# *De novo* self-assembling peptides possessing esterase properties


Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

Submitted to the Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by

Jason L. Heier

From Hague, ND, USA

October 2015

1st Reviewer: Prof. Dr. Beate Koksch

2nd Reviewer: Prof. Dr. Mathias Christmann

Date of defense: June 16, 2016

## Declaration

The work presented here was performed in the research group of Prof. Dr. Beate Koksch from November 2010 until October 2015 at the Institute of Chemistry and Biochemistry in the Department of Biology, Chemistry and Pharmacy of Freie Universität Berlin.

I hereby confirm that this dissertation entitled "*De novo* self-assembling peptides possessing esterase properties" is exclusively the result of my own autonomous work based on my research and published literature as cited. I also declare that no part of this dissertation has been prepared inappropriately or used in any other work in another higher education or research institution.

Berlin, October 2015 Jason L. Heier

## Publications

***ACS Chem. Biol. DOI: 10.1021/acschembio.5b00435***. Tailored Presentation of Carbohydrates on a Coiled Coil-Based Scaffold for Asialoglycoprotein Receptor Targeting. E. Zacco, J. Hütter, J. L. Heier, J. Mortier, P. H. Seeberger, B. Lepenies & B. Koksch, **2015**.

## Posters

**12th German Peptide Symposium, Darmstadt (Germany), March 18-21, 2015:** *De novo* peptide hexamer cast in new role as zinc metalloenzyme. J. L. Heier & B. Koksch.

**Center of Supramolecular Interactions 2nd General Meeting, Berlin (Germany), March 10-11, 2011:** Self-assembled metallo-supramolecular squares as templates for α-helical peptide bundles. J. L.Heier, C. A. Schalley & B. Koksch.

For reasons of data security, the *curriculum vitae* has been omitted from the published version.

For reasons of data security, the *curriculum vitae* has been omitted from the published version.

## Acknowledgements

I would like to thank Prof. Dr. Beate Koksch for her support as well as all the patience she has shown and taught. This project involved many aspects which were new to both of us. I appreciate her faith in me.

I thank Prof. Dr. Mathias Christmann for agreeing to be my second reviewer.

I thank the mass spectrometry and NMR services of the Chemistry Department at the FU-Berlin for their efficiency. I also thank the Material Management of Takustr. 3 for their friendly service.

I am grateful for the help of Dr. Sumati Bhatia and Dr. Manoj Kumar Muthyala.

I thank the entire group of AG Koksch for all their help and support. I thank “My Frient” Dr. Cosimo Cadicamo who was a mentor to me. Many thanks to Dr. Allison Berger for all her help and valuable suggestions. Sincere, helpful people can be hard to find. Thank you, Dr. Elsa Zacco. It was a pleasure working with you. I wish you all the best in your bright future. Thank you, Valentina and AnaRita. Anytime you need advice or just want to talk about peptides, call me.

Thanks for everyone in the research group of Jun. Prof. Dr. Annabelle Bertin for all your support.

I thank my family. I am sorry that I was not there. I owe everything to my parents, Jane and Tim Heier. I wish there was some form in which I could repay you. Thanks to my brothers and sisters, Janet, Allen, Randy, Shari, Brenda and Nora and your lovely families. I love and miss you all.

I thank you, Katie, for being Celina’s mom and always helping me.

And Katja, I hope you understand how much I appreciate all that you do and don’t do for me. I would be lonely and thin without you.

Most of all, I am thankful for you, Celina. I am so proud that you are my daughter. Thanks for taking care of me. You are the world to me.

## Zusammenfassung

Die Fähigkeit von Enzymen, Reaktionen unter Erhalt einer exzellenten Substrat- und Produktspezifität zu beschleunigen, interessiert Chemiker seit Jahrzehnten. Bei der Adaption von Enzymen für die Bedingungen der Synthesechemie entstanden jedoch zahlreiche Hindernisse, wie bspw. eine verringerte Löslichkeit und Thermostabilität sowie eine geringere Substratvielfalt. Für die Bewältigung dieser Probleme nutzt das protein engineering sowohl Methoden des rationalen Designs und gerichteter Evolution, als auch moderne computergestützte sowie kombinatorische Ansätze. Obwohl diese Techniken zur Entwicklung zahlreicher Biokatalysatoren erfolgreich eingesetzt wurden, bleibt das Verständnis über die Proteinfaltung und Wechselwirkungen, welche für die enzymatische Katalyse erforderlich sind, weiterhin begrenzt. Dies wird vor allem dadurch belegt, dass es unmöglich ist, katalytisch aktive Proteine *de novo* zu generieren, die ähnliche Fähigkeiten aufweisen wie Enzyme.

Um ein besseres Verständnis von der Enzymstruktur und Funktion zu ermöglichen, wird der „*bottom-up*"-Ansatz für das *de novo* Design von katalytisch aktiven Proteinen genutzt. Dieser minimalistische Ansatz nutzt einfache Peptide mit vorhersagbaren Selbstorganisationseigenschaften zu gut definierten dreidimensionalen Templaten. In diese Template wird katalytische Aktivität bzw. werden andere Funktionalitäten eingebaut. Die vereinfachten, dennoch proteinartigen Systeme ermöglichen ein verbessertes Verständnis darüber, wie einzelne Wechselwirkungen zu bestimmten enzymatischen Eigenschaften beitragen können.

Unter Einsatz der „*bottom-up"*-Technik wurde der enzymatische Apparat für eine Esterase-artige Aktivität in drei Modellpeptide eingeführt. Der Einbau von Histidin in das Carboxylat-reiche Peptid E3 führte zu einem *random coil* mit einer geringen katalytischen Aktivität. Als das Peptid mit K3 in einem heterodimeren *coiled coil* fixiert wurde, nahm die Aktivität ab, aber das Peptid zeigte eine enzymartige Sensitivität. Außerdem wurde festgestellt, dass Ac-IHIHQI-$NH_2$, welches ein Histidin-koordiniertes Zink für die Aktivierung von Wasser besitzt, eine höhere Selektivität für hydrophobe L-Aminosäureester besitzt, wenn die Wechselwirkung mit dem Substrat während des Prozesses der Selbstaggregation besitzt, als im Vergleich zu vollständig aggregiertem Zustand. Diese Beobachtung deutet auf die Wechselwirkung von hydrophoben Substratresten mit der hydrophoben Oberfläche der aggregierenden $\beta$-Faltblätter. Zum Schluss wurde Zink in ein hexameres $\alpha$-helikales *coiled coil* mit sechs Histidinresten im hydrophoben Kern eingeführt. Es zeigte sich, dass die Koordination von Zink in einem für Wasser zugänglichen und Substrat-aufnehmenden Spalt die Esterhydrolyse stark beschleunigte.

## Abstract

The ability of enzymes to accelerate reactions while maintaining excellent substrate selectivity and product specificity has interested chemists for decades. The adaptation of enzymes to the reaction conditions of chemical synthesis has, however, been met with numerous obstacles such as reduced solubility, thermostability and substrate scope. To overcome these obstacles, protein engineering employs methods of rational design and directed evolution as well as more modern computational and combinatorial approaches. Although these techniques have been successfully applied to alter biocatalysts, our understanding of protein folding and the interactions required for enzyme catalysis remains limited. This is best evidenced by the inability to design *de novo* catalytic proteins comparable to enzymes.

To develop a better understanding of enzyme structure and function, the “bottom-up” approach to the *de novo* design of catalytic proteins is applied. This minimalistic approach uses simple peptides which predictably self-assemble into well-defined three-dimensional templates, on which catalytic machinery or other functionalities believed to impart a particular property are introduced. The simplified, yet protein-like, environments of the self-assembling peptides allow for enhanced clarity of the individual interactions to assess how such interactions contribute to particular enzymatic properties.

Using the bottom-up technique, catalytic machinery for esterase-like activity was incorporated into three model peptide systems. The insertion of histidine into the carboxylate-rich E3 peptide resulted in a random coil with slight catalytic activity. This activity was muted when fixed in a heterodimeric coiled coil with K3 showing an enzyme-like sensitivity to its environment. In addition, the fibril-forming Ac-IHIHIQI-$NH_2$, which uses histidine-coordinated zinc to activate water, was found to be more selective for hydrophobic, L-amino acid esters while in the process of self-assembly than when applied as fully-formed fibrils. This finding points to the interaction of hydrophobic substrate moieties with exposed hydrophobic surfaces of the assembling β-sheets. Lastly, zinc was introduced to a hexameric α-helical coiled coil featuring six histidine residues in its hydrophobic interior. The coordination of zinc within a water-accessible, substrate accommodating cleft was found to greatly accelerate ester hydrolysis.

## List of abbreviations

| | |
|---|---|
| CA | Carbonic anhydrase |
| ZHN | Zinc-bound hydroxide nucleophile |
| NADH | Nicotinamide adenine dinucleotide (protonated) |
| DMSO | Dimethyl sulfoxide |
| DCE | Dichloroethane |
| CALB | *Candida antarctica* lipase B |
| PPL | Porcine pancreas lipase |
| PCR | Polymerase chain reaction |
| CLECs | Cross-linked enzyme crystals |
| CLEAs | Cross-linked enzyme aggregates |
| Taq | *Thermus aquaticus* |
| PAL | *Pseudomonas aeruginosa* lipase |
| ISM | Iterative saturation mutagenesis |
| ORBIT | Optimization of Rotamers by Itertive Techniques |
| *p*NPA | *p*-Nitrophenyl acetate |
| NADPH | Nicotinamide adenine dinucleotide phosphate (protonated) |
| bPP | Bovine pancreatic polypeptide |
| Z-Phe-ONp | *N*-Benzyloxycarbonyl-phenylalanine *p*-nitrophenyl ester |
| TEM | Transmission electron microscopy |
| CD | Circular dichroism |
| Tris | Tris(hydroxymethyl)aminomethane |
| CT | Catalytic triad |
| hPG | Hyper-branched polyglycerol |

## List of abbreviations (cont.)

| | |
|---|---|
| Cu-AAC | Copper-catalyzed alkyne-azide cycloaddition |
| Au-MPCs | Gold monolayer-protected clusters |
| TEG | Triethylene glycol |
| DNPB | 2,4-Dinitrophenylbutanoate |
| Z-Leu-ONp | *N*-Benzyloxycarbonyl-leucine-*p*-nitrophenyl ester |
| Z-Gly-ONp | *N*-Benzyloxycarbonyl-glycine -*p*-nitrophenyl ester |
| Boc-Asn-ONp | *N*-Butyloxycarbonyl-asparagine -*p*-nitrophenyl ester |
| Fmoc | 9-Fluorenylmethoxycarbonyl protecting group |
| SPPS | Solid phase peptide synthesis |
| RP-HPLC | Reverse-phase high-pressure liquid chromatography |
| MS | Mass spectrometry |
| *t*Bu | *tert*-Butyl protecting group |
| Trt | Triphenylmethyl protecting group |
| Boc | Butyloxycarbonyl protecting group |
| TGA | TentaGel-like composite resin for peptide acid synthesis |
| TGT | TentaGel-like composite resin with triphenylmethyl linker |
| MBHA | 4-Methylbenzhydrylamine hydrochloride salt |
| DMF | Dimethylformamide |
| DCM | Dichloromethane |
| NMP | *N*-Methyl-2-pyrolidone |
| DBU | 1,8-Diazabicyclo[5.4.0]undec-7-ene |
| DIC | *N,N'*-Diisopropylcarbodiimide |
| HOAt | 1-Hydroxy-7-azabenzotriazole |

## List of abbreviations (cont.)

| | |
|---|---|
| HOBt | *N*-Hydroxybenzotriazole |
| HCTU | *O*-(6-Chlorobenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate |
| TBTU | 2-(1*H*-Benzotriazol-1-yl)-1,1,3,3,-tetramethyluronium tetrafluoroborate |
| DIPEA | *N,N*-Diisopropylethylamine |
| TFA | Trifluoroacetic acid |
| TIPS | Triisopropylsilane |
| MeCN | Acetonitrile |
| Abz | 2-Aminobenzoic acid |
| PTFE | Polytetrafluoroethylene |
| ESI-ToF | Electrospray ionization-time of flight |
| PMMA | Poly (methyl methacrylate) |
| Mn | Number average molecular weight |
| PDI | Polydispersity index |
| MWCO | Molecular weight cut-off |
| EDTA | Ethylenediaminetetraacetic acid |
| GPC | Gel permeation chromatography |
| RI | Refractive index |
| DF | Dilution factor |
| UV-Vis | Ultraviolet-visible |

## List of abbreviations (cont.)

The three- and one-letter abbreviations used in this work to represent L-amino acids are consistent with biochemical nomenclature proposed by the IUPAC-IUB Commission (*Eur. J. Biochem*. **1984**, *138*, 9-37).

# Table of contents

# 1 Introduction

The number of diverse biochemical reactions required to maintain life are considerable. Even more staggering are the rates in which such events must proceed in order to meet biological demands. Nearly all reactions occurring within living systems are mediated by enzymes. Unlike other classes of protein, enzymes do not simply function by specifically binding other molecules, but binding them in such a manner which induces transition.[1] By stabilizing a high-energy transition state, enzymes act as catalysts, accelerating the rate of spontaneous reactions in which they themselves are not consumed. In contrast to conventional chemical catalysts (i.e., transition metals, acids and bases), enzymes are generally more efficient and operate under mild conditions with superior substrate specificity.[2,3]

Since enzymes catalyze reactions with remarkably high selectivity, their passage through cellular membranes (and often yet within intact cells) into the chemistry laboratory as biocatalysts[4] *only seems* natural. Enzymes impose their selectivity according to the functional groups present on substrates (i.e., chemo-selectivity), chirality (i.e., stereo-selectivity) and position (i.e., regio-selectivity). When effectively employed in synthetic chemistry,[5,6] this heightened selectivity is ultimately converted into uncomplicated protection schemes, fewer side reactions and simplified purification.[5,7-8] Moreover, biocatalysts are completely biodegradable and offer a much greener alternative to classical stoichiometric synthesis and to a lesser extent traditional chemical catalysis.[4,5,7]

The catalytic properties of enzymes have *traditionally* evolved within cells, not flasks or 96-well microtiter plates. An enzyme is evolved typically for a single substrate, not a broad class of reactants having a common functional group.[5] As a result, the high specificity which allows enzymes to function so adeptly within living organisms can be restrictive and less effective in chemical synthesis. Whereas life thrives on molecular fidelity, biocatalysis often settles for enzymatic promiscuity.[10-12] Furthermore, within living organisms, enzymes are responsible for maintaining chemical balance via cellular regulatory mechanisms known as feedback inhibition.[2,13] In the chemistry lab, this is called product inhibition and results in longer reaction times or reduced yields.

## 1. Introduction

In their natural form, enzymes are extremely complex molecules possessing intricate structure-function relationships, which are highly responsive to their environment. Therefore, the utility of enzymes as biocatalysts outside their familiar settings (e.g., in organic solvents, at high temperatures and under harsh pH conditions) in chemical synthesis has been considered limited.[5,6] Nonetheless, inspired by a partial, yet ever-unfolding knowledge of enzyme structure and function, efforts are continually made to expand the use of biocatalysis in chemical synthesis by engineering tailor-made enzymes,[14,15] and even enzyme mimics[16] and artificial enzymes.[17-21]

Recent developments in biotechnology and protein design have enhanced biocatalysis by broadening its substrate scope, fine-tuning selectivity, inducing structural stability under operational conditions and producing biocatalysts that are more robust and easier to recycle. The end result of such efforts is the production of catalysts which are not only more practical, but can be used in a cost-effective and an environmentally friendly fashion. To approach such challenges, protein engineering typically applies two techniques commonly referred to as rational design[22] and directed evolution.[23]

The rational design of biocatalysts from model enzymes was traditionally executed with site-directed mutagenesis,[24] a genetic method used to alter specific residues within a protein's amino acid sequence. However, in order for this alteration to be effective, a minimal knowledge of its structural and/or functional outcome is required. Such studies can be carried out systematically to elucidate the roles of specific residues within a protein's sequence. Unfortunately, since enzymes are so complex, the consequences of site-directed mutagenesis are often difficult to predict as even the slightest alteration can have a large effect in structure and function. Therefore, a great deal of time and effort is often required before achieving the properties of the biocatalyst desired.[7]

In contrast to rational design, directed evolution relies less on prior structural knowledge and more on mutagenesis and the screening of mutant libraries. In this technique, a gene encoding a model enzyme undergoes repetitive cycles of mutagenesis, expression and selection from mutant libraries. Selection is therefore based on the properties of the biocatalyst sought and its intended operational conditions. Surprisingly, the creation (even when random or blind) and screening of a large number of mutants is often the more effective method for rapidly securing

property-tailored proteins.[23] However, extensive deconvolution is needed to evaluate how each individual mutation contributes to the engineered properties.[7]

To obtain a less convoluted understanding of residue behavior in its larger protein-like periphery, *de novo* design can be advantageously implemented. For this purpose, peptides, particularly those which self-assemble into protein-like structures, provide a valuable platform for protein engineering. [25,26] By offering a simplified environment and, thereby, affording a higher resolution of the interactions involved, *de novo* peptide design takes a minimalist or "bottom-up" route[27] to understanding the fundamental aspects required for protein structure and function.[28] This bottom-up approach is, therefore, well suited for unravelling the minimal requirements of enzyme activity and integral to the *de novo* design of artificial enzymes.[29,30] When combined with molecular modeling and directed evolution, the once dreamed of design of novel enzymes from scratch becomes an actual possibility. Moreover, the application of *de novo* peptide design in catalysis establishes a logical bridge connecting the fields of organo- and biocatalysis.[27,31,32]

The postmillennial explosion of research in the field of organocatalysis[33-35] has revealed what is possible when small organic molecules behave like enzymes in terms of modes of activation and selectivity. Since amino acids,[36-38] amino acid derivatives[39-44] and oligopeptides[32,45-48] can serve as organocatalysts, there is little reason for the fields of organocatalysis and biocatalysis not to meet in the middle with the application of peptides which self-assemble into higher order structures. Consolidation via *de novo* peptide design could thereby result in a better understanding in biocatalysis and enzyme activity and broaden the practical scope of organocatalysis.

Hitherto, research in asymmetric organocatalysis has focused largely on bond-forming reactions involving aldehydes, ketones and ketenes as substrates and much less on reactions involving esters.[49] On the other hand, biocatalyzed stereoselective esterification and hydrolysis remain widely used in the production of fine-chemicals and the pharmaceutical synthesis of chiral drugs.[8] One advantage larger biocatalysts offer in comparison to small organic molecules is the provision of a molecular cavity.[50-53] A cavity allows for a multitude of additional interactions required for substrate recognition and transition state stabilization. In regard to selective ester hydrolysis, a biocatalyst or an artificial enzyme can differentiate between substrates

according to a property even if that property is well displaced from the point of hydrolysis. Such selectivity is less likely to be achieved by smaller organocatalysts.

This work delves into the field of designed catalytic proteins for reasons both fundamental and practical. Particular attention is focused on the “bottom-up” approach using *de novo* designed, self-assembling peptides as structural scaffolds to develop catalysts of ester hydrolysis. In doing so, peptides which adopt predictable α-helical and β-sheet assemblies provide a structural blueprint for the rational incorporation of catalytic activity and a clearer view of enzyme catalysis.

# 2

# Enzyme catalysis

Enzymes are extremely proficient biological catalysts which greatly accelerate reaction rates while discriminating between substrates and controlling the reaction outcome. With the exception of ribozymes,[54] catalytic RNA, enzymes are broadly considered as a class of proteins. As all catalysts, enzymes accelerate reaction rates by lowering the activation energy (Fig. 2.1). By lowering the energy of activation, catalysis affects the catalyzed reaction's kinetic properties without participating in the reaction's thermodynamics. Furthermore, as catalysts are not consumed in the course of the reaction, they can be recycled and are effective at low concentrations.

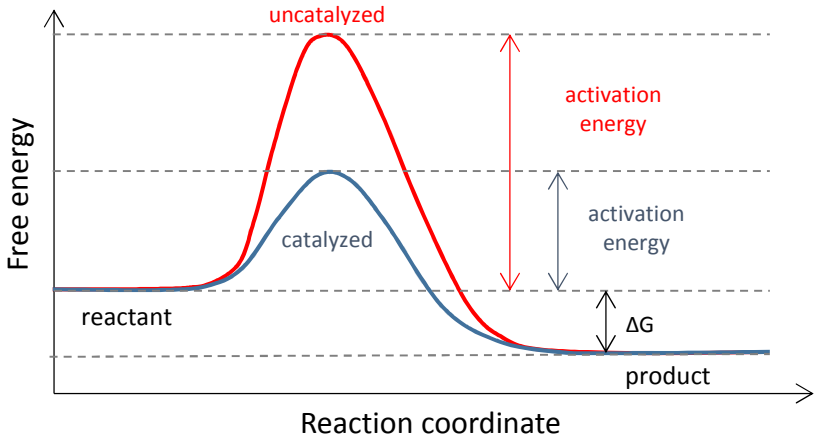



***Figure 2.1**: The effect of catalysis on the activation energy of a reaction.*

In order to increase the rate of a reaction, an enzyme like other catalysts must provide an alternative pathway to the reaction's highest energy transition state (i.e. activation energy). Catalytic mechanisms employed by enzymes to reduce energy barriers are commonly classified as acid-base catalysis, covalent catalysis, metal ion catalysis, catalysis through proximity and orientation effects, and preferred transition state binding.[3]

## 2.1 Catalytic mechanisms

### 2.1.1 Acid-base catalysis

Via proton transfer, acids and bases can activate electrophilic and nucleophilic reactants as well as stabilize transition states. The side chains of lysine, tyrosine, cysteine, histidine, glutamic acid and aspartic acid are capable of proton transfer under physiological conditions and are typically employed by enzymes in the acid-base catalysis of biochemical reactions. Due to its structural and functional versatility, an enzyme can create a local environment within its active site conducive to acid-base catalysis. Such effects are witnessed with the alteration of a particular residue's intrinsic pKa[55,56] (Table 2.1) or the arrangement of both acidic and basic residues around a substrate to promote concerted acid-base catalysis (Fig. 2.2).[3]

***Table 2.1:*** *Effects of local environment on pKa.*

| Local Microenvironment | Effect on pKa |
|---|---|
| Hydrophobic | Increase |
| Polar | Decrease |
| Adjacent to like charge | Increase |
| Adjacent to opposite charge | Decrease |
| Presence of salt bridges and hydrogen bonds | Decrease |

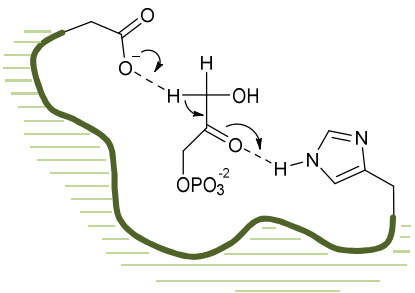



***Figure 2.2:*** *General acid-base catalysis as performed by triose-phosphate isomerase.*

### 2.1.2 Covalent catalysis

In covalent catalysis, a catalyst reduces the energy of a transition state by forming a transient covalent bond with the substrate. Enzymes perform covalent catalysis using activated groups within their active sites. Since enzymes typically initiate covalent catalysis with the nucleophilic side chains of serine, cysteine, lysine and histidine upon electrophilic substrates or cofactors, this mechanism is often referred to as nucleophilic catalysis. The transfer of electrons from the nucleophile to the substrate results in an electron sink, which allows for elimination. Upon elimination, the catalyst is restored to its original form.[3] An example of covalent catalysis is the cleavage of ester or amide bonds by serine proteases (Fig. 2.3).[13] In this example, an activated serine forms a covalent bond with the substrate resulting in an acyl-enzyme complex. The oxygen atom of serine serves as an electron sink in the subsequent hydrolysis. The mechanism presented in Figure 2.3 is an excellent example depicting how an enzyme may use multiple catalytic mechanisms to accelerate a reaction. In this case, general base catalysis is required to activate the nucleophile and stabilize the tetrahedral transition state. It is important to note that if the covalent bond formed by the catalyst is too stable, catalytic activity will be inhibited.

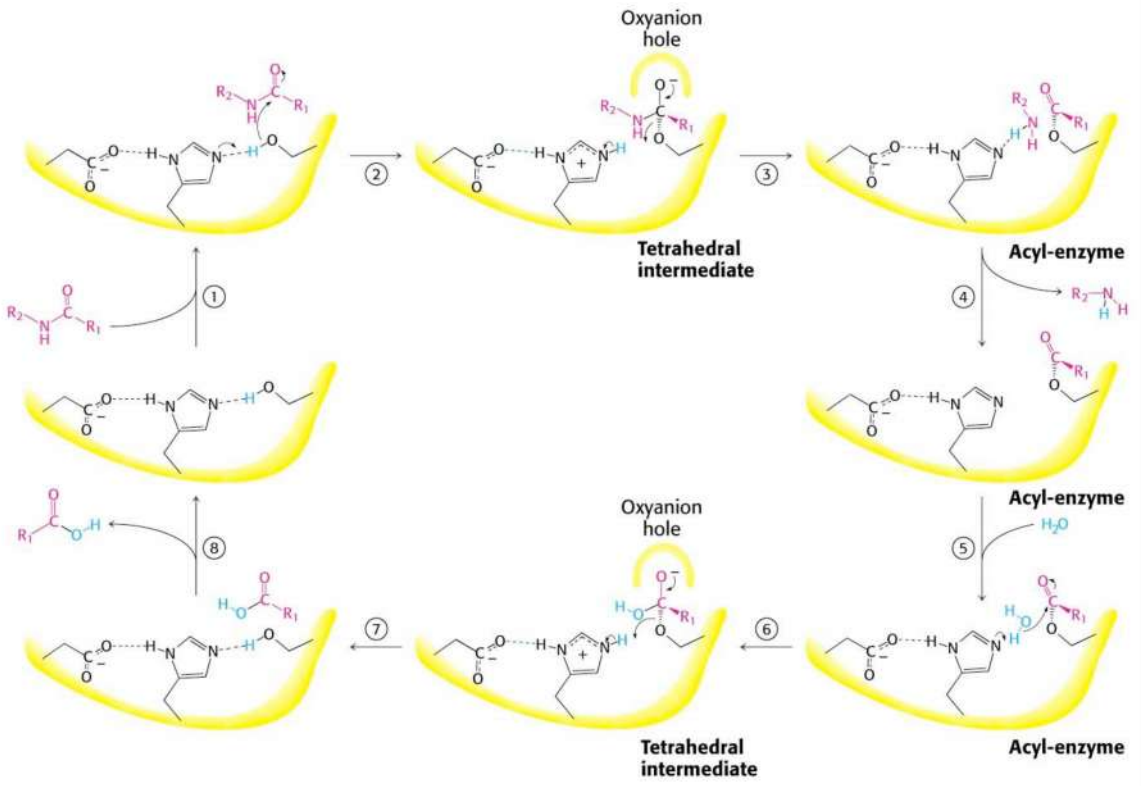



***Figure 2.3:*** *Mechanism of serine protease-catalyzed peptide bond cleavage. Adapted from Berg* et al.[13]

### 2.1.3 Metal ion catalysis

Many enzymes require metal ions as cofactors for catalytic activity. Metal ions impart activity either by playing a direct catalytic role in the active site or by stabilizing a precise three-dimensional structure required for catalysis.[57] Ions of transition metals iron, copper, manganese, and cobalt play important catalytic roles in active sites of metalloenzymes. Other metals ions such as $Na^+$, $K^+$ and $Ca^{2+}$ are often required to stabilize the active structure. $Zn^{2+}$ and $Mg^{2+}$ ions serve either structural or catalytic roles. Metal ions play a direct role in catalysis by binding substrates, mediating oxidation and reduction reactions and by stabilizing negative charges.[3] The metalloenzyme human carbonic anhydrase II (CA II) is a good example of how a metal ion can be utilized by an enzyme. CA II employs three histidine residues to coordinate $Zn^{2+}$, which in turn polarizes a water molecule (Fig. 2.4).[58,59] In CA II, this interaction lowers the pKa of water from 15.7 to 6.8. Since the lowered pKa value is associated with catalytic activity, it is a common parameter used to describe other enzymes or catalyst employing such a mechanism. The zinc-bound hydroxide nucleophile (ZHN) mechanism[60] is widely believed to be responsible for the CA II-mediated interconversion between carbon dioxide and a bicarbonate ion (Fig. 2.4B).[61]

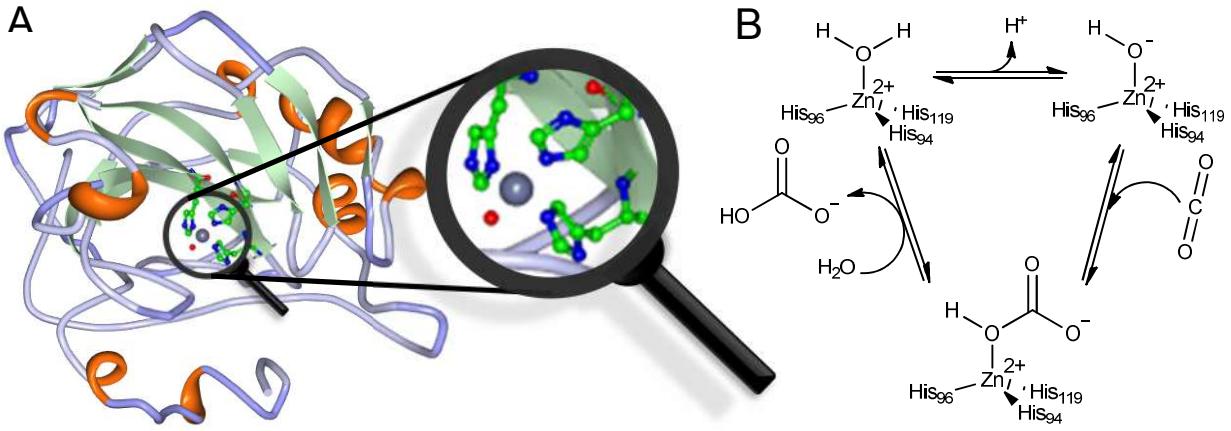



***Figure 2.4: (A)*** *Crystal structure of human CA II metal-binding motif. Active site shown with $Zn^{2+}$ (gray) coordinated by three histidine residues and a water molecule (oxygen, red). Crystallization performed by Avvaru* et al.[58] *and rendered with RCSB Protein Workshop.*[59] ***(B)*** *Proposed zinc-bound hydroxide nucleophile mechanism of carbonic anhydrase-catalyzed $CO_2$/bicarbonate ion interconversion.*

### 2.1.4 Catalysis through proximity and orientation effects

Like most proteins, enzymes provide particular groups and microenvironments in particular positions to favorably interact with their substrates. Rather than relying on random collisions in solution, enzymes lure their substrates into close proximity with catalytically reactive groups and/or other substrates. Furthermore, unlike traditional chemical catalysts, enzymes have been designed by evolution to stabilize their substrates in a precise orientation. By firmly holding its substrate in a reactive position, enzymes eliminate transitional and rotational motion. In other words, enzymes are capable of overcoming entropy loss associated with substrate organization by favorably binding its substrate and, with yet greater affinity, the reaction's transition state.

### 2.1.5 Catalysis though preferred transition state binding

Enzymes specifically bind substrates. However, this alone does not account for the remarkable turnover in which enzymes transform substrates into products. If a catalyst were to be so stably bound to its substrate, it would be unlikely for catalysis to occur. As is shown in Figure 2.1, the degree by which an enzyme enhances the rate of a reaction depends on its ability to stabilize the transition state in comparison to the substrate.[62] This has been evidenced by studies performed with transition-state analogs, which show preferential transition state binding.[63,64] Due to their chemical and structural flexibility, it is arguably by this mechanism that enzymes best exploit their advantage in terms of rate-enhancement over non-enzymatic catalysts.

## 2.2 Enzyme specificity and selectivity

Enzymes not only accelerate kinetically challenged reactions, but do so with remarkable substrate selectivity and reaction specificity. Once again these properties are due to the three dimensional structure of enzymes and in particular their active sites. Enzymes *select* which substrates to admit into their active sites. Selection can be limited by something as general as size or can be performed by more scrupulous chemoselective, regioselective or stereoselective means. In comparison to other types of catalysts, enzymes have various methods at their disposal to compensate for substrate desolvation.[65] Ultimately, the presence and placement of complementary hydrogen-bonding, electrostatic and/or hydrophobic groups in precise locations select which substrates are energetically feasible to desolvate.

Whereas selectivity deals with which substrate the enzyme chooses to interact with, enzyme specificity dictates which products are formed. Enzymes are stereospecific when they *specifically* catalyze the formation of one stereomeric reaction product in preference to another. An example of enzyme stereospecificity is shown in the lactate dehydrogenase-catalyzed transfer of a hydride from NADH to pyruvate (Fig. 2.5).[65] The orientation of both substrates within the three-dimension space of the enzyme active site results exclusively in a single enantiomer of lactate.

In addition to directing stereospecific products, enzymes are also reaction specific. Reactions catalyzed by enzymes in comparison with chemical catalysts generally occur with fewer side reactions. This arises from an enzyme's ability to promote a single pathway by thoroughly stabilizing one transition state.[3,65]

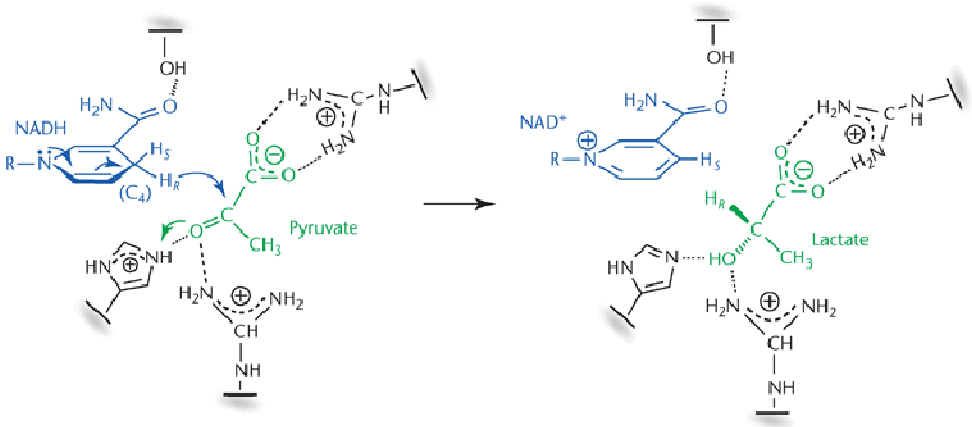



***Figure 2.5:*** *Mechanism of lactate dehydrogenase-mediated hydride transfer. Adapted from Hedstrom.*[65]

## 2.3 Enzyme kinetics

In order to study and assess the catalytic mechanisms enzymes impose upon their substrates, enzymologists commonly combine available structural knowledge with the discipline of enzyme kinetics. Enzyme kinetics is an invaluable tool used to characterize enzyme activity by determining the catalyzed rate of a reaction and how it responds to changes in experimental conditions. Just as enzyme kinetics is essential for our understanding of enzyme function and regulation under physiological conditions, it is also important for adapting enzyme behavior to non-physiological conditions.[66] Because enzyme kinetics is an extension of chemical kinetics, a brief summary of chemical kinetics and how it might be used to assess catalytic activity is made prior to a more detailed description of enzyme kinetics.

### 2.3.1 Chemical kinetics

In the chemical kinetics of a unimolecular or first-order reaction in which reactant S is converted to product P,

$$S \longrightarrow P$$

the velocity ($v$) is directly proportional to the concentration of reactant S

$$v = -d[S] / dt = k[S] \qquad [1]$$

in which the proportionality constant ($k$) is referred to as the rate constant. To obtain the first-order rate equation as a function of time, Equation 1 is rearranged

$$d[S] / [S] = -k\,dt \qquad [2]$$

and integrated from its initial concentration ($[S]_o$) to its concentration ([S]) at a specified time (t) resulting in a linear first-order rate equation.

$$\ln [S] = \ln [S]_o - kt \qquad [3]$$

A collision of two molecules S and B is a second-order reaction.

$$S + B \longrightarrow P$$

Its reaction rate depends on the concentration of both reactants S and B.

$$v = -d[S] / dt = -d[B] / dt = k[S][B] \quad [4]$$

To experimentally determine the rate constant of such a reaction, it is easier to employ one of the reactants in large excess, [B] >> [S]. Thus, the concentration of B would remain relatively constant in comparison to S. The concentration of B can therefore be treated as a constant and included with the rate constant ($k$) to become the apparent rate constant ($k'$). In this case, Equation 4 becomes

$$v = -d[S] / dt = k[S][B] = k'[S] \quad [5]$$

Since this reaction now resembles Equation 1 of the first-order reaction, it is referred to as a pseudo-first-order reaction. Following integration, the pseudo-first-order rate equation is obtained as

$$\ln [S] = \ln [S]_o - k't \quad [6]$$

Although hydrolysis reactions are of the second order (i.e. collision of water with ester), they are more conveniently studied under pseudo-first-order conditions with a large excess of water. This will be observed later in Section 8.2.3.

Chemical kinetics may also be applied to simple reactions catalyzed by enzymes.[67]

$$S \xrightarrow[E]{} P$$

In the reaction above, substrate S is converted into product P in the presence of enzyme E. This reaction can be treated as a first-order reaction as long as the substrate concentration is so low that it does not saturate the enzyme or catalytic species.[68-70] This limiting condition ensures that the reaction is first-order in [S] and not of the zeroth-order. In a simplified example of a catalyzed second-order reaction such as ester hydrolysis in which S represents the ester substrate,

$$S + H_2O \xrightarrow[E]{} P$$

pseudo-first-order conditions ([$H_20$] >> [S]) as shown in Equations 4-6 along with the limiting substrate concentration can be applied (Section 8.2.3).[71-75]

### 2.3.2 Enzyme kinetics and the Michaelis-Menten equation

Under biological conditions or conditions applied in synthetic chemistry, enzymes/catalysts are typically and most efficiently utilized in low concentrations in comparison to their substrates. When the concentration of substrate is much greater than enzyme ([S] >> [E]), the rate is zero order with respect to substrate and is limited by the concentration of enzyme. This clearly suggests that the enzyme and the substrate must form a complex (ES) as shown in the reaction scheme

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} P + E \qquad [7]$$

The first reaction in which the enzyme-substrate complex ES is formed is characterized by the forward ($k_1$) and reverse ($k_{-1}$) rate constants. The rate constant ($k_2$) is associated with the rate of the second reaction in which the enzyme-substrate complex is converted to product (P) and enzyme. For simplicity it is assumed that product formation is not reversible.

The Michaelis-Menten equation is derived to describe the rate of the enzyme-catalyzed Scheme 7 in terms of substrate concentration. From Equation 7, the velocity of product formation can be expressed as

$$v = d[\mathrm{P}] / dt = k_2[\mathrm{ES}] \qquad [8]$$

In turn, the formation of ES must be taken into account. This is performed through the summation of the individual rates of ES formation and decomposition

$$d[\mathrm{ES}] / dt = k_1[\mathrm{E}][\mathrm{S}] - k_{-1}[\mathrm{ES}] - k_2[\mathrm{ES}] \qquad [9]$$

To integrate Equation 9, Michaelis-Menten kinetics approximates that [ES] remains constant throughout the reaction. This is what is known as the steady-state approximation,

$$d[\mathrm{ES}] / dt \approx 0 \qquad [10]$$

which is only possible when an excess of substrate exists. For the majority of the catalyzed process in Equation 7, this approximation holds true except when the reaction is just begun (pre-steady-state) and once the substrate is sufficiently

exhausted (Fig. 2.6).[3] Therefore, [ES] is approximately constant and Equation 9 becomes

$$k_1[\mathrm{E}][\mathrm{S}] = k_{-1}[\mathrm{ES}] + k_2[\mathrm{ES}] \qquad [11]$$

Since neither [E] nor [ES] can be easily measured at any time during the reaction, the initial amount or total amount of enzyme $[\mathrm{E}]_\mathrm{T}$

$$[\mathrm{E}]_\mathrm{T} = [\mathrm{E}] + [\mathrm{ES}] \qquad [12]$$

is substituted in the form of $[\mathrm{E}] = [\mathrm{E}]_\mathrm{T} - [\mathrm{ES}]$ into Equation 11, which, following rearrangement, yields

$$\frac{([\mathrm{E}]_\mathrm{T} - [\mathrm{ES}])[\mathrm{S}]}{[\mathrm{ES}]} = \frac{k_{-1} + k_2}{k_1} \qquad [13]$$

At this point, the right side of Equation 13, which is made up of the individual rate constants of equation 7, can be combined to form the Michaelis constant ($K_M$).

$$K_M = \frac{k_{-1} + k_2}{k_1} \qquad [14]$$

Upon insertion of $K_M$ into Equation 13 and algebraic rearrangment, the value of [ES] is established as

$$[\mathrm{ES}] = \frac{[\mathrm{E}]_\mathrm{T}\,[\mathrm{S}]}{K_M + [\mathrm{S}]} \qquad [15]$$

With the insertion of [ES] from Equation 15 into the reaction velocity Equation 8 at t = 0, the equation for the initial velocity ($v_o$) is obtained.

$$v_0 = (d[\mathrm{P}] / dt)_{t=0} = \frac{k_2[\mathrm{E}]_\mathrm{T}\,[\mathrm{S}]}{K_M + [\mathrm{S}]} \qquad [16]$$

Earlier, the reaction scheme in Equation 7 was simplified by assuming that the reversible reaction of product back to substrate did not occur. This is experimentally possible by measuring the initial velocity. If [S] >> [E], as is likely at an early time point in the reaction, the steady-state approximation is valid. The initial velocity of an enzyme-catalyzed reaction should be measured within the first 10% of substrate conversion while product formation or substrate conversion is linear (zeroth-order) (Fig. 2.6) to avoid complications which may arise on account of reversible reactions or product inhibition.

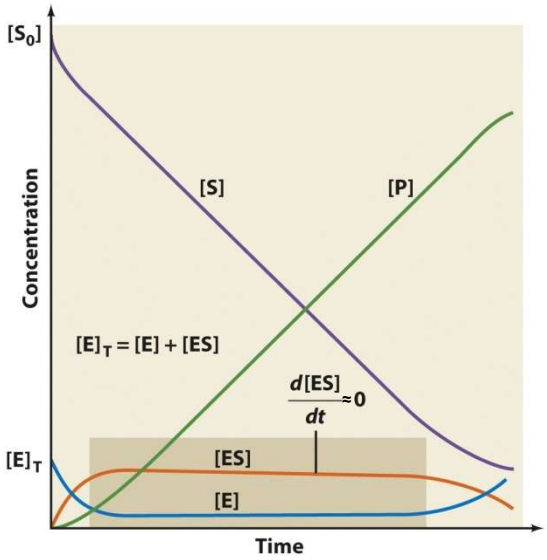



***Figure 2.6:*** *Relative concentration curves of substrate [S], enzyme [E], enzyme-substrate complex [ES] and product [P] during the course of an enzyme-catalyzed reaction. The shaded block shows the Michaelis-Menten steady-state approximation ($d$[ES]/$dt$) ≈ 0). After Voet, Voet & Pratt.*[3]

Once substrate concentration becomes high enough to saturate the enzyme ($[E]_T$ = [ES]), maximal reaction velocity ($V_{max}$) is approached.

$$V_{max} = k_2[E]_T \qquad [17]$$

Insertion of Equation 17 into Equation 16 results in what is generally referred to as the Michaelis-Menten equation plotted in Figure 2.7.[3,67]

$$v_o = \frac{V_{max}[S]}{K_M + [S]} \qquad [18]$$

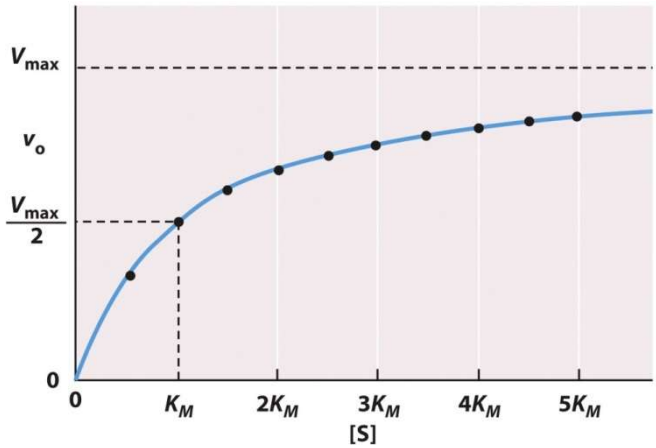



***Figure 2.7:*** *Initial velocity ($v_o$) versus substrate concentration [S] of an enzyme-catalyzed reaction according to the Michaelis-Menten equation. After Voet, Voet & Pratt.*[3]

### 2.3.3 Michaelis-Menten parameters

The Michaelis-Menten Equation 18 describes the kinetic behavior of enzymes and enzyme-like catalysts which exhibit saturation kinetics (Fig. 2.7). The parameters discussed below are commonly used by biochemists when relating and comparing the activities of enzymes.

The **Michaelis constant $K_M$** is mathematically defined as the substrate concentration when the initial velocity of a reaction is half the value of $V_{max}$ (Fig. 2.7). $K_M$ is inversely proportional to the fraction of total enzyme $E_T$ participating in the enzyme-substrate complex ES. This relationship has often led to the common misconception that the value $K_M$ is a measure of a substrate affinity to an enzyme. In the simple two-step process shown in Equation 7, $K_M$ can be used as an indicator for substrate affinity only when $k_2 << k_{-1}$. However, this is often not the case. The Michaelis constant is best interpreted by how the initial velocity $v_o$ of the catalyzed reaction responds as the concentration of the substrate [S] is increased (i.e. the steepness of the Michaelis-Menten plot, Fig. 2.7).[3,67]

As already mentioned in the derivation of the Michaelis-Menten equation, the **maximum velocity $V_{max}$** is the initial velocity of a reaction when the total enzyme is involved in an enzyme-substrate complex. In the simple reaction scheme of Equation 7, $V_{max}$ is defined as $k_2[E]_T$ and is the zeroth-order rate constant.[67]

The **turnover number**, $k_{cat}$, is a more applicable rate constant since it is independent of the enzyme concentration $[E]_T$. The turnover number is simply the rate at which a single enzyme active site converts substrate into product. In Equation 7, $k_{cat}$ is equal to $k_2$. However, in more complicated processes, it is the rate constant of the rate-limiting step or a more complicated combination of multiple rate-limiting steps. The turnover number is the first-order rate constant and is computed as $V_{max} / [E]_T$.[3,66,67]

When used alone, $k_{cat}$ indicates only the rate of substrate turnover and little in regard to the binding interaction between enzyme and substrate or the degree by which the enzyme accelerates a reaction.[65,66] By itself, the value of $K_M$ offers no information about the turnover rate.[65] However, when used together, $k_{cat}/K_M$ is the most conclusive constant used to compare the activity of different enzymes with different substrates. This is best shown by the Michaelis-Menten equation when [S] << $K_M$, which means that most of the enzyme is unbound, the initial rate becomes

$$v_0 = \frac{k_{\text{cat}}}{K_M}[\text{E}][\text{S}] = \frac{V_{\text{max}}}{K_M}[\text{S}] \qquad [19]$$

The parameter $k_{cat}/K_M$ is, therefore, the second-order rate constant for the conversion of free enzyme and substrate into product. Since $k_{cat}/K_M$ accounts for both substrate specificity and turnover, it is often referred to as a measure of catalytic efficiency[3,13] or the specificity constant.[66,67] The diffusion rate in which enzyme and substrate encounter one another imposes an upper limit ($10^8$-$10^9$ $M^{-1}$ $s^{-1}$) on values of $k_{cat}/K_M$. Michaelis-Menten parameters of several enzymes are listed in Table 2.2.[76]

Parameters $V_{max}$ and $K_M$ are in practice difficult to ascertain directly from the Michaelis-Menten Equation 18. This is often due to the high concentration of substrate required to approach $V_{max}$. Parameters $K_M$, $V_{max}$ and $k_{cat}$ can, however, be determined by transforming the Michaelis-Menten equation into the Lineweaver-Burk equation.

$$\frac{1}{v_o} = \left(\frac{K_M}{V_{\text{max}}}\right)\frac{1}{[\text{S}]} + \frac{1}{V_{\text{max}}} \qquad [20]$$

Parameters can be graphically determined by plotting experimental values $1/v_o$ versus 1/[S]. From the resulting y-intercept and slope of the linear Lineweaver-Burk plot, paramenters of $V_{max}$ and $K_M$ are determined as shown in Figure 2.8.[77]

***Table 2.2:*** *Values of $K_M$, $k_{cat}$ and $k_{cat}/K_M$ for various enzymes with corresponding substrate. From Voet, Voet & Pratt.* [75]

| Enzyme | Substrate | $K_M$ (M) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_M$ ($M^{-1} \cdot s^{-1}$) |
|---|---|---|---|---|
| Acetylcholinesterase | Acetylcholine | $9.5 \times 10^{-5}$ | $1.4 \times 10^{4}$ | $1.5 \times 10^{8}$ |
| Carbonic anhydrase | $CO_2$ | $1.2 \times 10^{-2}$ | $1.0 \times 10^{6}$ | $8.3 \times 10^{7}$ |
| | $HCO_3^-$ | $2.6 \times 10^{-2}$ | $4.0 \times 10^{5}$ | $1.5 \times 10^{7}$ |
| Catalase | $H_2O_2$ | $2.5 \times 10^{-2}$ | $1.0 \times 10^{7}$ | $4.0 \times 10^{8}$ |
| Chymotrypsin | *N*-Acetylglycine ethyl ester | $4.4 \times 10^{-1}$ | $5.1 \times 10^{-2}$ | $1.2 \times 10^{-1}$ |
| | *N*-Acetylvaline ethyl ester | $8.8 \times 10^{-2}$ | $1.7 \times 10^{-1}$ | 1.9 |
| | *N*-Acetyltyrosine ethyl ester | $6.6 \times 10^{-4}$ | $1.9 \times 10^{2}$ | $2.9 \times 10^{5}$ |
| Fumarase | Fumarate | $5.0 \times 10^{-6}$ | $8.0 \times 10^{2}$ | $1.6 \times 10^{8}$ |
| | Malate | $2.5 \times 10^{-5}$ | $9.0 \times 10^{2}$ | $3.6 \times 10^{7}$ |
| Urease | Urea | $2.5 \times 10^{-2}$ | $1.0 \times 10^{4}$ | $4.0 \times 10^{5}$ |

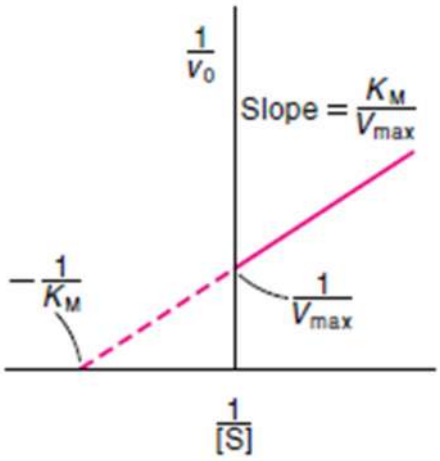



***Figure 2.8:*** *Lineweaver-Burk or double-reciprocal plot. From Chang.*[77]

# 3

# *Biocatalysis*

Enzymes are highly evolved catalysts with complex structures and remarkable properties. Regardless of our level of understanding, enzymes will continue to function with superior activity and selectivity within their natural environment. However, in order to better exploit the use of enzymes in biocatalysis, more knowledge is needed.

Biocatalysis is the application of whole cells or enzymes as catalysts in synthetic chemistry. Although more recently defined, biocatalysis predates recorded history when it was used to produce beer, wine and cheese by means of fermentation. A more scientific knowledge of biocatalysis arose in the middle of the 19$^{th}$ century when Louis Pasteur discovered biological enantioselectivity by resolving racemic tartaric acid with penicillin.[78] Later in the 19$^{th}$ century, Eduard Buchner fermented sugar with the use of cell-free yeast extracts, which revealed that living cells do not hold a monopoly on biological transformations.[79] Despite advances made in enzyme purification in the mid-20$^{th}$ century, the use of enzymes in synthetic chemistry lagged behind breakthroughs in enzyme crystallography and the elucidation of their mechanisms and involvement in biological pathways.[8,80] It was not until the late 1970s that biocatalysis became a trend due to the rapidly increasing accessibility of enzymes brought about by recombinant DNA methodology. By the 1990s, many enzymes became commercially available and widely used in industrial processes.[81]

## 3.1 The pros and cons of biocatalysis

Enzymes are complex molecules in comparison to traditional catalysts employed in organic synthesis and many of their means of biotransformation remain unclear and difficult to study. The synthesis of a product without knowing the mechanism may seem like mere hand-waving to many chemists. Moreover, enzymes possess properties that could be seen as both advantageous and disadvantageous when applied to organic synthesis. It would be unfair to say that many of the perceived disadvantages are based on misconceptions and prejudice without taking into account that many advantages are grounded on potential success or studies which require no shortage of effort. Therefore, this section could have just as fittingly been called *the potential and misconceptions of biocatalysis*.

### 3.1.1 Advantages of biocatalysis

*A) Catalytic efficiency*

Enzymes typically enhance reaction rates by factors of $10^6$-$10^{12}$ and as high as $10^{20}$ in comparison to non-catalyzed reactions.[3,5] In addition, they often provide effective catalysis when present in mole percentages of $10^{-3}$-$10^{-4}$.[80]

*B) Substrate selectivity*

For the most part, enzymes are *chemoselective* in that they often catalyze reactions involving a specific functional group. Furthermore, the three-dimensional structure and available binding surface within a cleft allow enzymes to *regioselectively* differentiate between the same functionality located at different positions on a substrate molecule. A good example of enzyme regioselectivity is shown in the subtilisin-catalyzed hydrolysis of benzyl esters of aspartic and glutamic acid (Fig. 3.1).[82] Since enzymes are sequences of L-amino acids, they are chiral catalysts and can distinguish between chiral substrates. Enzymes are *enantioselective* when they catalyze one enantiomeric form of a substrate at a higher velocity than the other. Therefore, enzymes are often applied to kinetically resolve racemic substrate mixtures. An example of enzymatic kinetic resolution's utility in synthesis is depicted in Figure 3.2.[83] Another example of enzymatic kinetic resolution is the convenient

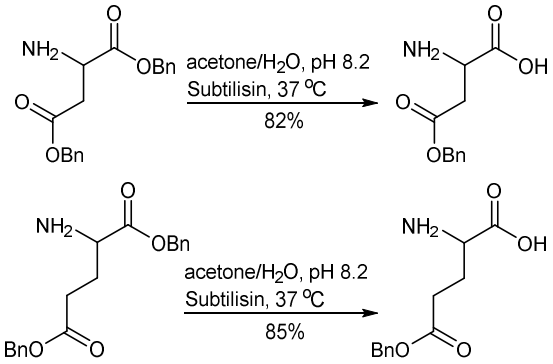



***Figure 3.1:*** *Regioselective hydrolysis of aspartic and glutamic acid benzyl esters catalyzed by subtilisin. Adapted from Chen & Wang.*[82]

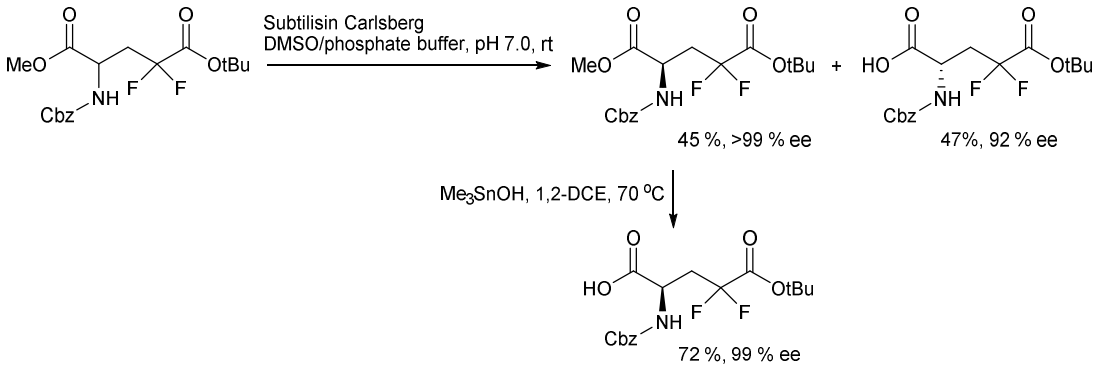



***Figure 3.2:*** *Enzymatic kinetic resolution of orthogonally protected D- and L- 4,4,-difluoroglutamic acid. From Li & Miller.*[83]

one-pot system introduced by Monteiro *et al*. to resolve racemic secondary alcohols (Fig. 3.3). In this system, fatty esters are used as both solvent and as an acylating shuttle between an enantioselective lipase-catalyzed transesterification and regenerative transesterification. Enantiopure alcohol is distilled after each transesterification.[84] Unfortunately, as described in Section 3.1.2A, superior enzyme selectivity is often viewed as a double-edged sword when applied in synthesis.

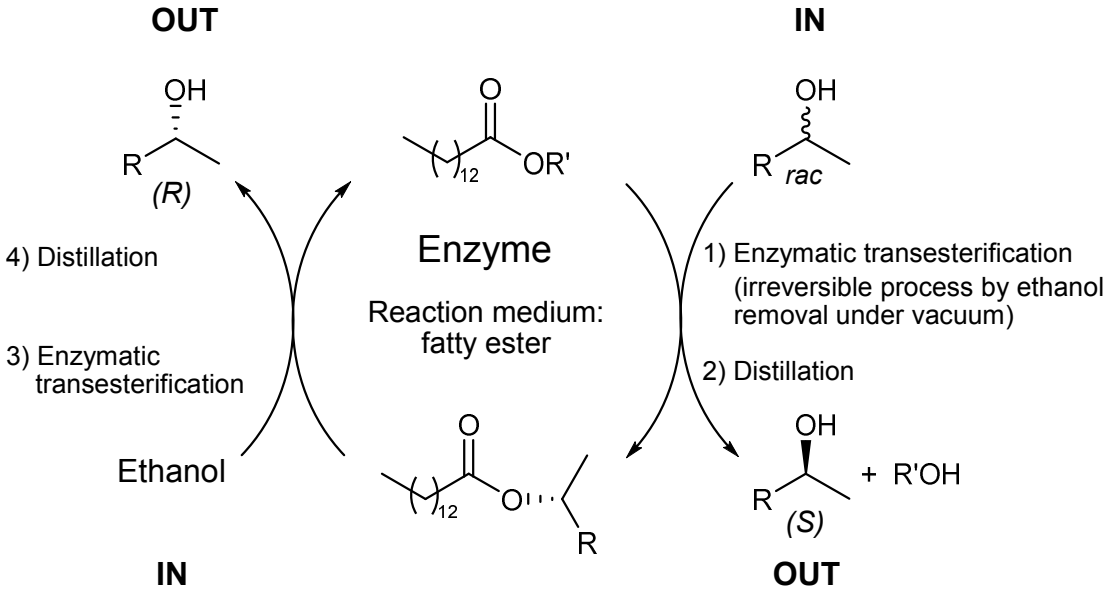



***Figure 3.3:*** *One-pot* Candida antarctica *lipase C-mediated kinetic resolution of secondary alcohols using fatty esters as solvent and acylating agent. R' = ethyl. Adapted from Monteiro* et al.[84]

*C) Product specificity*

Not only are enzymes regioselective and stereoselective, but the large enzyme to substrate surface interaction offers a high potential for *regio- and stereo-control*. In other words, within the cleft of an enzyme, a prochiral substrate is positioned so the catalytic machinery can encounter one area (regiocontrol) and one face (stereocontrol) of the substrate.[5,81] A biocatalytic example of product specificity is asymmetric transfer hydrogenation catalyzed by oxidoreductases (Fig. 2.5).[65]

*D) Catalytic activity under mild conditions*

Naturally, enzymes evolved to function under physiological conditions. Therefore, biocatalysts can be advantageously used on substrates sensitive to heat and harsh pH conditions. Enzymes function best at temperatures between 20-40$^{o}$C and at neutral pH values. However, when enzymes are not robust enough for certain environments, this property can also be viewed as a shortcoming.

*E) Nontoxic and environmentally acceptable*

Unlike metal or acid and base catalysts, enzymes are nontoxic and pose few health hazards. As is the case with all catalysis, biocatalysis greatly reduces the amount waste products resulting from stoichiometric chemistry, which has a tremendous impact at the industrial level. In addition, enzyme selectivity decreases the number of reactions required for orthogonal protection schemes. As enzymes are protein, they are completely biodegradable.[7,85]

As expressed above, it is difficult to argue that enzymes do not present potential advantages when employed in synthetic chemistry. However, these advantages are based only on the examples of established uses and the inevitable likelihood of discovering new uses. At present, a synthetic chemist would not rely on the everyday use of biocatalysts. This does not mean he could not, but simply that the field of biocatalysis has not approached a point in which it would be practical. As we begin the next section which focuses on the disadvantages of biocatalysis, it is important to keep in mind that many of these disadvantages are not directed at biocatalysis itself, but at the insufficient knowledge we have of enzymes, especially when removed from physiological conditions.

### 3.1.2 Disadvantages of biocatalysis

*A) Enzymes are* too *selective*

A common misconception about enzymes is that they only catalyze reactions involving their natural substrates. Since enzyme-functioning has evolved according to specific metabolites, it is not surprising that a preference for certain substrates exists. It is true that some enzymes possess much narrower substrate scopes than others.

However, enzymes often have a broad substrate range unless such a property is biologically detrimental. The catalytic activity of carbonic anhydrase is not limited to the interconversion of carbon dioxide and bicarbonate (Fig. 2.4), but can also be extended to ester hydrolysis (Fig. 3.14). The reversibility of enzymes should also not be ruled out. Under specific laboratory conditions, proteases have been shown to catalyze peptide synthesis.[86,87] Some remarkably promiscuous enzymes have been found to catalyze reactions in the laboratory with which they have no natural relation. For instance, *Candida antarctica* lipase B (CALB), known to catalyze ester hydrolysis in nature and as a biocatalyst (Fig. 3.3), catalyzes C-C bound forming aldol and Michael additions.[5,12,88] Furthermore, porcine pancreas lipase (PPL) has also been shown to catalyze asymmetric aldol reactions (Fig. 3.4).[89] It is perhaps incorrect to state that the biocatalytic substrate or reaction scope is limited, but a delicate balance between substrate selectivity and scope exists and is difficult to generalize according to the type of enzyme or reaction.

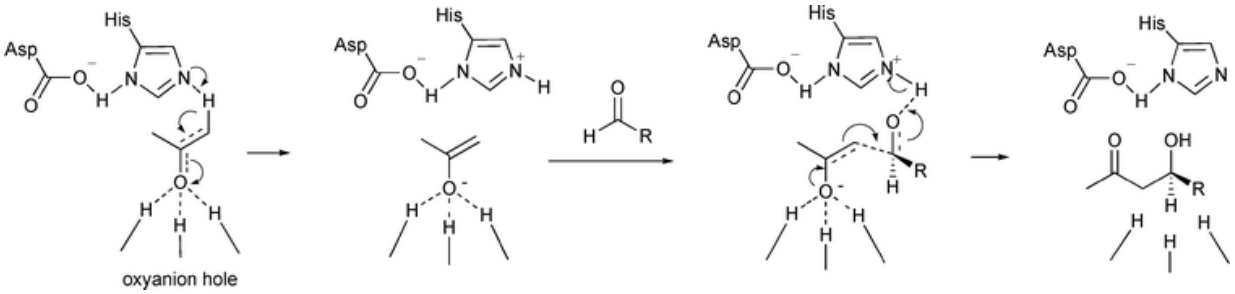



***Figure 3.4:*** *Proposed mechanism of lipase-catalyzed asymmetric aldol reaction. From Li et al.*[89]

*B) Enzymes are operative under few conditions*

Enzymes are developed by nature to operate under physiological conditions. In biocatalysis, many enzymes are functional under mild conditions. It is well known that the structure and activity of enzymes and protein in general is greatly affected by its environment. Enzymes generally become less catalytically active in organic solvents, at temperatures above 40°C and at pH conditions below 5 and above 9.[66,90] A common misconception is that enzymes are too sensitive or too unstable for use in synthetic conditions.[81] Nonetheless, if handled correctly, enzymes are far less sensitive than *t*-butyl lithium or Raney nickel.

Since many organic reactants are poorly dissolved or unstable in water, it would be beneficial to apply enzyme catalysis in an organic medium. Studies have shown that enzymes retain a portion of their activity in anhydrous media. Furthermore, in non-aqueous environments, pH is no longer an issue and enzymes often exhibit increased thermostability. Interestingly, enzymes have reportedly been used in organic solvents to impart increased stereocontrol, and even to direct stereocontrol and alter enantioselectivity according to the selection of organic solvent.[91]

The loss of enzymatic activity in anhydrous media is said to be largely due to increased enzyme rigidity. In fascinating studies by Klibanov[92] and others,[93,94] it was shown that enzymes have what is called a "molecular memory"[95] in anhydrous solvents (Fig. 3.5).[96] This effect is shown when an enzyme in aqueous solution with

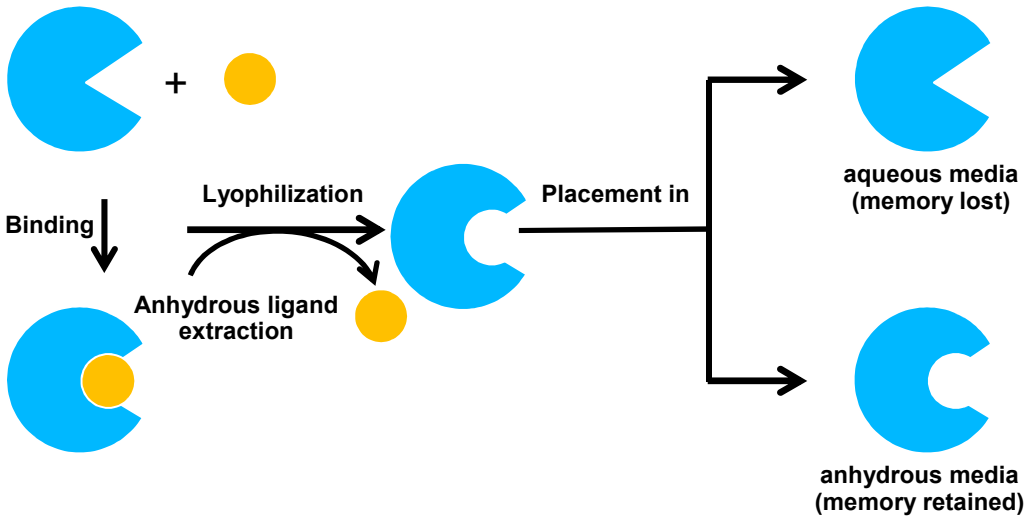



***Figure 3.5:*** *An enzyme's "molecular memory" of binding ligand in anhydrous solvent following lyophilization and ligand extraction. Adapted from Budisa.*[96]

particular ligands or a competitive inhibitor is lyophilized. After lyophilization and removal of the ligand by non-aqueous extraction, the enzyme by remaining rigid in anhydrous media possesses an increased selectivity for substrates resembling the initially “imprinted” ligand. Enzymes also maintain their molecular memory in supercritical carbon dioxide, an effective solvent gathering increased interest in biocatalysis.[96,97] Discoveries such as these greatly widen the operational scope of enzymes in synthetic chemistry.

*C) Substrate and product inhibition*

Within the cell, an elaborate network of metabolic pathways is required to maintain a fine chemical balance. Key components of such pathways are enzymes which do not operate by simply transforming substrate to product at high speeds, but by responding to substrate levels and the needs of the cell. Enzymes are, therefore, prone to both substrate and product inhibition. Substrate inhibition is a mechanism used to maintain a required metabolic flux.[98] In biocatalysis, substrate inhibition can often be circumvented by keeping the substrate concentration level low with incremental addition. Since products are in many cases similar to substrates, it is not surprising that they might produce inhibitory effects on enzymes. Product inhibition, or negative feedback, results when product concentrations become too high. Such a mechanism is physiologically important because it helps conserve the substrate when further conversion is not required. In biocatalysis, product inhibition presents a more complicated problem. To avoid product inhibition, product must be continually removed.[81]

Regardless of the current disadvantages or misconceptions of biocatalysis, it is becoming increasingly apparent that the use of biocatalysts represents a practical and efficient alternative in synthetic chemistry. Until this point, while explaining what biocatalysis is and its perceived advantages and disadvantages, specific examples have been given in which enzymes have been applied in synthetic chemistry. In some of the examples, the conditions of the reaction are changed to accommodate the enzyme. However, the complexity of enzymes offers many opportunities to optimize such properties as stability, substrate compatibility, enantioselectivity and stereocontrol. The next section focuses on how enzymes are modified to overcome limitations and bolster their application in synthetic chemistry.

## 3.2 Enzymes modified for biocatalysis

The first wave of biocatalysis began with Ludwig Rosenthaler's application of emulsin extracted from almonds in the asymmetric conversion of (R)-mandelnitrile from benzaldehyde and hydrogen cyanide and ended in the mid-1970s with the advent of recombinant DNA methods. During this initial period, biocatalysts were less accessible and, therefore, less affordable. In order to save on costs, biocatalysts were immobilized to make them more stable and reusable.[8,14]

An increased availability of enzymes due to recombinant DNA techniques ushered in the second wave of biocatalysis. In the mid-1970s, it became possible to overexpress desired enzymes and even mutated enzymes in host organisms. As a result, not only was there an increase in the use of biocatalysts, but also a means to study structure-function relationships in biocatalysis. The new field of protein engineering was amplified with the development of the polymerase chain reaction (PCR), which greatly simplified DNA synthesis and mutagenesis.[99,100] By the 1990s, the use of biocatalysts was no longer limited to industry, but becoming increasingly common in the production of pharmaceuticals and fine chemicals.[8,14]

In the early 1990s, mutagenesis of wild-type enzymes began to increase and expand outward from active-sites. It was soon discovered that a large fraction of residues were not necessary for general enzyme structure or function.[101] Site-directed mutagenesis is an adequate technique for learning about enzyme structure and function, but was believed by many to be an ineffective method for quickly adapting enzymes to synthetic needs (e.g. chiral pharmaceuticals). Before long, the third wave of biocatalysis[14] broke with the creation of large libraries of mutant enzymes, requiring increased screening for desired properties. This repetitive process is commonly referred to as directed or *in vitro* evolution. The engineering of biocatalysts no longer looked to overcome enzyme limitations, but to fine-tune biocatalytic activity according to process specifications.

### 3.2.1 Biocatalyst immobilization

Original efforts to immobilize enzymes were made to save money. Immobilization prolongs the lifetime and facilitates the reusability of enzymes. Today, enzymes are commonly immobilized for industrial use not only for cost efficiency and convenient purification (filtration), but also to reduce wastes. In addition, enzyme immobilization is becoming increasingly applied with multiple enzymes at different points on solid support or cross-linked to mirror pathway or compartmentalization processes. Enzymes are generally immobilized though solid support binding, entrapment or cross-linking.[102,103]

Enzymes can be immobilized to a solid support by means of adsorption or covalent bond. Supports consist of organic polymers (e.g. acrylic resins), biopolymers (e.g. water-insoluble polysaccharides) and inorganic solids (e.g. silica, zeolites). A commercially available example of a support-immobilized enzyme is Novozym 435, *C. antarctica* lipase B adsorbed to acrylic resin. As opposed to support binding, entrapment immobilizes enzymes by enclosing them within organic polymers, gelatinous networks or membrane-like structures. When an inert carrier is unwanted, immobilization can be performed by crosslinking enzymes. Enzymes can be cross-linked with an agent such as glutaraldehyde and maintain activity as crystals (cross-linked enzyme crystals, CLECs) or precipitate (cross-linked enzyme aggregates, CLEAs).[102,103] These strategies of enzyme immobilization along with efforts made in protein engineering are commonly applied to tailor biocatalysts.

### 3.2.2 Biocatalysts via traditional rational design

Early efforts to refine enzyme properties for biocatalysis adopted the rational design approach of protein engineering. This approach, however, is limited by the amount and quality of structural knowledge available.[104] Therefore, much of rational design research has focused on enzyme active sites,[105] although site-directed mutagenesis beyond the active site has been used successfully to increase stability.[106] A good example of the use of rational design outside of the active site is the 8-fold site-directed mutation of thermolysin, which maintained proteolytic activity at 100$^{o}$C as well as in the presence of denaturing agents.[107]

Provided with a three-dimensional map of the active site, site-directed mutagenesis can be effectively used to broaden substrate specificity[108] and directly alter the catalytic mechanism.[109,110] To a much lesser extent, site-directed mutagenesis has also been used to initiate enantioselectivity.[111] Rational design can also take a more piecewise approach.[105] Biocatalysts have been designed by incorporating the catalytic machinery of one enzyme into templates forming a type of hybrid enzyme. Such a template might provide advantageous solubility, binding or stability. An example of such modular rational design is the chimeric bacterial-human cytochrome P450 engineered to combine the high activity of the human enzyme with the beneficial solubility of the bacterial enzyme.[112]

As mentioned in Section 2.1.5, enzymes bind best with the catalyzed reaction's transition state. In the 1980s, another rational method was developed in which heptans analogous to a desired reaction's transition state were recruited to initiate an immune response.[113,114] Although catalytic antibodies, abzymes, actively bind the transition state, their catalytic activity is generally low due to an absence of necessary catalytic machinery.[8] Convergent rational design in which catalytic machinery is grafted into antibodies has shown promising results. This can be seen in research by Lee in which catalytic machinery in the form of glutamate, lysine and histidine were strategically grafted via site-directed mutagenesis into a structurally resolved antibody to introduce proteolytic activity (Fig. 3.6).[115]

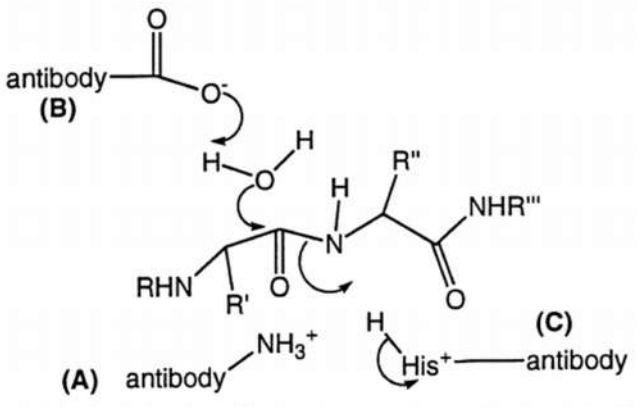



***Figure 3.6:*** *Antibody converted into a protease. Proposed mechanism resulting from the introduction of residues glutamate (B), lysine (A) and histidine (C) into an antibody via site-directed mutagenesis. In the reaction, glutamate activates a water molecule for nucleophilic attack. Histidine serves as a proton donor for the eliminated amine. Lysine stabilizes the oxyanion intermediate. Courtesy of Liu* et al.[115]

The schematic representation in Figure 3.7 summarizes possible strategies used by rational design to fine-tune biocatalysts from a template protein of resolved structure.[105] It is commonly said that the use of rational design is restricted by our knowledge of protein structure. Since design is based on structure, this view is partially correct. However, it could be more accurate to say that the true limitation of rational design is how little is known about folding dynamics. Most of the outcomes of rational design do not match expectations. A shift of a single angstrom within the active site or an amino acid substitution on a remote loop may make all the difference in regard to catalytic activity. As Daniel Koshland wrote, "perfecting the molecule requires such precision in very small changes that our theory and experiments are strained to make logical predictions."[116] Often it is more effective to make many small changes at random and select those that work.

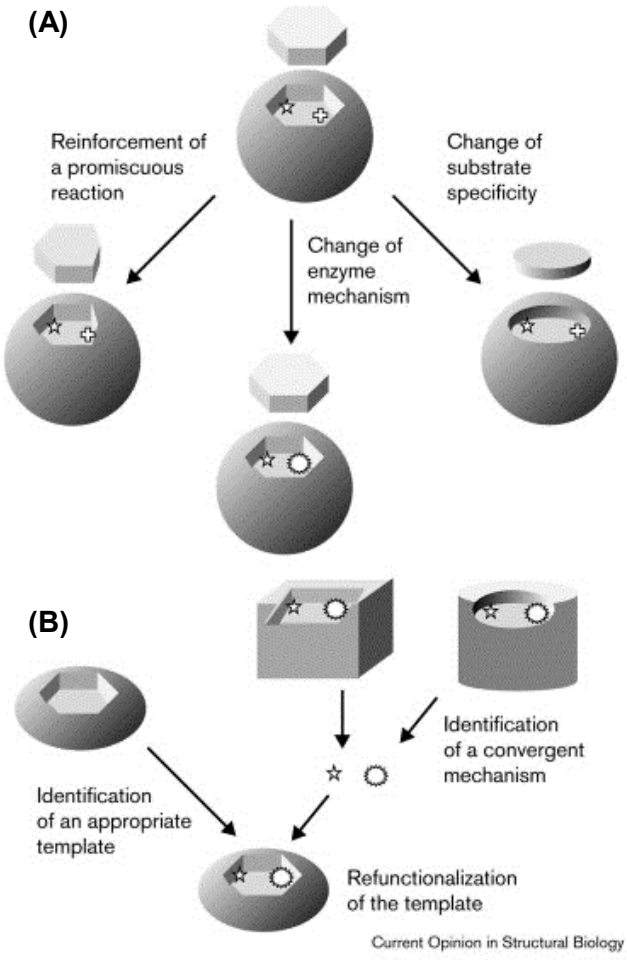



***Figure 3.7:*** *Possible strategies to rationally redesign biocatalysts.* ***(A)*** *Three routes to form a novel biocatalyst from a pre-existing enzyme.* ***(B)*** *Modular approach in which independent enzyme features are assembled to create a novel biocatalyst. The cross and star symbols represent amino acid functionalities required for a specific catalytic activity. Adapted from Cedrone* et al.[105]

### 3.2.3 Biocatalysts via directed evolution

Traditional rational design techniques such as site-directed mutagenesis effectively aid in the elucidation of catalytic mechanisms. However, when applied alone to create novel biocatalysts, rational design is often inefficient. Even if the complex structure of a protein were known, the results of mutation regardless of its preconceived simplicity are difficult to predict. For this reason, biocatalyst research relies on directed evolution.

As opposed to rational design, directed evolution focuses less on structure and the positioning of the mutation and more on screening or selection of large libraries of mutants. The general procedure (Fig. 3.8)[117] of directed evolution consists of mutagenesis, expression, and screening or selection. This procedure is repeated with selected mutants as many times as necessary to simulate natural evolution on a laboratory time-scale.

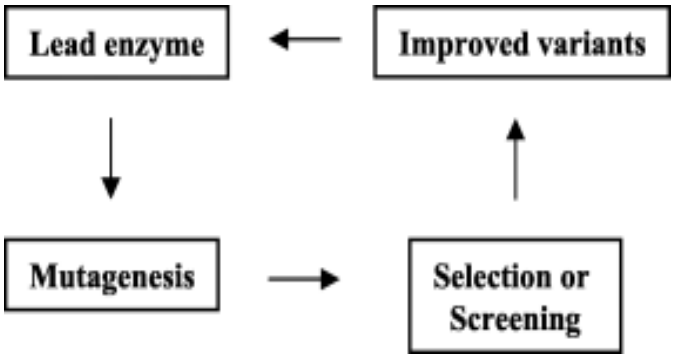



***Figure 3.8:*** *General scheme of directed evolution. Courtesy of Boersma* et al.[117]

*A) Mutagenesis*

Large mutant libraries of a protein of interest can be accumulated using a variety of strategies. Mutations are generally introduced via point mutation or recombination. The degree of randomness or “irrationality” and number of mutations depends on the property sought and pre-existing sequence or structural knowledge. Common strategies shown to be useful in the directed evolution of biocatalysts are error-prone PCR, saturation mutagenesis and DNA-shuffling.

Error-prone PCR is the most random method of mutagenesis as it targets the entire gene for point mutations. The error-rate of Taq DNA polymerase is adjusted by manipulating the conditions of PCR.[118] Since mutations are introduced at random, error-prone PCR requires no prior knowledge of structure. Error-prone PCR has been found useful in enhancing biocatalyst solubility in organic solvents[101] and thermostability.[119] Furthermore, results obtained from error-prone PCR studies may also identify potential "hot spots" which contribute to a particular property. Error-prone PCR was employed by Reetz in the first study[120] applying directed evolution to enhance the enantioselectivity of an enzyme, *Pseudomonas aeruginosa* lipase (PAL), for the kinetic resolution of racemic esters (Fig. 3.9A).[8] The success of this experiment and others to follow[121] was not only the enhancement of enantioselectivity (Fig. 3.9B), but also the elucidation of hot spots. These hot spots were the focal points of later studies by Reetz in which saturation mutagenesis was employed.[122]

Unlike error-prone PCR, saturated mutagenesis is an oligonucleotide-based approach which targets a specific location of expressed protein. Since a certain structural knowledge is required, this method is considered combinatorial (combines aspects of rational design and directed evolution). Mutant libraries are produced using degenerate codons substituted into DNA primers. In addition to enhancing the substrate scope or selectivity of enzymes by focusing on locations near to the active

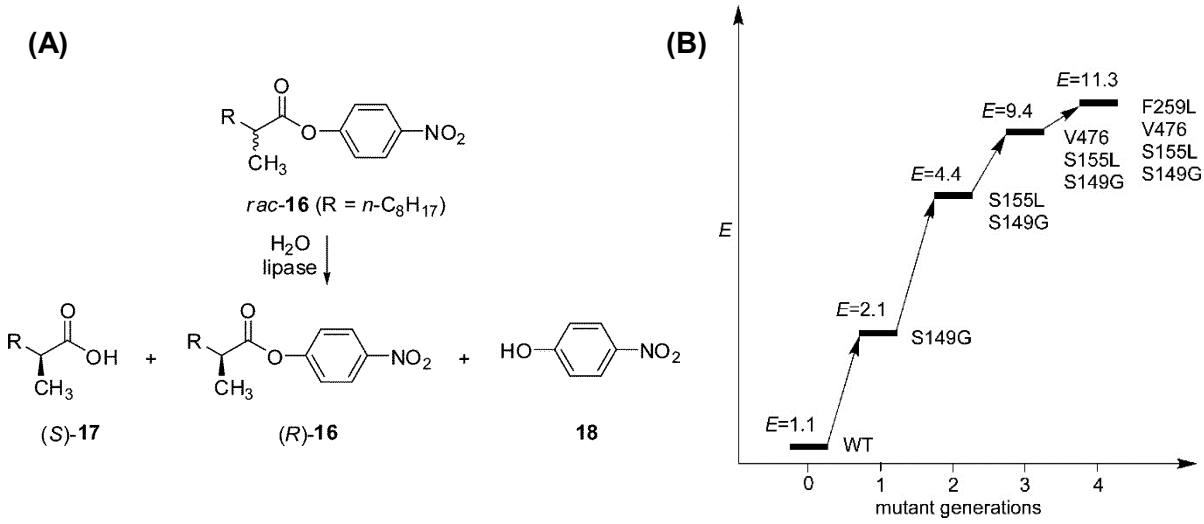



***Figure 3.9: (A)*** *Lipase-catalyzed enantioselective hydrolysis of p-nitrophenyl esters and* ***(B)*** *enhancement of enantioselectivity (E) following four successive generations of error-prone PCR starting with wild-type lipase from* Pseudomonas aeruginosa*. Courtesy of Reetz.*[8]

site, saturation mutagenesis can also be used to develop biocatalysts with improved stability. Properties such as thermostability, stability in denaturing agents and solubility in organic solvents are commonly improved by targeting locations shown by X-ray data to have high flexibility (high average B-factor)[123] or on sequences of low homology.

The benefit of saturation mutagenesis is that it targets specified critical positions, thereby reducing the size of mutant libraries and the amount of screening/selection, the bottleneck of directed evolution. In an effort to further curtail screening, Reetz *et al*. devised a systematic combinatorial approach known as iterative saturation mutagenesis (ISM).[124] With information obtained from previous error-prone PCR studies performed to enhance the enantioselectivity of *P. aeruginosa* lipase (Fig. 3.9) and X-ray data, three hot spots in the vicinity of the active site were chosen.[122] As shown in Figure 3.10a, a mutation map was created which contained six potential mutagenic pathways ending with 15 possible mutant libraries. Paths were then followed until the exceptional result of E = 594 (S) enantioselectivity was achieved (Fig. 3.10b). Furthermore, follow-up deconvolution studies revealed a strongly non-additive relationship between the individual mutations, emphasizing the importance of synergy in protein engineering.[125]

Another method commonly employed to create mutant libraries is DNA shuffling. DNA shuffling is a recombinant-based method in which a group of parent genes are cut into fragments with DNase. These fragments (10-50 base pairs) are then assembled into full-length genes using primerless PCR. Shuffling or crossover results when the fragment of one parent gene anneals with a fragment of another gene. In this method, parent genes can include any number of mutated or homologous genes.[126] DNA shuffling[127] and related recombinant methods[128,129] have been found to be effective in the directed evolution of biocatalysts.[130]

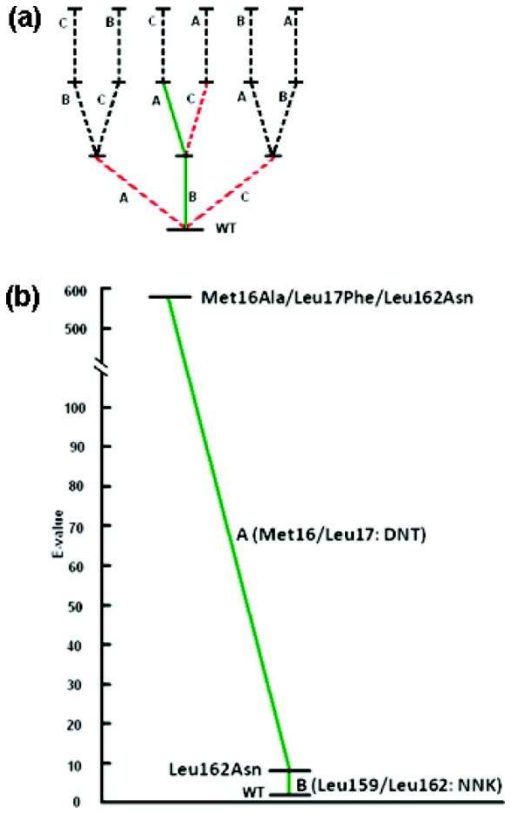



***Figure 3.10:*** *Iterative saturation mutagenesis (ISM) as applied by Reetz et al*[122] *to enhance* P. aeruginosa *lipase enantioselectivity (E).* ***(a)*** *Mutagenic pathway system showing all possible paths of saturated mutagenesis targeting three locations A, B and C. The solid green line represents the best pathway taken, the dotted red lines indicate attempts made with insignificant improvement, and the dotted black line denotes paths which were unnecessary and unexplored.* ***(b)*** *The best path found by ISM corresponding to actual mutations and resulting enantioselectivity.*

### *B) Screening or selection*

With large mutant libraries made available by biotechnology, the bottle-neck of directed evolution is passed on to the screening or selection process. Two general types of approaches are taken to reduce the time and energy spent in this process. The first involves simply reducing the number of mutants to be screened by producing what are commonly called smart libraries.[131] Smart libraries are normally created with a preexisting knowledge of structure or homology using combinatorial methods of mutagenesis. However, protein dynamics can be very unpredictable and, depending on the starting enzyme and the property desired, this may not be an option. Therefore, the second approach attacks the bottle-neck head on by attempting to increase the efficiency of the screening/selection process.

The main difference between screening and selection is that screening is performed on each individual mutant, whereas selection can be performed simultaneously on entire pools of mutants. From this it would seem obvious to avoid screening at all cost, but this is often not the case. Although processes of selection are less-time consuming and more closely resemble nature, they tend to be difficult to associate with or apply to the property desired. Moreover, selection for the property sought cannot be directly observed as in screening and like the biocatalyst itself, the selection technique must also be tailored to fit the task at hand.[117,132]

### 3.2.4 Biocatalysts via computational design and enzyme redesign

Driven by knowledge obtained from directed evolution and the time lost in screening mutant libraries, protein engineering is again taking on an increasingly rational form. Smarter libraries with "higher-quality" mutants are produced by combinatorial or semi-rational approaches and new and improved software is available for computational designs, which can subsequently be optimized by directed evolution. With the growth of structural protein databases and quantum mechanical models, novel enzymes are created by inserting active sites into previously non-enzymatic proteins and biocatalysts are designed for reactions that are not in the repertoire of natural enzymes.

The true meaning of *de novo* design in regard to proteins and peptides is debatable.[133] This section is written to establish a clearer delineation between current trends in rational enzyme design, in which catalytic machinery is engineered into native protein or protein sequences, and true *de novo* design (Section 4), which includes the design of not only the active site, but also the entire protein topology. In doing so, the remainder of this section will serve as a continuation of traditional rational design (Section 3.2.2) with examples of computational active site design in native proteins, which is often used in combination with directed evolution.[26,28,29,134,135]

In 2001, Bolon and Mayo reported the design of "protozymes."[136] Their strategy involved the use of the protein design software ORBIT to incorporate an active site into thioredoxin. This active site was modeled for nucleophilic catalysis (Section 2.1.2) with the histidine-mediated hydrolysis of *p*-nitrophenyl acetate (*p*NPA) (Fig.

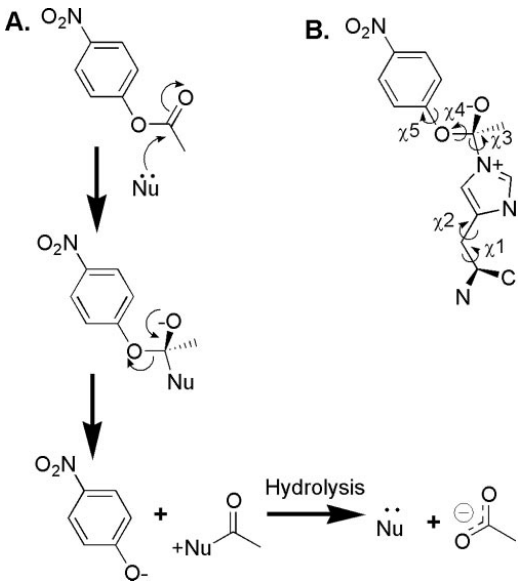



***Figure 3.11: A)*** *Histidine-mediated nucleophilic hydroylsis of pNPA* ***B)*** *Transition-state structure with torsion angles varied to generate rotamers used in the design calculations. From Bolon, Mayo* et al.[136]

3.11), a commonly used model substrate for kinetic ester hydrolysis studies. Although the catalytic results were not spectacular ($k_{cat}/K_M$ = 3 $M^{-1}s^{-1}$), their computational approach in combination with potential directed evolution represented wide-ranging possibilities for enzyme design.

By 2008, computational design had gained considerable ground with the use of the Rosetta software suite developed in the laboratory of David Baker. Biocatalysts engineered from scaffolds belonging to triose phosphate isomerase were found to accelerate a retro-aldol reaction (Fig. 3.12) by more than 2 x $10^4$-fold.[137] With later improvements and directed evolution, rate enhancements of almost 6 x $10^5$ were achieved.[138] That same year, the first unnatural reaction, the Kemp elimination of 5-nitrobenzisoxazole (Fig. 3.12), was also successfully modeled with Rosetta.[139] The best biocatalyst from this study was shown to have a second-order rate constant ($k_{cat}/K_M$) of 160 $M^{-1}s^{-1}$ (60,000 $M^{-1}s^{-1}$ following 13 rounds of directed evolution).[140] This finding inspired the use of an alternative scaffold for a more efficient Kemp eliminase (425 $M^{-1}s^{-1}$; 230,000 $M^{-1}s^{-1}$ after 17 rounds of mutagenesis and screening).[141,142]

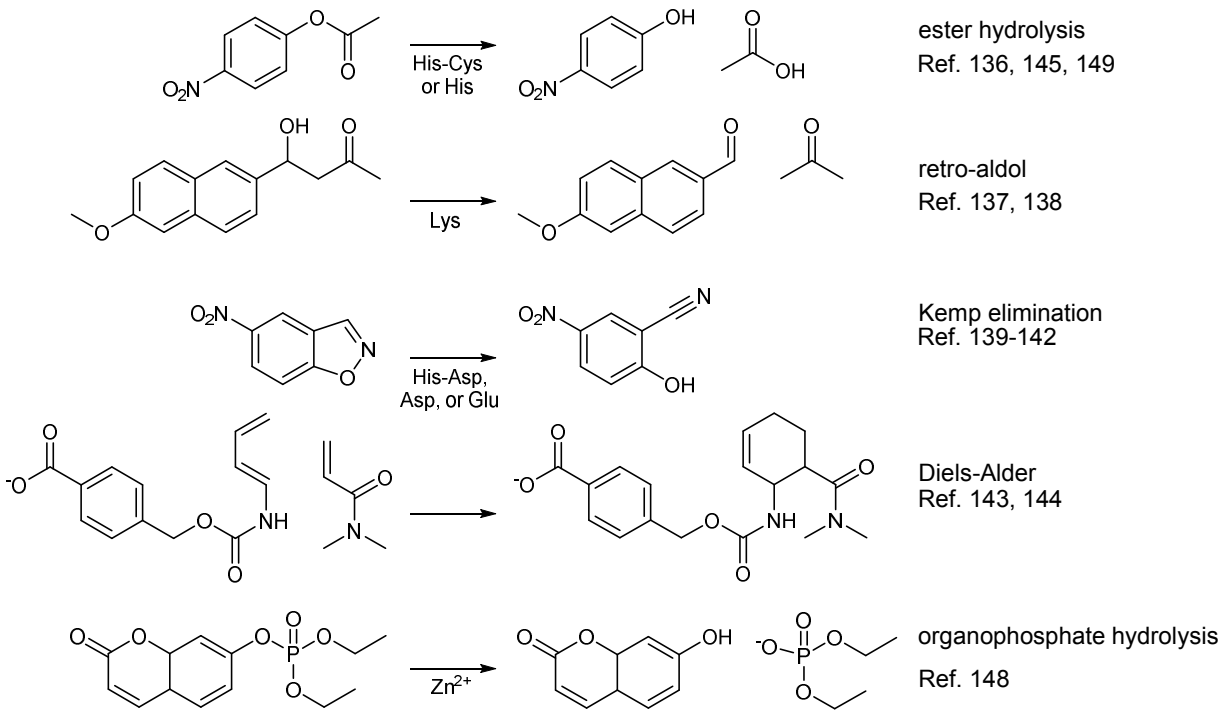



***Figure 3.12:*** *Target reactions of computational designed biocatalysts.*

Until 2010, no enzyme had been shown to catalyze bimolecular Diels–Alder reactions. The design of a stereoselective Diels-Alderase by Siegel, Baker and collaborators became a milestone in computational design.[143] As opposed to other biocatalysts developed with Rosetta, this design was the first to catalyze a C-C bond-forming reaction, which requires the accommodation of two substrates, a diene and a dienophile (Fig. 3.12). Although not overwhelmed with the activity ($k_{cat}/(K_{Mdiene}$ x $K_{Mdienophile}$) = 6 $s^{-1}M^{-1}M^{-1}$) of the original designs, a challenge presented to online enthusiasts of protein structure was answered with the insertion of a helix-turn-helix and an 18-fold increase in efficiency.[144]

As is common in biocatalysis research, computational design is often employed to develop hydrolases. Like Mayo had with ORBIT in 2001, Richter, Baker and collaborators applied Rosetta in an attempt to engineer an effective nucleophilic esterase *in silico*.[145] Unlike the previous study by Mayo (Fig. 3.11), this study focused on the importance of a histidine-cysteine catalytic dyad in the active site (Fig. 3.13). Following the directed evolution of their most active design, the second-order rate constant corresponding to catalyzed *p*NPA hydrolysis was increased from 34 $M^{-1}s^{-1}$ to 405 $M^{-1}s^{-1}$.

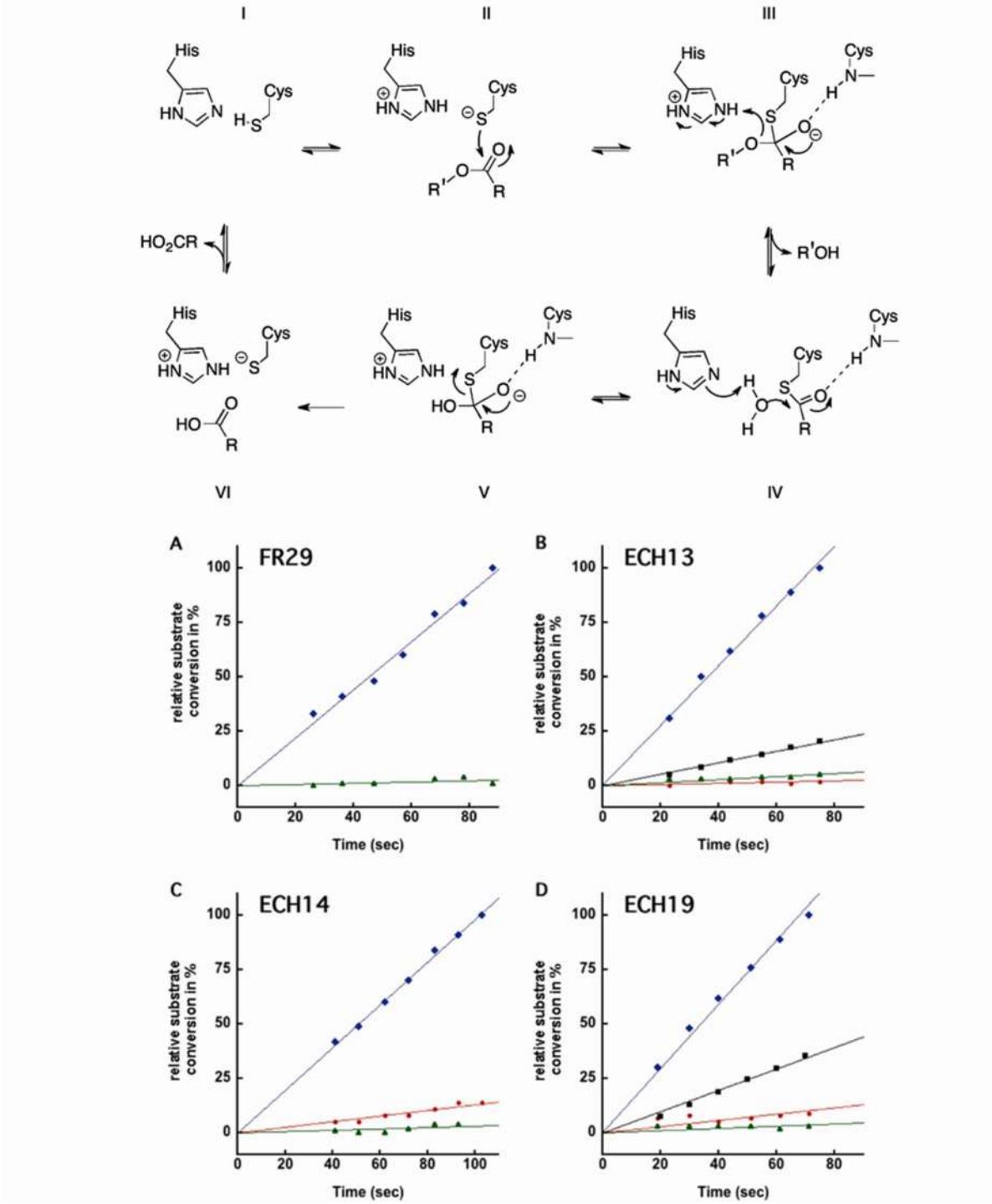



***Figure 3.13: Top)*** *Computational designed mechanism of ester hydrolysis by Cys-His dyad with an oxyanion binder.* ***Bottom)*** *Relative substrate (*p*-nitrophenyl ester) conversion catalyzed by four esterase designs. Blue progress curves correspond to initial designs possessing the complete dyad, red traces correspond to Cys-knockout variants, His-knockouts are shown in black and double Cys and His-knockouts are in green. Courtesy of Richter* et al.[145]

Another route used by enzymes to initiate ester hydrolysis is provided through the use of metal ion catalysis (Section 2.1.3). This is witnessed with the promiscuous ester hydrolytic activity of the zinc metalloenzyme carbonic anhydrase (Fig. 3.14). Zinc in particular plays an important role in numerous hydrolytic enzymes and is, therefore, commonly exploited in designer enzymes, especially those of *de novo* design (Section 4). Zinc serves both structural and catalytic roles in metalloenzymes. When catalytically active, zinc is bound in an open coordination sphere. This means that a position for coordination is occupied by solvent/water.[58] The geometry of zinc is most often tetrahedral, but trigonal pyramidal geometries, as is the case in adenosine deaminase, is also observed (Fig. 3.15).[146,147]

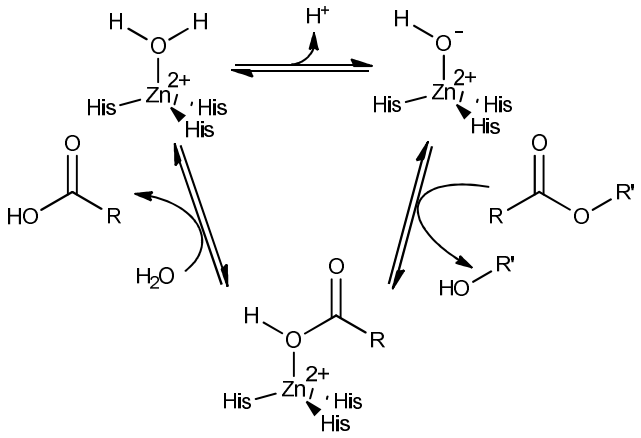



***Figure 3.14:*** *Promiscuous carbonic anhydrase-catalyzed ester hydrolysis via zinc-bound hydroxide mechanism.*

In yet another accomplishment, Baker and his team set out to repurpose the catalytic machinery of a functional zinc metalloenzyme.[148] Namely, murine adenosine deaminase was redesigned for organophosphate hydrolysis (Figs. 3.12 and 3.16). However, all zinc-coordinating residues (Fig. 3.15) were left intact. Following three rounds of *in vitro* evolution, the redeployed enzyme acquired a $10^7$-fold gain in $k_{cat}/K_M$ for organophosphate hydrolysis in comparison to the wild-type deaminase. On the other hand, deaminase activity dropped 50,000 fold.

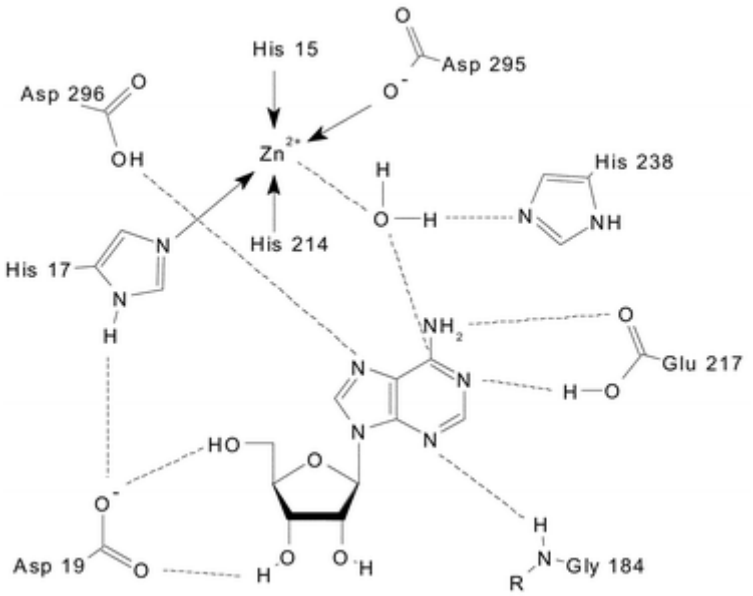



***Figure 3.15:*** *Trigonal bipyramidal zinc coordination and water activation within the active site of wild-type adenosine deaminase. From Gleeson* et al.[147]

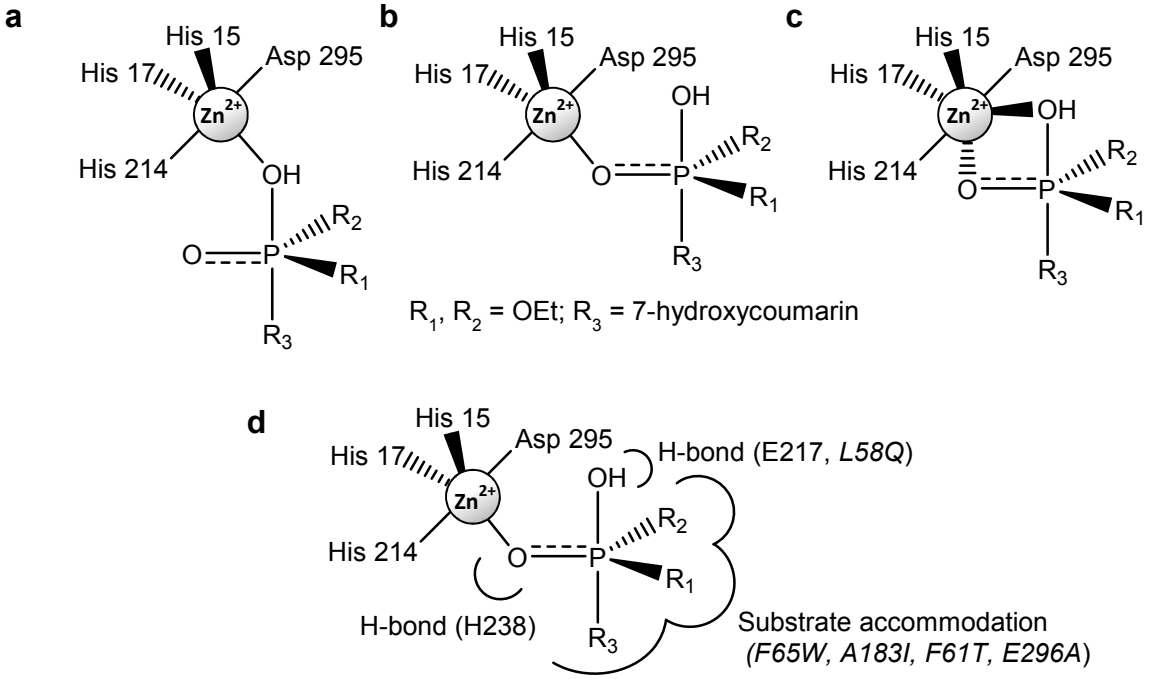



***Figure 3.16:*** *Computational adenosine deaminase active site redesign by Khare, Baker* et al.[148] *accounting for the mechanistic parameters:* ***a****) Zn-bound hydroxyl nucleophilic attack on the organophosphate,* ***b****) stabilization of the developing negative charge,* ***c****) simultaneous performance of both* ***a*** *and* ***b****; and* ***d****) stabilizing H-bond interactions and substrate accommodation. Mutations are shown in italics.*

As the ratio between the size of native protein scaffolds and the degree of mutagenesis decreases, it becomes increasingly difficult to clearly distinguish true *de novo* design from the broader rational design. Inevitably, an example was bound to arise in which a simple folding motif constructed of a short amino acid sequence was extracted from a native protein and modified by computational design to possess a specific function. This is exactly the case presented by the design of Der *et al.*[149] in which the Rab4-binding domain of rabenosyn[150] was used to derive the 46-residue MID1-zinc (Fig. 3.17). The initial intent behind MID1-zinc, was simply to show how zinc could induce helix-turn-helix monomers to form dimers (Fig. 3.17 B). However, the MID1-zinc crystal structure revealed an open coordination sphere, $(His)_3Zn$ (Fig. 3.17 C), typical of a zinc catalytic site as opposed to the expected closed $(His)_4Zn$. Furthermore, the co-crystallization of tartrate (Fig. 3.17 D) from the crystallization buffer within a cleft of the dimer suggested possible esterase activity. Following kinetic studies, unoptimized MID1-zinc was found to catalyze the hydrolysis of *p*NPA with an astounding rate enhancement of $10^5$ and $k_{cat}/K_M$ of 630 $M^{-1}s^{-1}$. Open site zinc coordination of water resulted in the lowering of its pKa to 8.2.

The findings of Der *et al.* highlight key concepts which could be very influential in future rational design and, more specifically, *de novo* design. Enzyme-like activity is not confined to large protein structures. Furthermore, the use of simple, self-assembling peptides creates protein interfaces resembling clefts of enzyme active sites. Simpler, well-defined peptide assemblies could, therefore, hold the key to understanding enzyme function. By applying this bottom-up approach to *de novo* design, we could determine the minimal requirements needed for enzyme catalysis. As knowledge of folding and function continue to grow, so will the complexity of *de novo designs.*

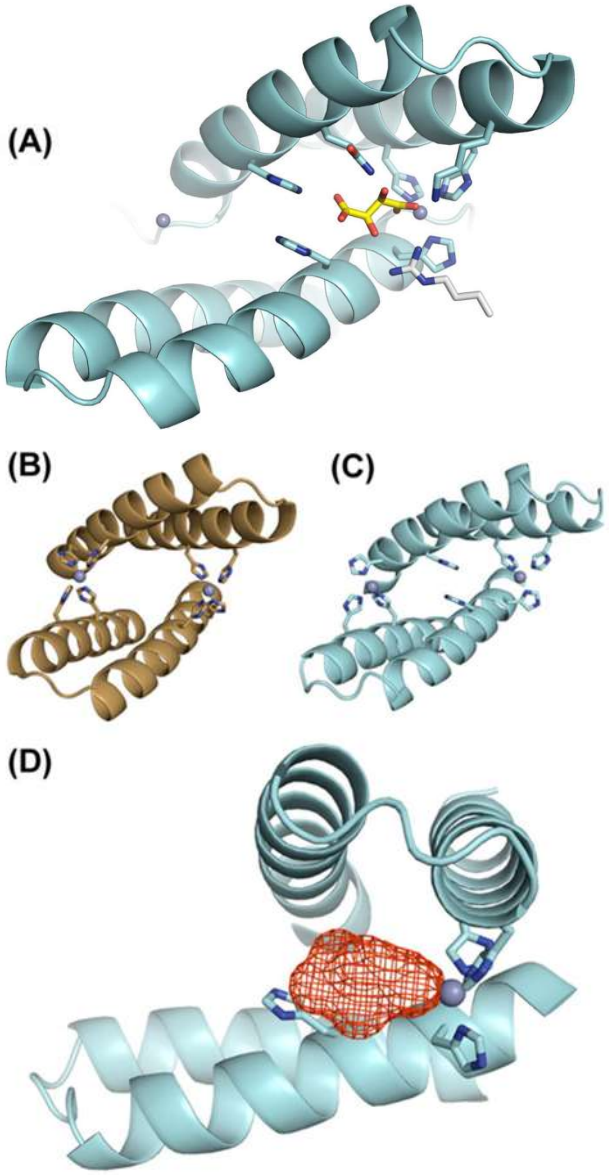



***Figure 3.17: (A)*** *Co-crystallization of tartrate from the crystallization buffer bound in the open coordination site of zinc.* ***(B)*** *Computational prediction showing closed zinc coordination,* $(His)_4Zn$*, typical of zinc structural sites.* ***(C)*** *Observed open coordination with three of the four histidine residues coordinating zinc.* ***(D)*** *Alternative angle revealing substrate accessible cleft at the open zinc coordination site. Adapted from Der* et al.[149]

# 4

# *De novo* catalytic proteins

If we really understood enzymes, we would be able to design them according to the reactions or processes desired. Although the field of enzyme design has experienced much advancement, a designed equivalent has yet to come close to the remarkable efficiency of native enzymes. The bottom-up or minimalist approach to *de novo* design allows researchers to break down complex enzyme interactions into their individual parts such as the presentation of catalytic machinery and substrate accommodation.

Through the years, surfactants,[151] cyclodextrins[51] and more complex supramolecular arrangements[152] have been employed to confine substrates in artificial, enzyme-like active sites. With a continually growing knowledge of the mechanisms which guide protein folding, the use of self-assembling, synthetic peptides in the truly *de novo* design of catalytic proteins has become a rapidly developing field. This has been witnessed by the steadily increasing catalytic efficiencies and complexity of designs achieved. Combined with advancements in computational methods, this trend in the bottom-up approach is likely to radiate outward to promote indirect interactions in the active-site periphery.

When using self-assembling peptides as scaffolds to support catalytic machinery, stability often becomes a major consideration. Functional groups which actively participate in catalysis are typically polar and prefer being exposed to aqueous surroundings. The active sites of enzymes commonly exist within clefts or at the interface of folds. The complex structures of enzymes allow for a multitude of energetically favorable interactions to stabilize such positioning. The number of stabilizing interactions of self-assembling peptides is much more limited. Therefore, the type of assembly and the positioning of catalytic machinery must be carefully chosen.

*De novo* peptide designs can take on numerous stable shapes and assemblies. These assemblies are typically classified according to the resulting secondary structure (Fig. 4.1). Both α-helical and β-sheet peptide designs are used in the truly *de novo* design of catalytic proteins.

## 4.1 α-Helical *de novo*-designed catalytic proteins

### 4.1.1 Coiled coils

Approximately 30% of all protein structure is comprised of short, α-helical amino acid sequences.[153] Many of these same sequences fail to adopt the same structure when removed from the larger protein environment. [154] Without the hydrophobic support of the parent protein structure, solitary α-helices are more likely to succumb to entropic effects.[155] However, when such sequences are of the coiled-coil motif, α-helical structure is more likely to be retained.

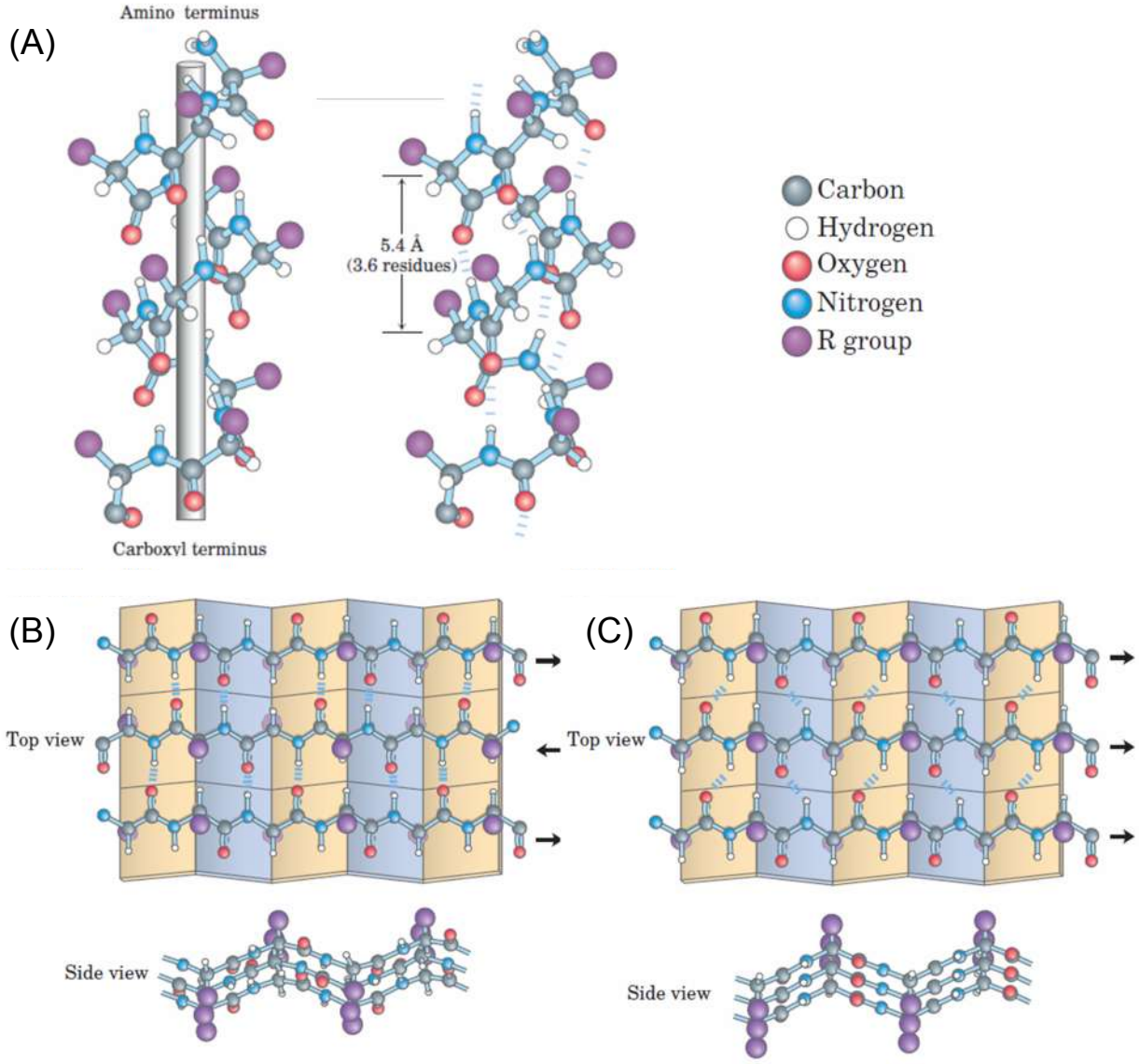



***Figure 4.1:*** *Secondary protein structure.* ***(A)*** *Right-handed α-helix showing central axis (left) and hydrogen bonds (right).* ***(B)*** *Antiparallel β-sheet* ***(C)*** *Parallel β-sheet. Structure-stabilizing hydrogen bonds are designated with dashed blue lines. Adapted from Nelson & Cox.*[66]

α-Helical coiled coils are protein structure domains in which two to six α-helices pack together and wrap around each other forming a superhelix. They are among the most ubiquitous folding motifs and are believed to be present in as much as 10% of all proteins.[156] The stability as well as the versatility of this fold, not only in native protein structures, but also in peptides, makes it an excellent candidate for *de novo* design.[157-160]

Sequences forming coiled-coil structures are characterized by a repeating heptad pattern $(abcdefg)_n$ in which *n* designates the number of repeats (Fig. 4.2).[161] The heptad's *a* and *d* positions comprise the hydrophobic interface and are therefore occupied by amino acids possessing hydrophobic side chains.[162,163] The stability of the coiled-coil fold is largely dictated by the ability of these "core" residues to pack tightly in what is referred to as "knobs into holes."[164,165] In coiled-coil dimers, trimers and tetramers, the heptad *e* and *g* positions are occupied by residues such as glutamate and lysine which participate in interhelical salt bridges for further stability.[166] The occupancy of *b*, *c* and *f* in dimers, trimers and tetramers is less restrictive. Since these positions are typically exposed to aqueous media, they are often occupied by polar residues.[167]

In addition to stability, α-helical coiled coils can also vary in the degree of oligomerization (dimer, trimer, etc.), the orientation of the helices relative to one another (parallel or antiparallel, Fig. 4.2) and specificity of oligomerization (homomeric or heteromeric). The degree of coiled-coil oligomerization or oligomerization number is determined by the side-chain geometry of residues and their packing at the hydrophobic interface. This oligomerization number is typically between two and six. Table 4.1, based on results collected over numerous studies, accurately summarizes basic principles in the *de novo* design of coiled coils. The orientation and the specificity of oligomerization are largely directed by forces of attraction or repulsion between charged residues in $g_n$ and $e_{n+1}'$ ($n + 1$ = one heptad repeat below; the prime signifies the involvement of the interacting helix).[168]

### 4.1.2 Examples of α-helical *de novo*-designed catalytic proteins

Enzymes are large molecules with complex topologies. However, as has been witnessed by the success of proline and other amino acids in the field of organocatalysis, complex structure is not a requirement for catalysis. As simple scaffolds, helical assemblies can predictably present an array of functionalities as long as stability is maintained.

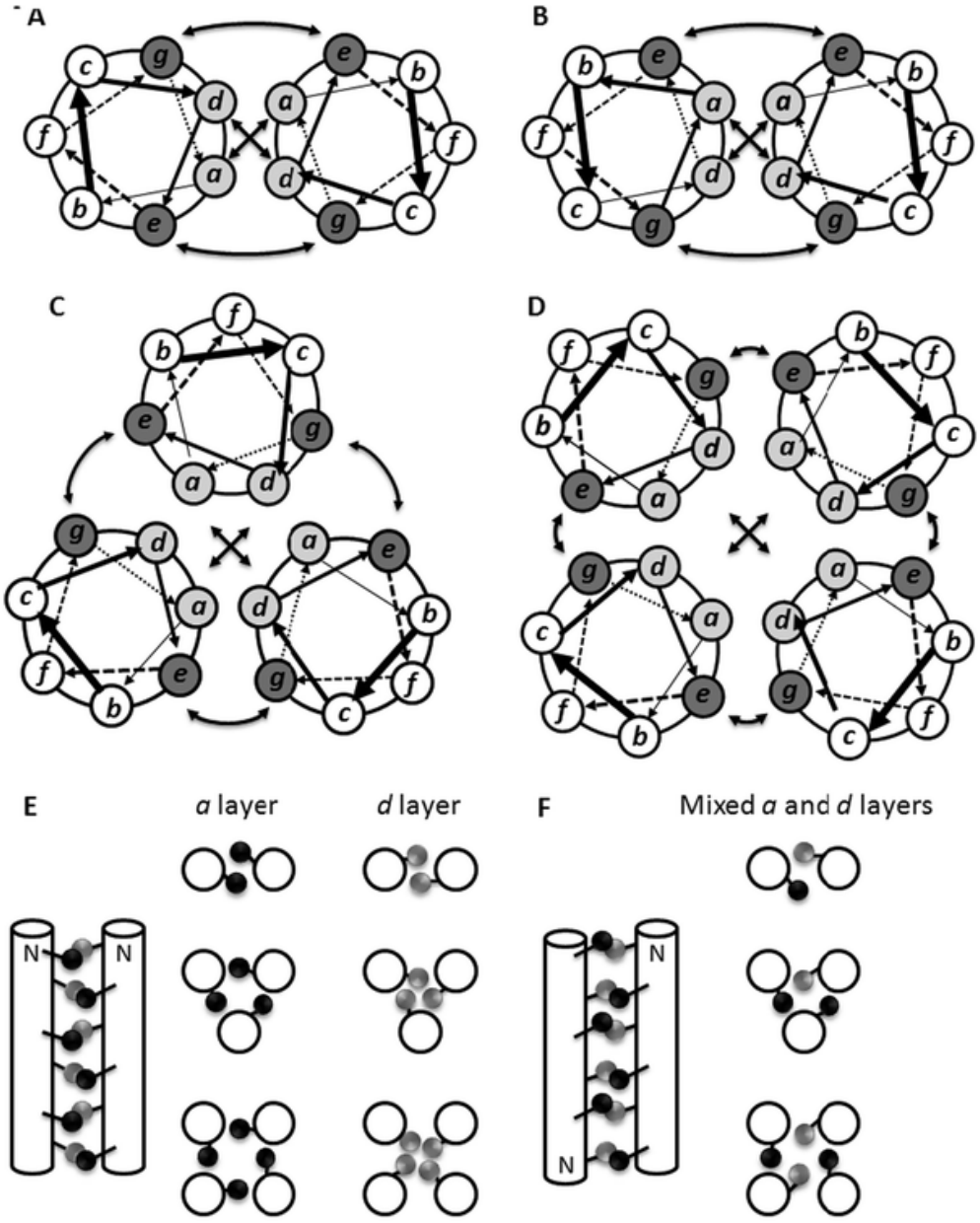



***Figure 4.2:*** *Helical wheel diagrams of coiled coils.* ***(A)*** *Parallel dimer.* ***(B)*** *Antiparallel dimer.* ***(C)*** *Parallel trimer.* ***(D)*** *Parallel tetramer. Curved arrows designate salt bridges and crossed arrows indicate interfacial hydrophobic interactions. Knobs into holes packing of* ***(E)*** *parallel coiled-coil dimers, trimers and tetramers and* ***(F)*** *antiparallel coiled coils. Courtesy of Apostolovic* et al.[161]

***Table 4.1:*** *Basic design principles of α-helical coiled coils.*[161]

| Amino acid sequence / heptad alignment | Oligomerization number | Parallel | Antiparallel |
|---|---|---|---|
| • Ile (*a*) and Leu (*d*)<br>• 4 Ala substitutions in potential tetramer core<br>• Insertion of polar residue in potential tetramer core<br>• Avoid Val and Asn in (*a*) (induce oligomerization > 2)<br>• Avoid hydrophobic residues at *e*, *g*, *b*, and *c* (induce oligomerization > 2) | 2 | • Preferred interactions: *a–a´* ; *d-d´* and $g_n$-$e_{n+1}$´<br>• $g_n$-$e_{n+1}$´ is Lys-Glu or Glu-Lys<br>• *a* (Leu ≈ Ile) and *d* (Leu) | • Preferred interactions: *a–d´* ; *d-a´* ; *g-g´* ; *e-e´*<br>• *(a-d´)* Ala-Ile or Leu-Arg<br>• *g-g´* ; *e-e´* are attractive interactions : ($g_n$-$e_{n+1}$´ are repulsive)<br>• Asn in core induces heterodimer |
| • Ile (*a* and *d*)<br>• Leu (*a* and *d*)<br>• Leu (*a*) and Val (*d*) | 3 | | • Trp in the core |
| • Leu (*a*) and Ile or Ala (*d*)<br>• Glu and Lys at *b* and *c* | 4 | • *a* (Met ≈ Ile >Val>Leu>Ala)<br>• *d* (Met>Ile>Glu) | • Leu (*a*) and Ala (*d*)<br>• Ala insertion in core of potential parallel tetramer |
| • Trp (*a* and *d*)<br>• Hydrophobic residues at all positions except (*f*) | 5 | • *a* (Leu>Ile)<br>• *d* (Met>Gln>Thr>Val) | |

Among the original *de novo* designed catalysts to incorporate α-helices was the synthetic hemeprotein, 'helichrome'.[169] Interestingly, the main structural scaffold component of helichrome was not the protein portion, but porphyrin. By covalently linking porphyrin to four short, amphiphilic peptides, Kaiser created an α-helical bundle with a hydrophobic core. In the presence of iron and cofactors, 7-acetylflavin

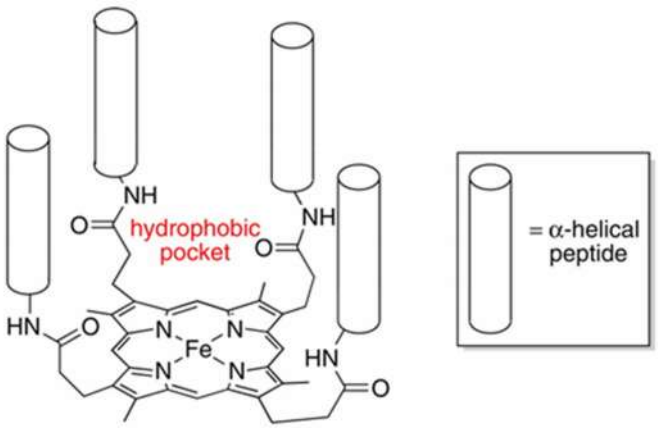



***Figure 4.3:*** *Proposed structure of Kaiser's helichrome.*[169] *Courtesy of Lewis.*[19]

and NADPH, aniline hydroxylase activity was observed as aniline was oriented in the hydrophobic pocket of the bundle, adjacent to the active Fe(III)-porphyrin complex (Fig 4.3).

Whereas helichrome offered a hydrophobic pocket to draw in substrate, another *de novo* design known as 'oxaldie' relied on coulombic binding interactions.[170] As an artificial oxaloacetate decarboxylase, oxaldie accelerated the conversion of oxaloacetate to pyruvate by 4 orders of magnitude. Like helichrome, oxaldie was a short (14 residues), amphiphilic peptide that assumed an α-helical, tetrameric bundle in its active state. The key to oxaldie's catalysis was the presentation of its exposed, lysine-lined cationic face. Upon luring oxaloacetate with its charged side-chain amines, the terminal amine, measured pKa of only 7.2, could initiate subsequent decarboxylation though an imine intermediate (Fig. 4.4). This concept was later exploited by Allemann with the insertion of the catalytic machinery of oxaldie into bovine pancreatic polypeptide (bPP) resulting in 'oxaldie 4' (Fig. 4.5 A and B).[171] The increased α-helical stability imparted by the linked proline-rich helix was translated into a three-fold increase in efficiency when compared to the original oxaldie. Allemann also applied the bPP scaffold to create 'Art-Est', an artificial esterase.[172] The Art-Est design promotes ester hydrolysis through the interplay of several residues (Fig. 4.5). Of these residues, two are histidine. One histidine among two arginine residues serves as a nucleophile with a perturbed pKa of 5.5. The second histidine is situated next to glutamate and acts as an acid to stabilize the developing negative charge of the intermediate.

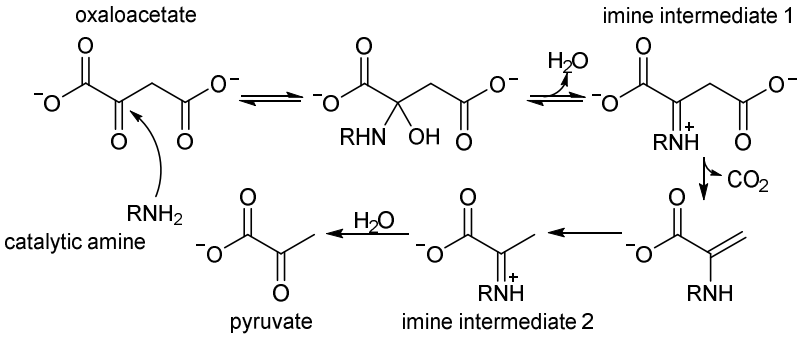



***Figure 4.4:*** *Amine-catalyzed decarboxylation of oxaloacetate*

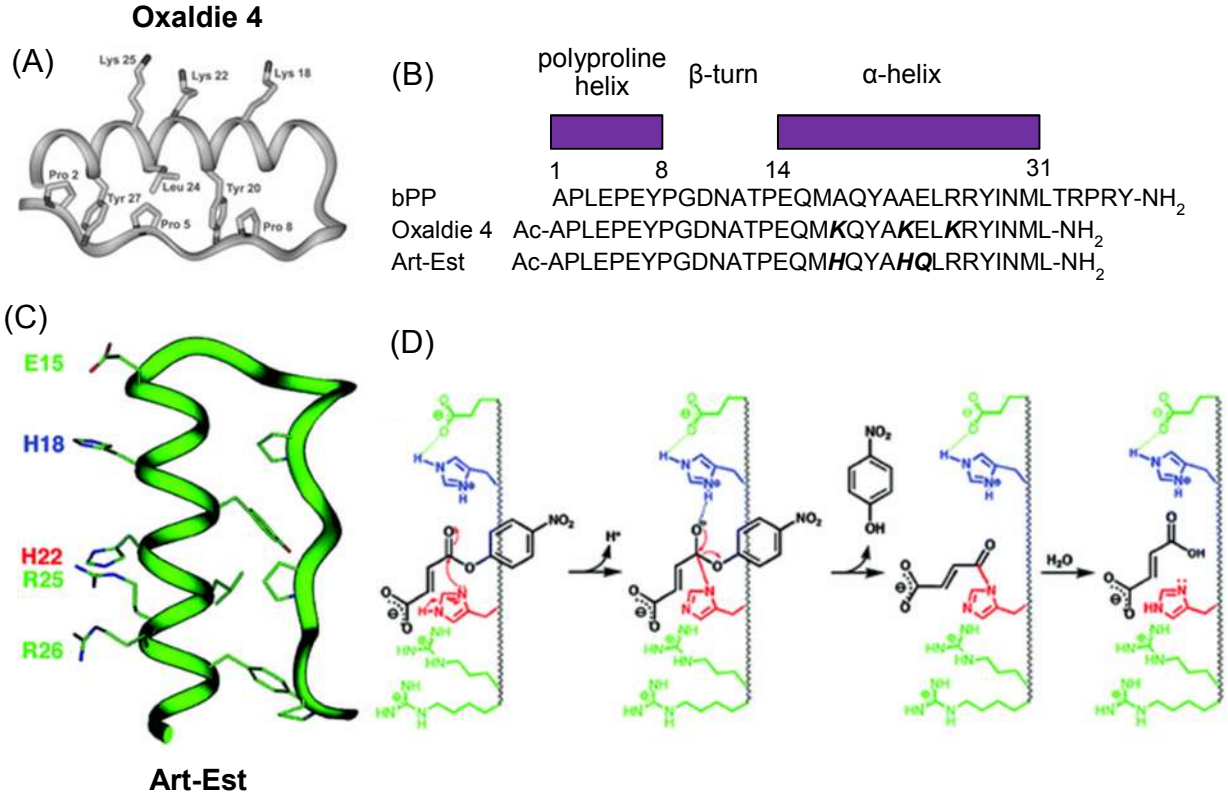



***Figure 4.5:*** ***(A)*** *Structure of oxaldie 4 modeled after bovine pancreatic polypeptide (bPP).* ***(B)*** *Sequence alignment and structures of bPP, oxaldie 4 and Art-Est.* ***(C)*** *Structure of Art-Est.* ***(D)*** *Proposed nucleophilic histidine-catalyzed ester hydrolysis of* p-*nitrophenyl fumarate by Art-Est. Adapted from Tayler* et al.[171] *and Nicoll* et al.[172]

Work discussed in the preceding paragraphs emphasizes the flexibility and potential of *de novo* peptide design. Essentially, Allemann and collaborators discovered the decarboxylation mechanism in oxaldie, incorporated it into bPP to form oxaldie 4, and lastly, repurposed that same scaffold for ester hydrolysis. A very similar cooperative nucleophilic and general acid mechanism was applied to ester hydrolysis in previous work performed in the research group of Baltzer.[69,173,174] The helix-loop-helix motif, KO-42, which dimerizes into a four-helix bundle, was shown to possess hydrolytic activity due to cooperative $HisH^+$-His pairings (Fig. 4.6). Like Allemann, Baltzer later incorporated a similar *i*, *i* + 3 and *i* + 4 arrangement of histidine, arginine and arginine, respectively, (Art-Est, Fig. 4.5 C and D) into the same helix-loop-helix motif to lower the histidine pKa for phosphoester hydrolysis (Fig. 4.7).[175,176]

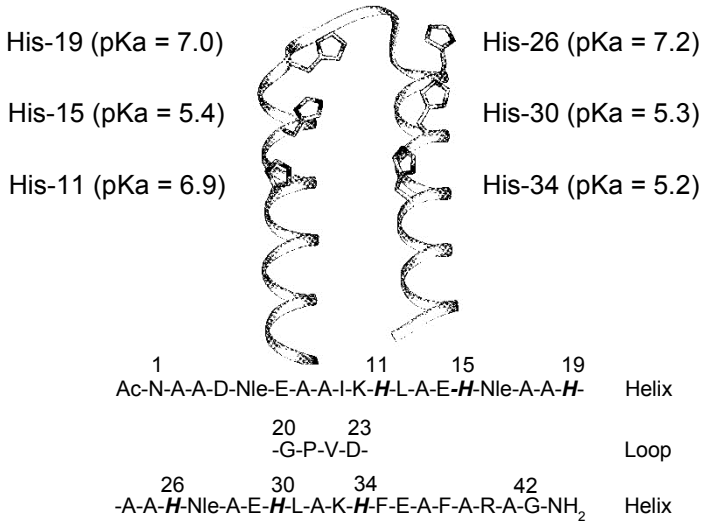



***Figure 4.6:*** *Helix-loop-helix motif and amino acid sequence of KO-42 monomer with* $HisH^+$ *-His acid-base pairings. Adapted from Broo* et al.[69]

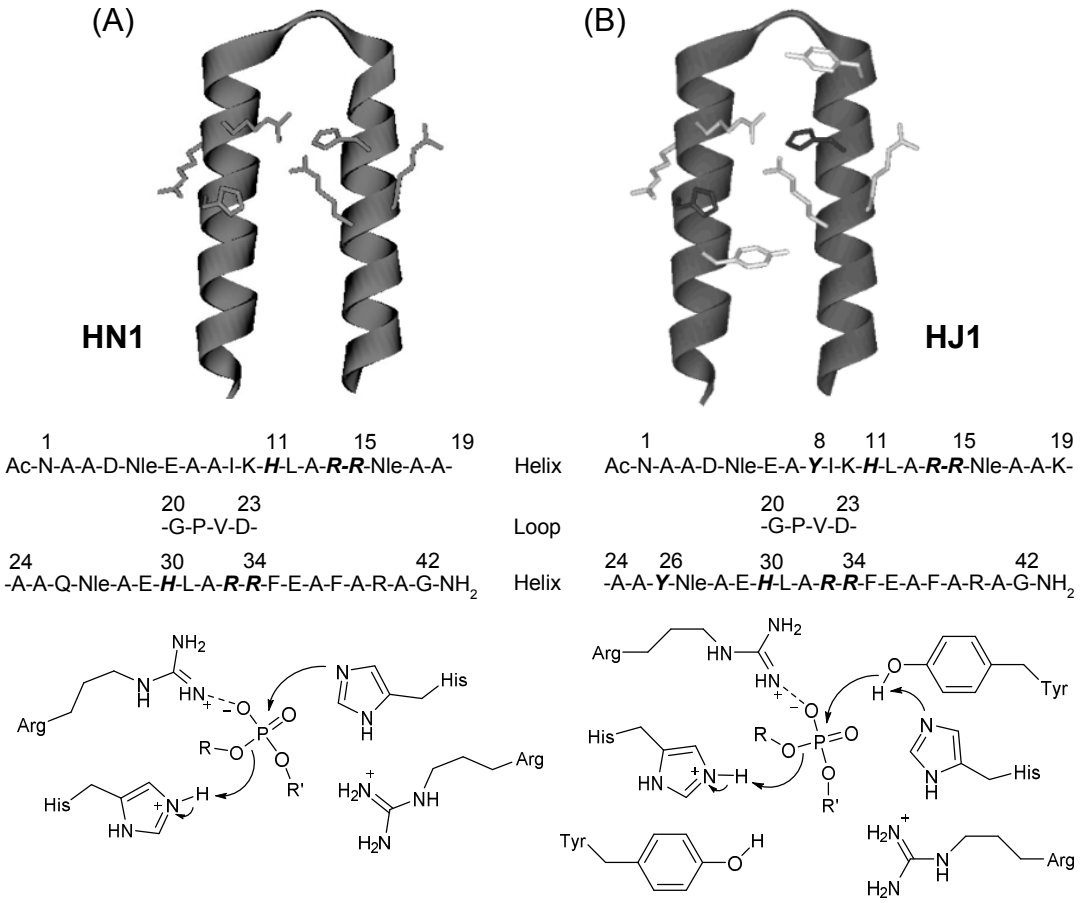



***Figure 4.7:*** *Helix-loop-helix motif as structural scaffolds for* de novo*-designed catalysts of phosphate diester hydrolysis. Only the monomers of the active four-helix bundles are shown* ***(A)*** *HN1 using histidine in a cooperative nucleophilic-general acid mechanism.* ***(B)*** *HJ1 using histidine in a general acid-base mechanism. Adapted from Razkin* et al.[175,176]

The most basic barrier of the *de novo* design of protein catalysts, particularly in the bottom-up approach, is the ability to introduce catalytic machinery into simple self-assembling peptide scaffolds without incurring a loss of structural stability. With the exception of helichrome, which required porphyrin for stability, other examples including oxaldie variants and the helix-loop-helix scaffolds of Baltzer presented catalytic functionalities to solvent-exposed environments. However, catalytic activity of native enzymes is not limited to the surface. A common source of stability and catalytic activity used by native proteins are metals.

Nearly half of all proteins found in nature are metalloproteins.[177] In nature, metal ions provide proteins with both structural stability and catalytic function. For such reasons, the incorporation of metal ions is also important for the *de novo* design of catalytic proteins. As will be seen in the following examples, the application of metals in the bottom-up approach of α-helical designs allows active site stabilization and activity within the hydrophobic core or cleft, a common feature of native metalloenzymes.[178]

Within the last decade, the increase in the efficiency of truly *de novo* designed catalytic proteins directly correlates with the incorporation of metal ions. In 2004, Kaplan and DeGrado applied their 'Due Ferri' concept of metalloproteins to bring about oxidative catalysis (Fig. 4.8).[25] In this study, a heterotetrameric α-helical bundle, $DF_{tet}$, was computationally designed to perform phenol oxidase activity by coordinating two iron (II) ions within a single active site (Fig. 4.9). What was most interesting about this design was that the two iron ions were coordinated by two histidine residues and four glutamates within the hydrophobic core of the α-helical bundle. The design was later optimized to provide quicker oxygen binding (Fig. 4.8 eqn. 1) and more space for substrate accommodation (Fig. 4.8 eqn. 2) by substituting glycine residues in place of a bulkier leucine and an alanine in two of the heteromers, resulting in a more active variant $G_4$-$DF_{tet}$. The heterotetrameric system provides an effective platform to find the optimal combination of subunits. Later, another design would join two subunits to form a more stable helix-loop-helix dimer DF3 with due ferri phenol oxidase activity.[179] The general applicability of *de novo* design was further demonstrated with single-chain due ferri G4DFsc as well as 3His-G4DFsc repurposed with an additional histidine for *N*-hydroxylation catalysis.[180]

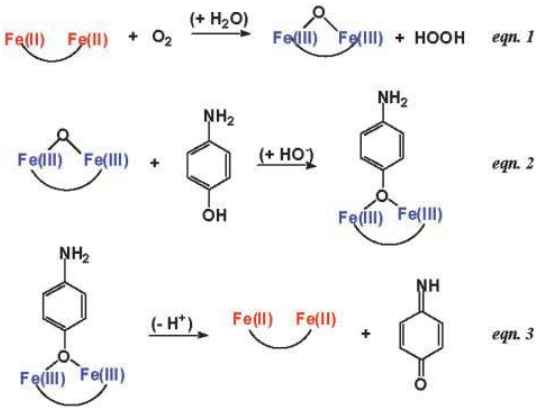



***Figure 4.8:*** *Proposed reaction mechanism of 4-aminophenol oxidation to benzoquinone monoimine catalyzed by diferric protein. Adapted from Kaplan & DeGrado.*[25]

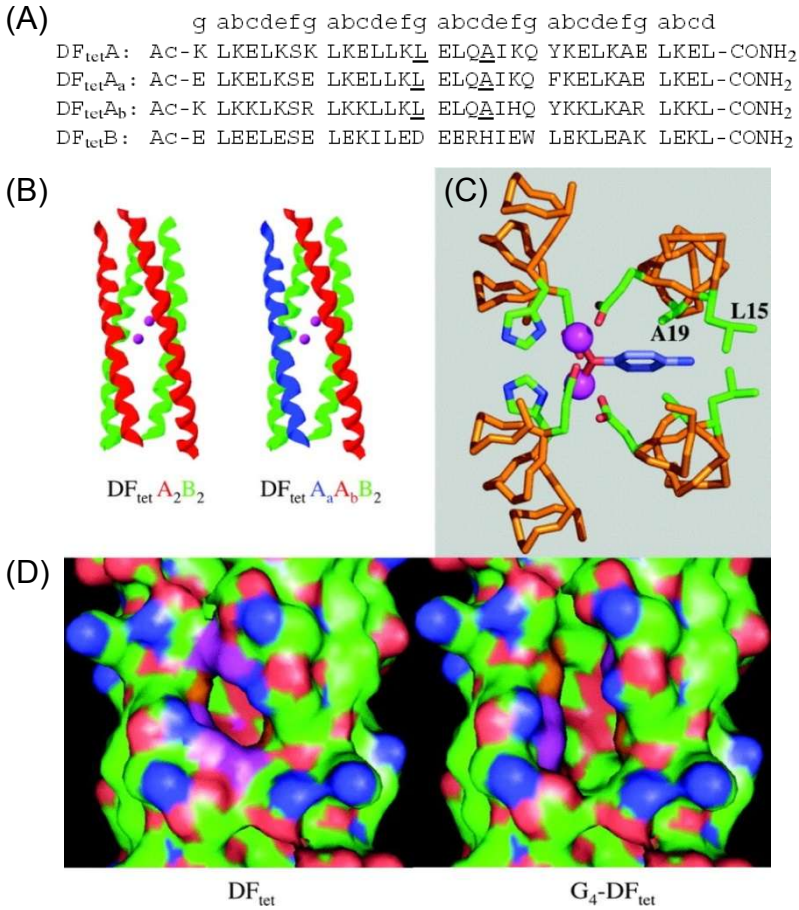
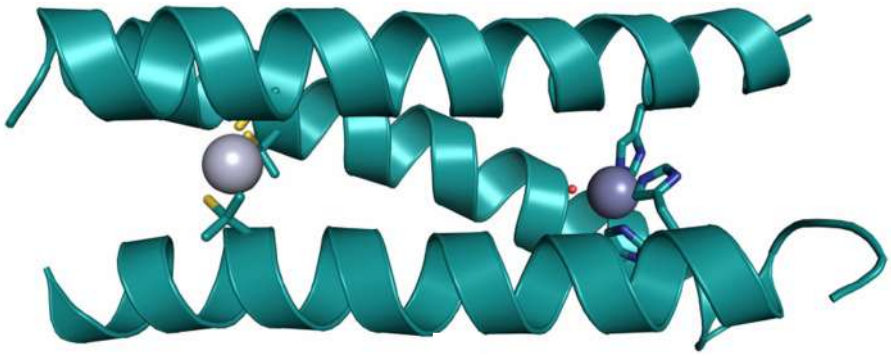



***Figure 4.9:*** ***(A)*** *Amino acid sequences with heptad register of peptide subunits designed to comprise the heterotetrameric bundle $DF_{tet}$. Underlined residues of the A subunits were substituted with glycine to create $G_4$-$DF_{tet}$.* ***(B)*** *Heteromeric subunit composition of $DF_{tet}$ with coordinated iron ions in violet.* ***(C)*** *$DF_{tet}$ active site showing coordinated di-iron (III) bound to 4-aminophenol substrate.* ***(D)*** *Comparative solvent-exposed surfaces showing the increased active site accessibility of $G_4$-$DF_{tet}$. Adapted from Kaplan & DeGrado.*[25]

Another *de novo* metalloprotein recently introduced is the Zn(II) TRI system by Pecoraro.[27] The Zn(II) TRI system is an α-helical trimeric coiled coil which positions three histidine residues within the hydrophobic core for zinc coordination. This open $(His)_3Zn$ alignment is analogous to the active site of carbonic anhydrase (Fig. 2.4) and MID1-zinc[149] (Fig. 3.17). Similar to carbonic anhydrase, the variant $Hg(II)_SZn(II)_N(TRIL9CL23H)_3$ was reported to catalyze both ester hydrolysis and the hydration of carbon dioxide.

The $Hg(II)_SZn(II)_N(TRIL9CL23H)_3$ variant of the TRI family is composed of a trimeric coiled-coil assembly of the 30-residue peptide TRIL9CL23H. To compensate for any destabilization that may result from the inclusion of histidine within the hydrophobic core, a second metal-binding site, $(Cys)_3Hg$, was designed (Fig. 4.10). This assembly was reported to possess a catalytic pKa of 8.8 and $k_{cat}/K_M$ values of 23 $M^{-1}s^{-1}$ for *p*-nitrophenyl acetate hydrolysis and 1.8 x $10^5$ $M^{-1}s^{-1}$ for $CO_2$ hydration at pH 9.5. The efficiency in which $Hg(II)_SZn(II)_N(TRIL9CL23H)_3$ hydrated $CO_2$ was found to be 70-fold greater than any reported model and within just 500-fold of the kinetically perfect CAII. In additional studies, it was shown that altering the position of the metal coordination sites within the core had an effect on the assembly's ability to catalyze ester hydrolysis. However, none of the variants, including one variant with switched metal coordination sites and another with a more centrally located zinc site, were



| peptide | | a b c d e f g | a b c d e f g | a b c d e f g | a b c d e f g | |
|---|---|---|---|---|---|---|
| TRIL9CL23H | Ac-G | L K A L E E K | C K A L E E K | L K A L E E K | H K A L E E K | G-$NH_2$ |
| CSL9PenL23H | Ac-E | W E A L E K K | Pen A A L E S K | L Q A L E K K | H E A L E H G | -$NH_2$ |

***Figure 4.10:*** *Trimeric coiled coil structure of $Hg(II)_SZn(II)_N(CSL9PenL23H)_3$ featuring the coordination of Hg (II) (gray) by penicillamine (Pen) sulfur (yellow) and Zn (II) (blue) by histidine and water (red). The peptide CSL9PenL23H used in crystallography studies is analogous to TRIL9CL23. The residues which participate in metal coordination are underlined. Adapted from Zastrow* et al.[27]

found to be more efficient.[181,182] Interestingly, by simply replacing zinc with redox-active copper, a very similar assembly, $Cu(I/II)(TRIL23H)_3^{+/2+}$ was found to accelerate the reduction of nitrite to nitric oxide.[183] Returning to carbon dioxide hydration, Pecoraro introduced the $(His)_3Zn$ machinery into the single-stranded three-helix (two-loop) bundle, $\alpha_3D$ modelled by DeGrado.[184] The resulting single-chain bundle, $\alpha_3DH3$, offers the advantage of being more stable and, therefore, requires no structural mercury coordination site.[185] Although, the initial design exhibited 5-fold less hydration activity as $Hg(II)_SZn(II)_N(TRIL9CL23H)_3$, it offers greater potential for future development of the secondary coordination sphere. This stabilization of the active site or primary metal coordination sphere with secondary shell residues and the difficulty in channeling water to hydrophobic reaction centers are two obstacles presently facing *de novo* hydrolases.

## 4.2 β-sheets as scaffolds for *de novo*-designed catalytic proteins

On paper, the design of β-sheet peptides is simple (Fig. 4.1). All that is needed is a simple alternating sequence of hydrophobic and hydrophilic residues. Since much of *de novo* protein design is predicated on the construction of protein interfaces, the inherent gregarious nature of β-sheet structures would at first seem like a perfect match. However, a disadvantage of β-sheet structures is that they tend to aggregate to the point of precipitation.[186] The difficulty encountered when designing stable β-sheet folds has reduced their utility as a platform for *de novo* catalyst design.[30] Nonetheless, the β-sheet blueprint of carbonic anhydrase (Fig. 2.4 A) and heightened interest in self-assembling nanostructures and hydrogels[187] has encouraged an interest in *de novo* β-sheet esterases.[188,189]

Taking a minimalistic route to mimic serine protease activity, Fukushima created poly(Asp-Leu-His-Leu-Ser-Leu) to obtain a β-sheet structure with its polar constituents, aspartate, histidine and serine performing as a catalytic triad (Fig. 2.3). The resulting peptide consisted of approximately ten repeats of the sequence DLHLSL and folded into a stable, monomeric β-sheet. Although the mechanism was not clearly revealed, poly(DLHLSL) did show slight esterase activity with a second order rate constant of 0.5 $M^{-1}s^{-1}$ (*p*NPA hydrolysis; pH 8) over the hexapeptide DLHLSL, which lacked a defined structure, a mixture of the individual amino acids

and imidazole.[73] In a separate study, poly(DLHLSL) was also found to exhibit selectivity (E = 1.5) of *N*-benzyloxycarbonyl-L-phenylalanine *p*-nitrophenyl ester (Z-L-Phe-ONp) over Z-D-Phe-ONp.[72] This was attributed to an interaction between the substrate and hydrophobic portion of the β-sheet.

More recently, Degrado and Korendovych attempted to model the active site of carbonic anhydrase from a minimalist β-sheet-forming heptapeptide (LKLKLKL).[188] Of the variants examined, the peptide of sequence Ac-IHIHIQI-$NH_2$ in the presence of $Zn^{2+}$ (Fig. 4.11) was found to most actively catalyze *p*NPA hydrolysis ($k_{cat}/K_M$ = 62 $M^{-1}s^{-1}$ at pH 8). This incredibly simple design is not only the most efficient truly *de novo* esterase to date, but also forms and remains active as amyloid fibrils (Fig. 4.12).

Further characterization revealed that $Zn^{2+}$ not only participated in catalysis (lowered the pKa of water to 9.3), but through histidine coordination also played a role in inducing amyloid formation. Catalytic assays performed after dialysis and again after filtration did, indeed, show that the fibrils were catalytically active. Furthermore, removal of blocking N-terminal acetyl group and C-terminal amide led to an inability of the peptides to form β-sheet structures and an absence of activity. This was likely the result of charge repulsion, which inhibited formation of the catalytically active parallel structure.

In a study by Zhang and collaborators, catalytic nanofibers were created from the self-assembling peptide Q11 ($H_2N$-QQKFQFQFEQQ-OH).[189] The aim of this study was to bring about ester hydrolysis with histidine through either general acid-base or nucleophilic catalysis. The best catalytic nanofiber was found to consist of a 10:1 ratio of Q11H ($H_2N$-HSGQQKFQFQFEQQ-$NH_2$) and Q11R ($H_2N$-RSGQQKFQFQFEQQ-$NH_2$). This is likely due to the interaction between arginine and histidine using a cooperative general acid/nucleophilic mechanism as seen earlier in the case of Art-Est (Fig. 4.5) and HN1 and HJ1 (Fig. 4.7). Although the second-order rate constant was only determined at pH 7.4 (*p*NPA, $k_{cat}/K_M$ = 0.15 $M^{-1}s^{-1}$), these nanofibers were shown to be biocompatible.

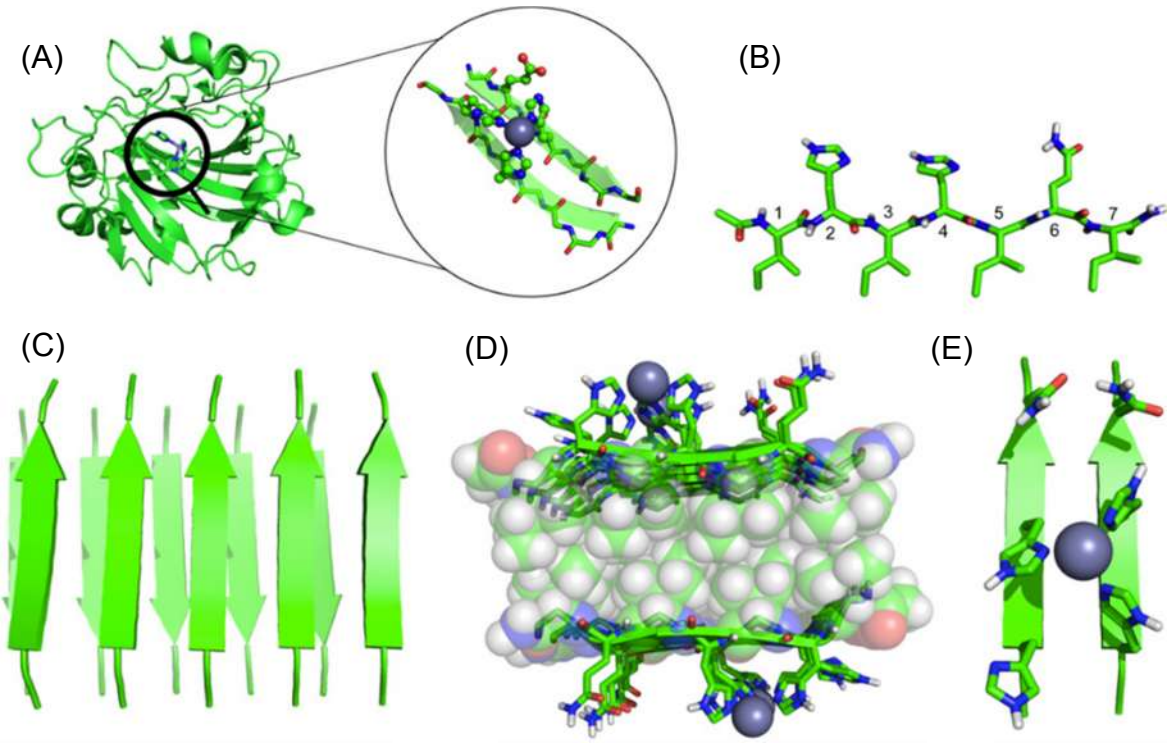



***Figure 4.11:*** ***(A)*** *The crystal structure and active site $Zn^{2+}$ coordination of carbonic anhydrase II.* ***(B)*** *Ac-IHIHIQI-$NH_2$ monomer.* ***(C)*** *Overall fold of Ac-IHIHIQI-$NH_2$.* ***(D)*** *Hydrophobic packing of isoleucine residues and* ***(E)*** *primary $Zn^{2+}$ coordination sphere of Ac-IHIHIQI-$NH_2$. Adapted from Rufo* et al.[188]

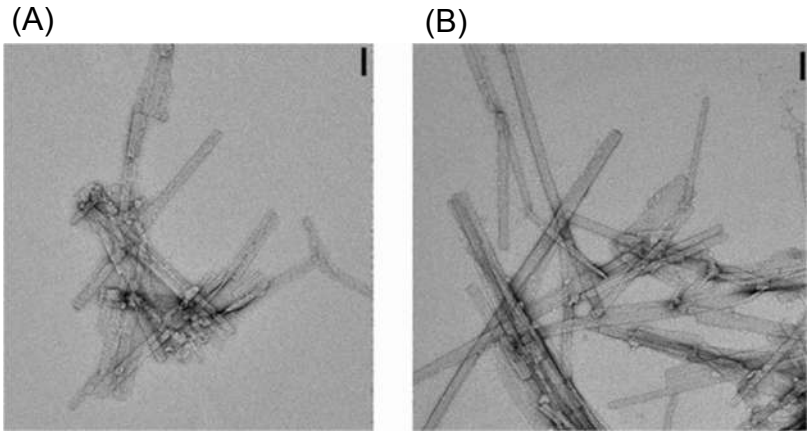



***Figure 4.12:*** *TEM images of Ac-IHIHIQI-$NH_2$ in the presence of $Zn^{2+}$* ***(A)*** *immediately following preparation and* ***(B)*** *after 72 hours of incubation. Adapted from Rufo* et al.[188]

To date, all β-sheet-based designs of *de novo* catalytic proteins have presented their catalytic machinery to the solvent-exposed surface. Such presentation is likely to result in less substrate specificity, which is an important aspect to consider in enzyme mimicry. It is unclear whether the tight packing which occurs at the hydrophobic core of fibrils can impart favorable or selective interactions with substrate molecules.

Future efforts are required to stabilize β-sheet design in solution and involve them in designs which are more three-dimensional to provide greater protein-substrate and secondary protein-protein interactions.

## 4.3 Histidine as simple, yet versatile catalytic machinery

Histidine plays an important role in the active site of nearly every enzyme and biocatalyst. This is due to its side-chain imidazole group. By itself, imidazole catalyzes the hydrolysis of *p*-nitrophenyl acetate with a second-order rate constant of 0.2 $M^{-1}s^{-1}$ at pH 8.[73] Depending on the environment of the active site, the imidazole of histidine (pKa ≈ 6.5) can be either protonated or deprotonated under physiological conditions (Section 2.1.1). When protonated, histidine acts as a general acid by donating its proton to a substrate or by stabilizing a transition state. When deprotonated, histidine is an effective base which can activate a nucleophilic species or serve directly as a nucleophile (Section 2.1.2 and Figs. 3.11, 4.5C and 4.7). As seen in proceeding sections, histidine is also an important ligand for the coordination of metal ions within an active site (Section 2.1.3). Histidine can readily participate in each of the catalytic mechanisms outlined in Section 2.1. However, the most important and often overlooked characteristic of histidine is perhaps its structural versatility. The attachment of imidazole C-5 to the amino acid allows histidine, via rotation, access to a large coverage area to perform its mechanistic roles (Fig. 4.13).[190]

Many of the studies thus far reported in enzyme redesign and the true *de novo* design of catalytic proteins have modeled hydrolase activity. As has been seen in the examples presented, histidine can play many different roles, which is especially the case in catalyzed hydrolysis. As far as flexibility is concerned, scaffolds designed or redesigned are not limited to only one specific substrate, reaction or even group of reactions. Numerous examples have been reported which show the chemical

versatility of scaffolds employed in enzyme-like catalysis. A scaffold with introduced catalytic machinery can be applied directly to another reaction or undergo modest redesign with alternative machinery or other slight adjustments, which have little effect on scaffold structure.

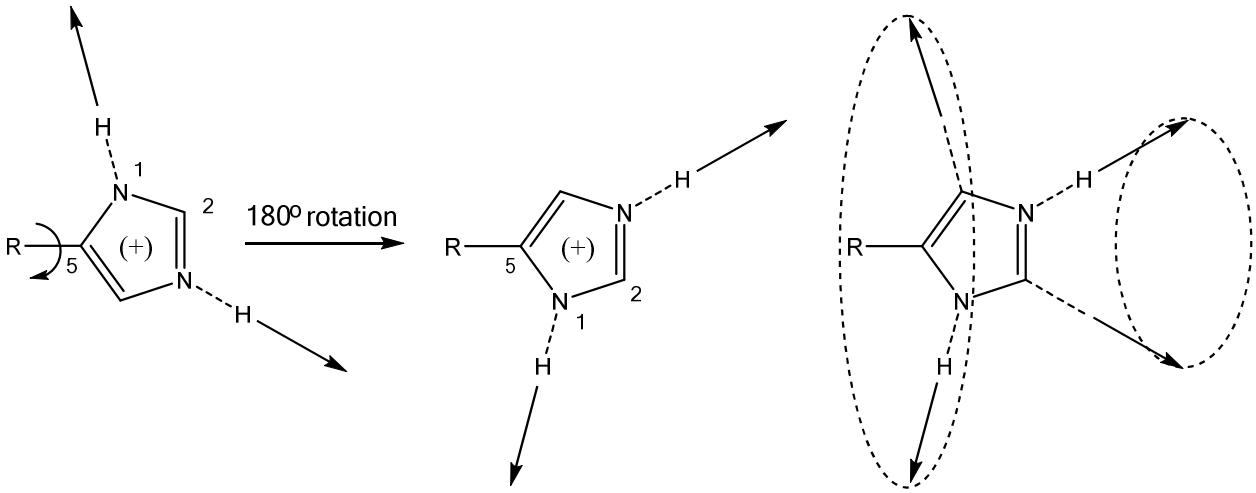



***Figure 4.13:*** *Spatial coverage of histidine upon rotation. Adapted from Rebek.*[190]

# 5 Project aim

In order to gain a clearer understanding of enzymes both in form and function, the bottom-up approach of *de novo* protein design is applied. In this case, catalytic machinery, namely histidine, is incorporated into peptides which self-assemble into predictable structures. This endeavor is pursued to develop catalytic proteins which behave in a manner similar to enzymes by enhancing reaction rates while providing substrate selectivity and the means to regulate activity. In order to gain a full breadth of possibilities, hydrolytic esterase activity is studied by employing three different peptide folding motifs, two different catalytic mechanisms and various ester substrates.

In the first study (Section 6.1), general acid-base catalytic machinery in the form of histidine, serine and aspartate as well as histidine in a glutamate-rich environment are to be introduced into one peptide of a heterodimeric coiled-coil. The goal of these experiments is to create a catalytic system whose activity can be regulated according to peptide structure.

In another study (Section 6.2), the catalytic ability and substrate selectivity of a previously reported β-sheet catalyst which aggregates into amyloid fibrils is examined for catalytic esterase proficiency as a fully-formed fibril and while in the process of self-assembly. This catalyst, which hydrolyzes esters using the zinc-bound-hydroxide mechanism, is also tested for substrate selectivity using substrates which varied in hydrophobicity and chirality. These experiments are performed to grasp the possible utility of such a β-sheet assembly if it could be integrated into a stable, non-amyloid forming design.

The final project (Section 6.3) employs a more complicated α-helical coiled-coil hexamer which also uses the zinc-bound-hydroxide mechanism of ester hydrolysis. The goal of this project is to create a *de novo* protein catalyst which offers its active site within a hydrophobic, yet water-accessible cleft. The main intention in each of these studies is to uncover minimal elements of enzyme form and function which can be incorporated into the design of more complex *de novo* catalytic proteins.

# 6
# Results and discussion

## 6.1 Regulation of random-coil peptide catalyst through coiled coil formation

In addition to providing a simple α-helical scaffold, the heterodimeric version of a coiled-coil model (Section 4.1.1; Fig. 4.2) offers yet another attractive feature as it helps clarify the importance of the structure/function relationship. In a heterodimeric system, the presence of a single peptide sequence forms an undefined secondary structure or random coil. Therefore, an interacting, partner sequence is required to induce the α-helical coiled-coil conformation.[191] By these means, a test is thereby established in which a function such as catalytic esterase activity can be attributed simply to the functional groups involved in the primary structure or to the interplay of chemical moieties in a higher-order fold. To best stress such applicability of model peptides, a very simple and well-characterized base-model was selected for these studies. The IAAL-E3/IAAL-K3 (or simply E3/K3) system meets the necessary structural requirements of a heterodimeric coiled-coil while consisting of only 21 residues per peptide and 5 different amino acids in the entire system (Fig. 6.1).[192]

Initial studies focused on the introduction of a catalytic triad (CT),[193] aspartate, histidine and serine, into both E3 and K3 of the E3/K3 scaffold to bring about ester hydrolysis and to observe the role structure plays in catalytic activity. Analogous to the studies of Fukushima,[72,73] in which serine protease-like activity was intended by the respective placement of aspartate, histidine and serine in positions *i, i + 2* and *i + 4* on the hydrophilic face of a β-sheet (Section 4.2), this study inserted the triad into the *i, i + 4* and *i + 8* positions of an α-helix. Selection of these positions was based on their close proximity when the peptide is folded in a coiled-coil motif. Variants E3CT and K3CT are shown in Figure 6.2. Catalytic assays were performed with peptides E3, K3, E3CT and K3CT as well as each heterodimeric combination (E3 and K3 variants in 1:1 ratio) using pseudo-first-order kinetics with *N*-benzyloxycarbonyl-L-phenylalanine *p*-nitrophenyl ester (Z-L-Phe-ONp) as substrate.

Surprisingly, the random coil form of E3CT showed the most esterase activity (Table 6.2). The mixture of E3CT/K3CT displayed relatively slight activity, but little coiled-coil formation according to the measured α-helicity. Although the combinations E3CT/K3 and E3/K3CT formed coiled coils, they exhibited no sign of activity. Unlike its random coil counterpart E3CT, K3CT was inactive. From these results, it appears that the formation of a coiled coil actually locks the peptides containing the catalytic triads into

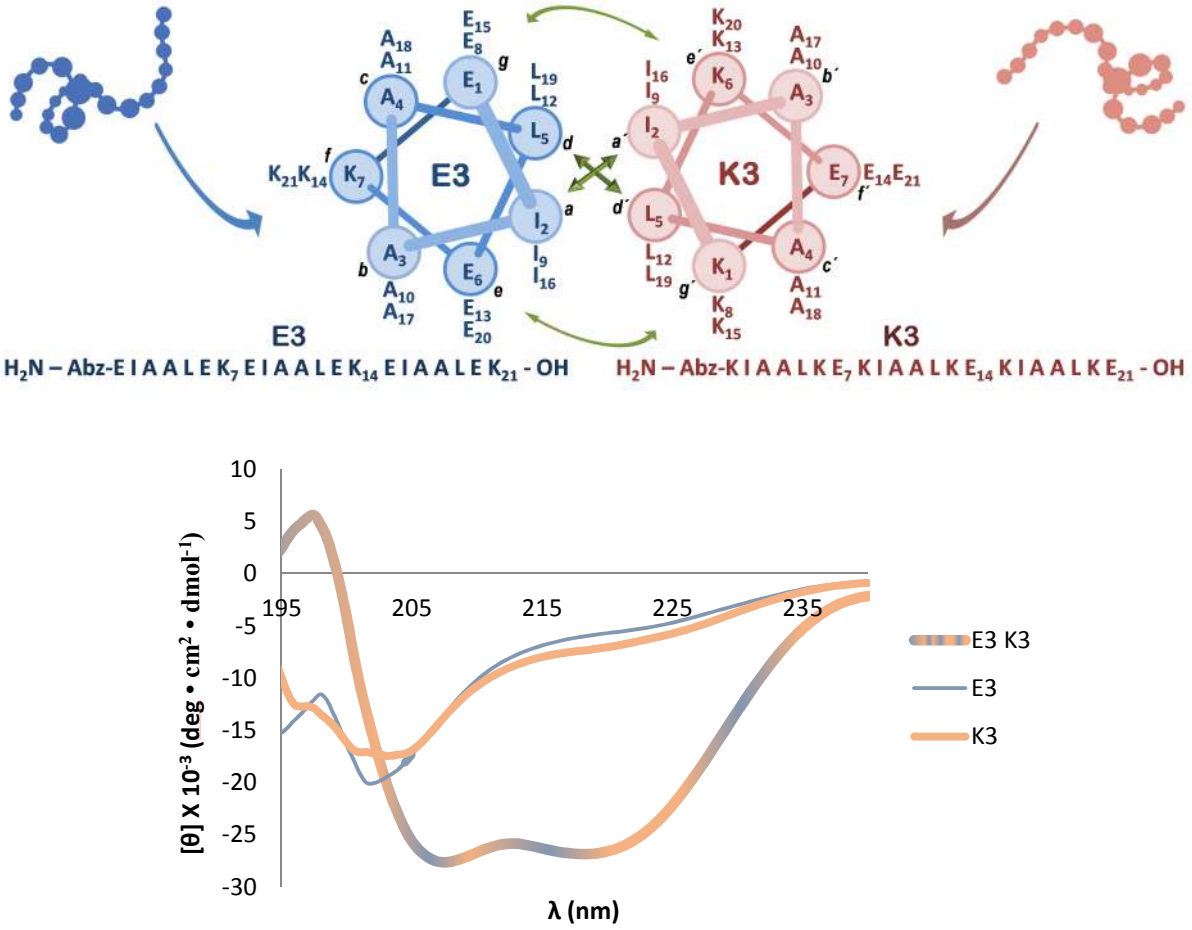



***Figure 6.1:*** *Sequence and helical wheel representation of parallel heterodimeric coiled-coil E3/K3 (top) and corresponding CD spectra (bottom). CD spectra taken at pH 7.5 (50 mM Tris/HCl) and 20°C with 100 µM peptide concentration.*

an inactive state. When designing the peptides, it was initially assumed that by locating individual residues of the catalytic triad in the relative pattern *i, i + 4, i + 8* that they would be close enough to one another to interact. However, the directionality of the side chains appears to keep the triad from forming the acid-base interaction required for charge relay (Fig. 6.3). Furthermore, the catalytic activity observed in the presence of E3CT was possibly due to the combination of histidine in a glutamate-rich environment and the allowed flexibility for histidine to interact with these glutamates. As can be seen in Figure 6.2, K3CT has just two glutamate residues and one aspartate in comparison to the five glutamates and one aspartate of E3CT. As the *total* peptide concentration was the same in all experiments, the reduced activity of the E3CT/K3CT mixture (25 µM each) was likely due to the

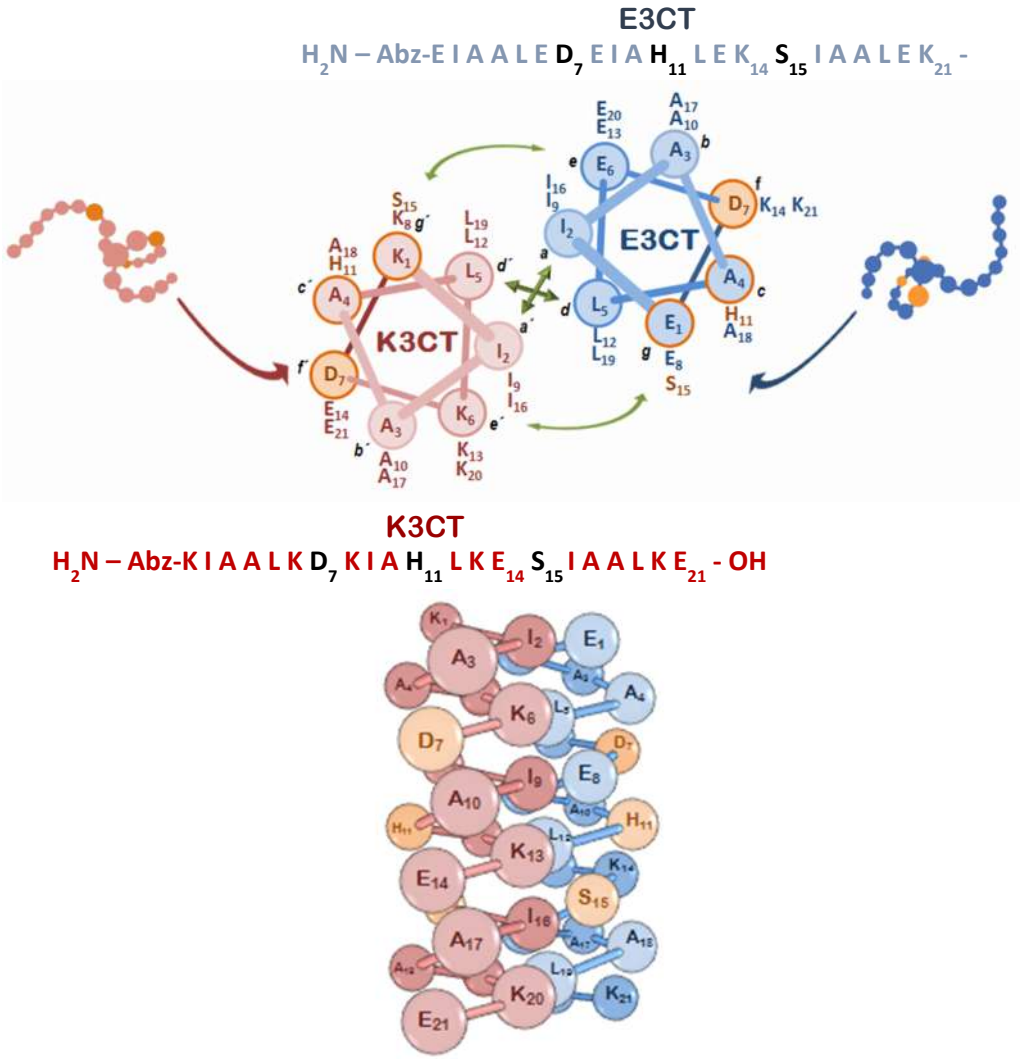



***Figure 6.2:*** *Parallel heterodimeric coiled-coil E3CT/K3CT. Sequence and helical wheel representation (top) and side-view (bottom). Residues corresponding to the catalytic triad are in shown in orange.*

concentration of E3CT present in random-coil form. The involvement of histidine and glutamate in catalysis is suggestive of two mechanisms. In each of the possible mechanisms, glutamate acts as a base to activate histidine. In turn, histidine can directly attack the ester as a nucleophile (Fig. 3.11) or can act as a general base to deprotonate water or another nucleophilic species (Fig. 6.4).[71,74,75,194]

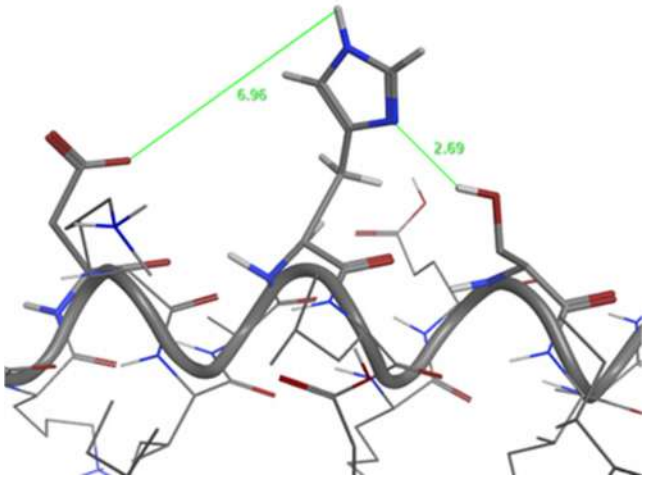



***Figure 6.3:*** *Molecular model showing an inadequate charge relay between the residues aspartate, histidine and serine when spaced in positions* i, i + 4, i + 8 *relative to one another in an α-helix. Molecular modeling performed by J. Mortier.*

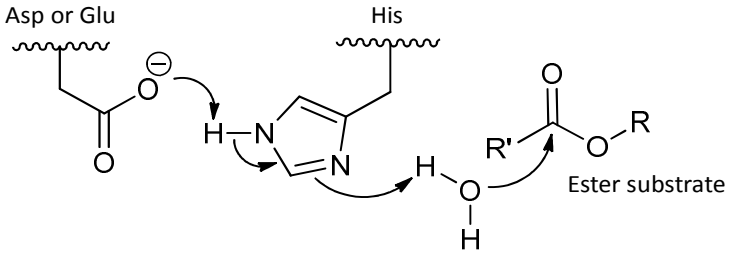



***Figure 6.4:*** *Cooperativity of carboxylic acid and histidine in general base-catalyzed ester hydrolysis.*

To further investigate the cause of catalytic activity by E3CT in its random-coil form, additional E3 variants E3H11 and E3S15 (Fig. 6.5) were assayed (Table 6.2). As before, each of the mixtures E3H11/K3 and E3S15/K3 formed α-helical structures, but lacked significant activity. As a random coil, the variant E3H11 was shown to be as active as E3CT. The variant E3S15 showed only minimal activity. These findings emphasize the importance of histidine and the less likely participation of serine in the catalytic activity of E3CT.

Additional tests were performed to better understand the role histidine and carboxylate-containing residues play in ester hydrolysis in the near neutral pH range. In doing so, the short peptide EEEHEEE of sequence $H_2N$-Abz-EEEHEEE-OH (same glutamate/histidine ratio as E3H11) and the peptide E3H11-Ala (Table 6.1) were synthesized and assayed along with L-histidine for catalytic activity. In E3H11-Ala,

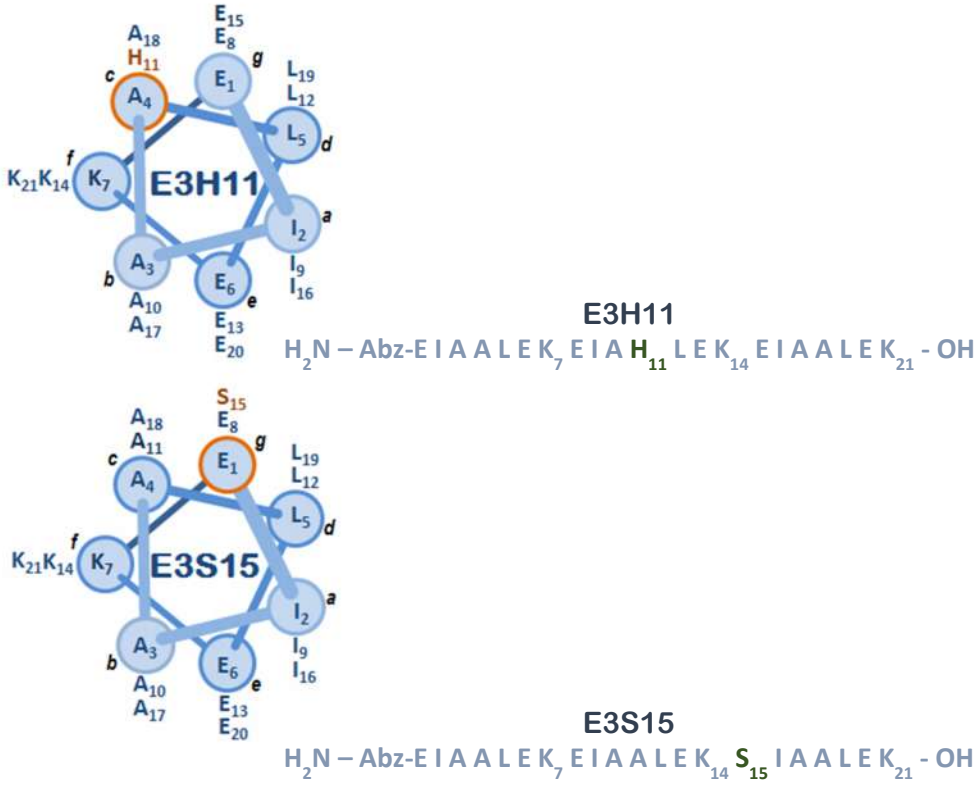



***Figure 6.5:*** *E3 variants E3H11 and E3S15 of the parallel heterodimeric coiled coil E3/K3.*

glutamate and histidine residues were left in the same position relative to each other as in E3H11. However, all other residues were substituted with alanine. Additional lysine residues were added to terminal positions to increase the net charge.

***Table 6.1:*** *Amino acid sequences of peptides E3, E3H11 and E3H11-Ala. Each peptide has unblocked amino and carboxyl termini.*

| | Peptide sequences and heptad register | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Peptide | | | | g | a | b | c | d | e | f | g | a | b | c | d | e | f | g | a | b | c | d | e | f | |
| E3 | | | Abz | E | I | A | A | L | E | $K_7$ | E | I | A | A | L | E | $K_{14}$ | E | I | A | A | L | E | $K_{21}$ | |
| E3H11 | | | Abz | E | I | A | A | L | E | $K_7$ | E | I | A | $H_{11}$ | L | E | $K_{14}$ | E | I | A | A | L | E | $K_{21}$ | |
| E3H11-Ala | Abz | K | A | E | A | A | A | A | E | A | E | A | A | H | A | E | A | E | A | A | A | A | E | A | K |

***Table 6.2:*** *Determined second-order rate constants of Z-L-Phe-ONp hydrolysis at 20 °C in 50 mM Tris/HCl pH 7.3 and 3.2 % acetonitrile in the presence of 50 µM peptide, 1:1 heteromeric peptide mixtures and L-histidine as well as the calculated α-helical content of peptide under the conditions specified.*

| Variant | $k_2$ ($M^{-1}min^{-1}$) | % α-Helix |
|---|---|---|
| E3 | 4 (± 11) | 18 |
| E3/K3 | 6 (± 4) | 70 |
| E3CT | 150 (± 20) | 13 |
| E3CT/K3 | 7 (± 11) | 58 |
| E3H11 | 186 (± 16) | 8 |
| E3H11/K3 | 3 (± 3) | 61 |
| E3S15 | 26 (± 8) | 14 |
| E3S15/K3 | 12 (± 10) | 51 |
| K3 | 3 (± 11) | 17 |
| K3CT | 6 (± 9) | 10 |
| E3/K3CT | 2 (± 12) | 45 |
| E3CT/K3CT | 36 (± 8) | 27 |
| L-Histidine | 20 (± 10) | |
| EEEHEEE | 50 (± 10) | |
| E3H11-Ala | 114 (± 18) | 10 |

Without neighboring carboxylic residues, the second-order rate constant ($k_2$) associated with Z-L-Phe-ONp in the presence of L-histidine was measured to be approximately 9-fold less than E3H11. With the positioning of three glutamate residues on both sides of histidine in the form of EEEHEEE, the second-order rate constant was elevated from 20 to 50 $M^{-1}min^{-1}$. The finding that EEEHEEE corresponds to a second-order rate constant which is just one-third of that belonging

to E3CT or E3H11 suggests the importance of spacing between glutamate and histidine residues within an extended sequence length. As expected, the peptide E3H11-Ala (Table 6.1), which has the same sequence distribution of glutamate and histidine residues as E3H11, was found to enhance ester hydrolysis with a $k_2$ more than twice the value corresponding to EEEHEEE. These results point to a cooperative interaction of histidine with distant, intramolecular glutamate residues in a flexible random coil. Although such contact between histidine and glutamate within an unstructured peptide are likely fleeting, they provide a minimal requirement for cooperative glutamate/histidine-catalyzed ester hydrolysis.

To test for enantioselectivity, the ester substrate Z-D-PheONp was hydrolysed in the presence of peptides E3CT and E3H11. Remarkably, E3CT and E3H11 showed little catalytic activity in the hydrolysis of the D-substrate with second-order rate constants of 2 and 19 $M^{-1}min^{-1}$, respectively. Unfortunately, these enantioselective results are difficult to explain without a defined peptide structure.

The activity shown by E3CT and E3H11 is modest, but could be greatly enhanced by increasing the local peptide concentration, which would induce contact between intermolecular histidine and glutamate residues. One way to increase the local concentration of catalytic peptide is to graft the peptide to a polymer or dendrimer.[195] In addition, conjugation to such a support would allow for easier recovery of catalyst from reaction media via dialysis. Due to its high loading capacity and solubility in water, hyper-branched polyglycerol (hPG) was selected (Fig. 6.6).[196,197] Immobilization of peptide E3H11 to hPG was achieved via copper-catalyzed alkyne–azide cycloaddition (Cu-AAC) (Fig. 6.7).[198] Prior to immobilization, alkyne

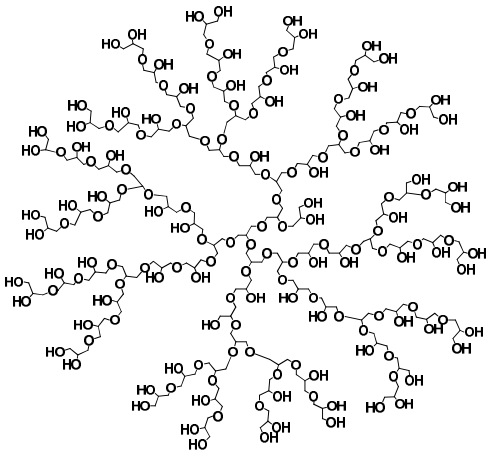


***Figure 6.6:*** *A hyper-branched polyglycerol (hPG) dendrimer*

functionalized peptides E3, E3H11 and E3H11-Ala, named E3yne, E3H11yne and E3H11-Alayne, respectively, were synthesized (8.1.1.5). Azide-functionalized hPG was provided by the research group of Professor Rainer Haag.

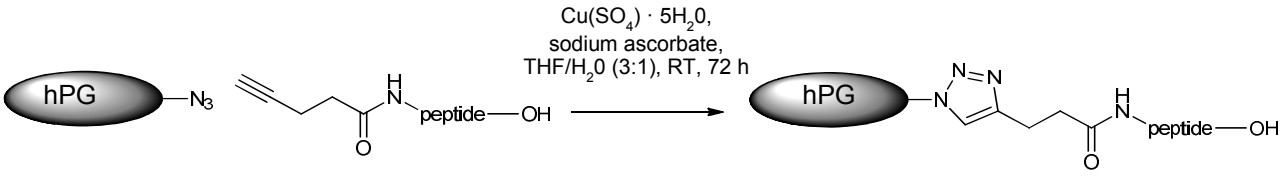



***Figure 6.7:*** *Immobilization of peptide to hyper-branched polyglycerol (hPG) via copper-catalyzed alkyne-azide cycloaddition.*

Although the grafting of peptide to hPG was achieved, further efforts are required to obtain a more complete conjugation. With complete or equal conjugation, the hPG-peptide variants (hPG-E3, hPG-E3H11 and hPG-E3H11-Ala) can then undergo comparative kinetic examination and structural studies as performed on their non-conjugated counterparts (Table 6.2). This includes measuring the α-helicity and the ability of these variants, including hPG-E3H11, to catalyze ester hydrolysis in the presence of ungrafted peptide K3.

In addition to polymers, the fuctionalization of gold monolayer-protected clusters (Au-MPC) with peptides such as E3CT or E3H11 could also be particularly effective in enhancing activity. Not only is the application of Au-MPCs likely to increase effective catalyst concentration, but is also expected to impart an increased affinity for hydrophobic substrates (esters) with the provision of a hydrophobic interface. Studies by Scrimin, Pasquato *et al*. have made use of HS-C8-TEG-ensconced gold nanoparticles to encapsulate hydrophobic substrates. [199] Furthermore, the core size of the nanoparticle plays a role in substrate selectivity as it dictates how tightly the protective monolayer molecules (HS-C8-TEG), shown in Figure 6.6, pack together as they radiate from the nanoparticle.[200] The end result of these studies, was the creation of a nanosystem capable of selective solute encapsulation based on solute size and hydrophobicity. Upon funtionalization with a simple random-coil peptide containing histidine, the $k_2$ corresponding to the hydrolysis of 2,4-dinitrophenylbutanoate (DNPB) was heightened by more than 60-fold at pH 8 (Fig. 6.8).[74] The increased activity under these conditions was attributed to the

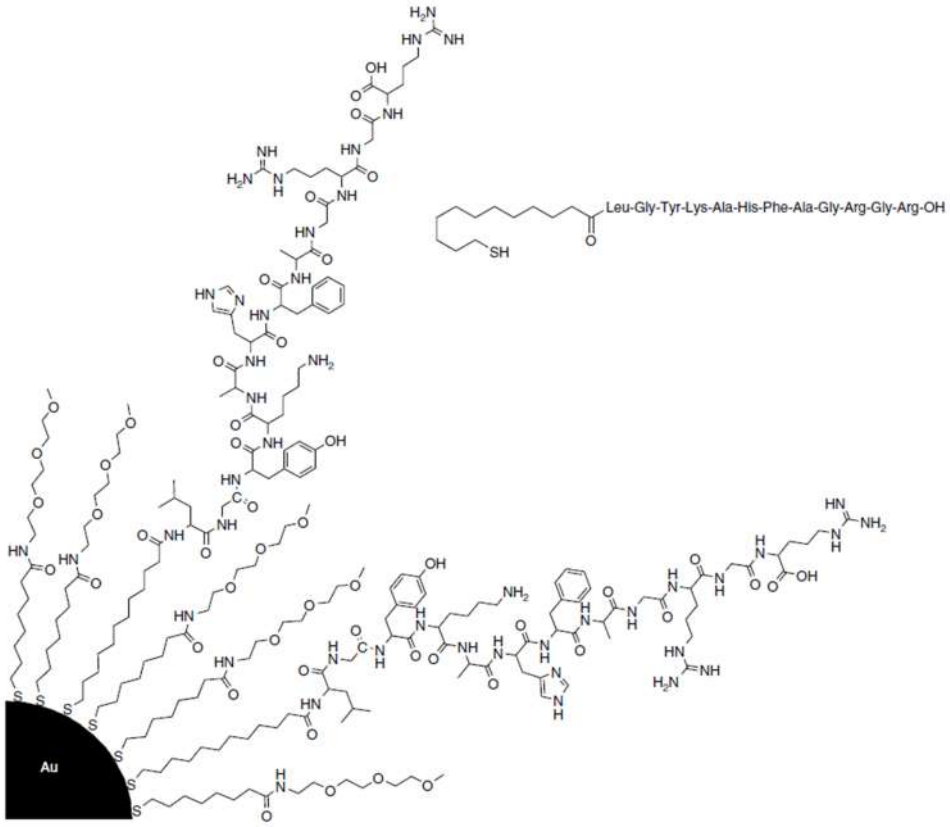



***Figure 6.8:*** *Peptide-functionalized gold monolayer-protected cluster possessing esterase activity. From Guarise et al.*[199]

interaction (general base; Fig. 6.4) between histidine and the terminal carboxyl group of the peptide. These “nanozymes” were also shown to be substrate selective. Although generally more stable, the more lipophilic substrate *N*-benzyloxycarbonyl-L-leucine-*p*-nitrophenyl ester (Z-L-Leu-ONp) was readily hydrolyzed in comparison to *N*-benzyloxycarbonyl-glycine-*p*-nitrophenyl ester (Z-Gly-ONp). It is therefore conceivable that either of the catalytic E3 varients E3CT or E3H11 presented in this work when applied in such a Au-MPC system would selectively accelerate ester hydrolysis. In addition to studies previously performed, the involvement of the simple heterodimeric E3/K3 coiled-coil system, would allow, in the case of E3CT/K3 or E3H11/K3 (Table 6.2; Fig. 6.9), activity regulation.

The merging of monolayer-protected gold nanoparticles with self-assembling peptides presents a promising approach to the construction of artifical enzymes. This modular approach allows the researcher to incorporate and tailor particular elements of natural enzymes into nanozymes. Such elements include substrate affinity and/or selectivity dictated by nanoparticle size and choice of monolayer at the interface and,

of course, catalytic activity according to peptide sequence and the insertion of catalytic machinery therein. The self-assembly of the peptides into specific structures is not only expected to have an effect on reaction acceleration, but on regulation.

Substrate selectivity, catalytic proficiency and regulation are all key features of natural enzymes which need better understanding in order to design catalytic proteins or artificial enzymes. This modular approach aids in the assessment of the individual elements of catalysis such as substrate affinity or catalytic turnover, a delineation made more complicated by the folding dynamics of enzymes and biocatalysts. Peptide-functionalized Au-MPCs could, therefore, be used directly to catalyze reactions or serve as a tool to help piece together more complicated *de novo* protein catalysts.

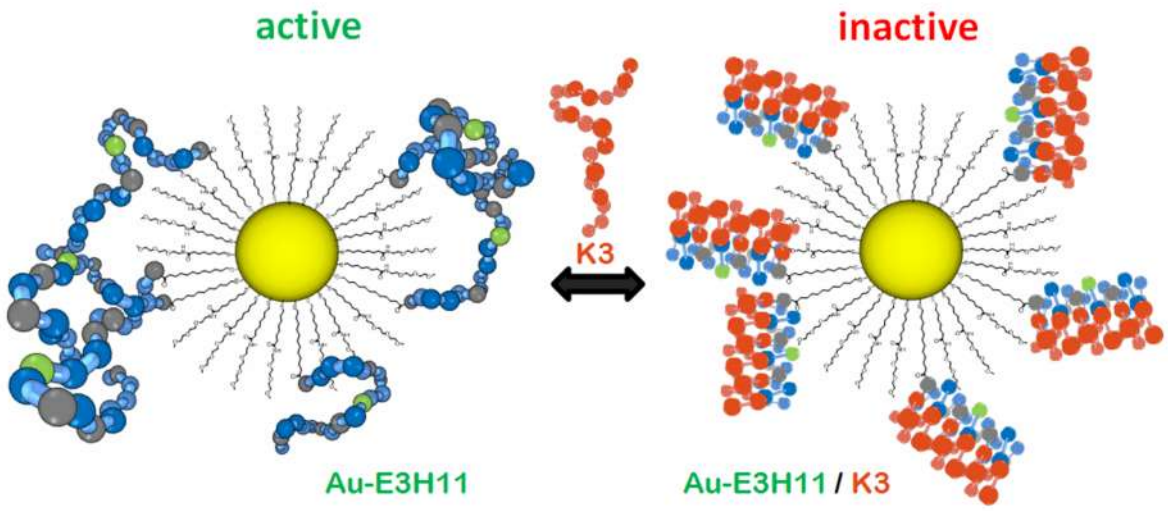



***Figure 6.9:*** *Proposed activity regulation of catalytic E3H11-functionalized monolayer-protected gold clusters via K3-induced heterodimeric coiled coil formation. Glutamate and histidine residues of E3H11 involved in general base-catalyzed ester hydrolysis are shown in gray and green, respectively.*

## 6.2 Amyloid-forming Ac-IHIHIQI-$NH_2$: *De novo* catalytic proficiency with selectivity

Korendovych and DeGrado showed that a *de novo*-designed catalytic protein constructed from a minimal β-sheet sequence can serve as a proficient catalyst of ester hydrolysis (Section 4.2; Fig. 4.11).[188] The peptide Ac-IHIHIQI-$NH_2$ was reported to possess a second-order rate constant for the hydrolysis of *p*NPA of 62 $M^{-1}s^{-1}$ at pH 8 and to lower the pKa of water to 9.3. In conditions which are slightly more basic, its catalytic efficiency approaches a respectable 360 $M^{-1}s^{-1}$. Furthermore, Ac-IHIHIQI-$NH_2$ was reported to catalyze ester hydrolysis via the zinc-bound hydroxide mechanism. Its active form for the hydrolysis of *p*NPA was characterized as an amyloid-like fibril.

The high catalytic activity of Ac-IHIHIQI-$NH_2$ coupled with its interesting structural aspects and aggregation behavior warrant further examination. As depicted in Figure 4.11, the catalytic machinery $(His)_3Zn$ is exposed to the surface. Therefore, it is easy to expect, with the absence of a three-dimensional cleft, little to no substrate selectivity. However, it should not be forgotten that the residues are chiral and well-organized in a tightly-packaged fibril. In addition, Ac-IHIHIQI-$NH_2$ was kept in a stock solution at pH 2 in which it was stable for more than a week prior to being added to the reaction mixture (Fig. 6.10). At this point, the question arises as to whether Ac-IHIHIQI-$NH_2$ would exhibit different catalytic behavior as a fully formed fibril or during the process of self-assembly and fibrillar growth.

Studies performed by Korendovych and DeGrado on their minimalist design Ac-IHIHIQI-$NH_2$ also employed the simple model ester *p*NPA as substrate (Fig. 6.11). As a model ester substrate, *p*NPA meets the most basic requirements, an acetyl acyl component and the standard nitrophenol (actually nitrophenolate, pKa = 7) as chromophore to monitor reaction progress. Because *p*NPA is achiral and lacks any significant functional group or hydrophobic portion, it is unlikely to participate in any specific interaction with the catalyst other than presenting its ester group for hydrolysis. To further study substrate selectivity of the catalyst Ac-IHIHIQI-$NH_2$ in its incubated, fibrillar form and while it is in the process of self-assembly, the work presented here applied both L- and D-enantiomers of benzyloxycarbonyl-protected phenylalanine *p*-nitrophenyl ester (Z-Phe-ONp) and butyloxycarbonyl-protected

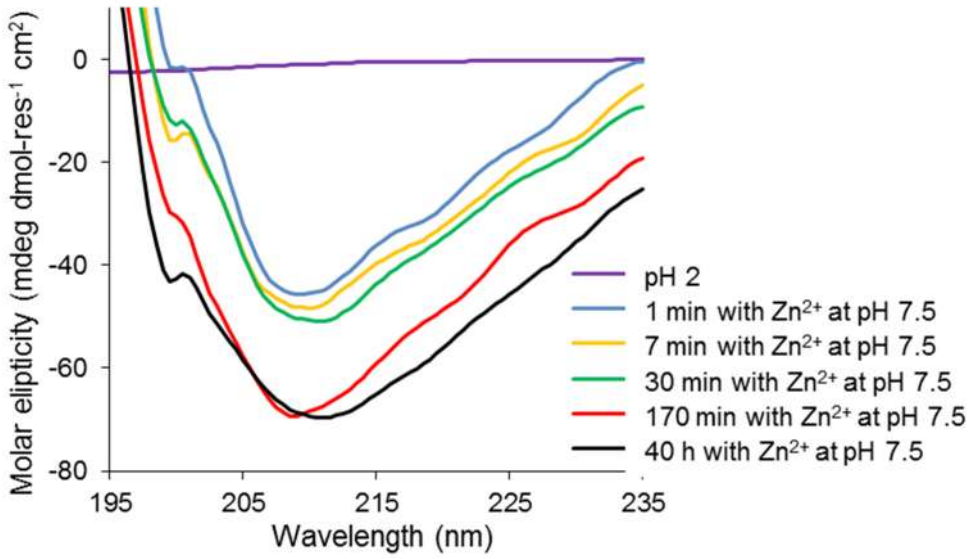



***Figure 6.10:*** *CD spectra of Ac-IHIHIQI-$NH_2$ (10 µM) at pH 2 (10 mM HCl) and at various time points following its addition to 50 mM Tris/HCl pH 7.5 with 100 µM $ZnCl_2$.*

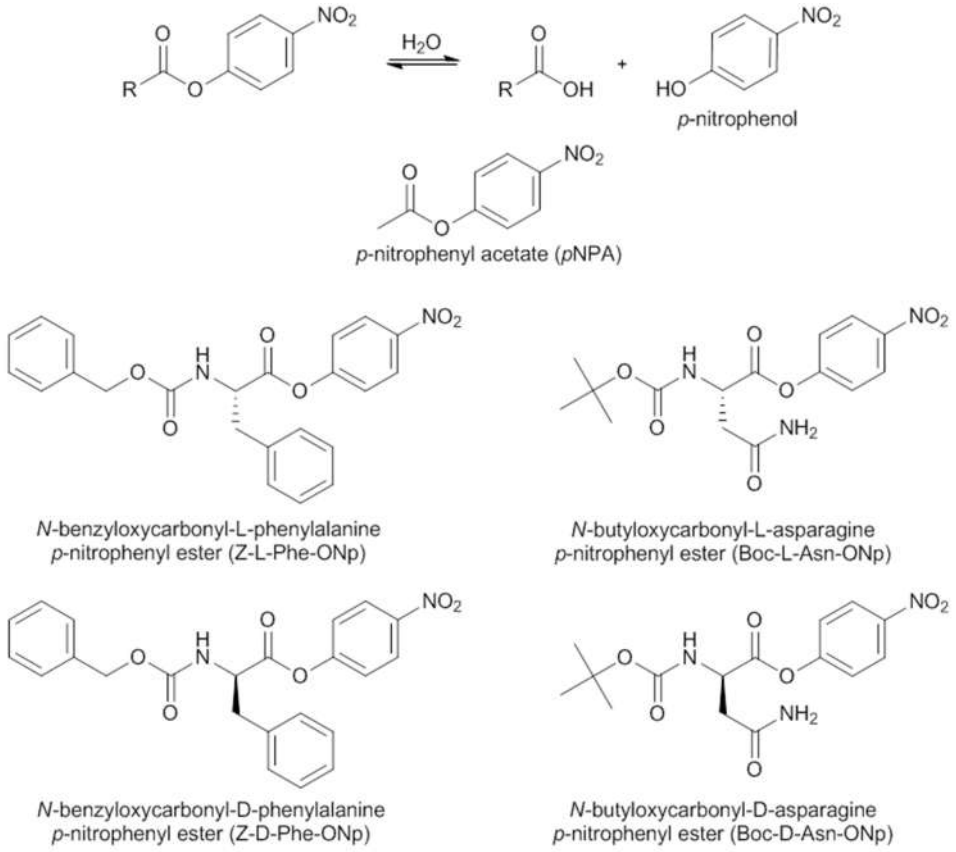



***Figure 6.11:*** *Nitrophenyl esters employed as substrates to study the kinetics of catalyzed ester hydrolysis.*

asparagine *p*-nitrophenyl ester (Boc-Asn-ONp) (Fig. 6.11).

Due to the limited solubility of substrate Z-Phe-ONp in water, pseudo-first-order kinetic experiments were performed. The production of *p*-nitrophenolate as mediated by Ac-IHIHIQI-$NH_2$ in its pre-incubated, fibrillar form and during the process of self-assembly is shown in Figures 6.12-6.14. Not surprisingly, the hydrolysis of *p*NPA was most rapid when treated with fibrils incubated with zinc 24 hours prior to the reaction (Fig. 6.12). The initial reaction velocity was more than 500 times greater in the presence of the fibrils than without catalyst. The second-order rate constant at pH 7.5 was determined to be 87 $M^{-1}s^{-1}$ (Table 6.3), based on the previous finding that two peptides constitute one active site.[188] On the other hand, the second-order rate constant corresponding to the actively assembling catalyst was just over 40 $M^{-1}s^{-1}$, which is in agreement with the results of Korendovych and DeGrado obtained via saturation kinetics.[188] The faster hydrolysis observed in the presence of the incubated fibrils was due to the greater number of organized and established active sites as opposed to catalytic peptides still in the process of aggregation.

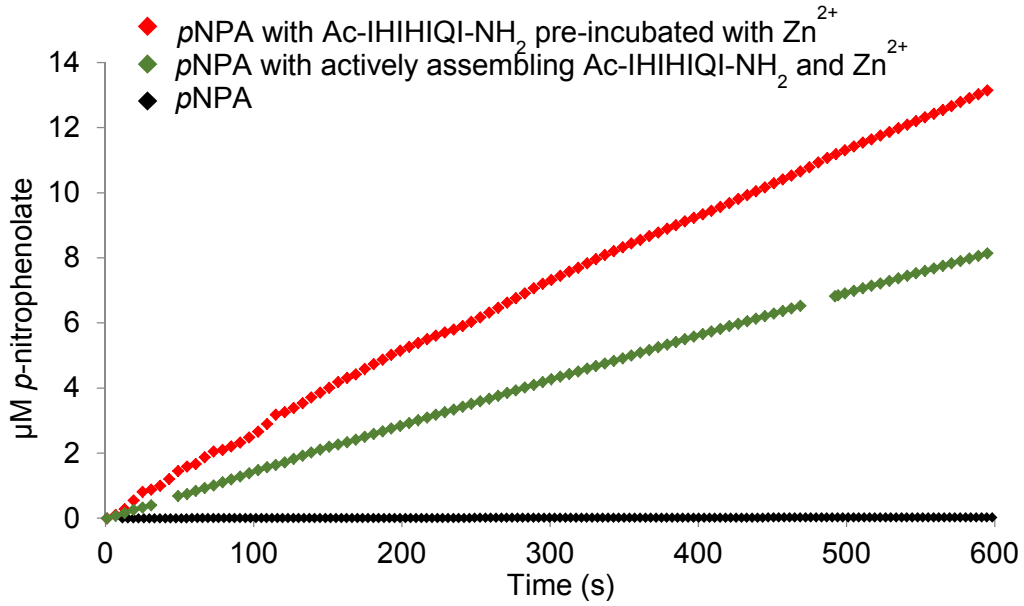



***Figure 6.12:*** *Production of* p*-nitrophenolate upon Ac-IHIHIQI-$NH_2$-mediated hydrolysis of* p*-nitrophenyl acetate (*p*NPA) in the presence of zinc. Reactions were performed in Tris/HCl (50 mM with 100 μM $ZnCl_2$) pH 7.5 at 20°C with 50 μM substrate and 10 μM Ac-IHIHIQI-$NH_2$.*

***Table 6.3:*** *Rate constants corresponding to the hydrolysis of esters pNPA and L- and D-enantiomers of Z-Phe-ONp and Boc-Asn-ONp in the presence of catalytic peptide Ac-IHIHIQI-$NH_2$ in its fully fibrullar form following 24 h incubation with zinc and with zinc in its actively self-assembling form. Reactions were perfomed at 20°C and pH 7.5 (50 mM Tris/HCl with 100 µM $ZnCl_2$) with 50 µM pNPA and 10 µM Z-Phe-ONp and Boc-Asn-ONp. A peptide concentration of 10 µM (2 peptides form 1 active site = 5 µM of catalyst) was used. The apparent $k_1$ is the observed pseudo-first order rate constant and does not account for uncatalyzed background hydolysis*

| | uncatalyzed | Ac-IHIHIQI-$NH_2$ pre-incubated with $Zn^{2+}$ | | Ac-IHIHIQI-$NH_2$ actively assembling with $Zn^{2+}$ | |
|---|---|---|---|---|---|
| Substrate | $k_1$ ($s^{-1}$) | Apparent $k_1$ ($s^{-1}$) | $k_2$ ($M^{-1}s^{-1}$) | Apparent $k_1$ ($s^{-1}$) | $k_2$ ($M^{-1}s^{-1}$) |
| *p*NPA | 7 (± 2) x $10^{-5}$ | 5.1 (± 0.2) x $10^{-4}$ | 87 ± 4 | 2.7 (± 0.2) x $10^{-4}$ | 41 ± 4 |
| Z-L-Phe-ONp | 1.53 (± 0.02) x $10^{-3}$ | 1.11 (± 0.05) x $10^{-2}$ | 1.9 (± 0.1) x $10^{3}$ | 1.13 (± 0.04) x $10^{-1}$ | 2.23 (± 0.08) x $10^{4}$ |
| Z-D-Phe-ONp | 1.53 (± 0.02) x $10^{-3}$ | 1.00 (± 0.05) x $10^{-2}$ | 1.7 (± 0.1) x $10^{3}$ | 3.87 (± 0.02) x $10^{-2}$ | 7.43 (± 0.04) x $10^{3}$ |
| Boc-L-Asn-ONp | 3.80 (± 0.02) x $10^{-3}$ | 1.21 (± 0.02) x $10^{-2}$ | 1.66 (± 0.04) x $10^{3}$ | 1.83 (± 0.03) x $10^{-2}$ | 2.90 (± 0.06) x $10^{3}$ |
| Boc-D-Asn-ONp | 3.80 (± 0.02) x $10^{-3}$ | 8.52 (± 0.02) x $10^{-3}$ | 9.44 (± 0.04) x $10^{2}$ | 9.49 (± 0.07) x $10^{-3}$ | 1.14 (± 0.01) x $10^{3}$ |

Remarkably, the opposite effect was observed when Z-L-Phe-ONp was employed as substrate. In this case, Ac-IHIHIQI-$NH_2$ was most active when self-assembly and the reaction occurred simultaneously (Fig. 6.13). The second-order rate constant of the actively assembling catalyst acting on Z-L-Phe-ONp at pH 7.5 was found to be 2.2 x $10^4$ $M^{-1}s^{-1}$, ten-fold higher than the 1.9 x $10^3$ $M^{-1}s^{-1}$ rate constant observed with incubated Ac-IHIHIQI-$NH_2$. The turnover rate (apparent first-order rate constant) of the freshly prepared variant was more than 70-fold greater than when uncatalyzed.

This finding is perhaps the result of favorable interactions between the hydrophobic domains of the substrate and hydrophobic portions of the self-assembling catalyst. As aggregation occurs and the active site is established, the substrate may, therefore, already be in the vicinity (Fig. 6.16). Such hydrophobic interactions would be less likely in the fully formed fibril, which has its hydrophobic isoleucine side chains well buried prior to the approach of substrate. Heightened catalytic efficiency during self-assembly was not the case when *p*NPA was employed as a substrate because it lacks the large hydrophobic groups present in Z-L-Phe-ONp.

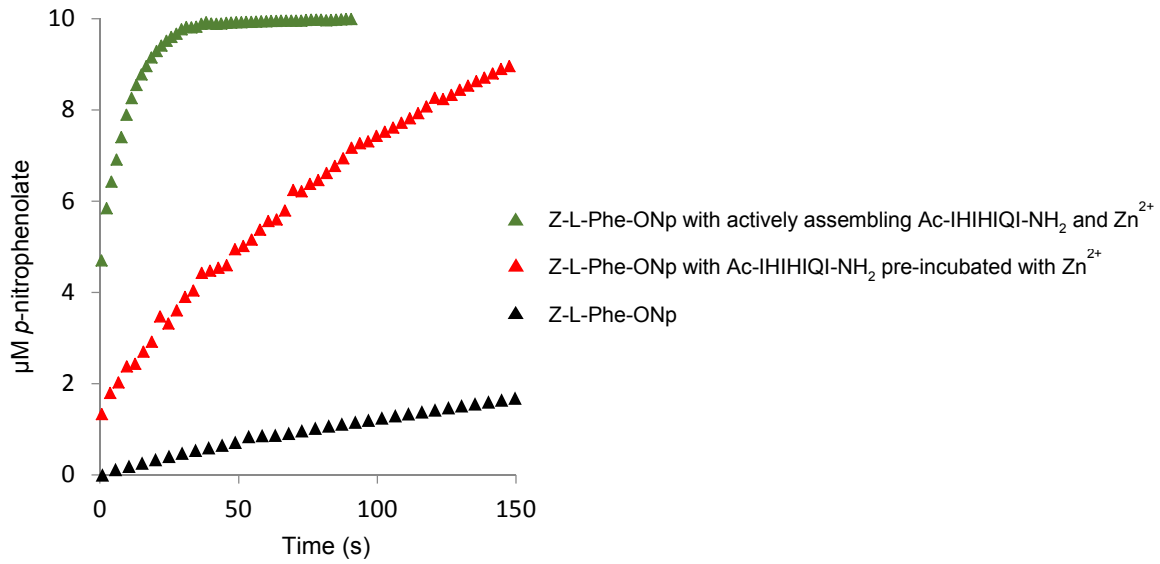



***Figure 6.13:*** *Production of* p*-nitrophenolate upon Ac-IHIHIQI-$NH_2$-mediated hydrolysis of* N*-benzyloxycarbonyl-L-phenylalanine* p*-nitrophenyl ester (Z-L-Phe-ONp) in the presence of zinc. Reactions were performed in Tris/HCl (50 mM with 100 µM $ZnCl_2$) pH 7.5 at 20°C with 10 µM substrate and 10 µM Ac-IHIHIQI-$NH_2$.*

The reduced catalysis produced by the fibrillar form of Ac-IHIHIQI-$NH_2$ upon the substrate Z-L-Phe-ONp may also be due to steric effects. The bulky benzyloxycarbonyl protecting group and benzyl side-chain of phenylalanine may not approach the broad surface of fully formed fibrils as easily as it would smaller assemblies.

Unlike Z-L-Phe-ONp, which was quickly cleaved by the self-assembling form of Ac-IHIHIQI-$NH_2$, hydrolysis of the more hydrophilic Boc-L-Asn-ONp occurred less rapidly (Fig. 6.14). In this case, the second-order rate constants corresponding to the incubated and freshly prepared forms were 1.6 x $10^3$ $M^{-1}s^{-1}$ and 2.7 x $10^3$ $M^{-1}s^{-1}$, respectively. Although hydrolysis of Boc-L-Asn-ONp is generally faster than Z-L-Phe-ONp, kinetic studies showed that Ac-IHIHIQI-$NH_2$, when self-assembling in the presence of $Zn^{2+}$, was clearly selective for the more hydrophobic Z-L-Phe-ONp (Fig. 6.15). The ability of Ac-IHIHIQI-$NH_2$ to exhaust the hydrophobic Z-L-Phe-ONp more than 5 times faster than the more reactive, yet more hydrophilic Boc-L-Asn-ONp clearly suggests that the hydrophobic nature of the substrate plays a role in the selectivity of the self-assembling catalytic peptide. Furthermore, substrate selectivity of Ac-IHIHIQI-$NH_2$ was greatly reduced when the catalyst was allowed to incubate with divalent zinc 24 hours prior to the initiation of the reaction (Fig. 6.17).[68,201]

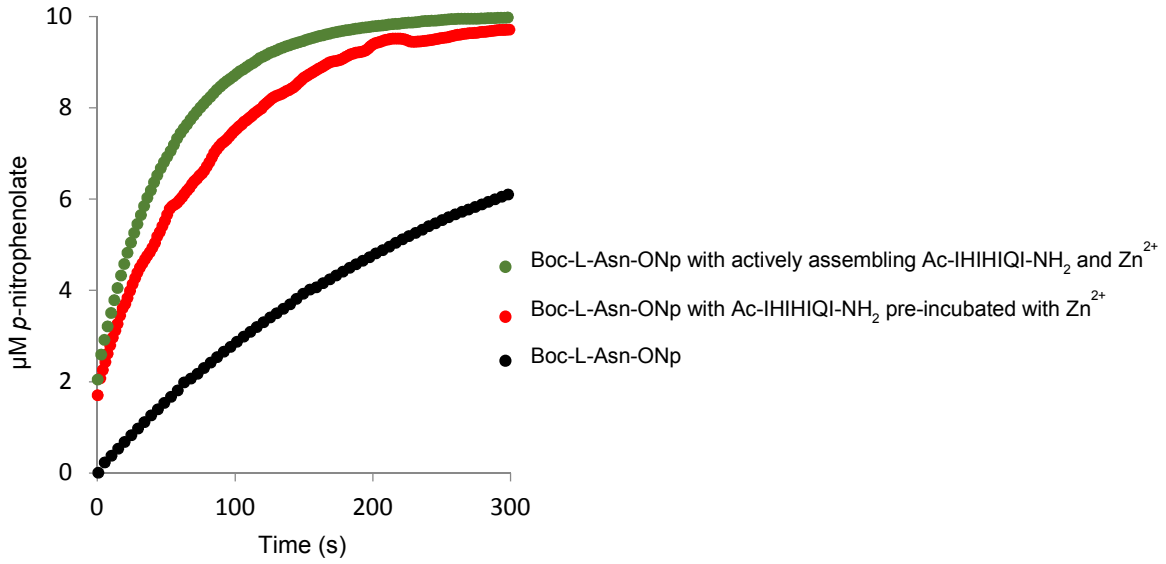



***Figure 6.14:*** *Production of* p*-nitrophenolate upon Ac-IHIHIQI-$NH_2$-mediated hydrolysis of* N*-butyloxycarbonyl-L-asparagine* p*-nitrophenyl ester (Boc-L-Asn-ONp) in the presence of zinc. Reactions were performed in Tris/HCl (50 mM with 100 µM $ZnCl_2$) pH 7.5 at 20°C with 10 µM substrate and 10 µM Ac-IHIHIQI-$NH_2$.*

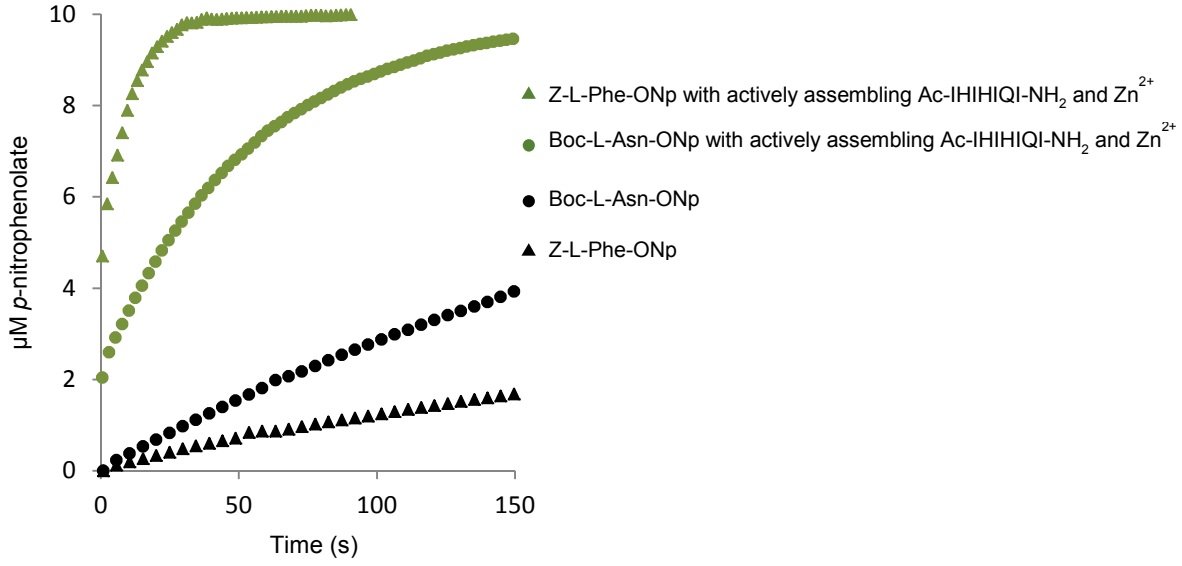



***Figure 6.15:*** *Production of* p*-nitrophenolate upon the hydrolysis of Z-L-Phe-ONp and Boc-L-Asn-ONp with Ac-IHIHIQI-$NH_2$ present in the process of fibril formation and growth with zinc. Reactions were performed in Tris/HCl (50 mM with 100 µM $ZnCl_2$) pH 7.5 at 20°C with 10 µM substrate and 10 µM Ac-IHIHIQI-$NH_2$.*

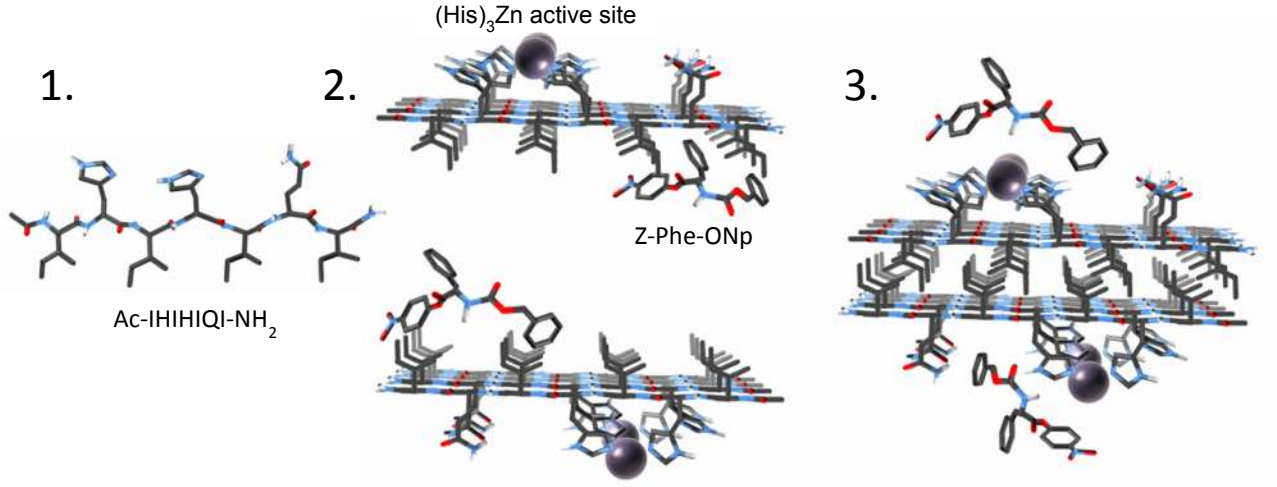



***Figure 6.16:*** *Primary Ac-IHIHIQI-$NH_2$ structure (1). Proposed interaction between hydrophobic substrate Z-Phe-ONp and exposed isoleucine residues of the catalytic peptide Ac-IHIHIQI-$NH_2$ during self-assembly (2). Tight packing of the isoleucine residues expels the substrate into the vicinity of the hydrolytic $(His)_3Zn$ active site (3).*

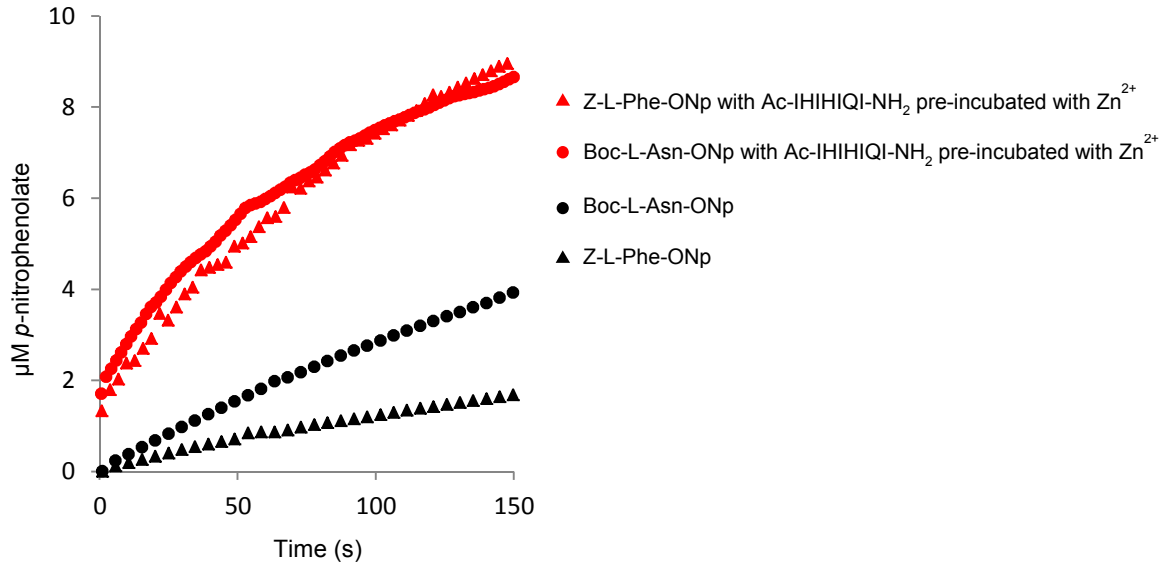



***Figure 6.17:*** *Production of* p*-nitrophenolate upon the hydrolysis of Z-L-Phe-ONp and Boc-L-Asn-ONp in the presence of fibrillar Ac-IHIHIQI-$NH_2$ incubated with zinc. Reactions were performed in Tris/HCl (50 mM with 100 µM $ZnCl_2$) pH 7.5 at 20°C with 10 µM substrate and 10 µM Ac-IHIHIQI-$NH_2$.*

In addition to substrate selectivity, both forms of Ac-IHIHIQI-$NH_2$ in the presence of zinc were examined for enantioselectivity using Z-D-Phe-ONp and Boc-D-Phe-ONp as substrates. While actively assembling, Ac-IHIHIQI-$NH_2$ in the presence of zinc was found to be selective for the L-enantiomer of both substrates (Fig. 6.18 and Fig. 6.19). The ratio of second-order rate constants ($k_{2L}$:$k_{2D}$) resulted in enantioselectivity values ($E_{L/D}$) of 3.0 and 2.5 for Z-Phe-ONp and Boc-Asn-ONp, respectively. However, enantioselectivity of Z-L-Phe-ONp ($E_{L/D}$ = 1.1) was greatly diminished when Ac-IHIHIQI-$NH_2$ was incubated with zinc prior to use (Fig. 6.20). To a lesser extent, this was also observed with Boc-Asn-ONp ($E_{L/D}$ = 1.7; Fig. 6.21).

Added to its catalytic capability, the simple β-sheet structure of Ac-IHIHIQI-$NH_2$ is also reasonably selective when carefully applied. The most encountered limitation of β-sheet structures in *de novo* design is solubility. The results shown here, hint at what might be possible, especially in terms of substrate selectivity, if such structures could be stabilized in solution. Future efforts to stabilize sequences similar to IHIHIQI between α-helices could afford additional stability and define a three-dimensional active site enhanced by secondary interactions and additional selectivity.

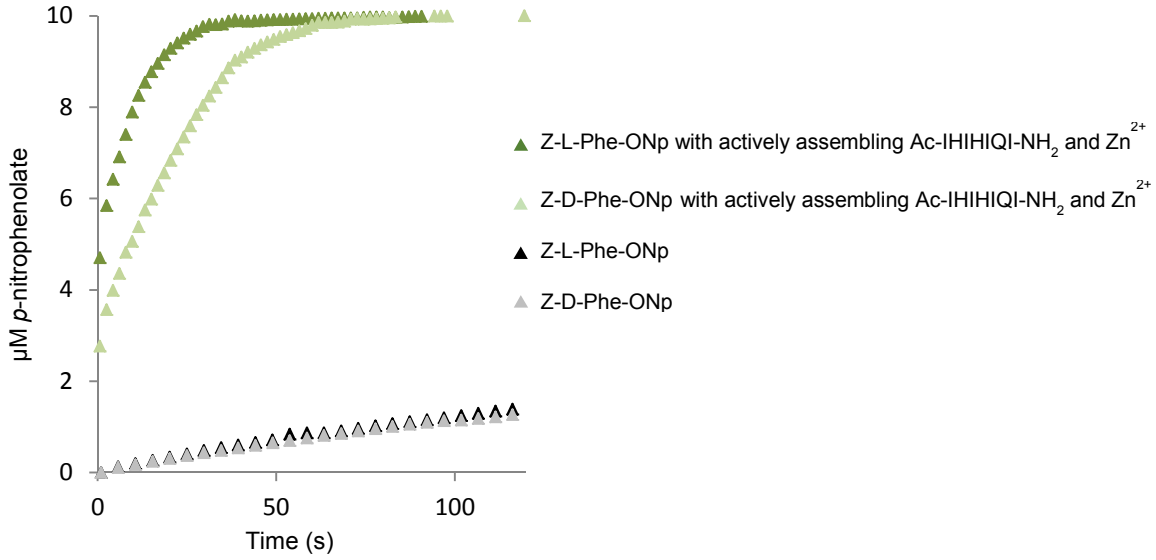



***Figure 6.18:*** *Production of* p*-nitrophenolate upon the hydrolysis of Z-L-Phe-ONp and Z-D-Phe-ONp with Ac-IHIHIQI-$NH_2$ present in the process of fibril formation and growth with zinc. Reactions were performed in Tris/HCl (50 mM with 100 µM $ZnCl_2$) pH 7.5 at 20°C with 10 µM substrate and 10 µM Ac-IHIHIQI-$NH_2$.*

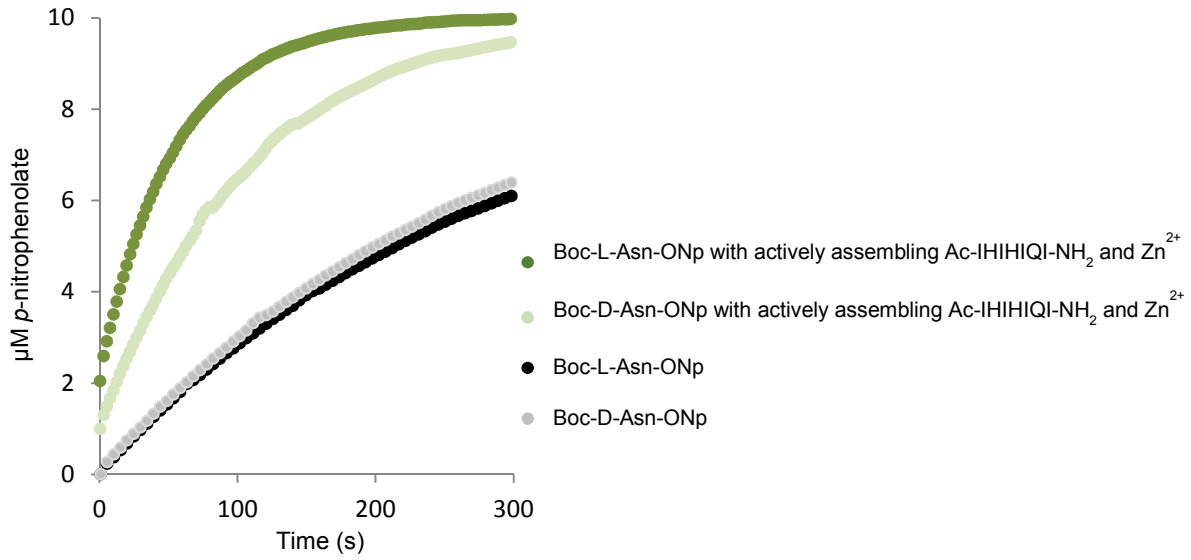



***Figure 6.19:*** *Production of* p*-nitrophenolate upon the hydrolysis of Boc-L-Asn-ONp and Boc-D-Asn-ONp with Ac-IHIHIQI-$NH_2$ present in the process of fibril formation and growth with zinc. Reactions were performed in Tris/HCl (50 mM with 100 µM $ZnCl_2$) pH 7.5 at 20°C with 10 µM substrate and 10 µM Ac-IHIHIQI-$NH_2$.*

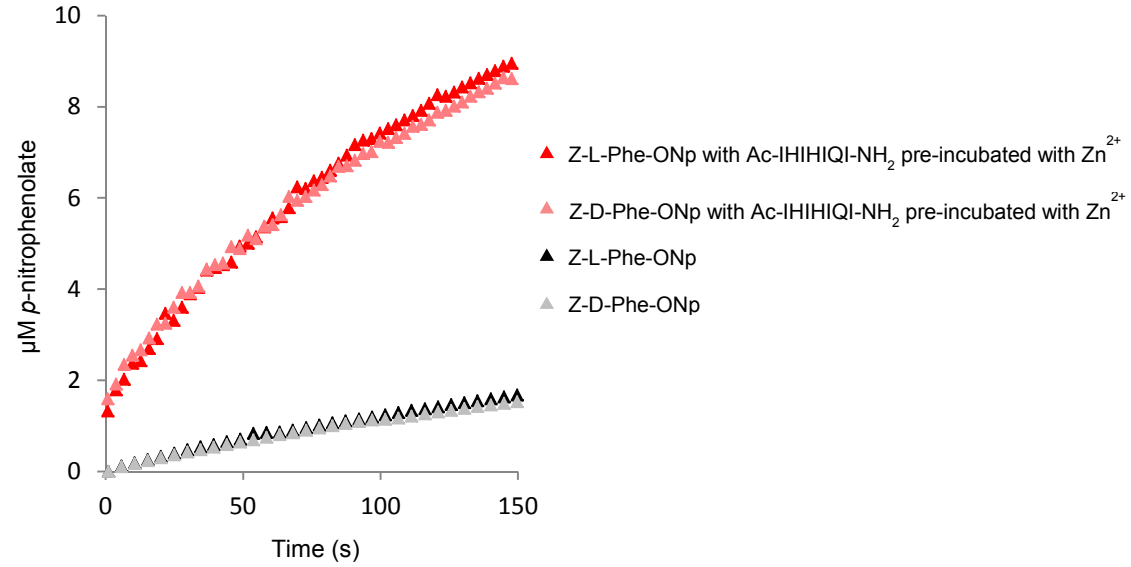



***Figure 6.20:*** *Production of* p*-nitrophenolate upon the hydrolysis of Z-L-Phe-ONp and Z-D-Phe-ONp in the presence of fibrillar Ac-IHIHIQI-$NH_2$ incubated with zinc. Reactions were performed in Tris/HCl (50 mM with 100 µM $ZnCl_2$) pH 7.5 at 20°C with 10 µM substrate and 10 µM Ac-IHIHIQI-$NH_2$.*

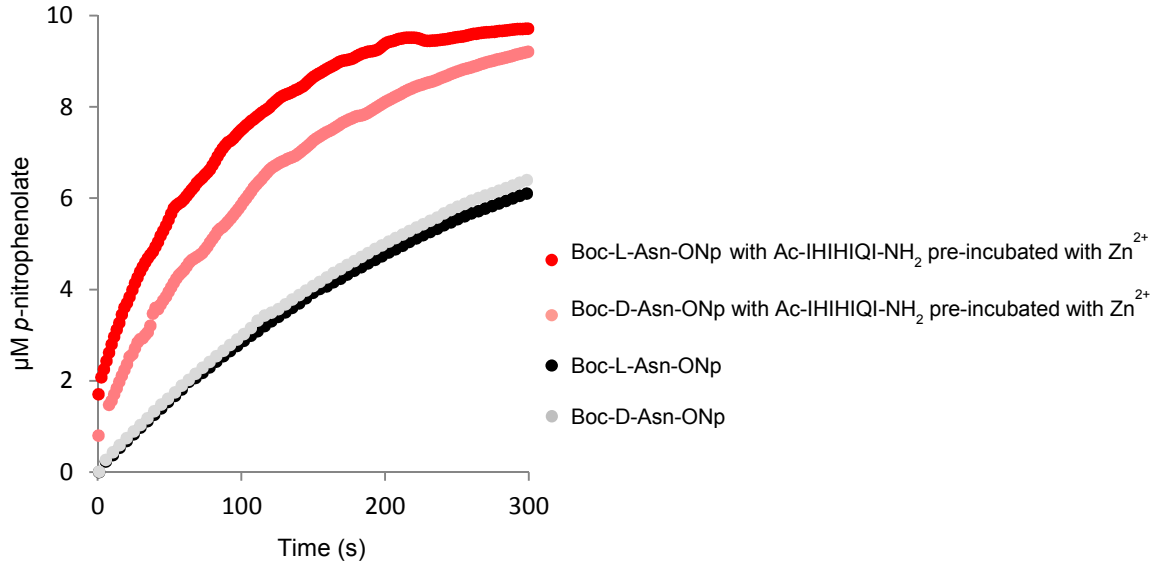



***Figure 6.21:*** *Production of* p*-nitrophenolate upon the hydrolysis of Boc-L-Asn-ONp and Boc-D-Asn-ONp in the presence of fibrillar Ac-IHIHIQI-$NH_2$ incubated with zinc. Reactions were performed in Tris/HCl (50 mM with 100 µM $ZnCl_2$) pH 7.5 at 20°C with 10 µM substrate and 10 µM Ac-IHIHIQI-$NH_2$.*

## 6.3 *De novo* peptide hexamer recast as metalloenzyme

Limitations facing the *de novo* design of catalytic proteins are largely centered on the inability to position a functional active site within a cleft which also provides sufficient substrate accommodation. Although peptides offer simple building blocks for bottom-up design, their small size reduces the number of stabilizing interactions needed to form and retain a proficiently catalytic species.

While focusing our attention on the *de novo* design of peptides possessing esterase activity, we have studied random-coil peptides (Section 6.1) which are just slightly catalytic as they contain catalytic machinery, but are unable to form a stable structure or to accommodate substrate. Also, we have studied the catalytic properties of small β-sheet peptides in the presence of zinc in the process of amyloid formation (Section 6.2). Such assemblies, though reasonably active, tend to lose the ability to accommodate substrates and, in turn, their substrate selectivity once the fibril is formed.

When researching possible *de novo* α-helical designs which to catalyze ester hydrolysis, particular attention was paid to whether or not the assembly could provide a cleft for the accommodation of both substrate and water as well as the presentation of catalytic machinery. To meet these criteria, a coiled-coil assembly featuring histidine within its hydrophobic core was sought. Since histidine within a hydrophobic core of a coiled-coil structure is likely to encourage destabilization, particularly when charged at acidic pH values, the catalytic mechanism employed would need to take place under basic conditions. Therefore, the use of zinc as applied in the zinc-bound hydroxide nucleophile mechanism (Figs. 2.4, 3.14) becomes a relevant choice. The employment of zinc in this case is expected not only to enhance esterase activity, as was the case in Section 6.2, but might also aid in structural stabilization through the coordination of histidine within the hydrophobic core at near-neutral pH conditions.

As pointed out by Zastrow *et al.,*[27] an obstacle facing *de novo* hydrolases based on the coiled-coil design is the inability to channel water to an active site nestled within the hydrophobic core. Taking this into account, a coiled-coil structure of a greater oligomerization state than the trimeric $Hg(II)_SZn(II)_N(TRIL9CL23H)_3$ (Fig. 4.10) reported by Zastrow, Pecoraro *et al.* was envisioned. Ultimately, the *de novo* peptide hexamer CC-HexH24 designed in the laboratories of D. N. Woolfson was selected.[202]

The peptide model CC-Hex (Table 6.4) reported by Zaccai *et al*., is the only parallel hexameric bundle featuring an interior channel. Although lined exclusively with hydrophobic leucine and isoleucine residues, the lumen of this coiled coil was shown by X-ray diffraction to be approximately 6 Å in diameter and water permeable. The design of this hexamer is based on the same principles discussed in Section 4.1.1. However, unlike previous models which place charged, salt-bridge-forming residues at positions *e* and *g*, CC-Hex has a neutral alanine at position *e* while relocating the salt-bridge-forming residue to position *b* (Fig. 6.22). This residue rearrangement results in the conversion of a coiled-coil tetramer into a hexamer. Furthermore, following the placement of a histidine residue in position *a* within the lumen, the peptide was still capable of parallel hexameric coiled-coil formation as shown by its crystal structure (Fig. 6.23). Unlike CC-Hex, the histidine-containing variant CC-HexH24 was not found to form an ideal hexameric species.[202]

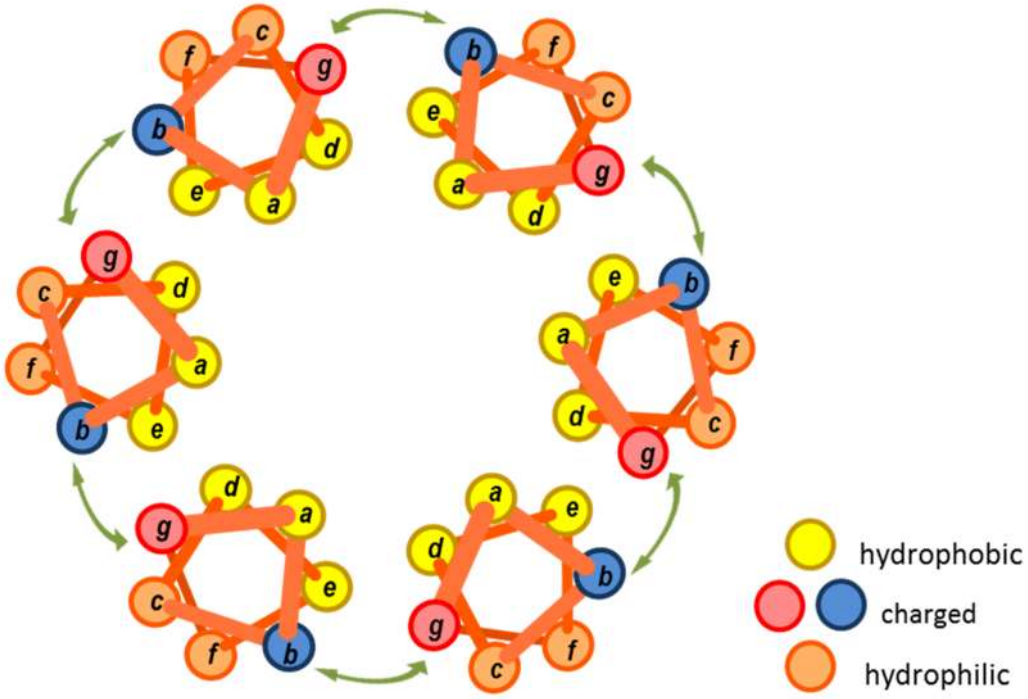



***Figure 6.22**: Helical wheel projection of the parallel hexameric coiled-coil motif CC-Hex. Yellow* a *and* d *positions are occupied, respectively, by hydrophobic leucine and isoleucine residues. Interhelical salt bridges occur between positions* g *(Glu) and* b *(Lys). Charge neutral alanine residues are located in position* e.

To study the catalytic esterase activity of CC-HexH24 (Table 6.4), the substrate *p*NPA was used. It is important to point out that the substrate *p*NPA is well-suited for the assessment of general esterase activity. Unlike many ester substrates, *p*NPA is quite soluble in aqueous solutions and its hydrolysis can be easily monitored and assessed by Michaelis-Menten parameters. However, as already discussed in the previous section, *p*NPA offers little information in regard to enzyme substrate affinity. Initial experiments were performed to test the catalytic ability of CC-HexH24 and the effect $Zn^{2+}$ has on the rate of *p*NPA hydrolysis and secondary peptide structure.

***Table 6.4:*** *Amino acid sequences and heptad register of CC-HexH24, CC-Hex and CC-HexH24 scrambled.*

| | Peptide sequences and heptad register |
|---|---|
| **Name** | *g a b c d e f g a b c d e f g a b c d e f g a b c d e f g* |
| CC-Hex-$H_{24}$ | $H_2N$- G E L K A I A Q E L K A I A K E L K A I A W E H K A I A Q G A G-OH |
| CC-Hex | $H_2N$- G E L K A I A Q E L K A I A K E L K A I A W E L K A I A Q G A G-OH |
| CC-Hex-$H_{24}$ scrambled | $H_2N$- E E K A H K L G A K I G L I Q E A I A A K A W E A A L Q A I K G-OH |

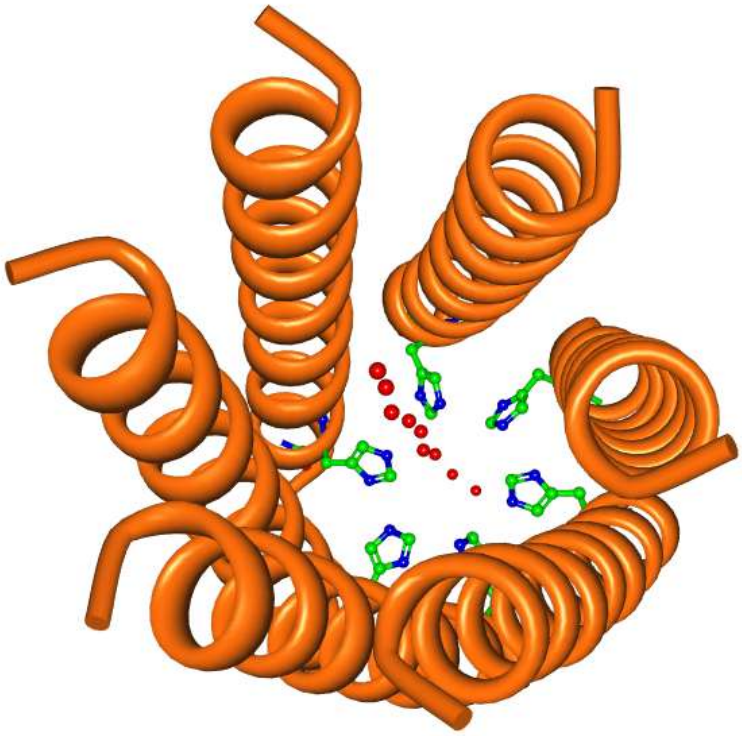


***Figure 6.23****: Crystal structure of parallel hexameric coiled-coil peptide CC-HexH24 showing the orientation of histidine residues and water (red spheres) within its lumen. Crystallization performed by Zaccai* et al.[202] *Image rendered with RCSB Protein Workshop.*[58]

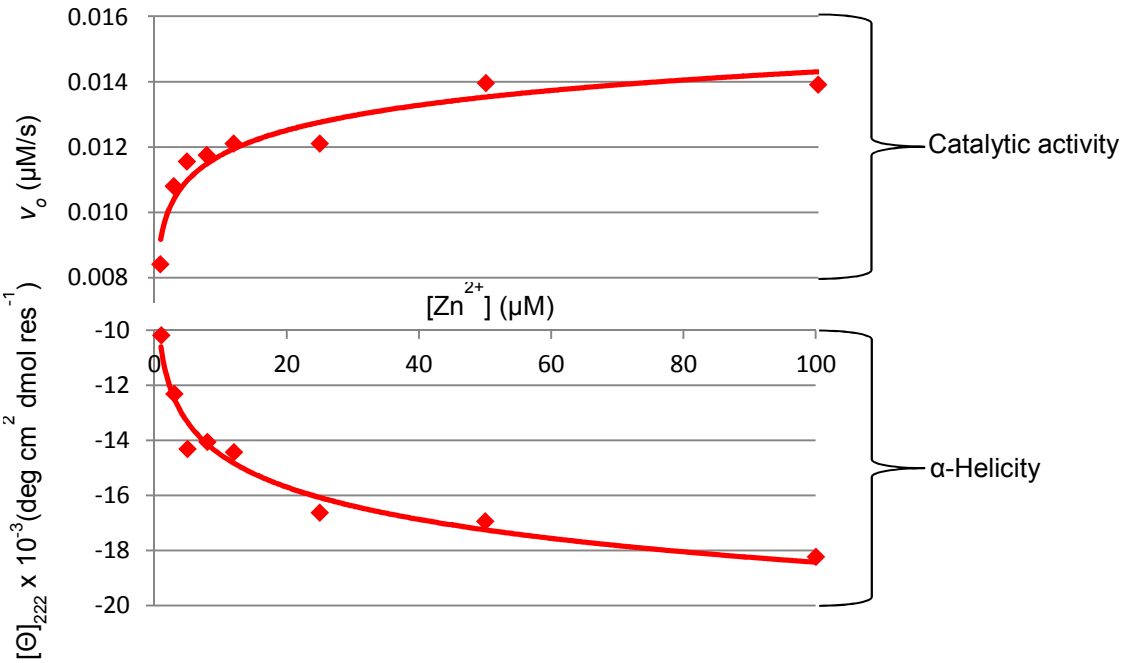



***Figure 6.24:*** *The mirrored effect of $Zn^{2+}$ in both increasing the rate of CC-HexH24-catalyzed* p*NPA hydrolysis and α-helicity at pH 7.5. The initial rates of reaction ($v_o$) were determined in the presence of 50 μM CC-HexH24 and with an initial substrate concentration of 1 mM. All reactions were carried out at 25°C in 50 mM Tris/HCl and 2 % acetonitrile. Values of molar ellipticity at 222 nm, $[\Theta]_{222}$, were determined under identical conditions.*

Interestingly, at pH 7.5, the effect that varying $Zn^{2+}$ concentrations had on CC-HexH24 activity (as shown by initial reaction rate $v_o$) was found to mirror the effect it had on α-helicity (as shown by molar ellipticity at 222 nm, $\Theta_{222}$, Fig. 6.24). These findings indicate that $Zn^{2+}$ not only directly affects the rate of *p*NPA hydrolysis when present within CC-HexH24, but may do so indirectly by increasing the α-helicity of the peptide structure. In addition, thermal denaturation studies showed that $Zn^{2+}$ enhances the thermal stability at pH 7.5 of CC-HexH24 (Fig. 6.25).

Under conditions of increased basicity, CC-HexH24 was found to assume an α-helical structure regardless of zinc present (Fig. 6.26). This is due to the increased likelihood of histidine being deprotonated or uncharged and, therefore, more stable within a hydrophobic environment. In other words, zinc coordination is no longer needed to provide structural stabilization. On the other hand, as the conditions become more basic, the esterase activity of CC-HexH24 with $Zn^{2+}$ increases (Fig. 6.27). Furthermore, Figure 6.25 shows that the initial reaction rates begin to plateau as the $Zn^{2+}$ concentration approaches 8.3 μM or in this case 1/6 molar equivalent of CC-HexH24, which suggests that the model's oligomerization is for the most part

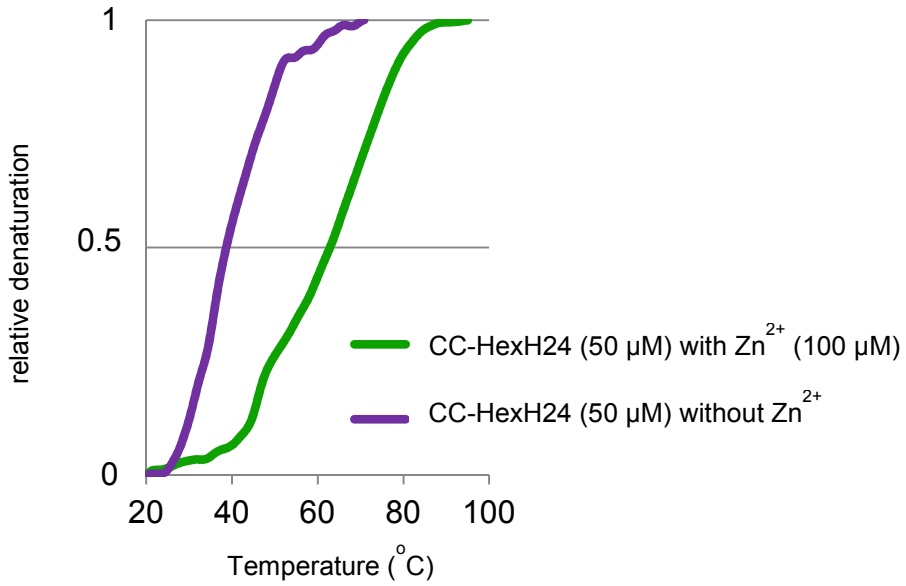



***Figure 6.25:*** Thermal denaturation of CC-HexH24 with and without $Zn^{2+}$ present at pH 7.5 as determined by CD-spectroscopy ($\theta_{222}$).

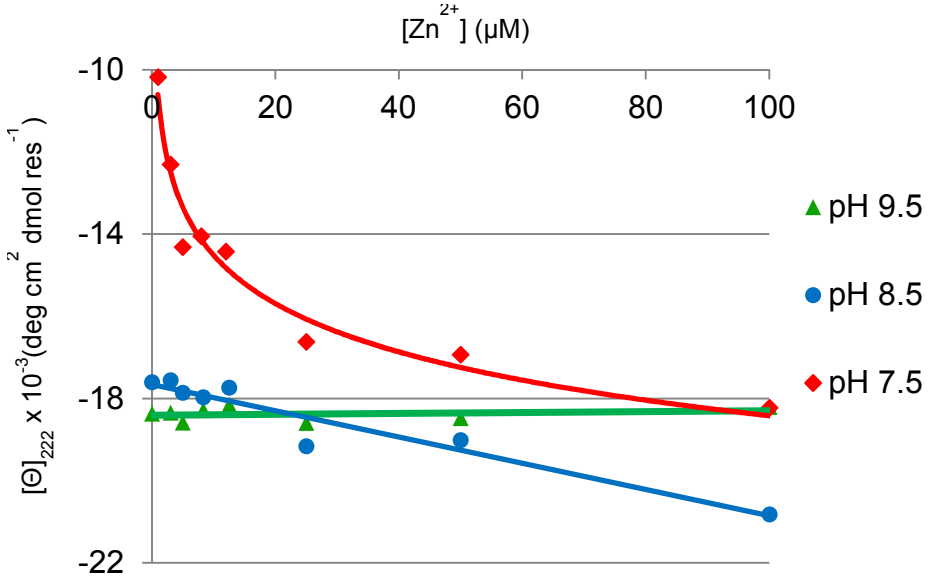



***Figure 6.26:*** *The effect of $Zn^{2+}$ concentration and pH on α-helicity of CC-HexH24 (50 μM) at 25°C. For pH 8.5 and 7.5, 50 mM Tris/HCl was used as buffer. For pH 9.5, 50 mM carbonate buffer was employed.*

maintained in the catalytically active metallo-form. As the mechanism for ester hydrolysis employs a zinc-bound-hydroxide nucleophile, the pKa of zinc-activated water was determined to be 9.4 according to catalytic activity as depicted in the sigmoidal plot of Figure 6.28. This is comparable to lowered water pKa values of 9.3 and 8.8 associated with Ac-IHIHIQI-$NH_2$ and $Hg(II)_SZn(II)_N(TRIL9CL23H)_3$, respectively.[27,188]

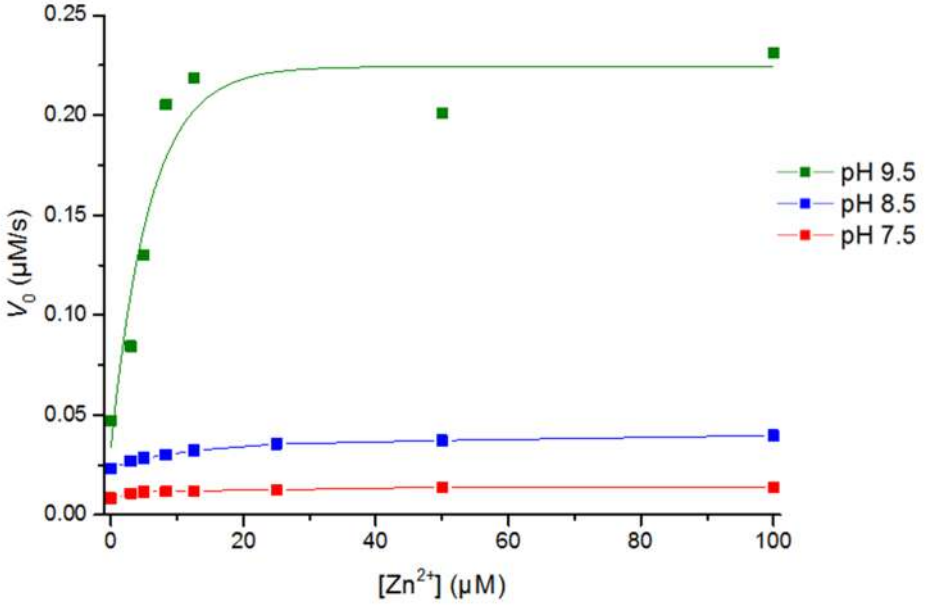



***Figure 6.27:*** *Effect of $Zn^{2+}$ concentration and pH on CC-Hex$H_{24}$-mediated (50 μM) ester (*p*NPA, 1 mM) hydrolysis (25°C). Reactions at pH 7.5 and 8.5 were performed in Tris/HCl (50 mM). Reactions at pH 9.5 were performed in 50 mM carbonate buffer.*

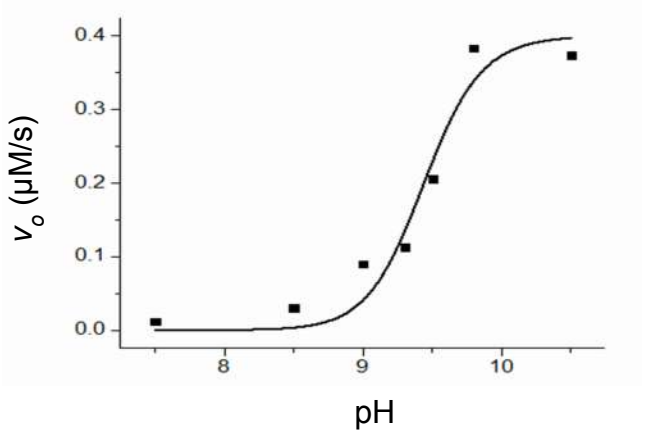



***Figure 6.28:*** *pH-dependency of* p*NPA (1 mM) hydrolysis at 25°C by CC-HexH24 (50 μM) in the presence of 8.3 μM $Zn^{2+}$. Reactions at pH 7.5 – 9 were performed in Tris/HCl (50 mM). Reactions above pH 9 were performed in 50 mM carbonate buffer.*

In order to show that the esterase activity of CC-HexH24 is due to the positioning of a functional histidine-coordinated zinc active site within a cleft, two control peptides, CC-Hex and CC-HexH24 scramble, were synthesized and subjected to kinetic and structural analysis (CD) at pH 9.5. CC-Hex (Table 6.4) is the model parallel hexameric coiled coil (Fig. 6.29) and lacks a histidine residue. The control peptide CC-HexH24 scramble consists of the same amino acids as CC-HexH24, but in a random sequence. In other words, it contains histidine, but is unlikely to fold into a

well-defined structure. As was observed in the case of CC-HexH24 (Fig. 6.26), the presence of zinc at pH 9.5 did not alter the α-helicity of either CC-Hex or CC-HexH24 scramble (Fig. 6.29). Figure 6.29 clearly shows that the peptide CC-Hex assumes a more α-helical structure than CC-HexH24. As reported by Zaccai *et al*.,[202] this is the result of a slight uncoiling (Fig. 6.30) observed at the carboxyl terminus of CC-HexH24 due to the incorporation of histidine. Furthermore, it has also been pointed out that unlike the model CC-Hex, CC-HexH24 does not form a single ideal hexamer as determined from analytical ultracentrifugation studies.[202] The presence of more than one oligomerization state is likely retained in the presence of zinc. Nonetheless, the CD spectra corresponding to CC-Hex and CC-HexH24 suggest the formation of α-helical coiled coils, whereas the spectrum for CC-HexH24 scramble corresponds, as expected, to a random coil.

To offer a clear view of activity, catalytically-cleaved *p*NPA, that is the concentration of *p*-nitrophenolate produced in the reaction with peptide present minus that caused by the random hydrolysis of an uncatalyzed reaction, was monitored during the course of the reaction (Fig. 6.31). As predicted, substrate turnover was noticeably lower (4-5x) in the presence of CC-Hex and CC-HexH24 scramble with and without zinc, as well as CC-HexH24 without zinc, than CC-HexH24 with zinc. This outcome clearly shows the catalytic advantage gained when catalytic machinery, particularly histidine and zinc, is placed within a relatively defined structure.

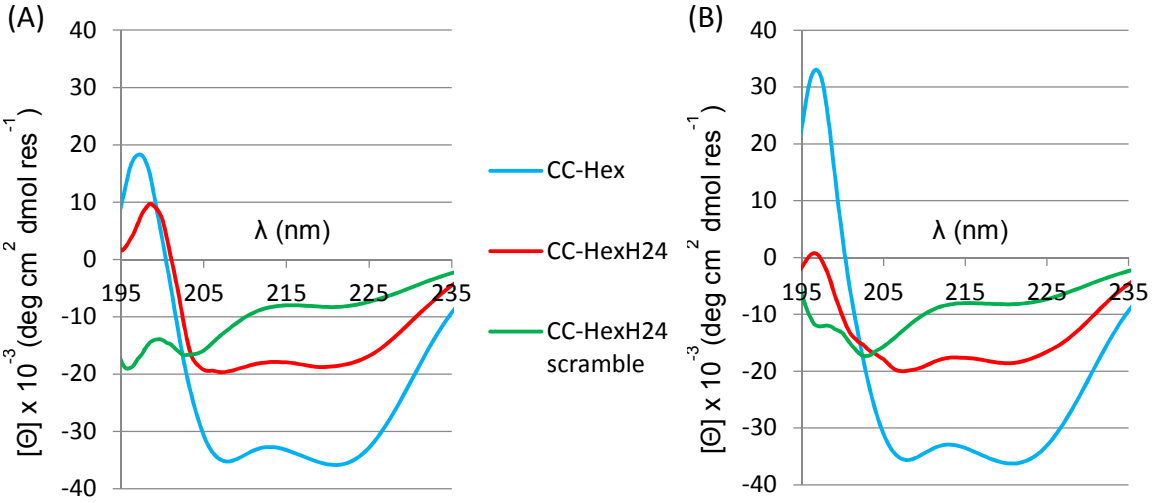



***Figure 6.29**: CD spectra of peptides CC-Hex, CC-HexH24 and CC-HexH24 scramble (50 μM) at pH 9.5 (50 mM $Na(CO_3)_2/NaHCO_3$ with 2% MeCN) and 25°C in the absence of $Zn^{2+}$ (A) and with $Zn^{2+}$ (8.3 μM) (B).*

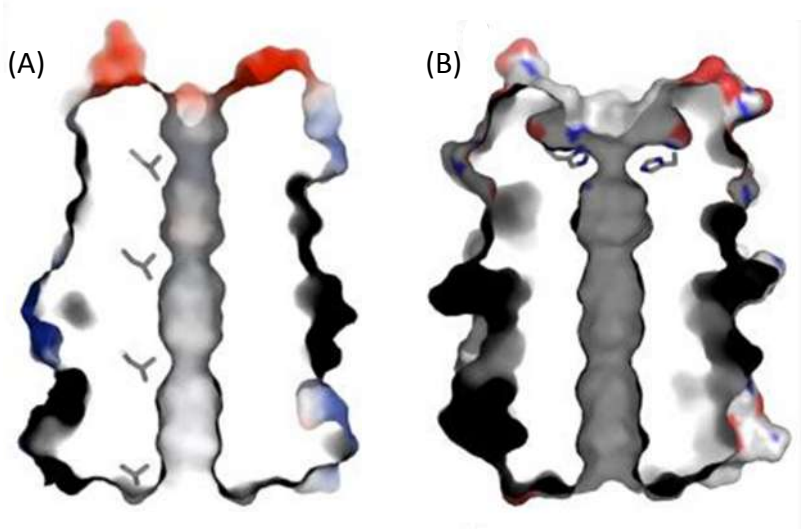



***Figure 6.30**: Cross sections of electrostatic surface of (A) CC-Hex with leucine rendered in heptad position a and (B) CC-HexH24 with histidine 24 occupying a heptad's a position. C-termini are shown at the top. Adapted from Zaccai* et al.[202]

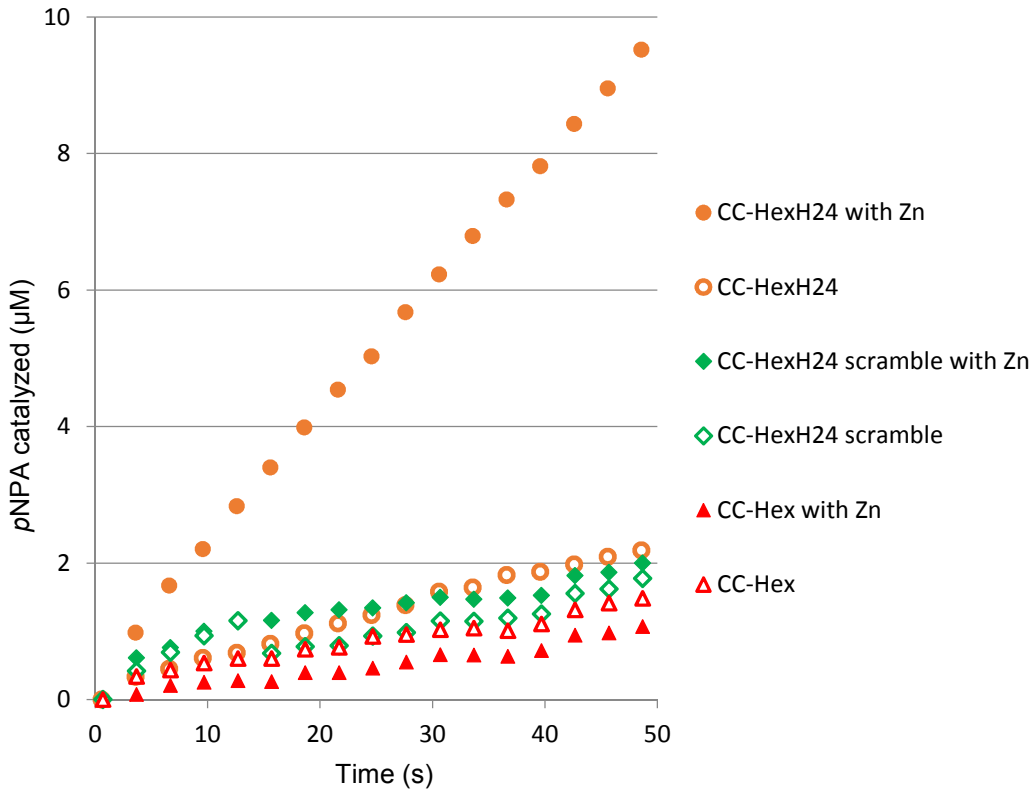



***Figure 6.31**: pNPA converted into p-nitrophenolate resulting from catalysis during the course of hydrolysis reactions performed in the presence of peptides CC-HexH24, CC-HexH24 scramble and CC-Hex (50 μM) with zinc chloride (8.3 μM) and without at pH 9.5 (50 mM $Na(CO_3)_2/NaHCO_3$ with 2% acetonitrile) and 25°C. Starting pNPA concentration was 1 mM.*

Michaelis-Menten enzyme kinetics was applied to determine the kinetic parameters of *p*NPA hydrolysis reactions catalyzed by CC-HexH24 in the presence of zinc. Since peptide CC-HexH24 does not assemble into a species possessing a single oligomerization state, quantification of its catalytic parameters becomes more complicated. However, by using CC-HexH24 without zinc as a blank to account for both random ester hydrolysis and hydrolysis attributed to histidine residues (general base or nucleophilic), catalyzed rate constants attributed only to the zinc-bound hydroxide mechanism can be calculated. As each catalyst using the zinc-bound hydroxide mechanism has one zinc atom, a limiting zinc concentration was used to obtain first-order ($k_{cat}$) and second-order ($k_{cat}/K_M$) rate constants.

In response to increased substrate concentrations, CC-HexH24 in the presence of zinc adhered to Michaelis-Menten kinetics (Fig. 2.7). The saturation of CC-HexH24 by substrate was observed as the measured values of initial velocity approached a maximum ($V_{MAX}$ = 0.55 μM/s) at elevated *p*NPA concentrations (Fig. 6.32). Double-reciprocal, Lineweaver-Burke plots were used to determine the Michaelis-Menten parameters $k_{cat}$, $K_M$ and $k_{cat}/K_M$, found to be 0.11 $s^{-1}$, 3.4 mM and 31.5 ± 0.9 $M^{-1}s^{-1}$, respectively. The second-order rate constant of 31.5 $M^{-1}s^{-1}$ at pH 9.5 compares favorably to the value reported by Zastrow *et al.* of 23.3 $M^{-1}s^{-1}$ corresponding to the trimeric coiled coil $Hg(II)_SZn(II)_N(TRIL9CL23H)_3$ (Fig. 4.10).[27]

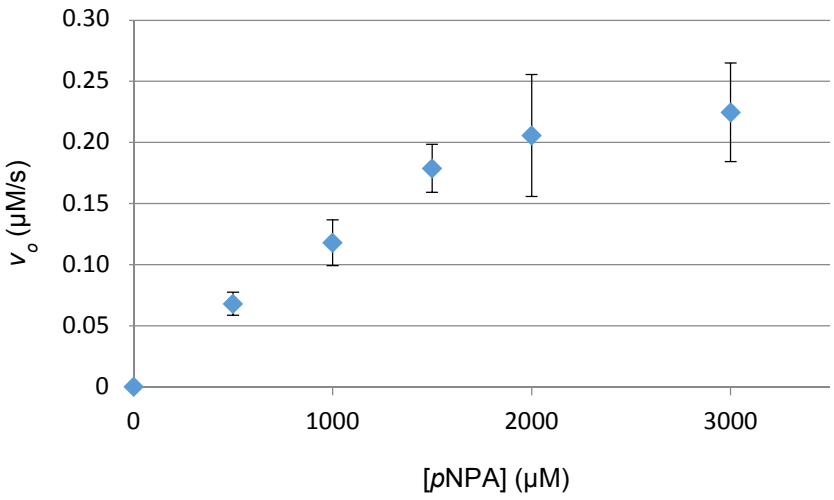



***Figure 6.32**: Esterase activity catalyzed by CC-HexH$_{24}$ (30 μM) in the presence of 5 μM $Zn^{2+}$ showing enzyme-like saturation at elevated substrate (pNPA) concentrations. Reactions were performed at pH 9.5 (50 mM $Na(CO_3)_2/NaHCO_3$ with 2% acetonitrile) and 25°C.*

The catalytic capacity of CC-HexH24 with zinc reported here is only the beginning. Research which addresses substrate selectivity using substrates of varying size, hydrophobicity and chirality remains to be performed. It would be especially interesting to observe the hydrolysis of the enantiomers of substrates such as Z-Leu-ONp or Z-Ile-ONp as they insert themselves within the substrate cleft rich with L-leucine and L-isoleucine residues. In addition, the catalyst CC-HexH24 has yet to be optimized. Mutations including those resulting from the repositioning of histidine along the inner channel of CC-HexH24 or from the substitution of other residues which interact with the active site would likely affect properties such as catalytic proficiency and substrate selectivity. Ultimately, the peptide CC-HexH24 as it functions with zinc represents a promising candidate for future design ideas in the progression toward more active and practical *de novo* protein catalysts.

# 7 Summary

The ability of enzymes to accelerate reactions while maintaining excellent substrate selectivity and product specificity has interested chemists for decades. The adaptation of enzymes to the reaction conditions of chemical synthesis has, however, been met with numerous obstacles such as reduced solubility, thermostability and substrate scope. To overcome these obstacles, protein engineering employs methods of rational design and directed evolution as well as more modern computational and combinatorial approaches. Although these techniques have been successfully applied to alter biocatalysts, our understanding of protein folding and the interactions required for enzyme catalysis remains limited. This is best evidenced by the inability to design *de novo* catalytic proteins comparable to enzymes.

To develop a better understanding of enzyme structure and function, the “bottom-up” approach to the *de novo* design of catalytic proteins is applied. This minimalistic approach uses simple peptides which predictably self-assemble into well-defined three-dimensional templates, on which catalytic machinery or other functionalities believed to impart a particular property are introduced. The simplified, yet protein-like, environments of the self-assembling peptides allow for enhanced clarity of the individual interactions to assess how such interactions contribute to particular enzymatic properties.

In the work presented here, the bottom-up technique was employed to introduce enzyme-like properties into three model peptide systems. Particular properties of interest include catalytic activity, substrate selectivity and affinity as well as the ability to regulate catalytic activity by altering conformation. As a model reaction, ester hydrolysis was selected due in large part to the ease by which such reactions can be kinetically monitored as well as the vast amount of studies and comparable data available in literature.

In an effort to design a catalytic peptide system in which the catalytic activity is regulated according to conformation, the heterodimeric coiled-coil model system E3/K3 was applied. The original design incorporated the catalytic triad (serine, histidine and aspartate) in proximal positions of the coiled coil to initiate a serine protease-like charge relay resulting in accelerated ester hydrolysis. Initial studies showed the moderate activity of the unfolded, carboxylate-rich E3CT peptide

containing the catalytic triad substitution. Almost no activity was observed in E3CT when folded into a coiled coil in the presence of K3. Further studies, particularly with the variant E3H11, in which only histidine is substituted, showed that the activity of the random coil is due to general base catalysis resulting from the interaction of histidine with glutamate. Such interactions are greatly reduced when peptides containing histidine are locked into a coiled coil. To develop the enzyme-like properties of E3H11, further efforts which include fixing it to a support such as hyper-branched polyglycerol (hPG) or gold monolayer-protected clusters (Au-MPC) may be taken.

In a separate study, the catalytically proficient, amyloid-forming peptide Ac-IHIHIQI-$NH_2$, which features the zinc-bound hydroxide mechanism (ZHN) of hydrolysis, was examined for properties of substrate selectivity. These studies revealed that Ac-IHIHIQI-$NH_2$ was especially selective for the hydrophobic substrate Z-L-Phe-ONp over the more hydrophilic Boc-L-Asn-ONp while in the process of fibril assembly. Furthermore, Ac-IHIHIQI-$NH_2$ showed increased enantioselectivity for the L-enantiomer of Z-Phe-ONp in the process of self-assembly as opposed to when it is in the form of a fully-established fibril. Such findings are likely the result of increased hydrophobic interactions occurring between hydrophobic moieties of the substrate and hydrophobic surfaces of the peptide, which experience greater exposure while in the process of fibril assembly. Future efforts to stabilize this β-sheet active site within a larger protein structure would provide interesting results.

In the final study, the ZHN mechanism applied in Ac-IHIHIQI-$NH_2$ was incorporated into the hexameric α-helical coiled coil CC-HexH24, which features six adjacent histidine residues in its hydrophobic interior. The coordination of zinc within a water-accessible, substrate accommodating cleft was found to greatly accelerate ester hydrolysis. However, further structural studies including crystallography should be pursued to obtain a detailed image of the active-site. In addition, substrate selectivity studies involving substrates varying in hydrophobicity and size would also contribute a wealth of information in regard to active-site design.

These studies represent a growing understanding in protein structure and function through the use of peptide design. As our knowledge of such molecules and their interactions develop so will the complexity of catalytic proteins designed *de novo* from the bottom up.

# 8

# Materials and methods

## 8.1 Peptide synthesis and verification

Peptides were synthesized using the Fmoc (fluorenylmethoxycarbonyl) method of solid phase peptide synthesis (SPPS),[203] purified via reverse-phase high-pressure liquid chromatography (RP-HPLC) and characterized with analytical HPLC and mass spectrometry. A list of peptides discussed in this work along with corresponding sequences and mass analyses is shown in Table 8.2.

### 8.1.1 Fmoc solid phase peptide synthesis

In solid phase peptide synthesis, amino acids are coupled to a growing peptide chain attached to a solid support, typically an organic polymer. Following each reaction, the peptide linked to the insoluble support can be easily purified from soluble reagents by filtration and repeated washing. Not only does this save time and labor, but allows the use of excess reagents, which translates into both increased peptide yield and purity.

Since the carboxyl group of first amino acid in linked to the support, synthesis of peptides on solid phase starts from the carboxyl terminus and proceeds in the direction of the amino terminus. The reaction scheme of peptide synthesis involves repetitive deprotection and coupling reactions (Fig. 8.1) which are separated by thorough washing.

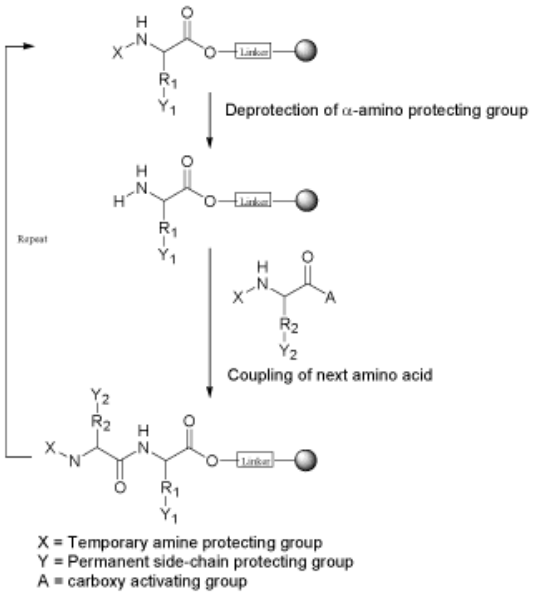



***Figure 8.1:*** *Reaction scheme of solid phase peptide synthesis. Courtesy of Chan & White.*[203]

Each of the peptides employed in this work were synthesized using the Fmoc or Fmoc/*t*Bu method of solid phase peptide synthesis. This strategy uses 9-fluorenylmethoxycarbonyl (Fmoc) as the temporary amine protecting group (X in Figure 8.1). Deprotection of Fmoc was carried out under basic conditions. Coupling reactions are performed upon the formation of active esters (Fig. 8.1 A) of the next amino acid. Once the final amino acid of the sequence has been coupled, the permanent side-chain protecting groups (Y in Fig. 8.1) are removed and the peptide is simultaneously cleaved from the solid support under acidic conditions.

#### 8.1.1.1 Materials for Fmoc SPPS

**Fmoc- and side-chain-protected amino acids:** Fmoc-L-Ala-OH·$H_2O$, Fmoc-L-Asp(O*t*Bu)-OH·$H_2O$, Fmoc-L-Glu(O*t*Bu)-OH·$H_2O$, Fmoc-L-Gln(Trt)-OH, Fmoc-L-Gly-OH, Fmoc-L-His(Trt)-OH, Fmoc-L-Ile-OH, Fmoc-L-Leu-OH, Fmoc-L-Lys(Boc)-OH, Fmoc-L-Ser(O*t*Bu)-OH and Fmoc-Trp(Boc)-OH were purchased from Orpegen Pharma.

**Solid support resins:** Fmoc-Glu(O*t*Bu)-NovaSyn® TGA was used in the syntheses of K3 variants and peptide EEEHEEE. Fmoc-Lys(Boc)-NovaSyn® TGA was employed in the syntheses of all E3 variants. Fmoc-Gly-NovaSyn® TGT was used in each of the CC-Hex variants and Rink Amide MBHA resin (100-200 mesh) for Ac-IHIHIQI-$NH_2$. Each of the resins was obtained from Novabiochem.

**Solvents:** Dimethylformamide (DMF, 99.9%; VWR), dichloromethane (DCM, ≥99%; Fischer) and *N*-methyl-2-pyrrolidone (NMP, 99%; ACROS) were used without further purification.

**Reagents for Fmoc deprotection:** Piperidine (99%; ACROS) and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU, ≥ 98%; Merck).

**Coupling reagents:** *N,N'*-Diisopropylcarbodiimide (DIC, 99%; Sigma Aldridge), 1-hydroxy-7-azabenzotriazole (HOAt, 99%; Carbolution), *N*-hydroxybenzotriazole (HOBt, 99%; Carbolution), *O*-(6-chlorobenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HCTU, 99%; Activotec), 2-(1*H*-benzotriazol-1-yl)-1,1,3,3,-tetramethyluronium tetrafluoroborate (TBTU, 99%;

Carbolution) and *N,N*-diisopropylethylamine (DIPEA, ≥ 98%; ACROS) were used as purchased.

**Reagents for resin cleavage and side-chain deprotection:** Trifluoroacetic acid (TFA, 99%; Merck) and triisopropylsilane (TIPS, 98%; Alfa Aesar) were used without further purification.

**HPLC Eluents:** Acetonitrile (MeCN, HPLC grade; VWR), trifluoroacetic acid (TFA, HPLC grade; Merck). Water was obtained from a Milli-Q® Advantage A10 purification system.

**Additional materials:** Boc-Abz-OH (Bachem), acetic anhydride (Fischer, distilled prior to use), 4-pentynoic acid (98%; Alfa Aesar), $CuSO_4 \cdot 5\ H_2O$ (≥ 98%; Sigma-Aldrich), sodium ascorbate (≥ 98%; Sigma-Aldrich), dialysis tubes (MWCO = 3.5-5 kD; Spectrum Labs®), diethyl ether (VWR, 99%). Water was obtained from a Milli-Q® Advantage A10 purification system.

#### 8.1.1.2 Manual Fmoc SPPS protocol

Resin (preloaded or not) was added to a polypropylene syringe reaction vessel furnished with a polytetrafluoroethylene (PTFE) frit and allowed to swell 20 minutes in DMF. As is always the case before proceeding to the next step, the solution or solvent was pushed through the frit. Fmoc was removed from the preloaded resin by shaking in a solution of 2% piperidine, 2% DBU in DMF for 10 min (3x). After sufficient washing with DMF, coupling was performed by one of the two methods shown below:

(1) DIC/HOBt or HOAt: Fmoc-amino acid (5 equiv.) and HOBt or HOAt (5 equiv.) are dissolved in approximately 6 mL DMF. Upon addition of 4.9 equiv. DIC, a 10 min. carboxylic acid activation period was allowed prior to being added to the resin. The vessel was agitated for at least 40 min.

(2) HCTU/HOBt/DIPEA: Fmoc-amino acid (5 equiv.), HOBt (5 equiv.) and HCTU (4.9 equiv.) were dissolved in approximately 6 mL DMF. Upon treatment with DIPEA (10 equiv.), the solution was added immediately to the resin. The vessel was agitated for at least 40 min.

To ensure completion, coupling reactions were performed twice. After coupling, the resin was again washed with DMF and the Fmoc group from the newly coupled amino acid was removed. Repeated Fmoc-removal, washing, coupling and washing cycles were carried out until the last amino acid of the sequence was added.

#### 8.1.1.3 Automated Fmoc SPPS protocol

Automated peptide syntheses of E3 and K3 variant peptides were performed by *SyroXP* manufactured by Multi-SynTec GmbH. The process of the automated cycle was very similar to that of the manual process (Fmoc-deprotection, washing, coupling, washing). However, the *SyroXP* is best suited for reaction scales of 0.5 µM performed in 10 mL PTFE fritted, polypropylene vessels. The reaction cycle of *SyroXP* including the reagents used and the time of each step is summarized in Table 8.1.

***Table 8.1:*** *Fmoc-deprotection and coupling procedure employed in a cycle by the fully-automated* SyroXP *by Multi-SynTech GmbH.*

| Process | Reagent | Time |
|---|---|---|
| **Fmoc-deprotection** | • 2 mL 20 % piperidine in DMF<br>• 2 mL 40 % piperidine in DMF | 2 x 5 min<br>2 x 5 min |
| **Washing** | 2.5 mL DMF | 6 x 1 min |
| **Coupling** | • 4 equiv. Fmoc-amino acid and 4 equiv. HOBt in DMF<br>• 4 equiv. TBTU in DMF<br>• 8 equiv. DIPEA in NMP | 30 min |
| **Washing** | 2.5 mL DMF | 1 x 1 min |
| **Coupling** | • 4 equiv. Fmoc-amino acid and 4 equiv. HOBt in DMF<br>• 4 equiv. TBTU in DMF<br>• 8 equiv. DIPEA in NMP | 30 min |
| **Washing** | 2.5 mL DMF | 6 x 1 min |

Syntheses of peptides CC-Hex and CC-HexH24 scramble were performed using the fully-automated peptide synthesizer *Activo-P11* manufactured by *ActivoTec* which is best suited for performing reaction up to 0.1 mmol scale. The reaction cycle of employed with *Activo-P11* is summarized in Table 8.2.

***Table 8.2:*** *Fmoc-deprotection and coupling procedure employed in a cycle by the fully-automated* Activo-P11 *by* ActivoTec*.*

| Process | Reagents | Time |
|---|---|---|
| Fmoc-deprotection | • 2 mL 2 % PIP / 2 % DBU in DMF | 2 x 5 min |
| | • 2 mL 2 % PIP / 2 % DBU in DMF | 2 x 10 min |
| Washing | 2.5 mL DMF | 5 x 1 min |
| Coupling | 8 eq. Fmoc-Amino acid-OH<br>8 eq. HOBt in DMF<br>8 eq. DIC in DMF | 60 min |
| Washing | 2.5 mL DMF | 5 x 1 min |
| Coupling | 8 eq. Fmoc-Amino acid-OH<br>8 eq. HOAt in DMF<br>8 eq. DIC in DMF | 60 min |
| Washing | 2.5 mL DMF | 5 x 1 min |

#### 8.1.1.4 N-terminal acetylation

Capping or acetylation is a common method employed to block the reactivity of the amino-terminus. Acetylation in the preparation of Ac-IHIHIQI-$NH_2$ was performed following the Fmoc-deprotection of the last isoleucine residue coupled. Thus, resin-attached peptide was treated with a solution of 10% freshly distilled acetic anhydride and 10% DIPEA in DMF twice for 10 min.

#### 8.1.1.5 N-terminal alkyne functionalization

In order to perform copper-catalyzed alkyne-azide cycloaddition, peptides E3, E3H11 and E3H11-Ala were each coupled to 4-pentynoic acid. Coupling of pentynoic acid took place at the N-terminus of the peptide following Fmoc cleavage of the last residue while the peptide was on solid support with side-chain protecting groups intact. 4-Pentynoic acid (5 equiv.), HOBt (5 equiv.) and HCTU (4.9 equiv.) were

dissolved in approximately 6 mL DMF. Upon treatment with DIPEA (10 equiv.), the solution was drawn immediately to the resin within the PTFE reaction vessel. The vessel was agitated for at least 40 min. This procedure was performed twice to ensure complete coupling. Following sufficient washing with DMF, the alkyne functionalized E3yne, E3H11yne and E3H11-Alayne were cleaved from the resin.

#### 8.1.1.6 Peptide cleavage from resin and trituration

Prior to cleavage and the deprotection of the side chains, the resin was washed with DCM in order to remove the less volatile DMF.[204] For each peptide, cleavage was carried out by treating the resin with a cocktail containing 95% TFA, 3% TIS and 2% $H_20$ (1 mL cocktail per 100 mg resin) for 4 h. The cocktail containing the cleaved peptide was then collected in a round-bottom flask and concentrated under a stream of argon. Crude peptide was collected via trituration in ice-cold diethyl ether and centrifugation. The resulting pellet containing the crude peptide was resuspended and centrifuged several times with cold ether. After decanting the ether from the final washing, the remaining ether was removed from the pellet *in vacuo* (400 mbar).

### 8.1.2 Peptide analysis and purification via HPLC

Crude peptide was dissolved in aqueous MeCN with 0.1% TFA and purified by RP-HPLC using a high-pressure gradient system supplied by Knauer. The system was outfitted with a Smartline Manager 5000 user interface with solvent degasser, two Smartline 1000 eluent pumps, an adjustable UV-detector 2500, eluent gradient mixer, and a 5 mL sample injection loop. As eluent, gradients of water and MeCN each with 0.1% TFA were used at a flowrate of 20 mL/min. Chromatography of peptide variants E3 and K3 were performed using the column Phenomenex® Gemini-NX C18 (10 µM; 250 x 21.20 mm) . Peptide variants of CC-Hex were purified with the column Phenomenex® Luna C8 (10 µM; 250 x 21.20 mm) and peptide Ac-IHIHIQI-$NH_2$ with Phenomenex® Jupiter C4 (10 µM; 250 x 21.20 mm). Fractions were collected corresponding to UV-absorption recorded with a Perkin Elmer 56 analog blotter. Collected fractions were then characterized by analytical HPLC and mass spectrometry.

Peptide purity was assessed using the analytical system Elite LaChrome from VWR-Hitachi or the system Chromaster from VWR-Hitachi. Elite LaChrome was furnished with two L-2130 pumps, diode array detector L-2485, autosampler L-2200, solvent degasser L7612 and 100 µL injection loop. Chromatography proceeded with a flowrate of 1 mL/min through the column Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm). The system Chromaster was outfitted with a single 5160 pump, diode array detector 5430, autosampler 5260, column oven 5310, solvent degasser and 20 µL injection loop. Chromatography was carried out with a flowrate of 0.6 mL/min through the column Purospher® STAR C18 (2 µM; 50 x 2.1 mm) from Merck. The software interface employed by both analytical systems was EZChrom Elite developed by Agilent.

Following purification, all peptides were lyophilized with a Christ Alpha 1-2 LD freeze dryer. All peptides were stored at -20°C when not in use.

### 8.1.3 Synthesis and purification details of individual peptides

**E3: $H_2N$-Abz-EIAALEKEIAALEKEIAALEK-OH**

The peptide E3 was synthesized on Fmoc-Lys(Boc)-NovaSyn® TGA resin (0.20 mmol/g loading). Peptide E3 was synthesized on separate occasions, once employing a 0.1 mmol scale using the manual coupling[a] method (8.1.1.2 with DIC and HOBt/HOAt) and twice using the automated *SyroXP* procedure[2] (Table 8.1) with a reaction scale of 0.05 mmol (8.1.1.3). When employing the automated procedure, the final residue, Boc-protected 2-aminobenzoic acid, was manually coupled using the standard procedure. The peptide was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC analysis was performed as described in Section 8.1.2 using the column Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm) with an eluent gradient which starts at 5% MeCN in water and increases to 70% over 30 minutes. Using this method, the retention time of E3 was 26.0 minutes. The same gradient was employed to purify E3 via preparative HPLC (8.1.2) using a Phenomenex® Gemini-NX C18 column (10 µM; 250 x 21.20 mm). The retention time for purification was 24.1 minutes. Individual fractions collected from

[1] Performed by Sabrina Lehmann under the author's supervision
[2] Performed by Jennifer Hoffmann under the author's supervision

preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10). Following HPLC and removal of solvent, yields of 86.9 mg resulting from manual synthesis as well as 18.3 mg and 24.6 mg from automated synthesis were obtained.

### K3: $H_2N$-Abz-KIAALKEKIAALKEKIAALKE-OH

The peptide K3 was synthesized on Fmoc-Glu(*t*Bu)-NovaSyn® TGA resin (0.24 mmol/g loading). K3 was synthesized on separate occasions, once using a 0.1 mmol scale using the manual coupling method[3] (8.1.1.2 with DIC and HOBt/HOAt) and twice using the automated *SyroXP* procedure[4] (Table 8.1) with a reaction scale of 0.05 mmol (8.1.1.3). When employing the automated procedure, the final residue, Boc-protected 2-aminobenzoic acid, was manually coupled using the standard protocol. The peptide was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC analysis was performed as described in Section 8.1.2 using the column Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm) with an eluent gradient which starts at 5% MeCN in water and increases to 70% over 30 minutes. Using this method the retention time of K3 was 19.8 minutes. The same gradient was employed to purify K3 via preparative HPLC (8.1.2) using a Phenomenex® Gemini-NX C18 column (10 µM; 250 x 21.20 mm). The retention time for purification was 18.6 minutes. Individual fractions collected from preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10). Following HPLC and removal of solvent, yields of 53.3 mg via manual synthesis as well as 45.4 mg and 32.1 mg by automated synthesis were obtained.

### E3CT: $H_2N$-Abz-EIAALEDEIAHLEKSIAALEK-OH

The peptide E3CT was synthesized on Fmoc-Lys(Boc)-NovaSyn® TGA resin (0.20 mmol/g loading). E3CT was synthesized on separate occasions, once employing a 0.1 mmol scale using the manual coupling method[b] (8.1.1.2 with DIC and HOBt/HOAt) and twice using the automated *SyroXP* procedure (Table 8.1) with a reaction scale of 0.05 mmol (8.1.1.3). When employing the automated procedure, the

[3] Performed by Sabrina Lehmann under the author's supervision
[4] Performed by Jennifer Hoffmann under the author's supervision

final residue, Boc-protected 2-aminobenzoic acid, was manually coupled using the standard procedure. The peptide was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC analysis was performed as described in Section 8.1.2 using the column Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm) with an eluent gradient which starts at 5% MeCN in water and increases to 70% over 30 minutes. Using this method the retention time of E3CT was 25.5 minutes. The same gradient was employed to purify E3CT via preparative HPLC (8.1.2) using the column Phenomenex® Gemini-NX C18 (10 µM; 250 x 21.20 mm). The retention time for purification was 23.9 minutes. Individual fractions collected from preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10). Following HPLC and removal of solvent, yields of 78.0 mg resulting from manual synthesis and 16.2 mg from automated synthesis were obtained.

**K3CT: $H_2N$-Abz-KIAALKDKIAHLKESIAALKE-OH**

The peptide K3CT was synthesized on Fmoc-Glu(*t*Bu)-NovaSyn® TGA resin (0.24 mmol/g loading). K3CT was prepared on separate occasions, once employing a 0.1 mmol scale using the manual coupling method[5] (8.1.1.2 with DIC and HOBt/HOAt) and once using the automated *SyroXP* procedure (Table 8.1) with a reaction scale of 0.05 mmol (8.1.1.3). When employing the automated procedure, the final residue, Boc-protected 2-aminobenzoic acid, was manually coupled using the standard procedure. The peptide was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC analysis was performed as described in Section 8.1.2 using the column Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm) with an eluent gradient which starts at 5% MeCN in water and increases to 70% over 30 minutes. Using this method the retention time of K3CT was 19.3 minutes. The same gradient was employed to purify K3CT via preparative HPLC (8.1.2) using the column Phenomenex® Gemini-NX C18 (10 µM; 250 x 21.20 mm). The retention time for purification was 18.2 minutes. Individual fractions collected from preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10).

[5] Performed by Sabrina Lehmann under the author's supervision

Following HPLC and removal of solvent, yields of 19.8 mg by manual synthesis and 46.7 mg resulting from automated synthesis were obtained.

### E3H11: $H_2N$-Abz-EIAALEKEIAHLEKEIAALEK-OH

The peptide E3H11 was synthesized on Fmoc-Lys(Boc)-NovaSyn® TGA resin (0.20 mmol/g loading). E3H11 was synthesized on separate occasions, once employing a 0.1 mmol scale using the manual coupling method (8.1.1.2 with DIC and HOBt/HOAt) and twice using the automated *SyroXP* procedure[6] (Table 8.1) with a reaction scale of 0.05 mmol (8.1.1.3). When employing the automated procedure, the final residue, Boc-protected 2-aminobenzoic acid, was manually coupled using the standard procedure. The peptide was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC analysis was performed as described in Section 8.1.2 using the column Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm) with an eluent gradient which starts at 5% MeCN in water and increases to 70% over 30 minutes. Using this method the retention time of E3H11 was 25.1 minutes. The same gradient was employed to purify E3H11 via preparative HPLC (8.1.2) using the column Phenomenex® Gemini-NX C18 (10 µM; 250 x 21.20 mm). The retention time for purification was 23.5 minutes. Individual fractions collected from preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10). Following HPLC and the removal of solvent, yields of 86.3 mg by manual synthesis as well as 12.7 mg and 15.2 mg by automated synthesis were obtained.

### E3S15: $H_2N$-Abz-EIAALEKEIAALEKSIAALEK-OH

The peptide E3S15 was synthesized on Fmoc-Lys(Boc)-NovaSyn® TGA resin (0.20 mmol/g loading) using the automated *SyroXP* procedure[f] (Table 8.1) with a reaction scale of 0.05 mmol (8.1.1.3). When employing the automated procedure, the final residue, Boc-protected 2-aminobenzoic acid, was manually coupled using the standard procedure. The peptide was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC analysis was performed as described in Section 8.1.2 using the column Phenomenex® Kinetix C18 (5 µM; 250 x

[6] Performed by Jennifer Hoffmann under the author's supervision

4.6 mm) with an eluent gradient which starts at 5% MeCN in water and increases to 70% over 30 minutes. Using this method the retention time of E3S15 was 25.9 minutes. The same gradient was employed to purify E3S15 via preparative HPLC (8.1.2) using the column Phenomenex® Gemini-NX C18 (10 µM; 250 x 21.20 mm). The retention time for purification was 24.0 minutes. Individual fractions collected from preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10). Following HPLC and the removal of solvent, a yield of 22.2 mg was obtained.

**E3H11-Ala: $H_2N$-Abz-KAEAAAAEAEAAHAEAAAAEAK-OH**

The peptide E3H11-Ala was synthesized on Fmoc-Lys(Boc)-NovaSyn® TGA resin (0.20 mmol/g loading) using the automated *SyroXP* procedure (Table 8.1) with a reaction scale of 0.05 mmol (8.1.1.3). When employing the automated procedure, the final residue, Boc-protected 2-aminobenzoic acid, was manually coupled using the standard procedure. The peptide was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC analysis was performed as described in Section 8.1.2 using the column Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm) with an eluent gradient which starts at 5% MeCN in water and increases to 70% over 18 minutes. Using this method the retention time of E3H11-Ala was 12.8 minutes. A gradient of 5 to 70% acetonitrile was employed over 30 minutes to purify E3H11-Ala via preparative HPLC (8.1.2) using the column Phenomenex® Gemini-NX C18 (10 µM; 250 x 21.20 mm). The retention time for purification was 22.5 minutes. Individual fractions collected from preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10). Following HPLC and solvent removal, a yield of 15.2 mg was obtained.

**EEEHEEE: $H_2N$-Abz-EEEHEEE-OH**

The peptide EEEHEEE was synthesized on Fmoc-Glu(*t*Bu)-NovaSyn® TGA resin (0.24 mmol/g loading) using the manual coupling procedure at a reaction scale of 0.1 mmol (8.1.1.2). The peptide was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC analysis was performed as described in Section 8.1.2 using the column Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm) with

an eluent gradient which starts at 0% MeCN in water and increases to 45% over 22 minutes. Using this method the retention time of EEEHEEE was 13.8 minutes. A gradient of 0 to 45% acetonitrile was employed over 30 minutes to purify EEEHEEE via preparative HPLC (8.1.2) using the column Phenomenex® Gemini-NX C18 (10 µM; 250 x 21.20 mm). The retention time for purification was 21.0 minutes. Individual fractions collected from preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10). Following HPLC and the removal of solvent, a yield of 48.4 mg was obtained.

**E3yne: Alkyne functionalized E3**

The peptide E3yne was synthesized on Fmoc-Lys(Boc)-NovaSyn® TGA resin (0.20 mmol/g loading) using the manual coupling procedure at a reaction scale of 0.2 mmol (8.1.1.2). Following the Fmoc deprotection of the N-terminus, the coupling of 4-pentynoic acid was carried out as described in Section 8.1.1.5. The peptide was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC analysis was performed as described in Section 8.1.2 using the column Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm) with an eluent gradient which starts at 5% MeCN in water and increases to 100% over 18 minutes. Using this method the retention time of E3yne was 14.4 minutes. A gradient of 15 to 100% acetonitrile was employed over 40 minutes to purify E3yne via preparative HPLC (8.1.2) using the column Phenomenex® Gemini-NX C18 (10 µM; 250 x 21.20 mm). The retention time for purification was 33.5 minutes. Individual fractions collected from preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10). Following HPLC and the removal of solvent, a yield of 127.4 mg was obtained.

**E3H11yne: Alkyne functionalized E3H11**

The peptide E3H11yne was synthesized on Fmoc-Lys(Boc)-NovaSyn® TGA resin (0.20 mmol/g loading) using the manual coupling procedure at a reaction scale of 0.2 mmol (8.1.1.2). Following the Fmoc deprotection of the N-terminus, the coupling of 4-pentynoic acid was carried out as described in Section 8.1.1.5. The peptide was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC analysis was performed as described in Section 8.1.2 using the column

Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm) with an eluent gradient which starts at 5% MeCN in water and increases to 100% over 18 minutes. Using this method the retention time of E3H11yne was 13.8 minutes. A gradient of 15 to 100% acetonitrile was employed over 40 minutes to purify E3H11yne via preparative HPLC (8.1.2) using the column Phenomenex® Gemini-NX C18 (10 µM; 250 x 21.20 mm). The retention time for purification was 32.0 minutes. Individual fractions collected from preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10). Following HPLC and solvent removal, a yield of 147.0 mg was obtained.

**E3H11-Alayne: Alkyne functionalized E3H11-Ala**

The peptide E3H11-Alayne was synthesized on Fmoc-Lys(Boc)-NovaSyn® TGA resin (0.20 mmol/g loading) using the manual coupling procedure at a reaction scale of 0.2 mmol (8.1.1.2). Following the Fmoc deprotection of the N-terminus, the coupling of 4-pentynoic acid was carried out as described in Section 8.1.1.5. The peptide was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC analysis was performed as described in Section 8.1.2 using the column Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm) with an eluent gradient which starts at 5% MeCN in water and increases to 100% over 18 minutes. Using this method the retention time of E3H11-Alayne was 10.7 minutes. A gradient of 15 to 100% acetonitrile was employed over 40 minutes to purify E3H11-Alayne via preparative HPLC (8.1.2) using the column Phenomenex® Gemini-NX C18 (10 µM; 250 x 21.20 mm). The retention time for purification was 27.5 minutes. Individual fractions collected from preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10). Following HPLC and solvent removal, a yield of 122.5 mg was obtained.

**Ac-IHIHIQI-$NH_2$**

The peptide Ac-IHIHIQI-$NH_2$ was synthesized on Rink Amide MBHA resin (100-200 mesh) resin (0.64 mmol/g loading) using the manual coupling procedure at a reaction scale of 0.2 mmol (8.1.1.2). Following the final Fmoc deprotection, the N-terminus underwent acetylation as mentioned in Section 8.1.1.4. The peptide was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC

analysis was performed as described in Section 8.1.2 using the column Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm) with an eluent gradient which starts at 10% MeCN in water and increases to 100% over 18 minutes. Using this method the retention time of Ac-IHIHIQI-$NH_2$ was 10.6 minutes. A gradient of 30 to 100% acetonitrile was employed over 30 minutes to purify Ac-IHIHIQI-$NH_2$ via preparative HPLC (8.1.2) using the column Phenomenex® Jupiter C4 (10 µM; 250 x 21.20 mm). The retention time for purification was 14.0 minutes. Individual fractions collected from preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10). Following HPLC and solvent removal, a yield of 95.2 mg was obtained.

**CC-HexH24: $H_2$N-GELKAIAQELKAIAKELKAIAWEHKAIAQGAG-OH**

The peptide CC-HexH24 was synthesized on Fmoc-Gly-NovaSyn® TGT resin (0.63 mmol/g loading) using the manual coupling procedure at a reaction scale of 0.2 mmol (8.1.1.2). The peptide was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC analysis was performed as described in Section 8.1.2 using the column Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm) with an eluent gradient which starts at 5% MeCN in water and increases to 95% over 18 minutes. Using this method the retention time of CC-HexH24 was 13.3 minutes. A gradient of 15 to 95% acetonitrile was employed over 30 minutes to purify CC-HexH24 via preparative HPLC (8.1.2) using the column Phenomenex® Luna C8 (10 µM; 250 x 21.20 mm). The retention time for purification was 22.5 minutes. Individual fractions collected from preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10). Following HPLC and solvent removal, a yield of 156.5 mg was obtained.

**CC-HexH24 scramble: $H_2$N-EEKAHKLGAKIGLIQEAIAAKAWEAALQAIKG-OH**

The peptide CC-HexH24 scramble was synthesized with *Activo-P11* on Fmoc-Gly-NovaSyn® TGT resin (0.63 mmol/g loading) using the automated ActivoTec procedure[7] (Table 8.2) employing a reaction scale of 0.1 mmol (8.1.1.3). The peptide

[7] Performed by Kumar Chaitanya Thota under the supervision of the author.

was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC analysis was performed as described in Section 8.1.2 using the column Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm) with an eluent gradient which starts at 5% MeCN in water and increases to 95% over 18 minutes. Using this method the retention time of CC-HexH24 scramble was 13.4 minutes. A gradient of 15 to 95% acetonitrile was employed over 30 minutes to purify CC-HexH24 scramble via preparative HPLC (8.1.2) using the column Phenomenex® Luna C8 (10 µM; 250 x 21.20 mm). The retention time for purification was 22.5 minutes. Individual fractions collected from preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10). Following HPLC and solvent removal, a yield of 70.1 mg was obtained.

**CC-Hex: $H_2$N-GELKAIAQELKAIAKELKAIAWELKAIAQGAG-OH**

The peptide CC-Hex was synthesized with *Activo-P11* on Fmoc-Gly-NovaSyn® TGT resin (0.63 mmol/g loading) using the automated ActivoTec procedure (Table 8.2) employing a reaction scale of 0.1 mmol (8.1.1.3). The peptide was cleaved from the resin, deprotected and triturated as described in Section 8.1.1.6. HPLC analysis was performed as described in Section 8.1.2 using the column Phenomenex® Kinetix C18 (5 µM; 250 x 4.6 mm) with an eluent gradient which starts at 5% MeCN in water and increases to 95% over 18 minutes. Using this method the retention time of CC-Hex was 14.1 minutes. A gradient of 15 to 95% acetonitrile was employed over 30 minutes to purify CC-Hex via preparative HPLC (8.1.2) using the column Phenomenex® Luna C8 (10 µM; 250 x 21.20 mm). The retention time for purification was 23.5 minutes. Individual fractions collected from preparative HPLC were further assessed for purity using analytical HPLC (Appendix: Section 10). Following HPLC and solvent removal, a yield of 62.8 mg was obtained.

### 8.1.4 Peptide characterization via mass spectrometry

Peptides were verified with ESI-ToF mass spectrometry using the spectrometer Agilent 6210 supplied by Agilent Technologies. Mass data (Table 8.3) was acquired and analyzed with the software Masshunter developed by Agilent. Mass spectra of the individual peptides are shown in the Appendix.

***Table 8.3:*** *Mass to charge ratios of synthetic peptides measured by ESI-ToF mass spectrometry and corresponding calculated ratios. The residue 1 represents the amide linkage of peptide to 4-pentynoic acid.*

| Peptide | Sequence | | Observed | Calculated |
|---|---|---|---|---|
| **E3** | $H_2N$-Abz-EIAALEKEIAALEKEIAALEK-OH | $[M+3]^{3+}$ | 801.4533 | 801.4484 |
| **E3CT** | $H_2N$-Abz-EIAALEDEIAHLEKSIAALEK-OH | $[M+3]^{3+}$ | 805.0942 | 805.0962 |
| **E3H11** | $H_2N$-Abz-EIAALEKEIAHLEKEIAALEK-OH | $[M+3]^{3+}$ | 823.4838 | 823.4557 |
| **E3S15** | $H_2N$-Abz-EIAALEKEIAALEKSIAALEK-OH | $[M+3]^{3+}$ | 787.4521 | 787.4449 |
| **E3H11-Ala** | $H_2N$-Abz-KAEAAAAEAEAAHAEAEAAAAEAK-OH | $[M+3]^{3+}$ | 791.4105 | 791.3796 |
| **EEEHEEE** | $H_2N$-Abz-EEEHEEE-OH | $[M+1]^{+}$ | 1049.3684 | 1049.3621 |
| **E3yne** | 1-EIAALEKEIAALEKEIAALEK-OH | $[M+3]^{3+}$ | 788.4217 | 788.4438 |
| **E3H11yne** | 1-EIAALEKEIAHLEKEIAALEK-OH | $[M+3]^{3+}$ | 810.4565 | 810.4510 |
| **E3H11-Alayne** | 1-KAEAAAAEAEAAHAEAEAAAAEAK-OH | $[M+3]^{3+}$ | 778.3559 | 778.3750 |
| **K3** | $H_2N$-Abz-KIAALKEKIAALKEKIAALKE-OH | $[M+3]^{3+}$ | 800.5047 | 800.5008 |
| **K3CT** | $H_2N$-Abz-KIAALKDKIAHLKESIAALKE-OH | $[M+3]^{3+}$ | 804.1370 | 804.1485 |
| **Ac-IHIHIQI-$NH_2$** | Ac-IHIHIQI-$NH_2$ | $[M+1]^{+}$ | 914.5619 | 914.5575 |
| **CC-HexH24** | $H_2N$-GELKAIAQELKAIAKELKAIAWEHKAIAQGAG-OH | $[M+4]^{4+}$ | 840.5129 | 840.4943 |
| **CC-HexH24 scramble** | $H_2N$-EEKAHKLGAKIGLIQEAIAAKAWEAALQAIKG-OH | $[M+4]^{4+}$ | 840.5141 | 840.4943 |
| **CC-Hex** | $H_2N$-GELKAIAQELKAIAKELKAIAWELKAIAQGAG-OH | $[M+4]^{4+}$ | 834.5010 | 833.9885 |

### 8.1.5 Copper-mediated cycloaddition of alkyne-functionalized peptide to azide-functionalized hyper-branched polyglycerol.

Azide-functionalized (15.5%) hyper-branched polyglycerol (Mn = 13.2 kD, PDI = 1.62, 26 azide/molecule PG) was synthesized and purified in the research group of Rainer Haag. Alkyne-functionalized peptide E3H11yne (35 mg, 14 µmol, 1.5 eq. alkyne) and azide-hPG (5 mg 0.37 µmol, 1 eq. azide) were dissolved in 3:1 mixture of THF/water, which was then purged for 10 minutes with argon. Sodium ascorbate (4 mg, 20 µmol, 2 eq.) and $CuSO_4 \cdot 5H_20$ (0.5 mg, 2 µmol, 0.2 eq.) were dissolved in water and added to the mixture which was stirred at room temperature. After 72 h, the solvent was removed from the reaction mixture and the residue was redissolved in 50 mM aqueous solution of Tris base to be purified via dialysis (MWCO = 3.5-5 kD). Dialysis was carried out further with a saturated solution of ethylenediaminetetraacetic acid followed by a 1:1 mixture or methanol/$H_2O$. RP-HPLC analysis (Phenomenex® Kinetix C18, 5%-100% MECN in 18 min) showed conjugated hPG-E3H11 product (retention time = 12.1 min) and remaining E3H11yne. Pure hPG-E3H11yne (6.5 mg) was obtained by preparative HPLC.

Characterization of the grafted product hPG-E3H11 was performed using gel permeation chromatography (GPC) coupled to a refractive index (RI) detector. The obtained hPG-E3H11 was found to possess an Mn of 41010 g/mol. Taking into account the molecular weight of the peptide and the Mn the hPG, this value corresponds to an average of 11 peptides per hPG. The resulting DPI of hPG-E3H11 was 1.09. These measurements were carried out under highly dilute conditions (5 mg/mL) with the use of an Agilent 1100 solvent delivery system with pump, manual injector, and an Agilent differential refractometer. Three 30 cm columns (PPS: Polymer Standards Service GmbH, Germany; Suprema 100 Å, 1000 Å, 3000 Å with 5 and 10 µm particle size) were used to separate aqueous polymer samples using water with 0.1 M $NaNO_3$ as the mobile phase at a flow rate of 1 mL/min. The columns were operated at room temperature with the RI detector at 50 °C.

## 8.2 Kinetic assays

### 8.2.1 Materials for kinetic assays

**Buffers:** Tris/HCl was prepared using 2-amino-2-hydroxymethyl-propane-1,3-diol (Merck) and concentrated hydrochloric acid (37%; Roth). Carbonate buffer was prepared using sodium carbonate (Merck) and sodium hydrogen carbonate (99.5%; Roth). Water was obtained from a Milli-Q® Advantage A10 purification system.

**Determination of peptide concentration:** $H_2N$-Abz-Gly HCl (Bachem), guanidine hydrochloride (Amresco).

**Substrates:** *p*-Nitrophenylacetate (pNPA, ≥98%; Alfa Aesar), *N*-benzyloxycarbonyl-L-phenylalanine-*p*-nitrophenyl ester (Z-L-Phe-ONp; Bachem), *N*-benzyloxycarbonyl-D-phenylalanine-*p*-nitrophenyl ester (Z-D-Phe-ONp; Bachem), *N*-butyloxycarbonyl-L-asparagine-*p*-nitrophenyl ester (Boc-L-Asn-ONp; Bachem), *N*-butyloxycarbonyl-D-asparagine-*p*-nitrophenyl ester (Boc-D-Asn-ONp; Bachem). All substrate stock solutions were prepared with acetonitrile (HPLC grade; VWR).

**Additional materials:** Zinc chloride (99%; Grüssing).

### 8.2.2 Determination of peptide concentration

Peptide concentrations are accurately determined by measuring their absorption of ultraviolet light. This measurement is greatly simplified with the presence of aromatic residues such as tryptophan, tyrosine and phenylalanine (ca. 280 nm). If an aromatic residue is not present, other chromophores such as 2-aminobenzoic acid (Abz; 320 nm) can be incorporated. The preparation of the peptide stock solutions and concentration determination are described in the following sections.

#### 8.2.2.1 Preparation of E3/K3 variant stock solutions

Peptide stock solutions were prepared by dissolving lyophilized peptides in 50 mM Tris/HCl buffer pH 7.3 (ca. 1 mg/mL). The pH was carefully determined with a WTW 526 pH-meter equipped with an InLab® microelectrode (Mettler Toledo) and adjusted with 50 mM Tris solution to pH 7.3. To determine the peptide concentration, 100 µL of the stock was diluted with 900 µL buffer (10-fold dilution factor). The absorbance ($A_{320}$) of 2-aminobenzoic acid located on the amino-terminus was measured at 320 nm (Cary 50 Bio, Varion) using 1 cm Plastibrand® PMMA cuvettes. After subtracting the background absorbance (50 mM Tris/HCl pH 7.3), the concentrations of the samples measured were calculated using a 2-aminobenzoic extinction coefficient ($\varepsilon_{320}$) of 2980 $Lmol^{-1}cm^{-1}$. This extinction coefficient was predetermined with various concentrations of $H_2N$-Abz-Gly-OH HCl in 50 mM Tris/HCl pH 7.3. The concentration of the diluted samples measured were then multiplied by the dilution factor (DF = 10) to obtain the concentration of the peptide stock solution ($C_{stock}$) as shown in Equation 21.

$$C_{stock} = (A_{320} \cdot DF)/\varepsilon_{320} \qquad [21]$$

#### 8.2.2.2 Preparation of CC-Hex variant stock solutions

Lyophilized peptides were dissolved in the appropriate buffer (ca. 4 mg/mL) to create stock peptide solutions. To determine the concentration, 20 µL of a stock solution was added to 980 µL (dilution factor, DF = 50) 6 M guanidine hydrochloride at pH 6.5. The absorbance ($A_{280}$) of this mixture was measured at 280 nm (tryptophan) in a 1 cm quartz cuvette. With consideration of background absorbance, the concentration of the stock solution was determined as shown in Equation 21, in which the known tryptophan extinction coefficient ($\varepsilon_{280}$) of 5690 $Lmol^{-1}cm^{-1}$ was used.[205]

#### 8.2.2.3 Preparation of Ac-IHIHIQI-$NH_2$ stock solution

Lyophilized peptide was dissolved in 10 mM HCl. To measure the peptide concentration of this solution, 10 µL were added to 990 µL 10 mM HCl (DF = 100) and the absorbance was measured at 214 nm in a 1 cm quartz cuvette. This

absorbance measured was corrected for background absorbance. The concentration of the original solution was determined as shown in Equation 21 with a calculated peptide Ac-IHIHIQI-$NH_2$ molar extinction coefficient of 17,956 $Lmol^{-1}cm^{-1}$.[206] The peptide solution was then diluted with 10 mM HCl to achieve a concentration of 1.1 mM pH2 stock Ac-IHIHIQI-$NH_2$ solution.

### 8.2.3 Pseudo-first-order kinetic assays

The substrate *N*-benzyloxycarbonyl-phenylalanine-*p*-nitrophenyl ester (Z-Phe-ONp) was selected for many of the kinetic assays because it provides reasonable hydrophobic groups for semi-selective peptide interaction. In addition, the amino acid ester also allows testing for catalytic enantioselectivity. However, due to the low solubility of Z-Phe-ONp, all assays involving this substrate employed pseudo-first-order kinetics (Section 2.3.1) as opposed to saturation kinetics. This includes the kinetic studies performed to determine the catalytic activity of the E3/K3 variants (Section 6.1) and the peptide Ac-IHIHIQI-$NH_2$ (Section 6.2). A catalyzed reaction can be treated as a pseudo-first-order reaction when the concentration of substrate is low in comparison to $K_M$.[207] Under such conditions, the rate of the reaction becomes directly proportional to the substrate concentration (Section 2.3; Equations 18 and 19).

#### 8.2.3.1 Kinetic assays of E3/K3 variants

Ester hydrolysis was monitored by measuring the absorbance of *p*-nitrophenolate at 400 nm as a function of time ($A_t$) using a UV-Vis spectrophotometer (Cary 50 Bio, Varion) thermostatted at 20$^{o}$C (Lauda RKS). The reactions were performed in stoppered 1 cm quartz QS cuvettes (3.5 mL; Hellma Analytics) with a total peptide concentration of 50 µM prepared from peptide stock solutions and additional buffer (50 mM Tris/HCl pH 7.3). Prior to adding substrate, the stock substrate solution and the contents of the reaction cuvette were allowed 30 min. for equilibration to the proper temperature. Absorbance readings were recorded upon the addition of ester (Z-L-PheONp or, Z-D-PheONp, dissolved in acetonitrile) to meet initial reaction conditions of 10 µM substrate and 3.2 % acetonitrile. As a blank, the assay was performed as just described without peptide. In addition, assays were also performed

using histidine (50 µM) in place of peptide. Upon completion of the reaction (>24 h), the final absorbance ($A_\infty$) was measured at 400 nm. The final absorbance reading is particularly important as it is proportional to the concentration of *p*-nitrophenolate when the reaction is finished, which in turn is proportional to the concentration of the substrate ($[S]_o$) at the beginning of the reaction.

$$[S]_o \sim A_\infty \quad [22]$$

Therefore, the concentration of Z-Phe-ONp at a given time, [S], is

$$[S] \sim A_\infty - A_t \quad [23]$$

With Equations 22 and 23 substituted into Equation 6, pseudo-first-order rate constants ($k_1$) were determined as the negative slope of the first-order plot of linear Equation 24.

$$\mathrm{Ln}\,(A_\infty - A_t) = -k_1 t + \mathrm{Ln}\,A_\infty \quad [24]$$

Second-order rate constants ($k_2$) were obtained by subtracting the pseudo-first-order rate constant of the blank ($k_{1blank}$) from the apparent first-order rate constants measured in the presence of peptide or histidine ($k_1$) and dividing by the initial concentration ($[Cat]_0$) of the catalyst (peptide or histidine) as shown in Equation 25.

$$k_2 = (k_1 - k_{1blank})/[Cat]_0 \quad [25]$$

### 8.2.3.2 Kinetic assays of Ac-IHIHIQI-$NH_2$

Ester hydrolysis reactions were performed in 1 cm quartz QS cuvettes (3.5 mL; Hellma Analytics) equipped with Teflon stopper. All reactions with the exception of the blanks were performed using 10 µM Ac-IHIHIQI-$NH_2$, which corresponds to 5 µM of the characterized active catalyst.[188] Each of the reactions including the blank took place at 20°C in 50 mM Tris/HCl buffer pH 7.5 with 100 µM zinc chloride and a starting substrate concentration of 10 µM for Z-Phe-ONp hydrolysis and 50 µM for *p*NPA hydrolysis. Reaction progress was monitored by measuring the formation of *p*-nitrophenolate at 405 nm as a function of time ($A_t$) using a UV-Vis spectrophotometer (Cary 50 Bio, Varion) thermostatted at 20°C (Lauda RKS). Upon completion of the reaction (>24 h), the final absorbance ($A_\infty$) was measured at 405 nm. The apparent

first-order rate constants $k_1$ of the catalyzed and uncatalyzed blank reaction as well as second-order rate constants $k_2$ were determined as described in Section 8.2.3.1.

For reactions performed with incubated Ac-IHIHIQI-$NH_2$, peptide was added from the pH 2 stock solution to the cuvette containing buffer with $ZnCl_2$ 24 h prior to the reaction. These contents were equilibrated at 20°C for 30 min. before the start of the reaction. The reaction was initiated and the absorbance was recorded with the addition of substrate dissolved in acetonitrile (2% acetonitrile in reaction mixture).

For reactions performed with actively assembling Ac-IHIHIQI-$NH_2$, the reaction cuvette containing only buffer with $ZnCl_2$ was allowed to equilibrate 30 min. before the reaction. The reactions were started with the simultaneous addition of peptide and substrate from their thermo-equilibrated stock solutions.

### 8.2.4 Michaelis-Menten kinetic assays

Michaelis-Menten kinetic assays were performed to evaluate the catalytic activity of the peptide CC-HexH24 in the presence of $Zn^{2+}$ involved in the hydrolysis of substrate *p*NPA. In doing so, the initial reaction rates ($v_o$) of catalyzed reactions over a range of substrate concentrations were determined. A range of substrate concentrations was selected in order to show a saturation effect as the maximum reaction rate ($V_{max}$) is approached (Section 2.3.3). Initial reaction velocities were determined by recording the absorbance at 405 nm of *p*-nitrophenolate produced until 10 % substrate turnover was reached. The absorbance was then converted into concentration using experimentally determined molar extinction coefficients of *p*-nitrophenolate corresponding to the reaction conditions. Molar extinction coefficients were found to be 12,700 and 16,030 $Lmol^{-1}cm^{-1}$ at pH 7.5 and 8.5, respectively. A molar extinction coefficient of 18,700 $Lmol^{-1}cm^{-1}$ was used corresponding to conditions in which the pH was 9 or greater. The initial reaction rates of catalyzed reactions were calculated by plotting the concentration of *p*-nitrophenolate vs. time and subtracting the obtained slopes from those corresponding blank reactions occurring under the same conditions. With values of $v_o$, all Michaelis-Menten parameters were determined as described in Section 2.3.

All reactions were performed in 1 cm quartz QS cuvettes equipped with Teflon stoppers (3.5 mL; Hellma Analytics) and monitored using UV-Vis spectrophotometer

(Cary 50 Bio, Varion) thermostatted at 25°C (Lauda RKS). The reaction cuvettes containing buffer, $ZnCl_2$, and peptide and, separately, the stock solution of substrate in acetonitrile were allowed 30 min. to equilibrate to 25°C prior to the start of the reaction. The reactions were initiated with the addition of substrate to the cuvette. The reaction cuvettes contained 2% acetonitrile following substrate addition.

To assess the dependence of CC-HexH24 activity on zinc, assays were performed to determine $v_o$ at pH 7.5, 8.5 (50 mM Tris/HCl) and 9.5 (50 mM $Na(CO_3)_2/NaHCO_3$). These tests were carried out with 50 µM peptide and a starting substrate concentration of 1 mM. $Zn^{2+}$ concentrations of 0, 2.5, 5.0, 8.3, 12.5, 25, 50, and 100 µM were tested.

In order to ascertain the effect of pH on CC-HexH24/$Zn^{2+}$ activity, assays were performed with 50 µM peptide in the presence of 8.3 µM $ZnCl_2$ with a starting concentration of 1 mM *p*NPA. Experiments were executed at pH 7.5, 8.5, 9 (50 mM Tris/HCl) and 9.3, 9.5, 9.8 and 10.5 (50 mM $Na(CO_3)_2/NaHCO_3$).

To determine Michaelis-Menten parameters, saturation kinetics were performed at pH 9.5 (50 mM $Na(CO_3)_2/NaHCO_3$) using 30 µM peptide and 5 µM $ZnCl_2$. Starting substrate *p*NPA concentrations of 0.5, 1.0, 1.5, 2.0 and 3.0 mM were selected. As the values of $k_{cat}$, $K_M$ and $k_{cat}/K_M$ were attributed solely to catalysis resulting from the zinc-bound hydroxide mechanism, reaction mixtures containing CC-HexH24 (30 µM) in the absence $ZnCl_2$ were employed as blanks.

## 8.3 Structural analysis

### 8.3.1 Circular dichroism spectroscopy

Most biological molecules are chiral. This includes 19 of the 20 canonical amino acids and their arrangement into higher-order protein structures. The specific behavior of chiral molecules is not limited to their interactions with other chiral molecules, but is also exhibited in their ability to differently absorb polarized light.[208,209] Because small amounts of analyte can be quickly measured and analyzed under a wide range of conditions, circular dichroism (CD) spectroscopy is widely exploited in the structural analysis of biological molecules.[210]

Circularly polarized light arises from the superposition of vertical and horizontal plane-polarized light waves with a relative phase shift of 90° (Fig. 8.2).[211] The resulting electric field vector maintains a constant amplitude as it rotates about the

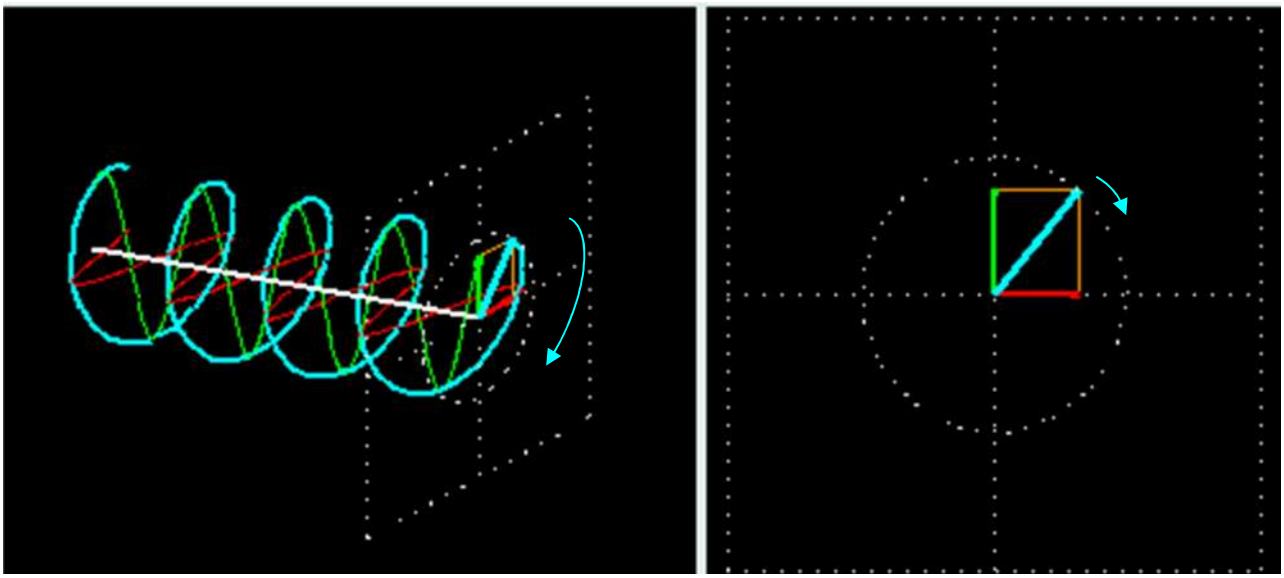


***Figure 8.2:*** *Right circularly polarized light (blue) resulting from the superposition of vertical plane-polarized (green) and vertical plane-polarized (red) light waves 90° out of phase. (Left) Three-dimension view and (right) front view of circular path traced by the electric field vector of the circularly polarized light (blue). Courtesy of Szilagyi.*[211]

axis of propagation. At a particular point of reference, this vector will trace one circle corresponding to one period. When facing the light source, the electric field vector of right circularly polarized light is observed tracing the circle in the clockwise direction, whereas that of the left counterpart will follow the circle in a counterclockwise manner. In turn, superposition of right and left circularly polarized light of equal amplitude results in the reformation of plane-polarized light.

Circular dichroism is the differential adsorption of right and left circularly polarized light by an optically active (in this case, CD-active) medium. This can be expressed as

$$\Delta A = A_L - A_R = \varepsilon_L \cdot c \cdot l - \varepsilon_R \cdot c \cdot l = \Delta\varepsilon \cdot c \cdot l \qquad [26]$$

In which A represents absorbance, c is the molar concentration of the CD-active substance, l is the path length of the light through the medium, and ε is the molar extinction coefficient of the substance for circularly polarized light.[212] Values of ε are dependent on the conditions in which the measurement is taken and on the wavelength of the light. The value Δε is empirically known as circular dichroism.

Although CD is typically measured according to the difference in absorption, it is commonly expressed in terms of ellipticity. The reasoning for the use of ellipticity becomes more obvious when the passage of circularly polarized light is observed through a CD-active medium (Fig. 8.3).

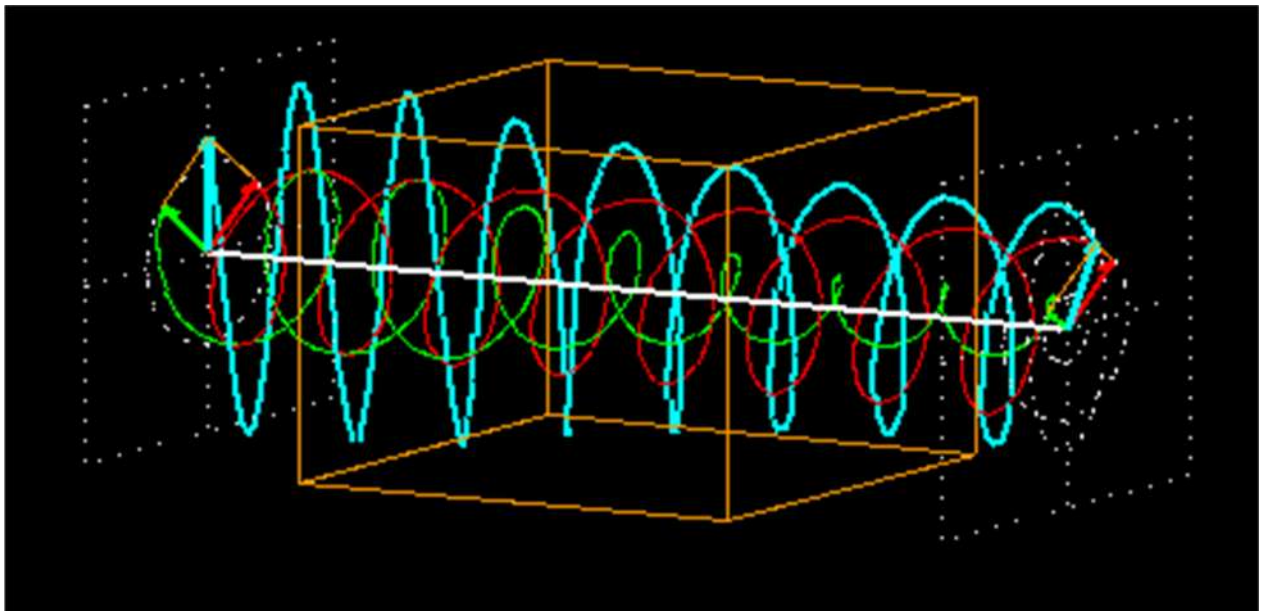


***Figure 8.3:*** *Plane-polarized light (blue) composed of equal amplitudes of right (green) and left (red) circularly polarized light passing through a CD-active medium (orange box). The additive electric field vectors of the circularly polarized light is transformed from a planar oscillation (blue) into an elliptical orbit due to the differential absorption of the left and right circularly polarized light. Courtesy of Szilagyi.*[211]

In Figure 8.3, it can be seen that as the plane-polarized light (blue), which is composed equally of left (red) and right (green) circularly polarized light, traverses the CD-active medium (orange box) the right circularly polarized light is absorbed to a greater extent than the left (Fig. 8.4). As a consequence, the incident plane-polarized light with an oscillating electric field is transformed into a detected light with an electric field tracing an elliptical pattern. The degree of the resulting ellipticity is described using the semimajor and semiminor elliptical axes (Fig. 8.5). The semimajor axis of an ellipse is the electric field vector of elliptically polarized light when both the left and right circularly polarized light are in phase, whereas the semiminor axis corresponds to left and right circularly polarized light when $180^{o}$ out of phase. As shown in Figure 8.5, the ratio of the semiminor and semimajor axes is

equal to the tangent of the ellipticity Θ. However, ellipticity is generally so small that tan Θ can be approximated as Θ in radians.

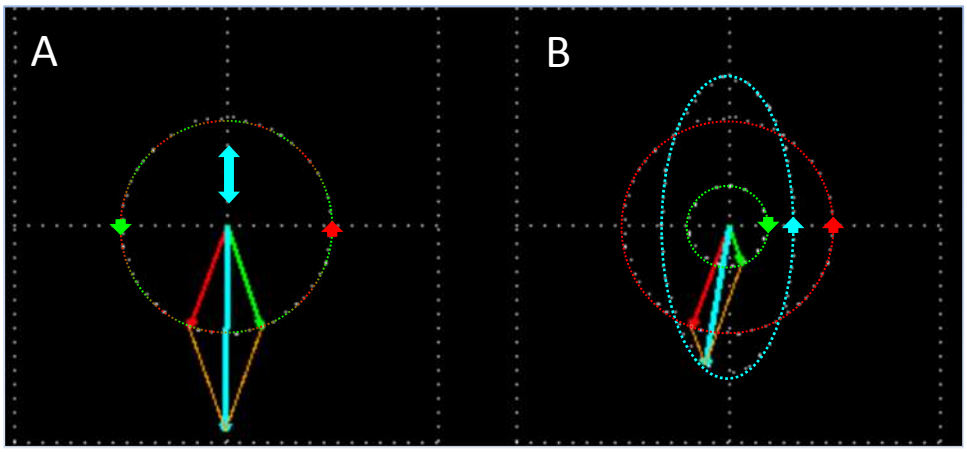



***Figure 8.4:*** *(A) Plane-polarized light (blue) composed of equal amplitudes of right (green) and left (red) circularly polarized light prior to passing through a CD-active medium. (B) After passing through a CD-active medium, The additive electric field vectors of the circularly polarized light is transformed from a planar oscillation (A:blue) into an elliptical orbit (B:blue) due to the differential absorption of the left (red) and right (green) circularly polarized light. Courtesy of Szilagyi.*[211]

$$\tan\Theta \approx \Theta\,(\text{rad}) = \frac{|E_L| - |E_R|}{|E_L| + |E_R|} \qquad [27]$$

Since intensity I is proportional to the square of the electric field vector E, Equation 27 becomes

$$\Theta\,(\text{rad}) = \frac{\sqrt{I_L} - \sqrt{I_R}}{\sqrt{I_L} + \sqrt{I_R}} \qquad [28]$$

By substituting $I = e^{A\ln 10}$ (Beer's law in natural log form),

$$\Theta\,(\text{rad}) = \frac{e^{\frac{A_L \ln 10}{2}} - e^{\frac{A_R \ln 10}{2}}}{e^{\frac{A_L \ln 10}{2}} + e^{\frac{A_R \ln 10}{2}}} = \frac{e^{\frac{\Delta A \ln 10}{2}} - 1}{e^{\frac{\Delta A \ln 10}{2}} + 1} \qquad [29]$$

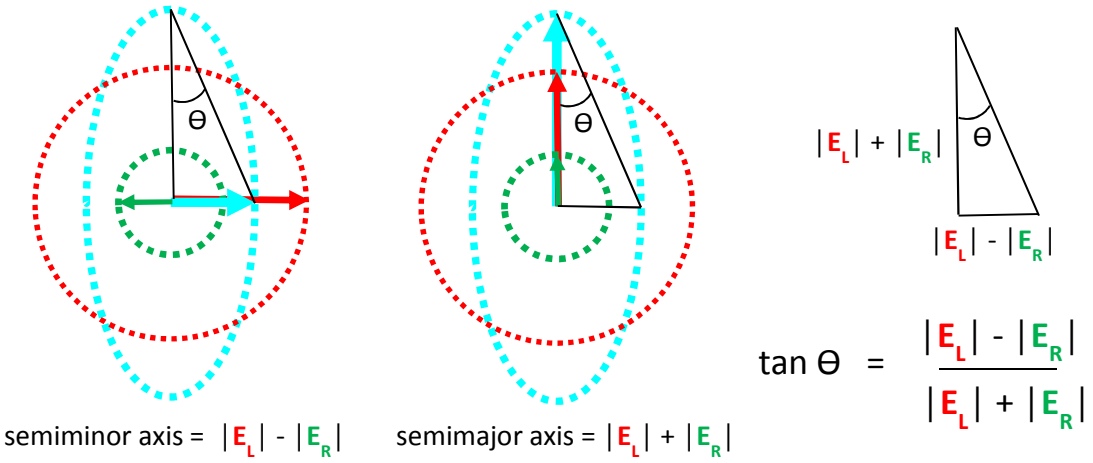



***Figure 8.5:*** *Circular dichroism characterized by ellipticity Θ. The blue ellipse represents the path of the electric field vector (blue) of elliptically polarized light resulting from the unequal absorption of left (red) and right (green) circularly polarized light. The tangent of ellipticity is determined by the ratio of the semiminor and semimajor axes.*

Following exponential expansion and the conversion of radians to degrees

$$\Theta\,(\mathrm{deg}) = \Delta A\left(\frac{\ln 10}{4}\right)\left(\frac{18\,0}{\pi}\right) \qquad [30]$$

Furthermore, molar ellipticity is defined as

$$[\Theta] = \frac{\Theta}{cl} \qquad [31]$$

When applied to peptides, molar ellipticity is commonly normalized according to the number of residues $n_r$ (ellipticity is often reported per residue, but should be per amide bond).

$$[\Theta] = \frac{\Theta}{cl} \cdot \frac{1}{n_r} \qquad [32]$$

As mentioned, CD spectroscopy is most commonly used to analyze the secondary structure of peptides and proteins. In doing so, the chiroptic character of the entire conformation as opposed to the sum of the chiral residues is acquired. Therefore, protein or peptides of the same residual constitution may possess different CD properties according to their conformation. Figure 8.6 shows CD spectra of secondary proteins structures with their signature bands.[213]

As shown in Figure 8.6, much of the CD-activity of protein and peptides takes place in the far UV region. It is in this energy range which electronic transitions of the amide group are observed (Fig. 8.7). Such transitions are the nπ* (ca. 220 nm) and

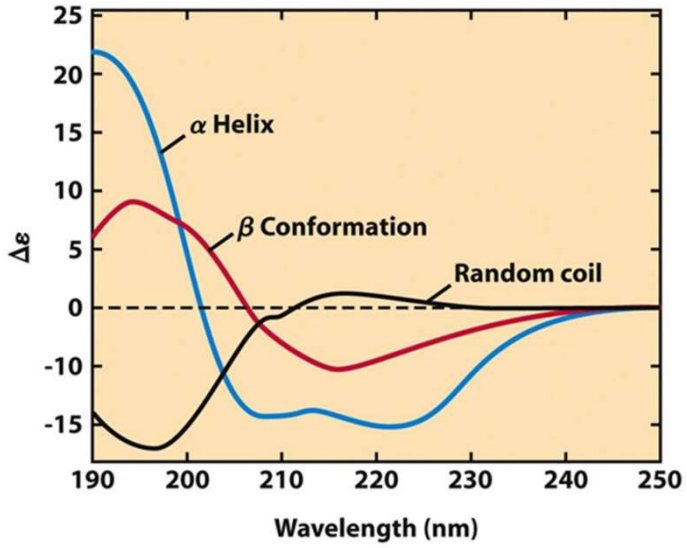



***Figure 8.6:*** *Characteristic CD-spectra of protein secondary structure. From Nelson & Cox.*[213]

the $\pi_o\pi^*$ (ca. 190 nm).[5] Characteristic signals of secondary structural elements shown in Figure 8.6 are summarized according to electronic transition in Table 8.4.

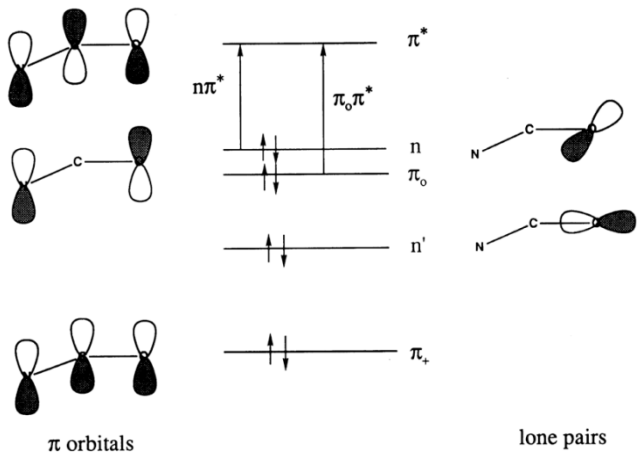



***Figure 8.7:*** *Molecular orbital and electronic transitions of the amide function. From Woody.*[212]

***Table 8.4:*** *Characteristic CD signals of protein secondary structural elements corresponding to electron transitions of the amide functionality.*

| Secondary structural | Signal | Electron transition | Wavelength (nm) |
|---|---|---|---|
| α-Helix | Positive | $\pi_0 \rightarrow \pi^*$ | 190-195 |
| | Negative | $\pi_0 \rightarrow \pi^*$ | 208 |
| | Negative | $n \rightarrow \pi^*$ | 222 |
| β-Sheet | Positive | $\pi_0 \rightarrow \pi^*$ | 195-200 |
| | Negative | $n \rightarrow \pi^*$ | 215-220 |
| Random coil | Negative | $\pi_0 \rightarrow \pi^*$ | ca. 200 |
| | Positive | $n \rightarrow \pi^*$ | 220 |

#### 8.3.1.1 Applied circular dichroism spectroscopy

Circular dichroism spectroscopy was performed using a *Jasco* J-810 spectropolarimeter with a *Jasco* PTC-348WI thermostat under a steady $N_2$ flow of 3 L/min. All samples of predetermined concentration were measured in a Quartz Suprasil cuvette (l = 0.1 cm) equipped with Teflon stopper from Hellma Analytics. Measurements on each sample were performed at least three times with the suitable background correction. Measurements were obtained in ellipticity units of mdeg which were converted in molar ellipticity (Equation 26) which were in turn normalized according to the number of amide bonds (Equation 27).

**Proportion α-helicity:** The proportion α-helicity of a sample was determined with the normalized, molar ellipticity $[\Theta]_{222}$ experimentally observed at 222 nm ($n \rightarrow \pi^*$; Fig. 8.6, Table 8.3). This calculation is based on the findings by Chen *et al.*[214] in which it was determined that the signal corresponding to the $n \rightarrow \pi^*$ transition of peptides depends on the number of residues.

$$[\Theta]_{222} = [\Theta]_{222}^{\infty} \cdot (1 - \frac{k}{n}) \qquad [33]$$

In this equation, $[\Theta]_{222}^{\infty}$ represents the ellipticity of an infinite chain of lysine residues corresponding to 100% α-helical content. The symbol k is what is called the chain-length-dependent factor, which is specific for the size-range of the peptide analyzed, and n is the number of residues. The value of k employed in this work was 4.6,[215]

which was also used to study the same E3/K3 heterodimeric peptide system applied here.[192]

**Thermal denaturation:** In addition to providing structural information of protein folding properties under various conditions, CD spectroscopy can also be effectively used to monitor structural transitions. The structural transition from native fold to random coil in the presence of increasing denaturing conditions is commonly used to assess the folding stability of a particular peptide or protein. Thermal stability is easily ascertained by monitoring the ellipticity in response to increasing temperature as the protein/peptide "melts" from its native to an undefined fold (random coil, Fig. 8.6).

In this work, ellipticity of the α-helical peptide CC-HexH24 at pH 7.5 was monitored as the temperature was gradually increased from 20-100$^{o}$C. This measurement was performed with and without $Zn^{2+}$ in order to determine the structure-stabilizing role of metal coordination.

# 9
# References

9. References

1. Pauling, L., Chemical Achievement and Hope for the Future, *Am. Sci.* **36**, 50-58 (1948).

2. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. & Walter, P. *Molecular Biology of the Cell* (Garland, 2007).

3. Voet, D., Voet, J.G. & Pratt, C.W. *Fundamentals of Biochemistry: Life at the Molecular Level* (Wiley, 2013).

4. Schmid, A., Dordick, J. S., Hauer, B., Keiner, A., Wubbolts, M. & Witholt, B. Industrial biocatalysis today and tomorrow. *Nature* **409**, 258-268 (2001).

5. Clouthier, C. M. & Pelletier, J. N. Expanding the organic toolbox: a guide to integrating biocatalysis in synthesis. *Chem. Soc. Rev.* **41**, 1585-1605 (2012).

6. Huisman, G. W. & Collier, S. J. On the development of new biocatalytic processes for practical pharmaceutical synthesis. *Curr. Opin. Chem. Biol*. **17**, 284-292 (2013).

7. Sheldon, R. A. Fundamentals of green chemistry: efficiency in reaction design. *Chem. Soc. Rev.* **41**, 1437-1451 (2012).

8. Reetz, M. T., Biocatalysis in Organic Chemistry and Biotechnology: Past, Present, and Future. *J. Am. Chem. Soc.* **135**, 12480-12496 (2013).

9. Johannes, T., Simurdiak, M. R. & Zhao, H. *Encyclopedia of Chemical Processing* (Taylor & Francis, 2006).

10. Khersonsky, O. & Tawfik, D. S. Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective. *Annu. Rev. Biochem.* **79**, 471-505 (2010).

11. Hult, K. & Berglund, P. Enzyme promiscuity: mechanism and applications. *Trends Biotechnol*. **25**, 231-238 (2007).

12. Busto, E., Gotor-Fernández, V. & Gotor, V. Hydrolases: catalytically promiscuous enzymes for the non-conventional reactions in organic synthesis*. Chem. Soc. Rev.* **39**, 4504-4523 (2010).

13. Berg, J. M., Tymoczko, J. L. & Stryer, L. *Biochemistry* (W.H. Freeman, 2006).

14. Bornscheuer, U. T., Huisman, G. W., Kazlauskas, R. J., Lutz, S., Moore, J. C. & Robins, K. Engineering the third wave of biocatalysis, *Nature* **485**, 185-194 (2012).

15. Arnold, F. H. Combinational and computational challenges for biocatalyst design. *Nature* **409**, 253-257 (2001).

16. Breslow, R. Biomimetic Chemistry and Artificial Enzymes: Catalysis by Design. *Acc. Chem. Res.* **28**, 146-153 (1995).

17. Breslow, R. Artificial Enzymes. *Science* **218**, 532-537 (1982).

18. Motherwell, W. B., Bingham, M. J. & Six, Y. Recent progress in the design and synthesis of artificial enzymes. *Tetrahedron* **57**, 4663-4686 (2001).

19. Lewis, J. C. Artificial Metalloenzymes and Metallopeptide Catalysts for Organic Synthesis. *ACS Catal*ysis **3**, 2954-2975 (2013).

20. Bos, J. & Roelfes, G. Artificial metalloenzymes for enantioselective catalysis. *Curr. Opin. Chem. Biol.* **19**, 135-143 (2014).

21. Sellergren, B., Karmalkar, R. N. & Shea, K. J. Enantioselective Ester Hydrolysis Catalyzed by Imprinted Polymers. *J. Org. Chem.* **65**, 4009-4027 (2000).

22. Bommarius, A. S., Blum, J. K. & Abrahamson, M. J. Status of protein engineering for biocatalysts: how to design an industrially useful biocatalyst. *Curr. Opin. Chem. Biol.* **15**, 194-200 (2011).

23. Reetz, M. T., Laboratory Evolution of Stereoselective Enzymes: A Prolific Source of Catalysis for Asymmetric Reactions. *Angew. Chem. Int. Ed.* **50**, 138-174 (2011).

24. Rotticci, D., Rotticci-Mulder, J.C., Denman, S., Norin, T. & Hult, K. Improved Enantioselectivity of a Lipase by Rational Protein Engineering. *ChemBioChem* **2**, 766-770 (2001).

25. Kaplan, J. & DeGrado, W. F. De novo design of catalytic proteins. *Proc. Nat. Acad. Sci.* **101**, 11566-11570 (2004).

26. Korendovych, I. V. & DeGrado, W. F. Catalytic efficiency of designed catalytic proteins. *Curr. Opin. Struct. Biol.* **27**, 113-121 (2014).

27. Zastrow, M. L., Peacock, A. F. A., Stuckey, J. A. & Pecoraro, V. L. Hydrolytic catalysis and structural stabilization in a designed metalloprotein. *Nat. Chem.* **4**, 118-123 (2012).

28. Kries, H., Blomberg, R. & Hilvert, D. *De novo* enzymes by computational design. *Curr. Opin. Chem. Biol.* **17**, 221-228 (2013).

29 Nanda, V. Do-it-yourself enzymes. *Nat. Chem. Biol.* **4**, 273-275 (2008)

30. Nanda, V. & Koder, R. L., Designing Artificial Enzymes by Intuition and Computation. *Nat. Chem.* **2**, 15-24 (2010).

31. Zastrow, M. L. & Pecoraro, V. L. Designing functional metalloproteins: from structural to catalytic metal sites. *Coord. Chem. Rev.* **257**, 2565-2588 (2013).

32. Davie, E. A. C., Mennen, S. M., Xu, Y. & Miller, S. J. Asymmetric Catalysis Mediated by Synthetic Peptides. *Chem. Rev.* **107**, 5759-5812 (2007).

33. List, B., Lerner, R. A. & Barbas, C. F. III, Proline-Catalyzed Direct Asymmetric Aldol Reactions. *J. Am. Chem. Soc*. **122**, 2395-2396 (2000).

34. Ahrendt, K. A., Borths, C. J. & MacMillan, D. W. C. New Strategies for Organic Catalysis: The First Highly Enantioselective Organocatalytic Diels-Alder Reaction. *J. Am. Chem. Soc.* **122**, 4243-4244 (2000).

35. MacMillan, D. W. C. The advent and development of organocatalysis. *Nature* **455**, 304-308 (2008).

36. Northrup, A. B. & MacMillan, D. W. C. The First Direct and Enantioselective Cross-Aldol Reaction of Aldehydes. *J. Am. Chem. Soc.* **124**, 6798-6799 (2002).

37. Yang, J. W., Stadler, M. & List, B. Proline-Catalyzed Mannich Reaction of Aldehydes with *N*-Boc-Imines. *Angew. Chem. Int. Ed.* **46**, 609-611 (2007).

38. Lam, Y., Houk, K. N., Scheffler, U. & Mahrwald, R. Stereoselectivities of Histidine-Catalyzed Asymmetric Aldol Additions and Contrasts with Proline Catalysis: A Quantum Mechanical Analysis. *J. Am. Chem. Soc.* **134**, 6286-6295 (2012).

39. Tang, Z., Jiang, F., Cui, X., Gong, L., Mi, A., Jiang, Y. & Wu, Y. Enantioselective direct aldol reactions catalyed by L-prolinamide derivatives. *Proc. Nat. Acad. Sci.* **101**, 5755-5760, (2004).

40. Mase, N., Nakai, Y., Ohara, N., Yoda, H., Takabe, K., Tanaka, F. and Barbas, C. F. III. Organocatalytic Direct Asymmetric Aldol Reactions in Water. *J. Am. Chem. Soc.* **128**, 734-735 (2006).

41. Tang, Z. & Marx, A. Proline-Modified DNA as Catalyst of the Aldol Reaction. *Angew. Chem. Int. Ed.* **46**, 7297-7300 (2007).

42. Mukherjee, S., Yang, J. W., Hoffmann, S. & List, B. Asymmetric Enamine Catalysis. *Chem. Rev.* **107**, 5471-5569 (2007).

43. Paradowska, J., Stodulski, M. & Mlynarski, J. Catalysts Based on Amino Acids for Asymmetric Reactions in Water. *Angew. Chem. Int. Ed.* **48**, 4288-4297 (2009).

44. Sutar, R. L. & Joshi, N. N. Systematic evaluation of a few proline derivatives as catalysts for a direct aldol reaction. *Tetrahedron: Asymmetry* **24**, 43-49 (2013).

45. Tsogoeva, S.B., Jagtap, S.B. & Ardemasova, Z.A. 4-*trans*-Amino-proline based di- and tetrapeptides as organic catalysts for asymmetric C-C bond formation reactions. *Tetrahedron: Asymmetry* **17**, 989-992 (2006).

46. Revell, J.D. & Wennemers, H. Peptidic catalysts developed by combinatorial screening methods. *Curr. Opin. Chem. Biol.* **11**, 269-278 (2007).

47. Jakobsche, C. E., Peris, G. & Miller, S. J. Functional Analysis of an Aspartate-Based Epoxidation Catalyst with Amide-to-Alkene Peptidomimetic Catalyst Analogues. *Angew. Chem. Int. Ed.* **47**, 6707-6711 (2008).

48. Fowler, B. S., Mikochik, P. J. & Miller, S .J. Peptide-Catalyzed Kinetic Resolution of Formamides and Thioformamides as an Entry to Nonracemic Amines. *J. Am. Chem. Soc.* **132**, 2870-2871 (2010).

49. Hao. L., Chen, S., Xu, J., Tiwari, B., Fu, Z., Li, T., Lim, J. & Chi, Y. R. Organocatalytic Activation of Alkylacetic Esters as Enolate Precursors to React with α,β–Unsaturated Imines. *Org. Lett.* **15**, 4956-4959 (2013).

50. Cram, D. J. The Design of Molecular Hosts, Guests, and Their Complexes. *Angew. Chem. Int. Ed.* **27**, 1009-1020 (1988).

51. Breslow, R. & Dong, S. D. Biomimetic Reactions Catalyzed by Cyclodextrins and Their Derivatives. *Chem. Rev.* **98**, 1997-2011 (1998).

52. Wulff, G. & Liu, J. Design of Biomimetic Catalysts by Molecular Imprinting in Synthetic Polymers: The Role of Transition State Stabilization. *Acc. Chem. Res.* **45**, 239-247 (2011).

53. Poznik, M. & König, B. Cooperative hydrolysis of aryl esters on functionalized membrane sufaces and in micellar solutions. *Org. Biolmol. Chem.* **12**, 3175-3180 (2014).

54. Doherty, E. A. & Doudna, J. A. Ribozyme Structures and Mechanisms. *Annul. Rev. Biophys. Biomol. Struct.* **30**, 457-475 (2001).

55. Edgcomb, S.P. & Murphy, K.P. Variability in the pKa of Histidine Side-Chains Correlates With Burial Within Proteins. *Proteins* **49**, 1-6 (2002).

56. Wong, S. S. & Jameson, D. M. *Chemistry of Protein and Nucleic Acid Cross-Linking and Conjugation* (CRC Press, 2011).

57. McCall, K. A., Huang, C. & Fierke, C. A. Function and Mechanism of Zinc Metalloenzymes. *J. Nutr.* **130**, 1437S-1446S (2000).

58. Avvaru, B. S., Kim, C. U., Sippel, K. H., Gruner, S. M, Agbandje-McKenna, M., Silverman, D. N. & McKenna, R. A Short, Strong Hydrogen Bond in the Active Site of Human Carbonic Anhydrase II. *Biochemistry* **49**, 249-251 (2010).

59. Moreland, J. L., Gramada, A., Buzko, O. V., Zhang, Q. & Bourne, P. E. The Molecular Biology Toolkit (MBT): a modular platform for developing molecular visualization applications. *BMC Bioinformatics* **6**, 6-21(2005).

60. Woolley, P. Models for metal ion function in carbonic anhydrase. *Nature* **258**, 677-682 (1975).

61. Vallee, B. L. & Auld, D. S. Active-site zinc ligands and activated $H_2O$ of zinc enzymes. *Proc. Natl. Acad. Sci.* **87**, 220-224 (1990).

62. Pauling, L. Nature of forces between large molecules of biological interest. *Nature* **161**, 707-709 (1948).

63. Wolfenden, R. Analog Approaches to the Structure of the Transition State in Enzyme Reactions. *Accounts Chem. Res.* **5**, 10-18 (1972).

64. Lienhard, G.E. Enymatic Catalysis and Transition-State Theory. *Science* **180**, 149-154 (1973).

65. Hedstrom, L. Enzyme Specificity and Selectivity. *Encyclopedia of Life Sciences* (Wiley, 2010).

66. Nelson, D. L. & Cox, M. M. *Lehninger: Principles of Biochemistry* (Freeman, 2013).

67. Pratt, C. & Cornely, K. *Essential Biochemistry* (Wiley, 2013).

68. Cleij, M. C., Drenth, W. & Nolte, R. J. M. Mechanism of Enantioselective Ester Cleavage by Histidine-Containing Dipeptides at a Micellar Interface. *J. Org. Chem.* **56**, 3883-3891 (1991).

69. Broo, K. S., Brive, L., Ahlberg, P. & Baltzer, L. Catalysis of Hydrolysis and Transesterification Reactions of p-Nitrophenyl Esters by a Designed Helix-Loop-Helix Dimer. *J. Am. Chem. Soc.* **119**, 11362-11372 (1997).

70. Zhang, B. & Breslow, R. Ester Hydrolysis by a Catalytic Cyclodextrin Dimer Enzyme Mimic with a Metallobipyridyl Linking Group. *J. Am. Chem. Soc.* **119**, 1676-1681 (1997).

71. Ueoka, R., Matsumoto, Y., Dŏzono, H., Yano, Y, Hirasa, H., Goto, K. & Kato, Y. Remarkable Substituent Effects on the Micellar Enantioselective Hydrolysis of Amino Acid Esters. *Tet. Lett.* **31**, 5311-5314 (1990).

72. Fukushima, Y. Enantioselectivity Enhancement of Ester Cleavage by a β-Sheet Polypeptide Containing Catalytic Triads in a Serine Protease. *Bull. Chem. Soc. Jpn.* **69**, 2269-2274 (1996).

73. Fukushima, Y. Stable β-Sheet Formation and Enhanced Hydrolytic Catalysis of a Sequential Alternating Amphiphilic Polypeptide Containing Catalytic Triads in a Serine Protease. *Bull. Chem. Soc. Jpn.* **69**, 431-439 (1996).

74. Pengo, P., Baltzer, L., Pasquato, L. & Scrimin, P. Substrate Modulation of the Activity of an Artificial Nanoesterase Made of Peptide-Functionalized Gold Nanoparticles. *Angew. Chem.* **119**, 404-408 (2007).

75. Zaramella, D., Scrimin, P. & Prins, L. J. Self-Assembly of a Catalytic Multivalent Peptide-Nanoparticle Complex. *J. Am. Chem. Soc.* **134**, 8396-8399 (2012).

76. Voet, D., Voet, J.G. & Pratt, C.W. *Fundamentals of Biochemistry* (Wiley, 2008).

77. Chang, R. *Physical Chemistry for the Biosciences* (Wiley, 2005).

78. Gal. J. The Discovery of Biological Enantioselectivity: Louis Pasteur and the Fermentation of Tartaric Acid, 1857-Areview and Analysis 150 Yr Later. *Chirality* **20**, 5-19 (2008).

79. Jaenicke, L. Centenary of the Award of a Nobel Prize to Eduard Buchner, the Father of Biochemistry in a Test Tube and Thus of Experimental Molecular Bioscience. *Angew. Chem. Int. Ed.* **46**, 6776-6782 (2007).

80. Scopes, R. K. Overview of Protein Purification and Characterization, *Curr. Protoc. Protein Sci.*, **1.1**, 1-6 (1995).

81. Faber, K., *Biotransformations in Organic Chemistry* (Springer, 1992).

82. Chen, S. T. & Wang, K. T. The Synthesis of β-Benzyl L-Aspartate and γ-Benzyl L-Glutamate by Enzyme-Catalyzed Hydrolysis. *Synthesis* **1987**, 581-582 (1987).

83. Li, Y. & Miller, S. J. Chemoenzymatic Synthesis of Each Enantiomer of Orthogonally Protected 4,4-Difluoroglutamic Acid: A Candidate Monomer for Chiral Bronsted Acid Peptide-Based Catalysts. *J. Org. Chem.* **76**, 9785-9791 (2011).

84. Monteiro, C. M., Lourenco, N. M. & Afonso, C. A. M. Separation of secondary alcohols via kinetic resolution using fatty esters as reusable acylating agents. *Tetrahedron: Asymmetry* **21**, 952-956 (2010).

85. Russo, M. E., Scialla, S., De Luca, V., Capasso, C., Olivieri, G. & Marzocchella, A. Immobilization of Carbonic Anhydrase for Biomimetic $CO_2$ Capture. *Chem. Eng. Trans.* **32**, 1867-1872 (2013).

86. Schellenberger, V., Jakubke, H. D., Zapevalova, N. P. & Mitin, Y. V. Protease-Catalyzed Peptide Synthesis Using Invers3e Substrates: The Influence of Reaction Conditions on the Trypsin Acyl Transfer Efficiency. *Biotechnol. Bioeng.* **38**, 104-108 (1991).

87. Ulijin, R. V., Baragana, B., Halling, P. J. & Flitsch, S. L. Protease-Catalyzed Peptide Synthesis on Solid Support. *J. Am. Chem. Soc.* **124**, 10988-10989 (2002).

88. Branneby, C., Carlqvist, P., Magnusson, A., Hult, K., Brinck, T. & Berglund, P. Carbon-Carbon Bonds by Hydrolytic Enzymes. *J. Am. Chem. Soc*. **125**, 874-875 (2003).

89. Li, C., Feng, X. W., Wang, N., Zhou, Y. J. & Yu, W. Q. Biocatalytic promiscuity: the first lipase-catalyzed asymmetric aldol reaction. *Green Chem.* **10**, 616-618 (2008).

90. Halgas, J. *Biocatalysis in Organic Synthesis* (Elsevier, 1992).

91. Klibanov, A. M. Improving enzymes by using them in organic solvents. *Nature.* **409**, 241-246 (2001).

92. Ke, T. & Klibanov, A. M. On Enzymatic Activity in Organic Solvents as a Function of Enzyme History. *Biotechnol. Bioeng.* **57**, 746-750 (1998).

93. Stahl, M., Jeppsson-Wistrand, U., Mansson, M. O. & Mosbach, K. Induced Stereoselectivity and Substrate Selectivity of Bio-Imprinted α-Chymotrypsin in Anhydrous Organic Media. *J. Am. Chem. Soc.* **113**, 9366-9368 (1991).

94. Rich, J. O. & Dordick, J. S. Controlling Subtilisin Activity and Selectivity in Organic Media by Imprinting with Nucleophilic Substrates. *J. Am. Chem. Soc.* **119**, 3245-3252 (1997).

95. Klibanov, A. M. Enzyme Memory: What is remembered and why? *Nature* **374**, 596 (1995).

96. Budisa, N. & Schulze-Makuck, D. Supercritical Carbon Dioxide and Its Potential as a Life-Sustaining Solvent in a Planetary Environment. *Life* **4**, 331-340 (2014).

97. Matsuda, T. Recent progress in biocatalysis using supercritical carbon dioxide. *J. Biosci. Bioeng.* **155**, 233-241 (2013).

98. Reed, M. C., Lieb, A. & Nijhout, H. F. The biological significance of substrate inhibition: A mechanism with diverse functions. *BioEssays* **32**, 422-429 (2010).

99. Glick, B. R., Pasternak, J. J. & Patten, C. L. *Molecular Biotechnology: Principles and Applications of Recombinant DNA* (ASM Press, 1998).

100. Zoller, M. J. New recombinant DNA methodology for protein engineering. *Curr. Opin. Biotechnol.* **3**, 348-354 (1992).

101. Chen. K. & Arnold, F. H. Tuning the activity of an enzyme for unusual environments: Sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. *Proc. Natl. Acad. Sci.* **90**, 5618-5622 (1993).

102. Sheldon, R. A. Enzyme Immobilization: The Quest for Optimum Performance. *Adv. Synth. Catal.* **349**, 1289-1307 (2007).

103. Datta, S., Christena, L. R. & Rajaram, Y. R. S. Enzyme immobilization: an overview on techniques and support materials. *3 Biotech* **3**, 1-9 (2013).

104. Woodley, J. M. Protein engineering of enzymes for process applications. *Curr. Opin. Chem. Biol.* **17**, 310-316 (2013).

105. Cedrone, F., Ménez, A. & Quéméneur. Tailoring new enzyme functions by rational redesign. *Curr. Opin. Struct. Biol.* **10**, 405-410 (2000).

106. Ó'Fágáin, C. Enzyme stabilization-recent experimental progress. *Enzyme Microb. Technol.* **33**, 137-149 (2003).

107. Van den Burg, B., Vriend, G., Veltman, O. R., Venema, G. & Eijsink, V. G. H. Engineering an enzyme to resist boiling. *Proc. Natl. Acad. Sci.* **95**, 2056-2060 (1998).

108. DeSantis, G. Shang, X. & Jones, J.B. Toward Tailoring the Specificity of the S1 Pocket of Subtilisin B. lentus: Chemical Modification of Mutant Enzymes as a Strategy for Removing Specificity Limitations. *Biochemistry* **38**, 13391-13397 (1999).

109. Hong, X., Luo, L., Taylor, K. L. & Dunaway-Mariano, D. Interchange of Catalytic Activity within the 2-Enoyl-Coenzyme A Hydratase/Isomerase Superfamily Based on a Common Active Site Template. *Biochemistry* **38**, 7638-7652 (1999).

110. Millard, C. B., Lockridge, O. & Broomfield, C. A. Organophosphorus Acid Anhydride Hydrolase Activity in Human Butyrylcholinesterase: Synergy Results in a Somanase. *Biochemistry* **37**, 237-247 (1998).

111. Ema, T., Kamata, S., Takeda, M., Nakano, Y. & Sakai, T. Rational creation of mutant enzyme showing remarkable enhancement of catalytic activity and enantioselectivity toward poor substrates. *Chem. Commun.* **46**, 5440-5442 (2010).

112. Shimoji, M., Yin, H., Higgins, L. & Jones, J. P. Design of a Novel P450: A Functional Bacterial-Human Cytochrome P450 Chimera. *Biochemistry* **37**, 8848-8852 (1998).

113. Lerner, R. A., Benkovic, S. J. & Schultz, P. G. At the Crossroads of Chemistry and Immunology: Catalytic Antibodies. *Science* **252**, 659-667 (1991).

114. Schultz, P. G. & Lerner, R. A. Completing the circle. *Nature* **418**, 485 (2002).

115. Liu. E., Prasad, L., Delbaere, L. T. J., Waygood, E. B. & Lee, J. S. Conversion of an antibody into an enzyme which cleaves the protein HPr. *Mol. Immunol.* **35**, 1069-1077 (1998).

116. Koshland, D. E. Jr. Confomational changes: How small is big enough? *Nat. Med*. **4**, 1112-1114 (1998).

117. Boersma, Y. L., Dröge, M. J. & Quax, W. J. Selection strategies for improved biocatalysts. *FEBS J.* **274**, 2181-2195 (2007).

118. Labrou, N. E. Random Mutagenesis Methods for In Vitro Directed Enzyme Evolution. *Curr. Protein Pept. Sci.* **11**, 91-100 (2010).

119. Giver, L. Gershenson, A., Freskgard, P-O. & Arnold, F. H. Directed evolution of a thermostable esterase. *Proc. Natl. Acad. Sci.* **95**, 12809-12813 (1998).

120. Reetz, M. A., Zonta, A., Schimossek, K., Liebeton, K. & Jaeger, K-E. Creation of Enantioselective Biocatalysts for Organic Chemistry by In Vitro Evolution. *Angew. Chem. Int. Ed.* **36**, 2830-2832 (1997).

121. Liebeton, K., Zonta, A., Schimossek, K., Nardini, M., Lang, D., Dijkstra, B.W., Reetz, M.T. & Jaeger, K-E. Directed evolution of an enantioselective lipase. *Chem. Biol*. **7**, 709-718 (2000).

122. Reetz, M. T., Prasad, S., Carballeira, J. D., Gumulya, Y. & Bocola, M. Iterative Saturation Mutagenesis Accelerates Laboratory Evolution of Enzyme Stereoselectivity: Rigorous Comparison with Traditional Methods. *J. Am. Chem. Soc.* **132**, 9144-9152 (2010).

123. Reetz, M. T. & Carballeira, J. D. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat. Protoc.* **2**, 891-903 (2007).

124. Reetz, M. T., Wang, L-W. & Bocola, M. Directed Evolution of Enantioselective Enzymes: Iterative Cycles of CASTing for Probing Protein-Sequence Space. *Angew. Chem. Int. Ed.* **45**, 1236-1241 (2006).

125. Reetz, M. T. The Importance of Additive and Non-Additive Mutational Effects in Protein Engineering. *Angew. Chem. Int. Ed*. **52**, 2658-2666 (2013).

126. Stemmer, W. P. C. DNA shuffling by random fragmentation and reassembly*: In vitro* recombination for molecular evolution. *Proc. Natl. Acad. Sci.* **91**, 10747-10751 (1994).

127. Stemmer, W. P. C. Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature* **370**, 389-391 (1994).

128. Coco, W. M., Levinson, W. E., Crist, M. J., Hektor, H. J., Darzins, A., Pienkos, P. T., Squires, C. H. & Monticello, D. J. DNA shuffling method for generating highly recombined genes and evolved enzymes. *Nat. Biotechnol.* **19**, 354-359 (2001).

129. Sieber, V., Martinez, C. A. & Arnold, F. H. Libraries of hybrid proteins from distantly related sequences. *Nat. Biotechnol.* **19**, 456-460 (2001).

130. Rubin-Pitel, S. B., Cho, C. M-H., Chen, W. & Zhao, H. Directed Evolution Tools in Bioproduct and Bioprocess Development. From *Bioprocessing for Value-Added Products from Renewable Resources* (Elsevier, 2007).

131. Chica, R.A., Doucet, N. & Pelletier, J. N. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. *Curr. Opin. Biotechnol.* **16**, 378-384 (2005),

132. Aharoni, A., Griffiths, A. D. & Tawfik, D. S. High-throughput screens and selections of enzyme-encoding genes. *Curr. Opin. Chem. Biol.* **9**, 210-216 (2005).

133. DeGrado, W. F., Summa, C. M., Pavone, V., Nastri, F. & Lombardi, A. De Novo Design and Structural Characterization of Proteins and Metalloproteins. *Annu. Rev. Biochem*. **68**, 779-819 (1999).

134. Lutz, S. Beyond directed evolution – semi-rational protein engineering and design. *Curr. Opin. Biotechnol.* **21**, 734-743 (2010).

135. Höhne, M. & Bornscheuer, U. T. Protein Engineering from “Scratch” Is Maturing. *Angew. Chem. Int. Ed.* **53**, 1200-1202 (2014).

136. Bolon, D. N. & Mayo, S. L. Enzyme-like proteins by computational design. *Proc. Natl. Acad. Sci.* **98**, 14274-14279 (2001).

137. Jiang, L., Althoff, E. A., Clemente, F. R., Doyle, L., Röthlisberger, D., Zanghellini, A., Gallaher, J. L., Betker, J. L., Tanaka, F., Barbas, C. F. III, Hilvert, D., Houk, K. N., Stoddard, B. L. & Baker, D. De Novo Computational Design of Retro-Aldol Enzymes. *Science* **319**, 1387-1391 (2008).

138. Althoff, E. A., Wang, L., Jiang, L., Giger, L., Lassila, J. K., Wang, Z., Smith, M., Hari, Kast, P., Herschlag, D., Hilvert, D. & Baker, D.Robust design and optimization of retroaldol enzymes. *Protein Sci.* **21**, 717-726 (2012).

139. Röthlisberger, D., Khersonsky, O., Wollacott, A. M., Jiang, L., DeChancie, J., Betker, J., Gallaher, J. L., Althoff, E. A., Zanghellini, A., Dym, O., Albeck, S., Houk, K. N., Tawfik, D. S. & Baker, D. Kemp Elimination catalysts by computational enzyme design. *Nature* **453**, 190-197 (2008).

140. Khersonsky, O., Kiss, G., Röthlisberger, D., Dym, O., Albeck, S. Houk, K. N., Baker, D. & Tawfik, D. S. Bridging the gaps in design methodologies by evolutionary optimization of the stability and proficiency of designed Kemp eliminase KE59. *Proc. Natl. Acad. Sci.* **109**, 10358-10363 (2012).

141. Privett, H. K., Kiss, G., Lee, T. M., Blomberg, R., Chica, R. A., Thomas, L. M., Hilvert, D., Houk, K. N., Mayo, S. L. Iterative approach to computational enzyme design. *Proc. Natl. Acad. Sci*. **109**, 3790-3795 (2012).

142. Blomberg, R., Kries, H., Pinkas, D. M., Mittl, P. R. E., Grütter, M. G., Privett, H. K., Mayo, S. L. & Hilvert, D. Precision is essential for efficient catalysis in an evolved Kemp eliminase. *Nature* **503**, 418-423 (2013).

143. Siegel, J. B., Zanghellini, A., Lovick, H. M., Kiss, G., Lambert, A. R., St. Clair, J. L., Gallaher, J. L., Hilvert, D., Gelb, M. H., Stoddard, B. L., Houk, K. N., Michael, F. E. & Baker, D. Computational Design of an Enzyme Catalyst for a Stereoselective Bimolecular Diels-Alder Reaction. *Science* **329**, 309-313 (2010).

144. Eiben, C. B., Siegel, J. B., Bale, J. B., Cooper, S., Khatib, F., Shen, C. B., Foldit Players, Stoddard, B. L., Popovic, Z. & Baker, D. Increased Diels-Alder Activity through Foldit Player Guided Backbone Remodeling. *Nat. Biotechnol.* **30**, 190-192 (2013).

145. Richter, F., Blomberg, R., Khare, S. D., Kiss, G., Kuzin, A. P., Smith, A .J. T., Gallahar, J., Pianowski, Z., Helgeson, R. C., Grjasnow, A., Xiao, R., Seetharaman, Su, M., Vorobiev, S., Lew, S., Forouhar, F., Kornhaber, G. J., Hunt, J. F., Montelione, G. T., Tong, L., Houk, K. N., Hilvert, D. & Baker, D. Computational Design of Catalytic Dyads and Oxyanion Holes for Ester Hydrolysis. *J. Am. Chem. Soc.* **134**, 16197-16206 (2012).

146. Gleeson, M. P., Burton, N. A. & Hillier, I. H. The mechanism of adenosine deaminase catalysis studied by QM/MM calculations: The role of histidine 238 and the activity of the alanine 238 mutant. *Phys. Chem. Chem. Phys.* **5**, 4272-4278 (2003).

147. Wilson, D. K., Rudolph, F. B. & Quiocho, F. A. Atomic Structure of Adenosine Deaminase Complexed with a Transition-State Analog: Understanding Catalysis and Immunodeficiency Mutations. *Science* **252**, 1278-1284 (1991).

148. Khare, S. D., Kipnis, Y., Greisen Jr., P., Takeuchi, R., Ashani, Y., Goldsmith, M., Song, Y., Gallaher, J. L., Silman, I., Leader, H., Sussman, J. L., Stoddard, B. L., Tawfik, D. S. & Baker, D. Redesign of a mononuclear zinc metalloenzyme for organophosphate hydrolysis. *Nat. Chem. Biol.* **8**, 294-300 (2014).

149. Der, D. S., Edwards, D. R. & Kuhlman, B. Catalysis by a de novo zinc-mediated interface: implications for natural enzyme and rational enzyme engineering. *Biochemistry* **51**, 3933-3940 (2012).

150. Eathiraj, S., Pan, X., Ritacco, C. J. & Lambright, D. G. Structural basis of family-wide Rab GTPase recognition by Rabenosyn-5. Nature 436, 415-419 (2005).

151. Dwars, T., Paetzold, E. & Oehme, G. Reactions in Micellar Systems. *Angew. Chem. Int. Ed*. **44**, 7174-7199 (2005).

152. Raynal, M., Ballester, P., Vidal-Ferran, A. & van Leeuwen, P. W. N. M. Supramolecular catalysis. Part 2: artificial enzyme mimics. *Chem. Soc. Rev*. **43**, 1734-1787 (2014).

153. Shepherd, N. E., Hoang, H. N., Abbenante, G. & Fairlie, D. P. Single Turn Peptide Alpha Helices with Exceptional Stability in Water. *J. Am. Chem. Soc*. **127**, 2974-2983 (2005).

154. Scholtz, J. M. & Baldwin, R. L. The Mechanism of α-Helix Formation by Peptides. *Annu. Rev. Biophys. Biomol. Struct.* **21**, 95-118 (1992).

155. Zimm, B. H. & Bragg, J. K. Theory of the Phase Transition between Helix and Random Coil in Polypeptide Chains. *J. Chem. Phys.* **31**, 526-535 (1959).

156. Moutevelis, E. & Woolfson, D. N. A Periodic Table of Coiled-Coil Protein Structures. *J. Mol. Biol.* **385**, 726-732 (2009).

157. Parry, D. A. D., Fraser, R. D. B. & Squire, J. M. Fifty years of coiled-coils and α-helical bundles: A close relationship between sequence and structure. *J. Struct. Biol.* **163**, 258-269 (2008).

158. Yu, Y. B. Coiled-coils: stability, specificity, and drug delivery potential. *Adv. Drug Deliv. Rev.* **54**, 1113-1129 (2002).

159. Burkhard, P., Stetefeld, J. & Strelkov, S. V. Coiled coils: a highly versatile protein folding motif. *Trends Cell Biol.* **11**, 82-88 (2001).

160. Gáspári, Z. & Nyitray, L. Coiled coils as possible models of protein structure evolution. *BioMol Concepts.* **2**, 199-210 (2011).

161. Apostolovic, B., Danial, M. & Klok, H-A. Coiled coils: attractive protein folding motifs for the fabrication of self-assembled, responsive and bioactive materials. *Chem. Soc. Rev.* **39**, 3541-3575 (2010).

162. Monera, O. D., Zhou, N. E., Lavigne, P., Kay, C. M. & Hodges, R. S. Formation of Parallel and Antiparallel Coiled-coils Controlled by the Relative Positions of Alanine Residues in the Hydrophobic Core. *J. Biol. Chem.* **271**, 3995-4001 (1996).

163. Zhou, N. E., Kay, C. M. & Hodges, R. S. Synthetic Model Proteins: Positional Effects of Interchain Hydrophobic Interactions on Stability of Two-stranded α-Helical Coiled-coils. *J. Biol. Chem.* **267**, 2664-2670 (1992).

164. Walshaw, J. & Woolfson, D. N. Extended knobs-into-holes packing in the classical complex coiled-coil assemblies. *J. Struct. Biol.* **144**, 349-361 (2003).

165. Walshaw, J. & Woolfson, D. N. SOCKET: A Program for Identifying and Analysing Coiled-coil Motifs Within Protein Structures. *J. Mol. Biol.* **307**, 1427-1450 (2003).

166. Lee, D. L., Ivaninskii, S., Bunkhard, P. & Hodges, R. S. Unique stabilizing interactions identified in the two-stranded α-helical coiled-coil: Crystal structure of a cortexillin I/GCN4 hybrid coiled-coil peptide. *Protein Sci.* **12**, 1395-1405 (2003).

167. Dahiyat, B. I., Gordon, D. B. & Mayo, S. L. Automated design of the surface positions of protein helices. *Protein Sci.* **6**, 1333-1337 (1997).

168. Mason, J. M. & Arndt, K. M. Coiled Coil Domains: Stability, Specificity, and Biological Implications. *ChemBioChem.* **5**, 170-176 (2004).

169. Sasaki, T. & Kaiser, E.T. Helichrome: Synthesis and Enzymatic Activity of a Designed Hemeprotein. *J. Am. Chem. Soc*. **111**, 380-381 (1989).

170. Johnsson, K., Allemann, R.K., Widmer, H. & Benner, S.A. Synthesis, structure and activity of artificial, rationally designed catalytic polypeptides. *Nature* **365**, 530-532 (1993).

171. Taylor, S. E., Rutherford, T. J. & Allemann, R. K., Design of a folded, conformationally stable oxaloacetate decarboxylase. *J. Chem. Soc., Perkin Trans. 2.* 751-755 (2002).

172. Nicoll, A. J. & Allemann, R. K. Nucleophilic and general acid catalysis at physiological pH by a designed miniature esterase. *Org. Biomol. Chem.* **2**, 2175-2180 (2004).

173. Broo, K. S., Nilsson, H., Nilsson, J., Flodberg, A. & Baltzer, L. Cooperative Nucleophilic and General-Acid Catalysis by the $HisH^+$-His Pair and Arginine Transition State Binding in Catalysis of Ester Hydrolysis Reactions by Designed Helix-Loop-Helix Motifs. *J. Am. Chem. Soc*. **120**, 4063-4068 (1998).

174. Baltzer, L., Broo, K. S., Nilsson, H. & Nilsson. J. Designed Four-helix Bundle Catalysts-the Engineering of Reactive Sites for Hydrolysis and Transesterification Reactions of *p*-Nitrophenyl Esters. *Bioorg. Med. Chem.* **7**, 83-91 (1999).

175. Razkin, J., Nilsson, H. & Baltzer, L. Catalysis of the Cleavage of Uridine 3´-2,2,2-Trichloroethylphosphate by a Designed Helix-Loop-Helix Motif Peptide. *J. Am. Chem. Soc*. **129**, 14752-14758 (2007).

176. Razkin, J., Lindgren, J., Nilsson, H. & Baltzer, L. Enhanced Complexity and Catalytic Efficiency in the Hydrolysis of Phosphate Diesters by Rationally Designed Helix-Loop-Helix Motifs. *ChemBioChem* **9**, 1975-1984 (2008).

177. Lu, Y., Yeung, N., Sieracki, N. & Marshall, M. Design of functional metalloproteins. *Nature* **460**, 855-862 (2009).

178. Yamashita, M. M., Wesson, L., Eisenman, G. & Eisenberg, D. Where metal ions bind in proteins. *Proc. Natl. Acad. Sci.* **87**, 5648-5652 (1990).

179. Faiella, M., Andreozzi, C., de Rosales, R.T.M., Pavone, V., Maglio, O., Nastri, F., DeGrado, W.F. & Lombardi, A. An artificial di-iron oxo-protein with phenol oxidase activity. *Nat. Chem. Biol.* **5**, 882-884 (2009).

180. Reig, A. J., Pires, M. M., Snyder, R. A., Wu, Y., Jo, H., Kulp, D. W., Butch, S. E., Calhoun, J. R., Szyperski, T., Soloman, E. I. & DeGrado, W. F. Altering the $O_2$-Dependent Reactivity of *de novo* Due Ferri Proteins. *Nat. Chem.* **4**, 900-906 (2012).

181. Zastrow, M. L. & Pecoraro, V. L. Influence of Active Site Location on Catalytic Activity in *de Novo*-Designed Zinc Metalloenzymes. *J. Am. Chem. Soc.* **135**, 5895-5903 (2013).

182. Zastrow, M. L. & Pecoraro, V. L. Designing Hydrolytic Zinc Metalloenzymes. *Biochemistry* **53**, 957-978 (2014).

183. Tegoni, M., Yu, F., Bersellini, M., Penner-Hahn, J. E. & Pecoraro, V. L. Designing a functional type 2 copper center that has nitrite reductase activity within α-helical coiled coils. *Proc. Natl. Acad. Sci.* 109, 21234-21239 (2012).

184. Bryson, J. W., Desjarlais, J. R., Handel, T. M. & DeGrado, W. F. From coiled coils to small globular proteins: Design of a native-like three-helix bundle. *Protein Sci.* **7**, 1404-1414 (1998).

185. Cangelosi, V. M., Deb, A., Penner-Hahn, J. E. & Pecoraro, V. L. A De Novo Designed Metalloenzyme for the Hydration of $CO_2$. *Angew. Chem. Int. Ed*. **53**, 7900-7903 (2014).

186. Hecht, M. H., *De novo* design of β-sheet proteins. *Proc. Natl. Acad. Sci.* **91**, 8729-8730 (1994).

187. Kopecek, J. & Yang, J. Smart Self-Assembled Hybrid Hydrogel Biomaterials. *Angew. Chem. Int. Ed.* **51**, 7396-7417 (2012).

188. Rufo, C. M., Moroz, Y. S., Moroz, O. V., Stöhr, J., Smith, T. A., Hu, X., DeGrado, W. F. & Korendovych, I. V. Short peptides self-assemble to produce catalytic amyloids. *Nat. Chem.* **6**, 303-309 (2014).

189. Zhang, C., Xue, X., Luo, Q., Li, Y., Yang, K., Zhuang, X., Jiang, Y., Zhang, J., Liu, J., Zou, G. & Liang, X-J. Self-Assembled Peptide Nanofibers Designed as Biological Enzymes for Catalyzing Ester Hydrolysis. *ACS Nano* **8**, 11715-11723 (2014).

190. Rebek Jr., J. On the Structure of Histidine and Its Role in Enzyme Active Sites. *Struct. Chem.* **1**, 129-131 (1989).

191. O'Shea, E. K., Lumb, K. J. & Kim, P. S. Peptide 'Velcro': design of a heterodimeric coiled coil. *Curr. Biol.* **3**, 658-667 (1993).

192. Litowski, J. R. & Hodges, R. S. Designing Heterodimeric Two-stranded α-Helical Coiled-coils. *J. Biol. Chem.* **277**, 37272-37279 (2002).

193. Hedstrom, L. Serine Protease Mechanism and Specificity. *Chem Rev*. **102**, 4501-4523 (2002).

194. Guarise, C., Manea, F., Zaupa, F., Pasquato, L., Prins, L. J. & Scrimin, P. Cooperative nanosystems. *J. Peptide Sci.* **14**, 174-183 (2008).

195. Caminade, A. M., Turrin, C.O., Laurent, R., Ouali, A. & Delavaux-Nicot, B. Eds. *Dendrimers: Towards Catalytic, Material and Biomedical Uses* (Wiley, 2011).

196. Pecchioli, T., Muthyala, M. K., Haag, R. & Christmann, M. Multivalent polyglycerol supported imidazolidin-4-one organocatalysts for enantioselective Friedel–Crafts alkylations. *Beilstein J. Org. Chem.* **11**, 730-738 (2015).

197. Sunder, A., Mülhaupt, R., Haag, R. & Frey, H. Hyperbranched Polyether Polyols: A Modular Approach to Complex Polymer Architectures. *Adv. Mater.* **12**, 235–239 (2000).

198. Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angew. Chem. Int. Ed*. **41**, 2596-2599 (2002).

199. Lucarini, M., Franchi, P.,Pedulli, G. F., Pengo, P., Scrimin, P., Pasquato, L. EPR Study of Dialkyl Nitroxides as Probes to Investigate the Exchange of Solutes between the Ligand Shell of Monolayers of Protected Gold Nanoparticles and Aqueous Solutions. *J. Am. Chem. Soc.* **126**, 9326-9329 (2004).

200. Lucarini, M., Franchi, P.,Pedulli, G. F., Gentilini, C., Polizzi, S., Pengo, P., Scrimin, P., Pasquato, L. Effect of Core Size on the Partition of Organic Solutes in the Monolayer of Water-Soluble Nanoparticles: An ESR Investigation. *J. Am Chem. Soc.* **127**, 16384-16385 (2005).

201. Ihara, Y. Kunikiyo, N. & Kunimasa, T. Stereoselective Micellar Catalysis, Part 4. Catalytic Hydrolysis of Enantiomeric Esters by Dipeptide Derivatives containing a Histidyl Residue. *J. Chem. Soc. Perkin Trans. II* 1741-1745 (1983).

202. Zaccai, N. R., Chi, B., Thomson, A. R., Boyle, A. L., Bartlett, G. J., Bruning, M., Linden, N., Sessions, R. B., Booth, P. J., Brady, R. L. & Woolfson, D. N. A de novo peptide hexamer with a mutable channel. *Nat. Chem. Biol*. **7**, 935-941 (2011).

203. Chan, W. C. & White, P. C. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach* (Oxford, 2004).

204. Dick, F. Acid Cleavage/Deprotection in Fmoc/*t*Bu Solid-Phase Peptide Synthesis from *Methods in Molecular Biology, vol 35 Peptide Synthesis Protocols* (Humana, 1994).

205. Edelhoch, H. Spectroscopic Determination of Tryptophan and Tyrosine in Proteins. *Biochemistry* **6**, 1948-1954 (1967).

206. Kuipers, B. J. H. & Gruppen, H. Prediction of Molar Extinction Coefficients of Proteins and Peptides Using UV Absorptions of the Constituent Amino Acids at 214 nm to Enable Quantitative Reverse Phase High-Performance Liquid Chromatography-Mass Spectrometry Analysis. *J. Agricult. Food Chem.* **55**, 5445-5451 (2007).

207. Schnell, S. & Mendoza, C. The condition for pseudo-first-order kinetics in enzymatic reactions is independent of the initial enzyme concentration. *Biophys. Chem.* **107**, 165-174 (2004).

208. Kelly, S.M. & Price, N.C. The Use of Circular Dichroism in the Investigation of Protein Structure and Function. *Curr. Protein Peptide Sci.* **1**, 349-384 (2000).

205. Whitmore, L. & Wallace, B. A. Protein Secondary Structure Analysis from Circular Dichroism Spectroscopy: Methods and Reference Databases. *Biopolymers* **89**, 392-400 (2007).

206. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protocol.* **1**, 2876-2890 (2006).

211. Szilagyi, A. Electromagnetic waves and circular dichroism: an animated tutorial. Institute of Enzymology of the Research Centre for Natural Sciences, Hungarian Academy of Sciences. Viewed 19 June, 2015. http://cddemo.szialab.org.

212. Fasman, G. D. Ed. Circular Dichroism and the Conformational Analysis of Biomolecules. (Plenum, 1996).

213. Nelson, D. L. & Cox, M. M. Lehninger Principles of Biochemistry (Freeman, 2008).

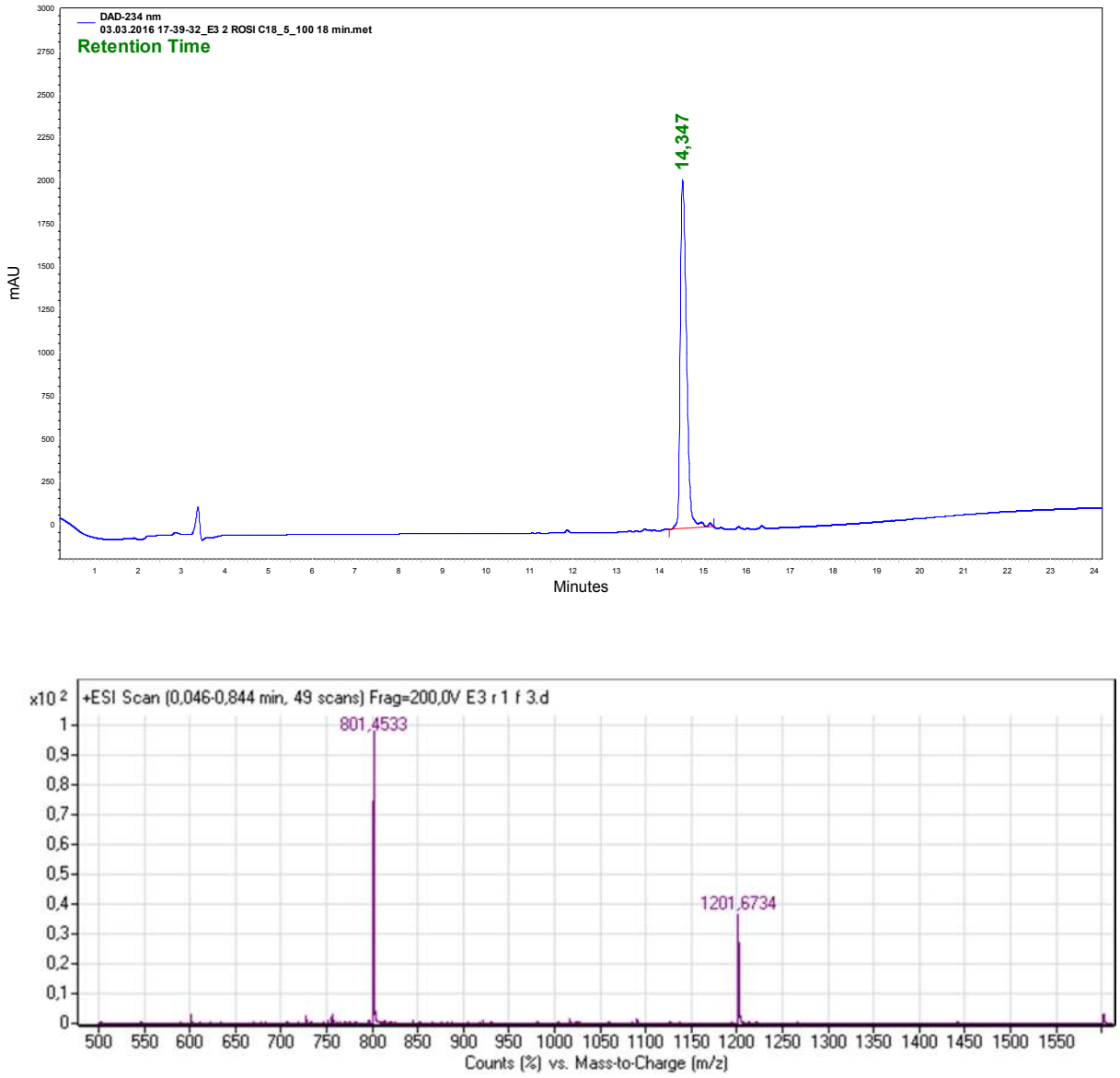
214. Chen, Y-H., Yang, J. T. & Chau, K. H. Determination of the Helix and β Form of Proteins in Aqueous Solution by Circular Dichroism. *Biochemistry* **13**, 3350-3359 (1974).

215. Gans, P. J., Lyu, P. C. Manning, M. C., Woody, R. W. & Kallenbach, N. R. The Helix-Coil Transition in Heterogeneous Peptides with Specific Side-Chain Interaction: Theory and Comparison with CD Spectral Data. *Biopolymers* **31**, 1605-1614 (1991).

# 10
# Appendix

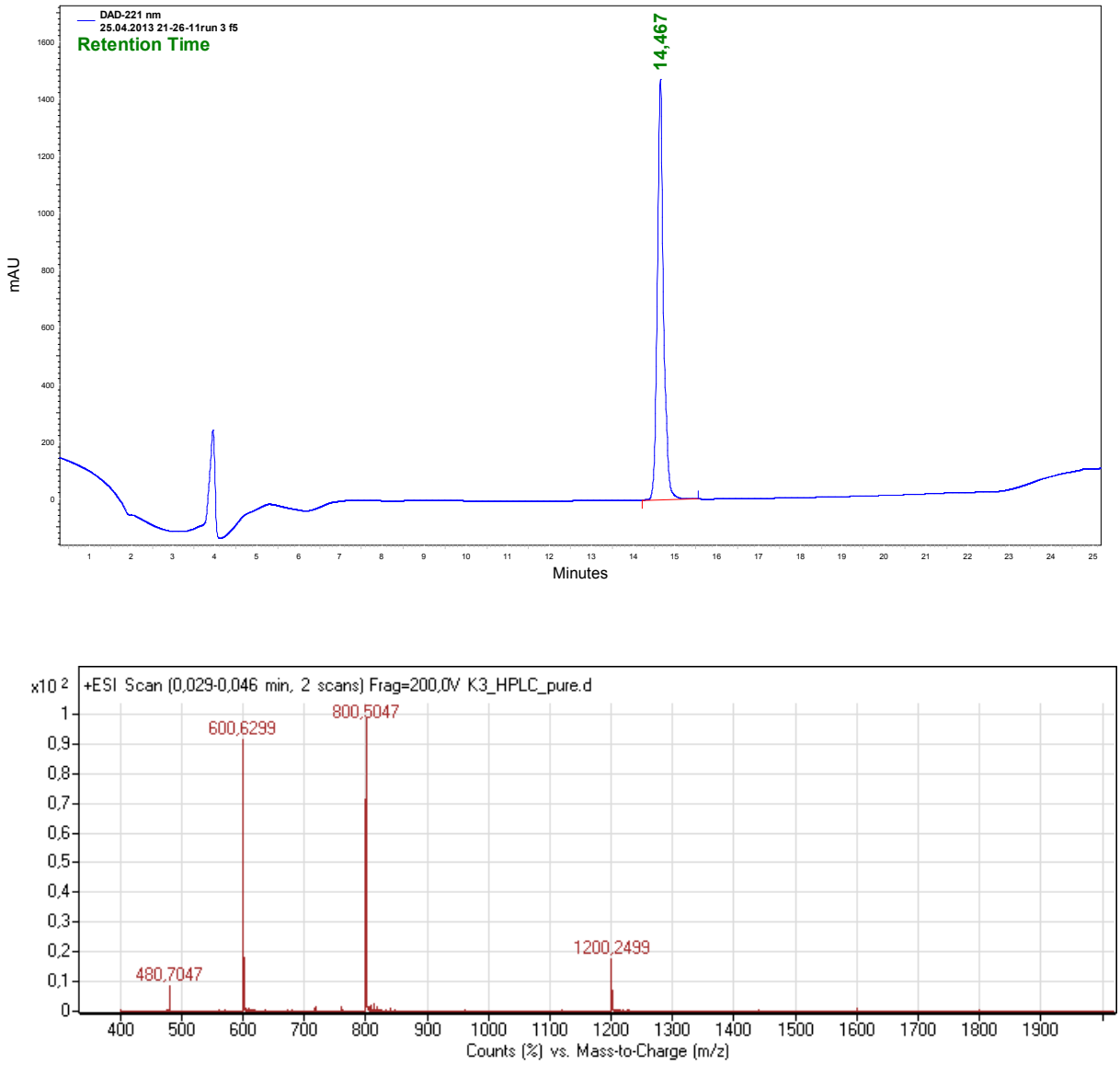
# Appendix: ESI-Tof and RP-HPLC peptide analysis

## 10.1 E3: $H_2N$-Abz-EIAALEKEIAALEKEIAALEK-OH



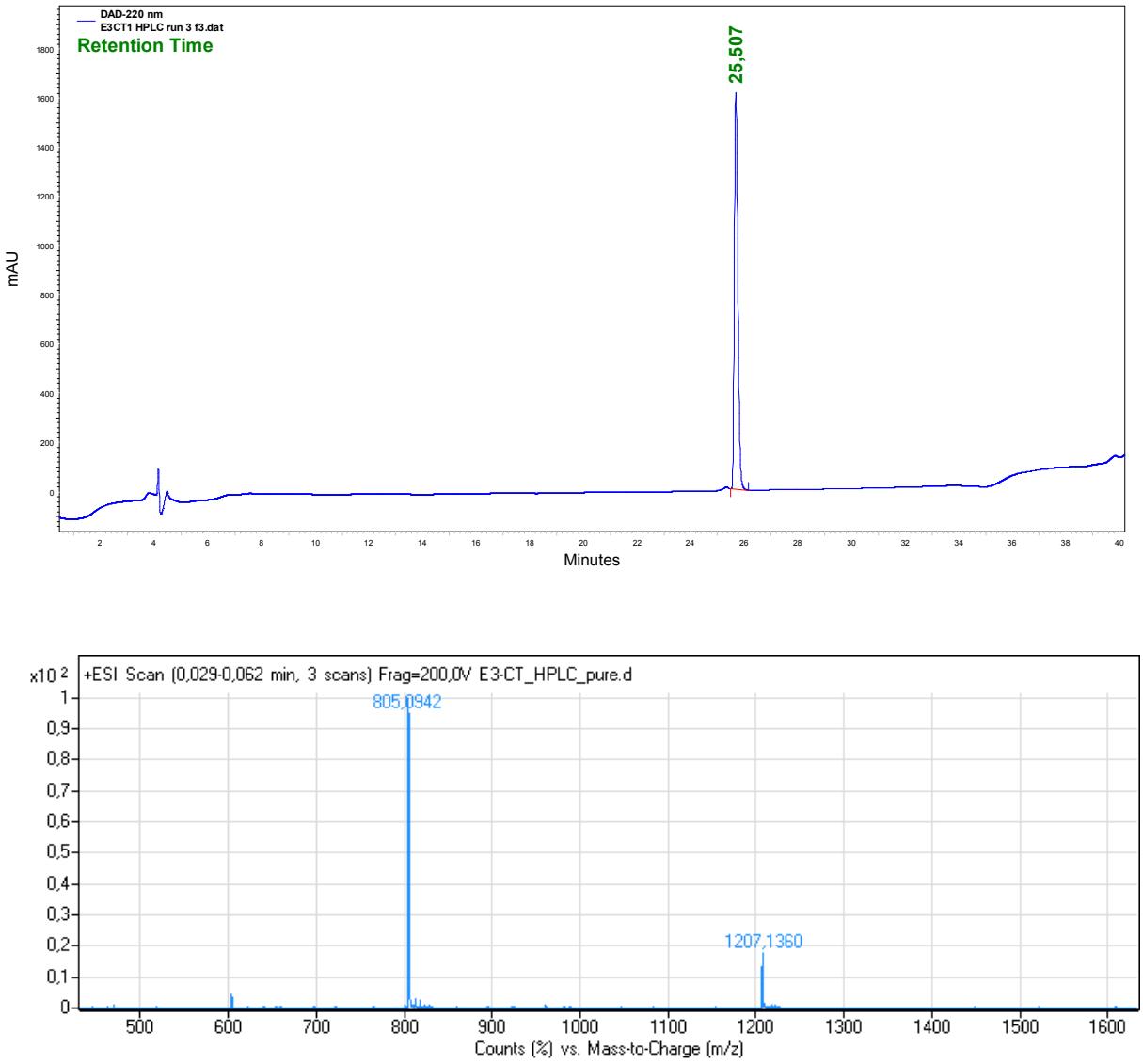

**App. 10.1:** (Top) RP-HPLC of peptide **E3** using a linear gradient of solvent A (95% $H_2O$, 5% MeCN, 0.1% TFA) and solvent B (100% MeCN, 0.1% TFA) over 18 min with Phenomenex® Kinetix C18 column. (Bottom) ESI-ToF mass spectrum of peptide **E3**.

## 10.2 K3: $H_2N$-Abz-KIAALKEKIAALKEKIAALKE-OH



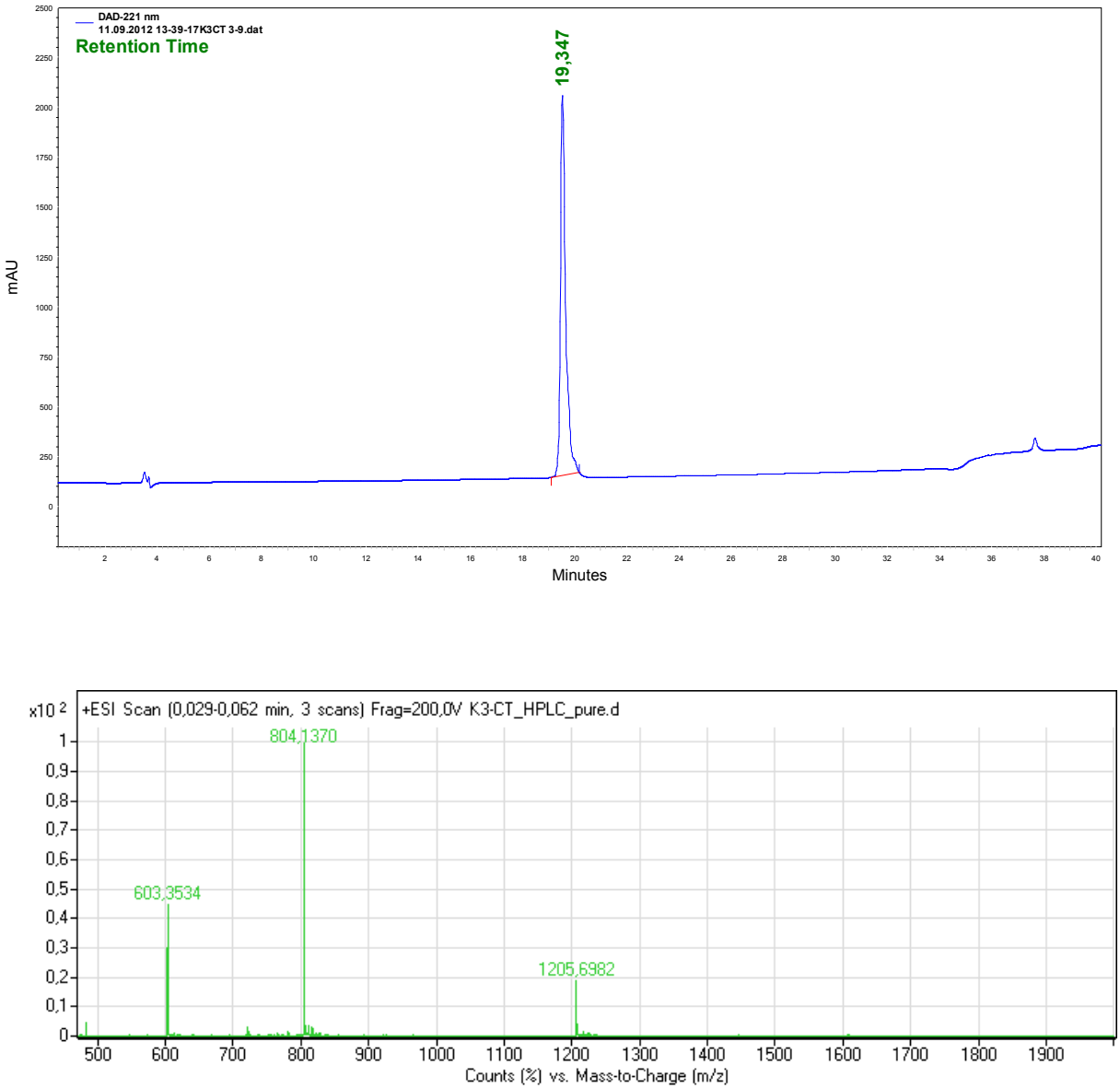

**App. 10.2:** (Top) RP-HPLC of peptide **K3** using a linear gradient of solvent A (95% $H_2O$, 5% MeCN, 0.1% TFA) and solvent B (70% MeCN, 30% $H_2O$, 0.1% TFA) over 18 min with Phenomenex® Kinetix C18 column. (Bottom) ESI-ToF mass spectrum of peptide **K3**.

## 10.3 E3CT: $H_2N$-Abz-EIAALEDEIAHLEKSIAALEK-OH



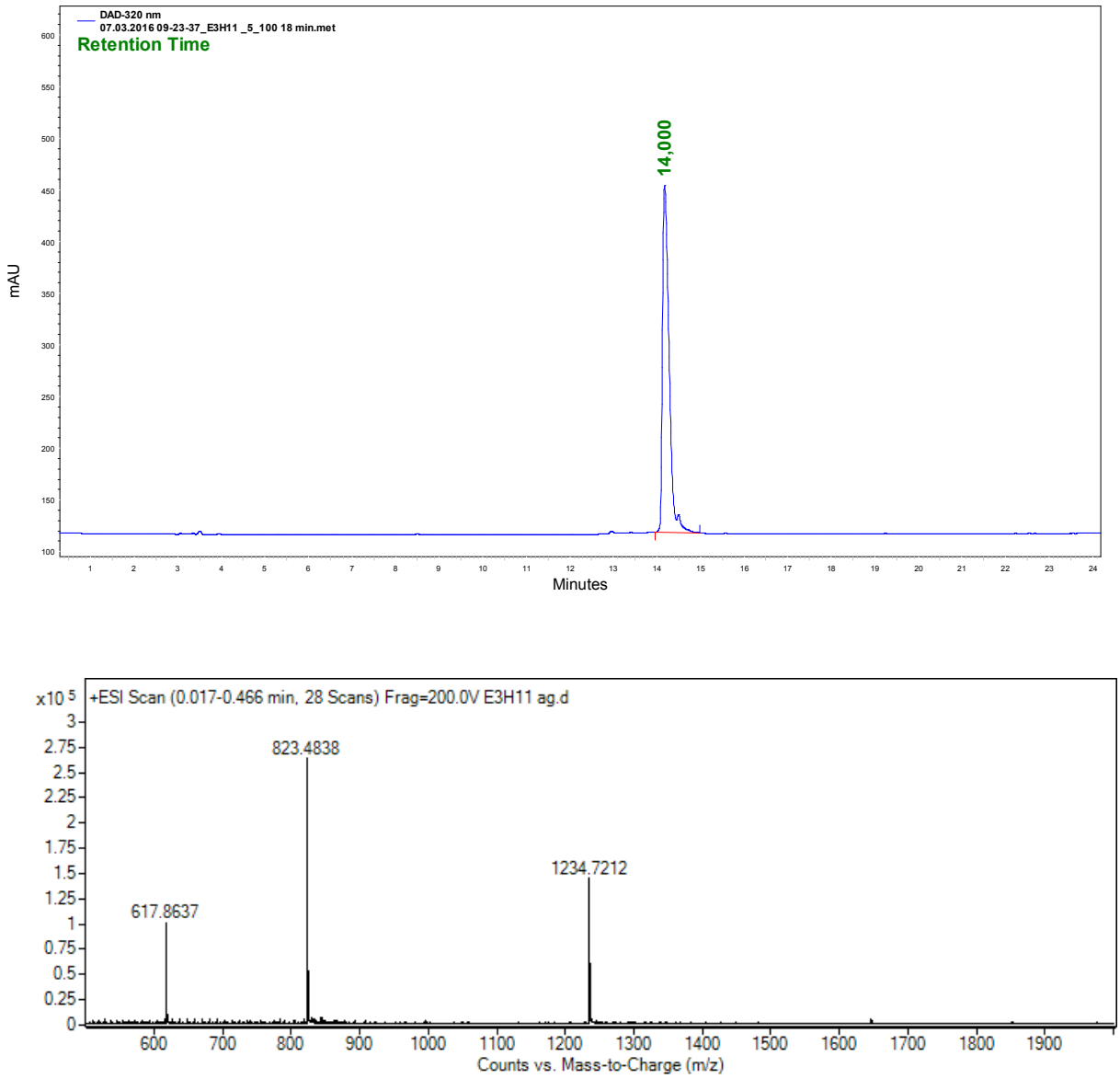

**App. 10.3:** (Top) RP-HPLC of peptide **E3CT** using a linear gradient of solvent A (95% $H_2O$, 5% MeCN, 0.1% TFA) and solvent B (70% MeCN, 30% $H_2O$, 0.1% TFA) over 30 min with Phenomenex® Kinetix C18 column. (Bottom) ESI-ToF mass spectrum of peptide **E3CT**.

## 10.4 K3CT: $H_2N$-Abz-KIAALKDKIAHLKESIAALKE-OH



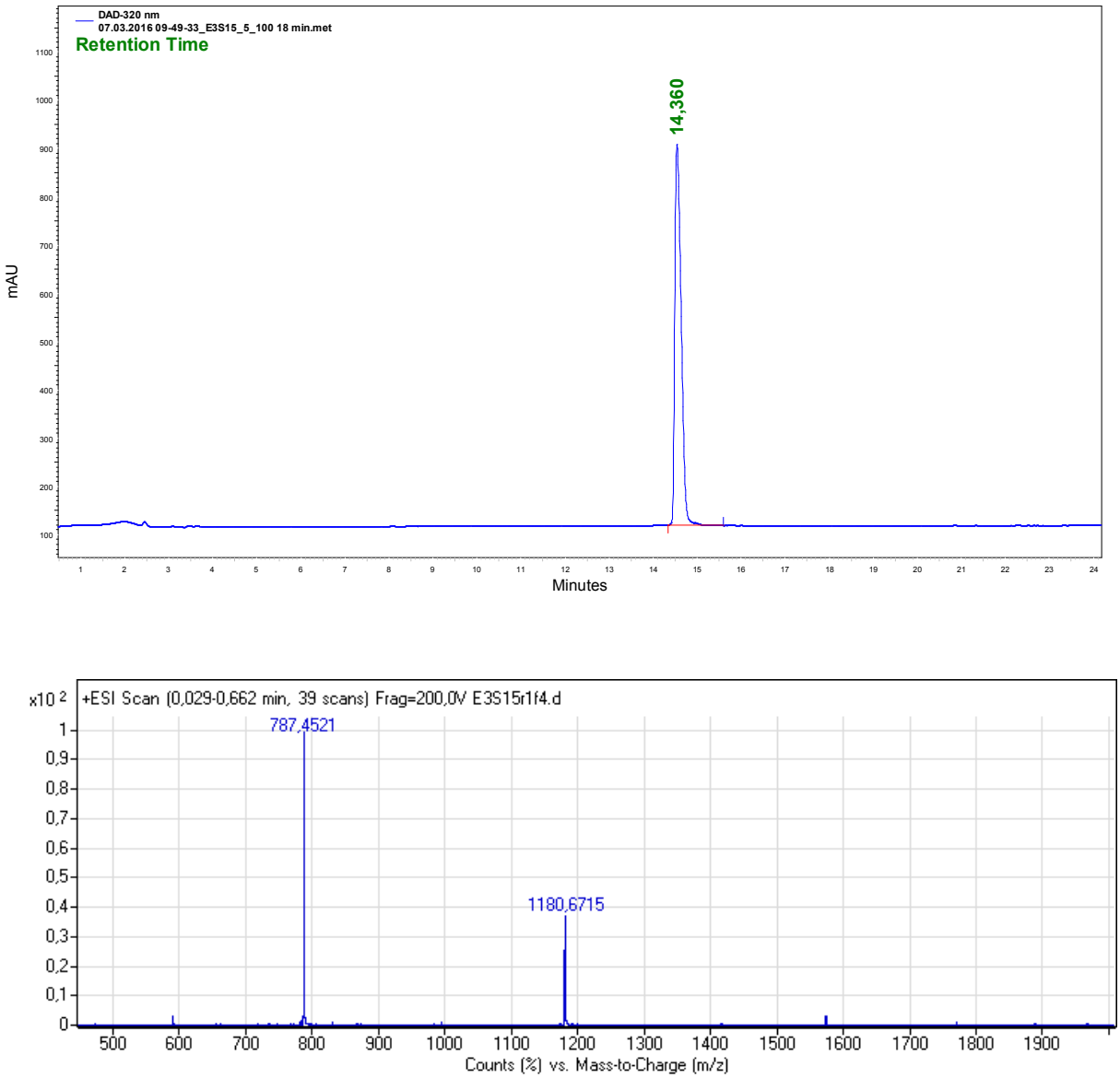

**App. 10.4:** (Top) RP-HPLC of peptide **K3CT** using a linear gradient of solvent A (95% $H_2O$, 5% MeCN, 0.1% TFA) and solvent B (70% MeCN, 30% $H_2O$, 0.1% TFA) over 30 min with Phenomenex® Kinetix C18 column. (Bottom) ESI-ToF mass spectrum of peptide **K3CT**.

## 10.5 E3H11: $H_2N$-Abz-EIAALEKEIAHLEKEIAALEK-OH



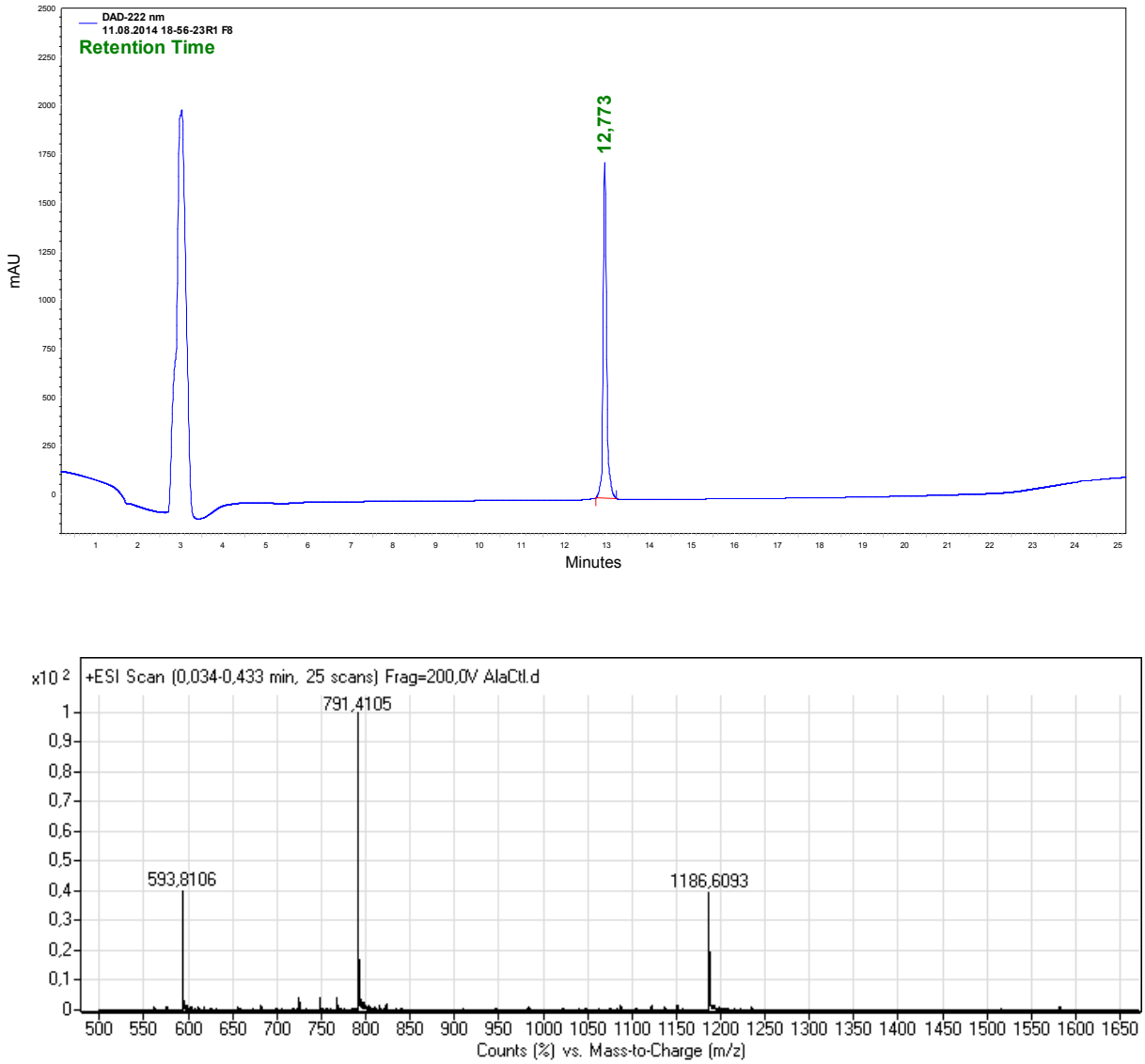

**App. 10.5:** (Top) RP-HPLC of peptide **E3H11** using a linear gradient of solvent A (95% $H_2O$, 5% MeCN, 0.1% TFA) and solvent B (100% MeCN, 0.1% TFA) over 18 min with Phenomenex® Kinetix C18 column. (Bottom) ESI-ToF mass spectrum of peptide **E3H11**.

## 10.6 E3S15: $H_2N$-Abz-EIAALEKEIAALEKSIAALEK-OH



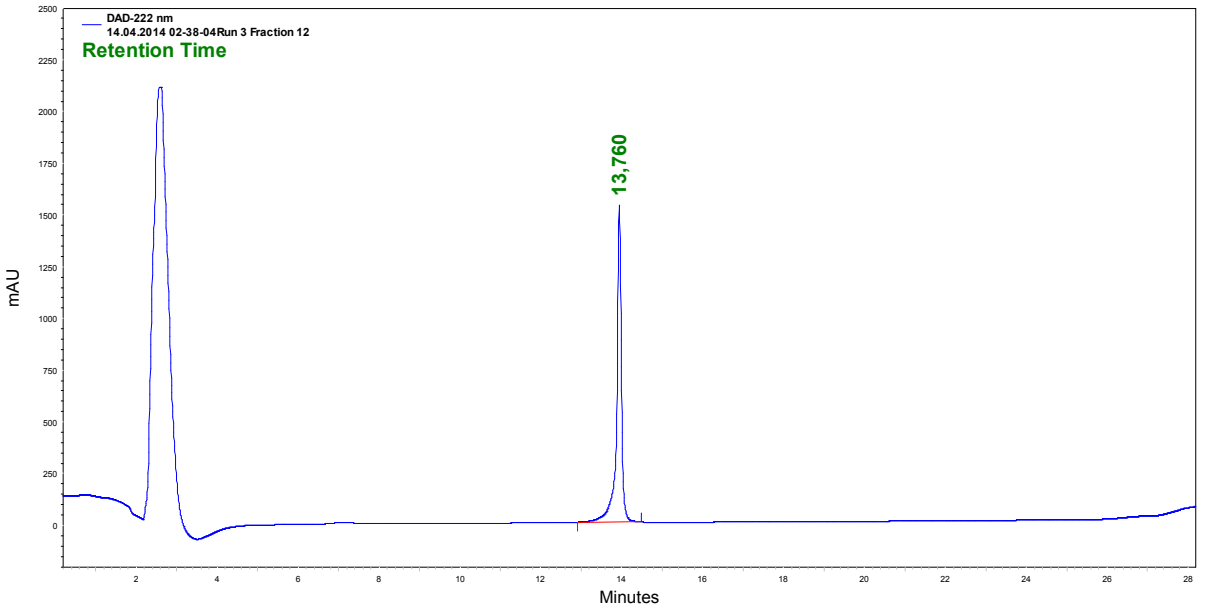

**App. 10.6:** (Top) RP-HPLC of peptide **E3S15** using a linear gradient of solvent A (95% $H_2O$, 5% MeCN, 0.1% TFA) and solvent B (100% MeCN, 0.1% TFA) over 18 min with Phenomenex® Kinetix C18 column. (Bottom) ESI-ToF mass spectrum of peptide **E3S15**.

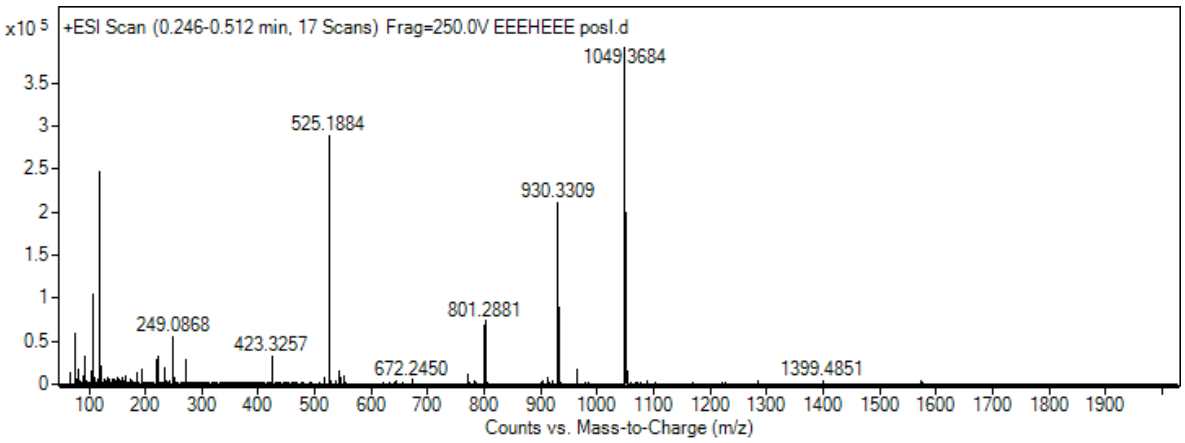
## 10.7 E3H11-Ala: $H_2N$-Abz-KAEAAAAEAEAAHAEAAAAEAK-OH




**App. 10.7:** (Top) RP-HPLC of peptide **E3H11-Ala** using a linear gradient of solvent A (95% $H_2O$, 5% MeCN, 0.1% TFA) and solvent B (70% MeCN, 30% $H_2O$, 0.1% TFA) over 18 min with Phenomenex® Kinetix C18 column. (Bottom) ESI-ToF mass spectrum of peptide **E3H11-Ala**.

## 10.8 EEEHEEE: $H_2N$-Abz-EEEHEEE-OH

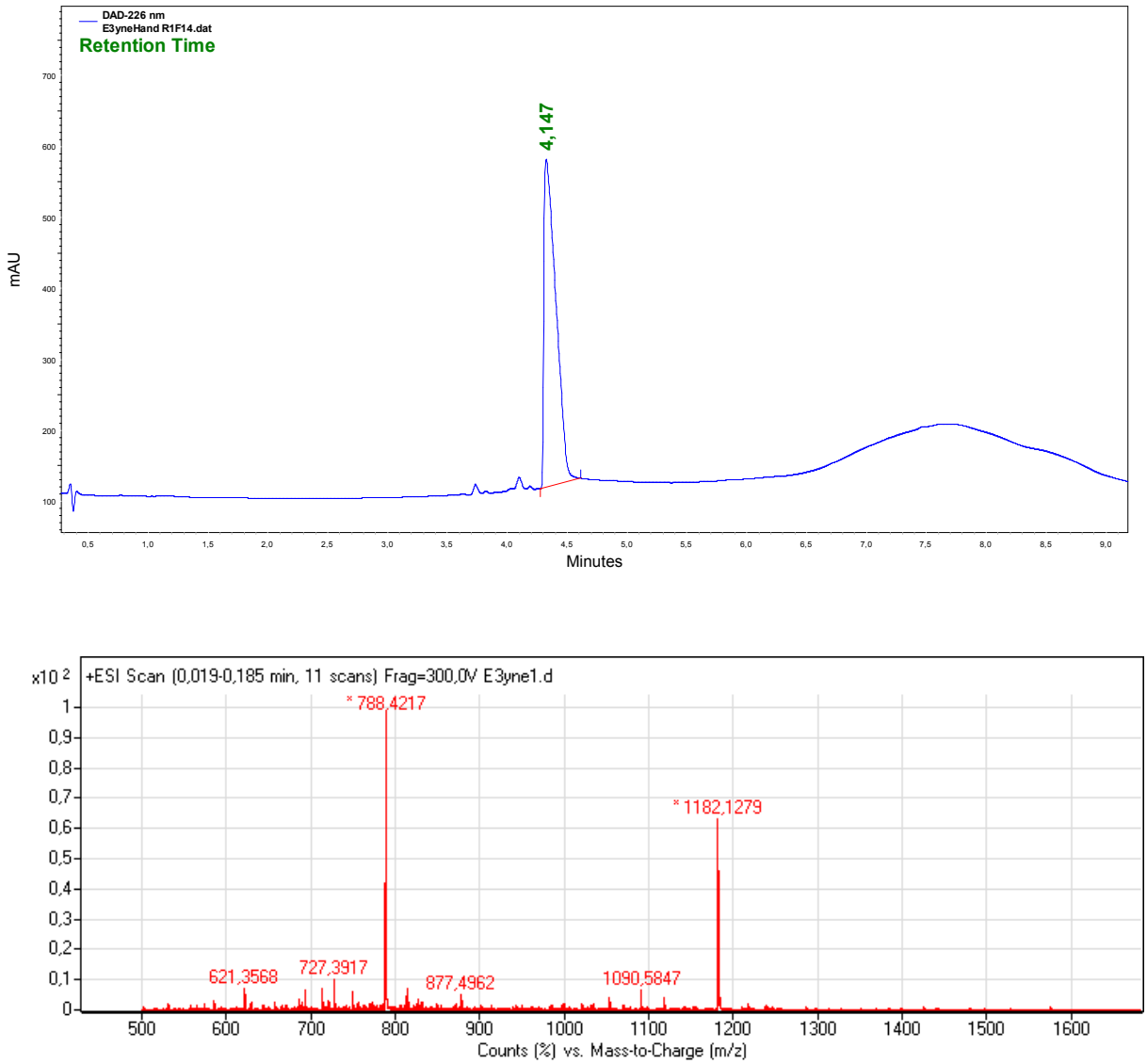



**App. 10.8:** RP-HPLC of peptide **EEEHEEE** using a linear gradient of solvent A (100% $H_2O$, 0.1% TFA) and solvent B (55% $H_2O$, 45% MeCN, 0.1% TFA) over 22 min with Phenomenex® Kinetix C18 column. (Bottom) ESI-ToF mass spectrum of peptide **EEEHEEE**.

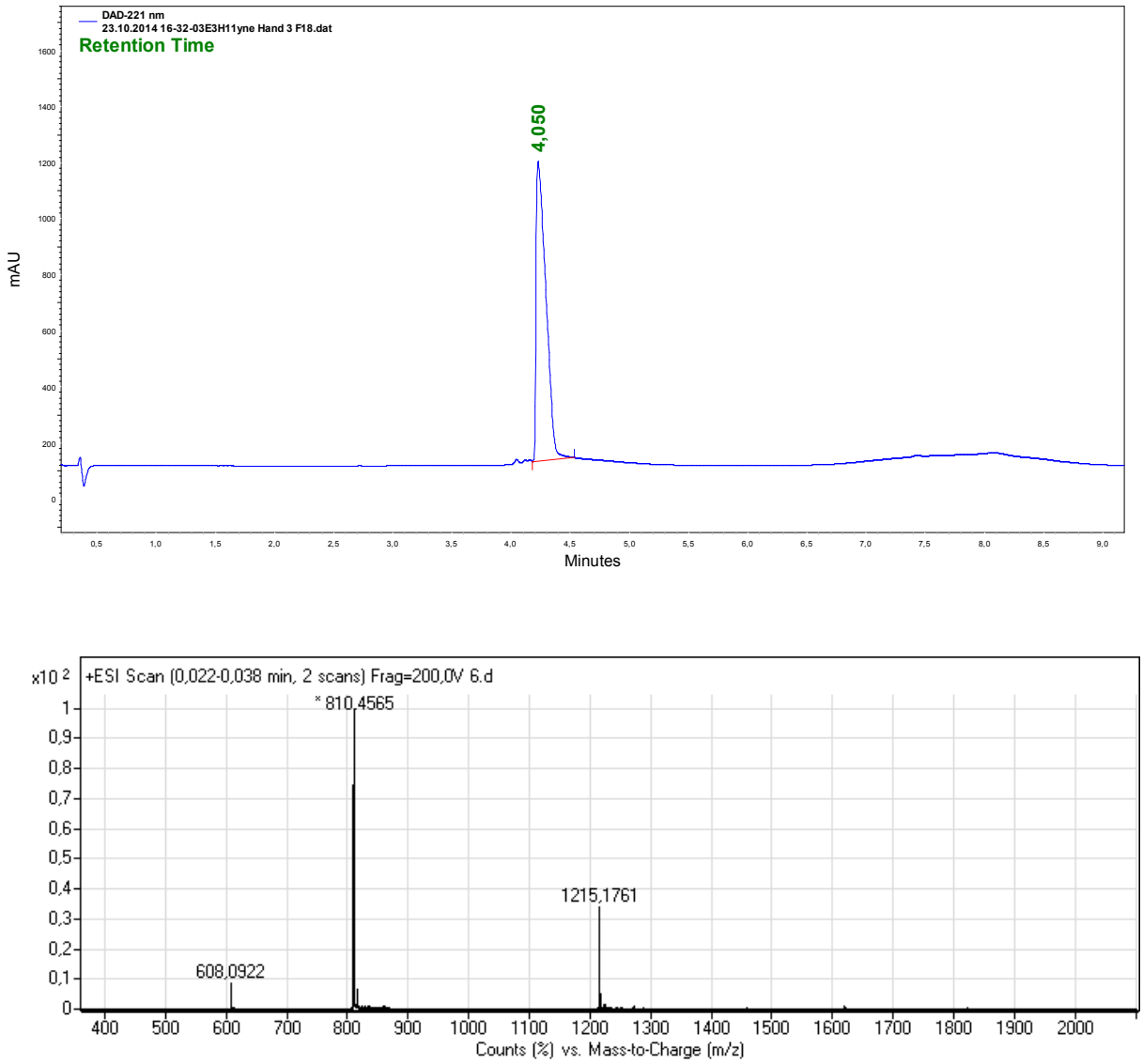
.

## 10.9 E3yne: Alkyne functionalized E3



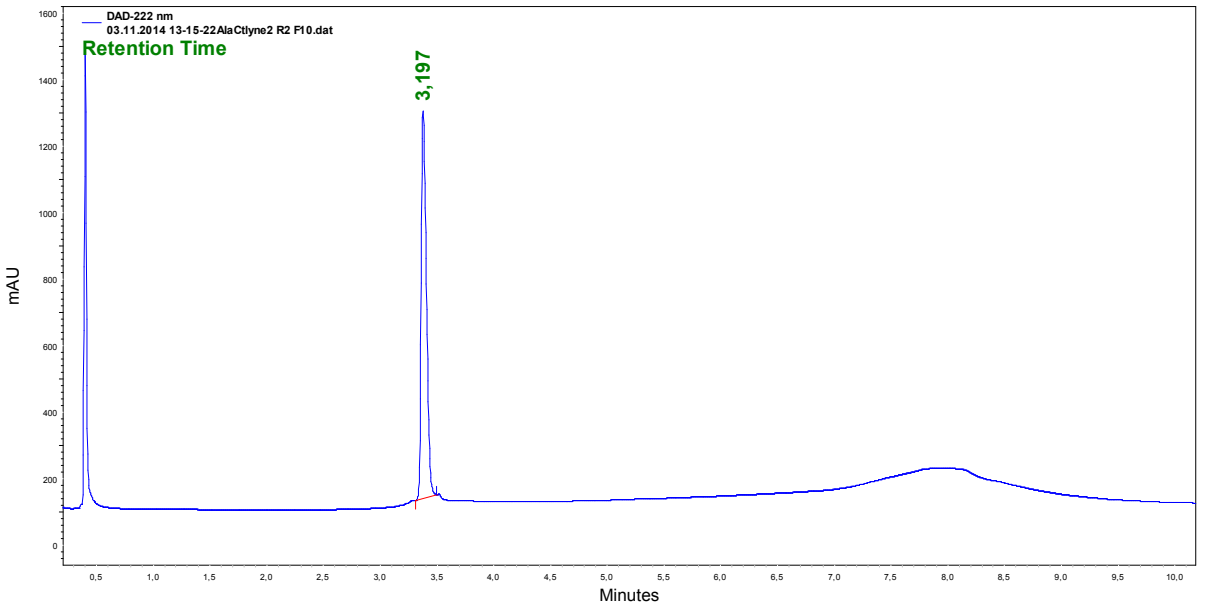

**App. 10.9:** (Top) RP-HPLC of peptide **E3yne** using a linear gradient of solvent A (70% $H_2O$, 30% MeCN, 0.1% TFA) and solvent B (70% MeCN, 30% $H_2O$, 0.1% TFA) over 5 min with column Purospher® STAR C18. (Bottom) ESI-ToF mass spectrum of peptide **E3yne**.

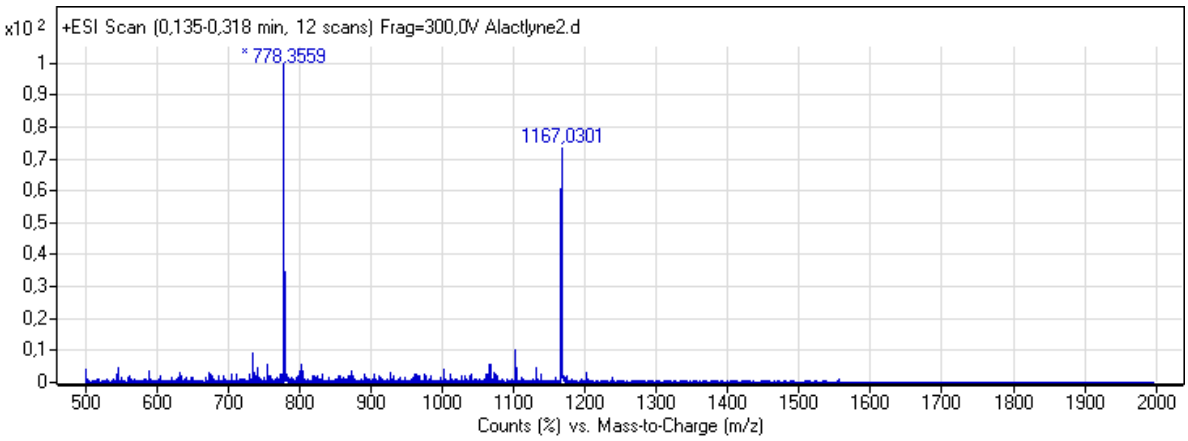
## 10.10 E3H11yne: Alkyne functionalized E3H11




**App. 10.10:** (Top) RP-HPLC of peptide **E3H11yne** using a linear gradient of solvent A (70% $H_2O$, 30% MeCN, 0.1% TFA) and solvent B (70% MeCN, 30% $H_2O$, 0.1% TFA) over 5 min with column Purospher® STAR C18. (Bottom) ESI-ToF mass spectrum of peptide **E3H11yne**.

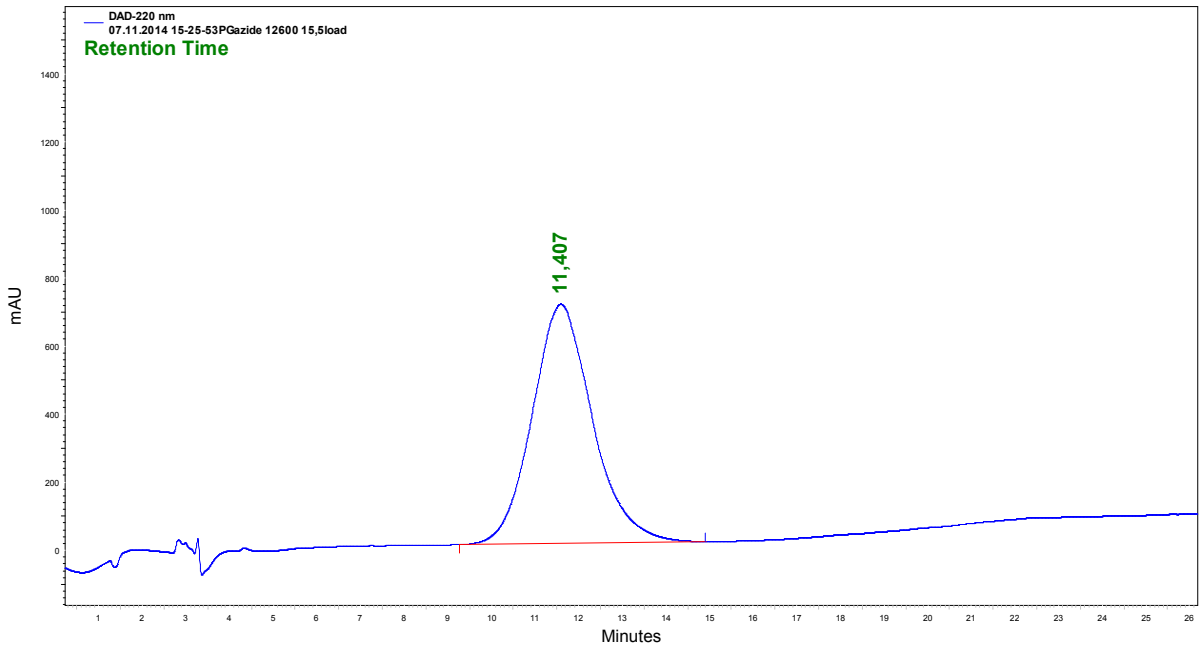
## 10.11 E3H11-Alayne: Alkyne functionalized E3H11-Ala



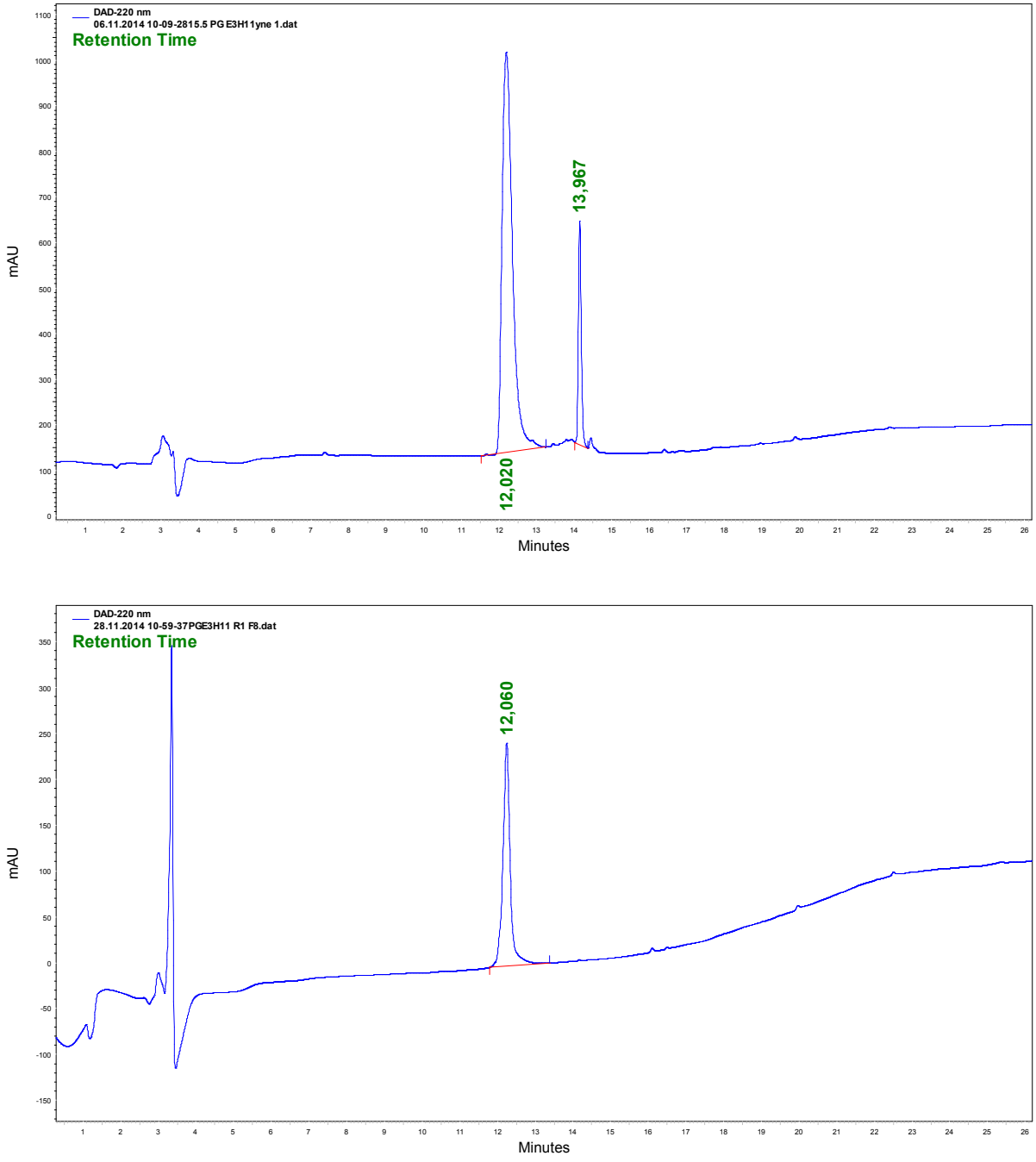

**App. 10.11:** (Top) RP-HPLC of peptide **E3H11-Alayne** using a linear gradient of solvent A (80% $H_2O$, 20% MeCN, 0.1% TFA) and solvent B (70% MeCN, 30% $H_2O$, 0.1% TFA) over 6 min with column Purospher® STAR C18. (Bottom) ESI-ToF mass spectrum of peptide **E3H11-Alayne**.

## 10.12 hPG-Azide: Azide-functionalized hyper-branched polyglycerol



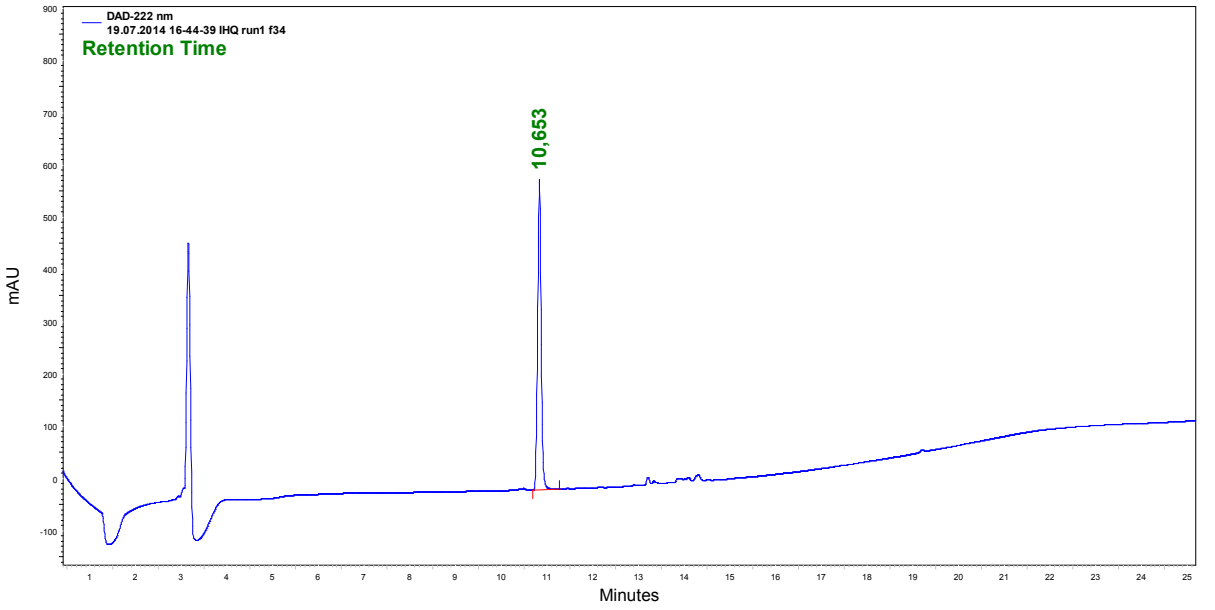

**App. 10.12:** RP-HPLC of azide-functionalized (15.5%) hyper-branched polyglycerol (Mn = 13.2 kD) using a linear gradient of solvent A (95% $H_2O$, 5% MeCN, 0.1% TFA) and solvent B (100% MeCN, 0.1% TFA) over 18 min with Phenomenex® Kinetix C18 column.

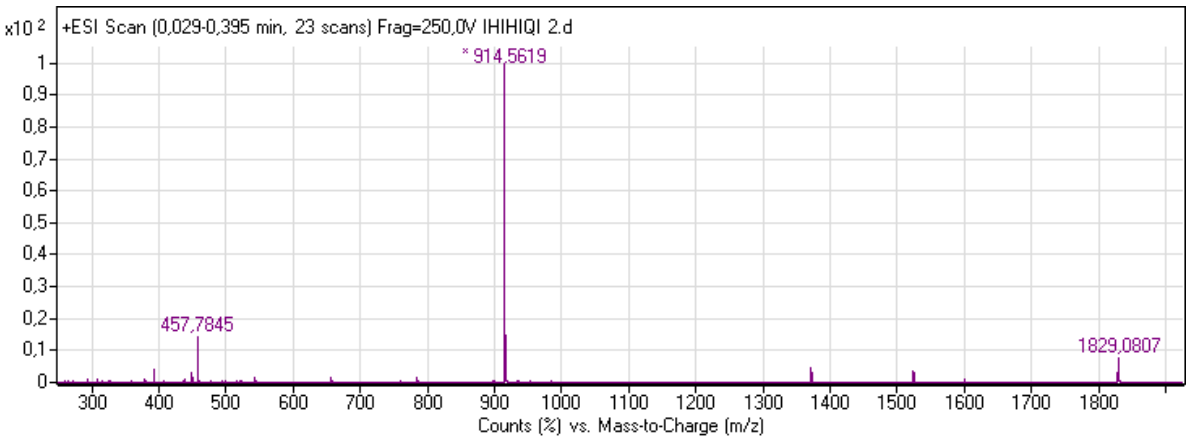
## 10.13 Immobilization of E3H11yne to hPG-Azide




**App. 10.13:** (Top) RP-HPLC of **hPG-E3H11** and residual **E3H11yne** following peptide immobilization and dialysis and (bottom) **hPG-E3H11** after preparative RP-HPLC. Chromatograms were obtained using a linear gradient of solvent A (95% $H_2O$, 5% MeCN, 0.1% TFA) and solvent B (100% MeCN, 0.1% TFA) over 18 min with Phenomenex® Kinetix C18 column.

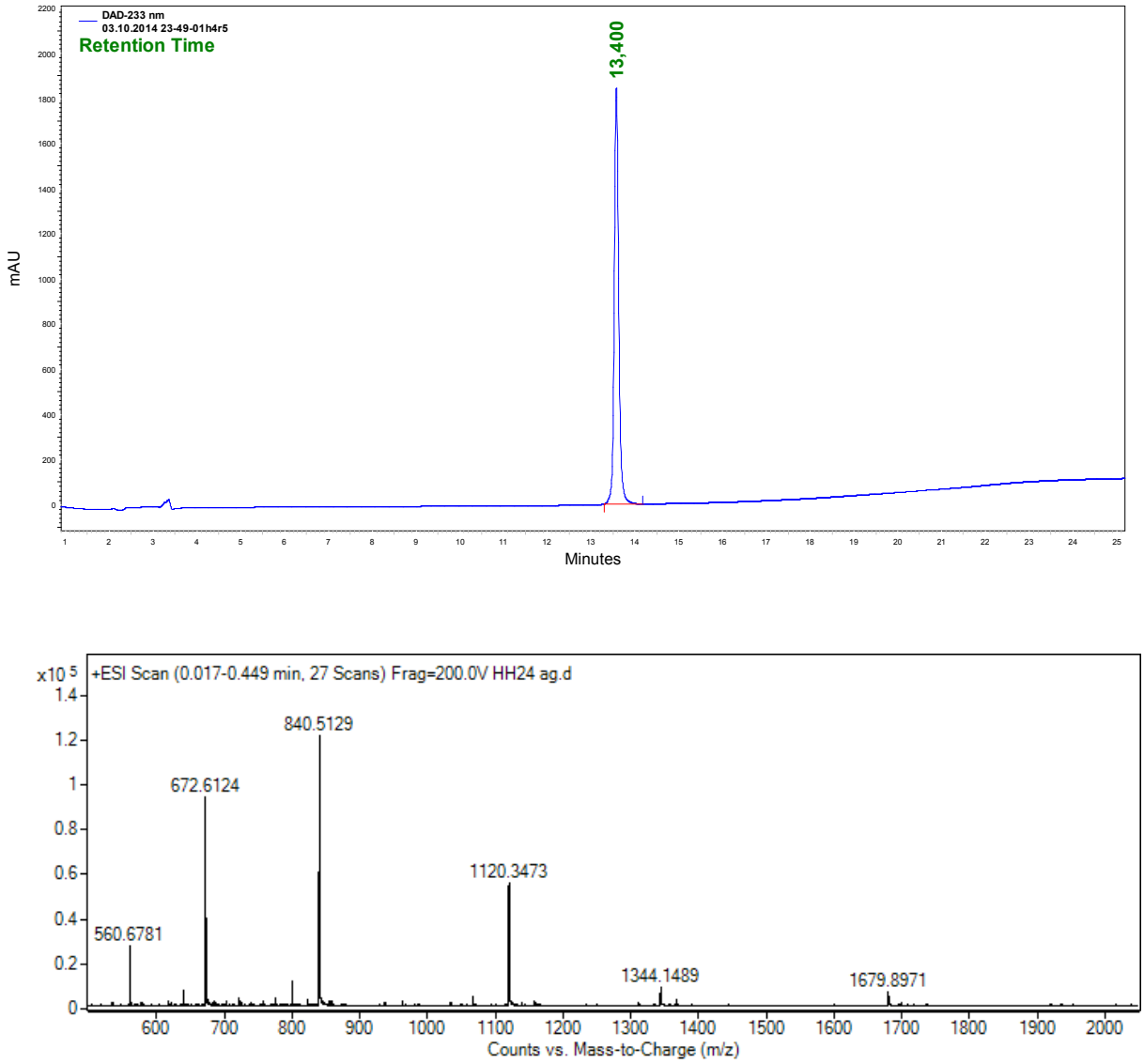
## 10.14 Ac-IHIHIQI-$NH_2$



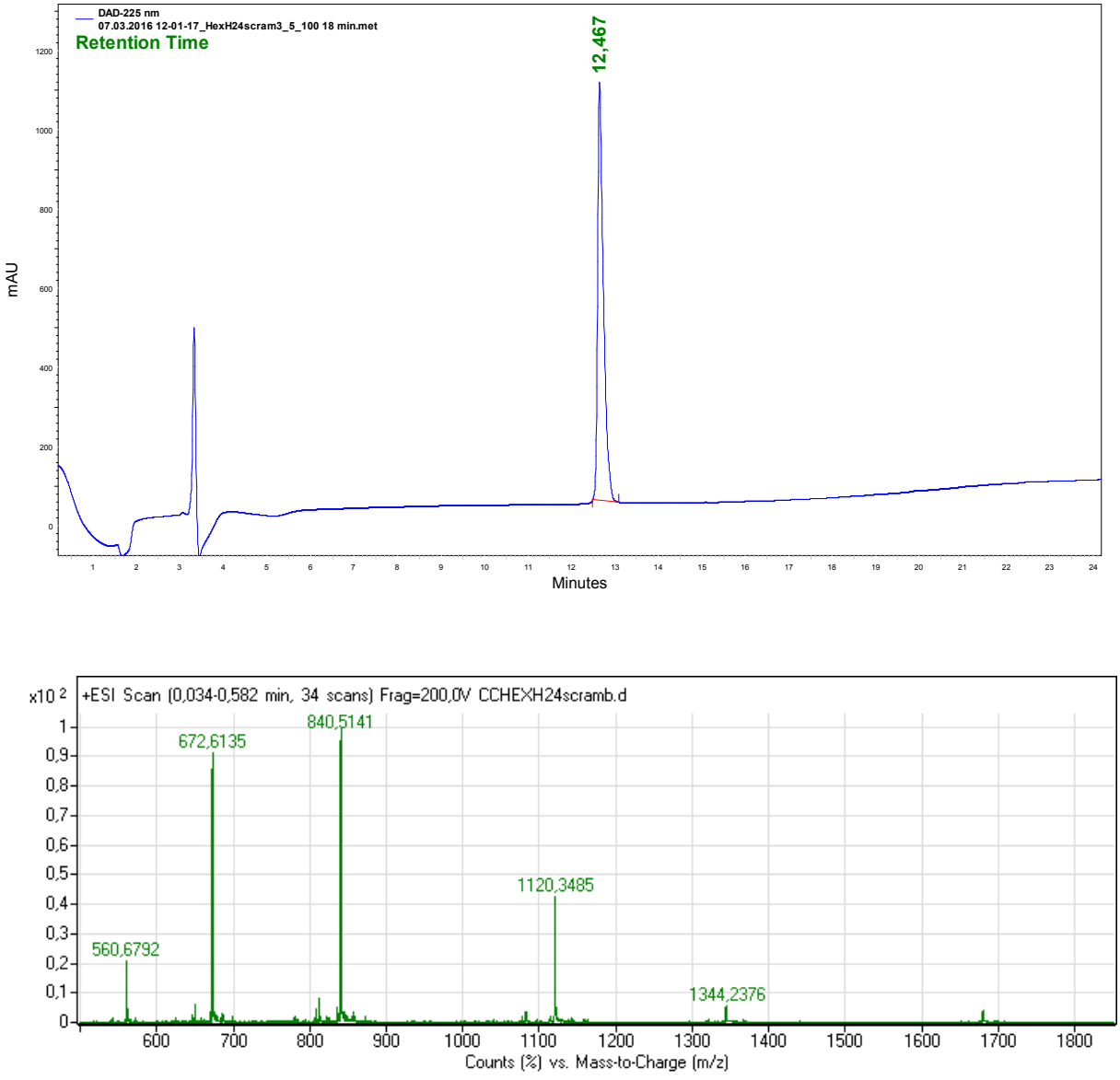

**App. 10.14:** (Top) RP-HPLC of peptide **Ac-IHIHIQI-$NH_2$** using a linear gradient of solvent A (90% $H_2O$, 10% MeCN, 0.1% TFA) and solvent B (100% MeCN, 0.1% TFA) over 18 min with Phenomenex® Kinetix C18 column. (Bottom) ESI-ToF mass spectrum of peptide **Ac-IHIHIQI-$NH_2$**.

## 10.15 CC-HexH24: $H_2N$-GELKAIAQELKAIAKELKAIAWEHKAIAQGAG-OH



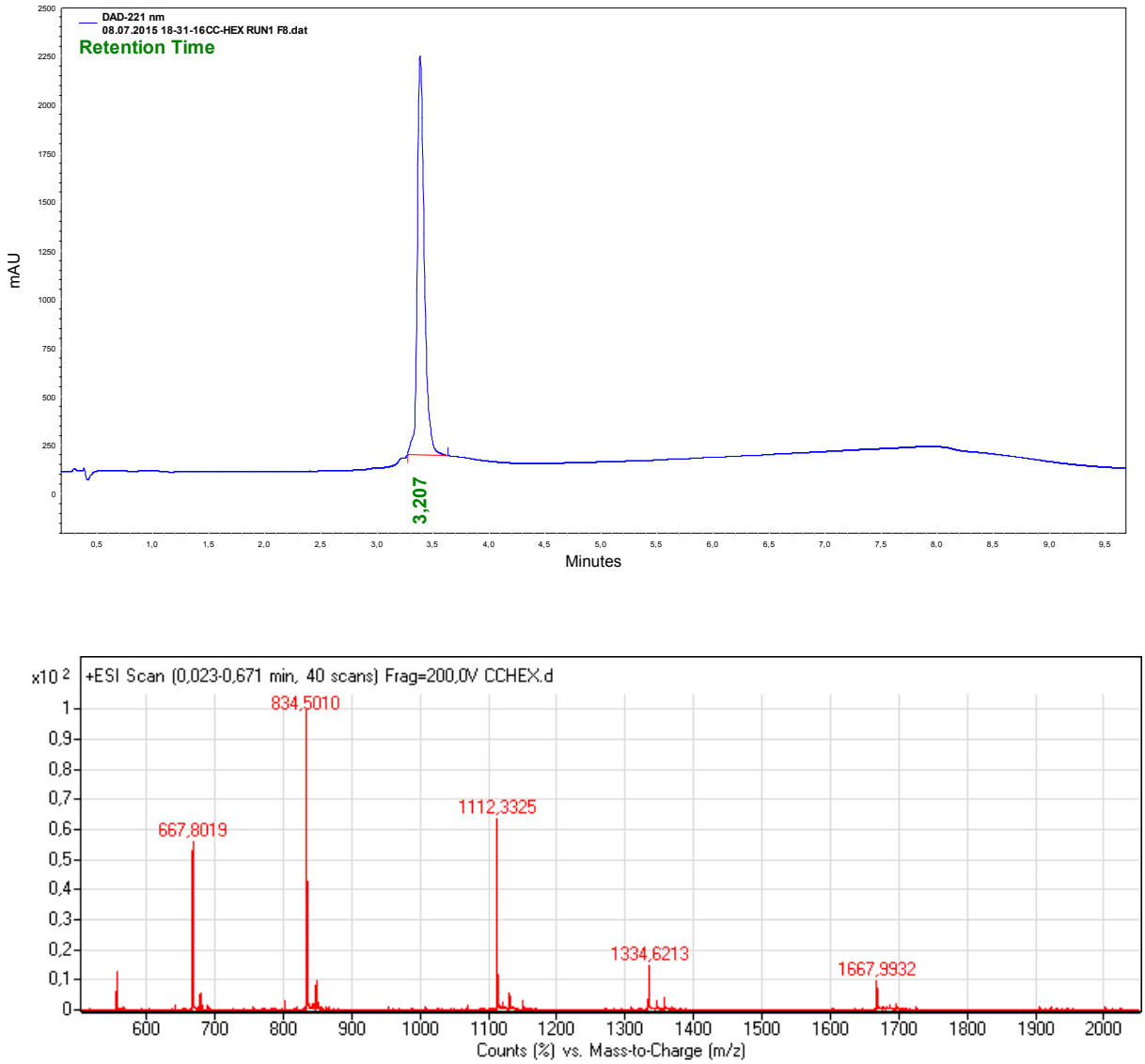

**App. 10.15:** (Top) RP-HPLC of peptide **CC-HexH24** using a linear gradient of solvent A (95% $H_2O$, 5% MeCN, 0.1% TFA) and solvent B (95% MeCN, 5% $H_2O$, 0.1% TFA) over 18 min with Phenomenex® Kinetix C18 column. (Bottom) ESI-ToF mass spectrum of peptide **CC-HexH24**.

## 10.16. CC-HexH24 scramble: $H_2N$-EEKAHKLGAKIGLIQEAIAAKAWEAALQAIKG-OH




**App. 10.16:** (Top) RP-HPLC of peptide **CC-HexH24 scramble** using a linear gradient of solvent A (95% $H_2O$, 5% MeCN, 0.1% TFA) and solvent B (100 % MeCN, 0.1% TFA) over 10 min with column Purospher® STAR C18. (Bottom) ESI-ToF mass spectrum of peptide **CC-HexH24 scramble**.

## 10.17 CC-Hex: $H_2N$-GELKAIAQELKAIAKELKAIAWELKAIAQGAG-OH




**App. 10.17:** (Top) RP-HPLC of peptide **CC-Hex** using a linear gradient of solvent A (60% $H_2O$, 40% MeCN, 0.1% TFA) and solvent B (100% MeCN, 0.1% TFA) over 6 min with column Purospher® STAR C18. (Bottom) ESI-ToF mass spectrum of peptide **CC-Hex**.