TECHNISCHE UNIVERSITÄT MÜNCHEN

DEPARTMENT CHEMIE

LEHRSTUHL FÜR BIOCHEMIE

The Red Sea Brine Pools and the Biotin-Streptavidin Technology as Sources for Highly Stable und Selective Biocatalysts

Anastassja Leyla Gespers

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Matthias Feige

Prüfer der Dissertation:

        1. Prof. Dr. Michael Groll

        2. Prof. Dr. Dirk Weuster-Botz

Die Dissertation wurde am 31.07.2018 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 17.09.2018 angenommen.

Acknowledgments

Abbreviations

Å	Ångström
ACN	Acetonitrile
ADH	Alcohol dehydrogenase
ADH/A1	Alcohol dehydrogenase of Atlantis II (old annotation)
ADH/A1a	Alcohol dehydrogenase of Atlantis II (new annotation)
ArM	Artificial metalloenzyme
Biot	Biotin
CAD	Cinnamyl alcohol dehydrogenase
CD	Circular dichroism
CDD	Conserved domain database
C-Terminal	Carboxy-terminal
bp	Base pair
Da	Dalton [g mol^{-1}]
dH$_2$O	Destilled water (or MilliQ water)
diRh	Dirhodium(II) complex
DMSO	Dimethylsulfoxide
DNA	Desoxyribonucleic acid
DNAse	Deoxyribonuclease
dNTP	Deoxyribonukleotide
EDTA	Ethylenediaminetetraacetic acid
E. coli	*Escherichia coli*
ee%	Enantiomeric access in percent (%)
eq.	Equivalents
ESI	Electrospray ionization
et al.	*et alii*
g	Gravity
GC	Gas chromatography

HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
his$_6$	Hexahistidine
HPLC	High performance liquid chromatography
Hfx. volcanii	*Haloferax volcanii*
IMAC	Immobilized metal affinity chromatography
INDIGO	Integrated Data Warehouse of Microbial Genomes
IPTG	Isopropylthiogalactoside
LB	Lysogeny broth
LC	Liquid chromatography
LDR	Long-chain dehydrogenase/reductase
MALDI	Matrix-assisted laser desorption/ionization
MDR	Medium-chain dehydrogenase/reductase
MeOH	Methanol
MS	Mass spectrometry
MTP	Modified tryptone medium
MW	Molecular weight
NAD(P)$^+$/H	Nicotinamidadenindinukleotide(phosphate)
NCBI	National Center of Biotechnology Information
N-terminal	Amino-terminal
OD$_{600}$	Optical density at 600 nm
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
pI	Isoelectric point
RMSD	Root-mean-square deviation
RNA	Ribonucleic acid
rpm	Rounds per minute
SAG	Single amplified genome
SAV	Streptavidin

SDS-PAGE	Sodium dodecyl sulfate
SDR	Short-chain dehydrogenase/reductase
SEC	Size exclusion chromatography
sp.	Species
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TOF	Time of flight
TON	Turnover number
Tris	Tris(hydroxymethyl)aminomethane
U	Unit
UAA	Unnatural amino acid
v/v	Volume per volume
w/v	Weight per volume
wt	Wild-type
ΔG	Free Gibbs energy/free entalphy
ε	Molar extinction coefficient
λ	Wavelenght

Summary

Highly stable and selective enzymes represent the foundation of the white biotechnology, which is forecasted to be the major driver for the global growth of the green chemical industry. New biotechnological products and processes require the constant discovery or development of novel biocatalysts with advanced functionalities and stabilities. Here, two different strategies were utilized to provide stable and selective biocatalysts with new reactivities.

The discovery of novel enzymes from nature is a major pool for biocatalysts with improved properties. Enzymes originating from hostile environments are often highly stable under industrial conditions and, therefore, are advantageous for the production of chiral compounds. Using single amplified genomes (SAGs), an alcohol dehydrogenase (ADH) encoding gene, *adh/a1a*, was identified of an uncultured species from the Atlantis II Deep Red Sea brine pool (16.8% salt, 63 °C, pH 5.3). ADH/A1a was highly stable and active at elevated temperatures and high salt concentrations (optimal activity at 70 °C and 4 M KCl), as well as in the presence of organic solvents. The polyextremophilic ADH/A1a exhibited broad substrate specificity for both primary and secondary alcohols, with a preference for cinnamyl alcohol and its derivatives. Interestingly, ADH/A1a was able to reduce cinnamyl-methyl-ketone and raspberry ketones in a reverse reaction, making it a possible candidate for the production of chiral secondary alcohols. Based on the substrate specificities and conserved domains, ADH/A1a was annotated in the cinnamyl alcohol dehydrogenase (CAD) family, one of the least studied microbial medium-chain ADH families. The structure analysis revealed a high abundance of acidic residues on the exposed surface at the expense of the basic residues, representing a unique feature of haloadaptation. The reduced surface hydrophobicity, however, resembles the structural adaptations of both halophilic and thermophilic enzymes. The extraordinary structural and catalytical properties of ADH/A1a suggest potential applications in the green chemical industry.

Summary

In parallel to the enzyme discovery, the development of artificial metalloenzymes (ArM) became an uprising tool for the creation of hybrid biocatalysts that combine a biological scaffold with an organometallic catalyst, providing novel reactivities that are not found in nature. Dirhodium(II) tetracarboxylate complexes have shown to be robust and efficient catalysts for carbene-transfer reactions, being active and stable under physiological conditions. Here, a dirhodium artificial metalloenzyme was designed based on the biotin-streptavidin technology, providing a hybrid biocatalyst for the enantioselective catalysis of the asymmetric cyclopropanation. The host scaffold streptavidin (SAV) was mutated in a rational approach using available structure information to create a hydrophobic binding pocket for the biotin-conjugated dirhodium(II) complex (Biot-diRh) and to introduce potential residues for the co-activation of substrates. Moreover, to improve the stability in organic solvents and to increase the surface hydrophobicity, the protein surface of SAV was modified by protein PEGylation. The SAV-dirhodium hybrid catalysts were highly stable in organic solvents up to 80% (v/v) and temperatures up to 100 °C, as well as under physiological aqueous conditions. The catalytic potential was validated in the intermolecular cyclopropanation reaction of a diazo substrate and methoxystyrene, showing enantioselective conversion. The optimization of the reaction conditions minimized the side product formation and increased the yield of the desired cyclopropane product considerably; furthermore, the enantiomeric access (ee%) was enhanced using SAV with altered hydrophobic key residues in the binding pockets. The created platform of highly stable and enantioselective dirhodium ArMs may be extended for the catalysis of carbene transfer reactions, including the C-H insertion, the amino-oxyarylation and the aziridination. Beside the implementation in the chemical synthesis of chiral compounds, several applications for the dirhodium ArM may be found in the chemical biology.

Zusammenfassung

Hochstabile und selektive Enzyme bilden das Fundament der weißen Biotechnologie, die sich zum Hauptantrieb des globalen Wachstums der industriellen „grünen" Chemie entwickelt hat. Neue biotechnologische Produkte und Verfahren erfordern die ständige Entdeckung oder Entwicklung von neuartigen Biokatalysatoren mit verbesserten Stabilitäten und Funktionalitäten. In dieser Arbeit werden zwei verschiedene Strategien verfolgt, um stabile, selektive und hochreaktive Biokatalysatoren zu entwickeln.

In der Natur befindet sich ein riesiges Repertoire an Biokatalysatoren mit verschiedensten Eigenschaften. Enzyme, die aus lebensfeindlichen Umgebungen stammen, sind oft sehr stabil und daher vorteilhaft für die industrielle Herstellung von chiralen chemischen Verbindungen. Unter Verwendung von sequenzierten Einzelzellgenomen (SAGs) wurde das Gen *adh/a1a* einer Alkoholdehydrogenase (ADH) identifiziert, welches von einer unkultivierten Spezies aus dem „Atlantis II Deep" Solebecken des Roten Meeres entstammt (16,8% Salzgehalt, 63 °C, pH 5,3). Die ADH/A1a war außerordentlich stabil und aktiv bei erhöhten Temperaturen und hohen Salzkonzentrationen (optimal bei 70 °C und 4 M KCl) sowie in Gegenwart von organischen Lösungsmitteln. Die polyextremophile ADH/A1a zeigte eine breite Substratspezifität für primäre und sekundäre Alkohole, mit einer Präferenz für Zimtalkohol und dessen Derivate. Interessanterweise war ADH/A1a in der Lage sowohl Cinnamyl-methyl-keton also auch Himbeerketon in der Rückreaktion zu reduzieren, was es zu einem möglichen Kandidaten für die Herstellung chiraler sekundärer Alkohole macht. Basierend auf den Substratspezifitäten und den konservierten Domänen wurde ADH/A1a als ein Mitglied der Zimtalkoholdehydrogenase (CAD) Familie, einer der am wenigsten untersuchten mikrobiellen mittelkettigen ADH-Familien, annotiert. Die Strukturanalyse offenbarte als Haloadaptation einen erhöhten Anteil an sauren und einen verringerten Anteil an basischen Resten auf der exponierten Proteinoberfläche. Die reduzierte Oberflächenhydrophobizität ist sowohl als strukturelle Adaptionen halophiler als auch thermophiler Enzymen bekannt. ADH/A1a ist aufgrund der außergewöhnlichen strukturellen und katalytischen Eigenschaften ein möglicher Kandidat für die Anwendungen in der industriellen „grünen" Chemie.

Parallel zur Entdeckung von neuartigen Enzymen ist die Entwicklung künstlicher Metalloenzyme (ArM), bei dem ein biomolekulares Gerüst mit einem metallorganischen Katalysator kombiniert wird, eine gezielter Ansatz für die Herstellung von Biokatalysatoren mit neuartige Reaktivitäten, die in der Natur nicht vorkommen. Dirhodium(II)tetracarboxylatkomplexe haben sich als robuste und effiziente Katalysatoren für Carben-transfer-Reaktionen erwiesen, die unter physiologischen Bedingungen aktiv und stabil sind. In dieser Arbeit wird ein künstliches Dirhodium-Metalloenzym basierend auf der Biotin-Streptavidin-Technologie entwickelt und als enantioselektiver Hybrid-Biokatalysator in der asymmetrischen Cyclopropanierung eingesetzt. Das als biomolekulares Grundgerüst verwendete Streptavidin (SAV) wurde durch gezielte Mutation verändert, um eine hydrophobe Bindungstasche für den Biotin-konjugierten Dirhodium(II)-komplex (Biot-diRh) bereit zu stellen und mögliche Reste für die Co-Aktivierung von den Substraten einzuführen. Um die Stabilität in organischen Lösungsmitteln zu verbessern und die Oberflächenhydrophobie zu erhöhen, wurde die Proteinoberfläche von SAV durch PEGylierung modifiziert. Die hergestellten SAV-Dirhodium-Hybridkatalysatoren waren in organischen Lösungsmitteln bis zu 80% (v/v) und Temperaturen bis zu 100 °C sowie unter physiologischen wässrigen Bedingungen höchst stabil. Das katalytische Potential wurde in der intermolekularen Cyclopropanierung eines Diazosubstrats und 4-Methoxystyrols mit enantioselektiver Umsetzung nachgewiesen. Die Optimierung der Reaktionsbedingungen minimierte die Nebenproduktbildung und erhöht die Ausbeute des gewünschten Cyclopropanproduktes erheblich. Außerdem wurde der Enantiomerenüberschuss (ee%) durch SAV Mutanten mit hydrophoben Resten an Schlüsselpositionen in den Bindungstaschen erhöht. Die entwickelten hochstabilen und enantioselektiven Dirhodium-ArMs könnten zusätzlich für die Katalyse von Carben-Transfer-Reaktionen, einschließlich der CH-Insertion, der Aminooxyarylierung und der Aziridinierung, verwendet werden. Neben dem Einsatz in der chemischen Synthese von chiralen Verbindungen könnten weitere nützliche Anwendungen für die Dirhodium-ArMs in der chemischen Biologie gefunden werde.

Table of contents

1 Introduction

1.1 General introduction

1.1.1 Catalysis

Catalysis plays a major role in both nature and industry since almost every reaction that produces or functionalizes a chemical, biomolecule or material depends on a specific catalyst [1]. The general definition of 'catalysis', a phenomenon by which a catalyst positively influences a reaction profile without undergoing any permanent change, was first defined by Wilhelm Ostwald. Hereby, the catalyst forms a temporary complex with the reactant and enables an alternative reaction mechanism with accelerated reaction rates, involving several stabilizing transition states. The overall activation energy of the catalytic reaction is decreased and is characterized by a reduced Gibbs free energy ($\Delta G'_{cat} < \Delta G'_{uncat}$), whereby the overall thermodynamics and equilibrium of the reaction remains the same ($\Delta G_{reaction}$) (Figure 1). The state of the catalyst at the beginning of the reaction is the same as at the end of the reaction, enabling the proceeding into the next catalytic cycle [2].
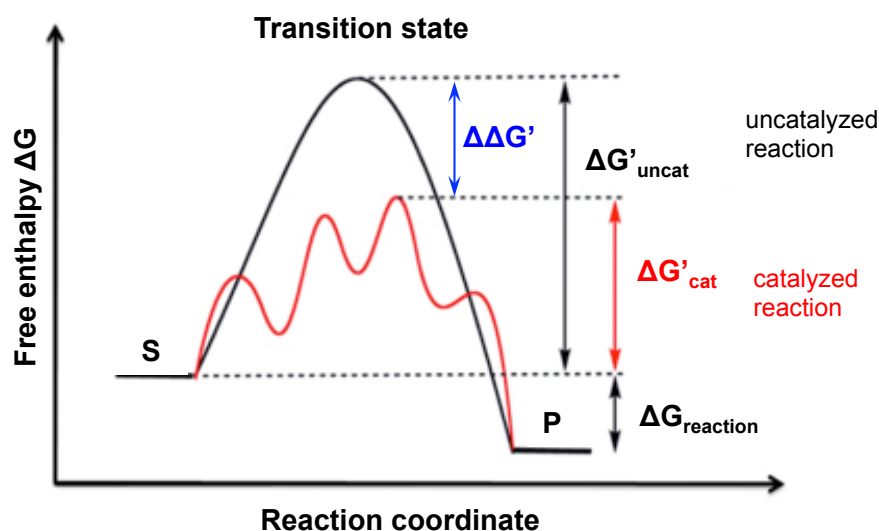
Figure 1: Depiction of the transition state of an uncatalyzed and catalyzed exothermic reaction of a substrate (S) to a product (P). The Gibbs free enthalphy $\Delta G'$ of the activation is shown for both reactions, as well as the overall free enthalphy of the reaction. The energetic difference in the activation energy between both reactions is shown by $\Delta\Delta G'$ ($\Delta\Delta G' = \Delta G'_{uncat} - \Delta G'_{cat.}$).

1

Since then, the field of catalysis was rising and became enormously important for today's industry. Several innovative approaches had led to the development and industrial application of a multitude of catalysts that contribute essentially in the production of numerous materials. Today, catalysis globally impacts the industry in all major domains such as health, nutrition, crop protection, energy and new materials, more than ever [3]. However, the chemical industry attained a very much bad public image for decades because of the generation of hazardous waste and pollution of both air and water. Furthermore, several chemical accidents had severely impacted societies and ecosystems [4]. The exploitation of non-renewable resources resulted in an escalation of costs for energy and raw material and resembles another major problem of the chemical industry [3]. Based on that, the concept of 'green chemistry' was initiated in 1990 that is seeking for waste minimization and application of sustainable non-hazardous alternative technologies that are contrary to the traditional chemistry [5].

1.1.2 Heterogeneous, homogeneous and biocatalysis

Catalysis is comprised by three major fields, the heterogeneous-, the homogeneous- and the biocatalysis. The heterogeneous catalysis refers to catalytic reactions in which reactants and catalysts are in different phases referring to gas, liquid or solid. [6]. Most of the heterogeneous catalysts are continuous solids such as oxides, sulfides, halides, and metals. Membrane catalysts and nanoporous solids represent a smaller group of heterogeneous catalysts, the latter containing open frameworks with immobilized metal ions, metal complexes or nanoparticles [7]. Heterogeneous catalysts have attracted a lot of attention in industrial use. They are especially used in (petro-) chemical, pharmaceutical and environmental processes to produce bulk intermediates and fine chemicals and also support depollution by converting poisonous molecules like carbon monoxide and nitrogen oxide into less harmful species [8].

In the homogeneous catalysis the reactants and the catalyst are present in the same phase means in liquid, solid or gas phase. Homogeneous catalysts comprehend a great variety, ranging from Brønstedt and Lewis acids over metal ions and organic compounds to

metal-, and organometallic complexes, the latter comprising of a metal possessing at least one polar bond to a carbon atom of an organic compound [9]. Most chemists nowadays refer to the 'homogeneous catalysis' as the liquid-phase catalysis by organometallic complexes (Figure 2). These organometallic catalysts become increasingly important in the modern fine-chemical and bulk-chemical industry. Particularly, transition metal complexes are increasingly applied due to their ability to form stable complexes [10]. Structurally, the catalytic active metal ion is directly bound to an ordered array of organic ligands that are forming the inner 'first coordination sphere'. The surrounding inner sphere plays a crucial role in the reactivity of the metal center and may influence the interaction of the metal with the substrate as well as the stereo- or site selectivity. More distal ligands are part of the 'secondary coordination sphere' and are also able to impart the catalyst reactivity [11]. By changing the ligands of the metal complex, one can change the catalysts properties. Notably, two of the seven Nobel Prices in chemistry in this century had been awarded in the field of the homogeneous catalysis for the asymmetric hydrogenation and oxidation by Barry Sharpless [12], William Knowels [13] and Ryoji Noyori [14] in 2001, and the asymmetric metathesis by Yves Chauvin [15], Robert Grubbs [16] and Richard Schrock [17] in 2005.
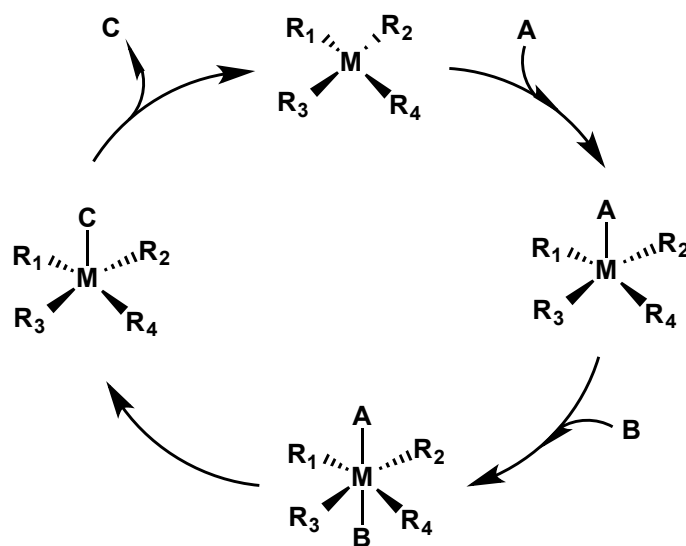
Figure 2: Elementary steps of the homogeneous catalysis by organometallic complexes. The organometallic catalyst (MR_{1-4}) forms an intermediate complex with the reactants A and B and catalyzes the reaction under formation of the product C.

The field of biocatalysis lies between the homogeneous and heterogeneous catalysis and comprises all reactions catalyzed by biological materials including enzymes, proteins, ribonucleoproteins, DNA and RNA, whole cells and cell extracts, or even plants. In terms of 'green chemistry', biocatalysis offers many attractive features. The catalyst is derived from a renewable resource and is biocompatible, biodegradable and non-hazardous to the environment. The catalytic reaction may be proceeded under physiological conditions and reaches, besides high activities, exceptional chemo-, region- and stereoselectivities, resulting in the higher purity of the products and less waste production compared with traditional chemical processes. Moreover, biocatalysts are easy and economically in production and often recyclable [18]. Biocatalysis is performed with isolated biocatalysts, mostly enzymes, or as whole-cell biotransformation. The advantage of isolated enzymes is the minimization of contamination and side-product formation by other enzymes of the host cells. On the other hand, the biotransformation in whole cells is less expansive and avoids laborious production and purification of enzymes. Furthermore, inner-cell multi-enzyme cascades and cofactor regenerations are envisaged, making the whole cell approach even more lucrative. However, the biocatalysis with isolated enzymes is often preferred as significant higher substrate concentrations may be used and, therefore, higher productivities are reached with superior purities and far less waste and water usage compared to fermentation processes [19].

1.1.3 Enzymes in biocatalysis

Enzymes are macromolecular biocatalysts that specifically accelerate chemical reactions of the cellular metabolism and hence coordinate various cellular functions. Enzymes are synthesized in all living cells and consist of polypeptide chains that are built from the 22 natural proteinogenic amino acids. The inner cell production of an enzyme is facilitated through the expression of the enzyme-encoding gene followed by the processing and folding of the polypeptide chain into a functional biocatalyst [20]. The biological activity and selectivity are defined by the unique three-dimensional structure of the enzyme. Compared to organometallic catalysts, the enzyme's structure provides a highly defined

catalytic site that is typically located in a binding pocket, providing efficient substrate orientation by the side residues and enabling high region-, stereo- and chemoselectivity [21]. Following the evolutionary concept, nature created a kingdom of enzymes exhibiting a great variety of macromolecular structures with diverse biological functionalities, reaction kinetics, substrate specificities or selectivities, to regulate all kinds of biochemical reactions and metabolisms. In the chemical industry enzymes with activities suited to catalyze reactions of major interest are increasingly employed in chemical manufacturing or environmental processing [20]. The advances in the biotechnology and protein engineering techniques in the last decades made it possible to obtain and create enzymes fitted for a specific purpose. Furthermore, the recombinant production of enzymes and utilization in large-scale made the use of enzymes in industrial scale economical [18].

Suitable enzymes may be explored in nature, whereby various biological sources are exploited ranging from single microbial cells to multicellular organisms that are derived from diverse environments. Adapted to specific conditions of their habitat or a particular biological task, these organisms provide enzymes with properties that may be tailored for a specific catalytic reaction. Culture-independent methods had been developed and optimized to seek, identify and predict promising enzymes [20]. Another approach to achieve the desired properties is the design and engineering of an enzyme, aiming in the modification of the molecular enzymatic structure to optimize the catalytic properties or reaction specificities for a particular purpose. For that, several methodical tools had been developed comprising a variety of rational design and directed evolution approaches [21]. A more recently established approach is the construction of enzymes with advanced properties and novel functionalities that are beyond nature's possibilities. Hereby, unnatural or synthetic compounds are used to extend the natural palette of building blocks for the enzyme. Several methods are available including the site-specific incorporation of unnatural amino acids (UAAs), the incorporation of non-native cofactors, transition metals or synthetic metal complexes, or the chemical modification of the enzyme surface, e.g., by polymer conjugates. Moreover, enzymes are often

immobilized to make them recoverable or recyclable, hence, to extend the life-period of the biocatalyst [20].

In general, all kinds of hydrolases including lipases, proteases, esterases, epoxide and nitrile hydrolases, and glycosidases, are broadly used in the industrial biocatalysis and cover around 75% of all used industrial enzymes [22]. Hydrolyses catalyze bond cleavage by reaction with water and usually show broad substrate specificity combined with high stereoselectivity. They are used among others in the detergent and food as well as in the pharmaceutical industry. Moreover, oxidoreductases are progressively employed in enantioselective redox-reactions for the synthesis of valuable chiral fine chemical products. Nowadays, about 30% of the biocatalytic processes in industry are oxidations or reductions. Most importantly oxidoreductases are used to produce chiral building blocks for the synthesis of complex therapeutic compounds in the pharmaceutical industry. Other important enzyme classes are dehalogenases that are employed for the selective halogenation or deracemization of halogenated compounds. Furthermore, hydantoinases and *N*-carbamoylases are used in combination for the production of D-p-Hydroxyphenylglycine, an intermediate for the synthesis of cephalosporins or penicillins, as well as for the production of D- or L-amino acids [19,23].

1.2   Introduction to enzymes from the Red Sea brine pools

1.2.1 Extremozymes

Most biocatalysts and products used in biotechnology are derived from mesophilic organisms that thrive in environments with moderate near physiological conditions [24]. However, industrial processes often require conditions comprising extreme temperatures and pH-values, high pressures, low water conditions, increased solvent contents and high concentrations of organic compounds, applied for reasons such as increased solubility of substrates or products, faster reaction rates or influence on enantioselectivity. Most enzymes are unstable under such industrial conditions and show loss in both catalytic activity and stability. Microorganisms that live in extreme environments evolved strategies to not only survive but thrive under harsh conditions, including extremes in

temperatures as found in hot springs, thermal vents, glaciers or the deep sea, extreme acidic or basic pH as found in wastewaters, high salinity from salt lakes or brine pools, or extreme desiccation of the deserts [25][26]. These extremophiles underwent structural and physiological adjustments and their evolved enzymes, called extremozymes, provide advanced properties regarding stability, activity, and specificity [27]. Over the years, numerous extremophiles have been isolated from different types of extremes and various extremozymes have been characterized and further utilized for industrial processes [24]. Major attention was drawn on thermophilic and hyperthermophilic enzymes that work optimally under high temperatures up to 80 °C and above [28]. The successful establishment of thermophilic DNA polymerases, such as Taq and Pfu, in the polymerase chain reaction (PCR) had an immense impact in biotechnology, making the automated PCR possible [29-31]. Several microorganisms coming from extremes had been implemented in processes including the biofuel production, the biomining of insoluble metals, production or processing in the food or cosmetic industry and medical applications [28]. Currently, enzymes that are highly stable and active and outperform those used in industry, such as hydrolases, amylases, cellulases, peptidases, and lipases are under investigation [32].

A major limit of the discovery and employment of extremophiles and their extremozymes is the difficult cultivation under their needed special conditions, making the production of these organisms on larger scales costly and unfeasible, or even impossible. Hence, the recombinant production of extremozymes in well-established mesophilic host organisms that are easy and efficient in production is widely established [24]. However, most of the extremozymes are not compatible with the mesophilic expression systems, as they need the inner-cell environment of their originated organisms for proper protein folding and modification. Therefore, novel extremophilic expression systems are recently developed using extremophilic organisms as a host to obtain high expression of soluble extremophilic enzymes. One example is the recently developed and optimized halophilic *Haloferax volcanii* (*Hfx. volcanii*) expression system, enabling the recombinant large-scale expression of halophilic enzymes of haloarchaea in a cost-effective manner [33,34].

Since the cost and accuracy of high throughput DNA sequencing were improved over the last decades, the exploration of novel extremozymes via the cumbersome sampling and cultivation approach is step by step replaced by culture independent methods. The collection of metagenomic data as well as the collection of genomic data of sampled single cells from remote environments provides a pool of extremozymes to be discovered for industrial applications [25]. The identification of feasible enzymes from genomic data of extremophiles is challenging, as the biochemistry of their unique proteins, metabolites and enzymes is still to be discovered. Therefore, the empirical work on the characterization of specific candidates improves the *in silico* annotation of extremozymes, leading to a broader knowledge of their sequence-structure-function relationship [35]. Finally, structure information expands the understanding of the robustness of extremozymes to a specific environmental parameter and might lead to new engineering techniques that improve mesophilic enzymes.

1.2.1 Halophilic enzymes from deep sea brine pools

The marine environment is the most abundant habitat on earth, offering a palette of niche habitats ranging from tropical sunlit surface waters to ocean trenches 11 km below sea level in the deep sea. The rich variety of microorganisms thriving in our oceans is adapted to various extreme conditions of diverse habitats [36]. The deep-sea anoxic brine pools are one of the most extreme environments known on earth [37]. In this habitat reign conditions like high salt and corresponding high density, elevated temperature, high hydrostatic pressure, heavy metal contents, the absence of light, anoxia and unnatural pH and substrates [38]. The adaptation of marine microorganisms to this harsh environment promoted the natural evolution of extreme hyperhalophilic enzymes that are stable and active under high salt concentrations up to saturation (≤ 5 M salt). Their extraordinary tolerance regarding high ionic strengths/low water conditions makes them additionally resistant against organic solvents, which might be beneficial for catalysis under water/solvent conditions [39]. Structure-wise, the core domains of most haloenzymes are similar to the mesophilic counterparts, containing functional motifs typical for the

specific enzyme class. Hence, the structural adaptation takes mainly place on the protein surface, where the residues are substituted in a way to enhance the stability of the enzyme in the hypersaline environment [40]. Generally, haloenzymes are found to exhibit increased ratios of acidic amino acids on the protein surface, allowing the formation of a protective hydration shell around the enzyme [41,42]. Besides their high stability and activity in low water conditions, most of the halophilic enzymes show additional an increased stability towards elevated temperatures. To date, few halophilic enzymes have been explored for biotechnological applications, although their superior properties are valuable for industrial biocatalysis [43].

One of the current studied anoxic deep-sea brines are the Red Sea brine pools, formed by the tectonic plate movement of the diverging African and Arabic plate. There are approximately 25 deep-sea anoxic brine pools in the Red Sea, all of them bearing several extreme physiochemical parameters, making some of them to the most inhospitable habitats on earth [37]. Recently, high attention was drawn to the discovery of novel extremophilic microbial communities of the brine pools Atlantis II Deep, Discovery Deep, Nereus, Erba and Kebrit [44]. These brine pools differ in depth (1.5 - 2.8 km), temperature (23 - 68 °C), salinity and thickness of brine layer (2.6 - 5.7 M NaCl, 75 - 200 m), pH values (5.5 - 7.9) and heavy metal contents [37,45]. One of the most interesting brine pools is the Atlantis II Deep, exhibiting the highest temperatures among the brine pools up to 68 °C, with a salinity of 25.7% and a pH value of 5.3. Its maximum depth is 2194 km with a total volume 17 - 20 km^2, divided into different layers with stepwise increasing temperature and salt concentration [46]. Atlantis II Deep is enriched in heavy metals, predominantly in copper and zinc, and has high levels of dissolved gases like N_2, methane, CO_2, ethane and H_2S [47]. The Red Sea Research Center provided sequencing data from single amplified genomes (SAGs) derived from water samples of several brine pools [15]. These SAGs were annotated and feed into the INDIGO Red Sea database and are open for exploration (Figure 3) [48]. Using a profile and pattern match algorithm (PPMA), these annotated SAGs were scanned for interesting enzymes with industrial potential [49]. Two genes of predicted alcohol dehydrogenases (ADHs) were identified

from the Atlantis II and the Discovery, the first one described in this work and the later one described in detail by Grotzinger *et al.* [34,50].
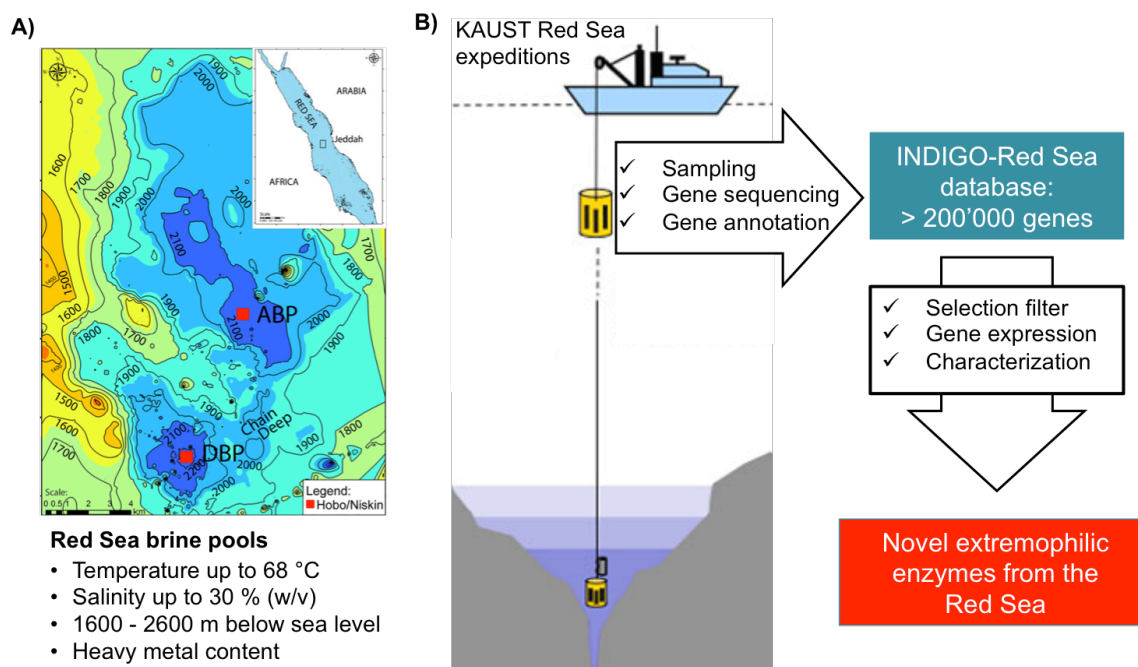
Figure 3: Discovery of novel extremozymes from the Red Sea brine pools. (A) Geographic map of the Atlantis II Deep and Discovery Deep brine pools of the Red Sea (inlet) [51], (B) Workflow from water sample to the extremozyme.

1.2.2 Alcohol dehydrogenases for chemical industry

In the modern synthetic chemistry the production of enantiomeric pure secondary alcohols by the asymmetric hydrogenation of ketones is of high importance. Chiral alcohols are valuable intermediates and building blocks for the production of fine chemicals, chiral pharmaceuticals, and fragrances [52]. The secondary alcohols can be produced either by conventional asymmetric hydrogenation using ruthenium based organometallic catalysts or by the complementary biocatalysis using ADHs [52,53]. The advantages of these biocatalysts, used as isolated enzymes or in whole cells, is their remarkably chemo-, stereo- and enantioselectivity. Furthermore, ADHs perform under mild conditions, are cheap in production and may be immobilized and reused many times

[54]. Naturally, ADHs catalyze the reversible dihydroxylation of primary and secondary alcohols to aldehydes and ketones (Figure 4). For the transfer of the hydride the expensive nicotinamide cofactor NAD(P)$^+$/NAD(P)H is commonly needed, which led to intensive research on the cofactor recycling [55]. Several cofactor regeneration systems were developed comprising chemical, electrochemical, photochemical and enzymatic methods, or utilizing the innate cell recycling mechanism in the whole-cell catalysis approach [54].
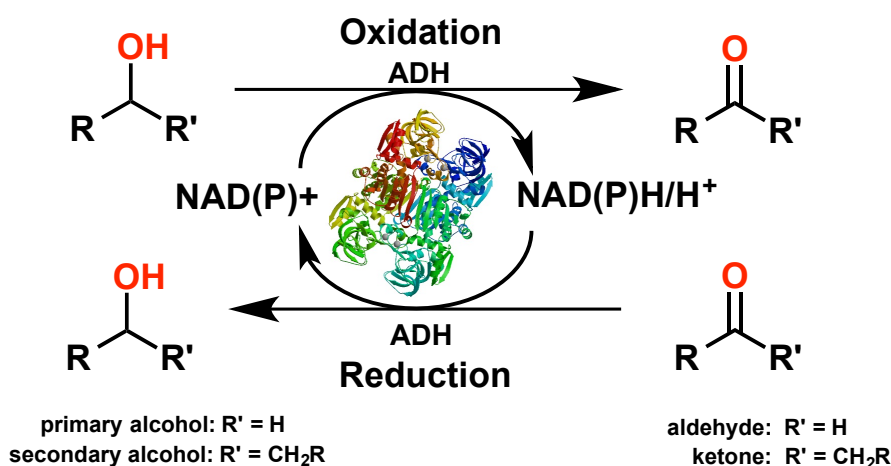
Figure 4: The catalytic cycle of an alcohol dehydrogenase (ADH). The ADH catalyzes the oxidation of alcohols as well as the reduction of aldehydes/ketones by consumption of the cofactors NAD(P)$^+$ and NAD(P)H. Schematic picture of a tetrameric ADH.

Today ADHs are broadly employed in industrial processes with rising numbers, aiming in the production of numerous chiral pharmaceuticals, like psychotropic drugs, anticancer drugs, anti-AIDS drugs, antiasthmatic drugs, and antibiotics [53]. A major issue in applications represents the low solubility of ketones in aqueous solutions; hence, the ADHs performance is limited by the accessibility of the organic compound. Furthermore, parameters that improve the solubility of organic compounds and accelerate reaction rates, as increased temperature and addition of solvents, are challenging for the enzyme's performance and may lead to inactivation [56]. Protein engineering and directed evolution approaches are widely used for the improvement of the performance of ADHs.

Extremophilic ADHs adapted to hostile environments provide naturally advanced stabilities and may, therefore, help to overcome these problems [32,38].

1.2.3 Classification of alcohol dehydrogenases

ADHs belong to the group of oxidoreductases and are classified according to their size and domains into three enzyme superfamilies, the short-chain (SDR), the medium-chain (MDR) and the long-chain dehydrogenase/reductase family (LDR). The Enzyme Commission number (EC) of ADHs is either EC1.1.1.1 for NAD(P)/H usage or EC1.1.1.2 for strict NADP/H usage. The majority of ADHs are oligomeric and form either homodimeric or homotetrameric quaternary structures. Common to all ADHs of the mentioned superfamilies is the presence of the Rossmann-fold motif in the dinucleotide-cofactor binding site [57]. The ADHs of the SDRs consist of about ~250 amino acid residues per subunit and contain a great functional diversity, which is also reflected in their low sequence identity (15 - 30%) among the group [58,59]. The MDR-ADHs consist of about ~350 amino acids and is a more homogenous group with two third of the enzymes sharing a sequence identity of about 90% [58]. The ADHs of the MDRs are zinc-binding metalloenzymes, containing two zinc ions per subunit, one catalytic zinc in the active site and one structural zinc at the dimerization site [60]. The ADHs of the LDR have a size of ~400 residues per subunit and comprise a heterogeneous group of enzymes, including iron-containing ADHs as well as ADHs without metal. So far, the prediction of the structure and function of LDRs is difficult caused by the inhomogeneity of the group [61].

In respect to this thesis, the group of ADHs of the MDRs (MDR-ADHs) will be reviewed in more detail. The first structural resolved ADH of the MDR family was the horse liver ADH1E that was crystallographically determined in 1976 [62]. The mammalian horse-liver MDR-ADH is dimeric that is in contrast to the corresponding tetrameric bacterial, fungal and yeast ADH structures. Structure-wise, the subunit of the MDR-ADHs is divided into two domains, the catalytic domain and the cofactor binding domain that are separated by a long deep cleft [63]. As mentioned above, the cofactor binding domain

consists of a Rossmann-fold that may be found in all kinds of dehydrogenases. The catalytic domain contains an N-terminal β-barrel structural motif surrounded by a few short helices that is called GroES-like domain due to its similarity to the structure group discovered in the chaperon GroES [64]. The second part of the catalytic domain consists of a C-terminal long α-helix and a βαβ-motif that is bound to the main N-terminal part of the domain. The catalytical and structural zinc ions are both bound by the catalytic domain, whereby the catalytical zinc is coordinated by a Cys-His-Cys binding motif and the structural zinc by four cysteine residues [63].

The ADH families in the MDR superfamily had been divided into the dimeric mammalian ADH family, the tetrameric ADH family (TADHs), and the cinnamyl ADH family (CAD). The dimeric ADH family is the most studied ADH family and include the six classes of mammalian ADHs (class I-VI) as the typical liver ADHs, as well as dimeric plant and vertebrate ADHs. The tetrameric TADHs were first discovered in yeast (previously called therefore YADH) and comprise tetrameric ADHs of bacteria, fungi, and non-vertebrates. The CAD family comprises a huge variety of ADHs, ranging from plant CADs to mannitol dehydrogenases. The CAD family was named after the first discovered CADs of plants (EC1.1.1.195) that reduce cinnamyl aldehydes and are important in the plant lignin biosynthesis. But CAD family is not limited to plants as related members of the CADs had been found in *Saccharomyces cerevisiae* (ADH6 and ADH7) as well as in bacteria like in *Escherichia coli* (*E. coli*) and *Mycobacteria*, although the function of these enzymes is still not resolved. However, the number of annotated and characterized ADHs rises continuously and gradually reveals ADHs with overlapping features across the ADH families [63].

1.3 Introduction to the design of a dirhodium artificial metalloenzyme

1.3.1 Artificial metalloenzymes

Homogeneous catalysis shows a broad diversity of catalytic reactions, however despite that it still lacks in several properties like selectivity, recovery, and feasibility in large-scale. On the contrary biocatalysis offers all the properties that homogeneous catalysis is missing like high enantio-, region- and chemoselectivity, cheap production and recyclability. However, biocatalysts are determined by their functionality given by nature [65]. The integration of a transition metal center within a biological scaffold is a rational approach for the generation of a hybrid catalyst, the so-called artificial metalloenzymes (ArMs) (Figure 5). ArMs were first introduced by Wilson and Whitesides [66] and aim to overcome obstacles of the homogeneous transition metal- and biocatalysis [67]. The created ArM should provide extended functionalities and catalyze the non-natural reaction of the homogeneous catalysis with advanced selectivities.
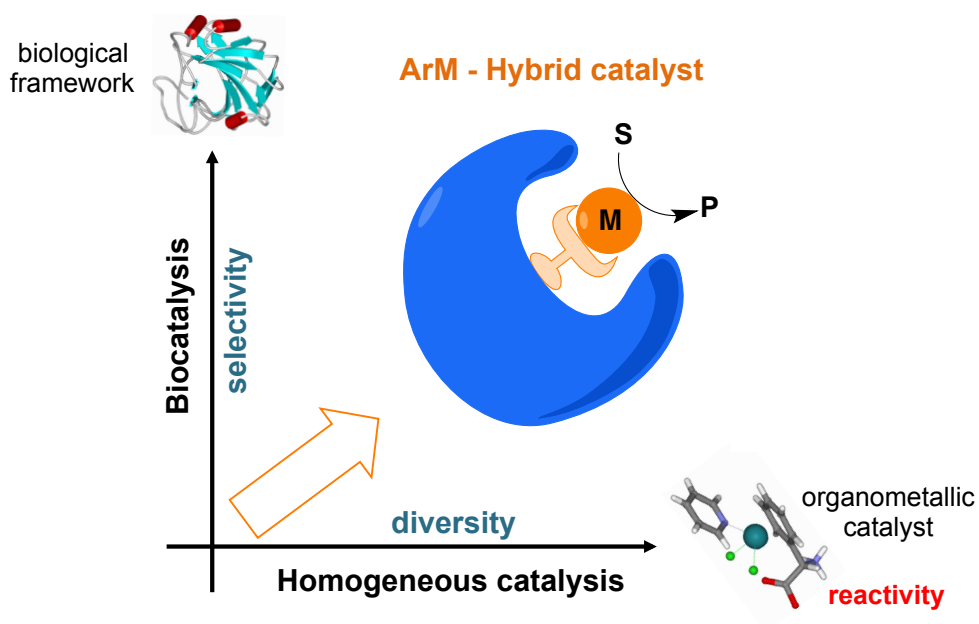
Figure 5: Concept of an artificial metalloenzyme (ArM). A catalytical transition metal from the homogeneous catalysis is linked to a biological scaffold-forming a hybrid catalyst.

The biomacromolecular scaffold of the ArM forms the secondary coordination sphere, a highly defined and functionalized surface that promotes substrate orientation and co-activation. The defined catalytic transition metal or metal complex that is incorporated via a specific anchoring method, determines the reactivity and the substrate activation [68,69]. Ideally, the pairing may result in high selectivity and rate acceleration in the reaction of interest.

The host scaffold can be derived from different groups of macromolecules. Mostly proteins or enzymes are used, but also peptides and nucleic acids such as deoxyribonucleic acid (DNA) have been described [70]. For the choice of the right biological host scaffold, one has to consider several properties that are needed for the establishment of an efficient ArM system. The ideal host scaffold should be well producible, non-toxic and preferably easy to quantify. Furthermore, it should be highly stable towards catalytical conditions including organic solvents and a range of temperatures. Optimally, a rigid structure is naturally provided with a deep cavity for metal incorporation. A reduced affinity for unspecific metal binding is an additional advantage as it minimizes background reactions. Different design strategies have been reported for different kind of host scaffolds (Figure 6).
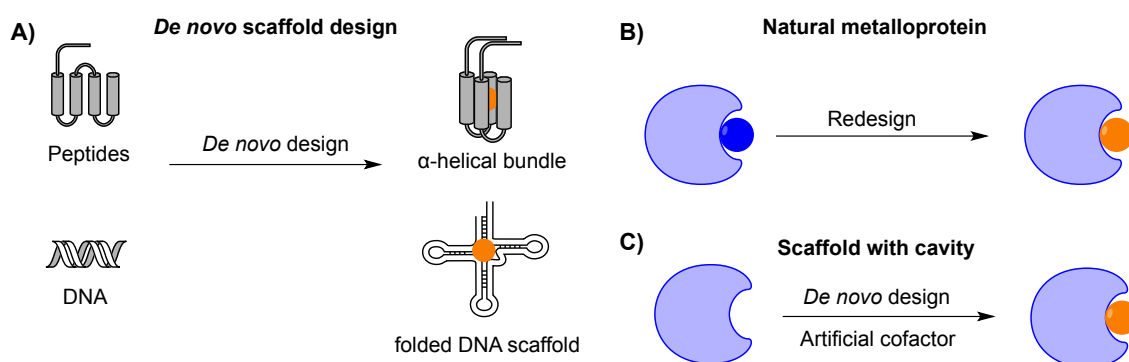
Figure 6: Design of an ArM using different natural biological scaffolds.

One is the *de novo* development of scaffolds by the use of computational science. Hereby, small peptides as well as DNA sequences that fold into well defined tertiary

structures are utilized to bind catalytical metal ions [71]. For example several metallopeptides have been created based on small α-helical peptides assembled into metal binding bundles that catalyzed reactions as the cyclopropanation or transfer hydrogenation with excellent enantiomeric access [69]. Similar to peptides, DNA folds into defined three-dimensional structures that can be utilized for the construction of artificial metalloenzymes. The asymmetric Cu(II)-catalyzed Diels-Alder and Friedl-Craft alkylation were successfully performed beside others, based on DNA scaffolds [72]. Taking into consideration that protein folding is very complex, the *de novo* design of a bigger and correctly folded protein scaffold requires enormous computational resources. Thus the *de novo* design of an artificial metalloenzyme is to date still beyond possibilities [69].

Therefore, scaffolds that already exist in nature represent a much simpler alternative [67]. The natural host scaffold may be a native metal-binding protein with a metal binding site, or a non-metal binding protein that provides a cavity. Regarding metalloenzyme, the metal binding site can be redesigned to integrate catalytic metal ions that differ from the natural ones and to change the reactivity of the enzyme [69]. Moreover, an artificial binding site may be designed *de novo* into a non-metal binding protein by the specific incorporation of coordinating natural amino acids (His, Cys, Glu, Asp, Ser, etc.), also known as 'dative anchoring' strategy (Figure 7A) [73]. The incorporation of unnatural amino acids (UAAs) with metal chelating groups had been utilized for the creation of ArM (Figure 7B). This approach allows the geometrically precise orientation of the moieties that provide the environment to complex the metal ion. The UAA BpyAla, incorporated into the catabolite activator protein serves as an example in this regard and showed successful catalysis in the copper(II)-oxide mediated DNA cleavage [74]

Another highly successful and commonly used strategy is the integration of artificial cofactors in a protein scaffold. This resulted into several innovative ArM platforms that had been utilized in a multitude of catalytic reactions as seen for streptavidin-based ArMs [65]. The artificial cofactor consists of a catalytic organometallic complex and a linked anchoring group that is either a reactive group used for covalent binding or an affinity label for the non-covalent binding to the protein scaffold (Figure 7C/D) [75].
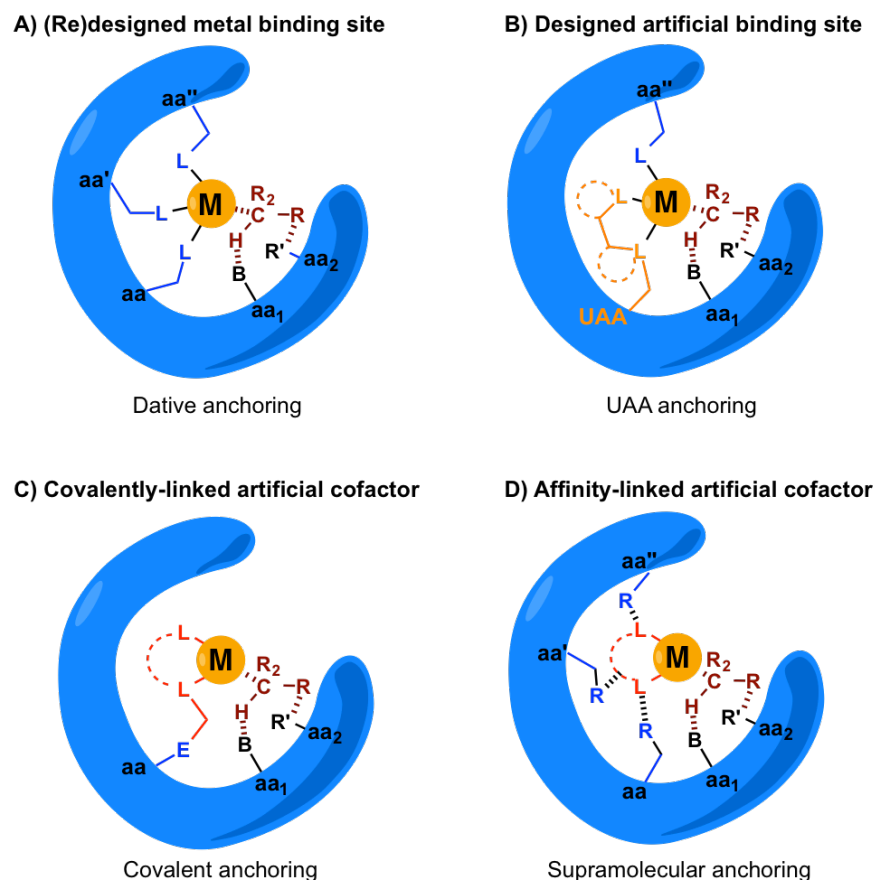
Figure 7: Anchoring strategies of the catalytic metal ion for the design of ArMs. (A) Dative anchoring by the use of natural amino acids. (B) Incorporation of a metal chelating UAA. (C) Use of an artificial cofactor integrated by covalent binding or (D) non-covalent affinity based binding. Coloring: blue = natural coordinating amino acids, black = substrate interacting amino acids, dark red = substrate, orange = unnatural amino acid, red = metal ion coordinating organometallic framework.

Based on that strategy, several homogeneous transition-metal complexes had been integrated into biological host scaffolds, providing a huge variety of hybrid catalysts for various reactions [75]. This so-called chemogenetic approach allows the independent engineering of the protein scaffold and the organometallic-complex and thus offers a much broader range of design opportunities for the generation of ArM with advanced catalytic properties [11,76].

1.3.2 The biotin-streptavidin technology

Streptavidin (SAV) is a protein with an extraordinarily high affinity to biotin and is originally found in *Streptomyces avidinii* (*S. avidinii*). It functions naturally as an antibiotic due to its ability to deplete biotin, also known as vitamin B_7, from the environment [77]. The SAV-biotin binding is one of the strongest non-covalent bonds known in nature with a dissociation constant (K_D) of 10^{-15} M, which makes the binding almost irreversible [78]. The structure of the native SAV with and without bound Biotin was first resolved 1987 [79]. The full-length SAV monomer consists of 159 amino acids with a molecular weight (MW) of ~16.7 kDa. Post-secretory processing in *S. avidinii*, moreover, leads to the 'core' SAV, a truncated version with proteolytic processed N and C termini that is about 32 amino acids shorter [80]. The polypeptide chain of a monomer consists of a β-barrel tertiary structure arranged by eight antiparallel β-strands, forming the biotin-binding cavity (Figure 8A). The quaternary structure consists of two dimers that form a homo-tetramer (Figure 8B), whereby the tetramers tend to aggregate to higher-order oligomeric forms [80]. The SAV structure provides extraordinary high rigidity and stability, maintaining the structure and functionality even under extremely harsh conditions. These conditions include high temperature up to 70 °C, extreme pH-values, high concentrations of denaturing agents like urea and surfactant like sodium dodecyl sulfate (SDS) [81,82]. The high affinity of biotin to SAV results from multiple interactions, including salt-bridges, hydrogen-bonds, and hydrophobic interactions between the biotin and the amino acids residues in the binding pocket (Figure 8A). Due to the biotin-binding, overall structural alterations in the quaternary structure of the tetramer are promoted, that wrap up the subunits more tightly and burry the biotin in the protein interior [83].
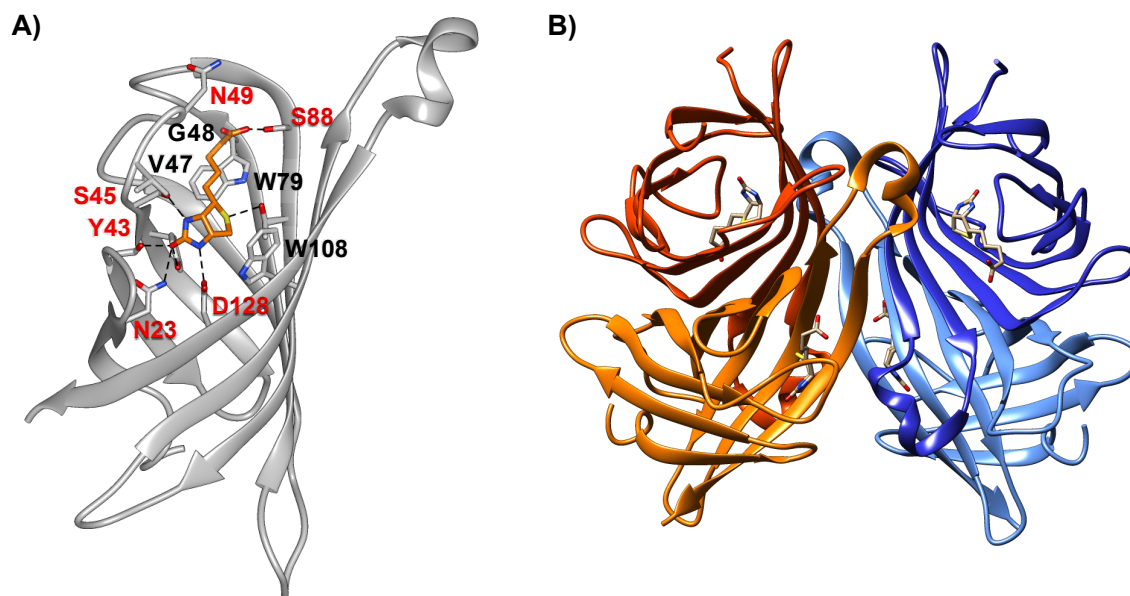
Figure 8: Structure of SAV with integrated biotin. (A) Monomer with bound biotin (orange). Key residues for the biotin-interaction are shown. Red-labeled residues are forming hydrogen bonds, whereas black-labeled residues form hydrophobic interactions. (B) Tetramer formed through to dimer-dimer interaction (orange and blue), holding one biotin per monomer.

Besides the structural and functional advantages, SAV is easily produced in *E. coli* with high yields and subsequently simply purified by affinity chromatography, utilizing its pH-dependent affinity to 2-iminobiotin [77]. Therefore, SAV is utilized for a broad range of biotechnological application, including clinical diagnostics, biotechnology and bioanalytics [84-86].

Based on the previous work of Whitesides [66], Ward successfully introduced the biotin-streptavidin technology for the generation of ArM [87]. Hereby, biotin is linked to a catalytical transition metal-complex through a suitable spacer group, anchoring the catalytic group into the protein cavity of the full-length SAV and creating an artificial catalytical active site (Figure 9). The secondary coordination sphere, build by the exposed loops of the SAV binding pocket, interacts with the substrate and catalyst and enables orientation during the catalytic cycle [88].
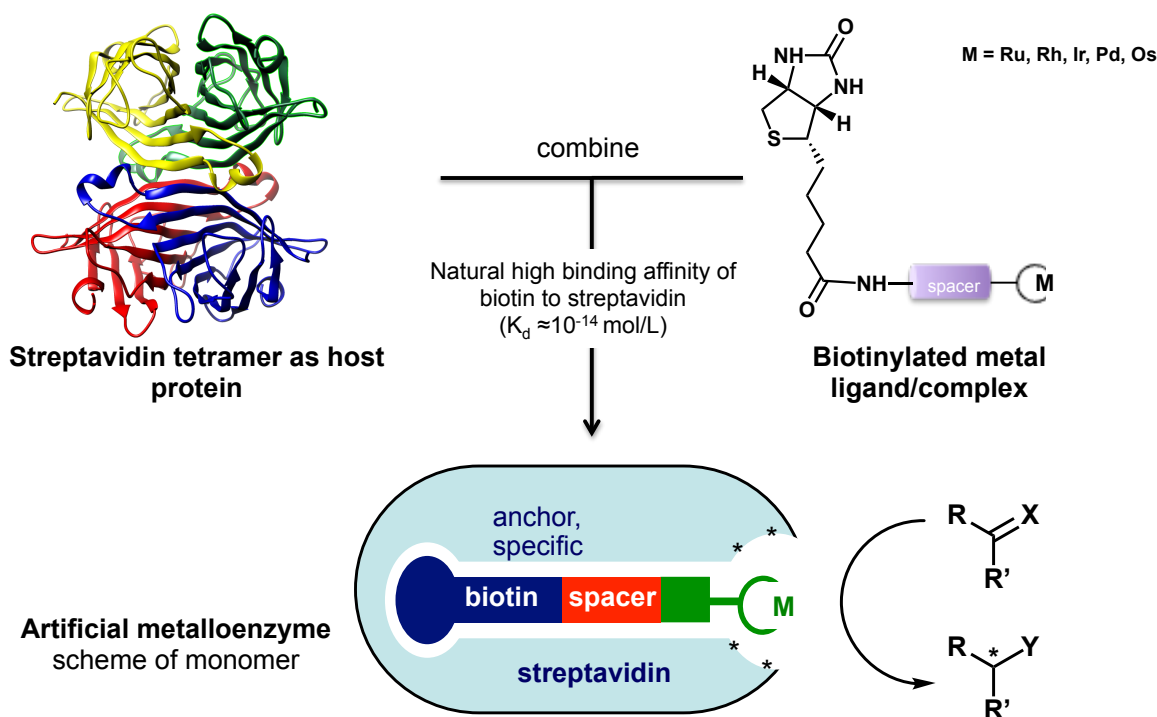
Figure 9: Schema of the biotin-streptavidin technology for the design of ArM. The catalytic active transition metal complex is linked to biotin and integrated to the SAV through supramolecular anchoring based on the natural affinity of biotin into the SAV binding moiety. The secondary coordination sphere of the binding pocket provides a chiral environment during catalysis.

Several ArMs were created based on the biotin-streptavidin technology, proving the system as a versatile tool for the construction of enantioselective catalysts with novel functionalities [89]. The integrated organometallic complexes define the functionality in catalytic reactions, including transition metals as ruthenium, palladium, iridium, rhodium, and osmium (Figure 10). Numerous reactions as hydrogenations, alcohol oxidations, ketone reductions and olefin metathesis have been catalyzed by SAV-based ArM comprise reactions [90-93]. Moreover, asymmetric catalysis was successfully performed including, among others, the asymmetric allylic alkylation, the asymmetric C-H activation, and the enantioselective Suzuki-Miyaura cross-coupling reaction [94-96]. Furthermore, the implementation of an NAD(P)H-dependent artificial transfer hydrogenase in a multienzymatic cascade was reported, enabling cofactor regeneration

and the simultaneous processing of intermediates by further enzymes [97]. Beyond that, the Ward group has demonstrated the possibility to assemble and use ArM in cell lysate as well as in vivo in the periplasm of *E. coli* [98]. Most recently, biotin-streptavidin based ArMs were implemented into living mammalian cells, catalyzing the uncaging of a hormone that triggers protein expression [99]. Uncovering the enormous potential of this ArM platform, continuous studies might lead to advanced applications in both medical chemistry and synthetic biology in the future [100].
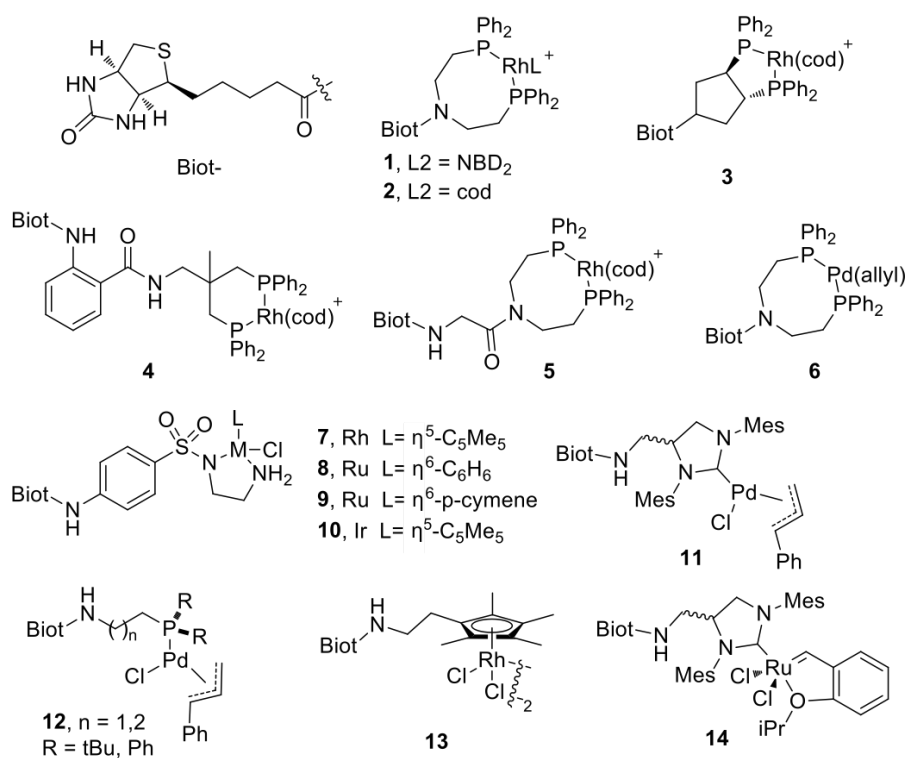
Figure 10: Supramolecular anchoring using biotin-linked transition metal complexes for the generation of SAV-based ArMs [101].

1.3.3 Dirhodium-carboxylate complexes in asymmetric catalysis

Dirhodium(II) carboxylates are highly efficient catalysts for diverse reactions and had been extensively studied and developed over the last few decades. The dirhodium catalyzed Merck synthesis developed in the 1980s for the production of the antibiotic thienamycin added considerable interest to further developments in this area [102]. Structure-wise, the unique features of paddlewheel structured dirhodium(II) carboxylates are their rigidity, the easy exchange of ligands, the two open diaxial coordination sites for Lewis bases, and their low oxidation potential (Figure 11) [103]. Based on their highly efficient operation under physiological conditions, several applications were found in the medical and chemical biology.
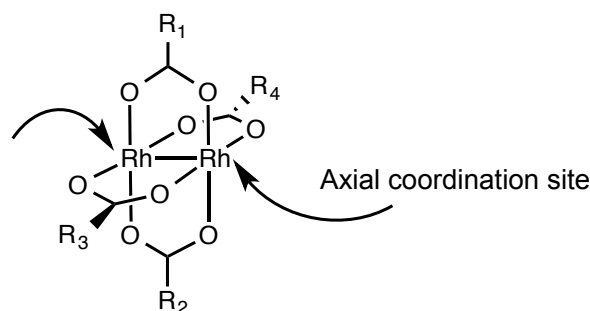
Figure 11: Paddlewheel structure of dirhodium(II) carboxylate complexes with two axial coordination sites.

In the homogeneous catalysis dirhodium complexes had been exploited in various reactions, showing exceptional activity and enantioselectivity for carbene-transfer reactions with diazocarbonyl compounds including the intra- and intermolecular cyclopropanation, the cyclopropenation, X-H insertions and Ylide transformation (Figure 12). Recently, the amino-oxyarylation and the aziridination of olefins was reported with the Du Bois catalyst ($Rh_2(esp)_2$); however, no selectivity was obtained yet [104]. The limitation of dirhodium catalysts lies in the high costs of the precious rhodium metal. Nevertheless, the reported high-turnover numbers of the catalyst make the use of tiny amounts necessary to generate bulk quantities of valuable products.
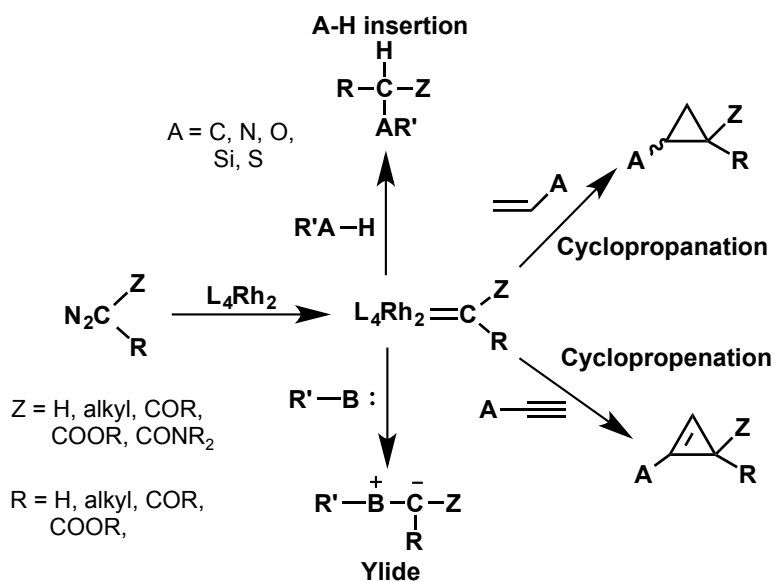
Figure 12: Application of dirhodium complexes (L$_4$Rh$_2$) in carbene transfer reactions with diazocarbonyl compounds.

In the biological chemistry, dirhodium complexes were utilized for the modification of DNA and the metal uptake within tumor cells [105,106]. The successful integration of a dirhodium(II) tetracarboxylate complex into a biological scaffold was first pioneered by Ball and coworkers. A dirhodium metallopeptide was developed that found application for protein modification and intracellular imagine [107]. Integration of a dirhodium carboxylate into a prolyl oligopeptidase (POP) scaffold generated a dirhodium ArM that efficiently catalyzed the asymmetric cyclopropanation with enantioselectivity and reduced levels of byproducts [108]. In figure 13 the reaction mechanism of the cyclopropanation over the formation of a carbenoid is shown. Chiral cyclopropanes are important building blocks for the production of valuable complex molecules with defined orientation of functional groups. Furthermore, they exhibit a large spectrum of biological functions including enzyme inhibition, insecticidal, herbicidal, antibacterial, antifungal, antiviral, and antitumor activities [103]. Following from that, the continuous development of highly selective and 'green' ArM with long life-time or recyclability is of major interest.

23
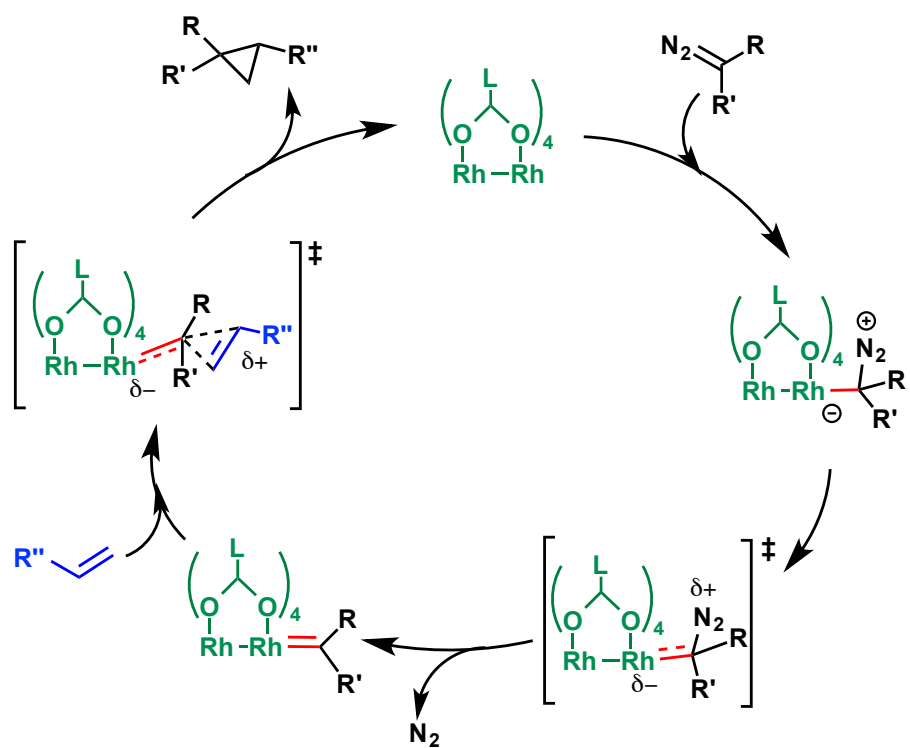
Figure 13: Reaction pathway for carbenoid formation and cyclopropanation [109].

2 Objectives

Novel biocatalysts are highly demanded in industry. Hence, the discovery and design of highly stable and enantioselective biocatalysts with novel functionalities is an ongoing research topic. In this thesis two approaches are followed to establish superior enzymes for industrial biocatalysis. In the first approach, nature's broad enzyme repertoire was utilized for the discovery of a polyextremophilic alcohol dehydrogenase that is highly stable and functional under industrial conditions. In the second approach, nature's portfolio of functionalities was extended by the design of a dirhodium artificial metalloenzyme that combines the functionalities of the homogeneous dirhodium catalyst with the enantioselectivity provided by the biological host scaffold.

Alcohol dehydrogenases (ADH) are implemented as biocatalysts in the production of chiral building blocks for pharmaceuticals or fine chemicals; hence, the continuous discovery of ADHs with advanced stabilities and extended substrate scopes is highly demanded. The Red Sea brine pools are one of the most remote environments on earth inhabiting extremophilic microorganisms with extremozymes that are not only stable but also active under harsh conditions. Here, a promising gene of a zinc-dependent ADH, named ADH/A1a, was identified of the Atlantis II Deep brine pool and chose it for extended characterization and evaluation. Thus, the aim of this thesis was to first implement a sufficient production and purification procedure for the ADH/A1a to gain a highly purified and functional enzyme by utilization of the available halophilic expression system *Hfx. volcanii*. The purified ADH/A1a should be characterized in terms of enzymatic functionality and stability with the special focus on the reaction optima and the substrate spectrum. The enantioselective formation of secondary chiral alcohols is of major interest; hence, the conversion to potential valuable products was investigated. The structure analysis of ADH/A1a should reveal unique features of the structure adaptation that explain the stability and activity of the polyextremophilic ADH in extreme

conditions. Moreover, the comparison of ADH/A1a with homologs should reveal the phylogenetic linkage to known ADHs and may uncover its metabolic function.

The objective in the creation of artificial metalloenzymes (ArM) is the extension of the natural functionality of enzymes by the incorporation of promising catalytical transition-metal complexes into enantioselective biological host scaffolds. Dirhodium(II) tetracarboxylate complexes have shown to be stable and active in the catalysis of carbene transfer reactions under physiological conditions, thus, are suitable as a catalytical unit in ArM. The aim of this work was the introduction of a biotin-linked dirhodium complex into a streptavidin (SAV) scaffold through supramolecular anchoring, forming a enantioselective hybrid-catalyst that may be functionalized for carbine transfer reactions. The synthesis of a suitable biotin-dirhodium complex should be led by docking studies of the ligand into the binding pocket of SAV. Furthermore, a small rational designed SAV mutant library should provide variants with hydrophobic binding pockets, providing a chiral reaction environment and enabling the co-activation of the organic substrates. The catalytic activity and enantioslectivity of the created SAV-dirhodium catalysts should be explored in the selective cyclopropanation of olefins, aiming to provide high yields and enantioselectivity for the desired cyclopropane product.

In summary, the objective was to contribute to the development of novel biocatalysts for industrial purposes. The discovery of polyextremophilic enzymes like the ADH/A1a provides novel structural and functional insights, which may be utilized for biocatalytic applications in future. On the other hand, the design of enantioselective ArM with novel catalytic functionalities, as the SAV-dirhodium hybrid catalyst, adds to the understanding of enzyme design for specific purposes and further broadens the scope of 'green' chemistry.

3 Materials and Methods

3.1 Materials

3.1.1 Chemicals

Unless stated otherwise, all chemicals were obtained from the following companies: AppliChem (Darmstadt, GER), Thermo Fisher Scientific (Waltham, USA), Fluka (Neu-Ulm, GER), Merck (Darmstadt, GER), Sigma-Aldrich (Steinheim, GER), Roth (Karlsruhe, GER), and VWR (Darmstadt, GER).

3.1.2 Media and solutions

Media used for *E. coli*

LB medium:

Peptone	1% (w/v)
Yeast extract	0.5% (w/v)
NaCl	0.5% (w/v)
For solid media: Agar	2% (w/v)

MTP medium:

Trypton	2% (w/v)
Yeast extract	1.5% (w/v)
NaCl	0.8% (w/v)
For solid media: Agar	2% (w/v)
Mix after autoclave:	
Na_2HPO_4	0.22% (w/v)
KH_2PO_4	0.1% (w/v)

Selective antibiotic:

Ampicillin	100 mg ml^{-1} (in dH_2O)
Kanamycin	50 mg ml^{-1} (in dH_2O)
Chloramphenicol	34 mg ml^{-1} (in 70% EtOH)

SOC medium:

Peptone	2% (w/v)
Yeast extract	0.5% (w/v)
$MgSO_4$	10 mM

NaCl	10 mM
KCl	2.5 mM
Glucose	20 mM

Media used for *Hfx. volcanii*

Hv-YPC medium (16.75% w/v salt):

Casamino acids	0.1% (w/v)
Peptone	0.1% (w/v)
Yeast extract	0.5% (w/v)
NaCl	14.4% (w/v)
$MgCl_2$	88.50 mM
KCl	37.84 mM
For solid media: Agar	2% (w/v)
Mix after autoclave:	
$CaCl_2$	3 mM
$MgSO_4$	85 mM
Tris-HCl pH 7.5	20 mM

Additional:	
Tymidine	40 ng L^{-1}

Induction solution (18% w/v salt):

Salt water (28% w/v)	60% (v/v)
L-Tryptophan	27 mM

Salt water (28% w/v salt):

NaCl	24% (w/v)
$MgCl_2$	1.4% (w/v)
$MgSO_4$	1.7% (w/v)
KCl	0.7 (w/v)
Tris-HCl pH 7.5	20 mM

Solutions used for the transformation of *Hfx. volcanii*

Buffered sphaeroplast solution [A]:

NaCl	1 M
KCl	27 mM
Saccharose	15% (w/v)
Tris-HCl pH 8.5	50 mM

Non-buffered sphaeroplast solution [B]:

NaCl	1 M
KCl	27 mM
Saccharose	15% (w/v)
1 M NaOH	adjust to pH 7.5

Spheroplast dilution solution [C]:

Salt water (28% w/v)	21% (v/v)
CaCl$_2$ (0.5 M stock)	3.75 mM
Saccharose	15% (w/v)

Regeneration solution [D]:

Salt water (28% w/v)	60% (v/v)
10 x YPC	10% (w/v)
Saccharose	15% (w/v)
CaCl$_2$ (0.5 M stock)	3 mM

Transformation dilution solution [E]:

Salt water	60% (v/v)
Saccharose	15% (w/v)
CaCl$_2$ (0.5 M stock)	3 mM

EDTA solution [F] (0.5 M):

EDTA	14.6% (w/v)
NaOH (10 M stock)	adjust to pH 8.0

10 x YPC solution:

Yeast extract	5% (w/v)
peptone	1% (w/v)
Casamino acids	1% (w/v)
1 M KOH	Adjust to pH 7.5

Buffers used for protein purification

Buffer A/B:

NaCl	11.7% (w/v)
HEPES	20 mM
Imidazol	10 mM (A) /300 mM (B)
Glycerol	10% (v/v)
1 M NaOH	adjust to pH 7.5

Dialysis buffer:

NaCl	11.7% (w/v)
HEPES	20 mM
Glycerol	10% (v/v)
1 M NaOH	adjust to pH 7.5

Storage solution:

NaCl	11.7% (w/v)
HEPES	20 mM
Glycerol	10% (v/v)
1 M NaOH	adjust to pH 7.5

Buffers used for activity measurements:

Reaction buffer (oxidation)

KCl	3 M
Glycine	50 mM
1 M NaOH	adjust to pH 10.0

Reaction buffer (reduction)

KCl	3 M
K_2HPO_4	50 mM
1 M citric acid	adjust to pH 6.0

3.1.3 Enzymes

*Bam*HI	New England Biolabs
*Nde*I	New England Biolabs
Phusion® High Fidelity DNA Polymerase	New England Biolabs
T4 DNA Ligase	New England Biolabs
Trypsin	Sigma-Aldrich

3.1.4 Primers

All oligonucleotides were dissolved in dH_2O to a final concentration of 100 pmol μl^{-1}. The working concentration of the primers was 10 pmol μl^{-1}. All primers were synthesized by Integrated DNA Technologies (Skokie, USA). All oligonucleotides of this work are listed in Table S1

3.1.5 Strains

Strain	Genotype	Supplier
E. coli cloning strains:		
E. coli One Shot TOP10	F$^-$ mcrA Δ(mrr- hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu) 7697 galU galK rpsL (Str^r) endA1 nupG λ-	Thermo Fisher Scientific - Invitrogen (Waltman, USA)
E. coli BL21(DE3) Gold	F$^-$ ompT hsdS(rB$^-$mB$^-$) dcm$^+$ gal λ(DE3) endA Hte	Stratagene (La Jolla, CA, USA)
E. coli XL1-Blue	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F$^-$ proAB lacI^{q}ZΔM15 Tn10 (Tet^r)]	Agilent Technologies (Santa Clara, USA)
E. coli expression strains:		
E. coli BL21-CodonPlus (DE3)-RIPL	F ompT hsdS(rB$^-$mB$^-$) dcm$^+$ Tet^r gal λ(DE3) endA Hte [argU proL Cam^r] [argU ileY leuW Strep/Spec^r]	Stratagene (La Jolla, CA, USA)
E. coli BL21(DE3) pLysS	F$^-$, *ompT*, hsdS(rB$^-$mB$^-$), dcm$^+$, *gal*, λ(DE3), pLysS, Cm^r	Thermo Fisher Scientific-Invitrogen (Waltman, MA, USA)
Hfx. volcanii strains:		
Hfx. volcanii H1424	ΔpyrE2 ΔhdrB Δmrr Nph-pitA cdc48d-Ct	Dr. T. Allers, University of Nottingham [33]
Hfx. volcanii H1895	ΔpyrE2 ΔhdrB Δmrr Nph-pitA cdc48d-Ct ΔpilB3C3	Dr. S. Groetzinger, Dr. E. Strillinger, TUM [34]

3.1.6 Plasmids

Vector name	Description	Supplier/Reference
pET-3a	Expression vector for *E. coli*, AmpR	Addgene, Cambridge
pTA963	Expression vector for *Hfx. volcanii*, containing a N-terminal his$_6$-tag, AmpR	Dr. T. Allers, University of Nottingam [33]
pET-3a-*adh/a1*	Expression of *adh/a1* in *E. coli*	Dr. Groetzinger, TUM
pTA963-*adh/a1*	Expression of *adh/a1* in *Hfx. volcanii*	Dr. Groetzinger, TUM [34]
pTA963-*adh/a1a*	Expression of *adh/a1a* in	In this work

	Hfx. volcanii	
pET11_T7_SAVwt	Expression of SAVwt	T. Ward, University of Basel
pET11_T7_SAVxx	Expression of SAV mutants (Supplementary Table S4)	In this work

3.1.7 Instruments

Equipment	*Modelle*	*Supplier*
Balance	LSM 200	PCE Group, Southampton Hampshire, UK
Centrifuge	Varifuge 3.0 R	Heraeus Sepatech
Centrifuge	Cool-Centrifuge 5417R	Eppendorf
Centrifuge	SIGMA 3-30K	SIGMA
Chromatography	prime plus	GE Healthcare
Cell disrupter	Constant Cell Disrupter System E1061	Constant Systems, Northants, GBR
Electroporation	PowerPac Universal Power Supply	BioRad, Hercules, USA
GC-MS	GC-MSD system 5975C, 7890A GC system	Agilent, Santa Clara, USA
Gel documentation	Gel Doc EZ system	Bio-Rad, Hercules, USA
Gel documentation	Bio-Rad Power Pac HV system	Bio-Rad , Hercules, USA
HPLC	1260 Infinity II LC system	Agilent Technologies
HPLC	UltiMate™ 3000 UHPLC	ThermoFisher
Imager	Versa Doc 4000MP	Bio-Rad, Hercules, USA
Mass spectrometer	microflex	Bruker
Mass spectrometer	Q-Exactive HF	Thermo Fisher Scientific
Microplate reader	TECAN INFINITE M1000	Tecan, ZürichSwitzerland
Microscope	DMi8	Leica
Pipetting robot	Oryx4 Crystallization Robot	Douglas Instruments, Newmarket, GBR
Pipetting robot	Phoenic Liquid Handling System	Art Robbins Instrument, Sunnyvale, USA
Plate reader	Infinite M1000	Tecan, Zurich, Switzerland
Peltier system	PCB 1500	DBS Vigonza, Italy
Rotary evaporator	WG-EV311-V-Plus Rotary evaporator	Wilmad-LabGlass, Vineland, NJ

Sonifier	Branson Digital Sonifier 250	G. Heinemann (Schwäbisch Gmünd, GER)
Spectrophotometer	NanoDrop 2000c	Thermo Scientific,
Spectrophotometer	Ultrospec 7000	GE healthcare, Fairfeld, USA
Spectropolarimeter	JASCO J-815	Jasco Applied Sciences, Nova Scotia, Canada
Thermomixer	Thermomixer Compact	Eppendorf
Vortexer	MS2 Minishaker	IKA Works Inc.

3.1.8 Softwares

ApE, A plasmid Editor v2.0	M. Wayne Davis
ChemBioDraw 14.0	Perkin Elmer (Cambridge, USA)
Endnote X7	Adept Scientific (Frankfurt, GER)
Graph Pad Prism 7	GraphPad Software Inc. (La Jolla, USA)
ImageJ 1.51	Wayne Rasband, NIH, USA
Jalview	Waterhouse et al., 2009
Microsoft Office for Mac 2011	Microsoft (Redmond, USA)
PrimeView 5.31	GE Healthcare (Chalfont St. Giles, GBR)
MacPyMOL 1.7	DeLano Scientific LLC, Schroedinger
MASCOT v2.3	Matrix Sciences Ltd, UK
SnapGene viewer 2.3.3	GSL Viotech LLC (Chicago, USA)
Spectra Manager v2.08.01	JASCO Applied Sciences (Nova Scotia, Canada)
UCSF Chimera 1.12	University of California
Yasara 14.7.17	Yasara Biosciences GmbH, Ellmar Krieger
XQuartz 2.7.11	Apple Inc.

3.2   Methods

3.2.1 Agarose gel electrophoresis

Agarose gel electrophoresis was used for the analytical and preparative separation of DNA fragments. For that purpose, DNA samples were mixed with DNA loading dye and loaded on 0.5 - 1% (w/v) agarose gels along with a DNA ladder (peqGOLD DNA Ladder Mix 100-10,000 bp, Peqlab, Erlangen, Germany). Electrophoresis was carried out in Tris-acetate-EDTA (TAE) buffer for 45 min at a constant voltage of 120 V. Gels were stained

for 10 min in a 100 ml staining solution containing 10 µl of SYBR Green I solution (Sigma Aldrich, St Louis, USA) in TAE-buffer. The DNA bands were visualized by UV light (254 nm) using the gel documentation system Imaging Versa Doc 4000MP System. For preparation DNA bands were excised from agarose gel using a UV- table and extracted using the peqGOLD Gel Extraction Kit.

3.2.2 Isolation and purification of DNA

For plasmid isolation from recombinant *E. coli* cultures the peqGOLD Plasmid Miniprep Kits I or II (Peqlab, Erlangen, GER) were used. DNA fragments from PCR or other reaction mixes were purified by the peqGOLD Cycle-Pure Kit (Peqlab, Erlangen, GER). Purified DNA was eluted in 25 µl dH$_2$O and stored at -20 °C. Instructions were followed according to manufacturer's manual.

3.2.3 Polymerase chain reaction

The polymerase chain reaction (PCR) was used to amplify the genes of interest. The reaction mixture with a total volume of 50 µl contained the following reagents and protocols from NEB (Ipswich, USA):

Quantity	Reagent
1.0 µl	Template DNA (50-100 ng µl^{-1})
0.5 µl	Phusion H.F. Polymerase (2 U µl^{-1})
2.5 µl	Forward Primer (10 pmol µl^{-1})
2.5 µl	Reverse Primer (10 pmol µl^{-1})
10 µl	5 x Phusion GC Buffer
1.5 µl	DMSO
1.0 µl	dNTP (10 mM)
32 µl	dH$_2$O

The reaction was carried out in a thermal cycler using a protocol designed after manufactures manual or using the Touchdown PCR protocol designed by the Optimase ProtocolWriter (Transgenomic Inc.). PCR products were analyzed by agarose gel electrophoresis.

3.2.4 DNA restriction digest and DNA ligation

Preparative restriction digest of DNA was done after the manufacturers manual. Double digestion with *NdeI* and *BamHI* was carried out in Cut-Smart Buffer (NEB) at 37 °C for 2 h. The restriction products were analyzed by agarose gel electrophoresis and purified as stated before.

DNA ligation was carried out in a total volume of 20 µl, filled up with dH_2O. In general, the insert was used in threefold excess compared to the vector. The reaction mixture was heated to 55 °C for 10 min to unwind the DNA and subsequently cooled on ice for 5 min. 2 µl T4 DNA ligation buffer (5x) and 0.5 µl T4 DNA ligase (1 U µl^{-1}) were added. The mixture was incubated over night at 16 °C. The ligation product was diluted 1:10 with dH_2O before electro-transformation.

Cloning strategies were developed with the software ApE (A plasmid Editor).

3.2.5 Transformation of *E. coli* by electroporation

Electrocompetent *E. coli* cells were used from transformation using electroporation. Competent cells (20 µl) were mixed with 0.5 - 1 µl of plasmid DNA and transferred into a 1 mm electroporation cuvette. An electric pulse (5.4 - 5.7 ms and 2.5 kV) was applied to the cuvette by an electroporation device. Cells were resuspended in 1 ml of SOC medium and incubated under shaking for 1 h at 37 °C. Finally, the 10 µl - 200 µl of the bacterial culture was streaked on agar plates containing the respective antibiotics and incubated over night at 37 °C.

3.2.6 Cultivation and storage of cells

For long-term storage, the cultures were centrifuged for 10 min at 5,000 x *g* and the resulting pellet was resuspended in 1 ml culture medium containing 30 % (v/v) glycerol. The cell suspension was transferred into a cryo tube and stored in liquid nitrogen.

3.2.7 Gibson cloning

Gibson cloning was done using the Gibson Assembly® Cloning Kit (NEB, Ipswich, USA) after manufacturers manual. Gibson cloning oligonucleotides were designed using the NEBuilder Assembly Tool provided by NEB (Ipswich, USA) and synthesized by Eurofins MWG Operon (Ebersberg, Germany).

3.2.8 DNA and protein concentration

DNA or protein concentrations were measured according to Lambert-Beer's rule with a NanoDrop2000c spectrophotometer at 260 nm or 280 nm, respectively. For proteins, the specific extinction coefficient ε and the molecular weight were calculated with the software Protparam (http://web.expasy.org/protparam/)[110].

3.2.9 Site-directed mutagenesis

Mutations in gene sequences were introduced *via* the QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies) after manufacturers manual. The used oligonucleotides were designed using the Agilent QuikChange Primer Design Program (https://www.genomics.agilent.com/ primerDesignProgram.jsp) and synthesized by Eurofins MWG Operon (Ebersberg, Germany).

3.2.10 DNA sequencing

DNA sequence analysis was performed using Sanger Sequencing by GATC Biotech AG, Konstanz or by the KAUST core facilities. The obtained results were compared to the expected DNA sequence using the ApE software.

3.2.11 SDS polyacrylamide gel electrophorese (SDS-PAGE)

The discontinuous SDS-PAGE was performed with Mini-PROTEAN TGX stain-free precast gels 4 - 20% (Bio-Rad Laboratories Inc., Hercules, USA). Gels were run in 1x running buffer (25 mM Tris-HCl, 192 mM glycine, 0.1% SDS, pH 8.3) at 35 mA per gel until the front line has passed the bottom of the gel (~30 min). The gel were visualized

stain-free using the Gel Doc EZ system (5 min incubation) or Coomassie-stained for 1 h under shaking using the QC Colloidal Coomassie Stain solution (Bio-Rad Laboratories Inc., Hercules, USA). Before the SDS-PAGE analysis, the samples were diluted 1:1 with 2x Laemmli sample buffer (Bio-Rad Laboratories Inc., Hercules, USA) and incubated at 95 °C for 10 min. As protein ladder the PageRuler™ Plus Prestained Protein Ladder, 10 to 250 kDa (Fisher Scientific, Hampton, USA)

3.2.12 Mass spectrometry analysis

Protein mass spectrometry was carried out on a BRUKER maXis HD ESI-TOF. For ESI-TOF 0.04 mM of pure protein sample was solved or dialyzed against H_2O and desalted with a zip tip 0.6 µl C_4 resin (Merck Millipore, Burlington, USA) according to manufacturer's manual and eluted in 30% ACN, 0.1% TFA. The protein sample was either directly injected or run over an HPLC (Agilent Technologies), C4 Column (Column volume 5 ml) before injection. The elution occurred with a flow rate of 0.5 µl min^{-1} and a gradient to 80% ACN, 0.1% FA. Subsequently fractions were recorded according to the standard protein medium mass analysis method of Bruker.

Further measurements were carried out on a BRUKER microflex LRF system. For MALDI-TOF measurements a pure 20 µl protein sample of 2 mg ml^{-1} in H_2O was desalted with a zip tip 0.6 µl C_4 resin (Merck Millipore, Burlington, USA) according to manufacturers manual and eluted in 30% ACN, 0.1% TFA. The protein sample was mixed 1:1 with a saturated α-Cyano-4-hydroxycinnamic acid matrix in 30% ACN and 0.1% TFA. From the mixture 1 µl was spotted on a MALDI-target, air-dried and subjected for analysis using the standard protein mass detection method of BRUKER. Before each measurements, the mass detector was calibrated using protein calibration standard II (BRUKER, Billerica, USA).

3.2.13 Protein crystallization

For crystallization, purified protein samples were concentrated to a concentration of 5 mg ml^{-1} (ADH/A1a) and 20 mg ml^{-1} (SAV) using an Amicon Ultracel® 30,000 MWCO (Millipore, Billerica, USA). All solutions were centrifuged at 14000 x g for 5 min before

crystallization. The protein purity (> 90 %) was checked by SDS-PAGE. Protein crystals were grown using the sitting drop vapor diffusion method, 96-Well Intelli-Plates, and the Phoenix or Oryx robots for pipetting. For initial screens, standard kits 1-12 from Qiagen were used. The diffraction data were collected at the beamline X06SA at the Paul Scherrer Institute, SLS, Villigen, Switzerland.

ADH/A1a: The crystallization screens were incubated at 20 °C and incubated for up to two weeks. The drops contained 0.2 - 0.3 µl of protein solution (5 mg ml^{-1}, in 10 mM HEPES, pH 7.5, and 0.2/ 0.5/ 1 M NaCl) and 0.2 µl reservoir solution of the screening kits.

SAV: The crystals were grown at 20 °C and incubated for up to two weeks. The drops contained 0.2 - 0.3 µl of protein solution (~20 mg ml^{-1}, in dH$_2$O) and 0.2 µl reservoir solution of the standard kits 3, 4, 8 and 9. The grown crystals of SAV apo were used for soaking experiment with the Biot-Ar(CO)$_2$(OAc)$_2$ ligand (stock: 20 mg ml^{-1} in DMSO). For that 0.1 µl of the diluted ligand (1:10 in reservoir solution) was added to the crystal and incubated for up to 30 min. The crystal was cryopotected in a suitable cryo buffer containing either 25% PEG200/400 or 30% glycerol and subsequently vitrified in liquid nitrogen. The SAV apo crystals were further utilized for seeding experiments for the co-crystalization screens of the SAV-diRh complex. For the co-crystallization the SAV (5 mg ml^{-1}, in dH$_2$O) was incubated 1:1 with the Biot-Ar(CO)$_2$(OAc)$_2$ ligand (stock: 20 mg ml^{-1} in DMSO), dialyzed against dH2O and concentrated to 20 mg ml^{-1} before the crystallization screens.

3.3 Detailed methods for ADH/A1a

3.3.1 Source, annotation and cloning of *adh/a1a*

The gene of the zinc-dependent ADH, *adh/a1,* was annotated from a SAG (AAA261F19) of an unclassified MSBL1 archaeon, sampled from the Red Sea Atlantis II Deep brine pool interphase (2036 m, 57 - 63 °C, 15.1 - 16.8%, pH 5.6) [44]. *Adh/a1* was identified and chosen from the INDIGO Red Sea database for characterization using the profile and

pattern matching algorithm (PPMA) [48,49]. The gene *adh/a1* was codon optimized by JCat using the codon usage of Halobacterium NRC-1, which is similar to the one of *Hfx. volcanii* [50]. The optimized gene was synthesized by Geneart and cloned into the expression plasmid pTA963 previously.

In this work, the gene *adh/a1a* was constructed using two PCR steps and Gibson cloning in order to delete a 20 amino acid long N-terminal nonsense region. In the first PCR the gene *adh/a1a* was amplified from the plasmid using the oligonucleotides P36/P38, excludes the nonsense DNA part of the gene. In the second PCR, the construct was amplified using two Gibson oligonucleotides (P39/P43) in a touchdown PCR. The plasmid backbone was generated by a restriction digest of pTA963 with *NdeI* and *BamHI* and purified by gel extraction. Both purified backbone and insert were fused together in the Gibson assembly after manufacturers protocol. The resulting construct pTA963-*adh/a1a*, was purified and transformed into *E. coli* BL21 (DE3) Gold (Stratagene, Agilent Technologies). The confirmed construct was transformed into the biofilm free halophilic expression strain *Hfx. volcanii* H1895 after the expression protocol of the halohandbook [111]. The *Hfx. volcanii* strain H1895 and the expression plasmid pTA963, were obtained from Thorsten Allers from the University of Nottingham, UK and Eva Strillinger, Technische Universität München, Germany [33,34]. The correctly annotated gene, named *adh/a1a*, was deposited into the NCBI's Genbank by Mwirichia *et al.* (Genbank: KXB02677) [44].

3.3.2 Recombinant ADH/A1a production in *Hfx. volcanii*

The overexpression of ADH/A1a in the shaking flask and the bioreactor was done after Strillinger *et al.* [34]. The recombinant *Hfx. volcanii* strain containing the expression plasmid was grown in baffled 5L-shaking flasks containing 1.8 L of Hv-YPC medium. The cultures were inoculated with an OD_{650} of 0.05 and incubated until OD_{650} 0.3 at 45 °C and 180 rpm orbital shaking. The gene expression was induced by addition of a L-Tryptophan solution to a final concentration of 0.6 mM and ADH/A1a was expressed at 42 °C (or 30 °C) for 16 h. After the 16 h incubation, an additional induction with 3 mM L-Tryptophan was done and the cells were incubation for another 4 h. The cells were

cooled down on ice and harvested by centrifugation (4,500 x g, 30 min, 4 °C). The cell pellets were stored at -80 °C.

The in-batch bioreactor expression was done in four parallel-controlled bench-top continuous stirred 2.5 L-bioreactor vessels (DasGip, Eppendorf, Germany), containing 1.8 L of Hv-YPC medium. The inoculation, the fermentation, the induction of the protein expression and the cell harvest were done as described for shaking flasks. The pH value was regulated automatically by addition of HCl and NaCl (2 M stock solutions) and hold constant at pH 7.4. The airflow was hold within 40 L h^{-1}, the dissolved oxygen concentration was hold > 50% and the stirring speed was 400 - 1200 U min^{-1}. Foaming was prevented by automatically addition of Antifoam A (0.05% v/v). For the OD detection a 1 ml sample was withdrawn and analyzed spectrophotometrically.

3.3.3 Purification of ADH/A1a

The frozen cell pellets (~5 g) were thawed and suspended in 25 ml of 50 mM HEPES, pH 7.5, 2 M NaCl, containing 10 mM Imidazole, 10% v/v glycerol (high salt buffer A). A final concentration of 1 mM of Pefabloc SC (Sigma-Aldrich) was added. The cells were lysed and the genomic DNA disrupted by sonication on ice, using an amplitude of 30% and a pulse time of 50%, with a total sonication time of 60 min. The resulting suspension was centrifuged twice with 60.000 x g for 30 min at 4 °C. The supernatant was filtered an applied to a 5 ml Histrap HP column (GE Healthcare, Little Chalfont, UK), which was loaded with nickel ions and equilibrated with buffer A (flow-rate 2 ml min^{-1}). The buffer was washed with buffer A (2 M NaCl, or 0.5 M NaCl) and a gradient of 10 - 300 mM imidazole was applied resulting in protein elution. The fractions were analyzed by SDS-PAGE and fractions containing ADH/A1a were combined and dialyzed against 50 mM HEPES, pH 7.5, 2 M NaCl, and 10% v/v glycerol (storage buffer) over night at 4 °C. The dialyzed protein was concentrated to ~2 -3 mg ml^{-1} using an Amicon Ultracel$^{®}$ 30,000 MWCO (Millipore, Billerica, USA). The concentrated protein was applied to a size exclusion chromatography (SEC) and loaded on a preparative column HiLoad 16/60 Superdex 75 prep grade (GE Healthcare, Little Chalfont, UK). The flow rate was set to 1 ml min^{-1} using high salt buffer A and the protein fractions were collected. The purified

protein was fractionated, frozen in liquid nitrogen and stored at -80 °C. The purity of ADH/A1a was determined by SDS-PAGE and analysis using ImageJ version1.51. The specific protein concentration was determined using an extinction coefficient of 26,930 M^{-1} cm^{-1} and a calculated molecular weight of 37.6 kDa.

3.3.4 Tryptic digest and LC-MS/MS analysis of ADH/A1a

The identification of ADH/A1a was performed by LC-MS/MS analysis, using a Coomassie-stained protein band at the according molecular size, excised from a SDS-PAGE, destained, and digested using the in-gel tryptic digest kit after manufacturers protocol (Thermo Fisher Scientific, Waltham, USA). Peptides were analyzed by the KAUST corelab facilities using a nano-pump UltiMate™ 3000 UHPLC binary HPLC system (Dionex, ThermoFisher) coupled to a Q-Exactive HF mass spectrometer (ThermoFisher, Germany). The peptides were identified using MASCOT.

3.3.5 Determination of ADH/A1a activity

The specific and relative activities of ADH/A1a as well as its kinetic parameters were determined by a photometrical activity assay. The assay is based on the formation or consumption of the cofactor NAD(P)H during the oxidation or reduction reaction, which can be followed by its specific absorbance at 340 nm using a spectrophotometer. The reactions were performed in 1 ml semi-micro PMMA cuvettes (Sigma-Aldrich, St. Louis, USA). Unless otherwise specified, the defined standard conditions were used as stated below.

Oxidation reaction mix: The standard reaction was performed in 3 M KCl, 50 mM glycine buffer, pH 10.0, with 10 mM $NAD(P)^+$, and 0.2% (v/v) cinnamyl alcohol.

Reduction reaction mix: The standard reaction was performed in 3 M KCl, 50 mM K_2HPO_4-citric acid buffer, pH 6.0, with 0.25 mM NAD(P)H, and 0.4% (v/v) cinnamyl aldehyde.

Standard reaction procedure: Before the reaction, the enzyme ADH/A1a (0.4 μM) was incubated with zinc sulfate (0.2 mM) in the reaction mixture at room temperature for 15 min. The reaction was induced by the addition of the substrate, the mix was vortexed

and subsequently followed spectrophotometrically at 340 nm at 60 °C for 30 min. The absorbance was followed at 350 nm when absorbing substrates as cinnamyl aldehyde and derivatives were used (specific NAD$^+$ spectrum in Figure S2). All experiments were performed in duplicates or triplicates. A blank reaction without enzyme was included for each measurement.

The specific activity of the ADH/A1a was calculated from the slope of the NAD(P)H absorption curve using the Lambert-Beer law. One unit of enzyme activity was as defined as the reduction/oxidation of 1 μmol NAD(P)$^+$/NAD(P)H per minute, whereas the specific molar extinction coefficient of NAD(P)H of 6.22 mM^{-1} cm^{-1} was used (Equation 1.1). The specific activity was calculated based on the enzyme quantity (Equation 1.2) [112]. The relative activity was calculated using the optimal enzyme activity at standard conditions as control (Equation 2.1). The Michaelis-Menten kinetic was calculated using the non-linear regression and the Michaelis-Menten analysis provided Graphpad Prism.

(1.1)
$$A = \frac{(\Delta E - \Delta E_{\,blank}) \times V_{total}}{V_{ADH} \times \varepsilon_{NAD(P)H} \times d}$$

(1.2)
$$A_{spec} = \frac{(\Delta E - \Delta E_{\,blank}) \times V_{total}}{V_{ADH} \times \varepsilon_{NAD(P)H} \times d} \times c_{ADH}^{-1}$$

(2.1)
$$A_{rel} = \frac{A_{spec}}{A_{control}} \times 100\%$$

A = enzyme activity, [U ml^{-1}]

ΔE = extinction change over time, [min^{-1}]

V_{total} = total reaction volume, [ml]

V_{ADH} = volume of the enzyme, [ml]

ε = molar extinction coefficient of cofactor, [L mol^{-1} cm^{-1}]

d = cuvette path length, [cm]

A_{spec} = specific enzyme activity, [U mg^{-1}]

c_{ADH} = enzyme concentration, [mol L^{-1}]

A_{rel} = relative enzyme activity, [%]

$A_{control}$ = specific activity of positive control [U mg^{-1}]

3.3.6 Determination of the ADH/A1a stability

For the determination of the stability the enzymes activity was determined before and after the stability test in the oxidation reaction using standard conditions.

The salt stability of ADH/A1a was determined by incubation of the enzyme in different NaCl concentrations for 24 h at 4 °C, followed by the measurement of the remaining enzyme activity. The regain of the lost activity was determined after subsequent dialysis against dialysis buffer (50 mM HEPES, pH 7.5, 2 M NaCl) at 4 °C for 24 h.

For solvent stability, ADH/A1a in dialysis buffer was mixed with indicated amounts of methanol, acetonitrile, and DMSO, and incubated over time at room temperature. At different time points, samples were withdrawn and the remaining activity was determined.

For the freeze-thaw stability of ADH/A1a, a sample was frozen and stored at -80 °C for 14 days. The sample was unfrozen on ice and the relative activity was determined.

The thermal stability and the melting temperature T_m of ADH/A1a was determined by differential static light scattering using a Stargazer 2 (Harbinger Biotechnology, Toronto, Canada). The melting curve of ADH/A1a was measured with a 1.5 mg ml^{-1} protein sample. ADH/A1a was pre-incubated with zinc sulfate for 30 min and dialyzed against dialysis buffer for 2 h at RT before measurement. The measurement was done in triplicates. The melting temperature T_m was calculated using the non-linear regression of GraphPad Prism (Graphpad Software Inc., La Jolla, USA).

3.3.7 Determination of the products by GC-MS

For the determination of the product formation and the yield, the reaction was performed in a 5 ml-scale using the previous mentioned reaction conditions. The reaction mixture was poured into a round bottom flask and was stirred at 700 rpm in a 50 °C water bath for 2 h. The reaction compounds were extracted using 3 x 5 ml ethyl acetate, dried over NaSO$_4$, filtered and analyzed by gas-chromatography-mass spectrometry. The measurement was done using the 78910A GC system with a 5975C inert XL EI/CI MSD

with a J&W HP-5ms Ultra inert column (30m, 0.25 mm, 0.25 µM, Agilent Technologies, Santa Clara, USA). For each sample 2 µl were injected and the sample was analyzed using the manufacturers default method.

3.3.8 Computational analysis

The analysis of the conserved sites and the identification of the enzyme family were done using the CD-Search tool of the Conserved Domain Database (CDD) of the National Center for Biotechnology Information (NCBI) [113,114]. The subdomains were identified using InterProScan [115]. The secondary structure elements of ADH/A1a were predicted using JPred 4 [116]. Homologous ADHs with resolved crystal structures were identified using SWISS-MODEL and I-Tasser [117,118]. The homology model was created using SWISS-MODEL and was evaluated by the GMQE and QMEAN scores [119]. The QMEANDisco method was used to validate the local quality of the model. The structure analysis and visualization was pursued with UCSF Chimera [120]. The structures files of the homolog ADHs were obtained from the Protein Data Base (PDB) [121]. The electrostatic surface charge was calculated using the Adaptive Poisson-Bolzmann Solver (APBS) of the PDB2PQR 2.0 server using default values and applied to the structures using the APBS plugin of UCSF Chimera [122]. The surface-accessible residues were calculated with the Swiss PDB viewer 4.1.1 using a surface accessibility of $\geq$ 30%. Substrate docking studies and energetic minimization were done with YASARA [123].

The experimentally verified ADH sequences of different subfamilies were selected using Uniprot [124]. The annotated sequences of the CAD2 subfamily were taken from the CDD [113]. The sequence alignments were done using MAFFT version 7 [125]. The phylogenetic tree was constructed with FastTree2 using the maximum-likelihood method and the Jones-Taylor-Thornton substitution model (1,000 bootstraps, default settings) [126]. The functional annotation and possible metabolic pathways were investigated using BlastKOALA of the KEGG [127].

3.4 Detailed methods for SAV wildtype and mutants

3.4.1 Protein-ligand docking studies

The structure analysis, modification and visualization was done using UCSF Chimera [120]. The ligand docking studies and the energetic minimization were done with YASARA [123]. The SAV template crystal structure was achieved from the PDB database [121]. The PDB2PQR tool of UCSF Chimera was used to prepare structures for further calculations by addition of missing atoms and hydrogens, assigning atomic charges and radii using the SWANSON force field [122]. The electrostatic potential of the surface was visualized after Coulomb's law using the Coulombic surface coloring tool of UCSF Chimera.

3.4.2 Mutagenesis and cloning of SAV mutants

The SAV mutants, containing single mutations at indicated positions, were created using site-directed mutagenesis with pET11_SAV_wt as template using the QuikChange Kit. A detailed outline of all Mutants and Primers is shown elsewhere (supplementary). The mutated plasmids were transformed into *E. coli* BL21 (DE3) Gold.

3.4.3 Recombinant production of streptavidin

For the production of the streptavidin variants the protocol of SAV expression and purification after Humbert *et al* [128] was used. The recombinant *E. coli* BL21(DE3) pLysS harboring the pET11_SAVwt or SAV mutant gene was grown in 5L-baffled shaking flasks containing 2 L MTP-medium supplemented with glucose (0.4% w/v), ampicillin (60 µg ml^{-1}), chloramphenicol (34 µg ml^{-1}) and 250 µl of Antifoam A (0.05% v/v). The culture was incubated at 37 °C and vigorous shaking (200 rpm) to an OD$_{600}$ of 1.8. After shifting the cultures to the respective expression temperature (20 °C, 30 °C or 37 °C), the protein expression was induced by addition of isopropyl-β-D-thiogalactoside (IPTG) to a final concentration of 0.8 mM and incubated for 4 h or overnight. The culture was cooled down on ice and harvested by centrifugation (4,500 x *g*, 30 min, 4 °C). The cell pellets were stored at -80 °C.

The in-batch bioreactor expression was done in four parallel-controlled bench-top continuous stirred 2.5 L-bioreactor vessels (DasGip, Eppendorf, Germany), each containing 1.8 L of MTP medium supplemented with glucose (0.4% w/v), antibiotics and 200 µl Antifoam A (pure). The induction was done at an OD_{600} of 10 with a final concentration of 0.6 mM IPTG. The fermentation was performed at 20 °C overnight (16 - 18 h). The pH value was regulated automatically by addition of HCl and NaCl (2 M stock solutions) and hold constant at pH 7.4. The airflow was set constant at 200 L h^{-1} and the stirring speed was adjusted manually at 800 - 1200 U min^{-1}. Foaming was prevented by automatically addition of Antifoam A (pure). The cell harvest was done as described above.

3.4.4 Purification of streptavidin

For the purification of SAV the protocol of Humbert *et al.* [128] was slightly adapted. The bacterial cell pellet (~5 mg) was unfrozen and suspended in 25 ml iminobiotin binding buffer (50 mM Na_2CO_3, pH 9.8, 0.5 M NaCl) including 25 mM $MgCl_2$ and 5 mM $CaCl_2$. To the cell suspension about 1 mg of DNAse I and 1 mM Pefabloc SC were added and the mix was stirred at RT for 20 min. The cells were disrupted by a constant cell disruptor with a pressure of 1.8 kbar. The resulting suspension was centrifuged twice at 60000 x *g* for 30 min and the supernatant was filtered and applied on a 20 ml iminobiotin column (2-Iminobiotin-Agarose, Sigma-Aldrich), which was equilibrated with iminobiotin binding buffer (flow-rate 3 ml min^{-1}). The column was washed with iminobiotin binding buffer (100 ml) and the protein was eluted using elution buffer I (50 mM Na-Acetate pH 4) followed by elution buffer II (0.1 M acetic acid pH 2.9). The collected 5 ml protein fractions were neutralized by addition of 250 µl of 2 M Tris-HCl, pH 7.5, pooled and dialyzed against 2 L of 10 mM Tris-HCl pH 7.5 overnight at 4 °C. The protein was subsequently dialyzed twice against 2 L of dH_2O and lyophilized for long time storage. The lyophilized SAV was stored at -20 °C. Purity was determined by SDS-PAGE. The specific protein concentration was determined using an extinction coefficient of 41,940 M^{-1} cm^{-1} and a calculated molecular weight of 16.4 kDa.

3.4.5 Determination of free binding pockets of SAV

The determination of free and functional binding pockets of SAV was done by a fluorescence titration assay after Ebner *et al.* [129] using the biotin-4-fluorescein (B4F) dye (Sigma-Aldrich), set up in a 96-well plate. A SAV solution of 10 nM was titrated (0 - 115 µl) to a row of wells containing fixed amount of 160 nM B4F (10 µl) and filled up to 200 µl with buffer (50 mM NaH_2PO_4, pH 7.5, 100 mM NaCl, 1 mM EDTA and 0.1 mg ml^{-1} bovine serum albumin). As control SAV from *Streptomyces avidinii* (Sigma-Aldrich) was used. The protein-dye mix was incubated for 10 min at room temperature and the fluorescence was measured with a TECAN infinite M1000 using an excitation of 490 nm and emission of 425 nm. The measurement was done in triplicates. The free binding pockets of SAV were determined by the calculation of the breakpoint between declining fluorescence and saturation, which resembles the functional SAV concentration. The functional SAV concentration was defined as the concentration of SAV monomers, which are capable of binding biotin.

3.4.6 Stability studies of SAV mutants

For stability studies, the secondary structure signature of the SAV mutants was analyzed by circular dichroism (CD) using a JASCO J-815 spectropolarimeter. For the measurement 300 µl of a 15 - 20 µM protein solution in HPLC-water filled in a quartz cell (Hellma Type no. 110 QS, 1mm light path) and measured between 190 - 250 nm. Each sample was measured at least in triplicates and the spectra were merged and smoothen.

The thermal stability was determined by heating the sample from 20 °C to 100 °C in increments of 2 °C while recording the corresponding CD spectra. A melting curve was obtained by plotting the ellipticity of the maximum at 234 nm against the temperature. The melting temperature T_m was calculated using the van't Hoff equation with the GraphPad Prism [130,131].

To determine the solvent stability 250 µM of SAV in dH_2O was mixed with 10% - 90% of organic solvents and incubated for 12 - 48 h at room temperature. Precipitated protein

was separated by centrifugation (10 min, 14,000 x g) and the residual protein concentration of the soluble fraction was determined. Changes in secondary structure were analyzed by CD measurements.

The pH stability of SAV with incorporated biotin-ligand was tested. For that 1 mg ml^{-1} SAV was solved in buffers (50 mM acetate, MES, MOPS, carbonate or boric-acid) with varying pH values and the protein concentration was spectrophotometrically determined. The biotin-di-rhodium ligand (20 mg ml^{-1}, in DMSO) was added in defined amount and incubated for 5 min at room temperature. The formed protein aggregates were removed by centrifugation (10 min, 14,000 x g) and the remaining protein concentration of the soluble fraction was determined. The background absorbance of the ligand in DMSO was measured in buffer and used as blank.

3.4.7 Chemical modification of SAV

For the covalent PEGylation of primary amines of SAV, the methyl-PEG-NHS-ester regents, MS(PEG)$_4$ or MS(PEG)$_{12}$ (Thermo Scientific Pierce), were used. The conjugation reaction was carried out with a 5 mg ml^{-1} SAV solution in 100 mM NaHCO$_3$, pH 8.3, mixed with a 105-fold excess of the methyl-PEG-NHS-ester reagent in DMF (15-fold access, 7 primary amines). The reaction mix was incubated for 24 h at 4 °C and the residual esters were removed by dialysis against dH$_2$O for at least 24 h. The PEGylated SAV was lyophilized and stored at -80 °C.

3.4.8 Synthesis of Biot-Ar(CO$_2$)$_2$Rh$_2$(OAc)$_2$

The ligand Biot-Ar(CO$_2$)$_2$Rh$_2$(OAc)$_2$ (3-f) was synthesized by Dr. Camille Chapelein, Chair of Biochemistry, Technische Universität München, and later by Dr. Yunfei Cai, King Abdullah University of Science and Technology. In Figure 14 an overview of the synthesis steps is shown. The results of the synthesis will be published elsewhere.
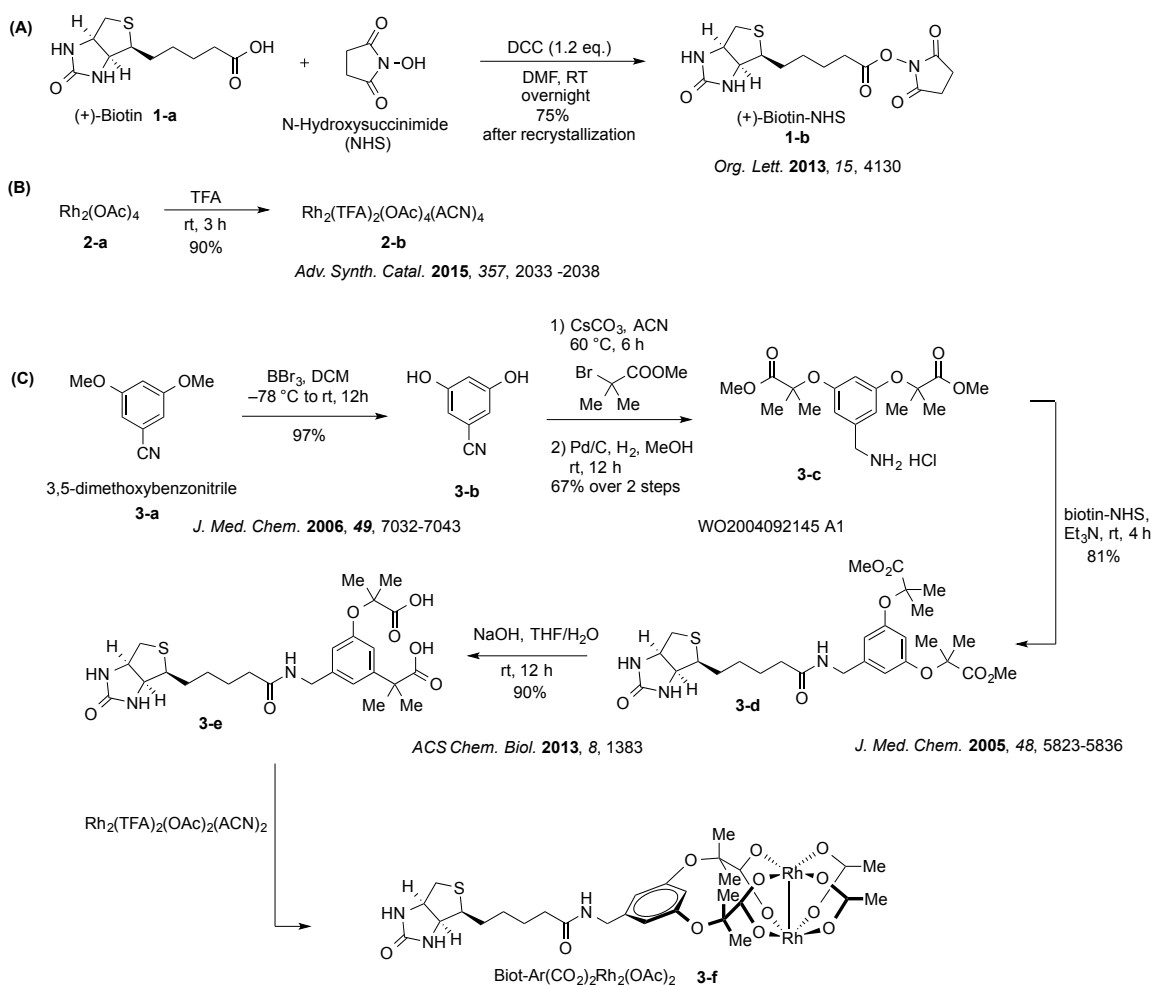
Figure 14: Synthesis of Biot-Ar(CO$_2$)$_2$Rh$_2$(OAc)$_2$. (A) Synthesis of (+)-biotin-NHS-ester; (B) Synthesis of dirhodium complex; (C) Complete synthesis of Biot-Ar(CO)$_2$(OAc)$_2$ ligand.

3.4.9 Synthesis of a diazo compound for the cyclopropanation

The diazoacetate **4-c** was synthesized by Dr. Yunfei Cai, King Abdullah University of Science and Technology after Chang *et al.* and Yang *et al.[132]*. To a solution of 2-(4-methoxyphenyl)acetic acid (10.0 g, 60.24 mmol) in MeOH (80 ml), conc. H$_2$SO$_4$ (5 ml) was added dropwise at 0 °C. The reaction was refluxed at 80 °C overnight, cooled down and distilled under reduced pressure. The residue was dissolved in water and neutralized (pH=7) using saturated aq. NaHCO$_3$. The product was extracted with EtOAc, dried over Na$_2$SO$_4$, and distilled under reduced pressure to provide the α-aryl-acetate **4-b**. The α-

aryl acetate **4-b** (3.0 mmol) and the 4-acetamidobenzenesulfonyl azide (p-ABSA, 864 mg, 3.6 mmol, 1.2 eq) were mixed in 10 ml CH_3CN and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU, 684 mg in 2 ml CH_3CN, 4.5 mmol, 1.5 eq) was added slowly at 0 °C. The reaction mixture was stirred at room temperature for 5 h, diluted with ether (10 ml), and washed with saturated aqueous NH_4Cl (20 ml), $NaHCO_3$ (20 ml) and NaCl (20 ml) in sequence. The organic phase was separated and dried with anhydrous Na_2SO_4. After evaporating the solvents, the residue was purified by column chromatography on silica gel (Hexanes:EtOAc = 20:1~10:1) to provide the pure diazo compound **4-c**.
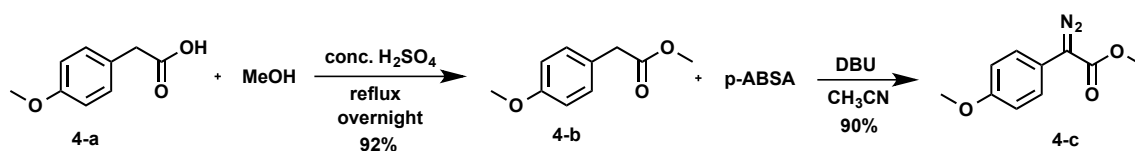
Figure 15: Synthesis of the diazo compound **4-c** for the cyclopropanation reaction.

3.4.10   Cyclopropanation reaction with SAV-dirhodium complex

The lyophilized SAV was solved in dH_2O or in reaction buffer and the protein concentration was adjusted to the desired functional SAV concentration. The Biot-$Ar(CO_2)_2Rh_2(OAc)_2$ ligand (20 mg ml^{-1} in DMSO) was added stoichiometrically 0.9:1 to the protein solution. The solution was slowly inverted and incubated for 2 min at room temperature and centrifuged at 14,000 x g for 10 min. The soluble fraction was withdrawn and dialyzed against reaction buffer for 2 h using 10 K slide-A-Lyzer mini dialysis units (Thermo Fisher Scientific) and was subsequently used for catalysis.

The cyclopropanation reaction was conducted after Srivastava *et al.* and adjusted for this approach [133]. If not stated otherwise, the catalyst (1 mol%) in reaction buffer (V = 1 ml) was mixed with 5 μmol of diazo compound **4-c** and 2.3 μl of methoxystyrene (3.5 eq., 17.5 μmol) in an HPLC-vial. The reaction was stirred with 1000 rpm at room temperature or 4 °C for 24 h. For extraction, the reaction mixture was diluted with water to 4 ml and extracted with DCM (5 ml, 6 times). The organic phase was dried with Na_2SO_4, filtered and the solvent was evaporated using a rotavap (100 mbar, 40 °C). The

product was solved in 0.96 ml of hexane/2-propanol 3/2, and 0.005 mmol of anisole was added as internal standard (in hexane/2-propanol 3/2, 0.125 M, 0.04 ml). The crude mixture was directly analyzed by HPLC with a Chiralpak AD-H column (Diacel corporation, Osaka, Japan) using hexane/2-propanol 98/2 with a flow of 1 ml min^{-1}. The compounds were analyzed at 210 nm. Both the ee% and yield of product were calculated using standard curves (Figure 16).

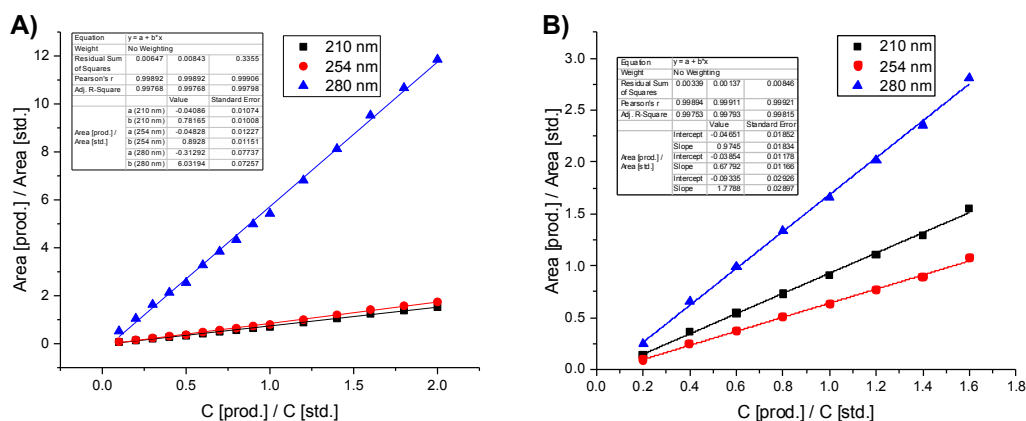**Figure 16:** Standard curve for the products of the cyclopropanation. (A) Cyclopropane **5-a** and (B) the benzyl alcohol **5-b.**

4 Characterization of a polyextremophilic zinc-dependent alcohol dehydrogenase from the Atlantis II Deep Red Sea brine pool

4.1 Results

4.1.1 Sequence analysis of the *adh/a1a* gene

The gene of the zinc-dependent ADH, *adh/a1,* was annotated from a SAG (AAA261F19) of an unclassified MSBL1 archaeon, sampled from the Red Sea Atlantis II Deep brine pool interphase (2036 m, 57 - 63 °C, 15.1 - 16.8%, pH 5.6) [44]. A*dh/a1* was identified and chosen from the INDIGO Red Sea database for characterization using the profile and pattern matching algorithm (PPMA) [48,49]. The in-depth analysis of the sequence revealed a 20 amino acid long N-terminal nonsense sequence. The sequence search against the non-redundant protein database did not reveal any homology of the nonsense part to other sequences. Furthermore the secondary structure prediction using Jpred [134] showed a random coil formation at this position. Deep analysis of the DNA sequence revealed the alternative start codon UUG, which was missed by the annotation algorithm. Alternative start codons are reported to be used in haloarchaea but are still often a source for incorrect annotations [135]. Based on this findings, the shortened construct, *adh/a1a*, was designed excluding the nonsense sequence and adding an N-terminal his$_6$-tag linked by a GSG-linker (Figure S1) [136].

ADH/A1a was annotated as a member of the zinc-binding medium-chain alcohol dehydrogenase (MDR) superfamily by the conserved domain database (CDD), exhibiting domains of the cinnamyl alcohol dehydrogenase (CAD) family [137]. Based on the conserved features of the CAD family, the catalytic and structural zinc-binding sites were identified, as well as putative NAD(P)$^+$ binding and substrate binding site (Figure 17). Further, the N-terminal GroeES-like domain (position 1 - 152) and the C-terminal NAD(P)H-binding domain (position 153 - 294), were identified and secondary structure elements were predicted.

```
                                                                  vvv
ADH/A1a   MKAMQLKEPKPVEQAPLEMVELKEPRPGPKEVKIDVQACGVCHTDLHTVEGELSLPKLPI  60
CAD2      MKAMVLEKPGPIEENPLRLTEVPVPEPGPGEVLIKVEACGVCRTDLHIVEGDLPPPKLPL
          **** *::*: *:*: **.:.*:  *.*** ** *.*:*****:**** ***:*. ****:

              v                                   v    v v v        v
ADH/A1a   VPGHEVVGVVEESGDEAEHFEVGDRVGVTWLYSSCGECKFCRRGQENLCEDPMFTGLHAD 120
CAD2      IPGHEIVGRVEAVGPGVTRFSVGDRVGVPWLGSTCGECRYCRSGRENLCDNARFTGYTVD
          :****:** **   *   . :*.*******.** *:****::** *:****::. ***  .*

                               v                               vvvvvv
ADH/A1a   GGYEESMVAKEDFVYPIPKNISDEDAAPLLCAGVIGYRSLRLSEVKPGQRLGLFGFGASA 180
CAD2      GGYAEYMVADERFAYPIPEDYDDEEAAPLLCAGIIGYRALKLAGLKPGQRLGLYGFGASA
          *** * ***.* *.*****:: .**:*******:****:*:*: :********:******

                    vv        v       v                        vv v
ADH/A1a   HIVIQLATDMGCEVYVFTRSEEHRRLARELGSAWEGSAKDDPPHRIDSGITFAPVGWIVK 240
CAD2      HLALQIARYQGAEVFAFTRSGEHQELARELGADWAGDSDDLPPEPLDAAIIFAPVGALVP
          *::.:*:*    *.**:.**** **:.******: * *.:.* **. :*:.* ***** :*

                  vv                        vvv
ADH/A1a   EALRDLEKGGTLAINAIHMTPIPELDYDLIYHEKKLRSVANVTREDAEGFLKIAGDIPVQ 300
CAD2      AALRAVKKGGRVVLAGIHMSDIPAFDYELLWGEKTIRSVANLTRQDGEEFLKLAAEIPIK
           *** ::*** :.: .***: ** :**:*:: **.:*****:**:*.* ***:*.:**::

ADH/A1a   TEVETFPLEEANRALRLLKDSKINGAGVLKVS 332
CAD2      PEVETYPLEEANEALQDLKEGRIRGAAVL---
          .****:******.**: **:.:*.**.**
```

v - Catalytical Zn binding site   → beta sheet
v - Structural Zn binding site   ▭ alpha helix
v - putative NAD(P) binding site   1 – 152 GroES-like domain
v - putative substrate binding site   153 – 294 NAD(P)-binding domain

Figure 17: Sequence analysis of ADH/A1a. ADH/A1a was aligned with the CAD2 consensus sequence. The conserved subdomains are highlighted and the conserved binding sites are marked. The secondary structure elements were predicted using JPred and are shown below the amino acid sequence. The sequence analysis was done using NCBI's CD-search of the CDD.

4.1.2 Protein production and purification

The expression of ADH/A1a was first attempted in *E. coli* under standard conditions (0.5 M NaCl) but resulted in an insoluble and inactive protein aggregate (data not shown). As expected, the enzyme was not successfully expressed and folded in *E. coli*

under low salt conditions, since the production of halophilic enzymes is more challenging and generally requires high intracellular salt concentrations [138,139]. Therefore, the efficient halophilic *Hfx. volcanii* H1895 strain was chosen for ADH/A1a expression at high salt concentrations (2 M NaCl), approximating the natural salt environment of the enzyme [33]. The expression of ADH/A1a in bioreactor scale was done after Strillinger *et al.* and could improve the yield significantly by 5.6-fold, yielding in 15.2 mg L^{-1} purified ADH/A1a (Figure 18A) [34]. Different expression temperatures (30 °C, 43 °C) were tried out to improve the expression yield, but no significant alteration in protein yields per gram cell weight was observed. However, at lower temperatures the cell growth is inhibited dramatically, leading to a decreased cell yield and, hence, decreased protein yield (Figure 18B).
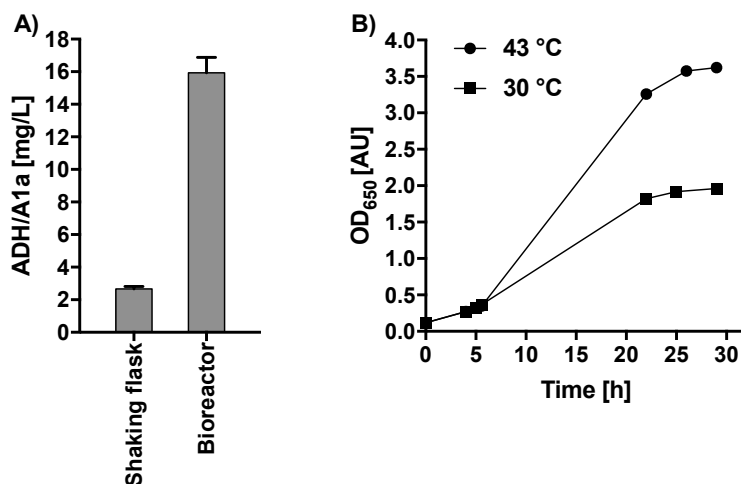

Figure 18: Expression of ADH/A1a in *Hfx. volcanii* 1895. (A) Yield of purified ADH/A1a expressed in both shaking flask and bioreactor. (B) Growth curve of *Hfx. volcanii* during expression under different temperatures. Protein expression was induced at OD 0.33 (after 5.5 h) with 6 mM L-Tryptophan and after 22 h with 3 mM L-Tryptophan.

ADH/A1a was purified under high-salt conditions using a combination of metal ion affinity chromatography (IMAC) and size exclusion chromatography (SEC). The Ni-affinity purification at high salt (1 - 2 M NaCl) and 10% (v/v) glycerol resulted in a protein purity of about 60 - 80% (Figure 19A). The purity was confirmed by SDS-PAGE, which showed a prominent band corresponding to ~38 kDa of the enzyme monomer

including the his$_6$-tag (Figure 19B). Different purification setups were tested for ADH/A1a using low salt concentration in the washing and elution buffer ($\leq$ 0.5 M NaCl), which resulted in an increased purity of > 95% (Figure 19C/D); however, the activity of the purified ADH/A1a decreased dramatically below 20% of the initial activity. Concluding the results, an affection of the affinity separation by high salt concentrations was seen, as contaminant proteins from the host organism tend to stick to the enzyme. The wash and gradient elution in low salt concentrations resulted in dissociation of the contaminant proteins and almost pure ADH/A1a, but also inactivated the enzyme irreversibly. Because of that, the enzyme was subjected to SEC purification.
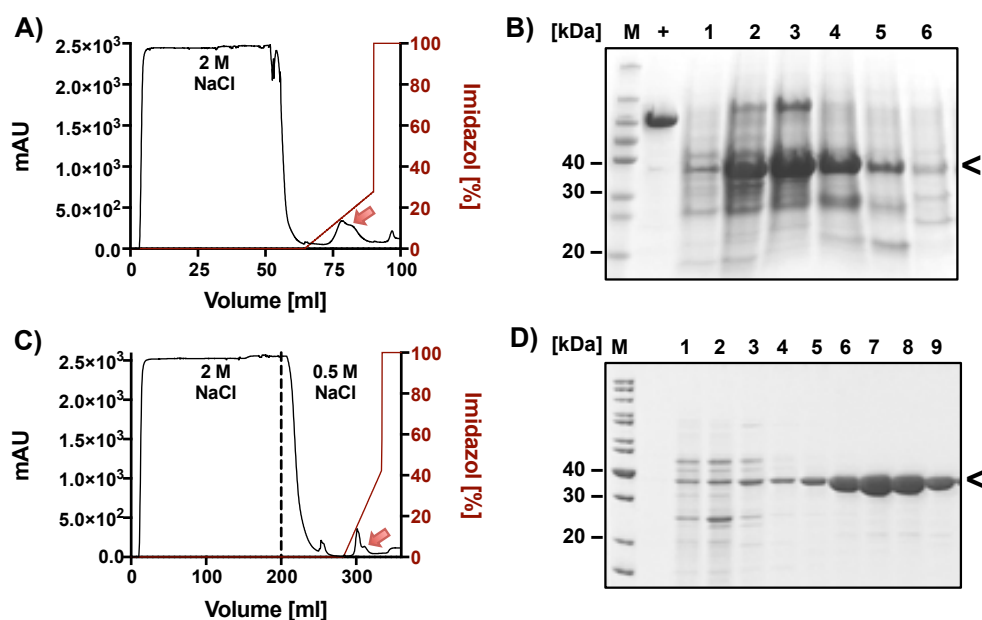

Figure 19: Ni-affinity purification of ADH/A1a. (A) Chromatogram of IMAC at 2 M salt; (B) Coomassie-stained SDS-PAGE of eluted fractions at 2 M salt; (C) Chromatogram of IMAC at 0.5 M salt; (D) Coomassie-stained SDS-PAGE of eluted fractions at 0.5 M salt. The elution peak is highlighted with a red arrow. The gel-band of ADH/A1a is indicated with a black arrow at ~38 kDa. M = Marker, (+) = his-tagged beta-galactosidase as positive control, (1 - 6) = protein elution fractions.

The IMAC purified ADH/A1a fractions were dialyzed, concentrated and applied onto a HiLoad 16/60 Superdex 75 prep grade preparative column (GE Healthcare, Little Chalfont, UK) with a flow rate of 1 ml min^{-1} using high salt buffer A. The SEC

55

purification resulted in an improved purity of > 90% with a total yield of 15 mg L^{-1} (Figure 20A/B). The identity of the ADH/A1a was confirmed by tryptic digest and LC-MS/MS analysis. A total of seven unique peptides throughout the whole sequence were identified and a sequence coverage of 66% was obtained (Figure 20C).
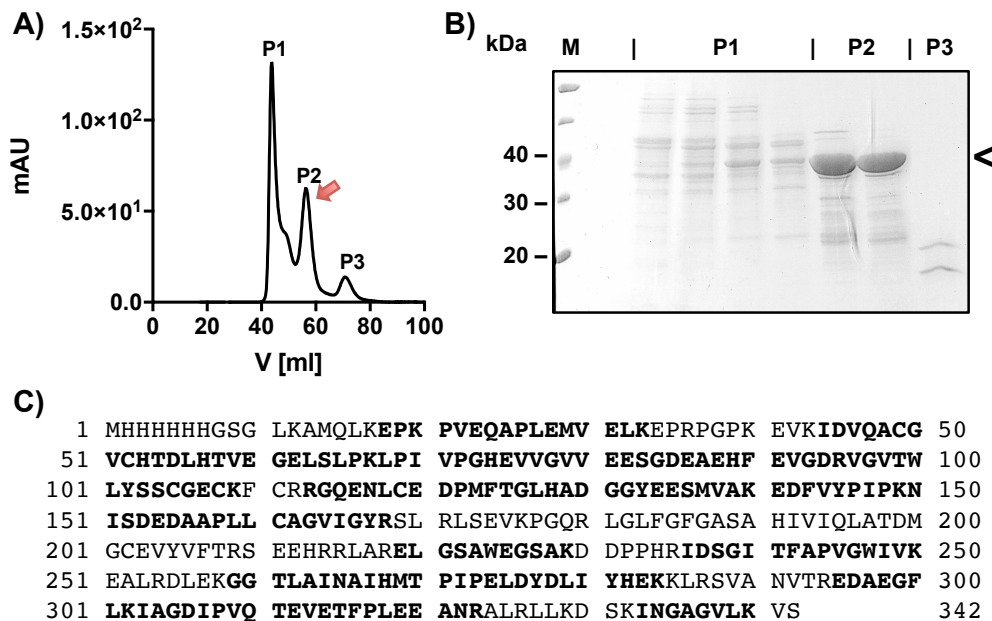

Figure 20: Gel-filtration purification and identification of ADH/A1a. (A) Chromatogram of SEC at 2 M NaCl; the elution peak of ADH/A1a is highlighted with a red arrow; (B) Coomassie-stained SDS-PAGE of eluted fractions. Band of ADH/A1a is indicated with a black arrow at ~38 kDa. M = Marker, (P1 - P3) = elution fractions of peak 1 - 3. (C) Identification of ADH/A1a by tryptic digest and LC-MS/MS analysis of an excised gel band. Matched peptides are shown in bold.

4.1.3 Effect of salt and temperature on the activity

The activity of ADH/A1a was tested over a broad range of NaCl and KCl concentrations and temperatures (Figure 21A/B). The enzyme was active in a salt range of 1.5 - 5 M with an optimum activity found at 4 M for both NaCl and KCl. Notably, higher activities were observed with KCl than with NaCl. The effect of the temperature on the activity was screened from 10 - 80 °C, with the optimum found at the relatively high temperature

of 70 °C. Substantial activity was measured at 80 °C, however, thermal aggregation of the protein occurred over time that is in agreement with the melting temperature T_m of ~73 °C (chapter 4.1.10). Based on the initial characterization, the standard reactions conditions were defined as 3 M KCl and 60 °C. This framework was used to avoid salt precipitation and buffer evaporation.
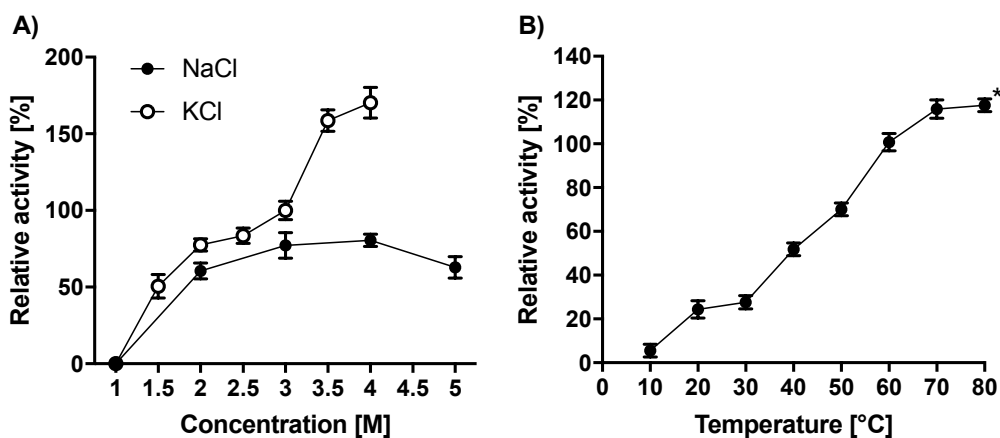

Figure 21: Effect of salt concentration and temperature on the activity of ADH/A1a. (A) Varying NaCl and KCl concentrations (saturation limit: 4 M for KCl, 5 M for NaCl), (B) varying temperatures. The activity was measured in the oxidation reaction using standard conditions and cinnamyl alcohol as substrate. * Protein precipitated during the reaction.

4.1.4 Effect of pH value

The effect of the pH-value on the activity was determined for the oxidation and the reduction reaction (Figure 22). ADH/A1a showed a very divergent activity profile with optima at pH 10.0 for the oxidation and pH 6.5 for the reduction reaction respectively. For the oxidation reaction rapid loss of activity was found below pH 10.0, with over 80% activity loss observed at pH 9.5 and below. The oxidation reaction was catalyzed in neutral and slightly acidic pH. However, the enzyme degraded at more acidic pH ($\leq$ pH 4.5).
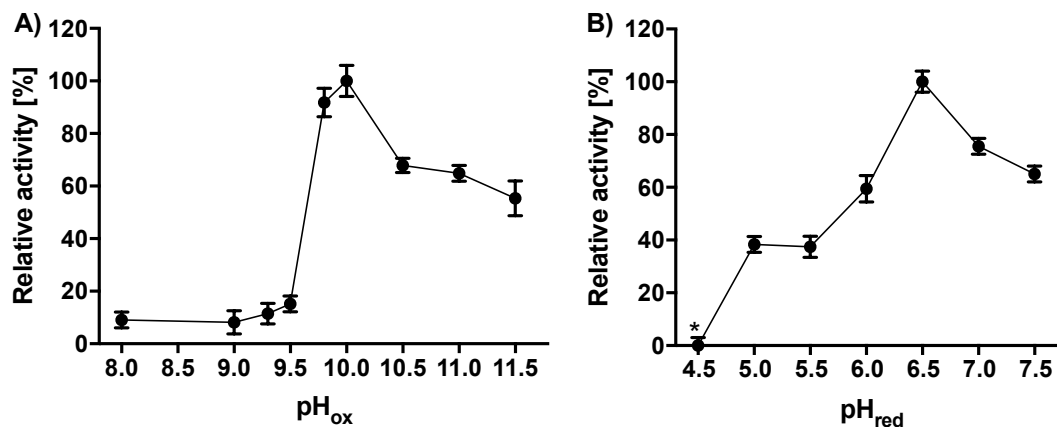
Figure 22: Effect of pH value on the activity of ADH/A1a. (A) Oxidation reaction, (B) Reduction reaction. The activity was measured using standard conditions and cinnamyl alcohol/aldehyde as substrates. * Protein precipitated during the reaction.

A *in situ* thermodynamic analysis of the reactions was done using the eQuillibrator [140]. The Gibbs free energy $\Delta_r G$ of the reaction was calculated using cinnamyl alcohol/NAD$^+$ and cinnamyl aldehyde/NADH as reactants, and an ionic strength of 3 M. The reduction reaction is overall favored over the oxidation reaction at pH values below 10; however, at basic pH values ($\geq$ pH 10) the oxidation reaction is slightly favored over the reduction reaction with higher negative $\Delta_r G^{'}$ values (Figure 23A). At physiological conditions (1 mM for all reactants) the Gibbs free energy for the oxidation reaction is positive, indicating a need of external activation energy for the reaction, whereas the reduction reaction is energetically favored with overall negative $\Delta_r G^{'m}$ values.
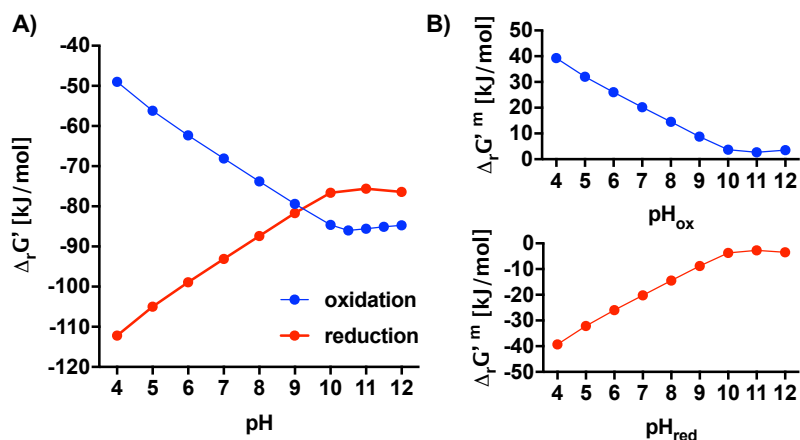
Figure 23: Thermodynamic analysis of the oxidation/reduction reaction at different pH values. The Gibbs free energy $\Delta_r G$ was calculated by eQuillibrator at different pH values (T = 25 °C, ionic strength: 3 M, reactants: cinnamyl alcohol/ NAD^+ and cinnamyl aldehyde/NADH). (A) Gibbs free energy $\Delta_r G'$ under reaction conditions (oxidation: 15.5 mM cinnamyl alcohol and 200 mM NAD^+, reduction: 31 mM cinnamyl aldehyde and 0.2 mM NADH). (B) Gibbs free energy $\Delta_r G'^m$ under physiological conditions (1 mM for all reactants).

4.1.5 Cofactor dependency

The ADH/A1a dependency on the cofactors was investigated. The cofactors NAD^+/NADH and $NADP^+$/NADPH were almost equally accepted (Figure 24A). The optimal cofactor concentration of NAD^+ for the oxidation was 10 mM (Figure 24B). Higher concentrations reduced the activity of the enzyme presumably due to an inhibitory effect. In the standard reactions, NAD^+ and NADH were used, as they are less expensive compared to $NADP^+$ and NADPH.
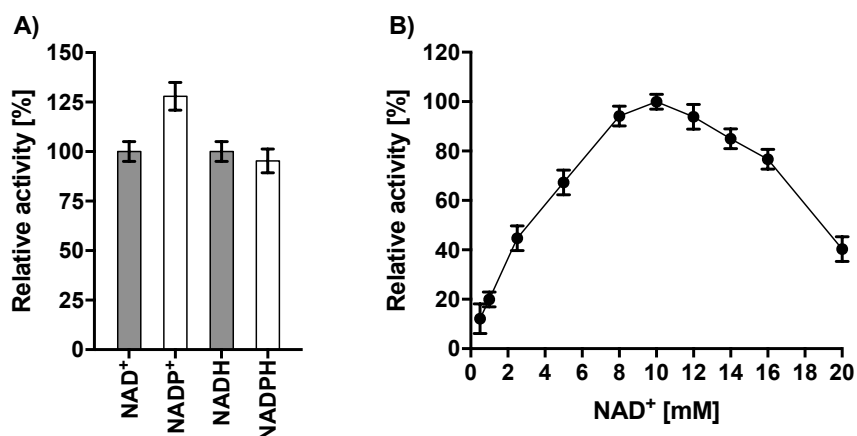
Figure 24: Cofactor dependency of ADH/A1a. (A) Activity of ADH/A1a using $NAD^+/NADP^+$ (10 mM) in the oxidation and NADH/NADPH (0.2 mM) in the reduction reaction; (B) Varying NAD^+ concentration in the oxidation reaction. The relative activity was measured using standard conditions.

4.1.6 Metal-ion dependency

The Atlantis II Deep brine pool contains high concentrations of heavy metals, of which several are suitable as catalytic active metal for ADH/A1a [37]. However, the sequence analysis of ADH/A1a predicted a zinc-dependent alcohol dehydrogenase with a catalytic zinc(II) binding site. To determine the catalytically active metal ion, ADH/A1a was incubated with different metal ions before catalysis and the effect on the enzyme activity was measured. The enzyme activity increased 3.5-fold with the addition of zinc(II) ions. The metals manganese(II) and magnesium(II) had no effect, and cobalt(II) and nickel(II) had adverse effects on the enzyme activity. The incubation with copper(II) or chromium(II) resulted in total loss of activity (Figure 25A). The metal-free ADH/A1a (dialyzed against EDTA) showed activity exclusively with the addition of zinc(II) ions, confirming the exclusive zinc dependency (Figure 25B).

The optimal zinc(II) concentration and incubation time were determined for the reactions (Figure 25C/D). The highest activity was measured with 0.2 mM $ZnSO_4$ (~100-fold access to protein) and 45 min preincubation time. No adverse effect was found for higher zinc(II) concentrations and incubation times. For the standard reactions, ADH/A1a was further on preincubated with 0.2 mM $ZnSO_4$ for 30 min before catalysis.
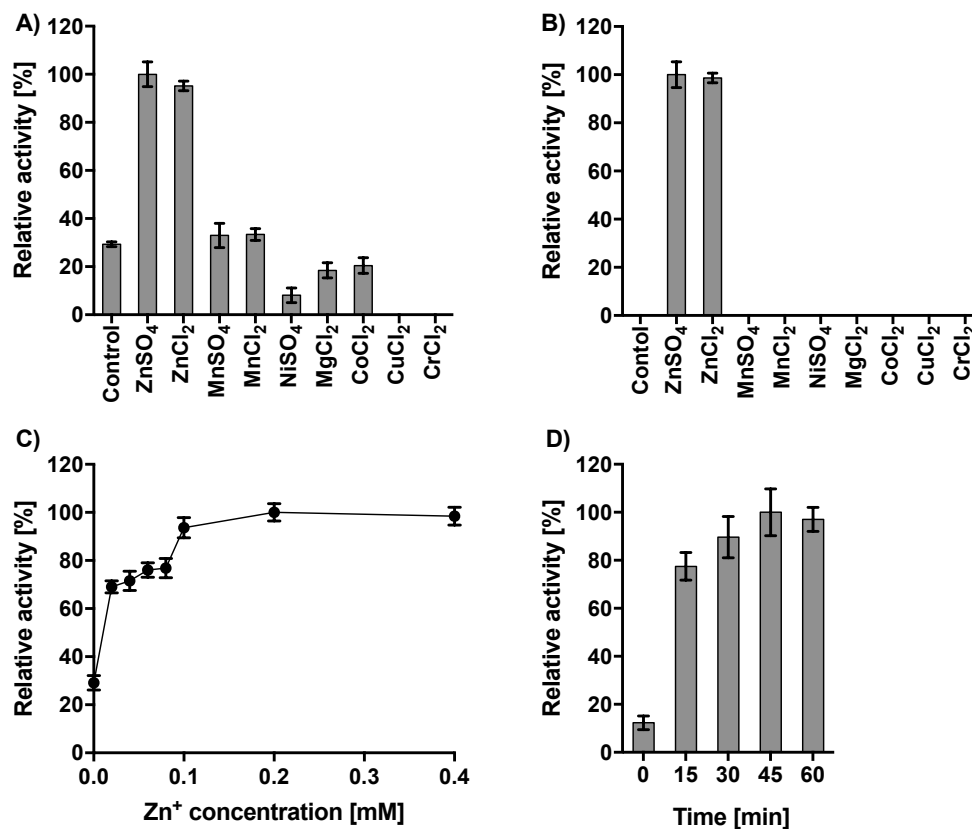
Figure 25: Effect of metal ions on the activity of ADH/A1a. (A) ADH/A1a incubated with various metal ions; (B) Metal-free ADH/A1a incubated with various metal ions; (C) ADH/A1a incubated with varying $ZnSO_4$ concentrations; (D) ADH/A1a with $ZnSO_4$ incubated for different times. If not stated otherwise, the enzyme was incubated with 0.2 mM metal ion solution in reaction buffer at room temperature for 30 min before measurement. The activity was measured in the oxidation reaction using standard conditions.

4.1.7 Substrate spectrum of ADH/A1a

To determine the substrate scope of ADH/A1a in the oxidation reaction, several primary and secondary alcohols were screened in a 96-well scale (Table 1). ADH/A1a oxidized a broad range of substrates including small-, medium- and long-chain alcohols with a preference for long primary alcohols. Remarkably, several phenol and terpene alcohols were converted. In the reduction reaction, exclusively conversion of cinnamyl aldehyde and derivatives was detected (Table 2).

Table 1: Substrate scope of ADH/A1a in the oxidation reaction tested in 96-well scale.

61

No.	Substrate	Activity	No.	Substrate	Activity
1	Methanol	-	12	3-Buten-2-ol	++
2	Ethanol	+	13	1-Pentanol	+++
3	Isopropanol	+	14	3-Pentanol	-
4	Glycerol	-	15	3-Methyl-3-Pentanol	-
5	1-Phenylethanol	-	16	1,5-Pentanediol	+++
6	1,3-Propandiol	+	17	1-Heptanol	+++
7	1-Propanol	++	18	3-Octanol	-
8	1-Butanol	+++	19	Sorbitol	-
9	2-Butanol	-	20	Pentaerythritol	-
10	2-3-Butandiol	-	21	Ethylen glycol	-
11	Prenol	++	22	Cinnamyl alcohol	+++

The activity was measured in the oxidation reaction with 0.2% (v/v) substrates at 45 °C using standard reaction conditions. (-) = no activity, (+) = low activity, (++) = medium activity, (+++) = high activity.

Table 2: Substrate scope of ADH/A1a in the reduction reaction tested in 1 ml scale.

No.	Substrate	Activity	No.	Substrate	Activity
1	Acetaldehyde	-	9	Cinnamyl-methyl ketone	+++
2	2-Butanone	-	10	2,4-Pentanedione	-
3	3-Buten-2-one	-	11	Isovaleraldehyde	-
4	2,3-Butanedione	-	12	Ethylcinnamat	-
5	Acetylacetone	-	13	3-Methylcrotonaldehyde	-
6	4-Isobutylacetophenone	-	14	2-Pentanone	-
7	Cinnamyl aldehyde	+++	15	2-Hexanone	-
8	Raspberry ketone	+++	16	2-Decanone	-

The activity was measured in the reduction reaction with 0.4% (v/v) substrates at 60 °C using standard reaction conditions. (-) = no activity, (+++) = high activity.

Based on the initial screens, the specific activities for several promising substrates were determined (Figure 26A/B). The highest reaction rates were achieved with cinnamyl alcohol (~0.49 U mg^{-1}), followed by 1,5-pentanediol and 1-heptanol (~0.29 U mg^{-1} for both). Remarkably, the secondary alcohol 3-buten-2-ol (~0.05 U mg^{-1}) was also oxidized, though at slow rates. In the reduction reaction, a conversion was exclusively found for cinnamyl aldehyde (0.31 U mg^{-1}), raspberry ketone (~0.25 U mg^{-1}), and cinnamyl-methyl-ketone (~0.40 U mg^{-1}).

The conversion of the substrates cinnamyl alcohol and hydroxycinnamyl alcohol to the corresponding aldehydes was confirmed in 5 ml scale by GC-MS analysis (Figure S3), though with low overall product yields.
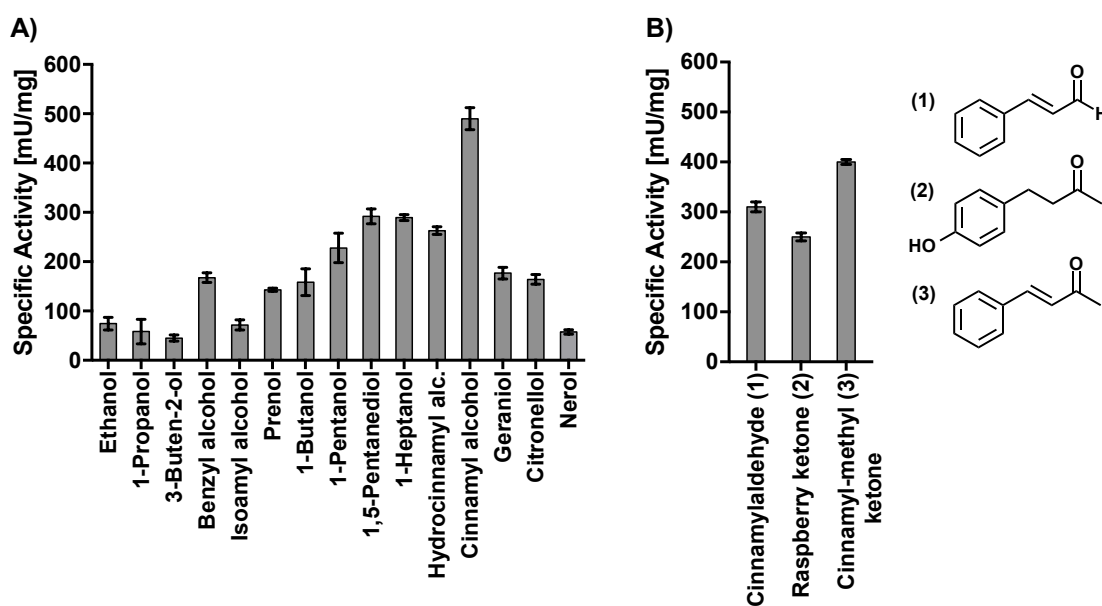

Figure 26: Substrate scope and specific activity of ADH/A1a for various substrates. (A) Substrates in the oxidation reaction (0.2% v/v substrate); (B) Substrates in the reduction reaction (0.4% v/v substrate); structure of converted aldehydes/ketones is shown on the right. The activity was measured using standard conditions.

4.1.8 Michaelis-Menten kinetics

The analysis of the Michaelis-Menten kinetics was done for the substrates cinnamyl alcohol and geraniol (Figure 27). The reaction was handled as a 1-substrate-1-enzyme reaction, as the cofactor was added in the access amount. The maximal velocity (V_{max}), the turnover number (k_{cat}) and the kinetic specificity (K_{cat}/K_m) for cinnamyl alcohol were higher than for geraniol (Table 3).
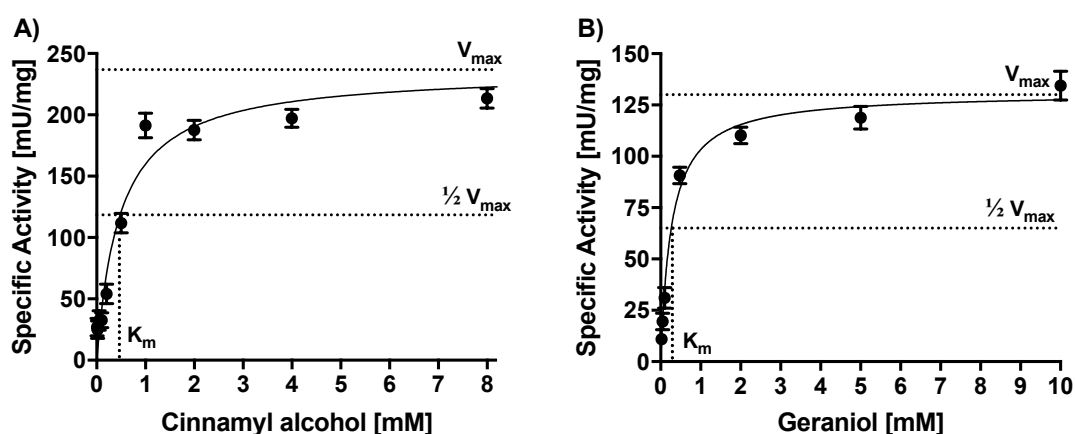

Figure 27: Michaelis-Menten kinetics of (A) cinnamyl alcohol and (B) geraniol as substrates. Enzymatic activities were measured with standard conditions.

Table 3: Kinetic parameters of ADH/A1a.

Substrate	V_{max} [mU mg^{-1}]	K_m (mM)	k_{cat} (s^{-1})	K_{cat}/K_m (s^{-1}mM^{-1})
Cinnamyl alcohol	236.8 ± 9.2	0.47 ± 0.09	0.144 ± 0.006	0.31 ± 0.01
Geraniol	130.9 ± 4.2	0.27 ± 0.04	0.080 ± 0.003	0.29 ± 0.08

4.1.9 Salt and solvent stability

The enzyme ADH/A1a was incubated at different salt concentrations for 24 h, and the remaining activity was determined. ADH/A1a was exceptionally stable at high salt concentrations ($\geq$ 0.8 NaCl); however, the enzyme stability significantly decreased at low salt concentrations ($<$ 0.8 M NaCl) resulting in loss of activity (Figure 28A). Nevertheless, almost one-third of the lost activity was regained when dialyzed against the high salt buffer (2 M NaCl, 50 mM HEPES, pH 7.5) (Figure 28B).
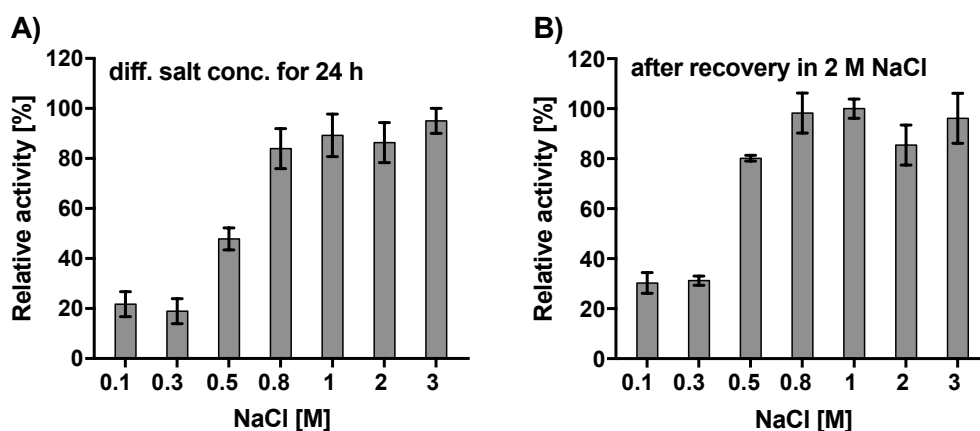

Figure 28: Stability of ADH/A1a at different salt concentrations, (A) Activity of ADH/A1a incubated at indicated salt concentrations for 24 h; (B) Recovered activity of ADH/A1a samples, incubated in dialysis buffer (2 M NaCl, 50 mM HEPES, pH 7.5). The activity was determined in the oxidation reaction using standard conditions.

The solvent stability was tested by incubation of the enzyme in varying concentrations of methanol, acetonitrile, and DMSO and subsequent analysis of the residual activity over time. ADH/A1a retained ~70 - 80% of its initial activity in 10% (v/v) acetonitrile, methanol, and dimethylsulfoxide after 32 h (Figure 29). At high solvent concentrations (30%, v/v), ADH/A1a was able to withstand dimethylsulfoxide and methanol, and retained almost half of its initial activity after 32 h, although, it lost its activity in acetonitrile after 1 h.
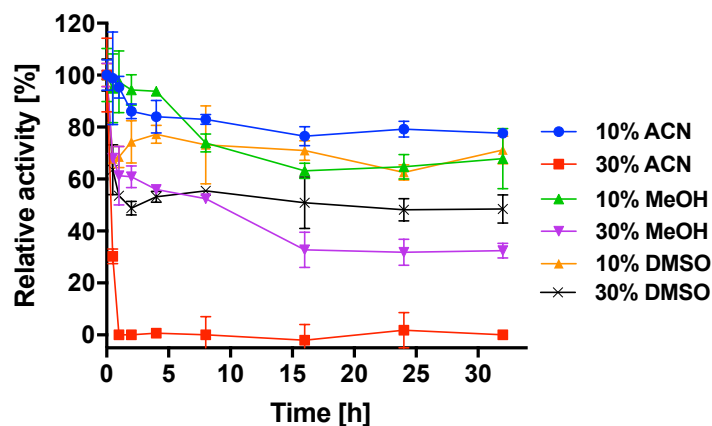
Figure 29: Stability of ADH/A1a in different solvents concentrations over time. ADH/A1a in dialysis buffer was incubated with indicated amounts of solvents, and the relative activity was measured at different time points. The activity was determined in the oxidation reaction using standard conditions.

4.1.10 Thermal stability

The thermal melting curve was determined for a temperature range of 50 - 90 °C using differential static light scattering. The metal-free ADH/A1a was pre-incubated with zinc sulfate before the measurement. The melting temperature T_m of ~73 °C was calculated using the non-linear regression of GraphPad Prism (Figure 30A).

The freeze-thaw stability was determined for an ADH/A1a sample in a storage buffer that was stored at -80 °C for 14 days. For that, the activity of ADH/A1a was determined before and after the freeze-thaw cycle. Almost the full enzyme activity was measured with the thawed ADH/A1a, suggesting high freeze-thaw stability (Figure 30B).
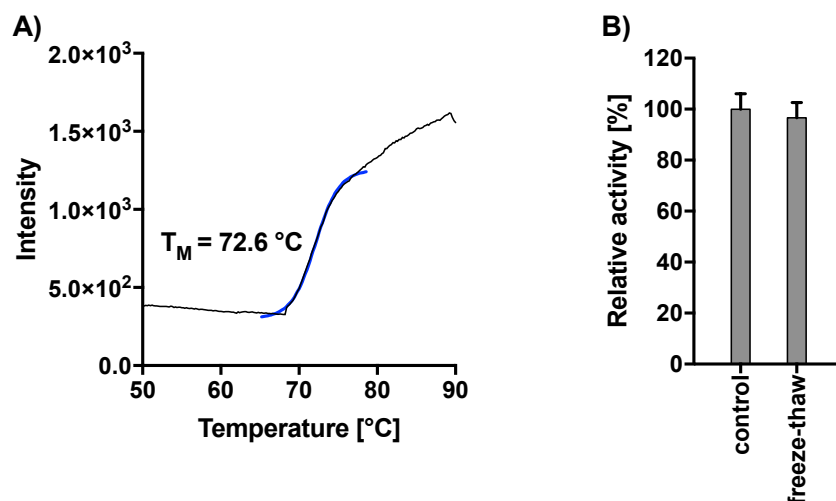
Figure 30: Thermal stability of ADH/A1a. (A) A thermal melting curve of ADH/A1a measured using differential static light scattering. (B) Freeze-thaw stability of ADH/A1a. The relative activity was measured after freeze-thaw of the sample. As a control, the activity of ADH/A1a was measured before freezing. The activity was measured in the oxidation reaction using standard conditions.

4.1.11 Homology model of ADH/A1a

Several crystallization screens were performed in order to get a protein crystal for structure analysis; however, no protein crystal was observed. Therefore, the homology model of ADH/A1a was build based on the template crystal structure of the thermophilic htADH of *Bacillus stearothermophilus* (PDB: 1rjw, 37% identity) [141]. The structure model of ADH/A1a revealed a typical bacterial zinc-dependent MDR structure, consisting of a homotetrameric quaternary structure bearing two zinc ions per subunit (Figure 31A). One subunit is built of two domains, a N-terminal catalytical active domain and a C-terminal cofactor-binding domain (Figure 31B). The N-terminal GroES-like domain encloses the catalytically active center and consists of four β-strands and a short α-helix that form a β-barrel-like structure [142]. The C-terminal cofactor NAD(P)H-binding domain contains a typical Rossmann-fold, composed of six parallel beta-strands flanked by four alpha-helices [57]. The catalytically active zinc ion is coordinated by C42, H64, and C151 and resembles the typical binding motif of the MDR superfamily

(Figure 32A). The structural zinc ion is coordinated by four cysteine residues (C95, C98, C101, C109) and located in the interface region between two monomers (Figure 32B).
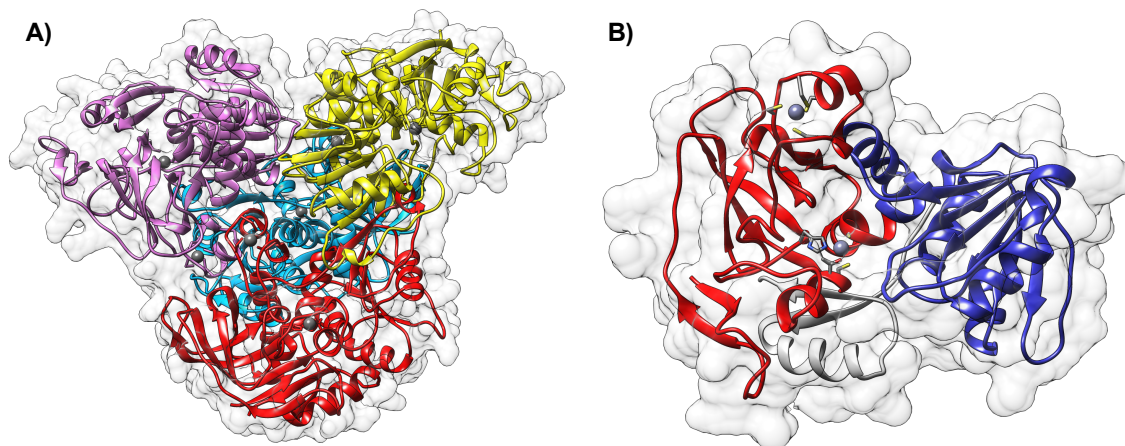

Figure 31: Homology model of ADH/A1a based on the htADH of *B. stearothermophilus* (PDB: 1rjw) [141]. (A) Ribbon diagram of the ADH/A1a homotetramer. The subunits are highlighted in different colors. The catalytical and structural zinc(II) are colored in gray. (B) Ribbon diagram of ADH/A1a monomer. The subdomains are highlighted red for the GroES-like domain and blue for the nucleotide-binding domain. The catalytical and structural zinc(II) are shown in gray.
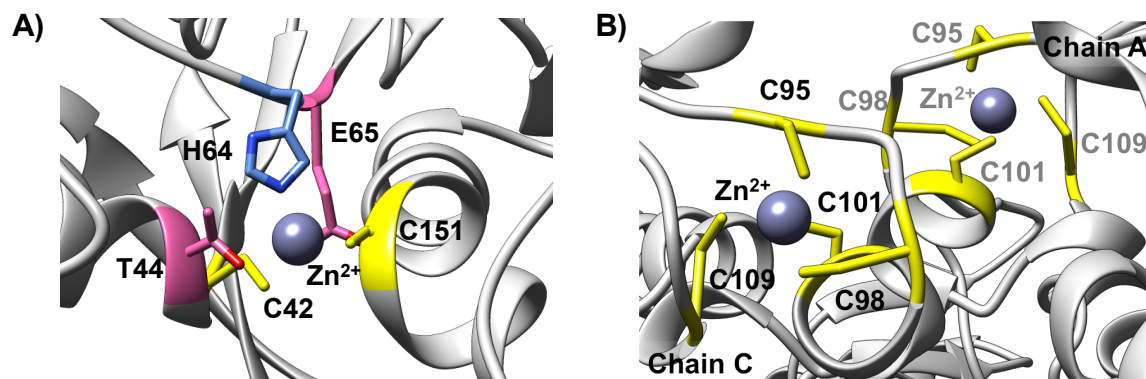
Figure 32: Coordination site of the catalytical and structural zinc. (A) Catalytical zinc binding site with Cys-His-Cys binding motif; conserved residues E65 and T44 are colored in pink (B) Structural zinc binding sites with four coordinating cysteines residues, located in the interface region between two monomers.

The homology model appeared to be reliable, with an overall QMEAN Z-score of -1.6 and a GMQE of 0.71. The local quality assessment of the model revealed high structural conservation in the protein core domains (QMEANDisco ~0.8 - 0.9), whereas few less

conserved regions were found in the surface loops (QMEANDisco < 0.7), indicating increased structural variance (Figure 33).
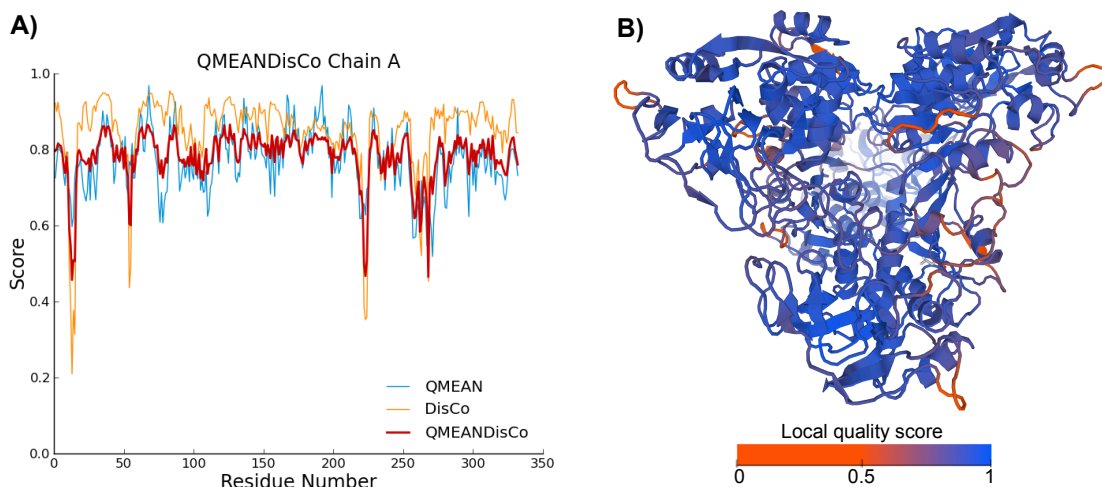

Figure 33: Quality assessment of the homology model. (A) Local quality scores for each amino acid position of the monomer chain A. QMEAN: Quality model energy analysis; Disco: distance constraint score; QMEANDisCo: local quality score, merged scores of QMEAN and Disco (values between 0 and 1, where values below 0.6 are of low quality). (B) Local quality score is shown on the structure model. The calculation and pictures were derived using QMEANDisco tool from QMEAN server [119].

4.1.12 Comparison of ADH/A1a with homologous ADH structures

To find structural features of ADH/A1a for the halo- and thermoadaption, the structure model was compared with homologous ADH structures. The closest identified homologous structures are the thermophilic htADH [141], the mesophilic FurX [143], and the psychrophilic MADH [144] (Table 4). A comparison of the homologs with the ADH/A1a model revealed high structural similarities with root mean square deviations (RMSD) of 0.63 - 0.7 Å, although the sequence identity was comparably low at 34 - 37%. The structure alignment of ADH/A1a with the homolog crystal structures was proper overlapping in almost all regions, except some surface regions (Figure 34).

Table 4: Information of closest homologous ADHs with known crystal structures

Name and PDB ID	Organism	Type	Swiss-model rank	I-Tasser rank	RMSD [Å]	Seq. ident. [%]	Ref.
htADH, 1rjw	*Bacillus stearothermophilus*	thermophile	1	1	0.63	37	[141]
FurX, 3s1l	*Cupriavidus necator*	mesophile	2	3	0.70	36	[143]
MADH, 4z6k	*Moraxella sp.* TAE123	psychrophile	4	2	0.68	34	[144]

The homologs were found using the sequence-based prediction approaches Swiss-model and I-Tasser. The route mean square deviation (RMSD) corresponds to the differences between the residues that are structurally aligned by TM-align of I-Tasser. The sequence identity to ADH/A1a is shown.
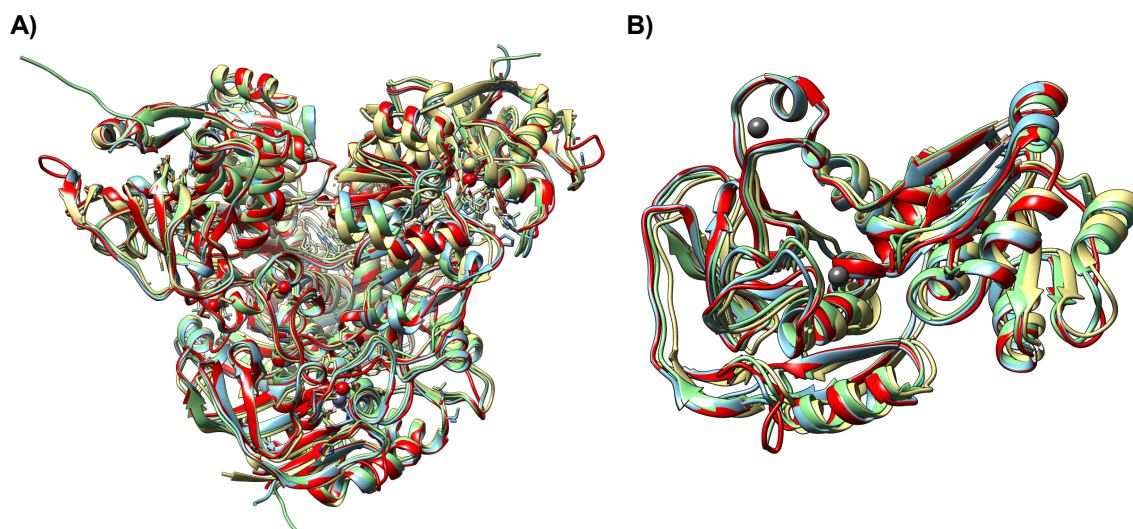
Figure 34: Structure alignment of ADH/A1a model and homologs as ribbon diagram. (A) Alignment of oligomeric structures. (B) Alignment of monomeric structures. The alignment of the structures was done using UCSF Chimera [120]; red: ADH/A1a, light blue: htADH, light yellow: FurX, light green: MADH.

The electrostatic surface potential and the composition of solvent-exposed amino acid of ADH/A1a and homologs were investigated. Among the homologous ADHs, ADH/A1a exhibits the most negative electrostatic surface potential (Figure 35), resulting from a great abundance of solvent-exposed acidic residues, particularly of glutamic acid

(Figure 36A/B). Similarly, but to a lesser extent, the cold-adapted MADH shows an increased negative surface charge, whereby the ratio of solvent-exposed glutamic and aspartic acid is balanced. Furthermore, both ADH/A1a and the thermophilic htADH exhibit increased amounts of charged surface-exposed residues, whereas the ratio of hydrophobic residues is reduced compared. However, ADH/A1a exhibits significant decreased amounts of surface-exposed lysine, in stark contrast to the over-exposure of lysine on the surface of the htADH.
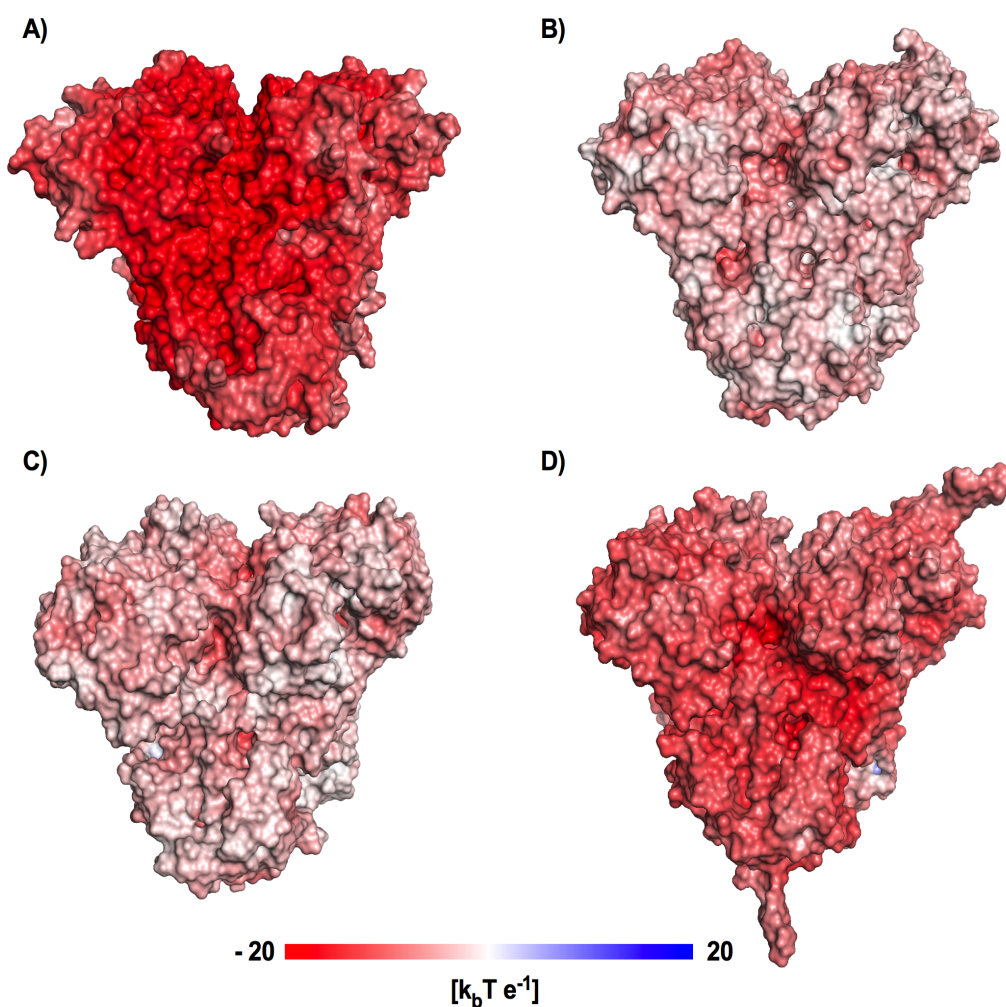
Figure 35: Electrostatic surface charge of ADH/A1a and structure homologs. The solvent-exposed surface of the enzymes is colored from red (negative) to blue (positive). (A) Structure model of ADH/A1a; (B) thermophilic htADH (PDB: 1rjw) from *Bacillus stearo-thermophilus*; (C) mesophilic FurX (PDB: 3s1l) from *Cupriavidus necator*; (D) psychrophilic MADH (PDB: 4z6k) from *Moraxella sp.* [144]. Unit: k_b= Boltzmann constant, T = temperature [K], e = charge of an electron. [K_bTe^{-1}]

Figure 36: Distribution of surface exposed amino acids of ADH/A1a and homologs. (A) Overall distribution of amino acids categorized into acidic, basic, polar and non-polar. (B) Ratio of acidic and basic amino acids of all charged surface residues. The amino acids on the solvent-exposed surface were determined using Swiss PDB viewer 4.1.1 with a surface accessibility of $\geq 30\%$.

4.1.13 Phylogenetic analysis

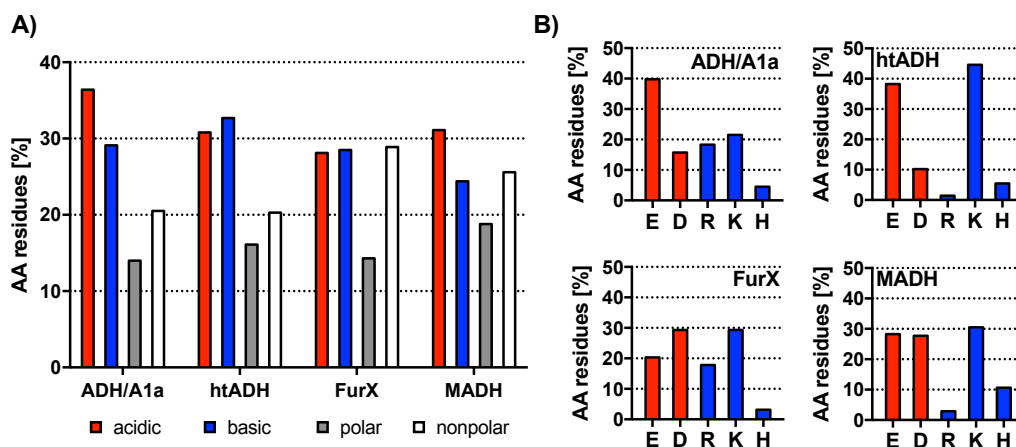The conserved domain database (CDD) of NCBI divided several enzyme families according to their evolutionary conserved protein domain patterns, aiming to provide a functional characterization of distinct subfamilies [114]. As mentioned earlier, ADH/A1a belongs to the CAD family of the MDRs [137]. The CAD family was split into four subfamilies, the CAD1, the CAD2, the CAD3, and the CAD-like subfamily. ADH/A1a was annotated as a member of the CAD2, a subfamily that has not been investigated to date. However, the classification of ADH/A1a as a propanol-preferring ADH (EC1.1.1.1, AdhP) was proposed by the functionality-based annotation tool KEGG (Kyoto Encyclopedia of genes and genomes). Several metabolic pathways were suggested, however, pathways of the classical CADs found in plants (EC1.1.1.195) were not proposed.

To gain deeper insight into the metabolic function of ADH/A1a, a phylogenetic analysis was conducted. A sequence alignment was done, including experimentally confirmed ADHs of the CAD family and of the propanol-preferring Adhp family, as well as uncharacterized hypothetical ADHs of the CAD2 subfamily. The alignment revealed that ADH/A1a shares conserved sequence patterns with both families (Figure 37).

72
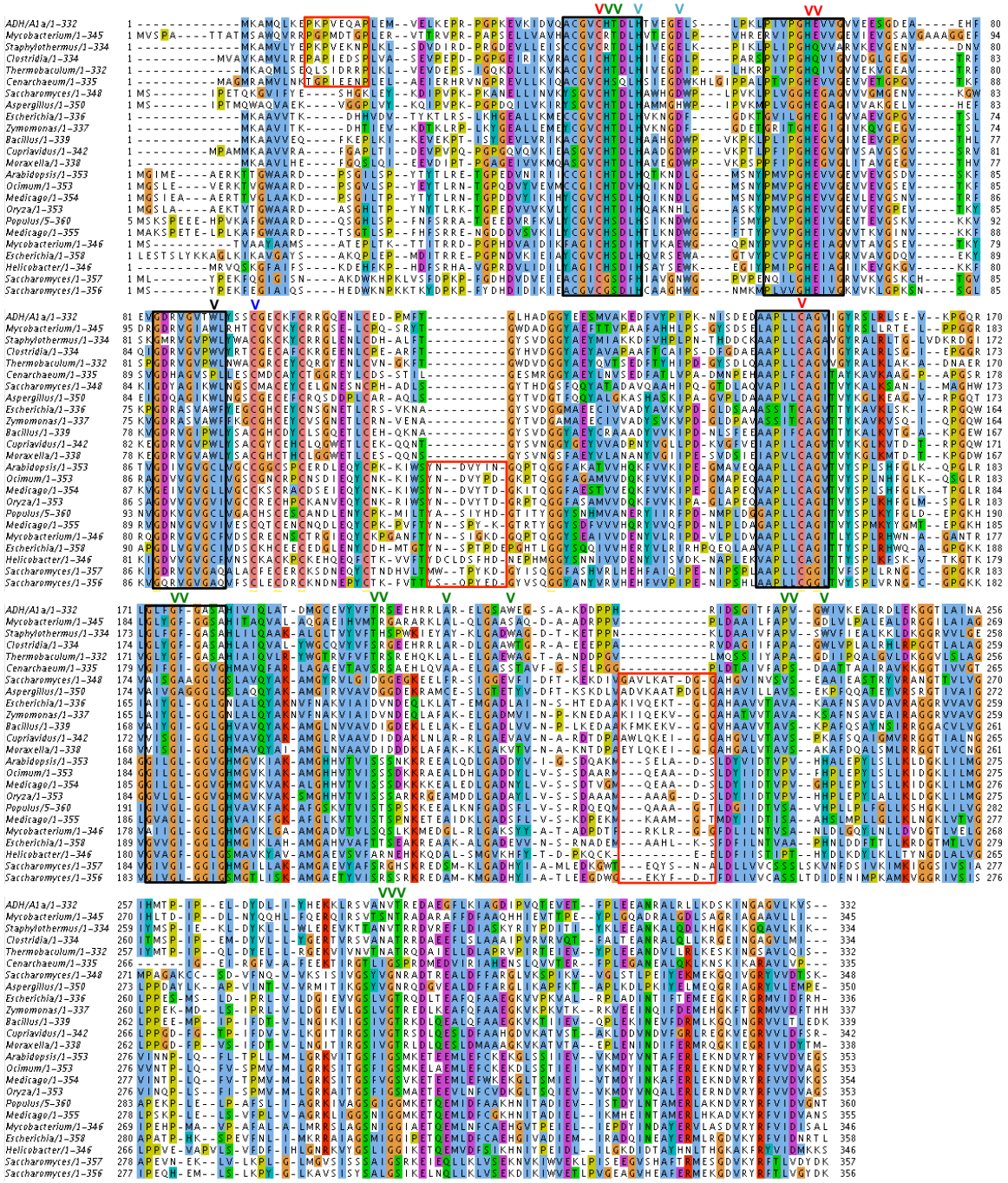
Figure 37: Sequence alignment of selected ADHs of the CAD family and propanol-preferring ADHs. The alignment was done using MAFFT and illustrated using Jalview 1.0 [145]. Highly conserved regions are framed black, unique regions are framed red. Conserved binding sites are highlighted by an arrow (black = substrate, red = catalytic zinc, blue = structural zinc, green = NAD(P)H. The amino acid coloring is according to the Clustal coloring scheme. (Blue = hydrophobic, red = positive charge, magenta = negative charge, green = polar, pink = cysteines, orange = glycines, yellow = prolines, cyan = aromatic, white = unconserved).

Interestingly, a unique extra loop of 4 amino acids at position 9 was exclusively found in the CAD2 subfamily including the ADH/A1a sequence. Small unique insertions (5 - 11 amino acids) were also found for the other subfamilies, as at position 115 in the CAD1 subfamily, and at position 224 in the CAD1, CAD3 and Adhp subfamily.

A phylogenetic tree was build based on the alignment using a maximum-likelihood algorithm, revealing the relations between the different subfamilies (Figure 38). According to that, the CAD2 is closer related to the CAD3 and the ethanol-active ADHs (confirmed propanol-preferring ADHs) than to the CAD1 subfamily. The structure homologs htADH, FurX, and MADH belong to the CAD3 subfamily and were also classified as propanol-preferring ADHs (EC1.1.1.1) by KEGG. CAD3 shares most of its features with ethanol active ADHs and is more closely related to CAD2 than to CAD1, as they form a joint branch. The CAD1 show more diversity as they form separate branches, one for yeast, one for bacteria and one for the classical CADs of plants, whereby the ADH of *Helicobacter pylori* is positioned between bacteria and plants.

The CAD2 subfamily is to date unexplored and comprises several uncharacterized homologous genes from genome annotations of bacteria and archaea, but not from plants or fungi. The investigation of these CAD2 sequences revealed that most of them were derived from extremophilic environments. This evidence corroborates the uniqueness of the CAD2, whose representatives look like propanol-preferring ADHs according to KEGG, but can convert substrates specific for both families, many of them being stable under extremophilic conditions.
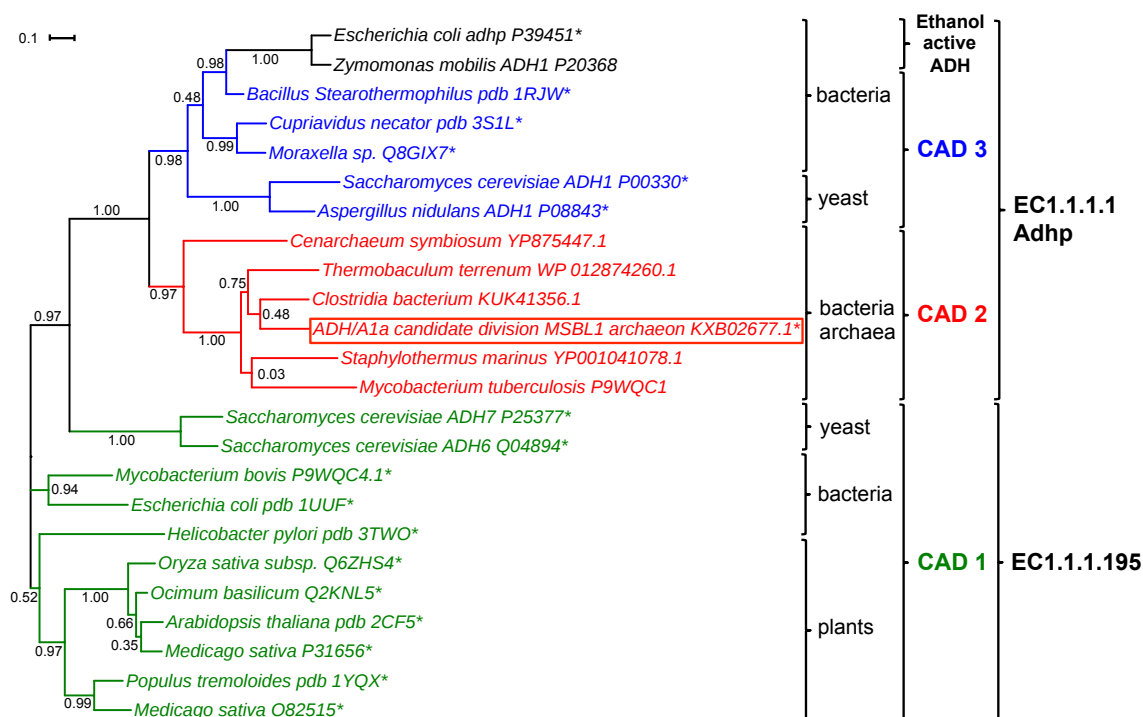
Figure 38: Phylogenetic tree of ADH/A1a and selected ADHs of closest subfamilies.

4.1.14 Proposed reaction mechanism of ADH/A1a

To analyze the reaction mechanism, the substrate and cofactor were modeled into the binding pocket of ADH/A1a followed by and energy minimization process. As a template the crystal structure of FurX with a bound substrate and NADH (PDB: 3s21) was used. The substrate cinnamyl aldehyde lies in the binding pocket of ADH/A1a and binds with the aldehyde group to the threonine T44 and the catalytic zinc(II) (Figure 39). The tryptophan W90 is in close proximity to the phenyl-ring of the substrate and likely attracts and holds the substrate in the binding pocket through hydrophobic interactions. The NADH is attached to the nucleotide-binding domain with the nicotinamide ring in close proximity to the functional aldehyde group of the cinnamyl aldehyde supported by hydrophobic interactions between the nicotinamide ring and the aliphatic substrate backbone. The cofactor-protein binding is mediated through several intermolecular forces. Several salt bridges are formed between the hydroxyl groups of the ribose and the

threonine T44 and histidine H47 as well as of the phosphate group and the histidine H43. Other interactions of NADH and the protein are the π - π stacking with tryptophan W237, the cation - π interaction with the arginine R199, and hydrophobic interactions between several aliphatic residues of the cofactor and protein.
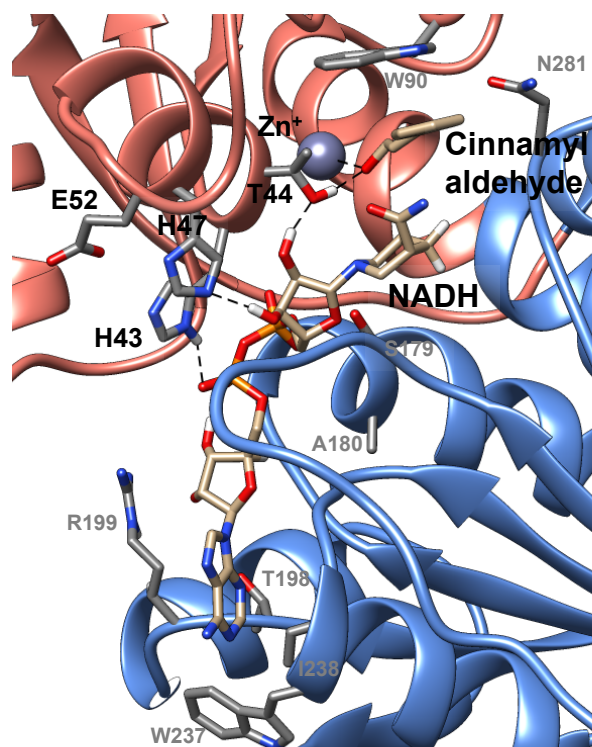
Figure 39: Ribbon diagram representing the ADH/A1a model structure with bound substrate cinnamyl aldehyde and cofactor NADH. The catalytic and nucleotide-binding sites are colored in red and blue respectively, the catalytical zinc(II) is colored in gray. The bound cinnamyl aldehyde and NADH are colored in beige. Important residues for the dihydroxylation are labeled in black, NADH binding residues in gray. The substrate docking studies were done using the substrate and NADH of the crystal structure of the homolog FurX (PDB: 3s21).

Based on the structure model and the proposed mechanism of homologous ADHs, a reaction mechanism for the catalysis of cinnamyl aldehyde to cinnamyl alcohol is proposed as shown in figure 40 [143,146]. In the open state of the ADH, the glutamine E65 coordinates the catalytical zinc(II) in addition to the C42, H64 and C151 binding sites (Figure 40, step **5**). The substrate cinnamyl aldehyde is diffusing into the binding pocket and forms a hydrogen bond with the hydroxyl side chain of the T44 (Figure 40,

step **1**). When the cofactor NADH is binding to the substrate-ADH complex, the ADH is transforming into its closed form (Figure 40, step **2**). The coordinating E65 dissociates from the zinc(II) and is replaced by the carbonyl group of the substrate, building a ternary complex. The T44, the H43 and the H47 form hydrogen bonds to NADH backbone, fixing the position of the NADH and enabling the hydride transfer from the cofactor to the substrate. The T44 and H43 positions are also involved in the proton relay mechanism and shuttle a proton from the aqueous solvent to the carbonyl group of the substrate. This proton shuttling is enabled over a hydrogen network, including E52 and the hydroxyl groups (O2' and O3') of the nicotinamide ribose of NADH. After the reduction of the aldehyde/ketone to the alcohol, the oxidized NAD^+ is leaving the binding site (Figure 40, step **3**) followed by the substrate (Figure 40, step **4**).
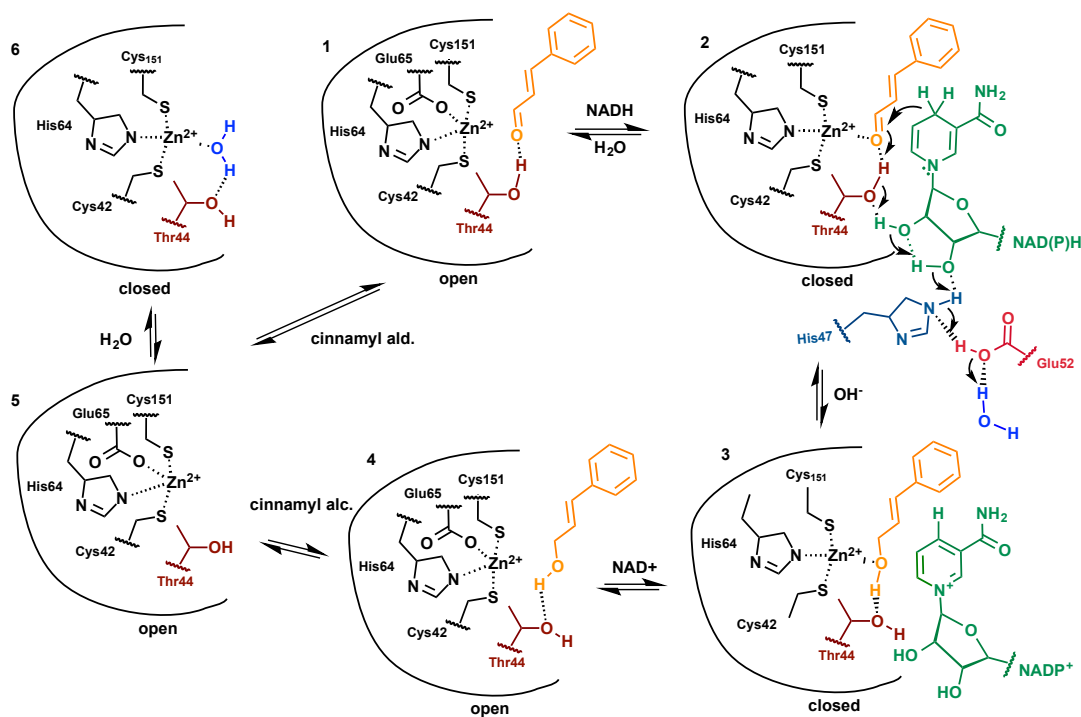
Figure 40: Proposed possible reaction mechanism for the ADH/A1a-mediated reduction of cinnamyl aldehyde into cinnamyl alcohol using the cofactor NADH. In the first step the substrate is binding into the open binding pocket of the enzyme (**1**), followed by the cofactor binding (**2**) and the proton transfer (**2** - **3**). The cofactor leaves the pocket (**3** - **4**) followed by the product (**5** - **6**). The open conformation binds a water molecule and closes (**5** - **6**). Notably, the cofactor may also bind first to the closed form of the ADH (**6**) and the substrate in a second step (not shown here).

4.2   Discussion

Stable extremozymes are in high demand in industrial biocatalysis [32]. The Red Sea brine pools are an attractive source for polyextremophilic enzymes and worth to be studied. However, microorganisms inhabiting and adapted to these extreme environments are barely cultivable. With the improvement of culture-independent methods based on DNA sequencing approaches, this "microbial dark matter" is now open for exploration [147]. The promising ADH gene *adh/a1a* was identified by utilizing annotated SAGs, derived from water samples of the Atlantis II Deep [48]. Already described enzymes, such as the nitrilase NitraS-ATII or the esterase EstATII, derived from the brine pool Atlantis II Deep, showed high thermostability, salt stability, and tolerance for heavy metal contents or organic solvents [50,148-150]. Therefore, the characterization of ADH/A1a was expected to deliver a valuable polyextremophilic ADH that might be utilized for industrial biocatalysis.

The first trial of the recombinant production of ADH/A1a in the mesophilic bacterium *E. coli* was not successful since the expression of halophilic enzymes is more challenging and generally requires high intracellular salt concentrations [138,139]. Haloarchaea in general accumulate KCl at high concentrations in the inner cell to provide osmotic balance to the hypersaline environments [151]; hence, halophilic enzymes are synthesized and natively folded under cytoplasmic high salt conditions. Therefore, the halophilic *Hfx. volcanii* expression system was used to express ADH/A1a under high-salt conditions and could obtain a soluble and active enzyme [34]. The purification of ADH/A1a was challenging as the high salt buffer conditions promoted contaminant host proteins to stick to the enzyme of interest and the column. The decrease of salt in the buffer resulted in highly purified protein; however, the activity was impaired and could not be recovered. The addition of 10% (v/v) glycerol to the purification buffer finally improved the purity of the enzyme respectively, enabling further characterization studies.

The broad characterization of the purified ADH/A1a revealed the halo-thermophilic character of the enzyme reflected by the activity optima at 4 M salt and 70 °C that resembles the natural conditions of the Atlantis II Deep brine pool. The preference for KCl over NaCl in the catalysis buffer was expected, as KCl is the predominant

78

cytoplasmic ion in cells [152]. The pH optima of the reactions performed with ADH/A1a were consistent with several reported ADHs [153]. The highly basic pH optimum of the oxidation reaction (pH 10) does not resemble the natural physiological conditions of a cell, but might be explained by the particular thermodynamics of the reaction. Interestingly, ADH/A1a equivalently converted the cofactors NAD(H) and NADP(H), despite the usual specificity of ADHs for one cofactor [154].

ADH/A1a showed remarkable stability in the presence of high salt concentrations and elevated temperatures. However, low salt conditions destabilized ADH/A1a that was already noticed during the low salt purification. The destabilization and inactivation of the enzyme might be induced by increased structural flexibility and partial unfolding that is promoted by the decline in external ionic forces [155]. Notably, the presence of salt is expected to play an important role in the structural protection of halophilic proteins and may be constantly be needed to maintain the structural robustness [43]. Furthermore, halophilic enzymes are reported to be stable in the presence of organic solvents, in contrast to mesophilic enzymes that generally tend to lose their activity [156,157]. As ADH/A1a is stable under high salt / low water conditions, the enzyme was also found to withstand exposure to organic solvents very well. The stability of ADH/A1a both in solvents and at high temperatures is valuable for biocatalytic reactions in various aspects, e.g., increased solubility of organic compounds or improved reaction rates [158]. ADH/A1a effectively oxidized a broad spectrum of alcohols, favoring long-chain aliphatic and aromatic primary alcohols. In the reduction reaction, exclusively cinnamyl aldehyde, cinnamyl-methyl-ketone, and raspberry ketone were reduced. Though, compared to homologous ADHs relatively slower reaction rates were obtained that might be explained by the fact that life in the depth of the Red Sea brine pools is very slow and high enzyme activities are not crucial for survival. The adaptation to the extreme conditions of the habitat might have shifted towards increased stabilities but simultaneously also to reduced activities [50].

The structural analysis of ADH/A1a was essential as it uncovered the unique features of the extremozyme, promoting the high stabilities to various conditions. Therefore, several crystallization attempts were carried out; however, they were not successful. The

crystallization trails might have failed due to the need of at least 4.8% NaCl (0.8 M) in the buffer to keep the protein structure intact. As protein crystallization generally needs non or minor salt contents (≤ 0.3 M), the successful crystallization might have been despaired either by unfolding of the enzyme at low salt concentrations or by the chaotropic NaCl at high salt concentrations [159]. Conversely, despite the low sequence identity to homologs, a highly reliable ADH/A1a homology model was build based on the significant structural similarity to the structures of homologous ADHs that show an RMSD below 1 Å. As observed already for other extremozymes, the enzyme core domains comprised by the active center and the nucleotide-binding site are highly conserved, whereas the solvent-exposed residues and some flexible regions are altered, enabling the adaptation of the enzyme to the extreme conditions. The solvent-exposed surface of the ADH/A1a tetramer displayed increased amounts of glutamic acid residues, resulting in a high negative electrostatic surface potential than compared with non-halophilic ADHs homologs. An increased negative surface charge is a known adaptation feature of halophilic enzymes and enables the binding of hydrated ions, which stabilize and increase the solubility of the protein under high-salt/low-water conditions [156]. A similar adaptation mechanism that enhances enzyme solvation under cold conditions was reported in psychrophilic enzymes and was also found for the homologous psychrophilic MADH [160]. Furthermore, ADH/A1a exhibits an increased ratio of charged surface residues and a decreased ratio of hydrophobic surface residues, which diminishes hydrophobic surface patches and adds to the haloadaptation [152]. An increased ratio of charged residues and decrease of hydrophobicity in the surface was also reported for thermophilic enzymes, as seen for the homologous htADH [161]. Concluding it may be assumed that the reduction of hydrophobic and the increase of charged residues on the surface adds probably to both, halo- and thermoadaptation of ADH/A1a, making the enzyme polyextremophilic.

The broad substrate scope of ADH/A1a with a preference for cinnamyl alcohol derivatives coincided with the substrates used by ADHs of the CAD family [162]. Based on its conserved domains, ADH/A1a was annotated as a member of the recently annotated CAD2 subfamily. Functional-wise, ADH/A1a was indicated to be of a

propanol-preferring ADH, rather than of a classical CAD. Phylogenetic analysis revealed the positioning of ADH/A1a between these ADH families, showing shared features with both, but also unique features specific for the CAD2 family. In plants, the function of CADs in the lignin biosynthesis and the plant defense mechanisms has been well characterized [146,163]; however, the role of CADs in microorganisms is not yet completely understood. The involvement in several metabolic pathways is considered including the Ehrlich pathway, $NAD(P)^+/H$ homeostasis, lipid metabolism, amino acid metabolism and lignin degradation [164-166]. Due to the lack of experimentally verified CAD2 sequences, further characterizations have to be carried out to investigate the metabolic function of this ADH subfamily. ADH/A1a is the first ADH of the CAD2 subfamily that has been described; thus, its characterization may give the first understanding of their catalytic functions.

Aromatic substrate preferring ADHs, such as CADs, are not well investigated to date, despite their great potential in the production of chiral aromatic precursors for chiral compounds [146], e.g., health-related lignans as podophyllotoxin [167,168]. Especially monoterpenes, including geraniol and nerol, are successfully used as starting material in the production of aroma chemicals used in the food and pharmaceutical industry [19]. The substrate specific reduction of aromatic ketones combined with the high stability makes ADH/A1a a possible candidate for the production of chiral compounds of high interest.

5 Design of a streptavidin-dirhodium ArM for the catalysis of enantioselective carbene transfer reactions

5.1 Results

5.1.1 Design of a biotin-dirhodium ligand and a small intelligent SAV library

The development of an artificial metalloenzyme includes the exploration of a suitable metal-center, which catalyzes reactions under aqueous conditions. Based on the dirhodium(II) Du Bois' catalyst, which was explored for aqueous catalysis by Gillingham, a biotinylated dirhodium catalyst was developed, eligible for the incorporation into the streptavidin host protein [106,169,170]. Ligand docking studies were required to find a suitable spacer for the dirhodium complex into the binding pocket. Thus, the crystal structure of SAV wild-type (wt) with an integrated biotin-ruthenium ligand (PDB: 2QCB) was used as a modeling template, resulting in the synthesis of the suitable Biot-Ar(CO$_2$)$_2$Rh$_2$(OAc)$_2$ ligand (Biot-diRh, **3-f**) (Figure 41A/B) [171]. Hereby, the organometallic Ar(CO$_2$)$_2$Rh$_2$(OAc)$_2$ complex was coupled to a biotin-NHS-ester in such a way that an additional C-atom lies between the anchoring biotin group and the catalytic unit, prolonging the linker-group and giving more space to the catalytic group into the SAV pocket.

The secondary coordination sphere of the binding pocket surrounds the catalytic metal ions and the amino acids in direct proximity influences the enantioselectivity of the reaction. In the last decades, Ward and coworkers have designed various artificial metalloenzymes (ArMs hereafter) based on the biotin-streptavidin technology. Several streptavidin (SAV) libraries were created, containing mutations of surface-exposed residues mainly at the entrance of the protein's binding pocket [89]. The effects of these mutations were well investigated on the efficiency and selectivity of these ArMs and crucial positions in the secondary coordination sphere of the protein were defined [171-173]. To design a small intelligent streptavidin library, the created SAV-ligand docking model was used to investigate the effects of single mutations on the ligand interaction.
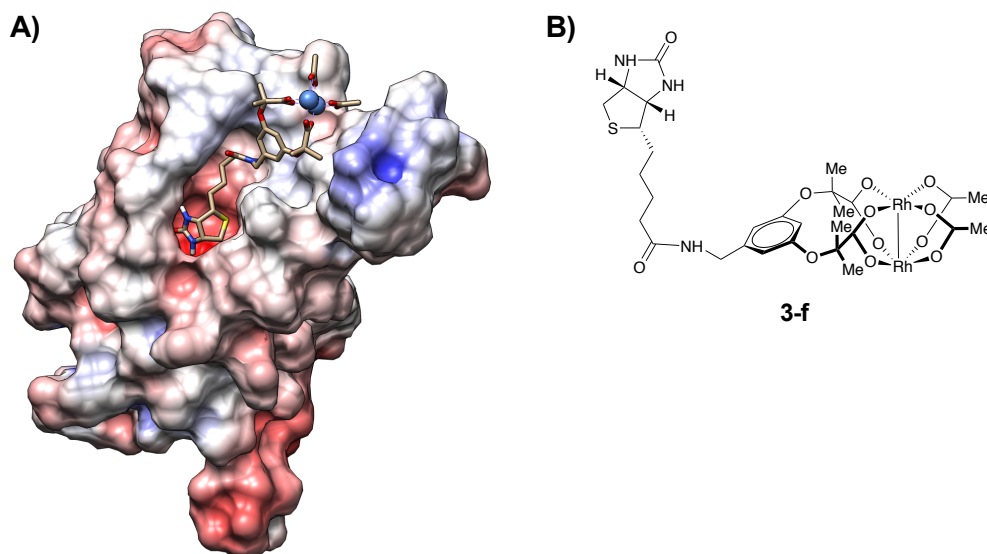
Figure 41: Design of a rhodium(II)-ArM, based on biotin-streptavidin technology. (A) Docking studies of the Biot-Ar(CO$_2$)$_2$Rh$_2$(OAc)$_2$ ligand into the SAV binding pocket. The Docking studies were performed based on the template PDB 2QCB. An energetic minimization of the SAV-Ligand complex was done using Yasara. Coloring: APBS electrostatic surface potential, blue = positive, red = negative, white = neutral (range: -10 to +10 k$_b$Te^{-1}). (B) Chemical structure of the catalytical active Biot-Ar(CO$_2$)$_2$Rh$_2$(OAc)$_2$ ligand (Biot-diRh, **3-f**).

The SAV library was designed to provide mutants with hydrophobic and chiral binding pockets to improve substrate binding and orientation of the organic compounds. A more hydrophobic binding pocket may enhance the local substrate concentration and enable catalysis at low concentrations (< 1 mM) [174,175]. The amino acid positions S112 and K121, located in direct proximity to the metal-complex, have the most influence on enantioselectivity and efficiency in catalysis (Figure 42A/B) [103]. Further position R84 was chosen for substitution based on the reported improvement of turnover number and reduction of substrate inhibition for an engineered Iridium-based ArM [175]. Taking the great size and the hydrophobicity of the Biot-diRh ligand into consideration, small hydrophobic amino acids were considered for substitution, but also residues with polar and charged characteristics were implemented. SAV constructs with single but also multiple mutations at several positions were created (Table 5). Optionally SAV variants with a histidine-free protein surface were created to reduce the unspecific metal binding affinity of the protein and increase the surface hydrophobicity. The two histidines H87

and H127 were substituted with the non-coordinating amino acids. As reported position H87 seems to be important for proper folding of the SAV [172,176]. Based on that, H87 was substituted with a polar asparagine, which might form the necessary hydrogen bond in the antiparallel beta-sheet, stabilizing the correct folding of the protein. The position N49 is important for the biotin binding and, hence, was not considered.
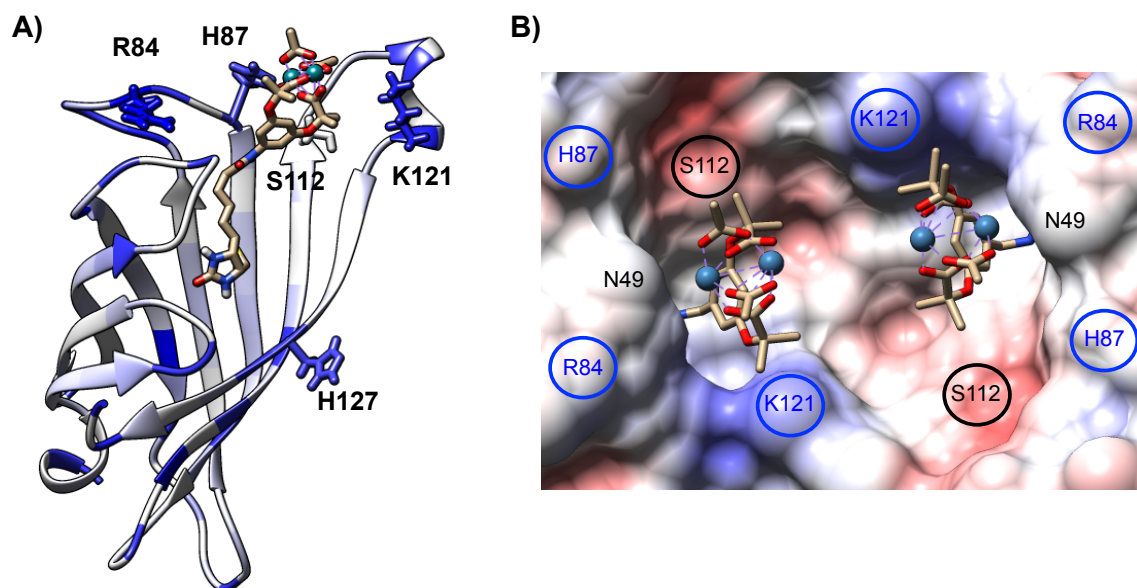
Figure 42: Mutated amino acid positions of SAV. (A) Ribbon diagram of docking model of the SAV monomer with Biot-diRh (beige/blue). The coloring was done with UCSF Chimera after the kdHydrophobicity of the amino acids, with values according to the hydrophobicity scale of Kyte and Doolittle (blue - white = hydrophilic - hydrophobic) [177]. (B) Surface of the docking model of SAV dimer binding pockets with Biot-diRh (beige/blue). The surface coloring was done with the UCSF Chimera colombic surface coloring (blue – basic, white – neutral, red – acidic, range: -10 to +10 kcal/(mol×e) at 298 K).

Table 5: Amino acid positions and mutations of streptavidin.

Position	R84	H87	S112	K121	H127
Mutation	A, Q	N	A, K, Y, F, W, H	A, L, D, E, W H	Y

Amino acids with different properties were integrated. SAV variants with single and multiple mutations were created.

5.1.2 Expression, purification and validation of streptavidin

The expression plasmid pET11 contains the mature SAV wild-type (SAV wt) gene with an N-terminal attached T7 bacteriophage leader sequence (T7-tag). The T7-tag enhances the solubility of the protein during expression and provides higher expression yields [178]. The optimal expression temperature (20 °C, 30 °C, 37 °C) was tested in small scale for SAV wt and several mutants, aiming to gain high yields of soluble protein and to prevent the formation of inclusion bodies (Table 5). The protein was purified by affinity chromatography using an iminobiotin agarose matrix utilizing the pH-dependent affinity of SAV to iminobiotin. The bound SAV was eluted from the column by acidic pH (pH 3 – 4) in one step followed by direct pH neutralization of the elution fractions (Figure 43A). The SDS-PAGE analysis of the purified protein revealed a contaminant band at ~ 20 kDa, which is presumably the biotin-containing acetyl-CoA carboxylase from *E. coli*. However, most of the contaminant protein could be excluded by precipitation during dialysis against dH$_2$O, yielding in a purity of ≥ 98% of soluble SAV (Figure 43B). The amount of free functional binding site per SAV tetramer was determined using the B4F titration assay. The mass of the protein was validated by mass spectrometry (Figure S4).

Good expression yields of soluble and functional SAV were obtained for SAV wt and SAV S112F in shaking flasks (100 - 130 mg L^{-1}), however, the yields of the other S112 and K121 mutants were decreased and showed a higher proportion of inclusion body formation (10 - 70 mg ml^{-1}) (Figure 43C). The yield of SAV, especially of hydrophobic variants, could be successfully improved by the in-batch expression in a 2 L bioreactor (Figure 43D). For the hydrophobic SAV S112A K121L and S112A K121A, a 5-fold and a 7-fold increase of protein yield was obtained. For SAV wt the yield increased by one third, whereas for SAV S112K the yield was almost the same as in the shaking flasks. The free functional binding pockets of SAV were between 60 - 100% for all SAV S112/K121 mutants.
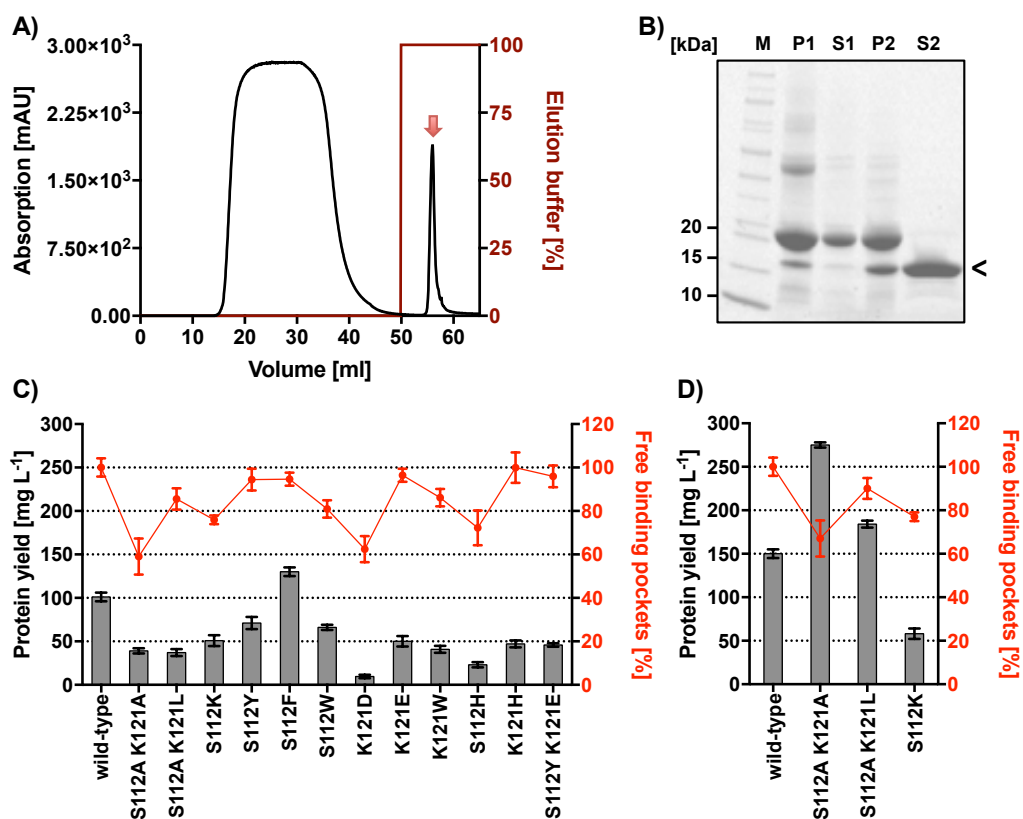
Figure 43: Expression and purification of SAV. (A) Chromatogram of Iminobiotin-affinity purification; the elution peak is highlighted by a red arrow; (B) SDS-PAGE of purified and dialyzed SAV; the SAV band appears at ~16 kDa highlighted by a black arrow; M = marker, P1 = insoluble protein after purification, S1 = soluble protein after purification, P2 = insoluble protein pellet after dialysis in dH_2O, S2 = soluble fraction after dialysis in dH_2O. (C) SAV expressed in a shaking flask. (D) SAV expressed in a bioreactor. Shown is the yield of purified protein and its functional binding sites. Bars - protein yield of purified and lyophilized SAV; points: free functional binding pockets per SAV tetramer determined by B4F titration assay.

The production of functional SAV dHis mutants was more challenging since the protein yield and the number of functional binding pockets significant dropped by every further added mutation (Figure S5). All SAV mutants with altered position R84 showed very low expression yields and the binding of biotin was strongly impaired, making the host scaffold unfeasible. Consequently, several SAV dHis variants were impractical for catalysis and were excluded. The SAV mutants dHis, dHis S112A K121A and dHis S112K provided acceptable functionality, containing 52%, 34% and 43% functional

binding pockets per tetramer with a yield of 35, 9 and 14 mg L^{-1}. However, in this work, these mutants were excluded from catalysis.

5.1.3 Characterization of streptavidin mutants by Circular Dichroism (CD)

The SAV wt shows a distinctive CD-spectrum, consisting of a maximum at 195 nm and 234 nm, and a minimum at 215 nm [179]. The point mutations in the SAV scaffold and binding pocket can alter the secondary structure of SAV, and therefore change the overall stability and biotin-binding affinity of the protein. Extreme structural changes caused by mutation can be revealed by changes in the CD-spectrum. The promising hydrophobic SAV mutants S112A K121A/L and the hydrophilic S112K were analyzed for structural alterations and showed only minor changes in secondary structure (Figure 44A). However, the addition of the double histidine mutations resulted in reduced maxima at 234 nm and changes in the minima at 215 nm, especially those with multiple mutations (Figure 44B). The SAV mutants with altered position R84 showed a significant drop of the maximum at 234 nm, probably resulting from a major structural change, disabling the biotin binding to the SAV pocket (Figure S5). The thermal stability was further investigated by thermal meltdown experiments (chapter 5.1.6).
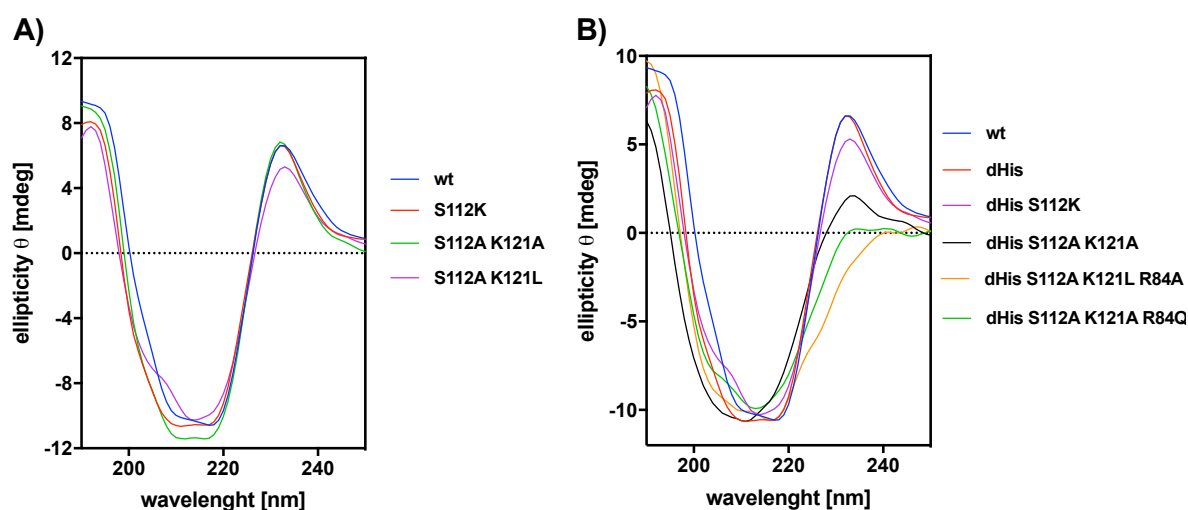

Figure 44: CD-spectra of SAV wt and several mutants. (A) Spectra of SAV with mutations at the binding pocket (S112, K121); (B) Spectra of SAV with dHis mutation (H87, H127) and additional mutation in the binding pocket (S112, K121, R84). As control the spectrum of SAV wt is shown.

87

5.1.4 Chemical modification of streptavidin by PEGylation

To improve the stability and solubility of the SAV scaffold, chemical modification of the protein shell was conducted. The conjugation with polyethylene glycol (PEGs) improves the protein solubility in water as well as in organic solvents, hence, is well suited for protein surface modifications [180,181]. Therefore, activated MS-PEG$_{4,12}$-NHS-esters (Thermo Fisher, Waltham, USA) was utilized for the conjugation of primary amines of the outer SAV surface (Figure 45). The mild PEGylation reaction was carried out in neutral pH overnight.
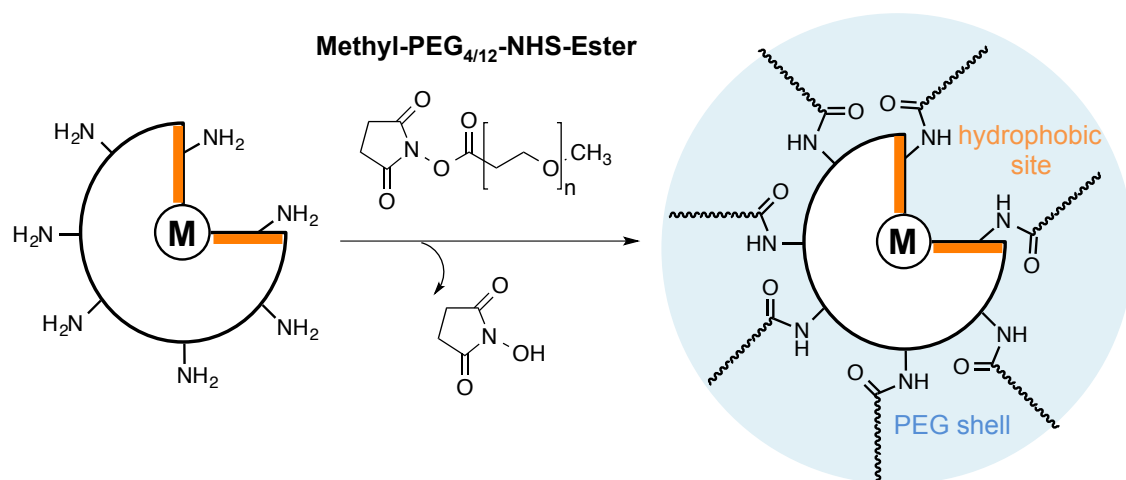
Figure 45: PEGylation reaction of SAV performed in aqueous buffer. The activated MS-PEG$_n$-NHS-esters are reacting with the exposed primary amines of the protein. SAV exhibits seven primary amines, six of exposed lysines and one of the N-terminus. The formed PEG shell around the protein enhances the solubility in water as well as in solvent.

The PEGylation of SAV wt was confirmed by SDS-PAGE and MALDI-TOF analysis (Figure 47). Due to the modification, the tertiary structure of the PEGylated SAV partially maintained under the denaturing conditions of the SDS-PAGE. This fact demonstrates already the enhanced stability of the protein gained by PEGylation. The SDS-PAGE analysis revealed protein bands at the corresponding masses of the SAV-PEG$_4$ and SAV-PEG$_{12}$ monomer as well as of the dimer. For the PEG$_{12}$-modified SAV additionally protein bands at the corresponding masses of the tetramer and oligomer were visible. The blurred protein bands of the SAV-PEG$_{12}$ hint on an incomplete PEGylation

of the protein (Figure 46) as the mass of the protein is dispersed. The corresponding molecular mass was confirmed by MALDI-TOF analysis (Figure 47). The conjugation with PEG_4 added seven PEG-chains to the SAV monomer (17.9 kDa for monomer, 35.8 kDa for dimer), whereas the conjugation with PEG_{12} was incomplete and added between three to five PEG-chains per monomer (18.1 kDa - 19.2 kDa for monomer, 38.5 kDa for dimer). The biotin binding of the PEGylated SAV was tested using the B4F titration assay, revealing an unchanged biotin binding functionality (data not shown).
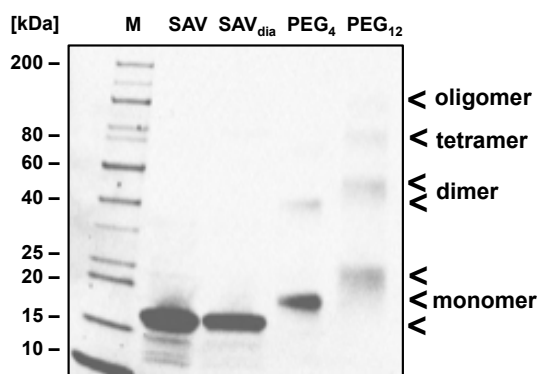
Figure 46: Denaturing SDS-PAGE of PEGylated SAV. M = Protein Marker, SAV = SAV wt, SAV_{dia} = SAV wt dialyzed in reaction buffer, PEG_4 = PEGylated SAV wt with PEG_4 (monomer at ~18 kDa, dimer at ~35 kDa), PEG_{12} = PEGylated SAV wt with PEG_{12} (monomer at ~20 kDa, dimer at ~40 kDa, tetramer at 80 kDa, oligomer < 100 kDa).
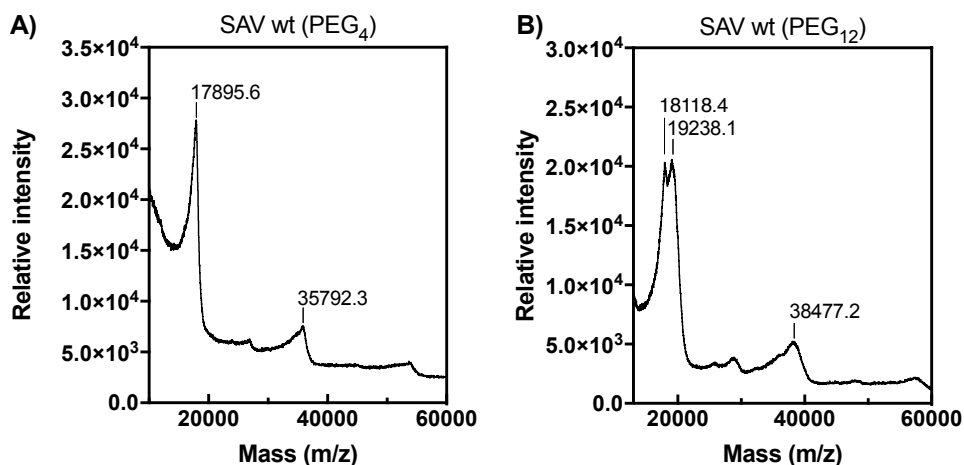
Figure 47: MALDI-TOF analysis of PEGylated SAV wt, Right: SAV wt (PEG_4) shows a major peak around the expected mass for the addition of 7 x MS-PEG_4, a minor peak for the dimer, Left: SAV wt (PEG_{12}) shows two major peaks around the expected mass for the addition of 3 x MS-PEG_{12} and 5 x MS-PEG_{12}, a broad minor peak for the dimer.

89

5.1.5 Solvent stability of PEGylated streptavidin

The solvent stability of SAV and SAV-PEG$_4$ was determined and compared in 20 - 80% of tetrahydrofuran (THF) (complete measurement in Figure S7). The unfolding of the protein structure was followed by CD analysis. The SAV wt was surprisingly stable in dH$_2$O mixed with 20 - 60% (v/v) THF for up to 48 h; however, slow unfolding and precipitation of the protein was visible at 80% (v/v) THF (Figure 48A). In comparison, the SAVwt-PEG$_4$ could withstand 80% (v/v) THF up to 48 h without any structural changes or precipitation, showing that the PEGylated protein is more solvent stable than the SAV wt protein (Figure 48B).
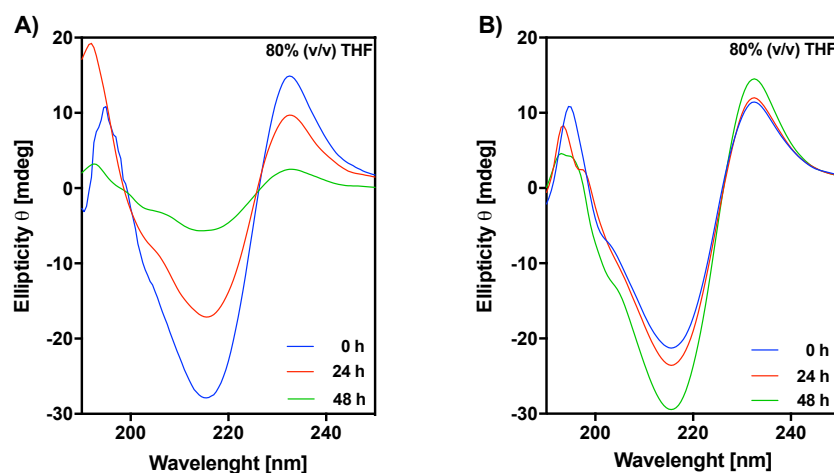

Figure 48: CD-spectra of SAV incubated in 80% THF for 48 h. (A) SAV wt at different times points; (B) PEGylated SAV wt-PEG$_4$ at different times.

5.1.6 Thermal stability of streptavidin mutants

The temperature stability of SAV was determined in thermal denaturation experiments analyzed by CD measurements. The melting temperatures (T$_m$) were withdrawn from the corresponding melting curves of the proteins plotted at 234 nm. The promiscious SAV S112A K121A and S112K variants showed almost no decrease in the T$_m$ compared to the SAV wt (Table 6). However, the functional SAV dHis mutants showed decreased thermal stability, resulting in a drop in the T$_m$ of ~ 25 - 31 °C that indicates an overall destabilization of the protein structure.

Table 6: Melting temperatures of SAV mutants.

SAV	wt	S112A K121A	S112A K121L	S112K	H87N H127Y (dHis)	dHis S112A K121A	dHis S112K
Tm [°C]	72	68	65	72	43	41	47

The chemical modified SAV-PEG$_4$ revealed a significant improvement of the thermal stability, reaching an impressive T_m of 92 °C for the SAV wt and 90 °C for the SAV dHis respectively (Figure 49). Notably, the T_m of the destabilizing dHis mutant was improved by 2.1-fold, proving an overall stabilization effect of the PEGylation on the protein structure. The addition of the Biot-diRh ligand to the PEGylated SAV wt as well as SAV dHis increased the thermal stability exceptionally above ~100 °C.
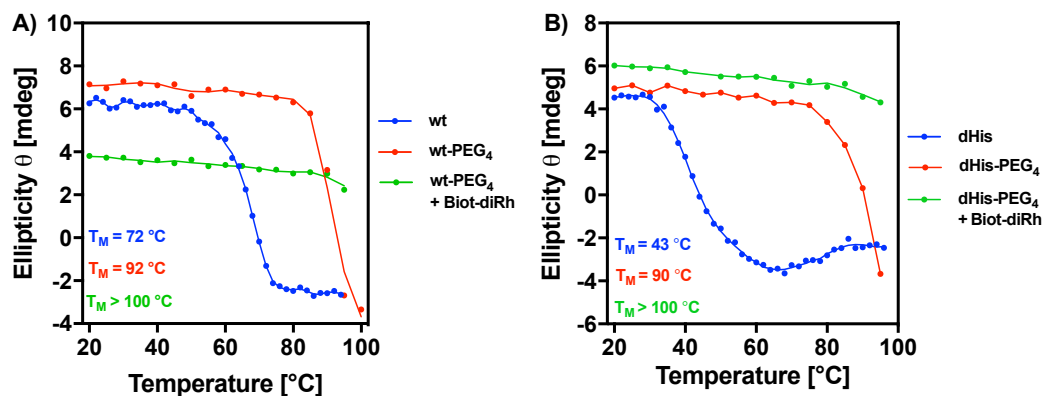
Figure 49: Melting curves of SAV in apo-form, PEGylated, and PEGylated with added Biot-Ar(CO$_2$)$_2$Rh$_2$(OAc)$_2$ ligand (Biot-diRh) measured for (A) SAV wt and (B) SAV H87N H127Y (dHis).

5.1.7 Stability of the streptavidin-dirhodium complex

The stability and solubility of the SAV-dirhodium (SAV-diRh) complex was determined in order to guarantee an intact biocatalyst. The binding of the biotin-dirhodium ligand was tested using the biotin-4-fluorescein (B4F) titration assay [129]. For that SAV was incubated with different amounts of the Biot-diRh ligand (**3-f**). Subsequently, the formed SAV-diRh complex was titrated with the B4F dye in a 96-well scale, incubated for

15 min at room temperature, and the fluorescence emission was measured. The titration assay confirmed the binding of the Biot-diRh ligand (**3-f**) to SAV and further showed no dissociation and displacement with B4F (Figure 50).
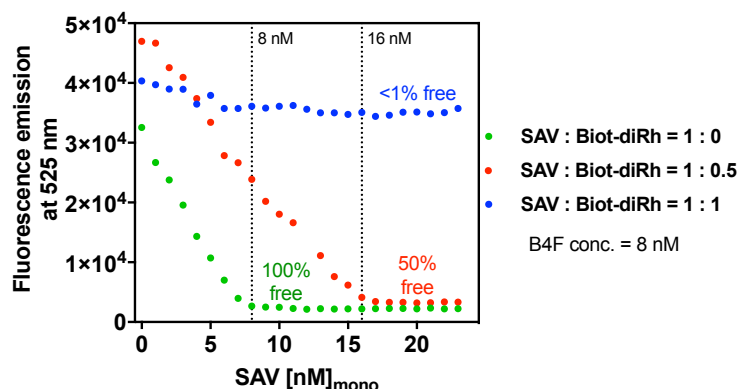
Figure 50: Titration assay of SAV wt with different amounts of Biot-diRh (**3-f**) (protein : ligand = 1 : 0, 1 : 0.5, 1 : 1). Reverse titration with a fixed B4F concentration of 8 nM, titrated with increasing amount of SAV-diRh complex.

The solubility of the SAV-diRh complex was tested in several buffers. Due to the high hydrophobicity of the Biot-diRh ligand (**3-f**) and the increased hydrophobicity of the outer binding pocket of several mutants, the SAV-diRh complex was prone to precipitation in aqueous solutions. The solubility of the SAV-diRh complex was screened in several buffers with varying pH for different SAV mutants (Figure 51A). The SAV wt as well as the hydrophobic SAV S112A K121A/L mutants showed increased precipitation at neutral pH-values, close to the theoretical isoelectric point (pI) of the protein. However, high solubility of the SAV-diRh complex was observed at basic pH 10 for most of the mutants (Figure 51B), except the mutants K121W and S112Y K121E that needed at least pH 11 (Figure S9). Notably, the stepwise addition of the Biot-diRh ligand to the SAV showed a gradual decrease of the solubility at neutral pH, proofing that the protein becomes insoluble upon binding of the ligand (Figure S10).
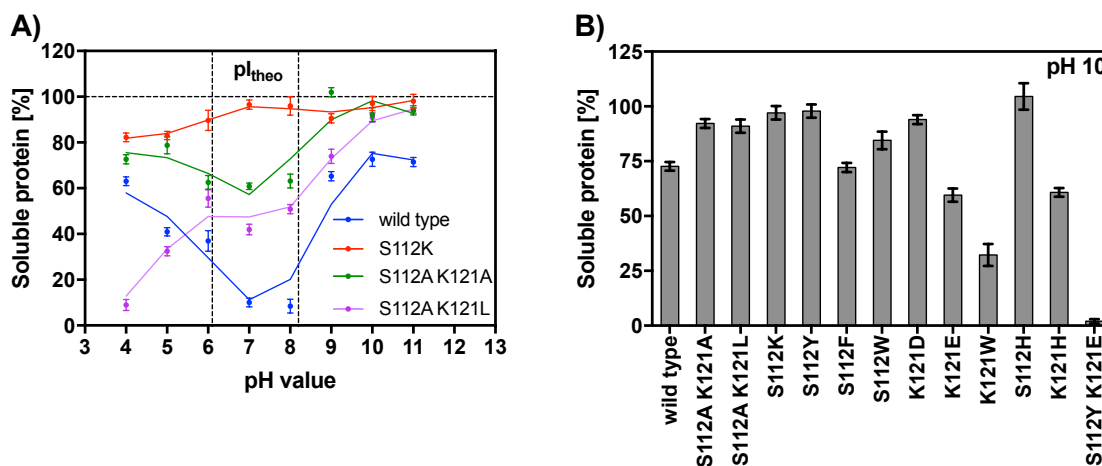
Figure 51: Solubility studies of the SAV-diRh complex. (A) The solubility of SAV mutants incubated with Biot-diRh (**3-f**) (1:1 ratio) at different pH values. (B) The solubility of the SAV-diRh complex of several mutants at pH 10. SAV solution was incubated 1:1 with Biot-diRh (**3-f**) in pH 10 buffer. Soluble fraction was determined by concentration measurements.

Due to some minor precipitation during the catalytic reaction, the stability of the precipitated SAV-diRh complex was checked, to rule out a racemic background reaction catalyzed by released Biot-diRh ligand. For the measurement, the SAV wt was equimolar mixed with the Biot-diRh ligand (**3-f**) in aqueous buffer pH 7, resulting in almost 90% insoluble protein precipitate. The precipitate and the soluble fraction were separated by centrifugation, and the rhodium content of both was determined by ICP-OES measurements. The precipitate contained ~ 88% of the total amount of rhodium versus 7% found in the soluble fraction, proofing the protein-catalyst-complex to be stable even after precipitation (Figure 52).

93
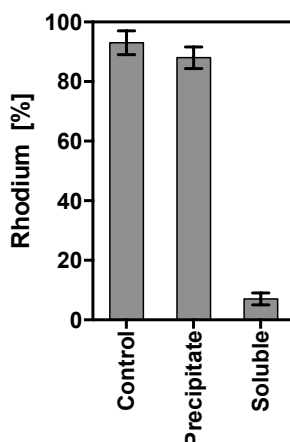
Figure 52: ICP-OES measurement of SAV-diRh complex in pH 7 buffer. The SAV wt in buffer pH 7 was incubated 1:1 with Biot-Ar(CO_2)$_2$Rh$_2$(OAc)$_2$. The formed precipitate and the soluble fraction were separated by centrifugation and analyzed separately by ICP-OES. Control = buffer with Biot-Ar(CO_2)$_2$Rh$_2$(OAc)$_2$ without protein; Precipitate = aggregated SAV-diRh complex; soluble = soluble fraction of SAVwt-Rh$_2$ complex.

5.1.8 Crystallization trials of the SAV-dirhodium complex

The crystallization of the SAV with integrated Biot-diRh ligand was attempted but unfortunately no crystal of the complex was obtained. Several crystallization set-ups were tried including co-crystallization of the SAV-diRh complex with and without crystal seeds as well as crystal soaking of SAV apo crystals with the Biot-diRh ligand, using various crystallization buffers. The co-crystallization of SAV with bound Biot-diRh ligand with and without crystal seeds did unfortunately not result into any protein crystal. The SAV wt apo crystals were grown within 2 days, and various SAV mutants were grown within one or two weeks, providing crystal with the typical rectangular shape with a final size of up to 0.5 mm (Figure 53A). However, in the soaking experiment with the Biot-diRh ligand the crystals were destroyed upon ligand binding, leaving disrupted crystals without a functional diffraction pattern (Figure 53B).
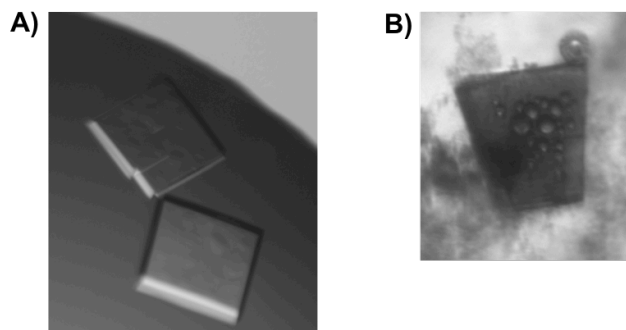
Figure 53: Crystallization trials of SAV with Biot-diRh. (A) Two SAV wt apo crystals. (B) SAV wt crystal after soaking with the Biot-diRh ligand.

5.1.9 Asymmetric cyclopropanation with the SAV-dirhodium complex

The following results of the catalytic reactions were done together with Dr. Yunfei Cai, King Abdullah University of Science and Technology.

The SAV-diRh catalyst was tested in the asymmetric cyclopropanation of the diazo compound *methyl 2-diazo-2-(4-methoxyphenyl) acetate* **4-c** and the 4-methoxystyrene (Figure 54). The reaction was modified following the protocol of Srivastava *et al.* [133]. Beside the desired cyclopropane *methyl 1,2-bis(4-methoxyphenyl)cyclopropane-1-carboxylate* **5-a**, the benzyl alcohol *methyl 2-hydroxy-2-(4-methoxyphenyl)acetate* **5-b** and the corresponding ketone *methyl 2-(4-methoxyphenyl)-2-oxoacetate* **5-c** were formed as side products. The reaction was carried out in 1 ml of the aqueous buffer using 1 mol% of the SAV-diRh complex (1:0.9) with 5 - 10 μmol of the diazo compound **4-c** and an access amount 4-methoxystyrene (3.5 eq.) while stirring. After the reaction the organic compounds were extracted and analyzed by chiral HPLC (Figure 55). The yield was determined using the corresponding standard curves with anisol as an internal standard. In the following, several reaction conditions including buffer /solvent composition, pH-value, temperature and compound concentrations were optimized for the reaction. The effect of the created SAV mutants on the enantioselectivity of the reaction was further determined and compared.

95
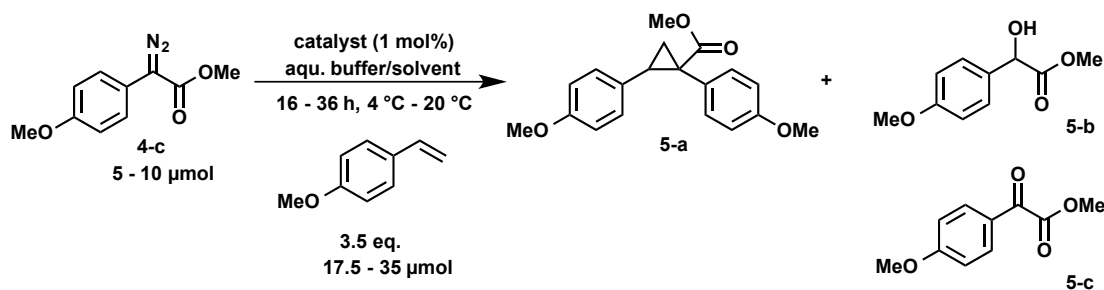
Figure 54: Cyclopropanation reaction with diazo compound **4-c** and 4-methoxystyrene. In the reaction, three products are formed, whereas the cyclopropane **5-a** is the desired compound. The side products benzyl alcohol **5-b** and corresponding ketone **5-c** may also be formed.
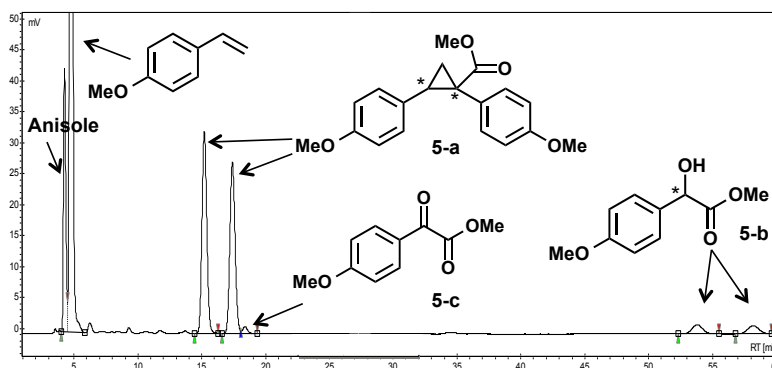
Figure 55: HPLC-chromatogram of separation of the cyclopropane product **5-a** and side-product **5-b**. The access methoxystyrene and the standard anisole are also visible. For the analysis, the compounds were separated by chiral HPLC using a Chiralpak AD-H column.

5.1.10 Effect of solvent and pH-value on the cyclopropanation

The reaction was tested in different ratios of water/THF mixtures, and the side-product formation was analyzed using both Biot-Ar(CO_2)$_2$Rh$_2$(OAc)$_2$ and SAV-diRh complex as catalyst (Table 7). For all catalyzed reaction 100% conversion was reached almost after 3 h. Looking into the yield of the formed products, an unexpected solvent effect of THF was seen, promoting the formation of side products **5-b** and **5-c** (Table 7). With the reduction of the THF to water ratio, the formation of the side products **5-b** and **5-c** decreased and the yield of the desired cyclopropane **5-a** increased significantly. The highest yield was obtained in pure water, reaching 80% yield with the Biot-Ar(CO_2)$_2$Rh$_2$(OAc)$_2$ and 53% yield with the SAVwt-diRh respectively. The influence of

the solvent addition on the product formation was tested with acetonitrile, toluene, hexane, 2-propanol, and methanol; however, an adverse solvent effect was seen for all tested solvents except methanol, which showed comparable results as the catalysis in pure water (Table S6).

Table 7: Product yields of the cyclopropanation at different THF/water ratios.
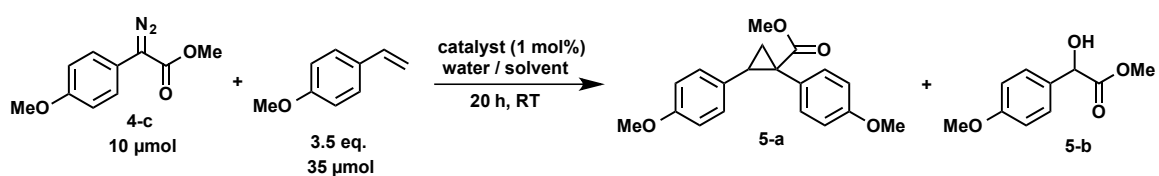
Catalyst	Condition	Yield of (5a)[%]	Yield of (5b) [%]	Yield (5a)/(5b)
Rh_2OAc_4	H_2O	**57**	22	2.59:1
Biot-Ar$(CO_2)_2Rh_2(OAc)_2$	THF	41.1	14.2	2.89:1
Biot-Ar$(CO_2)_2Rh_2(OAc)_2$	80% THF	4.7	47.7	0.1:1
Biot-Ar$(CO_2)_2Rh_2(OAc)_2$	60% THF	9.3	45.4	0.2:1
Biot-Ar$(CO_2)_2Rh_2(OAc)_2$	40% THF	15.7	50.3	0.31:1
Biot-Ar$(CO_2)_2Rh_2(OAc)_2$	20% THF	40.5	46.5	0.87:1
Biot-Ar$(CO_2)_2Rh_2(OAc)_2$	H_2O	**80**	16	**5:1**
SAVwt-diRh	60% THF	3.4	50	0.07:1
SAVwt-diRh	40% THF	5	46.2	0.11:1
SAVwt-diRh	20% THF	5.4	64.2	0.08:1
SAVwt-diRh	10% THF	17.7	59.9	0.30:1
SAVwt-diRh	H_2O	**53.3**	26.4	**2.02:1**
No catalyst	H_2O	25	19.4	1.1:1

The reaction was catalyzed using Biot-Ar$(CO_2)_2Rh_2(OAc)_2$ or SAVwt-diRh complex; Rh_2OAc_4 was used as positive control. All reactions showed 100% conversion. The yield of product **5-c** was < 1%. No enantioselectivity was seen for any reaction.

The effect of the pH-value on the cyclopropanation reaction was investigated using suitable biological buffers ranging from pH 4 to pH 10 and with SAVwt-diRh as catalyst

(**Table 8**). The pH value did not influence the conversion as 100% conversion was obtained for all reactions, which were also seen during the reaction performed with Biot-diRh (Table S7). The overall yield of the diazo compound **5-a** was between ~50% - 60%, except of reactions in borate as the extraction was more problematic and the product was lost. However, the basic pH showed a significant reduction of the side product yield to below 1%.

In acidic and neutral pH the SAVwt-diRh precipitated during protein-metal complex preparation and during the reaction over time. The solubility studies for the protein-metal complex were performed as explained previously (chapter 5.1.7) and revealed a stable and soluble complex in basic pH. Hence, reaction buffers with basic pH 10 or pH 11 were chosen for further reaction studies.

Table 8: Effect of pH-value on the cyclopropanation reaction using different aqueous buffers.
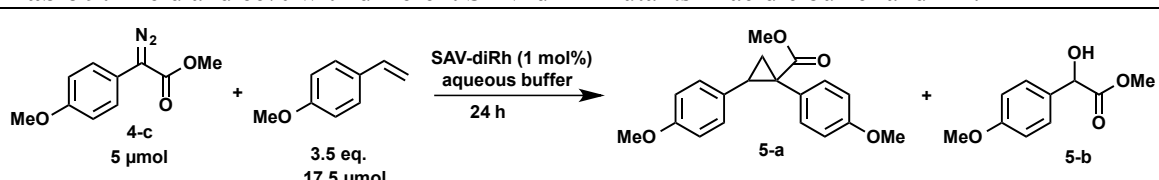
Catalyst	Solvent	Yield **5a** [%]	Yield **5b** [%]	Yield [**5a**]/[**5b**]
SAVwt-diRh	H_2O, pH ~6	53.3	26.4	2.0 : 1
SAVwt-diRh	Na-acetate, pH 4	51.7	36.1	1.4 : 1
SAVwt-diRh	Na-acetate, pH 5	48.9	26.6	1.8 : 1
SAVwt-diRh	Na-carbonate, pH 6	59.5	26.1	2.3 : 1
SAVwt-diRh	Na-carbonate, pH 7	54.3	11.9	4.5 : 1
SAVwt-diRh	Na-carbonate, pH 8	61.5	29.2	2.1 : 1
SAVwt-diRh	Borate, pH 9	21.3*	20.3	1.1 : 1
SAVwt-diRh	Borate, pH 10	30*	0	**1.0 : 0**
SAVwt-diRh	Na-carbonate, pH 10	50	0	**1.0 : 0**

The reaction was performed in 100 mM buffer with varying pH, using 1 mol% SAVwt-diRh as catalyst (24 h, at room temperature). All reaction showed 100% conversion. No enantioselectivity was observed. * Product lost during extraction.

5.1.11 Effect of various SAV mutants on the enantioselectivity

The enantioselective conversion was first tested for the SAV wt and the S112K and S112A K121A/L mutants in acetate buffer pH4. Enantiomeric access (ee%) was found for the cyclopropane **5-a** with all SAV mutants, although just 8 - 15 ee% was obtained (Table 9). Notably, protein precipitation was observed during reaction presumably due to solubility issues, as mentioned before.

Table 9: Yield and ee% with different SAV-diRh mutants in acidic buffer and RT.

Cat. SAV-diRh	Condition	Yield of **5a** (%)	Yield [**5a**]/[**5b**]	ee of **5a** [%]
wt	Acetate, pH 4, RT*	51	1.28 : 1	0
wt-PEG$_4$	Acetate, pH 4, RT*	52	1.41 : 1	0
S112K	Acetate, pH 4, RT*	39	0.66 : 1	9
S112A K121A	Acetate, pH 4, RT*	44	0.92 : 1	**15**
S112A K121A-PEG$_4$	Acetate, pH 4, RT*	43	0.91 : 1	10
S112A K121L	Acetate, pH 4, RT*	31	0.54 : 1	**15**
S112A K121L-PEG$_4$	Acetate, pH 4, RT*	35	0.56 : 1	**17**

The reaction was performed in 50 mM aqueous buffer for 24 h. For the reaction 1 mol% SAV-diRh was used. All reaction showed 100% conversion. The yield of product **5-c** was < 1%. * Protein precipitation occurred during preparation and reaction.

To improve the enantioselectivity, several parameters were changed including buffer, temperature and salt concentration using SAV S112A K121L (Table 10). The change of the buffer to carbonate/borate pH 10 decreased the side product formation significantly (<1 %) and no protein precipitation was observed, as seen for SAVwt-diRh. However, the change to pH 10 did not change the ee%.

The change of the reaction temperature 4 °C had a major effect on the enanatioselectivity of the reaction and successfully increased the ee% for the diazo compound **5-a**, reaching 27 ee% at pH 10 and even 49% at pH 11. However, the addition of NaCl or NaBr had an adverse effect on the ee%, though the overall TON was improved. The PEGylation of the protein did not show a major impact on the overall catalytic performance of the catalyst.

Table 10: Yield and ee% with SAV S112A K121L mutants under various conditions.

Cat. SAV-diRh	Condition	Yield of **5a** (%) / TON	ee of **5a** [%]
S112A K121L	Carbonate, pH 10, RT	55	15
S112A K121L	Borate, pH 10, RT	30	19
S112A K121L	Carbonate, pH 10, 4 °C	23	27
S112A K121L	Borate, pH 10, 4 °C	25	26
S112A K121L-PEG$_4$	Borate pH 10, 4 °C	24	22
S112A K121L	Borate pH 11, 4 °C	23	**49**
S112A K121L	Carbonate pH 10, 4 °C, 1 M NaBr	43	10
S112A K121L	Borate, pH 10, 4 °C, 1 M NaBr	57	10

The reaction was performed in 50 mM aqueous buffer for 24 h. For the reaction 1 mol% SAV-diRh was used. All reaction showed 100% conversion. The yield of product **5-b** and **5-c** was < 1%. Therefore the yield of **5-a** is also the TON (TON = reactant/catalyst x yield).

To investigate the influence of SAV mutations in the binding pocket, the enantioselectivity was tested for several SAV mutants containing hydrophobic but also hydrophilic residues at position S112/K121. The catalytic reactions were carried out in borate buffer pH 10 or 11, as the SAV-diRh complexes were tested to be soluble under basic conditions and the formation of side products **5-b** and **5-c** is < 1%. The catalysis with hydrophilic SAV mutants increased the TON up to 83 (S112K); however, no ee% was obtained. The hydrophobic mutants showed lower TONs compared to the hydrophilic variants, but the desired enantioselectivity for the product was obtained.

Concluding, the highest ee% were obtained with the hydrophobic double mutants S112A K121A/L (Table 11).

Table 11: Yield and ee% in the cyclopropanation catalyzed by different SAV-diRh mutants.
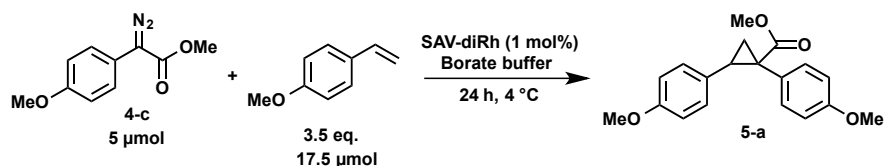
Cat. SAV-Rh$_2$	Residue properties	pH	Yield of **5a** (%) / TON	ee of **5a** [%]
wt	hydrophilic	pH 10	27	0
S112K	hydrophilic	pH 10	**83**	0
S112H	hydrophilic	pH 10	64	1
K121D	hydrophilic	pH 10	22	0
K121E	hydrophilic	pH 11	30	0
K121H	hydrophilic	pH 11	15	0
S112Y	hydrophobic	pH 10	40	-16
S112F	hydrophobic	pH 11	41	-14
S112W	hydrophobic	pH 10	39	-7
K121W	hydrophobic	pH 11	30	10
S112A K121A	hydrophobic	pH 10	17	28
S112Y K121E	hydrophobic	pH 11	10	-10
S112A K121L	hydrophobic	pH 11	23	**49**

The reaction was performed with 1 mol% SAV-diRh in 50 mM borate buffer pH 10/11 for 24 h at 4 °C. All reaction showed 100% conversion. The yield of the side products (**5-b**, **5-c**) were < 1%. Therefore the yield of **5-a** is also the TON (TON = reactant/catalyst x yield)

5.1.12 Effect of reactant concentrations and ratio

Different amounts of starting material were used and the effect on the yield and ee% was studied. The increase of the diazo compound **4-c** and decrease of the methoxystyrene resulted in an overall increase of the yield and of the ee% of the desired cyclopopane **5-a** (Table 12).

Table 12: Yield and ee% with different ratios of starting material.
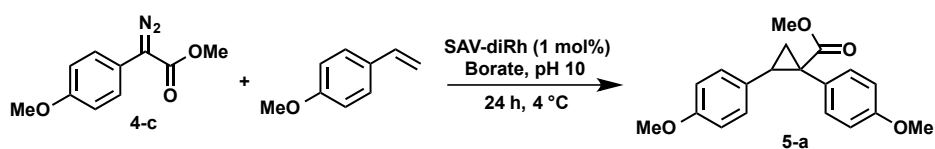
Cat. SAV-Rh$_2$	Diazo compound (4-c) [μmol]	Ratio of (4-c) / methoxystyrene	Yield of 5a [%] / TON	ee of 5a [%]
S112A K121L	5	1/3.5	23	25
S112A K121L	15	3/1	36	30
S112A K121L	25	5/1	40	33

The reaction was performed with 1 mol% SAV-diRh in 50 mM borate buffer pH 10 for 24 h at 4 °C. All reaction showed 100% conversion. The yield of the side products (**5-b**, **5-c**) were < 1%. Therefore the yield of 5-a is also the TON (TON = reactant/catalyst x yield)

5.2   Discussion

Dirhodium(II) carboxylates are highly stable and efficient homogeneous catalysts in carbene transfer reactions and functional under physiological conditions. Especially their low oxidative potential and catalytic functionality inside of cells make them interesting for the utilization for biological chemistry. Therefore, dirhodium(II) carboxylates are attractive complexes for the creation of artificial metalloenzymes (ArM), whereby a biological scaffold is used as secondary coordination sphere for an integrated catalytical dirhodium complex. In this work, the biotin-streptavidin technology was used for the development of an artificial dirhodium cyclopropanase and successfully employed the created ArM in the asymmetric intermolecular cyclopropanation of a diazo-compound.

Starting with the design of the organometallic catalyst, the biotin-linked dirhodium complex $Biot\text{-}Ar(CO_2)_2Rh_2(OAc)_2$ was developed based on ligand-docking studies into the crystal structure of the SAV wt binding pocket. With two Biot-diRh cofactors bound to the SAV dimer, the catalytic axial coordination sites of the two complexes lie face-to-face towards each other in the SAV cavity (Figure 56). The binding pocket of SAV was rationally engineered based on a calculated SAV-diRh model to provide a versatile secondary coordination sphere for the dirhodium complex.

The SAV mutants were screened for soluble expression, functional biotin binding, and protein stability. The focus was set on the promising variants S112A K121A/L as a suitable scaffold for the Biot-diRh ligand, offering enough space and hydrophobicity for the organometallic complex and substrates in the cavity. The expression of the hydrophobic SAV mutants resulted in an increased formation of insoluble inclusion bodies and hence lower protein yields. However, the biotin binding of the soluble fraction was not impaired and functional. Fermentation in a bioreactor significantly improved the yields, making the use of the hydrophobic SAV variants feasible. All SAV mutants with mutations at the propitious positions S112 and K121 showed similar expression yields and possessed approximately 3-4 free binding pockets, which is well acceptable for the catalytic purpose.

The SAV variants with double histidine mutations showed a structural instability resulting in lower expression yield and biotin binding capability that dropped with each

added mutation. As previously reported, the position H87 may be critical for proper folding [175]. Additional mutation R84A/Q resulted in total loss of biotin binding, which excluded those mutants as host scaffold. The multiple point mutations probably resulted in a collapse of the loop and the short α-helix at the opening of the binding pocket. The structural distortion was observed via CD-measurement indicated by a reduced signal at 234 nm, whereby the typical minimum at 219 nm for β-sheets was still intact.

The stability of the SAV scaffold was successfully enhanced by the chemical conjugation of the exposed primary amines with PEG_4/PEG_{12}-chains. The PEGylation significantly increased the solvent stability and the thermostability of SAV mutants, whereas the binding of the Biot-diRh ligand was not hindered. The PEGylated SAV was extraordinary stable in 80% (v/v) THF for up to 48 h as well as in temperatures up to 92 °C and under denaturing conditions of an SDS-PAGE. Notably, the melting temperature of the fragile SAV dHis mutant (H87N H127Y) increased from 42 °C up to 90 °C. Upon binding of the Biot-diRh ligand, the PEGylated SAV-ligand complex withstood even a temperature of 100 °C, probably promoted through the additional stability caused by the closed form of the SAV [182].

The Biot-diRh ligand binds completely to all four binding pocket of the SAV wt tetramer and the high-affinity complex did not dissociate over time, resembling a stable artificial metalloenzyme with four catalytic dirhodium sites. However, the high hydrophobicity of the organometallic complex and the additionally increased hydrophobicity of several SAV mutants caused precipitation of the SAV-diRh complex at operational concentrations at pH values near the pI of the protein. Therefore, the solubility of all SAV mutants with bound Biot-diRh were screened and found to be most soluble at basic pH values around pH 10 or 11. The insoluble precipitate of SAV-diRh was additionally tested and found to be stable as the Biot-diRh remained in the SAV cavity.

Unfortunately, the crystallization trials did not provide a crystal of the SAV-diRh complex. The high hydrophobicity of the Biot-diRh ligand as well as the flexibility in the binding pocket might hinder the successful co-crystallization. Soaking experiments of SAV crystals with the Biot-diRh ligand resulted in cracked crystals, probably due to a major structural change in the crystal upon binding. However, other ArMs with bound

dirhodium complexes seem to have similar crystallization issues, as none crystal structure was resolved yet [183,184].

To test the catalytic functionality of the created ArMs, the SAV-diRh was employed in the asymmetric cyclopropanation reaction. The reaction conditions were optimized regarding temperature, ratio of the reactants and reaction medium. The highest yields of the desired cyclopropane **5-a** were found in aqueous solution using SAV-diRh, Biot-diRh or Rh_2OAc_4. The formation of the side products **5-b** and **5-c** were promoted with rising organic solvent concentrations. The optimal pH-value for the reaction was at alkaline pH 10 and 11, as the SAV-diRh complex is most soluble and the yield of the side products was diminished to < 1%. Furthermore, the reduction of the reaction temperature to 4 °C increased the formation of enantiomers, while using selective SAV-diRh variants.

Several SAV mutants with substituted positions S112 and K121 were tested for enantioselective conversion. As expected, the hydrophobic mutations promoted enantioselectivity, with the promising SAV S112A K121A/L showing the highest ee, with 49% for S112A K121L and 28% for S112A K121A respectively. The S112Y and S112F mutants reached an ee of -16% and -14% for the corresponding enantiomeric opponent. Surprisingly, hydrophilic mutants resulted in higher yields of the cyclopropane **5-a**; however, no or minor enantioselectivity was obtained. Furthermore, the 5-fold access of the diazo compound **4-c** to 4-methoxystyrene improved the yield as well as the ee% of the desired product **5-a**.

The enantioselectivity of the SAV mutants was significantly influenced by the positions S112/K121 and the hydrophobicity of the cavity (Figure 56). The organometallic complex seems to be quite flexible and sticks out of the cavity due to the prolonged linker to the biotin. The SAV wt is composed of hydrophilic positively and negatively charged residues that might promote the flexibility of the complex, resulting in racemic catalysis. In contrast, the mutants S112A K121A/L provide a more hydrophobic and more negatively charged binding pocket that might interact with the organometallic complex through hydrophobic interactions, causing a more rigid and selective catalytic site. Moreover, the S112A K121L mutant has a smaller cavity than the S112A K121A variant that promoted higher enantioselectivity. On the other side, the enantioselective S112Y/F

mutants promoter the formation of the opposite enantiomer probably due to reduced cavity size and a stabilization of the organometallic complex through hydrophobic interaction with the phenyl groups of the residues.
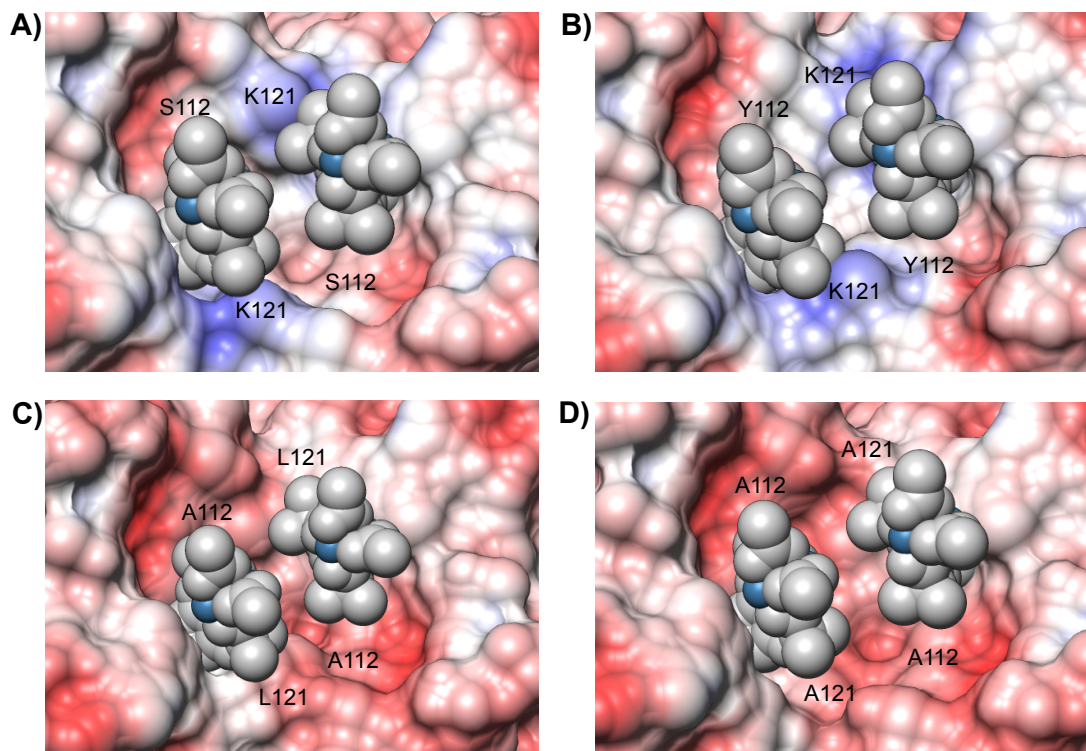
Figure 56: Binding pockets of modeled SAV-diRh variants. (A) SAV wt, (B) SAV S112Y, (C) SAV S112A K121L , and (D) SAV S112A K121A. The surface of SAV dimers with bound Biot-diRh that is shown in spheres (gray/blue). The surface coloring was done with the UCSF Chimera colombic surface coloring (blue – basic, white – neutral, red – acidic, range: -10 to +10 kcal/(mol×e) at 298 K).

In comparison to the recently described SAV-diRh complex designed by *Zhao et al.*, the here described system provides enantioselectivity that could not be delivered by the reported system. The major difference between both designs is the structure of the Biot-diRh cofactor, as the here described ligand provides an additional C-atom in the linker between the biotin and the dirhodium(II) carboxylate complex. Hence, the $Ar(CO_2)_2Rh_2(OAc)_2$ has more space into the SAV cavity and shows a vertical ordered conformation, providing just one free axial coordination site per dirhodium complex than compared with the more horizontally ordered cofactor of *Zhao et al.* [184].

The reported artificial cyclopropanase that is based on the Pfu POP scaffold with a covalently linked dirhodium cofactor provided high enantioselectivity up 94 ee% after directed evolution by random mutagenesis, with yields similar to results described here [108,183]. The evolution of the SAV scaffold may therefore further improve the enantioselectivity of the SAV-diRh complex. However, the easy production and effortless coupling of the dirhodium complex through supramolecular anchoring as well as the tunable enantioselectivity makes the created dirhodium ArM to a powerful hybrid catalyst for carbene transfer reactions that can easily compete and outperform existing systems. The performance in alternative reaction as the asymmetric C-H insertion, the aziridination of olefins or the intramolecular C-H insertion of diazo-compounds would be of major interest for future applications.

6 Conclusion

In this work, two different approaches were successfully utilized to gain novel biocatalyst with promising properties. The discovery of novel extremozymes based on SAGs derived from the hypersaline brine pools of the Red Sea has proven to be an abundant method for the implementation of highly stable biocatalysts [50]. Here, a novel polyextremophilic zinc-dependent medium-chain alcohol dehydrogenase was investigated that showed high stability under various industrial conditions and a versatile substrate scope that may be utilized for the synthesis of chiral compounds. Furthermore, structural adaptations of the halo- and thermoadaptation were uncovered and phylogenetic relations to similar ADHs investigated.

The engineering of a dirhodium ArM based on the biotin-streptavidin technology provided a tunable catalyst for the asymmetric carbene transfer reactions. The streptavidin host scaffold was rationally designed for the enantioselective catalysis using a biotin-anchored dirhodium(II) carboxylate complex as a catalytical unit. The engineered SAV-dirhodium catalyst was applied in the asymmetric cyclopropanation reaction and showed enantioselectivity, whereby it outperformed existing SAV-based systems. The structural stability of SAV was further increased by PEGylation of the protein surface, making the artificial dirhodium(II) enzyme extraordinary stable under high temperatures and organic solvents. The easy affinity-based coupling of the dirhodium cofactor, the high production yield and the tunability of the secondary coordination sphere in the SAV cavity makes the here developed dirhodium ArM system competitive and versatile.

7 References

1. Ertl G KH, Schüth F, Weitkamp J (2008) Handbook of heterogeneous catalysis. Vol. **2nd**. Weinheim: *Wiley-VCH*.
2. Anastas P. T. CRH (2009) Handbook of green chemistry Vol. **1**. *Wiley-VCH*.
3. Rothenberg G (2008) Introduction. in *Catalysis - Concepts and green applications*, ed. G. Rothenberg, John Wiley & Sons. p. 1-42.
4. Souza Porto MF and Freitas CM (1996) Major Chemical Accidents in Industrializing Countries: The Socio‐Political Amplification of Risk. *Risk Anal* **16**, 19-29.
5. Clark JH (2006) Green chemistry: today (and tomorrow). *Green Chem* **8**, 17-21.
6. Rothenberg G (2008) Basics of catalysis. in *Catalysis: Concepts and green applications*, ed., John Wiley & Sons: Weinheim, Germany. p. 43 - 88.
7. Thomas JM and Williams RJ (2005) Catalysis: principles, progress, prospects. *Philos Trans A Math Phys Eng Sci* **363**, 765-791.
8. Weckhuysen BM (2002) Snapshots of a working catalyst: possibilities and limitations of in situ spectroscopy in the field of heterogeneous catalysis. *Chem Commun* 97-110.
9. Crabtree RH (2009) The organometallic chemistry of the transition metals. *Wiley-VCH*
10. Rothenberg G (2008) Homogeneous catalysis. in *Catalysis - Concepts and green applications*, ed. G. Rothenberg, John Wiley & Sons: Weinheim. p. 89-136.
11. Lewis JC (2013) Artificial metalloenzymes and metallopeptide catalysts for organic synthesis. *ACS Catalysis* **3**, 2954-2975.
12. Sharpless KB (2002) Searching for new reactivity (Nobel lecture). *Angew Chem Int Ed Engl* **41**, 2024-2032.
13. William SK (2002) Asymmetric hydrogenations (Nobel Lecture). *Angew Chem Int Ed* **41**, 1998-2007.
14. Noyori R (2002) Asymmetric catalysis: science and opportunities (Nobel lecture). *Angew Chem Int Ed Engl* **41**, 2008-2022.
15. Chauvin Y (2006) Olefin metathesis: the early days (Nobel Lecture). *Angew Chem Int Ed Engl* **45**, 3740-3747.
16. Grubbs RH (2006) Olefin-metathesis catalysts for the preparation of molecules and materials (Nobel Lecture). *Angew Chem Int Ed Engl* **45**, 3760-3765.
17. Schrock RR (2006) Multiple metal-carbon bonds for catalytic metathesis reactions (Nobel Lecture). *Angew Chem Int Ed Engl* **45**, 3748-3759.
18. Sheldon RA (2016) Biocatalysis and green chemistry. in *Green biocatalysis*, ed. R.N. Patel, Wiley-VCH p. 1-15.
19. Alcántara A, Hernáiz M, and V. Sinisterra J (2011) Biocatalyzed production of fine chemicals. in *Comprehensive Biotechnology*, ed. M. Moo-Young, Elsevier B.V. p. 309-331.
20. Pyne M, Abedi D, Zhang L, and Perry Chou C (2011) Enzyme biocatalysis. in *Comprehensive biotechnology*, ed. M. Moo-Young, Elsevier. p. 15-24.
21. Jaeger K-E and Eggert T (2004) Enantioselective biocatalysis optimized by directed evolution. *Curr Opin Biotechnol* **15**, 305-313.
22. Li S, Yang X, Yang S, Zhu M, and Wang X (2012) Technology prospecting on enzymes: application, marketing and engineering. *Computational and Structural Biotechnology Journal* **2**, e201209017.
23. Tao J and Xu JH (2009) Biocatalysis in development of green pharmaceutical processes. *Curr Opin Chem Biol* **13**, 43-50.
24. R Berlemont CG (2011) Extremophiles. in *Comprehensive biotechnology*, ed. M. Moo-Young, Elsevier B.V.: Burlington. p. 229-242.
25. Ferrer M, Golyshina O, Beloqui A, and Golyshin PN (2007) Mining enzymes from extreme environments. *Curr Opin Microbiol* **10**, 207-214.

26. Sarmiento F, Peralta R, and Blamey JM (2015) Cold and hot extremozymes: Industrial relevance and current trends. *Frontiers in Bioengineering and Biotechnology* **3**, 148.
27. Burton SG, Cowan DA, and Woodley JM (2002) The search for the ideal biocatalyst. *Nat Biotech* **20**, 37-45.
28. Coker JA (2016) Extremophiles and biotechnology: current uses and prospects. *F1000Research* **5**, F1000 Faculty Rev-396.
29. Tindall KR and Kunkel TA (1988) Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* **27**, 6008-6013.
30. Lundberg KS, Shoemaker DD, Adams MW, Short JM, Sorge JA, and Mathur EJ (1991) High-fidelity amplification using a thermostable DNA polymerase isolated from Pyrococcus furiosus. *Gene* **108**, 1-6.
31. Ishino S and Ishino Y (2014) DNA polymerases as useful reagents for biotechnology - the history of developmental research in the field. *Front Microbiol* **5**, 465.
32. Dumorne K, Cordova DC, Astorga-Eló M, and Renganathan P (2017) Extremozymes: a potential source for industrial applications. *J Microbiol Biotechnol* **27**, 649-659.
33. Allers T, Barak S, Liddell S, Wardell K, and Mevarech M (2010) Improved strains and plasmid vectors for conditional overexpression of His-tagged proteins in *Haloferax volcanii. Appl Environ Microbiol* **76**, 1759-69.
34. Strillinger E, Grotzinger SW, Allers T, Eppinger J, and Weuster-Botz D (2016) Production of halophilic proteins using *Haloferax volcanii* H1895 in a stirred-tank bioreactor. *Appl Microbiol Biotechnol* **100**, 1183-95.
35. Ye Y, Osterman A, Overbeek R, and Godzik A (2005) Automatic detection of subsystem/pathway variants in genome analysis. *Bioinformatics* **21 Suppl 1**, i478-86.
36. Dionisi HM, Lozada M, and Olivera NL (2012) Bioprospection of marine microorganisms: biotechnological applications and methods. *Rev Argent Microbiol* **44**, 49-60.
37. Antunes A, Ngugi DK, and Stingl U (2011) Microbiology of the Red Sea (and other) deep‐sea anoxic brine lakes. *Environ Microbiol Rep* **3**, 416-433.
38. De Vitis V, Guidi B, Contente ML, Granato T, Conti P, Molinari F, Crotti E, Mapelli F, Borin S, Daffonchio D, and Romano D (2015) Marine Microorganisms as source of stereoselective esterases and ketoreductases: Kinetic resolution of a prostaglandinintermediate. *Mar Biotechnol* **17**, 144-152.
39. Flam F (1994) The chemistry of life at the margins. *Science* **265**, 471-473.
40. Mevarech M, Frolow F, and Gloss LM (2000) Halophilic enzymes: proteins with a grain of salt. *Biophys Chem* **86**, 155-164.
41. Lanyi JK (1974) Salt-dependent properties of proteins from extremely halophilic bacteria. *Bacteriol Rev* **38**, 272.
42. Oren A (2013) Life at high salt concentrations, intracellular KCl concentrations, and acidic proteomes. *Front Microbiol* **4**, 315.
43. Sinha R and Khare SK (2014) Protective role of salt in catalysis and maintaining structure of halophilic proteins against denaturation. *Front Microbiol* **5**, 165.
44. Mwirichia R, Alam I, Rashid M, Vinu M, Ba-Alawi W, Anthony Kamau A, Kamanda Ngugi D, Goker M, Klenk HP, Bajic V, and Stingl U (2016) Metabolic traits of an uncultured archaeal lineage -MSBL1- from brine pools of the Red Sea. *Sci Rep* **6**, 19181.
45. Hendricks RL, Reisbick FB, Mahaffey EJ, Roberts DB, and Peterson MN (1969) Chemical composition of sediments and interstitial brines from the Atlantis II, Discovery and Chain Deeps. in *Hot brines and recent heavy metal deposits in the Red Sea*, ed., Springer. p. 407-440.
46. Miller AR, Densmore CD, Degens ET, Hathaway JC, Manheim FT, McFarlin PF, Pocklington R, and Jokela A (1966) Hot brines and recent iron deposits in deeps of the Red Sea. *Geochim Cosmochim Acta* **30**, 341-359.
47. Backer H and Schoell M (1972) New deeps with brines and metalliferous sediments in the Red Sea. *Nature Physical Science* **240**, 153-158.

48. Alam I, Antunes A, Kamau AA, Ba Alawi W, Kalkatawi M, Stingl U, and Bajic VB (2013) INDIGO - Integrated data warehouse of microbial genomes with examples from the red sea extremophiles. *PLoS One* **8**, e82210.

49. Grotzinger SW, Alam I, Ba Alawi W, Bajic VB, Stingl U, and Eppinger J (2014) Mining a database of single amplified genomes from Red Sea brine pool extremophiles-improving reliability of gene function prediction using a profile and pattern matching algorithm (PPMA). *Front Microbiol* **5**, 134.

50. Grotzinger SW, Karan R, Strillinger E, Bader S, Frank A, Al Rowaihi IS, Akal A, Wackerow W, Archer JA, Rueping M, Weuster-Botz D, Groll M, Eppinger J, and Arold ST (2018) Identification and experimental characterization of an extremophilic brine pool alcohol dehydrogenase from single amplified genomes. *ACS Chem Biol* **13**, 161-170.

51. Wang Y, Cao H, Zhang G, Bougouffa S, Lee OO, Al-Suwailem A, and Qian PY (2013) Autotrophic microbe metagenomes and metabolic pathways differentiate adjacent Red Sea brine pools. *Sci Rep* **3**, 1748.

52. Zanotti-Gerosa A, Hems W, Groarke M, and Hancock F (2005) Ruthenium-catalysed asymmetric reduction of ketones. *Platinum Met Rev* **49**, 158-165.

53. Zheng Y-G, Yin H-H, Yu D-F, Chen X, Tang X-L, Zhang X-J, Xue Y-P, Wang Y-J, and Liu Z-Q (2017) Recent advances in biotechnological applications of alcohol dehydrogenases. *Appl Microbiol Biotechnol* **101**, 987-1001.

54. Patel RN (2013) Biocatalytic synthesis of chiral alcohols and amino acids for development of pharmaceuticals. *Biomolecules* **3**, 741-777.

55. Musa MM and Phillips RS (2011) Recent advances in alcohol dehydrogenase-catalyzed asymmetric production of hydrophobic alcohols. *Catalysis Science & Technology* **1**, 1311-1323.

56. Zhang R, Xu Y, and Xiao R (2015) Redesigning alcohol dehydrogenases/reductases for more efficient biosynthesis of enantiopure isomers. *Biotechnol Adv* **33**, 1671-1684.

57. Rossmann MG, Moras D, and Olsen KW (1974) Chemical and biological evolution of nucleotide-binding protein. *Nature* **250**, 194-9.

58. Kallberg Y, Oppermann U, Jornvall H, and Persson B (2002) Short-chain dehydrogenases/reductases (SDRs). *Eur J Biochem* **269**, 4409-17.

59. Kavanagh KL, Jörnvall H, Persson B, and Oppermann U (2008) Medium- and short-chain dehydrogenase/reductase gene and protein families: The SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. *Cell Mol Life Sci* **65**, 3895-3906.

60. Auld DS and Bergman T (2008) Medium- and short-chain dehydrogenase/reductase gene and protein families : The role of zinc for alcohol dehydrogenase structure and function. *Cell Mol Life Sci* **65**, 3961-3970.

61. Persson B, Jeffery J, and Jornvall H (1991) Different segment similarities in long-chain dehydrogenases. *Biochem Biophys Res Commun* **177**, 218-23.

62. Eklund H, Nordström B, Zeppezauer E, Söderlund G, Ohlsson I, Boiwe T, Söderberg B-O, Tapia O, Brändén C-I, and Åkeson Å (1976) Three-dimensional structure of horse liver alcohol dehydrogenase at 2.4 Å resolution. *J Mol Biol* **102**, 27-59.

63. Eklund H and Ramaswamy S (2008) Medium- and short-chain dehydrogenase/reductase gene and protein families : Three-dimensional structures of MDR alcohol dehydrogenases. *Cell Mol Life Sci* **65**, 3907-3917.

64. Murzin AG (1996) Structural classification of proteins: new superfamilies. *Curr Opin Struct Biol* **6**, 386-394.

65. Durrenberger M and Ward TR (2014) Recent achievments in the design and engineering of artificial metalloenzymes. *Curr Opin Chem Biol* **19**, 99-106.

66. Wilson ME and Whitesides GM (1978) Conversion of a protein to a homogeneous asymmetric hydrogenation catalyst by site-specific modification with a diphosphinerhodium(I) moiety. *J Am Chem Soc* **100**, 306-307.

67. Rosati F and Roelfes G (2010) Artificial Metalloenzymes. *ChemCatChem* **2**, 916-927.

68. Letondor C and Ward TR (2006) Artificial metalloenzymes for enantioselective catalysis: recent advances. *ChemBioChem* **7**, 1845-52.

69. Bos J and Roelfes G (2014) Artificial metalloenzymes for enantioselective catalysis. *Curr Opin Chem Biol* **19**, 135-43.

70. Ward TR (2009) Bio-inspired catalysts. Vol. **25**. *Springer Science & Business Media*.

71. Chakraborty S, Hosseinzadeh P, and Lu Y (2011) Metalloprotein design and engineering. in *Encyclopedia of Inorganic and Bioinorganic Chemistry*, ed. L.R. MacGillivray and C.M. Lukehart, John Wiley & Sons, Ltd. p. 1-51.

72. Boersma AJ, de Bruin B, Feringa BL, and Roelfes G (2012) Ligand denticity controls enantiomeric preference in DNA-based asymmetric catalysis. *Chem Commun (Camb)* **48**, 2394-2396.

73. Jeschek M, Panke S, and Ward TR (2018) Artificial Metalloenzymes on the verge of new-to-nature metabolism. *Trends Biotechnol* **36**, 60-72.

74. Lee HS and Schultz PG (2008) Biosynthesis of a site-specific DNA cleaving protein. *J Am Chem Soc* **130**, 13194-13195.

75. Schwizer F, Okamoto Y, Heinisch T, Gu Y, Pellizzoni MM, Lebrun V, Reuter R, Köhler V, Lewis JC, and Ward TR (2018) Artificial metalloenzymes: Reaction scope and optimization strategies. *Chem Rev* **118**, 142-231.

76. Pordea A and Ward TR (2008) Chemogenetic protein engineering: an efficient tool for the optimization of artificial metalloenzymes. *Chem Commun (Camb)* 4239-4249.

77. Green NM (1990) Avidin and streptavidin. *Methods Enzymol* **184**, 51-67.

78. Chaiet L and Wolf FJ (1964) The properties of streptavidin, a biotin-binding protein produced by Streptomycetes. *Arch Biochem Biophys* **106**, 1-5.

79. Weber PC, Cox MJ, Salemme FR, and Ohlendorf DH (1987) Crystallographic data for Streptomyces avidinii streptavidin. *J Biol Chem* **262**, 12728-12729.

80. Bayer EA, Ben-Hur H, Hiller Y, and Wilchek M (1989) Postsecretory modifications of streptavidin. *Biochem J* **259**, 369-376.

81. Gonzalez M, Bagatolli LA, Echabe I, Arrondo JL, Argarana CE, Cantor CR, and Fidelio GD (1997) Interaction of biotin with streptavidin. Thermostability and conformational changes upon binding. *J Biol Chem* **272**, 11288-11294.

82. Kurzban GP, Bayer EA, Wilchek M, and Horowitz PM (1991) The quaternary structure of streptavidin in urea. *J Biol Chem* **266**, 14470-14477.

83. Weber P, Ohlendorf D, Wendoloski J, and Salemme F (1989) Structural origins of high-affinity biotin binding to streptavidin. *Science* **243**, 85-88.

84. Sano T, Vajda S, and Cantor CR (1998) Genetic engineering of streptavidin, a versatile affinity tag. *Journal of Chromatography B: Biomedical Sciences and Applications* **715**, 85-91.

85. Schetters H (1999) Avidin and streptavidin in clinical diagnostics. *Biomol Eng* **16**, 73-8.

86. Skerra A and Schmidt TG (1999) Applications of a peptide ligand for streptavidin: the Strep-tag. *Biomol Eng* **16**, 79-86.

87. Collot J, Gradinaru J, Humbert N, Skander M, Zocchi A, and Ward TR (2003) Artificial metalloenzymes for enantioselective catalysis based on biotin−avidin. *J Am Chem Soc* **125**, 9030-9031.

88. Ward TR (2005) Artificial metalloenzymes for enantioselective catalysis based on the noncovalent incorporation of organometallic moieties in a host protein. *Chemistry–A European Journal* **11**, 3798-3804.

89. Heinisch T and Ward TR (2016) Artificial metalloenzymes based on the biotin–streptavidin technology: Challenges and opportunities. *Acc Chem Res* **49**, 1711-1721.

90. Gérard K, Nicolas H, Julieta G, Anita I, François G, E. RU, and R. WT (2005) Tailoring the active site of chemzymes by using a chemogenetic‐optimization procedure: Towards substrate‐specific artificial hydrogenases based on the biotin‐avidin technology. *Angew Chem Int Ed* **44**, 7764-7767.

91. Letondor C, Humbert N, and Ward TR (2005) Artificial metalloenzymes based on biotin-avidin technology for the enantioselective reduction of ketones by transfer hydrogenation. *Proc Natl Acad Sci U S A* **102**, 4683-4687.

92. Lo C, Ringenberg MR, Gnandt D, Wilson Y, and Ward TR (2011) Artificial metalloenzymes for olefin metathesis based on the biotin-(strept)avidin technology. *Chem Commun (Camb)* **47**, 12065-12077.

93. Thomas CM, Letondor C, Humbert N, and Ward TR (2005) Aqueous oxidation of alcohols catalyzed by artificial metalloenzymes based on the biotin–avidin technology. *J Organomet Chem* **690**, 4488-4491.

94. Julien P, Christophe M, Marc C, Julieta G, Ines H, Anita I, Alessia S, and R. WT (2008) Artificial metalloenzymes for asymmetric allylic alkylation on the basis of the biotin–avidin technology. *Angew Chem* **120**, 713-717.

95. Hyster TK, Knorr L, Ward TR, and Rovis T (2012) Biotinylated Rh(III) complexes in engineered streptavidin for accelerated asymmetric C-H activation. *Science* **338**, 500-3.

96. Chatterjee A, Mallin H, Klehr J, Vallapurackal J, Finke AD, Vera L, Marsh M, and Ward TR (2016) An enantioselective artificial Suzukiase based on the biotin-streptavidin technology. *Chem Sci* **7**, 673-677.

97. Okamoto Y, Köhler V, Paul CE, Hollmann F, and Ward TR (2016) Efficient in situ regeneration of NADH mimics by an artificial metalloenzyme. *ACS Catalysis* **6**, 3553-3557.

98. Jeschek M, Reuter R, Heinisch T, Trindler C, Klehr J, Panke S, and Ward TR (2016) Directed evolution of artificial metalloenzymes for in vivo metathesis. *Nature* **537**, 661-665.

99. Okamoto Y, Kojima R, Schwizer F, Bartolami E, Heinisch T, Matile S, Fussenegger M, and Ward TR (2018) A cell-penetrating artificial metalloenzyme regulates a gene switch in a designer mammalian cell. *Nature Communications* **9**, 1943.

100. Rebelein JG and Ward TR (2018) In vivo catalyzed new-to-nature reactions. *Curr Opin Biotechnol* **53**, 106-114.

101. Steinreiber J and Ward TR (2008) Artificial metalloenzymes as selective catalysts in aqueous media. *Coord Chem Rev* **252**, 751-766.

102. Reider PJ and Grabowski EJJ (1982) Total synthesis of thienamycin: a new approach from aspartic acid. *Tetrahedron Lett* **23**, 2293-2296.

103. Adly FG and Ghanem A (2014) Chiral dirhodium(II) carboxylates and carboxamidates as effective chemzymes in asymmetric synthesis of three-membered carbocycles. *Chirality* **26**, 692-711.

104. Jat JL, Paudyal MP, Gao H, Xu Q-L, Yousufuddin M, Devarajan D, Ess DH, Kürti L, and Falck JR (2014) Direct stereospecific synthesis of unprotected N-H and N-Me aziridines from olefins. *Science* **343**, 61-65.

105. Kiril T, Kristina S, Daniel H, and G. GD (2012) Structure‐selective catalytic alkylation of DNA and RNA. *Angew Chem Int Ed* **51**, 12000-12004.

106. G. BD, J. SP, N. GS, Antoinette C, Wolf‐Dietrich W, and G. GD (2015) Modular ligands for dirhodium complexes facilitate catalyst customization. *Adv Synth Catal* **357**, 2033-2038.

107. Minus MB, Kang MK, Knudsen SE, Liu W, Krueger MJ, Smith ML, Redell MS, and Ball ZT (2016) Assessing the intracellular fate of rhodium(ii) complexes. *Chem Commun* **52**, 11685-11688.

108. Srivastava P, Yang H, Ellis-Guardiola K, and Lewis JC (2015) Engineering a dirhodium artificial metalloenzyme for selective olefin cyclopropanation. *Nat Commun* **6**, 7789.

109. Pirrung MC, Liu H, and Morehead AT (2002) Rhodium chemzymes: Michaelis−Menten kinetics in dirhodium(II) carboxylate-catalyzed carbenoid reactions. *J Am Chem Soc* **124**, 1014-1023.

110. Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, and Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. in *The Proteomics Protocols Handbook*, ed. J.M. Walker, Humana Press: Totowa, NJ. p. 571-607.

111. Holmes M, Kamekura M, Lam W, Nuttall S, Woods W, Jablonski P, Serrano J, Ngui K, Antón J, and Allers T (2008) The halohandbook: protocols for haloarchaeal genetics, ed. M. Dyall-Smith. Vol. **7**. 144.

112. Walker J (1992) Spectrophotometric determination of enzyme activity: alcohol dehydrogenase (ADH). *Biochem Mol Biol Educ* **20**, 42-43.

113. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, and Bryant SH (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res* **43**, D222-D226.

114. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Geer LY, and Bryant SH (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* **45**, D200-D203.

115. Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang H-Y, Dosztányi Z, El-Gebali S, Fraser M, Gough J, Haft D, Holliday GL, Huang H, Huang X, Letunic I, Lopez R, Lu S, Marchler-Bauer A, Mi H, Mistry J, Natale DA, Necci M, Nuka G, Orengo CA, Park Y, Pesseat S, Piovesan D, Potter SC, Rawlings ND, Redaschi N, Richardson L, Rivoire C, Sangrador-Vegas A, Sigrist C, Sillitoe I, Smithers B, Squizzato S, Sutton G, Thanki N, Thomas PD, Tosatto Silvio C E, Wu CH, Xenarios I, Yeh L-S, Young S-Y, and Mitchell AL (2017) InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res* **45**, D190-D199.

116. Drozdetskiy A, Cole C, Procter J, and Barton GJ (2015) JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* **43**, W389-94.

117. Arnold K, Bordoli L, Kopp J, and Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195-201.

118. Roy A, Kucukural A, and Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function predictioA. *Nat Protoc* **5**, 725-38.

119. Benkert P, Kunzli M, and Schwede T (2009) QMEAN server for protein model quality estimation. *Nucleic Acids Res* **37**, W510-4.

120. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, and Ferrin TE (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12.

121. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, and Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Res* **28**, 235-242.

122. Unni S, Huang Y, Hanson R, Tobias M, Krishnan S, Li WW, Nielsen JE, and Baker NA (2011) Web servers and services for electrostatics calculations with APBS and PDB2PQR. *J Comput Chem* **32**, 1488-1491.

123. Land H and Humble MS (2018) YASARA: A tool to obtain structural guidance in biocatalytic investigations. *Methods Mol Biol* **1685**, 43-67.

124. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, and Yeh LS (2004) UniProt: the universal protein knowledgebase. *Nucleic Acids Res* **32**, D115-119.

125. Katoh K, Rozewicki J, and Yamada KD (2017) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* bbx108.

126. Price MN, Dehal PS, and Arkin AP (2010) FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490.

127. Kanehisa M, Sato Y, and Morishima K (2016) BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* **428**, 726-731.

128. Humbert N, Schurmann P, Zocchi A, Neuhaus JM, and Ward TR (2008) High-yield production and purification of recombinant T7-tag mature streptavidin in glucose-stressed E. coli. *Methods Mol Biol* **418**, 101-110.

129. Ebner A, Marek M, Kaiser K, Kada G, Hahn CD, Lackner B, and Gruber HJ (2008) Application of biotin-4-fluorescein in homogeneous fluorescence assays for avidin, streptavidin, and biotin or biotin derivatives. *Methods Mol Biol* **418**, 73-88.

130. Hsu CK and Park S (2010) Computational and mutagenesis studies of the streptavidin native dimer interface. *J Mol Graph Model* **29**, 295-308.

131. Lim KH, Huang H, Pralle A, and Park S (2013) Stable, high-affinity streptavidin monomer for protein labeling and monovalent biotin detection. *Biotechnol Bioeng* **110**, 57-67.

132. Yang Z, Rongjian B, Xiaolu Z, Ruwei Y, Lihua Q, Xiaoguang B, and Xinfang X (2016) Catalyst‐free S‐S bond insertion reaction of a donor/acceptor‐free carbene by a radical process: A combined experimental and computational study. *Eur J Org Chem* **2016**, 3872-3877.

133. Srivastava P, Yang H, Ellis-Guardiola K, and Lewis JC (2015) Engineering a dirhodium artificial metalloenzyme for selective olefin cyclopropanation. *Nature Communications* **6**, 7789.

134. Drozdetskiy A, Cole C, Procter J, and Barton GJ (2015) JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* **43**, W389-W394.

135. Hering O, Brenneis M, Beer J, Suess B, and Soppa J (2009) A novel mechanism for translation initiation operates in haloarchaea. *Mol Microbiol* **71**, 1451-1463.

136. Chen X, Zaro J, and Shen W-C (2013) Fusion protein linkers: Property, design and functionality. *Adv Drug Del Rev* **65**, 1357-1369.

137. Persson B, Hedlund J, and Jörnvall H (2008) Medium- and short-chain dehydrogenase/reductase gene and protein families: the MDR superfamily. *Cell Mol Life Sci* **65**, 3879-3894.

138. Timpson LM, Liliensiek AK, Alsafadi D, Cassidy J, Sharkey MA, Liddell S, Allers T, and Paradisi F (2013) A comparison of two novel alcohol dehydrogenase enzymes (ADH1 and ADH2) from the extreme halophile *Haloferax volcanii. Appl Microbiol Biotechnol* **97**, 195-203.

139. Ortega G, Lain A, Tadeo X, Lopez-Mendez B, Castano D, and Millet O (2011) Halophilic enzyme activation induced by salts. *Sci Rep* **1**, 6.

140. Flamholz A, Noor E, Bar-Even A, and Milo R (2012) eQuilibrator—the biochemical thermodynamics calculator. *Nucleic Acids Res* **40**, D770-D775.

141. Ceccarelli C, Liang Z-X, Strickler M, Prehna G, Goldstein BM, Klinman JP, and Bahnson BJ (2004) Crystal structure and amide H/D exchange of binary complexes of alcohol dehydrogenase from *Bacillus stearothermophilus*: insight into thermostability and cofactor binding. *Biochemistry* **43**, 5266-5277.

142. Taneja B and Mande SC (1999) Conserved structural features and sequence patterns in the GroES fold family. *Protein Eng* **12**, 815-8.

143. Kang C, Hayes R, Sanchez EJ, Webb BN, Li Q, Hooper T, Nissen MS, and Xun L (2012) Furfural reduction mechanism of a zinc-dependent alcohol dehydrogenase from *Cupriavidus necator* JMP134. *Mol Microbiol* **83**, 85-95.

144. Tsigos I, Velonia K, Smonou I, and Bouriotis V (1998) Purification and characterization of an alcohol dehydrogenase from the Antarctic psychrophile *Moraxella sp.* TAE123. *Eur J Biochem* **254**, 356-362.

145. Clamp M, Cuff J, Searle SM, and Barton GJ (2004) The Jalview Java alignment editor. *Bioinformatics* **20**, 426-427.

146. Youn B, Camacho R, Moinuddin SG, Lee C, Davin LB, Lewis NG, and Kang C (2006) Crystal structures and catalytic mechanism of the *Arabidopsis* cinnamyl alcohol dehydrogenases AtCAD5 and AtCAD4. *Org Biomol Chem* **4**, 1687-1697.

147. Rashid M and Stingl U (2015) Contemporary molecular tools in microbial ecology and their application to advancing biotechnology. *Biotechnol Adv* **33**, 1755-1773.

148. Sonbol SA, Ferreira AJS, and Siam R (2016) Red Sea Atlantis II brine pool nitrilase with unique thermostability profile and heavy metal tolerance. *BMC Biotechnol* **16**, 14.

149. Mohamed YM, Ghazy MA, Sayed A, Ouf A, El-Dorry H, and Siam R (2013) Isolation and characterization of a heavy metal-resistant, thermophilic esterase from a Red Sea brine pool. *Sci Rep* **3**, 3358.
150. Elbehery AHA, Leak DJ, and Siam R (2017) Novel thermostable antibiotic resistance enzymes from the Atlantis II Deep Red Sea brine pool. *Microbial Biotechnology* **10**, 189-202.
151. Oren A (2002) Adaptation of halophilic archaea to life at high salt concentrations. in *Salinity: Environment - Plants - Molecules*, ed. A. Läuchli and U. Lüttge, Springer Netherlands: Dordrecht. p. 81-96.
152. Danson MJ and Hough DW (1997) The structural basis of protein halophilicity. *Comp Biochem Physiol A Physiol* **117**, 307-312.
153. Loderer C, Wagner D, Morgenstern F, Spieß A, and Ansorge-Schumacher MB (2018) Discovery of a novel thermostable Zn2+ -dependent alcohol dehydrogenase from *Chloroflexus aurantiacus* through conserved domains mining. *J Appl Microbiol* **124**, 480-490.
154. Weckbecker A, Groger H, and Hummel W (2010) Regeneration of nicotinamide coenzymes: principles and applications for the synthesis of chiral compounds. *Adv Biochem Eng Biotechnol* **120**, 195-242.
155. Takahashi M, Takahashi E, Joudeh LI, Marini M, Das G, Elshenawy MM, Akal A, Sakashita K, Alam I, Tehseen M, Sobhy MA, Stingl U, Merzaban JS, Di Fabrizio E, and Hamdan SM (2018) Dynamic structure mediates halophilic adaptation of a DNA polymerase from the deep-sea brines of the Red Sea. *FASEB J* **32**, 3346-3360.
156. Karan R, Capes MD, and DasSarma S (2012) Function and biotechnology of extremophilic enzymes in low water activity. *Aquatic Biosystems* **8**, 4-4.
157. Karan R, Singh SP, Kapoor S, and Khare SK (2011) A novel organic solvent tolerant protease from a newly isolated *Geomicrobium* sp. EMB2 (MTCC 10310): production optimization by response surface methodology. *N Biotechnol* **28**, 136-145.
158. Adlercreutz P (2008) Fundamentals of biocatalysis in neat organic solvents. in *Organic synthesis with enzymes in non-aqueous media*, ed. G. Carrera and S. Riva, WILEY-VCH Verlag GmbH, Weinheim. p. 3-24.
159. Bonneté F (2007) Colloidal approach analysis of the Marseille protein crystallization database for protein crystallization strategies. *Crystal Growth & Design* **7**, 2176-2181.
160. Siddiqui KS and Cavicchioli R (2006) Cold-adapted enzymes. *Annu Rev Biochem* **75**, 403-33.
161. Dominy BN, Minoux H, and Brooks CL (2004) An electrostatic basis for the stability of thermophilic proteins. *Proteins* **57**, 128-141.
162. Mee B, Kelleher D, Frias J, Malone R, Tipton KF, Henehan GTM, and Windle HJ (2005) Characterization of cinnamyl alcohol dehydrogenase of *Helicobacter pylori. FEBS J* **272**, 1255-1264.
163. Bagniewska-Zadworna A, Barakat A, Lakomy P, Smolinski DJ, and Zadworny M (2014) Lignin and lignans in plant defence: insight from expression profiling of cinnamyl alcohol dehydrogenase genes during development and following fungal infection in *Populus. Plant Sci* **229**, 111-121.
164. Wilkin J-M, Soetaert K, Stélandre M, Buyssens P, Castillo G, Demoulin V, Bottu G, Laneelle M-A, Daffe M, and De Bruyn J (1999) Overexpression, purification and characterization of *Mycobacterium bovis* BCG alcohol dehydrogenase. *Eur J Biochem* **262**, 299-307.
165. Larroy C, Rosario Fernández M, González E, Parés X, and Biosca JA (2003) Properties and functional significance of *Saccharomyces cerevisiae* ADHVI. *Chem-Biol Interact* **143-144**, 229-238.
166. Pick A, Rühmann B, Schmid J, and Sieber V (2013) Novel CAD-like enzymes from *Escherichia coli* K-12 as additional tools in chemical production. *Appl Microbiol Biotechnol* **97**, 5815-5824.

167.   Beutner KR and von Krogh G (1990) Current status of podophyllotoxin for the treatment of genital warts. *Semin Dermatol* **9**, 148-51.

168.   Canel C, Moraes RM, Dayan FE, and Ferreira D (2000) Podophyllotoxin. *Phytochemistry* **54**, 115-20.

169.   Bachmann DG, Schmidt PJ, Geigle SN, Chougnet A, Woggon W-D, and Gillingham DG (2015) Modular ligands for dirhodium complexes facilitate catalyst customization. *Adv Synth Catal* **357**, 2033-2038.

170.   Espino CG, Fiori KW, Kim M, and Du Bois J (2004) Expanding the scope of C-H amination through catalyst design. *J Am Chem Soc* **126**, 15378-15379.

171.   Creus M, Pordea A, Rossel T, Sardo A, Letondor C, Ivanova A, Letrong I, Stenkamp RE, and Ward TR (2008) X-ray structure and designed evolution of an artificial transfer hydrogenase. *Angew Chem Int Ed Engl* **47**, 1400-1404.

172.   Reetz MT, Peyralans JJ, Maichele A, Fu Y, and Maywald M (2006) Directed evolution of hybrid enzymes: Evolving enantioselectivity of an achiral Rh-complex anchored to a protein. *Chem Commun (Camb)* 4318-4320.

173.   Mazurek S, Ward TR, and Novic M (2007) Counter propagation artificial neural networks modeling of an enantioselectivity of artificial metalloenzymes. *Mol Divers* **11**, 141-152.

174.   Matsuo T, Fukumoto K, Watanabe T, and Hayashi T (2011) Precise Design of Artificial Cofactors for Enhancing Peroxidase Activity of Myoglobin: Myoglobin Mutant H64D Reconstituted with a "Single-Winged Cofactor" Is Equivalent to Native Horseradish Peroxidase in Oxidation Activity. *Chemistry – An Asian Journal* **6**, 2491-2499.

175.   Schwizer F, Köhler V, Dürrenberger M, Knörr L, and Ward TR (2013) Genetic Optimization of the Catalytic Efficiency of Artificial Imine Reductases Based on Biotin–Streptavidin Technology. *ACS Catalysis* **3**, 1752-1755.

176.   Kohler V, Mao J, Heinisch T, Pordea A, Sardo A, Wilson YM, Knorr L, Creus M, Prost JC, Schirmer T, and Ward TR (2011) OsO4.streptavidin: a tunable hybrid catalyst for the enantioselective cis-dihydroxylation of olefins. *Angew Chem Int Ed Engl* **50**, 10863-10866.

177.   Kyte J and Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-132.

178.   Gallizia A, de Lalla C, Nardone E, Santambrogio P, Brandazza A, Sidoli A, and Arosio P (1998) Production of a Soluble and Functional Recombinant Streptavidin inEscherichia coli. *Protein Expression Purif* **14**, 192-196.

179.   Lees JG, Miles AJ, Wien F, and Wallace B (2006) A reference database for circular dichroism spectroscopy covering fold and secondary structure space. *Bioinformatics* **22**, 1955-1962.

180.   Konieczny S, Leurs M, and Tiller JC (2015) Polymer enzyme conjugates as chiral ligands for sharpless dihydroxylation of alkenes in organic solvents. *ChemBioChem* **16**, 83-90.

181.   Vasconcelos T, Sarmento B, and Costa P (2007) Solid dispersions as strategy to improve oral bioavailability of poor water soluble drugs. *Drug Discov Today* **12**, 1068-1075.

182.   González Mn, Argaraña CE, and Fidelio GD (1999) Extremely high thermal stability of streptavidin and avidin upon biotin binding. *Biomol Eng* **16**, 67-72.

183.   Yang H, Swartz AM, Park HJ, Srivastava P, Ellis-Guardiola K, Upp DM, Lee G, Belsare K, Gu Y, Zhang C, Moellering RE, and Lewis JC (2018) Evolving artificial metalloenzymes via random mutagenesis. *Nat Chem* **10**, 318.

184.   Zhao J, Bachmann DG, Lenz M, Gillingham DG, and Ward TR (2018) An artificial metalloenzyme for carbene transfer based on a biotinylated dirhodium anchored within streptavidin. *Catalysis Science & Technology* **8**, 2294-2298.

8 Appendix

8.1 Characterization of a polyextremophilic zinc-dependent alcohol dehydrogenase from the Atlantis II Deep Red Sea brine pool

Table S 1: Oligonucleotide-sequence of primers used for cloning of ADH-A1 short

No	Name	Sequence
P19	pTA963_ptnaA_F	5'-gttcgaaccgccctttcc-3'
P20	pTA963_REV	5'-atgaccatgattacgccaag-3'
P36	AZA constr.1 longPrim2 for	5'-accaccaccacggcagtggcctcaaggccatgcagctca-3'
P38	AZA longPrim rev	5'-cttagctgaccttgaggacgcc-3'
P39	AZA Gib for1	5'-cgcctgccgattacttcacattcgcggacctattgcgcat atgcaccaccaccaccaccggcagtggcctcaaggcc-3'
P43	AZA Gib rev	5'-cgatggtccagaggtgcggccgctctagaactagtggatc cttagctgaccttgaggacgc-3'

(A) Codon optimized gene sequence of *adh/a1*

```
    NdeI
  1 CATATGCACCACCACCACCACCACATGCGCATGGAGCTCAAGTTCTTCCACCAGGAGGTC    60
  1    M  H  H  H  H  H  H  M  R  M  E  L  K  F  F  H  Q  E  V    20
 61 TACCTCGAGGCCGACGGCCGCGAGCCCATCCTCAAGGCCATGCAGCTCAAGGAGCCCAAG... 120
 21  Y  L  E  A  D  G  R  E  P  I  L  K  A  M  Q  L  K  E  P  K ...  40
```

(B) Gene sequence of *adh/a1a*

```
    NdeI
  1 CATATGCACCACCACCACCACCACGGCAGTGGCCTCAAGGCCATGCAGCTCAAGGAGCCC    60
  1    M  H  H  H  H  H  H  G  S  G  L  K  A  M  Q  L  K  E  P    20
 61 AAGCCCGTCGAGCAGGCCCCCCTCGAGATGGTCGAGCTCAAGGAGCCCCGCCCCGGCCCC... 120
 21  K  P  V  E  Q  A  P  L  E  M  V  E  L  K  E  P  R  P  G  P ...  40
```

Figure S 1: DNA and amino acid sequence of *adh/a1* gene and shortened *adh/a1a* gene. The two sequences (A) and (B) show the beginning of the gene before and after the deletion of the nonsense sequence (gray). The his6-tag is shown in yellow and the linker in blue before. The *NdeI* restriction site is highlighted. The beginning of the gene is highlighted in bold red.
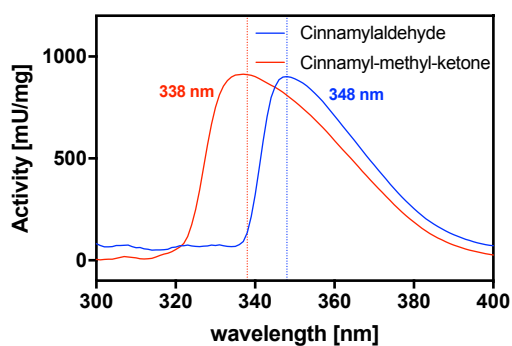
Figure S 2: Wavelength scan of NADPH (0.2 mM) absorption. Before measurement the reaction buffer with substrate (50 mM) was used as blank.
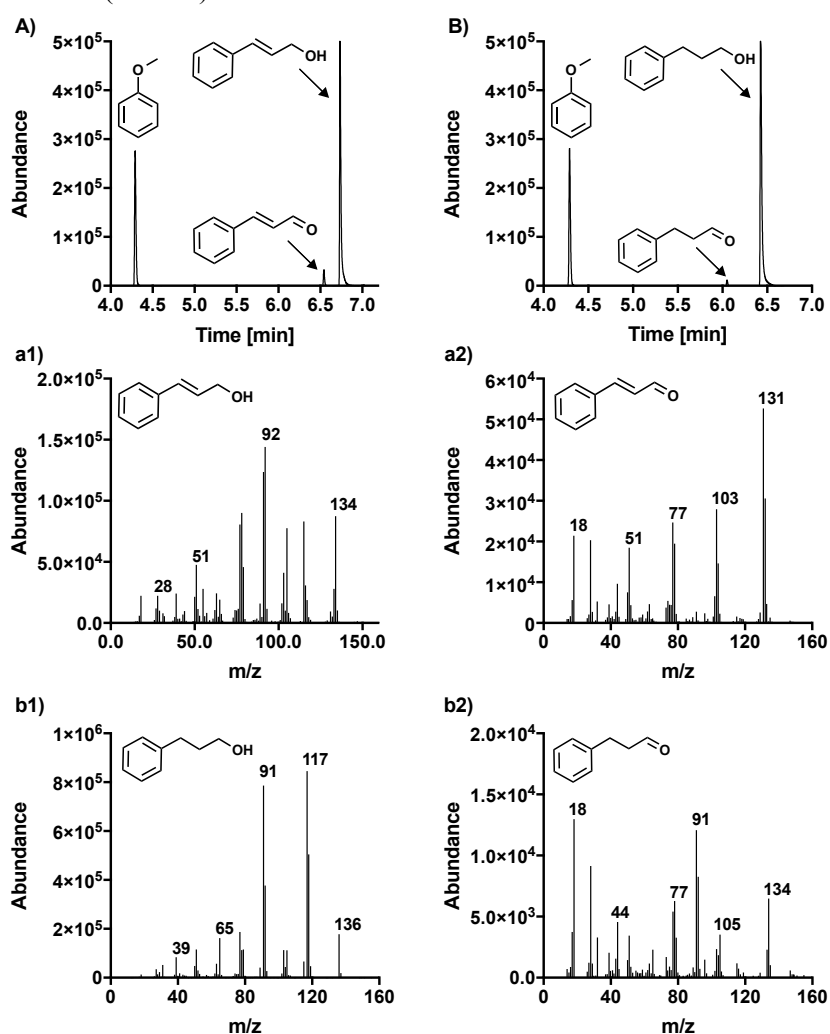

Figure S 3: GC-MS analysis of product formation. GC of the reaction with (A) cinnamyl alcohol and (B) hydrocinnamyl alcohol as substrate. Anisol was used as standard. MS of the substrates (a1) cinnamyl alcohol /(b1) hydrocinnamyl alcohol and products (a2) cinnamyl aldehyde /(b2) hydrocinnamyl aldehyde. Reaction was done in 5 ml scale for 3 h at 45 °C in standard buffer using standard concentrations while stirring.

Table S 2: ADHs used for the phylogenetic analysis of ADH/A1a.

Accession no.	Organism, name	Converted substrates, cofactor or sampling site	ADH subfamily
Pdb 2CF5	*Arabidopsis thaliana*, CADH5	coniferaldehyde, 5-hydroxyconifer-aldehyde, sinapaldehyde, 4-coumar-aldehyde and caffeoyl aldehyde, NADP	CAD1, CAD
Pdb 1YQX	*Populus tremuloides*	hydroxycinnamaldehyde substrates, NADP	CAD1, CAD
Q6ZHS4	*Oryza sativa* subsp.	Coniferaldehyde, sinapaldehyde, NADP	CAD1, CAD
Q2KNL5	*Ocimum basilicum*	cinnamyl alcohol geraniol, cinnamyl alcohol, phenylpropanol, hexanol, NADP	CAD1, CAD
P31656	*Medicago sativa* CAD2	Cinnamyl aldehyde, coniferyl aldehyde, sinapaldehyde Specific, NADP	CAD1, CAD
O82515	*Medicago sativa* CAD1	Cinnamyl aldehyde, coniferyl aldehyde, sinapaldehyde, valeraldehyde, 3-methoxybenzylaldehyde, NADP	CAD1, CAD
P25377	*Saccharomyces cerevisiae*	ADHVII, Cinnamyl alcohol, cinnamyl aldehyde, pentanol, hexanol, pentanal, NADP	CAD1, CAD
3TWO	*Helicobacter pylori*	Cinnamyl alcohol, coniferyl alcohol, benzyl alcohol, cinnamylaldehyde, NADP	CAD1, CAD
P9WQC4.1	*Mycobacterium bovis*	Benzaldehyde, 3-methoxybenzaldehyde, octanal, coniferaldehyde, cinnamyl alcohol/aldehyde, NADP	CAD1, CAD
P75691	*Escherichia coli*, YahK	furfural, benzaldehyde, isobutyr-aldehyde, acetaldehyde…., NADP	CAD1, CAD
P00330	*Saccharomyces cerevisiae*, YADH1	ethanol, cinnamyl alcohol, allyl alcohol, NAD	CAD3, Adhp
P08843	*Apergillus nidulans*, Adhp	short-chain alcohols and ketones, propanol, ethanol, cinnamyl alcohol very good, NAD	CAD3, Adhp
P39451	*E. coli*, Adhp	short-chain alcohols, ethanol/acetaldehyde	PRK, Adhp
P20368	*Zymomonas mobilis*, Adhp	ethanol/acetaldehyde, not confirmed, NAD	PRK, Adhp
1RJW	*Bacillus stearothermophilus*, htADH	2-propanol, ethanol, 2-butanol,benzyl alcohol, isopropyl alcohol, (cinnamyl alcohol, cinnamyl aldehyde), NAD	CAD3, Adhp
	Cupriavidus necator, FurX	NAD, Furfural, Ethanol	CAD3, Adhp
4Z6K	*Moraxella* sp. TAE123, MADH	NAD, Ethanol, Propanol, Butanol	CAD3, Adhp
YP001041078	*Staphylothermus marinus*	substrate/cofactor unknown, derived from Volcano, hydrothermal vent, 85°C, pH 6.5	CAD2, Adhp

YP875447	*Cenarchaeum symbiosum*	substrate/cofactor unknown, derived from a marine sponge, 7 -19˚C	CAD2, Adhp
WP012874260	*Thermobaculum terrenum*	substrate/cofactor unknown, derived from thermal soil, Yellowstone park, pH 3.9, 65 – 92 °C	CAD2, Adhp
KUK41356.1	*Clostridia bacterium*	substrate/cofactor unknown, derived from oil reservoir, Kuparuk, Alaska North Slope, 65 – 80 °C, 2.1% salt	CAD2, Adhp

8.2 Extending the scope of the biotin-streptavidin technology for the design of complex biocatalysts

Table S 3: Oligonucleotide-sequence of primers used for mutation of SAV variants

No	Name	Sequence
P1	SAV Mutation K121A_for	5'-gaccagcgtggacgcccaggcgttggcc-3'
P2	SAV Mutation K121A_rev	5'-ggccaacgcctgggcgtccacgctggtc-3'
P3	SAV Mutation S112A_for	5'-ggtggtgccggcggtcagcagcc-3'
P4	SAV Mutation S112A_rev	5'-ggctgctgaccgccggcaccacc-3'
P5	SAV Mutation H87Y_for	5'-cacgtggtcgcggaataggcgttgcggtagt-3'
P6	SAV Mutation H87Y_rev	5'-actaccgcaacgcctattccgcgaccacgtg-3'
P7	SAV Mutation R84A_for	5'-cggagtgggcgttggcgtagttattcttccaggc-3'
P8	SAV Mutation R84A_rev	5'-gcctggaagaataactacgccaacgcccactccg-3'
P9	SAV Mutation H127Y_for	5'-cttggtgaaggtgtcatagccgaccagcgtgga-3'
P10	SAV Mutation H127Y_rev	5'-tccacgctggtcggctatgacaccttcaccaag-3'
P11	SAV Mutation H87N_for	5'-gtggtcgcggagttggcgttgcggtag-3'
P12	SAV Mutation H87N_rev	5'-ctaccgcaacgccaactccgcgaccac-3'
P13	SAV Mutation K121L_for	5'-gccgaccagcgtggataaccaggcgttggcctc-3'
P14	SAV Mutation K121L_rev	5'-gaggccaacgcctggttatccacgctggtcggc-3'
P15	SAV Mutation R84Q_for	5'-cgcggagtgggcgttctggtagttattcttccagg-3'
P16	SAV Mutation R84Q_rev	5'-cctggaagaataactaccagaacgcccactccgcg-3'
P17	SAV Mutation S112K_for	5'-ggcctcggtggtgcccttggtcagcagccactg-3'
P18	SAV Mutation S112K_rev	5'-cagtggctgctgaccaagggcaccaccgaggcc-3'
P21	SAV Mutation R84A_for	5'-cggagttggcgttggcgtagttattcttccaggc-3'
P22	SAV Mutation R84A_rev	5'-gcctggaagaataactacgccaacgccaactccg-3'

P23	SAV Mutation R84Q_for	5'-tcgcggagttggcgttctggtagttattcttccagg-3'
P24	SAV Mutation R84Q_rev	5'-cctggaagaataactaccagaacgccaactccgcga-3'
P44	SAV Mutation S112Y_for	5'-cctcggtggtgccataggtcagcagccactg-3'

Table S 4: List of all cloned SAV variants and plasmid names

Plasmid No	Mutations	Plasmid No	Mutations
WT	-	11.2	H87N H127Y S112A K121A R84A
0.2	H87N	12.1	H87N H127Y S112A K121A R84Q
0.1	H127Y	13.1	H87N H127Y S112A K121L R84A
0.5	S112A	14.1	H87N H127Y S112A K121L R84Q
0.3	K121A	15.1	S112Y
0.4	K121L	16.1	S112F
4	S112K	17.1	S112W
1	H87N H127Y	18.1	K121D
2.5	S112A K121A	19.1	K121E
3	S112A K121L	20.2	K121W
5.1	H127Y S112A K121A	21.1	S112H
6.9	H127Y S112A K121L	22.2	K121H
7.1	H127Y S112K	23	S112Y K121D
8.1.1	H87N H127Y S112A K121A	24.1	S112Y K121E
9.1.1	H87N H127Y S112A K121L	25	S112H K121H
10.1.1	H87N H127Y S112K		

Table S 5: Determined optimal expression temperatures for soluble SAV after 16 h.

SAV	wt	2.5	3	4
T [°C]	30	30	30	30

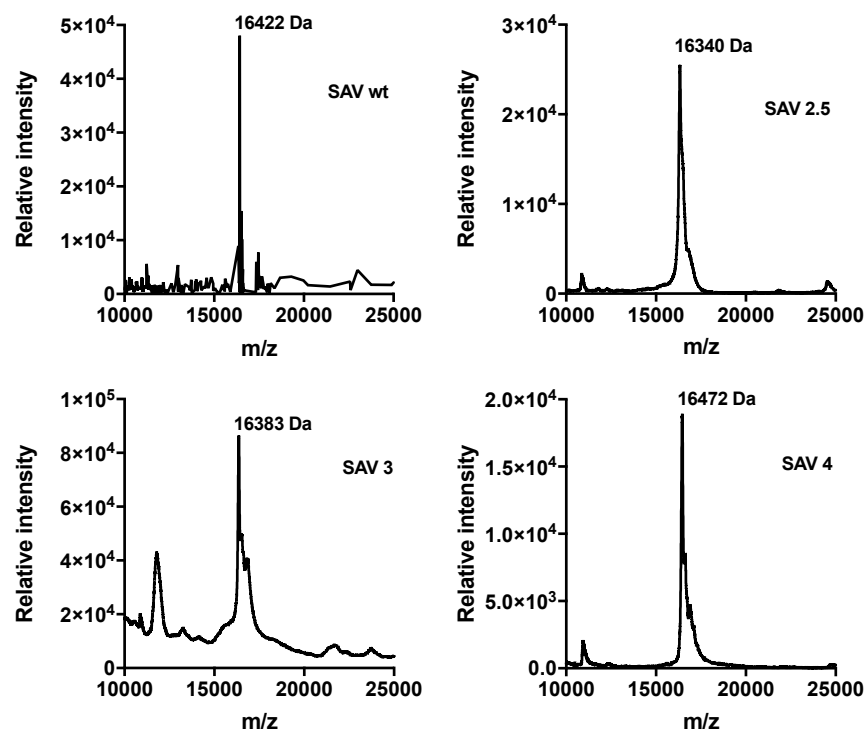| SAV	1	8.1.1	9.1.1	10.1.1	11.2	12.1	13.1	14.1
T [°C]	30	20	20	30	20	20	20	20

Figure S 4: Mass analysis of selected SAV variants. The mass was analyzed by ESI-TOF (first graph) or by MALDI-TOF (other graphs).
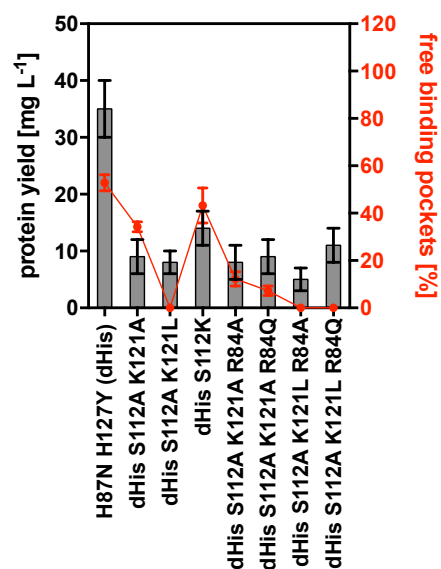

Figure S 5: Protein yield and free binding pockets of SAV dHis mutants expressed in shaking flasks. Bars - protein yield of purified and lyophilized SAV; points: free functional binding pockets per SAV tetramer determined by B4F titration assay
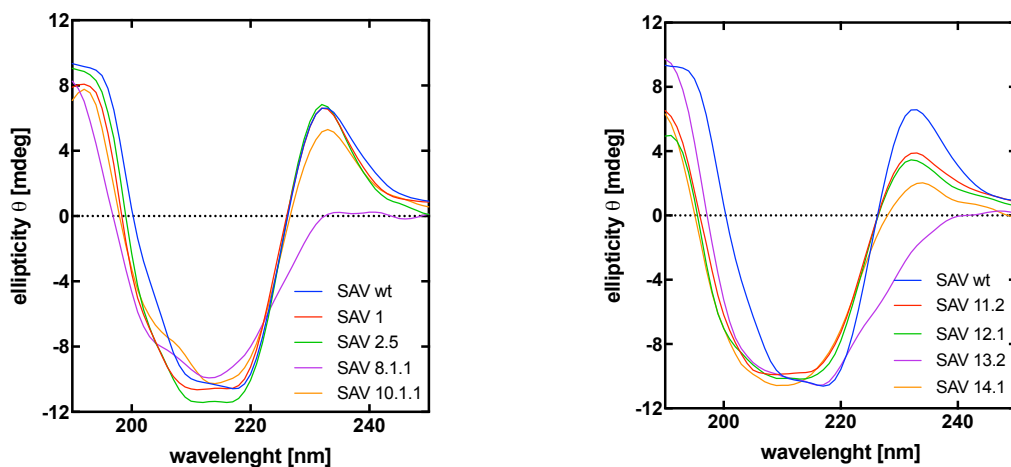
Figure S 6: Comparison of CD-curves of streptavidin variants. Mutations in the loop region of the binding pockets alters the CD-signal and lowers the maximum at 234 nm.
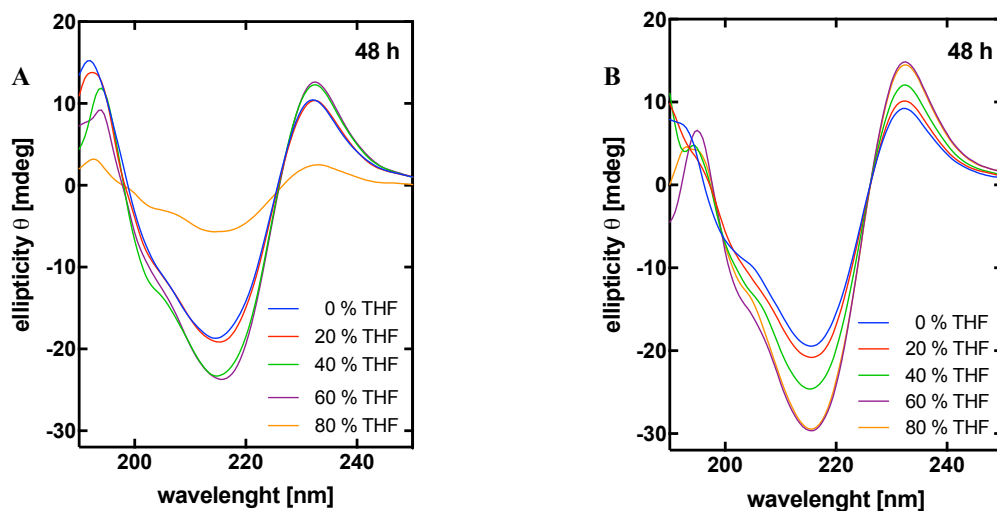

Figure S 7: CD-spectra of SAV incubated in different ratios of THF for 48 h. (A) CD-spectra of SAV wt after 48 h; (B) CD-spectra of PEGylated SAV wt (PEG4) after 48 h.
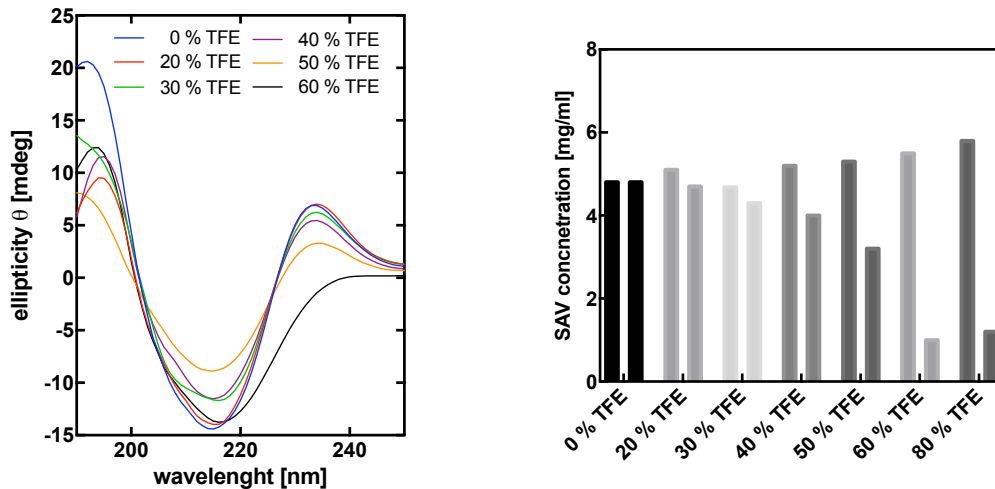
Figure S 8: Stability test of SAV wt in TFE, (A) CD-spectra of SAV wild-type incubated in different ratios of TFE for 72 h; (B) Concentration of samples measured at t = 0 h and 72 h.
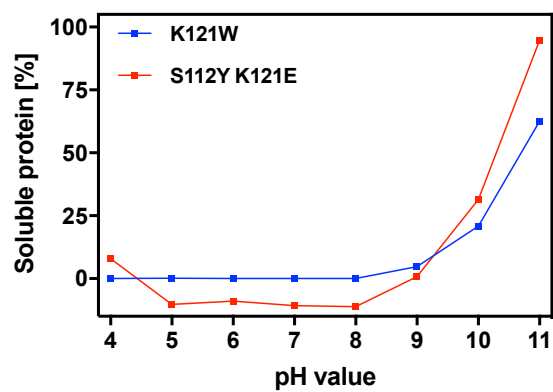

Figure S 9: Solubility test of SAV-Rh(II)-complex of SAV K121W and S112Y K121E at different pH-values. The SAV apo protein was incubated 1:1 with Biot-diRh and the concentration of the soluble fraction was determined.
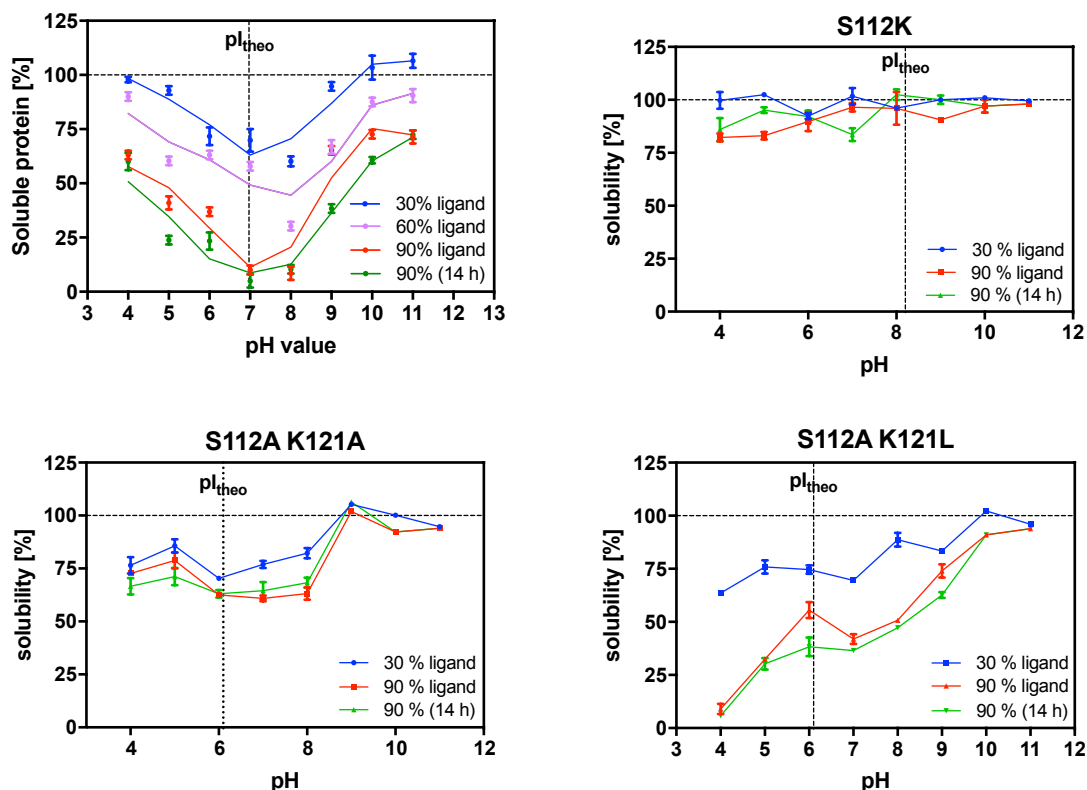
Figure S 10: Solubility of SAV-diRh complex. SAV mutants were incubated stepwise with different amounts of Biot-diRh ligand (1: 0.3/0.6/0.9) and the soluble protein concentration was determined. The sample with 90% ligand was incubated for 14 h at room temperature and measured again.

Table S 6: Solvents effect on cyclopropanation reaction using different organic solvents.

Catalyst	Solvent	Conv. [%]	Yield **5a** [%]	Yield **5b** [%]	Yield [**5a**]/[**5b**]
SAVwt-diRh	10% ACN	100	14	60	0.23 : 1
SAVwt-diRh	10% Toluene	not complete	4	53	0.08 : 1
SAVwt-diRh	10% Hexane	100	3	92	0.03 : 1
SAVwt-diRh	10% 2-Propanol	100	18	73	0.25 : 1
SAVwt-diRh	10% Methanol	100	48	43	1.12 : 1
SAVwt-diRh	-	100	50	47	1.06 : 1
no	-	not complete	41	24	1.7 : 1

The reaction was performed in acetate buffer, pH 4, with 10% solvent additive, using 1 mol% SAVwt-diRh as catalyst (24 h, at room temperature). No enantioselectivity was observed.

Table S 7: Effect of pH-value on the cyclopropanation reaction.

Catalyst	Solvent	Conv. [%]	Yield of **5a** [%]
Biot- Rh$_2$	Na-acetate pH 4	100	51.7
Biot- Rh$_2$	Na-acetate pH 5	100	48.9
Biot- Rh$_2$	Na-carbonate pH 6	100	59.5
Biot- Rh$_2$	Na-carbonate pH 7	100	54.3
Biot- Rh$_2$	Na-carbonate pH 8	100	61.5
Biot- Rh$_2$	Borate pH 9	100	21.3
Biot- Rh$_2$	Na-carbonate pH 9	100	55.2

The reaction was performed in 100 mM buffer with varying pH 4, using 1 mol% Biot-diRh as catalyst (24 h, at room temperature).

9 Publications

Fist author

Akal A, Karan R, Hohl A, Inthikhab A, Vogler M, Grotzinger SW, Eppinger J, Rüping M (2018) A polyextremophilic alcohol dehydrogenase from the Atlantis II Deep Red Sea brine pool. *FEBS Open Bio*, (accepted)

Co-author

Grotzinger SW, Karan R, Strillinger E, Bader S, Frank A, Al Rowaihi IS, Akal A, Wackerow W, Archer JA, Rueping M, Weuster-Botz D, Groll M, Eppinger J, and Arold ST (2018) Identification and Experimental Characterization of an Extremophilic Brine Pool Alcohol Dehydrogenase from Single Amplified Genomes. *ACS Chem Biol* **13**, 161-170.

Declaration of Authorship

I, Anastassja Leyla Gespers, née Akal, hereby declare that I independently prepared the present thesis, using only the references and resources stated. This work has not been submitted to any examination board yet. Parts of this work have been or will be published in scientific journals. The permission to reuse published text and figures for this thesis was granted by the publishers.

München, July 2018