DISS. ETH N° 21658

Expanding the scope of enzyme catalysis with β-amino acids and metal cofactors

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZÜRICH

(Dr. sc. ETH Zurich)

presented by

CLEMENS MAYER

Dipl. Ing., Technische Universität Graz

born on 06.05.1985

citizen of Austria

accepted on the recommendation of

Prof. Dr. Donald Hilvert, examiner

Prof. Dr. Helma Wennemers, co-examiner

2014

I would take the awe of understanding over the awe of ignorance any day.

Douglas Adams

ACKNOWLEDGMENTS

Experiencing the screams, manic laughter, somber tragedies, and the sense of awe and agony that underpins the doctorial life span would not have been such a pleasure without the help and support of the kind people around me. This Acknowledgment section provides me with an opportunity to say *thank you* to a few of them.

I am grateful to my PhD supervisor, Don Hilvert, who kept a sense of humor when I had lost mine and attempted – with his maddening attention to detail – to teach me the English language and how to punctuate prose. His constant support and the granted freedom were the cornerstones to the success of my PhD thesis.

I would also like to thank Helma Wennemers for agreeing to co-examine my thesis and Peter Kast for his advice to every biological problem I encountered over the past 5½ years. I am indebted to Sam Gellman for providing his unmatched knowledge about α/β-peptides to the work described in Chapter 4 and his group for providing cyclic building blocks when they were needed. I would like to further acknowledge Novartis for financial support.

Particular thanks go to my undergraduate students Andreas Küchler, Raoul Rosenthal, Anja Pabst, Thomas Wucherpfennig, Marjorie Corrales Carnero, and Benita Heiz for their contributions and their interest in my research.

Over the years, I benefited tremendously from the great atmosphere in the Hilvert group. Foremost, I am grateful to Dennis Gillingham and Manuel Müller

for teaching me the basics of science & life in the Hilvert group and for showing me the ins and outs of the day & nightlife in Zürich. From the plethora of undergrads, PhD students, postdocs, visiting scientists, etc., that I had the pleasure to meet, I would like to thank those who accompanied me for most of my time at ETH: Hajo Kries for his *open tap* policy concerning lunch & coffee invitations, when I once again did not have any money, Nathalie Preiswerk for her honesty and friendship, Richard Obexer for his delicious cooking, Cindy Schulenburg and An Vandemeulebroucke for taking the time to listen and cheer me up, David Niquille and Sabine Studer for continuing some of the projects described in this thesis, Reinhard Zschoche for providing milk for my coffee, and Clement Dince for making day-to-day lab life and the after-work parties such an enjoyable experience.

I am deeply thankful for all the wonderful friendships in Zürich and Austria that make my life such a pleasure to live. In this respect, special thanks goes to Johannes Reiter and Tobias Abel for having put up with me for such a long time.

Lastly, I would like to thank my family for all the support and my Natalie for always being there for me.

Parts of this thesis have been published

A genetically-encodable ligand for transfer hydrogenation
Mayer, C.; Hilvert, D.
Eur. J. Org. Chem., **2013**, 17, 3427-3431

An artificial metalloenzyme for olefin metathesis
Mayer, C.; Gillingham, D. G.; Ward, T. R.; Hilvert, D.,
Chem. Comm. **2011,** *47* (44), 12068-12070

Contents

Abstract

Enzymes are remarkable catalysts. Over eons, nature has fine-tuned these protein molecules to catalyze the plethora of reactions necessary to sustain life with unmatched rates and selectivities. Chemists and biologists alike have long sought to harness the power of enzyme catalysis for non-natural reactions. However, the design of catalysts that display made-to-order activities has proven to be a formidable challenge. This thesis describes studies to expand the scope of enzyme catalysis with artificial metal cofactors and β-amino acids.

One possible means of generating novel activities in proteinogenic scaffolds emulates nature's use of metal ions for catalysis by placing a metal complex, which catalyzes a non-natural transformation, in a protein binding pocket. In principle, these *artificial metalloenzymes* could potentially combine the high rates and selectivities achieved by enzymes with the broad reaction and substrate scope displayed by transition metal catalysts. Toward this end, Chapter 2 describes the use of a Grubbs-Hoveyda-type olefin metathesis catalyst, equipped with an electrophilic bromoacetamide group, to modify a cysteine-containing variant of a small heat shock protein from *Methanocaldococcus jannaschii*. The resulting artificial metalloenzyme catalyzes a ring closing metathesis reaction, despite a plethora of potentially inactivating functional groups present within the protein scaffold. Nevertheless, anchoring the organometallic moiety on the protein via a site-selective, but flexible linker hampers effective preorganization of substrate and catalyst in a well-defined binding pocket. To overcome this hurdle, Chapter 3 explores new strategies to create hybrid enzymes by anchoring the small-

molecule catalyst directly on the protein backbone. In a proof-of-concept study, we demonstrated that small, unmodified peptides allow for large ligand acceleration effects in aqueous transfer hydrogenation reactions. For example, the tripeptide Gly-Gly-Phe in combination with a water-soluble iridium complex catalyzes reduction of diverse ketones, aldehydes, imines, and NAD^+ by formate with high turnover numbers. With the only prerequisites for efficient catalysis being an unmodified N-terminus and a glycine as penultimate amino acid, it should be possible to append this catalytic motif to the N-terminus of virtually any protein as a simple coordination tag. By this means, the capsid-forming heat shock protein used in Chapter 2 could potentially be utilized as a nanoreactor for dynamic kinetic resolutions in a productive reaction cascade.

Another possibility to expand the properties of protein catalysts is to utilize building blocks different than the canonical α-amino acids for enzyme catalysis. Non-natural oligomers, often referred to as *foldamers*, can form distinct conformations in solution that display a larger number of secondary structures and improved stability compared to natural peptides. In order to take advantage of these properties for enzyme catalysis, we investigated foldamers constructed from more than one type of building block (α- and β-amino acids) as surrogates for a helical segment in a split chorismate mutase (Chapter 4). More specifically, by systematically substituting α-amino acids with their β-counterparts we created seven chimeric chorismate mutases containing up to five β-amino acids, including an essential catalytic residue. Two of these variants rival the activity of the wild-type enzyme. From this study we deduced valuable guidelines that may facilitate the introduction of non-natural building blocks into a wider range of different proteins. To make use of the unique

properties of foldamers for enzyme function, the evolvability of such chimeric enzymes is of utmost importance. In Chapter 5, we report an evolutionary approach to adapt the natural binding partner of the split chorismate mutase to the foldamer surrogates. Following parallel protein production and purification protein variants that better accommodate the nonstandard building blocks can be conveniently identified by microtiter plate screening for chorismate mutase activity. Although the proof-of-concept study in Chapter 5 has not yet yielded chimeric chorismate mutases that display significant improvements compared to the starting hybrid catalysts, it is conceivable that a more in-depth evolutionary study will afford such variants.

The results of this thesis highlight means to expand the scope of enzyme catalysis and to evolve the acquired functions. In the future, experiments guided by our findings could generate a new breed of hybrid enzymes that, inspired by nature, take full advantage of the properties that these artificial functionalities confer.

Zusammenfassung

Enzyme sind bemerkenswerte Katalysatoren. Diese Proteinmoleküle wurden über Jahrmilliarden hinweg evolviert, um mit unerreichter Geschwindigkeit und Selektivität die Reaktionen zu katalysieren, welche die Grundlage des Lebens bilden. Chemiker und Biologen versuchen seit geraumer Zeit, das Potenzial enzymatischer Katalysatoren auch für unnatürliche Reaktionen zu nutzen. Die Entwicklung von Enzymen mit massgeschneiderten Aktivitäten hat sich jedoch als grosse Herausforderung erwiesen. Die vorliegende Arbeit verfolgt daher das Ziel, die Möglichkeiten der enzymatischen Katalyse mit Hilfe künstlicher Metallkofaktoren und β-Aminosäuren zu erweitern.

Enzyme nutzen eine große Vielfalt an Metallen als Kofaktoren für die Katalyse. Dieses Prinzip kann man sich zu Eigen machen, indem man einen synthetischen Metallkomplex, welcher eine unnatürliche Reaktion katalysiert, in die Bindungstasche eines Proteins platziert. Grundsätzlich könnten diese „künstlichen Metallenzyme“ die hohe Reaktionsgeschwindigkeit und Selektivität von Enzymen mit dem breiten Substrat- und Reaktionsspektrum von Übergangsmetallkatalysatoren vereinen. Gemäss dieser Strategie beschreibt Kapitel 2 die kovalente Modifikation eines kleinen Hitzeschockproteins von *Methanocaldococcus jannaschii* mit einem Grubbs-Hoveyda-Olefinmetathesekatalysator. Der Metallkomplex ist mit einer elektrophilen Bromoacetamidgruppe versehen, welche mit einer nukleoophilen Thiolgruppe eines Cysteinerestes im Hitzeschockprotein reagiert. Das daraus resultierende künstliche Metallenzym katalysiert trotz der zahlreichen potentiell inaktivierend wirkenden funktionellen Gruppen des

Proteins eine Ringschlussmetathesereaktion. Das flexible Verbindungsstück, welches Metallkatalysator und Protein kovalent verknüpft, verhindert jedoch eine effektive Präorganisation von Substrat und Katalysator. Damit ist der Einfluss des Proteins auf den Übergangsmetallkomplex gering. Kapitel 3 beschreibt Versuche, den niedermolekularen Katalysator direkt an das Rückgrat des Proteins zu binden, um dieses Problem zu umgehen. In einer Machbarkeitsstudie konnte gezeigt werden, dass kurze, unmodifizierte Peptidliganden wässrige Transferhydrierungen wesentlich beschleunigen. Wasserlösliche Iridiumkomplexe katalysieren etwa in Verbindung mit dem Tripeptid Gly-Gly-Phe die Reduktion einer Vielzahl von Ketonen, Aldehyde, Imine und NAD^+ mit Format als Reduktionsmittel mit hohem Durchsatz. Da als einzige Voraussetzung für eine effiziente Katalyse ein unmodifizierter N-Terminus und ein Glycin als zweiter Rest benötigt werden, sollte es möglich sein, dieses katalytische Motiv an den N-Terminus jedes beliebigen Proteins anzubringen. Gemäss dieser Strategie könnte das in Kapitel 2 beschriebene kapsidbildende Hitzeschockprotein als Nanoreaktor für dynamische kinetische Racematspaltungen eingesetzt werden.

Eine andere Möglichkeit, die Eigenschaften von Proteinkatalysatoren zu erweitern, ist die Verwendung von Bausteinen, welche über die 20 kanonischen α-Aminosäuren hinausgehen. Unnatürliche Oligomere, die häufig als „Foldamere“ bezeichnet werden, bilden in Lösung ebenfalls charakteristische Sekundärstrukturen, welche natürliche Peptide bezüglich struktureller Vielfalt und Stabilität sogar übertreffen. Um von diesen Eigenschaften für die Enzymkatalyse Gebrauch zu machen, untersuchen wir Foldamere mit zwei verschiedenen Bausteinen (α- und β-Aminosäuren) als möglichen Ersatz für ein helikales Fragment einer heterodimeren Chorismatmutase (Kapitel 4). Durch

das systematische Ersetzen von α-Aminosäuren mit den entsprechenden unnatürlichen Bauteilen wurden sieben verschiedene Chorismatmutasechimären produziert, welche bis zu fünf β-Aminosäuren besitzen, unter anderem auch eine Seitenkette, die für die Katalyse besonders wichtig ist. Zwei dieser Varianten zeigten vergleichbare Aktivitäten wie das ursprüngliche Enzym. Aus diesen Experimenten konnten wir wertvolle Richtlinien ableiten, die den Einbau von unnatürlichen Bauteilen in ein breites Spektrum von Proteinen ermöglichen könnten. Damit Enzyme sich die einzigartigen Eigenschaften von Foldameren zu Nutze machen können, ist es besonders wichtig, dass diese Hybridkatalysatoren evolvierbar sind. In Kapitel 5 wird eine evolutionäre Strategie vorgestellt, welche es ermöglichen könnte, den natürlichen Teil der heterodimerischen Chorismatmutase an das künstliche Gegenstück anzupassen. Dafür haben wir ein Auswahlsystem etabliert, das es uns ermöglicht, aus parallel aufgereinigten Proteinbibliotheken des natürlichen Fragments diejenigen Proteine zu bestimmen, welche die gewünschte Reaktion am besten katalysieren. Obwohl in dieser Machbarkeitsstudie bislang keine verbesserte Hybridchorismatmutase gefunden wurde, ist zu erwarten, dass umfangreichere Evolutionsexperimente diese schlussendlich hervorbringen werden.

Die vorliegende Arbeit beschreibt Strategien, um die Möglichkeiten der enzymatischen Katalysatoren zu erweitern und die erworbenen Fähigkeiten anschliessend zu evolvieren. Mit Hilfe unserer Ergebnisse könnten zukünftige Experimente in der Lage sein, eine neue Klasse von Hybridenzymen zu erzeugen, die die Vorteile der künstlichen Funktionalitäten in vollem Masse ausschöpfen.

1. Introduction

1.1. The origin of life

Roughly 4 billion years ago, life emerged from the primordial soup in the form of polymers that, shaped by prebiotic chemistry, were capable of self-replication from available precursors (**Figure 1**).[1] In order to persist (or survive) in a changing environment, these early "life forms" acquired rudimentary metabolic pathways to compensate for the increasing demand for precursors and the ability of compartmentalization to retain the synthesized molecules in close proximity to the replicator. Although the nature of the initial heritable polymer remains elusive, there is evidence that, when life increased in complexity, RNA became dominant in early life forms.[2-3] This scenario, often referred to as the RNA world, is plausible because RNA is able to maintain genetic information (i.e. retroviruses) and perform catalytic tasks, among others, necessary for replication.[4-5] The degree of complexity of RNA-based organisms remains unclear, but was likely limited by the prowess of ribozymes and their ability to maintain a stable genome upon self-replication.[1,6] As a consequence, peptides and proteins were ultimately added as a new class of biopolymers over the next 400 million years, initially to improve the catalytic performance of ribozymes and subsequently as freestanding catalysts. Similarly, DNA was adopted as the genetic material because of its higher stability compared to RNA. Some 3.6 billion years ago, life based on these three biopolymers started to diversify and conquer most niches of the planet.

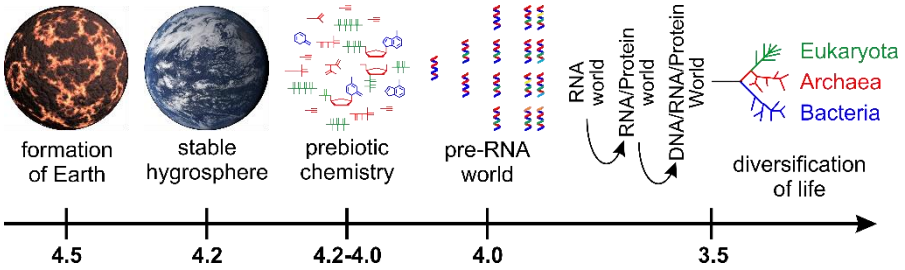


Figure 1: Timeline of plausible events in the early history of life on Earth, with approximate dates in billions of years before the present (adapted from reference [1]).

1.2. From RNA to protein catalysts

1.2.1. The origin of translation

In 1976, Crick famously described the problem of the origin of protein synthesis as *notoriously difficult*.[7] Because evolution is opportunistic and the selection of catalytic RNA fast and relatively straightforward,[8-9] a key question concerning the evolution of protein catalysts is why would an evolved ribo-organism venture into the protein world?[10] It is tempting to suggest that the greater versatility of amino acid side chains simply allows for constructing more efficient catalysts. However, evolution is also a myopic process and does not select for traits that will become advantageous in the far future. The idea that a mechanism for translation spontaneously came into existence can also be dismissed, because even in its most primitive form such a machinery requires multiple components to function.[11] Thus, we must conclude that the first replication apparatus was assembled from existing parts. Moreover, such a rudimentary apparatus probably did not synthesize functional enzymes but rather catalyzed the formation of small peptides, quite possibly of ill-defined

length and sequence. Nevertheless, in order for a primitive translation machinery to persist in a ribo-organism, these peptides needed to convey an immediate evolutionary advantage.[12]

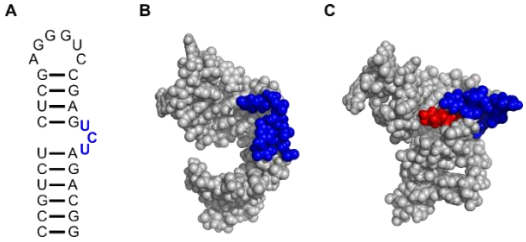


Figure 2: The formation of a hairpin structure in TAR RNA from the HIV virus can be triggered by addition of arginine amide. **A**: Sequence of the TAR RNA hairpin with a bulge region critical for the structural rearrangement in blue. **B**: Structure of the HIV TAR RNA in the absence of a ligand or binding protein (PDB 1ANR). **C**: Structure of the same RNA upon addition of arginine amide (red). The bulge region (blue) undergoes a dramatic rearrangement compared to the apo structure (PDB: 1ARJ).

Benefits of recruiting amino acids and short peptides to an RNA molecule can be deduced from experiments. The ability of RNA to adopt complex structures by itself is limited,[13] but can be significantly expanded by providing adequate binding partners. For example, ribozymes from in vitro evolution experiments typically display higher activities in the presence of divalent metal ions that stabilize the RNA structures.[14] Conversely, natural ribonucleotide-containing proteins (RNPs) rarely retain their fold when stripped of their native binding proteins.[15] Interestingly, for some RNPs this process can be reversed by adding short peptides or amino acids. The hairpin-like structure of the trans-activation response (TAR) element of the HIV RNA, for example, only folds upon binding to the Tat protein, an event crucial for viral function.[16] After identifying an arginine-rich sequence in Tat as the main interaction partner, it was demonstrated that the addition of short, arginine-rich peptides or even arginine amide in high concentrations is sufficient to trigger hairpin formation

of TAR (**Figure 2**).[17-18] Thus, it is conceivable that amino acids and peptides in an RNA world could have been critical elements that expanded the structural diversity of functional RNA molecules. Similarly, amino acids and peptides may have been recruited to ribozymes to engage in catalytic activities with their side chains.[19] By both means, peptides could eventually have become essential for survival, consequently favoring ribo-organisms with the ability to synthesize these protein precursors.

Furthermore, the ability of ribozymes to catalyze reactions associated with the synthesis of peptides has been demonstrated by in vitro selection of such catalysts.[20-22] Notably, if amino acids in an RNA world were recruited by ribozymes via an RNA handle, the evolution of a template mechanism for protein synthesis, such as the one found in the ribosome today, would become imaginable. However, the details of how the genetic code has evolved, how the 20 standard amino acids were chosen, and why the genetic code is almost universally shared throughout all organisms, remain poorly understood.[10-11, 23-24] In order to understand how enzymes have evolved, it suffices to accept the premise that amino acids and small peptides originally served as structural and possibly catalytic cofactors. Thus, the biosynthesis of these protein precursors offered a significant evolutionary advantage to their hosts, ultimately enabling the evolution of a sophisticated translation machinery, namely the ribosome.

1.2.2. RNA → RNP → Protein

Although the reign of the RNA world ended more than 3.6 billion years ago, relics from this period can still be found in modern organisms.[25] The ribosome,

an ancient ribozyme that is held together by an armada of binding proteins, is one example.[26] Ribonucleotides attached to enzymatic cofactors such as NADH, FAD, or SAM are another.[11] Moreover, ribozymes and RNPs that persisted through eons of evolution allow us to reconstruct how the transition from ribozymes to enzymes might have occurred.

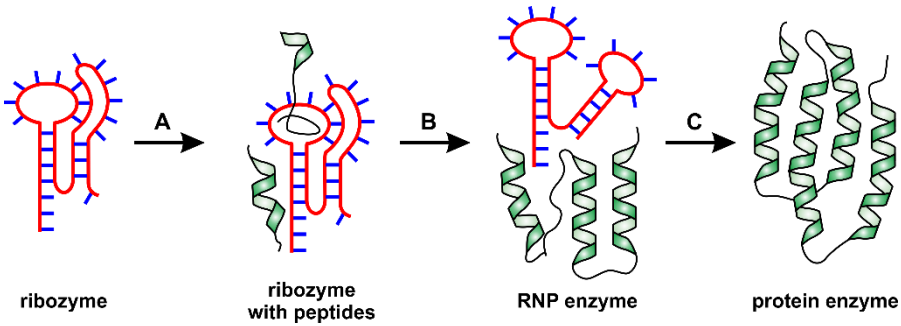


Figure 3: Model for the stepwise evolution of protein catalysts. **A**: Ribozymes from the RNA world (left) would first recruit peptides to increase their performance. **B**: When peptide synthesis became more efficient, the proteinogenic part increased in length at the expense of the original ribozyme. **C**: Ultimately, the RNA segment would become nonessential and the protein enzyme therefore solely responsible for catalysis (adapted from reference [27]).

In general, it is believed that ribozymes were replaced by enzymes in a stepwise fashion, with the transition going from ribozymes to RNP catalysts and ultimately to enzymes (**Figure 3**).[6, 12, 27] As previously described, ribozymes presumably recruited peptides initially for structural and catalytic purposes. Putatively, as peptide synthesis became more proficient, the role of RNA in catalysis shrank at the expense of the more efficient and versatile protein fragment and in the last stage disappeared completely when enzymes became solely responsible for catalysis. However, for some ribozymes this transition

was not completed, leaving unique traces of the evolution of protein catalysts. The ribosome, RNase P,[28] telomerases,[29] the splicosome,[30] group I and group II introns,[31] and signal recognition particles[32] are hallmark model systems for an incomplete transition.

RNase P, for example, generates the mature 5' end of tRNA from different precursors by site-selective cleavage of a phosphodiester bond.[28] Furthermore, RNase P is an ancient riboprotein catalyst (RNP enzyme) that consists of an RNA fragment and 1 – 9 additional proteins that assist in substrate/product discrimination and stabilize the fold of the ribozyme. The RNA portion of RNase P is highly conserved and participates in catalysis in all kingdoms of life, albeit to different extents.[33] While it provides all functional groups necessary for catalysis in bacteria, in archaea and eukaryotes the active site is formed jointly by RNA and proteins. Consequently, when stripped of its protein partners, the core RNA from some bacteria retains the ability to cleave the tRNA precursor in the presence of divalent ions at high concentration, while the corresponding eukaryotic RNA often loses its ribozyme activity.[34] The recent discovery that mitochondrial RNase P does not harbor any RNA molecule completes, in principle, the transition from RNA → RNP → protein catalysts.[35] However, sequence analysis of the three proteins responsible for tRNA maturation activity in mitochondria show that they are not related to any proteins usually associated with RNase P.[36] This result suggests that the catalytic scaffold was assembled from existing protein catalysts within the genome.

The more common path by which RNA eventually became nonessential for the activity of RNP catalysts can be reconstructed from the evolution of signal recognition particles (SRPs). SRPs bind to the signaling sequence of secreted

proteins and are RNP enzymes in all kingdoms of life. Portions of the RNA structure and an associated GTPase are highly conserved.[37] However, unlike RNase P or the ribosome, the RNA in SRP does not provide residues essential for activity, but contributes to recognition of the signaling sequence and facilitates the binding of SRP to its receptors. Because proteins are routinely capable of performing these functions themselves, it is not surprising that the SRP that targets light-harvesting chlorophyll proteins in plants was found to contain a homologous GTPase but no RNA.[38] Therefore, this system provides an example for the last step of the gradual substitution of ribozymes with more powerful protein catalysts.[27]

Overall, the evolution of enzymes was a logical and retraceable consequence of the emergence of a translation machinery. Over the last 3.6 billion years, constant tinkering with the amino acid sequence of primordial enzymes boosted their initial activities far beyond the capabilities of ribozymes or RNP enzymes. As a consequence, enzymes have greatly facilitated the evolution of highly adapted species that spread over the planet.

1.3. How enzymes work

1.3.1. General principles of enzyme catalysis

Enzymes are linear polypeptides that upon folding into complex three dimensional structures catalyze most reactions necessary for survival.[39] From a historical perspective, Emil Fischer suggested the “lock and key” model as one of the first hypotheses for how enzymes work.[40] In the absence of structural information, he surmised that a substrate molecule binds to an active species

(in modern terms the active site) within the enzyme and as a result undergoes activation. In this model the enzyme is the lock that does not quite fit the substrate (the key) but exerts a certain strain on it, an effect often referred to as ground state stabilization.[41] This theory was later refined by Koshland who argued that the high selectivity of enzymes is a result of the active site only forming in the presence of the substrate.[42] In this theory, motions in the enzyme's loop regions are responsible for creating a productive enzyme-substrate complex. Triosephosphate isomerase is a notable example in which such a lid mechanism has been identified as crucial for accelerating the interconversion of dihydroxyacetone phosphate to D-glyceraldehyde-3-phosphate at rates close to the diffusion limit while avoiding the formation of side products.[43]

Nevertheless, the "induced fit" model of Koshland has no explanatory power for how enzymes achieve their large rate accelerations. It was Linus Pauling who suggested that enzymes, like all catalysts, speed up transformations by stabilizing the transition state of a given reaction.[44] This notion becomes more evident when a catalytic cycle in the presence and the absence of the enzyme is considered (**Figure 4A**). By utilizing transition-state theory, this cycle can be translated to a reaction coordinate with the corresponding Gibbs free energies.[45] From **Figure 4B** one can easily deduce that catalysis only occurs if an enzyme binds the transition state tighter than the substrate. Comparing the enzyme-catalyzed with the non-catalyzed reaction (k_{cat}/k_{uncat}) provides an estimate for the achieved rate accelerations. Staphylococcus nuclease, for example, speeds up the rate for hydrolysis of an ionized phosphate monoester by 10^{15} to 10^{17} fold compared to the spontaneous reaction.[46] Transition state theory can also be used to calculate the hypothetical affinity of the enzyme

active site to a transition state. From the thermodynamic cycle in **Figure 4A** theoretical dissociation constants (K_{TS}) of up to 10^{-23} M for the decarboxylation of orotic acid have been deduced.[47]

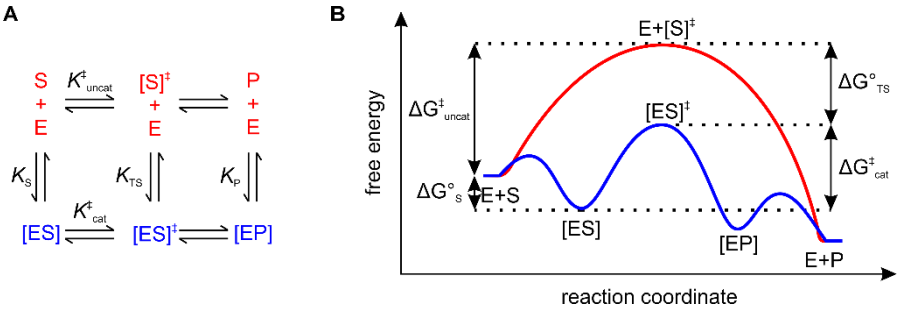


Figure 4: General principles of enzyme catalysis: **A**: Thermodynamic cycle of a reaction in the presence (blue) and in the absence (red) of the enzyme, with S being the substrate, E the enzyme and P the product. Complexes between S and P with enzymes are shown in brackets. **B**: The reaction coordinate with Gibbs free energies of the catalyzed and uncatalyzed reaction. Note that $\Delta G^{\ddagger}_{cat} < \Delta G^{\ddagger}_{uncat}$ is only true if $\Delta G^{\circ}_{TS} > \Delta G^{\circ}_{S}$.

The catalytic power of enzymes will always appear as a result of increased transition-state stabilization for the enzymatic process relative to the spontaneous reaction. The specific forces that give rise to the large rate accelerations include electrostatic, steric, hydrogen bonding, or differential solvation effects.[41] Much of the work in mechanistic enzymology over the last decades has been devoted to quantifying the contributions of accurately positioned catalytic groups and protein motions to transition state stabilization by biophysical characterization and computationally-assisted calculations. In the following section, we will discuss the enzyme chorismate mutase as a model enzyme in more detail.

1.3.2. Chorismate mutase – a case study

The intramolecular rearrangement of chorismate to prephenate is the first committed step in the biosynthesis of phenylalanine and tyrosine via the shikimate pathway (**Figure 5**).[48] As a rare example of a pericyclic reaction in biology, this transformation has been intensively studied over the last decades.[49] The [3+3]-sigmatropic rearrangement is strongly exergonic and occurs spontaneously at an appreciable rate in the absence of enzyme via a concerted but asynchronous pathway involving a chair-like transition state.[50-51] In bacteria, fungi, and higher plants the enzyme chorismate mutase speeds up the rearrangement by about a million fold. Over the last decades, biophysical characterization and calculation have been used to identify to which extent conformationally constraining the flexible substrate, specific hydrogen bonding to the transition state, and electrostatic interactions contribute to the catalytic performance of this enzyme.

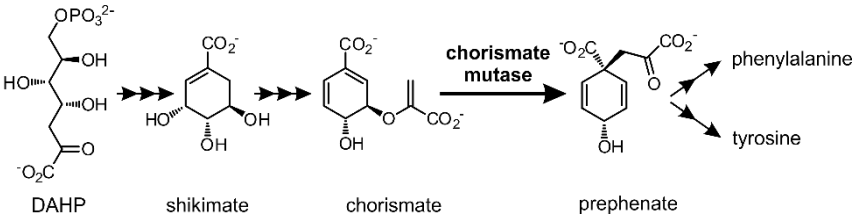


Figure 5: The shikimate pathway. In bacteria, fungi, and higher plants 3-deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) is converted to the amino acids phenylalanine and tyrosine through a series of enzyme catalyzed reactions. Notable intermediates are shown with the transformation catalyzed by chorismate mutase in the center.

The enzymatic reaction, like its uncatalyzed counterpart, proceeds via a chair like transition state. With (Z)-[9-^{3}H]chorismate as a substrate both

reactions afford the (S)-[9-^{3}H]prephenate isomer, a result that effectively rules out a boat-type transition state (for numbering see **Figure 6A**).[51] A large ^{18}O kinetic isotope effect on C-O bond cleavage and a smaller ^{13}C effect on C-C bond formation for the catalyzed reaction provide evidence for a concerted but asynchronous pathway.[52] As in the uncatalyzed rearrangement, C(1)-C(9) bond formation lags considerably behind C(5)-O(5) bond cleavage, yielding a polarized species in the transition state (**Figure 6A**). Furthermore, structural data is consistent with a mechanism in which the flexible chorismate molecule is fixed in the pseudo-diaxial conformation at the active site of the enzyme.[53-54] As deduced from available atomic resolution data, stabilization of the developing partial charge at the transition state also appears important for catalysis, as a positively charged side chain is commonly found in close vicinity to the ether oxygen (**Figure 6B** shows a schematic representation of the active site of a chorismate mutase from *Bacillus subtilis*). Thus, an alternative mechanism, in which an enzymatic nucleophile assists in heterolytically cleaving the chorismate ether bond can be ruled out.[52]

Dissecting the relative contributions of substrate constraint and electrostatic effects to catalysis has proven challenging. The active site of the *B. Subtilis* chorismate mutase, for example, displays only a 20-fold higher affinity for an isosteric transition state analog (TSA, **Figure 6B**) than for the substrate.[55] What at first sight might indicate that transition state stabilization is indeed less important for catalysis - given that TSAs are usually bound much tighter at the active site than the substrate - must be viewed with caution. The TSA, although isosteric, is not polarized like the actual transition state and therefore might not accurately reflect the binding constant for the true transition-state. In this regard, the performance of catalytic antibodies that

were raised against the TSA is notable.[56] Although these man-made catalysts display excellent binding for the hapten, they are poor catalysts. In the crystal structure of one of these catalytic antibodies (1F7), the strictly conserved positively charged side chain proximal to the ether oxygen is absent, underscoring its importance for transition-state stabilization.

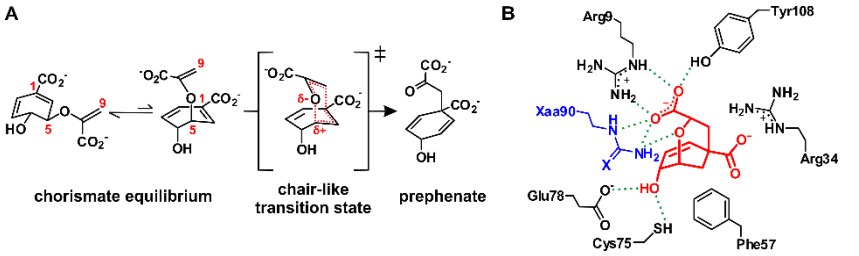


Figure 6: Principles for the enzyme catalyzed transformation of chorismate to prephenate: **A**: The chorismate reaction with the pre-equilibrium between the extended diequitorial and the compact diaxial form of chorismate is shown. The numbering of positions involved in catalysis is in red. The chair-like transition state is shown with its partial charge highlighted in red. **B**: Schematic view of the active site of the chorismate mutase from *B. subtilis*. Interactions with the TSA (red) are shown in green. For the wild-type enzyme the blue residue at position 90 corresponds to arginine (X = NH_2^+), while for the variant containing citrulline (X = O) (adapted from reference [55]).

Quantifying the contribution of electrostatics to catalysis has been attempted by experiment. In line with its crucial role in catalysis, even conservative mutations of the cationic residue in *B. subtilis* chorismate mutase, Arg90, drastically decrease catalytic efficacy.[57-59] However, these mutations are typically accompanied by large increases in K_m values, making it complicated to distinguish between ground state and transition state effects. To escape this catch-22 situation, Hilvert and coworkers prepared a *B. subtilis* chorismate mutase variant containing citrulline, an isosteric but neutral arginine analog, at position 90 by expressed protein ligation.[55] Substituting the positively charged guanidinium group with the neutral urea derivative caused

a >10^4 fold decrease in k_{cat} (a destabilizing effect of about 6.5 kcal/mol) with the K_m value being only 2.7 fold increased. Similarly, the inhibition constant for the TSA increased only 5.7 fold as a result of this mutation, corresponding to a 1.1 kcal/mol less favorable free energy of binding. These values underscore and quantify the importance of stabilizing the polarized transition state and demonstrate that shape complementary to the reactive substrate conformer is not sufficient for efficient catalysis of the chorismate mutase reaction.

Although computationally-assisted calculations have long underestimated the role of electrostatics in the mechanism of chorismate mutase,[60] more recent high-level QM/MM simulations agree well with the results obtained by experiment.[61] Multiple profiles for the enzymatic reaction from chorismate to prephenate estimated an average stabilization of 4.7 kcal/mol for the electrostatic term. Notably, when adding the calculated difference in energy for forming a reactive pseudo-diaxial conformation in solution compared to the enzyme (3.8 – 5 kcal/mol), the sum yields a value that is comparable to the experimentally observed catalytic effect ($\Delta\Delta G^{\ddagger}$ = 9.1 kcal/mol). Thus, it appears that conformational effects – such as sampling of the diaxial conformer from the solution (equilibrium ≈10:1 in favor of the extended form)[62] – and electrostatic transition-state stabilization relative to the bound substrate contribute roughly equally to catalysis in chorismate mutase.

1.3.3. From understanding to design

Overall, mechanistic studies of the model enzyme chorismate mutase demonstrate that our understanding of enzyme catalysis is steadily increasing. Starting from general principles, we are beginning to quantify the different effects responsible for the large rate accelerations that are achieved. This

steady progress is also evident from the increasing efficacy of designed enzymes.[63] At an early stage, protein catalysts could only be created by taking advantage of the immune system to enrich tight binders for transition-state analogs.[56] The resulting catalytic antibodies, however, performed poorly compared to natural enzymes. The advent of computation has significantly expanded our attempts to create artificial enzymes. Transition states and different mechanisms to stabilize them can be quantitatively calculated in silico and the resulting *theozyme* directly transplanted into an existing protein scaffold.[63] Catalysts that perform promisingly in silico can be expressed in vivo and tested for activity. Although initial activities are again inferior to those achieved by natural catalysts, laboratory evolution can be exploited to boost the performance of these artificial enzymes to levels that rival their natural counterparts.[64] In principle, careful analysis of the evolved catalyst can uncover shortcomings in the design algorithms, thereby yielding valuable insights toward a more complete understanding of enzyme catalysis.[63]

1.4. Expanding the scope of enzyme catalysis – learning from nature

1.4.1. Limits of enzyme catalysis

Over the course of evolution, enzymes replaced ribozymes as nature's first choice of catalysts.[6, 10] Polypeptides offer a greater diversity in folding and functional groups for catalysis. While it is true that the 20 common amino acids with their charged, polar and hydrophobic side chains clearly contain a wider variety of functional groups when compared to the 4 building blocks of RNA,

only a subset of these 20 participates directly in catalysis.[65] As judged by the functionalities highlighted in **Figure 7**, enzymes are clearly well-equipped for general acid-base and nucleophilic catalysis, but to a lesser extent for electrophilic catalysis and redox reactions. Nevertheless, the translational machinery and the possibility for exact control of the placing of side chains in a protein structure allowed for the evolution of different strategies that expand the reaction scope of enzymes vastly and enable them to catalyze the plethora of reactions necessary to sustain life. In principle, mechanisms to broaden the reaction scope of enzymes include the use of genetically encoded amino acids that differ from the 20 standard ones,[66] post-translational modification of enzyme active sites to create reactive functionalities,[67] and recruiting organic and inorganic to the reactive center.[39]

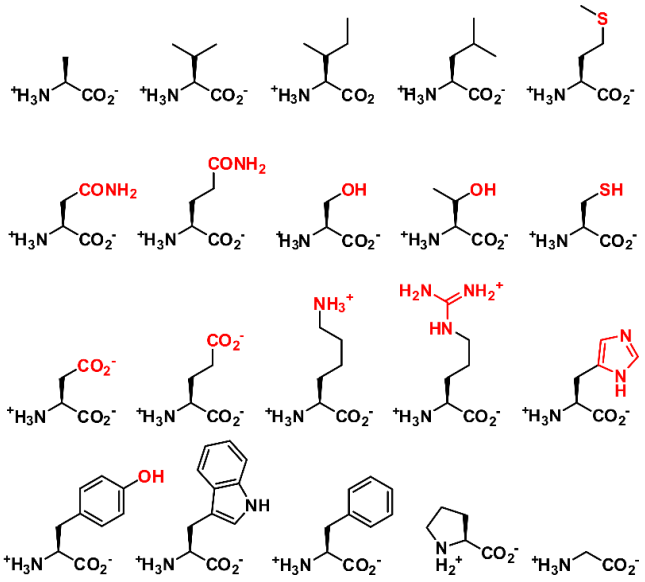


Figure 7: The 20 standard amino acids. Functional groups that are often employed for catalysis of transformations are highlighted in red.

1.4.2. Co-translational and post-translational mechanisms

At the time of its discovery, the genetic code – and with it the choice of the 20 standard amino acids – was considered a "frozen accident" and not a perfect solution.[68-69] Indeed, the high redundancy within the genetic code and the structural similarities of certain amino acids, such as valine, leucine, and isoleucine, support this argument.[70] However, the discovery of the 21st amino acid, selenocysteine, and more recently pyrrolysine as the 22nd addition, has demonstrated that the genetic code is still evolving.[66, 71] Both amino acids are biosynthesized from available precursors in the cell and are incorporated into growing peptide chains by the ribosome. Selenocysteine and pyrrolysine harbor unique functional groups and consequently are almost exclusively found as essential residues at the active sites of enzymes.

Selenocysteine (Sec, U) is a close homolog of cysteine in which the thiol side chain is replaced by a selenol (**Figure 8A**).[72] In bacteria, its co-translational introduction proceeds via suppression of the stop codon UGA and requires a specialized tRNA, a unique elongation factor, and a stem-loop structure (SECIS element) downstream of the stop codon in the mRNA. When compared to its close relative cysteine, the p$\underline{K}_a$ of the selenol group in Sec is lower and thus, a better nucleophile under physiological conditions. Moreover, Sec is more easily oxidized. These properties make this amino acid valuable for catalyzing the formation and rearrangement of disulfide bonds.[73-74] Furthermore, while cysteine-containing proteins can undergo permanent inactivation when the thiol side chain is oxidized to sulfenic or sulfonic acids, this process is reversible for Sec-containing proteins.[75] Consequently, selenocysteine is also utilized for cellular antioxidant defense.

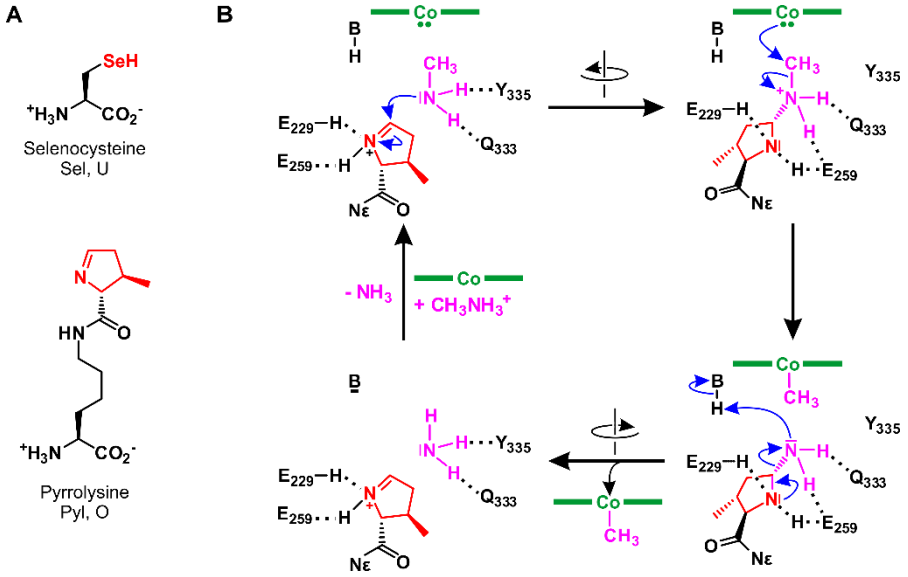


Figure 8: Expanding the scope of enzyme catalysis by co-translationally incorporated amino acids. **A:** Structures of selenocysteine and pyrrolysine with their unique functional groups highlighted in red. **B**: Schematic representation of the catalytic cycle of a methyltransferase, with methylamine as substrate (magenta). The cyclic amine is highlighted in red and the Co-containing corrinoid cofactor in green.

While selenocysteine is found in all three kingdoms of life, pyrrolysine (Pyl, O, **Figure 8A**) so far has only been identified in methanogenic archaea and in one Gram positive bacterium.[76] For these organisms Pyl is essential because it facilitates the generation of coenzyme M, a key intermediate in the energy cycle of methane-producing organisms. In cells, Pyl is biosynthesized from two lysines and incorporated into proteins by decoding the stop codon UAG by a process that is not fully understood yet.[77] The cyclic imine in the side chain of Pyl is a highly reactive electrophile that has been shown to be the corner stone at the active site of methyltransferases.[78] These enzymes catalyze the transfer

of a methyl group from an amine substrate to a corrinoid cofactor. In the catalytic cycle (**Figure 8B**), Pyl is protonated by a glutamic acid side chain which facilitates the nucleophilic attack of the substrate amine onto the imine bond. In the next step, rotation of the pyrrolidine side chain positions the methyl group for abstraction by a cobalt-containing corrinoid cofactor. In subsequent steps, the methylated cofactor and the product are released, and the catalytic cycle is completed upon binding another substrate molecule.

It is conceivable that a more detailed analysis of available sequence information and advances in crystallography will identify further additions to the genetic code. As an alternative to this co-translational strategy of expanding the scope of enzyme catalysis, nature has also evolved mechanisms that rely on post-translational modification of protein residues.[67] Recent access to high resolution crystal structures has revealed a series of unanticipated side chain modifications within some protein active sites that have forced a reevaluation of the catalytic mechanism of the enzyme in question. A prominent example in which active site remodeling endows a protein scaffold with a unique function is the green fluorescent protein (GFP).[79] In this protein, which has become indispensable for a wide variety of applications in chemical biology, a fluorescent chromophore is formed autocatalytically inside a β-barrel from adjacent tryptophan and glycine residues (**Figure 9A**). A post-translational modification that proceeds via a mechanism that is reminiscent of fluorophore formation in GFP has been identified at the active site of phenyl ammonia lyases (PALs). Atomic resolution data from the *Pseudomonas putida* PAL revealed the presence of 4-methylidene-imidazole-5-one (MIO, **Figure 9B**). It is believed that this species forms in an autocatalytic cyclization of an alanine and glycine residue, followed by dehydration of a serine residue. For PAL activity, it has

been suggested that MIO enables an electrophilic attack on the phenyl group of the substrate, phenylalanine. For this step the loss of aromaticity at the benzene ring is compensated by the concomitant aromatization of MIO and as a result, readily enables proton abstraction at the β-position of the substrate.

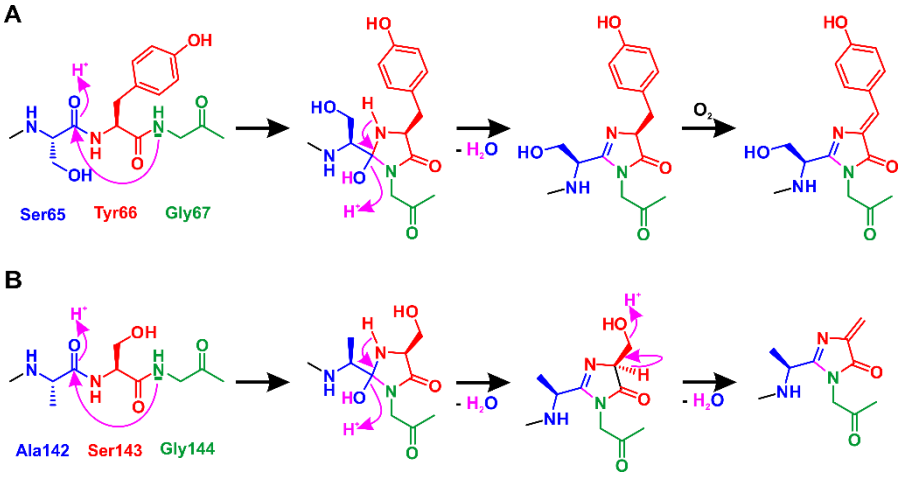


Figure 9: The autocatalytic formations of the fluorophore in the green fluorescence protein (**A**) and of 4-methylidene-imidazole-5-one at the active site phenyl ammonia lyases (**B**, adapted from reference [67]).

Over the last decades, a multitude of active site modifications that introduce novel functionalities into a protein scaffold have been uncovered. In many cases, these modifications involve tyrosine, tryptophan, or cysteine residues and activities that are introduced are often useful for electrophilic catalysis or in redox reactions.[67] With mass spectrometry becoming a more useful means to determine such rearrangements, future studies will not only uncover the full range of activities that these modifications enable but also shed light on how frequently nature uses them to expand the scope of enzyme catalysis.

1.4.3. Metal ions and metal cofactors

While the gamut of functional groups present in proteins is only of limited utility for catalyzing the plethora of reactions necessary to sustain life, accurate positioning of amino acid side chains readily enables binding of small molecules and metal ions with novel activities of their own.[65] Consequently, enzymes often recruit organic and inorganic cofactors to their active site in order to significantly broaden their reaction scope. Organic cofactors, for example, typically are equipped with functional groups that endow a proteinogenic scaffold with a wide variety of useful activities.[39] A small selection of these structures and the transformations that their introduction enables are depicted in **Figure 10**.

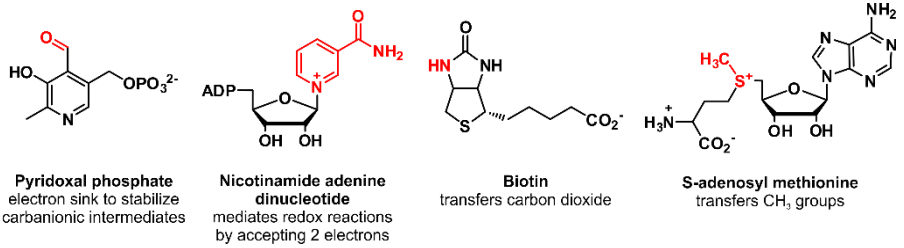


Figure 10: Examples of organic cofactors that are recruited by enzymes to expand their reaction scope. The active portion of the cofactor is highlighted in red.

Compared with organic cofactors, the incorporation of metal ions represents a more versatile strategy for enzymes to expand their catalytic palette.[80] Indeed, bioinformatic studies suggest that roughly half of all known proteins bear a metal ion within their scaffold.[81] More importantly, one third of all enzymes are metalloenzymes. These enzymes bind metal ions in their active site either via coordinating side chains or via non-proteinogenic, metalated ligands to catalyze a vast array of transformations. More specifically,

metals confer function to a protein scaffold by providing strong electrophilic centers, specific ligand affinities (i.e.: O_2, CO_2, or N_2, ATP, and others), multiple stable oxidation states, and other means.[82] Metals utilized by enzymes include most of the first row transition-metals but also magnesium, calcium, sodium, potassium, tungsten, molybdenum, and even cadmium have been identified.[83] Because these metals bind to the gamut of functional groups present in proteins with different propensities and sometimes require similar coordination geometries, the supply of metal ions in cells is tightly regulated to avoid mis-metalation.[84]

Metalloenzymes catalyze transformations that are among the most challenging reactions known in chemistry. The reduction of nitrogen to ammonia, for example, requires enormous amounts of energy to proceed in organic chemistry, but occurs under mild conditions in nitrogen fixing organisms that utilize molybdenum-iron-dependent nitrogenases.[85] Not surprisingly, chemists and biologists have long sought to understand the origin of the large rate accelerations and enviable selectivities of metalloenzymes. Mechanistic studies that take advantage of the unique spectroscopic and magnetic properties of metalloenzymes have elucidated some of the strategies that nature uses to exploit metal ions for catalysis.[80-81] In general, all metalloenzymes bind metal ions (or a metal-containing prosthetic group) at their active site with one labile coordination site. This strategy conveniently dictates where catalysis takes place and enables proteins to bind substrates in close proximity to the reactive center. Metalloenzyme are represented in all six enzyme classes and, as previously noted, utilize a variety of metals. Consequently, a concise overview of the plethora of structures and functions of these catalysts will not be attempted. Instead, the following paragraphs will

discuss two prominent examples of iron-dependent oxygenases and the wide spread class of zinc hydrolases to demonstrate our current understanding of catalysis in metalloenzymes.[86-87]

Iron, as the most abundant metal in the earth's crust, is a logical choice for an ion to be used in metalloenzymes. Under aerobic conditions iron's oxophilicity is not only exploited in hemoglobin to supply cells with oxygen but is also utilized to catalyze a wide variety of different oxidation reactions. Iron-dependent oxygenases can be categorized into enzymes that bind the ion through the heme prosthetic group and catalysts that recruit the metal ion directly via coordinating amino acids.[86] Independent of the source of the ion, oxidation is believed to be catalyzed by an iron(IV)oxo species that forms in the catalytic cycle after activation by molecular oxygen.

Non-heme iron-containing enzymes are further subdivided into monoxygenases and dioxygenases.[88] While the former transfer only one oxygen from O_2 to the substrate with the other converted to H_2O, the latter can incorporate both oxygen atoms into one or two substrate(s). In non-heme iron dependent oxygenases the iron is usually coordinated via a facial triad consisting of a pair of histidines and a carboxylate side chain (**Figure 11,** center). A well-understood class of such dioxygenases are enzymes that use α-ketoglutarate (α-KG) as a co-substrate.[89] Among other reaction types, α-KG-dependent dioxygenases (α-KDDs) facilitate hydroxylation of unreactive C-H bonds, a reaction valuable to organic chemistry. A typical catalytic cycle for such an oxidation catalyzed by α-KDD is depicted in **Figure 11**. In the resting state of these enzymes, iron is bound to the active site via the facial triad and three water ligands.[90] Upon consecutive binding of α-KG and the substrate all water molecules are displaced from the catalytic center to

yield a vacant coordination site. This position is subsequently filled by molecular oxygen to form an iron(III)superoxo species. In this intermediate the terminal oxygen of the superoxo group attacks the keto position of the bound α-KG to form a bicyclic ring system which upon eliminating CO_2 is believed to afford the highly reactive iron(IV)oxo complex. Abstraction of a proton from the substrate generates a radical that in an *oxygen rebound* mechanism yields an Fe^{2+} complex and the hydroxylated product. Upon release of the converted substrate, CO_2, and succinate, the resting state of the catalyst is restored by addition of three water ligands.

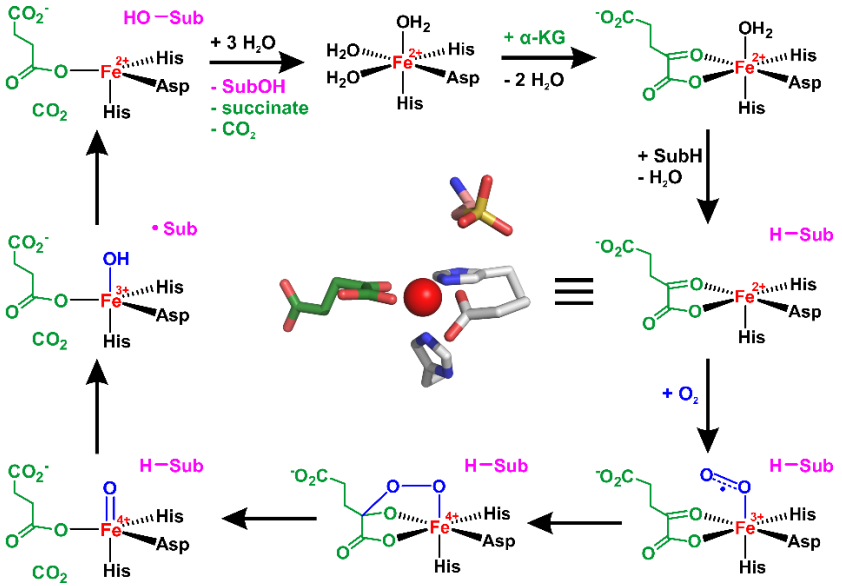


Figure 11: Schematic overview of the catalytic cycle (starting at the top middle position) of an α-ketoglutarate-dependent dioxygenase. The substrate is denoted as Sub (in magenta), α-ketoglutarate is shown in green, and molecular oxygen in blue. The intermediate in the middle is from the crystal structure of a taurine dioxygenase (PDB 1GY9). The iron is shown in red, protein ligands in grey, and taurine above the catalytic metal ion. The individual steps of the catalytic cycle are described in the text (adapted from reference [90]).

In contrast to the first class of iron-dependent oxygenases, which include α-KDDs, metalloenzymes that utilize heme (**Figure 12A**) are strictly monooxygenases.[86] A particularly versatile class of these enzymes are cytochrome P450s, commonly termed P450s. [91] These metalloenzymes are widely studied as a consequence of their multitude of physiological functions in organisms, and also for their biotechnological potential.[92] The catalytic cycle of P450s starts from an Fe^{3+} species, and consequently requires a redox partner to form an oxophilic Fe^{2+} center. The iron(IV)oxo species is formed after one of the oxygens from molecular O_2 is converted to H_2O. Subsequent oxygen insertion into an unreactive C-H bond in the substrate regenerates the original Fe(III) center. P450s are notable for their high regio- and stereoselective oxidation of non-activated C-H bonds, abilities that are only rarely obtained with traditional small molecule catalysts. As a consequence, understanding the mechanistic details of these complex enzymes has become increasingly important in order to exploit P450s for synthetic purposes.[92]

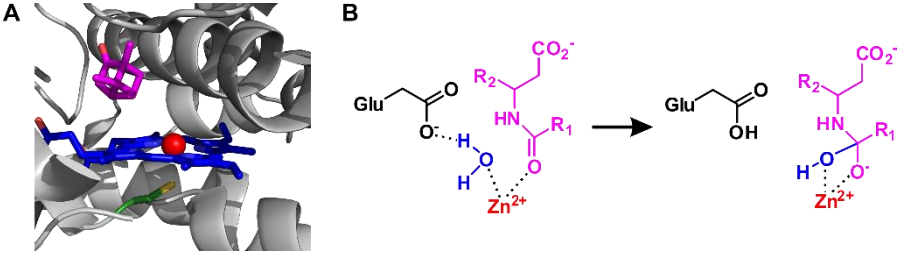


Figure 12: The use of metal ions for catalysis. **A**: Crystal structure of the P450cam with bound heme-cofactor. The porphyrin ring is shown in blue and the iron in red, the coordinating cysteine side chain in green, and bound substrate in magenta (PDB 4L4E). **B**: Initial step of the hydrolysis of an amide substrate (magenta) by zinc proteases. Coordinating amino acids that position the zinc (red) in the active site are omitted for clarity.

While mechanistic studies on α-KDDs and P450s highlight how nature uses the multiple oxidation states of iron for catalysis, other important transition-metals like zinc are not redox active. Metalloenzymes utilize Zn^{2+} ions predominantly for its high Lewis acidity.[93] Zinc-dependent enzymes often take advantage of these properties to coordinate water and lower its pK_a to generate reactive hydroxide ions that can be used to hydrolyze esters, amides, and phosphoesters (**Figure 12B**).[87] The fact that binding of Zn^{2+} is a straightforward means to endow proteinogenic scaffolds with hydrolytic activity has recently been demonstrated by the serendipitous design of a zinc hydrolase in the laboratory. Kuhlman et al. computationally installed a zinc-binding site in a coiled-coil structure that facilitated the dimerization of the protein.[94] Although not intended to, the hydrophobic cleft at the dimer interface was able to bind activated esters adjacent to the zinc center and thereby facilitate hydrolysis.[95] In spite of the low efficiency of this artificial enzyme, future improvements in the design algorithms and a more complete understanding of enzyme catalysis may allow for the creation of better mimics of natural metalloenzymes.[63]

1.5. Aims of this thesis

Over the course of evolution, nature has identified enzymes as its preferred catalysts to accelerate the pool of transformations that enables life. Although the recent past has witnessed tremendous progress toward a solid qualitative understanding of how enzymes work, emulating nature's strategies to create catalysts that display made-to-order activities has proven to be a formidable challenge.[63] In this thesis, two methods for expanding the scope of enzyme

catalysis by chemical means and evolving the acquired functions by directed evolution are explored.

Transition-metal complexes that accelerate a wide variety of different reactions are the result of man's attempts to exploit metal ions for catalysis.[96-97] Small-molecule catalysts that are routinely used for the chemical synthesis of organic compounds often utilize metals that are not found in natural metalloenzymes. Consequently, the incorporation of such metal catalysts into a proteinogenic scaffold represents one means of creating artificial enzymes that display activities for reactions with no natural counterpart.[98] In Chapters 2 and 3 we investigate the feasibility of using this approach to generate hybrid catalysts for olefin metathesis and transfer hydrogenation reactions. More specifically, we set out to covalently modify a capsid-forming protein with a Grubbs-Hoveyda-type catalyst and fill the luminal space of the cage-like structure via a ring-opening metathesis polymerization (Chapter 2). In Chapter 3, we investigate the potential of small, unmodified peptides for possible ligand acceleration effects in the transition-metal-catalyzed transfer hydrogenation of ketones and imines. We aim to discover genetically-encodable tags that, when appended onto a protein sequence, can readily endow a scaffold with transfer hydrogenation activity.[99] By creating such hybrid species we hope to identify strategies that allow for the introduction of novel, synthetically useful activities into a wide variety of proteins.[100] Although the activities of such hybrid enzymes may pale in comparison to the rate accelerations achieved by natural metalloenzymes, we anticipate that it will be possible to improve *artificial metalloenzymes* by evolutionary algorithms.[101] By identifying beneficial effects exerted by the

protein scaffold, these improved catalysts could become valuable for synthetic purposes.

The folding of linear, sequence-specific peptide chains into well-defined structures is a prerequisite for enzyme catalysis. However, over the last decades the study of foldamers – non-natural oligomers that adopt compact conformation in solution – has vividly demonstrated that peptides comprised of α-amino acids are not unique in their ability to fold.[102] In fact, foldamers can adopt a diverse set of highly stable secondary structures that could be interesting in the context of expanding enzyme function. Nevertheless, the construction of protein-like catalysts entirely made from nonstandard amino acids is an extremely difficult task.[103] Instead, we suggest a stepwise approach that approach that gradually converts an enzyme into an artificial catalyst. In Chapter 4, we ask the question which tasks of enzyme catalysis can and cannot be mimicked by non-natural amino acids. More specifically we describe attempts to replace a critical α-helical segment in a heterodimeric chorismate mutase with β-amino acid-containing foldamer prostheses.[104] Because such α/β-peptides have been demonstrated to adopt structures that closely resemble α-helices, we surmise that this strategy can lead to hybrid catalysts that retain the chorismate mutase activity of the parent enzyme. To compensate for inevitable changes that arise from introducing such backbone mutations, Chapter 5 evaluates the feasibility of adapting the natural segment of the heterodimeric chorismate mutase to the foldamer prostheses and thereby identify more efficient hybrid catalysts. By demonstrating the feasibility of creating and evolving enzymes with non-natural building blocks, we pave the way to take full advantage of the desirable properties of foldamers to confer functions to a wide variety of different enzymes.

2. An artificial metalloenzyme for olefin metathesis

2.1. Introduction

2.1.1. Merging the best of two worlds – artificial metalloenzymes

Over countless rounds of natural selection, a vast number of enzymes has evolved to catalyze a plethora of reactions. Since the pool of proteinogenic amino acids is limited, enzymes routinely recruit organic or inorganic cofactors to expand and enhance their activities (Chapter 1.4.3.).[39] For example, an estimated 50% of all proteins harbor metal ions or metal complexes in their scaffolds.[81] By placing these reactive species into the well-defined environment of an active site, natural metalloenzymes are able to promote a diverse set of chemical transformations with impressive efficiencies and enviable selectivities.[80] Independently, organic chemists have also exploited the potential of metal ions for catalysis. Most prominently, fine-tuning the properties of transition-metals with small molecule ligands has afforded efficient catalysts for a wide variety of transformations unknown to nature.[96-97] Some of these reactions, such as palladium-catalyzed cross-coupling reactions,[105] olefin metathesis,[106] dihydroxylation,[107] or hydrogenation,[108] have become indispensable for the synthesis of small molecules in research and industry.

Intriguingly, the individual strengths of enzymes and transition-metal catalysts are complementary (**Figure 13**).[100] While natural catalysts are often praised for their superior efficiencies and exacting selectivities, small-molecule catalysts usually display broader substrate scope and allow for a more diverse set of transformations. In an attempt to meld the individual benefits into a single framework and create enzymes with non-natural activities, chemists and biologists have begun to incorporate synthetic transition-metal complexes into protein scaffolds.[81, 98, 109-110] Ideally, such hybrid enzymes – often referred to as artificial metalloenzymes – harness the intrinsic reactivity provided by the artificial cofactor, while the protein environment steers reactivity and selectivity by providing binding interactions for a substrate. In principle, the initial activity of such an artificial metalloenzyme can be further refined by laboratory evolution and ultimately afford tailored catalysts for reactions that lack a biological counterpart.[111]

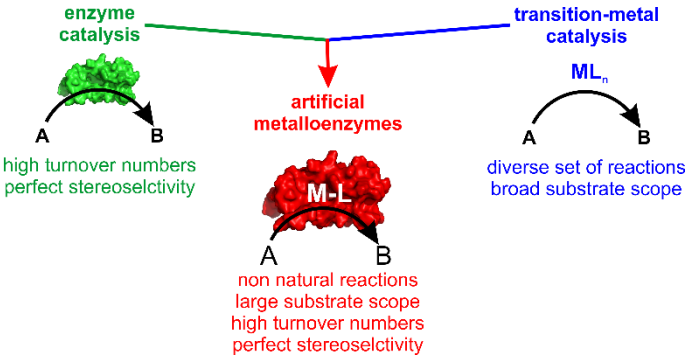


Figure 13: Rationale for artificial metalloenzymes. The individual benefits of enzymes and transition-metal catalysts can be merged by introducing a non-natural cofactor into a protein scaffold (adapted from reference [100]).

Reliable anchoring of the synthetic cofactor into a protein binding pocket is crucial for the synthesis of artificial metalloenzymes. Although efforts toward

this goal have been diverse in the past, anchoring approaches can generally be divided into three distinct strategies: Covalent, dative, and supramolecular anchoring (**Figure 14A-C**).[112]

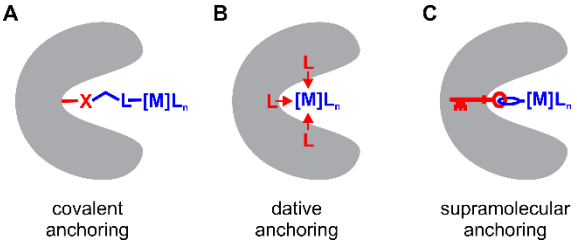


Figure 14: Anchoring strategies for creating artificial metalloenzymes. **A**: Covalent anchoring to a uniquely reactive group (X i.e. the thiol of a cysteine side chain) within the protein scaffold. **B**: Dative anchoring uses coordinating groups from the protein to bind a transition-metal via dative bonds. **C**: An active catalyst is recruited to the protein scaffold by appending a small molecule (represented as a key) that displays high affinity for a binding pocket (the lock).

Covalent anchoring takes advantage of a uniquely reactive group within a protein scaffold to site-selectively attach a transition-metal complex (**Figure 14A**). Ever since the preparation of flavopapain by Kaiser in 1977,[113] alkylation of cysteine residues has remained the mainstay in site-selective protein modification.[114] For the creation of artificial metalloenzymes, an organometallic moiety is first equipped with an appropriate electrophilic group and subsequently linked covalently to the nucleophilic thiol functionality. This strategy is applicable to a wide variety of different protein scaffolds, as unique cysteine residues can be introduced at will by standard cloning procedures. However, covalent anchoring depends on the use of a linker moiety between the electrophilic site and the metal ligand; a factor that severely complicates burying of the artificial cofactor within a protein cavity. As a consequence, the protein of choice rarely participates in the catalytic cycle and artificial

metalloenzymes created by covalent anchoring strategies usually display low selectivities and do not give rise to rate enhancements for the target reaction.[115-118] Using different natural and non-natural side chains for the modification reaction[101, 119] or rigidifying the resulting protein-cofactor adduct via multiple attachment points,[120] may help to overcome these limitations, but have thus far failed to significantly improve the prowess of artificial metalloenzymes.

In dative anchoring strategies (**Figure 14B**) non-natural metal ions are placed within an active site of a protein via dative interactions from amino acid side chains. In the first successful demonstration of this approach, Kaiser et al. took advantage of the promiscuity of the zinc binding site in carboxypeptidase A for copper to generate an artificial metalloenzyme with novel oxidase activity.[121] Since then, substituting the native metal at binding sites in a variety of proteins has become a means to endow hybrid catalysts with sulfoxidation,[122] epoxidation,[123] hydrogenation,[124] or hydroformylation[125] activities. More recently, computational algorithms with the ability to carve metal binding sites into a given scaffold have significantly expanded the utility of dative anchoring.[126-128] These programs have been successfully used to arrange constellations of residues in geometries appropriate for binding a range of metals ions, including zinc, copper and iron.[95, 129-130] De novo proteins have also been equipped with unique metal-binding capabilities and, as a result, appreciable hydrolytic and redox activities.[131-133] Nonetheless, the design of functional metalloenzymes from scratch remains challenging. Avoiding unproductive interactions of the metal ions with residues elsewhere in the protein is one challenge. Tailoring the second and third coordination environments around the metal ion to fine-tune reactivity is another.

Engineering a substrate binding site in proximity to the metal ion is a further hurdle.[81]

For the third approach, supramolecular anchoring, an organometallic complex is recruited to the scaffold protein via an appended small-molecule that exhibits unusually high affinity for a specific protein pocket (**Figure 14C**). A prominent example for such an interaction is the biotin-(strept)avidin system (K_a = 10^{15} M). In a seminal study, Whitesides et al. introduced an achiral, biotinylated rhodium-complex into the secondary coordination sphere of avidin.[134] The resulting artificial metalloenzyme induced modest stereoselectivities (e.e. 44%) in the hydrogenation of N-acetamidoacrylic acid. Twenty five years later, Ward suggested that the selectivity of this hybrid catalyst might be further improved by tailoring the protein environment around the rhodium catalyst.[135] Indeed, using a *chemogenetic* engineering approach in which different biotinylated catalysts were tested in combination with a small number of streptavidin mutants, he identified hybrid enzymes that catalyze the hydrogenation of acrylic acid derivatives efficiently and give rise to either enantiomer in >95% e.e..[136] Within a few years, the Ward group greatly expanded the utility of the biotin-(strept)avidin system, creating enantioselective artificial metalloenzymes for transfer hydrogenation,[137-138] oxidation reactions,[139] and allylic alkylation.[140]

Over the last few decades, the field of artificial metalloenzymes has vividly demonstrated that organometallic complexes are able to expand the scope of enzyme catalysis. With the biotin-(strept)avidin system leading the way, these chimeric enzymes have been shown to be evolvable and able to steer the selectivity of a given reaction. Nevertheless, artificial metalloenzymes generally do not take full advantage of the interactions provided by the protein

environment. The recent report of an enantioselective, rhodium catalyzed C-H insertion in the streptavidin scaffold suggests that synergistic effects from amino acid side chains can significantly boost the performance of artificial metalloenzymes (**Figure 15**).[141]

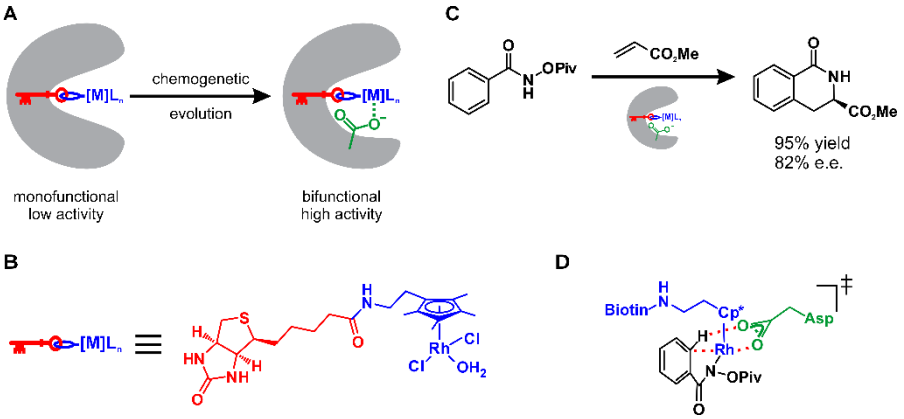


Figure 15: An artificial metalloenzyme for accelerated, asymmetric C-H activation. **A**: Introducing an aspartate into the active side of streptavidin yielded a highly active, bifunctional artificial metalloenzyme for rhodium-catalyzed C-H insertions. **B**: A rhodium catalyst was attached to biotin via the cyclopentadienyl ligand (Cp*). **C**: The depicted C-H insertion reaction was accelerated by 100-fold in the presence of the bifunctional artificial metalloenzyme. **D**: Postulated transition state for the C-H activation step. Critical interactions made by the aspartate side chain are depicted in red (adapted from reference [141]).

In this study the introduction of a carboxylate side chain at the active site speeds up the coupling of benzamides and alkenes by nearly 100-fold. Such bifunctional artificial metalloenzymes resemble much more their natural counterparts in which catalytic tasks are routinely shared between the reactive metal center and the surrounding protein environment. In order to further exploit the potential of artificial metalloenzymes, expanding the limited reaction scope of these hybrid catalysts is an important task. In this chapter,

we evaluate olefin metathesis, an important transformation for the total synthesis of organic compounds, as a new type of reaction that can be catalyzed by artificial metalloenzymes.

2.1.2. Olefin metathesis

The discovery of the olefin metathesis reaction, through which pairs of carbon-carbon double bonds are reorganized via two metallocyclobutane intermediates (**Figure 16A**), has led to fundamentally new strategies for natural product synthesis and polymer chemistry.[106, 142-143] Olefin metathesis can be classified into three categories (**Figure 16B**). Cross metathesis reactions join a pair of alkene substrates; in ring opening metathesis a cyclic olefin reacts with a linear one to generate an acyclic diene; and ring closing metathesis yields a cyclic olefin from a substrate with two terminal alkenes. Catalysts used for these transformations are either molybdenum (**1**) or ruthenium (**2**) based (**Figure 16C**). While the former, pioneered by Schrock, usually displays higher turnover numbers, the latter, introduced by Grubbs and Hoveyda, shows a broader functional group tolerance.

Besides the extensive use of olefin metathesis in organic synthesis, cross metathesis reactions are also an emerging tool for chemical biology, providing a powerful bio-orthogonal method for site-selective protein modification.[144-145] Although Grubbs-Hoveyda-type catalysts (**2**, **Figure 16C**) demonstrate some tolerance for the gamut of functional groups present in proteins, a large excess of catalyst over protein is typically required for such applications. We therefore wondered whether true multi-turnover catalysis could be achieved in a proteinaceous environment and set out to create an

artificial metalloenzyme by covalently tethering a suitable ruthenium complex to a protein scaffold. Such a strategy is complementary to approaches that take advantage of the affinity of (strept)avidin for a biotinylated guest to anchor a metal complex near a protein binding site.[146] Covalent modification is applicable to nearly any protein and therefore enables a systematic exploration of protein ligand interactions and their effects on catalysis. Here, we use a Grubbs-Hoveyda type catalyst, equipped with an electrophilic bromoacetamide group, to modify a cysteine-containing variant of a small heat shock protein and as a result, create an artificial metalloenzyme for olefin metathesis.

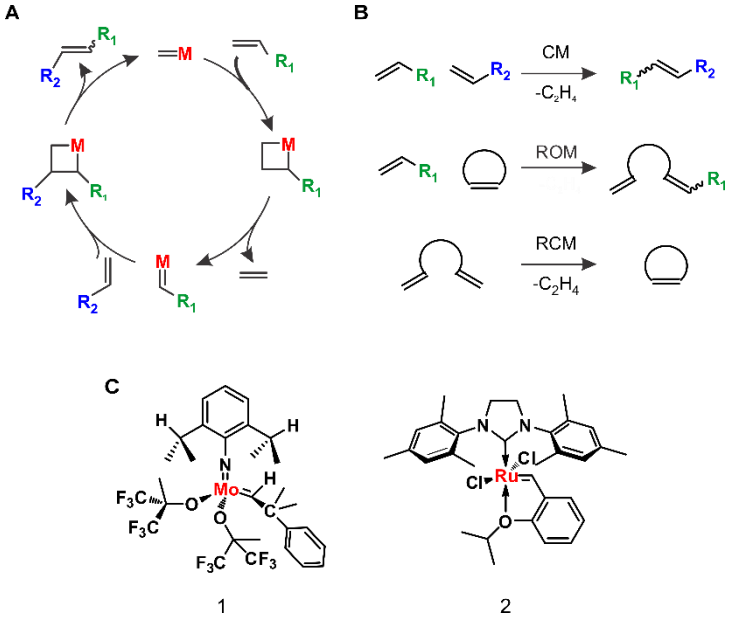


Figure 16: First principles of olefin metathesis. **A**: The catalytic cycle proceeds via two metallocyclobutane intermediates that facilitate the scrambling of the double bonds of one or two substrate(s). **B**: Reaction scope of olefin metathesis from top to bottom: Cross metathesis (CM), ring opening metathesis (ROM), and ring closing metathesis (RCM). **C**: Frequently used catalysts for olefin metathesis. A molybdenum-based Schrock-type catalyst (**1**) on the left and a Grubbs-Hoveyda-type ruthenium catalyst (**2**) on the right.

2.2. Results

A number of water-soluble ruthenium-based metathesis catalysts have been recently reported in the literature.[147-149] They typically contain an N-heterocyclic carbene (NHC) ligand, which remains associated to the metal throughout the catalytic cycle and hence can be used to link the catalyst permanently to a protein. An NHC equipped with an electrophilic bromoacetamide, for instance, would allow site-selective alkylation of cysteine thiols. To prepare such a catalyst (i.e., **5**), we modified ruthenium complex **4**, which contains a pendant amino group and is readily prepared from NHC ligand **3** by slight modification of a published route (**Figure 17A,** see Materials and Methods for details).[148] Compound **4** was acylated with bromoacetic acid using benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBop) as a coupling agent and diisopropylethylamine as base. Although deprotonation of the ammonium group of **4** causes considerable decomposition, the desired complex **5** was isolated in 29% yield after silica gel chromatography.

As a robust scaffold for attachment of the metathesis catalyst we chose the *Methanocaldococcus jannaschii* small heat shock protein 16.5 (MjHSP). It forms a 24-subunit spherical capsid with 4:3:2 symmetry; the assembled structure is 12 nm in diameter and has large 3 nm pores, allowing for free exchange of small molecules between the interior and bulk solution.[150-151] The capsid is stable up to 70 °C over a pH range of 4 to 11. At higher temperatures or lower pH, the protein capsid dissociates into smaller oligomeric structures that are thought to be the active species in cell protection.[152-153] Nevertheless, the MjHSP

subunits remain folded, even at pH 2.2 (**Figure 22A** in the Materials and Methods section).

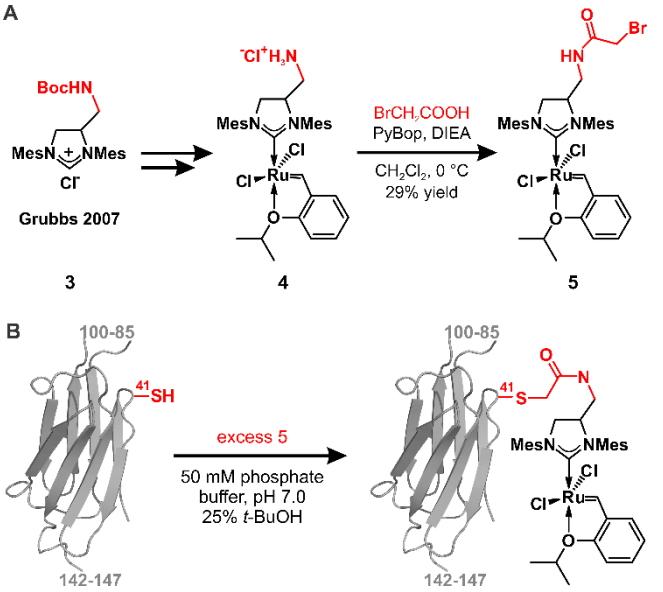


Figure 17: Preparation of an artificial metalloenzyme for olefin metathesis. **A**: Synthesis of the ruthenium complex **5** bearing an electrophilic motif for covalent protein modification. **B**: Modification reaction of G41C MjHSP (monomer shown). A loop region that participates in subunit interactions (residues 85-100) and the C-terminus (residues 142-147) are omitted from the graphic for simplicity; neither is expected to interact significantly with the appended ruthenium complex.

Modification experiments were carried out with the G41C variant of MjHSP, which contains a single reactive cysteine that was previously employed for selective attachment of small molecules to the interior surface of the cage structure.[154-155] The protein was treated with a three-fold excess of **5** in degassed potassium phosphate buffer (50 mM, pH 7) containing 25% *t*-butanol as a cosolvent (**Figure 17B**). The extent of reaction was conveniently monitored by titration of free thiols with 5,5'-dithiobis-(2-nitrobenzoic

acid).[156] After 4 hours at room temperature, more than 95% of the cysteines had reacted. Excess ruthenium catalyst was removed by ultrafiltration and the sample was exhaustively dialyzed against potassium phosphate buffer at pH 2.2. Reverse-phase HPLC analysis of the alkylated product, [Ru]MjHSP, confirmed that the dialyzed sample contained less than 2% free **5** (**Figure 22B**). The modified capsid eluted largely as a single peak from a Phenomenex BioSep size exclusion column, contaminated by a small amount of higher order aggregates (**Figure 18A**); it has an absorption band at 395 nm, consistent with covalent addition of the ruthenium complex to the protein. Moreover, its ESI-MS spectrum shows a peak at 17,121.0 Da (**Figure 18B**), which corresponds to the mass expected for [Ru]MjHSP after loss of two chloride ligands from the complex. It was not possible to resolve whether fragmentation of the complex occurred during the ionization process or results from prior ligand exchange with a coordinating group on the protein.

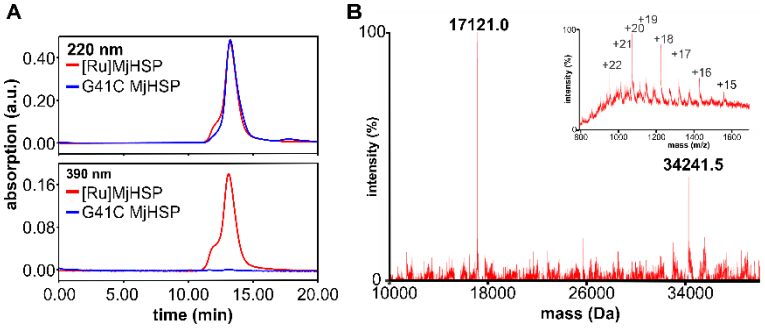
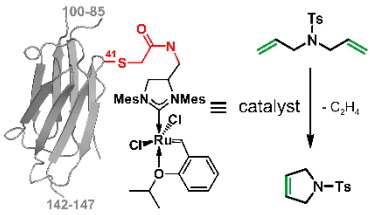


Figure 18: Biophysical characterization of [Ru]MjHSP. **A**: Comparison of SEC-HPLC chromatograms of G41C MjHSP (blue traces) and [Ru]MjHSP (red traces) at 220 nm and 390 nm. [Ru]MjHSP elutes in 20 mM MOPS buffer, pH 7.0 predominantly as a capsid with a leading peak corresponding to an uncharacterized aggregate. **B**: ESI-MS analysis of [Ru]MjHSP. The raw data are shown in the insert. Mass calculated [Ru]MjHSP -HCl_2: 17,122.8 Da, mass found 17,121.0 Da. A dimer peak is observed at 32,241.5 Da (calc. 34245.6 Da).

A



B

entry	catalyst	loading	pH	buffer	TON
1	**5**	2%	7.0	50 mM	2 ± 0.2
2	**5**	2%	3.9	50 mM MES	16 ±0.4
3	**5**	2%	2.0	10 mM HCl	33 ± 0.5
4	**[Ru]MjHSP**	4%	7.0	50 mM	3 ± 0.1
5	**[Ru]MjHSP**	4%	3.9	50 mM MES	12 ± 1.5
6	**[Ru]MjHSP**	4%	2.0	10 mM HCl	25 ± 2.1

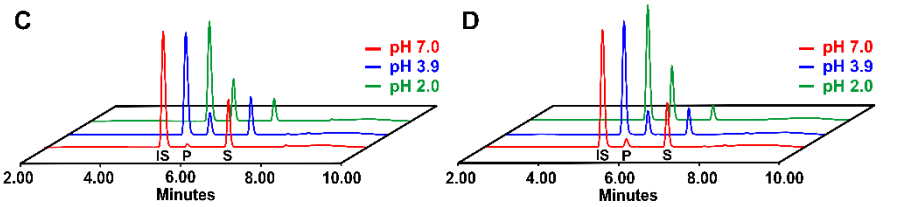


Figure 19: Kinetic characterization of [Ru]MjHSP and **5. A**: Model RCM reaction, used in activity assays for the artificial metalloenzyme (monitored by HPLC). **B**: Table for the RCM of N,N-diallyl-4-toluenesulfonamide. Reactions were carried out at 5 mM substrate at 45 °C for 12 hrs. Twenty percent *t*-butanol was included in the reactions to ensure catalyst solubility; control experiments with and without this co-solvent confirmed that it does not have an adverse effect on reaction efficiency. Turnover number (TON) was determined as an average of two separate experiments, both carried out in triplicate. **C**: RP-HPLC chromatograms for activity assays with **5**. Reactions were carried out as described in **B**. *p*-Bromophenol (IS), N-(*p*-toluenesulfonyl)-3-pyrroline (P), and N-diallyl-4-toluenesulfonamide (S). **D**: Activity assays with [Ru]MjHSP, also performed as described in **B**.

The catalytic properties of [Ru]MjHSP were examined using the model ring closing metathesis reaction (RCM) of N,N-diallyl-4-toluenesulfonamide to give N-(p-toluenesulfonyl)-3-pyrroline in aqueous solution (**Figure 19A**). As a control, analogous reactions were carried out in parallel with the free ruthenium complex **5**. In both cases, reactions were monitored by RP-HPLC (**Figure 19 C and D**). At neutral pH, neither catalyst afforded significant amounts of product with catalyst loadings of 2% and 4%, respectively

(**Figure 19B**, entries 1 and 4). This finding is in line with previous observations that coordinating solvents and Lewis basic groups adversely affect metathesis reactions.[157-159] Since activation can often be achieved by simple protonation of the coordinating species, we lowered the pH of the reaction medium. As summarized in **Figure 19B**, the turnover number (TON) of both catalysts increased significantly and to similar extents as pH was decreased. The highest yields were observed at pH 2, with **5** and [Ru]MjHSP affording 33 ± 1 and 25 ± 2 turnovers, respectively. These results show that the hybrid enzyme functions as a true catalyst for olefin metathesis. Nevertheless, the protein structure itself exerts little influence on the reaction, presumably because the ruthenium complex is relatively exposed on the surface of the scaffold rather than embedded in a pocket that can also accommodate substrate. However, if the quaternary capsid structure could be stabilized at low pH by mutagenesis, ring opening metathesis polymerization reactions could conceivably be used to fill the luminal space and generate monodisperse polymer particles.

2.3. Discussion

In summary, these experiments establish the feasibility of performing multi-turnover RCM catalysis with a protein-bound ruthenium complex. Although the protein provides water solubility and prevents aggregation, these hybrid catalysts are not yet practical. Exploring a wider range of protein scaffolds and identifying conditions that allow for metathesis activity under neutral pH will be crucial for fine-tuning of reactivity and selectivity of these artificial metalloenzymes.

Since our proof-of-concept study, covalent modification of chymotrypsin and the β-barrel protein FhuA with Grubbs-Hoveyda-type catalysts gave rise to artificial metalloenzymes with metathesis activity.[160-161] Notably, both hybrid catalysts remain active under neutral conditions when high concentrations of KCl are present in the buffer. Presumably, the addition of large amounts of chloride ions prevents the addition of water (or hydroxide ions) to the metal center; a mechanism that has been demonstrated to reduce the activity of metathesis catalysts.[162] Under these conditions the artificial metalloenzyme derived from FhuA gave rise to 375 turnovers for the ring-opening polymerization of a water-soluble substrate.[161] If these activities could be reproduced in the MjHSP capsid, it is conceivable that the luminal space of the protein cage could be filled with this polymer and as a result, create particles of well-defined size and shape.

Despite the higher activities that were identified for the FhuA metathesis enzyme, it will be crucial to devise evolutionary approaches to boost the performance of these artificial metalloenzymes to levels that render them practical for synthetic purposes.[101] The broad substrate and reaction scope of metathesis catalysts are an ideal starting point for such developments. For example, cross metathesis reactions could be employed to append fluorescent molecules to a substrate and facilitate a screening system for cross metathesis enzymes. Moreover, because Grubbs-Hoveyda-type catalysts display a reasonable tolerance to the gamut of functional groups present within a cell, an evolutionary approach in vivo becomes conceivable. Selection systems in which central metabolites necessary for survival of a cell are synthesized by the hybrid metathesis enzyme would presumably be a powerful means to rapidly

generate artificial metalloenzymes that outperform their small-molecule counterparts.

2.4. Materials and Methods

2.4.1. General procedures

All reactions were conducted in oven-dried (135 °C) or flame-dried glassware under an inert atmosphere of dry N_2 unless otherwise stated. All solvents and buffers were degassed prior to use. NMR spectra were recorded on Bruker AV600 (^{1}H 600 MHz, ^{13}C 150.9 MHz), ARX 300 (^{1}H 300 MHz, ^{13}C 75 MHz), or Varian Gemini 300 (^{1}H 300 MHz, ^{13}C 75 MHz) NMR spectrometers. All ^{13}C-NMR spectra are ^{1}H-broadband decoupled and were measured at room temperature; the ^{13}C resonance of chloroform was used as an internal standard ($^{13}CDCl_3$: 77.16). Chemical shifts are reported in ppm from tetramethylsilane with the solvent resonance resulting from incomplete deuteration as the internal standard ($CDCl_3$: 7.26). Data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, br = broad, m = multiplet or combinations thereof), coupling constants, and integration. Multiplicity was assigned according to a method described by Hoye.[146] Compounds **3**, **6** (**Figure 20**) and N,N-diallyl-4-toluenesulfonamide were synthesized as previously described in the literature.[163],[164] For anaerobic silica gel chromatography, columns were first purged with N_2, and all eluents degassed with a generous N_2 sparge (at least 30 min). Product was then eluted under N_2 and collected in a round-bottom flask already purged with N_2. Solvents were then removed in vacuo (not by rotary evaporation). Analytical

RP-HPLC runs were performed with a Luna C18 3 µm 50 x 2.00 mm using a linear gradient from 5% to 60% ACN containing 0.08% TFA in 0.1% aqueous TFA for 6 min with a flow rate of 700 µL/min. SEC-HPLC were carried out with a Phenomenex BioSep-SEC-S 4000 300 x 7.8 mm, employing an isocratic flow of 20 mM 3-(N-morpholino)propanesulfonic acid (MOPS) buffer pH 7.0 with a flow rate of 1 ml/min.

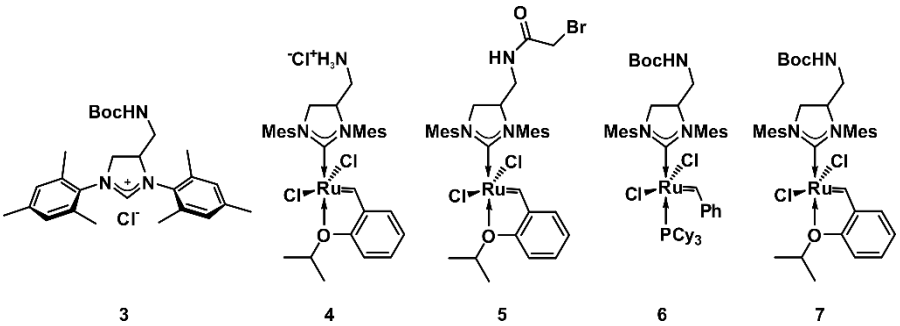


Figure 20: NHC ligand 3 and ruthenium complexes 4 - 7 used in this study.

2.4.2. Synthesis of 4

A flame dried round-bottom flask equipped with a magnetic stir-bar, a septum and purged with N_2 was charged with **6**[163] (138 mg, 0.141 mmol), 2-isopropoxystyrene (34.2 mg, 0.211 mmol) and CuCl (27.9 mg, 0.282 mmol). The seal of the flask was reinforced with Teflon tape and dry degassed dichloromethane (5 ml) was added. The reaction mixture was heated to 40 °C for 90 min. In the course of the reaction the magenta suspension turned dark green. Conversion was followed by TLC (15% EtOAc in hexane), monitoring the disappearance of **6**. After complete conversion, the reaction mixture was concentrated to 1 ml and loaded on an anaerobic silica gel column (elution with

15% EtOAc in hexane). The dark green band was collected and the solvent was removed in vacuo to obtain the Boc protected ruthenium complex **7** as a dark green solid. This material was subsequently dissolved in 5 ml degassed methanol and hydrochloride acid (37% in water, 2 ml) was added dropwise. The reaction mixture was stirred for 30 min at room temperature and after complete deprotection of the Boc-group (judged by TLC) the solvent was removed in vacuo. The green residue was taken up in 5 ml degassed water and frozen. Lyophilization yielded ruthenium complex **4** as a green solid (65 mg, 0.086 mmol, 61% yield over 2 steps). An NMR sample of **4** was prepared under N_2 in degassed d_4-methanol.

^{1}H-NMR (300 MHz, CD_3OD): δ 16.46 (s, 1H), 7.59 (t, J= 7.5 Hz, 1H), 7.13 (bs, 4H), 7.00 (d, J= 8.4 Hz, 1H), 6.88 (m, 2H), 4.95 (m, 1H), 4.77 (m, 1H), 4.46 (t, J= 10.8 Hz, 1H), 4.13 (t, J= 6.6 Hz, 1H), 3.21 (m, 1H), 2.70 – 2.28 (bd, 18H), 1.21 (d, J= 5.7 Hz, 6H); **^{13}C-NMR (75 MHz, CD_3OD):** 299.3, 217.1, 153.4, 146.3, 140.7, 140.4, 140.1, 131.2, 130.6, 130.5, 130.3, 123.2, 122.9, 114.3, 76.6, 62.7, 57.9, 43.5, 21.9, 21.5, 20.4. **HRMS (MALDI, -HCl):** calc. 655.1660, found 655.1659

2.4.3. Synthesis of 5

A dry round-bottom flask equipped with a magnetic stir-bar and purged with N_2 was charged with PyBOP (108 mg, 0.29 mmol), bromoacetic acid (40.2 mg, 0.29 mmol) and 5 ml dry degassed dichloromethane. Hünig's base (55 µL, 41.1 mg, 0.318 mmol) was added at 0 °C and the reaction mixture was allowed to stir for 10 min to form the activated ester. A solution of **4** (20 mg, 0.029 mmol) in 2 ml dry degassed dichloromethane was added

dropwise over a time period of 20 min with stirring. The reaction was allowed to warm to room temperature and was stirred for a further 15 min. The resulting dark green solution was concentrated to 1 ml by evaporating the solvent under a constant N_2 flow. The dark green mixture was subjected to anaerobic silica-gel chromatography (elution with 20% EtOAc in hexane) and the green band was collected. The solvent was removed in vacuo to afford **5** (6.4 mg, 8.2 µmol, 29% yield) as a green solid. An NMR sample of **5** was prepared under N_2 in degassed chloroform.

^{1}H NMR (600 MHz, $CDCl_3$): 7.51 – 7.47 (m, 1H), 7.14 – 7.04 (m, 4H), 6.95 – 6.81 (m, 3H), 6.81 (d, J = 8.3 Hz, 1H), 4.97 – 4.89 (m, 1H), 4.60 (m, 1H), 4.29 (t, J = 10.9 Hz, 2H), 3.94 – 3.88 (m, 2H), 3.81 (m, 2H), 3.62 (bs, 3H), 2.91 – 2.28 (m, 18H), 1.40 – 1.27 (m, 6H); **^{13}C NMR (151 MHz, $CDCl_3$):** δ 298.3, 216.3, 190.2, 152.4, 145.2, 139.4, 139.2, 135.8, 130.6, 130.1, 130.0, 129.7, 128.3, 122.9, 122.4, 120.4, 114.0, 112.9, 75.1, 63.2, 55.6, 42.7, 28.7, 22.0, 21.2, 21.1; **HRMS (MALDI, MH^+):** calc. 777.0864, found 777.0865

2.4.4. Molecular cloning of G41C MjHSP

The gene encoding the small heat shock protein 16.5 (MjHSP) was amplified from *M. jannaschii* genomic DNA with the following primers (restriction sites for NdeI and SpeI underlined): forward: 5-GGA GAT ATA CAT ATG TTC GGA AGA GAC CC-3'; reverse: 5'-CAG CTG ACT AGT CAT TAT TAT TCA ATG TTG ATT CCT TTC-3'. Polymerase chain reaction (PCR)-mediated site directed mutagenesis was employed to introduce the desired G41C muation into MjHSP as described earlier.[165] The resulting insert encoding G41C MjHSP was cloned into the SpeI / NdeI restriction site of the pMG211 vector,[166] yielding the

pMG211-G41C-MjHSP plasmid. After transformation into XL1 Blue *E. coli*, cells were grown on LB_{Amp} (containing 150 µg/mL ampicillin) plates overnight at 37 °C, plasmid DNA isolated (Eppendorf Mini Prep Kit), and the DNA insert sequenced. After sequence verification, the pMG211-G41C-MjHSP was subcloned into BL21(DE3) *E. coli* for protein production.

2.4.5. Protein production and purification

The G41C MjHSP gene was expressed from plasmid pMG211-G41C-MjHSP in BL21(DE3) *E. coli* in LB_{Amp}. A densely grown overnight culture (750 µL) was used to inoculate 750 mL LB_{Amp}. After incubating at 37 °C and 250 rpm for 4 hrs an OD of 0.6 was reached and protein production was induced with IPTG (250 µM final concentration). After 6 hours of induction, cells were harvested by centrifugation at 4000 g at 4 °C for 5 min. The cell pellet was resuspended in ice-cold cell lysis buffer (20 ml, 50 mM HEPES, pH 8.0) and subsequently lysed by sonication (0.5 s cycles, amplitude 60, 5 x 1 min with a 1 min break between cycles). Cell debris was removed by centrifugation (14,000 g at 4 °C for 30 min) and the supernatant was heated for 15 min to 60 °C, thereby denaturing many heat labile *E. coli* proteins. The suspension was centrifuged at 14,000 g at 4 °C for 30 min and tris(2-carboxyethyl)phosphine (TCEP, final concentration 1 mM) was added to the supernatant to fully reduce G41C MjHSP. The crude protein solution was further purified by anion exchange chromatography (MonoQ, Biorad Duoflow, 20 mM potassium phosphate buffer, pH 7.0 with a linear gradient from 0 to 1 M NaCl). The purity of G41C MjHSP was estimated by HPLC to be >95 % and the mass was confirmed by LC/MS analysis (found 16,498 Da, calc. 16,498 Da, **Figure 21**). The yield of G41C MjHSP was determined to be 25 mg/L culture.

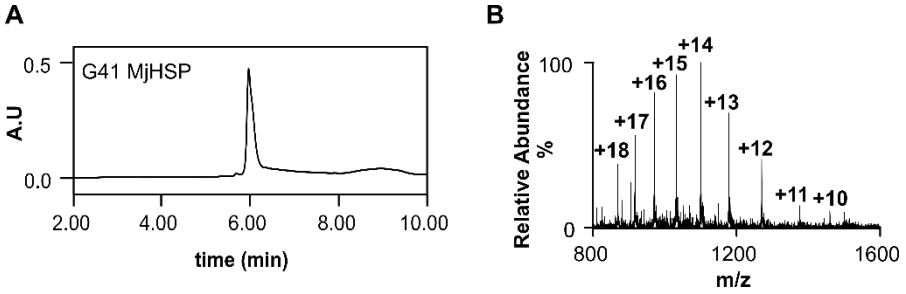


Figure 21: Characterization of G41 MjHSP. **A**: HPLC trace of the purified heat shock protein. **B**: Obtained mass spectrum for G41 MjHSP. The raw data is depicted. Deconvolution of the obtained spectrum yields the expected mass of G41 MjHSP (16,498 Da).

2.4.6. Alkylation of G41C MjHSP

All steps were carried out in degassed buffers under an N_2 atmosphere. 400 µL *t*-BuOH was added to 3 ml of freshly purified G41C MjHSP (3 mg/mL, 182 µM, in 50 mM potassium phosphate buffer pH 7.0). The reaction mixture was stirred for 10 min at room temperature prior to dropwise addition of a solution of **5** (1.27 mg, 3.0 eq.) in 600 µL *t*-BuOH. After complete addition, the green solution was allowed to stand for another 4 hours at room temperature. Reaction progress was monitored by titrating free thiols in the reaction mixture with 5,5'-dithiobis-(2-nitrobenzoic acid).[156] After more than 95% of cysteine thiols had reacted, 10 mL of nanopure water was added and the protein mixture was concentrated employing Amicon Ultra-15 centrifugal filters (10 kDa cutoff). Two further cycles of dilution with nanopure water followed by concentration were performed until a 2 mL [Ru]MjHSP solution was obtained. Subsequently, 13 mL of 50 mM potassium phosphate buffer (pH 2.2) was added and the reaction mixture was allowed to stand for 15 min after mixing on ice. Amicon Ultra-15 Centrifugal filters (3 kDa cutoff) were used to concentrate the

protein solution to 3 mL, followed by another addition of 12 mL potassium phosphate buffer pH 2.2 and concentrating to 3 mL. The green solution was transferred to a Slide-A-Lyzer dialysis chamber (7 kDa cutoff) and dialyzed against 50 mM potassium phosphate buffer pH 2.2 overnight to remove loosely associated **5** from the protein solution. After dialysis an analytical sample (100 µL) was removed for analytical HPLC, SEC-HPLC, and ESI-MS analyses (**Figure 18 and 22A**). Nanopure water (12 mL) was added to the remaining green solution and Amicron Ultra-15 Centrifugal filters (3 kDa cutoff) were used to concentrate the [Ru]MjHSP solution to 1.5 mL. The protein concentration of the solution was determined by Bradford assay in which a G41C MjHSP stock solution of known concentration (determined by UV, ε = 9,322 $M^{-1}cm^{-1}$)[165] was used for the calibration. The concentration of the ready-to-use [Ru]MjHSP solution was typically between 3.5 and 4.5 mg/mL.

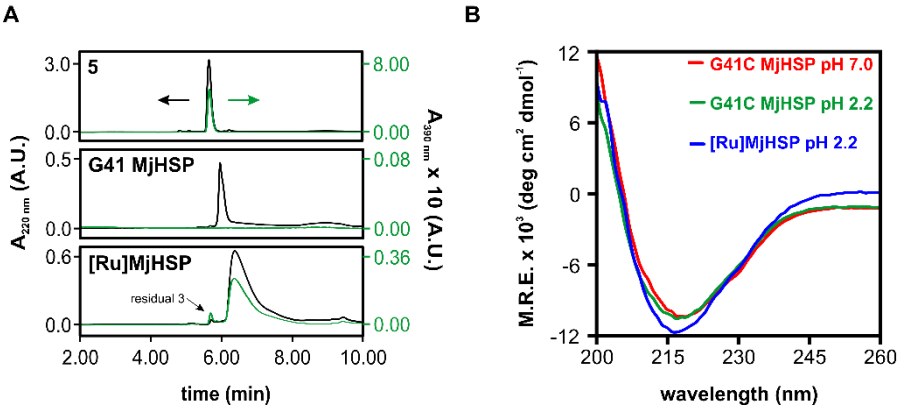


Figure 22: Biophysical characterization of [Ru]MjHSP. **A**: Comparison of analytical HPLC runs of G41C MjHSP, 3 and [Ru]MjHSP at 220 nm (black traces) and 370 nm (green traces). The amount of residual catalyst (integration of peaks at 370 nm) was found to be < 2 %.**B**: Comparison of CD spectra of G41 MjHSP and [Ru]MjHSP at 10 µM in 10 mM potassium phosphate buffer pH 7.0 and 2.2, respectively.

2.4.7. CD spectroscopy

Far UV spectra of G41C MjHSP and [Ru]MjHSP were recorded at 25 °C at an Aviv 202 spectropolarometer from 200 to 260 nm in 1 nm steps with an averaging time of 3 seconds at 10 μM protein concentrations and in 10 mM potassium phosphate buffer (pH 7.0 and 2.2). Five scans were averaged and a buffer blank was subtracted (**Figure 22B)**

2.4.8. Representative activity assay with [Ru]MjHSP

A freshly prepared [Ru]MjHSP solution (100 μL, 4.29 mg/mL, 260 μM) was mixed with 20 μL of 325 mM buffer stock or 65 mM HCl solution. After mixing, 10 μL of a N,N-diallyl-4-toluenesulfonamide solution (65 mM in *t*-BuOH) was added and the reaction vessels were incubated for 12 hrs at 45 °C. After the reaction was complete 20 μL *p*-bromophenol (internal standard, 65 mM in *t*-BuOH) was added and the reaction mixture extracted 3 times with 700 μL diethyl ether. The solvent was removed in vacuo and the residue was redissolved in 520 μL acetonitrile by vortexing. A 10 μL aliquot of the resulting solution was analyzed by HPLC, and conversion was determined according to a previously performed calibration (N-(p-toluenesulfonyl)-3-pyrroline, P, was obtained from ABCR-Chemicals). Turnover numbers (TONs) in **Figure 19B** represent averages from two separate experiments, each carried out in triplicate.

3. A genetically-encodable ligand for transfer hydrogenation

3.1. Introduction

3.1.1. Metallopeptides – from catalysis to genetically-encodable ligands

Metalloenzymes are a benchmark for efficient, enantioselective transformations and, as such, a constant source of inspiration for the design of new catalysts.[81] The previous chapter highlighted current efforts to mimic metalloenzymes by introducing transition-metal catalysts into a protein scaffold and thereby create artificial metalloenzymes with non-natural activities.[110] In principle, placing an artificial cofactor within a proteinogenic environment allows for additional interactions with amino acid side chains that can steer the reactivity and selectivity of the transition-metal complex. Nevertheless, such hybrid catalysts are typically much less effective than natural metalloenzymes with respect to turnover number and selectivity. These shortcomings can largely be attributed to trade-offs made in the design process.[98, 109] The organometallic complex is usually attached to the scaffold protein via a flexible linker that is equipped either with a reactive chemical group for site-selective protein modification (covalent anchoring) or a small ligand like biotin with unusually high affinity for a specific protein pocket (supramolecular anchoring). The flexibility of the linker may limit accurate

placement of the metal center in a defined environment, hampering effective preorganization of substrate and catalyst. Unlike natural metalloenzymes, residues at the active site of these hybrid constructs rarely interact productively with the cofactor (recent work on C-H activation represents a notable exception[141]), making fine-tuning of metal properties difficult. As a consequence, the unbound metal complex usually displays higher activity than its bound form. Moreover, optimization of the initial activities of artificial metalloenzymes by directed evolution has proven to be a formidable challenge.[101] Because small-molecules, proteins, and nucleic acids from the cell can competitively chelate the transition-metal complex, screening approaches typically rely on purified proteins.[167] Thus, only a small number of variants and a marginal fraction of available sequence space can be tested in each evolution round for improved variants.

Since controlling the proteinogenic environment proximal to the transition metal complex in artificial metalloenzymes has proven difficult, chemists have begun to evaluate simpler, metal-binding peptides for catalysis.[99] Both the peptidic backbone and amino acid side chains are useful for coordinating metal ions and libraries of potential ligands are readily available by parallel automated peptide synthesis. Moreover, fluorescent metallopeptides, including arsenic- or lanthanide-binding peptides, have already been identified by combinatorial synthesis and were subsequently appended to protein sequences.[168-169] The high selectivity of these fluorescent tags facilitates the study of protein function in vivo without observing competitive chelation from cellular material. In the same fashion, it is conceivable that catalytic metallopeptides with a high selectivity for a given transition metal can be incorporated into proteins.[170] Ultimately, artificial metalloenzymes that result

from such a bottom-up strategy could be more malleable in directed evolution experiments.

3.1.2. Catalytic metallopeptides from non-natural amino acids

Peptides are an appealing choice in the quest for chiral ligands that in combination with (transition-)metals efficiently catalyze asymmetric reactions of value to synthesis. Chiral natural and non-natural amino acids are readily available building blocks that can be joined by synthesis in a combinatorial fashion.[171] Moreover, peptides containing more than 10 building blocks or constructs that are conformationally constrained by cyclization, disulfide formation, or other means readily adopt stable conformations in solution that can further boost reaction rates and selectivities.

Initial efforts in the combinatorial synthesis of metal-binding ligands took advantage of non-proteinogenic, chelating amino acids or introduced specific capping units to boost metal coordination. Francis and Jacobsen, for example, pursued an on-bead screening approach to identify peptide-based ligands for a variety of transition metals from an initial combinatorial, metal-binding library.[172] Catalytically active compounds were further optimized to yield epoxidation catalysts that displayed modest stereoselectivities (20% e.e, **Figure 23A**). By investigating peptide-based Schiff bases with different O, N, or P donors in the presence of several transition states, the Hoveyda group identified stereoselective catalysts for a series of C-C bond forming reactions. Notable examples include the Ti^{IV}-catalyzed addition of cyanide to epoxides and imines[173-174], a dialkyl zinc addition to imines facilitated by Zr^{IV},[175] or copper-

catalyzed allylic substitution reactions (**Figure 23B**).[176] For all transformations, highly specific ligands could be identified by fine-tuning the individual components (metal source, peptidic ligand, Schiff base used, reaction conditions, etc.) of the initially obtained catalytic system. In another early example Gilbertson et al. utilized β-turn forming peptides harboring a phosphine ligand to enable palladium-catalyzed, enantioselective alkylations (e.e. 85%, **Figure 23C**).[177] An important aspect of this system was the selection for peptide-based phosphine ligands that showed high substrate specificity, a property typically associated with enzyme catalysis.

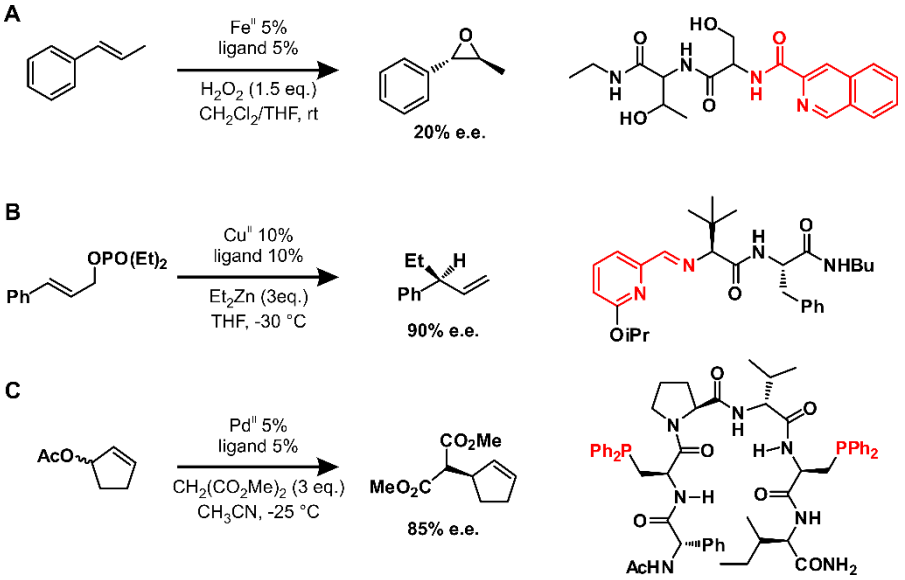


Figure 23: Examples of reactions catalyzed by modified peptides: **A**: Fe^{2+}-catalyzed cyclopropanation reaction with the peptidic ligand harboring a non-natural C-cap (red). **B**: Cu^{2+}-catalyzed allylic substitution with a peptide featuring a Schiff base at the N-terminus (red). **C**: Pd^{2+}-catalyzed allylic substitution with a β-hairpin forming peptide that is modified with two phosphine ligands (red).

Despite their relative simple structures, metallopeptides typically catalyze transformations with efficiencies and selectivities that are comparable to most artificial metalloenzymes. Nevertheless, all of these transformations rely on heavily modified amino acids and only function in organic solvents. As a consequence, transplantation of these catalytic motifs into a more sophisticated proteinogenic environment is almost impossible to achieve.

Peptide synthesis is a highly efficient production process and has become a valuable tool for protein synthesis.[178] Compared to protein production in vivo, solid phase peptide synthesis provides a straightforward means to introduce non-natural building blocks into a given scaffold. For the synthesis of catalytic metallopeptides – or rather artificial metalloenzymes – amino acids equipped with pyridine, bipyridine or phenanthroline side chains are of particular interest.[99] These non-canonical amino acids efficiently bind to transition metals such as copper or osmium and have been demonstrated to be effective catalysts in aqueous solution. Indeed, introducing these amino acids into protein scaffolds via solid phase peptide synthesis affords active hybrid catalysts. Recently, incorporation of a copper chelating phenantroline side chain into the hydrophobic dimeric interface of the transcription factor Lactococcal multidrug resistance regulator yielded an enantioselective artificial metalloenzyme that catalyzes a bimolecular Diels-Alder reaction as well as the hydration of a conjugated alkene (**Figure 24**).[179-180] The possibility of incorporating bipyridine side chains genetically into proteins via stop codon suppression should greatly extend this approach.[181] Unfortunately, this technique often suffers from premature termination in the translation process or the incorporation of a standard amino acid instead of the non-natural building block, and as a consequence, gives rise to the desired proteins in low

yields. Nevertheless, an artificial metalloenzyme was recombinantly produced by stop-codon suppression that cleaves a DNA substrate site-selectively in vitro in the presence of copper.[182]

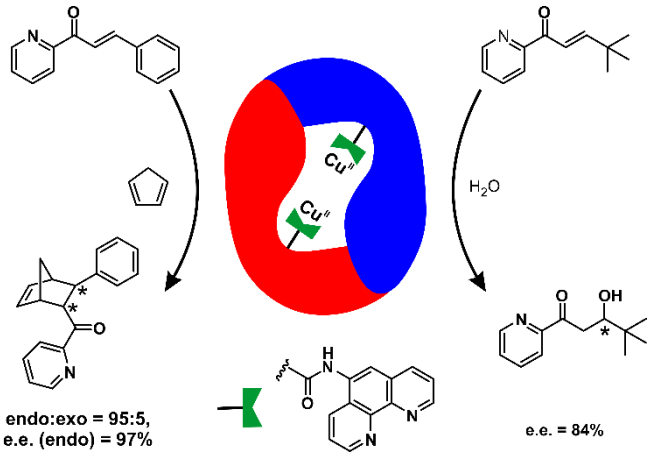


Figure 24: Schematic representation of the modified Lactococcal multidrug resistance regulator dimer. Introducing a phenanthroline side chain at the dimeric interface yields an enantioselective artificial metalloenzyme for a Diels-Alder reaction and hydration of α/β-unsaturated ketones (adapted from reference [180].

Taken together, these studies demonstrate that peptidic motifs can be incorporated into protein scaffolds to create selective artificial metalloenzymes. Nevertheless, activity largely depends on the introduction of an artificial amino acid, a fact that entails a considerable chemical investment or the low-yielding use stop-codon suppression.

3.1.3. Catalytic metallopeptides from canonical amino acids

While the advent of combinatorial peptide synthesis has allowed for the critical discovery of asymmetric metal catalysts that feature non-natural amino

acids, the ability of canonical peptides to bind transition metals has been largely overlooked.[99] This trend is surprising considering that natural metalloenzymes almost exclusively use the peptidic backbone and amino acid side chains for exact coordination of metals as well as fine-tuning of their catalytic properties. An example of a canonical catalytic metallopeptide is the small fragment Gly-Gly-His, a motif commonly found at the N-terminus of serum albumins, which binds to copper and nickel efficiently.[183] Interestingly, nickel binding was found to endow these short peptide motifs with nuclease activity with a preference for A/T rich DNA regions. When appended onto the N-terminus of a variety of DNA-binding proteins, site-selective DNA cleavage was achieved. Moreover, when all Xaa1-Xaa2-His variants were synthesized, peptide sequences were obtained that cleave DNA more efficiently than the parent peptide without losing the inherent preference for A/T rich sequences.[184]

Recently, the Ball group reported the first catalytic metallopeptide for non-natural transition metals.[185] They realized that secondary structures of short peptides can be regulated by binding of a dirhodium complex to amino acids bearing carboxylate residues (**Figure 25A**).[186] More specifically, chelating of this non-natural complex promoted α-helix formation from a random coil when glutamic or aspartic acids in the sequence were appropriately spaced (i and i+4, a binding pattern that places the carboxylates on the same face of the alpha helix). Conversely, a helical peptide harboring the same amino acids at positions i and i+7 was destabilized upon binding the transition metal complex to afford random coils. The Ball group additionally took advantage of the catalytic characteristics of the dirhodium species to confer helical peptides with a variety of different functions. When the metal complex was introduced into

natural and engineered coiled-coil structures, proximity of the rhodium complex to an amino acid of the other strand allowed for catalytic, site-selective protein modification by C-H insertion.[187] Tryptophan was the preferred substrate, although a large number of canonical amino acids could be targeted. These modification reactions have also been shown to work on natural substrates in cell lysates, an indication that medium-throughput screening approaches for dirhodium metalloenzymes could become feasible in the future. Furthermore, on-bead screening of peptide libraries has already afforded catalysts for Si-H and C-H activation.[188-189] Libraries were prepared by automated solid phase peptide synthesis and ligands were identified for the enantioselective cyclopropanation of styrene (**Figure 25B**). Notably, the reaction proceeded in the presence of small amounts of catalyst and afforded a single stereoisomer in excellent yields (96% conversion, 93% e.e.).

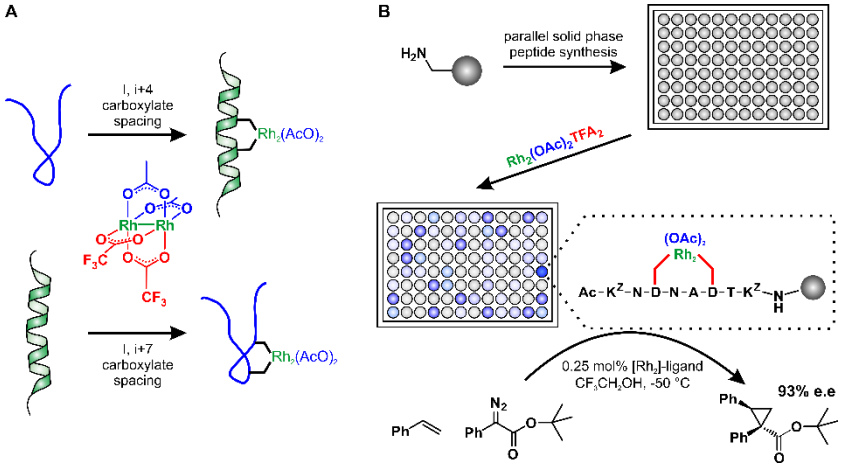


Figure 25: **A**: The secondary structure of peptides is regulated by a dirhodium complex. Whether an alpha helix is formed or destabilized depends on positioning of carboxylate side chains in the primary sequence. **B**: Schematic representation of the parallel synthesis and screening approach that afforded efficient, stereoselective catalysts for the cyclopropanation of styrene (adapted from references [185] and [189]).

Studies on dirhodium-peptide complexes have demonstrated that natural peptides can yield powerful ligands for non-natural transition metals. The inherent catalytic activity of such dirhodium complexes can be used for asymmetric catalysis, affording catalysts in which activities and selectivities are stirred by the surrounding residues. As discussed before, modular and genetically-encodable elements, such peptides should be easily embedded as tags in larger foldameric structures, potentially providing greater control over metal ion placement and reactivity.[185] However, the scope of peptide ligands that in combination with transition metals efficiently catalyze transformations is fairly limited so far. Consequently, in this work we aim to identify novel ligands from the set of canonical amino acids that allow to further evaluate a bottom-up strategy for the design of artificial metalloenzymes.

3.1.4. Transfer hydrogenation

Transfer hydrogenation reactions are important synthetic transformations in which unsaturated substrates are reduced by hydrogen donors other than molecular hydrogen.[190-191] They are attractive reactions for exploring the previously outlined strategy for the design of artificial metalloenzymes, because transfer hydrogenation reactions proceed in water under mild conditions. In a typical catalytic cycle (**Figure 26**) the sacrificial hydrogen donor, formate, displaces a water ligand from the catalyst which upon elimination of CO_2 yields a metal hydride, the reducing species. Subsequently, the hydride is transferred from the metal to an unsaturated substrate. This transfer can be accelerated by a ligand that hydrogen-bonds to the substrate in the transition state. The catalytic cycle is completed upon re-protonation of the ligand by water and addition of formate to the metal center. A myriad of

transition metal complexes with amino acids, proline derivatives, and pseudopeptides, have been reported to promote transfer hydrogenation in water.[192] In this work, we show that short peptide sequences also function as versatile ligands for iridium-based transfer hydrogenation catalysts, providing large ligand-induced accelerations. The biocompatibility of such complexes could be advantageous for a range of biotechnological applications.

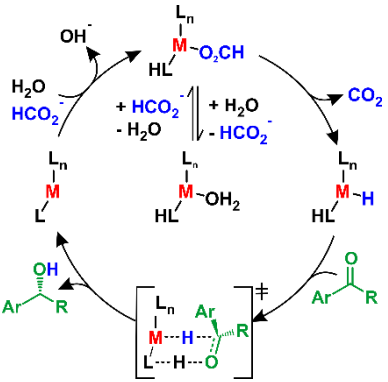


Figure 26: Catalytic cycle of transfer hydrogenation with formate as sacrificial hydrogen donor (blue) and a bifunctional catalyst. The metal is shown in red and the ketone substrate in green.

3.2. Results

3.2.1. Model system

Despite the well-established capabilities of peptides to bind metals,[193-194] aqueous transfer hydrogenations have not been systematically studied with unmodified peptide ligands. Here, the tripeptide Gly-Gly-Phe was chosen as a simple test ligand. It contains a free amino group at its N-terminus, a

carboxylate at its C-terminus, and an aromatic chromophore to facilitate purification. Experiments were carried out in water at neutral pH and 40 °C in the presence of 50 μM catalyst, 10 mM substrate, and 50 mM sodium formate (**Figure 27**).

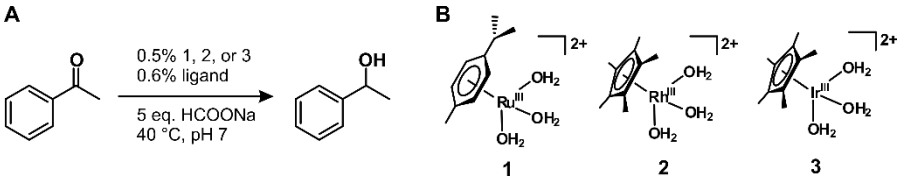


Figure 27: Model reaction (**A**) and catalysts (**B**) studied in this work. The metal is highlighted in red and the sacrificial hydrogen donor, formate, in blue.

3.2.2. Evaluation of peptide ligands

In the absence of peptide, the catalysts displayed no detectable activity within 60 minutes (**Figure 28A,** Table 1, entries 1-3,). Addition of 1.2 equivalents of Gly-Gly-Phe to ruthenium and rhodium complexes **1** and **2** marginally increased the reaction rates (**Table 1**, entries 4-5). In contrast, adding the tripeptide to iridium complex **3** gave quantitative conversions and more than 140 turnovers per hour (**Table 1**, entry 6, **Figure 28B**). Even higher turnover numbers were achieved by increasing the formate concentration. Thus, doubling the number of reducing equivalents afforded 200 turnovers per hour at pH 8 (entry 7, **Figure 33A** in the Materials and Methods section). At pH 7 and a substrate-to-catalyst ratio of 2000:1, the reduction of acetophenone proceeded to >96% conversion after 16 h, corresponding to >1900 total turnovers (**Figure 33B**). Acetylation of the N-terminus of the peptide ligand completely abrogated activity (entry 8), whereas replacement of the C-terminal carboxylate with an amide did not significantly affect catalyst performance

(**Figure 28C,** entry 9). Scrambling the tripeptide sequence also diminished catalytic efficiency. Thus, Phe-Gly-Gly exhibits 35% lower activity than Gly-Gly-Phe, whereas Gly-Phe-Gly affords basal levels only (**Figure 28D,** entries 10 and 11).

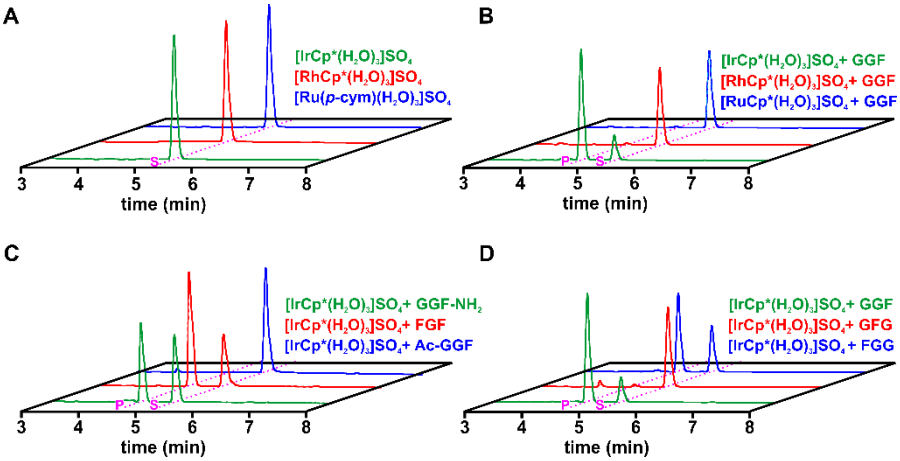


Figure 28: Representative HPLC chromatograms for reactions in this study. Retention times for substrate (acetophenone, S) and product (2-phenylethanol, P) are 5.3 and 4.7 minutes. Reaction progress in the absence (**A**) and in the presence (**B**) of Gly-Gly-Phe with 1, 2, 3 at a S/C ratio of 200:1 for 60 minutes. Reaction progress with a variety of ligands at a S/C ratio of 200:1 after 30 minutes (**C**) and 60 minutes (**D**).

3.2.3. Mechanistic considerations

Taken together, these results suggest that tripeptides may act as bifunctional Noyori-type catalysts, binding to **3** via the N-terminal amine and adjacent amide group (**Figure 29**).[190] Analogous complexes have been previously characterized by X-ray crystallography but never tested as catalysts for transfer hydrogenation.[195] The first and third amino acids are tolerant to

substitution, but steric considerations probably dictate the requirement for glycine as the second amino acid. Additional bulk at this position could hinder either formation of the peptide-iridium adduct or substrate binding. Consistent with this model, Phe-Gly-Phe still gives rise to large rate accelerations, even outperforming the original Gly-Gly-Phe ligand (**Table 1**, entry 12, **Figure 28C**). Even more active catalysts might be generated by further sequence optimization.

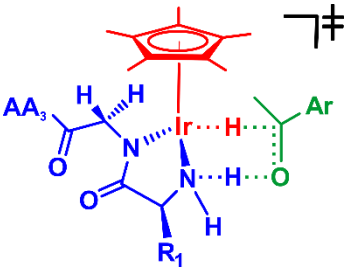


Figure 29: Proposed transition-state for the iridium-peptide-catalyzed transfer hydrogenation. The metal catalyst is shown in red, the peptide ligand in blue and the substrate in green.

3.2.4. pH dependence

Transfer hydrogenation of acetophenone is pH dependent. In the presence of Gly-Gly-Phe and 50 mM sodium formate, the pH-rate profile for the iridium-catalyzed reaction is bell-shaped with a broad maximum around pH 8 (**Figure 30**). Turnover frequency exceeds 100 h^{-1} between pH 6 and 9. As proposed by Xiao,[196] protonation of the ligand, the active metal hydride, or formate decreases activity. Above pH 9, replacement of a catalyst-bound water molecule by hydroxide prevents formate binding, thereby shutting down the catalytic cycle.

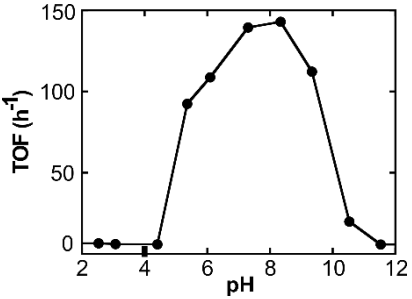


Figure 30: pH dependence of the transfer hydrogenation of acetophenone with peptide-iridium complex.

Table 1: Transfer hydrogenation of acetophenone with **1**, **2**, and **3** with peptidic ligands.

Entry	Complex	Ligand	TOF [a] (h^{-1})
1	$[Ru(p\text{-}cymene)(H_2O)_3]SO_4$	none	< 1
2	$[RhCp^*(H_2O)_3]SO_4$	none	< 1
3	$[IrCp^*(H_2O)_3]SO_4$	none	< 1
4	$[Ru(p\text{-}cymene)(H_2O)_3]SO_4$	Gly-Gly-Phe	2
5	$[RhCp^*(H_2O)_3]SO_4$	Gly-Gly-Phe	1
6	$[IrCp^*(H_2O)_3]SO_4$	Gly-Gly-Phe	142
7[b]	$[IrCp^*(H_2O)_3]SO_4$	Gly-Gly-Phe	200
8[c]	$[IrCp^*(H_2O)_3]SO_4$	Ac- Gly-Gly-Phe	< 1
9[c]	$[IrCp^*(H_2O)_3]SO_4$	Gly-Gly-Phe-NH_2	149
10	$[IrCp^*(H_2O)_3]SO_4$	Gly-Phe-Gly	3
11	$[IrCp^*(H_2O)_3]SO_4$	Phe-Gly-Gly	92
12[c]	$[IrCp^*(H_2O)_3]SO_4$	Phe-Gly-Phe	193

Acetophenone/formate/complex: 200/1000/1; 1.2 eq. of ligands used in the reaction; [a] turnover frequencies: determined after 60 minutes by RP-HPLC. [b] carried out at pH 8 with 10 equivalents of formate for 30 minutes. [c] determined after 30 min.

3.2.5. Biocompatibility

The ligand acceleration effect observed here is similar in magnitude to that of highly active ligand complexes with diamines, amino-sulfonamides, and

amino alcohols.[192, 197] Unlike many highly stereoselective transfer hydrogenation catalysts, however, the peptide-iridium catalysts generate racemic products despite the presence of chiral phenylalanine(s) in the ligand. Embedding the catalytic tripeptide motif into larger foldameric structures will therefore be necessary to control substrate access to the metal center and achieve enantioselective hydride transfer. Toward this end, the catalyst must form and retain high activity in the presence of other protein functional groups. In fact, when one equivalent of **3** is added to a mixture of the tripeptide (60 μM) and BSA (0.1 mg/mL), a catalytically active complex is formed that converts substrate to product with a TOF of 77 h^{-1} (**Figure 34**). In the absence of tripeptide, no conversion is detected. If the amount of BSA is increased to 0.5 mg/ml, the TOF drops to 11 h^{-1}, indicative of competitive chelation by the protein (**Figure 31A**).[134] Since the BSA-iridium complex is inactive, though, the concentration of the iridium species can simply be increased to overcome this problem. For example, addition of five equivalents of **3** to the mixture of tripeptide and 0.5 mg/ml BSA affords a TOF of 88 h^{-1} (**Figure 31B**). Such protein compatibility is unusual for many transition-metal catalysts and supports the feasibility of using this approach to construct artificial metalloenzymes.[198-199]

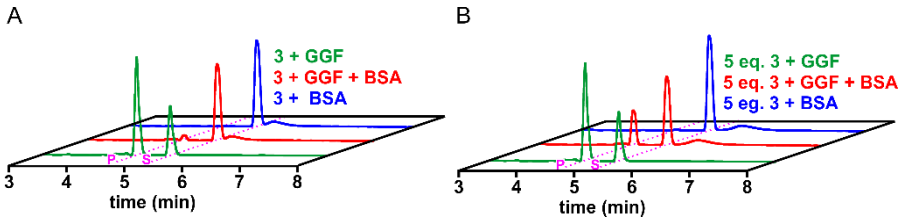


Figure 31: Comparison of the reaction progress for the reduction of acetophenone in the presence of 0.5 mg/mL BSA with 1 eq. (**A**) and 5 eq. (**B**) **3** after 30 min at a S/C ratio of 200:1 at pH 8. Substrate (S) and product (P) as in **Figure 28.**

Table 2: Turnover frequencies (TOFs) for transfer hydrogenation of different substrates catalyzed by **3** in the presence of Gly-Gly-Phe.

Entry	Substrate	TOF (h^{-1})	Rel. act.[a]
1	Acetophenone	200	1
2	4'-Chloroacetophenone	391	1.96
3	2'-Chloroacetophenone	361	1.81
4	2,2,2-Trifluoroacetophenone	366	1.83
5	4'-Methylacetophenone	198	0.99
6	4'-Methoxyacetophenone	154	0.77
7	2'-Methylacetophenone	138	0.69
8	4-Phenyl-2-butanone	98	0.49
9	Benzaldehyde	273	1.37
10[b]	1-Methyl-3,4-dihydroisoquinoline	197	0.99

Reaction conditions: Substrate/formate/catalyst = 200/2000/1, 1.2 eq. of Gly-Gly-Phe, pH 8.0, 40 °C, 600 rpm. [a] relative activity compared to acetophenone as substrate. [b] pH 6.25

3.2.6. Substrate scope

Broad substrate scope is another notable attribute of the Gly-Gly-Phe peptide-iridium catalyst. Initial turnover frequencies for a variety of aldehydes, ketones, and imines are summarized in **Table 2**. Benzaldehyde (entry 9) and activated acetophenone derivatives containing electron-withdrawing aryl substituents or a trifluoromethyl group adjacent to the carbonyl are particularly good substrates (entries 1-4), whereas derivatives containing electron-donating groups and aliphatic ketones are converted less rapidly than acetophenone itself (entries 5-8). Although transfer hydrogenation of imines is well documented,[200] the cyclic imine 1-methyl-3,4-dihydroisoquinoline proved to be a poor substrate for the peptide-iridium complex under our standard conditions (<1 h^{-1}). Lowering the pH to 6.25, however, afforded a

large increase in activity (197 h^{-1}), presumably due to protonation of the imine. The resulting rate is comparable to that for reduction of acetophenone under our standard conditions.

3.2.7. Regeneration of NADH

Regeneration of dihydronicotinamide adenine dinucleotide (NADH), an essential coenzyme in biological redox reactions, from nicotinamide adenine dinucleotide (NAD^+) is a biotechnologically important reduction reaction.[201-202] Although nicotinamide cofactors can be regenerated enzymatically, non-enzymatic approaches involving organometallic redox mediators represent increasingly attractive alternatives.[203-206]

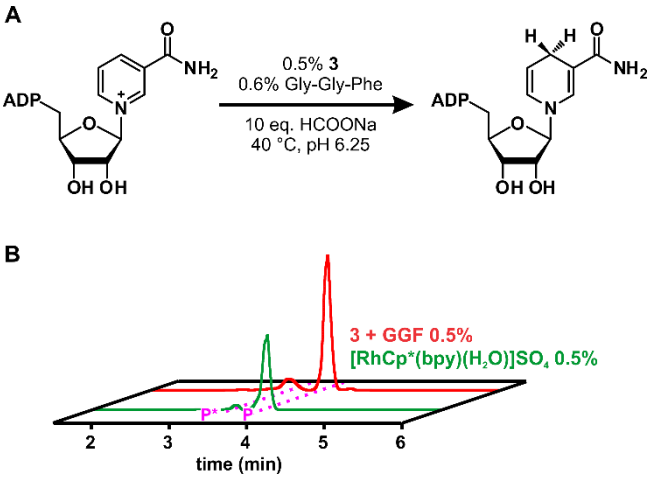


Figure 32: Regeneration of NADH. **A**: Reaction conditions for the reduction of NAD+ by formate catalyzed by **3** in the presence of Gly-Gly-Phe. **B**: Comparison of **3** and [RhCp*(bpy)(H2O)]SO4 under the conditions in A. Chromatograms at 340 nm, which show the formation of NADH (P, retention time 3.9 min) versus a non-natural NADH isomer (P*, 3.7 min).

The iridium complex with Gly-Gly-Phe efficiently reduces NAD^+ at pH 6.25 with a TOF of 225 h^{-1} (**Figure 32A**). The reaction proceeds regioselectively with a >10:1 preference for the biologically relevant 1,4-dihydronicotinamide isomer over the 1,2- and 1,6-isomers. The iridium catalyst is thus >2-fold more efficient than $[Rh(bpy)H_2O]SO_4$ (bpy = 2,2'-bipyridine), which has been previously used for the regeneration of NADH (TOF = 96 h^{-1}); the rhodium catalyst is somewhat more selective, though (20:1 in favor of 1,4-NADH, **Figure 32B**). These results are particularly notable as iridium catalysts often display significantly lower activities than corresponding rhodium and ruthenium complexes.[203] The combination of high turnover frequencies and good protein compatibility makes this system attractive for multistep chemoenzymatic processes.

3.3. Discussion

Our results show that peptide-iridium catalysts, which are readily formed under mild conditions in aqueous buffer, efficiently catalyze transfer hydrogenation of diverse ketones, aldehydes and imines. Simple tripeptides like Gly-Gly-Phe significantly enhance the reactivity of a water soluble d^6-piano stool iridium complex, providing turnover frequencies that rival or surpass those of established high-performance ligand systems.[192] They are thus interesting alternatives to amino acids[207-208] and pseudopeptides[209] for transfer hydrogenations under biological conditions.

The peptides likely bind the iridium complex via their free N-terminal amine and the adjacent amide to afford bifunctional Noyori-type catalysts. Because the C-terminal carboxylate is not required for activity, it should be possible to

append the tripeptides to the N-terminus of virtually any protein or foldameric structure as simple coordination tags to create artificial transfer hydrogenases. Formation of a productive complex, even in the presence of an excess of competing protein functionality, augurs well for transplantation of this catalytic motif into larger, more complex structures. Although the peptides tested to date confer no selectivity, placing the organometallic complex within the chiral environment of a protein would be expected to enable enantioselective transformations as observed for biotinylated transfer hydrogenation catalysts bound to streptavidin.[138, 210] In contrast to biotinylated systems, though, a wide range of structures, tailored to specific applications, can be considered as scaffolds. Utilization of a short peptide as opposed to a flexible linker to anchor the iridium complex to a protein is also likely to provide greater control over the placement of the catalytic center at the active site as well as a means of fine-tuning affinity and reactivity through systematic sequence variation. The defined structure of such complexes may facilitate computational design efforts, as well.

The biocompatibility of the peptide-iridium complexes and the low background activity observed in the absence of peptide are important prerequisites for in vitro and in vivo applications. Like biotin-streptavidin-based artificial transfer hydrogenases,[138, 210] it may be possible to combine these organometallic complexes with natural biocatalysts to generate productive reaction cascades. The fact that tripeptides are genetically encodable ligands also opens the possibility of generating artificial metalloenzymes in vivo. Coupling the activity of such species to the survival of a cell might then enable genetic selection of highly active and selective transfer hydrogenation catalysts.[211]

3.4. Materials and Methods

3.4.1. General procedures

NMR spectra were recorded on an AVIII 400 (^{1}H 400 MHz, ^{13}C 100 MHz) spectrometer. All ^{13}C-NMR spectra were ^{1}H-broadband decoupled and measured at room temperature. Chemical shifts are reported in ppm from tetramethylsilane with the solvent resonance resulting from incomplete deuteration. Data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, br = broad, m = multiplet or combinations thereof), coupling constants, and integration. Multiplicity was assigned according to a method described by Hoye.[212] Chemicals were purchased from Sigma, Acros, ABCR, Merck, Aldrich or Fluka and were used without further purification, unless otherwise noted. Buffer salts were purchased from Sigma, J.T. Baker, or Fluka and used without further purification.

Compounds **1**,[213] **2**,[214] **3**,[215] and [Rh(bpy)(H_2O)]SO_4[216] were synthesized as previously described in the literature. An XBridge C18 column (4.6 x 75 mm; particle size 2.5 µm) was employed for analytical RP-HPLC analysis. Compounds were typically eluted in 8 minute runs using a linear gradient from 5% acetonitrile containing 0.08% TFA to 95% containing 0.1% aqueous TFA at a flow rate of 1200 µL/min. For transfer hydrogenation of 1-methyl-3,4-dimethoxyisoquinoline and NAD^+ a linear gradient from 0% to 5% acetonitrile in 50 mM sodium acetate over 7 minutes was used. Conversion rates for all reactions were determined by consumption of starting material. The identity of the product was confirmed by comparison to an authentic standard.

3.4.2. Solid phase peptide synthesis of tripeptide ligands

Tripeptide ligands were synthesized manually by Fmoc solid phase peptide synthesis on a 0.25 mM scale in commercially available reaction vessels with Wang resin preloaded with either phenylalanine or glycine (Novabiochem); resin loadings were 0.62 mmol/g and 0.65 mmol/g, respectively. A ChemMatrix® Rink-amide resin (resin loading 0.55 mmol/g), preloaded with Fmoc-Phe as previously reported,[217] was used for the synthesis of Gly-Gly-Phe-NH_2. The resin was initially washed 5x with CH_2Cl_2 and 5x with DMF and then treated with 20% piperidine in DMF (1x2 min and 2x8 min). For preactivation of Fmoc-protected amino acids, 4.9 eq. of O-(6-chlorobenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate, 4.9 eq. of 2-(hydroxyimino)-3-nitrilopropanic acid ethyl ester, and 10 eq. of diisopropylamine were added to a solution of 5 eq. Fmoc-protected amino acid (0.3 M) in DMF. After preactivation for 2 minutes, the mixture was added to the resin and coupled for 45 minutes under vigorous shaking. The resin was subsequently washed 5x with DMF and, after deprotection of the Fmoc protecting group, with 20% piperidine in DMF before adding the next Fmoc-protected amino acid. Upon complete assembly, the peptide was globally deprotected by treatment of the dried resin with 5 mL of a deprotection cocktail (90% TFA, 5% CH_2Cl_2, 2.5% TIS, 2.5% water) for 2 hrs. Excess TFA was evaporated and the crude peptide was precipitated from ice-cold ether. The precipitate was recovered by centrifugation, washed again with ice-cold ether, dissolved in 30% acetonitrile in water, and purified by RP-HPLC. Product containing fractions were pooled and lyophilized. Yields, NMR analyses, and mass spectroscopic characterization of the synthesized tripeptides are summarized below.

Gly-Gly-Phe: 30.6 mg, 43.9% yield; ^{1}H NMR (400 MHz, CD_3OD) δ 7.39 – 7.12 (m, 5H), 4.69 (dd, J = 8.5, 5.2 Hz, 1H), 3.97 (d, J = 16.8 Hz, 1H), 3.90 (d, J = 16.8 Hz, 1H), 3.72 (s, 2H), 3.23 (dd, J = 13.9, 5.1 Hz, 1H), 3.00 (dd, J = 13.9, 8.5 Hz, 1H). ^{13}C NMR (101 MHz, CD_3OD) δ 173.23, 169.41, 166.39, 136.97, 128.92, 128.05, 126.41, 53.92, 41.52, 40.07, 37.06. HRMS (ESI, MH^+): calc. 280.1292, found 280.1296

Gly-Phe-Gly: 31.7 mg, 44.7% yield; ^{1}H NMR (400 MHz, CD_3OD) δ 7.36 – 7.27 (m, 4H), 7.27 – 7.18 (m, 1H), 4.76 (dd, J = 9.4, 5.2 Hz, 1H), 3.92 (d, J = 0.6 Hz, 2H), 3.72 (d, J = 16.1 Hz, 1H), 3.58 (d, J = 16.1 Hz, 1H), 3.24 (dd, J = 14.0, 5.2 Hz, 1H), 2.92 (dd, J = 4.0, 9.4 Hz, 1H). ^{13}C NMR (101 MHz, CD_3OD) δ 172.24, 171.44, 165.79, 161.51, 136.90, 128.87, 128.11, 126.46, 115.39, 54.66, 40.53, 40.00, 37.59. HRMS (ESI, MNa^+): calc. 302.1111, found 302.1116

Phe-Gly-Gly: 27.1 mg, 38.8% yield; ^{1}H NMR (400 MHz, CD_3OD) δ 7.45 – 7.29 (m, 5H), 4.19 (t, J = 7.3 Hz, 1H), 4.05 (d, J = 16.7 Hz, 1H), 3.96 (s, 2H), 3.83 (d ,J = 16.7 Hz, 1H), 3.26 (dd, J = 13.9, 7.0 Hz, 1H), 3.14 (dd, J = 13.9, 7.6 Hz, 1H). ^{13}C NMR (101 MHz, MeOD) δ 172.10, 170.15, 169.20, 169.12, 134.10, 129.17, 128.83, 127.60, 54.58, 41.78, 40.61, 36.96. HRMS (ESI, MNa^+): calc. 302.1111, found: 302.1111

N-Acetyl-Gly-Gly-Phe: For the last coupling step N-acetyl-glycine was used. 50.5 mg, 62.9 % yield; ^{1}H NMR (400 MHz, DMSO) δ 12.76 (bs, 1H), 8.14 (m, 2H), 8.05 (t, J = 5.8 Hz, 1H), 7.35 – 7.14 (m, 5H), 4.42 (td, J = 8.8, 5.1 Hz, 1H), 3.79 - 3.58 (m, 4H), 3.05 (dd, J = 3.8, 5.0 Hz, 1H), 2.89 (dd, J = 13.8, 9.0 Hz, 1H), 1.86 (s, 3H). ^{13}C NMR (101 MHz, DMSO) δ 173.21, 170.24, 169.72, 169.13, 137.94, 129.59, 128.68, 126.92, 53.96, 42.59, 42.05, 37.24, 22.95. HRMS (ESI, MNa^+): calc. 344.1217, found 344.1215

Gly-Gly-Gly-NH$_2$: 34.9 mg, 50.2% yield. ^{1}H NMR δ 7.37 – 7.17 (m, 5H), 4.65 (dd, J = 9.0, 5.5 Hz, 1H), 3.96 (d, J = 16.7 Hz, 1H), 3.85 (d, J = 16.7 Hz, 1H), 3.73 (s, 2H), 3.20 (dd, J = 13.9, 5.5 Hz, 1H), 2.92 (dd, J = 13.9, 9.0 Hz, 1H).^{13}C NMR (101 MHz, MeOD) δ 174.60, 169.49, 166.59, 137.06, 128.87, 128.06, 126.39, 54.41, 41.79, 40.05, 37.55. HRMS (ESI, MNa$^+$): calc. 301.1271, found 302.1274

Phe-Gly-Phe: 37.1 mg, 40.0% yield. ^{1}H NMR (400 MHz, MeOD) δ 7.46 – 7.12 (m, 10H), 4.69 (dd, J = 8.6, 5.1 Hz, 1H), 4.10 (dd, J = 8.2, 6.3 Hz, 1H), 4.01 (d, J = 16.7 Hz, 1H), 3.75 (d, J = 16.7 Hz, 1H), 3.25 (d, J = 5.7 Hz, 1H), 3.22 (d, J= 5.7 Hz, 1H), 3.09 – 2.95 (m, 2H). ^{13}C NMR (101 MHz, MeOD) δ 173.00, 169.20, 168.70, 136.91, 134.23, 129.05, 128.92, 128.71, 128.08, 127.45, 126.45, 54.44, 53.84, 41.44, 37.08, 37.05. HRMS (ESI, MNa$^+$): calc. 392.1586, found 392.1589

3.4.3. Typical procedure for transfer hydrogenation

Stock solutions of **3** (2.0 mM in water) and the corresponding peptide (2.4 mM in water) were freshly prepared, and 25 μL of each was added to 550 μL water in a 1.5 mL screw-cap Eppendorf tube. The resulting solution was allowed to incubate in a thermo mixer at 40 °C for 10 minutes while shaking (600 rpm) to facilitate formation of the peptide-iridium catalyst. Then, 200 μL of 250 mM phosphate buffer at the appropriate pH containing 500 mM sodium formate and 200 μL of substrate (50 mM in water) were added. The mixture was allowed to react for 30 to 60 minutes at 40 °C while shaking before stopping the reaction by cooling the samples to 4 °C. Aliquots were periodically removed and analyzed by RP-HPLC (5 μL injection) to monitor reaction progress and estimate initial turnover frequencies.

3.4.4. Transfer hydrogenation in the presence of BSA

BSA (10 or 50 µL of a 10 mg/mL stock solution in 250 mM sodium phosphate buffer, pH 8) was mixed with the Gly-Gly-Phe tripeptide (60 µM). Following addition of **3** (25 to 125 µL of a 2.0 mM stock solution in water) and incubation at 40 °C for 10 minutes, the resulting mixture was used for transfer hydrogenation as described for the standard procedure.

3.4.5. Total turnover number experiments

The total turnover number for the catalytic reduction of acetophenone was determined by treating 50 mM substrate with 25 µM **3**, 30 µM Gly-Gly-Phe, and 500 mM sodium format in 250 mM sodium phosphate buffer at pH 7 for 16 hours.

3.4.6. Representative HPLC chromatograms

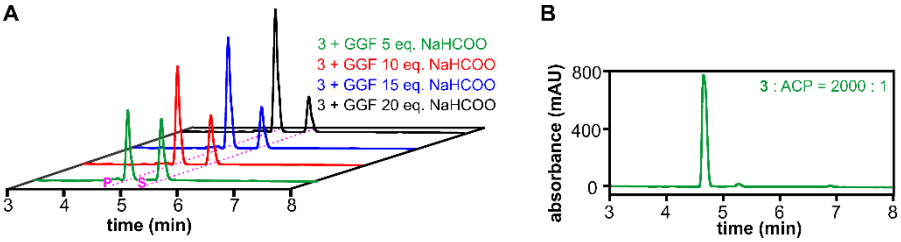


Figure 33: **A**: Reaction progress with increasing formate concentrations at a S/C ratio of 200:1 at pH 8 after 30 min. Substrate (S) and product (P) as in **Figure 6**. **B**: Chromatogram of the transfer hydrogenation of acetophenone with 10 eq. sodium formate after 16 hours at a S/C ratio of 2000:1 at pH 7.

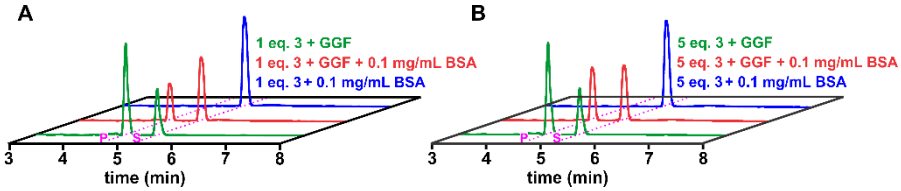


Figure 34: Comparison of the reaction progress for reactions in the presence of 0.1 mg/mL BSA with 1 eq. (**A**) and 5 eq. (**B**) 3 after 30 min at a S/C ratio of 200:1 at pH 8.

4. Building efficient enzymes with foldamer prostheses

4.1. Introduction

4.1.1. Why did nature choose alpha amino acids?

Before ribo-organisms ventured into the protein world, it is believed that these life forms took advantage of amino acids and small peptides to expand the structural and catalytic diversity of RNA molecules (Chapter 1.2.1).[10] However, this scenario requires that amino acids were present at reasonable concentrations on an early earth. Indeed, experiments with the aim to mimic prebiotic, chemical processes and samples from exobiological origins, such as carbonaceous meteorites, suggest that many of the proteinogenic amino acids were readily available. Miller and coworkers, for example, generated a wide variety of molecules of biological interest from simple inorganic precursors (H_2, CO, NH_3, CH_4) using an electronic discharge as energy source.[218] Moreover, similar concentrations of potential protein building blocks were identified in a carbonaceous meteorite that impacted near Murchison, Australia.[219-220] In both studies, glycine, alanine, glutamic acid, valine and proline were found to be the most abundant molecules with concentrations of more than 1 µg/g. Notably though, the list of identified organic compounds also contained a small number of β-amino acids and N-substituted amino acids. In light of the fact that even modern organisms utilize these non-natural amino acids for a wide variety

of different functions,[221] one has to wonder whether the selection of the canonical amino acids as basic building blocks for proteins was arbitrary, or whether α-peptides confer a significant advantage over their non-natural counterparts.

4.1.2. Mimicking protein functions with non-natural oligomers

The folding of linear, sequence specific peptide chains into well-defined structures endows proteins with a remarkable diversity of functions.[222] Chemists and biologists have long sought to mimic the breadth of structure and function of proteins with non-natural oligomers.[102, 223] Such molecules, often referred to as "foldamers", use building blocks other than the canonical α-amino acids, arranged in sequences enabling them to adopt distinct conformations in solution. Over the last two decades, efforts toward this goal have revealed a surprisingly large number of polymers that fold under biologically relevant conditions.[224-227] From the growing list of foldamers, β-peptides have proven particularly amenable to mimic protein functions.[228] Judiciously chosen β-amino acid sequences have been exploited to design biomimetic functions, including self-assembly,[229] protein binding,[230-231] and cell entry or lysis.[232-233] However, mimicking the complex interplay of structure and function in enzymes has thus far proven to be more challenging.

Our group has previously designed a β-peptidic catalyst for a retroaldol reaction by decreasing the pK_a of an amine side chain through Coulombic interactions with nearby cations; a strategy often employed by natural enzymes (**Figure 35**).[103] Although only 10 residues long and designed according to very

basic principles, the self-aggregating β-peptide achieved reasonable rate accelerations for the cleavage of a model substrate (k_{cat}/k_{uncat} = 3000).

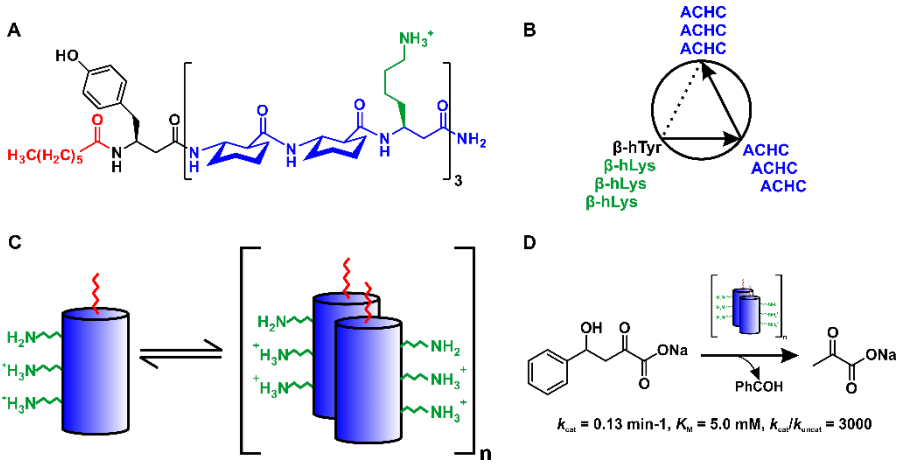


Figure 35: A rationally designed aldolase foldamer: **A**: Sequence and structure of the designed foldamer catalyst. The design features a threefold repeat of two cyclically-constrained β-amino acids (blue) and a β-lysine residue (green). A β-tyrosine was introduced at the C-terminus to simplify biophysical characterization and a hydrophobic tail group was incorporated (red) to facilitate self-aggregation of the catalyst. **B**: Helical wheel representation of the aldolase foldamer: Introduction of cyclically constrained building blocks favors formation of a 3.14 helix that places three residues in a 120° projection in one turn of the helix. **C**: Aggregation of the foldamer via the hydrophobic tail group results in catalytically active but ill-defined species. **D**: The catalyst in its aggregated form accelerates the retro-aldol cleavage of β-hydroxy ketones. Catalytic parameters for the catalyzed reaction are shown (adapted from reference [103]).

However, this proof-of-principle study also highlights shortcomings of such de novo designed enzyme mimics. Besides the forming an ill-defined, catalytic species, the foldamer also lacks a well-defined binding pocket for the substrate as indicated by the high K_m value, limiting its proficiency and selectivity. Thus, we reasoned that pursuing a top-down approach – in which foldamers are

evaluated within an enzymatic scaffold – might reveal design criteria that could ultimately facilitate the construction of productive active sites in de novo designed enzyme mimics.

A first step toward this end requires testing the consequences of introducing non-natural elements in a natural context. For example, an amphiphilic helix in the protein hormone interleukine-8 could successfully be mimicked by a foldamer surrogate in spite of the inverted handedness and dipole orientation of the β-peptide fragment.[234] Similarly, studies from the Raines group have demonstrated that replacing a reverse turn in bovine pancreatic ribonuclease (RNase A) with synthetic analogs (i.e. β^2-homoalanine-β^3-homoalanine) afforded active enzymes, despite the non-natural hydrogen bonding patterns of the turn surrogate.[235] However, increasing the number of non-natural building blocks in RNase A severely reduced the activity of the hybrid catalysts.[236] Zuckermann and colleagues introduced only three peptoid residues into a helix containing catalytically important groups, severely compromising both hydrophobic packing and native hydrogen-bonding networks. As a result, largely unfolded variants of the protein were obtained that were 100-times less active than the wild-type enzyme.

In order to incorporate non-natural building blocks in segments crucial for catalysis, we envisioned taking advantage of the ability of α-helices to accommodate the substitution of up to one third of their canonical amino acids with β^3-amino acids[237]. Such "α/β-peptides", if properly designed, act as powerful, alpha-helical inhibitors for protein interactions, indicating that the parent structure is not severely distorted upon introduction of these non-natural backbone units. The Gellman group has pioneered the use of foldamers with heterogeneous backbone for constructing inhibitors of protein

interactions for a variety of biologically relevant targets.[237] When compared to their natural counterparts, mixed α/β-peptides exhibit enhanced conformational stability and superior protease resistance. In this work, we utilized the informational mimicry of these α/β-peptides to design foldamer prostheses that complement an otherwise inactive protein fragment and, as a result, create highly efficient hybrid enzymes.

4.2. Results

4.2.1. Model system

An engineered heterodimeric chorismate mutase (hdCM, **Figure 36A**),[104] which catalyzes the sigmatropic rearrangement of chorismate to prephenate served as a protein scaffold for the assembly of hybrid enzymes. hdCM was originally constructed by dissecting the helical-bundle chorismate mutase from Methanocaldococcus jannaschii into a short N-terminal helix ($MjCM_{1-21}$) and a 3-helix segment ($MjCM_{22-93}$) and attaching an antiparallel leucine zipper dimerization domain to the individual fragments (yielding $MjCM_{1-21}$-NZ and CZ-$MjCM_{22-93}$). This split enzyme assembles spontaneously from the separately inactive fragments and catalyzes the conversion from chorismate to prephenate with efficiencies comparable to the parent homodimeric enzyme. When compared to other split enzymes,[238] hdCM combines a unique set of desirable traits for the construction of foldamer-containing catalysts. The 52 amino acid long one-helix fragment, $MjCM_{1-21}$-NZ, is readily accessible by synthesis, thus facilitating the introduction of β^3-amino acids into the N-terminal helix. A methionine-to-cysteine mutation in the linker sequence of

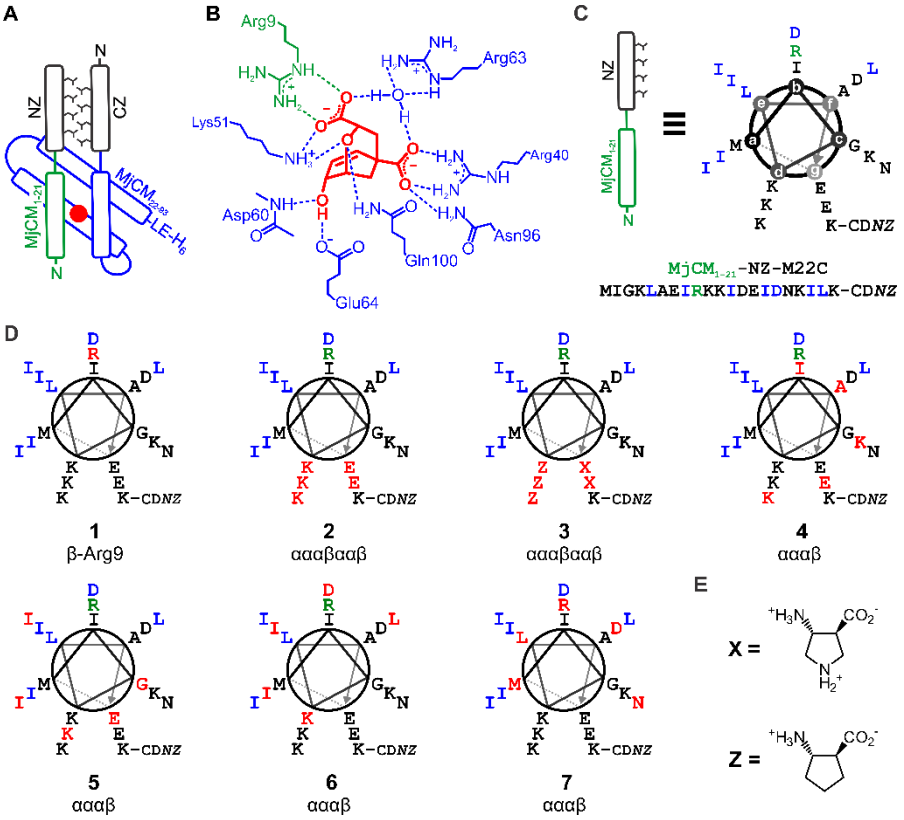


Figure 36: Design principles of α/β-hdCM. **A**: Schematic drawing of hdCM: antiparallel leucine zipper in black, 3-helix domain ($MjCM_{22-93}$) in blue, one-helix segment ($MjCM_{1-21}$) in green, active site in red. **B:** Schematic view of the putative hdCM active site with bound transition state analog in red; polar contact residues from $MjCM_{22-93}$, in blue, from $MjCM_{1-21}$ in green. **C:** Sequence and helical wheel representation of $MjCM_{1-21}$-NZ-M22C: the catalytic Arg in green, residues at the helical interface in blue. **D**: Design of chimeric foldamer one-helix fragments: β-amino acids are displayed in red with the color code for other amino acids as before. **E:** Structures of cyclic building blocks used in **3.** Trans-4-aminopyrrolidine-3-carboxylic acid (**X**) and trans-2-aminocyclopentane-carboxylic acid (**Z**).

NZ and $MjCM_{1-21}$ enables modular attachment of different designs to the invariable leucine zipper portion by native chemical ligation.[239-240] Additionally, all key interactions at the helical interface of hdCM are made through its side chains, alleviating some of the constraints on backbone modification. Lastly, the active site of hdCM is formed jointly by both fragments (**Figure 36B**). In the active complex, $MjCM_{1-21}$-NZ not only provides hydrophobic interactions that ensure tight packing of the 4-helix bundle hdCM but also an essential active side residue, Arg9 (**Figure 36C**). Consequently, enzymatic activity offers a sensitive read-out of the ability of foldamer-containing $MjCM_{1-21}$-NZ variants to mimic their natural counterpart.

4.2.2. Design

We designed foldamer prostheses for hdCM by systematically introducing $\alpha \rightarrow \beta^3$ substitutions in the N-terminal helix yielding seven α/β-peptides that differ in number, location, and type of β^3-amino acids placed within the $MjCM_{1-21}$ segment (**Figure 36D**).

Design **1** bears a single $\alpha \rightarrow \beta^3$ substitution at the catalytically crucial Arg9 to assess whether a side chain projecting from the artificial backbone unit is able to adopt an active conformation. Foldamers **2** and **3** are based on a heterogeneous ααβαααβ backbone with five mutations clustered at the solvent exposed face of $MjCM_{1-21}$, thereby minimizing perturbations in the orientation of side chains at the helical interface or the active site. While the side chains in **2** remain unchanged, we introduced helix stabilizing cyclic β-amino acids[237] in **3** (**Figure 36E**) to probe whether rigidifying the helix is beneficial for complex formation and catalysis. Designs **4-7** feature four permutations of the αααβ

repeat, and thus deviate from the regular heptad periodicity of alpha helices causing the five β^3-amino acids to spiral around the helix. We surmised that the impact of the backbone modifications would be least severe in **4** and most dramatic in **7**. While the mutations in **4** occur predominantly on the solvent-exposed face of the N-terminal helix, **5** and **6** harbor two and four mutations at residues important for hydrophobic packing, respectively. In addition to the substitution of the catalytic arginine, β^3-amino acids in foldamer **7** cluster around critical regions for helical bundle formation.

4.2.3. Synthesis

Helical $MjCM_{1-21}$-NZ were prepared by Fmoc-based solid phase peptide synthesis. Introducing a native chemical ligation site in the linker sequence of NZ and $MjCM_{1-21}$ via a methionine to cysteine mutation yields a modular system, in which different designs can be joined to the invariable leucine zipper portion (**Figure 37**). Syntheses of $MjCM_{1-21}$ variants as C-terminal acyl hydrazides and S-*tert*-butyl protected NZ afforded peptides in good purities (**Figures 43** and **44** in the Materials and methods). In the following one-pot, two-step ligation the hydrazide moiety is first converted to the corresponding C-terminal acyl azide, which undergoes a clean native chemical ligation in the presence of 4-mercaptophenylacetic acid and NZ.[239] Following this strategy all foldamer containing one-helical fragments (**1-7**) as well as the corresponding α-peptide $MjCM_{1-21}$-NZ containing the M22C mutation were synthesized in excellent purities with overall yields of up to 8% (**Figure 38** and **Table 4** in the Materials an Methods section). In contrast, the 115 amino acids long three helix fragment CZ-$MjCM_{22-93}$ was produced recombinantly in a chorismate mutase deficient

E. coli strain, thereby excluding the possibility of contaminations with endogenous enzyme.

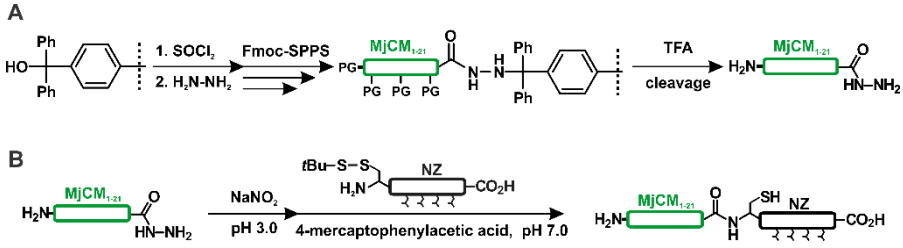


Figure 37: Synthesis strategy for MjCM$_{1\text{-}21}$-NZ fragments. **A**: Starting from a Trityl-OH resin, standard Fmoc-based solid phase peptide synthesis afforded MjCM$_{1\text{-}21}$ variants as C-terminal acyl hydrazides. **B**: One-pot, two-step native chemical ligation with the invariable leucine zipper portion.

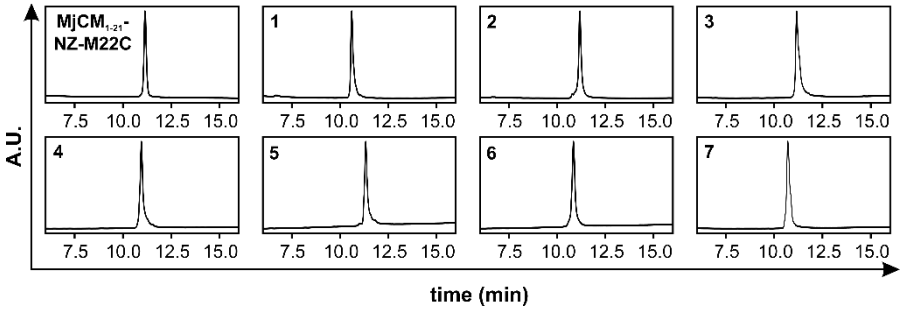


Figure 38: HPLC chromatogram for MjCM$_{1\text{-}21}$-NZ fragments. Overall yields and HI-RES-MS results can be found in **Table 4** in the Materials and Methods.

4.2.4. Biophysical characterization

Introducing β-amino acids in MjCM$_{1\text{-}21}$-NZ does not adversely affect folding of this segment. In circular dichroism measurements all variants are helical at 10 µM; foldamer prostheses containing five building blocks (**2**-**7**) display a more pronounced minimum at 208 nm, which has been previously observed

for α/β-peptides (**Figure 39A**).[241] We also found large portions of the one-helix fragment folded when recording 1D-^{1}H-NMR spectra for MjCM$_{1-21}$-NZ-M22C as well as for foldamers **2** and **3**, as indicated by good peak dispersion in the amide regions (**Figure 39B**).

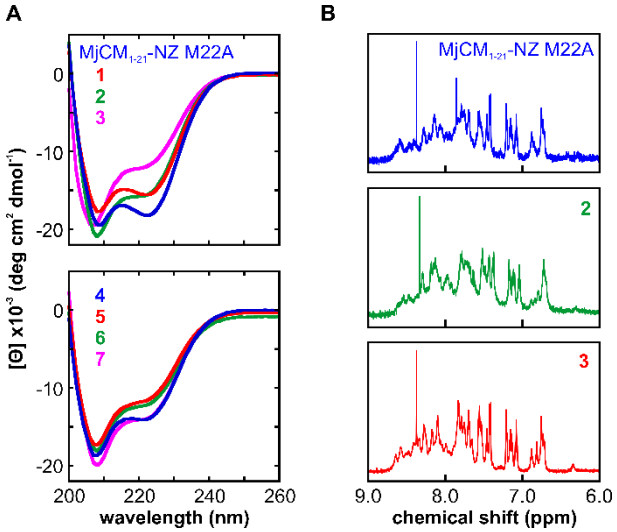


Figure 39: Biophysical characterization of MjCM$_{1-21}$-NZ fragments. **A**: Far-UV CD spectra for one-helix segments at 10 μM concentration. The color code for the curves is provided in the plots. **B**: 1D-^{1}H-NMR spectra for the parent variant (blue), **2** (green), and **3** (red) recorded at 50 μM fragment concentration.

Notably though, for all spectra well-defined peaks and regions with broad, unresolved signals are observed. Such a behavior is indicative of a dynamic structure.[242] This trend was confirmed when we attempted to directly monitor complex formation of all chimeric proteins by size-exclusion chromatography. In fact, when MjCM$_{1-21}$-NZ variants (50 μM) are incubated with equimolar concentration of CZ-MjCM$_{22-93}$, chromatograms display a peak at higher

molecular weights compared to the individual fragments, consistent with heterodimer formation (**Figure 47**). However, multiple peaks for the one-helix fragments and broad peaks in the other chromatograms demonstrate the dynamic nature of these constructs and make an assignment of the species difficult. The flexibility of hdCM and its components can presumably be attributed to the leucine zipper feature. Although the formation of the heterodimer is preferred, the individual parts are able to interconvert between monomeric and multimeric states and as a result, yield a complex equilibrium. However, it is conceivable that a well-defined species forms upon addition of the substrate, as similar effects have been observed for a structurally disordered chorismate mutase.[243-244]

4.2.5. Kinetic characterization

The prowess of foldameric prostheses was evaluated by their ability to catalyze the chorismate mutase reaction in the presence of recombinantly produced CZ-MjCM$_{22\text{-}93}$. Reference values are provided by the catalytic parameters of the parent variant, which displays wild-type activities (MjCM$_{1\text{-}21}$-NZ-M22C, k_{cat} = 6.7 s^{-1}, K_m = 150 µM and k_{cat}/K_m = 45,000 $M^{-1}s^{-1}$), and MjCM$_{1\text{-}21}$-NZ-R9A (k_{cat}/K_m = 1.7 $M^{-1}s^{-1}$), a knock-out variant that highlights the importance of Arg9 for catalysis and does not show saturation kinetics. Upon mixing synthetic MjCM$_{1\text{-}21}$-NZ constructs with CZ-MjCM$_{22\text{-}93}$ we observed apparent second-order rate constants (k_{cat}/K_m) for chorismate consumption that span four orders of magnitude, dependent on the location of the introduced β^3-amino acids (**Figure 40**, for a concise overview for all variants see **Figure 42A-C** and **Table 3**).

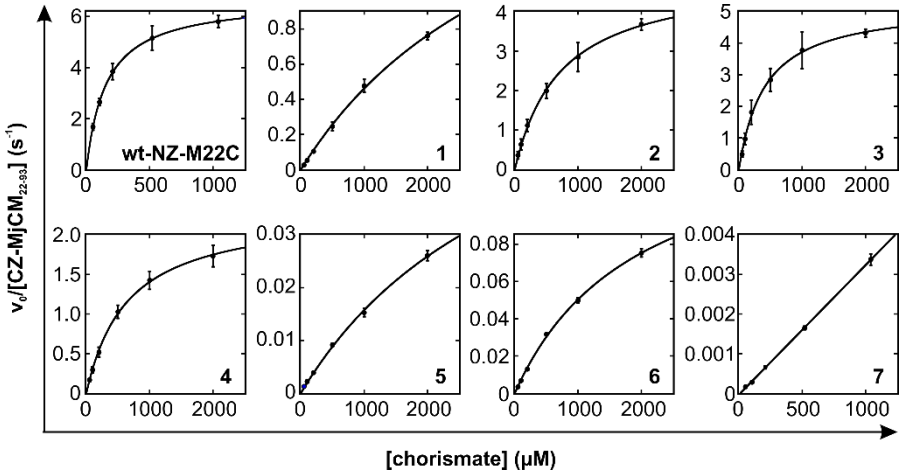


Figure 40: Kinetic characterization of hdCM variants. $MjCM_{1\text{-}21}$-NZ-M22C and foldamers **1** – **6** were fitted non-linearly according to the Michaelis-Menten equation. Variant **7** did not show saturation kinetics and a linear fit was used to estimate the k_{cat}/K_m value.

Surprisingly, the isolated α→β^3 substitution of Arg9 in **1** only has a modest effect on k_{cat} (2.2 s^{-1}). However, the K_m value (3830 μM) is 25 fold higher than for the parent construct, suggesting an energetic penalty for orienting the side chain of the artificial building blocks in a conformation favorable for interacting with the enol pyruvate side chain carboxylate group of the substrate (and the transition state). Segregation of non-natural backbone units away from the helical interface and active site in **2** and **3** afforded highly efficient hybrid enzymes. Turnover numbers for these variants (5.0 s^{-1} and 5.2 s^{-1}, respectively) rival the k_{cat} value obtained for the parent variant (6.7 s^{-1}), and k_{cat}/K_m values are within one order of magnitude of natural chorismate mutases.[166],[245] All four designs in the αααβ series catalyze the anticipated reaction, albeit with diverse activities. While changes proximal to but not within the active site and the helical interface are well-tolerated (**4**, k_{cat} = 2.3 s^{-1} k_{cat}/K_m = 3450 $M^{-1}s^{-1}$),

substitutions of residues within these critical regions yield inferior catalysts. Designs **5** and **6** with altered helical interfaces display catalytic efficiencies that are 2200 and 600 fold lower (20 and 76 $M^{-1}s^{-1}$), respectively, than the parent enzyme. Additional replacement of the catalytic arginine with β^3-hArg drastically diminishes activity; **7** does not display saturation kinetics (k_{cat}/K_m = 3.4 $M^{-1}s^{-1}$) and is only roughly twice as efficient as the R9A mutant.

Because the introduction of β^3-amino acids does not adversely affect folding of the helical $MjCM_{1\text{-}21}$-NZ fragment (**Figure 39A**), the varying efficiency of foldamer prostheses presumably reflects successful formation of an active complex with CZ-$MjCM_{22\text{-}93}$ as well as accurate projection of the catalytically crucial arginine side chain into the binding pocket. While the observed k_{cat} values provide information on the latter, the dissociation constant (K_d) of a given design reports on the former. We determined dissociation constants for the complexes by adding increasing amounts of the $MjCM_{1\text{-}21}$-NZ variants to fixed concentrations of CZ-$MjCM_{22\text{-}93}$ and measuring reaction rates. Dissociation constants for variants not sufficiently active for titration were estimated by monitoring inhibition of the formation of an active complex with **2** (**Figure 41**, **Figure 42D**, and **Table 3**).

Homologation of Arg9 alone (**1**) or residues at the boundaries of critical regions (**4**) afforded molecules that are readily capable of forming catalytically active complexes, although dissociation constants (580 nM for **1** and 3.6 μM for **4**) are elevated when compared to $MjCM_{1\text{-}21}$-NZ-M22C (260 nM). As anticipated, the introduction of cyclic building blocks in **3** (K_d = 760 nM) generates a prosthesis that displays a four-fold higher affinity for CZ-$MjCM_{22\text{-}93}$ with respect to **2** (2.9 μM), which lacks these stabilizing residues. Consistent with an improved pre-organization of the N-terminal helix, **3** unfolds at higher

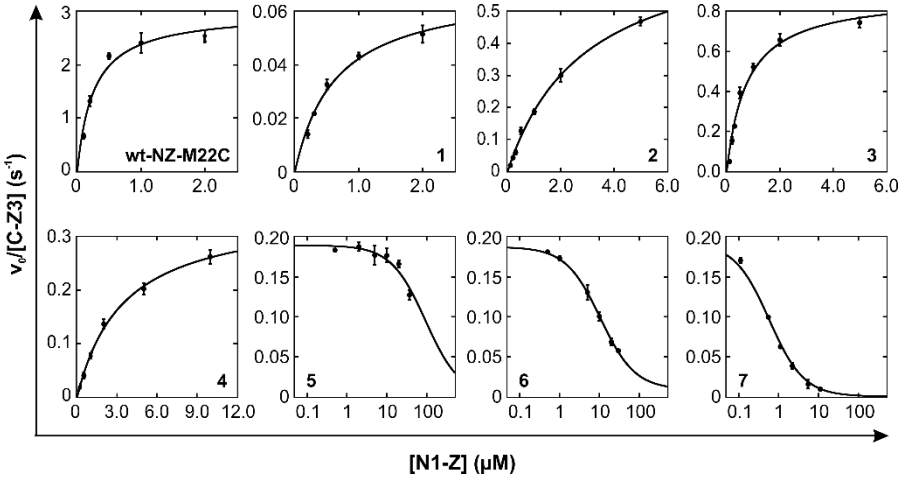


Figure 41: Determination of dissociation constants and IC_{50} values for $MjCM_{1-21}$-NZ variants. K_d values for $MjCM_{1-21}$-NZ-M22C and foldamers **1** – **4** were determined by adding increasing amounts of the one-helix fragment to CZ-$MjCM_{22-93}$ and nonlinear fitting of the resulting hyperbolic binding curves to the dimeric binding model. For foldamers **5** -**7**, IC_{50} values were obtained by monitoring inhibition of the formation of an active complex with **2** and fitting the data points to the Hill equation. K_d values for **5** – **7** were subsequently estimated by the Cheng-Prusoff equation[246] (see Materials and Methods).

temperatures than the parent variant or **2** (**Figure 46**). K_d values for **5** and **6** are significantly higher (70.0 μM and 7.9 μM) demonstrating that introduction of β^3-amino acids at the helical interface severely compromises complex formation. As a result of this handicap, we substantially underestimated the performance of both prostheses in our previous kinetic characterization that was carried out at 2 μM CZ-$MjCM_{22-93}$ and a 5 fold excess of one-helix fragments. Upon extrapolation, k_{cat} values for **5** and **6** increase by a factor of 8.2 (0.60 s^{-1}) and 1.9 (0.28 s^{-1}) with respect to the initially obtained values (**Table 3**). Surprisingly, **7** displays the most efficient complex formation (K_d = 430 nM) of all foldamer prostheses. Taking into account the low catalytic efficiency of this

design, it is possible that the adaptation of an unproductive side chain conformation for Arg9 in **7** has a beneficial effect on formation of the heterodimer. The active site of chorismate mutase is highly polarized, and consequently binding energy could potentially be gained by avoiding the electrostatic repulsion for Arg9 in this environment.

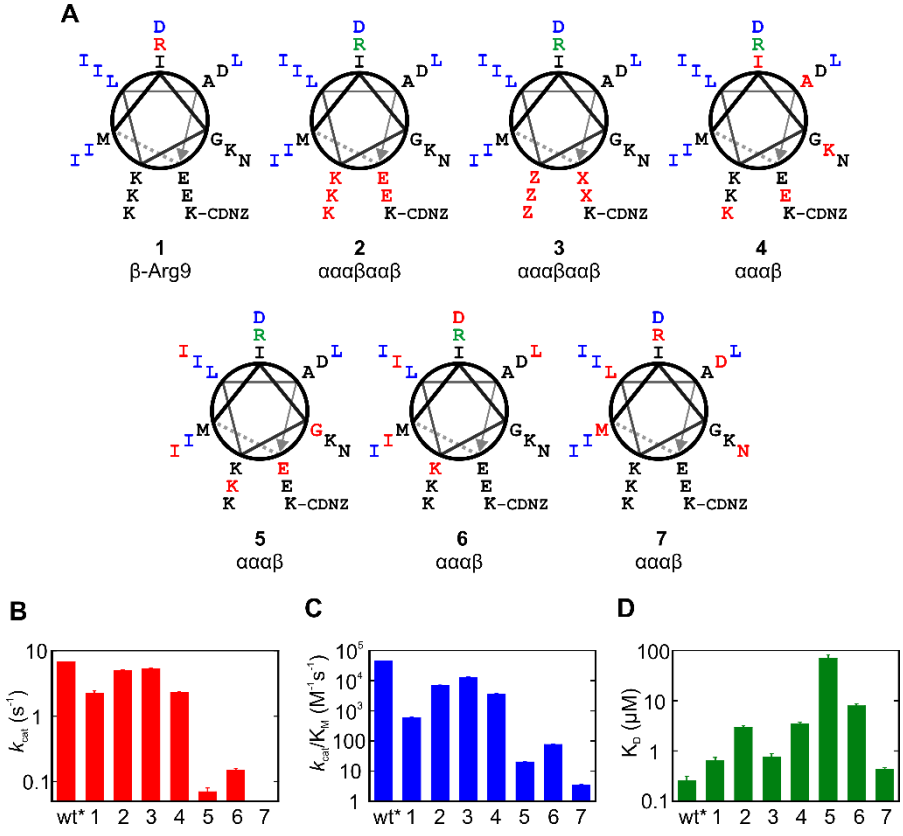


Figure 42: Kinetic characterization of α/β-heterodimeric chorismate mutases. **A:** Foldameric MjCM$_{1\text{-}21}$-NZ variants, color code as in **Figure 36**. **B-D:** Comparison of observed turnover numbers (k_{cat}), apparent second-order rate constants (k_{cat}/K_m), and dissociation constants (K_d), wt* refers to MjCM$_{1\text{-}21}$-NZ-M22C.

Table 3 Catalytic parameters of heterodimeric chorismate mutases investigated in this chapter.

MjCM$_{1-21}$-NZ fragment	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)	K_d (μM)
MjCM$_{1-21}$-NZ-M22C	6.69 ± 0.03	150 ± 2	44,600 ± 700	0.26 ± 0.06
1	2.2 ± 0.2	3800 ± 500	580 ± 30	0.64 ± 0.11
2	5.0 ± 0.1	720 ± 40	6,900 ± 300	2.93 ± 0.30
3	5.2 ± 0.1	410 ± 30	13,000 ± 800	0.76 ± 0.13
4	2.3 ± 0.1	650 ± 40	3600 ± 200	3.49 ± 0.31
5	0.07 ± 0.01	3700 ± 470	20 ± 1	70 ± 12.5 [b]
6	0.15 ± 0.01	2000 ± 150	76 ± 4	7.9 ± 0.8 [b]
7[a]	-	-	3.4 ± 0.2	0.4 ± 0.04 [b]
MjCM$_{1-21}$-NZ-R9A [a]	-	-	1.5 ± 0.3	n.d.

[a] MjCM$_{1-21}$-NZ-R9A and **7** did not show saturation kinetics. [b] K_d values were determined by measuring inhibitor constants for an active one-helix fragment (**2**).[c] extrapolated values from the obtained K_d values. n.d.: not determined

4.3. Discussion

Previous studies have shown that the introduction of individual non-natural building blocks into an enzymatic scaffold is well-tolerated.[235] However, increasing the number of nonstandard amino acids in protein catalysts has thus far only afforded variants with activities that are orders of magnitude lower than those displayed by natural enzymes.[236] In this chapter, we demonstrate

for the first time the feasibility of constructing highly active hybrid catalysts by replacing a catalytically-crucial secondary structure element with an α/β peptide. We have generated chimeric chorismate mutases containing up to 5 β-amino acids, including an essential catalytic residue, to create variants that rival the activity of the wild-type enzymes.[104]

Altogether, our results provide key insights on how to construct efficient enzymes containing foldamer prosthetics employing a top-down approach**.** Design activities are dependent upon adequate patterning of the foldamer. Backbone mutations are well-tolerated when bundled at solvent exposed sites, yielding highly active hybrid chorismate mutases. Preorganization with conformationally constrained building blocks further improves foldamer-protein assembly, affording proficient chimeric enzymes that display activities similar to naturally occurring chorismate mutases[245]. In contrast, introducing multiple non-natural building blocks within the helix bundle interface yields inferior or, in combination with an active site mutation, crippled catalysts.

Such design rules will be crucial for enabling a straightforward introduction of non-natural modules into a wide variety of enzymatic scaffolds. The number of accessible secondary structures and the improved conformational stability of foldamers could provide unprecedented control over biological functions of proteins. Because structures of non-natural polymers are highly predictable, imperfect prostheses, such as **5**-**7**, can presumably be optimized by computationally-aided laboratory evolution. Ultimately, catalysts that take full advantage of the superior properties of foldamers could allow for the introduction of novel functionality that is not accessible with natural enzymes.

4.4. Materials and Methods

4.4.1. General procedures

All reagents, resins, amino acids, and solvents were obtained from commercial suppliers and used without further purification, unless otherwise noted. 1-Hydroxy-7-azabenzotriazole (HOAt) was a generous donation from Prof. Jeffrey Bode. Manual peptide synthesis was performed in SPPS reaction vessels (15 mL) at room temperature using O-(6-chlorobenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluoro-phosphate (HCTU) and ethyl cyanoglyoxylate-2-oxime (Oxyma) as coupling reagents for α-amino acids and O-(7-azabenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HATU) and HOAt for β-amino acids. Completion of couplings was verified by the TNBS test, whereas deprotection was monitored by UV (304 nm) on a NanoDrop spectrophotometer (Witeg AG). Crude peptides were purified by preparative HPLC (Waters) on a Nucleosil 100-7 C18 column (Machery-Nagel) with a gradient of 5-60% acetonitrile (0.08% TFA) in water (0.1% TFA) at a flow rate of 10 mL/min. Analytical HPLC (Waters) was performed employing the same gradient on a XBridge C18 column (Waters, 4.6 x 150 mm; particle size 2.5 μM) at a flow rate of 1 mL/min.

4.4.2. Preparation of Fmoc-Lys(Boc)-Trityl-NHNH$_2$-resin

In an oven-dried 100 mL round-bottom flask, 1 g of Trityl-OH resin (ChemMatrix, 0.47 mmol/g) was added and swelled in 15 mL dry DCM under inert atmosphere. After 30 min, 300 μL $SOCl_2$ (4.11 mmol) was added dropwise and the reaction mixture shaken overnight at room temperature. The resin was

transferred into an SPPS vessel, drained, and washed with dichloromethane (DCM, 5 x 15 mL), 2% N-methylmorpholine in DCM (3 x 15 mL), and N,N-dimethylformamide (DMF, 5 x 15 mL). After the final washing step, the resin was drained and transferred into a clean 100 mL round-bottom flask. After addition of 10 mL DMF the suspension was cooled to 0 °C while shaking and a solution of triethylamine (198 µL, 1.41 mmol) and hydrazine hydrate (228 µL, 4.7 mmol) in 1 mL DMF added dropwise. The reaction mixture was shaken for 30 min on ice, following another addition of triethylamine and hydrazine hydrate. The suspension was allowed to warm to room temperature and stirred for another 30 minutes. The resin was subsequently transferred to a new SPPS vessel, drained, and washed with DMF, DCM, and DMF (5 x 15 mL for each washing step). The resulting Trityl-$NHNH_2$ resin was used for loading of the first amino acid without further characterization.

For the loading procedure, Fmoc-Lys(Boc)-OH (2200 mg, 4.7 mmol), HCTU (1847 mg, 4.47 mmol) and Oxyma (635 mg, 4.47 mmol) were dissolved in 10 mL DMF. The amino acid was activated by addition of N,N-diisopropylethylamine (DIPEA, 1626 µL, 9.4 mmol). The reaction mixture was stirred for 2 minutes and added to an SPPS vessel containing freshly prepared Trityl-$NHNH_2$ resin. The coupling step was performed for 60 minutes while shaking at room temperature. After washing the resin with DMF (5 x 15 mL, 1 minute for each washing step), the activation and coupling procedure was repeated and the reaction mixture shaken for an additional 60 minutes at room temperature. Subsequently, the resin was drained and rinsed with DMF (5 x 15 mL). Unreacted Trityl-$NHNH_2$ was capped by adding 10 mL of a solution containing 1 mL acetic anhydride and 1 mL DIPEA in 10 mL DMF to the resin. After 10 minutes at room temperature the resin was

thoroughly washed with DMF, DCM, and ether (5 x 15 mL for each washing step). Lastly, the resin was drained and dried in vacuo overnight to give 1 g of Fmoc-Lys(Boc)-Trityl-$NHNH_2$ resin with typical amino acid loadings of 0.15-0.22 mmol/g. Dried resin was stored for up to two weeks at -20 °C and was used in the synthesis of all $MjCM_{1\text{-}21}\text{-}NHNH_2$ variants.

4.4.3. Synthesis of $MjCM_{1\text{-}21}\text{-}NHNH_2$ variants

Fmoc-Lys(Boc)-Trityl-$NHNH_2$ (0.035 mmol, 1eq.) was placed in a 15 mL SPPS vessel and swelled for 30 minutes in 5 mL DMF. The resin was subsequently drained and amino acids are coupled following standard Fmoc-SPPS methods. In brief, appropriately protected Fmoc-α-amino acids (0.21 mmol, 6 eq.) were dissolved in 500 µL DMF. For activation, 500 µL of a solution containing HCTU and Oxyma (0.4 M in DMF, 5.71 eq.) and 350 µL 1 M DIPEA in DMF (10 eq.) were consecutively added, and the reaction mixture shaken for 2 minutes prior to addition to the resin. Couplings were performed for 60 minutes while shaking. The resin was subsequently washed with DMF (5 x 5 mL for 1 min). After each coupling step, unreacted amines were capped for 5 minutes by addition of DIPEA and acetic anhydride (50 eq. each) in 5 mL DMF. The resin was then drained, washed with DMF (5 x 5 mL for 1 min), and the Fmoc protecting group was removed by addition of 20% piperidine in DMF (5 mL) for 18 minutes (1 x 2 minutes and 2 x 8 minutes). Couplings with β^3-amino acids were carried out using 3 eq. of Fmoc-β^3-amino acid, a solution of HATU and HOAt (0.4 M in DMF, 2.9 eq.), and 6 eq. DIPEA (1 M in DMF) for 90 minutes. After capping unreacted amines, the Fmoc protecting group was removed by addition of 20% piperidine in DMF (5 mL) for 24 minutes (3 x 8 minutes). After complete chain assembly, the resin was thoroughly

washed with DMF (5 x 5 mL) and DCM (5 x 10 mL) and subsequently dried in vacuo at room temperature overnight. $MjCM_{1\text{-}21}$-$NHNH_2$ variants were then cleaved from the resin by addition of a cleavage cocktail containing 85 % TFA, 5% thioanisole, 5% 1,2-ethanedithiol (EDT), 2.5% water, and 2.5% DCM (10 mL) at room temperature under inert atmosphere (N_2). After two hours the resin was washed with TFA (3 x 5 mL) and the combined flowthroughs were concentrated by rotary evaporation. Crude peptides were obtained by precipitation from ice-cold ether and centrifugation. The crude peptides were dissolved in 30% acetonitrile in water containing 0.1% TFA, lyophilized, and subsequently purified by preparative HPLC. Individual fractions were analyzed by analytical HPLC and pure fractions were pooled and lyophilized. $MjCM_{1\text{-}21}$-$NHNH_2$ variants were typically obtained in 6-17% overall yields and good purities (see **Figure 43** for HPLC traces, yields, and results of mass spectrometry); the only appreciable impurity was the corresponding MjCM1-21-CO_2H fragment, presumably the result of incomplete conversion of Trityl-Cl to the Trityl-$NHNH_2$ resin.

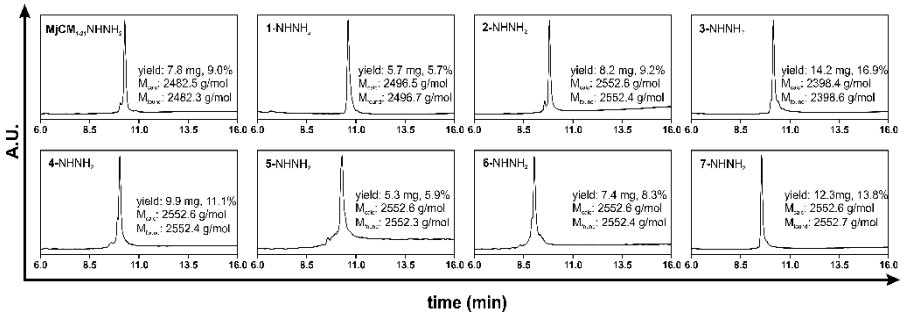


Figure 43: HPLC chromatograms of N-terminal helix fragments $MjCM_{1\text{-}21}$, C-terminally activated as acyl hydrazides. Yields and masses are displayed next to the traces (masses were determined by LC/MS).

4.4.4. Synthesis of *S*-*t*Bu-CD-NZ

S-tBu-CD-NZ was synthesized on an HMPB resin (0.22 mmol/g ChemMatrix) preloaded with Fmoc-Gln(Trt)OH by methods previously reported,[104] on a 433A peptide synthesizer (Applied Biosystems) on a 0.1 mmol scale with Fmoc-protected amino acids following standard procedures. After complete chain assembly, the peptide was cleaved from the resin by incubation in 95% TFA 2.5% H_2O, and 2.5% triisopropylsilane (TIS) for 2 hours. The resin was filtered off, washed with TFA (3 x 5 mL), excess TFA removed by rotary evaporation, and the peptide precipitated with cold ether. The crude peptide was recovered by centrifugation and dried under vacuum. The residue was dissolved in 30% ACN in water containing 0.1% TFA and purified by preparative HPLC. Product containing fractions were analyzed by analytical HPLC, and pure fractions were pooled and lyophilized. The identity of the peptide was confirmed by LC/MS (**Figure 44** for HPLC chromatogram and MS results). S-*t*But-NZ was isolated in excellent purity and an overall yield of 98 mg (26.5%).

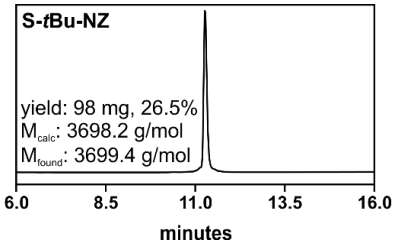


Figure 44: HPLC chromatogram for S-*t*Bu-NZ with yield and obtained mass (mass was determined by LC/MS).

4.4.5. Native chemical ligation

$MjCM_{1\text{-}21}\text{-}NHNH_2$ variants were ligated to *S*-tBu-NZ in a previously described one-pot, two-step native chemical ligation with slight modification.[239] In a

5 mL glass vial, $MjCM_{1-21}$-$NHNH_2$ variants (~2 mM. 1 eq.) and s-tBu-NZ (1.2 eq.) were dissolved in 700 µL 0.2 M Na_2HPO_4 buffer containing 6 M guanidine hydrochloride at pH 3.0. The resulting solution was cooled to -10 °C. Then, 70 µL of $NaNO_2$ (200 mM in H_2O) was added dropwise and the reaction mixture was stirred for 20 min. Afterwards, 700 µL of 0.2 M Na_2HPO_4 buffer containing 6 M guanidine hydrochloride and 200 mM 4-mercaptophenylacetic acid (MPAA) at pH 7.0 were added and the pH of the mixed solution was adjusted to 7.0 with NaOH (2 M in H_2O). The glass vial was then sealed with Teflon tape, and the reaction mixture stirred for 48 – 72 hours at room temperature, monitoring the progress of the reaction by LC/MS. Upon complete conversion, 200 µL of TCEP (1 M in neutral ligation buffer) was added. The reaction mixture was incubated for 10 min at room temperature, after which peptides and small molecules were separated by gravitational gel chromatography (NAP-20 column). Peptide-containing fractions were pooled and purified by preparative HPLC. Individual fractions were analyzed by analytical HPLC and pure fractions were pooled and lyophilized. The resulting $MjCM_{1-21}$-NZ variants were obtained in good yields and high purities (**Table 4** and **Figure 38**). Identity of the peptides was confirmed by HI-RES mass spectroscopy or LC/MS. After lyophilization the peptides were dissolved in 700 µL 10 mM TFA and stored at -20 °C.

4.4.6. Molecular cloning

E. coli strains KA12[104] and XL-1Blue (Stratagene) were used for cloning purposes and KA13, a KA12 derivative, protein production. Nucleic acid manipulations were performed following standard procedures. Plasmids were purified from 5 mL overnight cultures with JETquick kits on a miniprep scale.

DNA concentrations were measured with a NanoDrop spectrophotometer. DNA sequencing was performed by Microsynth (Balgach, Switzerland) using the flanking primers T7 and T7term. All primers were synthesized and purified by Microsynth and are listed in **Table 5**. All plasmids used in this work have been described before.[166]

Table 4: Yields and masses obtained after native chemical ligation of foldamers to the invariant leucine zipper portion.

MjCM$_{1-21}$.NZ variant	Yield (mg)	Yield (%)	Mass expected (Da)	Mass found (Da)
MjCM$_{1-21}$-NZ-R9A	6.2	51.2	6060.47	6060.46
1	4.7	33.6	6074.48	6074.45
2	5.8	47.3	6130.54	6130.52
3	5.7	45.4	5976.42	5976.40
4	9.3	38.9	6130.54	6130.6[a)]
5	6.5	50.5	6130.54	6130.6[a)]
6	8.5	47.8	6130.54	6130.6[a)]
7	5.1	41.6	6130.54	6130.54

[a)] Masses were determined by LC/MS but no high resolution masses were obtained.

Table 5: Primers used in this work. Nucleobases in bold show mutations from the original sequence.

Primer	DNA sequence (5' → 3')
T7	TAA TAC GAC TCA CTA TAG G
T7term	AAG ACC CGT TTA GAG GCC CCA A
R9A_fw	GCT AGC TGA AAT T**GC G**AA GAA GAT TGA TGA GA TTG AC
R9A_rv	CTC ATC AAT CTT CTT **CGC** AAT TTC AGC TAG CTT CCC

4.4.7. Site-directed mutagenesis

The knock-out mutant of the one-helix fragment, $MjCM_{1-21}$-NZ-R9A, was prepared by site-directed mutagenesis. A gene containing the desired mutation (R9A) was constructed by standard overlap extension PCR from the isolated plasmids with the primer pairs listed in **Table 5**. Full length genes were assembled using the primers T7 and T7term. The linear PCR products and a pmG211 vector containing a stuffer fragment were digested using NdeI and XhoI (New England Biolabs), gel purified, and subsequently ligated using T4 DNA ligase (New England Biolabs). The resulting plasmids were transformed into XL1-Blue competent cells and the mixture was plated on LB agar containing 150 µg/mL ampicillin. Two clones were randomly picked and sequenced to confirm the desired mutation in the gene. An overnight culture with clones having the correct sequence was grown, the plasmid purified, and transformed into *E. coli* KA 13 cells. The transformation mixture was again plated onto LB agar containing 150 µg/mL ampicillin and single colonies were used for protein production.

4.4.8. Protein purification

Flasks containing 750 mL LB-medium with 150 µg/mL ampicillin were inoculated with 750 µL of a densely grown overnight culture of KA13 cells harboring a plasmid encoding the appropriate plasmid (pMG211-hdC3[104] or pMG-$MjCM_{1-21}$-NZ-R9A). The cells were incubated at 37 °C and 230 rpm until an OD_{600} of 0.6 was reached. Gene expression was induced with IPTG at a final concentration of 0.25 mM and incubation was continued for 20 hours at 37 °C, after which cells were harvested by centrifugation. The cell pellet was

resuspended in 25 mL PBS buffer and incubated in the presence of lysozyme (1 mg/mL) and DNAse I (0.1 mg/mL) for 30 min at 4 °C and subsequently lysed by sonication (0.5 s cycles, amplitude 60, 5 x 1 min with a 1 min break between cycles). After removing cell debris by centrifugation at 14,000 g and 4 °C for 30 min, the cleared lysate was purified by Ni-NTA chromatography (Sigma). His-tagged proteins were eluted with PBS buffer containing 250 mM imidazole. Protein-containing fractions, identified by SDS-PAGE, were pooled and further purified by preparative HPLC on a Nucleosil 100-7 C18 column (Machery-Nagel) with a gradient of 5 - 60% acetonitrile (0.08% TFA) in water (0.1% TFA). Pure fractions containing the desired protein were pooled, lyophilized, and the resulting protein dissolved in 10 mM TFA and stored at -20 °C. Protein purity was assessed by HPLC chromatography and protein masses were confirmed by mass spectrometry (**Figure 45**). Concentrations were determined by measuring the absorbance at 280 nm on a NanoDrop spectrophotometer with calculated extinction coefficients ($MjCM_{1-21}$-NZ-R9A: 5,500 $M^{-1}cm^{-1}$; CZ-$MjCM_{22-93}$: 9,970 $M^{-1}cm^{-1}$). Yields were 7 mg for $MjCM_{1-21}$-NZ-R9A and 11 mg CZ-$MjCM_{22-93}$, respectively, per liter of culture.

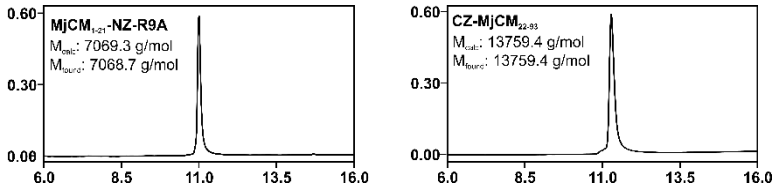


Figure 45: HPLC chromatograms and obtained masses for $MjCM_{1-21}$-NZ-R9A and CZ-$MjCM_{22-93}$.

4.4.9. CD spectroscopy

Far-UV CD spectra of protein samples were recorded at 25 °C on an Aviv 202 spectropolarometer from 200 to 260 nm in 1 nm steps with an averaging time

of 5 s in PBS pH 6.5 at 10 µM protein concentration. Four scans were averaged and a PBS blank was subtracted. Thermal denaturation curves for $MjCM_{1-21}$-NZ-M22C, **2**, and **3** were obtained by measuring ellipticity of a 10 µM peptide solution at 222 nm between 15 and 90 °C. After a 5 minute initial equilibration at 15 °C, the sample was heated at a rate of 5 °C/min in 0.5 °C steps with 1 min equilibration time and 1 min signal averaging time every point.

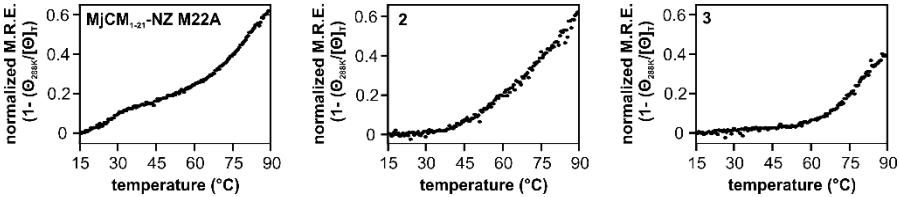


Figure 46: Thermodenaturation curves for the variants, recorded at 0.5 °C steps with equilibration times and measuring time of 1 min

4.4.10. Size-exclusion FPLC

Size-exclusion FPLC was performed on a Waters HPLC system equipped with a Superdex 75 HR 10/30 column and isocratic elution with 10 mM phosphate buffer (pH 6.5) containing 150 mM NaCl at a flow rate of 600 µL/min. A total of 50 µL containing $MjCM_{1-21}$-NZ variants, the three-helix segment, or equimolar mixtures (all at 50 µM) was injected. The obtained chromatograms are shown in **Figure 47**.

4.4.11. Kinetic assay

Kinetic assays were performed by measuring the consumption of chorismate at either 274 nm (ε_{274} 2630 $M^{-1}cm^{-1}$) or 310 nm (ε_{310} 370 $M^{-1}cm^{-1}$) with a UV-VIS spectrophotometer (Perkin-Elmer Lambda 35) according to

previously published procedures.[104] All measurements were conducted at 25 °C in PBS buffer at pH 6.5 in the presence of 0.1 mg/mL BSA. Kinetic measurements were carried out at enzyme concentrations ranging from 0.2 µM to 2 µM, depending on the activity of the MjCM$_{1\text{-}21}$-NZ variant. Initial velocities were determined for a given CZ-MjCM$_{22\text{-}93}$ concentration with 10 µM MjCM$_{1\text{-}21}$-NZ variant present. hdCM was pre-equilibrated in a 25x stock solution in the presence of 1 mg/mL BSA in PBS pH 6.5 in a glass vial overnight prior to measurement. For measurement, the premix was added to the temperature-equilibrated cuvette containing buffer and 50 to 2000 µM chorismate. The average initial velocities that were obtained from at least three independent measurements are depicted in **Figure 40**. Data for MjCM$_{1\text{-}21}$-NZ-R9A not shown.

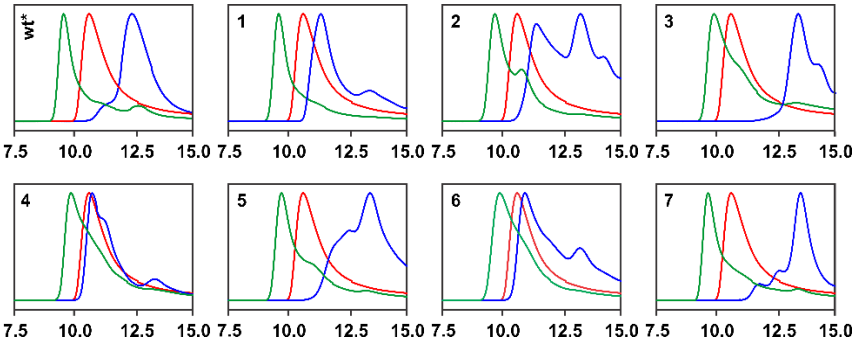


Figure 47: Comparison of size-exclusion FPLC chromatograms for MjCM$_{1\text{-}21}$-NZ-M22C (wt*) and foldameric variants. The plots show an overlay of traces obtained from one-helix segments (blue), CZ-MjCM$_{1\text{-}21}$-NZ, and equimolar mixtures of both (green, all 50 µM).

4.4.12. K_d measurements

Dissociation constants for the MjCM$_{1\text{-}21}$-NZ variants were determined by two different methods depending on the activity of the constructs. For

MjCM$_{1\text{-}21}$-NZ-M22C and foldamers **1** – **4**, K_d values were calculated from a plot of observed initial velocities for fixed concentrations of CZ-MjCM$_{22\text{-}93}$ (0.1 µM for the parent variant, **2**, and **3**; 0.2 µM for **1** and **4**) and increasing concentrations of the one-helix fragments. The hyperbolic binding curves were fit nonlinearly to the dimeric binding model[104] according to equation (*1*), in which c is a constant relating the concentration of the dimer hdCM to the activity, [MjCM$_{1\text{-}21}$-NZ] and [CZ-MjCM$_{22\text{-}93}$] the concentrations of the fragments, and K_d the dissociation constant.

$$\frac{v_{obs}}{[\text{CZ-MjCM}_{22\text{-}93}]} = c - \left\{ \frac{[\text{MjCM}_{1\text{-}21}\text{-NZ}] + [\text{CZ-MjCM}_{22\text{-}93}] + K_d}{2} - \sqrt{\frac{([\text{MjCM}_{1\text{-}21}\text{-NZ}] + [\text{CZ-MjCM}_{22\text{-}93}] + K_d)^2}{4} - [\text{MjCM}_{1\text{-}21}\text{-NZ}][\text{CZ-MjCM}_{22\text{-}93}]} \right\} \quad (1)$$

Dissociation constants for relatively inactive foldamer variants **5** – **7** were determined indirectly by monitoring inhibition of the formation of an active complex with **2**. IC$_{50}$ values were obtained by plotting the resulting initial velocities for fixed concentrations of **2** and CZ-MjCM$_{22\text{-}93}$ (both 1 µM) as a function of increasing amounts of MjCM$_{1\text{-}21}$-NZ fragments. To determine dissociation constants from the observed IC$_{50}$ values, we used the Cheng-Prusoff equation (*2*),[246] in which K_d* is the K_d for the desired foldamer, [MjCM$_{1\text{-}21}$-NZ] the concentration of variant **2**, and K_d the dissociation constant for **2.** Plots shown in **Figure 41** are the results from at least two independent measurements.

$$K_d^* = \frac{\text{IC}_{50}}{1 + \frac{[\text{MjCM}_{1-21}\text{-NZ}]}{K_d}} \quad (2)$$

5. Toward the directed evolution of hybrid enzymes

5.1. Introduction

Over billions of years, nature has optimized the primary sequence of polypeptides to catalyze the plethora of biological reactions with unmatched rate accelerations and selectivities.[65] Since their discovery 150 years ago, these biocatalysts have been a source of inspiration for chemists and biologists alike. Although in-depth mechanistic, structural, and computational studies have provided a qualitative understanding of how enzymes work, the complex and subtle structure-function relationship in protein scaffolds generally prevents a quantitative accounting of enzyme efficiency (Chapter 1.3).[41, 43, 47] This incomplete understanding of enzyme catalysis is reflected in the limited performance of protein catalysts designed by man.[63] Even for reactions that require less than a handful of accurately-positioned residues for catalysis, such protein engineering endeavors have afforded artificial enzymes with rate accelerations that pale in comparison to their natural counterparts.[247-248] In order to boost the performance of artificial enzymes one can adapt the Darwinian algorithm - nature's solution to the design problem – for use in the laboratory (**Figure 48**).[249]

The activities and selectivities of natural and artificial biocatalysts can be altered or improved through iterative cycles of mutagenesis and amplification

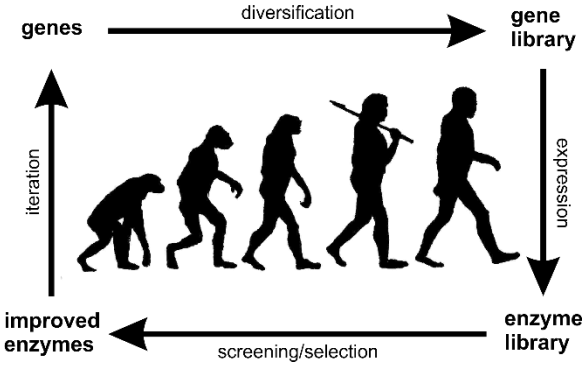


Figure 48: The iterative cycle of laboratory evolution. Genes are first diversified by DNA manipulation techniques to yield a gene library. Expression of the gene library results in an ensemble of proteins. In a selection or screening approach improved variants are identified and separated from the inactive. Lastly, the genetic material for the improved mutants is used as a starting point for a new cycle (adapted from reference [249].

of promising variants.[211] The advent of efficient protocols for DNA manipulation to create partially randomized gene libraries and the possibility of evolving a single biomolecule instead of a whole organism have reduced the time scale for protein evolution to weeks in the laboratory instead of millennia in nature.[250-252] The greatest challenge facing the directed evolution of enzymes is identifying the relative small number of variants that display the desired characteristic and separate them from the large pool of molecules that do not. Toward this end, protein libraries are either evaluated by testing each member individually (screening) or by simultaneously assessing the fitness of all mutants under conditions that allow only desirable clones to appear (selection).[211, 253] While screening approaches are typically less prone to accumulate false positives or other artefacts, selection systems usually allow for a larger number of mutants to be analyzed per round. Both strategies have been successfully applied to the improvement of natural and promiscuous

activities of enzymes, as well as for tailoring specific properties of biocatalysts including stability, substrate scope, or tolerance to organic cosolvents.[211, 253] A notable example is the directed evolution of a transaminase that catalyzes the final synthetic step leading to sitagliptin, a blockbuster diabetes drug (**Figure 49**).[254] In this seminal study, the enzyme that was obtained after 11 rounds of mutagenesis displayed a roughly ~40,000 fold higher activity than the starting variant together with an 99.95% e.e. for the (R)-configured amine. When compared to the previously established metal-catalyzed transformation, the biocatalytic process increases the overall yield by 13% and the productivity by 53%.[255]

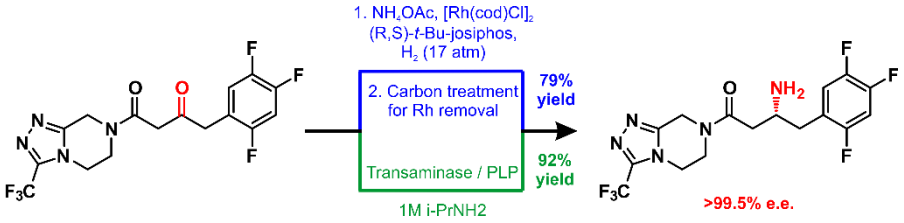


Figure 49: Comparison of the chemical (blue) and the biotechnological (green) process for the synthesis of sitagliptin from a common precursor. Abbreviations: cod = 1,5-cyclooctadiene, josiphos = a chiral diphosphine ligand, and PLP = pyridoxal-5'-phosphate (adapted from reference [255]).

As shown by the sitagliptin transaminase example, tailoring enzymatic properties by the evolutionary algorithm is a manageable task if a functional catalytic machinery is already in place in a proteinogenic scaffold.[111] However, creating non-natural activities in a protein scaffold by directed evolution has proven far more challenging as sequence space is dominated by non-catalytic polypeptides (for example, only an estimated 1 out of 10^{23} sequences are chorismate mutases).[256] Consequently, robust strategies to introduce even

low catalytic activities into a protein scaffold where none are naturally present are crucial to generate evolutionary stepping stones. In this respect, Chapters 2 and 3 have already highlighted the design and evolution of artificial metalloenzymes.[101, 110] Other possibilities involve the incorporation of non-natural amino acids that carry unique functionalities or computationally-assisted redesign of protein scaffolds. While functions that have been generated by the former approach have rarely been refined, computationally-designed enzymes have been extensively optimized by laboratory evolution.[111] Active catalysts obtained from design algorithms (Chapter 1.3.3) are usually equipped with a minimal catalytic machinery and give rise to modest rate accelerations (k_{cat}/k_{uncat} = 10^2-10^5).[257] In recent years, however, our laboratory has demonstrated that directed evolution provides an effective means to boost the catalytic prowess of these first generation variants to levels that rival those displayed by natural enzymes.[63] For example, after 17 rounds of mutagenesis and screening the turnover number of an enzyme that catalyzes a Kemp elimination – a proton transfer reaction unknown to nature - was improved from 0.7 s^{-1} to 700 s^{-1}.[64] The latter k_{cat} value is comparable to the rate constants measured for triosephosphate isomerase (430 s^{-1}), which converts dihydroxyacetone phosphate to R-glyceraldehyde-3-phosphate, a reaction that like the Kemp elimination proceeds through a proton abstraction step.[43]

The previous chapter highlighted a novel strategy to expand the scope of enzyme catalysis, namely the use of foldamer prostheses. Foldamers are versatile, non-natural oligomers that could potentially confer unique control over biological functions to proteins (Chapter 4.1.2).[223, 235] However, in order to take full advantage of their superior properties it will be crucial to identify strategies for readily optimizing their properties. For example, whether this

new class of artificial catalyst can be improved by a directed evolution approach like natural biopolymers is an interesting question and, if so, to what extent. In this chapter, a robust and practical screening system is presented that promises to facilitate the optimization of foldamer-containing catalysts. It may serve as the basis to explore the evolution of these non-natural hybrid systems.

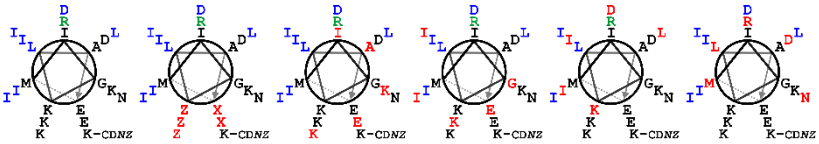
5.2. Results

As demonstrated in the previous chapter, the introduction of β^3-amino acids into the N-terminal helix, $MjCM_{1\text{-}21}$, of a heterodimeric chorismate mutase afforded hybrid catalysts. Notably, the activities for foldamer prostheses depended on the specific α/β patterning employed. More specifically, when the nonstandard building blocks were clustered at the solvent exposed face of $MjCM_{1\text{-}21}$, highly efficient hybrid enzymes were obtained (**Figure 50**, foldamers **3** and **4**). In contrast, insertion of non-natural backbone units at the dimer interface yielded variants with inferior catalytic parameters, in part because they form the anticipated heterodimer less efficiently (as indicated by elevated dissociation constants, **5** and **6**). In order to compensate for deleterious changes at the helical interface in the latter, we wondered whether engineering of the CZ-$MjCM_{22\text{-}93}$ binding partner would be a feasible strategy to recover activities for these inefficient hybrid catalysts.

5.2.1. Establishing a screening system

For the anticipated engineering approach, a reliable method to identify improved CZ-$MjCM_{22\text{-}93}$ variants is of utmost importance. In principle, a powerful selection system based on the complementation of chorismate

mutase activity in an *E. coli* strain lacking the natural enzyme is available in our lab.[58] However, the inability of the ribosome to introduce multiple β-amino acids into growing peptide chains severely complicates the use of an in vivo selection.[258] Attempts to circumvent this limitation using cell-penetrating tags[259] to shuttle MjCM$_{1\text{-}21}$-NZ variants across the cell membrane have also proven ineffective. As a consequence, we have to rely on a screening approach for identifying improved foldamer-containing chorismate mutases.



Variant	MjCM$_{1\text{-}21}$-NZ-M22C	3	4	5	6	7
AA-pattern	all α	αααβαααβ	αααβ	αααβ	αααβ	αααβ
k_{cat} (s^{-1})	6.87 ± 0.03	5.65 ± 0.14	3.14 ± 0.07	0.60 ± 0.06	0.28 ± 0.01	-
K_m (M^{-1})	150 ± 2	410 ± 30	650 ± 40	3700 ± 470	2000 ± 150	-
k_{cat}/K_m (M^{-1} s^{-1})	44,600 ± 700	13,800 ± 900	4800 ± 230	160 ± 9	140 ± 7	3.6 ± 0,2
K_d (μM)	0.26 ± 0.06	0.76 ± 0.13	3.49 ± 0.31	70 ± 12.5	7.9 ± 0.8	0.43 ± 0.04

Figure 50: Helical wheel representation of the parent one-helix fragment, MjCM$_{1\text{-}21}$-NZ-M22C and four foldamer prostheses. Amino acid patterns, kinetic parameters and affinity constants are taken from **Table 3** in Chapter 4. Color code is the same as in **Figure 36**.

Chorismate mutase activity can be followed in parallel fashion by using UV-permeable multi-well plates. We obtained a linear response for chorismate consumption when varying amounts of CZ-MjCM$_{22\text{-}93}$ (0.5 – 10 nM) were mixed with a constant concentration of MjCM$_{1\text{-}21}$-NZ (5 μM, **Figure 51A**). Because compounds in cell lysates strongly absorb at 274 nm – the preferred wavelength to measure chorismate mutase activity – the three-helix fragments need to be separated from these molecules. We, therefore, tested purification of wild-type CZ-MjCM$_{22\text{-}93}$ and a knock-out mutant, CZ-jCM$_{22\text{-}93}$R40A, by parallel His-tag affinity chromatography after production in 24-well plates. Concentrating the fractions obtained upon acidic elution (50 mM TFA, see

Materials and Methods for details) afforded both variants in reasonable purity (**Figure 51B**) and concentrations of ca. 50 μM (judged by Bradford assay). When 1000-fold dilutions of the protein-containing solutions were added to $MjCM_{1-21}$-NZ (5 μM), we observed a 30-fold higher rate of chorismate consumption for the wild-type enzyme compared to the knock-out variant (**Figure 51C**). Notably, this ratio is an order of magnitude smaller than that determined from kinetic measurements with HPLC-purified proteins. Nevertheless, the results from the parallel purification augur well for detecting activities of less efficient foldamer prostheses, as the concentration of the purified three-helix fragments can be increased in the assay.

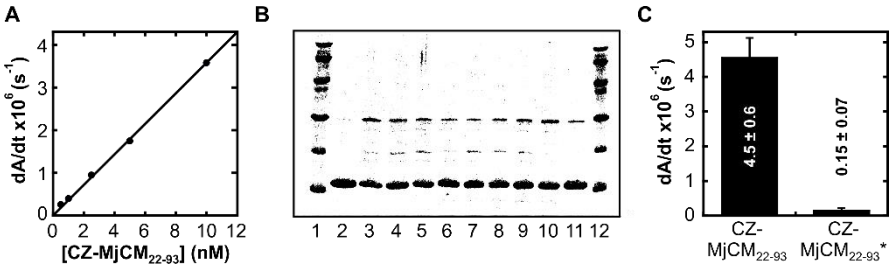


Figure 51: Establishing of a screening assay for foldamer-containing chorismate mutases. **A**: Observed velocities with varying CZ-$MjCM_{22-93}$ concentrations with constant amounts of $MjCM_{1-21}$-NZ (5 μM). **B**: SDS-PAGE analysis of 8 randomly chosen CZ-$MjCM_{22-93}$ variants (lanes 2-10) from parallel His-tag chromatography purification of protein samples produced in 24-well plate formate. Lanes 1 and 12: Low-molecular weight markers (the lowest bands corresponds to 14.4 kDa, the size expected for the three-helix fragment is 13.8 kDa). Lanes 2 and 11: HPLC-purified CZ-$MjCM_{22-93}$ variants. **C**: Averaged rates for the wild-type and the R40A knock-out mutant (depicted as CZ-$MjCM_{22-93}$*) of the three-helix fragment.

5.2.2. Library construction

With a robust screening method in hand, three CZ-$MjCM_{22-93}$ libraries were designed. Compared to high-throughput selection systems or screening assays,

the need to partially purify the three-helix segments drastically decreases the number of clones that can be tested in one round of directed evolution. As a consequence, we chose to specifically target residues in CZ-MjCM$_{22\text{-}93}$ that are important for hydrophobic packing of hdCM and preferably not in close proximity to the active site (**Figure 52A**). We surmised that conservative changes to hydrophobic amino acids at these positions would be best suited to enhance activities of inefficient MjCM$_{1\text{-}21}$-NZ variants by compensating for backbone mutations at the helical interface in these designs. From a model of hdCM, we identified eight positions in the three-helix fragment that we randomized with degenerate codons, allowing only three new amino acids in addition to the original residue (**Figure 52B**). A total of 3 libraries was constructed, with libraries 1 and 2 harboring four mutations (256 members) and library 3 combining all eight (65,536 members).

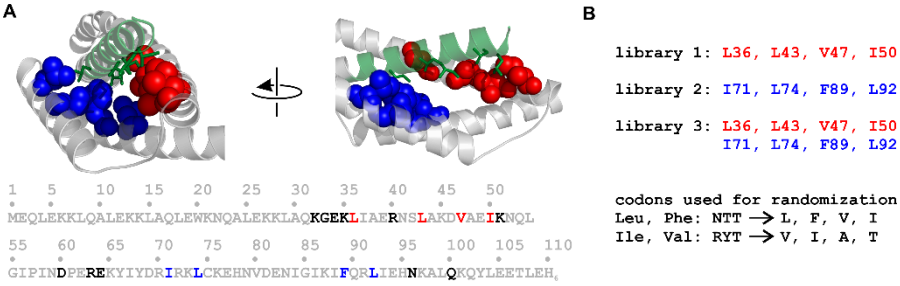


Figure 52: Principles of library construction for CZ-MjCM$_{22\text{-}93}$. **A**: Location of the libraries in a fold-it model of wild-type hdCM and in the sequence of the three helix fragment. The one-helix segment is shown in green, residues important for hydrophobic packing in CZ-MjCM$_{22\text{-}93}$ in blue and red. Active site residues and the engineered linker sequence (residues 32-35) are highlighted in black **B**: Libraries of CZ-MjCM$_{22\text{-}93}$ with the degenerate codons used in this study. The color code for the mutations of the library is the same as in the fold-it model. In the randomized codons, N specifies equimolar mixtures of all four bases, R only purine bases (A,G), and Y only pyrimidine bases (C,T).

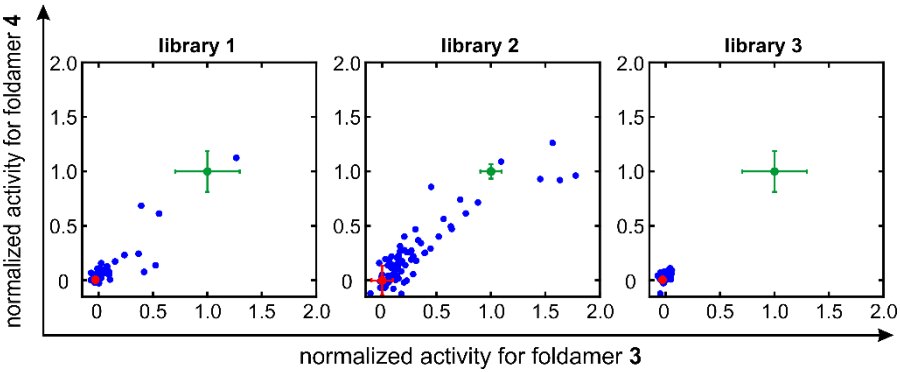


Figure 53: Results from screening libraries 1 – 3 with foldamer prostheses **3** and **4**. Plots show the activity of each tested clone (blue) normalized to the average value obtained for CZ-MjCM$_{22\text{-}93}$ (green, four measurements). The average rate from four measurements of chorismate consumption for the knock-out variant is highlighted in red.

5.2.3. Library screening

Production of a total of 166 clones (+16 controls) in 96-well formate afforded variants in concentrations of roughly 5 µM after purification. In order to evaluate the performance of the individual library members, we measured chorismate consumption by adding 1:50 solutions of the obtained CZ-MjCM$_{22\text{-}93}$ variants to α/β-peptides **3** or **4** (5 µM each). By normalizing the measured activities for individual variants against the value obtained for the wild-type three-helix fragment, the fitness of libraries can be assessed (**Figure 53**). In library 1, only one out of 44 clones displayed activities similar to the parent CZ-MjCM$_{22\text{-}93}$. In contrast, for the second library (88 variants were tested) multiple clones were identified that display elevated or similar activities compared to the wild-type. The number of library members that performed significantly better than the knock-out mutant was also increased for library 2

compared to library 1. Clearly, the mutational burden that results from randomizing all eight positions in library 3 was too heavy, as not a single clone catalyzed chorismate consumption at rates that are significantly higher than those obtained for the inactive variant CZ-MjCM$_{22\text{-}93}$-R40A.

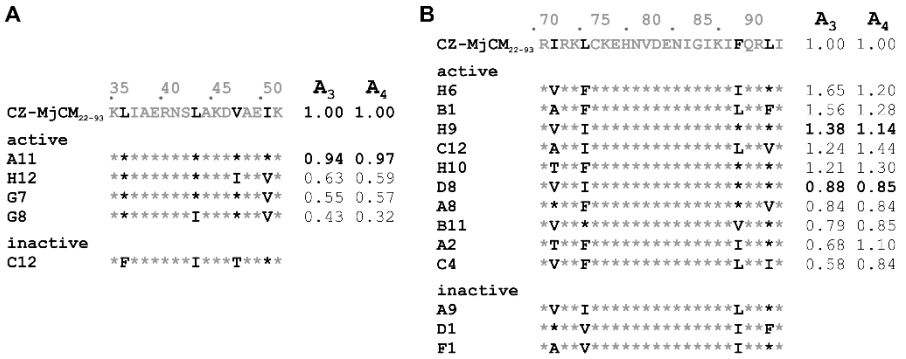


Figure 54: Sequence alignment for active and inactive variants from library 1 (**A**) and library 2 (**B**). Activities in the presence of foldamer prostheses **3** and **4** are shown..

5.2.4. Sequencing

A subset of 14 active and a small number of inactive variants from library 1 and 2 were sequenced to explore the cause for the differences in fitness of these two protein ensembles. Notably, despite the conservative substitutions we introduced, changes at Leu36 appear to abrogate activities in the first library (**Figure 54A**). A possible explanation for this trend could be the proximity of this residue to the leucine zipper feature (**Figure 52A**). Given that the linker sequence that directly follows Leu36 was optimized by directed evolution for the parent heterodimer, it is conceivable that changes at this position have a negative effect on helix-bundle formation. Other residues in library 1 tolerate

mutations, although the addition or removal of a single methyl group reduces activity roughly two-fold. The re-isolation of the wild-type sequence (clone A11) as the most active member indicates that this library is unlikely to contain mutations that significantly improve the performance of chimeric chorismate mutases.

Trends for library 2 are less pronounce even though a larger number of variants was sequenced (**Figure 54B**). Unlike library 1, all positions subjected to randomization tolerate mutation without reducing activity relative to wild-type CZ-MjCM$_{22\text{-}93}$. While the original residues for F89 and L92 are still present to a significant extent (10 out of 20) in the active clones, positions 70 and 74 are mutated in 18 out of 20 cases. The enrichment for bulkier side chains at L74 (predominantly F, less frequently I) is notable, because the corresponding L74F mutation for a homodimeric chorismate mutase from *M. jannaschii* has been previously identified to have a beneficial effect when large portions of this enzyme were randomized.[256] Intriguingly, the same study has also shown that the segment containing the mutations of library 2 was more sensitive to changes than the helix we targeted with library 1. The apparent difference for chimeric hdCMs suggests that the hydrophobic packing for the two helices that were partially randomized in library 2 is not yet optimized. Consequently, this library should be an excellent starting point for selecting improved foldameric chorismate mutases. Consistent with this hypothesis, we identified multiple clones that display activities that rival or outperform the wild-type α/β-hdCM. However, the obtained differences are not necessarily significant, as indicated by two selected variants that share the same DNA sequence (H9 and DA8, **Figure 54B**) but displayed reaction rates that differ by ≈50%. Nevertheless, it will be interesting to purify three-helix segments that display higher activities

in the screen and determine whether these clones indeed show improvements over the wild-type CZ-MjCM$_{22\text{-}93}$ in a more detailed kinetic characterization. Ultimately, these variants could serve as starting points for the next round of directed evolution.

5.3. Discussion

In this work, we have developed a robust screening system that should facilitate the directed evolution of foldamer-containing chorismate mutases. Libraries of CZ-MjCM$_{22\text{-}93}$ variants can be purified in a medium-throughput fashion and tested for improvements against multiple foldamer prostheses. In this proof-of-concept study we have evaluated more than 150 clones from three libraries that harbor different mutations at the heterodimer interface. Our results highlight the different levels of fitness for these protein ensembles and identify three-helix fragments with similar or possibly improved activities with respect to the wild-type CZ-MjCM$_{22\text{-}93}$. In the future, improved hybrid enzymes can readily be identified by combining purified CZ-MjCM$_{22\text{-}93}$ variants with all available α/β-surrogates. It is conceivable that mutations at the hydrophobic core of the three-helix segment will be better suited to compensate for deleterious changes at the helical interface that were introduced by the nonstandard building blocks in less proficient foldameric MjCM$_{1\text{-}21}$-NZ variants, such as **5** and **6**.

Compared to other strategies that expand the scope of enzyme catalysis, the introduction of foldamers into an existing enzymatic scaffold affords hybrid catalysts with the catalytic machinery of the parent enzyme largely intact.[63] Therefore, the directed evolution of foldamer-containing enzymes should

proceed more readily than for other artificial catalysts that often require substantial reconstruction of their minimal active sites.[260] Using chorismate mutase as a model system, the findings of an in-depth evolutionary study may yield further insights into how α/β-peptides can be reliably incorporated into a wide variety of enzymatic scaffolds. Demonstrating that these hybrid catalysts are evolvable is crucial in order to take full advantage of the appealing properties of foldamers, including their higher stability and increased number of possible secondary structures.

5.4. Materials and Methods

5.4.1. Construction of CZ-MjCM$_{22\text{-}93}$-R40A

A knock-out mutant for CZ-MjCM$_{22\text{-}93}$ was prepared by site-directed mutagenesis. A gene containing the desired mutations (R40A) was constructed by standard overlap extension PCR from the isolated plasmids with the primer pairs listed in **Table 6**. Full length genes were assembled using the primers T7 and T7term. The linear PCR products and a pmG211[166] vector containing a stuffer fragment were digested using NdeI and XhoI (New England Biolabs), gel purified and subsequently ligated using T4 DNA ligase (New England Biolabs). The resulting plasmids were transformed into XL1-Blue competent cells and the mixture was plated on LB agar containing 150 µg/mL ampicillin. Two clones were randomly picked and sequenced to confirm the desired mutation in the gene. An overnight culture with clones having the correct sequence was grown, the plasmid purified, and transformed into *E. coli* KA 13 cells. The

transformation mixture was again plated onto LB agar containing 150 μg/mL ampicillin and single colonies were used for protein production.

5.4.2. Library construction

Genes for the three libraries used in this work were constructed by site-directed mutagenesis with the primer pairs listed in **Table 6**. The residues targeted for mutagenesis were encoded with degenerate NTT and RYT codons, where N specifies equimolar mixtures of all four bases, R only purine bases (A,G), and Y only pyrimidine bases (C,T). Full length genes were assembled from the appropriate fragments by overlap extension PCR using the flanking primers T7 and T7term. The linear libraries and a pmG211 vector containing a stuffer fragment were digested with NdeI and XhoI (New England Biolabs) for 2 hours. The resulting fragments were gel purified and ligated with T4 DNA ligase (New England Biolabs) using a 4 fold excess of insert over the acceptor plasmid. Products from the ligation were desalted using a PCR extraction kit and transformed into freshly prepared electro-competent XL1-blue *E. coli* cells. Immediately after electroporation, 900 μL SOC medium was added and the cells were allowed to recover at 37 °C for 60 min while shaking (230 rpm). Cells were subsequently pelleted by centrifugation (4000 g, 5 min), the SOC medium discarded, and the cells resuspended in 5 mL LB medium containing 150 μg/mL ampicillin. A total of 50 μL of 1:10 and 1:100 dilutions of these cell mixtures were plated on LB Amp plates to determine the transformation efficiency. The rest of the cells were incubated at 37 °C overnight in a shaker (230 rpm) and plasmids with the libraries were isolated by miniprep (typical concentrations of 200 - 600 ng/μL from 2.5 mL cell suspension). Library sizes were 5.5×10^5 (library 1), 2.6×10^5 (library 2), and 6.5×10^5 (library 3). The quality of the library

was assessed by DNA sequencing. For protein production the libraries were transformed into *E. coli* KA13 cells applying standard protocols.

Table 6: Table of primers used in this study. Codons that deviate from the wild-type sequence are highlighted in bold.

Primer	DNA sequence (5' → 3')
R40A_fw	GGC GAA AAG CTT ATT GCT GAA **GCA** AAT AGT TTA GCT AAG GAT GTA GC
R40A_rv	CAG CAA TAA GCT TTT CGC C
T7_fw	TAA TAC GAC TCA CTA TAG G
T7_rv	AAG ACC CGT TTA GAG GCC CCA A
Lib1_fw[a]	GCA CAG AAA GGC GAA AAG **NTT** ATT GCT GAA AGA AAT AGT **NTT** GCT AAG GAT **RYT** GCT GAG **RYT** AAA AAT CAG CTT GGT ATT CC
Lib1_rv	CTT TTC GCC TTT CTG TGC
Lib2_1_fw[a]	CCA GAA AGA GAA AAA TAT ATA TAC GAT AGA **RYT** AGA AAA **NTT** TGT AAA GAA CAT AAC G
Lib2_1_rv	TCT ATC GTA TAT ATA TTT TTC TCT TTC TGG
Lib2_2_fw[a]	GGA GGG CCT TAT TAT GCT CTA TAA **NAC** GTT GAA **NTA** TTT TAA TGC CAA TAT TTT CAT CAA CG
Lib2_2_rv	TGCTAGTTATTGCTCAGCGG

[a] in the randomized codons, N specifies equimolar mixtures of all four bases, R only purine bases (A,G), and Y only pyrimidine bases (C,T).

5.4.3. Protein production

CZ-MjCM$_{22\text{-}93}$-R40A was produced starting from single colonies of KA13 *E. coli* cells containing the appropriate plasmid under the same conditions as the wild-type (Chapter 4.4.8.). Yields after HPLC purification were 9 mg/L of culture.

5.4.4. Parallel protein purification

For parallel protein production, 7 mL (24-deep-well plates) or 1.5 mL (96-deep-well plates) LB medium containing 150 µg/mL ampicillin were inoculated with single colonies of *E. coli* KA13 carrying the appropriate plasmid. The cultures were incubated at 37 °C overnight while shaking (230 rpm) and gene expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG, final concentration 1 mM). The cultures were grown for 12 – 14 hours at 37 °C (230 rpm) before harvesting the cells by centrifugation. The supernatant was removed and the cell pellets stored overnight at -80 °C.

Parallel protein purification was carried out in Protino 96 Ni-IDA plates. In brief, cell pellets from the parallel production were thawed and resuspended in 1 mL of lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 7.0) containing 1 mg/mL lysozyme. The cells were incubated for 30 minutes at room temperature and lysis was achieved by two freeze-thaw cycles (-80 °C → room temperature). Cell debris was removed by centrifugation (45 min, 6000 rpm) and the supernatant was loaded onto Protino 96 Ni-IDA plates pre-equilibrated in lysis buffer. The wells were allowed to drain by gravity and the plates washed twice with 800 µL lysis buffer. The wells were washed an additional 3 times with water before proteins were eluted with 50 mM TFA. The flow-through was collected in a clean 96-deep-well plate and frozen at -80 °C. The liquid was removed by lyophilization, the dry protein dissolved in 150 µL 10 mM TFA in water, and the resulting solution transferred to 96-well round bottom plates. The plates containing the peptides were stored at -20 °C until further use.

5.4.5. Establishing a chorismate mutase assay in 96-well plate formate

Kinetic assays were performed by measuring the consumption of chorismate at 274 nm with a plate reader (Molecular Devices) at 25 °C in PBS buffer at pH 6.5 in the presence of 0.1 mg/mL BSA. To establish the screening conditions, reactions were carried out with 5 µM MjCM$_{1-21}$-NZ in the presence of 0.1 – 10 nM CZ-MjCM$_{22-93}$. hdCM was pre-equilibrated in a 10x stock solution in the presence of 1 mg/mL BSA in PBS pH 6.5 in 96-well round bottom plates for 30 minutes at room temperature. For measurement, 20 µL of the pre-equilibrated stock was added to a temperature-equilibrated 96-well UV plate containing buffer and 100 µM chorismate (180 µL). Initial velocities displayed in **Figure 51A** were obtained from linear regression of the obtained slopes.

5.4.6. Test screen with CZ-MjCM$_{22-93}$ and CZ-MjCM$_{22-93}$-R40A

Concentrations of protein fractions (from 24 well protein purification) were determined to be around 50 µM by Bradford assay. For preparation of premixes, 5 µL of a 1:10 dilution of the protein fractions were added to 45 µL PBS buffer containing 25 µM MjCM1-21-NZ and 1 mg/mL BSA. Premixes were equilibrated for 30 minutes at room temperature prior to transfer to a 96-well UV plate containing buffer and 200 µM chorismate. For the values displayed in **Figure 51C**, the initial velocities obtained with forty CZ-MjCM$_{22-93}$ clones and eight CZ-MjCM$_{22-93}$-R40A clones were averaged and corrected for chorismate consumption in the absence of enzyme.

5.4.7. Screening of libraries

Concentrations of protein fractions (from 96 well protein purification) were determined to be around 5 μM by Bradford assay. For preparation of premixes, 10 μL of the protein fractions were added to 20 μL PBS buffer containing 25 μM **3** or **4** and 1 mg/mL BSA. Premixes were equilibrated for 30 minutes at room temperature prior to transfer to a 96-well UV plate containing buffer and 100 μM chorismate. Initial velocities for library members were determined by linear regression of the obtained curves. The values displayed in **Figure 53** are corrected for the background rates.

6. Conclusions and Outlook

The design of protein catalysts that expand the natural repertoire of enzyme catalysis is a formidable challenge that we attempted to meet with the studies conducted in this thesis. In Chapter 2, for example, we appended a metathesis catalyst onto a small heat shock protein from *M. jannaschii* to create an artificial metalloenzyme for olefin metathesis. Although the protein provided water solubility and prevents aggregation of the Grubbs-Hoveyda-type complex, the hybrid enzyme, [Ru]MjHSP, displayed appreciable activities only under acidic conditions. Because the capsid structure of the heat shock protein dissociates at low pH we dismissed our initial plan to fill the luminal space of the protein cage via a ring-opening metathesis polymerization (ROMP). However, recent studies from other groups report that similar metathesis catalysts, when appended to proteins, display high turnover numbers (up to 375 were achieved for a ROMP of a water-soluble substrate) in the presence of high concentrations of KCl at neutral pH.[160, 162] If under the reported conditions comparable activities could be achieved in [Ru]MjHSP, filling the protein cage with a polymer becomes conceivable. Such experiments could create particles of well-defined size that could be of value to applications in material science. In separate efforts the high tolerance of Grubbs-Hoveyda-type catalysts to the gamut of functional groups present inside cells could pave the way toward screening or even selection systems with the aim to improve the catalytic properties of artificial metalloenzymes with metathesis activity.[167]

In Chapter 3, we identified simple tripeptides as versatile ligands for iridium-catalyzed transfer hydrogenations affording large ligand acceleration

effects. Notable properties of the peptide-iridium complex include broad substrate scope, sufficient robustness to catalyze the reaction over long periods of time, and formation even in the presence of high concentrations of BSA. With the only prerequisite for activity being an unmodified N-terminus and a glycine as penultimate amino acid, it should be possible to append this genetically-encodable tag to virtually any protein. The enzyme 4-oxalocrotonate tautomerase (4-OT)[261] is an appealing choice to enable stereoselective transfer hydrogenations with our iridium complex that otherwise only affords racemic mixtures of products. The N-terminus of 4-OT is embedded in the active site of this protein and placing the organometallic moiety within this chiral environment would be expected to enable enantioselective transformations.[137] Another possibility is taking advantage of the unselective characteristics of the iridium catalyst in productive reaction cascades.[199] For example, the N-terminus of the heat shock protein we utilized in Chapter 2 is located in the luminal space of this capsid-forming protein. If the tripeptide were appended to the interior of this cage-like structure, the protein could potentially serve as a nanoreactor for dynamic kinetic resolutions.[262] Co-encapsulation of a stereoselective amine oxidase, for example, could give rise to a powerful system that, due to the proximity of the two catalysts, allows for an effective synthesis of chiral amines from racemic mixtures.

Chapters 4 and 5 demonstrated the use of α/β-foldamers as prostheses to complete an otherwise inactive fragment of a heterodimeric chorismate mutase and suggest that the resulting hybrid catalysts can be evolved by evolutionary algorithms. Although further increasing the share of non-natural building blocks in chorismate mutase and creating foldamer-containing enzymes that surpass the performance displayed by the parent enzyme are valuable goals, it

will be more important to apply the design rules that were elucidated to expand this strategy to other scaffolds. In this regard, a computationally-designed hydrolase, MID1, is especially intriguing.[95] Its small size of only 46 amino acids and its simple architecture, a dimer of coiled-coils, makes it an ideal target for creating novel foldamers by solid phase peptide synthesis. Moreover, MID1 has not undergone an evolutionary refinement like hdCM and consequently closely reflects a primordial enzyme that is by no means optimized. Thus, it is less likely that the introduction of β^3-amino acids will entail detrimental changes to coiled-coil formation, increasing the chance of endowing this scaffold with the desirable properties of foldamers. Although solid phase peptide synthesis presumably allows only for the synthesis of small libraries of zinc-dependent hydrolases, a directed evolution approach for β-amino acid-containing MID1 is in principle conceivable. Moreover, improving the enzyme in the absence of non-natural building blocks by the means of traditional evolution could enable a fascinating comparison of evolvability of natural and artificial catalysts. Similar to studies that aim to elucidate whether a thermostable or a molten globule scaffold is a better starting point to afford highly active enzymes,[263] an evolutionary race between artificial and natural catalysts could answer the question whether nature has made the optimal choice in selecting α-amino acids as the sole backbone unit for enzymes.

The design of enzymes that display non-natural activities is a fascinating area that, over decades, has inspired chemists and biologists to devise a diverse set of strategies to create made-to-order catalysts. Until further advances in computation allow for a complete understanding of the complex and subtle structure-function relationship of enzymes, we rely on the Darwinian algorithm to improve the prowess of the designed catalysts. This thesis has attempted to

expand the scope of enzyme catalysis by chemical intuition and highlighted biological means to evolve the acquired functions.

7. Bibliography

[1] G. F. Joyce (2002), The antiquity of RNA-based evolution. *Nature*, 418:214.

[2] A. Lazcano, S. L. Miller (1996), The origin and early evolution of life: Prebiotic chemistry, the pre-RNA world, and time. *Cell*, 85:793.

[3] L. E. Orgel (2004), Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol. Biol.*, 39:99.

[4] W. Gilbert (1986), Origin of life - the RNA world. *Nature*, 319:618.

[5] G. F. Joyce (1989), RNA evolution and the origins of life. *Nature*, 338:217.

[6] A. M. Poole, D. C. Jeffares, D. Penny (1998), The path from the RNA world. *J. Mol. Evol.*, 46:1.

[7] F. H. C. Crick, S. Brenner, A. Klug, G. Pieczenik (1976), Speculation on Origin of Protein-Synthesis. *Origins Life Evol. Biosph.*, 7:389.

[8] D. S. Wilson, J. W. Szostak (1999), *In vitro* selection of functional nucleic acids. *Annu. Rev. Biochem.*, 68:611.

[9] R. R. Breaker (1997), *In vitro* selection of catalytic polynucleotides. *Chem. Rev.*, 97:371.

[10] E. Szathmary (1999), The origin of the genetic code - amino acids as cofactors in an RNA world. *Trends Genet.*, 15:223.

[11] C. R. Woese (1967) *The genetic code: The molecular basis for genetic expression*. Harper & Row, New York.

[12] H. F. Noller (2004), The driving force for molecular evolution of translation. *RNA*, 10:1833.

[13] W. Saenger (1984) *Principles of nucleic acid structure*. Springer, New York.

[14] A. M. Pyle (2002), Metal ions in the structure and function of RNA. *J. Biol. Inorg. Chem.*, 7:679.

[15] J. A. Doudna, T. R. Cech (2002), The chemical repertoire of natural ribozymes. *Nature*, 418:222.

[16] T. M. Rana, K. T. Jeang (1999), Biochemical and functional interactions between HIV-1 Tat protein and TAR RNA. *Arch. Biochem. Biophys.*, 365:175.

[17] B. J. Calnan, S. Biancalana, D. Hudson, A. D. Frankel (1991), Analysis of arginine-rich peptides from the Hiv Tat protein reveals unusual features of RNA-protein recognition. *Genes Dev.*, 5:201.

[18] J. D. Puglisi, R. Y. Tan, B. J. Calnan, A. D. Frankel, J. R. Williamson (1992), Conformation of the TAR RNA-arginine complex by NMR-spectroscopy. *Science*, 257:76.

[19] E. Szathmary, J. M. Smith (1997), From replicators to reproducers: the first major transitions leading to life. *J. Theor. Biol.*, 187:555.

[20] A. J. Hager, J. D. Pollard, J. W. Szostak (1996), Ribozymes: Aiming at RNA replication and protein synthesis. *Chem. Biol.*, 3:717.

[21] N. Lee, Y. Bessho, K. Wei, J. W. Szostak, H. Suga (2000), Ribozyme-catalyzed tRNA aminoacylation. *Nat Struct Biol*, 7:28.

[22] B. L. Zhang, T. R. Cech (1997), Peptide bond formation by *in vitro* selected ribozymes. *Nature*, 390:96.

[23] M. Di Giulio (2005), The origin of the genetic code: Theories and their relationships, a review. *Biosystems*, 80:175.

[24] A. L. Weber, S. L. Miller (1981), Reasons for the occurrence of the twenty coded protein amino acids. *J. Mol. Evol.*, 17:273.

[25] D. C. Jeffares, A. M. Poole, D. Penny (1998), Relics from the RNA world. *J. Mol. Evol.*, 46:18.

[26] T. R. Cech (2000), The ribosome is a ribozyme. *Science*, 289:878.

[27] T. R. Cech (2009), Crawling out of the RNA world. *Cell*, 136:599.

[28] D. N. Frank, N. R. Pace (1998), Ribonuclease P: Unity and diversity in a tRNA processing ribozyme. *Annu. Rev. Biochem.*, 67:153.

[29] F. Qiao, T. R. Cech (2008), Triple-helix structure in telomerase RNA contributes to catalysis. *Nat. Struct. Mol. Biol.*, 15:634.

[30] S. Valadlkhan (2007), The spliceosome: A ribozyme at heart? *Biol. Chem.*, 388:693.

[31] R. Saldanha, G. Mohr, M. Belfort, A. M. Lambowitz (1993), Group-I and Group-II Introns. *FASEB J.*, 7:15.

[32] R. T. Batey, M. B. Sagar, J. A. Doudna (2001), Structural and energetic analysis of RNA recognition by a universally conserved protein from the signal recognition particle. *J. Mol. Biol.*, 307:229.

[33] S. C. Walker, D. R. Engelke (2006), Ribonuclease P: The evolution of an ancient RNA enzyme. *Crit. Rev. Biochem. Mol. Biol.*, 41:77.

[34] J. J. Kim, A. F. Kilani, X. Y. Zhan, S. Altman, F. Y. Liu (1997), The protein cofactor allows the sequence of an RNase P ribozyme to diversify by maintaining the catalytically active structure of the enzyme. *RNA*, 3:613.

[35] J. Holzmann, P. Frank, E. Loffler, K. L. Bennett, C. Gerner, W. Rossmanith (2008), RNase P without RNA: Identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell*, 135:462.

[36] S. C. Walker, D. R. Engelke (2008), A protein-only Rnase P in human mitochondria. *Cell*, 135:412.

[37] R. J. Keenan, D. M. Freymann, R. M. Stroud, P. Walter (2001), The signal recognition particle. *Annu. Rev. Biochem.*, 70:755.

[38] D. Schuenemann, S. Gupta, F. Persello-Cartieaux, V. I. Klimyuk, J. D. G. Jones, L. Nussaume, N. E. Hoffman (1998), A novel signal recognition particle targets light-harvesting proteins to the thylakoid membranes. *Proc. Natl. Acad. Sci. U. S. A.*, 95:10312.

[39] T. Bugg (2012) *Introduction to enzyme and coenzyme chemistry*. Wiley, Hoboken, NJ.

[40] E. Fischer (1894), Einfluss der Configuration auf die Wirkung der Enzyme. *Ber. Dtsch. Chem. Ges.*, 27:

[41] S. J. Benkovic, S. Hammes-Schiffer (2003), A perspective on enzyme catalysis. *Science*, 301:1196.

[42] D. E. Koshland (1958), Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. U. S. A.*, 44:98.

[43] J. R. Knowles (1991), Enzyme catalysis: Not different, just better. *Nature*, 350:121.

[44] L. Pauling (1946), Molecular architecture and biological reactions. *Chem. Eng. News*, 24:1375.

[45] G. E. Lienhard (1973), Enzymatic catalysis and transition-state theory. *Science*, 180:149.

[46] R. Wolfenden, M. Snider, C. Ridgway, B. Miller (1999), The temperature dependence of enzyme rate enhancements. *J. Am. Chem. Soc.*, 121:7419.

[47] R. Wolfenden, M. J. Snider (2001), The depth of chemical time and the power of enzymes as catalysts. *Acc. Chem. Res.*, 34:938.

[48] K. M. Herrmann, L. M. Weaver (1999), The shikimate pathway. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 50:473.

[49] W. J. Guilford, S. D. Copley, J. R. Knowles (1987), On the mechanism of the chorismate mutase reaction. *J. Am. Chem. Soc.*, 109:5013.

[50] P. Kast, Y. B. Tewari, O. Wiest, D. Hilvert, K. N. Houk, R. N. Goldberg (1997), Thermodynamics of the conversion of chorismate to prephenate: Experimental results and theoretical predictions. *J. Phys. Chem. B*, 101:10976.

[51] S. G. Sogo, T. S. Widlanski, J. H. Hoare, C. E. Grimshaw, G. A. Berchtold, J. R. Knowles (1984), Stereochemistry of the rearrangement of chorismate to prephenate - chorismate mutase involves a chair transition-state. *J. Am. Chem. Soc.*, 106:2701.

[52] D. J. Gustin, P. Mattei, P. Kast, O. Wiest, L. Lee, W. W. Cleland, D. Hilvert (1999), Heavy atom isotope effects reveal a highly polarized transition state for chorismate mutase. *J. Am. Chem. Soc.*, 121:1756.

[53] A. Y. Lee, P. A. Karplus, B. Ganem, J. Clardy (1995), Atomic-structure of the buried catalytic pocket of *Escherichia coli* chorismate mutase. *J. Am. Chem. Soc.*, 117:3627.

[54] Y. M. Chook, H. M. Ke, W. N. Lipscomb (1993), Crystal-structures of the monofunctional chorismate mutase from *Bacillus subtilis* and its complex with a transition-state analog. *Proc. Natl. Acad. Sci. U. S. A.*, 90:8600.

[55] A. Kienhofer, P. Kast, D. Hilvert (2003), Selective stabilization of the chorismate mutase transition state by a positively charged hydrogen bond donor. *J. Am. Chem. Soc.*, 125:3206.

[56] D. Hilvert (2000), Critical analysis of antibody catalysis. *Annu. Rev. Biochem.*, 69:751.

[57] S. T. Cload, D. R. Liu, R. M. Pastor, P. G. Schultz (1996), Mutagenesis study of active site residues in chorismate mutase from Bacillus subtilis. *J. Am. Chem. Soc.*, 118:1787.

[58] P. Kast, M. AsifUllah, N. Jiang, D. Hilvert (1996), Exploring the active site of chorismate mutase by combinatorial mutagenesis and selection: The importance of electrostatic catalysis. *Proc. Natl. Acad. Sci. U. S. A.*, 93:5043.

[59] P. Kast, C. Grisostomi, I. A. Chen, S. L. Li, U. Krengel, Y. F. Xue, D. Hilvert (2000), A strategically positioned cation is crucial for efficient catalysis by chorismate mutase. *J. Biol. Chem.*, 275:36832.

[60] X. D. Zhang, X. H. Zhang, T. C. Bruice (2005), A definitive mechanism for chorismate mutase. *Biochemistry*, 44:10443.

[61] F. Claeyssens, K. E. Ranaghan, F. R. Manby, J. N. Harvey, A. J. Mulholland (2005), Multiple high-level QM/MM reaction paths demonstrate transition-state stabilization in chorismate mutase: Correlation of barrier height with transition-state stabilization. *Chem. Comm.*, 5068.

[62] S. D. Copley, J. R. Knowles (1987), The conformational equilibrium of chorismate in solution - implications for the mechanism of the nonenzymatic and the enzyme-catalyzed rearrangement of chorismate to prephenate. *J. Am. Chem. Soc.*, 109:5008.

[63] D. Hilvert (2013), Design of Protein Catalysts. *Annu. Rev. Biochem.*, 82:447.

[64] R. Blomberg, H. Kries, D. M. Pinkas, P. R. E. Mittl, M. G. Grutter, H. K. Privett, S. L. Mayo, D. Hilvert (2013), Precision is essential for efficient catalysis in an evolved Kemp eliminase. *Nature*, advance online publication:

[65] D. E. Metzler, C. M. Metzler (2001) *Biochemistry: The chemical reactions of living cells*. Harcourt/Academic Press, San Diego, CA.

[66] A. Ambrogelly, S. Palioura, D. Soll (2007), Natural expansion of the genetic code. *Nat. Chem. Biol.*, 3:29.

[67] N. M. Okeley, W. A. van der Donk (2000), Novel cofactors via post-translational modifications of enzyme active sites. *Chem. Biol.*, 7:R159.

[68] F. H. C. Crick (1967), The origin of genetic code. *Nature*, 213:119.

[69] F. H. Crick, S. Brenner, Watstobi.Rj, L. Barnett (1961), General nature of the genetic code for proteins. *Nature*, 192:1227.

[70] E. V. Koonin, A. S. Novozhilov (2009), Origin and evolution of the genetic code: The universal enigma. *IUBMB Life*, 61:99.

[71] M. A. S. Santos, G. Moura, S. E. Massey, M. F. Tuite (2004), Driving change: The evolution of alternative genetic codes. *Trends Genet.*, 20:95.

[72] N. Metanis, J. Beld, D. Hilvert, in *The Chemistry of Selenocysteine in PATAI'S Chemistry of Functional Groups*, John Wiley & Sons, Ltd, **2009**.

[73] J. Beld, K. J. Woycechowsky, D. Hilvert (2007), Selenoglutathione: Efficient oxidative protein folding by a diselenide. *Biochemistry*, 46:5382.

[74] N. Metanis, D. Hilvert (2012), Strategic use of non-native diselenide bridges to steer oxidative protein folding. *Angew. Chem. Int. Ed.*, 51:5585.

[75] G. W. Snider, E. Ruggles, N. Khan, R. J. Hondal (2013), Selenocysteine confers resistance to inactivation by oxidation in thioredoxin reductase: comparison of selenium and sulfur enzymes. *Biochemistry*, 52:5472.

[76] Y. Zhang, P. V. Baranov, J. F. Atkins, V. N. Gladyshev (2005), Pyrrolysine and selenocysteine use dissimilar decoding strategies. *J. Biol. Chem.*, 280:20740.

[77] J. A. Krzycki (2013), The path of lysine to pyrrolysine. *Curr. Opin. Chem. Biol.*, 17:619.

[78] J. A. Krzycki (2004), Function of genetically encoded pyrrolysine in corrinoid-dependent methylamine methyltransferases. *Curr. Opin. Chem. Biol.*, 8:484.

[79] R. Y. Tsien (1998), The green fluorescent protein. *Annu. Rev. Biochem.*, 67:509.

[80] A. Messerschmidt (2001) *Handbook of metalloproteins*. Wiley, Chichester. NY.

[81] Y. Lu, N. Yeung, N. Sieracki, N. M. Marshall (2009), Design of functional metalloproteins. *Nature*, 460:855.

[82] K. D. Karlin (1993), Metalloenzymes, structural motifs, and inorganic models. *Science*, 261:701.

[83] T. W. Lane, M. A. Saito, G. N. George, I. J. Pickering, R. C. Prince, F. M. M. Morel (2005), A cadmium enzyme from a marine diatom. *Nature*, 435:42.

[84] K. J. Waldron, J. C. Rutherford, D. Ford, N. J. Robinson (2009), Metalloproteins and metal sensing. *Nature*, 460:823.

[85] B. K. Burgess, D. J. Lowe (1996), Mechanism of molybdenum nitrogenase. *Chem. Rev.*, 96:2983.

[86] S. P. de Visser (Ed.) (2011) *Iron-containing enzymes: Versatile catalysts of hydroxylation reactions in nature*. Royal Society of Chemistry, Cambridge, UK.

[87] J. Weston (2005), Mode of action of bi- and trinuclear zinc hydrolases and their synthetic analogues. *Chem. Rev.*, 105:2151.

[88] M. J. Ryle, R. P. Hausinger (2002), Non-heme iron oxygenases. *Curr. Opin. Chem. Biol.*, 6:193.

[89] T. D. H. Bugg (2003), Dioxygenase enzymes: Catalytic mechanisms and chemical models. *Tetrahedron*, 59:7075.

[90] J. M. Bollinger, J. C. Price, L. M. Hoffart, E. W. Barr, C. Krebs (2005), Mechanism of taurine α-ketoglutarate dioxygenase (TauD) from *Escherichia coli. Eur. J. Inorg. Chem.*, 4245.

[91] P. R. Ortiz de Montellano (2005) *Cytochrome P450: Structure, mechanism, and biochemistry*. Kluwer Academic / Plenum Publishers, New York, NY.

[92] R. Bernhardt (2006), Cytochromes P450 as versatile biocatalysts. *J. Biotechnol.*, 124:128.

[93] W. N. Lipscomb, N. Strater (1996), Recent advances in zinc enzymology. *Chem. Rev.*, 96:2375.

[94] B. S. Der, M. Machius, M. J. Miley, J. L. Mills, T. Szyperski, B. Kuhlman (2012), Metal-mediated affinity and orientation specificity in a computationally-designed protein homodimer. *J. Am. Chem. Soc.*, 134:375.

[95] B. S. Der, D. R. Edwards, B. Kuhlman (2012), Catalysis by a *de novo* zinc-mediated protein interface: Implications for natural enzyme evolution and rational enzyme engineering. *Biochemistry*, 51:3933.

[96] R. H. Crabtree (2009) *The organometallic chemistry of the transition metals*. Wiley, Hoboken, N.J.

[97] E. N. Jacobsen, A. Pfaltz, H. Yamamoto (editors) (1999) *Comprehensive asymmetric catalysis*. Springer, Berlin, NY.

[98] T. Heinisch, T. R. Ward (2010), Design strategies for the creation of artificial metalloenzymes. *Curr. Opin. Chem. Biol.*, 14:184.

[99] G. Roelfes (2011), Metallopeptides for enantioselective catalysis. *ChemCatChem*, 3:647.

[100] M. R. Ringenberg, T. R. Ward (2011), Merging the best of two worlds: Artificial metalloenzymes for enantioselective catalysis. *Chem. Comm.*, 47:8470.

[101] M. T. Reetz, M. Rentzsch, A. Pletsch, M. Maywald, P. Maiwald, J. J. P. Peyralans, A. Maichele, Y. Fu, N. Jiao, F. Hollmann, *et al.* (2007), Directed evolution of enantioselective hybrid catalysts: A novel concept in asymmetric catalysis. *Tetrahedron*, 63:6404.

[102] D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore (2001), A field guide to foldamers. *Chem. Rev.*, 101:3893.

[103] M. M. Muller, M. A. Windsor, W. C. Pomerantz, S. H. Gellman, D. Hilvert (2009), A rationally designed aldolase foldamer. *Angew. Chem. Int. Ed.*, 48:922.

[104] M. M. Muller, H. Kries, E. Csuhai, P. Kast, D. Hilvert (2010), Design, selection, and characterization of a split chorismate mutase. *Prot. Sci.*, 19:1000.

[105] A. d. Meijere, F. o. Diederich (2004) *Metal-catalyzed cross-coupling reactions*. Wiley-VCH, Weinheim, Germany.

[106] A. H. Hoveyda, A. R. Zhugralin (2007), The remarkable metal-catalysed olefin metathesis reaction. *Nature*, 450:243.

[107] H. C. Kolb, M. S. Vannieuwenhze, K. B. Sharpless (1994), Catalytic asymmetric dihydroxylation. *Chem. Rev.*, 94:2483.

[108] J. G. d. Vries (2007) *Handbook of homogeneous hydrogenation*. WILEY-VCH, Weinheim.

[109] F. Rosati, G. Roelfes (2010), Artificial metalloenzymes. *ChemCatChem*, 2:916.

[110] T. R. Ward (editor) (2009) *Bio-inspired catalysts*. Springer, Berlin.

[111] E. M. Brustad, F. H. Arnold (2011), Optimizing non-natural protein function with directed evolution. *Curr. Opin. Chem. Biol.*, 15:201.

[112] J. Steinreiber, T. R. Ward (2008), Artificial metalloenzymes as selective catalysts in aqueous media. *Coord. Chem. Rev.*, 252:751.

[113] H. L. Levine, Y. Nakagawa, E. T. Kaiser (1977), Flavopapain: Synthesis and properties of semi-synthetic enzymes. *Biochem. Biophys. Res. Commun.*, 76:64.

[114] R. L. Lundblad (2005) *Chemical reagents for protein modification*. CRC Press, Boca Raton, FL.

[115] D. F. Qi, C. M. Tann, D. Haring, M. D. Distefano (2001), Generation of new enzymes via covalent modification of existing proteins. *Chem. Rev.*, 101:3081.

[116] L. Panella, J. Broos, J. F. Jin, M. W. Fraaije, D. B. Janssen, M. Jeronimus-Stratingh, B. L. Feringa, A. J. Minnaard, J. G. de Vries (2005), Merging homogeneous catalysis with biocatalysis: Papain as hydrogenation catalyst. *Chem. Comm.*, 5656.

[117] W. Laan, B. K. Munoz, R. den Heeten, P. C. J. Kamer (2010), Artificial metalloenzymes through cysteine-selective conjugation of phosphines to photoactive yellow protein. *ChemBioChem*, 11:1236.

[118] P. Haquette, B. Talbi, L. Barilleau, N. Madern, C. Fosse, M. Salmain (2011), Chemically engineered papain as artificial formate dehydrogenase for NAD(P)H regeneration. *Org. Biomol. Chem.*, 9:5720.

[119] C. A. Kruithof, H. P. Dijkstra, M. Lutz, A. L. Spek, M. R. Egmond, R. J. M. K. Gebbink, G. van Koten (2008), Non-tethered organometallic phosphonate inhibitors for lipase inhibition: Positioning of the metal center in the active Site of cutinase. *Eur. J. Inorg. Chem.*, 4425.

[120] J. R. Carey, S. K. Ma, T. D. Pfister, D. K. Garner, H. K. Kim, J. A. Abramite, Z. L. Wang, Z. J. Guo, Y. Lu (2004), A site-selective dual anchoring strategy for artificial metalloprotein design. *J. Am. Chem. Soc.*, 126:10812.

[121] K. Yamamura, E. T. Kaiser (1976), Studies on oxidase activity of copper(II) carboxypeptidase A. *J. Chem. Soc., Chem. Commun.*, 830.

[122] M. Ohashi, T. Koshiyama, T. Ueno, M. Yanase, H. Fujii, Y. Watanabe (2003), Preparation of artificial metalloenzymes by insertion of chromium(III) Schiff base complexes into apomyoglobin mutants. *Angew. Chem. Int. Ed.*, 42:1005.

[123] A. Fernandez-Gacio, A. Codina, J. Fastrez, O. Riant, P. Soumillion (2006), Transforming carbonic anhydrase into epoxide synthase by metal exchange. *ChemBioChem*, 7:1013.

[124] Q. Jing, K. Okrasa, R. J. Kazlauskas (2009), Stereoselective hydrogenation of olefins using rhodium-substituted carbonic anhydrase - a new reductase. *Chem. Eur. J.*, 15:1370.

[125] M. Marchetti, G. Mangano, S. Paganelli, C. Botteghi (2000), A protein-rhodium complex as an efficient catalyst for two-phase olefin hydroformylation. *Tetrahedron Lett.*, 41:3717.

[126] N. D. Clarke, S. M. Yuan (1995), Metal search: A computer program that helps design tetrahedral metal-binding sites. *Proteins*, 23:256.

[127] H. Fazelinia, P. C. Cirino, C. D. Maranas (2009), OptGraft: A computational procedure for transferring a binding site onto an existing protein scaffold. *Prot. Sci.*, 18:180.

[128] H. W. Hellinga, F. M. Richards (1991), Construction of new ligand-binding sites in proteins of known structure: I. Computer-aided modeling of sites with predefined geometry. *J. Mol. Biol.*, 222:763.

[129] N. Yeung, Y. W. Lin, Y. G. Gao, X. Zhao, B. S. Russell, L. Y. Lei, K. D. Miner, H. Robinson, Y. Lu (2009), Rational design of a structural and functional nitric oxide reductase. *Nature*, 462:1079.

[130] D. E. Benson, M. S. Wisz, H. W. Hellinga (2000), Rational design of nascent metalloenzymes. *Proc. Natl. Acad. Sci. U. S. A.*, 97:6292.

[131] J. Kaplan, W. F. DeGrado (2004), De novo design of catalytic proteins. *Proc. Natl. Acad. Sci. U. S. A.*, 101:11566.

[132] M. Faiella, C. Andreozzi, R. T. M. de Rosales, V. Pavone, O. Maglio, F. Nastri, W. F. DeGrado, A. Lombardi (2009), An artificial di-iron oxo-protein with phenol oxidase activity. *Nat. Chem. Biol.*, 5:882.

[133] M. L. Zastrow, A. F. A. Peacock, J. A. Stuckey, V. L. Pecoraro (2012), Hydrolytic catalysis and structural stabilization in a designed metalloprotein. *Nat. Chem.*, 4:118.

[134] M. E. Wilson, G. M. Whitesides (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.

[135] J. Collot, J. Gradinaru, N. Humbert, M. Skander, A. Zocchi, T. R. Ward (2003), Artificial metalloenzymes for enantioselective catalysis based on biotin-avidin. *J. Am. Chem. Soc.*, 125:9030.

[136] U. E. Rusbandi, C. Lo, M. Skander, A. Ivanova, M. Creus, N. Humbert, T. R. Ward (2007), Second generation artificial hydrogenases based on the biotin-avidin technology: Improving activity, stability and selectivity by introduction of enantiopure amino acid spacers. *Adv. Synth. Catal.*, 349:1923.

[137] M. Durrenberger, T. Heinisch, Y. M. Wilson, T. Rossel, E. Nogueira, L. Knorr, A. Mutschler, K. Kersten, M. J. Zimbron, J. Pierron, *et al.* (2011), Artificial transfer hydrogenases for the enantioselective reduction of cyclic imines. *Angew. Chem. Int. Ed.*, 50:3026.

[138] C. Letondor, N. Humbert, T. R. Ward (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.

[139] A. Pordea, M. Creus, J. Panek, C. Duboc, D. Mathis, M. Novic, T. R. Ward (2008), Artificial metalloenzyme for enantioselective sulfoxidation based on vanadyl-loaded streptavidin. *J. Am. Chem. Soc.*, 130:8085.

[140] J. Pierron, C. Malan, M. Creus, J. Gradinaru, I. Hafner, A. Ivanova, A. Sardo, T. R. Ward (2008), Artificial metalloenzymes for asymmetric allylic alkylation on the basis of the biotin-avidin technology. *Angew. Chem. Int. Ed.*, 47:701.

[141] T. K. Hyster, L. Knorr, T. R. Ward, T. Rovis (2012), Biotinylated Rh(III) complexes in engineered streptavidin for accelerated asymmetric C-H activation. *Science*, 338:500.

[142] R. H. Grubbs (2003) *Handbook of metathesis*. Wiley-VCH, Weinheim, Germany.

[143] T. M. Trnka, R. H. Grubbs (2001), The development of $L_2X_2Ru=CHR$ olefin metathesis catalysts: An organometallic success story. *Acc. Chem. Res.*, 34:18.

[144] B. G. Davis, Y. A. Lin, J. M. Chalker (2010), Olefin cross-metathesis on proteins: Investigation of allylic chalcogen effects and guiding principles in metathesis partner selection. *J. Am. Chem. Soc.*, 132:16805.

[145] B. G. Davis, Y. A. Lin, J. M. Chalker, N. Floyd, G. J. L. Bernardes (2008), Allyl sulfides are privileged substrates in aqueous cross-metathesis: Application to site-selective protein modification. *J. Am. Chem. Soc.*, 130:9642.

[146] T. R. Hoye, H. Y. Zhao (2002), A method for easily determining coupling constant values: An addendum to "a practical guide to first-order multiplet analysis in ^{1}H-NMR spectroscopy". *J. Org. Chem.*, 67:4014.

[147] K. Grela, D. Burtscher (2009), Aqueous olefin metathesis. *Angew. Chem. Int. Ed.*, 48:442.

[148] R. H. Grubbs, J. P. Jordan (2007), Small-molecule N-heterocyclic-carbene-containing olefin-metathesis catalysts for use in water. *Angew. Chem. Int. Ed.*, 46:5152.

[149] L. Gulajski, A. Michrowska, J. Naroznik, Z. Kaczmarska, L. Rupnicki, K. Grela (2008), A highly active aqueous olefin metathesis catalyst bearing a quaternary ammonium group. *ChemSusChem*, 1:103.

[150] S. H. Kim, K. K. Kim, R. Kim (1998), Crystal structure of a small heat-shock protein. *Nature*, 394:595.

[151] R. Kim, K. K. Kim, H. Yokota, S. H. Kim (1998), Small heat shock protein of *Methanococcus jannaschii*, a hyperthermophile. *Proc. Natl. Acad. Sci. U. S. A.*, 95:9129.

[152] M. P. Bova, Q. L. Huang, L. L. Ding, J. Horwitz (2002), Subunit exchange, conformational stability, and chaperone-like function of the small heat shock protein 16.5 from *Methanococcus jannaschii. J. Biol. Chem.*, 277:38468.

[153] A. N. Cao, Z. Wang, L. H. Lai (2009), High activity of MjHSP16.5 under acidic condition. *Sci. China Ser. B.*, 52:325.

[154] T. Douglas, M. L. Flenniken, L. O. Liepold, B. E. Crowley, D. A. Willits, M. J. Young (2005), Selective attachment and release of a chemotherapeutic agent from the interior of a protein cage architecture. *Chem. Comm.*, 447.

[155] Z. Varpness, J. W. Peters, M. Young, T. Douglas (2005), Biomimetic synthesis of a H_2 catalyst using a protein cage architecture. *Nano Lett.*, 5:2306.

[156] G. L. Ellman (1959), Tissue sulfhydryl groups. *Arch. Biochem. Biophys.*, 82:70.

[157] R. H. Grubbs, B. K. Keitz, J. Bouffard, G. Bertrand (2011), Protonolysis of a ruthenium-carbene bond and applications in olefin metathesis. *J. Am. Chem. Soc.*, 133:8498.

[158] R. H. Grubbs, J. P. Gallivan, J. P. Jordan (2005), A neutral, water-soluble olefin metathesis catalyst based on an N-heterocyclic carbene ligand. *Tetrahedron Lett.*, 46:2577.

[159] R. H. Grubbs, M. S. Sanford, L. M. Henling, M. W. Day (2000), Ruthenium-based four-coordinate olefin metathesis catalysts. *Angew. Chem. Int. Ed.*, 39:3451.

[160] T. Matsuo, C. Imai, T. Yoshida, T. Saito, T. Hayashi, S. Hirota (2012), Creation of an artificial metalloprotein with a Hoveyda-Grubbs catalyst moiety through the intrinsic inhibition mechanism of alpha-chymotrypsin. *Chem. Comm.*, 48:1662.

[161] F. Philippart, M. Arlt, S. Gotzen, S. J. Tenne, M. Bocola, H. H. Chen, L. L. Zhu, U. Schwaneberg, J. Okuda (2013), A hybrid ring-opening

metathesis polymerization catalyst based on an engineered variant of the β-barrel Ppotein FhuA. *Chem. Eur. J.*, 19:13865.

[162] T. Matsuo, T. Yoshida, A. Fujii, K. Kawahara, S. Hirota (2013), Effect of added salt on ring-closing metathesis catalyzed by a water-soluble Hoveyda–Grubbs-type complex to form N-containing heterocycles in aqueous media. *Organometallics*, 32:5313.

[163] R. H. Grubbs, J. P. Jordan (2007), Small-molecule N-heterocyclic-carbene-containing olefin-metathesis catalysts for use in water. *Angew. Chem. Int. Ed.*, 46:5152.

[164] C. Che, W. Z. Li, S. Y. Lin, J. W. Chen, J. Zheng, J. C. Wu, Q. X. Zheng, G. Q. Zhang, Z. Yang, B. W. Jiang (2009), Magnetic nanoparticle-supported Hoveyda-Grubbs catalysts for ring-closing metathesis reactions. *Chem. Comm.*, 5990.

[165] M. L. Flenniken, D. A. Willits, S. Brumfield, M. J. Young, T. Douglas (2003), The small heat shock protein cage from *Methanococcus jannaschii* is a versatile nanoscale platform for genetic and chemical modification. *Nano Lett.*, 3:1573.

[166] S. Sasso, C. Ramakrishnan, M. Gamper, D. Hilvert, P. Kast (2005), Characterization of the secreted chorismate mutase from the pathogen *Mycobacterium tuberculosis. FEBS J.*, 272:375.

[167] T. Heinisch, K. Langowska, P. Tanner, J. L. Reymond, W. Meier, C. Palivan, T. R. Ward (2013), Fluorescence-based assay for the optimization of the activity of artificial transfer hydrogenase within a biocompatible compartment. *ChemCatChem*, 5:720.

[168] K. J. Franz, M. Nitz, B. Imperiali (2003), Lanthanide-binding tags as versatile protein coexpression probes. *ChemBioChem*, 4:265.

[169] B. Albert Griffin, S. R. Adams, J. Jones, R. Y. Tsien, in *Methods Enzymol., Vol. Volume 327* (Eds.: S. D. E. Jeremy Thorner, N. A. John), Academic Press, **2000**, pp. 565.

[170] G. Licini, P. Scrimin (2003), Metal-ion-binding peptides: From catalysis to protein tagging. *Angew. Chem. Int. Ed.*, 42:4572.

[171] K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski, R. J. Knapp (1991), A new type of synthetic peptide library for identifying ligand-binding activity. *Nature*, 354:82.

[172] M. B. Francis, E. N. Jacobsen (1999), Discovery of novel catalysts for alkene epoxidation from metal-binding combinatorial libraries. *Angew. Chem. Int. Ed.*, 38:937.

[173] K. D. Shimizu, B. M. Cole, C. A. Krueger, K. W. Kuntz, M. L. Snapper, A. H. Hoveyda (1997), Search for chiral catalysts through ligand diversity: Substrate-specific catalysts and ligand screening on solid phase. *Angew. Chem. Int. Ed.*, 36:1703.

[174] J. R. Porter, W. G. Wirschun, K. W. Kuntz, M. L. Snapper, A. H. Hoveyda (2000), Ti-catalyzed regio- and enantioselective synthesis of unsaturated alpha-amino nitriles, amides, and acids. Catalyst identification through screening of parallel libraries. *J. Am. Chem. Soc.*, 122:2657.

[175] J. R. Porter, J. F. Traverse, A. H. Hoveyda, M. L. Snapper (2001), Enantioselective synthesis of arylamines through Zr-catalyzed addition of dialkylzincs to imines. Reaction development by screening of parallel libraries. *J. Am. Chem. Soc.*, 123:984.

[176] S. J. Degrado, H. Mizutani, A. H. Hoveyda (2001), Modular peptide-based phosphine ligands in asymmetric catalysis: Efficient and enantioselective Cu-catalyzed conjugate additions to five-, six-, and seven-membered cyclic enones. *J. Am. Chem. Soc.*, 123:755.

[177] S. R. Gilbertson, S. E. Collibee, A. Agarkov (2000), Asymmetric catalysis with libraries of palladium β-turn phosphine complexes. *J. Am. Chem. Soc.*, 122:6522.

[178] V. Y. Torbeev, S. B. H. Kent (2007), Convergent chemical synthesis and crystal structure of a 203 amino acid "covalent dimer" HIV-1 protease enzyme molecule. *Angew. Chem. Int. Ed.*, 46:1667.

[179] J. Bos, A. Garcia-Herraiz, G. Roelfes (2013), An enantioselective artificial metallo-hydratase. *Chem. Sci.*, 4:3578.

[180] J. Bos, F. Fusetti, A. J. M. Driessen, G. Roelfes (2012), Enantioselective artificial metalloenzymes by creation of a novel active site at the protein dimer interface. *Angew. Chem. Int. Ed.*, 51:7472.

[181] J. M. Xie, W. S. Liu, P. G. Schultz (2007), A genetically encoded bidentate, metal-binding amino acid. *Angew. Chem. Int. Ed.*, 46:9239.

[182] H. S. Lee, P. G. Schultz (2008), Biosynthesis of a site-specific DNA cleaving protein. *J. Am. Chem. Soc.*, 130:13194.

[183] C. Harford, B. Sarkar (1997), Amino terminal Cu(II)- and Ni(II)-binding (ATCUN) motif of proteins and peptides: Metal binding, DNA cleavage, and other properties. *Acc. Chem. Res.*, 30:123.

[184] X. F. Huang, M. E. Pieczko, E. C. Long (1999), Combinatorial optimization of the DNA cleaving Ni(II)•Xaa-Xaa-His metallotripeptide domain. *Biochemistry*, 38:2160.

[185] Z. T. Ball (2013), Designing enzyme-like catalysts: A rhodium(II) metallopeptide case study. *Acc. Chem. Res.*, 46:560.

[186] A. N. Zaykov, B. V. Popp, Z. T. Ball (2010), Helix induction by dirhodium: Access to biocompatible metallopeptides with defined secondary structure. *Chem. Eur. J.*, 16:6651.

[187] B. V. Popp, Z. T. Ball (2011), Proximity-driven metallopeptide catalysis: Remarkable side chain scope enables modification of the Fos bZip domain. *Chem. Sci.*, 2:690.

[188] R. Sambasivan, Z. T. Ball (2010), Metallopeptides for asymmetric dirhodium catalysis. *J. Am. Chem. Soc.*, 132:9289.

[189] R. Sambasivan, Z. T. Ball (2012), Screening rhodium metallopeptide libraries "on bead": Asymmetric cyclopropanation and a solution to the enantiomer problem. *Angew. Chem. Int. Ed.*, 51:8568.

[190] S. Gladiali, E. Alberico (2006), Asymmetric transfer hydrogenation: Chiral ligands and applications. *Chem. Soc. Rev.*, 35:226.

[191] S. Hashiguchi, A. Fujii, J. Takehara, T. Ikariya, R. Noyori (1995), Asymmetric transfer hydrogenation of aromatic ketones catalyzed by chiral ruthenium(II) complexes. *J. Am. Chem. Soc.*, 117:7562.

[192] X. F. Wu, J. L. Xiao (2007), Aqueous-phase asymmetric transfer hydrogenation of ketones - a greener approach to chiral alcohols. *Chem. Comm.*, 2449.

[193] K. Severin, R. Bergs, W. Beck (1998), Bioorganometallic chemistry transition metal complexes with alpha-amino acids and peptides. *Angew. Chem. Int. Ed.*, 37:1635.

[194] H. Sigel, R. B. Martin (1982), Coordinating properties of the amide bond - stability and structure of metal-ion complexes of peptides and related ligands. *Chem. Rev.*, 82:385.

[195] R. Kramer, M. Maurus, K. Polborn, K. Sünkel, C. Robl, W. Beck (1996), Organometallic half-sandwich complexes promote the formation of linear oligopeptides from amino acid esters. *Chem. Eur. J.*, 2:1518.

[196] X. F. Wu, X. G. Li, F. King, J. L. Xiao (2005), Insight into and practical application of pH-controlled asymmetric transfer hydrogenation of aromatic ketones in water. *Angew. Chem. Int. Ed.*, 44:3407.

[197] C. Wang, X. F. Wu, J. L. Xiao (2008), Broader, greener, and more efficient: Recent advances in asymmetric transfer hydrogenation. *Chem. Asian J.* 3:1750.

[198] J. Lutz, F. Hollmann, T. V. Ho, A. Schnyder, R. H. Fish, A. Schmid (2004), Bioorganometallic chemistry: Biocatalytic oxidation reactions with biomimetic NAD^+/NADH co-factors and $[Cp^*Rh(bpy)H]^+$ for selective organic synthesis. *J. Organomet. Chem.*, 689:4783.

[199] V. Kohler, Y. M. Wilson, M. Durrenberger, D. Ghislieri, E. Churakova, T. Quinto, L. Knorr, D. Haussinger, F. Hollmann, N. J. Turner, *et al.* (2013), Synthetic cascades are enabled by combining biocatalysts with artificial metalloenzymes. *Nat. Chem.*, 5:93.

[200] C. Wang, B. Villa-Marcos, J. L. Xiao (2011), Hydrogenation of imino bonds with half-sandwich metal catalysts. *Chem. Comm.*, 47:9773.

[201] H. K. Chenault, G. M. Whitesides (1987), Regeneration of nicotinamide cofactors for use in organic synthesis. *Appl. Biochem. Biotechnol.*, 14:147.

[202] W. A. van der Donk, H. M. Zhao (2003), Recent developments in pyridine nucleotide regeneration. *Curr. Opin. Biotechnol.*, 14:421.

[203] J. Canivet, G. Suss-Fink, P. Stepnicka (2007), Water-soluble phenanthroline complexes of rhodium, iridium and ruthenium for the regeneration of NADH in the enzymatic reduction of ketones. *Eur. J. Inorg. Chem.*, 4736.

[204] F. Holmann, A. Kleeb, K. Otto, A. Schmid (2005), Coupled chemoenzymatic transfer hydrogenation catalysis for enantioselective reduction and oxidation reactions. *Tetrahedron: Asymmetry*, 16:3512.

[205] E. Steckhan, S. Herrmann, R. Ruppert, J. Thommes, C. Wandrey (1990), Continuous generation of NADH from NAD^+ and formate using a homogeneous catalyst with enhanced molecular-weight in a membrane reactor. *Angew. Chem. Int. Ed.*, 29:388.

[206] F. Hollmann, K. Hofstetter, A. Schmid (2006), Non-enzymatic regeneration of nicotinamide and flavin cofactors for monooxygenase catalysis. *Trends Biotechnol.*, 24:163.

[207] H. Y. Rhyoo, H. J. Park, Y. K. Chung (2001), The first Ru(II)-catalysed asymmetric hydrogen transfer reduction of aromatic ketones in aqueous media. *Chem. Comm.*, 2064.

[208] K. Ahlford, H. Adolfsson (2011), Amino acid derived amides and hydroxamic acids as ligands for asymmetric transfer hydrogenation in aqueous media. *Catal. Commun.*, 12:1118.

[209] I. M. Pastor, P. Vastila, H. Adolfsson (2003), Employing the structural diversity of nature: Development of modular dipeptide-analogue ligands for ruthenium-catalyzed enantioselective transfer hydrogenation of ketones. *Chem. Eur. J.*, 9:4031.

[210] M. Durrenberger, T. Heinisch, Y. M. Wilson, T. Rossel, E. Nogueira, L. Knorr, A. Mutschler, K. Kersten, M. J. Zimbron, J. Pierron, *et al.* (2011), Artificial Transfer Hydrogenases for the Enantioselective Reduction of Cyclic Imines. *Angew. Chem. Int. Ed.*, 50:3026.

[211] C. Jackel, P. Kast, D. Hilvert (2008), Protein design by directed evolution. *Annu. Rev. Biophys.*, 37:153.

[212] T. R. Hoye, H. Y. Zhao (2002), A method for easily determining coupling constant values: An addendum to "a practical guide to first-order multiplet analysis in H-1 NMR spectroscopy". *J. Org. Chem.*, 67:4014.

[213] M. Jahncke, G. Meister, G. Rheinwald, H. StoeckliEvans, G. SussFink (1997), Dinuclear (arene)ruthenium hydrido complexes: Synthesis, structure, and fluxionality of $(C_6Me_6)_2Ru_2H_3BH_4$. *Organometallics*, 16:1137.

[214] A. Nutton, P. M. Bailey, P. M. Maitlis (1981), Pentamethylcyclopentadienyl-rhodium and iridium Complexes. Part 29. Syntheses and X-ray structure determinations of $[(Rh(C_5Me_5))_2(OH)_3]OH•11H_2O$ and $[(Ir(C_5Me_5))_2(OH)_3]OH•14H_2O$ and related complexes. *J. Chem. Soc., Dalt. Trans.*, 1997.

[215] S. Ogo, N. Makihara, Y. Watanabe (1999), pH-dependent transfer hydrogenation of water-soluble carbonyl compounds with $[Cp^*Ir^{III}(H_2O)_3']^{2+}$ as a catalyst precursor and HCOONa as a hydrogen donor in water. *Organometallics*, 18:5470.

[216] F. Hollmann, B. Witholt, A. Schmid (2002), $[Cp^*Rh(bpy)(H_2O)]^{2+}$: a Versatile tool for efficient and non-enzymatic regeneration of nicotinamide and flavin coenzymes. *J. Mol. Cat. B,* 19:167.

[217] K. Gorska, J. Beyrath, S. Fournel, G. Guichard, N. Winssinger (2010), Ligand dimerization programmed by hybridization to study multimeric ligand-receptor interactions. *Chem. Comm.*, 46:7742.

[218] S. L. Miller (1953), A production of amino acids under possible primitive earth conditions. *Science*, 117:528.

[219] M. H. Engel, S. A. Macko (1997), Isotopic evidence for extraterrestrial non-racemic amino acids in the Murchison meteorite. *Nature*, 389:265.

[220] Kvenvold.K, J. Lawless, K. Pering, E. Peterson, J. Flores, Ponnampe.C, I. R. Kaplan, C. Moore (1970), Evidence for extraterrestrial amino acids and hydrocarbons in Murchison meteorite. *Nature*, 228:923.

[221] R. L. Lundblad, F. Macdonald (2010) *Handbook of biochemistry and molecular biology*. CRC Press, Boca Raton, Fl.

[222] G. A. Petsko, D. Ringe (2004) *Protein structure and function*. New Science Press, London, UK.

[223] S. H. Gellman (1998), Foldamers: A manifesto. *Acc. Chem. Res.*, 31:173.

[224] K. Kirshenbaum, A. E. Barron, R. A. Goldsmith, P. Armand, E. K. Bradley, K. T. V. Truong, K. A. Dill, F. E. Cohen, R. N. Zuckermann (1998), Sequence-specific polypeptoids: A diverse family of heteropolymers with stable secondary structure. *Proc. Natl. Acad. Sci. U. S. A.*, 95:4303.

[225] A. D. Bautista, C. J. Craig, E. A. Harker, A. Schepartz (2007), Sophistication of foldamer form and function in vitro and in vivo. *Curr. Opin. Chem. Biol.*, 11:685.

[226] R. P. Cheng, S. H. Gellman, W. F. DeGrado (2001), β-Peptides: From structure to function. *Chem. Rev.*, 101:3219.

[227] D. Seebach, A. K. Beck, D. J. Bierbaum (2004), The world of β- and γ-peptides comprised of homologated proteinogenic amino acids and other components. *Chem. Biodiversity*, 1:1111.

[228] D. Seebach, J. L. Matthews (1997), β-Peptides: A surprise at every turn. *Chem. Comm.*, 2015.

[229] D. S. Daniels, E. J. Petersson, J. X. Qiu, A. Schepartz (2007), High-resolution structure of a β-peptide bundle. *J. Am. Chem. Soc.*, 129:1532.

[230] D. G. Udugamasooriya, S. P. Dineen, R. A. Brekken, T. Kodadek (2008), A peptoid "antibody surrogate" that antagonizes VEGF receptor 2 activity. *J. Am. Chem. Soc.*, 130:5744.

[231] J. A. Kritzer, O. M. Stephens, D. A. Guarracino, S. K. Reznik, A. Schepartz (2005), β-Peptides as inhibitors of protein-protein interactions. *Bioorg. Med. Chem.*, 13:11.

[232] M. Rueping, Y. Mahajan, M. Sauer, D. Seebach (2002), Cellular uptake studies with β-peptides. *ChemBioChem*, 3:257.

[233] N. Umezawa, M. A. Gelman, M. C. Haigis, R. T. Raines, S. H. Gellman (2002), Translocation of a β-peptide across cell membranes. *J. Am. Chem. Soc.*, 124:368.

[234] R. David, R. Gunther, L. Baurmann, T. Luhmann, D. Seebach, H. J. Hofmann, A. G. Beck-Sickinger (2008), Artificial chemokines: Combining chemistry and molecular biology for the elucidation of interleukin-8 functionality. *J. Am. Chem. Soc.*, 130:15311.

[235] U. Arnold, B. R. Huck, S. H. Gellman, R. T. Raines (2013), Protein prosthesis β-peptides as reverse-turn surrogates. *Prot. Sci.*, 22:274.

[236] B. C. Lee, R. N. Zuckermann (2011), Protein sde-chain translocation mutagenesis via incorporation of peptoid residues. *ACS Chem. Biol.*, 6:1367.

[237] L. M. Johnson, S. H. Gellman (2013), Alpha-helix mimicry with α/β-peptides. *Methods in Protein Design*, 523:407.

[238] S. S. Shekhawat, I. Ghosh (2011), Split-protein systems: Beyond binary protein-protein interactions. *Curr. Opin. Chem. Biol.*, 15:789.

[239] G. M. Fang, Y. M. Li, F. Shen, Y. C. Huang, J. B. Li, Y. Lin, H. K. Cui, L. Liu (2011), Protein chemical synthesis by ligation of peptide hydrazides. *Angew. Chem. Int. Ed.*, 50:7645.

[240] P. E. Dawson, T. W. Muir, I. Clarklewis, S. B. H. Kent (1994), Synthesis of proteins by native chemical ligation. *Science*, 266:776.

[241] Z. E. Reinert, G. A. Lengyel, W. S. Horne (2013), Protein-like tertiary folding behavior from heterogeneous backbones. *J. Am. Chem. Soc.*, 135:12528.

[242] M. Butz, M. Neuenschwander, P. Kast, D. Hilvert (2011), An N-terminal protein degradation tag enables robust selection of highly active enzymes. *Biochemistry*, 50:8594.

[243] G. MacBeath, P. Kast, D. Hilvert (1998), Redesigning enzyme topology by directed evolution. *Science*, 279:1958.

[244] K. Vamvaca, B. Vogeli, P. Kast, K. Pervushin, D. Hilvert (2004), An enzymatic molten globule: Efficient coupling of folding and catalysis. *Proc. Natl. Acad. Sci. U. S. A.*, 101:12860.

[245] D. H. Calhoun, C. A. Bonner, W. Gu, G. Xie, R. A. Jensen (2001), The emerging periplasm-localized subclass of AroQ chorismate mutases, exemplified by those from Salmonella typhimurium and Pseudomonas aeruginosa. *Gen. Biol.*, 2:

[246] Y. Cheng, W. H. Prusoff (1973), Relationship between inhibition constant (K_i) and concentration of inhibitor which causes 50% inhibition (IC_{50}) of an enzymatic reaction. *Biochem. Pharmacol.*, 22:3099.

[247] D. Baker (2010), An exciting but challenging road ahead for computational enzyme design. *Prot. Sci.*, 19:1817.

[248] G. Kiss, N. Celebi-Olcum, R. Moretti, D. Baker, K. N. Houk (2013), Computational Enzyme Design. *Angew. Chem. Int. Ed.*, 52:5700.

[249] N. Preiswerk (2013), *Application of modern enzyme design technologies and strategies for their refinement.* PhD thesis, ETH Zürich.

[250] A. Crameri, S. A. Raillard, E. Bermudez, W. P. C. Stemmer (1998), DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature*, 391:288.

[251] L. Yuan, I. Kurek, J. English, R. Keenan (2005), Laboratory-directed protein evolution. *Microbiol. Mol. Biol. Rev.*, 69:373.

[252] A. V. Shivange, J. Marienhagen, H. Mundhada, A. Schenk, U. Schwaneberg (2009), Advances in generating functional diversity for directed protein evolution. *Curr. Opin. Chem. Biol.*, 13:19.

[253] F. H. Arnold, G. Georgiou (Editors) (2003) *Directed enzyme evolution : Screening and selection methods*. Humana Press, Totowa, N.J.

[254] C. K. Savile, J. M. Janey, E. C. Mundorff, J. C. Moore, S. Tam, W. R. Jarvis, J. C. Colbeck, A. Krebber, F. J. Fleitz, J. Brands, *et al.* (2010), Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science*, 329:305.

[255] A. A. Desai (2011), Sitagliptin manufacture: A compelling tale of green chemistry, process intensification, and industrial asymmetric catalysis. *Angew. Chem. Int. Ed.*, 50:1974.

[256] S. V. Taylor, K. U. Walter, P. Kast, D. Hilvert (2001), Searching sequence space for protein catalysts. *Proc. Natl. Acad. Sci. U. S. A.*, 98:10596.

[257] H. Kries, R. Blomberg, D. Hilvert (2013), De novo enzymes by computational design. *Curr. Opin. Chem. Biol.*, 17:221.

[258] R. Maini, D. T. Nguyen, S. X. Chen, L. M. Dedkova, S. R. Chowdhury, R. Alcala-Torano, S. M. Hecht (2013), Incorporation of beta-amino acids into dihydrofolate reductase by ribosomes having modifications in the peptidyltransferase center. *Bioorg. Med. Chem.*, 21:1088.

[259] J. P. Richard, K. Melikov, E. Vives, C. Ramos, B. Verbeure, M. J. Gait, L. V. Chernomordik, B. Lebleu (2003), Cell-penetrating peptides - a reevaluation of the mechanism of cellular uptake. *J. Biol. Chem.*, 278:585.

[260] L. Giger, S. Caner, R. Obexer, P. Kast, D. Baker, N. Ban, D. Hilvert (2013), Evolution of a designed retro-aldolase leads to complete active site remodeling. *Nat. Chem. Biol.*, 9:494.

[261] M. D. Toscano, K. J. Woycechowsky, D. Hilvert (2007), Minimalist active-site redesign: Teaching old enzymes new tricks. *Angew. Chem. Int. Ed.*, 46:3212.

[262] H. Pellissier (2011), Recent developments in dynamic kinetic resolution. *Tetrahedron*, 67:3769.

[263] K. J. Woycechowsky, A. Choutko, K. Vamvaca, D. Hilvert (2008), Relative Tolerance of an Enzymatic Molten Globule and Its Thermostable Counterpart to Point Mutation. *Biochemistry*, 47:13489.