a protein scaffold due to their sensitivity to the surrounding environments. Okamoto, Sato and coworkers demonstrated that the photocatalytic activity of $[\text{Ru}(\text{bpy})_2\text{dppz}]^{2+}$ (bpy = 2,2'-bipyridine, dppz = dipyrido[3,2-a:2',3'-c]phenazine) can be switched by incorporating it into a riboflavin-binding protein^[34]. They assembled the ArM by hydrophobic intercalation of the dppz ligand into a riboflavin binding site of the protein with submicromolar affinity. This ArM catalyzed photocatalytic NADH oxidation more slowly than the free cofactor and also catalyzed singlet oxygen production, unlike the free cofactor under the tested conditions. They proposed that the limited access of NADH to the dppz ligand, as well as the hydrophobic environment surrounding the ligand, prevented the non-emissive decay, thereby switching the reactivity from electron transfer photocatalysis to energy transfer catalysis. They demonstrated that this switching behavior can be applied for selective protein labeling using 1-methyl-4-arylurazole **I-29**.

Lewis and coworkers also explored the effect of $[\text{Ru}(\text{bpy})_3]^{2+}$ binding in the POP scaffold in its photochemical properties^[35]. They hypothesized that the large and closed cavity of the protein allows for the control in the catalyst's surroundings. They found that both covalent linking with SPAAC and noncovalent binding without any linker took place. Mutation in the binding pocket improved the cofactor binding, presumably via charge compensation, with the best protein host exhibited a dissociation constant (K_d) of 1.08 μM without any linker. In these ArMs, the hydrophobic environment elongated the luminescence lifetime of $[\text{Ru}(\text{bpy})_3]^{2+}$. They demonstrated that both the photoreductive cyclization of olefin **I-30** and the cycloaddition of the styrenes **I-31** and **I-32** were accelerated when $[\text{Ru}(\text{bpy})_3]^{2+}$ was embedded within the

protein host. This enhanced catalytic activity was attributed to the higher local concentration of both the substrate and the photocatalyst and the longer-lived excited state.

Similarly, they reported that the photochemical properties of iridium(III) polypyridine complexes could be tuned with the POP scaffold^[36]. Incorporation of the cofactors into the binding pocket of the protein caused elongation of the phosphorescence lifetime and an increase in the triplet excited state energy. Additionally, one of the cofactors underwent partial resolution upon binding to the protein pocket, resulting in improved e.e. compared to the free cofactor (14.5% and -0.1%, respectively).

1.2.5 Oxyacid Reduction

Iron-sulfur clusters, along with heme, represent important biological redox-active cofactors. Lu and coworkers reported on an artificial sulfite reductase based on a cytochrome c peroxidase scaffold^[37]. By incorporating a [4Fe-4S] cluster and a heme b into the protein cavity, they constructed a structural and functional model of a natural sulfite reductase. The ArM exhibited the activity in sulfite reduction in the presence of an artificial electron donor, methylviologen. They rationally engineered the protein scaffold by introducing cationic residues to stabilize the sulfite and amine residues around the [4Fe-4S] cluster. The final variant, W51K/H52R/P145K-SiRCcP.3, showed 63-fold increase in activity compared to the initial variant, representing the first substantially active functional model of natural sulfite reductases.

Ward and coworkers reported an ArM that catalyzes the Fischer-Tropsch reaction based on an embedded [4Fe-4S] cluster in a cavity of Sav^[38]. The synthesized cofactor contained a single [4Fe-4S] cluster with two biotin linkers for double anchoring of the complex in the binding cavity. They established that 2 equivalents of the cofactor were bound to the Sav tetramer based on HRMS and CD titration. Cyclic voltammetry indicated that the protein environment within the active site influenced the tuning of the redox potential of the cluster. The resulting ArM catalyzed CO₂ reduction, yielding short-chain alkanes from C1 to C4, with the better TTON than only the free cofactor or the free cofactor in the presence of bovine serum albumin. Chemical and genetic optimization of the ArM by changing the spacer identity and single mutations suggested the possibility of fine-tuning the activity upon further optimization.

1.2.6 Transfer Hydrogenation

The search for new protein-anchor pairs is the key to broadening the diversity of ArMs. Pordea, Jäger and coworkers introduced a computational screening workflow aimed at identifying a suitable supramolecular anchor for an alcohol dehydrogenase from *Thermoanaerobacter brockii* (TbADH)^[39]. The resultant cofactor, bearing an anchoring group and an iridium piano-stool complex, bound to the wild type enzyme with the K_d up to 52 μM . The constructed ArM catalyzed the transfer hydrogenation of salsolidine (**I-33**), albeit with a lower TON than the free cofactor and lacking enantioselectivity. This study underscores the capacity of computational screening in efficiently discovering a new protein-anchor pair for ArMs.

1.2.7 Conclusions

Since the first report of ArMs in the late 1970s^[40], the field has been growing in terms of the variety of chemical reactions, protein-anchor pairs, metal complexes, and the reaction media. The advancement in DNA synthesis and molecular biology has facilitated the efficient engineering of protein scaffolds. A large set of data obtained by such engineering can be nowadays used to train a machine learning program to extract the key patterns in protein mutation and accelerate the engineering process^[41]. Moreover, recent notable progress in *de novo* protein design holds the potential to greatly expand the scope of ArMs^[42].

As outlined in this review, achieving *in vivo* catalysis with ArMs represents a crucial milestone in the field, and we anticipate that increasing efforts will be made toward this goal. Advances in both organometallic catalysis and computational science, including protein design, can help push the boundaries of ArMs into the complex biological contexts.

1.2.8 References

- [12] a) F. Schwizer, Y. Okamoto, T. Heinisch, Y. Gu, M. M. Pellizzoni, V. Lebrun, R. Reuter, V. Köhler, J. C. Lewis, T. R. Ward, *Chem. Rev.* **2018**, *118*, 142-231; b) J. Bos, G. Roelfes, *Curr. Opin. Chem. Biol.* **2014**, *19*, 135-143; c) M. T. Reetz, *Acc. Chem. Res.* **2019**, *52*, 336-344; d) B. J. Bloomer, D. S. Clark, J. F. Hartwig, *Biochemistry* **2023**, *62*, 221-228; e) C. Van Stappen, Y. Deng, Y. Liu, H. Heidari, J. X. Wang, Y. Zhou, A. P. Ledray, Y. Lu, *Chem. Rev.* **2022**, *122*, 11974-12045; f) K. Oohora, A. Onoda, T. Hayashi, *Acc. Chem. Res.* **2019**, *52*, 945-954; g) K. Vong, T. Yamamoto, K. Tanaka, *Small* **2020**, *16*, 1906890; h) C. Marchi-Delapierre, L. Rondot, C. Cavazza, S. Ménage, *Isr. J. Chem.* **2015**, *55*, 61-75; i) Y. Lu, N. Yeung, N. Sieracki, N. M. Marshall, *Nature* **2009**, *460*, 855-862; j) F. Yu, V. M. Cangelosi, M. L. Zastrow, M. Tegoni, J. S. Plegaria, A. G. Tebo, C. S. Mocny, L. Ruckthong, H. Qayyum, V. L. Pecoraro, *Chem. Rev.* **2014**, *114*, 3495-3578; k) B. Large, N. G. Baranska, R. L. Booth, K. S. Wilson, A.-K. Duhme-Klair, *Curr. Opin. Green Sustainable Chem.* **2021**, *28*, 100420; l) A. Lombardi, F. Pirro, O. Maglio, M. Chino, W. F. DeGrado, *Acc. Chem. Res.* **2019**, *52*, 1148-1159; m) Y. Yu, X. Liu, J. Wang, *Acc. Chem. Res.* **2019**, *52*, 557-565; n) W. J. Song, F. A. Tezcan, *Science* **2014**, *346*, 1525-1528; o) T. Ueno, S. Abe, N. Yokoi, Y. Watanabe, *Coord. Chem. Rev.* **2007**, *251*, 2717-2731.
- [13] S. D. Fried, S. G. Boxer, *Annu. Rev. Biochem.* **2017**, *86*, 387-415.
- [14] Y. Wang, P. Xue, M. Cao, T. Yu, S. T. Lane, H. Zhao, *Chem. Rev.* **2021**, *121*, 12384-12444.
- [15] M. Wittwer, U. Markel, J. Schiffels, J. Okuda, D. F. Sauer, U. Schwaneberg, *Nat. Catal.* **2021**, *4*, 814-827.
- [16] Y. Gu, S. N. Natoli, Z. Liu, D. S. Clark, J. F. Hartwig, *Angew. Chem. Int. Ed.* **2019**, *58*, 13954-13960.
- [17] B. J. Bloomer, S. N. Natoli, M. Garcia-Borràs, J. H. Pereira, D. B. Hu, P. D. Adams, K. N. Houk, D. S. Clark, J. F. Hartwig, *Nat. Catal.* **2023**, *6*, 39-51.
- [18] C. Rumo, A. Stein, J. Klehr, R. Tachibana, A. Prescimone, D. Häussinger, T. R. Ward, *J. Am. Chem. Soc.* **2022**, *144*, 11676-11684.
- [19] I. Kalvet, M. Ortmayer, J. Zhao, R. Crawshaw, N. M. Ennist, C. Levy, A. Roy, A. P. Green, D. Baker, *J. Am. Chem. Soc.* **2023**, *145*, 14307-14315.
- [20] S. Eda, I. Nasibullin, K. Vong, N. Kudo, M. Yoshida, A. Kurbangalieva, K. Tanaka, *Nat. Catal.* **2019**, *2*, 780-792.
- [21] A. Tracz, M. Matczak, K. Urbaniak, K. Skowerski, *Beilstein J. Org. Chem.* **2015**, *11*, 1823-1832.
- [22] I. Nasibullin, H. Yoshioka, A. Mukaimine, A. Nakamura, Y. Kusakari, T.-C. Chang, K. Tanaka, *Chem. Sci.* **2023**, *14*, 11033-11039.
- [23] T. Wu, X. Chen, H. G., Y. Deng, A. Yang, X. Feng, Y. Bai, *ChemRxiv* **2022**. <https://doi.org/10.26434/chemrxiv-2021-b6098>
- [24] Z. Zhou, G. Roelfes, *Nat. Catal.* **2020**, *3*, 289-294.
- [25] S. Chordia, S. Narasimhan, A. Lucini Paioni, M. Baldus, G. Roelfes, *Angew. Chem. Int. Ed.* **2021**, *60*, 5913-5920.
- [26] S. Kato, A. Onoda, U. Schwaneberg, T. Hayashi, *J. Am. Chem. Soc.* **2023**, *145*, 8285-8290.
- [27] D. M. Upp, R. Huang, Y. Li, M. J. Bultman, B. Roux, J. C. Lewis, *Angew. Chem. Int. Ed.* **2021**, *60*, 23672-23677.
- [28] S. Gutiérrez, M. Tomás-Gamasa, J. L. Mascareñas, *Angew. Chem. Int. Ed.* **2021**, *60*, 22017-22025.
- [29] Y. Gu, B. J. Bloomer, Z. Liu, R. Chen, D. S. Clark, J. F. Hartwig, *Angew. Chem. Int. Ed.* **2022**, *61*, e202110519.
- [30] K. Yu, Z. Zou, N. V. Igareta, R. Tachibana, J. Bechter, V. Köhler, D. Chen, T. R. Ward, *J. Am. Chem. Soc.* **2023**, *145*, 16621-16629.
- [31] Y. Liu, K. A. Harnden, C. Van Stappen, S. A. Dikanov, Y. Lu, *Proceedings of the National Academy of Sciences* **2023**, *120*, e2308286120.

- [32] J. Serrano-Plana, C. Rumo, J. G. Rebelein, R. L. Peterson, M. Barnet, T. R. Ward, *J. Am. Chem. Soc.* **2020**, *142*, 10617-10623.
- [33] L. Longwitz, R. B. Leveson-Gower, H. J. Rozeboom, A.-M. W. H. Thunnissen, G. Roelfes, *ChemRxiv* **2023**. <https://doi.org/10.26434/chemrxiv-2023-10vcv>
- [34] Y. Okamoto, T. Mabuchi, K. Nakane, A. Ueno, S. Sato, *ACS Catal.* **2023**, *13*, 4134-4141.
- [35] Y. S. Zubi, B. Liu, Y. Gu, D. Sahoo, J. C. Lewis, *Chem. Sci.* **2022**, *13*, 1459-1468.
- [36] B. Liu, Y. S. Zubi, J. C. Lewis, *Dalton Trans.* **2023**, *52*, 5034-5038.
- [37] E. N. Mirts, I. D. Petrik, P. Hosseinzadeh, M. J. Nilges, Y. Lu, *Science* **2018**, *361*, 1098-1101.
- [38] V. Waser, M. Mukherjee, R. Tachibana, N. V. Igarreta, T. R. Ward, *J. Am. Chem. Soc.* **2023**, *145*, 14823-14830.
- [39] Floriane L. Martins, A. Pordea, C. M. Jäger, *Faraday Discuss.* **2022**, *234*, 315-335.
- [40] M. E. Wilson, G. M. Whitesides, *J. Am. Chem. Soc.* **1978**, *100*, 306-307.
- [41] a) P. Kouba, P. Kohout, F. Haddadi, A. Bushuiev, R. Samusevich, J. Sedlar, J. Damborsky, T. Pluskal, J. Sivic, S. Mazurenko, *ACS Catal.* **2023**, 13863-13895; b) K. K. Yang, Z. Wu, F. H. Arnold, *Nat. Methods* **2019**, *16*, 687-694.
- [42] S. L. Lovelock, R. Crawshaw, S. Basler, C. Levy, D. Baker, D. Hilvert, A. P. Green, *Nature* **2022**, *606*, 49-58.

2 Goal of the Thesis

As introduced in Chapter 1, stereoselective C–H hydroxylation is challenging. In green chemistry context, development of a catalyst that operates under aqueous conditions with minimal amount of waste is desired. Thus, capitalizing on the activity of the Fe-bTAML catalyst in aqueous conditions in the presence of H₂O₂, we set out our investigation in the development of an artificial peroxidase (ArPase).

The initial goal of this project was the development of an ArPase with the following properties:

- 1) Utilizes H₂O₂ as the oxidant
- 2) Exhibits high TTON (> 100)
- 3) Catalyzes enantioselective C–H hydroxylation

To this end, we constructed an ArPase by incorporation of a Fe-bTAML catalyst into a human carbonic anhydrase II protein. Upon confirming that the ArPase is formed, we conducted directed evolution to enhance the catalytic activity to achieve the aforementioned goals 2 and 3. To facilitate this process, we established a screening assay on cell lysates.

Additionally, with the aim of enhancing the catalytic activity of the ArPase, we investigated a rationally designed library of chimeric hCAII proteins. The computationally designed protein scaffolds were investigated whether they could be the alternative to the conventional hCAII mutants reported for ArMs.

3 Directed Evolution of an Artificial Peroxidase Based on a Human Carbonic Anhydrase Protein²

Iori Morita, Adriana Faraone, Kailin Zhang, Roman P. Jakob, Timm Maier, and Thomas R. Ward

3.1 Abstract

The assembly of artificial metalloenzymes provides a second coordination sphere around a metal catalyst. Such a well-defined microenvironment can lead to enhancing the activities and selectivity of the catalyst. Herein, we present the development of artificial peroxidase (ArPase) by embedding a Fe-TAML (TAML = Tetra Amide Macrocyclic ligand) catalyst into a human carbonic anhydrase II (hCAII). Incorporation of the Fe-TAML catalyst ([BS-Fe-bTAML]⁻) within hCAII enhanced the total turnover number (TTON) for the hydroxylation of a benzylic C–H bond. After selecting a thermostable variant of hCAII (hCAII^{TS}), the resulting ArPase, [BS-Fe-bTAML]⁻ · hCAII^{TS}, was subjected to directed evolution using cell lysates in a 384-well format. After three rounds of laboratory evolution, the best-performing variants exhibited 36-fold enhancement in the initial rate and 2.8-fold enhancement in the TTON for the hydroxylation of a benzylic C–H bond. We surmise that an arginine residue introduced in the course of the directed evolution engage in hydrogen bonding with [BS-Fe-bTAML]⁻. This study highlights the potential of relying on a thermostable host protein, to expand the catalytic repertoire of the hCAII-based ArMs.

3.2 Introduction

Direct hydroxylation of C–H bonds presents a promising strategy for the late stage diversification of diverse compounds in a step-economical fashion. Despite the challenges

² This chapter was written as a manuscript for publication in a scientific journal.

associated with the ubiquity and inertness of C–H bonds, numerous transition metal catalysts have been reported to catalyze this transformation^[2, 7, 43]. A significant proportion of them are biomimetic, relying on iron- or manganese-based catalysts to generate a high-valent metal-oxo species which is capable of abstracting a hydrogen atom^[44]. The resulting carbon-centered radical can engage in the formation of C–O bonds, either through a rebound mechanism or by utilizing an alternative source of oxygen atoms^[45]. Within the green chemistry framework, regio- and stereoselective hydrocarbon oxidation should proceed under mild conditions and produce minimal waste. In this context, H₂O₂ stands out as an ideal oxidant in light of its atom-economy and practicality. A versatile complex that catalyzes such reactions is derived from the so-called Fe-TAML family (TAML = Tetra Amide Macrocyclic Ligand). These catalysts were pioneered by Collins and coworkers^[8]. In the last 20 years, multiple variations of the TAML ligand have been explored in the pursuit of oxidation catalysts with significantly improved performance.

An alternative approach to enhance the catalytic activity of metal catalysts involves the development of artificial metalloenzymes (ArMs)^[12]. Combining the attractive features of both homogeneous catalysis and biocatalysis, ArMs leverage a unique and well-defined environment provided by the protein scaffold upon incorporating a non-native cofactor. Our previous work has revealed that a biotinylated-[Fe-TAML] catalyst embedded within streptavidin leads to remarkable enantioselectivity (>98% e.e.) for benzylic hydroxylation via kinetic resolution^[32]. While this study highlighted the positive effect of streptavidin on enantioselectivity, the turnover number (TON) was only moderately affected upon incorporation of the biotinylated-[Fe-TAML] cofactor within streptavidin.

To overcome this limitation, ArMs can be subjected to directed evolution, involving iterative rounds of mutation and screening for the desired activity^[12c, 14]. Multiple enzymes have been repurposed and improved through directed evolution^[42, 46]. Given the typically large number of clones screened during this process, there is a need for high-throughput methods to accelerate such laborious procedures. Although whole-cell or cell lysate-based screening methods are commonly employed, challenges such as interference from cellular components or limited cellular uptake of the cofactor or substrates often impede their application to ArMs^[12d, 15].

Surface-display of proteins on cell membranes is a promising approach to circumvent such challenges. In this context, the development of ArMs based on monomeric proteins is desirable to minimize uncertainties regarding the oligomeric state of the surface-displayed protein^[47].

In this study, we aimed to improve the catalytic performance of an Fe-TAML complex through the assembly of an artificial peroxidase (ArPase) relying on a monomeric protein, human carbonic anhydrase II (hCAII, Figure 1). The activity of the ArPase was optimized through directed evolution facilitated by an assay performed on *E. coli* cell lysates.

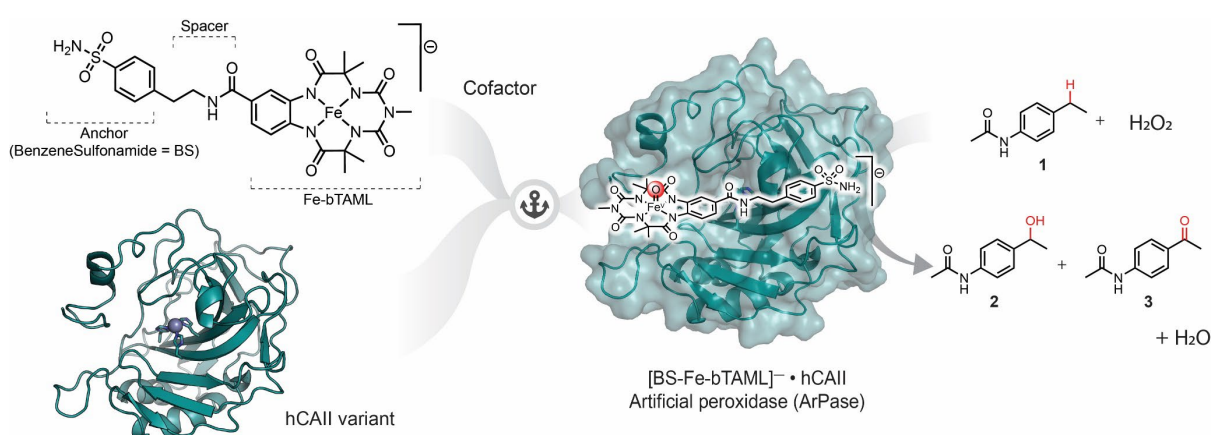


Figure 1. Assembling an artificial peroxidase (ArPase). A $[\text{Fe-bTAML}]^-$ catalyst is linked to a sulfonamide anchor moiety with an ethylene spacer, favoring the tight anchoring of the cofactor and human carbonic anhydrase II (hCAII).

3.3 Results and Discussion

3.3.1 Cofactor synthesis

Based on its stability and activity towards challenging C–H bonds, we selected the 5th generation Fe-TAML catalyst (biuret TAML, bTAML hereafter) as a starting point for the assembly of the ArPase^[48]. As a high-affinity moiety for hCAII binding, we selected an arylsulfonamide anchor^[49]. Based on preliminary docking studies, we included an ethylene spacer between the TAML moiety and the sulfonamide anchor, Figure 1.

The cofactor [BS-Fe-bTAML]⁻ (BS = benzenesulfonamide) was synthesized following a previously reported procedure (See Supporting Information). Following the characterization by high-resolution mass spectrometry, the cofactor was tested towards the benzylic C–H hydroxylation of the ethylbenzene derivative **1** in the presence of H₂O₂ as oxidant (Table 1). Overoxidation of ethylbenzene **1** to the corresponding ketone **3** was observed ([alcohol]/[ketone] = 2.2, Table 1, entry 1). Upon incorporation into a wild-type hCAII (hCAII^{WT}), the activity of [BS-Fe-bTAML]⁻ increased by 12% compared to the free cofactor (Total TurnOver Number (TTON) = 1044 ± 32 vs. 931 ± 59, Table 1, entries 1 and 2). This enhancement may stem from interactions between the catalyst and close-lying amino acids. In addition, we hypothesize that the protein scaffold suppresses the deleterious dimerization of the catalyst via a μ-oxo-bridge. Indeed, dimerization of Fe-TAML catalysts, facilitated through comproportionation between a Fe(III) species and a Fe(V)-oxo species leading to catalytically inactive (Fe(IV))₂-μ-oxo species, has been established^[10, 48b, 50].

Next, we determined the affinity of the cofactor to hCAII^{WT} (Table S4) using both (1) an inhibitory assay of the esterase activity of hCAII^[51], and (2) a DNSA displacement assay^[51-52]. Method (1) and (2) afforded a $K_d = 48 \pm 20$ nM (Figure S8) and a $K_d = 35 \pm 10$ nM (Figure S9), respectively. We thus conclude that the affinity of [BS-Fe-bTAML]⁻ for hCAII^{WT} is sufficiently high for directed evolution purposes, working with μM concentrations of protein and [BS-Fe-bTAML]⁻.

Entry	hCAII variant	TTON	[alcohol]/[ketone]
1	-	931 ± 59	2.2
2	WT	1044 ± 32	2.2
3	TS	1850 ± 50	1.6

Table 1. Scaffold optimization of the ArPase. Conditions: ethylbenzene derivative **1** (10 mM), cofactor (4 μM), hCAII variant (10 μM), ZnSO₄ (10 μM), H₂O₂ (20 mM), potassium phosphate buffer (pH = 7.2, 10 mM), MeCN (5%v/v), 25 °C, 18 hours. All reactions were performed in triplicate.

3.3.2 Establishing a cell lysate assay

To facilitate the directed evolution campaign, we set out to establish a screening protocol that does not require any protein purification step. Initial attempts to screen using *E. coli* whole cells failed, possibly due to the degradation of H₂O₂ by naturally-occurring catalases or an alkyl hydroperoxide reductase present in *E. coli*, combined with the limited diffusion of the negatively-charged cofactor [BS-Fe-bTAML]⁻ through both *E. coli* membranes. To circumvent these challenges, we hypothesized that the use of a thermostable hCAII might enable screening peroxidase activity on heat-treated *E. coli* cell lysates.

Therefore, we sought a thermostable alternative to hCAII^{WT}. Thermostability of the protein leads to increased stability during the purification procedure and under the reaction conditions. Although there are multiple carbonic anhydrases from thermophilic bacteria, we did not pursue these as they exist as oligomers^[53]. Instead, we selected a monomeric thermostable sextuple mutant of hCAII, initially patented (the 2006 Canadian Patent 2,541,986, and the United States Patent 7,521,217) and later studied by Mehl and Karplus^[54]. The reported crystal structure closely resembles that of hCAII^{WT} (backbone root-mean-square-deviations (RMSDs) of 0.36 Å to hCAII^{WT}, PDB: 3KS3), suggesting its potential as a substitute for hCAII^{WT}.

The gene of the thermostable variant of hCAII (hCAII-A65T-L100H-K153L-L223S-L239P-A247T, hCAII^{TS} hereafter) was cloned via sequential site-directed mutagenesis of hCAII^{WT} using polymerase chain reaction (PCR). The protein was then expressed in an *E. coli* BL21(DE3) strain and purified by affinity chromatography relying on sulfonamide-decorated agarose beads (See Supporting Information).

The activity of [BS-Fe-bTAML]⁻ · hCAII^{TS} was evaluated in the hydroxylation of the ethylbenzene derivative **1**. Gratifyingly, the activity of [BS-Fe-bTAML]⁻ · hCAII^{TS} was significantly improved compared to both the free cofactor and the wild-type ArPase (2 fold for the free cofactor and 1.8 fold for [BS-Fe-bTAML]⁻ · hCAII^{WT}, Table 1). This extent of activity enhancement was not observed with the previously-reported biotinylated-[Fe-TAML] · streptavidin system. The affinity of [BS-Fe-bTAML]⁻ for hCAII^{TS} was comparable to hCAII^{WT}

(52 ± 8 nM, Table S4, Figure S10). Differential scanning fluorimetry confirmed that the melting temperature of hCAII^{TS} increased by 14 °C compared to hCAII^{WT} (i.e. 71 °C and 57 °C, Figure S11). Accordingly, we selected hCAII^{TS} for the rest of this study.

Next, we investigated whether the hydroxylation proceeds in a cell lysate using *E. coli* BL21(DE3) cells harboring a plasmid encoding hCAII^{TS}. Gratifyingly, upon cell lysis and thermal treatment at 55 °C for 30 minutes, followed by addition of [BS-Fe-bTAML]⁻, the ethylbenzene derivative **1** (2500 equiv. vs. cofactor) and H₂O₂ (5000 equiv. vs. cofactor), we were delighted to detect the alcohol **2** and the ketone **3**, albeit in significantly lower yield compared to the purified ArPase (17 vs. 2140 TTON, Figure S12). Upon increasing the number of equivalents of H₂O₂ (50,000 equiv. vs. cofactor), yields (1929 TTON) comparable to those of the purified ArPase were obtained. We speculated that the background activity resulting from catalysis with the empty plasmid cell-lysates and the free cofactor (1020 TTON, corresponding to ca. 53% of the TTON for ethylbenzene **1**) would remain constant during the directed evolution campaign, thus enabling to detect improved ArPase activity resulting from [BS-Fe-bTAML]⁻ · hCAII^{TS} mutants. Accordingly, we set out to evolve the ArPase performance using heat-treated *E. coli* cell lysates.

With the objective of increasing the throughput of the ArPase screening assays, we selected a substrate that affords a product that can be quantified by optical spectroscopy (Figure 2a). Inspired by the literature reports by Schwaneberg^[55] and Arnold,^[56] we selected an ether-bearing substrate which, upon hydroxylation affords a hemiacetal. This intermediate spontaneously hydrolyzes, releasing an alcohol which may be detected by absorbance or fluorescence. Accordingly, we designed the substrate **4**, endowed with a *p*-nitrophenol, as the caged chromophore. Following its synthesis and characterization, we subjected ether **4** to the ArPase using a borate buffer at pH = 8.5. Monitoring the reaction by UV/vis spectroscopy, the growth of an absorbance at $\lambda = 410$ nm was observed for [BS-Fe-bTAML]⁻, [BS-Fe-bTAML]⁻ · hCAII^{WT}, and [BS-Fe-bTAML]⁻ · hCAII^{TS} in this order (Figure S13a). We confirmed that the TONs based on the *in situ* monitoring and high-performance liquid chromatography (HPLC) analysis exhibited good agreement (Figure S13b). Additionally, the TON determined by HPLC

for both *p*-nitrophenol and *p*-formylbenzoic acid correlated well, indicating negligible extent of oxidation of *p*-formylbenzoic acid to terephthalic acid. For high-throughput screening purposes, we tested whether the hydroxylation of caged chromophoric substrate **4** proceeds in cell lysate in 384-well plate format. In the presence of excess H₂O₂ (8000 equiv. vs. [BS-Fe-bTAML]⁻), comparable initial rates to those obtained in the presence of purified ArPase were obtained (Figure S14).

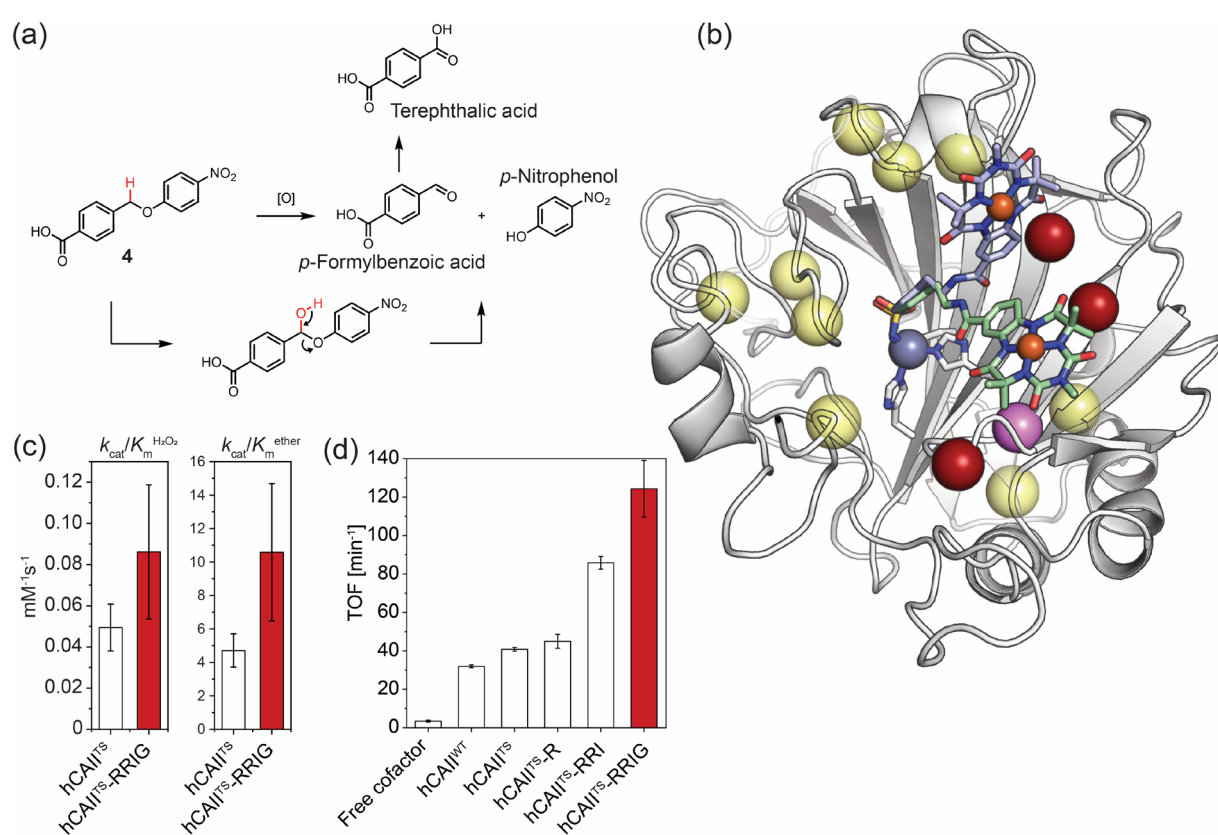


Figure 2. Directed evolution of an ArPase relying on a colorimetric assay. (a) The benzylic hydroxylation of the pro-chromogenic substrate **4** leads to the release of *p*-nitrophenol and *p*-formylbenzoic acid. *p*-Formylbenzoic acid can be further oxidized to afford terephthalic acid. (b) Superposition of the two crystal structures of [BS-Fe-bTAML]⁻ · hCAII^{TS} (PDB: 9F2E, 9F2F). Only one of the protein scaffolds is displayed as a cartoon model. The [BS-Fe-bTAML]⁻ from each datasets are represented as sticks (carbon chains in pale blue or pale green with color-coded heteroatoms; N = blue, O = red, S = yellow) with the Fe as orange sphere. The zinc atom is displayed as a violet sphere and the three coordinating histidine residues as sticks (C = white, N = blue). The Ca atoms of the mutated

positions are displayed as spheres (red: targeted and beneficial, yellow: targeted but not beneficial, purple: accidental F66I mutation). (c) The comparison of the catalytic efficiency for H_2O_2 ($k_{\text{cat}}/K_{\text{m}}^{\text{H}_2\text{O}_2}$) and substrate **4** ($k_{\text{cat}}/K_{\text{m}}^{\text{ether}}$). Conditions: substrate **4** (31.25 ~ 500 μM), hCAII variant (4 μM), [BS-Fe-bTAML]⁻ (0.25 μM), H_2O_2 (4 ~ 128 mM), ZnSO_4 (40 μM), borate buffer (pH = 8.5, 10 mM), DMSO (0.5%v/v), 25 °C. (d) Improvement of the initial rate of the reaction. Conditions: substrate **4** (0.5 mM), hCAII variant (4 μM), [BS-Fe-bTAML]⁻ (0.25 μM), H_2O_2 (128 mM), ZnSO_4 (40 μM), borate buffer (pH = 8.5, 10 mM), DMSO (0.5%v/v), 25 °C. The slope of the initial linear segment of the reaction profile (typically around 2 minutes) was used for the determination of the TOF.

3.3.3 Genetic engineering of ArPase performance

Starting from [BS-Fe-bTAML]⁻ · hCAII^{TS} as parent, we set out to optimize ArPase activity through directed evolution. For this purpose, we solved the X-ray crystal structure of the [BS-Fe-bTAML]⁻ · hCAII^{TS}. Two datasets were collected from distinct crystals resulting from different crystallization conditions. The overall structure of the protein scaffold are very similar (backbone RMSD = 0.27). The benzenesulfonamide moiety was bound to the zinc atom of hCAII^{TS} at nearly the same location. However, the conformation of the ethylene spacer differed, resulting in two distinct poses for the Fe-bTAML moiety (Figure 2b).

To include the observed flexibility of the cofactor in the directed evolution campaign, we selected twelve amino acids surrounding the active site cavity of the protein for site-directed mutagenesis (Figure 2b). Each amino acid was randomized individually with NDT degenerate codons using site-directed mutagenesis with PCR. The resulting *E. coli* clones from the single mutant library were oversampled (≥ 44 clones per targeted position) to ensure a library coverage of > 97%. The expression and preparation of the cell lysate were conducted in a 96-well format, and the hydroxylation assays were carried out in a 384-well format.

Following the initial screening, L60R, G63R, E69G, E69R, I91R, and G131R single variants were identified as most promising. Subsequently, their activity was validated using purified proteins, confirming the activity trend observed in the 384-well screening based on cell lysates (Figure S15). This correlation reveals that the screening assay, conducted on a 50 μL scale,

is sufficiently reliable to identify positive hits. The best-performing variant, hCAII^{TS}-E69R (hCAII^{TS}-R hereafter), was used as the parent for the second round of directed evolution. The remaining eleven positions were individually randomized, as described above. This led to the identification of a G63R mutation in all the top five variants. Interestingly, this position did not afford significant improvement in the 1st round (Figure S16), highlighting the epistatic effect of the mutations. Furthermore, the best hit among the top five variants contained an additional accidental mutation, F66I. Although the contribution of this mutation to the catalytic activity is modest (Figure S17), the accidental mutation was retained. We randomized the ten remaining positions for the third round of directed evolution, using hCAII^{TS}-E69R-G63R-F66I (hCAII^{TS}-RRI hereafter) as parent. This led to the identification of hCAII^{TS}-E69R-G63R-F66I-**I91G** (hCAII^{TS}-RRIG hereafter) as the best-performing variant.

3.3.4 Characterization of the evolved ArPase

Next, we compared the Michaelis-Menten saturation kinetic profiles of [BS-Fe-bTAML]⁻ · hCAII^{TS} with that of the evolved variant [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRIG. We observed an enhanced catalytic efficiency ($k_{\text{cat}}/K_{\text{m}}$) for both H₂O₂ ([BS-Fe-bTAML]⁻ · hCAII^{TS}: $k_{\text{cat}}/K_{\text{m}}^{\text{H}_2\text{O}_2} = 0.05 \pm 0.01 \text{ mM}^{-1}\cdot\text{s}^{-1}$, [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRIG: $k_{\text{cat}}/K_{\text{m}}^{\text{H}_2\text{O}_2} = 0.09 \pm 0.03 \text{ mM}^{-1}\cdot\text{s}^{-1}$) and the substrate **4** ([BS-Fe-bTAML]⁻ · hCAII^{TS}: $k_{\text{cat}}/K_{\text{m}}^{\text{ether}} = 4.7 \pm 1.0 \text{ mM}^{-1}\cdot\text{s}^{-1}$, [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRIG: $k_{\text{cat}}/K_{\text{m}}^{\text{ether}} = 10.6 \pm 4.1 \text{ mM}^{-1}\cdot\text{s}^{-1}$) (Figure 2c, Figure S18). Although the K_{m} values for both H₂O₂ and substrate **4** increased in the evolved variant, the k_{cat} value increased by more than threefold ([BS-Fe-bTAML]⁻ · hCAII^{TS}: $k_{\text{cat}} = 1.8 \pm 0.1 \text{ s}^{-1}$, [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRIG: $k_{\text{cat}} = 5.7 \pm 0.6 \text{ s}^{-1}$). This improvement is reflected in the improvement in turnover frequency (TOF) for the uncaging reaction (Figure 2d, [BS-Fe-bTAML]⁻ · hCAII^{TS}: 41 min⁻¹, [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRIG: 124 min⁻¹). For the evolved variant, the overall enhancement of the initial rate is 36 fold compared to [BS-Fe-bTAML]⁻.

Next, we evaluated the catalytic performance of the evolved ArPase with several benzylic substrates: **1**, **5**, **8**, and **11** (Table 2, see Table S5 for a comprehensive list). We initially validated both [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRI and [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRIG under the

standard validation conditions (condition A: 10 mM phosphate buffer, pH 7.2, 5~20%v/v MeCN). However, [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRIG precipitated. Although it did not precipitate when different conditions similar to the screening were employed (condition B: 10 mM borate buffer, pH 8.5, 0.5% DMSO), these conditions led to lower TTON than [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRI under condition A (Table 2 entries 4 vs. 2, Table S5 entries 12 vs. 11 and 14 vs. 13). Therefore, the following activity assays concern [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRI under condition A for the most part.

Hydroxylation of the ethylbenzene derivative **1** in the presence of [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRI resulted in a TTON of 2629 ± 412, albeit with a low e.e. (Table 2, entry 2). When [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRIG was tested under condition B, less than half of the TTON compared to [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRI and slightly improved e.e. (12%) were observed (Table 2, entry 4). Reducing the reaction time to 1 hour led to a 20% e.e. (Table 2, entry 3). The slightly higher e.e. observed after short reaction times suggests that the oxidation of the alcohol **2** to the corresponding ketone **3** proceeds with a slight preference for the major enantiomer of the alcohol **2**, thus leading to an erosion of the e.e. over time.

Inspired by a publication by Bach and coworkers^[57], we tested lactam **5** with the evolved ArPases. We hypothesized that the rigidity of the bicyclic lactam, combined a polar handle enabling propitious placement of the prochiral methylene moiety within the active site of the ArPase, might lead to higher enantioselectivity. Gratifyingly, [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRI afforded the corresponding alcohol **6** in 53% e.e. (Table 2, entry 5). Extending the reaction time to 18 hours led to an erosion of e.e. and of [alcohol]/[ketone] ratio (Table 2, entry 6). This supports the hypothesis that the overoxidation step consumes preferentially the major enantiomer of the alcohol **6**. Additionally, we tested ethylbenzene derivatives with different substituents on the aromatic ring, **8** and **11** (Table 2, entries 7-8). Benzoic acid **8** exhibited a lower TTON after the overnight reaction than substrates **1** and **5**, possibly due to the charge repulsion between the carboxyl group and the negatively-charged Fe(V)=O intermediate. The very low conversion observed for ethylbenzene **11** is attributed to its very limited solubility in

aqueous media. Hydroxylation of cyclohexane with $[\text{BS-Fe-bTAML}]^- \cdot \text{hCAII}^{\text{TS-RRI}}$ yielded a TTON of < 1 (Table S5, entry 15).

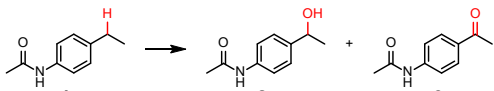
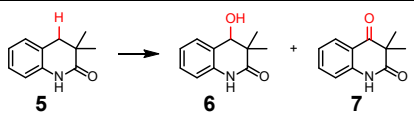
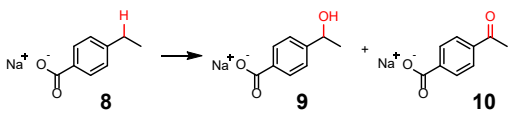
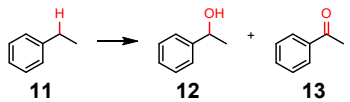
Entry	Reaction	hCAII variant	TTON ^a	e.e. %	[alcohol] / [ketone]
1 ^b		TS	1850 ± 50	-3	1.6
2 ^b		TS-RRI	2629 ± 412	5	0.9
3 ^{c,d}		TS-RRIG	643 ± 60	20 ± 0	1.4
4 ^c		TS-RRIG	1110 ± 33	12 ± 1	0.7
5 ^{d,e}		TS-RRI	486 ± 7	53 ± 1	6.4
6 ^e		TS-RRI	1041 ± 23	29 ± 3	1.4
7 ^{b,f}		TS-RRI	265 ± 3	10 ± 1	7.1
8 ^g		TS-RRI	44 ± 3	-8	2.8

Table 2. Activity of the evolved ArPases. The products were analyzed using either supercritical fluid chromatography (SFC, substrates **1,5,8**) or GC-FID (ethylbenzene **11**). The reactions (200 μL) were quenched by addition of HCl aq (1 M, 15 μL) and a suitable internal standard was added. For the HPLC analysis, the protein was removed by the addition of MeOH (200 μL), followed by centrifugation. For the SFC and GC-FID analysis, the organic compounds were extracted with EtOAc (400 $\mu\text{L} \times 2$). The combined organic fractions were dried over magnesium sulfate.

^aTTON corresponds to the sum of the TONs determined for both the alcohol and the ketone.

^bSubstrate (10 mM), hCAII variant (10 μM), $[\text{BS-Fe-bTAML}]^-$ (4 μM), H_2O_2 (20 mM), ZnSO_4 (10 μM), phosphate buffer (pH = 7.2, 10 mM), MeCN (5%v/v), 25 $^\circ\text{C}$, 18 hours.

^cSubstrate (5 mM), hCAII variant (10 μM), $[\text{BS-Fe-bTAML}]^-$ (4 μM), H_2O_2 (10 mM), ZnSO_4 (10 μM), borate buffer (pH = 8.5, 10 mM), DMSO (0.5%v/v), 25 $^\circ\text{C}$, 18 hours.

^d1-hour reaction.

^eSubstrate (5 mM), hCAII variant (10 μM), $[\text{BS-Fe-bTAML}]^-$ (4 μM), H_2O_2 (10 mM), ZnSO_4 (10 μM), phosphate buffer (pH = 7.2, 10 mM), MeCN (20%v/v), 25 $^\circ\text{C}$, 18 hours.

^fMeCN was not used.

^gSubstrate (10 mM), hCAII variant (10 μ M), [BS-Fe-bTAML]⁻ (4 μ M), H₂O₂ (20 mM), ZnSO₄ (10 μ M), phosphate buffer (pH = 7.2, 10 mM), MeCN (20%v/v), 25 °C, 18 hours.

3.3.5 Structural analysis

To elucidate the structural basis contributing to the enhanced catalytic activity of the ArPase, we solved the X-ray crystal structure of the evolved variants. Unfortunately, we were unable to obtain the crystal structure of [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRIG. We were however able to solve the X-ray crystal structure of [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRI at 1.93 Å resolution (Figure 3). The general structure of [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRI agrees well with [BS-Fe-bTAML]⁻ · hCAII^{TS} (backbone RMSD < 0.29 Å). The [BS-Fe-bTAML]⁻ cofactor is located at one of the positions observed in the structure of [BS-Fe-bTAML]⁻ · hCAII^{TS} (Figure 3a, Figure 2b). The aromatic ring of Fe-bTAML moiety has a π - π interaction with the F130 residue (Figure S19). Notably, the amide carbonyl group in the spacer moiety of the cofactor forms a hydrogen bond with the amide of Q92 residue (Figure 2b). Furthermore, one of the carbonyl groups within the TAML ring is engaged in two hydrogen bonds: one with the guanidine group of R69, and the other with a water molecule. The water molecule further interacts with side chain of D72 and the backbone nitrogen of the F130 residue. This structure suggests that the E69R mutation may be involved in activating the cofactor through hydrogen bonding, together with the neighboring Q92, D72, and F130 residues. Both Brønsted and Lewis acids have been proposed to interact with the carbonyl groups of Mn- or Fe-TAML complexes, thereby enhancing the oxidizing power of the corresponding metal(V)-oxo complexes^[58]. Given that the guanidine group is protonated, we hypothesize that the arginine side chain serves as a Brønsted acid, enhancing the oxidation activity of the Fe(V)-oxo species.

The influence of G63R mutation is not obvious from this crystal structure due to the large distance between the mutation and [BS-Fe-bTAML]⁻. Additionally, due to the lack of a crystal structure of [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRIG, the role of the I91G mutation also remains elusive. Given the spatial proximity of E69R and I91 and the size difference between the side chains of glycine and isoleucine, it may affect the orientation of the amino acid side chains that form hydrogen bonds with the cofactor.

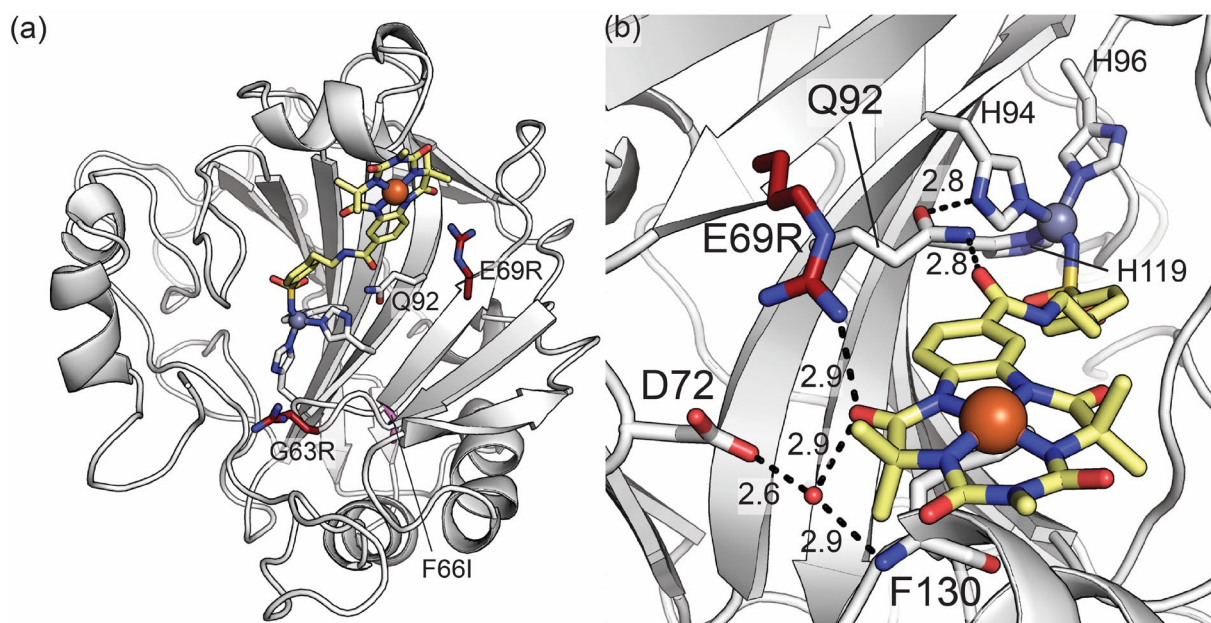


Figure 3. Crystal structure of $[\text{BS-Fe-bTAML}]^- \cdot \text{hCAII}^{\text{TS}}\text{-RR1}$ (PDB: 9F15). The protein is represented as a cartoon model in grey. Selected amino acids and the cofactor are represented as sticks (with color-coded heteroatoms; N = blue, O = red, S = yellow) and Fe as an orange sphere. The E69R and G63R mutations, the F66I mutation, and the cofactor are highlighted in red, purple, and yellow, respectively. The other amino acids including the 3 histidines coordinating to the Zn are colored in white. The Zn atom is represented as violet sphere. (a) Overview of the structure of $[\text{BS-Fe-bTAML}]^- \cdot \text{hCAII}^{\text{TS}}\text{-RR1}$. The structure is viewed from the same angle as in Figure 2b. (b) Close-up view of the structure of $[\text{BS-Fe-bTAML}]^- \cdot \text{hCAII}^{\text{TS}}\text{-RR1}$. A water molecule is displayed as a red sphere. The hydrogen bonds among the selected residues, the cofactor, and the water molecule are displayed as black dashed lines with close contact distances displayed in Å.

3.4 Conclusion

In summary, we have engineered an ArPase by anchoring an Fe-bTAML catalyst with a thermostable variant of hCAII. hCAII^{TS} turned out to be a beneficial scaffold for the hydroxylation on benzylic C–H bonds, partly due to the suppression of the dimerization of the cofactor. Subsequent engineering of the protein scaffold through directed evolution, relying on the caged substrate, yielded the enhanced variants, hCAII^{TS}-RR1 and hCAII^{TS}-RRIG. The enhancement of the initial rate was 36 fold, and that of the TTON for the ethylbenzene derivative **1** was 2.8 fold compared to $[\text{BS-Fe-bTAML}]^-$, which underscores the advantage of

the development and genetic optimization of ArPases. The evolved variants exhibited improved enantioselectivity, albeit modest, for hydroxylation of lactam **5**. We attribute the enhanced hydroxylation activity to hydrogen bonds with the arginine side chain introduced through directed evolution. It may act as a key Brønsted acid to activate the catalyst in combination with other neighboring amino acids.

Overall, this study underscores the potential to enhance the catalytic activity of TAML catalysts through genetic optimization, leading to the development of a highly active thermostable artificial peroxidase capable of operating under mild conditions.

3.5 Supporting Information

3.5.1 General aspects

3.5.1.1 Materials and methods

All commercially-available chemicals were purchased from Thermo Scientific, Sigma-Aldrich, Acros Organics, TCI Europe, Fluka, Fluorochem, Lancaster Synthesis and used without further purification. The most important chemicals are listed below:

4-Formylbenzoic acid (Sigma-Aldrich, 124915)

Terephthalic acid (Sigma-Aldrich, 185361)

4-Nitrophenol (Fluka, 73562)

Benzyltriethylammonium bromide (Sigma-Aldrich, 147125)

Indole (Sigma-Aldrich, I3408)

4-Ethylbenzoic acid (Alfa Aesar, A10235)

4-Acetylbenzoic acid (Alfa Aesar, A15239)

Bibenzyl (Sigma-Aldrich, B33706)

Ethylbenzene (Fluka, 03080)

Acetophenone (Fluke, 00790)

DL-sec-Phenethyl alcohol (Acros Organics, 304280250)

(*R*)-(+)-1-Phenylethanol (Lancaster Synthesis, 19296)

Biphenyl (Sigma-Aldrich, B34656)

Cyclohexane (Sigma-Aldrich, 34855)

Cyclohexanol (Sigma-Aldrich, 105899)

Cyclohexanone (Sigma-Aldrich, 29140)

Dry solvents were purchased from Acros Organics and used directly without further purification. The water used for all biological and catalytic experiments was purified with a Milli-Q Advantage system. The cofactors were purified by reversed-phase chromatography on a preparative HPLC. All enzymes used for PCRs and cloning were purchased from NEB, New England Biolabs.

GenElute™ Plasmid Miniprep Kit (Merck) was used for isolating the plasmids. Lysozyme was purchased from GoldBio (from egg white, L-040-100). DNaseI was purchased from PanReac AppliChem (A3778).

The 96-well transparent plates were purchased from Greiner (655101, Flat-bottom, polystyrene). The black 96-well plates were purchased from Thermo Scientific (Nunc™, Nunclon Delta-treated, flat-bottom, 137101). The 384-well transparent plates were purchased from Thermo Scientific (Nunc™, non-treated, flat-bottom)

3.5.1.2 Instrumentation

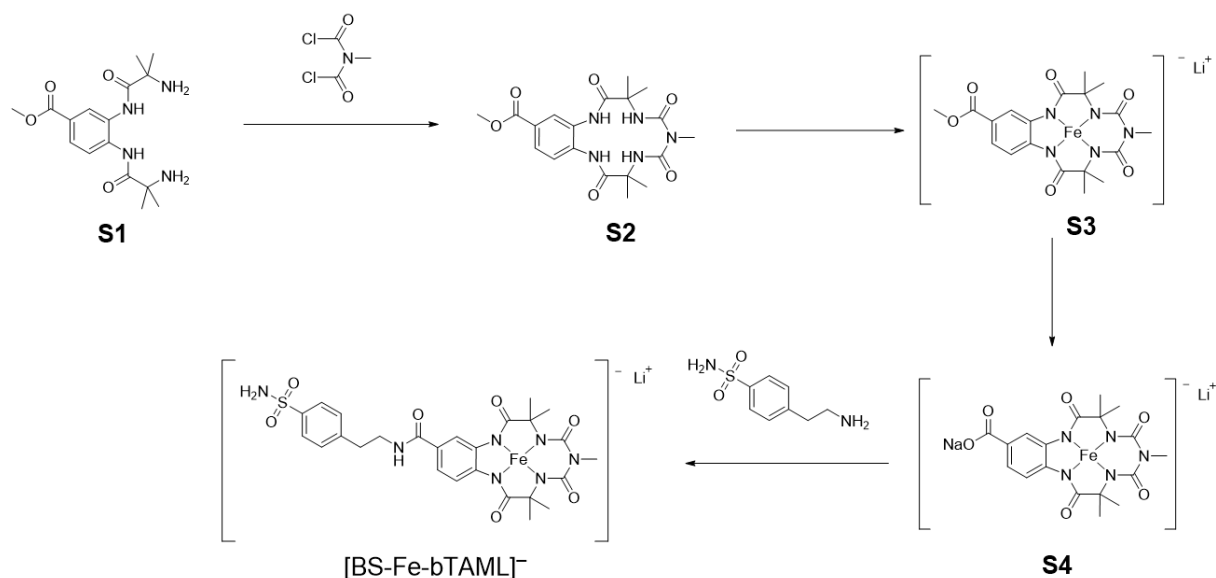
¹H and ¹³C spectra were recorded on a Bruker 500 MHz or 600 MHz at room temperature. Chemical shifts are reported in ppm (parts per million) relative to TMS ($\delta = 0.00$ ppm for ¹H and ¹³C). Signals are quoted as s (singlet), d (doublet), t (triplet), bs (broad singlet) and m (multiplet). Spectra were analyzed on MestReNova and calibrated relative to the residual solvent peak. 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 4G QTOF ESI mass spectrometer. Supercritical fluid chromatography (SFC) was carried out on an Acquity UPC² system from Waters using CO₂ as the mobile phase with addition of a cosolvent. Absorbance or fluorescence assays on microtiter plates were performed with The Infinite® M1000 (Tecan). HPLC analysis was performed with 1290 Infinity II LC System (Agilent). GC analysis of the catalysis experiments was performed on a GC-FID equipped with a CP-Chirasil-Dex CB column (25m 0.25mm 0.25 μ m) using He as carrier gas and biphenyl as internal standard. Flash chromatography was performed on a Biotage Isolera or a Büchi Pure chromatography system. No unexpected or unusually high safety hazards

were encountered. Nano-DSF measurements of thermal stability were performed using a Prometheus NT.48 (NanoTemper) instrument in nanopure water, with standard capillaries.

3.5.2 Synthesis of cofactors and substrates

3.5.2.1 Cofactors

The diamine **S1** was synthesized according to a reported procedure^[32].



Scheme S1: Synthetic pathways followed for the synthesis of the cofactor [BS-Fe-bTAML]⁻

Tetraamide ligand S2. An oven-dried two-necked flask was charged with dry CH₂Cl₂ (300 mL) under nitrogen. The diamine **S1** (723 mg, 2.15 mmol) was dissolved in dry CH₂Cl₂ (20 mL), of dry Et₃N (604 μL, 4.3 mmol, 2 equiv.) were added, and this mixture was loaded into a syringe. *N,N*-dichloroformylmethylamine (246 μL, 2.37 mmol, 1.1 equiv.) was dissolved in dry CH₂Cl₂ (20 mL) and loaded to a second syringe. Both syringes were placed on a syringe pump and added via a cannula to the flask containing the diamine **S1** under nitrogen simultaneously over 1.5 hours. The mixture was stirred for 36 hours at room temperature. The solvent was removed under reduced pressure to yield a pale yellow solid, which was purified by silica gel column chromatography using EtOAc and DCM as mobile phase (gradient method: EtOAc 12% for 1 column volume (CV); 12% to 100% over 15 CV; 100% over 3 CV). The tetraamide ligand **S2** was obtained as a white powder (280 mg, 0.67 mmol, 31%).

^1H NMR (500 MHz, DMSO) δ 9.26 (s, 1H), 9.12 (s, 1H), 8.06 (s, 1H), 7.98 (d, J = 2.1 Hz, 1H), 7.84 (d, J = 8.5 Hz, 1H), 7.81 (s, 1H), 7.76 (dd, J = 8.5, 2.1 Hz, 1H), 3.78 (s, 3H), 2.91 (s, 3H), 1.46 (s, 6H), 1.45 (s, 6H).

^{13}C NMR (126 MHz, DMSO) δ 174.13, 174.06, 166.00, 157.47, 156.28, 136.73, 129.82, 128.06, 127.43, 126.20, 124.09, 59.51, 59.19, 52.64, 32.11, 26.00, 25.55.

HRMS (ESI positive mode, m/z): calc. for $\text{C}_{19}\text{H}_{25}\text{N}_5\text{NaO}_6$ $[\text{M}+\text{Na}]^+$ 442.1697; found 442.1702.

Methyl ester S3. The tetraamide ligand **S2** (200 mg, 0.477 mmol) was dissolved in dry THF (40 mL) and an N_2 atmosphere was established. Lithium hexamethyldisilazide was added dropwise (2.86 mL of a 1 M solution in THF, 2.86 mmol, 6 equiv.) on ice, and the solution was stirred for 10 minutes. Then, FeCl_3 (118 mg, 0.716 mmol, 1.5 equiv.) dissolved in dry THF (2 mL) was added, and the mixture was stirred overnight under an N_2 atmosphere at room temperature to afford a precipitate. The mixture was transferred into a falcon tube (15 mL), and the solid was separated by centrifugation (4G, 3 min). The liquid was discarded and the solid was washed twice with dry THF (5 mL). Dry MeOH (5 mL) was added to dissolve the complex, and the remaining undissolved material was discarded. The solvent was removed under reduced pressure, and the product was further purified by alumina column chromatography using DCM and MeOH as mobile phase (gradient method: MeOH 10% for 3 CV; 10% to 70% over 10 CV; 70% over 5 CV). The methyl ester product **S3** was obtained as orange solid (188 mg, 0.399 mmol, 84%) after drying under vacuum.

HRMS (ESI negative mode, m/z): calc. for $\text{C}_{19}\text{H}_{21}\text{FeN}_5\text{O}_6$ $[\text{M}]^-$ 471.0847; found 471.0856.

Carboxylate S4. In a round-bottom flask, the methyl ester **S3** (60 mg, 0.127 mmol) was dissolved in $\text{H}_2\text{O}/\text{MeOH}$ (1:1, 10 mL) under an N_2 atmosphere. Then NaOH aq (305 μL of 5 M solution, 1.52 mmol, 12 equiv.) was added. The mixture was stirred (5 hours, 40 $^\circ\text{C}$). The reaction progress was monitored with thin-layer chromatography. After completion (typically ~5 hours), the solvent was removed under reduced pressure, and the residue was

concentrated to yield the crude orange carboxylate **S4**, which was used further purification for the next step.

HRMS (ESI negative mode, m/z): calc. for $C_{18}H_{19}FeN_5O_6 [M]^-$ 457.0690; found 457.0699.

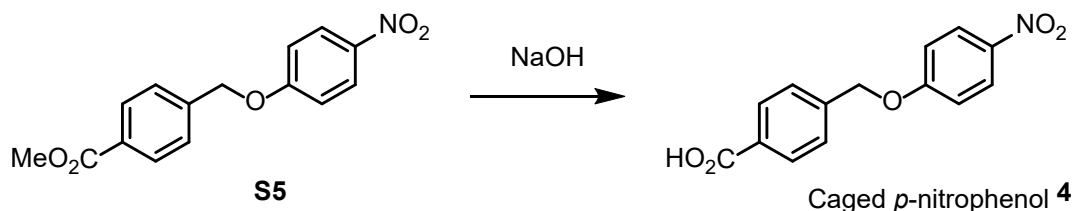
[BS-Fe-bTAML]⁻Li⁺. The carboxylate **S4** (170 mg, 0.355 mmol) and *p*-(2-aminoethyl)benzenesulfonamide (87.1 mg, 0.426 mmol, 1.2 equiv.) were loaded in a round-bottom flask (25 mL) under an N_2 atmosphere. Dry DMF (15 mL) was added, and the mixture was cooled to 0 °C. HCTU (297 mg, 0.717 mmol, 2.02 equiv.) was dissolved in a minimal amount of dry DMF and added to the reaction mixture dropwise. The reaction was initiated by adding *N,N*-diisopropylethylamine (176 μ L, 1.07 mmol, 3 equiv.), and the solution was stirred overnight at room temperature. The reaction progress was monitored with ESI-MS. After the consumption of the starting material was confirmed, the solvent was removed under reduced pressure. The crude was dissolved in MeOH (4 mL), and then Et_2O (35 mL) was added to get a pale yellow precipitate. The suspension was transferred to a falcon tube (15 mL), centrifuged (4000g, 5 minutes), and the supernatant was removed. The precipitate was washed with Et_2O (25 mL). The solid was dried under vacuum.

The product was purified by reverse phase chromatography using a Biotage SNAP Ultra C18 column. H_2O (no additive) and MeOH (no additive) were used as the mobile phase. The separation method starts from 5% MeOH for 3 CV followed by a gradient from 5 to 80% over 10 CV. This purification was performed twice. Only pure fractions containing **[BS-Fe-bTAML]⁻Li⁺** (confirmed by ESI-MS) were collected. Pure **[BS-Fe-bTAML]⁻Li⁺** (6.38 mg, 0.00987 mmol, 3%, orange powder) was obtained after removing the solvent under reduced pressure.

HRMS (ESI negative mode, m/z): calc. for $C_{26}H_{29}FeN_7O_7S [M]^-$ 639.1204; found 639.1213.

3.5.2.2 Substrates

Ethylbenzene derivative **1**, alcohol **2**, ketone **3**, methyl ester **S5**, lactam **5**, alcohol **6**, ketone **7**, alcohol **9** were synthesized according to a reported procedure^[59].



Caged *p*-nitrophenol 4. The methyl ester **S5** (240 mg, 0.84 mmol) was dissolved in MeOH (20 mL). NaOH aq (1M, 1.67 mL, 2 equiv.) was added, and the mixture was heated to reflux (1.5 hour). The solvent was removed under reduced pressure. The oily residue was acidified with HCl aq (1M) on ice. The white precipitate was then filtered and collected. The white solid was then purified with silica gel column chromatography (DCM and EtOAc containing 0.1% TFA as the mobile phase, gradient method: EtOAc 5% for 1 CV; 5% to 40% over 15 CV; 40% over 3 CV). The solvent was removed under reduced pressure to afford pure caged *p*-nitrophenol **4** as a white powder (146 mg, 0.53 mmol, 64%)

¹H NMR (500 MHz, DMSO) δ 8.27 – 8.19 (m, 2H), 8.01 – 7.95 (m, 2H), 7.62 – 7.55 (m, 2H), 7.28 – 7.21 (m, 2H), 5.37 (s, 2H).

¹³C NMR (126 MHz, DMSO) δ 167.01, 163.40, 141.10, 140.96, 130.47, 129.55, 127.57, 125.92, 115.42, 69.48.

HRMS (ESI positive mode, m/z): calc. for C₁₄H₁₁NNaO₅ [M+Na]⁺ 296.0529; found 296.0531.

3.5.3 Cloning, expression and purification of the proteins

3.5.3.1 Cloning and expression

Name	Sequence (5' to 3')
Fw hCAII A65T	TCA TAC CTT CAA CGT GGA GTT TG
Rv hCAII A65T	TGA AGG TAT GAC CAT TGT TGA GG
Fw hCAII L100H	TTC ACA TGA TGG ACA AGG TTC AG
Rv hCAII L100H	CAT CAT GTG AAC CCC AGT GAA AG
Fw hCAII K153L	CGC TCT GCC GGG CCT TCA GAA AG
Rv hCAII K153L	CCG GCA GAG CGC TGC CAA CCT TC
Fw hCAII L223S	GGT GAG CAA ATT CCG TAA ACT TA
Rv hCAII L223S	ATT TGC TCA CCT GCT CGC TGC TG
Fw hCA L239P	AGA ACC GAT GGT GGA CAA CTG G
Rv hCA L239P	CCA TCG GTT CTT CGG GTT CAC C
Fw hCA A247T	CCC AAC CCA GCC ACT GAA GAA CAG
Rv hCA A247T	CTG GGT TGG GCG CCA GTT GTC C

Table S1. Sequence of the primers used for mutagenesis. Fw = forward primer, Rv = reverse primer.

pACYCDuet-1 vector was double-digested at NcoI and XhoI sites, and the gene fragment of hCAII^{WT} was incorporated. The plasmid encoding hCAII^{TS} was generated by sequential PCR mutagenesis. A suspension of *E. coli* BL21 (DE3) chemically competent cells (50 µL) were transformed with the plasmid of interest. After heat-shock treatment (42 °C, 45 sec.), the cells were placed on ice (3 min.). Super Optimal Broth (SOC) medium (500 µL) was added to the tube, and the cells were incubated with a shaker (1h, 300 rpm at 37 °C). The suspension (50 µL) was plated out on Lysogeny Broth (LB) agar plates supplemented with chloramphenicol (30 µg/mL). The plates were incubated (37 °C, 16 hours). For protein expression (200 mL culture), a preculture was prepared by inoculating a fresh single colony in LB medium (3 mL) in a culture tube (13 mL, Sarstedt AG), and incubated (37 °C, 300 rpm, 16 hours). The preculture (2 mL) was inoculated in the main culture (200 mL, LB medium containing 30 µg/mL chloramphenicol and 1% w/v glucose) in a baffled flask (1 L) and further incubated (37 °C, 160 rpm). The rest of the overnight preculture was used to prepare a glycerol stock (25%v/v final

glycerol content). The OD_{600} of the culture was monitored with a UV/vis spectrometer. When the OD_{600} is optimal (between 0.6 and 0.8, typically after 2-3 hours after the start of the main culture), isopropyl- β -D-thiogalactopyranoside (IPTG, 0.5 mM) was added. The culture was shaken (25 °C, 20 hours). After expression, the cell culture was transferred to a centrifuge bottle and the cells were harvested by centrifugation (4000 rpm for 15 minutes). The cell pellet was stored at -20 °C for 16 hours or more before the following application.

3.5.3.2 Gene sequence of the pACYC plasmid harboring the hCAII variants

>pACYC-hCAII^{WT} (4508 bp, hCAII^{WT} is highlighted in bold)

```
GGGGAATTGTGAGCGGATAACAATTCCCCTGTAGAAATAATTTTGTTTAACTTTAATAAG
GAGATATACCATGGCCCATCACTGGGGGTACGGCAAACACAACGGACCTGAGCACTG
GCATAAGGACTTCCCATTGCCAAGGGAGAGCGCCAGTCCCCTGTTGACATCGACAC
TCATACAGCCAAGTATGACCCTTCCTGAAGCCCCTGTCTGTTTCCTATGATCAAGCA
ACTTCCCTGAGAATCCTCAACAATGGTCATGCTTTCAACGTGGAGTTTGATGACTCTCA
GGACAAAGCAGTGCTCAAGGGAGGACCCCTGGATGGCACTTACAGATTGATTCAGTT
TCACTTTCCTGAGGTTCACTTGATGGACAAGGTTTCAGAGCATACTGTGGATAAAAAG
AAATATGCTGCAGAACTTCACTTGGTTCACTGGAACACCAAATATGGGGATTTTGGGA
AAGCTGTGCAGCAACCTGATGGACTGGCCGTTCTAGGTATTTTTTTGAAGGTTGGCAG
CGCTAAACCGGGCCTTCAGAAAGTTGTTGATGTGCTGGATTCCATTAACAAGGGC
AAGAGTGCTGACTTCACTAACTTCGATCCTCGTGGCCTCCTTCCTGAATCCCTGGATTA
CTGGACCTACCCAGGCTCACTGACCACCCCTCCTCTTCTGGAATCAGTGACCTGGATT
GTGCTCAAGGAACCCATCAGCGTCAGCAGCGAGCAGGTGTTGAAATTCCGTAACTT
AACTTCAATGGGGAGGGTGAACCCGAAGAACTGATGGTGGACAACCTGGCGCCCAGCT
CAGCCACTGAAGAACAGGCAAATCAAAGCTTCCTTCAAATAACTCGAGTCTGGTAAA
GAAACCGCTGCTGCGAAATTTGAACGCCAGCACATGGACTCGTCTACTAGCGCAGCT
TAATTAACCTAGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCT
CTAAACGGGTCTTGAGGGGTTTTTTGCTGAAACCTCAGGCATTTGAGAAGCACACGGT
```


**CACACTGCTTCCGGTAGTCAATAAACCGGTAAACCAGCAATAGACATAAGCGGCTATT
TAACGACCCTGCCCTGAACCGACGACCGGGTGAATTTGCTTTTGAATTTCTGCCATT
CATCCGCTTATTATCACTTATTCAGGGCTAGCACCAGGGCTTTAAGGGCACCAATAACT
GCCTTAAAAAATTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCATTAAG
CATTCTGCCGACATGGAAGCCATCACAGACGGCATGATGAACCTGAATCGCCAGCGGC
ATCAGCACCTTGTCGCCTTGCGTATAATATTTGCCCATAGTGAAAACGGGGGCGAAGAA
GTTGTCCATATTGGCCACGTTTAAATCAAACCTGGTGAAACTCACCC**

AGGGATTGGCTGAGACGAAAAACATATTCTCAATAAACCCTTTAGGGAAATAGGCCAGG
TTTTACCGTAACACGCCACATCTTGCGAATATATGTGTAGAACTGCCGGAAATCGTC
GTGGTATTCACTCCAGAGCGATGAAAACGTTTCAGTTTGCTCATGGAAAACGGTGTAAC
AAGGGTGAACACTATCCCATATCACCCAGCTCACCGTCTTTCATTGCCATACGGAACTCC
GGATGAGCATTATCAGGCGGGCAAGAATGTGAATAAAGGCCGGATAAACTTGTGCTT
ATTTTTCTTTACGGTCTTTAAAAAGGCCGTAATATCCAGCTGAACGGTCTGGTTATAGGT
ACATTGAGCAACTGACTGAAATGCCTCAAATGTTCTTTACGATGCCATTGGGATATATC
AACGGTGGTATATCCAGTGATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAATCTC
GATAACTCAAAAATACGCCCCGGTAGTGATCTTATTTTATTATGGTGAAAGTTGGAACCT
CTTACGTGCCGATCAACGTCTCATTTTCGCCAAAAGTTGGCCCAGGGCTTCCCGGTATC
AACAGGGACACCAGGATTTATTTATTCTGCGAAGTGATCTTCCGTACAGGTATTTATTC
GGCGCAAAGTGCGTCGGGTGATGCTGCCAACTTACTGATTTAGTGTATGATGGTGTTTT
TGAGGTGCTCCAGTGGCTTCTGTTTCTATCAGCTGTCCCTCCTGTTTACGCTACTGACGG
GGTGGTGCCTAACGGCAAAGCACCGCCGGACATCAGCGCTAGCGGAGTGTATACTG
GCTTACTATGTTGGCACTGATGAGGGTGTGAGTGAAGTGTTCATGTGGCAGGAGAAAA
AAGGCTGCACCGGTGCGTCAGCAGAATATGTGATACAGGATATATTCCGCTTCCTCGCT
CACTGACTCGCTACGCTCGGTGTTTCGACTGCGGCGAGCGGAAATGGCTTACGAACGG
GGCGGAGATTTCTGGAAGATGCCAGGAAGATACTTAACAGGGAAGTGAGAGGGCCGC
GGCAAAGCCGTTTTTCCATAGGCTCCGCCCCCTGACAAGCATCACGAAATCTGACGCT
CAAATCAGTGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGC
GGCTCCCTCGTGCGCTCTCCTGTTTCTGCTTTTCGGTTTACCGGTGTCATTCCGCTGTT

ATGGCCGCGTTTGTCTCATTCCACGCCTGACACTCAGTTCGGGTAGGCAGTTCGCTCC
AAGCTGGACTGTATGCACGAACCCCCGTTTCAGTCCGACCGCTGCGCCTTATCCGGTA
ACTATCGTCTTGAGTCCAACCCGGAAAGACATGCAAAGCACCACTGGCAGCAGCCACT
GGTAATTGATTTAGAGGAGTTAGTCTTGAAGTCATGCGCCGGTTAAGGCTAAACTGAAA
GGACAAGTTTTTGGTACTGCGCTCCTCCAAGCCAGTTACCTCGGTTCAAAGAGTTGGTA
GCTCAGAGAACCTTCGAAAAACCGCCCTGCAAGGCGGTTTTTTTCGTTTTTCAGAGCAAGA
GATTACGCGCAGACCAAACGATCTCAAGAAGATCATCTTATTAATCAGATAAAATATTT
CTAGATTTTCAGTGCAATTTATCTCTTCAAATGTAGCACCTGAAGTCAGCCCCATACGATA
TAAGTTGTAATTCTCATGTTAGTCATGCCCCGCGCCCACCGGAAGGAGCTGACTGGGTT
GAAGGCTCTCAAGGGCATCGGTGAGATCCCGGTGCCTAATGAGTGAGCTAACTTACA
TTAATTGCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTCGTGCCAGCTGCA
TTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGG
TTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGA
GAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGA
TGTTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACC
GAGATGTCCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGCCCAGC
GCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTACGATTTG
CATGGTTTGTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCT
GAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGAC
AGAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGATGC
TCCACGCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTG
GTCAGAGACATCAAGAAATAACGCCGGAACATTAGTGACAGGCAGCTTCCACAGCAATG
GCATCCTGGTCATCCAGCGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAA
GATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACC
ACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCG
CGTGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCCGCA
GTTGTTGTGCCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTT
TCCCGCGTTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGAT

AAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACC
CTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTC
GATGGTGTCCGGGATCTCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAATTAATA
CGACTCACTATA

>pACYC-hCAII^{TS} (783 bp, only the hCAII^{TS} fragment is shown)

**ATGGCACATCACTGGGGGTACGGCAAACACAACGGACCTGAGCACTGGCATAAGGA
CTTCCCCATTGCCAAGGGAGAGCGCCAGTCCCCTGTTGACATCGACACTCATACAGC
CAAGTATGACCCTTCCCTGAAGCCCCTGTCTGTTTCCTATGATCAAGCAACTTCCCTGA
GAATCCTCAACAATGGTCATACCTTCAACGTGGAGTTTGATGACTCTCAGGACAAAGC
AGTGCTCAAGGGAGGACCCCTGGATGGCACTTACAGATTGATTCAGTTTCACTTTCAC
TGGGGTTCAcatGATGGACAAGGTTTCAGAGCATACTGTGGATAAAAAGAAATATGCTG
CAGAACTTCACTTGGTTCACTGGAACACCAAATATGGGGATTTTGGGAAAGCTGTGCA
GCAACCTGATGGACTGGCCGTTCTAGGTATTTTTTTGAAGGTTGGCAGCGCTCTGCCG
GGCCTTCAGAAAGTTGTTGATGTGCTGGATTCCATTAACAAGGGCAAGAGTGCTG
ACTTCACTAACTTCGATCCTCGTGGCCTCCTTCCCTGAATCCCTGGATTACTGGACCTAC
CCAGGCTCACTGACCACCCCTCCTTCTGGAATCAGTGACCTGGATTGTGCTCAAGG
AACCCATCAGCGTCAGCAGCGAGCAGGTGAGCAAATTCCGTAACTTAACTTCAATG
GGGAGGGTGAACCCGAAGAACCGATGGTGGACAACCTGGCGCCCAACCCAGCCACTG
AAGAACAGGCAAATCAAAGCTTCCTTCAAATAA**

3.5.3.3 Purification

The frozen pellet of *E. coli* cells expressing the protein was thawed at room temperature. The pellet the expression culture (200 mL culture volume) was resuspended in a lysis buffer (20 mL, lysozyme (0.1 mg/mL), DNaseI (0.01 mg/mL) in the binding buffer: ZnSO₄ (0.5 mM), tris-sulfate (50 mM), pH = 8.6). The suspension was shaken (37 °C, 300 rpm, 1 hour). The lysed suspension was sonicated (Fisherbrand™ Model 705 Sonic Dismembrator with a Microtip, amplitude 50, 10 on and off alternately). The sonicated suspension was centrifuged to remove

cell debris (4 °C, 25000g, 30 minutes). The supernatant was collected. A small amount (10 µL) of the supernatant was put aside for SDS-PAGE analysis.

Pre-packed *p*-aminomethylbenzenesulfonamide-agarose beads or synthesized arylsulfonamide-decorated Sepharose beads^[60] in an open column (30 mL bed volume) were equilibrated with the binding buffer (3 CV). The clear supernatant of the lysate was filtered through a syringe filter (Filtropur S, PES, pore size: 0.2 µm, Sarstedt), and loaded onto the resin. A small amount of the flow-through fraction was put aside for SDS-PAGE analysis. The column was washed with the binding buffer (2 CV) followed by the wash buffer (2 CV, Na₂SO₄ (50 mM), NaClO₄ (50 mM), tris-sulfate (25 mM), pH = 8.8). The bound protein was eluted with the elution buffer (3 CV, NaClO₄ (200 mM), sodium acetate (100 mM), pH adjusted to 5.6 with dilute sulfuric acid). The purity of the protein was confirmed by SDS-PAGE. Only the pure fractions were combined in a dialysis bag (Spectra/Por® 1 Dialysis membrane, MWCO 6-8 kD) and dialyzed against MilliQ water (4 °C). The water was changed several times so that the final salt concentration was below 10 nM. The dialyzed protein solution was transferred to falcon tubes and centrifuged (4000 g for 30 minutes, 4 °C) to remove any precipitate. The supernatant was flash-frozen with liquid nitrogen and lyophilized. The protein was obtained as white cotton-like solid.

3.5.3.4 Cell lysate preparation

The frozen pellet of *E. coli* cells expressing the protein was thawed at room temperature. The pellet from the expression culture (50 mL culture volume) was resuspended in a lysis buffer (15 mL, lysozyme (0.1 mg/mL), DNaseI (0.01 mg/mL) in borate buffer (pH = 8.5, 10 mM)) so that the OD₆₀₀ of the suspension is around 10. The suspension was shaken (37 °C, 250 rpm, 30 minutes). The lysate was then frozen (-20 °C, 2 hours). The frozen lysate was then thawed at room temperature and aliquoted into Eppendorf tubes (1.5 mL). The tubes were shaken for heat shock (55 °C, 600 rpm, 30 minutes). The precipitate was spun down by centrifugation (21100 g, 20 minutes, 4 °C). The clear supernatant was aliquoted into new Eppendorf tubes

(0.5 mL). The lysate was flash-frozen with liquid nitrogen and stored at $-80\text{ }^{\circ}\text{C}$. The lysate was used within 1 month after the preparation.

3.5.4 Catalytic assays

3.5.4.1 General procedure

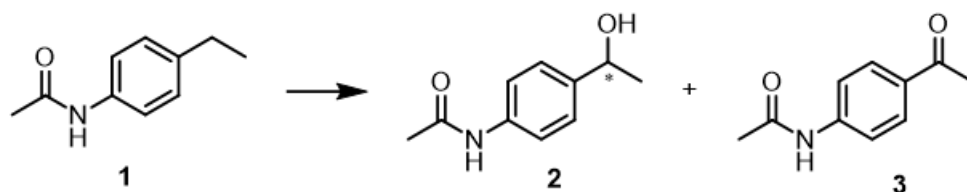
For the reactions with the purified proteins, the stock solution of the proteins was prepared by dissolving the lyophilized protein in MilliQ water or the reaction buffer. The concentration of the stock solution was determined by the absorbance (280 nm, NanoDrop1000 Spectrophotometer, Thermo Fischer Scientific) and the calculated extinction coefficient of the protein (extinction coefficient at 280 nm = 50420 [a.u. · mol⁻¹ · cm⁻¹] for all the mutants tested, calculated based on a reported equation^[61]). For the reactions with the cell lysate, the frozen lysate was thawed at room temperature. The stock solution of ethylbenzoic acid **8** was prepared by dissolving the acid with an equimolar amount of NaOH aq.

The reactions (200 µL) were typically performed in an HPLC glass vial (1.5 mL) or a transparent 96-well microtiter plate. MilliQ water, buffer, ZnSO₄, the protein, [BS-Fe-bTAML]⁻, and the substrate (in water, MeCN or DMSO) was added to a suitable reaction vessel. The reaction was initiated by addition of H₂O₂ (1 ~ 256 equiv. vs. substrate) to the mixture. The reaction in the HPLC vial was shaken with Eppendorf ThermoMixer[®] C (750 rpm, 25 °C). For the reactions on a 96-well plate, the evolution of the absorbance at 410 nm was determined by measuring each well every 30 seconds with a plate reader. The plate was shaken (2 seconds) right before every measurement.

The reaction was quenched by adding HCl aq (1 M, 15 µL) and analyzed with a suitable analytical instrument (See Section 3.5.4.2 Analysis methods and calibration curves).

3.5.4.2 Analysis methods and calibration curves

3.5.4.2.1 Ethylbenzene derivative 1



After quenching the reaction, the internal standard (indole, 4 μmol for 200 μL a reaction mixture) was added, and the mixture was extracted with EtOAc (400 μL x 2). The organic phases were dried over MgSO_4 and subjected to SFC for quantification. Analytical method: Chiralpak IH[®], 4.6 \times 250 mm; gradient method from iPrOH 5% for 3 min, 5 to 20% over the next 10 min, 20% for the next 5 min, 2 mL/min. Injection volume = 10 μL . For the calibration curve, indole (5 mM) and the product standards (from 0.3125 mM to 10 mM) were dissolved in EtOAc. The peak areas of the compounds were integrated at specific wavelengths: indole = 265 nm, ketone **3** = 284 nm, alcohol **2** = 245 nm.

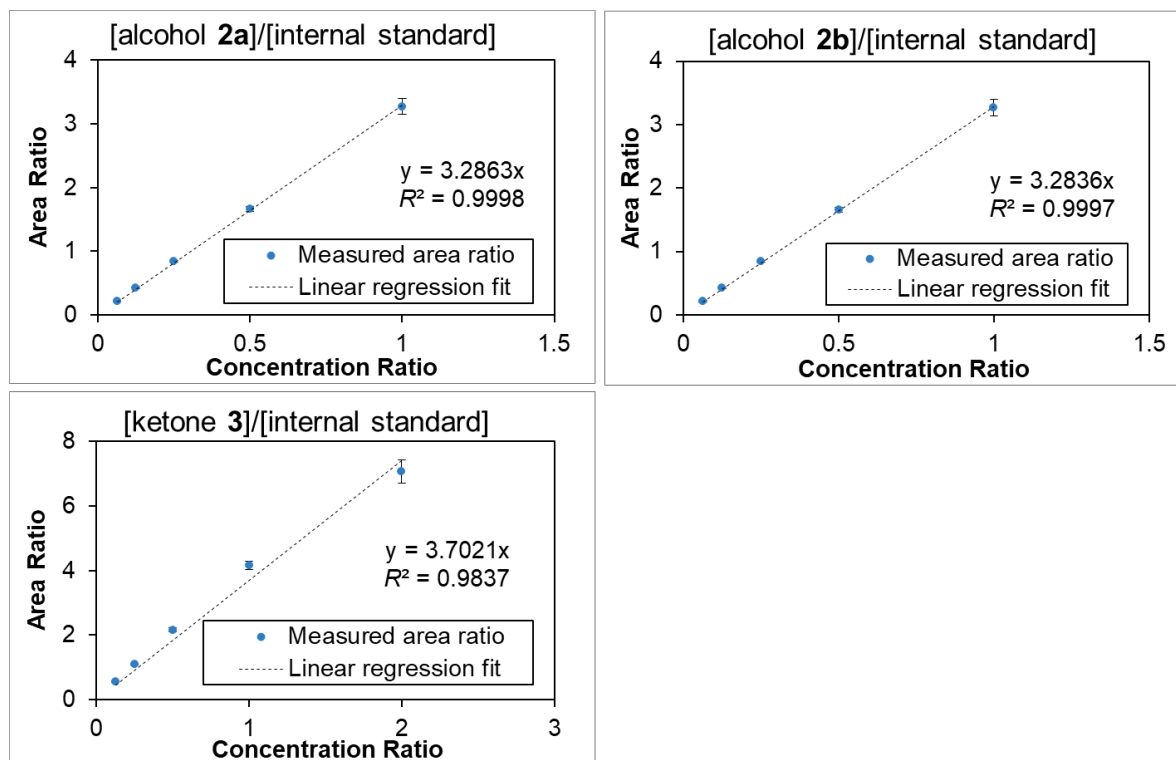
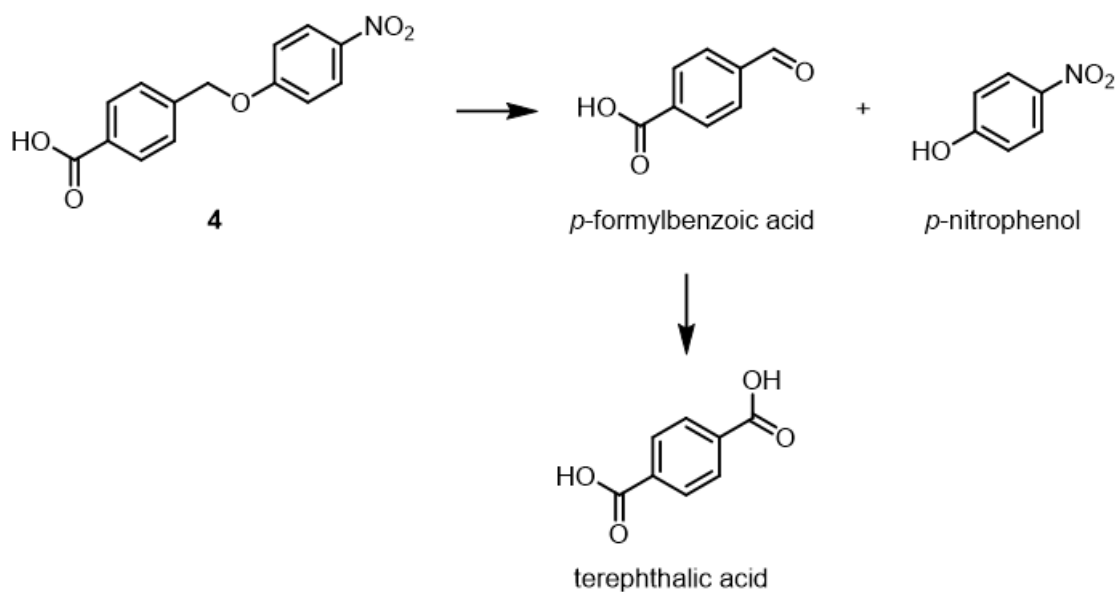


Figure S1. Calibration curves for the reaction with ethylbenzene derivative 1. Alcohols **2a** and **2b** are the two enantiomers of alcohol **2**. The equation and the R^2 value of each linear regression are displayed in each plot area. Internal standard = indole.

3.5.4.2.2 Caged *p*-nitrophenol **4**

The evolution of the absorbance at 410 nm was determined by measuring each well every 30 seconds with a plate reader. The plate was shaken (2 seconds) right before every measurement. After quenching the reaction, internal standard (benzyltriethylammonium bromide, 0.4 μmol for a 200 μL reaction). The protein in the reaction mixture was precipitated with analytical grade MeOH (200 μL) and removed by centrifugation. The clear supernatant was subject for HPLC analysis. Analytical method: Poroshell 120 EC-C18, 2.1 x 50 mm, 1.9 μm ; a gradient method between solvent A (H₂O containing 3% MeCN and 0.1% TFA) and B (MeCN containing 3% H₂O and 0.1% TFA) from 10% solvent B for 0.5 min., 10% to 60% over 0.5 min, 60% for 2 min., 60% to 90% over 0.5 min., 90% for 0.5 min. Injection volume = 0.5 μL . For the calibration line, benzyltriethylammonium bromide (1 mM) and the analytes (from 0.0625 mM to 1 mM) were mixed. The peak areas of the compounds were integrated at specific wavelengths: benzyltriethylammonium bromide, terephthalic acid, *p*-formylbenzoic acid, and the substrate **4** = 254 nm (ref. 500 nm), *p*-nitrophenol = 280 nm (ref. 500 nm).

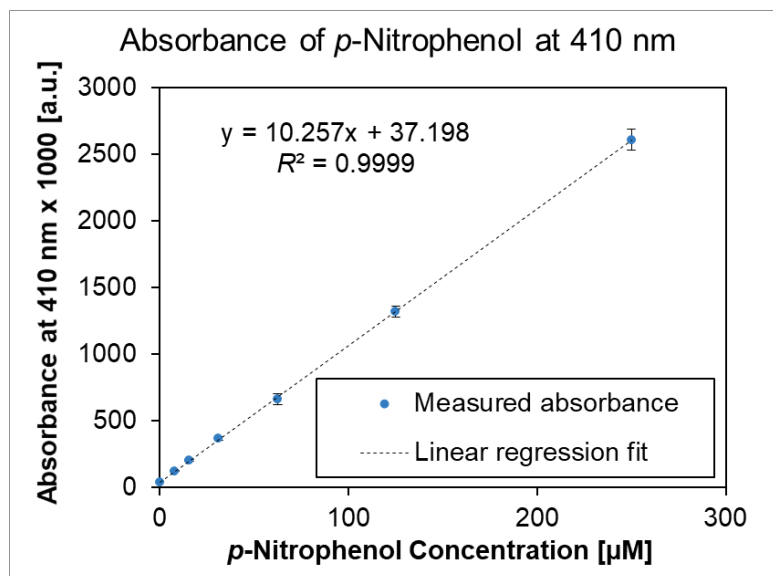


Figure S2. Calibration curve for the microtiter-plate assays for the uncaging of *p*-nitrophenol using ArPases. The equation and the R^2 value of each linear regression fit are displayed in the plot area.

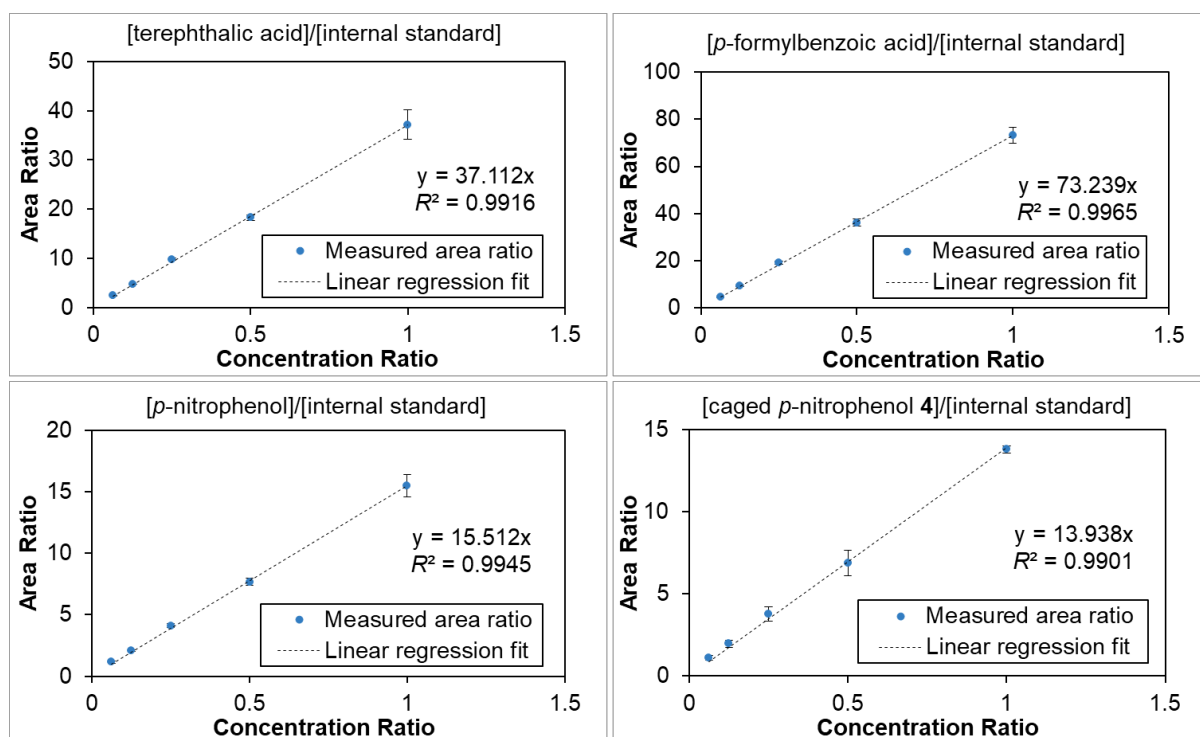
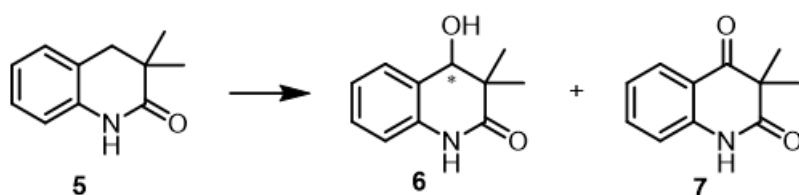


Figure S3. Calibration curves for the uncaging of *p*-nitrophenol using ArPases. The equation and the R^2 value of each linear regression fit are displayed in each plot area. Internal standard = benzyltriethylammonium bromide.

3.5.4.2.3 Lactam **5**

After quenching the reaction, the internal standard (indole, 4 μmol for 200 μL a reaction mixture) was added, and the mixture was extracted with EtOAc (400 μL x 2). The organic phases were dried over MgSO_4 and subjected to SFC for quantification. Analytical method: Chiralpak IF[®], 4.6 \times 250 mm; isocratic method with iPrOH 15% for 20 minutes, 2 mL/min. Injection volume = 10 μL . For the calibration line, indole (5 mM) and the product standards (from 0.15625 mM to 2.5 mM) were dissolved and mixed in analytical grade EtOAc. The peak areas of the compounds were integrated at the following wavelengths: indole = 265 nm, ketone **7** = 332 nm, alcohol **6** = 250 nm.

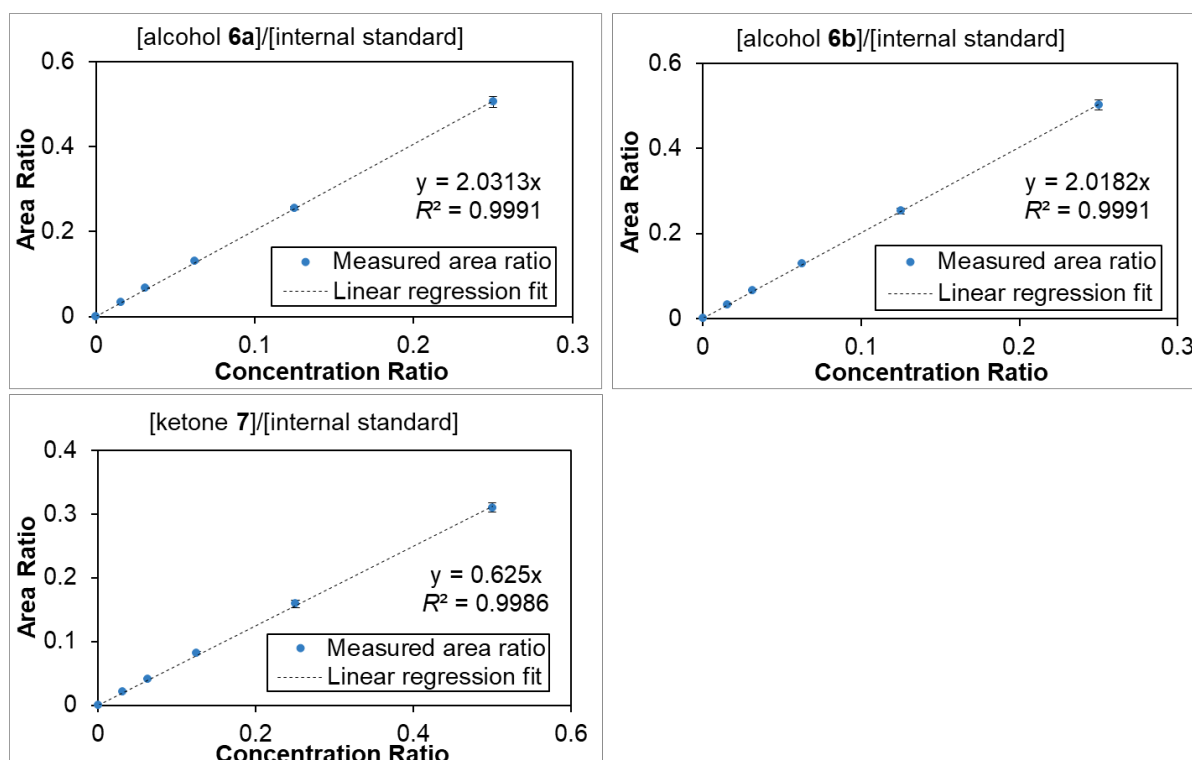
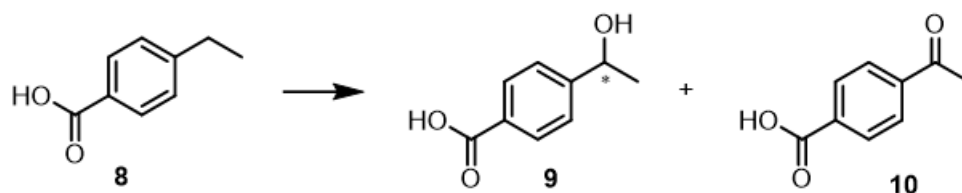


Figure S4. Calibration curves for the reaction with lactam **5**. Alcohols **6a** and **6b** are the two enantiomers of alcohol **6**. The equation and the R^2 value of each linear regression fit are displayed in each plot area. Internal standard = indole.

3.5.4.2.4 Ethylbenzoic acid **8**

After quenching the reaction, the internal standard (bibenzyl, 4 μmol for 200 μL a reaction mixture) was added, and the mixture was extracted with EtOAc (400 μL x 2). The organic phases were dried over MgSO_4 and subjected to SFC for quantification. Analytical method: Chiralpak IG[®], 4.6 \times 250 mm; isocratic method with MeOH 10% for 20 minutes, 2 mL/min. Injection volume = 10 μL . For the calibration curve, bibenzyl (5 mM) and the product standards (from 0.3125 mM to 10 mM) were dissolved and mixed in EtOAc. The peak areas of the compounds were integrated at the following wavelengths: bibenzyl = 260 nm, ketone **10** = 245 nm, alcohols **9** = 235 nm.

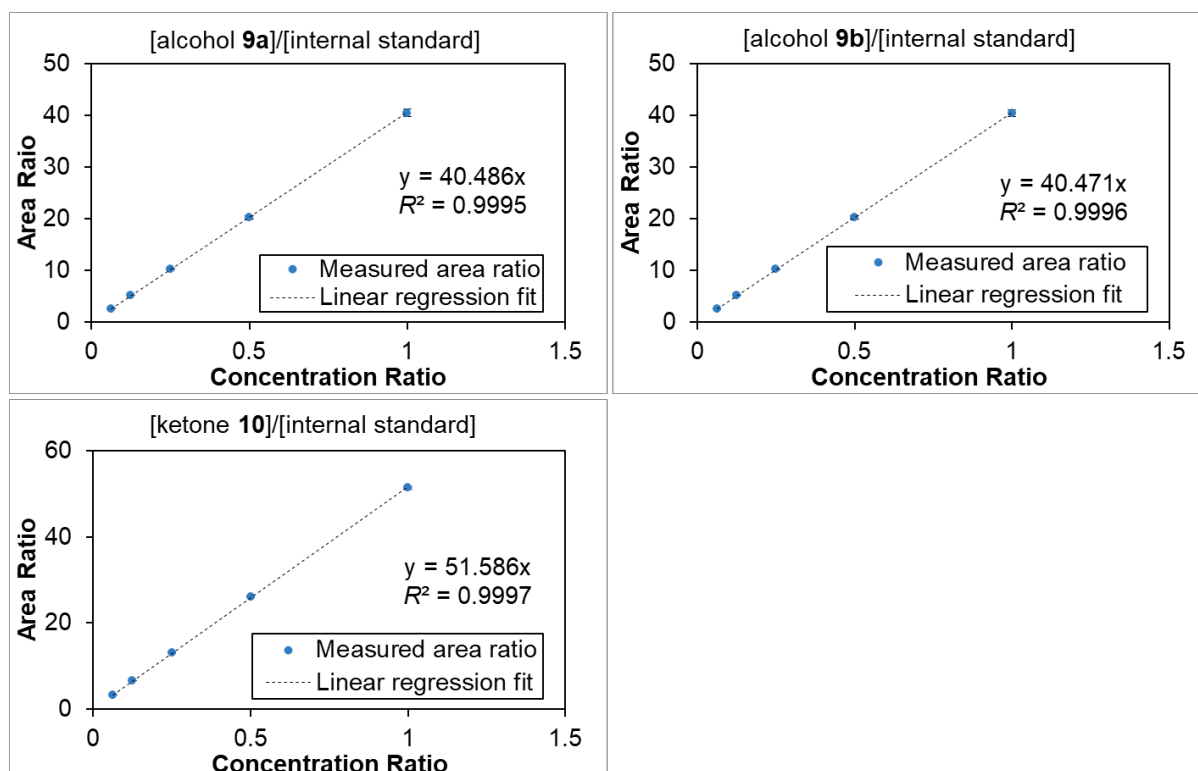
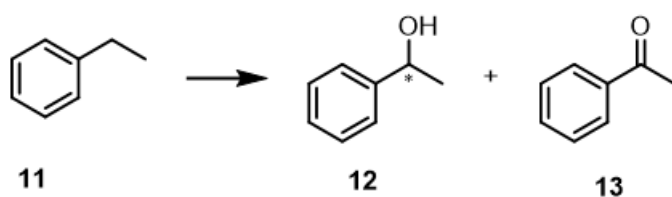


Figure S5. Calibration curves for the reaction with ethylbenzoic acid **8**. Alcohols **9a** and **9b** are the two enantiomers of alcohol **9**. The equation and the R^2 value of each linear regression fit are displayed in each plot area. Internal standard = bibenzyl.

3.5.4.2.5 Ethylbenzene **11**

After quenching the reaction, the internal standard (biphenyl, 0.4 μmol for 200 μL a reaction mixture) was added, and the mixture was extracted with EtOAc (400 μL x 2). The organic phases were dried over MgSO_4 and subjected to GC-FID for quantification. Column oven temperature: 105 $^\circ\text{C}$ for 2 min, 105 $^\circ\text{C}$ to 108 $^\circ\text{C}$ at 2 $^\circ\text{C}/\text{min}$, 108 $^\circ\text{C}$ to 116 $^\circ\text{C}$ at 0.75 $^\circ\text{C}/\text{min}$, 116 to 190 at 10 $^\circ\text{C}/\text{min}$, 190 $^\circ\text{C}$ for 3.5 min, total 25.07 minutes method. For the calibration line, biphenyl (2.5 mM) and the product standards (from 0.15625 mM to 2.5 mM) were dissolved and mixed in analytical grade EtOAc.

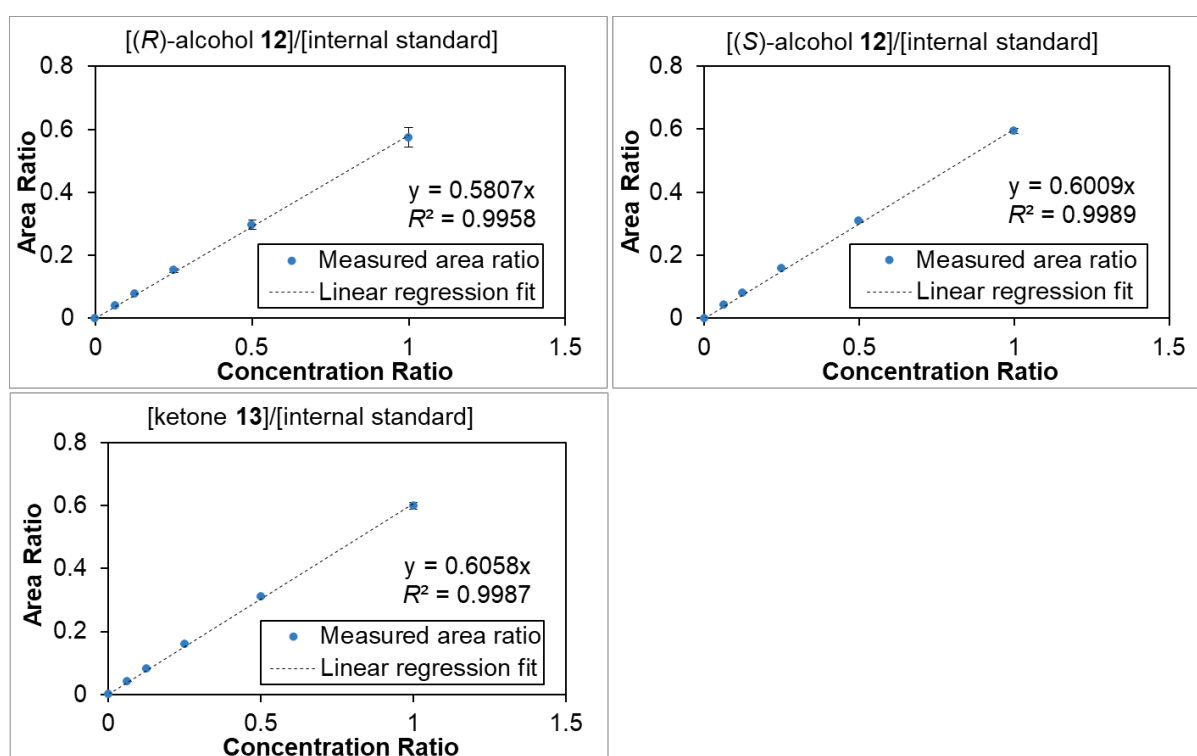
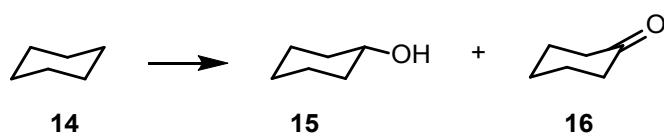


Figure S6. Calibration curves for the reaction with ethylbenzene **11.** The equation and the R^2 value of each linear regression fit are displayed in each plot area. Internal standard = biphenyl.

3.5.4.2.6 Cyclohexane **14**

After quenching the reaction, the internal standard (biphenyl, 2 nmol for 200 μ L a reaction mixture) was added, and the mixture was extracted with EtOAc (200 μ L x 2). The organic phases were dried over $MgSO_4$ and subjected to GC-FID for quantification. Column oven temperature: 50 $^{\circ}C$ to 85 $^{\circ}C$ at 20 $^{\circ}C/min$, 85 $^{\circ}C$ for 2 min, 85 $^{\circ}C$ to 86 $^{\circ}C$ at 1 $^{\circ}C/min$, 86 $^{\circ}C$ for 2 min, 86 $^{\circ}C$ to 87 $^{\circ}C$ at 1 $^{\circ}C/min$, 87 $^{\circ}C$ to 190 $^{\circ}C$ at 20 $^{\circ}C/min$, 190 $^{\circ}C$ for 3 min, total 15.9 minutes method. For the calibration curve, biphenyl (10 μ M) and the product standards (from 6 μ M to 97 μ M) were dissolved and mixed in analytical grade EtOAc.

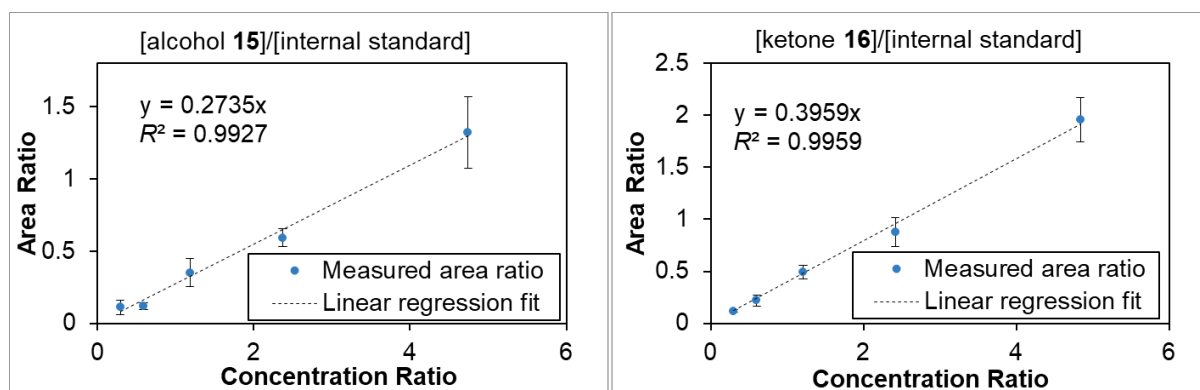


Figure S7. Calibration curves for the reaction with cyclohexane **14.** The equation and the R^2 value of each linear regression fit are displayed in each plot area. Internal standard = biphenyl.

3.5.4.3 Michaelis-Menten kinetics

The Michaelis-Menten kinetic analysis for $[BS-Fe-bTAML]^- \cdot hCAI^{TS}$ and $[BS-Fe-bTAML]^- \cdot hCAI^{TS}-RRIG$ was performed according to the general procedure in the previous sections (Sections 3.5.4.1 General procedure and 3.5.4.2.2 Caged *p*-nitrophenol **4**). Conditions: caged *p*-nitrophenol **4** (31.25 ~ 500 μ M), a hCAI variant (4 μ M), $[BS-Fe-bTAML]^-$ (0.25 μ M), H_2O_2 (4 ~ 128 mM), $ZnSO_4$ (40 μ M), borate buffer (pH = 8.5, 10 mM), DMSO (0.5%v/v), 25 $^{\circ}C$. The

initial reaction rate in each well was calculated according to the calibration line, plotted, and fitted with Origin with the following equations.

$$v = \frac{V^{\text{app}} \cdot a}{K_m^{\text{app}} + a} \quad (1)$$

$$V^{\text{app}} = \frac{V_{\text{max}} \cdot b}{K_m + b} \quad (2)$$

v : Measured initial rate

V^{app} : Apparent V_{max} at a fixed concentration of B

K_m^{app} : Apparent Michaelis constant of substrate A at a fixed concentration of B

V_{max} : Maximum velocity of the reaction

K_m : Michaelis constant of substrate B

a, b : concentration of substrate A or B, respectively

The two substrates, H_2O_2 and cage p -nitrophenol **4**, are represented here as A and B (interchangeable). First, the initial rates v at a fixed concentration of substrate B were plotted over the concentration of substrate A and fitted with eq 1. The obtained V^{app} at each concentration of B are then plotted and fitted with eq 2 to obtain V_{max} and the Michaelis constant (K_m) for substrate B.

3.5.5 Binding assays of the cofactors and hCAII

3.5.5.1 Esterase inhibition assay^[51]

In each well of a 96-well transparent plate, [BS-Fe-bTAML]⁻ (200 μ M to 0.0125 μ M, serial dilution, and 1 well without it), hCAII (1 μ M), potassium phosphate buffer (pH 8.2, 50 mM), and water were mixed. The reaction was initiated by addition of *p*-nitrophenylacetate (in MeCN, final concentration = 0.5 mM). The hydrolysis was monitored based on the absorption of the released *p*-nitrophenol using a plate reader (348 nm, every 30 seconds). The initial rate (typically during the first 2 minutes) of the reaction was calculated, plotted, and fitted using Origin with the formula described in the literature^[62].

3.5.5.2 DNSA displacement assay^[51-52]

First, the dissociation constant of dansylamide (K_{d_DNSA}) was determined by its change in fluorescence using a one-to-one binding to the protein. Based on the K_d value obtained, a competitive binding assay with DNSA and [BS-Fe-bTAML]⁻ was performed.

One-to-one binding assay for DNSA and hCAII. To 15 wells on a black flat-bottom 96-well plate, potassium phosphate buffer (pH = 8.2, 50 mM), hCAII variant (100 nM), DNSA solution in DMSO (10 μ L, final DNSA concentration = 0.565 nM to 100 μ M, serial dilution from well to well), and water was added. The final volume was 200 μ L. The fluorescence in each well was measured (excitation: 280 nm, measurement: 470 nm). The datasets were obtained in triplicate, and they were fitted with a curve using the solver add-in in Excel software according to the following equation^[63].

$$F_{obs} = F_{free} + (F_{bound} - F_{free}) \frac{(C_{hCAII} + C_{DNSA} + K_{d_DNSA}) - \sqrt{(C_{hCAII} + C_{DNSA} + K_{d_DNSA})^2 - 4 \cdot C_{hCAII} \cdot C_{DNSA}}}{2 \cdot C_{hCAII}}$$

F_{obs} : observed fluorescence intensity

F_{free} : fluorescence intensity of the protein when not bound to DNSA

F_{bound} : fluorescence intensity of the protein-DNSA complex

C_{hCAII} : the total concentration of hCAII

C_{DNSA} : the total concentration of DNSA

$K_{\text{d_DNSA}}$: the dissociation constant of DNSA

DNSA displacement assay. To 15 wells on a black flat-bottomed 96-well plate, potassium phosphate buffer (pH = 8.2, 50 mM), the hCAII variant (100 nM), DNSA solution in DMSO (10 μL , final DNSA concentration = 2.25 μM), [BS-Fe-bTAML]⁻ (0.565 nM to 100 μM , serial dilution from well to well) and water was added. The final volume is 200 μL . The fluorescence in each well was measured (excitation: 280 nm, measurement: 470 nm). The datasets were obtained in triplicate, and were fitted with a curve using the solver add-in in Excel software according to the reported equations^[64].

3.5.6 Differential scanning fluorimetry

Thermal denaturation of the protein (1 mg/mL) was monitored by fluorescence intensity (280 nm excitation, 330 nm and 350 nm emission) during a continuous thermal ramp (20 to 95 °C, 1 °C/min). In parallel, apparent absorbance of excitation light due to light-scattering was measured by back-reflection optics, to monitor formation of aggregates.

The ratio of emission intensities at the detection wavelengths (350 nm / 330 nm) was analysed automatically using the “PR. Stability analysis” software, to determine the onset temperature (T_{on}) and midpoint temperature (T_m) for thermal denaturation. The onset temperature marks the departure of the signal from the baseline in the first observable denaturation transition, and the midpoint is the inflexion point of the sigmoidal curve (the maximum in the first derivative). The T_{on} for the wild-type protein was adjusted manually to account for a shallow denaturation transition (between 40 and 50 °C) that was not detected by the software.

3.5.7 Directed evolution of the ArPase

3.5.7.1 Site-directed mutagenesis

Name	Sequence (5' to 3')
Fw H4X	ATGGCACAT NDT TGGGGGTACGGCAAACAC
Rv H4X	GTACCCCA AHN ATGTGCCATGGTATATCT
Fw F20X	CATAAGGAC NDT CCCATTGCCAAGGGAGA
Rv F20X	GCAATGGG AHNG TCCTTATGCCAGTGCT
Fw L60X	TCCCTGAGAAT NDT AACAATGGTCATACCTTCAAC
Rv L60X	ATGACCATTGTT AHNG ATTCTCAGGGAAGTTGCTT
Fw L60X for the 3 rd round	TCCCTGAGAAT NDT AACAATCGTCATACCTTCAAC
Rv L60X for the 3 rd round	ATGACGATTGTT AHNG ATTCTCAGGGAAGTTGCTT
Fw G63X	CAAT NDT CATACCTTCAACGTGG
Rv G63X	GGTATG AHN ATTGTTGAGGATTCTC
Fw G63X for the 2 nd round	AACAAT NDT CATACCTTCAACGTGC
Fw I91X	TTACAGATTG NDT CAGTTTCACTTTCACTGG
Rv I91X	AGTGAACTG AHN CAATCTGTAAGTGCCATC
Fw G131X	TGGGGATTTT NDT AAGCTGTGCAGCAACC
Rv G131X	GCACAGCTTT AHN AAAATCCCCATATTTGGT
Fw V134X	TGGGAAAGCT NDT CAGCAACCTGATGGACT
Rv V134X	CAGGTTGCTG AHN AGCTTTCCCAAATCCCC
Fw Q135X	AAAGCTGTG NDT CAACCTGATGGACTGGC
Rv Q135X	CATCAGGTTG AHN CACAGCTTTCCCAAATC
Fw K169X	TCCATTA AACAN DT TGGCAAGAGTGCTGACTTC
Rv K169X	AGCACTCTTGCC AHNT GTTTTAATGGAATCCAG
Fw P200X	CACC NDT CCTCTTCTGGAATCAG
Rv P200X	GAGG AHNG TGGTCAGTGAGCCT
Fw L202X	ACCCCTCCT NDT CTGGAATCAGTGACCT
Rv L202X	TGATTCCAG AHN AGGAGGGGTGGTCAGTGA
Fw Gibson Assembly NDT library	TCCAGCTGAACGGTCTGGTTATAGGT
Rv Gibson Assembly NDT library	GACCGTTCAGCTGGATATTACGGCCT

Table S2. List of primers designed and purchased for site-directed randomization of the targeted positions. Fw = forward primer, Rv = reverse primer.

Up to 12 positions (H4, F20, L60, G63, E69, I91, G131, V134, Q135, K169, P200, and L202) were randomized with PCR mutagenesis using degenerate NDT codons. The primers are collected in Table S2. The resulting gene fragments were assembled by Gibson Assembly. The reaction mixture was then used for transforming *E. coli* Top10 or DH5 α cells, following the manufacturer's protocol.

The obtained colonies were harvested together, and the plasmids were extracted with a mini-prep kit. The pooled plasmid library was sequenced by Sanger sequencing to determine its diversity.

3.5.7.2 Protein expression and cell lysate preparation

The pooled plasmid library was used for the transformation of *E. coli* BL21(DE3) chemically-competent cells, and the cells were treated as described in Section 3.1 Cloning and Expression . After obtaining the colonies, at least 46 were picked with sterilized toothpicks and inoculated into an LB medium (0.3 mL), supplemented with chloramphenicol (30 μ g/mL) in a 96-deep-well plate. The glycerol stocks of *E. coli* cells harboring the plasmid of the parent variant or the empty vector were inoculated into at least 1 well per plate as controls. The plates were sealed with sealing membrane (Breathe Easier sealing membrane for multiwell plates, Merck) and shaken overnight (37 $^{\circ}$ C, 300 rpm). The next morning, the preculture from each well (10 μ L) was transferred to LB medium (1 mL) supplemented with chloramphenicol (30 μ g/mL) on a 96-deep-well U-bottom plate. The plate was shaken (37 $^{\circ}$ C, 300 rpm, 2 hours) until the OD₆₀₀ reached 0.6 – 0.8. Then, IPTG (0.5 mM) was added to initiate the overexpression, and the plate was shaken (20 $^{\circ}$ C at 300 rpm, 20 hours). The remainder of the preculture was used for the preparation of the glycerol stock (25%v/v glycerol). After overexpression, the cells were harvested by centrifugation (4400g, 10 minutes at 4 $^{\circ}$ C). The supernatant was discarded by decantation, and the plate was laid upside-down on a paper towel (ca. 15 seconds). The plates were sealed with an aluminium foil seal (VWR) and stored at –20 $^{\circ}$ C overnight.

For the cell lysate preparation, the cell pellets were thawed at room temperature. The cells were suspended in lysis buffer (150 μ L, lysozyme (1 mg/mL) and 1 μ g/mL DNaseI) in the reaction buffer) and the plate was shaken (37 $^{\circ}$ C at 300 rpm for 30 min.). The lysed cells were frozen (-20 $^{\circ}$ C, 2 hours). The lysate was thawed again at room temperature and transferred into a 96-well PCR plate. The lysate was heated (55 $^{\circ}$ C, 30 minutes) using a PCR cycler. After cooling to room temperature, the lysate was flash-frozen in liquid nitrogen and stored at -80 $^{\circ}$ C. The lysate was used for the following screening assay within a week.

3.5.7.3 Screening assays

Cofactor stock solution in water (10 μ M, 5 μ L) was distributed into a 384-well plate. The clear cell lysate (25 μ L, thawed at room temperature) was added from well to well and mixed well with the cofactor. After incubation at room temperature (10 min), a mixture of the reaction components (20 μ L, substrate **4** (0.5 mM) in DMSO, H₂O₂ (32 mM), ZnSO₄ (1 μ M), borate buffer (pH=8.5, 10 mM), water, the final reagent concentrations present in the reaction mixture are listed) was added and mixed quickly and gently. The evolution of the absorbance at 410 nm was determined by measuring each well every 30 seconds with a plate reader.

The initial rates of the reactions from each well were compared. Up to 5 wells with improved activity vs. the positive control (i.e. the parent) was detected for every mutated position. The corresponding glycerol stock was sequenced. The hit mutants were further validated using purified hCA samples.

3.5.8 Crystallographic characterization of [BS-Fe-bTAML]⁻ · hCAII^{TS} and [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRI

Lyophilized hCAII^{TS} and hCAII^{TS}-RRI were dissolved in ultrapure water (18.2 MΩ·cm, MilliQ, Millipore Corporation, Burlington, USA) separately. The 96 well crystallization screens were set up by mixing precipitation buffer (0.15 μL) with protein solution (0.15 μL) with a robot (Crystal Gryphon, Art Robins Instruments, USA; MRC 2-well plates, SWISSCI, Switzerland), equilibrated against the precipitation buffer (60 μL), and stored (20 °C).

The hCAII^{TS} (stock concentration of ~5 mg/mL, ~200 μM) was co-crystallized with a 2.5-fold excess of [BS-Fe-bTAML]⁻ (final DMSO concentration of 2%v/v) at room temperature. The crystals in space group R3 were obtained with the precipitation buffer containing magnesium acetate tetrahydrate (0.1 M), MOPS/NaOH (pH 7.5, 0.1 M), PEG 8K (12% w/v). Thin plate-like crystals in space group C2 were grown with the precipitation buffer containing sodium acetate trihydrate (0.2 M), Bis-Tris/HCl (pH 5.5, 0.1M), PEG 3,350 (25% w/v). The crystals were frozen in liquid nitrogen after a quick soak in reservoir containing ethylene glycol (25%v/v).

For hCAII^{TS}-RRI, a more concentrated stock solution was used (~15 mg/mL). Apo crystals were grown in sodium acetate (0.2 M), sodium cacodylate (0.1 M, solution pH = 6.5), PEG 8K (30 %w/v), Shotgun SG1, Molecular Dimensions. For the cofactor soaking, a solution of [BS-Fe-bTAML]⁻ (0.5 μL, 10 mM in DMSO) was mixed with the precipitation buffer (4.5 μL). This mixed solution (2 μL) was added to the drop containing the crystals. After soaking (overnight, 20 °C), the crystals were cryo-protected with ethylene glycol (25%) and flash-frozen in liquid nitrogen before data collection.

For [BS-Fe-bTAML]⁻ · hCAII^{TS}, the data collection was carried out at the Swiss Light Source beam line X06SA at a wavelength of 1.0000 Å. For [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRI, the data collection was carried out at the Diamond Light Source beam line I03 at a wavelength of 0.9763 Å. XDS^[65] was used for crystal indexing, integration and AIMLESS^[66] for scaling, within the graphical interface CCP4i2^[67] of the CCP4 suite^[68]. The structures were solved by molecular replacement using PHASER-MR^[69] and an AlphaFold^[70] generated search model. Refinement

was carried out by REFMAC5^[71] and for structure modeling and electron-density visualization COOT^[72] was used. Ligand restraints were generated using eLBOW^[73]. Figures were generated with PyMOL (the PyMOL Molecular Graphics System, Version 2.5.0, Schrödinger, LLC). Data collection and refinement statistics are listed in Table S3.

For both [BS-Fe-bTAML]⁻ · hCAII^{TS} and [BS-Fe-bTAML]⁻ · hCAII^{TS}-RRI, one monomer per asymmetric unit was found. Residual electron density in the Fo-Fc map was observed in the “active site” of the protein for both structures. Anomalous dispersion density was observed for the Zn²⁺ coordinated to His94, His96 and His119. For modeling of [BS-Fe-bTAML]⁻, we used the anomalous density peak of iron. The 2Fo-Fc map showed a weaker electron density for the iron-binding part of the cofactor, suggesting higher flexibility.

For hCAII^{TS}-RRI, after modeling [BS-Fe-bTAML]⁻, there was still unexplained density next to His4 and His64 within this hCAII protomer, and His36 on the neighboring protomer. A Zn²⁺ ion and a sulfonamide group perfectly fitted the density and were placed.

hCAII ^{TS} variant	hCAII ^{TS}	hCAII ^{TS}	hCAII ^{TS} -RRI
Cofactor	[BS-Fe-bTAML] ⁻	[BS-Fe-bTAML] ⁻	[BS-Fe-bTAML] ⁻
PDB Code	9F2F	9F2E	9F15

Data Processing Statistics

Resolution Range (Å)	44.79 - 1.52 (1.55 - 1.52)	41.72 - 2.01 (2.09 - 2.01)	38.75 - 1.93 (1.999 - 1.93)
Cell Parameters - a, b, c (Å) - α, β, γ (°)	179.17, 179.17 109.71, 90.0, 90.0, 120.0	85.39, 47.20, 71.09, 90.00, 113.68, 90.00	46.34, 70.66, 146.35, 90.00, 90.00, 90.00
Space group	R 3 :H	C 1 2 1	C 2 2 2 1
Unique reflections	202072 (20121)	17306 (1683)	17297 (1563)
Rmerge (%)	13.2 (201.2)	14.6 (131.9)	16.0 (130.3)
Multiplicity	21.1 (20.4)	6.9 (6.9)	13.1 (12.3)
Mean I/Sig(I)	13.0 (1.7)	9.5 (1.3)	12.3 (2.1)
Completeness (%)	100 (99.9)	99.8 (98.2)	93.4 (85.8)
CC (1/2)	0.999 (0.717)	0.997 (0.510)	0.998 (0.711)

Structure Refinement Statistics

R _{work} /R _{free}	0.149/0.168	0.182/0.220	0.18/0.24
RMS deviation			
-Bond length (Å)	0.015	0.010	0.066
-Bond angles (°)	1.59	1.35	3.28
-Ramachandran favored (%)	96.2	96.5	96.5
Average B-factors (Å ²)			
-Protein	24.0	36.6	28.2
-Ligands	23.5	30.6	35.0
-Solvent	39.8	42.6	31.4

Table S3. Data processing and crystal structure refinement statistics

3.5.9 Supporting Tables

	K_d	Method
hCAII ^{WT}	47.8 ± 20.5 nM	Esterase inhibition
	34.5 ± 10 nM	DNSA displacement
hCAII ^{TS}	51.5 ± 8 nM	DNSA displacement

Table S4. Summary of the affinity assays. [BS-Fe-bTAML]⁻ exhibited comparable affinity to both hCAII^{WT} and hCAII^{TS}.

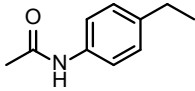
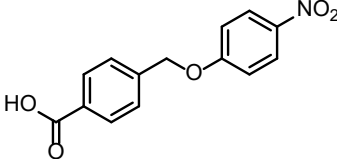
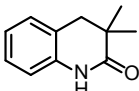
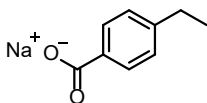
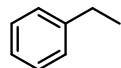

Entry	Substrate	hCAII variant	TTON ^a	e.e. %	[alcohol]/ [ketone]
1 ^{b,c}		TS-RRI	1722 ± 53	4	1.3
2 ^b		TS-RRI	2629 ± 412	5	0.9
3 ^{c,d}	1	TS-RRIG	643 ± 60	20 ± 0.4	1.4
4 ^d		TS-RRIG	1110 ± 33	12 ± 0.5	0.7
5 ^e		TS-RRIG	830 ± 65 ^f	-	-
	4				
6 ^{c,g}		TS-RRI	486 ± 7	53 ± 1	6.4
7 ^g		TS-RRI	1041 ± 23	29 ± 3	1.4
8 ^{c,h}	5	TS-RRI	348 ± 28	50 ± 5	9.9
9 ^h		TS-RRI	459 ± 37	46 ± 1	7.1
10 ^{b,c,i}		TS-RRI	243 ± 9	8	7.0
11 ^{b,i}		TS-RRI	265 ± 3	10 ± 1	7.1
	8				
12 ^d		TS-RRIG	244 ± 14	8	3.7
13 ^j		TS-RRI	44 ± 3.4	-8	2.8
14 ^d	11	TS-RRIG	23 ± 3.4	- ^k	0.7
15 ^l		TS-RRI	0.3	-	1.3
	14				

Table S5. Summary of the enzymatic assays with the evolved ArPases. The products were analyzed using either SFC (**1,5,8**), HPLC (**4**), or GC-FID (**11, 14**). The reactions (200 μM) were quenched by addition of HCl aq (1 M, 15 μM), and a suitable internal standard was added. For the

HPLC analysis, the protein was removed by adding MeOH (200 μ M), followed by centrifugation. For the SFC and GC-FID analysis, the organic compounds were extracted with EtOAc (400 μ M x 2). The combined organic fractions were dried over magnesium sulfate.

^aTTON corresponds to the sum of the TONs determined for both the alcohol and the ketone.

^bSubstrate (10 mM), hCAII variant (10 μ M), [BS-Fe-bTAML]⁻ (4 μ M), H₂O₂ (20 mM), ZnSO₄ (10 μ M), phosphate buffer (pH = 7.2, 10 mM), MeCN (5%v/v), 25 °C, 18 hours.

^c1-hour reaction.

^dSubstrate (5 mM), hCAII variant (10 μ M), [BS-Fe-bTAML]⁻ (4 μ M), H₂O₂ (10 mM), ZnSO₄ (10 μ M), borate buffer (pH = 8.5, 10 mM), DMSO (0.5%v/v), 25 °C, 18 hours.

^eSubstrate (0.5 mM), hCAII variant (4 μ M), [BS-Fe-bTAML]⁻ (0.5 μ M), H₂O₂ (32 mM), ZnSO₄ (10 μ M), borate buffer (pH = 8.5, 10 mM), DMSO (0.5%v/v), 25 °C, 18 hours.

^fSum of *p*-nitrophenol and terephthalic acid

^gSubstrate (5 mM), hCAII variant (10 μ M), [BS-Fe-bTAML]⁻ (4 μ M), H₂O₂ (10 mM), ZnSO₄ (10 μ M), phosphate buffer (pH = 7.2, 10 mM), MeCN (20%v/v), 25 °C, 18 hours.

^hSubstrate (5 mM), hCAII variant (10 μ M), [BS-Fe-bTAML]⁻ (4 μ M), H₂O₂ (10 mM), ZnSO₄ (10 μ M), borate buffer (pH = 8.5, 10 mM), MeCN (20%v/v), 25 °C, 18 hours.

ⁱMeCN was not used.

^jSubstrate (10 mM), hCAII variant (10 μ M), [BS-Fe-bTAML]⁻ (4 μ M), H₂O₂ (20 mM), ZnSO₄ (10 μ M), phosphate buffer (pH = 7.2, 10 mM), MeCN (20%v/v), 25 °C, 18 hours.

^kThe peak area of the alcohol products was too small for an accurate determination of the e.e.

^lSubstrate (100 μ L, 0.94 mmol, 4.7×10^8 equiv. vs. cofactor), hCAII variant (10 μ M), [BS-Fe-bTAML]⁻ (10 μ M), H₂O₂ (100 mM), ZnSO₄ (10 μ M), phosphate buffer (pH = 7.2, 10 mM), MeCN (20%v/v), 25 °C, 18 hours.

3.5.10 Supporting Figures

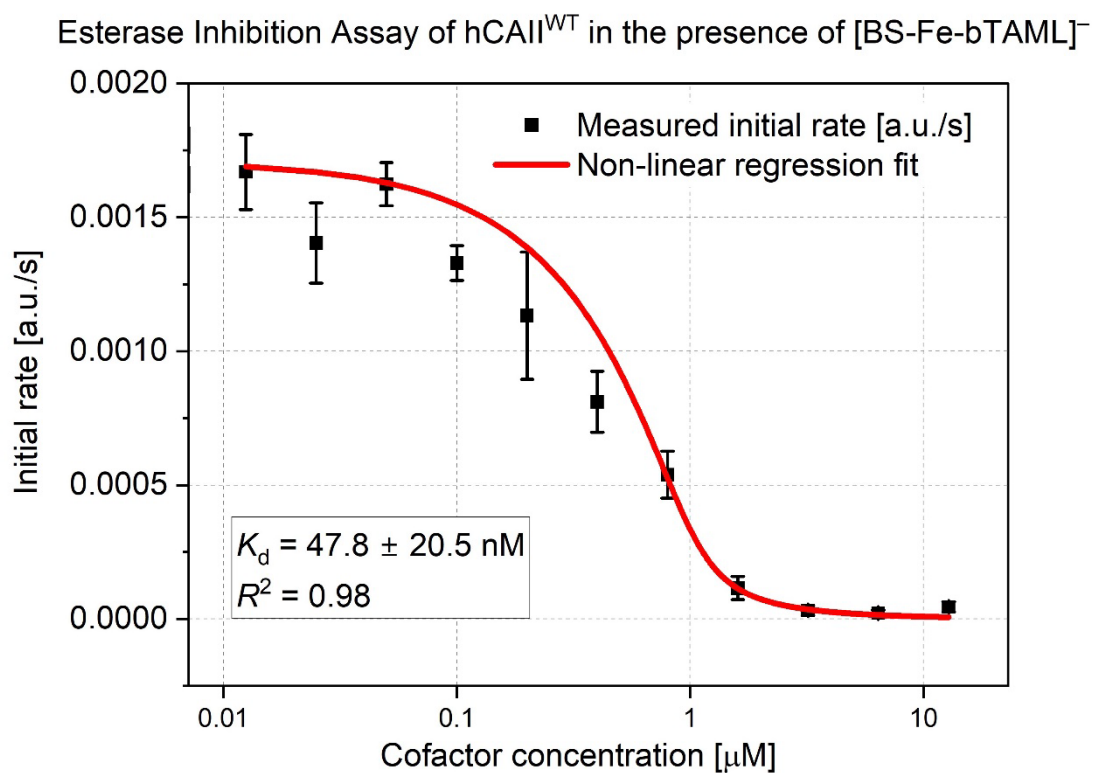
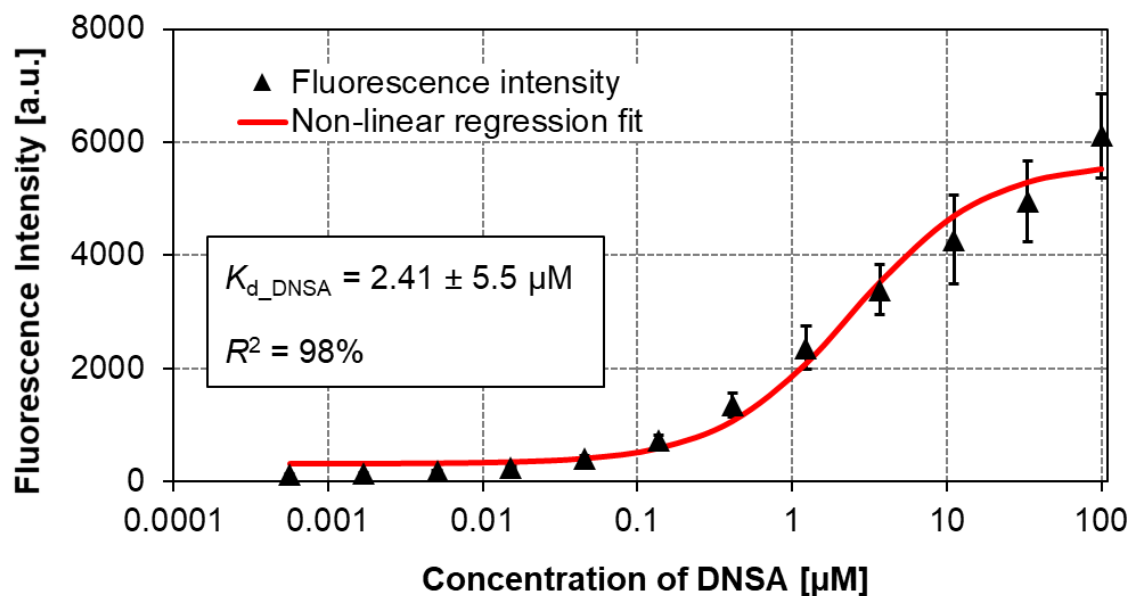


Figure S8. Esterase inhibition assay of hCAII^{WT} in the presence [BS-Fe-bTAML]⁻.

One-to-one binding assay for DNSA and hCAII^{WT}

Competitive binding assay

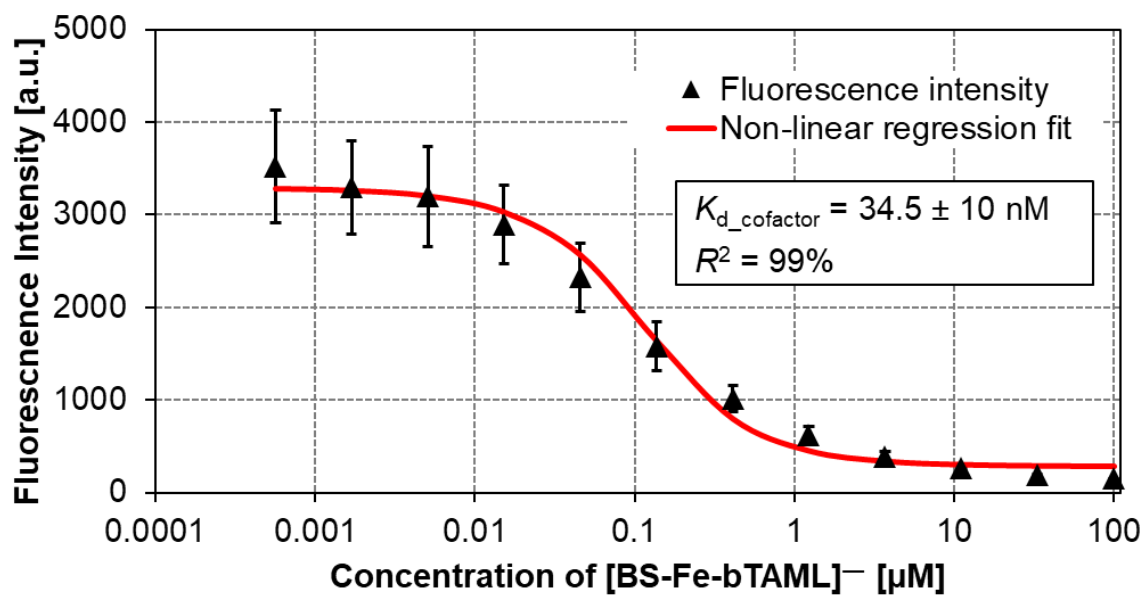


Figure S9. Displacement binding assay for hCAII^{WT} and [BS-Fe-bTAML]⁻. The determined K_d for the cofactor matches with the esterase inhibition assay.

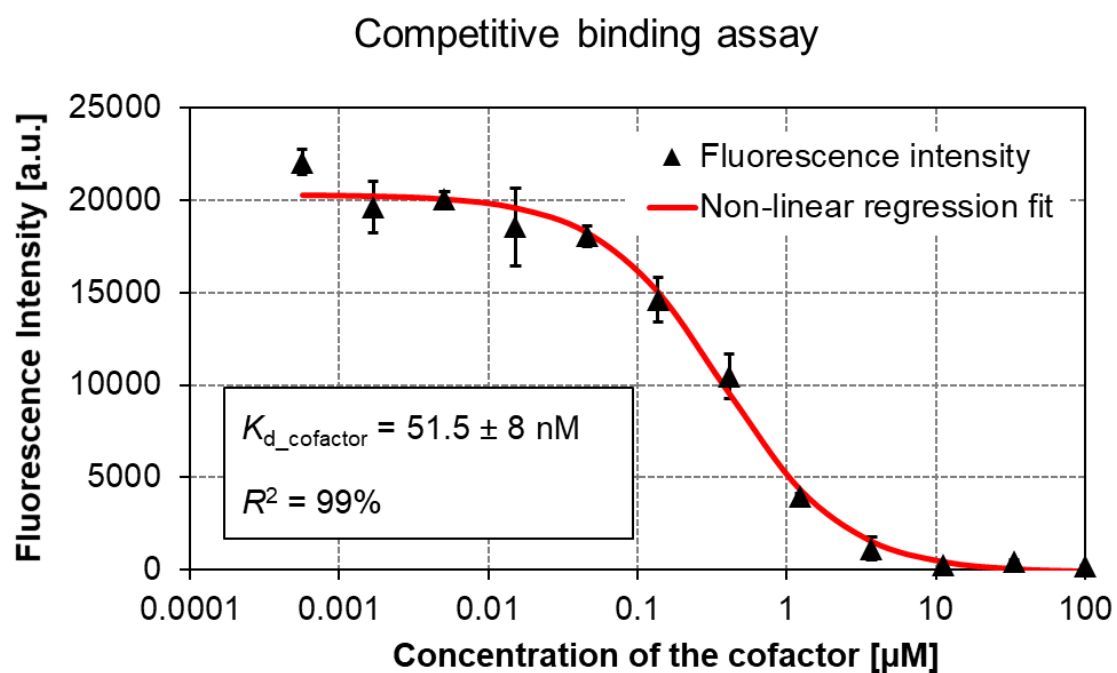
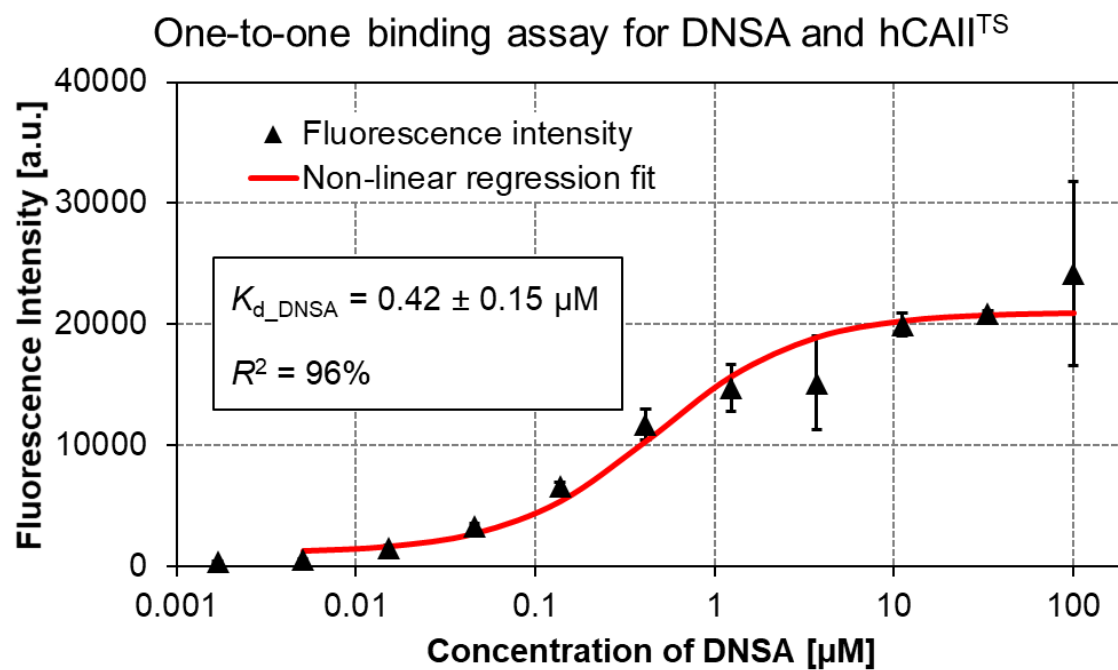
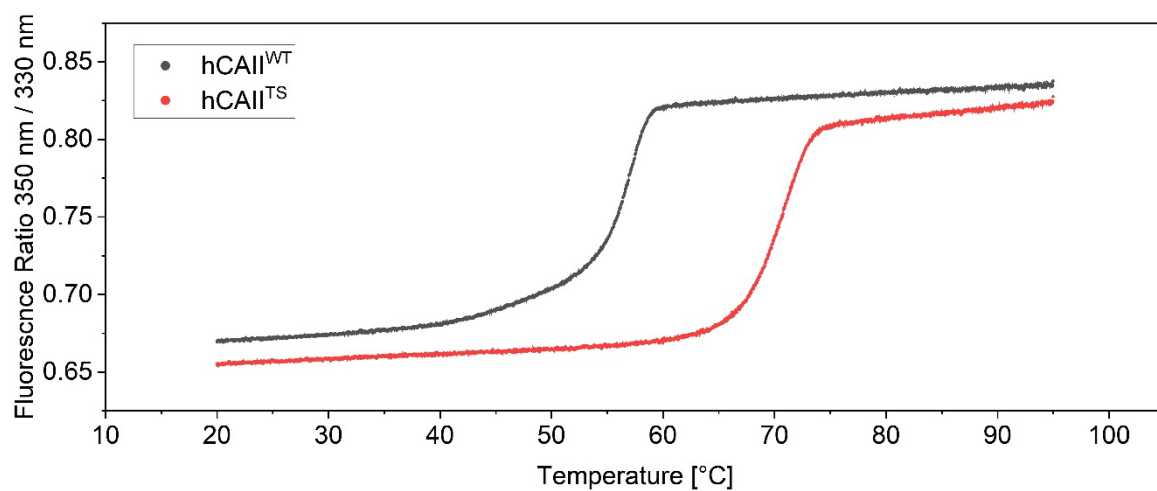


Figure S10. Displacement binding assay for hCAII^{TS} and [BS-Fe-bTAML]⁻. The determined $K_{d_cofactor}$ was comparable to hCAII^{WT}.

(a)



(b)

	T_m (°C)	T_{on} (°C)
hCAII ^{WT}	56.64 ± 0.01	56.64 ± 0.01
hCAII ^{TS}	70.56 ± 0.01	70.56 ± 0.01

Figure S11. Melting temperature of hCAII^{WT} and hCAII^{TS} determined with differential scanning fluorimetry. (a) Plot of the ratio of the fluorescence intensity (350 nm / 330 nm). (b) Determined midpoint (T_m) and onset (T_{on}) temperature of the hCAII variants.

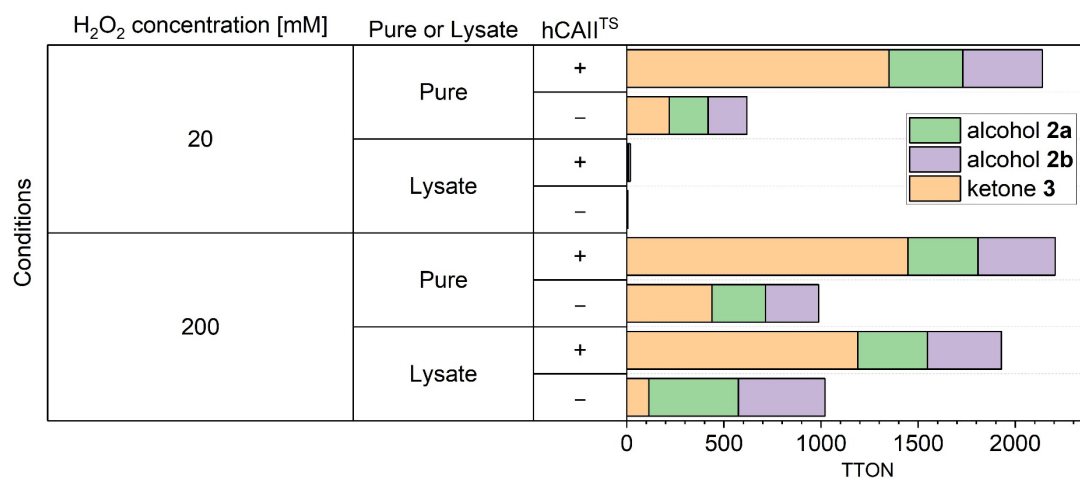
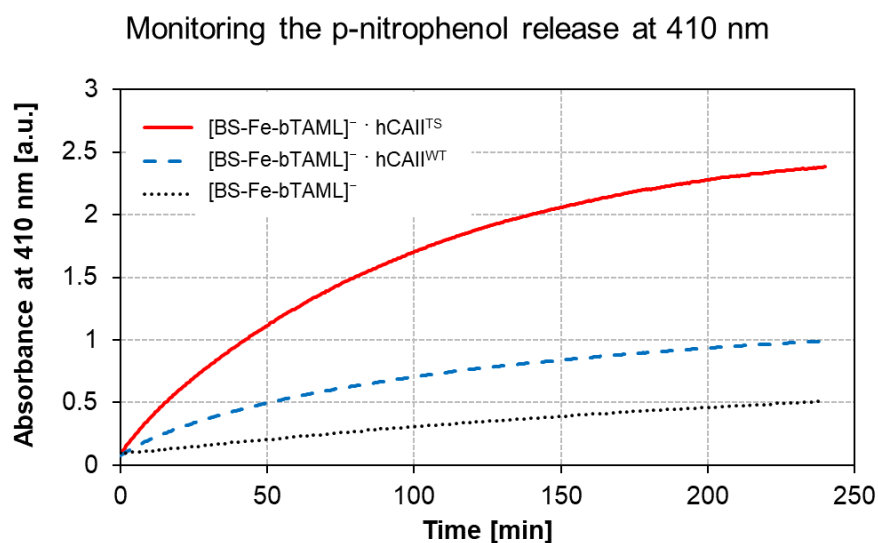


Figure S12. Comparison of TTONs between pure hCAII^{TS} and cell lysate. Conditions: ethylbenzene derivative **1** (10 mM), [BS-Fe-bTAML]⁻ (4 μM), hCAII^{TS} (50 μM from a pure stock in MilliQ water or 100 uL of cell lysate), H₂O₂ (20 or 200 mM), MeCN (5%v/v), phosphate buffered saline, 25 °C, 18 hours. [BS-Fe-bTAML]⁻ was added in all the reactions. The reactions using cell lysate without hCAII^{TS} were performed with a cell lysate prepared from *E. coli* cells that harbor an empty expression vector (i.e. no hCAII^{TS} present).

(a)



(b)

	TTON		
	<i>p</i> -formylbenzoic acid (HPLC)	<i>p</i> -nitrophenol (HPLC)	<i>p</i> -nitrophenol (Plate reader)
hCAII ^{WT}	152	154	173
hCAII ^{TS}	405	422	434
[BS-Fe-bTAML] ⁻	59	62	82

Figure S13. Comparison of both quantification methods for the *p*-nitrophenol uncaging with either (a) a plate reader assay and (b) an HPLC. Conditions: caged *p*-nitrophenol **4** (0.5 mM), [BS-Fe-bTAML]⁻ (1 μM), hCAII variant (6.5 μM), H₂O₂ (1 mM), DMSO (0.5%v/v), borate buffer (pH 8.5, 10 mM), 25 °C, 5 hours. The reaction mixture (200 μL) was mixed on a transparent 96-well plate. The evolution of the absorbance at 410 nm was determined by measuring each well every 30 seconds with a plate reader. The reaction was quenched with HCl aq (1 M, 15 μL). The protein was removed by adding MeOH (100 μL) and centrifugation. The clear fraction was subject to HPLC analysis.

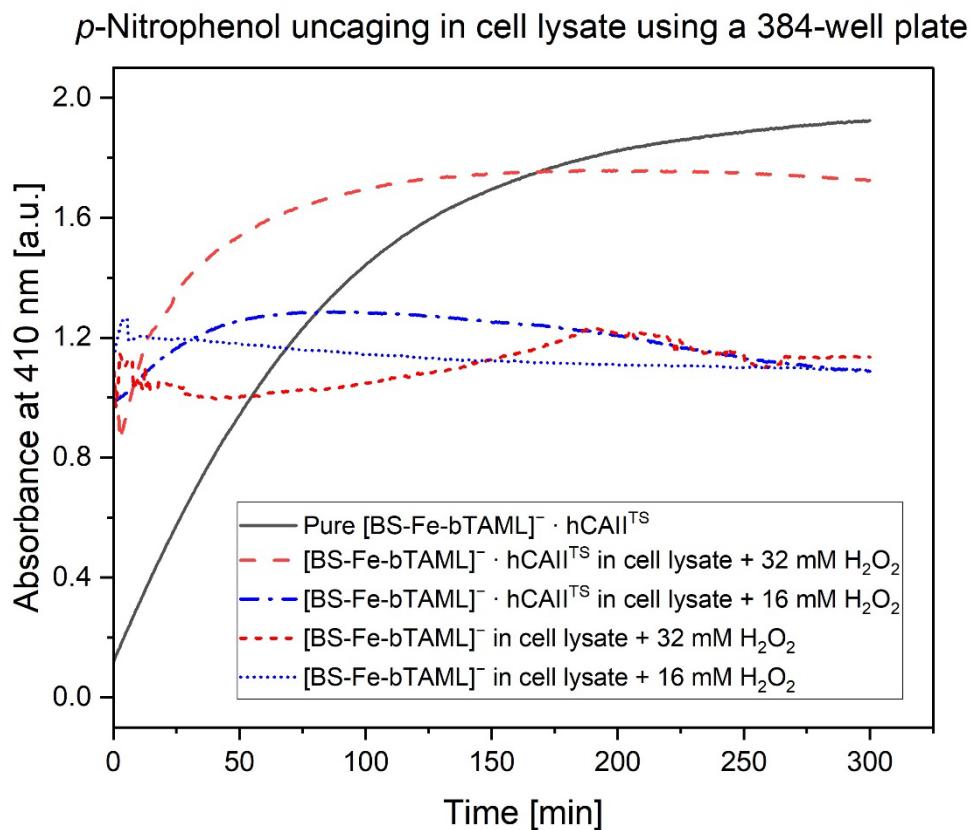


Figure S14. *p*-Nitrophenol uncaging in cell lysate using a 384-well plate. Conditions: caged *p*-nitrophenol **4** (0.5 mM), [BS-Fe-bTAML]⁻ (1 μM), hCAII^{TS} (6.5 μM from a pure stock in MilliQ water or half the reaction volume of cell lysate), H₂O₂ (1 mM if not indicated otherwise), DMSO (0.5%v/v), borate buffer (pH 8.5, 10 mM), 25 °C, 5 hours. The reaction mixture (200 μL) was mixed on a transparent 96-well plate. The evolution of the absorbance at 410 nm was determined by measuring each well every 30 seconds with a plate reader. The reaction with the cell lysate in a 384-well plate showed higher background absorption than with the pure protein. We found that addition of the excess amount of H₂O₂ (32,000 equiv. vs. [BS-Fe-bTAML]⁻) enables the differentiation between [BS-Fe-bTAML]⁻ and [BS-Fe-bTAML]⁻ · hCAII^{TS}.

Comparison of the turnover frequency between the screening and the validation in the presence of [BS-Fe-bTAML]⁻ and hCAII^{TS} single mutants

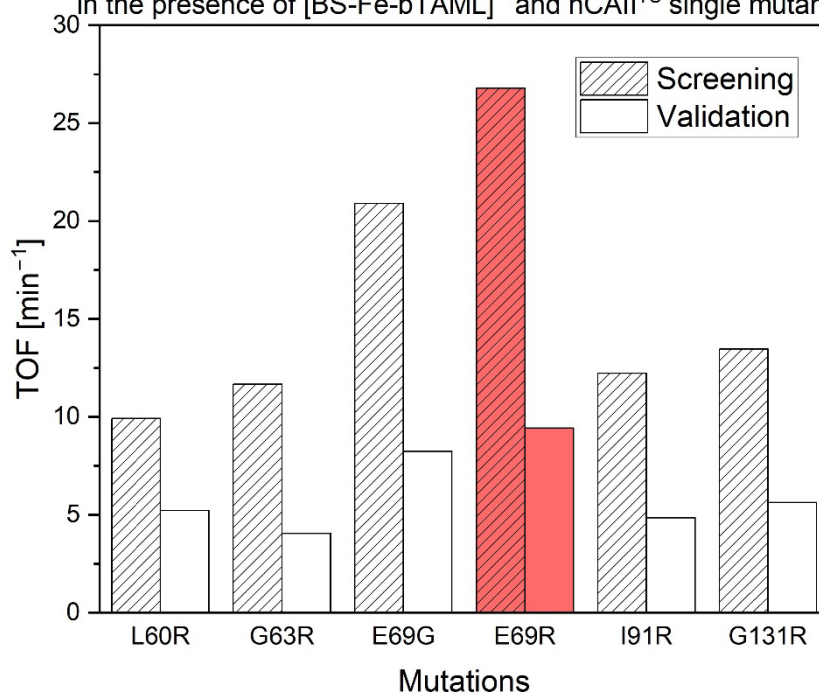


Figure S15. Comparison of the turnover frequency between the screening and the validation in the presence of [cofactor] and hCAII^{TS} single mutants. The cell lysate screening and the validation using the pure hCAII^{TS} were performed in the presence of different amount of H₂O₂ (32 mM, and 1 mM, respectively). This difference accounts for the different TOFs between the screening and the validation. The hCAII^{TS}-E69R is highlighted in red.

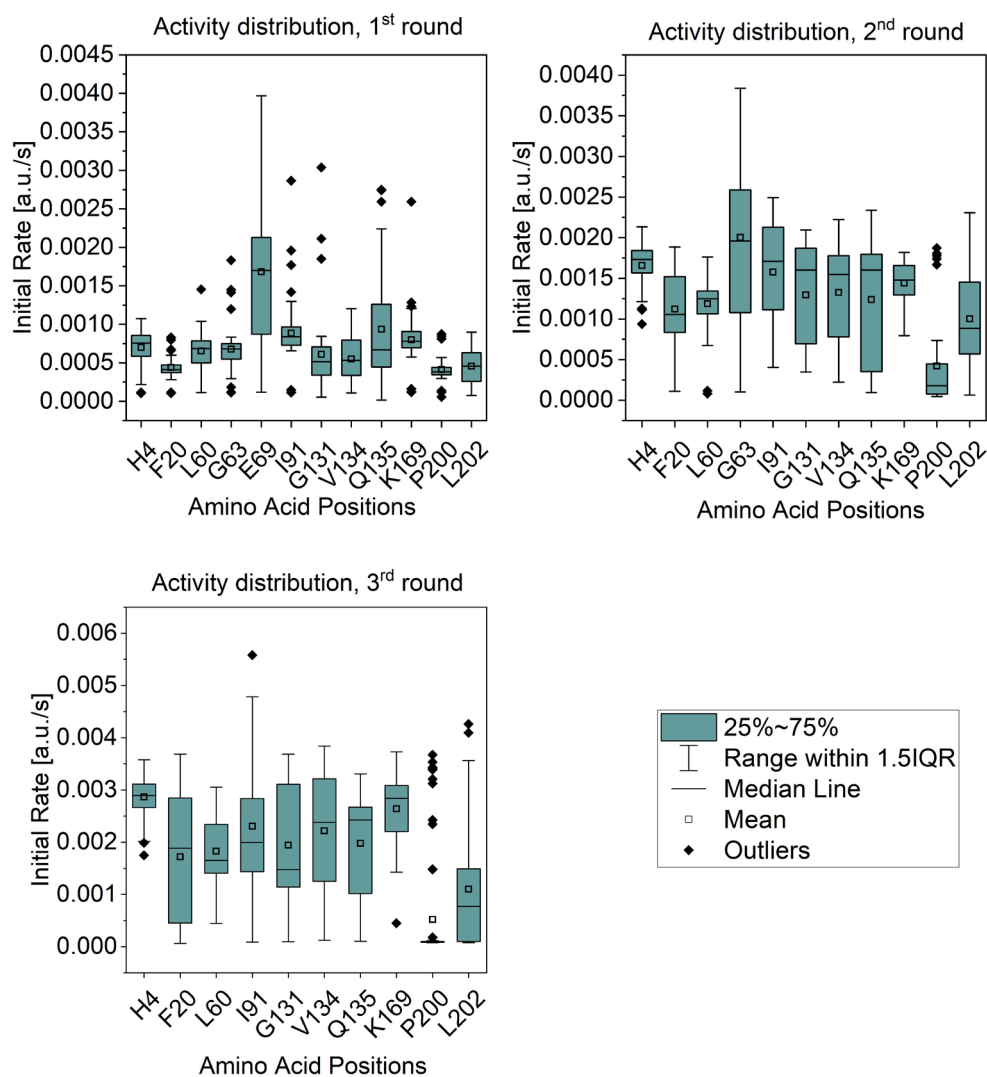


Figure S16. Distribution of the initial rate during the assays for each round of directed evolution.

The calculated initial rates during the cell lysate assays are plotted for each targeted position and displayed as the box and whisker chart. Although G63 did not appear to be a promising position during the 1st round, it emerged as the hit position in the 2nd round. The 3rd round of the screening in most of the positions did not result in a notable improvement except for position I91. IQR = interquartile range.

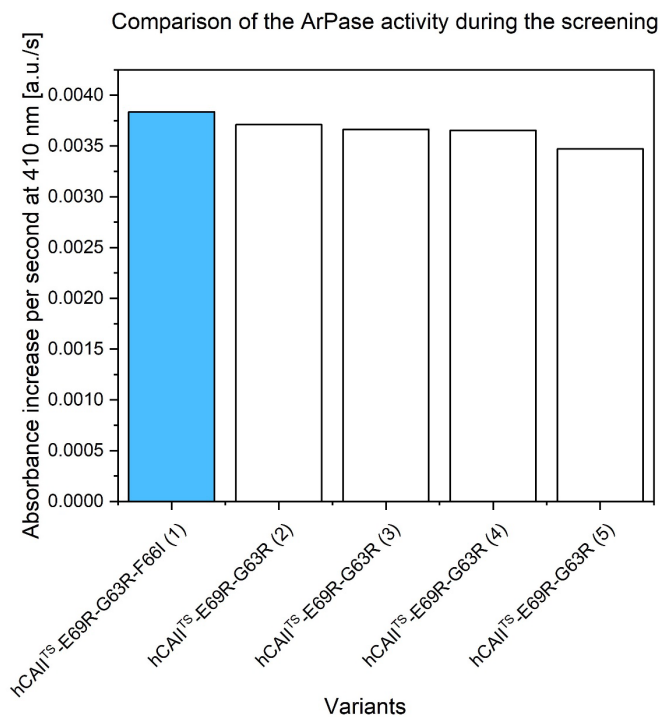


Figure S17. Comparison of the ArPase activity between the E69R-G63R and E69R-G63R-F66I mutants. The initial rates of the top five hits in the 2nd round of the screening are compared in a bar chart. Based on the sequencing results from the 2nd round of the screening, we found that the best-performing variant (colored in blue) contained the accidental mutation, F66I.

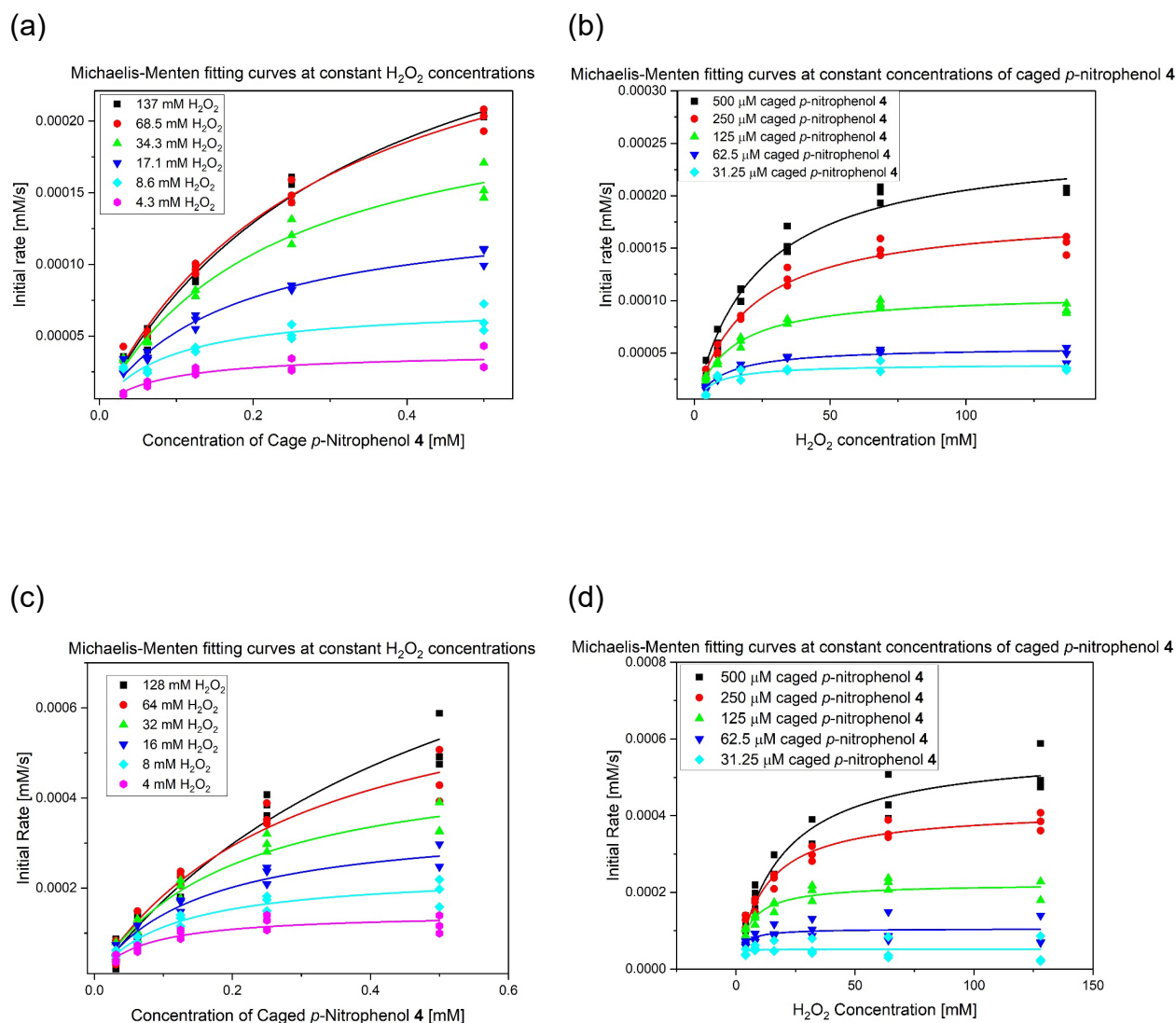


Figure S18. Michaelis Menten kinetics. Initial rates of hCAIT^S ((a),(b)) and hCAIT^S-RRIG ((c), (d)) were plotted over the concentration of the caged *p*-nitrophenol 4 or H_2O_2 . (a),(c) The plots at the constant H_2O_2 concentrations were fitted with Eq.1 (see Section 3.5.4.3 Michaelis-Menten kinetics) to obtain V^{app} values at each H_2O_2 concentration. Subsequently, the V^{app} values were plotted over the concentration of H_2O_2 and fitted with eq 2 to obtain V_{max} and K_m for H_2O_2 . The same fitting was performed for the caged *p*-nitrophenol 4 with plots (b) and (d). The poor solubility of caged *p*-nitrophenol 4 in the reaction medium impeded the experiments at concentration higher than 500 μ M.

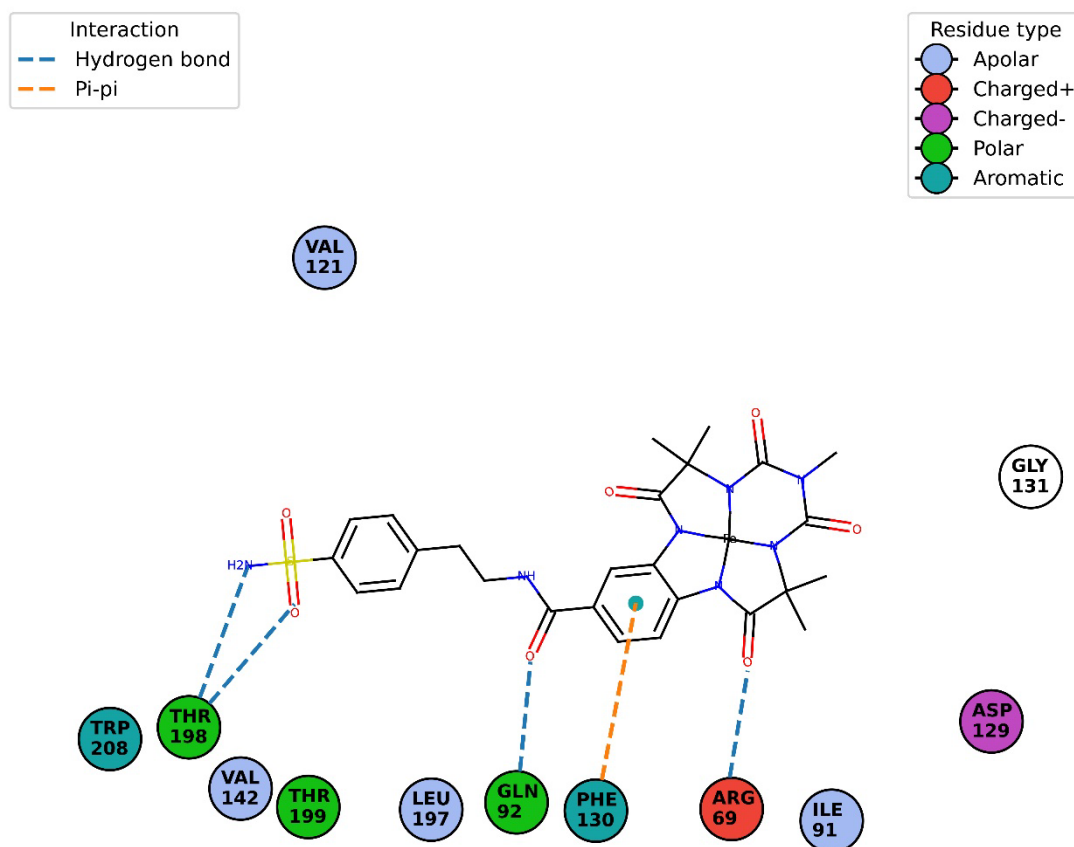


Figure S19. Schematic representation of the interactions between [BS-Fe-bTAML]⁻ and hCAII^{TS}-RRI. The figure was created with PlexView and modified.

3.5.11 Spectra

Kuerzel IM
Gruppe Ward
Nummer MOI-002
pure

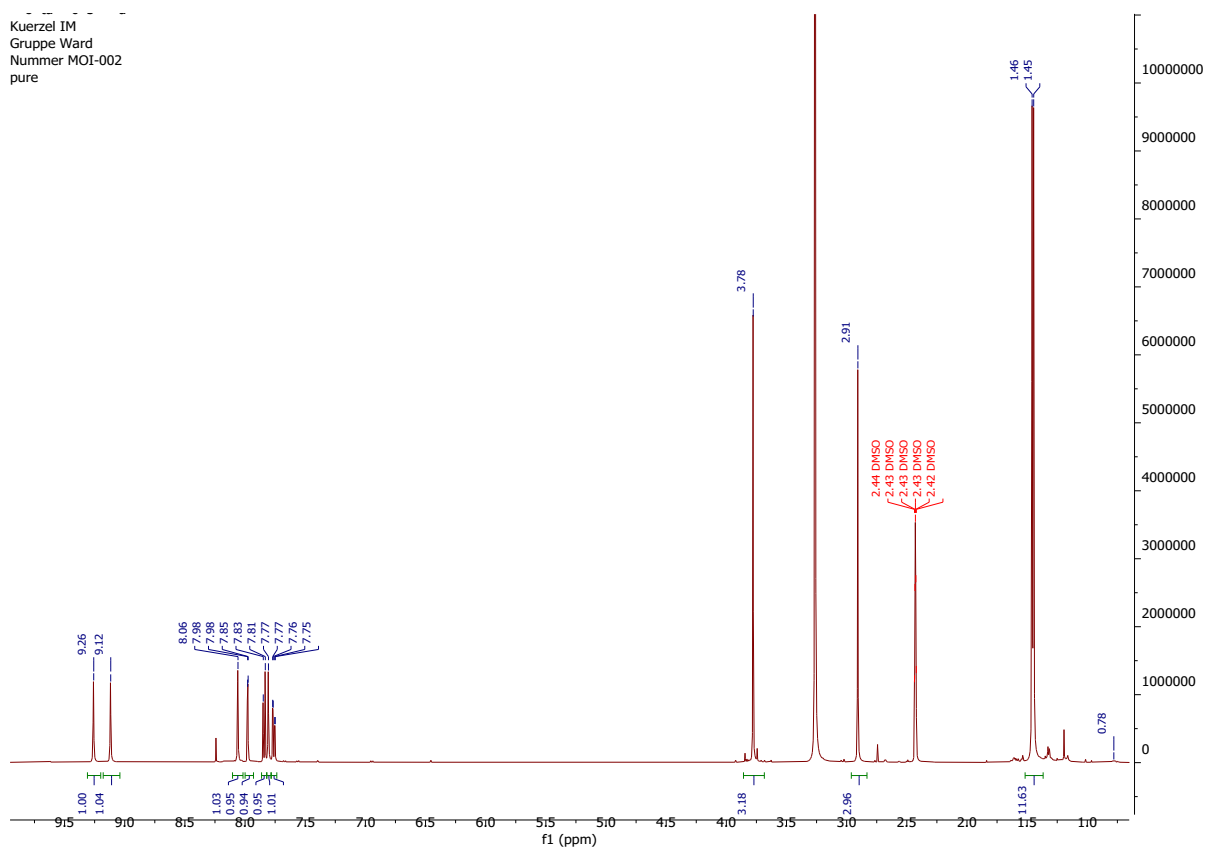


Figure S20: ^1H NMR spectrum of ligand S2.

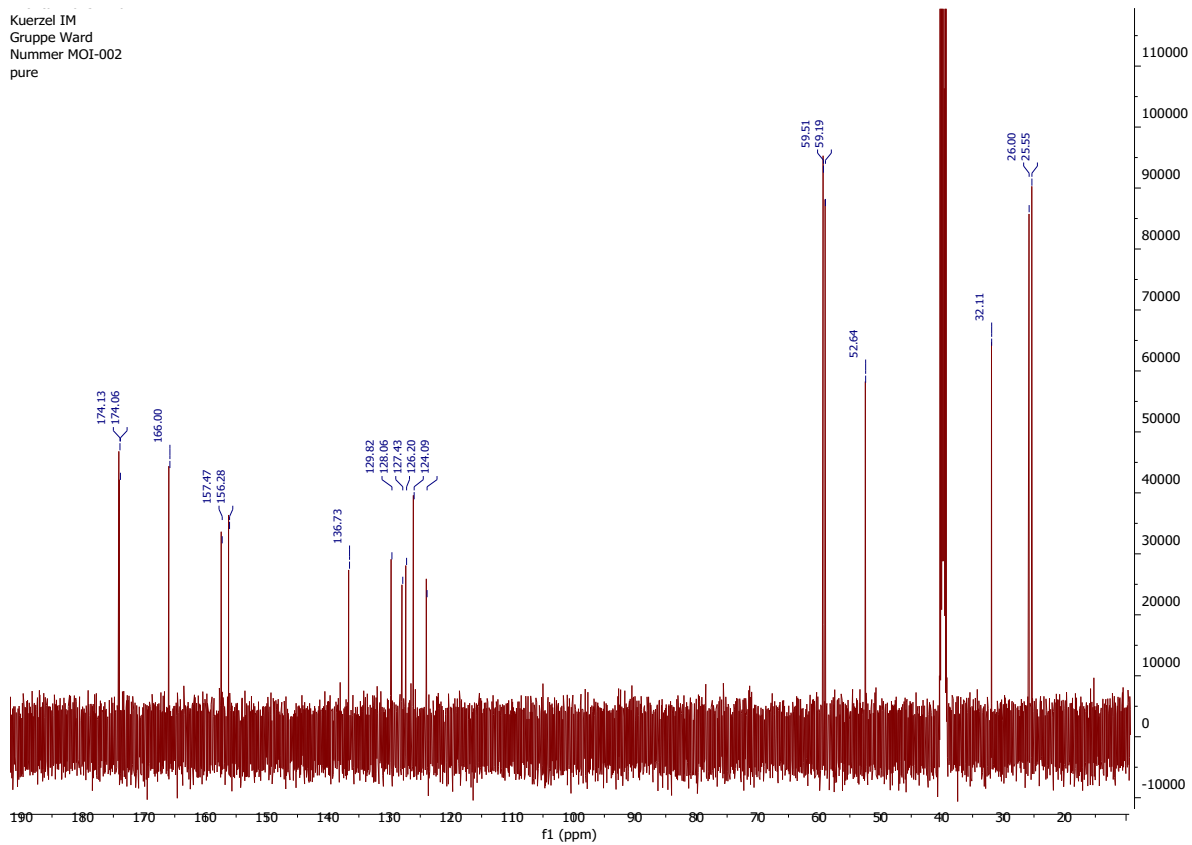


Figure S21. ^{13}C NMR spectrum of ligand S2.

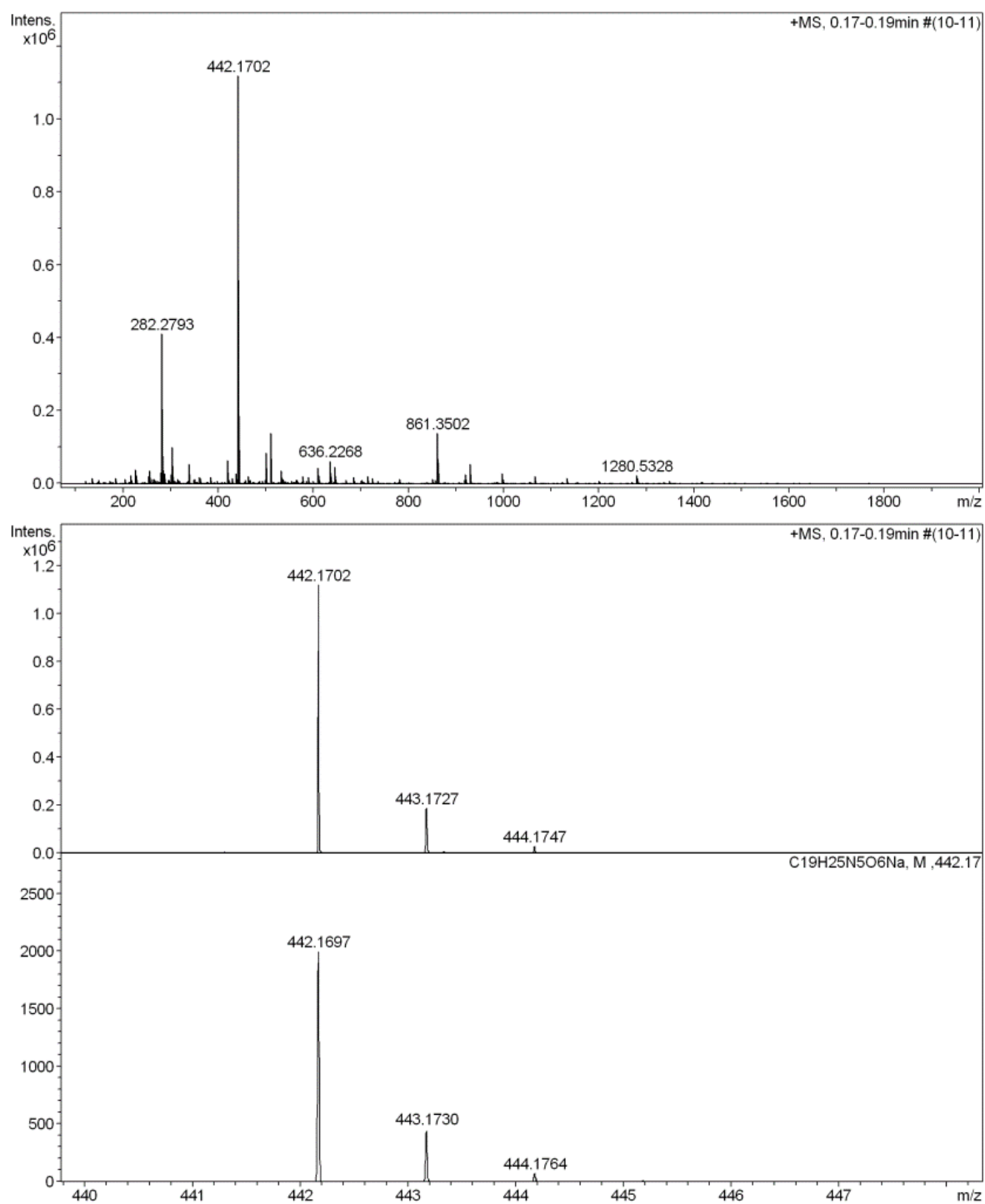


Figure S22. HRMS of ligand S2.

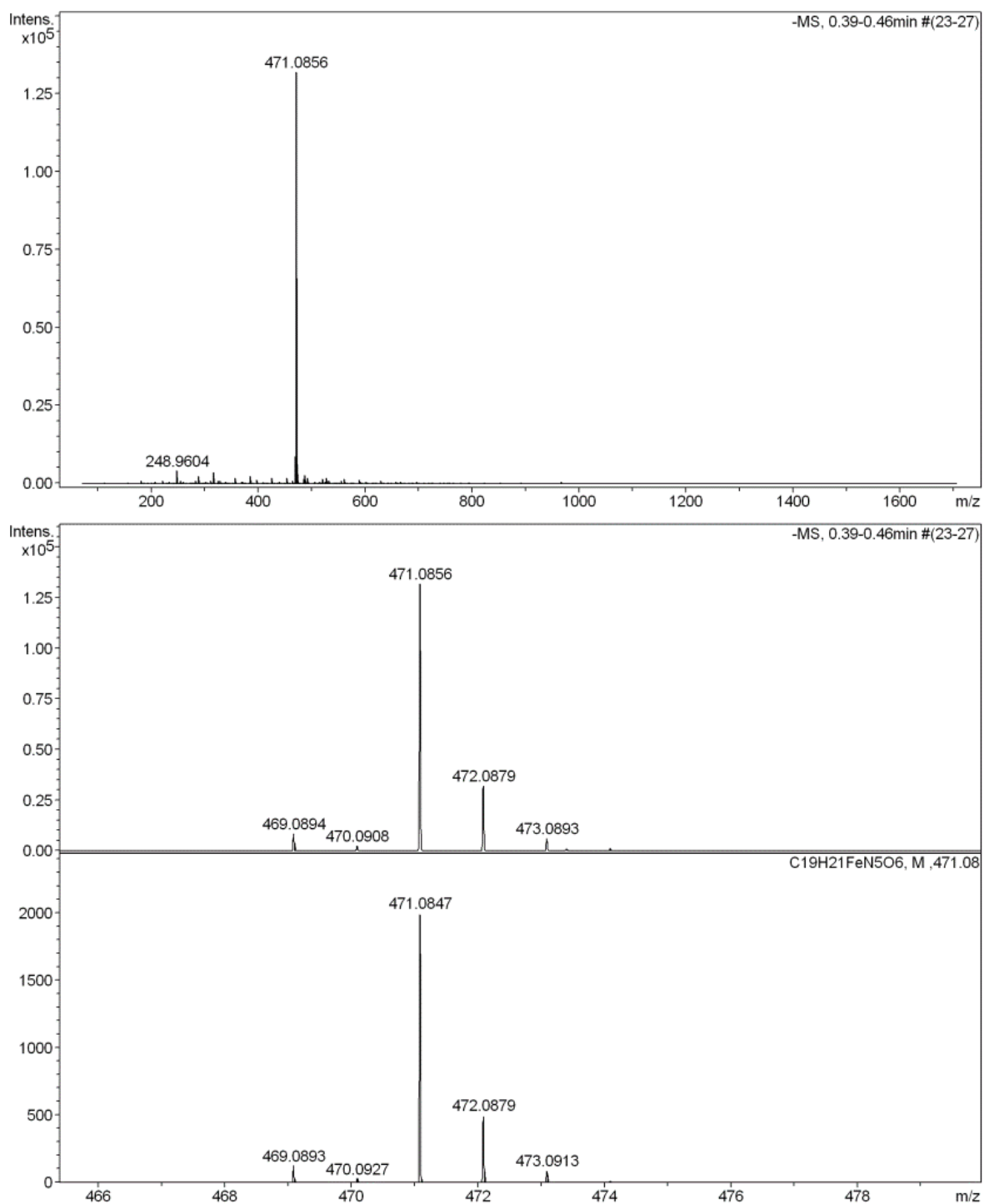


Figure S23. HRMS of methyl ester S3.

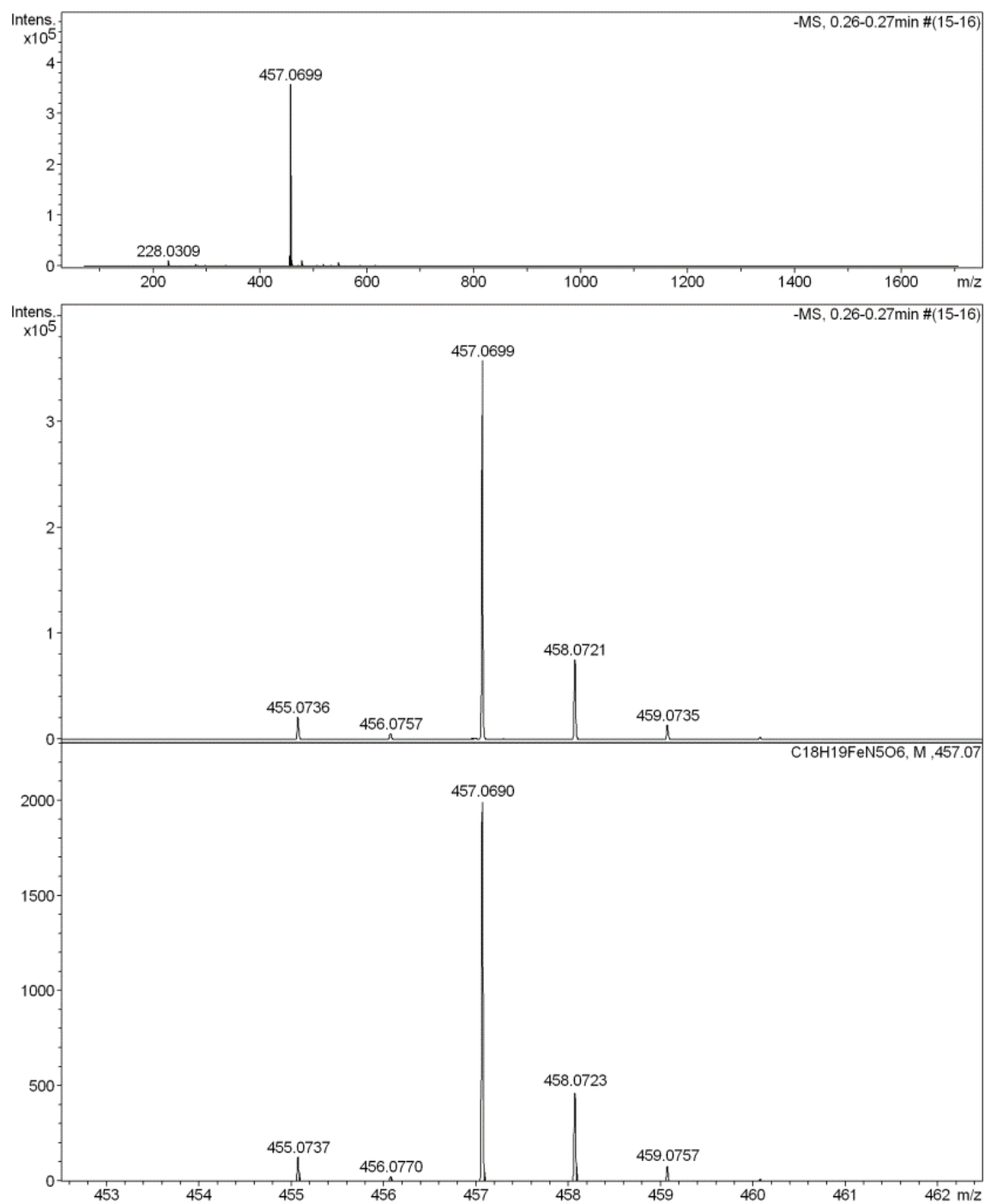


Figure S24. HRMS of carboxylate S4.

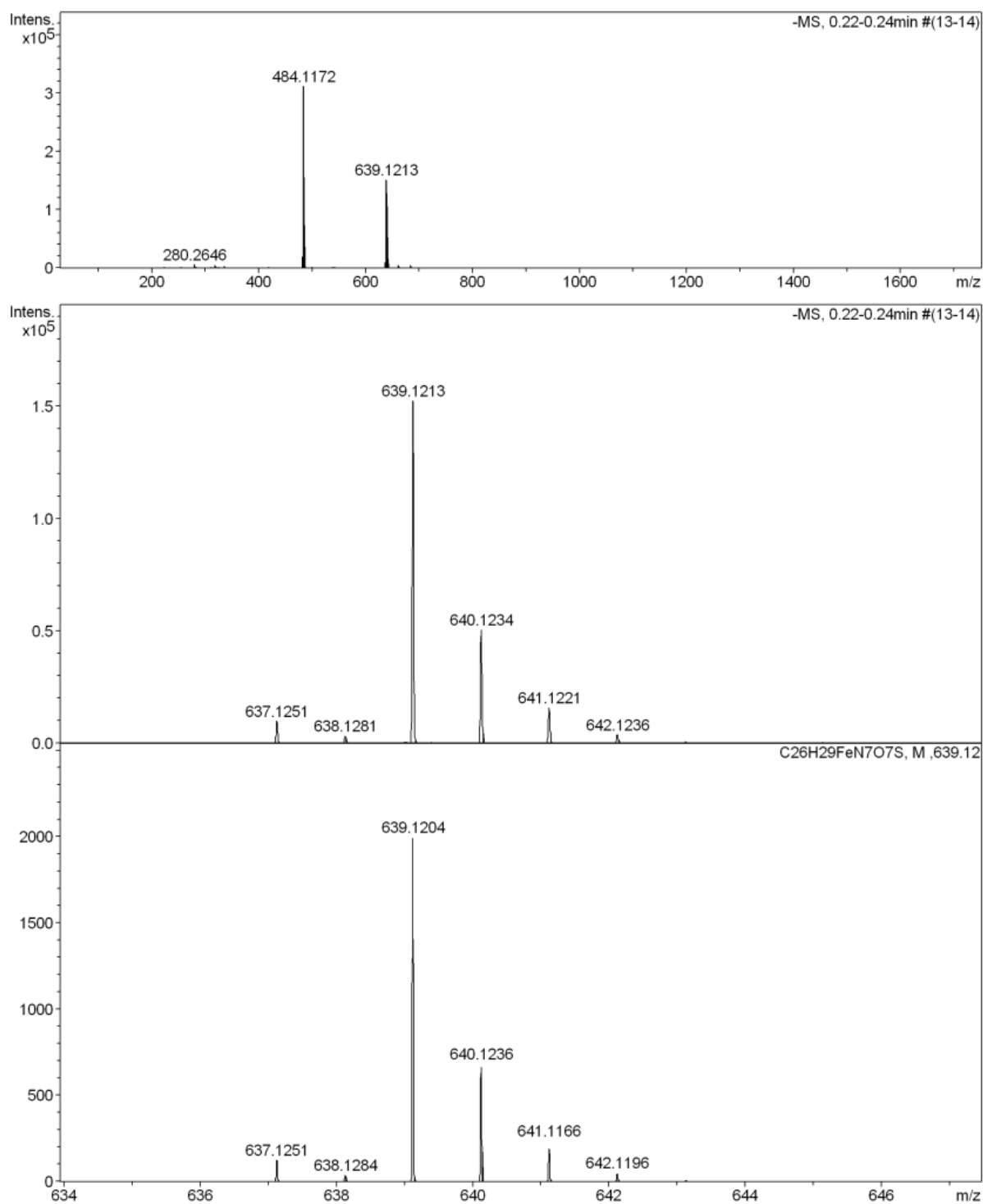


Figure S25. HRMS of [BS-Fe-bTAML]⁻.

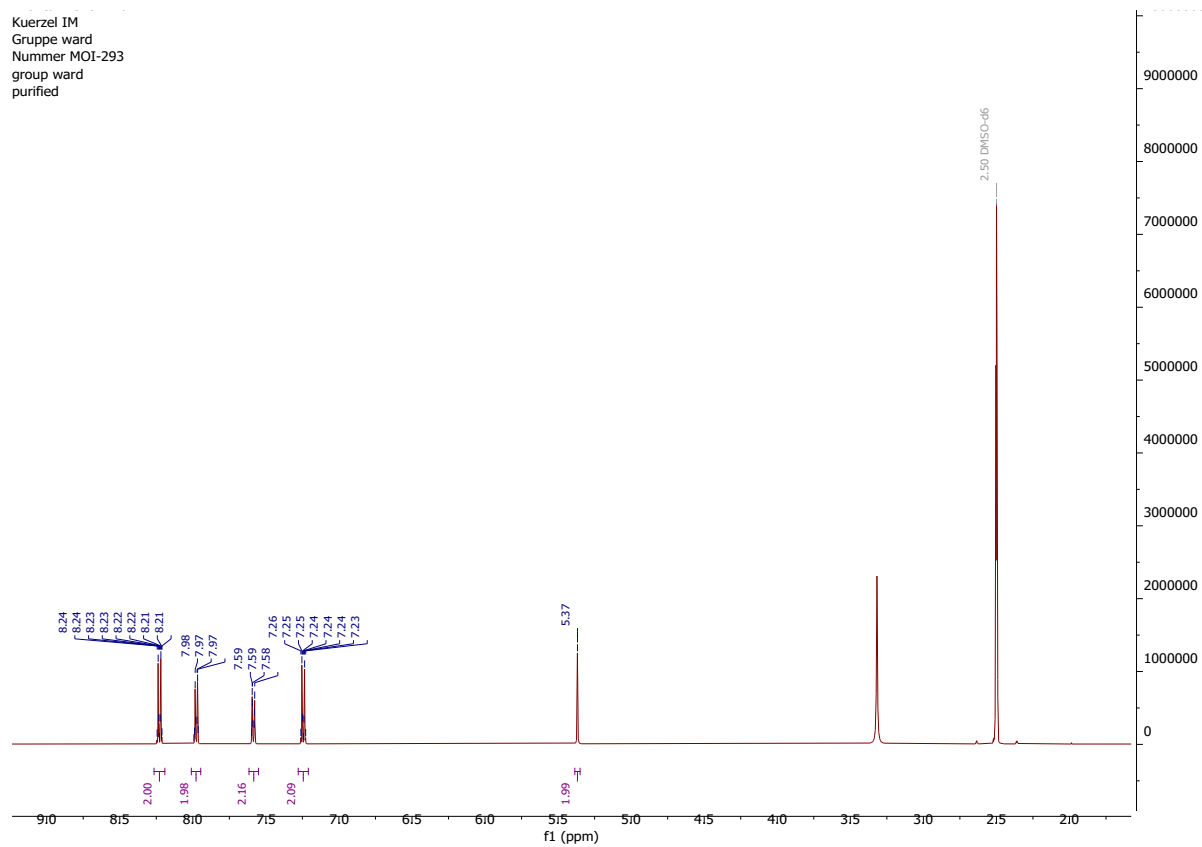


Figure S26. ¹H NMR spectrum of caged *p*-nitrophenol 4.

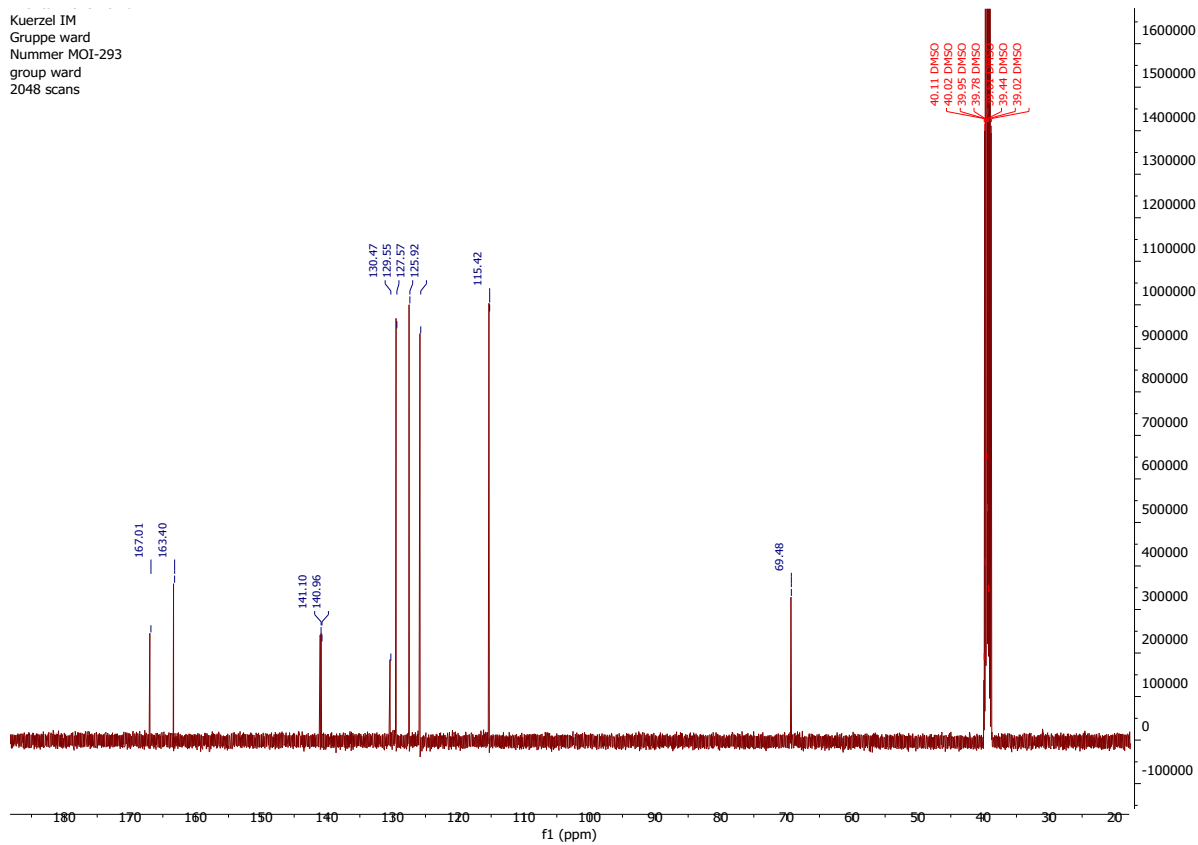


Figure S27. ^{13}C NMR spectrum of caged *p*-nitrophenol 4.

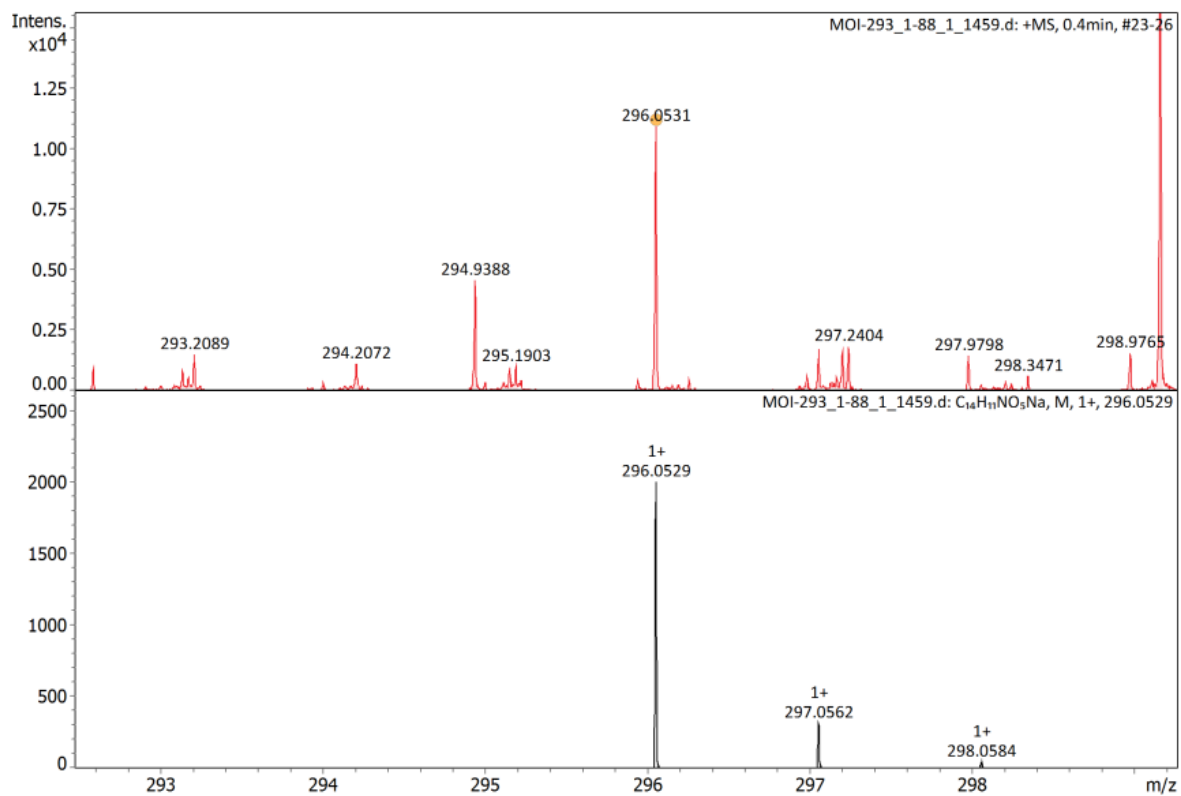


Figure S28. HRMS of caged *p*-nitrophenol 4.

3.6 References

- [2] a) M. C. White, J. Zhao, *J. Am. Chem. Soc.* **2018**, *140*, 13988-14009; b) M. Costas, *Coord. Chem. Rev.* **2011**, *255*, 2912-2932; c) X. Engelmann, I. Monte-Pérez, K. Ray, *Angew. Chem. Int. Ed.* **2016**, *55*, 7632-7649.
- [7] a) J. Chen, Z. Jiang, S. Fukuzumi, W. Nam, B. Wang, *Coord. Chem. Rev.* **2020**, *421*; b) M. Milan, M. Bietti, M. Costas, *Chem. Commun.* **2018**, *54*, 9559-9570.
- [8] T. J. Collins, A. D. Ryabov, *Chem. Rev.* **2017**, *117*, 9140-9162.
- [10] S. Kundu, J. V. Thompson, A. D. Ryabov, T. J. Collins, *J. Am. Chem. Soc.* **2011**, *133*, 18546-18549.
- [12] a) F. Schwizer, Y. Okamoto, T. Heinisch, Y. Gu, M. M. Pellizzoni, V. Lebrun, R. Reuter, V. Köhler, J. C. Lewis, T. R. Ward, *Chem. Rev.* **2018**, *118*, 142-231; b) J. Bos, G. Roelfes, *Curr. Opin. Chem. Biol.* **2014**, *19*, 135-143; c) M. T. Reetz, *Acc. Chem. Res.* **2019**, *52*, 336-344; d) B. J. Bloomer, D. S. Clark, J. F. Hartwig, *Biochemistry* **2023**, *62*, 221-228; e) C. Van Stappen, Y. Deng, Y. Liu, H. Heidari, J. X. Wang, Y. Zhou, A. P. Ledray, Y. Lu, *Chem. Rev.* **2022**, *122*, 11974-12045; f) K. Oohora, A. Onoda, T. Hayashi, *Acc. Chem. Res.* **2019**, *52*, 945-954; g) K. Vong, T. Yamamoto, K. Tanaka, *Small* **2020**, *16*, 1906890; h) C. Marchi-Delapierre, L. Rondot, C. Cavazza, S. Ménage, *Isr. J. Chem.* **2015**, *55*, 61-75; i) Y. Lu, N. Yeung, N. Sieracki, N. M. Marshall, *Nature* **2009**, *460*, 855-862; j) F. Yu, V. M. Cangelosi, M. L. Zastrow, M. Tegoni, J. S. Plegaria, A. G. Tebo, C. S. Mocny, L. Ruckthong, H. Qayyum, V. L. Pecoraro, *Chem. Rev.* **2014**, *114*, 3495-3578; k) B. Large, N. G. Baranska, R. L. Booth, K. S. Wilson, A.-K. Duhme-Klair, *Curr. Opin. Green Sustainable Chem.* **2021**, *28*, 100420; l) A. Lombardi, F. Pirro, O. Maglio, M. Chino, W. F. DeGrado, *Acc. Chem. Res.* **2019**, *52*, 1148-1159; m) Y. Yu, X. Liu, J. Wang, *Acc. Chem. Res.* **2019**, *52*, 557-565; n) W. J. Song, F. A. Tezcan, *Science* **2014**, *346*, 1525-1528; o) T. Ueno, S. Abe, N. Yokoi, Y. Watanabe, *Coord. Chem. Rev.* **2007**, *251*, 2717-2731.
- [14] Y. Wang, P. Xue, M. Cao, T. Yu, S. T. Lane, H. Zhao, *Chem. Rev.* **2021**, *121*, 12384-12444.
- [15] M. Wittwer, U. Markel, J. Schiffels, J. Okuda, D. F. Sauer, U. Schwaneberg, *Nat. Catal.* **2021**, *4*, 814-827.
- [32] J. Serrano-Plana, C. Rumo, J. G. Rebelein, R. L. Peterson, M. Barnet, T. R. Ward, *J. Am. Chem. Soc.* **2020**, *142*, 10617-10623.
- [42] S. L. Lovelock, R. Crawshaw, S. Basler, C. Levy, D. Baker, D. Hilvert, A. P. Green, *Nature* **2022**, *606*, 49-58.
- [43] a) P. Gandeepan, T. Müller, D. Zell, G. Cera, S. Warratz, L. Ackermann, *Chem. Rev.* **2019**, *119*, 2192-2452; b) B. Liu, A. M. Romine, C. Z. Rubel, K. M. Engle, B.-F. Shi, *Chem. Rev.* **2021**, *121*, 14957-15074; c) T. Ishizuka, H. Sugimoto, S. Itoh, T. Kojima, *Coord. Chem. Rev.* **2022**, *466*, 214536; d) H. Srour, P. Le Maux, S. Chevance, G. Simonneaux, *Coord. Chem. Rev.* **2013**, *257*, 3030-3050.
- [44] a) V. A. Larson, B. Battistella, K. Ray, N. Lehnert, W. Nam, *Nat. Rev. Chem.* **2020**, *4*, 404-419; b) A. S. Borovik, *Chem. Soc. Rev.* **2011**, *40*, 1870-1874; c) S. Kal, S. Xu, L. Que Jr., *Angew. Chem. Int. Ed.* **2020**, *59*, 7332-7349.
- [45] K.-B. Cho, H. Hirao, S. Shaik, W. Nam, *Chem. Soc. Rev.* **2016**, *45*, 1197-1210.
- [46] a) Y. Yang, F. H. Arnold, *Acc. Chem. Res.* **2021**, *54*, 1209-1225; b) X. Ren, R. Fasan, *Curr. Opin. Green Sustainable Chem.* **2021**, *31*, 100494; c) D. Mondal, H. M. Snodgrass, C. A. Gomez, J. C. Lewis, *Chem. Rev.* **2023**, *123*, 10381-10431.
- [47] A. Baiyoumy, J. Vallapurackal, F. Schwizer, T. Heinisch, T. Kardashliev, M. Held, S. Panke, T. R. Ward, *ACS Catal.* **2021**, *11*, 10705-10712.
- [48] a) C. Panda, M. Ghosh, T. Panda, R. Banerjee, S. Sen Gupta, *Chem. Commun.* **2011**, *47*, 8016-8018; b) M. Ghosh, K. K. Singh, C. Panda, A. Weitz, M. P. Hendrich, T. J. Collins, B. B. Dhar, S. Sen Gupta, *J. Am. Chem. Soc.* **2014**, *136*, 9524-9527.
- [49] a) V. M. Krishnamurthy, G. K. Kaufman, A. R. Urbach, I. Gitlin, K. L. Gudiksen, D. B. Weibel, G. M. Whitesides, *Chem. Rev.* **2008**, *108*, 946-1051; b) F. W. Monnard, T.

- Heinisch, E. S. Nogueira, T. Schirmer, T. R. Ward, *Chem. Commun.* **2011**, 47, 8238-8240.
- [50] A. Ghosh, F. T. de Oliveira, T. Yano, T. Nishioka, E. S. Beach, I. Kinoshita, E. Munck, A. D. Ryabov, C. P. Horwitz, T. J. Collins, *J. Am. Chem. Soc.* **2005**, 127, 2505-2513.
- [51] R. Iyer, A. A. Barrese, S. Parakh, C. N. Parker, B. C. Tripp, *J. Biomol. Screen.* **2006**, 11, 782-791.
- [52] a) T. T. Baird, A. Waheed, T. Okuyama, W. S. Sly, C. A. Fierke, *Biochemistry* **1997**, 36, 2669-2678; b) S. C. Wang, D. B. Zamble, *Biochem. Mol. Biol. Educ.* **2006**, 34, 364-368.
- [53] M. S. Mesbahuddin, A. Ganesan, S. Kalyanamoorthy, *Protein Eng. Des. Sel.* **2021**, 34.
- [54] K. M. Kean, J. J. Porter, R. A. Mehl, P. A. Karplus, *Protein Sci.* **2018**, 27, 573-577.
- [55] U. Schwaneberg, C. Schmidt-Dannert, J. Schmitt, R. D. Schmid, *Anal. Biochem.* **1999**, 269, 359-366.
- [56] U. Schwaneberg, C. Otey, P. C. Cirino, E. Farinas, F. H. Arnold, *J. Biomol. Screen.* **2001**, 6, 111-117.
- [57] F. Burg, M. Gicquel, S. Breitenlechner, A. Pöthig, T. Bach, *Angew. Chem. Int. Ed.* **2018**, 57, 2953-2957.
- [58] a) S. S. Xue, X. X. Li, Y. M. Lee, M. S. Seo, Y. Kim, S. Yanagisawa, M. Kubo, Y. K. Jeon, W. S. Kim, R. Sarangi, S. H. Kim, S. Fukuzumi, W. Nam, *J. Am. Chem. Soc.* **2020**, 142, 15305-15319; b) S. Hong, Y. M. Lee, M. Sankaralingam, A. K. Vardhaman, Y. J. Park, K. B. Cho, T. Ogura, R. Sarangi, S. Fukuzumi, W. Nam, *J. Am. Chem. Soc.* **2016**, 138, 8523-8532.
- [59] a) Q.-Y. Meng, T. E. Schirmer, A. L. Berger, K. Donabauer, B. König, *J. Am. Chem. Soc.* **2019**, 141, 11393-11397; b) T. Wang, P. M. Stein, H. Shi, C. Hu, M. Rudolph, A. S. K. Hashmi, *Nat. Commun.* **2021**, 12, 7029; c) F. H. Blindheim, M. B. Hansen, S. Evjen, W. Zhu, E. E. Jacobsen, *Catalysts* **2018**, 8, 516; d) J. A. Marko, A. Durgham, S. L. Bretz, W. Liu, *Chem. Commun.* **2019**, 55, 937-940; e) Q. M. Malik, S. Ijaz, D. C. Craig, A. C. Try, *Tetrahedron* **2011**, 67, 5798-5805; f) D. K. Winter, A. Drouin, J. Lessard, C. Spino, *J. Org. Chem.* **2010**, 75, 2610-2618; g) J. P. N. Papillon, C. M. Adams, Q.-Y. Hu, C. Lou, A. K. Singh, C. Zhang, J. Carvalho, S. Rajan, A. Amaral, M. E. Beil, F. Fu, E. Gangl, C.-W. Hu, A. Y. Jeng, D. LaSala, G. Liang, M. Logman, W. M. Maniara, D. F. Rigel, S. A. Smith, G. M. Ksander, *J. Med. Chem.* **2015**, 58, 4749-4770; h) J. Bergman, A. Brynolf, E. Vuorinen, *Tetrahedron* **1986**, 42, 3689-3696; i) J. Liu, X. Qiu, X. Huang, X. Luo, C. Zhang, J. Wei, J. Pan, Y. Liang, Y. Zhu, Q. Qin, S. Song, N. Jiao, *Nat. Chem.* **2019**, 11, 71-77.
- [60] M. Bozdog, S. Isik, S. Beyaztas, O. Arslan, C. T. Supuran, *J. Enzyme Inhib. Med. Chem.* **2015**, 30, 240-244.
- [61] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, *Protein Sci.* **1995**, 4, 2411-2423.
- [62] D. K. Srivastava, K. M. Jude, A. L. Banerjee, M. Haldar, S. Manokaran, J. Kooren, S. Mallik, D. W. Christianson, *J. Am. Chem. Soc.* **2007**, 129, 5528-5537.
- [63] X. Li-Blatter, L. Zweifel, T. Sharpe, in *Protein-Ligand Interactions*, 3 ed. (Eds.: T. Daviter, C. M. Johnson, S. H. McLaughlin, M. A. Williams), Humana New York, NY, New York, **2021**, pp. 47-79.
- [64] Z.-X. Wang, *FEBS Lett.* **1995**, 360, 111-114.
- [65] W. Kabsch, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, 66, 125-132.
- [66] P. R. Evans, G. N. Murshudov, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, 69, 1204-1214.
- [67] L. Potterton, J. Agirre, C. Ballard, K. Cowtan, E. Dodson, P. R. Evans, H. T. Jenkins, R. Keegan, E. Krissinel, K. Stevenson, A. Lebedev, S. J. McNicholas, R. A. Nicholls, M. Noble, N. S. Pannu, C. Roth, G. Sheldrick, P. Skubak, J. Turkenburg, V. Uski, F. von Delft, D. Waterman, K. Wilson, M. Winn, M. Wojdyr, *Acta Crystallogr., Sect. D: Struct. Biol.* **2018**, 74, 68-84.
- [68] J. Agirre, M. Atanasova, H. Bagdonas, C. B. Ballard, A. Basle, J. Beilsten-Edmands, R. J. Borges, D. G. Brown, J. J. Burgos-Marmol, J. M. Berrisford, P. S. Bond, I. Caballero, L. Catapano, G. Chojnowski, A. G. Cook, K. D. Cowtan, T. I. Croll, J. E. Debreczeni, N.

- E. Devenish, E. J. Dodson, T. R. Drevon, P. Emsley, G. Evans, P. R. Evans, M. Fando, J. Foadi, L. Fuentes-Montero, E. F. Garman, M. Gerstel, R. J. Gildea, K. Hatti, M. L. Hekkelman, P. Heuser, S. W. Hoh, M. A. Hough, H. T. Jenkins, E. Jimenez, R. P. Joosten, R. M. Keegan, N. Keep, E. B. Krissinel, P. Kolenko, O. Kovalevskiy, V. S. Lamzin, D. M. Lawson, A. A. Lebedev, A. G. W. Leslie, B. Lohkamp, F. Long, M. Maly, A. J. McCoy, S. J. McNicholas, A. Medina, C. Millan, J. W. Murray, G. N. Murshudov, R. A. Nicholls, M. E. M. Noble, R. Oeffner, N. S. Pannu, J. M. Parkhurst, N. Pearce, J. Pereira, A. Perrakis, H. R. Powell, R. J. Read, D. J. Rigden, W. Rochira, M. Sammito, F. Sanchez Rodriguez, G. M. Sheldrick, K. L. Shelley, F. Simkovic, A. J. Simpkin, P. Skubak, E. Sobolev, R. A. Steiner, K. Stevenson, I. Tews, J. M. H. Thomas, A. Thorn, J. T. Valls, V. Uski, I. Uson, A. Vagin, S. Velankar, M. Vollmar, H. Walden, D. Waterman, K. S. Wilson, M. D. Winn, G. Winter, M. Wojdyr, K. Yamashita, *Acta Crystallogr., Sect. D: Struct. Biol.* **2023**, *79*, 449-461.
- [69] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658-674.
- [70] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Židek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, *Nature* **2021**, *596*, 583-589.
- [71] G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 355-367.
- [72] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Cryst.* **2010**, *66*, 486-501.
- [73] N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2009**, *65*, 1074-1080.

4 Investigation of the Impact of a Grafted Motif near the Active Site

Active Site

Elias Salvisberg, Iori Morita, Kailin Zhang, Anastassia Vorobieva, Bruno Correia, and Thomas R. Ward

4.1 Introduction

As described in Chapter 3, the ArPase exhibited modest enantioselectivity. We attributed it to the exposed cofactor upon anchoring within hCAII. Therefore, we hypothesized that incorporating an additional bulky structure near the active site could enhance the peroxidase's selectivity through specific interactions, such as hydrogen bonding with either the substrate or the cofactor. We thus sought to identify a suitable chimeric scaffold with an additional insert proximal to the ArPase active site. To achieve this, we designed a library of sixteen chimeric hCAII based on computational screening. In this section, we present the screening of rationally designed chimeric hCAII^{TS}, and mutation studies conducted concurrently with the studies detailed in Chapter 3.

4.2 Results and Discussion

4.2.1 Segment screening

Within hCAII, residues 17 to 24 were selected as the site for grafting, leading to 16 chimeric structures (hCAII^{TS}-S_x, x = 1 to 16) generated through *in silico* screening (see Supplementary Information for design procedure). Among the sixteen designed constructs, we were able to clone thirteen of them into cytosolic *E. coli* expression vectors. The initial tests revealed that nine were stable during cell lysate preparation. Subsequently, seven variants (hCAII^{TS}-S_{1,2,3,4,9,12}, and 13) were obtained, following affinity purification and lyophilization, as pure white solid and used for further investigation.

These variants were evaluated for their affinity toward [BS-Fe-bTAML]⁻ using the DNSA displacement assay (Table 3, Figure S32 and Figure S33). All the variants exhibited higher K_d values than hCAII^{TS}, with apparent affinity differences mostly falling within the error range. We excluded hCAII^{TS}-S13 due to its relatively low affinity, and the other six variants were selected for further investigation.

Next, we determined the thermal stability of hCAII^{TS}-Sx with differential scanning fluorimetry (Table 3, Figure S34). Both the midpoint and onset temperature of thermal denaturation (T_m and T_{on} , respectively) of hCAII^{TS} were higher than all the hCAII^{TS}-Sx variants tested. We concluded that these variants were equally stable, and all of them were investigated in the following study.

Variant	K_d [nM]	T_m (°C)	T_{on} (°C)
hCAII ^{TS}	51.5 ± 8	71	63
hCAII ^{TS} -S1	152 ± 23	66	59
hCAII ^{TS} -S2	186 ± 34	66	58
hCAII ^{TS} -S3	182 ± 24	66	60
hCAII ^{TS} -S4	249 ± 43	66	60
hCAII ^{TS} -S9	209 ± 59	66	62
hCAII ^{TS} -S12	207 ± 32	66	61
hCAII ^{TS} -S13	396 ± 80	n.d.	n.d.

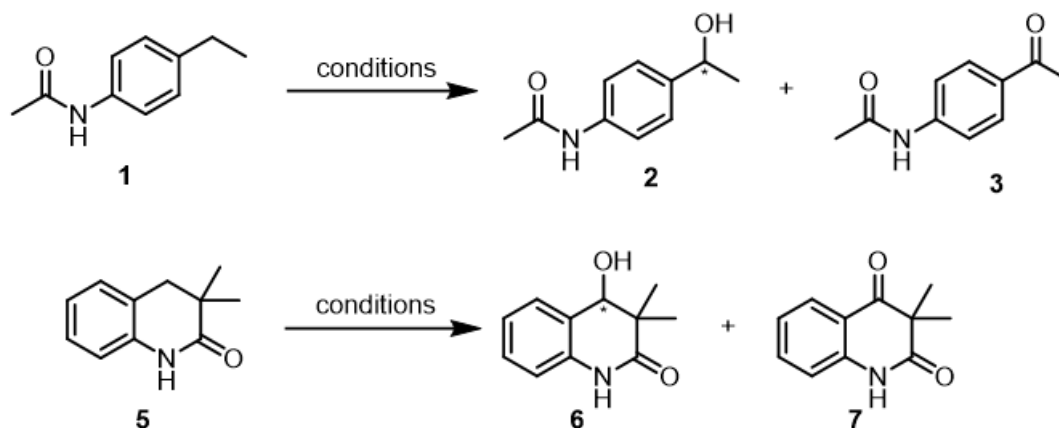
Table 3. The affinity of [BS-Fe-bTAML]⁻ to hCAII^{TS} variants and thermal stability of the hCAII^{TS} variants. K_d values were determined using the DNSA displacement assay in duplicate. T_m and T_{on} were determined with differential scanning fluorimetry in triplicate. n.d. = not determined.

4.2.2 Catalytic assays

Initially, we evaluated the activity of the ArPases for the hydroxylation of ethylbenzene derivative **1** (Scheme 1, top). All the chimeric variants exhibited higher TTON after the overnight reactions (Table 4). However, the [alcohol]/[ketone] ratios were more pronounced compared to the reactions with hCAIIT^{TS}, indicating a greater preference for overoxidation. Additionally, e.e. of the products are less than 5% across all the cases. Accordingly, we directed our focus towards lactam **5** (Scheme 1, bottom), aiming to enhance enantioselectivity of the hydroxylation reaction with the hCAIIT^{TS}-Sxs variants. The reactions were conducted for 1 hour in order to suppress overoxidation to avoid the erosion of the e.e. of the product. In accordance with the hydroxylation of ethylbenzene derivative **1**, we observed higher TTON for all the hCAIIT^{TS}-Sxs variants than hCAIIT^{TS}. However, the e.e. of the product was less than 5% in all cases.

Despite the observed low enantioselectivity, the larger [alcohol]/[ketone] ratio observed can be an advantage for the development of an enantioselective hydroxylation catalyst. Therefore, we proceeded with the engineering of an hCAIIT^{TS}-Sx using lactam **5** as substrate. Among all the chimeric hCAIIT^{TS} scaffolds, we selected hCAIIT^{TS}-S4 for further investigation based on the catalytic activity and the structure prediction confidence by ColabFold^[74] (Figure S35). However, subsequent mutant screening efforts with lactam **5** did not yield an improved variant. This outcome may be attributed to either i) poor reliability of the screening method (See Supporting Information) or ii) the negligible influence of the selected positions on catalytic activity. The former reason is apparent from high background activity observed in the negative control during the screening assay. However, the latter reason still remains reasonable considering the plausible flexibility of [BS-Fe-bTAML]⁻ and the grafted motif (Figure 2b, *vide infra*).

From here on, the residues numbers within the inserted motif are defined as 'n^S', where n ranges from 1 to 28 from the N to C termini of the inserted segment, respectively.

Scheme 1. The hydroxylation reactions performed with $[\text{BS-Fe-bTAML}]^- \cdot \text{hCAII}^{\text{TS-Sxs}}$.

Entry	hCAII variant	Ethylbenzene derivative 1 ^a		Lactam 5 ^b	
		TON	[alcohol]/[ketone]	TON	[alcohol]/[ketone]
1	TS ^c	1980	1.6	163	> 11
2	TS-S1	2916	1.2	176	11
3	TS-S2	2598	1.2	201	13
4	TS-S3	2439	1.1	238	7.5
5	TS-S4	3071	0.8	253	8.5
6	TS-S9	2663	1.3	203	6.5
7	TS-S12	2368	1.8	206	6.5
8	-	476	7	139	>9

Table 4. Reactivity of the ArPases based on hCAII^{TS-Sxs} toward hydroxylation of ethylbenzene derivative **1 and lactam **5**.** All the reactions were analyzed with SFC. e.e. < 5% for all reactions.

^aEthylbenzene **1** (10 mM), $[\text{BS-Fe-bTAML}]^-$ (4 μM), hCAII variant (10 μM), H_2O_2 (20 mM), phosphate buffered saline, MeCN (5%v/v), 25 $^\circ\text{C}$, 3 hours. ^bLactam **5** (5 mM), $[\text{BS-Fe-bTAML}]^-$ (4 μM), hCAII variant (10 μM), H_2O_2 (10 mM), phosphate buffered saline, MeCN (20%v/v), 25 $^\circ\text{C}$, 1 hour. ^c $[\text{hCAII}^{\text{TS}}] = 4 \mu\text{M}$.

4.2.3 X-ray crystallography

In order to get insight into the structure of $[\text{BS-Fe-bTAML}]^- \cdot \text{hCSII}^{\text{TS-S4}}$, we solved its X-ray crystal structure at 2.02 Å resolution (Figure 4). Four monomers are present within the unit cell. We did not observe clear electron density of the first 44 residues at the N terminus including the grafted motif, suggesting that the structure of the motif is disordered. Upon overlaying the crystal structures with the predicted structure, the inserted motifs from two monomers clearly overlap, highlighting a significant deviation of the predicted vs. X-ray structure of the insert. These findings suggest that the grafted motif exhibits considerable flexibility.

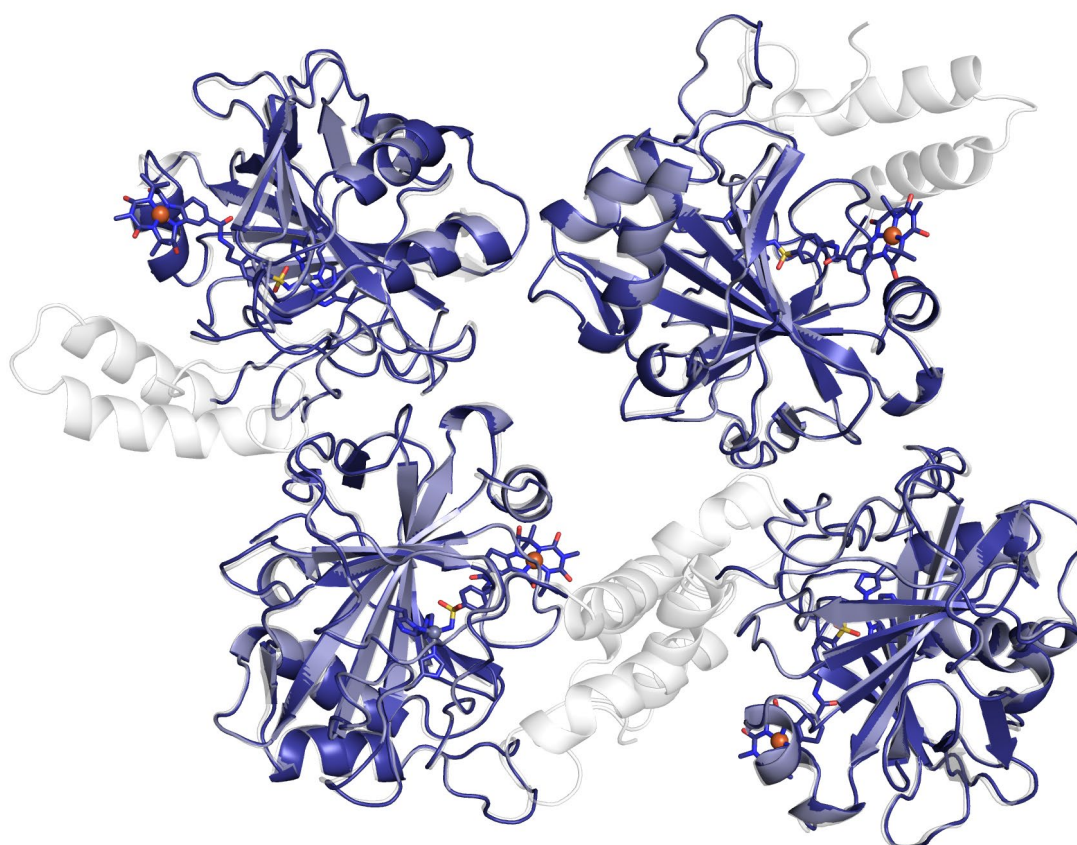


Figure 4. X-ray crystal structure of $[\text{BS-Fe-bTAML}]^- \cdot \text{hCAIITS-S4}$. The unit cell is displayed. The protein backbone is represented as cartoon. The crystal structure is displayed in dark blue, and the predicted structure by ColabFold is displayed in white at 50% transparency. The cofactor and histidine side chains coordinating with the zinc are displayed as stick (dark blue with color-coded heteroatoms; O = red, N = blue, S = yellow) with the Fe as an orange sphere. The zinc atom is represented as violet sphere.

4.2.4 Introducing disulfide bridges

Additional preliminary investigations have been conducted to enhance the rigidity of the grafted motif. We hypothesized that the introduction of a disulfide bridge between two cysteine residues within the motif may constrain its flexibility, thereby increasing the chance of establishing specific interactions with the substrate during the catalytic reaction. Based on the predicted structure of hCAII^{TS}-S4, we identified four pairs of positions for potential disulfide installation using Disulfide by Design v2.12 (Figure 5). We incorporated the E69R and G63R mutations identified in Chapter 3 along with the additional two cysteine mutations into the hCAII^{TS}-S4 scaffold. The four double cysteine mutants (ECAC = E1^{SC}-A27^{SC}, ACAC = A4^{SC}-A23^{SC}, LCAC = L7^{SC}-A20^{SC}, YCLC = Y11^{SC}-L16^{SC}) were expressed in *E. coli* BL21, and their thermal stability was evaluated using cell lysates (Figure S36). Quantitative analysis based on SDS-PAGE revealed that the introduction of cysteine residues at the most basal position, ECAC, led to better thermal stability than the distal positions (ACAC, LCAC, YCLC).

Accordingly, we purified the most stable variant, hCAII^{TS}-S4-E69R-G63R-E1^{SC}-A27^{SC} (hCAII^{TS}-S4-RR-ECAC) and evaluated its performance in the hydroxylation of lactam **5**. As a result, the TON was comparable to [BS-Fe-bTAML]⁻ and the e.e. was slightly better than hCAII^{TS}-S4 (12%, Figure S37, “No additives”). To ensure that the intramolecular disulfide bridge is formed, we conducted pre-incubation of the protein before the reaction with either glutathione disulfide (GSSG) or Aldrithiol-2 (i.e. 2,2'-dipyridyldisulfide). Notably, the e.e. improved from 12% to 30%, indicating the enhanced interaction between the protein scaffold and the substrate during the reaction. Since we observed precipitation during the reaction, we performed the reaction in the presence of a lower concentration of MeCN (5% v/v). However, this did not lead to the enhanced e.e. The moderate enhancement of e.e. in the hydroxylation of lactam **5** in the presence of the additives suggests that the designed disulfide bridge may contribute to increased enantioselectivity by rigidifying the conformation of the grafted motif. Although the previously observed effect of E69R-G63R mutations in e.e. enhancement cannot

be seen in the hCAIIT^S-S4-RR-ECAC variant, this scaffold may serve as a promising starting point for directed evolution.

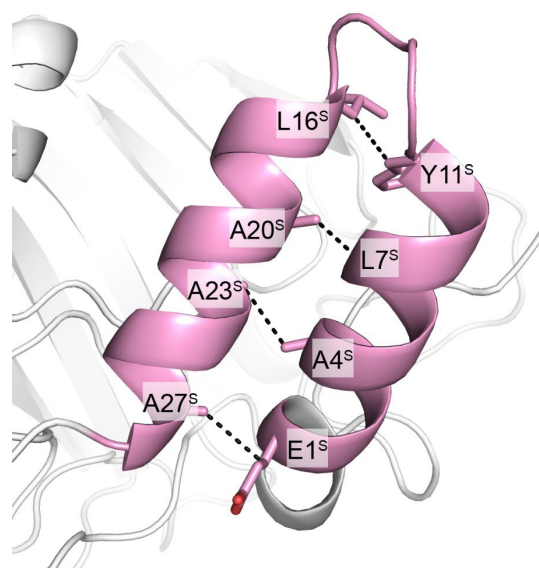


Figure 5. Selected positions for the introduction of disulfide bridges. The protein backbone is represented as a cartoon model, and the inserted motif of hCAIIT^S-S4 is highlighted in pink. The amino acid residues at the selected positions are represented as stick models (with O atoms in red) and labeled. Each pair of potentially beneficial cysteines is connected with the black dotted line.

4.3 Conclusion

In this chapter, we presented our investigation of the rationally-designed chimeric hCAIs as the scaffold for the ArPase. The binding assays along with the catalytic assays revealed that the designed proteins serve as the scaffolds for ArPases. The increased TONs indicate that the grafted motifs are beneficial for the reaction, but the low e.e. values indicate that the motifs do not have well-defined interactions with the substrate. We attributed the low selectivity to the flexibility of the grafted motif. Therefore, we surmise that rigidifying the motif with disulfide bridges may be a viable strategy for further engineering of the ArPases for improved stereoselectivity.

4.4 Supporting Information

4.4.1 General aspects

4.4.1.1 Materials and methods

All commercially available chemicals were purchased from Thermo Scientific, Sigma-Aldrich, Acros Organics, TCI Europe, Fluka, Fluorochem and used without further purification. Indole was purchased from Sigma-Aldrich (I3408).

The water used for all biological and catalytic experiments was purified with a Milli-Q Advantage system. [BS-Fe-bTAML]⁻ and the substrates were synthesized as described in section 3.5.2 Synthesis of cofactors and substrates. All enzymes used for PCRs and cloning were purchased from NEB, New England Biolabs.

GenElute™ Plasmid Miniprep Kit (Merck) was used for isolating the plasmids. Lysozyme was purchased from GoldBio (from egg white, L-040-100). DNaseI was purchased from PanReac AppliChem (A3778).

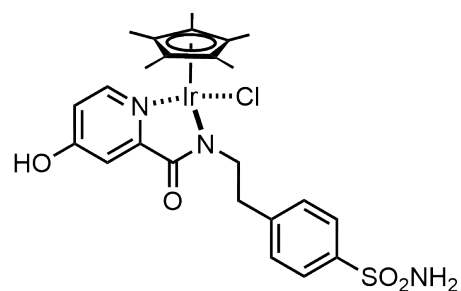
The 96-well transparent plates were purchased from Greiner (655101, Flat-bottom, polystyrene). The black 96-well plates were purchased from Thermo Scientific (Nunc™, Nunclon Delta-treated, flat-bottom, 137101).

4.4.1.2 Instrumentation

Supercritical fluid chromatography (SFC) was carried out on an Acquity UPC² system from Waters using CO₂ as the mobile phase with addition of a cosolvent. Fluorescence assays on microtiter plates were performed with The Infinite® M1000 (Tecan). HPLC analysis was performed with 1290 Infinity II LC System (Agilent). No unexpected or unusually high safety hazards were encountered. Nano-DSF measurements of thermal stability were performed using a Prometheus NT.48 (NanoTemper) instrument in nanopure water, with standard capillaries.

4.4.2 Computational design of chimeric hCAII variants

The structure design was conducted based on the crystal structure of hCAII^{WT} (PDB: 6qfu) bound to an Ir piano-stool complex (right figure). Two positions close to the Ir complex were selected for insertion of the segment: residues 17-24 and 127-132. The additional motifs were grafted using SegmentLookupWorkflow



Ir piano stool complex in PDB 6qfu

(<https://github.com/vorobieva/SegmentLookupWorkflow>). The initial attempt to graft segments between residues 127-132 yielded more than 3000 solutions, but majority of the designed motifs did not have promising contacts with the cofactor. Therefore, we selected the other position for further computation. After a series of filtration of the candidate structures, we obtained final sixteen rationally designed structures.

design number	Amino acid sequence (from N to C termini)	Length of the insertion
1	ADLAKKLENYNPEWAEATLKGAN	23
2	TELAQGEMENGNPYEFAKAMRDAAD	25
3	YDAARKLIDEYRGTEFEKMAKLALEAAN	28
4	EDLAKKLVEEYRGTSLEKYAQMAAEGAN	28
5	EDMLRSLAGTLKTYAETMLTALSHPDEDYSPDTAGELAKWM DEATKAAT	50
6	VDTMRKLAENASDSTEEKLYRLAAKAAN	28
7	KDLAQKHRSEMMEWCFTTNMESVPKDTRSEDNYPRDWLDCS SMSYINNAAT	51
8	VELAKEFTKNIDPSSIRPDDYKYAKAAT	28
9	LDLALKDSSAENWNEKVARAAN	22
10	EEMAGNMPSEYKDWAQAKKGAN	23
11	KDLATELAKQTGNKFFKEAVKAAN	24
12	KDVMKEIMNKDSSDKLAEWSAKAAD	25
13	EELLKKFVELTDKDDSWLAETAKLGAKYAT	30
14	VVLTYASSDYEPKITDETLRTATKQYRENPLPENPAN	37
15	TDLAKTMVKQFSKYVEIAEKSGLPELANRLRKLAEAMAKKVAK AAK	45
16	EKVIEKMWEEVYRYAQLYNDEYAVKLAKIAKKALEAAT	38

Table S6. Amino acid sequences of the grafted motifs. These sequences were grafted in place of residues 17 to 24.

Structure prediction of hCAII^{TS}-S1,2,3,4,9,12, and 13 was performed using ColabFold^[74].

The prediction of the potential disulfide bridge installation was performed by Disulfide by Design v2.12 (<http://cptweb.cpt.wayne.edu/DbD2/>).

4.4.3 Cloning, expression and purification of the proteins

4.4.3.1 Cloning and expression of hCAII^{TS} variants

hCAII^{WT}-Sx genes were purchased from IDT and cloned into a pET24a vector. The six mutations of hCAII^{TS} were introduced by Gibson Assembly using a PCR-amplified fragment of pACYC-hCAII^{TS} (see section 3.5.3.1 Cloning and Expression) as the insert. We obtained the desired plasmids of hCAII^{TS}-S1,2,3,4,5,6,7,9,12,13,14,15,16. The double cysteine mutations for disulfide bridges as well as E69R-G63R mutations were introduced by PCR mutagenesis using the primers listed in Table S7. A suspension of *E. coli* BL21 (DE3) chemically competent cells (50 μ L) were transformed with the plasmid of interest. After heat-shock treatment (42 °C, 45 sec.), the cells were placed on ice (3 min.). Super Optimal Broth (SOC) medium (500 μ L) was added to the tube, and the cells were incubated with a shaker (1h, 300 rpm at 37 °C). The suspension (50 μ L) was plated out on Lysogeny Broth (LB) agar plates supplemented with a suitable antibiotics (chloramphenicol (30 μ g/mL) for pACYC, kanamycin (50 μ g/mL) for pET24a). The plates were incubated (37 °C, 16 hours). For protein expression (200 mL culture), a preculture was prepared by inoculating a fresh single colony in LB medium (3 mL) in a culture tube (13 mL, Sarstedt AG), and incubated (37 °C, 300 rpm, 16 hours). The preculture (2 mL) was inoculated in the main culture (200 mL, LB medium containing the antibiotics and 1% w/v glucose) in a baffled flask (1 L) and further incubated (37 °C, 160 rpm). The rest of the overnight preculture was used to prepare a glycerol stock (25%v/v final glycerol content). The OD₆₀₀ of the culture was monitored with a UV/vis spectrometer. When the OD₆₀₀ is optimal (between 0.6 and 0.8, typically after 2-3 hours after the start of the main culture), isopropyl- β -D-thiogalactopyranoside (IPTG, 0.5 mM) was added. The culture was shaken (25 °C, 20 hours). After expression, the cell culture was transferred to a centrifuge bottle and the cells were harvested by centrifugation (4000 rpm for 15 minutes). The cell pellet was stored at -20 °C for 16 hours or more before the following application.

Name	Sequence (5' to 3')
Fw hCAII ^{TS} -S4-E17C	CACTGGTGTGATCTGGCGAAAAAACTGGTG
Rv hCAII ^{TS} -S4-E17C	CAGATCACACCAGTGCTCAGGTCGGTT
Fw hCAII ^{TS} -S4-A20C	GATCTGTGTAAAAAACTGGTGGAAGAATATCGC
Rv hCAII ^{TS} -S4-A20C	TTTTTTACACAGATCTTCCCAGTGCTCAGG
Fw-hCAII ^{TS} -S4-L23C	AAAAAATGTGTGGAAGAATATCGCGGCACC
Rv-hCAII ^{TS} -S4-L23C	TTCCACACATTTTTTCGCCAGATCTTCCCAGTG
Fw-hCAII ^{TS} -S4-Y27C-L32C	TCGCGGCACCAGCTGTGAAAAATATGCG
Rv-hCAII ^{TS} -S4-Y27C-L32C	GCTGGTGCCGCGACATTCTTCCACCA
Fw-hCAII ^{TS} -S4-A36C	TGTCAGATGGCGGCGGAAGG
Rv-hCAII ^{TS} -S4-A36C	CGCCGCCATCTGACAATATTTTTCCAGGCTG
Fw-hCAII ^{TS} -S4-A39C	CAGATGTGTGCGGAAGGCGCGAAT
Rv hCAII ^{TS} -S4-A39C	TTCCGCACACATCTGCGCATATTTTTCCAG
Fw hCAII ^{TS} -S4-A43C	GAAGGCTGTAATGGAGAGCGCCAGTCC
Rv hCAII ^{TS} -S4-A43C	TCCATTACAGCCTTCCGCCGCCATCT
Fw hCAII ^{TS} -S4-E69R	AACGTGCGCTTTGATGACTCTCAG
Rv hCAII ^{TS} -S4-E69R	ATCAAAGCGCACGTTGAAGGTATG
Fw-hCAII ^{TS} -S4-E69R-G63R	AACAATCGCCATACCTTCAACGTGCGC
Rv-hCAII ^{TS} -S4-E69R-G63R	GGTATGGCGATTGTTGAGGATTCTCAG

Table S7. Sequences of the primers for the installation of disulfide bridges and the E69R-G63R mutations to hCAII^{TS}-S4. Fw = Forward primer, Rv = Reverse primer.

4.4.3.2 Gene sequence of the pET24a plasmids harboring the hCAII^{TS}-Sxs

>pET24a-hCAII^{TS}-S1 (6105 bp, the hCAII^{TS}-S1 gene is highlighted in bold)

TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGGTGTGGTGGTTAC
GCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTGCTTTCTTC
CCTTCCTTTCTCGCCACGTTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCC
TTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTG
ATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAG
TCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCG
GTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTTAAAAAATGAGC
TGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATTCAGGTGG

CACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAAT
ATGTATCCGCTCATGAATTAATTCTTAGAAAACTCATCGAGCATCAAATGAACTGCAA
TTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGA
GAAACTCACCGAGGCAGTTCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCC
GACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAG
TGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAGTTTATGCATTT
CTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAATCACTCGCATCAA
CCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTA
AAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCAT
CAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGG
GGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTC
GGAAGAGGCATAAATTCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATTG
GCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCGGGCTTCCCATACAA
TCGATAGATTGTGCGACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAA
ATCAGCATCCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATAT
GGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTTCATGACC
AAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAA
AGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACC
ACCGCTACCAGCGGTGGTTTGTGGCCGGATCAAGAGCTACCAACTTTTTTCCGAAGG
TAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTA
GGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTT
ACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGA
TAGTTACCGGATAAGGCGCAGCGGTCCGGGCTGAACGGGGGGTTCGTGCACACAGCCC
AGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAA
GCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCCG
GAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCC
TGTCGGGTTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGG
CGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTG

GCCTTTTGCTCACATGTTCTTTCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTAC
CGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTC
AGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGC
GGTATTTACACCGCATATATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGCATAGT
TAAGCCAGTATACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACAC
CCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACA
GACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTACCGTCATCACC
GAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGGTCGTGAAGCGATTCACAG
ATGTCTGCCTGTTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTG
GCTTCTGATAAAGCGGGCCATGTAAAGGGCGGTTTTTCTGTTTGGTCACTGATGCCT
CCGTGTAAGGGGGATTTCTGTTTCATGGGGTAATGATACCGATGAAACGAGAGAGGAT
GCTCACGATACGGGTTACTGATGATGAACATGCCCGTTACTGGAACGTTGTGAGGGTA
AACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAATGCCA
GCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATG
CAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAAC
ACGGAAACCGAAGACCATTTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGT
CGCTTCACGTTTCGCTCGCGTATCGGTGATTCATTCTGCTAACCAGTAAGGCAACCCCGC
CAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCATGCGCACCCGTGGGGCCGCC
ATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGTGACGA
AGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGT
CGCGCTCCAGCGAAAGCGGTCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTG
TCCTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCC
GCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTGAGATC
CCGGTGCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTC
CAGTCGGGAAACCTGTTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAG
GCGGTTTTCGTATTGGGCGCCAGGGTGGTTTTTCTTTTACCAGTGAGACGGGCAACA
GCTGATTGCCCTTACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGT
TTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACGGCGGGATATAACATGAGC

TGTCTTCGGTATCGTTCGTATCCCCTACTACCGAGATATCCGCACCAACGCGCAGCCCGGA
CTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCA
GTGGGAACGATGCCCTCATTGAGCATTGTCATGGTTTGTGAAAACCGGACATGGCACT
CCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGATATTTATGCC
AGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGA
TTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCAGTCGCGTACCGTCTTCATG
GGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAAATAACGCCGGAA
CATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATG
ATCAGCCCCTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGA
CGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAGA
TTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAAC
GCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAAT
TCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTTCGCAGAAACGTGGCTGGCC
TGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGT
ATAACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATG
CCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTT
ATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCC
GCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGG
CCTGCCACCATAACCACGCCGAAACAAGCGCTCATGAGCCCAGAGTGGCGAGCCCGAT
CTTCCCCATCGGTGATGTGCGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGT
GATGCCGGCCACGATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAA
TACGACTCACTATAGGGGAATTGTGAGCGGATAACAATCCCCTCTAGAAATAATTTTGT
TTAACTTTAAGAAGGAGATATACATATGGCCCATCACTGGGGGTACGGCAAACACAAC
GGACCTGAGCACTGGGCGGATCTGGCGAAAAAACTGGAAAATTATAATCCGGAATGG
GCCGAAGCGACCCTGAAAGGCGCGAATGGAGAGCGCCAGTCCCCTGTTGACATCGA
CACTCATAACAGCCAAGTATGACCCTTCCCTGAAGCCCCTGTCTGTTTCCTATGATCAA
GCAACTTCCCTGAGAATCCTCAACAATGGTCATACCTTCAACGTGGAGTTTGATGACT
CTCAGGACAAAGCAGTGCTCAAGGGAGGACCCCTGGATGGCACTTACAGATTGATTC

AGTTTCACTTTCACTGGGGTTCACATGATGGACAAGGTTTCAGAGCATACTGTGGATAA
AAAGAAATATGCTGCAGAACTTCACTTGGTTCCTGGAACACCAAATATGGGGATTTT
GGGAAAGCTGTGCAGCAACCTGATGGACTGGCCGTTCTAGGTATTTTTTTGAAGGTTG
GCAGCGCTCTGCCGGGCCTTCAGAAAGTTGTTGATGTGCTGGATTCCATTAACAACAA
GGCAAGAGTGCTGACTTCACTAACTTCGATCCTCGTGGCCTCCTCCTGAATCCCTG
GATTACTGGACCTACCCAGGCTCACTGACCACCCCTCCTCTTCTGGAATCAGTGACCT
GGATTGTGCTCAAGGAACCCATCAGCGTCAGCAGCGAGCAGGTGAGCAAATTCCGTA
AACTTAACTTCAATGGGGAGGGTGAACCCGAAGAACCGATGGTGGACAACCTGGCGCC
CAACCCAGCCACTGAAGAACAGGCAAATCAAAGCTTCCTTCAAATAATAGGGATCCG
AATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCT
GAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGA
GCAATAACTAGCATAACCCCTTGGGGCCTCTAACGGGTCTTGAGGGGTTTTTTGCTGA
AAGGAGGAACTATATCCGGAT

> pET24a-hCAI^{TS}-S2 (834 bp, only the hCAI gene)

ATGGCCCATCACTGGGGGTACGGCAAACACAACGGACCTGAGCACTGGACCGAACT
GGCGCAGGGCGAAATGGAAAATGGCAATCCGTATGAATTTGCGAAAGCGATGCGCG
ACGCGGCGGATGGAGAGCGCCAGTCCCCTGTTGACATCGACACTCATACAGCCAAGT
ATGACCCTTCCCTGAAGCCCCTGTCTGTTTCCTATGATCAAGCAACTTCCCTGAGAATC
CTCAACAATGGTCATACCTTCAACGTGGAGTTTGTGACTCTCAGGACAAAGCAGTGC
TCAAGGGAGGACCCCTGGATGGCACTTACAGATTGATTCAGTTTCACTTTCACTGGGG
TTCatGATGGACAAGGTTTCAGAGCATACTGTGGATAAAAAGAAATATGCTGCAGAA
CTTCACTTGGTTCCTGGAACACCAAATATGGGGATTTTGGGAAAGCTGTGCAGCAAC
CTGATGGACTGGCCGTTCTAGGTATTTTTTTGAAGGTTGGCAGCGCTCTGCCGGGCCT
TCAGAAAGTTGTTGATGTGCTGGATTCCATTAACAACAAAGGGCAAGAGTGCTGACTTC
ACTAACTTCGATCCTCGTGGCCTCCTCCTGAATCCCTGGATTACTGGACCTACCCAG
GCTCACTGACCACCCCTCCTCTTCTGGAATCAGTGACCTGGATTGTGCTCAAGGAACC
CATCAGCGTCAGCAGCGAGCAGGTGAGCAAATTCCGTAACTTAACTTCAATGGGGA

**GGGTGAACCCGAAGAACCGATGGTGGACAACCTGGCGCCCAACCCAGCCACTGAAGA
ACAGGCAAATCAAAGCTTCCTTCAAATAA**

> pET24a-hCAII^{TS}-S3 (843 bp, only the hCAII gene)

**ATGGCCCATCACTGGGGGTACGGCAAACACAACGGACCTGAGCACTGGTATGATGCG
GCGCGCAAACCTGATTGATGAATATCGCGGCACCGAATTTGAAAAATGGCGAAACTG
GCGCTGGAAGCGGCGAATGGAGAGCGCCAGTCCCCTGTTGACATCGACACTCATA
GCCAAGTATGACCCTTCCTGAAGCCCCTGTCTGTTTCCTATGATCAAGCAACTTCCT
GAGAATCCTCAACAATGGTCATACCTTCAACGTGGAGTTTGATGACTCTCAGGACAAA
GCAGTGCTCAAGGGAGGACCCCTGGATGGCACTTACAGATTGATTCAGTTTCACTTTC
ACTGGGGTTCAcatGATGGACAAGGTTTCAGAGCATACTGTGGATAAAAAGAAATATGC
TGCAGAACTTCACTTGGTTCACTGGAACACCAAATATGGGGATTTTGGGAAAGCTGTG
CAGCAACCTGATGGACTGGCCGTTCTAGGTATTTTTTTGAAGGTTGGCAGCGCTCTGC
CGGGCCTTCAGAAAGTTGTTGATGTGCTGGATTCCATTAACAAGGGCAAGAGTG
CTGACTTCACTAACTTCGATCCTCGTGGCCTCCTTCCTGAATCCCTGGATTACTGGACC
TACCAGGCTCACTGACCACCCCTCCTCTTCTGGAATCAGTGACCTGGATTGTGCTCA
AGGAACCCATCAGCGTCAGCAGCGAGCAGGTGAGCAAATTCGGTAACTTAACTTCA
ATGGGGAGGGTGAACCCGAAGAACCGATGGTGGACAACCTGGCGCCCAACCCAGCCA
CTGAAGAACAGGCAAATCAAAGCTTCCTTCAAATAA**

> pET24a-hCAII^{TS}-S4 (843 bp, only the hCAII gene)

**ATGGCCCATCACTGGGGGTACGGCAAACACAACGGACCTGAGCACTGGGAAGATCT
GGCGAAAAAATGGTGAAGAATATCGCGGCACCGCCTGGAAAAATATGCGCAGAT
GGCGGCGGAAGGCGCGAATGGAGAGCGCCAGTCCCCTGTTGACATCGACACTCATA
CAGCCAAGTATGACCCTTCCTGAAGCCCCTGTCTGTTTCCTATGATCAAGCAACTTC
CCTGAGAATCCTCAACAATGGTCATACCTTCAACGTGGAGTTTGATGACTCTCAGGAC
AAAGCAGTGCTCAAGGGAGGACCCCTGGATGGCACTTACAGATTGATTCAGTTTCACT
TTCCTGGGGTTCAcatGATGGACAAGGTTTCAGAGCATACTGTGGATAAAAAGAAATAT
GCTGCAGAACTTCACTTGGTTCACTGGAACACCAAATATGGGGATTTTGGGAAAGCTG**

TGCAGCAACCTGATGGACTGGCCGTTCTAGGTATTTTTTTGAAGGTTGGCAGCGCTCT
GCCGGGCCTTCAGAAAGTTGTTGATGTGCTGGATTCCATTAACAAGGGCAAGAG
TGCTGACTTCACTAACTTCGATCCTCGTGGCCTCCTTCTGAATCCCTGGATTACTGGA
CCTACCCAGGCTCACTGACCACCCCTCCTTCTGGAATCAGTGACCTGGATTGTGCT
CAAGGAACCCATCAGCGTCAGCAGCGAGCAGGTGAGCAAATTCCGTAACTTACTT
CAATGGGGAGGGTGAACCCGAAGAACCGATGGTGGACAACCTGGCGCCCAACCCAGC
CACTGAAGAACAGGCAAATCAAAGCTTCCTTCAAATAA

> pET24a-hCAII^{TS}-S5 (909 bp, only the hCAII gene)

ATGGCCCATCACTGGGGGTACGGCAAACACAACGGACCTGAGCACTGGGAAGATAT
GTTACGCAGCCTGGCGGGCACCCTGAAAACCTATGCGGAAACCATGCTGACCGCGCT
GAGCAAGCATCCGGATGAAGATTATAGCCCGGATACCGCGGGCGAACTGGCGAAAT
GGATGGACGAAGCGACCAAAGCGGCGACCGGAGAGCGCCAGTCCCCTGTTGACATC
GACTCATAACAGCCAAGTATGACCCTTCCCTGAAGCCCCTGTCTGTTTCCTATGATC
AAGCAACTTCCCTGAGAATCCTCAACAATGGTCATACCTTCAACGTGGAGTTTGATGA
CTCTCAGGACAAAGCAGTGCTCAAGGGAGGACCCCTGGATGGCACTTACAGATTGAT
TCAGTTTCACTTTCACTGGGGTTCAcatGATGGACAAGGTTTCAGAGCATACTGTGGATA
AAAAGAAATATGCTGCAGAACTTCACTTGGTTCACTGGAACACCAAATATGGGGATTT
TGGGAAAGCTGTGCAGCAACCTGATGGACTGGCCGTTCTAGGTATTTTTTTGAAGGTT
GGCAGCGCTCTGCCGGGCCTTCAGAAAGTTGTTGATGTGCTGGATTCCATTAACA
AGGGCAAGAGTGCTGACTTCACTAACTTCGATCCTCGTGGCCTCCTTCTGAATCCCT
GGATTACTGGACCTACCCAGGCTCACTGACCACCCCTCCTTCTGGAATCAGTGACC
TGGATTGTGCTCAAGGAACCCATCAGCGTCAGCAGCGAGCAGGTGAGCAAATTCCGT
AACTTAACTTCAATGGGGAGGGTGAACCCGAAGAACCGATGGTGGACAACCTGGCGC
CCAACCCAGCCACTGAAGAACAGGCAAATCAAAGCTTCCTTCAAATAA

> pET24a-hCAII^{TS}-S6 (843 bp, only the hCAII gene)

ATGGCCCATCACTGGGGGTACGGCAAACACAACGGACCTGAGCACTGGGTGGATAC
CATGCGCAAACCTGGCGGAAAATGCGAGCGATAGCACCGAAGAAAACTGTATCGCCT

GGCGGCGAAAGCGGCCAATGGAGAGCGCCAGTCCCCTGTTGACATCGACACTCATAC
AGCCAAGTATGACCCTTCCCTGAAGCCCCTGTCTGTTTCCTATGATCAAGCAACTTCC
CTGAGAATCCTCAACAATGGTCATACCTTCAACGTGGAGTTTGATGACTCTCAGGACA
AAGCAGTGCTCAAGGGAGGACCCCTGGATGGCACTTACAGATTGATTCAGTTTCACTT
TCACTGGGGTTCAcatGATGGACAAGGTTTCAGAGCATACTGTGGATAAAAAGAAATAT
GCTGCAGAACTTCACTTGGTTCACTGGAACACCAAATATGGGGATTTTGGGAAAGCTG
TGCAGCAACCTGATGGACTGGCCGTTCTAGGTATTTTTTTGAAGGTTGGCAGCGCTCT
GCCGGGCCTTCAGAAAGTTGTTGATGTGCTGGATTCCATTAACAAGGGCAAGAG
TGCTGACTTCACTAACTTCGATCCTCGTGGCCTCCTTCTGAATCCCTGGATTACTGGA
CCTACCCAGGCTCACTGACCACCCCTCCTCTTCTGGAATCAGTGACCTGGATTGTGCT
CAAGGAACCCATCAGCGTCAGCAGCGAGCAGGTGAGCAAATTCCGTAACTTAACTT
CAATGGGGAGGGTGAACCCGAAGAACCGATGGTGGACAACCTGGCGCCCAACCCAGC
CACTGAAGAACAGGCAAATCAAAGCTTCTTCAAATAA

> pET24a-hCAI^{TS}-S7 (912 bp, only the hCAI gene)

ATGGCCCATCACTGGGGGTACGGCAAACACAACGGACCTGAGCACTGGAAAGATCTG
GCGCAGAAACATCGCAGCGAAATGATGGAATGGTGCTTTACCACCAATATGGAAAGC
GTGCCGAAAGATACCCGCAGTGAAGATAATTATCCGCGCGATTGGCTGGATTGCAGC
AGCATGAGCTATATTAATAATGCGGCGACCGGAGAGCGCCAGTCCCCTGTTGACATC
GACTCATAACAGCCAAGTATGACCCTTCCCTGAAGCCCCTGTCTGTTTCCTATGATC
AAGCAACTTCCCTGAGAATCCTCAACAATGGTCATACCTTCAACGTGGAGTTTGATGA
CTCTCAGGACAAAGCAGTGCTCAAGGGAGGACCCCTGGATGGCACTTACAGATTGAT
TCAGTTTCACTTTCACTGGGGTTCAcatGATGGACAAGGTTTCAGAGCATACTGTGGATA
AAAAGAAATATGCTGCAGAACTTCACTTGGTTCACTGGAACACCAAATATGGGGATTT
TGGGAAAGCTGTGCAGCAACCTGATGGACTGGCCGTTCTAGGTATTTTTTTGAAGTT
GGCAGCGCTCTGCCGGGCCTTCAGAAAGTTGTTGATGTGCTGGATTCCATTAACA
AGGGCAAGAGTGCTGACTTCACTAACTTCGATCCTCGTGGCCTCCTTCTGAATCCCT
GGATTACTGGACCTACCCAGGCTCACTGACCACCCCTCCTCTTCTGGAATCAGTGACC
TGGATTGTGCTCAAGGAACCCATCAGCGTCAGCAGCGAGCAGGTGAGCAAATTCCGT

**AAACTTAACTTCAATGGGGAGGGTGAACCCGAAGAACCGATGGTGGACAACCTGGCGC
CCAACCCAGCCACTGAAGAACAGGCAAATCAAAGCTTCCTTCAAATAA**

> pET24a-hCAI^{TS}-S9 (825 bp, only the hCAI gene)

**ATGGCCCATCACTGGGGGTACGGCAAACACAACGGACCTGAGCACTGGCTGGATCTG
GCGCTGAAAGATAGCAGCGCGGAAAATTGGAATGAAAAAGTTGCGCGCGCGGGCGAA
TGGAGAGCGCCAGTCCCCTGTTGACATCGACACTCATACAGCCAAGTATGACCCTTCC
CTGAAGCCCCTGTCTGTTTCCTATGATCAAGCAACTTCCCTGAGAATCCTCAACAATG
GTCATACCTTCAACGTGGAGTTTGATGACTCTCAGGACAAAGCAGTGCTCAAGGGAG
GACCCCTGGATGGCACTTACAGATTGATTCAGTTTCACTTTCACTGGGGTTCAcatGAT
GGACAAGGTTTCAGAGCATACTGTGGATAAAAAGAAATATGCTGCAGAACTTCACTTG
GTTCACTGGAACACCAAATATGGGGATTTTGGGAAAGCTGTGCAGCAACCTGATGGA
CTGGCCGTTCTAGGTATTTTTTTGAAGGTTGGCAGCGCTCTGCCGGGCCTTCAGAAAG
TTGTTGATGTGCTGGATTCCATTAACAAGGGCAAGAGTGCTGACTTCACTAACTT
CGATCCTCGTGGCCTCCTTCTGAATCCCTGGATTACTGGACCTACCCAGGCTCACTG
ACCACCCCTCCTTCTGGAATCAGTGACCTGGATTGTGCTCAAGGAACCCATCAGCG
TCAGCAGCGAGCAGGTGAGCAAATTCCGTAACTTAACTTCAATGGGGAGGGTGAAC
CCGAAGAACCGATGGTGGACAACCTGGCGCCCAACCCAGCCACTGAAGAACAGGCAA
ATCAAAGCTTCCTTCAAATAA**

> pET24a-hCAI^{TS}-S12 (834 bp, only the hCAI gene)

**ATGGCCCATCACTGGGGGTACGGCAAACACAACGGACCTGAGCACTGGAAAGATGT
GATGAAAGAAATTATGAATAAAGATAGCAGCGATAAACTGGCGGAATGGAGCGCGAA
AGCGGCGGATGGAGAGCGCCAGTCCCCTGTTGACATCGACACTCATACAGCCAAGTA
TGACCCTTCCCTGAAGCCCCTGTCTGTTTCCTATGATCAAGCAACTTCCCTGAGAATCC
TCAACAATGGTCATACCTTCAACGTGGAGTTTGATGACTCTCAGGACAAAGCAGTGCT
CAAGGGAGGACCCCTGGATGGCACTTACAGATTGATTCAGTTTCACTTTCACTGGGGT
TCacatGATGGACAAGGTTTCAGAGCATACTGTGGATAAAAAGAAATATGCTGCAGAAC
TTCACCTTGGTTCACTGGAACACCAAATATGGGGATTTTGGGAAAGCTGTGCAGCAACC**

TGATGGACTGGCCGTTCTAGGTATTTTTTTGAAGGTTGGCAGCGCTCTGCCGGGCCTT
CAGAAAGTTGTTGATGTGCTGGATTCCATTA AAACAAAGGGCAAGAGTGCTGACTTCA
CTAACTTCGATCCTCGTGGCCTCCTTCCTGAATCCCTGGATTACTGGACCTACCCAGG
CTCACTGACCACCCCTCCTCTTCTGGAATCAGTGACCTGGATTGTGCTCAAGGAACCC
ATCAGCGTCAGCAGCGAGCAGGTGAGCAAATTCCGTAAACTTAACTTCAATGGGGAG
GGTGAACCCGAAGAACCGATGGTGGACAACCTGGCGCCCAACCCAGCCACTGAAGAA
CAGGCAAATCAAAGCTTCCTTCAAATAA

> pET24a-hCAII^{TS}-S13 (849 bp, only the hCAII gene)

ATGGCCCATCACTGGGGGTACGGCAAACACAACGGACCTGAGCACTGGGAAGAACT
GCTGAAAAAATTTGTGGAAGTACCGATAAAGATGATAGCTGGCTGGCGGAAACCGC
GAAACTGGGCGCGAAATATGCGACCGGAGAGCGCCAGTCCCCTGTTGACATCGACAC
TCATACAGCCAAGTATGACCCTTCCTGAAGCCCCTGTCTGTTTCCTATGATCAAGCA
ACTTCCTGAGAATCCTCAACAATGGTCATACCTTCAACGTGGAGTTTGATGACTCTC
AGGACAAAGCAGTGCTCAAGGGAGGACCCCTGGATGGCACTTACAGATTGATTCAGT
TTCACITTTCACTGGGGTTCA^{cat}GATGGACAAGGTTTCAGAGCATACTGTGGATAAAAAG
AAATATGCTGCAGAACTTCACTTGGTTCACTGGAACACCAAATATGGGGATTTTGGGA
AAGCTGTGCAGCAACCTGATGGACTGGCCGTTCTAGGTATTTTTTTGAAGGTTGGCAG
CGCTCTGCCGGGCCTTCAGAAAGTTGTTGATGTGCTGGATTCCATTA AAACAAAGGGC
AAGAGTGCTGACTTCACTAACTTCGATCCTCGTGGCCTCCTTCCTGAATCCCTGGATTA
CTGGACCTACCCAGGCTCACTGACCACCCCTCCTCTTCTGGAATCAGTGACCTGGATT
GTGCTCAAGGAACCCATCAGCGTCAGCAGCGAGCAGGTGAGCAAATTCCGTAAACTT
AACTTCAATGGGGAGGGTGAACCCGAAGAACCGATGGTGGACAACCTGGCGCCCAAC
CCAGCCACTGAAGAACAGGCAAATCAAAGCTTCCTTCAAATAA

> pET24a-hCAII^{TS}-S14 (870 bp, only the hCAII gene)

ATGGCCCATCACTGGGGGTACGGCAAACACAACGGACCTGAGCACTGGGTGGTGCT
GACCTATGCGAGCAGCGATTATGAACCGAAAATTACCGATGAAACCCTGCGCACCGC
GACCAAACAGTATCGCGAAAATCCGCTGCCGGAAAATCCGGCGAATGGAGAGCGCC

AGTCCCCTGTTGACATCGACTCATACAGCCAAGTATGACCCTTCCCTGAAGCCCCT
GTCTGTTTCCTATGATCAAGCAACTTCCCTGAGAATCCTCAACAATGGTCATACCTTCA
ACGTGGAGTTTGATGACTCTCAGGACAAAGCAGTGCTCAAGGGAGGACCCCTGGATG
GCACTTACAGATTGATTCAGTTTCACTTTCACTGGGGTTCAcatGATGGACAAGGTTCA
GAGCATACTGTGGATAAAAAGAAATATGCTGCAGAACTTCACTTGGTTCACTGGAACA
CCAAATATGGGGATTTTGGGAAAGCTGTGCAGCAACCTGATGGACTGGCCGTTCTAG
GTATTTTTTTGAAGGTTGGCAGCGCTCTGCCGGGCCTTCAGAAAGTTGTTGATGTGCT
GGATTCCATTAACAAAGGGCAAGAGTGCTGACTTCACTAACTTCGATCCTCGTGGC
CTCCTTCTGAATCCCTGGATTACTGGACCTACCCAGGCTCACTGACCACCCCTCCTC
TTCTGGAATCAGTGACCTGGATTGTGCTCAAGGAACCCATCAGCGTCAGCAGCGAGC
AGGTGAGCAAATTCCGTAACTTAACTTCAATGGGGAGGGTGAACCCGAAGAACCGA
TGGTGGACAACTGGCGCCCAACCCAGCCACTGAAGAACAGGCAAATCAAAGCTTCTCT
TCAAATAA

> pET24a-hCAII^{TS}-S15 (894 bp, only the hCAII gene)

ATGGCCCATCACTGGGGGTACGGCAAACACAACGGACCTGAGCACTGGACCGATCTG
GCGAAAACCATGGTGAAACAGTTTAGCAAATATGTGGAAATTGCGGAAAAAAGCGGC
CTGCCGGAACCTGGCCAATCGCCTGCGCAAATTGGCGGAAATGGCCAAGAAAGTGGC
CAAAGCCGCCAAAGGAGAGCGCCAGTCCCCTGTTGACATCGACTCATACAGCCAA
GTATGACCCTTCCCTGAAGCCCCTGTCTGTTTCCTATGATCAAGCAACTTCCCTGAGA
ATCCTCAACAATGGTCATACCTTCAACGTGGAGTTTGATGACTCTCAGGACAAAGCAG
TGCTCAAGGGAGGACCCCTGGATGGCACTTACAGATTGATTCAGTTTCACTTTCACTG
GGGTTCAcatGATGGACAAGGTTTCAAGCATACTGTGGATAAAAAGAAATATGCTGCA
GAACTTCACTTGGTTCACTGGAACACCAAATATGGGGATTTTGGGAAAGCTGTGCAGC
AACCTGATGGACTGGCCGTTCTAGGTATTTTTTTGAAGGTTGGCAGCGCTCTGCCGGG
CCTTCAGAAAGTTGTTGATGTGCTGGATTCCATTAACAAAGGGCAAGAGTGCTGAC
TTCATAACTTCGATCCTCGTGGCCTCCTTCTGAATCCCTGGATTACTGGACCTACCC
AGGCTCACTGACCACCCCTCCTCTTCTGGAATCAGTGACCTGGATTGTGCTCAAGGAA
CCCATCAGCGTCAGCAGCGAGCAGGTGAGCAAATTCCGTAACTTAACTTCAATGGG

**GAGGGTGAACCCGAAGAACCGATGGTGGACAACTGGCGCCCAACCCAGCCACTGAA
GAACAGGCAAATCAAAGCTTCCTTCAAATAA**

> pET24a-hCAII^{TS}-S16 (873 bp, only the hCAII gene)

**ATGGCCCATCACTGGGGGTACGGCAAACACAACGGACCTGAGCACTGGGAAAAAGT
GATTGAAAAAATGTGGGAAGAAGTGTATCGCTATGCGCAGCTGTATAATGATGAATAT
GCGGTGAAACTGGCGAAAATTGCGAAAAAAGCGCTGGAAGCGGCGACCGGAGAGCG
CCAGTCCCCTGTTGACATCGACACTCATAACAGCCAAGTATGACCCTTCCCTGAAGCCC
CTGTCTGTTTCCTATGATCAAGCAACTTCCCTGAGAATCCTCAACAATGGTCATACCTT
CAACGTGGAGTTTGTGACTCTCAGGACAAAGCAGTGCTCAAGGGAGGACCCCTGGA
TGGCACTTACAGATTGATTCAGTTTCACTTTCACTGGGGTTCAcatGATGGACAAGGTT
CAGAGCATACTGTGGATAAAAAGAAATATGCTGCAGAACTTCACTTGGTTCACTGGAA
CACCAAATATGGGGATTTTGGGAAAGCTGTGCAGCAACCTGATGGACTGGCCGTTCT
AGGTATTTTTTTGAAGGTTGGCAGCGCTCTGCCGGGCCTTCAGAAAGTTGTTGATGTG
CTGGATTCCATTAACAAAGGGCAAGAGTGCTGACTTCACTAACTTCGATCCTCGTG
GCCTCCTTCTGAATCCCTGGATTACTGGACCTACCCAGGCTCACTGACCACCCCTCC
TCTTCTGGAATCAGTGACCTGGATTGTGCTCAAGGAACCCATCAGCGTCAGCAGCGA
GCAGGTGAGCAAATTCCGTAACTTAACTTCAATGGGGAGGGTGAACCCGAAGAACC
GATGGTGGACAACTGGCGCCCAACCCAGCCACTGAAGAACAGGCAAATCAAAGCTTC
CTTCAAATAA**

4.4.3.3 Purification

The frozen pellet of *E. coli* cells expressing the protein was thawed at room temperature. The pellet the expression culture (200 mL culture volume) was resuspended in a lysis buffer (20 mL, lysozyme (0.1 mg/mL), DNaseI (0.01 mg/mL) in the binding buffer: ZnSO₄ (0.5 mM), tris-sulfate (50 mM, pH = 8.6)). The suspension was shaken (37 °C, 300 rpm, 1 hour). The lysed suspension was sonicated (Fisherbrand™ Model 705 Sonic Dismembrator with a Microtip, amplitude 50, 10 on and off alternately). The sonicated suspension was centrifuged to remove

cell debris (4 °C, 25000g, 30 minutes). The supernatant was collected. A small amount (10 µL) of the supernatant was put aside for SDS-PAGE analysis.

Pre-packed *p*-aminomethylbenzenesulfonamide-agarose beads or synthesized arylsulfonamide-decorated Sepharose beads^[60] in an open column (30 mL bed volume) were equilibrated with the binding buffer (3 CV). The clear supernatant of the lysate was filtered through a syringe filter (Filtropur S, PES, pore size: 0.2 µm, Sarstedt), and loaded onto the resin. A small amount of the flow-through fraction was put aside for SDS-PAGE analysis. The column was washed with the binding buffer (2 CV, dithiothreitol (1 mM) was supplemented for the double cysteine mutants) followed by the wash buffer (2 CV, Na₂SO₄ (50 mM), NaClO₄ (50 mM), tris-sulfate (25 mM, pH = 8.8), dithiothreitol (1 mM, for the double cysteine mutants)). The bound protein was eluted with the elution buffer (3 CV, NaClO₄ (200 mM), sodium acetate (100 mM), dithiothreitol (1 mM, for the double cysteine mutants), pH adjusted to 5.6 with dilute sulfuric acid). The purity of the protein was confirmed by SDS-PAGE. Only the pure fractions were combined in a dialysis bag (Spectra/Por® 1 Dialysis membrane, MWCO 6-8 kD) and dialyzed against MilliQ water (4 °C). The water was changed several times so that the final salt concentration was below 10 nM. The dialyzed protein solution was transferred to falcon tubes and centrifuged (4000 g for 30 minutes, 4 °C) to remove any precipitate. The supernatant was flash-frozen with liquid nitrogen and lyophilized. The protein was obtained as white cotton-like solid.

4.4.3.4 Cell lysate preparation

The frozen pellet of *E. coli* cells expressing the protein was thawed at room temperature. The pellet from the expression culture (50 mL culture volume) was resuspended in a lysis buffer (15 mL, lysozyme (0.1 mg/mL), DNaseI (0.01 mg/mL) in borate buffer (pH = 8.5, 10 mM)) so that the OD₆₀₀ of the suspension is around 10. The suspension was shaken (37 °C, 250 rpm, 30 minutes). The lysate was then frozen (-20 °C, 2 hours). The frozen lysate was then thawed at room temperature and aliquoted into Eppendorf tubes (1.5 mL). The tubes were shaken for

heat shock (55 °C, 600 rpm, 30 minutes). The precipitate was spun down by centrifugation (21100 g, 20 minutes, 4 °C). The clear supernatant was aliquoted into new Eppendorf tubes (0.5 mL). The lysate was flash-frozen with liquid nitrogen and stored at –80 °C. The lysate was used within 1 month after the preparation.

4.4.4 Catalytic assays

4.4.4.1 General procedure

For the reactions with the purified proteins, the stock solution of the proteins was prepared by dissolving the lyophilized protein in MilliQ water or the reaction buffer. The concentration of the stock solution was determined by the absorbance (280 nm, NanoDrop1000 Spectrophotometer, Thermo Fischer Scientific) and the calculated extinction coefficient of the protein (Table S8). For the reactions with the cell lysate, the frozen lysate was thawed at room temperature.

Protein	Extinction coefficient at 280 nm [a.u. · mol ⁻¹ · cm ⁻¹]
hCAII ^{TS}	50420
hCAII ^{TS} -S1	57410
hCAII ^{TS} -S2	51910
hCAII ^{TS} -S3	53400
hCAII ^{TS} -S4	53400
hCAII ^{TS} -S9	55920
hCAII ^{TS} -S12	55920
hCAII ^{TS} -S13	57410
hCAII ^{TS} -S4-RR-ECAC	53650

Table S8. Extinction coefficients of the purified proteins at 280 nm. The values were calculated based on a reported equation^[61].

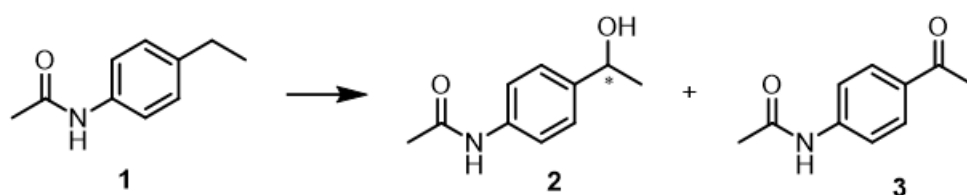
For the reactions with disulfide additives (GSH or Aldrithiol-2), the protein was mixed with an equimolar amount of the additive in MilliQ water and incubated (25°C, 30 min). Precipitate formed during the incubation was removed by centrifugation.

The reactions (200 µL) were typically performed in an HPLC glass vial (1.5 mL). MilliQ water, buffer, ZnSO₄, the protein, [BS-Fe-bTAML]⁻, and the substrate (in water, MeCN or DMSO) was added to a suitable reaction vessel. The reaction was initiated by addition of H₂O₂ (2 equiv. vs. substrate) to the mixture. The HPLC vials were shaken with Eppendorf ThermoMixer® C (750 rpm, 25 °C).

The reaction was quenched by adding HCl aq (1 M, 15 μ L), and the internal standard (indole, 4 μ mol for 200 μ L a reaction mixture) was added. The mixture was extracted with EtOAc (400 μ L x 2). The organic phases were combined, dried over MgSO₄ and subjected to SFC for quantification.

4.4.4.2 Analysis methods and calibration curves

4.4.4.2.1 Ethylbenzene derivative 1



Analytical method: Chiralpak IH[®], 4.6 × 250 mm; gradient method from iPrOH 5% for 3 min, 5 to 20% over the next 10 min, 20% for the next 5 min, 2 mL/min. Injection volume = 10 μ L. For the calibration curve, indole (5 mM) and the product standards (from 0.3125 mM to 10 mM) were dissolved in EtOAc. The peak areas of the compounds were integrated at specific wavelengths: indole = 265 nm, ketone **3** = 284 nm, alcohol **2** = 245 nm.

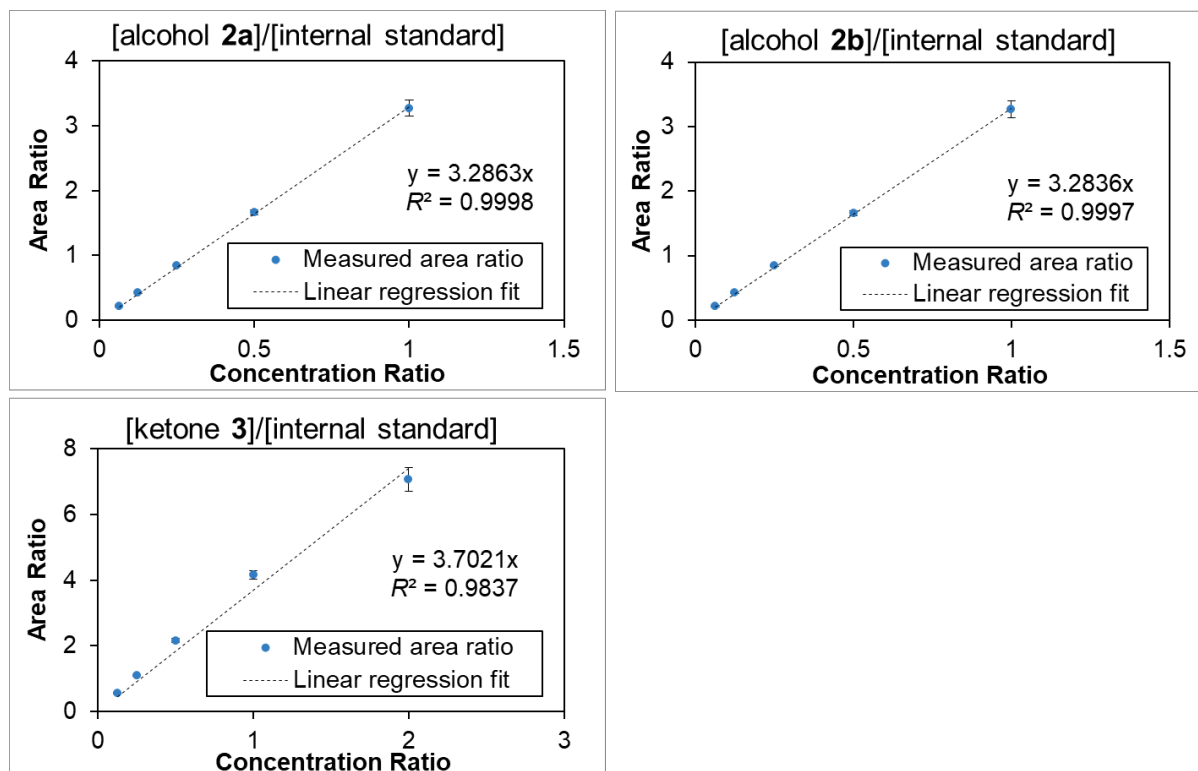
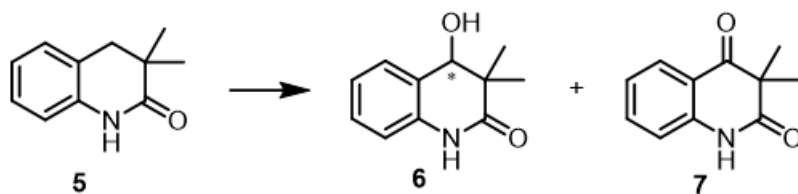


Figure S29. Calibration curves for the reaction with ethylbenzene derivative **1**. Alcohols **2a** and **2b** are the two enantiomers of alcohol **2**. The equation and the R^2 value of each linear regression are displayed in each plot area. Internal standard = indole.

4.4.4.2.2 Lactam **5**



Analytical method: Chiralpak IF[®], 4.6 × 250 mm; isocratic method with iPrOH 15% for 20 minutes, 2 mL/min. Injection volume = 10 μL. For the calibration line, indole (5 mM) and the product standards (from 0.15625 mM to 2.5 mM) were dissolved and mixed in analytical grade EtOAc. The peak areas of the compounds were integrated at the following wavelengths: indole = 265 nm, ketone **7** = 332 nm, alcohol **6** = 250 nm.

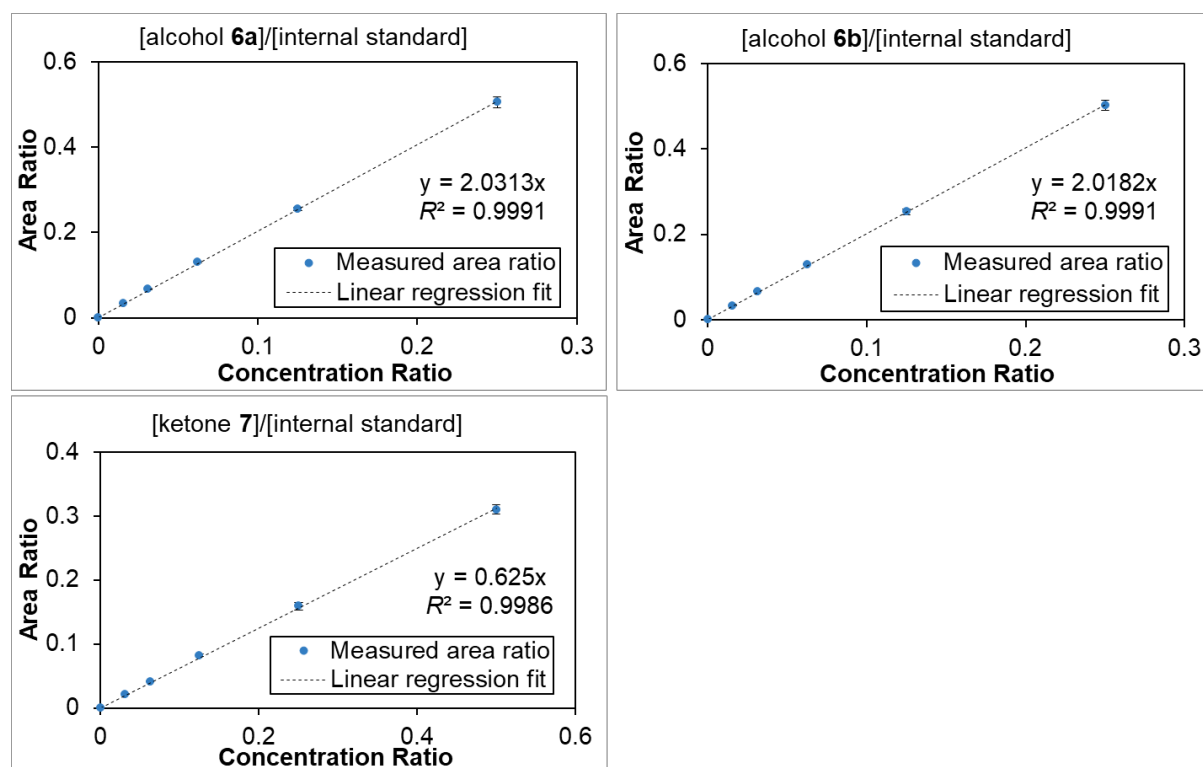


Figure S30. Calibration curves for the reaction with lactam 5. Alcohols **6a** and **6b** are the two enantiomers of alcohol **6**. The equation and the R^2 value of each linear regression fit are displayed in each plot area. Internal standard = indole.

4.4.5 DNSA displacement assay of the cofactors and hCAII^[51-52, 63-64]

First, the dissociation constant of dansylamide (K_{d_DNSA}) was determined by its change in fluorescence using a one-to-one binding to the protein. Based on the K_d value obtained, a competitive binding assay with DNSA and [BS-Fe-bTAML]⁻ was performed.

One-to-one binding assay for DNSA and hCAII. To 15 wells on a black flat-bottom 96-well plate, phosphate buffered saline, hCAII variant (100 nM), DNSA solution in DMSO (10 μ L, final DNSA concentration = 0.565 nM to 100 μ M, serial dilution from well to well), and water was added. The final volume was 200 μ L. The fluorescence in each well was measured (excitation: 280 nm, measurement: 470 nm). The datasets were obtained in triplicate, and they were fitted with a curve using the solver add-in in Excel software according to the following equation^[63].

$$F_{obs} = F_{free} + (F_{bound} - F_{free}) \frac{(C_{hCAII} + C_{DNSA} + K_{d_DNSA}) - \sqrt{(C_{hCAII} + C_{DNSA} + K_{d_DNSA})^2 - 4 \cdot C_{hCAII} \cdot C_{DNSA}}}{2 \cdot C_{hCAII}}$$

F_{obs} : observed fluorescence intensity

F_{free} : fluorescence intensity of the protein when not bound to DNSA

F_{bound} : fluorescence intensity of the protein-DNSA complex

C_{hCAII} : the total concentration of hCAII

C_{DNSA} : the total concentration of DNSA

K_{d_DNSA} : the dissociation constant of DNSA

DNSA displacement assay. To 15 wells on a black flat-bottomed 96-well plate, phosphate buffered saline, the hCAII variant (100 nM), DNSA solution in DMSO (10 μ L, final DNSA concentration = 2.25 μ M), [BS-Fe-bTAML]⁻ (0.565 nM to 100 μ M, serial dilution from well to well) and water was added. The final volume is 200 μ L. The fluorescence in each well was measured (excitation: 280 nm, measurement: 470 nm). The datasets were obtained in triplicate,

and were fitted with a curve using the solver add-in in Excel software according to the reported equations^[64].

4.4.6 Differential scanning fluorimetry

Thermal denaturation of the protein (1 mg/mL) was monitored by fluorescence intensity (280 nm excitation, 330 nm and 350 nm emission) during a continuous thermal ramp (20 to 95 °C, 1 °C/min). In parallel, apparent absorbance of excitation light due to light-scattering was measured by back-reflection optics, to monitor formation of aggregates.

The ratio of emission intensities at the detection wavelengths (350 nm / 330 nm) was analysed automatically using the “PR. Stability analysis” software, to determine the onset temperature (T_{on}) and midpoint temperature (T_m) for thermal denaturation. The onset temperature marks the departure of the signal from the baseline in the first observable denaturation transition, and the midpoint is the inflexion point of the sigmoidal curve (the maximum in the first derivative).

4.4.7 Directed evolution of the ArPase

4.4.7.1 Site-directed mutagenesis

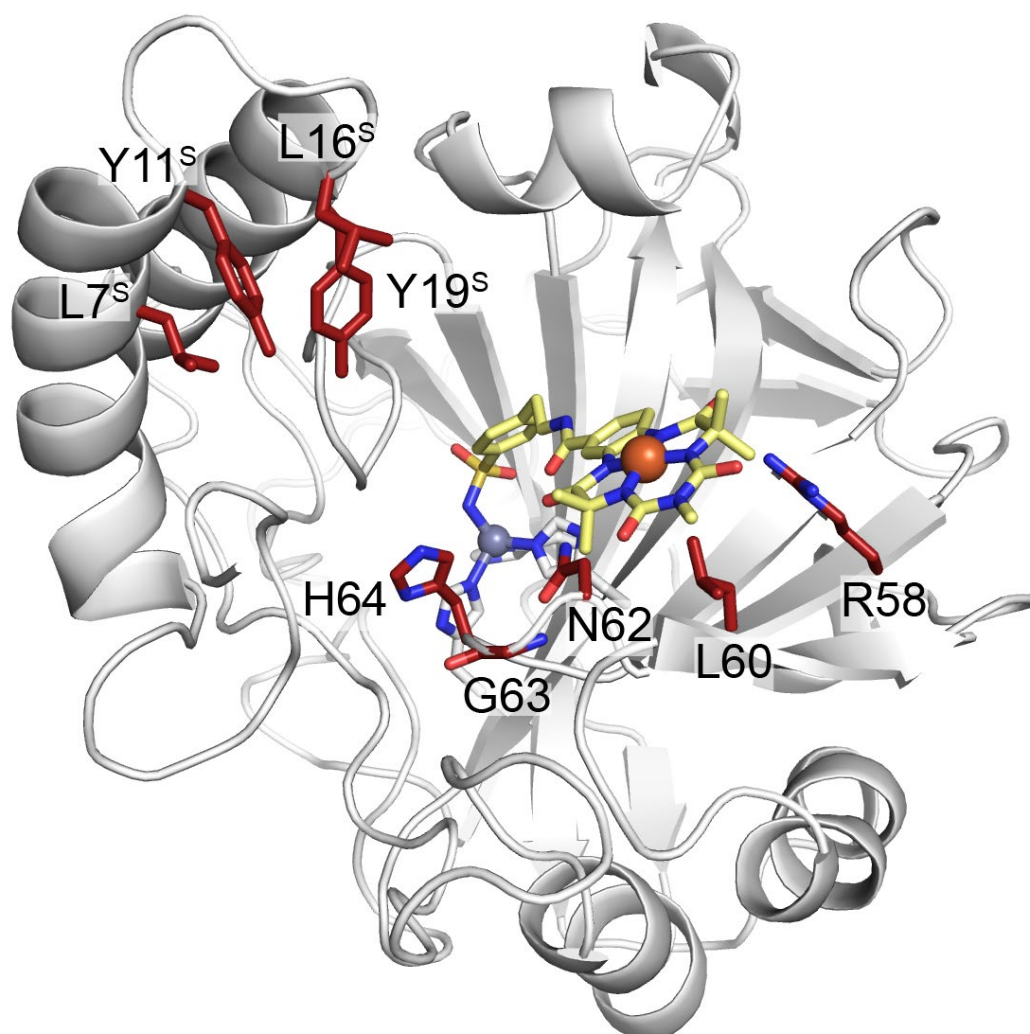


Figure S31. Targeted positions for the mutant screening of hCAII^{TS}-S4. The ColabFold prediction of hCAII^{TS}-S4 and a crystal structure of [BS-Fe-bTAML]⁻ · hCAII^{TS} are overlaid. The backbone of hCAII^{TS}-S4 prediction is represented as a cartoon model. [BS-Fe-bTAML]⁻, the zinc atom, and the histidines coordinating to the zinc atom are taken from the crystal structure and represented as stick models ([BS-Fe-bTAML]⁻ in yellow and the histidines in white with color-coded heteroatoms: O = red, N = blue, S = yellow) with the Fe as an orange sphere. The zinc atom is represented as a violet sphere. The mutated residues are represented as stick models in red with the color-coded heteroatoms. The residue numbers in the inserted motif (from 1^s to 28^s) are independent of the numbering in the parental scaffold.

Name	Sequence (5' to 3')
Fw L7 ^S X	CTGGCGAAAAA NDT GTGGAAGAATATCGC
Rv L7 ^S X	ATATTCTTCCACA HN TTTTTTTCGCCAGATCT
Fw Y11 ^S X	GTGGAAGA NDT CGCGGCACCAGC
Rv Y11 ^S X	GGTGCCGCGA HN TTCTTCCACCAG
Fw L16 ^S X	GGCACCAGC NDT GAAAAATATGCGCAG
Rv L16 ^S X	ATATTTTTCA HN GCTGGTGCCGCGATA
Fw Y19 ^S X	CTGGAAAA NDT GCGCAGATGGCGG
Rv Y19 ^S X	CATCTGCGCA HN TTTTTTCCAGGCTGG
Fw R58X	GCAACTTCCCTG NDT ATCCTCAACAATGGTCAT
Rv R58X	ATTGTTGAGGATA HN CAGGGAAGTTGCTTGATC
Fw L60X	TCCCTGAGAATC NDT AACAATGGTCATACCTTC
Rv L60X	ATGACCATTGTTA HN GATTCTCAGGGAAGTTGC
Fw N62X	AGAATCCTCAAC NDT GGTCATACCTTCAACGTG
Rv N62X	GAAGGTATGACCA HN GTTGAGGATTCTCAGGGA
Fw G63X	ATCCTCAACAAT NDT CATACCTTCAACGTG
Rv G63X	GTTGAAGGTATGA HN NATTGTTGAGGATTCT
Fw H64X	CTCAACAATGGT NDT ACCTTCAACGTGGAGTTT
Rv H64X	CACGTTGAAGGTA HN ACCATTGTTGAGGATTCT
Fw pET24a backbone	GCAGACAGTTTTATTGTTTCATGACCA
Rv pET24a backbone	AATAAACTGTCTGCTTACATAAAC

Table S9. Purchased primers for site-directed randomization of pET24a-hCAII^{TS}-S4. Fw = Forward primer, Rv = Reverse primer.

For the engineering of hCAII^{TS}-S4 by single point mutagenesis, we selected nine positions as depicted in Figure S31.

Each position was randomized with PCR mutagenesis using degenerate NDT codons. The primers are collected in Table S9. The resulting gene fragments were assembled by Gibson Assembly. The reaction mixture was then used for transforming *E. coli* Top10 or DH5 α cells, following the manufacturer's protocol. The obtained colonies were harvested together, and the plasmids were extracted with a mini-prep kit. The pooled plasmid library was sequenced by Sanger sequencing to determine its diversity.

4.4.7.2 Protein expression and cell lysate preparation

The pooled plasmid library was used for the transformation of *E. coli* BL21(DE3) chemically-competent cells, and the cells were treated as described in section 4.4.3.1 Cloning and Expression of hCAII^{TS}. After obtaining the colonies, at least 46 were picked with sterilized toothpicks and inoculated into an LB medium (0.3 mL), supplemented with a suitable antibiotics (chloramphenicol (30 µg/mL) for pACYC, kanamycin (50 µg/mL) for pET24a) in a 96-deep-well plate. The glycerol stocks of *E. coli* cells harboring the plasmid of the parent variant or the empty vector were inoculated into at least 1 well per plate as controls. The plates were sealed with sealing membrane (Breathe Easier sealing membrane for multiwell plates, Merck) and shaken overnight (37 °C, 300 rpm). The next morning, the preculture from each well (10 µL) was transferred to LB medium (1 mL) supplemented with chloramphenicol (30 µg/mL) on a 96-deep-well U-bottom plate. The plate was shaken (37 °C, 300 rpm, 2 hours) until the OD₆₀₀ reached 0.6 – 0.8. Then, IPTG (0.5 mM) was added to initiate the overexpression, and the plate was shaken (20 °C at 300 rpm, 20 hours). The remainder of the preculture was used for the preparation of the glycerol stock (25%v/v glycerol). After overexpression, the cells were harvested by centrifugation (4400g, 10 minutes at 4 °C). The supernatant was discarded by decantation, and the plate was laid upside-down on a paper towel (ca. 15 seconds). The plates were sealed with an aluminium foil seal (VWR) and stored at –20 °C overnight.

For the cell lysate preparation, the cell pellets were thawed at room temperature. The cells were suspended in lysis buffer (150 µL, lysozyme (1 mg/mL) and 1 µg/mL DNaseI) in the reaction buffer) and the plate was shaken (37 °C at 300 rpm for 30 min.). The lysed cells were frozen (–20 °C, 2 hours). The lysate was thawed again at room temperature and transferred into a 96-well PCR plate. The lysate was heated (55 °C, 30 minutes) using a PCR cycler. After cooling to room temperature, the lysate was flash-frozen in liquid nitrogen and stored at –80 °C. The lysate was used for the following screening assay within a week.

4.4.7.3 96-well HPLC assay for lactam **5**

The clear lysate was transferred to each well of a 96-well microtiter plate (PP, 1 mL). A catalysis master mix was prepared by mixing phosphate buffered saline (60%v/v), MeCN (40%v/v), lactam **5** (21 mM) and the cofactor (8.4 μ M). The master mix (95 μ L) was dispensed into each well of the 96-well plate and the plate was shaken (10 min, 300 rpm). H₂O₂ stock solution (5 μ L, 4 M in phosphate buffered saline, pH 7.4, 20 eq.) was added to each well and the reaction mixture was shaken (25°C, 1 hour). The reactions were then quenched by addition of HCl aq. (1M, 15 μ L) and indole stock solution (5 μ L, 800 mM in DMSO) was added as internal standard. The protein in the reaction mixture was removed by addition of MeOH (500 μ L) followed by centrifugation (4400 rpm, 1 hour, 4 °C). The clear supernatant from each well (120 μ L) was transferred to a new 96-well plate and analyzed with HPLC.

Analytical method: Poroshell 120 EC-C18, 2.1 x 50 mm, 1.9 μ m; a gradient method between solvent A (H₂O containing 3% MeCN and 0.1% TFA) and B (MeCN containing 3% H₂O and 0.1% TFA) from 10% solvent B for 0.5 min., 10% to 40% over 0.5 min, 40% for 2 min., 40% to 90% over 0.5 min., 90% for 0.5 min. Flow rate = 0.5 mL/min. Injection volume = 0.5 μ L.

The peak areas of the compounds were integrated at specific wavelengths: indole = 265 nm, alcohol **6** = 254 nm, ketone **7** = 230 nm.

Based on the TON and the [alcohol]/[ketone] ratio, best-performing wells were identified. The corresponding glycerol stock was sequenced. The hit mutants were further validated with cell lysate or purified proteins. However, the screening method resulted in a large error in the activity, as mentioned in the main text. Possible causes of the error are:

- Oxidation of indole by residual H₂O₂ before HPLC injection, thereby decreasing the reliability of the internal standard
- The expression level of the single mutants in the 96-well format was not enough for the screening condition
- Potential unknown side products eluted with one of the reaction products

4.4.8 Crystallographic characterization of [BS-Fe-bTAML]⁻ · hCAII^{TS}-S4

Lyophilized hCAII^{TS}-S4 was dissolved in ultrapure water (18.2 MΩ·cm, MilliQ, Millipore Corporation, Burlington, USA). The 96 well crystallization screens were set up by mixing precipitation buffer (0.15 μL) with protein solution (0.15 μL) with a robot (Crystal Gryphon, Art Robins Instruments, USA; MRC 2-well plates, SWISSCI, Switzerland), equilibrated against precipitation buffer (60 μL), and stored at 20 °C.

Apo crystals were grown by mixing the stock solution of hCAII^{TS}-S4 (~15 mg/mL) with precipitation buffer (PEG 3350 (20 %w/v), NaF (0.2 M), Bis-Tris propane (0.1 M, pH 7.5), PACT premierTM HT, Molecular Dimensions). For the cofactor soaking, a solution of [BS-Fe-bTAML]⁻ (0.5 μL, 10 mM in DMSO) was mixed with the precipitation buffer (4.5 μL). Then, this mixed solution (2 μL) was added to the drop containing the crystals. After soaking (overnight, 20 °C), the crystals were cryo-protected with ethylene glycol (25%) and flash-frozen in liquid nitrogen prior to data collection.

The data collection was carried out at the Swiss Light Source beam line X06SA at a wavelength of 1.0000 Å. XDS^[65] was used for crystal indexing, integration and AIMLESS^[66] for scaling, within the graphical interface CCP4i2^[67] of the CCP4 suite^[68]. The structures were solved by molecular replacement using PHASER-MR^[69] and an AlphaFold^[70] generated search model. Refinement was carried out by REFMAC5^[71] and for structure modeling and electron-density visualization COOT^[72] was used. Ligand restraints were generated using eLBOW^[73]. Figures were generated with PyMOL (the PyMOL Molecular Graphics System, Version 2.5.0, Schrödinger, LLC). Data collection and refinement statistics are listed in Table S10.

Four monomers were found per asymmetric unit. The first 44 residues at the N-terminus including the grafted motif were not well structured on the electron density map. This is presumably due to high flexibility of this 2-α helix loop. On the other hand, SDS-PAGE gel proved that the loop was not degraded.

Residual electron density in the Fo-Fc map was observed in the “active site” of the protein. Anomalous dispersion density was observed for the Zn²⁺ coordinated to His94, His96 and His119 residues. For modeling of [BS-Fe-bTAML]⁻, we used the anomalous density peak of iron. The 2Fo-Fc map showed a weaker electron density for the iron-binding part of the cofactor, suggesting higher flexibility.

Data Processing Statistics

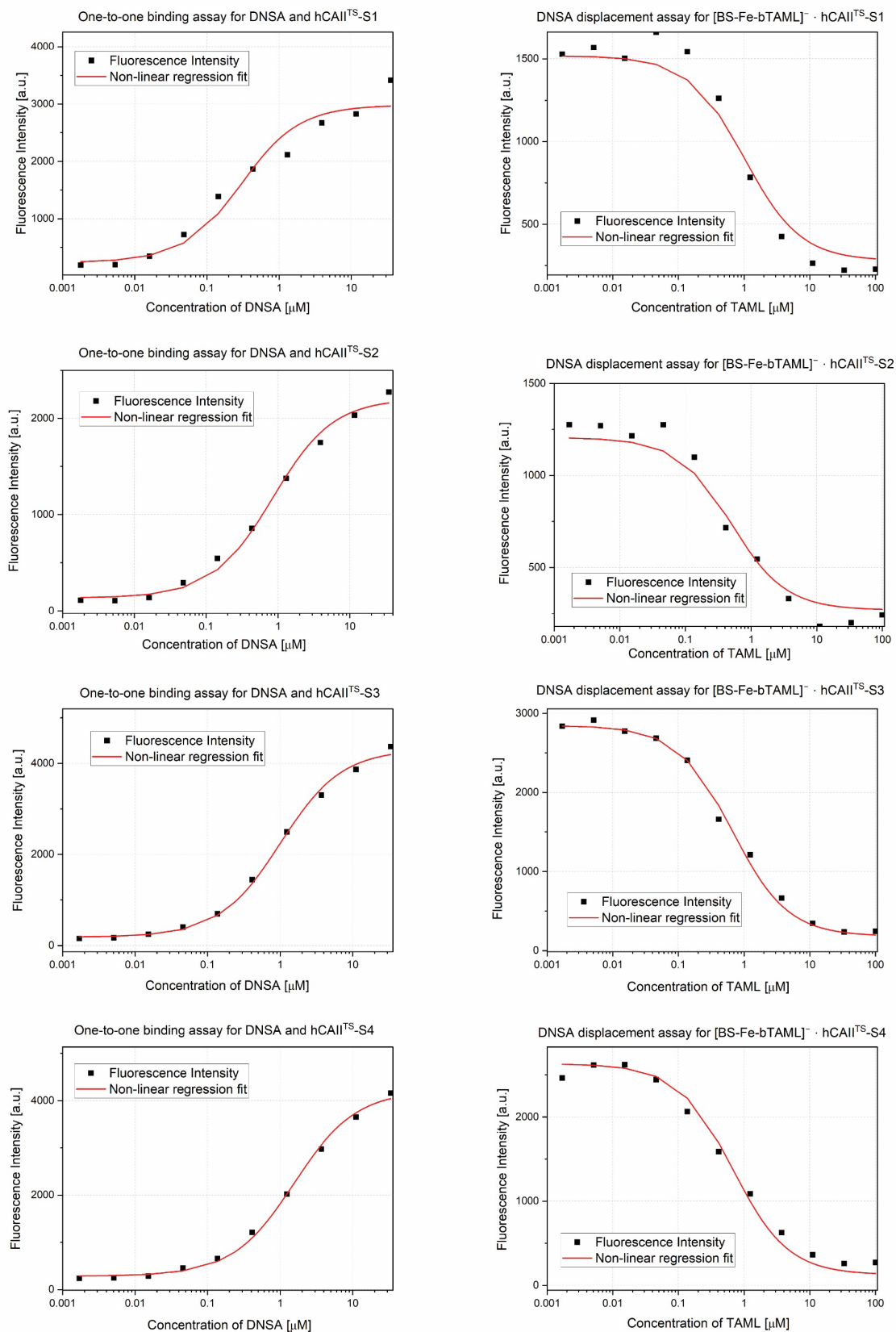
Resolution Range (Å)	49.53 - 2.024 (2.097 - 2.024)
Cell Parameters	
- a, b, c (Å)	55.13, 72.44 79.38
- α, β, γ (°)	95.96, 104.88, 109.12
Space group	P 1
Unique reflections	69640 (6691)
Rmerge (%)	10.3 (62.0)
Multiplicity	1.9 (1.9)
Mean I/Sig(I)	6.3 (1.1)
Completeness (%)	97.3 (93.3)
CC (1/2)	0.981 (0.589)

Structure Refinement Statistics

R _{work} /R _{free}	0.20/0.25
RMS deviation	
-Bond length (Å)	0.069
-Bond angles (°)	3.19
-Ramachandran favored (%)	96.2
Average B-factors (Å ²)	
-Protein	31.1
-Ligands	62.1
-Solvent	31.0

Table S10. Data processing and crystal structure refinement statistics

4.4.9 Supporting Figures

Figure S32. Displacement binding assay for hCAII^{TS}-S1,2,3,4 and [BS-Fe-bTAML]⁻.

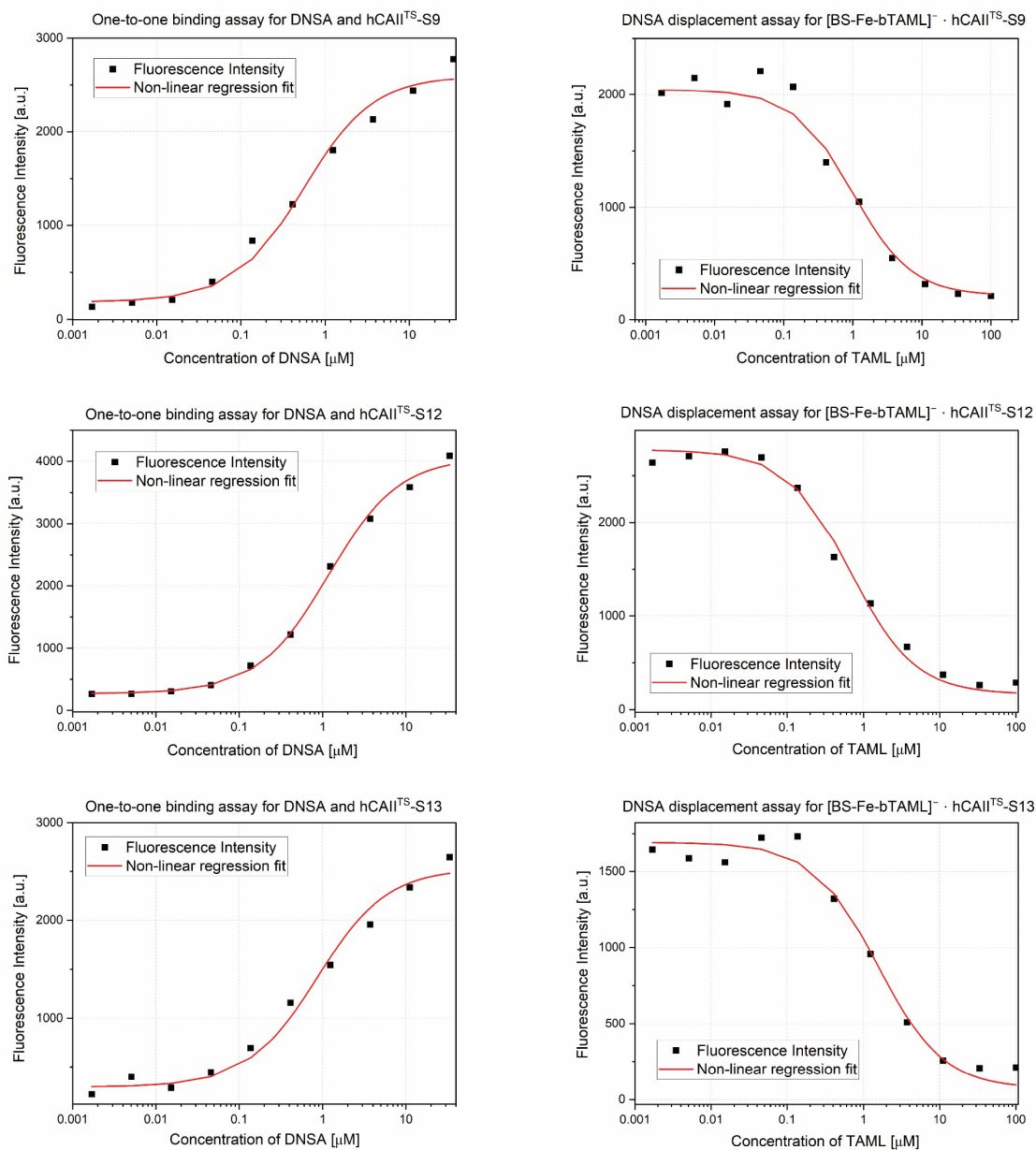
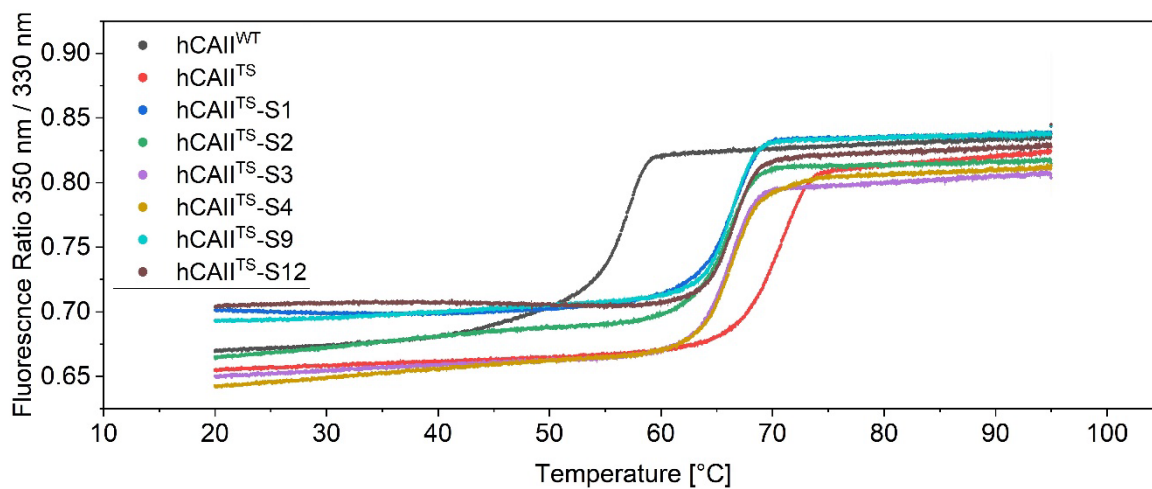


Figure S33. Displacement binding assay for hCAII^{TS}-S9,12,13 and [BS-Fe-bTAML]⁻.

(a)



(b)

Protein	T_m (°C)	T_{on} (°C)
WT	56.64 ± 0.01	38.55 ± 1.46
TS	70.56 ± 0.01	63.20 ± 0.13
TS-S1	66.36 ± 0.02	58.95 ± 0.11
TS-S2	65.91 ± 0.01	58.20 ± 0.22
TS-S3	66.05 ± 0.02	60.22 ± 0.33
TS-S4	66.29 ± 0.05	60.11 ± 0.04
TS-S9	66.25 ± 0.03	61.56 ± 0.08
TS-S12	66.42 ± 0.02	61.33 ± 0.00

Figure S34. Differential scanning fluorimetry data for T_m determination of hCAII-Sx. (a) Plot of the ratio of the fluorescence intensity (350 nm / 330 nm). (b) Calculated midpoint (T_m) and onset (T_{on}) temperature of the hCAII variants.

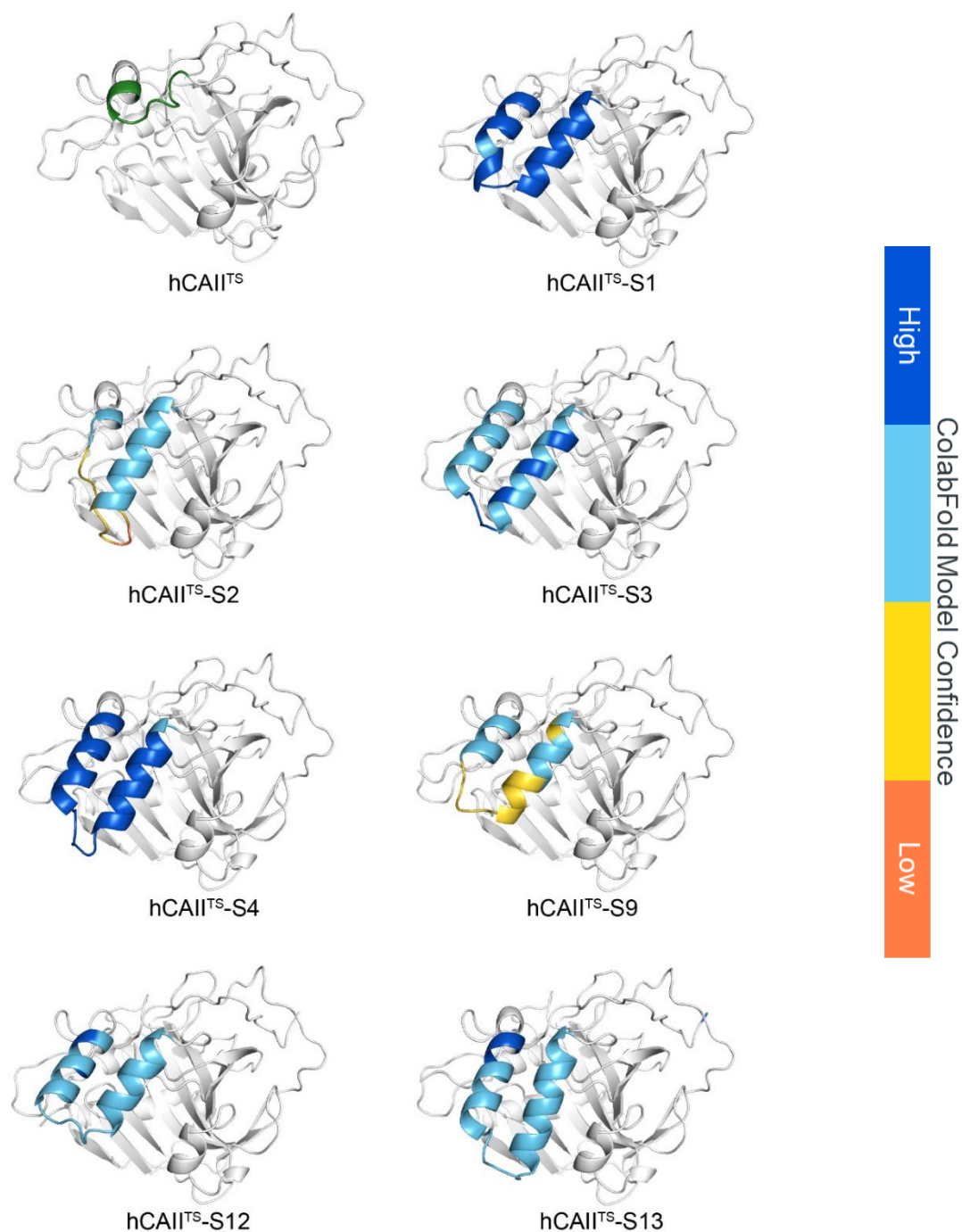


Figure S35. Structures of the hCAII^{TS}-Sxs predicted by ColabFold. The prediction was performed with ColabFold^[74]. The protein backbones are represented as cartoon models. The structure of hCAII^{TS} was taken from X-ray crystallography and the residues 17 to 24 are highlighted in green. For the other predicted structures, the inserted motifs are color coded based on the prediction confidence of ColabFold.

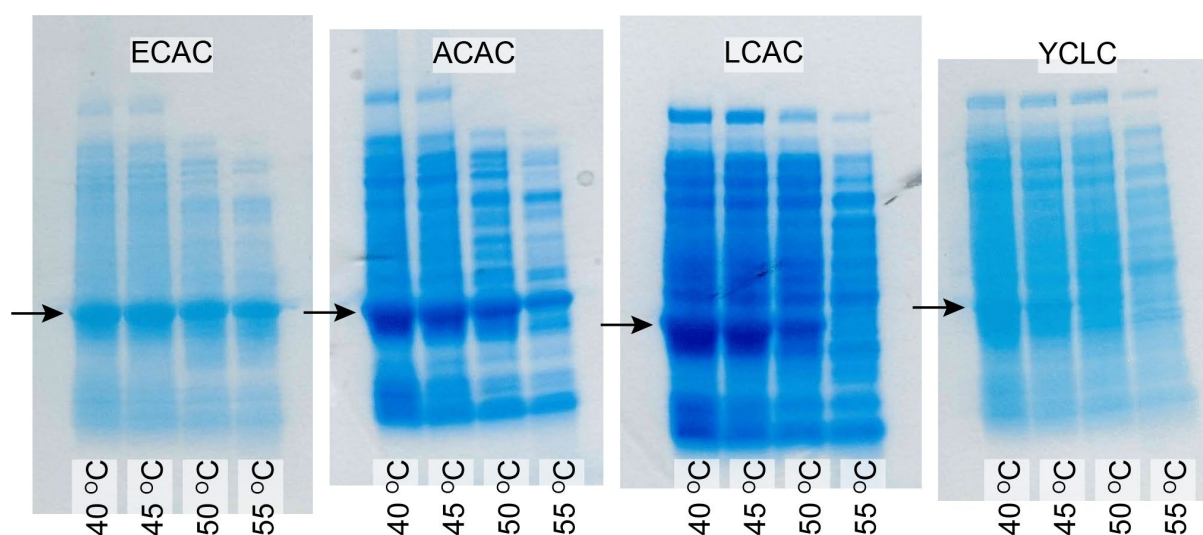


Figure S36. The thermal stability of the disulfide mutants were tested using cell lysate. The cell lysate was heated to the specified temperature for 30 minutes and the precipitate was spun down. The treated lysate was subjected to SDS-PAGE analysis. The bands of the desired proteins are indicated with the black arrows. ECAC = hCAII^{TS}-S4-RR-ECAC, ACAC = hCAII^{TS}-S4-RR-ACAC, LCAC = hCAII^{TS}-S4-RR-LCAC, YCLC = hCAII^{TS}-S4-RR-YCLC.

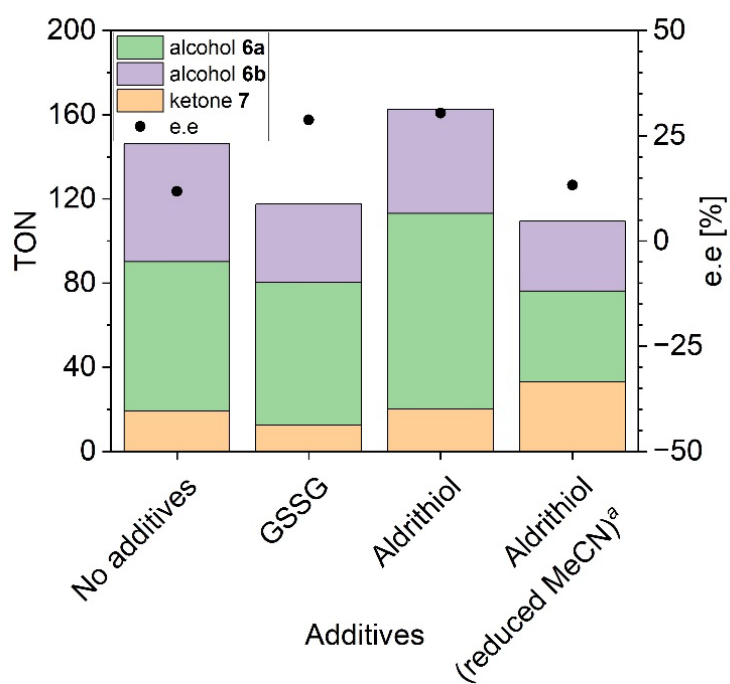


Figure S37. Hydroxylation of lactam 5 with [BS-Fe-bTAML]⁻ · hCAII^{TS}-S4-RR-ECAC. Conditions: substrate (5 mM), cofactor (4 μM), hCAII^{TS}-S4-RR-ECAC (>10 μM), H₂O₂ (10 mM), phosphate buffered saline, MeCN (20%v/v), 25 °C, 1 hour. The protein was incubated with the additive prior to the reactions (25 °C, 30 minutes). All the reactions were analyzed with SFC. ^a[substrate] = 1.25 mM, [MeCN] = 5%v/v.

4.5 References

- [51] R. Iyer, A. A. Barrese, S. Parakh, C. N. Parker, B. C. Tripp, *J. Biomol. Screen.* **2006**, *11*, 782-791.
- [52] a) T. T. Baird, A. Waheed, T. Okuyama, W. S. Sly, C. A. Fierke, *Biochemistry* **1997**, *36*, 2669-2678; b) S. C. Wang, D. B. Zamble, *Biochem. Mol. Biol. Educ.* **2006**, *34*, 364-368.
- [60] M. Bozdogan, S. Isik, S. Beyaztas, O. Arslan, C. T. Supuran, *J. Enzyme Inhib. Med. Chem.* **2015**, *30*, 240-244.
- [61] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, *Protein Sci.* **1995**, *4*, 2411-2423.
- [63] X. Li-Blatter, L. Zweifel, T. Sharpe, in *Protein-Ligand Interactions*, 3 ed. (Eds.: T. Daviter, C. M. Johnson, S. H. McLaughlin, M. A. Williams), Humana New York, NY, New York, **2021**, pp. 47-79.
- [64] Z.-X. Wang, *FEBS Lett.* **1995**, *360*, 111-114.
- [65] W. Kabsch, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125-132.
- [66] P. R. Evans, G. N. Murshudov, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1204-1214.
- [67] L. Potterton, J. Agirre, C. Ballard, K. Cowtan, E. Dodson, P. R. Evans, H. T. Jenkins, R. Keegan, E. Krissinel, K. Stevenson, A. Lebedev, S. J. McNicholas, R. A. Nicholls, M. Noble, N. S. Pannu, C. Roth, G. Sheldrick, P. Skubak, J. Turkenburg, V. Uski, F. von Delft, D. Waterman, K. Wilson, M. Winn, M. Wojdyr, *Acta Crystallogr., Sect. D: Struct. Biol.* **2018**, *74*, 68-84.
- [68] J. Agirre, M. Atanasova, H. Bagdonas, C. B. Ballard, A. Basle, J. Beilsten-Edmands, R. J. Borges, D. G. Brown, J. J. Burgos-Marmol, J. M. Berrisford, P. S. Bond, I. Caballero, L. Catapano, G. Chojnowski, A. G. Cook, K. D. Cowtan, T. I. Croll, J. E. Debreczeni, N. E. Devenish, E. J. Dodson, T. R. Drevon, P. Emsley, G. Evans, P. R. Evans, M. Fando, J. Foadi, L. Fuentes-Montero, E. F. Garman, M. Gerstel, R. J. Gildea, K. Hatti, M. L. Hekkelman, P. Heuser, S. W. Hoh, M. A. Hough, H. T. Jenkins, E. Jimenez, R. P. Joosten, R. M. Keegan, N. Keep, E. B. Krissinel, P. Kolenko, O. Kovalevskiy, V. S. Lamzin, D. M. Lawson, A. A. Lebedev, A. G. W. Leslie, B. Lohkamp, F. Long, M. Maly, A. J. McCoy, S. J. McNicholas, A. Medina, C. Millan, J. W. Murray, G. N. Murshudov, R. A. Nicholls, M. E. M. Noble, R. Oeffner, N. S. Pannu, J. M. Parkhurst, N. Pearce, J. Pereira, A. Perrakis, H. R. Powell, R. J. Read, D. J. Rigden, W. Rochira, M. Sammito, F. Sanchez Rodriguez, G. M. Sheldrick, K. L. Shelley, F. Simkovic, A. J. Simpkin, P. Skubak, E. Sobolev, R. A. Steiner, K. Stevenson, I. Tews, J. M. H. Thomas, A. Thorn, J. T. Valls, V. Uski, I. Uson, A. Vagin, S. Velankar, M. Vollmar, H. Walden, D. Waterman, K. S. Wilson, M. D. Winn, G. Winter, M. Wojdyr, K. Yamashita, *Acta Crystallogr., Sect. D: Struct. Biol.* **2023**, *79*, 449-461.
- [69] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658-674.
- [70] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, *Nature* **2021**, *596*, 583-589.
- [71] G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 355-367.
- [72] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Cryst.* **2010**, *66*, 486-501.
- [73] N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2009**, *65*, 1074-1080.
- [74] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, M. Steinegger, *Nat. Methods* **2022**, *19*, 679-682.

5 Concluding Remarks

5.1 Summary and Outlook

In this PhD thesis, we delved into the development of ArPases based on an Fe-bTAML catalyst and hCAII proteins. We found that the investigated protein scaffolds generally have a positive effect in TONs for the tested hydroxylation reactions. We evolved an ArPase relying on a caged chromophore as substrate with an easy readout. After the three rounds of directed evolution, the evolved variant exhibited enhanced catalytic efficiency, TTON, and enantioselectivity. We propose that these enhancements are the result of interactions between the cofactor and arginine side chain introduced in the course of the directed evolution campaign. To explore the possibility of further enhancing the enantioselectivity, we tested a rationally-designed library of chimeric hCAIIs. We found that the grafted motifs enhanced the TON, although the enantioselectivity did not improve. The introduction of the disulfide bridges may be a promising approach for further development of stereoselective ArPases.

The former part of the thesis underscores the effect of the interaction between the second coordination sphere and a metal complex in enhancing its catalytic activity. The latter part highlighted the challenge in establishing specific interactions between the substrate and the protein scaffold. Overall, initial goals 1 (H_2O_2) and 2 (high TTON) were achieved, although there is a room for improvement for goal 3 (enantioselectivity).

This challenge stems from the current design strategy of ArMs. In most cases, the design of ArMs is centered around a specific binding between a natural protein and a cofactor. This 'bottom-up' approach limits the repertoire of the available protein scaffolds. Accordingly, the architecture of the active site may not be optimal for the reaction of interest.

A promising, yet ambitious solution for this challenge is a top-down approach, where a suitable protein scaffold is built for the specific catalyst and substrates. Recent progress in computational protein design methods enables the generation of a *de-novo* small-molecule binding protein with the structure of the small molecule as the input^[75]. Accordingly, it is

possible to build a *de-novo* protein scaffold around a metal complex or a transition state configuration of the catalyst and the substrate. Although many challenges are foreseeable (e.g. the affinity and the specificity), this could be a potential path for tailor-made development of ArMs.

5.2 Reference

- [75] R. Krishna, J. Wang, W. Ahern, P. Sturmfels, P. Venkatesh, I. Kalvet, G. R. Lee, F. S. Morey-Burrows, I. Anishchenko, I. R. Humphreys, R. McHugh, D. Vafeados, X. Li, G. A. Sutherland, A. Hitchcock, C. N. Hunter, A. Kang, E. Brackenbrough, A. K. Bera, M. Baek, F. DiMaio, D. Baker, *Science*, 0, ead12528