Reconstruction of the Emergence of New Enzyme Functions by Protein Engineering: Examples From Secondary and Xenobiotic Metabolism

DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.) DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN DER UNIVERSITÄT REGENSBURG

vorgelegt von
Markus Richard Busch

aus
Landshut

im Jahr
2022

Das Promotionsgesuch wurde eingereicht am:
22.08.2022

Die Arbeit wurde angeleitet von:
Prof. Dr. Reinhard Sterner

Unterschrift:

Contents

List of Figures

List of Tables

List of Abbreviations

α-KG	$alpha$-ketoglutarate
ACP	Aryl carrier protein
Amp	Ampicillin
ATP	Adenosine triphosphate
BPD	2-2'-Bipyridine
Cam	Chloramphenicol
CAS	Chromeazurol S
CD	Circular dichroism
CT	6-Chloro-1,3,5-triazine-2,4(1H,3H)-dione
2,3-DHB	2,3-Dihydroxybenzoate
2,3-DHDHB	2,3-Dihydro-2,3-dihydroxybenzoate
dsDNA	Double standed DNA
CH	Chorismate
cPCR	Colony PCR
CV	Column volume
DMG	N^2,N^2-Dimethylguanine
EDTA	Ethylenediaminetetraacetic acid
EtA	N-ethylammelide
FMN	Flavin mononucleotide
GDH	Glutamate dehydrogenase
Gua	Guanine
HA	Hydroxyatrazine
HPLC	High performance liquid chromatography
IAD	Innovation-Amplification-Divergence
IHL	Isochorismatase-like hydrolase
IMAC	Immobilized metal ion affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IsoA	N-isopropylammelide
LUCA	Last universal common ancestor
MSA	Multiple sequence alignment
MW	Molecular weight
PP	Potassium phosphate
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
$NADP^+$	Nicotinamide adenine dinucleotide phosphate (oxidized form)
NADH	Nicotinamide adenine dinucleotide (reduced form)
NAD^+	Nicotinamide adenine dinucleotide (oxidized form)
N-Et	Ethylamine

N-Iso	Isopropylamine
N-Me_2	Dimethylamine
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
RBS	Ribosome binding site
RT	Room temperature
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
SLS	Swiss Light Source
ssDNA	Single stranded DNA
TEA	triethylammonium acetate
UAc	(Z)-3-Ureidoacrylate
UPro	Ureidopropionate

1 Introduction

1.1 Principles of Enzyme Evolution

Evolutionary fitness of a certain organism is, to a significant part, determined by its metabolic capabilities, which, in turn, are largely defined by the enzymes coded for by its genome. Enzyme evolution, the gradual change of enzymes towards novel functional features, thus constitutes a significant factor in evolution as a whole.
Three central processes drive the emergence and spread of novel enzymatic functions. New enzymes can arise either *de novo* from a previously non-coding DNA sequence [1] [2] or through divergence of a pre-existing gene. After any such event, the novel enzymatic function can be spread to other species by horizontal gene transfer (HGT) [3]. In recent times, enzyme evolution *via* duplication and divergence of existing genes is likely the most important mechanism by which new catalytic functions have evolved [4], plausibly due to the enormous number of highly diverse pre-existing enzymes [5] which can potentially serve as progenitors for novel ones.
Various evolutionary models exist [6] [7] [8] that attempt to explain the emergence of new enzymes from pre-existing ancestors and particularly how the trade-off between functional innovation and increased metabolic cost can be balanced in a biological system. The most sophisticated of these models is the Innovation-Amplification-Divergence (IAD) model, proposed by Bergthorsson and co-workers [9] (see Figure 1.1). This model poses that a specific gene randomly acquires a novel function (innovation) while maintaining its original one. If the novel function is beneficial to the organism, gene duplication is favored and multiple gene copies are maintained (amplification). Some gene copies are then free to acquire mutations that lead to enhancement of and specialization for the novel gene function (diversification) while the original function is retained in other copies.

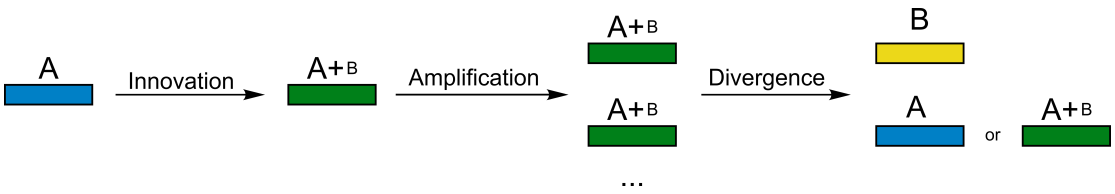


Figure 1.1: The Innovation-Amplification-Divergence model
Proposed by Bergthorsson and co-workers [9], the Innovation-Amplification-Divergence (IAD) model describes the emergence of new genes by a mechanism in which each of the steps is favored by natural selection. An ancestral gene (blue) with a physiologically relevant function A randomly acquires a secondary function B (innovation). This bi-functional gene (green; A+B) maintains the original function and provides a selection advantage through its novel function. Function B is initially enhanced by gene duplication (amplification) leading to multiple copies of the bi-functional gene. Duplication removes the selection pressure for maintenance of the original function, allowing for specialization to occur in the different gene copies (divergence). This process ultimately leads to a new gene, exclusively possessing the new function B (yellow). The original function A is maintained in the form of a bi-functional gene which may additionally undergo loss-of-function of the secondary function B.

In the context of enzyme evolution, the gene functions discussed in the IAD model equate to distinct catalytic activities. Accordingly, the process of innovation describes the emergence of a promiscuous side activity for a different than the native substrate. Promiscuity is actually a common feature among enzymes [10] [11] [12] implying significant pre-existing potential for enzyme evolution in Nature.
Gene duplication, underlying the amplification step of the IAD model, is relatively common [13] [14] and notably occurs more often than point mutations per nucleotide [15]. Typically, gene duplications are selected against, due to the increased cost of duplicate gene expression and resulting metabolic imbalance [15]. However, the resulting increase in gene dosage may also be beneficial in the case of a previously limiting gene function, leading to long-term maintenance of the duplicated gene copies [16]. Similarly, subfunctionalization, specific loss-of-function mutations in each of the gene copies, may preserve gene duplications [17].
The process of divergence on the enzyme level can proceed either through modifications to substrate binding or catalytic residues. Continuous diversification of enzymes over the course of evolution has lead to the emergence of superfamilies of homologous enzymes catalyzing chemically similar reactions and of functionally distinct suprafamilies which consist of homologous enzymes with different reactivities acting on structurally similar substrates [18]. Furnham and co-workers [19] have reconstructed the ancestral functions of a number of enzymes at the level of enzyme commission (EC) numbers. Their findings indicate that most functional changes in evolution were within the same EC class but changes across classes were also relatively common.

1.2 Scope of This Work

The central aim of this project was to understand how new enzyme functions can arise from pre-existing enzymes. Specifically, it was to be attempted to establish novel enzymatic activities or improve existing side activities in certain enzyme scaffolds by a limited number of residue exchanges. This process, mimicking natural evolution, would provide evolutionary trajectories which might indicate the molecular mechanisms underlying the functional conversion of the respective enzyme. These findings would reveal how new enzyme functions could have emerged in Nature over the course of natural evolution and could further be exploited for the targeted design of specific desired enzyme functions. This work can be separated into three main projects, which are described in detail below.

1.2.1 Evolutionary Origin of the Enterobactin Pathway Proteins EntB and EntA

The primary metabolic enzymes ureidoacrylate amidohydrolase (RutB) and 3-oxoacyl-[acyl-carrier-protein] reductase (FabG) were supposed to be converted into variants with the secondary metabolic activities of isochorismatase (EntB) and 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EntA), respectively. As rational design strategies had previously failed [20], a directed evolution approach was employed here. To this end, a suitable selection system for enterobactin producing strains was to be established. This required the identification of medium conditions in which bio-availability of iron ions was limited by chelating agents to such a degree, that enterobactin non-producing strains were prevented from growth while enterobactin producers were reasonably unimpeded. Full randomization of *rutB* and *fabG* as well as targeted mutagenesis of *rutB* would produce various gene libraries, which would be searched for variants capable of *in vivo* complementation of chromosomal deletions of *entB-IC* and *entA*, respectively. Variants of *rutB* and *fabG* either identified by this selection or independently predicted by bioinformatical methods would then be characterized for their capacity for *in vivo* complementation under controlled conditions as well as for the *in vitro* activities of their gene products. The most active enzyme variants would then have served as starting states for attempts to further increase their respective activity, possibly resulting in evolutionary trajectories indicating the mechanisms by which the enzyme functions of EntB and EntA had originally evolved.

1.2.2 Structural and Functional Characterization of the Ureidoacrylate Amidohydrolase RutB From *E. coli*

The structure and catalytic mechanism of ureidoacrylate amidohydrolase (RutB) were to be investigated in detail to possibly assist the evolutionary investigation of this enzyme (see chapter 3). A protocol for the synthesis and purification of the RutB substrate ureidoacrylate (UAc) was first to be established as this substance is not commercially available. Crystallization conditions that produced high quality RutB protein crystals were to be identified to allow for structure elucidation by X-ray crystallography. Co-crystallization of RutB and catalytically inactive variants of RutB with the substrate analogue ureidopropionate (UPro) and UAc, respectively, was conceptualized as a way to elucidate the respective complex structures of RutB and its variants. This structural information would then inform mutational studies of RutB.
Quantification of the effect of certain amino acid exchanges on the ureidoacrylate amidohydrolase activity of RutB required the establishment of a glutamate dehydrogenase (GDH) coupled assay. Enzyme kinetics of RutB variants would then serve to formulate a detailed catalytic mechanism of the ureidoacrylate amidohydrolase activity of RutB, with the RutB complex structures informing the substrate binding mode and the mutational studies elucidating the role of specific active site residues.

1.2.3 Experimental Reconstruction of the Evolution of the Recently Emerged Hydroxyatrazine Ethylaminohydrolase AtzB

The recent evolution of hydroxyatrazine ethylaminohydrolase (AtzB) was to be investigated in detail to explain how new enzymatic functions can arise within short time scales. This effort was expanding upon a previously described guanine deaminase side activity of AtzB and a rudimentary evolutionary trajectory [21]. Characterization of the closest known homologues of AtzB (AtzB_Hom_Hal and AtzB_Hom_Pleo) was to be undertaken to gain a point of reference for the hypothetical endpoint of the evolutionary trajectory of AtzB. The substrate range of these enzymes was to be investigated by substrate screening and their catalytic activity was to be determined by steady-state enzyme assays further to be established.
Recent mutations, possibly responsible for the establishment of AtzB activity, were to be identified by comparison of multiple sequence alignments of AtzB and the closest sequence homologues. Stepwise introduction of residue exchanges thus selected and characterization of the resulting enzyme variants was conceptualized as a way to establish a plausible evolutionary trajectory linking AtzB to a variant with strong guanine deaminase activity, which would constitute a pseudo-progenitor. Similarly, the inverse residue exchanges of this trajectory

were to be introduced into AtzB_Hom_Hal and AtzB_Hom_Pleo to generate variants of these enzymes with increased AtzB activity. This would generate artificial evolutionary trajectories towards alternative AtzB enzymes not realized by natural evolution.
HPLC based assays would further quantify the specificity of relevant enzyme variants for hydroxyatrazine ethylaminohydrolase or hydroxyatrazine isopropylaminohydrolase activity. Establishing a protocol for the synthesis of the two resulting products N-ethylammelide and N-isopropylammelide was required, as these substances are not commercially available.
The sum of information thus gained would then allow for a specific, mechanistic explanation of the functional conversion between the two enzymatic functions.

2 Methods

2.1 Materials

2.1.1 Devices

CD spectrometer	
Jasco J-815	JASCO GmbH, Groß-Umstadt
Centrifuges and rotors	
Centrifuge 5810R	EPPENDORF, Hamburg
A-4-81 rotor	EPPENDORF, Hamburg
Sorvall RC 2B	DU PONT, Bad Homburg
SS34 rotor	DU PONT, Bad Homburg
Avanti J-26 S XP	BECKMAN COULTER, Krefeld
JLA-8.1000 rotor	BECKMAN COULTER, Krefeld
Centrifuge 5415 D	EPPENDORF, Hamburg
Centrifuge 5415 R	EPPENDORF, Hamburg
Centrifuge 5417 R	EPPENDORF, Hamburg
FPLC columns	
HisTrap FF crude 5 ml	GE HEALTHCARE, München
HisTrap excel 5 ml	GE HEALTHCARE, München
Superdex S75 10/300GL	GE HEALTHCARE, München
Superdex S200 10/300GL	GE HEALTHCARE, München
HiLoad 26/600 Superdex 75 pg	GE HEALTHCARE, München
Mono Q HR 16/10	GE HEALTHCARE, München
FPLC systems	
Amersham FPLC 900er	AMERSHAM, Little Chalfont, UK
ÄKTA prime/prime plus	GE HEALTHCARE, München
ÄKTA purifier 10	GE HEALTHCARE, München
ÄKTA pure	GE HEALTHCARE, München
Gel electrophoresis devices	
Gel electrophoresis chamber	HOEFER Pharmacia Biotech, USA
Power supply EPC-301	AMERSHAM, Little Chalfont, UK

Heating blocks	
HBP-2 131	HEAP LABOR CONSULT, Bovenden
HPLC columns	
Eclipse XDB-C18 (4.6x150)	AGILENT, Santa Clara, USA
HPLC system	
Agilent 1100 series	AGILENT, Santa Clara, USA
Incubators	
Binder (9010-0086)	BINDER, Tuttlingen
Lyophilizer	
Lyovac GT2	LEYBOLD-HERAEUS, Köln
Shakers	
Certomat H	B. BRAUN, Melsungen
Multitron	Infors HT, Bottmingen, Switzerland
Spectrophotometers	
Jasco V-650	JASCO GmbH, Groß-Umstadt
NanoDrop One	THERMO SCIENTIFIC, Waltham, USA
Ultrospec 10	AMERSHAM, Little Chalfont, UK
Thermal cyclers	
Mastercycler personal	EPPENDORF, Hamburg
Mastercycler gradient S	EPPENDORF, Hamburg
Ultrapure water systems	
MilliQ Q-POD	MILLIPORE GmbH, Schwalbach
Biopak Polisher filter	MILLIPORE GmbH, Schwalbach
Millipak filter	MILLIPORE GmbH, Schwalbach
UV imager	
Reprostar	Camag, Berlin
GelDoc Go	BioRad, Hercules, USA
Vacuum pump	
VP 2 Autovac	VWR, Leuven, Belgium

2.1.2 Consumables

Centrifugal filter units	
Amicon Ultra 0.5 (10 kDa)	MERCK MILLIPORE, Darmstadt
Amicon Ultra 4 (10 kDa)	MERCK MILLIPORE, Darmstadt
Amicon Ultra 15 (10 kDa)	MERCK MILLIPORE, Darmstadt
Amicon Ultra 15 (30 kDa)	MERCK MILLIPORE, Darmstadt
Centrifuge tubes	
10 ml/15 ml/50 ml	SARSTEDT, Nümbrecht
Column material	
C18 bonded silica (Versa Flash)	SUPELCO, Bellefonte, USA
Cryo tubes	
Cryo Pure Tube (1.8 ml)	SARSTEDT, Nümbrecht
Cuvettes	
UV Cuvettes	SARSTEDT, Nümbrecht
Electroporation cuvettes (2 mm)	MOLECULAR BIOPRODUCTS, San Diego, USA
Desalting columns	
NAP-5 columns	GE HEALTHCARE, München
Dialysis tubes	
Visking, type 27/32 (14 kDa)	ROTH GmbH & Co., Karlsruhe
Membrane filters	
ME24 Ø 47 mm; 0.2 μm pores	WHATMAN, Dassel
Micro plates	
UV-STAR® 96-well	Greiner Bio-One, Frickenhausen
CELLSTAR® 96-well	Greiner Bio-One, Frickenhausen
Micro tubes	
1.5 ml/2 ml	SARSTEDT, Nümbrecht
PCR tubes	
0.2 ml	PEQLAB, Erlangen
Syringe filters	
Filtropur 0.2 μm/0.45 μm pores	SARSTEDT, Nümbrecht

Phenex 0.2 μm/0.45 μm pores	PHENOMENEX, Aschaffenburg

2.1.3 Reagents

2.1.3.1 Chemicals

All reagents that were commercially available, were purchased from the following companies at highest purity:

- APPLICHEM, Darmstadt
- FLUKA, Buchs, Switzerland
- GE HEALTHCARE, München
- MERCK, Darmstadt
- MP BIOCHEMICALS, Illkirch, France
- RIEDEL-de HAËN AG, Seelze
- ROCHE, Mannheim
- ROTH, Karlsruhe
- SERVA, Heidelberg
- SIGMA-ALDRICH, Deisenhofen
- VWR, Leuven, Belgium

The chorismate producing strain *Escherichia coli KA12* was employed for the production of said substance. A protocol derived from [22] and [23] was used to obtain chorismate of high purity. The method is described in detail in subsubsection 2.3.2.1.
Ureidoacrylate was synthesized via a two step reaction from maleic anhydride and trimethylsilyl azide. A detailed description of this synthesis is given in subsection 2.4.1.
N-ethylammelide and N-isopropylammelide were synthesized from 6-chloro-1,3,5-triazine-2,4(1H,3H)-dione as described in subsubsection 2.5.2.1.

2.1.3.2 Buffers and Solutions

Saturated brine
360 g/l NaCl

Equilibration buffer for C-18 flash column

10 mM	Ammonium acetate (pH 6.0)

Saline for P1 transduction

145 mM	NaCl
50 mM	Trisodium citrate

Equilibration buffer for HisTrap runs (Buffer A)

100 mM	Tris/HCl (pH 7.5)
300 mM	KCl
10 mM	Imidazole

Elution buffer for HisTrap runs (Buffer B)

100 mM	Tris/HCl (pH 7.5)
300 mM	KCl
1 M	Imidazole

Size exclusion chromatography (SEC) and protein storage buffer

50 mM	Tris/HCl (pH 7.5)
50 mM	KCl

Protein sample buffer for SDS-PAGE

10% (w/v)	Glycerol
2% (w/v)	Sodium dodecyl sulfate (SDS)
0.01% (v/v)	Bromphenol blue
5% (v/v)	2-Mercaptoethanol
1.25 M	Tris/HCl (pH 6.8)

CD buffer

50 mM	potassium phosphate (PP) buffer (pH 7.0)

HPLC solvent 1

0.1% (v/v)	Formic acid (in H_2O)

HPLC solvent 2

0.1% (v/v)	Formic acid (in acetonitrile)

Indicator solution for chromeazurol S plates

0.1 mM	$FeCl_3$
1 mM	HCl
1 mM	Chromeazurol S
5 mM	Hexadecyl-trimethyl-ammonium bromide

TBE buffer

89 mM	Tris
89 mM	Boric acid
2.5 mM	EDTA

Transformation buffer I (TFB I)

100 mM	KCl
50 mM	$MnCl_2$
30 mM	Potassium acetate
10 mM	$CaCl_2$
15%	Glycerol
pH	6

Transformation buffer II (TFB II)

10 mM	Tris
10 mM	KCl
75 mM	$CaCl_2$
15%	Glycerol
pH	7

2.1.3.3 Growth Media

Vogel-Bonner salts (50x)

10 g/l	$MgSO_4 \cdot 7\ H_2O$
100 g/l	Citric acid monohydrate
500 g/l	K_2HPO_4
175 g/l	$NaH(NH_4)PO_4 \cdot 4\ H_2O$

Growth medium for *E. coli* KA12

2 g/l	Casamino acids
2 g/l	Yeast extract
41 mg/l	Tryptophan
20 ml	Vogel-Bonner salts
1.6 g/l	Glucose

Accumulation medium for *E. coli* KA12

12.8 g/l	Na_2HPO_4
1.36 g/l	KH_2PO_4
18 g/l	Glucose
2.7 g/l	NH_4Cl
20 mg/l	$MgCl_2 \cdot 6H_2O$
2 mg/l	Tryptophan

LB medium

10 g/l	NaCl
10 g/l	Tryptone
5 g/l	Yeast extract

LB agar

10 g/l	NaCl
10 g/l	Tryptone
5 g/l	Yeast extract
15 g/l	Agar

SOB medium

0.58 g/l	NaCl
20 g/l	Tryptone
5 g/l	Yeast extract
2.5 mM	KCl
10 mM	$MgCl_2$
10 mM	$MgSO_4$

M9 medium

12.8 g/l	Na_2HPO_4 x $7H_2O$
3.0 g/l	KH_2PO_4
0.5 g/l	NaCl
1.0 g/l	NH_4Cl
2 mM	$MgSO_4$
100 μM	$CaCl_2$
0.4%	Glucose (w/v)

M9-BPD medium

	M9 Medium
75 μM	2,2'-bipyridine (BPD)

M9-EDTA medium

	M9 Medium
1 μM	ethylenediaminetetraacetic acid (EDTA)

M9-BPD agar

10 g/l	NaCl
10 g/l	Tryptone
5 g/l	Yeast extract
15 g/l	Agar

2.1.3.4 Kits

Gene JET Gel Extraction Kit	Thermo Scientific, Waltham, USA
Gene JET Plasmid Miniprep Kit	Thermo Scientific, Waltham, USA
Crystallization Kit MD1-30	Molecular Dimensions, Rotherham, UK
Crystallization Kit MD1-37	Molecular Dimensions, Rotherham, UK
Crystallization Kit MD1-46	Molecular Dimensions, Rotherham, UK
Crystallization Kit PEGRx HT	Hampton Research, Aliso Viejo, USA
SelenoMet™	Molecular Dimensions, Rotherham, UK

2.1.3.5 Primers

Nucleotide primers were synthesized by BIOMERS.NET (Ulm) or Metabion (Planegg). Table 8.1 shows all primers used in this work.

2.1.3.6 Synthetic Genes

Genes from organisms other than *E. coli* and such mutants that could not reasonably be created in the laboratory were ordered as gene strings from Geneart (Regensburg). The sequences were optimized for gene expression in *E. coli*, moderate GC content, and lack of BsaI restriction sites. The lyophilized gene strings were solved in water and used as template in a PCR reaction. Synthetic genes used in this work are listed in Table 2.1. Full sequences of these genes are shown in Table 8.2.

Table 2.1: Synthetic genes used in this work
Required genes were ordered from Geneart (Regensburg) and received as lyophilized samples, which served as template DNA for PCR and later subcloning.

Name	Origin	Length
rutB FuncLib MIMD	*E. coli* // FuncLib [24]	0.69 kb
*fabG*2*entA*16	*E. coli* // JANUS [25]	0.74 kb
*rutB*2*entB*15	*E. coli* // JANUS [25]	0.69 kb
atzB	*Pseudomonas sp. ADP*	1.4 kb
*atzB*_Hom_Hal	*Haliea sp. SAOS-164*	1.4 kb

Name	Origin	Length
*atzB*_Hom_Pleo	*Pleomorphomonas oryzae*	1.4 kb

2.1.3.7 Plasmids

Vectors needed for gene expression were created from the pET21a and pUR22 plasmids by subcloning of PCR amplicons of genomic DNA or synthetic gene strings. For *in vivo* studies, relevant genes or gene libraries were subcloned into the pExp vector. Site directed or random mutagenesis yielded expression vectors for protein variants or gene libraries, respectively. pKD3 and pKD4 carry a chloramphenicol and a kanamycin resistance cassette, respectively, that were used to delete specific regions of the *E. coli* genome (see subsubsection 2.2.3.10). All plasmids utilized in this work are listed in Table 2.2.

Table 2.2: Plasmids used in this work
pET21a and pUR22 were the basis for all expression vectors. Vectors for *in vivo* studies are based on the constitutively active plasmid pExp. The resistance cassettes amplified from pKD3 and pKD4 were used for chromosomal gene deletion in *E. coli*.

Name	Resistance	Reference
pET21a	Ampicillin	NOVAGEN - MERCK MILLIPORE, Darmstadt
pUR22	Ampicillin	[26]
pExp	Ampicillin	[26] (named pTNA)
pKD3	Ampicillin, Chloramphenicol	[27]
pKD4	Ampicillin, Kanamycin	[27]

2.1.3.8 Proteins

DNA polymerases

Go-Taq polymerase	PROMEGA, Madison, USA
Q5 polymerase	NEW ENGLAND BIOLABS, Ipswich, USA

Miscellaneous enzymes	
BsaI-HF	NEW ENGLAND BIOLABS, Ipswich, USA
T4 polynucleotide kinase	Thermo Fisher Scientific, Waltham MA, USA
T4 DNA ligase	Thermo Fisher Scientific, Waltham MA, USA
T4 DNA ligase	ROCHE, Mannheim
Glutamate dehydrogenase	ROCHE, Mannheim
EntB	Maximilian Plach, Institute for
EntC	Biophysics and Physical Biochemistry,
TrpE	University of Regensburg
EntA	[20]
RutB	[20]
FabG	[20]

2.1.3.9 *E. coli* Strains

All plasmids were amplified in the NEB Turbo strain. Genes cloned into pUR22 and pUR23 were expressed in either BL21 Gold (DE3) or a BW25113 derived deletion strain to avoid contamination with the respective wild-type gene product. Genes cloned into pET21a were expressed in BL21 Gold (DE3). The strain DY329 was used for the targeted deletion of chromosomal genes (see subsubsection 2.2.3.10). Chorismate production was achieved with strain KA12, which was engineered for that purpose [22]. Deletion strains of *entA* and *entB-IC* (with and without additional deletion of the *fec* operon) were used for *in vivo* complementation experiments. All *E. coli* strains used in this work are listed in Table 2.3.

Table 2.3: Bacterial strains used in this work
The strain NEB Turbo was used to amplify all plasmids in this work. Both BL21 Gold (DE3) and BW25113 derived strains were used for protein production. The strains DY329 and BW25113 are required for the targeted chromosomal gene deletion protocol (see subsubsection 2.2.3.10). KA12 is a chorismate producing strain, graciously provided by the group of Prof. Dr. Donald Hilvert and Prof. Dr. Peter Kast, ETH Zürich. Strains carrying deletions of the genes *entA*, *entB*, *fecIRABCDE*, and combinations thereof were used for the screening of gene libraries.

Strain	Phenotype	Reference
E. coli NEB Turbo	F' $proA^+B^+$ $lacI^q$ $\Delta lacZM15/fhuA2$ $\Delta(lac\text{-}proAB)$ $glnV$ $galK16$ $galE15$ $R(zgb\text{-}210::Tn10)$Tet^S $endA1$ $thi\text{-}1$ $\Delta(hsdS\text{-}mcrB)5$	New England Biolabs

Strain	Phenotype	Reference
E. coli BL21 Gold (DE3)	F$^-$ $ompT$ $hsdS$(r_B^- m_B^-) dcm^+ Tet' gal $\lambda(DE3)endA$ Hte	Agilent Technologies
E. coli DY329	W3110 $\Delta lacU169$ $nadA::Tn10$ $gal490$ $\lambda cI857\Delta$(*cro-bioA*)	[28]
E. coli KA12	F$^-$ λ^- Δ(*pheA-tyrA-aroF*) thi-1 $endA1$ $hsdR17\Delta(argF-lac)205(U169)$ $supE44\Delta$(*srlR-recA*)306::$Tn10$	[22]
E. coli BW25113	F$^-$ Δ(*araD-araB*)567 $\Delta lacZ4787$ (::rrnB-3 λ^- rph-1 Δ(*rhaD-rhaB*)568 $hsdR514$	[27]
BW25113 Δ*entA*	BW25113 *entA*::*Cam*	Markus Busch Master's Thesis
BW25113 Δ*entB*	BW25113 *entB*::*Kan*	[29] (JW0587-1)
BW25113 Δ*entB-IC*	BW25113 *entB(1-561)*::*Cam*	this work
BW25113 Δ*entA* Δ*fec*	BW25113 *entA*::*Cam fecIRABCDE*::*Kan*	this work
BW25113 Δ*entB-IC* Δ*fec*	BW25113 *entB(1-561)*::*Cam fecIRABCDE*::*Kan*	this work
BW25113 Δ*guaD*	BW25113 *guaD*::*Kan*	[29] (JW5466-3)

2.1.3.10 Phages

P1 phage lysate	Markus Busch, Master's Thesis [20]

2.1.4 Software

2.1.4.1 Local Applications

ÄKTA Unicorn Version 5.31	© GE HEALTHCARE, 2011
AlphaFold	[30]
Chemdraw Professional 20	© Perkin Elmer, 1998-2020
CLC Main Workbench 8.1	© QIAGEN

Cytoscape 3.8.0	© Cytoscape Consortium [31]
Doc-It 2.3.0	© 2003 UVP Inc.
JalView 2.8.0b1	© Waterhouse A. 2014
JANUS	[25]
The PyMOL Molecular Graphics System, Version 2.3.3	© Schrodinger, LLC
Sigma Plot Version 12.5	© SYSTAT SOFTWARE, Inc., 2011
Spectra Manager 2.10.01	© JASCO COPERATION, 2012

2.1.4.2 Server Based Applications

BLAST *blast.ncbi.nlm.nih.gov/*	[32]
CASTp 3.0 *http://sts.bioe.uic.edu/castp/index.html*	[33]
Dali *http://ekhidna2.biocenter.helsinki.fi/dali/*	[34]
Clustal Omega *https://www.ebi.ac.uk/Tools/msa/clustalo/*	[35]
Enzyme Similarity Tool *http://efi.igb.illinois.edu/efi-est/*	[36]
ExPASy ProtParam tool *http://web.expasy.org/protparam/*	[37]
ExPASy Translate tool *http://web.expasy.org/translate/*	[37]
FuncLib *http://funclib.weizmann.ac.il/bin/steps*	[24]
NCBI *https://www.ncbi.nlm.nih.gov/*	NCBI, Bethesda, MD
PISA *http://www.ebi.ac.uk/pdbe/prot_int/pistart.html*	[38]
WegLogo 3 *http://weblogo.threeplusone.com/create.cgi*	[39]

2.2 Standard Methods

This section contains all methods that apply to multiple or all distinct projects of this work. Methods used in only one of the projects can be found in section 2.3, section 2.4, and section 2.5.

2.2.1 Bioinformatics

2.2.1.1 Sequence Acquisition

Protein and DNA sequences of interest were retrieved from the respective databases provided by the National Center for Biotechnology Information (NCBI) by text search. Homologues to sequences of interest were identified with the BLAST [32] toolkit provided by the NCBI (*https://blast.ncbi.nlm.nih.gov/*).

2.2.1.2 Generation of Multiple Sequence Alignments

Multiple sequence alignment (MSA) describes the arrangement of homologous sequences in such a way, that identical or similar regions overlap. MSAs were generated with Clustal Omega [35], using standard parameters. Visualization of specific regions of an MSA in the form of sequence logos was achieved with WebLogo [39].

2.2.2 Microbiological Methods

2.2.2.1 Preparation of Equipment and Solutions

Thermostable solutions, media, glassware, and consumables were sterilized by autoclaving (20 min; 121°C; 2 bar). All solid objects were subsequently dried (50°C) and glassware was additionally heat sterilized (4 h; 200°C). Heat-labile solutions were sterilized by ultrafiltration with a membrane (0.2 μM pore size) or syringe filter (0.2 or 0.45 μM pore size). Solvents intended for use in chromatographic systems were degassed by stirring in vacuum (30 min) or by sonication (15 min).

2.2.2.2 Transformation of Chemically Competent *E. coli* Cells

Chemically competent cells can easily be transformed with plasmid DNA, which is required for plasmid amplification, gene expression, or *in vivo* experiments. Cells were grown in 500 ml SOB medium to an OD_{600} of 0.6 before being cooled on ice (15 min), harvested (3220 g; 4°C; 10 min), washed with 50 ml transformation buffer I, and resuspended in 5 ml transformation buffer II. Aliquots of this cell suspension were stored at -80°C. Chemically competent *E. coli* cells were transformed by mixing a 100 μl aliquot of cell suspension with 100 - 200 ng of DNA on ice. Cells were then incubated according to the following scheme:

0°C	20 min
42°C	45 s
0°C	20 min

addition of 900 μl LB medium

37°C 20 min

Then, either an aliquot of the resulting suspension was added to liquid medium or cells were pelleted by centrifugation (1 min; 3200 g), resuspended in 100 μl of the resulting supernatant, and plated onto an agar plate.

2.2.2.3 Cultivation and Storage of *E. coli* Strains

All liquid cultures of *E. coli*, except for strain DY329, were grown at 37°C while shaken at either 100 - 140 rpm (1 liter cultures) or 220 rpm (5, 50, and 250 ml cultures). Strain DY329 was grown at 30°C at the same shaking intervals. All media were supplemented with applicable antibiotics at concentrations of 150 μg/ml, 30 μg/ml, 50 μg/ml, and 12.5 μg/ml, for ampicillin, chloramphenicol, kanamycin, and tetracycline, respectively. Plates were either incubated at 37°C for up to 24 h or at RT for longer periods of time. For storage, colonies grown on plates were stored at 4°C. For long term storage of specific strains, glycerol stocks were created by mixing equal volumes of liquid culture and glycerol (80%), which were kept at -80°C.

2.2.2.4 Disposal of Microorganisms

Bacterial cultures, agar plates, and contaminated equipment were autoclaved for 20 min before disposal. Surfaces and other contaminated objects unsuitable for autoclaving were sterilized with ethanol (70%).

2.2.3 Molecular Biology Methods

2.2.3.1 Standard Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) is routinely employed for amplification of a specific fragment from a DNA template [40]. Standard PCR entails thermic denaturation of double stranded DNA (dsDNA), annealing of short oligonucleotide primers to the resulting single stranded DNA (ssDNA), and synthesis of a new DNA strand from dNTP substrates by a DNA polymerase, restoring the original dsDNA. Cyclic repetition of these three steps leads to exponential amplification of the DNA fragment determined by the binding sites of the primers.
Primers used in this work are listed in Table 8.1. DNA fragments were amplified by Q5 DNA polymerase, possessing 3' $\rightarrow$ 5' exonuclease activity. This proof reading capacity suppresses DNA mutations and is particularly important for the amplification of long DNA fragments. Standard PCRs were performed in a total volume of 50 μl in a thermal cycler (lid temperature 110°C). Reaction mixtures contained 5-150 ng template DNA, 1 U Q5 polymerase, 0.5 μM of each primer, 200 nM dNTPs, and 1x Q5 reaction buffer. In cases where this reaction

composition produced no amplicon, an additional 3% DMSO was included. The standard PCR program consisted of the following steps:

1. Initialization	95°C	5 min
2. Denaturation	95°C	45 s
3. Annealing	T_A	45 s
4. Extension	72°C	t_E
5. Repetition of 2.-4.	34x	
6. Final elongation	72°C	10 min
7. Storage	8°C	∞

The extension time (t_E) for each reaction was dependent on the length of the DNA fragment to be amplified. The speed of the Q5 polymerase (20-30 s/kb) was provided by the manufacturer. The resulting theoretical t_E was typically increased by an additional 30 s. For each primer a theoretical melting point (T_M) was calculated in CLC Workbench. The annealing temperature (T_A) for each standard reaction was then set to be 2°C higher than the lowest T_M of the primers involved. When DMSO was included in the reaction mixture, T_A was chosen to be 2°C lower than the lowest T_M.

2.2.3.2 Colony PCR

Colony PCR (cPCR) refers to the amplification of DNA from biomatter, particularly bacterial colonies. cPCR was employed to verify the success of cloning, transformation, and chromosomal deletion by using small amounts of colonies from selective agar plates as template for a PCR reaction. The cheaper, non-proofreading polymerase GoTaq was used for cPCR as the resulting DNA is typically only used for analytic agarose gel electrophoresis. The reaction mixture (1 U GoTaq DNA polymerase; 1x Green GoTaq reaction buffer; 200 nM dNTPs; 0.5 μM of forward/reverse primers from Table 8.1) in a total volume of 50 μl was placed in a thermal cycler (lid temperature 110°C) running the following program:

1. Initialization	95°C	5 min
2. Denaturation	95°C	45 s
3. Annealing	T_A	45 s
4. Extension	72°C	t_E
5. Repetition of 2.-4.	32x	
6. Final elongation	72°C	10 min
7. Storage	8°C	∞

Extension times (t_E) and annealing temperatures (T_A) were chosen in the same way as for regular PCR (see subsubsection 2.2.3.1).

2.2.3.3 Site-Directed Mutagenesis

Site-directed mutagenesis was used to introduce selected point mutations into plasmids which would then generate specific protein variants. In this work, a modified version of the QuikChange protocol [41] was used. Two DNA primers with specific mismatches at their 5' ends were designed in such a way that their 5' ends would precisely meet when hybridized with the template DNA, thus leading to the amplification of a linearized version of the entirety of the target plasmid in a PCR reaction. The resulting linear amplicon was treated with T4 polynucleotide kinase and T4 DNA ligase (5 U; 10 U; 30 min 37°C; 30 min RT) in 1x T4 ligase buffer in a total volume of 50 μl. The resulting ligation solution, containing circular, mutated plasmids, was used for the transformation of competent *E. coli* cells (see subsubsection 2.2.2.2) without further purification.

2.2.3.4 Agarose Gel Electrophoresis

Agarose gels (1% agarose; 0.01% ethidium bromide; 0.5x TBE-buffer) were used for purification or analysis of DNA products or mixtures. DNA samples in 1x DNA loading dye (NEW ENGLAND BIOLABS, Ipswich, USA) were applied to the gel and separated by electrophoresis (190 V; 200 mA; 20 min). During this process, the negatively charged DNA migrates towards the anode, with larger fragments being more strongly retained by the gel matrix, resulting in distinct bands. DNA was detected under UV-light (λ= 302 nm) and the size of each band could be estimated by comparison to the GeneRuler^{TM} 1 kb Plus DNA length standard (Thermo Fisher Scientific, Waltham MA, USA).

2.2.3.5 Isolation of DNA From Agarose Gels

Relevant DNA bands were excised from an agarose gel under UV light ($\lambda = 302$ nm) and purified according to the Gene JET Gel Extraction Kit protocol. DNA that was to be used for electrotransformation was eluted with 50 μl of sterile water instead of the elution buffer included in the kit. Resulting DNA solutions were stored at -20°C.

2.2.3.6 Subcloning of dsDNA Utilizing BsaI Restriction Endonuclease

Subcloning of DNA fragments into a vector was achieved by the protocol published by Rohweder and co-workers [26]. This protocol uses the special properties of the restriction endonuclease BsaI, which cuts dsDNA at a defined position outside of its recognition sequence, creating 4 nt 5'-terminal sticky ends. Properly designing the sequences of these overhangs yields restriction products with selectively complementary termini, allowing for efficient and oriented ligation and even assembly of multiple DNA fragments in a one-pot reaction. Furthermore the process of ligation in this protocol is irreversible, as the BsaI recog-

nition sites are located outside of the DNA regions of interest and are thus not retained in the ligation product. BsaI recognition sites and desired overhang sequences are easily introduced by PCR with a properly designed primer.
The standard reaction for BsaI mediated subcloning contained 100 ng of the target vector, a fivefold molar excess of linear amplicon, 0.2 mg/ml bovine serum albumin (BSA; NEW ENGLAND BIOLABS, Ipswich, USA), 5 U T4 DNA ligase, and 20 U BsaI-HF in 1x T4 ligase buffer (Thermo Fisher Scientific, Waltham MA, USA or ROCHE, Mannheim) in a total volume of 20 μl. The reaction proceeded in a thermo cycler which alternated between 5 min at 37°C and 5 min at 16°C for 50 cycles, followed by a final inactivation step of 10 min at 65°C. The resulting reaction mixture was used for the transformation of competent *E. coli* cells (see subsubsection 2.2.2.2) without further purification.

2.2.3.7 Isolation of Plasmid DNA From Bacterial Culture

Plasmid DNA was extracted from *E. coli* cultures utilizing the Gene JET Plasmid Miniprep Kit. Cells were harvested from 5 ml of stationary-phase culture by centrifugation (10 min; 3200 g; 4°C) and treated according to the official protocol. The final elution fraction, containing plasmid DNA, was stored at -20°C.

2.2.3.8 Determination of DNA Concentration

DNA concentration in solution was determined photometrically in a NanoDrop One. Measurement of absorption at λ=260 nm and an assumed specific extinction coefficient of 20 cm^2/mg gave the mass concentration of DNA according to Equation 2.1.

$$\beta = \frac{A_{260}}{A_{260}^{0.1\%} \cdot d} \tag{2.1}$$

β: mass concentration [mg/ml]
A_{260}: absorbance at 260 nm
$A_{260}^{0.1\%}$: specific extinction coefficient at 260 nm (20 cm^2/mg)
d: pathlength [cm]

DNA solutions were considered pure if there was no measurable absorption above 300 nm and the OD_{260}/OD_{280} ratio was at least 1.8.

2.2.3.9 DNA Sequencing

All subcloning, as well as relevant gene deletion and DNA amplification processes were ultimately evaluated by DNA sequencing (SEQLAB, Göttingen). Standard sequencing mixtures contained up to 1 μg of DNA preparation and 30 pmol of a region-specific sequencing primer (see Table 8.1) in a total volume of 15 μl.

2.2.3.10 Deletion of Chromosomal Genes in *E. coli*

E. coli deletion strains, lacking a specific region of genomic DNA, often display distinct phenotypes due to the lack of one or more functional gene products. Deletion strains were either purchased from the Coli Genetic Stock Center (CGSC) at Yale University or produced in our lab.
Gene deletion was achieved by producing a PCR amplicon containing a chloramphenicol (Cam) or kanamycin (Kan) resistance gene as encoded by plasmids pKD3 or pKD4, respectively, as well as terminal sequences (50 nt) identical to the 5'- and 3'-regions outside of the region to be deleted. The amplicons were purified and transformed into electrocompetent *E. coli* DY329 cells which were cured in LB medium (16 h; RT), plated on LB_{Cam} or LB_{Kan} plates, and incubated at 30 °C for 48 h. *E. coli* strain DY329 contains a defective λ prophage and is thus able to integrate linear DNA fragments into its genome via homologous recombination [28]. Colonies exhibiting the desired antibiotic resistance were cultivated in liquid medium and treated with pre-existing P1 phage lysate (6 h; RT). Non-lyzed cells were destroyed with chloroform treatment and cell debris removed by centrifugation (10 min; 3200 g; 4°C). The resulting supernatant, containing P1 phages stochastically carrying DNA fragments of the desired gene deletion, was then used to infect a liquid culture of *E. coli* BW25113 or of a pre-existing BW25113 deletion strain. This transfer of the gene deletion into a different strain was necessary, due to several experimental limitations of the *E. coli* DY329 strain. Infection of the target strain was achieved by supplementing 0.8 ml liquid culture with 2.45 mM $CaCl_2$ and adding 0.4 ml of the previously generated phage lysate. After incubation (25 min; 37 °C) infection was halted by addition of 5 ml P1 saline and cells were cured (1 h; 37 °C), harvested (8 min; 3220 g; 4 °C), and resuspended in 100 μl P1 saline before being plated on selective LB agar plates. Colonies exhibiting the desired antibiotic resistance were further evaluated by colony PCR and sequencing of the relevant DNA region. Deletion strains obtained from the CGSC were similarly assayed for the integrity of their gene deletions. Colonies with properly deleted gene regions were kept as glycerol stocks.

2.2.4 Protein Biochemistry

2.2.4.1 Recombinant Expression of Genes in *E. coli*

Recombinant expression of plasmid coded genes yielded desired proteins and protein variants required for *in vitro* applications. Expression strains used to this end were *E. coli* BL21(DE3) Gold, BW25113 Δ*entA*, and BW25113 Δ*guaD*. Cells were made chemically competent, transformed with expression vector, and grown in LB medium supplemented with applicable antibiotics. These precultures were used to inoculate varying amounts of 1 l batches of selective LB medium to an OD_{600} of 0.1. Cells were grown (37°C; 140 rpm) to an OD_{600}

of 0.6-0.8 before supplementing the culture with 0.5 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) to induce gene expression. Cells were further shaken (20°C; 16 h), harvested (JFA 8.1; 16 000 g; 20 min; 4°C), and resuspended in 20 ml buffer A per liter of total culture volume. Cells were disrupted by sonication (60% amplitude; 3 min; 2 s intervals; 2 s pauses) while kept on a freezing mixture of ice and NaCl. Cell debris was removed by centrifugation (16 000 g; 45 min; 4°C) and the recombinant proteins were purified from the supernatant as described below.

2.2.4.2 Immobilized Metal Ion Affinity Chromatography

Modifying a protein sequence to contain multiple adjoining histidines, creates a local structure referred to as polyhistidine-tag, His-tag, or His_n-tag (where n is the number of histidines). This structure displays reversible, high affinity binding to Ni^{2+} ions, which is routinely used to purify His-tagged proteins by immobilized metal ion affinity chromatography (IMAC). In this method, tagged proteins are binding to Ni^{2+} ligands immobilized to a Sepharose resin, after which imidazole is used to elute bound protein.
After equilibrating the FPLC system and the IMAC column in equilibration buffer, the crude extract described in subsubsection 2.2.4.1 was filter sterilized (pore size 0.45 μm) and applied to the column. The column was washed with 15 column volumes (CVs) of equilibration buffer to remove non-binding protein, after which the protein of interest was eluted with a gradient to elution buffer over 15 CVs (0-100%). Elution of protein and impurities was monitored by continuous measurement of OD_{280}, OD_{260}, and OD_{333}. Elution fractions were analyzed by SDS-PAGE and relevant ones pooled for further purification by size exclusion chromatography (see subsubsection 2.2.4.3) or dialysis (subsubsection 2.2.4.4). IMAC columns were washed with elution buffer after each run and stored in 20% ethanol at 4°C when not in use.

2.2.4.3 Preparative Size Exclusion Chromatography (SEC)

Various SEC or gel filtration methods exist, all of which make use of porous column materials which function as inverse molecular sieves. Molecules differ in their tendency to enter the porous beads and accordingly in their retention time within the column material when a continuous solvent flow is applied. This effect is mainly dependent on the size of the molecular species, with larger molecules being eluted first. The relationship between the elution time and molecular weight of a molecular species can be described by a logarithmic function.

The SEC protocol used in this work consisted of the application of up to 10 ml of HisTrap eluate (see subsubsection 2.2.4.2) to a HiLoad 26/60 Superdex 75 or a HiLoad 26/60 Superdex 200 column equilibrated with SEC buffer (see subsubsection 2.1.3.2) and elution with 1.2 CV of SEC buffer at a flow rate of

1.5 ml/min. Runs were conducted at 4°C and fractions of 3.5 ml were collected. Protein elution was monitored by continuous measurement of OD_{280}. Relevant fractions were further analyzed by SDS-PAGE before select fractions being pooled and stored at -80°C.

2.2.4.4 Dialysis of Protein Solutions

In the case of low protein yields, no SEC purification was performed but HisTrap eluate was dialyzed, mainly to remove imidazole. Protein solutions were dialyzed against an at least 200-fold volume of dialysis buffer (SEC buffer) while placed in a dialysis tube (14 kDa molecular weight cutoff). Dialysis proceeded for at least 3 h at 4°C before the protein solutions were stored at -80°C.

2.2.4.5 Concentrating Protein Solutions

Protein solutions were concentrated by ultrafiltration with centrifugal filter devices (Amicon Ultra; molecular cutoff: 10/30 kDa). These devices were centrifuged (4000 g; 4 °C) until the desired concentration was achieved.

2.2.4.6 Storage of Protein Solutions

Purified protein solutions were added dropwise to liquid nitrogen and stored at -80°C in sterile tubes.

2.2.5 Analytical Methods

2.2.5.1 SDS-Polyacrylamide Gel Electrophoresis

Proteins treated with the detergent sodium dodecyl sulfate (SDS) become unfolded and bind SDS at a fixed ratio of approximately one molecule SDS per 1.4 amino acid residues. The resulting uniform mass-to-charge ratio allows for solely size-dependent separation of proteins in an SDS-polyacrylamide gel subjected to an electric field. In this system, the migration speed of a particle is inversely proportional to the logarithm of its mass.
Protein samples were mixed with an equal volume of 2x SDS sample buffer and heated (95°C; 5 min), causing protein denaturation and SDS binding. Treated samples were applied to SDS-polyacrylamide gels (13.5% acrylamide) and subjected to an electric field (35 min; 50 mA; 300 V). Gels were stained with Coomassie Brilliant Blue R250 (0.2% w/v; 15 min) and excess dye was removed by boiling the gels in water until acceptable contrast was achieved. The detection limit for this staining method is 200-500 ng/mm^2. The molecular weight of the resulting protein bands was estimated by comparison to protein mass standard LMW (Thermo Fisher Scientific, Waltham MA, USA) applied to the same gel.

2.2.5.2 Determination of Protein Concentration by UV-Absorption Spectroscopy

A protein's absorbance spectrum at near-UV wavelengths is determined by the number of its tryptophan (Trp) and tyrosine (Tyr) residues as well as by the number of disulfide bonds between two cysteines (cystine). Since the cytoplasm in *E. coli* is a reducing environment [42] cystine formation is not a factor for the proteins in this work. Accordingly, the extinction coefficient of a protein can be determined solely based on its sequence (see Equation 2.2).

$$\epsilon_{280} = \sum \text{Trp} \cdot 5500 + \sum \text{Tyr} \cdot 1490 + (\sum \text{Cystine} \cdot 125) \tag{2.2}$$

ϵ_{280}: molar extinction coefficient at 280 nm [$M^{-1}cm^{-1}$]

To determine the concentration of a protein solution, its OD_{280} value was measured and the concentration was calculated according to Lambert-Beer's law (see Equation 2.3).

$$A_{280} = \epsilon_{280} \cdot c \cdot d \tag{2.3}$$

A_{280}: absorbance at 280 nm
ϵ_{280}: molar extinction coefficient at 280 nm [$M^{-1}cm^{-1}$]
c: protein concentration [M]
d: pathlength [cm]

2.2.5.3 Circular Dichroism Spectroscopy

Linear polarized light, containing equal intensities of left and right circular polarized light, is affected in a unique way by optically active substances. Such substances exhibit differential absorption of left and right circular polarized light, thus changing the composition of the source light, turning it into elliptically polarized light [43]. This property of optically active substances is called circular dichroism (CD).
Protein solutions were thawed and protein aggregates were removed by centrifugation (20 min; 16 000 g; 4°C). Buffer was exchanged to 50 mM potassium phosphate (PP) buffer (pH 7.0) with a NAP-5 column or by dilution. Protein solutions were diluted with PP buffer so their minimal ellipticity signal between 185 and 260 nm was approximately -15 mdeg. CD spectra were measured between 185 and 260 nm at a scanning speed of 20 nm/min with three repetitions. The ellipticity values (θ) from the recorded spectra were converted into normalized θ_{MRW} values (see Equation 2.4).

$$\theta_{MRW} = \frac{\theta}{c \cdot d \cdot N_R} \tag{2.4}$$

θ: observed ellipticity [mdeg]
c: concentration [mol/l]
d: pathlength [cm]
N_R: Number of protein residues

2.2.5.4 Thermal Denaturation of Proteins

Thermal unfolding of proteins usually proceeds in a cooperative manner, characterized by a sudden, sigmoidal shift from folded to unfolded state with rising temperatures. The midpoint of this transition is considered the melting point (T_M) of the respective protein. T_M values generally indicate the overall stability of a protein or protein variant. The CD signal at a fixed wavelength is a suitable indicator of the amount of folded protein in solution, as its intensity depends on the sum of protein conformation in solution. Protein solutions were prepared as described above (see subsubsection 2.2.5.3) and heated from 20°C to 95°C at a fixed rate of 1°C/min. During the gradient, the ellipticity signals at λ = 208 nm and λ = 220 nm, correlating to α-helix and β-sheet content, respectively, were continuously observed. The resulting thermal denaturation curve was fitted with the following sigmoidal function, revealing the T_M values of the respective protein (see Equation 2.5).

$$\theta_0 + \frac{\Delta\theta}{1+e^{-\frac{(T-T_M)}{b}}} \tag{2.5}$$

θ_0: initial ellipticity [mdeg]
$\Delta\theta$: amplitude of ellipticity change [mdeg]
e: Euler's number
T: temperature (variable) [°C]
T_M: melting point [°C]
b: curvature parameter

2.2.5.5 C-18 Reverse Phase Chromatography

High performance liquid chromatography (HPLC) encompasses a number of powerful analytical techniques for the identification and quantification of substances in solution. In this work HPLC based C-18 reverse phase chromatography methods were employed to screen for very low promiscuous enzyme activities and changes in product composition.
In vitro enzymatic reactions were conducted as described below (see subsection 2.3.6 and subsubsection 2.5.4.1). Enzymes were removed from the reaction mixture by ultrafiltration through an Amicon Ultra 0.5 centrifugal filter unit (MERCK MILLIPORE, Darmstadt). Substance only controls, containing no protein in the reaction mixture, were not filtrated. 5 μl of filtrate were injected onto

an Eclipse XDB-C18 reverse phase column (AGILENT, Santa Clara, USA; 4.6x150 mm) connected to an Agilent 1100 series HPLC system and preequilibrated with a solution of 5% HPLC solvent 2 in HPLC solvent 1. Separation proceeded at a constant flow rate of 1.0 ml/min with an initial 5 min of 5% HPLC solvent 2 and a gradient from 5% to 100% HPLC solvent 2 over 5 or 15 min. During the run the absorption spectrum between 210 and 320 nm was continuously recorded. Substances contained in the analyte solution were identified by comparison of retention times and absorption spectra to those of genuine substance samples. Relative quantification of components was achieved by integration of the respective peak in the chromatogram. The column was cleaned with 100% acetonitrile and stored in 40% acetonitrile (in H_2O) when not in use.

2.2.5.6 Mass Spectrometry

HPLC-ESI-MS runs of specific samples were conducted by the central analytics division of the Faculty of Chemistry and Pharmacy at the University of Regensburg to confirm the identity of the products of chemical reactions.

2.3 Enterobactin Pathway Methods

2.3.1 Bioinformatics

2.3.1.1 FuncLib

The software FuncLib [24] offers computational design of new stable residue networks within enzyme active sites. This strategy can guide the exploration of sequence space and has been employed to design and improve novel enzyme functions [24].
The web server of FuncLib (*https://funclib.weizmann.ac.il/bin/steps*) was provided with the required crystal structures (RutB: chapter 4; FabG: 1Q7B) and residue positions to be mutated (RutB: 29, 35, 41, 43, 72, 74, 92, 136, 162, 165, 191; FabG: 90, 92, 138, 140, 148, 183, 186, 188, 189, 192). The active site residues of the respective enzymes (RutB: D24, K133, C166; FabG: 110, 151, 155) as well as the $NADP^+$ cofactor of FabG were kept unchanged. The enzyme variants predicted to be most stable were selected for production and further characterization.

2.3.1.2 JANUS

JANUS [25] uses phylogenetic and optional structural information of two homologous enzyme families to specifically suggest residue exchanges for the conversion between the two distinct enzyme activities. JANUS was run locally with standard parameters, providing either MSAs of RutB and EntB alongside the crystal structure of RutB (see chapter 4) or MSAs of FabG and EntA alongside

the crystal structure of FabG (PDB: 1Q7B). Best rated residue exchanges for RutB and FabG were compiled into genes coding for the respective enzymes, which were produced and characterized.

2.3.2 Chemical Methods

2.3.2.1 Purification of Chorismate

E. coli KA12 has previously been engineered as a chorismate producing strain by deletion of the chorismate mutase gene, resulting in accumulation and eventual secretion of chorismate into the culture medium [22]. KA12 cells were cultivated overnight in LB_{Tet} medium (12.5 μg/ml tetracycline) then used to inoculate 500 mL of growth medium. Growth in this medium continued in a shaker (30°C; 150 rpm) until an OD_{600} of 1.6 to 1.9 was reached. After centrifugation (20 min; 4000 g; 4°C), the cell pellet was resuspended in 500 ml of accumulation medium and shaking was continued (30°C; 120 rpm; 16 h). Cells were removed by centrifugation as before and the chorismate containing supernatant was purified as follows.
HCl (5 M) was added to the supernatant to reach pH 1.5 and extraction with ethyl acetate followed (2x 180 ml). The organic fractions were pooled, filtered, and extracted with saturated brine. The organic phase was evaporated by means of a solvent resistant pump (40 mbar; 20 h), resulting in a solid residue, tainted with a red, oily substance. A flash column was filled with 50 g of spherical C-18 reverse phase material and flushed with equilibration buffer. After application of the solid residue solved in equilibration buffer to the column, elution proceeded with equilibration buffer and 10 ml fractions were collected. Fractions were analyzed by HPLC analysis (see subsubsection 2.2.5.5), allowing for detection of chorismate as well as 4-hydroxybenzoate impurities, which could be distinguished by their absorption maxima of 275 nm and 244 nm, respectively, as well as by their retention times in relation to genuine substance standards. Sufficiently pure fractions were pooled and lyophilized. The resulting residue was dissolved in sterile water, aliquoted, and stored at -80°C.

2.3.2.2 Characterization of Chorismate Preparation

The purity of the chorismate preparation described above, was determined by HPLC analysis (see subsubsection 2.2.5.5). Chorismate was identified by its retention time and spectrum by comparison to a genuine substance standard. The concentration of chorismate was determined photometrically (ϵ(275 nm) = 2630 $M^{-1}cm^{-1}$; [44]).
Furthermore, the substance purified was unambiguously shown to be chorismate by an enzymatic assay employing anthranilate synthase (TrpE) from *E. coli*. TrpE converts chorismate and NH_3 to anthranilate and possesses no known

side activity [45]. A standard assay contained the following substances in a total volume of 150 μl:

50 mM	Bicine buffer (pH 8.5)
50 mM	$MgCl_2$
200 mM	NH_4Cl
10 μM	TrpE
500 μM	Chorismate (putative)

The reaction mixture was incubated at RT for 1 h, before being ultrafiltrated through an Amicon Ultra 0.5 centrifugal filter unit. The filtrate was analyzed on a C-18 reverse phase HPLC column as described (see subsubsection 2.2.5.5). Anthranilate was identified by its retention time and spectrum by comparison to a genuine substance standard.

2.3.3 Microbiological Methods

2.3.3.1 Selection Conditions for Enzymatic Complementation

E. coli uses the siderophore enterobactin as main uptake mechanism for iron when this essential nutrient is scarce or otherwise limited. The strains used in the following selection experiments carry chromosomal gene deletions that render them unable to produce enterobactin. This enterobactin deficiency hinders growth of the strains under minimal media conditions and prevents growth entirely when iron in the medium is complexed by a chelating agent like 2,2'-bipyridine (BPD) [46]. Growth can be restored by complementing the function of the gene deleted with an isofunctional gene, thus restoring enterobactin production and iron uptake. In this work, the deletion strains BW25113 *entA*::*Cam* and *entB-IC*::*Cam* were created as described (see subsubsection 2.2.3.10 to identify such variants of *rutB* and *fabG* that produced enzymes with isochorismatase activity and 2,3-dihydro-2,3-dihydroxy-benzoate dehydrogenase activity, respectively.
Deletion strains were made electrocompetent and transformed with the applicable gene library as described in subsubsection 2.3.4.3. Transformants were grown in M9 medium to stationary phase (37°C; 120 rpm) before being used to inoculate a volume of M9 medium containing 75 μM BPD or 1 μM EDTA to an OD_{600} of 0.01. After at least two rounds of duplication (37°C; 120 rpm), cells were plated onto LB_{AmpCam} agar plates and grown overnight (37°C). Multiple colonies from the same plate were cultivated separately and their respective plasmids were sequenced as described in subsubsection 2.2.3.7 and subsubsection 2.2.3.9, respectively.

2.3.3.2 Quantitative Growth Assay

Selected gene variants isolated from the *fabG* or *rutB* gene libraries or predicted by FuncLib (see subsubsection 2.3.1.1) were tested for their potential to increase the growth rate of *entA* or *entB* deletion strains under iron limited conditions, respectively, to gauge the effect of each individual variant on growth rate. To this end both BW25113 *entA*::*Cam* and BW25113 *entA*::*Cam fec*::*Kan* were transformed with pExp vectors each carrying a different *fabG* variant and BW25113 *entB-IC*::*Cam* cells were transformed with *rutB* variants cloned into the pExp plasmid. Transformants were grown overnight (37°C; 120 rpm) in M9 medium containing the applicable antibiotics, as well as 20 μM 2,3-dihydroxybenzoate (2,3-DHB) to improve initial growth. 10 μl of each of those cultures were then used to inoculate 5 ml volumes of M9 medium containing the applicable antibiotics which were then shaken at 37°C for 48 h. The resulting cultures were used to inoculate volumes of 1 ml M9 medium containing the applicable antibiotics and 20 μM BPD. These final cultures were grown in a 24 well plate in a TECAN plate reader (37°C; 140 rpm) for up to 70 h with the OD_{600} being continuously measured.

2.3.4 Molecular Biology Methods

2.3.4.1 Random Mutagenesis

Random mutagenesis of a certain gene can create a great number of variants with variable amounts of mutations, depending on the conditions applied. Expression of this gene library in a suitable selection strain then allows for the selection of a certain feature of the gene product. As randomized gene libraries potentially explore the entirety of a region of sequence space, they are usually employed when no other, specific mutational strategy is viable.
In this work, the wild-type *E. coli rutB* and *fabG* genes were randomly mutated over the entirety of their open reading frames (ORFs) by error prone PCR (epPCR). epPCRs were performed in a total volume of 50 μl in a thermal cycler (lid temperature 110 °C). Reaction mixtures contained 5-150 ng template DNA, 15 U GoTag, 0.5 μM of each primer, 1 mM $MgCl_2$, 0.35 mM dATP, 0.40 mM dCTP, 0.20 mM dGTP, 1.35 mM dTTP, and 1x GoTaq reaction buffer. Additionally, reaction mixtures contained varying concentrations of $MnCl_2$, which were 0.9 mM in the case of the *rutB* gene library and 0.8 mM in the case of *fabG*. The PCR program was identical to the one described in subsubsection 2.2.3.1
In RutB, residues Y29, Y35, and F41 may be relevant to substrate binding, prompting the creation of a small, focused gene library. Full randomization of the associated codons was achieved by adhering to the protocol for site-directed mutagenesis (see subsubsection 2.2.3.3), utilizing the wild-type *rutB* gene as template and two partly degenerate oligonucleotides (FlexLoop-Rand1 and FlexLoop-Rand2 in Table 8.1) as primers. The two primers were fully ran-

domized (NNN) at the codons selected for randomization, translating to the residue exchanges Y29X Y35X F41X in RutB.

2.3.4.2 Efficient Subcloning of dsDNA

Due to the large number of possible variants resulting from full random mutagenesis of a certain gene, subcloning, as well as all other subsequent steps, have to be highly efficient to preserve the full diversity of gene variants. In the case of the focused gene library of *rutB*, containing only a limited number of possible variants, the standard protocol for subcloning of dsDNA (see subsubsection 2.2.3.6) was employed.
Efficient ligation was achieved by mostly adhering to the protocol described in subsubsection 2.2.3.6 with the products of epPCR (see subsubsection 2.3.4.1) serving as the linear amplicon and pExp being the target vector. Furthermore, T4 ligase produced by ROCHE was employed here, owing to its higher activity compared to T4 ligase distributed by Thermo Fisher Scientific. To further increase the yield of successfully subcloned mutants, 25 ligation reactions were conducted in parallel.

2.3.4.3 Transformation of Electrocompetent *E. coli* Cells

Since maintaining the full spectrum of sequence variants is essential in the creation of a gene library, electrotransformation rather than transformation of chemically competent cells (see subsubsection 2.2.2.2) was employed, due to the increased efficiency of the former method. Similarly, electrotransformation is preferrable when transforming a gene library into a selection strain. *E. coli* NEB Turbo was used for the initial generation of a gene library, while BW25113 *entA*::*Cam* and *entB-IC*::*Cam* strains (see subsubsection 2.1.3.9) were used for growth under selection conditions (see subsubsection 2.3.3.1).
To make *E. coli* cells electrocompetent, 50 ml of LB culture, grown to stationary phase, were harvested by centrifugation (8 min; 3200 g; 4°C). The pellet was then sequentially washed with H_2O (30 ml, 2x 1 ml) at 4°C, before being resuspended in 1 ml H_2O. Between each washing step, cells were pelleted like before.
The DNA solutions used in electrotransformation were dialyzed against an approximately 500-fold volume of water for 30 min. 250 μl of electrocompetent cells were mixed with 600 ng of DNA and electroporated in an Eppendorf Electroporator 2510 (2500 V; 25 μF; 200 Ω) with electroporation cuvettes featuring 2 mm gaps. Immediately after electroporation, the cell suspension was pipetted into a reservoir of LB medium and eventually cured (1-2 h; 37°C).

2.3.4.4 Preparation of Gene Libraries

Gene libraries created by random mutagenesis (see subsubsection 2.3.4.1) and subcloning into a target plasmid (see subsubsection 2.3.4.2) needed to be amplified to gain enough copies of each variant for subsequent selection or screening experiments.
E. coli NEB Turbo cells transformed with a newly generated gene library (see subsubsection 2.3.4.3) were allowed to grow on LB_{Amp} agar plates (16 h; 37°C). Multiple plates were used to allow for uniform and unimpeded growth of each transformant. The number of actual transformants was determined by serial dilution and growth on selective agar plates. Colonies were scrubbed off the plates and resuspended in LB medium. Plasmid DNA was extracted from an aliquot of this cell suspension as described in subsubsection 2.2.3.7 to give the final gene library preparation.

2.3.5 Protein Purification

Proteins were produced and purified as described in subsection 2.2.4. Expression strains were BL21 (DE3) for EntA and EntB, BW25113 Δ*entA* for all variants of FabG, and BW25113 Δ*entB* for all variants of RutB. The deletion strains were employed to prevent any contaminations with EntA or EntB, to avoid false positive results.

2.3.6 Qualitative Enzyme Assays

In order to confirm novel enzymatic activities in the enzyme variants extracted from different gene libraries (see subsubsection 2.3.3.1) or derived from bioinformatic analysis (see subsection 2.3.1), a qualitative assay based on HPLC analysis was employed, that was previously established in the author's masters thesis [20] (see Figure 2.1).

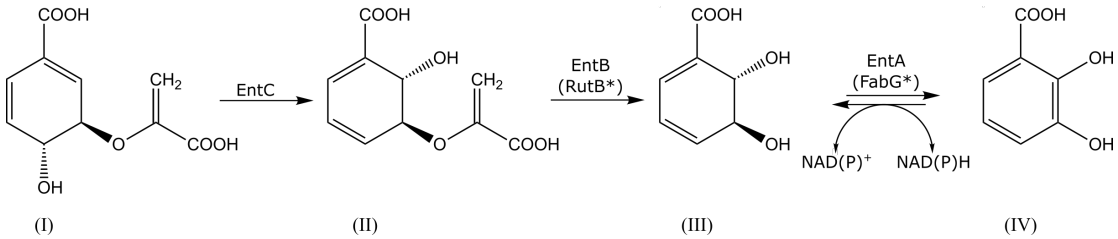


Figure 2.1: Reaction scheme of the enterobactin pathway enzymes EntC, EntB, and EntA
The reactions shown constitute the initial steps of the enterobactin pathway from *E. coli*, which were replicated *in vitro* to identify variants of ureidoacrylate amidohydrolase (RutB) and 3-oxoacyl-[acyl-carrier-protein] reductase (FabG) with novel enzymatic activities. Chorismate (I) is converted into isochorismate (II) by isochorismate synthase (EntC). Isochorismatase (EntB) then converts (II) into 2,3-dihydro-2,3-dihydroxybenzoate (III) by hydrolytic cleavage of the pyruvate group. Finally, (III) is oxidized by 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EntA) utilizing the cofactor NAD^+ to give 2,3-dihydroxybenzoate (IV). Enzyme variants of RutB (RutB*) and FabG (FabG*) possessing EntB and EntA activity, respectively, were supposed to be identified in this work. Variants of FabG might potentially utilize $NADP^+$ as a cofactor or favor the inverse reaction.

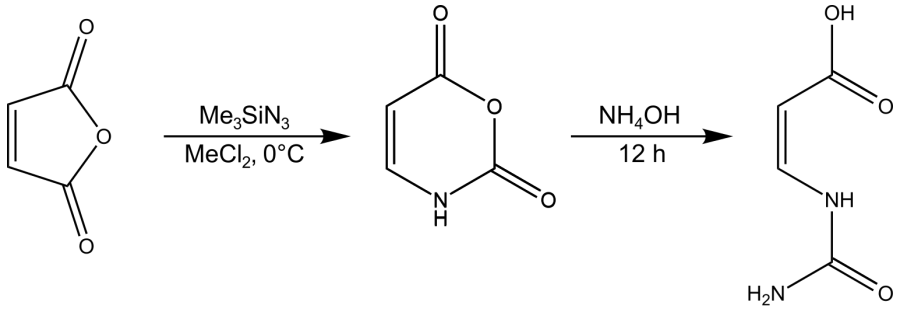
All enzymatic assays contained 500 μM chorismate in a total volume of 150 μl potassium phosphate (PP) buffer (50 mM; pH 7.0) as well as different combinations of enzymes, according to the activity to be tested. Reaction mixtures containing EntC were supplemented with 500 μM $MgCl_2$ to achieve optimal EntC activity [47]. To test for isochorismatase (EntB) activity, 10 μM EntC were added to produce the substrate of EntB, isochorismate, *in situ*. 1 μM EcEntB served as a positive control while the enzyme variants to be tested were added up to a concentration of 150 μM. The expected reaction product was 2,3-dihydro-2,3-dihydroxybenzoate (2,3-DHDHB).
To test for 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EntA) activity, 10 μM EntC and 1 μM EcEntB were added to produce the substrate of EntA, 2,3-DHDHB, *in situ*. 10 μM EcEntA together with 10 mM NAD^+ served as a positive control while the enzyme variants to be tested were added up to concentrations of 150 μM with either 10 mM NAD^+ or 10 mM $NADP^+$. The expected reaction product was 2,3-dihydroxybenzoate (2,3-DHB).
Additionally, the reverse reaction of EntA was assayed in a total volume of 150 μl potassium phosphate buffer (50 mM; pH 7.0) with 500 μM 2,3-DHB as a substrate. A positive control for this reaction does not exist. The enzyme variants to be tested were added up to a concentration of 150 μM with either 10 mM NADH or 10 mM NADPH. The expected reaction product was 2,3-DHDHB.

All assays were incubated for 16-20 h at RT, at which point all enzymes were removed by ultrafiltration and the filtrate was analyzed by HPLC as described (see subsubsection 2.2.5.5).

2.4 RutB Methods

2.4.1 Synthesis of Ureidoacrylate

The two known substrates of RutB, (Z)-3-ureidoacrylate and (Z)-3-ureidoacrylate peracid were not commercially available, necessitating in-house synthesis. As the peracid is known to be unstable under reaction conditions [48], only the synthesis of (Z)-3-ureidoacrylate was established. In the following, (Z)-3-ureidoacrylate is referred to simply as ureidoacrylate. In a first step, 3-oxauracil was synthesized by nitrogen insertion into maleic anhydride, based on the protocols published by Washburne and coworkers [49] as well as Rehberg and Glass [50]. 3-Oxauracil was then converted into ureidoacrylate by a ring-opening ammonolysis, as described by Farkaš and coworkers [51] as well as Kim and coworkers [52] (see Figure 2.2).



(I) (II) (III)

Figure 2.2: Two-step reaction scheme for the synthesis of (Z)-3-ureidoacrylate
In a first reaction step, maleic anhydride (I) in dry dichloromethane was reacted with trimethylsilyl azide until cessation of gas formation. The resulting product 3-oxauracil (II) was extracted from the reaction mixture by lyophilization and washed repeatedly with ethanol. (II) was then dissolved in 4 M NH_4OH and incubated at RT for 12 h. The final product (Z)-3-ureidoacrylate (III) was attained from the lyophilizate of the reaction mixture by methanol extraction.

Maleic anhydride was solved in dichloromethane dried over anhydrous $CaCl_2$ and reacted with a slight excess of trimethylsilyl azide at 0°C, with gentle formation of N_2 indicating the reaction progress. After cessation of gas formation, the reaction mixture was completely lyophilized and the residue was repeatedly washed with a volume of ethanol equal to that of the reaction mixture. Ethanol

was removed by lyophilization to give the intermediate product 3-oxauracil, the purity of which was tested via HPLC analysis.
100 mg of 3-oxauracil were dissolved in 3 ml ammonia solution (4 M) and incubated for 12 h at RT. 1 ml of NaOH (1 M) was added to the reaction mixture before lyophilization and extraction of the residue with methanol. The methanolic fraction was again lyophilized and the residue further purified by anion exchange chromatography. To this end, a Mono Q column (HR 16/10; GE HEALTHCARE) attached to an FPLC system was equilibrated with triethylammonium acetate (TEA) buffer (50 mM). The residue of the methanolic extraction was dissolved in 200 μl of TEA buffer and injected onto the column. Elution was performed by a gradient from 50 mM to 1 M of TEA buffer over 20 to 40 column volumes with a constant flow rate of 5 ml/min. Fractions, containing ureidoacrylate were identified by HPLC analysis (see subsubsection 2.2.5.5) and a RutB assay subsection 2.4.3. In HPLC experiments, purity was determined as the relative area under the curve at 280 nm. Fractions of sufficient purity were combined and lyophilized. The residue was dissolved in water and stored at -80°C after HPLC analysis and concentration measurement (ϵ_{266} = 18300 M^{-1} cm^{-1}; ϵ_{278} = 5360 M^{-1} cm^{-1}; in 40 mM Tris/HCl, pH 8.2).
Ammonium chloride buffer was not suitable for the chromatographic process, as incomplete removal of ammonia severely impeded a glutamate dehydrogenase-coupled assay (see subsubsection 2.4.3.2).

2.4.2 Protein Purification

Proteins were produced and purified as described in subsection 2.2.4. Selenomethionine-labeled RutB was produced in *E. coli* B834 (DE3) cells, using the SelenoMet™ kit (Molecular Dimensions, Rotherham, UK) following the provided protocol.

2.4.3 RutB Steady-State Enzyme Assays

2.4.3.1 Spectrophotometric Measurement

The hydrolysis of ureidoacrylate catalyzed by RutB can be followed by a direct photometric assay. The standard assay contained various concentrations of ureidoacrylate in 40 mM Tris/HCl buffer (pH 8.2) at 25°C. The reaction was started by addition of RutB enzyme and the reaction progress was followed by measuring absorption at 278 nm (ϵ_{278} = 5360 M^{-1} cm^{-1}). Due to inconsistencies in this assay, the glutamate dehydrogenase-coupled assay described below was preferred for all kinetic measurements.

2.4.3.2 Glutamate Dehydrogenase-Coupled Assay

RutB hydrolyzes ureidoacrylate into (Z)-3-aminoacrylate and carbamate. Each of these products then undergoes spontaneous hydrolysis yielding ammonia, for a total of two molecules of ammonia per molecule of ureidoacrylate hydrolyzed. The production of ammonia is coupled to the reaction of glutamate dehydrogenase (GDH), leading to NADH consumption at a rate proportional to the hydrolysis of ureidoacrylate (see Figure 2.3).

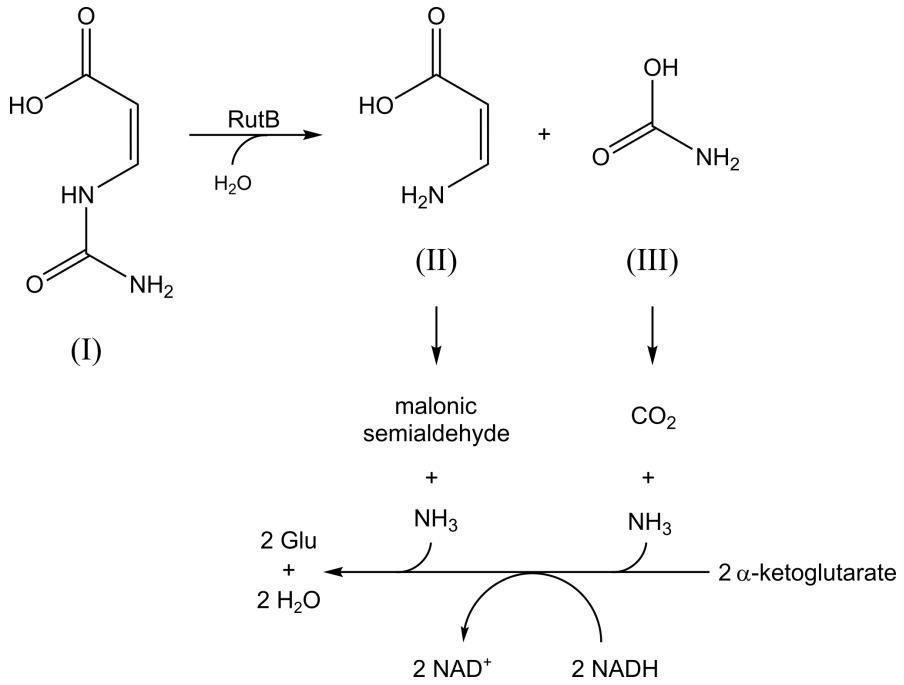


Figure 2.3: GDH-coupled assay for ureidoacrylate amidohydrolase activity of RutB
RutB catalyzes the hydrolysis of (Z)-3-ureidoacrylate (I) into (Z)-3-aminoacrylate (II) and carbamate (III). Both (II) and (III) spontaneously undergo hydrolysis, yielding one molecule of ammonia each, as well as malonic semialdehyde and carbon dioxide, respectively. Ammonia production is routinely coupled to NADH consumption by glutamate dehydrogenase (GDH), thus producing a photometric signal at 340 nm (ϵ_{340}(NADH) = 6200 M^{-1} cm^{-1}).

The standard assay for the ureidoacrylate amidohydrolase activity of RutB variants contained 600 μM NADH, 2.5 mg/ml GDH, 10 mM α-ketoglutarate (α-KG), and 2-12000 μM ureidoacrylate in a total volume of 300 μl of 100 mM Tris/HCl buffer (pH 8.2) at 25°C. The reaction was started by addition of RutB enzyme with the reaction progress being accessible through the decrease in absorption at 340 nm (ϵ_{340}(NADH) = 6200 M^{-1} cm^{-1}).
Initial rates of change of E_{340} were determined by linear regression and converted into the rate constant (v/E_0) considering the stoichiometry of 2:1 for

NADH and ureidoacrylate turnover. Rate constants were then plotted against substrate concentration and the resulting graph was fitted either with a linear function, revealing k_{cat}/K_M or with a hyperbolic function, revealing both k_{cat} and K_M.

2.4.4 Protein Crystal Structure Determination

2.4.4.1 Protein Crystallization

Solution of protein crystal structures requires protein crystals of sufficient size and purity. Optimal crystallization conditions often differ drastically between proteins and have to be empirically determined. Purified protein solutions of RutB and RutB variants were concentrated to approximately 45 mg/ml. For initial screening of crystallization conditions, 200 nl of concentrated RutB solution were mixed with 200 nl of each crystallization solution. The crystallization screens employed were MD1-30, MD1-48 (Molecular Dimensions, Rotherham, UK), and PEGRx HT (Hampton Research, Aliso Viejo, USA). Each mixture was placed as a hanging drop over a reservoir of 100 μl of the respective crystallization solution and stored at 18°C.
Crystallization conditions in which crystal formation was observed, were recreated from appropriate stock solutions and gradually modified to identify optimal additive concentrations. A drop consisting of 1 μl of protein solution (45 mg/ml) and 1 μl of crystallization solution was placed as a hanging drop over a reservoir of 500 μl of the respective crystallization solution and stored at 18°C. Additionally, wild-type RutB and RutB C166S were co-crystallized with ureidopropionate and (Z)-3-ureidoacrylate, respectively. In these instances, drops were supplemented with 40 mM and 5 mM of the respective substance. Multiple, identical hanging drops were prepared according to the best crystallization condition identified to grow a sufficient amount of crystals for data collection.

2.4.4.2 Data Collection

Data collection of protein crystals was performed by Dr. Chitra Rajendran at the Swiss Light Source (SLS) synchrotron.

2.5 AtzB Methods

2.5.1 Generation of Protein Structure Models

Structural models of AtzB and AtzB variants were generated by Dr. Julian Nazet with a local implementation of AlphaFold [30]. Resulting structures were then relaxed with the AMBER package for molecular simulation [53].

2.5.2 Chemical Methods

2.5.2.1 Synthesis of s-Triazine Compounds

The s-triazine compounds N-ethylammelide and N-isopropylammelide were required as genuine substance standards for HPLC-based qualitative enzyme assays. Since neither of the substances were available, they had to be chemically synthesized.
6 mg of 6-Chloro-1,3,5-triazine-2,4(1H,3H)-dione (CT) were solved in 300 μl KOH (0.2 M) and mixed with either 2.6 μl ethylamine solution (70%) or 3.5 μl of isopropylamine. The reaction proceeded for 5 min at 90°C. The reaction mixtures were analyzed by HPLC (see subsubsection 2.2.5.5) and mass spectrometry (see subsubsection 2.2.5.6) to confirm the identity of the respective substance.

2.5.3 Protein Purification

Proteins were produced and purified as described in subsection 2.2.4. The expression strain BW25113 Δ*guaD* was employed for all variants of AtzB to prevent any contamination with GuaD. GuaD, AtzB_Hom_Hal, and AtzB_Hom_Hal were purified from cultures of *E. coli* BL21.

2.5.4 Analytical Methods

2.5.4.1 Qualitative Enzyme Assays

In order to confirm the activity of an enzyme variant towards a certain substrate or identify product composition, qualitative assays based on HPLC analysis were established. All enzymatic assays contained 500 μM (or close to the saturation limit) of the respective substrate in a total volume of 150 μl of potassium phosphate buffer (50 mM; pH 7.0) as well as 2-10 μM of the enzyme variant to be tested. Genuine substance standards were set up in the same way but without any enzyme present. All assays were incubated for 16-20 h at RT, at which point all enzymes were removed by ultrafiltration and the filtrate was analyzed by HPLC as described (see subsubsection 2.2.5.5). Substrates tested in this manner included atrazine, dechloroatrazine, cytosine, isocytosine, adenine, guanine, isoguanine, melamine, ammeline, and ammelide.

2.5.4.2 Steady-State Enzyme Assays

2.5.4.2.1 Hydroxyatrazine Ethylaminohydrolase Assay

Quantification of enzyme activity was essential to assess the effects of the various amino acid exchanges introduced. AtzB hydrolyzes hydroxyatrazine to give ethylamine and N-isopropylammelide. This conversion produces a photometric signal with a differential extinction coefficient of $\Delta\epsilon_{242}$= 10957$M^{-1}cm^{-1}$. This

value was derived from the differential spectrum of a) genuine hydroxyatrazine and b) the reaction mixture of complete conversion of hydroxyatrazine by AtzB (adjusted for the absorption of the AtzB enzyme).
A standard assay contained 5-50 μM of hydroxyatrazine in a total volume of 300 μl of potassium phosphate buffer (50 mM; pH 7.5) as well as 1 nM-1 μM of the enzyme variant to be tested. Reactions proceded at 25°C. Photometric measurements were taken either in a cuvette with a pathlength of 1 mm or in a 96 well plate (with a resulting pathlength of 8.3 mm). Initial rates of change of E_{242} were determined by linear regression and converted into the rate constant (v/E_0). Rate constants were then plotted against substrate concentration and the resulting graph was fitted either with a linear function, revealing k_{cat}/K_M or with a hyperbolic function, revealing both k_{cat} and K_M.

2.5.4.2.2 Guanine Deaminase Assay

The guanine deaminase activity of certain AtzB variants was quantified by direct photometric measurements. Hydrolysis of guanine, resulting in xanthine and ammonia, produces a photometric signal with a differential extinction coefficient of $\Delta\epsilon_{253} = 5023 M^{-1}cm^{-1}$. This value was derived from the differential spectrum of genuine guanine and xanthine in aqueous solution.
A standard assay contained 10-300 μM of guanine in a total volume of 300 μl of potassium phosphate buffer (100 mM; pH 7.5) as well as 0.1-5 μM of the enzyme variant to be tested. Reactions proceded at 25°C. Photometric measurements were taken either in a cuvette with a pathlength of 1 mm or in a 96 well plate (with a resulting pathlength of 8.3 mm).
Initial rates of change of E_{253} were determined by linear regression and converted into the rate constant (v/E_0). Rate constants were then plotted against substrate concentration and the resulting graph was fitted either with a linear function, revealing k_{cat}/K_M or with a hyperbolic function, revealing both k_{cat} and K_M.

2.5.4.2.3 N^2,N^2-Dimethylguanine Dimethylaminohydrolase Assay

N^2,N^2-Dimethylguanine dimethylaminohydrolase activity, a novel enzyme activity discovered in this work, was assayed by direct photometric measurements. Hydrolysis of N^2,N^2-dimethylguanine, resulting in xanthine and dimethylamine, produces a photometric signal with a differential extinction coefficient of $\Delta\epsilon_{294} = 3187\ M^{-1}cm^{-1}$. This value was derived from the differential spectrum of genuine N^2,N^2-dimethylguanine and xanthine in aqueous solution.
A standard assay contained 50-8000 μM of N^2,N^2-dimethylguanine in a total volume of 300 μl of potassium phosphate buffer (100 mM; pH 7.5) as well as 50 nM of the enzyme variant to be tested. Reactions proceded at 25°C. Photometric measurements were taken in a cuvette with a pathlength of 1 mm.

Initial rates of change of E_{294} were determined by linear regression and converted into the rate constant (v/E_0). Rate constants were then plotted against substrate concentration and the resulting graph was fitted either with a linear function, revealing k_{cat}/K_M or with a hyperbolic function, revealing both k_{cat} and K_M.

3 Evolutionary Origin of the Enterobactin Pathway Proteins EntB and EntA

3.1 Introduction

3.1.1 Primary and Secondary Metabolism

The entirety of all metabolic reactions found in Nature can roughly be separated into primary and secondary metabolism [54]. Primary metabolism encompasses all reactions necessary for the synthesis of compounds that are essential for survival and procreation of an organism. Such essential metabolites are the building blocks of cellular life (e.g. fatty acids, amino acids, nucleotides) and intermediates of energy metabolism (e.g. sugars, ATP).
The enzymes and enzymatic pathways of primary metabolism are assumed to have emerged early during evolution [55], either by retrograde evolution [6] or through specialization of ancestral, low-specificity enzymes [8]. Consequently, the last universal common ancestor (LUCA) of all cellular life likely already possessed a rather complex primary metabolism [56] [57].
Later in evolution, many organisms have evolved secondary metabolic pathways, producing a wide array of advantageous but non-essential compounds. These secondary metabolites can have anti-bacterial [58], anti-viral [59] [60], and anti-fungal [61] effects, act as repellants or toxins [62] [63], as acquisition systems for nutrient metals [46] [64] [65], or as regulators of symbiotic [66] and parasitic [67] [68] relationships. Production of each of these functional molecules provides a specific selection advantage in certain environments or under certain growth conditions, justifying their metabolic cost.
Due to their effect on human physiology, certain secondary metabolites have long been used in the form of natural medicine and have more recently become the focus of pharmaceutical research. Many secondary metabolites feature functional architectures that have been the basis for the development of various pharmaceuticals like antibiotics [69], anti-tumor drugs [70] [71], and analgesics [72] [63] [73].
It is generally accepted, that most novel enzymatic functions arise from pre-existing enzymes by gene duplication and divergence (see section 1.1). Following a gene duplication event, one of the copies can maintain its original function,

while genetic drift and natural selection can introduce or improve a novel activity in the other copy. Over the years, different models have been devised and gradually refined to more accurately describe the underlying principles of the emergence of new enzyme functions. The Innovation-Amplification-Divergence (IAD) model proposed by Bergthorsson and coworkers [9] poses that an ancestral gene first randomly acquires a physiologically relevant side activity. In this context, increased gene dosage is beneficial, leading to stable gene duplication, which then allows for the divergence of the different gene copies. Eventually, one gene copy will acquire an increased activity of the novel function while the ancestral function is preserved in another gene copy.
In the context of the initial emergence of secondary metabolism during evolution, this translates to the recruitment of pre-existing primary metabolic genes by duplication and divergence, leading to genes coding for new secondary metabolic enzymes [74] [75]. Generally, different evolutionary dynamics apply to primary and secondary metabolic enzymes. The latter generally possess lower activities [76] and substrate specificities [77], which may serve the exploration of metabolic space and maintain evolvability of these enzymes [75]. Organization of multiple functionally related secondary metabolic genes into biosynthetic gene clusters [78] [74] gives rise to modular enzyme pathways and facilitates horizontal gene transfer of entire pathways or functionally discrete sub-clusters [74]. This process has occurred in different species at different times during evolution, leading to relatively narrow phylogenetic distributions of orthologous groups of secondary metabolic enzymes and pathways [79].
The evolutionary relationship between homologous primary and secondary metabolic enzymes has been investigated in some cases. The emergence of the secondary metabolic quinate dehydrogenase (QDH) function from a primary metabolic shikimate dehydrogenase (SDH) was investigated by ancestral sequence reconstruction and QDH activity could be established in an extant SDH with only a single residue exchange [80]. Similarly, a primary metabolic kaurene synthase, involved in the synthesis of the essential phytohormone gibberelin, was converted into a secondary metabolic pimaradiene synthase with a single residue exchange [81]. Recently our group achieved the conversion of primary metabolic anthranilate synthase (TrpE) from *E. coli* into a secondary metabolic isochorismate synthase (EntC) by only two residue exchanges [82], prompting analogous projects aiming at the establishment of isochorismatase (EntB) and 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EntA) activity in their respective primary metabolic homologues, which are described in this work.

3.1.2 The Enterobactin Pathway

Almost all organisms require iron as a protein cofactor and for electron transfer [83]. As bioavailability of iron is naturally limited due to its low solubility (c[Fe^{3+}] 10^{-18} M at pH 7.4) [84]), various strategies for iron acquisition have

evolved [85] [86], which are considered as part of the secondary metabolism due to their non-essential nature. Notably, such iron acquisition strategies play an important role in pathogenic species, which must compete with various iron sequestering agents employed by their hosts [87] [88] [89] [90] [91].
Siderophores are specific high affinity compounds for iron complexation found in bacteria [85], fungi [92], and plants [93] [94]. These small molecular substances are typically produced intracellularly and exported into the surrounding where they tightly bind available iron. The resulting ferric complexes are then reimported into the organism where iron is released from the siderophore.
One particular, well researched siderophore is enterobactin, which is employed by *E. coli* [95] and a number of other bacterial species [96] [97]. Enterobactin production (see Figure 3.1) proceeds from chorismate, which is also required in the primary metabolic synthesis of aromatic amino acids [98] and folic acid [99]. In the enterobactin (Ent) pathway, chorismate is first reversibly converted to isochorismate by isochorismate synthase (EntC). Successive conversion of isochorismate by isochorismatase (EntB) and 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EntA) gives 2,3-dihydroxybenzoate (2,3-DHB) [46], a compound which can already function as a siderophore [100]. Three molecules of 2,3-DHB are combined with three molecules of serine by non-ribosomal peptide synthesis to give the trilactone enterobactin [101]. This process is catalyzed by the cooperation of the aryl carrier domain of EntB, as well as EntD, EntE, and EntF [102] [103].

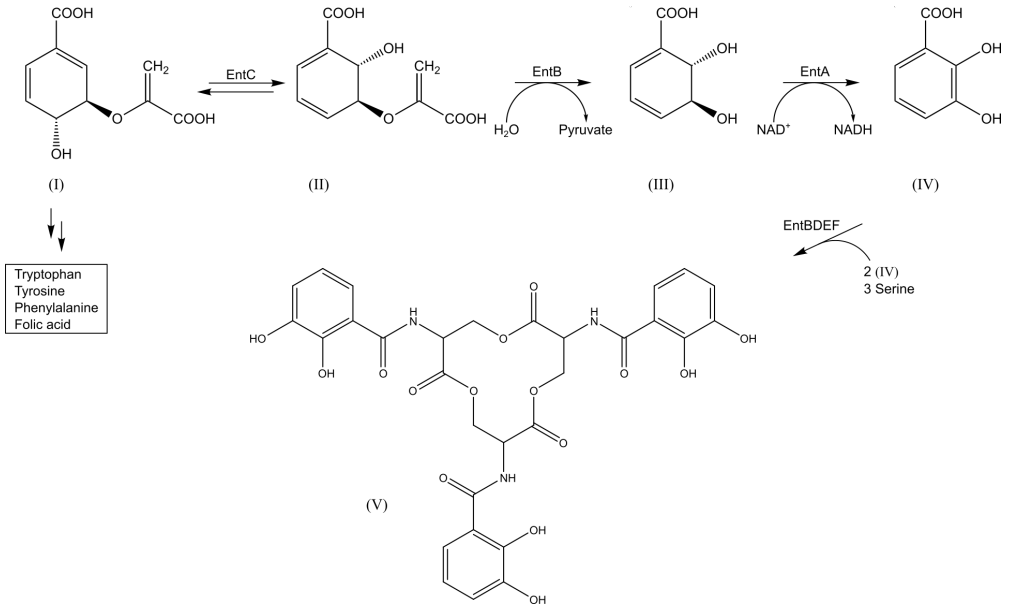


Figure 3.1: The enterobactin pathway
Enterobactin is synthesized by the enzymes of the enterobactin (Ent) pathway [46] by a series of enzymatic conversions (EntABC), followed by non-ribosomal peptide synthesis (EntBDEF). The primary metabolite chorismate (I), which is also an intermediate in the production of the aromatic amino acids and folic acid, is reversibly converted into isochorismate (II) by the action of isochorismate synthase (EntC). Isochorismatase (EntB) converts (II) into 2,3-dihydro-2,3-dihydroxybenzoate (III) with pyruvate as a side product. (III) is further oxidized by 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EntA) in an NAD^+-dependent reaction, producing 2,3-dihydroxybenzoate (IV). Three molecules each of (IV) and serine are assembled into one molecule of enterobactin (V) by a non-ribosomal peptide synthesis mechanism involving the aryl carrier domain of EntB, 2,3-dihydroxybenzoylserine synthetase (EntD), and the multi-domain proteins EntE and EntF [103] [104].

Like other siderophores, enterobactin is exported into the environment where it tightly binds iron ions ($K_a = 10^{51}M^{-1}$ [105]) and the resulting complex is selectively reimported into the cell by a number transport proteins [106] [107]. In the cytoplasm, ferric enterobactin is partially hydrolyzed by the esterase Fes [108] leading to iron release. The breakdown products can then further serve as lower-affinity siderophores [109].

3.1.3 Primary Metabolic Homologues of Enterobactin Pathway Enzymes

Operating under the assumption that secondary metabolic enzymes evolved from primary metabolic progenitors (see subsection 3.1.1), the closest homo-

logues within an organism are likely to share a common ancestor. Ureidoacrylate amidohydrolase (RutB) and 3-oxoacyl-[acyl-carrier-protein] reductase (FabG) were identified as the closest primary metabolic homologues of EntB and EntA, respectively, in *E. coli* [20].

3.1.3.1 Ureidoacrylate Amidohydrolase (RutB)

Ureidoacrylate amidohydrolase (RutB) is part of the recently discovered Rut pathway [110], that enables growth of *E. coli* on pyrimidines as a sole nitrogen source [52]. The enzyme was also characterized in detail in this work (see chapter 4). RutB catalyzes the hydrolysis of ureidoacrylate into aminoacrylate and carbamate [52]. The enzyme shares its isochorismatase-like hydrolase (IHL) fold with EntB (see Figure 3.2), although it lacks the latter's aryl carrier domain (ArCP). RutB employs a catalytic triad, consisting of D24, K133, and C166, with the cysteine residue initiating the reaction by a nucleophilic attack to the ureido group of ureidoacrylate (see Figure 4.10), forming a covalent intermediate. After the release of the aminoacrylate moiety of the molecule, the remaining carbamate portion is hydrolytically cleaved from C166, restoring the starting state.

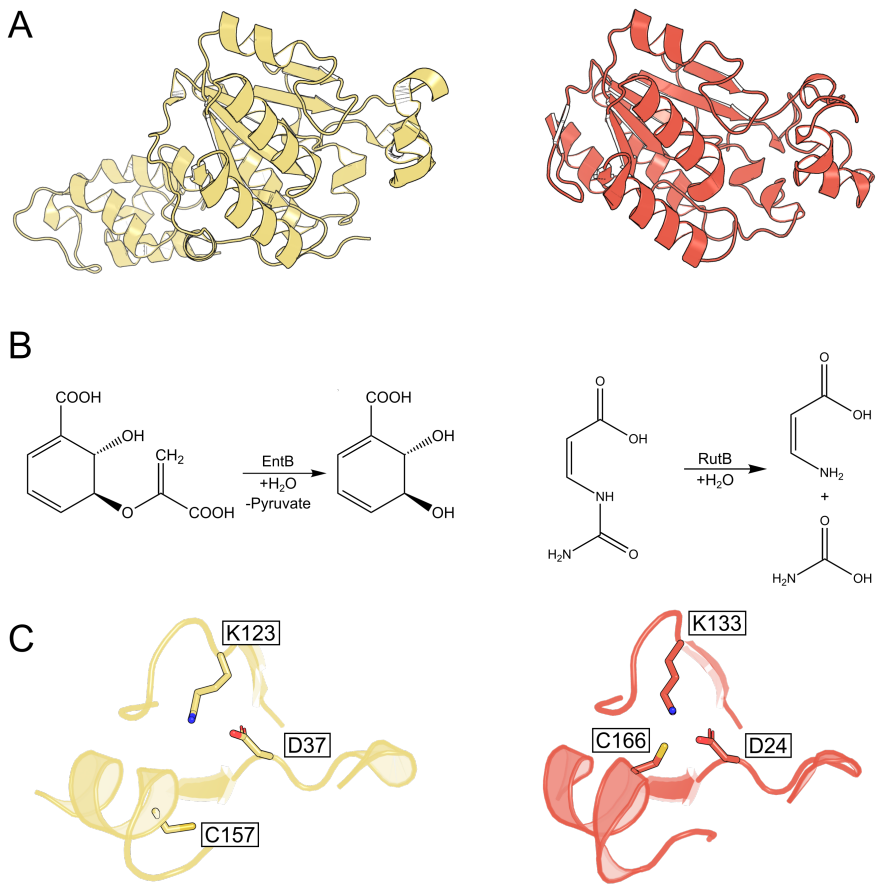


Figure 3.2: Comparison of EntB and RutB
(A) Isochorismatase (EntB; yellow; PDB: 2FQ1; [111]) and ureidoacrylate amidohydrolase (RutB; red; this work) share the isochorismatase-like hydrolase fold. (B) The reactions catalyzed feature different chemistries but share a hydrolytic character. (C) EntB and RutB may share a catalytic triad consisting of D37, K123, C157 and D24, K133, C166, respectively.

EntB shares two residues of the catalytic triad of RutB (D37, K123; [111]) but the closest cysteine residue (C157) is found in a position different from C166 in RutB and may not actually take part in the reaction. Accordingly, the reaction mechanism of EntB may be more accurately be represented by the homologous isochorismatase PhzD [112] which employs a catalytic aspartate (D38) for the hydrolysis of the vinyl ether portion of isochorismate.

3.1.3.2 3-Oxoacyl-[Acyl-Carrier-Protein] Reductase (FabG)

The closest primary metabolic homologue to EntA in *E. coli* is the 3-oxoacyl-[acyl-carrier-protein (ACP)] reductase (FabG). This enzyme is part of the bacterial fatty acid synthase [113], which is essential for membrane formation. FabG

reduces ACP-bound 3-oxoacyl compounds in an NADPH-dependent reaction to their respective (3R)-3-hydroxyacyl-ACP products [114]. Crystal structures of FabG from *E. coli* exist as both the *apo* (PDB: 1I01; [115]) and the *holo* form (PDB: 1Q7B; [116]) and evidence has been found for allosteric communication in the FabG homotetramer [115] [117]. The interaction of FabG with ACP is essential *in vivo* [113] but a number of *in vitro* activities on coenzyme A bound 3-oxoacyl substrates have been described [118] [119], indicating that ACP is not inherently essential to FabG activity. The molecular basis of binding of ACP to FabG has been investigated in detail [120].
Both FabG and EntA are part of the short-chain dehydrogenase/reductase (SDR) superfamily and share a Rossmann fold [121] (see Figure 3.3). The reactions catalyzed by FabG and EntA differ significantly in that they proceed in the reductive and the oxidative direction, respectively, *in vivo*. While the existence of a catalytic triad consisting of S138, Y151, and K155 has previously been proposed for FabG, based on inference from other SDR enzymes [115], a recent inquiry has shown that none of these residues is absolutely required for FabG activity [122], implying a more complicated reaction mechanism than encountered in other SDR enzymes. EntA shares the previously assumed catalytic triad of FabG (EntA: S131, Y144, K148).

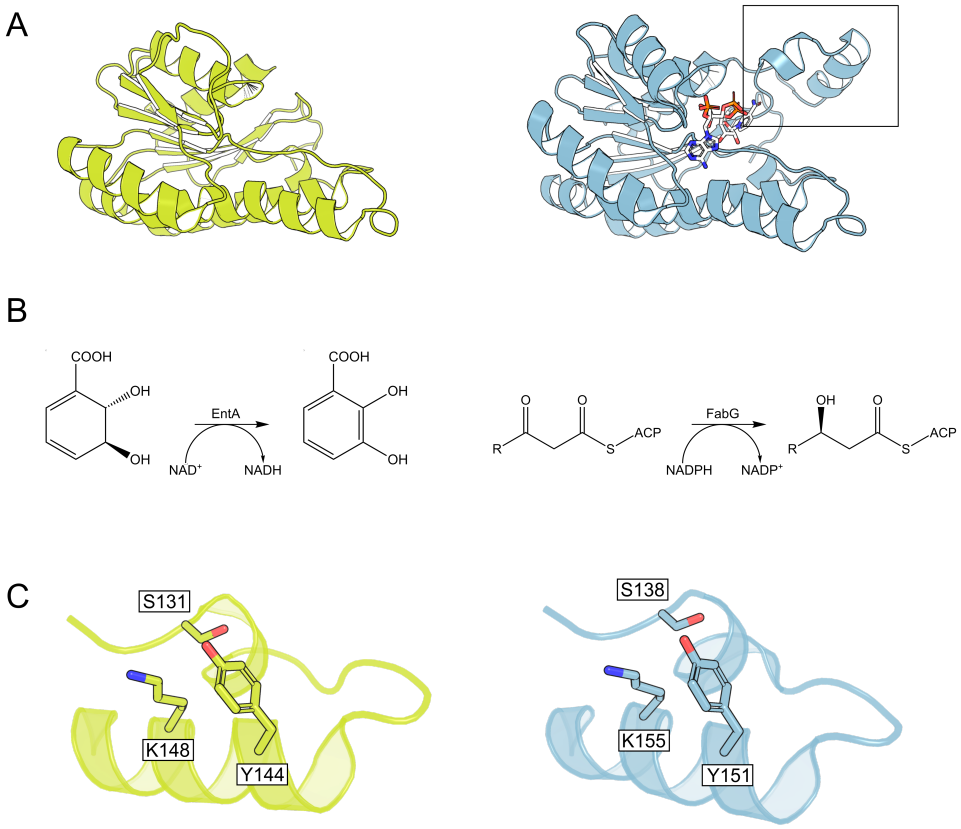


Figure 3.3: Comparison of EntA and FabG
(A) 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EntA; green; PDB: 2FWM; [123]) and 3-oxoacyl-[acyl-carrier-protein (ACP)] reductase (FabG; blue; PDB: 1Q7B; [116]) both belong to the short-chain dehydrogenase/reductase (SDR) superfamily and exhibit a Rossmann fold, which is often encountered in dinucleotide-utilizing enzymes [124]. A loop structure indicated in FabG (black box) was not resolved in the structure of EntA. (B) Both enzymes catalyze redox reactions, although the co-factor utilized and the oxidative direction differ. (C) A catalytic triad, typical for SDR enzymes, exists in both EntA (S131, Y144, K148) and FabG (S138, Y151, K155), although it is not absolutely required for catalysis in FabG [122].

Following the approach of Plach and coworkers [82], functional conversion of the two primary metabolic enzymes into their secondary metabolic homologues was attempted in this work.

3.2 Results

3.2.1 Bioinformatics

Various bioinformatic methods offer the potential to design specific properties of an enzyme. Two of those were employed to achieve the desired functional conversion of RutB and FabG into EntB and EntA, respectively. FuncLib [24] utilizes phylogenetic information and Rosetta design to suggest a number of stable active site mutations, which may potentially alter the substrate range and activity of the targeted enzyme. JANUS [25] compares multiple sequence alignments (MSAs) of two homologous enzyme families to specifically identify and rank such residue exchanges that could bring about functional conversion when introduced.
For the analysis of RutB, FuncLib (see subsubsection 2.3.1.1) was provided its crystal structure (see chapter 4) and the residue positions to be designed (29, 35, 41, 43, 72, 74, 92, 136, 162, 165, 191) while the catalytic triad (D24, K133, C166) was set to be conserved. For the design of FabG, the PDB structure 1Q7B [116] was used. The residue positions to be altered were 90, 92, 138, 140, 148, 183, 186, 188, 189, and 192 while the previously assumed catalytic triad (S138, Y151, K155) as well as the ligand $NADP^+$ were kept unchanged. For RutB, a number of enzyme variants with five residue exchanges each were predicted. 17 of the 18 highest ranked variants were chosen for analysis (see Table 8.3). In the case of FabG, FuncLib similarly proposed a number of enzyme variants, with the number of residue exchanges ranging from three to five. Here, the 20 highest ranked variants were selected for analysis (see Table 8.3). All selected enzyme variants were analyzed by an *in vivo* assay (see subsubsection 2.3.3.2) for their ability to remedy specific deletions in the *ent* operon (see subsubsection 3.2.4.3). Variants restoring growth under selective conditions were expressed, purified, and tested by an HPLC based assay (see subsection 2.3.6) for their respective target activities (see subsection 3.2.7).
JANUS (see subsubsection 2.3.1.2) was provided with MSAs of RutB and EntB or FabG and EntA together with the structures of RutB (see chapter 4) or FabG (PDB: 1Q7B), respectively. The best 15 residue exchanges suggested for the conversion of RutB into EntB were V23H, A28Y, Y29F, F70T, G73A, W74Q, L108G, A162Y, N164H, V165I, C166G, H189A, Q190D, A191F, and G192S. The 16 most highly rated residue exchanges for the conversion of FabG into EntA were I21V, N59D, N86A, T90L, D92L, I106T, I107F, T109V, L111V, I136V, V139N, G144P, Q148M, N150A, F183S, and I184T. All of these residue exchanges were jointly introduced into the respective enzyme, giving rise to the variants RutB2EntB15 and FabG2EntA16 which were purified and tested by an HPLC based assay (see subsection 2.3.6) for their respective target activities (see subsection 3.2.7).

3.2.2 Generation of Deletion Strains

Specific deletion strains were required as expression strains and as selection strains for complementation studies. In the first case, the deletion of a chromosomal gene of *E. coli* was introduced to entirely prevent the respective gene product from being produced and possibly persisting throughout protein purification. This measure was employed for the production of RutB and FabG variants possibly exhibiting a miniscule EntA or EntB activity, respectively. The same strains were used for *in vivo* studies to test whether a deletion of *entB* or *entA* can be rescued by a functional RutB or FabG variant, respectively. Deletion of the *fec* operon, responsible for citrate-mediated iron uptake, was devised to prevent complementation *via* this mechanism.
Required deletion strains were generated by homologous recombination in the strain *E. coli* DY329, followed by P1 phage transduction into *E. coli* BW25113 [27], as described in subsubsection 2.2.3.10. In this process, each targeted genomic region was selectively replaced by either a chloramphenicol O-acetyltransferase or an aminoglycoside phosphotransferase, making the strain resistant to the antibiotic chloramphenicol or kanamycin, respectively. Deletions were designed in such a way that the entire open reading frame of a gene was removed, except for its first codon and a brief 3' terminal region which may contain the ribosome binding site (RBS) of the following gene. The deletion of the *fec* operon (*fecIRABCDE*) would similarly leave only the initial codon of *fecI* and the 3' terminal region of *fecE*. In the case of the *entB(1-561)*::*Cam* deletion, only the portion of the gene coding for the isochorismatase domain of EntB was removed and an optimized RBS was introduced upstream of the remaining portion of the *entB* gene to allow for the standalone expression of the aryl carrier domain of EntB. The strains created in our lab were BW25113 *entA*::*Cam* (Δ*entA*), BW25113 *rutB*::*Cam* (Δ*rutB*), BW25113 *entB(1-561)*::*Cam* (Δ*entB-IC*), BW25113 *fecIRABCDE*::*Kan* (Δ*fec*), BW25113 *entA*::*Cam fecIRABCDE*::*Kan* (Δ*entA*Δ*fec*), and BW25113 *entB(1-561)*::*Cam* (Δ*entB-IC*Δ*fec*), while strain BW25113 *entB*::*Kan* (Δ*entB*) existed as part of the KEIO collection [29] and was purchased from the Coli Genetic Stock Center (CGSC) at Yale University.
In all strains, presence and integrity of the respective deletion was confirmed by colony PCR (see subsubsection 2.2.3.2) and sequencing of amplified genomic sequences. Colony PCRs were negative for all deleted genes and positive for genes adjacent to the respective deletion. Sequencing revealed that all deletions had occurred exactly as intended, replacing the desired portion of DNA with a resistance cassette for chloramphenicol or kanamycin and leaving intact the genomic regions directly adjacent to the insertion site. This latter fact is of particular importance for *in vivo* complementation studies, as enterobactin pathway genes only have very brief intergenic regions between their open reading frames and even overlap by one nucleotide in the case of *entB* and *entA* and partial deletion of the adjacent genes would unpredictably disrupt the enterobactin pathway.

3.2.3 Generation of Gene Libraries

3.2.3.1 Fully Randomized Gene Libraries

Gene libraries in which the entirety of the open reading frame (ORF) of a gene is randomized offer the potential to sample through a large number of enzyme variants when combined with a suitable selection system. This approach may lead to unexpected new solutions to the underlying question.
Full randomization of the genes *rutB* and *fabG* from *E. coli* was achieved by error prone PCR (epPCR), caused by addition of $MnCl_2$ and the use of an unevenly distributed mixture of dNTPs in the reaction mixture (see subsubsection 2.3.4.1). Amplicons were cloned into the low level constitutive expression plasmid pExp (see subsubsection 2.3.4.2) and amplified in *E. coli* NEB Turbo. The number of distinct colonies in this step was determined by serial dilution and the frequency of nucleotide mutations introduced was identified by DNA sequencing of eight plasmids.
The fully randomized gene library of *rutB* produced 3.2e8 distinct colonies. Sequencing of eight variants revealed seven to 14 nucleotide mutations per gene copy (∅10.3) which translate to five to eight amino acid exchanges (∅6.5). Two genes contained a single STOP codon each and one nucleotide deletion was present in another variant. Two of the plasmids sequenced contained no ORF, leaving only three of eight plasmids with a viable gene, coding for a full-length RutB variant. The total number of distinct gene variants coding for full-length proteins can thus be given as 1.2e8.
In the case of the gene library of *fabG*, the number of distinct clones grown was 2.0e7. Based on the sequencing results of eight variants, the nucleotide mutations per gene copy ranged from five to 17 (∅12). These mutations correlate to four to 15 amino acid exchanges (∅8.7). Five of the nucleotide exchanges caused an internal STOP codon to appear and three nucleotide deletions were observed. Additionally one of the plasmids sequenced contained no ORF. Again, only three of the eight plasmids can be expected to produce a full-length protein. Based on this limited sample, the number of distinct viable enzyme variants coded for by the gene library can be given as 7.5e6.
Overall, the limited number of viable plasmids contained within the gene libraries is to be expected from a random mutagenesis method and the plasmid variants not coding for full-length proteins should not disrupt the later selection process as they would simply be lost due to being unproductive. The number of residue exchanges per enzyme is likely high enough to allow for the functional changes desired. The number of viable variants in the *rutB* gene library is close to what can be expected from this method [26]. While the *fabG* gene library contains significantly fewer viable variants, it is comparable in size to successfully employed gene libraries created by similar methods [125]. Overall, the size and composition of the fully randomized gene libraries described here are satisfactory for further experiments.

3.2.3.2 Focused Randomization of RutB

RutB contains a set of three aromatic residues (Y29, Y35, F41) that, together with other nearby residues (L36, Y136), form a hydrophobic pocket as part of the active site cavity. Taking into account the proximity and direction of the three aromatic residues to the bound substrate (chapter 4) as well as variations of these positions encountered in other isochorismatase-like hydrolase (IHL) enzymes [126] implied that the identity of these residues might modulate the substrate specificity of RutB.
Focused randomization of the three correlating codons was achieved by executing the standard mutagenesis protocol (see subsubsection 2.2.3.3) with two primers carrying a total of three degenerated codons (FlexLoop-Rand1 and FlexLoop-Rand2 in Table 8.1). The gene library was transformed into NEB Turbo resulting in the growth of 2.3e4 distinct colonies. Four distinct plasmids were sequenced, revealing nucleotide exchanges at each of the targeted codons.
Overall, the focused mutagenesis of positions Y29, Y35, and F41 was successful, with the number of distinct plasmid variants produced likely sufficient to capture the full chemical diversity of possible residue combinations. The gene library was expressed in BW25113 Δ*entB* and the resulting mixture of enzyme variants was analyzed by an HPLC based assay for EntB activity (see subsection 2.3.6).

3.2.4 Selective Growth Conditions

3.2.4.1 Establishment of Selection Conditions for Enterobactin Producing Strains

Establishing medium conditions that selectively allow growth of enterobactin producing strains is a prerequisite for later *in vivo* complementations experiments. In the initial identification of the enterobactin pathway genes, the chelating agent 2,2'-bipyridine (BPD) has been used to reduce the bioavailability of iron in minimal media [46] and prevent growth of enterobactin non-producing strains. Other chelators, like ethylenediaminetetraacetic acid (EDTA), might potentially function in a similar manner even though antimicrobial effects have been observed at higher concentrations [91].
To identify optimal conditions for selection and complementation experiments, 5 ml portions of M9 minimal medium (see subsubsection 2.1.3.3) were supplemented with different concentrations of either BPD or EDTA, with wild-type *E. coli* BW25113 and enterobactin deficient deletion strains (see subsection 3.2.2) serving as positive and negative controls, respectively. Optimal chelator concentrations, that did not significantly inhibit growth of the wild type but fully inhibited growth of the deletion strains were 75 μM BPD and 1 μM EDTA.
Metabolic complementation of the *entA* gene deletion was achieved in M9-BPD medium (subsubsection 2.1.3.3) by addition of 20 μM of 2,3-dihydroxybenzoate (2,3-DHB), the product of the EntA reaction. Supplementing this metabolite un-

der iron limiting conditions has previously been shown to restore enterobactin production by serving as a substrate for the lower part of the enterobactin pathway [46]. 2,3-DHB did not restore growth of Δ*entB*, because *entB* also encodes an aryl carrier protein (ArCP) domain required downstream in the pathway. Consequently, an *entB-IC* knockout strain, containing only the deletion of the region of the *entB* gene coding for the isochorismatase domain of EntB (residues 1-187) was produced, which could be rescued by 2,3-DHB. Successful *in vivo* complementation was achieved in each deletion strain by transforming it with a pExp plasmid carrying the respective gene deleted from the genome. The results of all growth conditions are summarized in Table 3.1.

Table 3.1: Growth of strains in iron deficient liquid media
Growth of different *E. coli* strains under various conditions was assayed in a volume of 5 ml at 37°C while shaking at 120 rpm for 24 h. The wild-type *E. coli* BW25113 strain (see Table 2.3) displayed growth in M9 minimal medium, reaching an OD_{600} of approximately 1.0 after 24 h (+++). Strains with any deletion in the enterobactin pathway (Δ*entB*, Δ*entB-IC*, Δ*entA*) grew slightly worse, reaching OD_{600} values of approximately 0.8 (++).
Under iron limiting conditions, induced by addition of a chelator (75 μM 2,2'-bipyridine (BPD) or 1 μM ethylenediaminetetraacetic acid (EDTA)), the enterobactin producing strains still reached an OD_{600} of approximately 1.0 after 24 h, while enterobactin non-producing strains remained at an $OD_{600} < 0.05$ (-). Strains lacking *entA* or the part of *entB* coding for the isochorismatase domain of EntB (Δ*entB-IC*) could be rescued by addition of 20 μM 2,3-dihydroxybenzoate (2,3-DHB). Growth of deletion strains could also be restored by a plasmid borne copy of the chromosomally deleted gene but not by the null vector (pExp) or by vectors carrying the wild-type sequences of *rutB* or *fabG*. Additional deletion of the *fec* operon (Δ*fec*) did not have a measurable effect on the growth behavior of the deletion strains.

Strain	Plasmid	Chelator	2,3-DHB	Growth
BW25113	-	-	-	+++
Δ*entB*	-	-	-	++
Δ*entB-IC*	-	-	-	++
Δ*entA*	-	-	-	++
BW25113	-	+	-	+++
Δ*entB*	-	+	-	-
Δ*entB-IC*	-	+	-	-
Δ*entA*	-	+	-	-
Δ*entB*	-	+	+	-
Δ*entB-IC*	-	+	+	+++
Δ*entA*	-	+	+	+++

Strain	Plasmid	Chelator	2,3-DHB	Growth
Δ*entB* Δ*fec*	-	+	+	-
Δ*entB-IC* Δ*fec*	-	+	+	+++
Δ*entA* Δ*fec*	-	+	+	+++
Δ*entB*	*pExp::entB*	+	-	+++
Δ*entB-IC*	*pExp::entB-IC*	+	-	+++
Δ*entA*	*pExp::entA*	+	-	+++
Δ*entB* Δ*fec*	*pExp::entB*	+	-	+++
Δ*entB-IC* Δ*fec*	*pExp::entB-IC*	+	-	+++
Δ*entA* Δ*fec*	*pExp::entA*	+	-	+++
Δ*entB*	*pExp*	+	-	-
Δ*entB-IC*	*pExp*	+	-	-
Δ*entA*	*pExp*	+	-	-
Δ*entB-IC*	*pExp::rutB*	+	-	-
Δ*entA*	*pExp::fabG*	+	-	-

BPD and EDTA could also be utilized to suppress growth of enterobactin non-producing strains on solid growth medium, although higher concentrations of the chelators were required. Enterobactin non-producing strains were prevented from growing on M9 agar containing 200 μM BPD, while growth of the wild-type strain was normal under these conditions. Enterobactin deficient strains grew at BPD concentrations of 150 μM and lower, although they could still be separated from the faster growing wild-type strains under the right growth conditions (37°C; 16 h). Growth of enterobactin non-producing strains was visibly hindered on M9 agar plates with EDTA concentrations starting from 50 μM. At higher concentrations of EDTA, growth of both wild-type and enterobactin deficient strains was gradually reduced. The best conditions to separate wild-type and enterobactin deficient strains were 500 μM EDTA with the right growth conditions (37°C; 16 h). Growth of enterobactin non-producing strains was still observed at longer incubation times or when growing close to enterobactin producing colonies. The latter effect is probably due to diffusion of enterobactin through the agar [127]. Wild-type strains no longer grow at EDTA concentrations exceeding 2 mM.

Overall the growth conditions identified in this work are a robust way to select for enterobactin producing strains and thus to search a gene library for variants able to complement a specific deletion in the enterobactin pathway. BPD is more suitable as a selection factor than EDTA, because the latter also inhibits the growth of wild-type *E. coli* at concentrations necessary for full suppression of the deletion strains, probably by chelation of other essential metal ions in the medium [91].

3.2.4.2 Gene Library Selection

The fully randomized gene libraries based on *rutB* and *fabG* created in this work (see subsubsection 3.2.3.1) were transformed into *E. coli* BW25113 Δ*entB-IC* and Δ*entA*, respectively. Efficient transformation was achieved by electroporation (see subsubsection 2.3.4.3).
E. coli BW25113 Δ*entB-IC* cells were transformed with the *rutB* gene library (1.4e10 transformants; i.e. 44-fold more transformants than distinct variants contained in the library) before being cultivated on M9 agar plates containing either 50 μM or 500 μM EDTA (see subsubsection 3.2.4.1). Single colonies grew on the plates after incubation at 37°C for 7 days but often failed to yield a sequencing read of their respective plasmid or revealed a null vector. Most *rutB* variants that could be sequences contained an early STOP codon or nucleotide deletion, implying a truncated gene product. A total of eight *rutB* variants could be identified and are listed in Table 3.2.

Table 3.2: Variants isolated from the *rutB* gene library.
Variants of *rutB* isolated from the fully randomized gene library (see subsubsection 3.2.3.1) that lead to growth of BW25113 Δ*entB-IC* under iron limiting conditions (M9 agar + 50/500 μM EDTA) were identified by DNA sequencing. The table shows the corresponding amino acid exchanges. Synonymous exchanges are shown in square brackets. The first three variants were isolated from M9 agar plates containing 50 μM EDTA and the remaining variants stem from plates containing 500 μM EDTA. Mutations after the first STOP codon or deletion are not shown.

ID	Mutations
50-1	T55DEL
50-2	Q26H, N52D, T58DEL
50-3	[S44S], M65V, W69STOP
500-1	Q26H, L38F, [S44S], [P48P], G83R, G85D, [N93N], A94T, P124S, [L148L], G160D, F202S, W210STOP
500-2	D37V, Y78STOP
500-3	Q26H, M65V, [G83G], [L95L], [L104L], Q117L, F158I, [C166C], A191V
500-4	Q26R, V43D, A60T, P84H, R99C, Q101STOP
500-5	P32DEL

E. coli BW25113 Δ*entA* cells were transformed with the *fabG* gene library (2.4e6 transformants; 12% coverage) and cultivated in six separate liquid precultures before being subjected to iron limiting conditions (M9-BPD medium). Following growth for 48 h, each culture was streaked on agar plates and incubated at 37°C. The plasmids of individual colonies were sequenced, which revealed

that *fabG* variants isolated from a given selection culture were mostly identical, indicating a selection advantage of the respective gene product. All variants isolated from the *fabG* gene library in this manner are listed in Table 3.3.

Table 3.3: Variants isolated from the *fabG* gene library.
Variants of *fabG* isolated from the fully randomized gene library (see subsubsection 3.2.3.1) that lead to improved growth of the BW25113 Δ*entA* strain under iron limiting conditions were identified by DNA sequencing. The table shows the corresponding amino acid exchanges. Synonymous exchanges are shown in square brackets. The first group of variants were the most abundant in their respective liquid culture, while the second group was less prevalent. Mutations after the first STOP codon or deletion are not shown.

ID	Mutations
A1	K72E, N85D, [N86N], [I89I], V114I, [S118S], R123H, M143I, A152D, S162C, R172C, M188L, E214G, E226K
A2	V139A, G182D
A3	V32D, A42S, G50D, K54STOP
A4	[A44A], N93S, R129C, N150D, A166T, [T175T], D187E, R208H, I215F, E226D, E233V
A5	L9M, V80E, K99E, S138P, A166T, A180V, T189S, A191V, Q203H, [A220A]
A6	S38N, A64T, R74C, D100V, L111P, M126T, M143L, T186M, R190H, A202E, L209P
B1	G18S, [E23E], S38I, N93S, E101K, N145I, I159F, A166V, V178G, L192Q, [A198A]
B2	A13T, R15C, G16R, T24A, G41D, [T142T], Q213L
B3	T11I, [N59N], F77L, G78S, A154S, Q203H, L209P, E214V

Selection of the *rutB* gene library on iron limited agar plates was in principle successful. However, due to the almost ubiquitous presence of deletions and STOP codons in the gene variants isolated, other effects than the desired *in vivo* complementation may have been responsible for restoring growth.
Selection of the *fabG* library was also generally successful, as each separate culture resulted in the prevalence of a single gene variant after application of iron limiting conditions. Less abundant variants were observable in some cultures and may either have provided a minor advantage compared to the respective main variant or be a false positive. Variants of *rutB* and *fabG* identified by the selection methods that could be expected to produce a full-length protein, were further analyzed by a quantitative growth assay (see subsubsection 3.2.4.3).

3.2.4.3 Quantitative Growth Assay

Iron limiting liquid media previously established to prevent growth of enterobactin non-producing strains (see subsubsection 3.2.4.1) was employed to quantify the effect of certain variants of *rutB* and *fabG* derived either from FuncLib (see subsection 3.2.1) or gene library selection (see subsubsection 3.2.4.2) on the growth of applicable deletion strains. To this end, transformants were grown in 1 ml of M9-BPD medium in a plate reader (37°C; 140 rpm) with growth being monitored by continuous measurement of OD_{600}. Precultures in M9 medium supplemented with 20 μM 2,3-DHB were used to inoculate each well to an initial OD_{600} of 0.01.
Enterobactin producing strains (e.g. Δ*entA* + pExp:*entA*) typically grew to stationary phase within 16 h under these conditions, reaching OD_{600} values of approximately 0.4. Enterobactin non-producing strains (e.g. Δ*entA*) displayed some initial growth, probably due to transfer of intracellular iron or 2,3-DHB from the precultures. However, these strains would typically enter a stationary phase after around 10 h, reaching OD_{600} values of no more than 0.1 and display no further growth until at least 60 h after the start of the experiment.
Gene variants of *rutB* and *fabG* to be tested were subcloned into pExp and transformed into the *E. coli* strains Δ*entB-IC* and Δ*entA*, respectively. Some of the *fabG* variants but none of the *rutB* variants induced growth after around 40 h of incubation (see Table 3.4). The proteins coded for by these variants of *fabG* were purified (see subsection 2.3.5) and tested for possible EntA activity (see subsection 3.2.7.

Table 3.4: *fabG* variants inducing *in vivo* growth advantages of BW25113 *entA::Cam*
Cells of various *E. coli* strains were grown in M9-BPD medium, that is non-permissive for enterobactin non-producing strains. Experiments were started with an OD_{600} of 0.01 and growth proceeded at 37°C and 140 rpm in a plate reader. Under these conditions, wild-type BW25113 reached an OD_{600} of 0.4 after 16 h, while enterobactin non-producing BW25113 Δ*entA* did not exceed an OD_{600} of 0.05 after more than 60 h. Some *fabG* variants (cloned into the pExp plasmid) improved growth of BW25113 Δ*entA*. The first group of variants is derived from FuncLib predictions (see subsection 3.2.1), while the second group is derived from gene library selection (see subsubsection 3.2.4.2.)

Strain	Variant	Growth
BW25113	-	16 h; 0.4
Δ*entA*	-	>60 h; <0.05
Δ*entA*	FabG S138A V140L F183W T186S M188L	14 h; 0.10
Δ*entA*	FabG T90S S138A V140L	55 h; 0.35

Strain	Variant	Growth
Δ*entA*	FabG T90V D92E S138A F183Y M188L	55 h; 0.35
Δ*entA*	FabG T90V D92E S138A V140L F183Y	50 h; 0.35
Δ*entA*	FabG T90V S138A M188L	51 h; 0.35
Δ*entA*	A1 (see Table 3.3)	54 h; 0.23
Δ*entA*	A2	55 h; 0.30
Δ*entA*	A5	38 h; 0.20
Δ*entA*	A6	50 h; 0.22
Δ*entA*	B2	44 h; 0.25

3.2.5 Purification of Chorismate

Chorismate was required for the enzymatic assays employed for the detection of isochorismatase (EntB) activity and 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EntA) activity (see subsection 2.3.6). The substrates of EntB and EntA were not commercially available and thus had to be generated *in situ* through the actions of EntC, or both EntC and EntB, respectively. The protocol for the purification of chorismate (see subsubsection 2.3.2.1) was previously established in the author's master's thesis [20] due to supply problems, which have since been alleviated. However, most preparations of chorismate commercially available feature low purity and have been reported to give unpredictable results (Sandra Schlee, personal communications), prompting the continued employment of the purification protocol.
Chorismate was purified from culture supernatant of *E. coli KA12*, which had previously been developed as a chorismate producing strain [22] [23]. Extraction of the supernatant with ethyl acetate and C-18 reverse phase chromatography was conducted as described in subsubsection 2.3.2.1. HPLC analysis (subsubsection 2.2.5.5) allowed for the selection of sufficiently pure fractions and indicated a purity of the final product of 98.6% (based on peak areas at 280 nm; see Figure 3.4), far exceeding that of commercially available chorismate preparations. HPLC analysis of the products of enzymatic conversion of chorismate by anthranilate synthase (see subsubsection 2.3.2.2) indicated an even higher purity of 99.7% (based on the amount of anthranilate produced). For unknown reasons, the achieved yield of 23 mg pure chorismate per 1 l of culture supernatant was significantly lower than yields reported by Grisostomi *et al.* [22] but still sufficient for the further experiments.

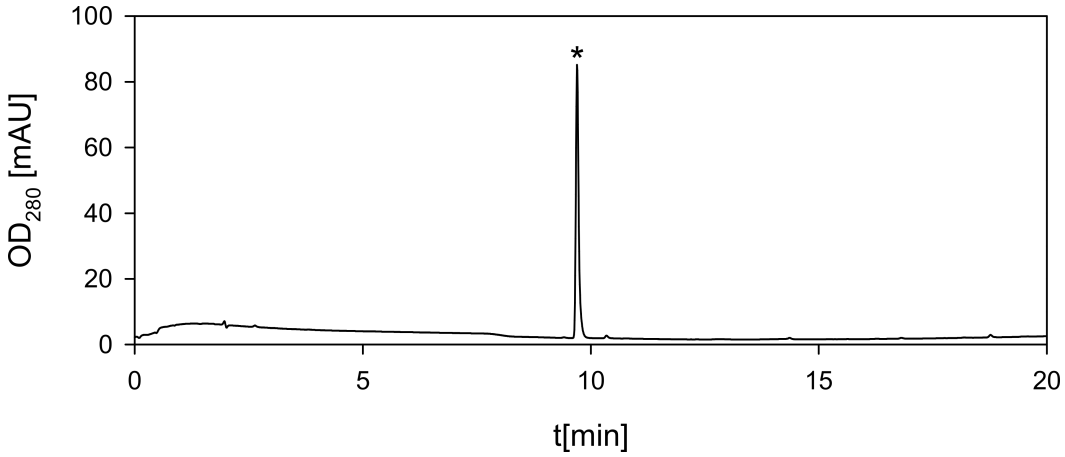


Figure 3.4: HPLC analysis of chorismate preparation
HPLC chromatogram of the chorismate preparation used in this work. Separation proceeded as described in subsubsection 2.2.5.5 and absorption at 280 nm was measured over the elution time. The peak corresponding to chorismate is indicated by an asterisk. Integration of all peak areas revealed a fraction of 98.6% of the chorismate peak, indicating high purity.

3.2.6 Protein Purification

Required expression plasmids based on pET21a and pUR22 (see Table 2.2) were generated as described (see subsubsection 2.2.3.3). Enzyme variants of RutB and FabG were produced in BW25113 Δ*entB* and BW25113 Δ*entA* (see Table 2.3), respectively, to avoid contamination with the respective wild-type Ent enzyme. Proteins were purified to homogeneity by immobilized metal ion affinity chromatography (see subsubsection 2.2.4.2) and preparative size exclusion chromatography (see subsubsection 2.2.4.3). EntB, EntC and TrpE were previously purified by Maximilian Plach in our group (see subsubsection 2.1.3.8). Preparations of wild-type EntA, RutB, and FabG were previously generated in the author's master's thesis [20]. Yields for variants of RutB and FabG varied drastically depending on the residue exchanges present. The sets of mutations suggested by JANUS [25] (see subsection 3.2.1) were generally well tolerated, leading to yields of 12 mg and 21 mg of protein per liter of culture medium for variants RutB2EntB15 and FabG2EntA16, respectively, which is close to the yields of the respective wild-type enzyme (12 mg/l and 17 mg/l for RutB and FabG, respectively; [20]). Mutations suggested by FuncLib [24] were similarly tolerated with yields ranging from 19 mg/l to 41 mg/l for the variants produced (see subsubsection 3.2.4.3), while variants isolated from the gene libraries expressed significantly worse with protein yields as low as 1.1 mg/l (FabG A4; see Table 3.2).

The purification of FabG2EntA16, exemplifies the overall success and applicability of the purification protocol (see Figure 3.5). The respective purification steps of the remaining enzyme variants are shown in subsection 8.2.2.

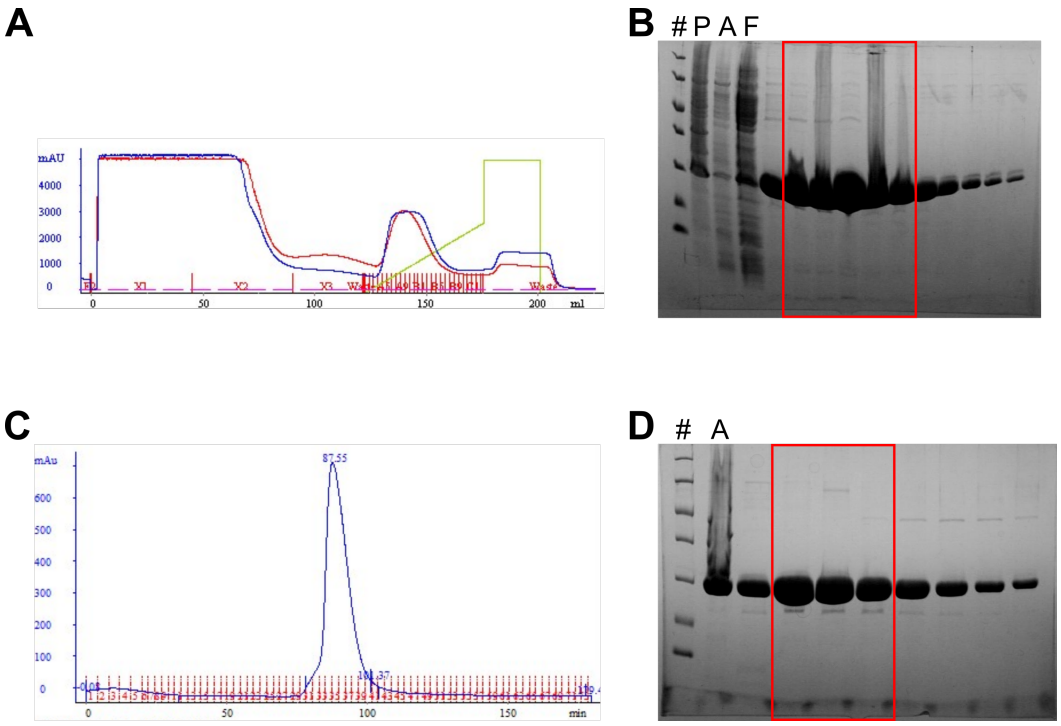


Figure 3.5: Purification of FabG2EntA16
FabG2EntA16, containing 16 residue exchanges suggested by JANUS [25] (see subsection 3.2.1), was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. FabG2EntA16 appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.5 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) S75 elution profile of FabG2EntA16. (D) SDS-PAGE (13.5% acrylamide) of S75 elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

3.2.7 Qualitative *in vitro* Enzyme Assays

Long incubation times and sensitive HPLC analysis have the potential to detect even minimal side activities of a given enzyme. The HPLC based assay initially established in the author's master's thesis [20] was employed in this work to

test specific RutB and FabG variants for possible EntB and EntA activities, respectively (see subsection 2.3.6).
The two *rutB* variants isolated from the *rutB* gene library (see Table 3.2) that could be expected to produce full-length or close to full-length RutB variants (500-1, 500-3) shared the residue exchange Q26H, prompting the production and investigation of RutB Q26H, which did not produce a measurable amount of the expected isochorismatase product 2,3-DHDHB. Due to the setup of the assay, in which isochorismate exists in an equilibrium with chorismate due to the isochorismate synthase activity of EntC, a chorismate hydrolase activity can equally be ruled out, since no product peak was observable at all.
A *rutB* gene library, in which three positions close to the active site were randomized (see subsubsection 3.2.3.2), was expressed, resulting in a mixture of different RutB enzyme variants separating into two fractions during size exclusion chromatography. One of the fractions consisted of variants forming tetramers, the other of variants forming dimers. Both fractions were tested for their ability to hydrolyze isochorismate into 2,3-DHDHB but no product could be identified after incubation. Again, a chorismate hydrolase activity can be ruled out as well.
A number of FabG enzyme variants derived from gene library selection (see Table 3.3) and FuncLib (see Table 3.4) or JANUS predictions (see subsection 3.2.1) were produced and tested with either NAD^+ or $NADP^+$ for their ability to convert 2,3-dihydro-2,3-dihydroxybenzoate (2,3-DHDHB) into 2,3-dihydroxybenzoate (2,3-DHB) and with either NADH or NADPH for conversion of DHB into DHDHB (see Figure 2.1), representing the regular and a hypothetical inverse activity of EntA. No detectable amount of the respectively expected product was found in the reaction mixtures after incubation (RT; 24 h).
In summary, none of the enzyme variants produced displayed any measurable target activity. Due to the high sensitivity of the HPLC-based method, any undetected activity would have to be exceedingly small.

3.3 Discussion

It is generally accepted that over the course of evolution, secondary metabolic functions arose from primary metabolic progenitor enzymes [75]. In this work, it was attempted to retrack this evolutionary process for the secondary metabolic enterobactin pathway proteins isochorismatase (EntB) and 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EntA) from *Escherichia coli*. Their closest respective primary metabolic homologues, ureidoacrylate amidohydrolase (RutB) and 3-oxoacyl-[acyl-carrier-protein] reductase (FabG), were chosen as model progenitors. Establishing and understanding the functional conversion between these two pairs of enzymes could have demonstrated, how secondary metabolic enzymes have emerged during evolution and how new enzyme functions can arise in general.

3.3.1 A Viable Selection System for Enterobactin Producing Strains

The functional conversion of FabG into a construct with EntA activity has previously been attempted in the authors master's thesis with the focus on rational mutations [20]. As this strategy was insufficient to achieve functional conversion, the focus in this work was shifted towards a gene library approach with the additional inclusion of the conversion of RutB into EntB. Random mutagenesis of the entire open reading frame of a gene produces a gene library with a large number of variants which holds the potential to find mutations that would not be considered in a rational approach. Identifying gene variants coding for enzymes with the desired properties requires the establishment of stringent selective growth conditions.

The central result of this project was the establishment of robust conditions for the selective growth of enterobactin producing *E. coli* strains. Owing to the essential nature of iron as a cofactor for many enzymes [83] [128] [129] and its generally low bioavailability [84], many organisms have developed specific uptake mechanisms for this essential nutrient [85] [86]. *Escherichia coli* employs enterobactin, its breakdown products, and citrate as endogenously synthesized siderophores [85] [130]. Other mechanisms of iron uptake exist in certain pathogenic *E. coli* strains [131] [85] [130] [132] but these do not play a role here.

The central idea of the selection system for enterobactin producers, was to limit the bioavailability of iron in the medium to such a degree that enterobactin production becomes essential for cell growth. This has previously been achieved by the addition of the iron chelating agent 2,2'-bipyridine (BPD) [46] [133], which tightly binds iron ions in the medium and cannot be imported into the cell. Due to its extremely high affinity for iron ($K_a = 10^{51} M^{-1}$; [105]), enterobactin can abstract iron from the ferric BPD complex and restore iron uptake. Chromosomal deletions of *ent* genes can be expected to interrupt the enzymatic pathway and prevent production of enterobactin *in vivo*.

Iron limitation was achieved by addition of 75 μM BPD or 1 μM ethylenediaminetetraacetic acid (EDTA) to M9 minimal media (see subsubsection 3.2.4.1). Under these media conditions, only enterobactin producing strains (BW25113; Δ*entA+pExp::entA*; etc.) could grow unimpeded, while enterobactin non-producing strains stagnated early into the exponential phase. Consistent lack of growth of *ent* deletion strains proved, that no other metabolic route for the production of enterobactin or any other sufficient mechanism for iron acquisition exists in *E. coli*.

Growth of enterobactin non-producing strains could generally be restored under these conditions by addition of 2,3-dihydroxybenzoate (2,3-DHB) into the growth medium, as previously noted [46]. 2,3-DHB constitutes the product of the EntA reaction and, accordingly, the downstream reactions of the enterobactin pathway can proceed normally resulting in production of enterobactin.

These findings prove the general integrity of the non-deleted components of the enterobactin pathway. Notably, this was not the case for the *entB* deletion strain, as this protein carries an additional function in the form of its aryl carrier protein (ACP) domain, that is required for the assembly of enterobactin by non-ribosomal peptide synthesis (see Figure 3.1). Thus, the *entB* deletion strain is not suitable for *in vivo* complementation experiments with variants of *rutB*, as RutB lacks an ACP domain. As the ACP domain of EntB was previously shown to be functional as a standalone construct *in vitro* [102], the partial *entB* deletion *entB(1-561)::Cam*, in which only the isochorismatase domain of EntB (residues 1-187) is deleted, was devised and introduced into BW25113. The growth of the resulting strain (Δ*entB-IC*) under iron limiting conditions could be restored either by a plasmid carrying a gene for the isochorismatase domain of EntB (*pExp::entB-IC*) or addition of 2,3-DHB, making it a suitable selection strain for gene libraries of *rutB* as well as demonstrating the standalone functionality of the ACP domain of EntB *in vivo*.
Deletion of the *fec* operon (*fecIRABCDE*) prevents iron uptake via iron citrate, conceivably placing further pressure on the restoration of enterobactin competence. *fec* deletions were combined with the deletion of either *entB(1-561)* or *entA*, with the resulting strains behaving effectively identically to the respective single deletion strains. This observation implies, that iron uptake in the form of iron citrate plays no role under the growth conditions tested, probably because citrate is not able to abstract iron from the metal BPD complex.
Growth of BW25113 Δ*entB-IC* and Δ*entA* under iron limiting conditions could be restored by transformation with a plasmid borne copy of the respectively deleted gene but not with the null plasmid (*pExp*) or the respective primary metabolic homologue (*pExp::rutB*; *pExp::fabG*). These successful *in vivo* complementations of the respective chromosomal gene deletions show that BW25113 Δ*entB-IC* and Δ*entA* are suitable selection strains for the gene libraries of *rutB* and *fabG*, respectively. Variants of *rutB* and *fabG* coding for gene products with EntB and EntA activity, respectively, can be expected to restore growth of the deletion strains under iron limiting conditions. Notably, even low enzymatic activities of the RutB and FabG variants should be sufficient to restore growth, due to the potency of enterobactin and the relatively low amount of iron ions required per cell division (10^5-10^6; [83]).

3.3.2 Isolation and Characterization of RutB and FabG Variants

The central strategies for functional conversion of RutB and FabG into variants with EntB and EntA activity, respectively, were fully randomized gene libraries, targeted gene libraries, and bioinformatic analysis tools. Random mutagenesis of the full open reading frames (ORFs) of the *rutB* and *fabG* genes has the potential to identify variants of RutB or FabG with a minimal EntB or EntA activity

(see subsubsection 3.2.3.1), which could then serve as a starting point for the further increase of the respective enzyme activity by directed evolution. Similar promising starting states for further evolution could be identified by the targeted randomization of substrate binding residues (see subsubsection 3.2.3.2) and utilization of computational tools for the prediction of substrate specificity altering residue exchanges (see subsection 3.2.1). Based on further directed evolution, an evolutionary trajectory could be constructed that could explain the emergence of EntB or EntA activity from primary metabolic progenitors.
Gene libraries were produced by error prone PCR of wild-type *rutB* and *fabG*, achieving an average of 10.3 and 12 nucleotide mutations per ORF, corresponding to 6.5 and 8.7 amino acid exchanges, respectively (see subsubsection 3.2.3.1). The *rutB* gene library produced in this work contained a total of 3.2e8 unique variants, with a theoretical 1.2e8 of them coding for full-length gene products (i.e. not containing STOP codons or deletions). The gene library of *fabG* consisted of 2e7 unique variants, of which 7.5e6 could be expected to produce a full-length gene product. The size of the *rutB* gene library approaches the upper limit of what has been reported for the method employed [26] while the significantly smaller *fabG* gene library may still be sufficient, as gene libraries of similar size have previously been successfully employed for *in vivo* complementation experiments [125].
The deletion strains and iron limiting growth conditions previously discussed (see subsection 3.3.1) were used for *in vivo* complementation experiments to select for variants of *rutB* and *fabG* able to restore growth and further to quantify the growth advantage of variants thus identified. Gene libraries of *rutB* and *fabG* were transformed into *entB* and *entA* deletion strains, respectively, and the resulting transformants were subjected to growth under iron limiting conditions. Transformants of Δ*entB* were grown on M9 agar plates containing either 50 μM or 500 μM EDTA and transformants of Δ*entA* were grown in separate batches of M9 medium containing 75 μM BPD. In both selection methods, single gene variants were isolated due to growth as single colonies on agar plates or due to them outperforming other variants in mixed liquid culture (see subsubsection 3.2.4.2). Some of the *fabG* but none of the *rutB* variants reproducibly provided a growth advantage under iron limiting conditions, that is discussed in detail below (see subsection 3.3.3). Considering that the *rutB* variants isolated from the fully randomized gene library failed to restore growth of the Δ*entB-IC* deletion strain under iron limited conditions and the prevalence of variants carrying STOP codons and deletions, implies them to be false-positive hits. It is conceivable that a rare spontaneous mutation restored enterobactin production or iron uptake by an unknown mechanism in these cases. Alternatively, colony growth in the gene library selection might have arisen through localized depletion of EDTA.
Iron limited growth conditions were further employed to screen gene variants predicted by FuncLib (see Table 8.3) for their ability to complement deletions of

entB and *entA*. Five of the FuncLib variants of *fabG* displayed a slight growth advantage (see Table 3.4) and were further analyzed for a possible *in vitro* activity. An HPLC-based discontinuous assay was employed to screen enzyme variants for EntB or EntA activity (see subsection 3.2.7). Variants tested in this way were the FabG variants derived from gene library selection (see Table 3.3) and a selection of FuncLib derived variants that were previously shown to enhance growth of an *entA* deletion strain under iron limited growth conditions (see Table 3.4). Further enzymes assayed were variants of RutB and FabG generated by JANUS (see subsection 3.2.1), as well as RutB Q26H, as this residue exchange was present in multiple of the RutB variants isolated from gene bank selection (see Table 3.2). Furthermore, mixed expression of a small *rutB* gene library (see subsubsection 3.2.3.2) produced the protein mixture RutB Y29X Y35X F41X, which was similarly tested. None of the variants of RutB tested displayed any detectable activity of EntB and due to the setup of the assay, a chorismate hydrolase activity can equally be ruled out. None of the FabG variants assayed displayed an activity for either the substrate or the product of EntA, in the presence of either $NAD(P)^+$ or NAD(P)H cofactors, respectively.
None of the enzyme variants tested displayed any measurable new activity. This lack of a suitable starting point for the functional conversion of RutB and FabG into EntB and EntA, respectively, prohibited any further directed evolution experiments and eventually lead to the discontinuation of the project. Notably however, some variants of *fabG* reproducibly lead to increased growth rates of the enterobactin non-producing *entA* deletion strain under iron limiting conditions, which is discussed in detail in subsection 3.3.3.

3.3.3 *In vivo* Growth Advantages of *fabG* Variants

While none of the *fabG* variants isolated from the fully randomized gene library of *fabG* or predicted by FuncLib produced an enzyme with measurable *in vitro* EntA activity (see subsection 3.2.7), some did convey a growth advantage under iron limiting conditions (see subsubsection 3.2.4.3). The BW25113 *entA*::*Cam* strain transformed with any of these *fabG* variants in a pExp plasmid grows faster (stationary phase after 14-55 h) than the same strain transformed with an empty pExp plasmid (no growth after 60 h), albeit slower than an enterobactin producing strain (BW25113 *entA*::*Cam* + pExp:*entA*; stationary phase after 12 h). The same general trend can be observed for the strain BW25113 *entA*::*Cam fec*::*Kan* in which the uptake machinery for iron-citrate is missing (not shown).
The mode of action by which the *fabG* variants convey the observed *in vivo* advantage is currently not clear but seems to be unrelated to the uptake of iron citrate, which is employed as an alternative iron requisition mechanism in *E. coli*, as the growth rate is comparable in *entA*::*Cam* and *entA*::*Cam fec*::*Kan* strains. Other iron uptake strategies of *E. coli* also do not play a role under the growth

conditions applied as they would require an exogenous siderophore source [85] [130]. While none of the proteins coded for by these *fabG* variants possessed a measurable EntA activity *in vitro*, it is conceivable that the respective enzyme variants are functional *in vivo* due to the action of an unknown substrate, co-factor, or interaction partner.
It has previously been observed, that attempts at directed evolution can give rise to unanticipated metabolic adaptations. Novel protein constructs can act as transcriptional regulators for enzymes or pathways with applicable moonlighting activities [134] [135]. Furthermore, the significant potential of metabolic repair in *E. coli* has previously been noted. In one case, interruption of β-alanine biosynthesis resulted in three independent pathways, that restored β-alanine production by different metabolic routes [136]. Taking these findings into consideration, the *fabG* variants isolated from the fully randomized gene library may code for proteins that either act as transcription factors or possess novel catalytic activities (other than EntA activity), that restore the production of 2,3-DHB *via* an alternative metabolic route, based on either a promiscuous activity of an up-regulated enzyme or the respective FabG variant. This hypothesis is somewhat supported by the locations of the residue exchanges in the FabG variants isolated (see Figure 3.6). The core structure of the Rossmann fold features relatively few mutations, implying that a selection for well-folded protein variants has occurred. In contrast, a short α-helix close to the active site (residues 186-192) is highly mutated, which might give rise to altered substrate binding affinities of the protein variants.

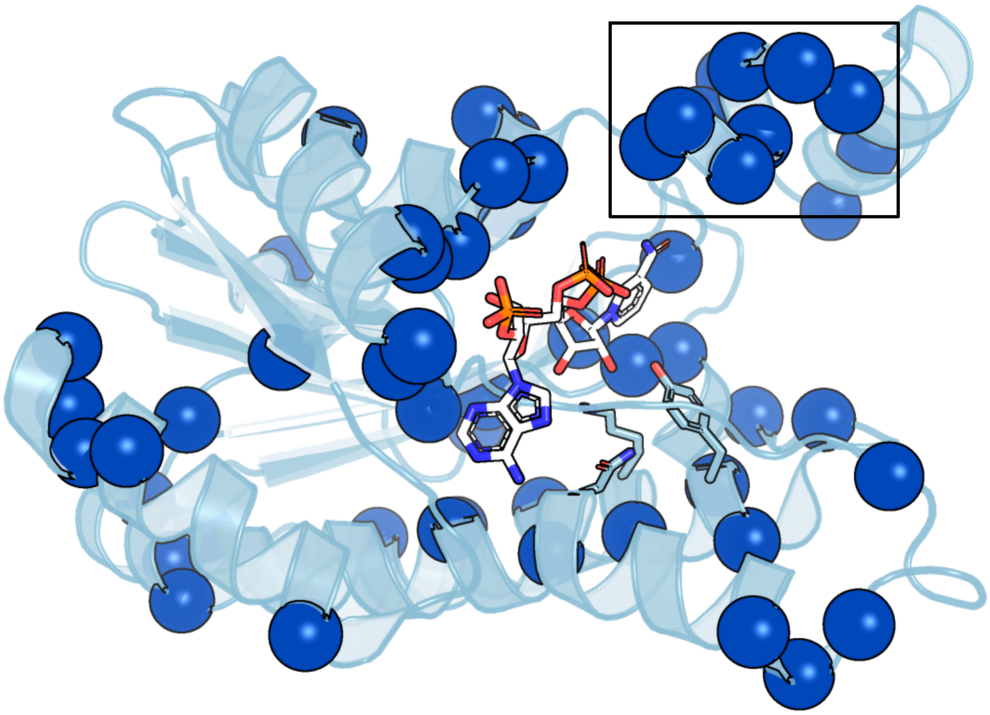


Figure 3.6: Locations of mutations of FabG variants isolated from the fully randomized gene library
Shown in this figure is the crystal structure of FabG (PDB: 1Q7B) with the $NADP^+$ cofactor (white) and the catalytic triad (S138, Y151, K155; blue) shown in stick representation. Dark blue spheres indicate where residue exchanges occurred in the FabG variants isolated from the random gene library of *fabG* (see Table 3.3). Variant A3 is ignored here as this variant would not produce a full-length gene product. Residue exchanges are relatively rare in the core structure of the Rossmann fold but frequent around the cofactor binding site, particularly in a small α-helical element near the active site cavity (residues 186-192; black box), indicating that selection may have favored well-folded protein variants with distinct new substrate affinities.

Due to the involvement of wild-type FabG in fatty acid synthesis [113] an effect of these FabG variants on membrane composition is conceivable. These changes might hypothetically allow for the transmembranous uptake of the iron BPD complex and later intracellular iron release.
In summary, the causal relationship between the *fabG* variants analyzed and the observed *in vivo* growth advantage under iron limiting conditions is unclear. Elucidation of the underlying effects would require in-depth metabolomic and transcriptomic efforts, that lie outside the scope of this work. Possible directions for future research will be discussed in section 3.4.

3.4 Outlook

A continuation of the attempts at functional conversion of RutB and FabG into EntB and EntA, respectively, can not be recommended, as the extant enzymes are likely too far diverged to allow for a direct conversion. A further characterization of the protein variants isolated in this work as well as their *in vivo* effects might however yield some interesting insights.
A qualitative method for the detection of siderophore producing strains has previously been established in the author's master's thesis [20]. Agar plates containing the dye chromeazurol S (CAS) exhibit a localized color change in response to siderophore production by a bacterial colony [137]. Establishing a more sensitive CAS based assay [138] might help detect and quantify a possible minor siderophore production in *entA* deletion strains transformed with the *fabG* variants determined to improve growth under iron limiting conditions (see Table 3.4). Proof of siderophore production would hint towards an *in vivo* activity of the FabG variants, that could then be characterized by in-depth metabolomics and transcriptomics studies.
In the case of RutB, focusing on other evolutionary targets might be a promising avenue. As discussed in detail below (see subsection 4.3.2), RutB is structurally similar to both a triuret hydrolase (TrtA; PDB: 6XJE; [139]) and a biuret hydrolase (BiuH; PDB: 6AZN; [140]), which are involved in the degradation of the anthropogenic substances triuret and cyanuric acid, respectively. Investigating the evolutionary relationship between RutB and either TrtA or BiuH, might reveal the evolutionary potential of isochorismatase-like hydrolases and how enzymes of this superfamily can be integrated into novel catabolic pathways. The overall mutational strategies outlined in this work could be similarly applied to this new project but the establishment of selective growth conditions for the utilization of either triuret or cyanuric acid would be required first. Alternatively, rational mutations could be devised for RutB, TrtA, and BiuH based on comparison of their respective MSAs in a similar approach as described for the interconversion of hydroxyatrazine ethylaminohydrolase (AtzB) into a guanine deaminase (see subsection 5.2.2).
Considering RutB as an evolutionarily ancient enzyme due to its formally primary metabolic function may have been an error of judgement. The Rut pathway in its entirety shares a number of features with evolutionarily young catabolic operons for the utilization of specific nutrients. It features a narrow phylogenetic spread and is only essential for survival under specific conditions [52]. Accordingly it likely emerged in more recent evolutionary times, raising in turn the question of the evolutionary history of the enzymes of the Rut pathway.

4 Structural and Functional Characterization of the Ureidoacrylate Amidohydrolase RutB From *E. coli*

4.1 Introduction

In mammals [141], plants [142], and some bacteria [143] [144], pyrimidines are degraded by a well characterized reductive pathway [145] *via* an initial reduction to 5,6-dihydropyrimidines by dihydropyrimidine dehydrogenase, ultimately providing β-alanine and ammonia as its products. In some species, β-alanine can further be utilized as an energy [146] or nitrogen source [147]. A less prevalent oxidative pyrimidine degradation pathway has so far only been identified in certain bacterial species [148]. Here, degradation proceeds *via* an initial oxidation of pyrimidines to barbituric acid derivatives [149] [150], eventually giving malonate and urea as final products.
Escherichia coli appears to possess neither the reductive nor the oxidative pathway [151], although an isolated dihydropyrimidine dehydrogenase activity has been identified [152] [153]. Instead, *E. coli* relies on the Rut pathway (derived from py**r**imidine **ut**ilization) for the degradation of uracil and thymine [154] [52]. The Rut pathway has only recently been discovered and enables growth of *E. coli* on uracil as sole nitrogen source at temperatures up to 22°C [52]. Degradation of pyrimidines proceeds *via* a flavin mononucleotide-dependent pyrimidine monooxygenase (RutA; [48]) which acts in conjunction with a flavin reductase (RutF), ureidoacrylate/ureidoacrylate peracid amidohydrolase (RutB), aminoacrylate peracid reductase (RutC), aminoacrylate hydrolase (RutD), and 3-hydroxy propionic acid dehydrogenase (RutE) to liberate both ring nitrogens in the form of ammonia [52] (see Figure 4.1). RutC and RutD catalyze reactions that can proceed spontaneously *in vitro* but are required *in vivo* for growth on uracil as the sole nitrogen source [52]. Similarly, the RutE reaction is not required for ammonia release but serves to remove the toxic intermediate malonic semialdehyde *in vivo* [52]. RutG is a pyrimidine base transporter, not directly involved in the pathway [52].

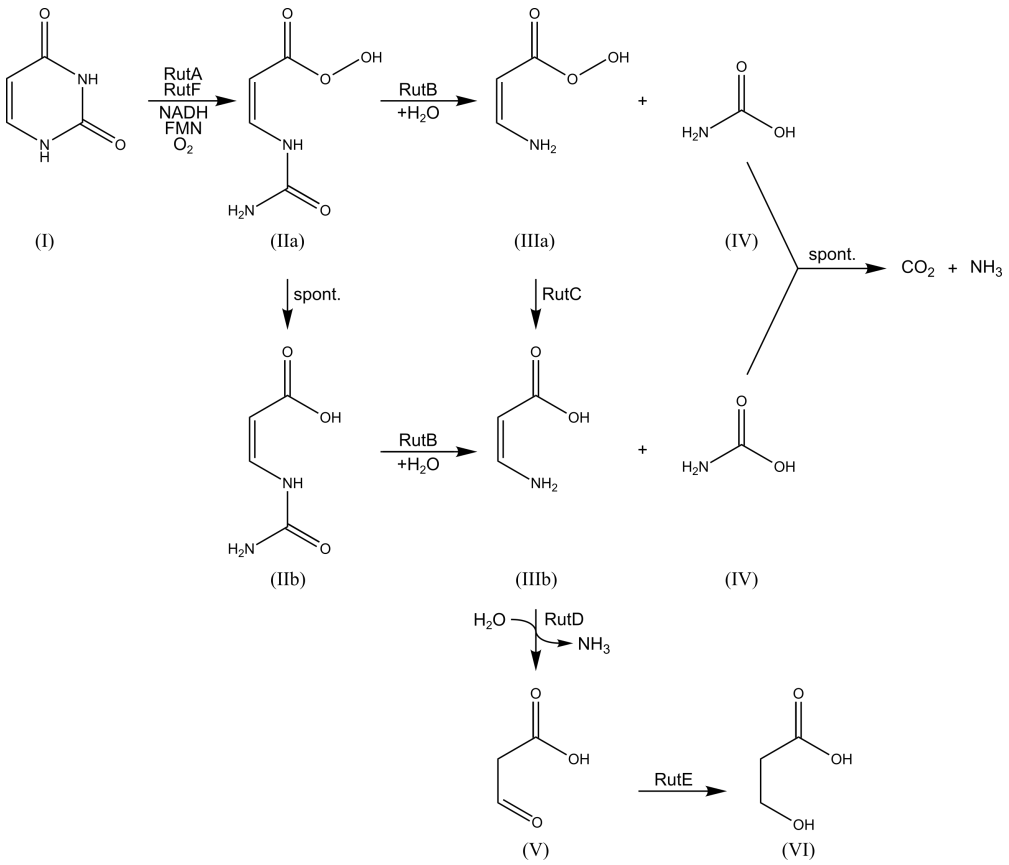


Figure 4.1: The Rut pathway
The enzymes of the Rut pathway catalyze the degradation of uracil (I) and thymine (not shown) for utilization as a nitrogen source [52]. RutA is a flavin mononucleotide-dependent pyrimidine monooxygenase that acts in conjunction with the flavin reductase RutF to catalyze a ring-opening reaction, giving ureidoacrylate peracid (IIa) as its product. This peracid can undergo spontaneous hydrolysis, giving ureidoacrylate (IIb). The ureidoacrylate/ureidoacrylate peracid amidohydrolase RutB catalyzes the hydrolysis of both (IIa) and (IIb), producing aminoacrylate peracid (IIIa) and aminoacrylate (IIIb), respectively, as well as carbamate (IV). (IIIa) is converted into (IIIb) by aminoacrylate peracid reductase (RutC) while (IV) undergoes spontaneous hydrolysis to NH_3 and CO_2. Another molecule of ammonia is released from (IIIb) by the action of aminoacrylate hydrolase (RutD), further yielding malonic semialdehyde (V). This substance is further reduced by 3-hydroxy propionic acid dehydrogenase (RutE) to give 3-hydroxy propionic acid (VI) to avoid toxicity effects of (V) [52].

Most of the Rut enzymes have been characterized, with crystal structures being available for RutA [48] [155], RutC [156], RutD [157], and the transcriptional regulator RutR [158]. However, the structure and reaction mechanism of RutB have not been described up to now, although two close homologues, biuret hy-

drolase (BiuH; [140]) and triuret hydrolase (TrtA; [139]), have recently been characterized in detail.
An in-depth characterization of RutB would close this knowledge gap regarding the enzymes of the Rut pathway and further open up an avenue for investigations into the evolutionary origin of this enzyme. As discussed in section 3.4, the Rut pathway may be relatively young, possibly allowing for the identification and enhancement of the respective progenitor activities of the Rut enzymes, similarly to the strategy applied to the recently evolved hydroxyatrazine ethylaminohydrolase AtzB (see chapter 5).
Furthermore, RutB is an interesting target for characterization as it has previously been utilized for the enantiopure synthesis of Vince lactam (2-azabicyclo-[2.2.1]hept-5-en-3-one), exploiting the enzyme's promiscuous (+)-γ-lactamase side activity [159]. This process is of pharmaceutical importance, as Vince lactam is an important precursor for a number of antiviral drugs [160] [161]. At least two other γ-lactamases from the isochorismatase-like hydrolase (IHL) superfamily have been identified [162] [163], one of which has since been engineered for improved enantioselectivity and thermostability [164].

4.2 Results

4.2.1 Ureidoacrylate Synthesis

Ureidoacrylate (UAc), required for co-crystallization experiments and enzyme assays, had to be synthesized in-house as it was not available commercially. Synthesis of UAc was achieved by a two-step reaction from maleic anhydride and trimethylsilyl azide (see subsection 2.4.1). Product yield was approximately 30% in each of the reaction steps, even though higher yields have previously been reported [50]. Purification by anion exchange chromatography yielded sufficient quantities (approximately 30 mg) of high purity UAc (see Figure 4.2). No significant degradation was observed after prolonged storage at -80°C (not shown).

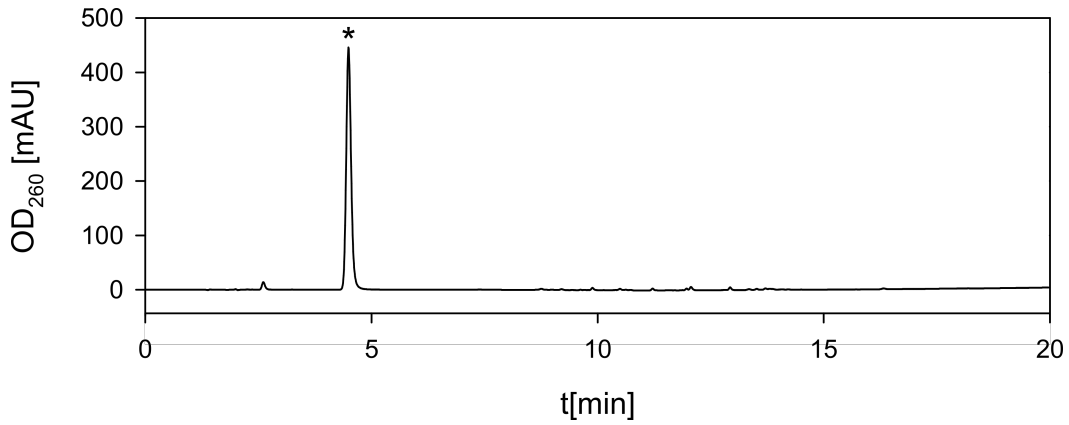


Figure 4.2: HPLC analysis of ureidoacrylate preparation
HPLC chromatogram of the ureidoacrylate preparation used in this work. Separation proceeded as described in subsubsection 2.2.5.5. Absorption at 260 nm was measured over the elution time. The peak corresponding to ureidoacrylate is indicated by an asterisk. Integration of all peak areas revealed a ratio of >95% of the ureidoacrylate peak, indicating high purity.

4.2.2 Protein Purification

Required expression plasmids based on pET21a and pUR22 (see Table 2.2) were generated as described (see subsubsection 2.2.3.3). Wild-type RutB and variants containing single amino acid exchanges were produced in the *E. coli* expression strain BL21 (DE3) Gold (see Table 2.3) and purified to homogeneity by immobilized metal ion affinity chromatography (see subsubsection 2.2.4.2) and preparative size exclusion chromatography (see subsubsection 2.2.4.3). Yield for wild-type RutB was approximately 60 mg per liter of culture medium and the respective values of most variants were between 30 and 50 mg/l. The purification of wild-type RutB (see Figure 4.3) exemplifies the overall success and applicability of the purification protocol. The respective purification steps of the RutB variants are shown in subsection 8.3.1.

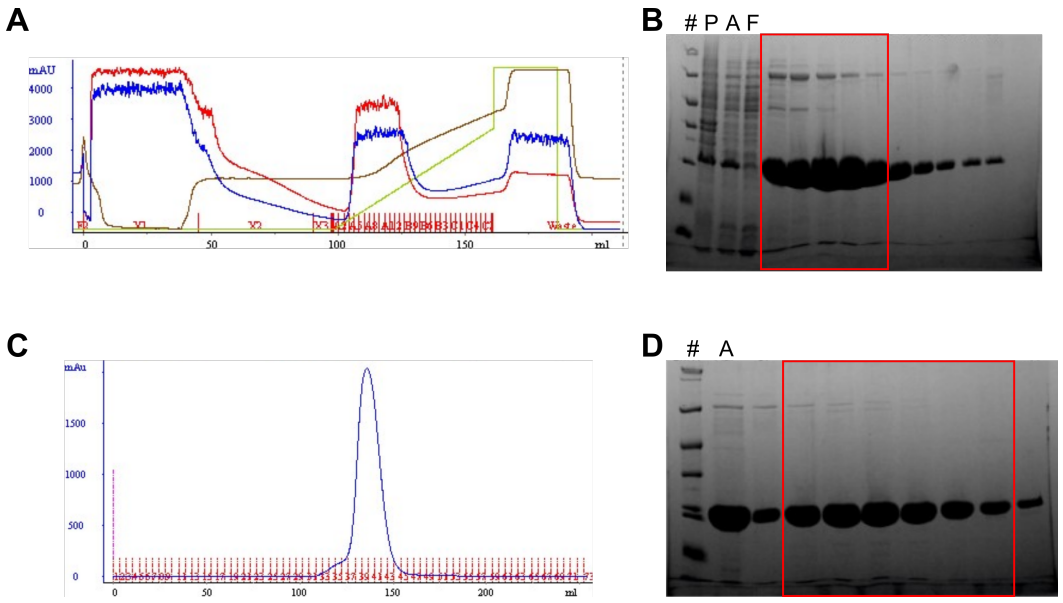


Figure 4.3: Purification of wild-type RutB
Purification of wild-type RutB proceeded by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) S75 elution profile of RutB. (D) SDS-PAGE (13.5% acrylamide) of S75 elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

4.2.3 Circular Dichroism

Circular dichroism (CD) spectra of all RutB variants were measured as described (see subsubsection 2.2.5.3) to determine effects of the residue exchanges introduced on the secondary structural content of the protein. Generally the spectra are similar, pointing to the overall fold being conserved (see Figure 4.4). Some exchanges of active site residues (D24A, D24N, K133A, C166K, C166S) as well as N72A, Y136A, and Y136F lead to a slight shift towards increased α-helical or β-strand content and RutB Y35F displayed slightly reduced overall CD signal, indicative of a partial loss of secondary structure. All CD spectra are shown separately in subsection 8.3.2

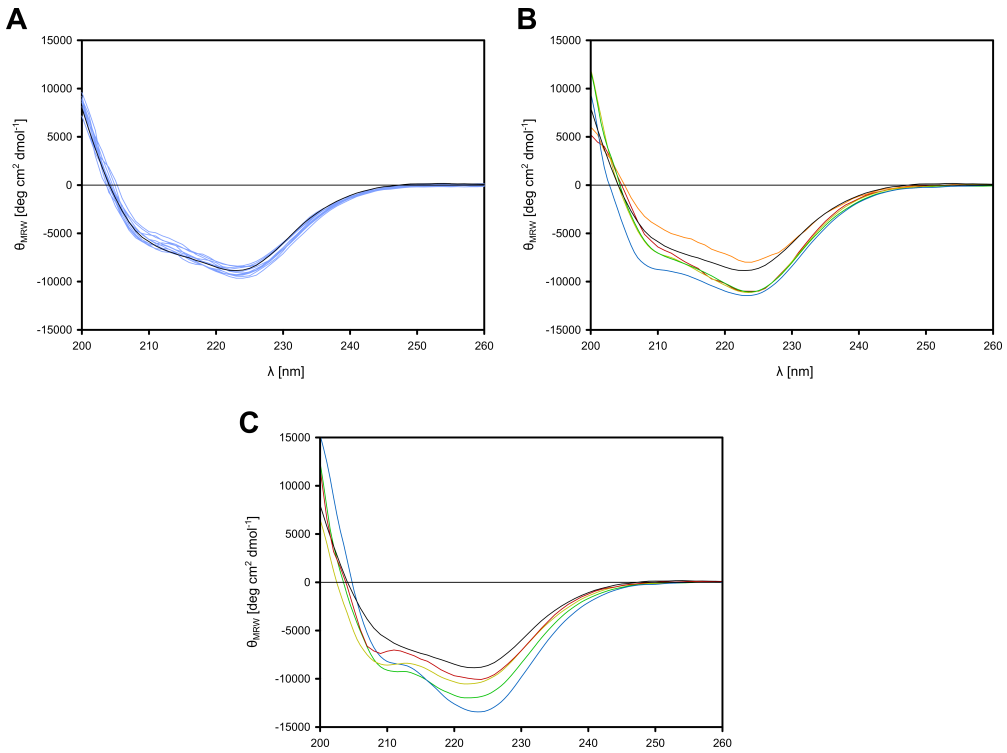


Figure 4.4: Superposition of CD spectra of RutB variants
(A) Most RutB variants (Y29F, W74A, W74F, E80A, E80D, S92A, S92T, S92Y, Q105A, C166A, C166T; blue) show similar secondary structural content as the wild type (black). (B) RutB Y35F (orange) displays significant loss of secondary structure, compared to the wild type (black). Variants C166S (red), D24N (yellow), N72A (green), and Y136A (blue) show a significantly increased signal at 224 nm, indicating an increased amount of β-strands. (C) Y136F (red), K133A (yellow), C166K (green), and D24A (blue) produce overall increased CD spectra, with relatively increased signals at 210 nm, indicating a shift towards α-helical content [165].

4.2.4 Crystal Structure Determination

Protein crystals of RutB were grown by the hanging drop method (see subsubsection 2.4.4.1). Initial screens revealed five promising crystallization conditions (MD1-30 B5; MD1-37 2-32; MD1-46 1-14; MD1-46 2-8; PEGRx D10; see subsubsection 2.1.3.4). Optimization of MD1-30 B5 (1 M NaAc; 0.1 M imidazole; pH 6.5) produced the best results with large, high quality protein crystals growing within 48 h (see Figure 8.57). Ideal crystallization conditions were determined to be 1.1-1.2 M NaAc, 0.1 M imidazole pH 5.75 for the wild-type protein (see Figure 8.58).
Selenomethionine-labeled RutB (RutB SeMet) was produced as described (see subsection 2.4.2) and crystallized for initial structure determination. RutB vari-

ants D24N and C166S were catalytically inactive (see Table 4.2) and thus chosen for co-crystallization with UAc. Furthermore, wild-type RutB was co-crystallized with the substrate analogue ureidopropionate (UPro). Resulting crystals were analyzed at the *Swiss Light Source* synchrotron by Dr. Chitra Rajendran. Parameters of data acquisition are presented in Table 4.1.

Table 4.1: Crystallographic and refinement data of RutB and RutB variants
Data collection and refinement statistics for the protein structures solved in this work. Crystal analysis and refinement was performed by Dr. Chitra Rajendran. Statistics for the highest-resolution shell are shown in parentheses.

	RutB SeMet	RutB (UPro)	RutB D24N	RutB C166S
Wavelength [Å]	1.00	0.979	0.99	1.00
Resolution range [Å]	47.67 - 1.9 (1.96 - 1.9)	47.89 - 1.4 (1.45 - 1.4)	41.08 - 1.88 (1.94 - 1.88)	47.63 - 1.54 (1.59 - 1.54)
Space group	P 1 2_1 1	P 1 2_1 1	P 1 2_1 1	P 1
Unit cell a,b,c [Å]	121.59 110.89 128.71	122.92 111.61 128.98	58.11 125.10 123.26	69.45 108.60 124.34
α, β, γ [°]	90 99.12 90	90 99.12 90	90 93.21 90	85.66 83.21 74.34
Total reflections	1401020 (863)	7349015 (367681)	90 93.21 90	85.66 83.21 74.34
Unique reflections	223328 (714)	661819 (64589)	139494 (13587)	472783 (47452)
Multiplicity	6.3 (1.1)	11.1 (5.7)	6.8 (6.3)	3.3 (3.3)
Completeness [%]	83.44 (2.71)	98.29 (96.14)	97.77 (95.82)	87.81 (92.53)
Mean I/sigma(I)	5.50 (0.08)	11.03 (1.00)	7.25 (1.36)	7.20 (1.06)
Wilson B-factor	27.49	15.05	23.34	16.45
CC1/2	0.98 (0.19)	0.99 (0.4)	0.995 (0.584)	0.974 (0.39)
CC*	0.99 (0.57)	1 (0.75)	0.999 (0.859)	0.993 (0.75)
R-work	0.259 (0.537)	0.169 (0.299)	0.254 (0.356)	0.211 (0.329)
R-free	0.313 (0.602)	0.212 (0.340)	0.296 (0.394)	0.241 (0.348)
Number of non-hydrogen atoms	28686	32562	14641	30893
macromolecules	27442	27547	13716	27440
ligands	-	122	-	224
solvent	1244	4913	925	3325
Protein residues	3552	3567	1776	3552
RMS (bonds)	0.011	0.005	0.009	0.012
RMS (angles)	1.49	0.76	1.10	1.17
Ramachandran favored [%]	97.27	98.19	96.48	97.84
Ramachandran allowed [%]	2.67	1.81	3.47	2.16

	RutB SeMet	RutB (UPro)	RutB D24N	RutB C166S
Ramachandran outliers [%]	0.06	0.00	0.06	0.00
Average B-factor	34.48	23.05	25.63	21.95
macromolecules	34.51	20.40	25.44	20.84
ligands	-	22.29	-	17.35
solvent	33.72	37.91	28.53	31.22

The RutB structure was initially determined from the anomalous scattering of the crystals of selenomethionine-labeled RutB and structures of the remaining RutB variants were solved by molecular replacement based on this structure. Protein structures could be determined with a resolution of 1.4 to 1.9 Å. Monomeric assemblies of RutB consist of 222 residues with M1 and the 15 C-terminal residues (including the His_6 tag) not resolved.
RutB exhibits a typical IHL fold (see Figure 4.5A). A central β-sheet, consisting of six parallel strands (β-5, β-4, β-3, β-6, β-7, β-8), surrounded by five α-helices (α-3, α-9, α-10, α-11, α-12) forms the enzyme's core. α-1 and α-2 are part of a hydrophobic loop containing a number of aromatic and hydrophobic residues (Y29, Y35, L36, F41) directed towards the active site cavity. Residues 73-116 form an extensive loop structure with a number of short, interspersed α-helical elements (α-4, α-5, α-6, α-7, α-8), covering the active site. The first 13 N-terminal residues of RutB form a two strand antiparallel β-hairpin (β-1, β-2).
CASTp analysis [33] reveals an active site cavity with a solvent-accessible surface of 45 Å^2, which is roughly delimited by the central β-sheet, the β-6-α-10 loop, the hydrophobic loop containing α-1 and α-2, and the long loop containing α-4-8 (see Figure 4.5B). Residues involved in the formation of the active site, as predicted by CASTp, are Y29, Y35, L36, F41, N72, W74, N88, S92, N93, A94, L109, R135, Y136, A162, V165, C166, and A191. The residues constituting the presumed catalytic triad (D24, K133, C166) are located at the C-terminal ends of β-3 and β-5 as well as at the N-terminal end of α-10, respectively.
RutB crystallizes as a dimer-of-dimers with two distinct dimerization interfaces (see Figure 4.5C). PISA analysis [38] calculates an area of 1831 Å^2 and a stabilization energy Δ^iG = -19.9 kcal/mol for the main dimerization interface. The secondary interface is significantly smaller (467.2 Å^2) and contributes less to the stabilization of the complex (Δ^iG = -4.7 kcal/mol). The main dimerization interface relies on stacking of α-10 and α-11 of each monomer, involving residues E168, S169, R172, and D173 (α-10), as well as F195, K198, F202, and N203 (α-11). Furthermore, contacts of α-5/α-6 and the β-5-α-9 loop to the C-terminal ends of α-10/α-11 and the N-terminal hairpin structure (β-1, β-2) of the neighboring monomer contributes to dimerization across the main interface. Specific residue pairs in this interaction are K91-E9', K96-E205' (ionic bond), Y136-F208' (π-stacking), R135-F175'-F176' (cation-π), and N92-T206' (H-bond), where a dash indicates residues of the respective monomer part-

ner. The secondary dimerization interface consists of interactions between α-9 and α-4/α-4-α-5 loop as well as portions of the β-5-α-9 loops of neighboring monomers. Residue interactions in this interface are restricted to a number of hydrogen bonds (V79-R151', R135-S146', S86-S150', and R135-S150').

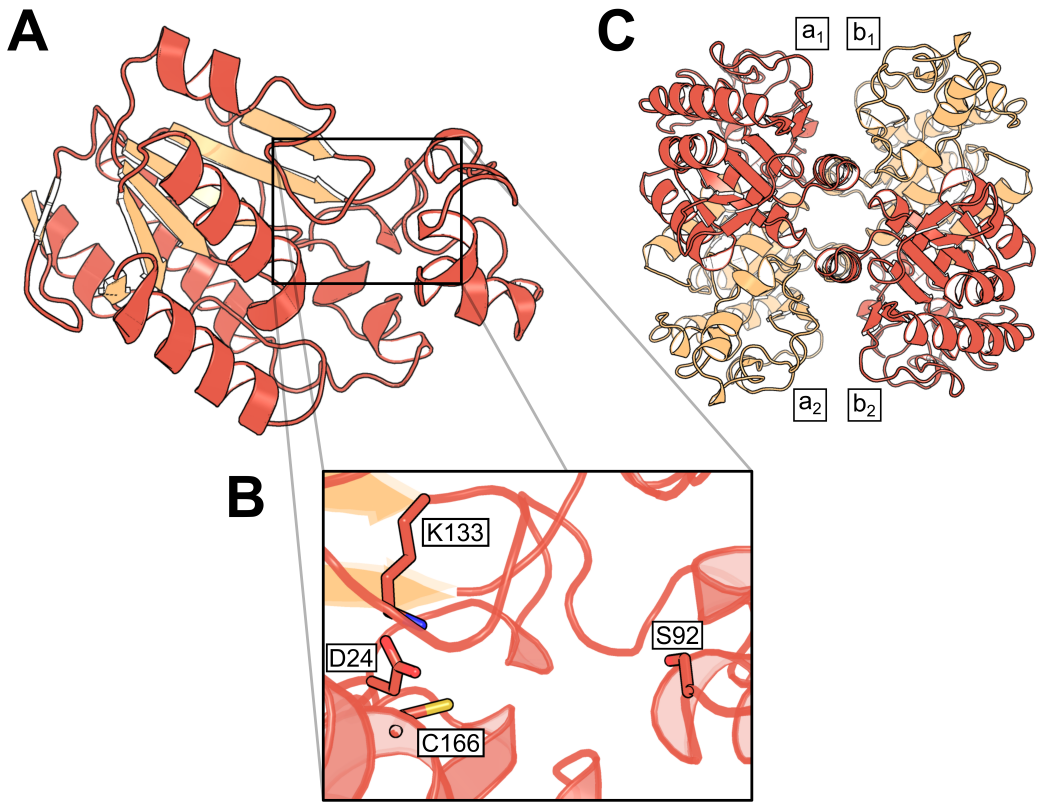


Figure 4.5: Structural Features of RutB
(A) RutB exhibits a typical isochorismatase-like hydrolase (IHL) fold. A central β-sheet, consisting of six parallel strands, is surrounded by five α-helices. The active site is indicated by a black box. (B) D24, K133, and C166 constitute a catalytic triad often observed in IHL enzymes [126]. As discussed below (see subsection 4.3.2), S92, located across the active site cavity from the catalytic triad, plays an important role in substrate binding. (C) RutB displays a tetrameric organization in the form of a dimer-of-dimers with differently sized dimerization interfaces. The main dimerization interface lies between monomers a_1 and a_2 as well as between b_1 and b_2. A minor dimerization interface connects a_1 and a_2 to b_1 and b_2, respectively.

After residue assignment, additional electron density was observed in the active site of enzyme variants co-crystalized with UAc or UPro. Detailed investigation by Dr. Chitra Rajendran revealed the presence of UPro in several monomers of wild-type RutB and two molecules of acetate in each active site of RutB C166S

(see Figure 4.6). The implication of these findings will be discussed below (see subsection 4.3.2).

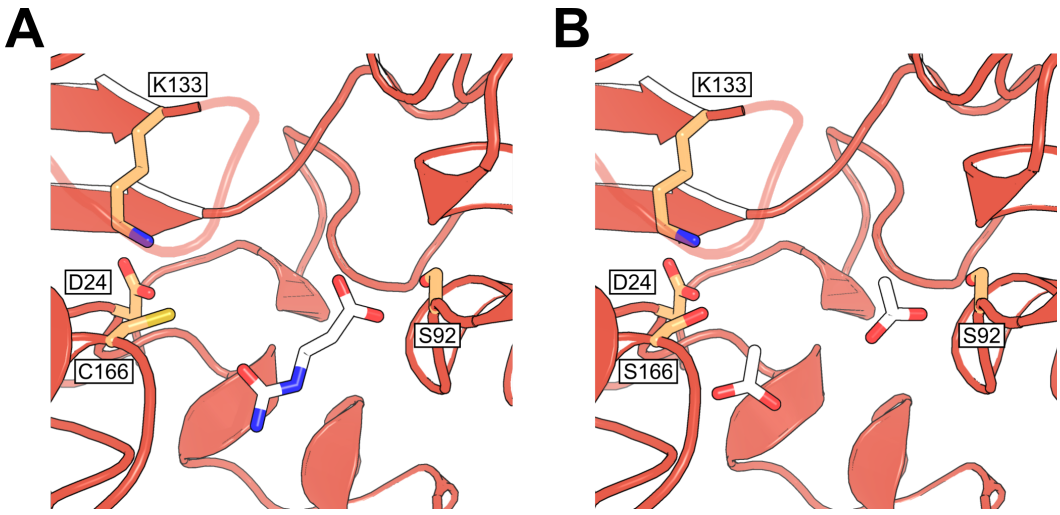


Figure 4.6: Ligands of wild-type RutB and RutB C166S
Structure determination of wild-type RutB (A) and RutB C166S (B) revealed electron density that points to the presence of ureidopropionate (UPro) and two molecules of acetate, respectively. Active site residues (D24, K133, C/S166) and possible substrate binding residue S92 are shown (yellow sticks). Carboxy groups of the ligands are found coordinated to the serine residues S92 and S166. UPro was not identified in all monomers of the ensemble. RutB C166S was co-crystallized with ureidoacrylate but excess acetate (>1 M) was present in the crystallization solution which may have been favorably bound due to the C166S mutation.

The protein structures solved in this work (see Table 4.1) display no significant conformational changes (all-atom RMSD <= 0.633), despite the presence of different residue exchanges and ligands. Conformational changes observed in CD spectroscopy (see subsection 4.2.3) apparently do not translate into the crystal structure.

4.2.5 Establishment of Steady-State Enzyme Assays

A steady-state enzyme assay was established to quantify the effects of the introduced amino acid exchanges on the kinetic parameters for the RutB reaction. Direct photometric tracking of the reaction process by following the conversion of ureidoacrylate into aminoacrylate (see Figure 4.1) at 278 nm was not possible due to a strong absorption signal of the RutB enzyme itself (see Figure 4.7).

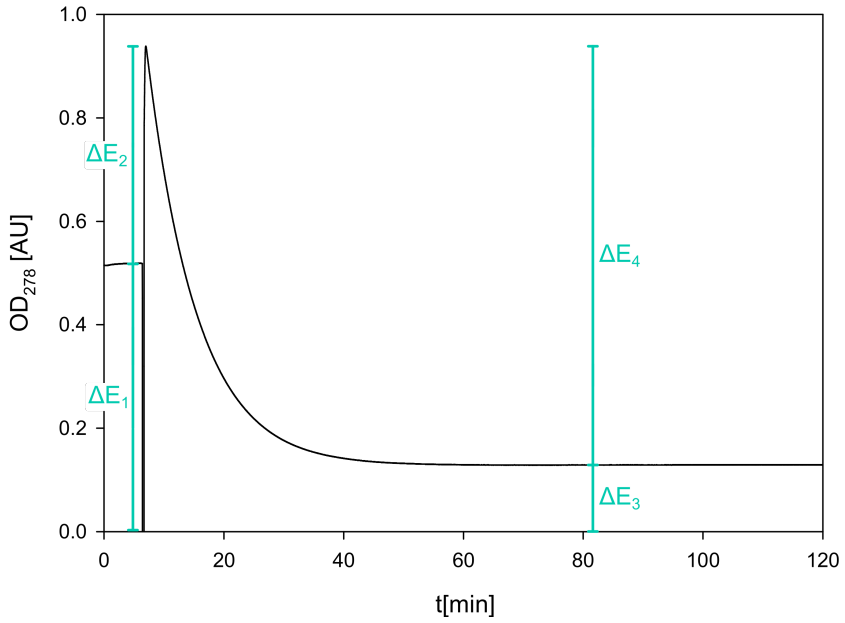


Figure 4.7: Direct photometric measurement of the conversion of ureidoacrylate by wild-type RutB
The assay shown was conducted as described in subsubsection 2.4.3.1. The reaction was started by addition of enzyme at 7 min and absorption at 278 nm was tracked over the reaction time. This direct measurement of enzyme activity was not suitable for determination of kinetic parameters due to a strong absorption signal after addition of enzyme (ΔE_2), that is not adequately explained by its native absorption, which is apparent after the reaction has proceeded (ΔE_3). Accordingly, the total decrease in absorption signal over the reaction (ΔE_4) is higher than would be expected from the initial baseline, when only ureidoacrylate is present (ΔE_1). This discrepancy may indicate a strongly absorbing complex or rearrangement of aromatic residues in the enzymatic reaction, which could lead to unpredictable effects in different enzyme variants. A glutamate dehydrogenase-coupled assay was used for kinetic measurements instead (see Figure 4.8).

Instead, the reaction of RutB was coupled to glutamate dehydrogenase (GDH) by the formation of ammonia that follows the hydrolysis of UAc. The resulting NADH consumption allowed for monitoring of the RutB reaction at 340 nm (see Figure 4.8).

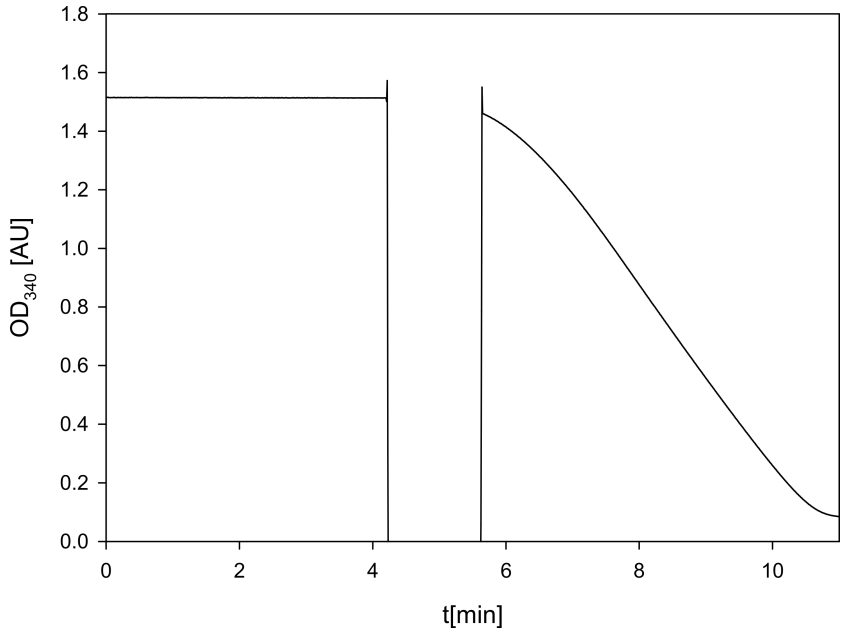


Figure 4.8: Glutamate dehydrogenase-coupled measurement of the conversion of ureidoacrylate by wild-type RutB
The assay shown was conducted as described in subsubsection 2.4.3.2. Spontaneous production of ammonia, following hydrolysis of ureidoacrylate (see Figure 2.3), was coupled to the NADH-dependent reaction of glutamate dehydrogenase (GDH), producing a photometric signal. The reaction was started by addition of 2.5 μM RutB at 5 min and absorption at 340 nm was tracked over the reaction time. The reaction displays an initial lag phase, that could be alleviated by reduction of enzyme concentration to 10 nM (not shown). This lag phase might correlate to the observed increase in absorption at 278 nm in the direct assay (see Figure 4.7) and might indicate a slow pre-steady state process of RutB.

Steady-state enzyme kinetics were measured in triplicate in a plate reader in UV-transparent plates (UV-STAR® 96-well; see subsubsection 2.4.3.2). Initial slopes were plotted against substrate concentrations and the resulting data points were fitted with the Michaelis-Menten equation, yielding the catalytic parameters k_{cat} and K_M from which the catalytic efficiency (k_{cat}/K_M) was calculated. In variants with exceedingly high K_M values, where only the initial slope of the Michaelis-Menten curve could be observed, a linear fit was applied, which solely revealed the catalytic efficiency k_{cat}/K_M.

4.2.6 Mutational Studies

Based on the crystal structure of RutB, a number of active site residues were chosen for mutational studies (see Figure 4.9). Residue exchanges Y29F and Y35F are located within the short hydrophobic loop near the active site and

may reveal the relevance of its exposed aromatic residues. N72A, W74A, W74F, Y136A, and Y136F lie in close proximity to the catalytic triad and may prove instrumental in disrupting substrate binding or hydrolysis. Residue S92 is positioned across the active site from the catalytic triad and could conceivably act to precisely position the substrate towards these catalytic residues. S92A was introduced to investigate the importance of this position, while residue exchanges S92T and S92Y were conceived to explore the possibility of alternative modes of substrate binding. Mutation of positions E80 and Q105 (E80A, E80D, Q105A) were designed as negative controls due to their distance from the active site. Various mutations of the presumed catalytic triad (D24A, D24N, K133A, C166A, C166K, C166S, C166T) were investigated. The location of all residue exchanges characterized in this work are outlined in Figure 4.9.

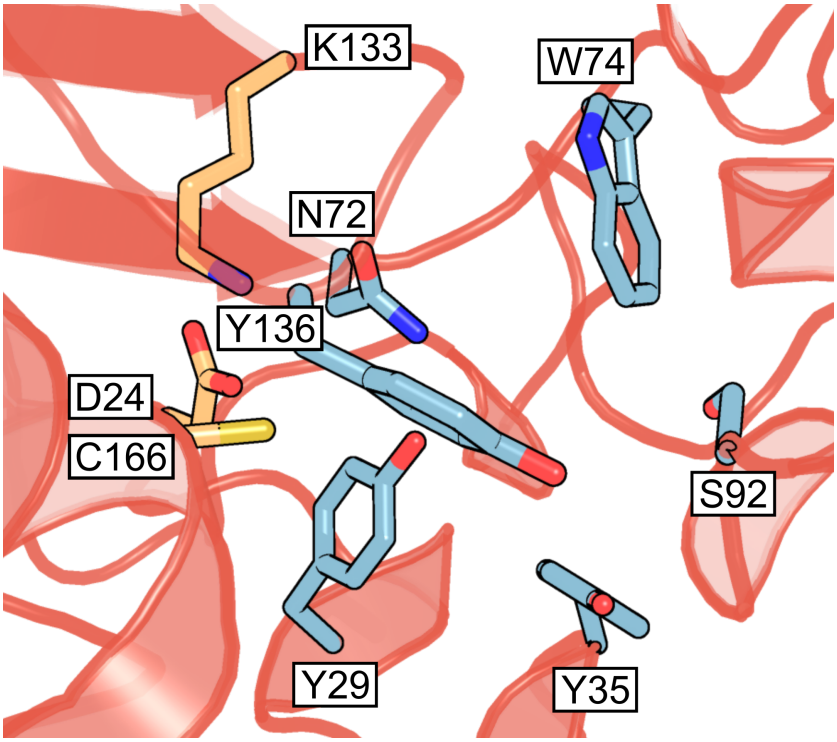


Figure 4.9: Locations of residue exchanges introduced into RutB
Representation of the active site cavity of RutB, defined by the presumed catalytic triad (D24, K133, C166; yellow sticks). The importance of these residues was investigated by applicable residue exchanges (D24A, D24N, K133A, C166S, C166T, C166Y). Residues potentially involved in the mechanism of RutB (blue sticks) were mutated to determine their respective effect on catalytic parameters (Y29F, Y35F, N72A, W74A, W74F, S92A, S92T, S92Y, Y136A, Y136F). The residue exchanges E80A, E80D, and Q105A (not shown) were introduced as a point of reference.

Catalytic parameters of the resulting enzyme variants are given in Table 4.2. Kinetics of each enzyme variant are shown in subsection 8.3.4. The catalytic parameters of wild-type RutB for ureidoacrylate are k_{cat}=0.86 s^{-1} and K_M=52 μM, resulting in a k_{cat}/K_M of 1.7e4 $M^{-1}s^{-1}$. Any mutation of the presumed catalytic triad (D24A, D24N, K133A, C166S, C166T, C166Y) lead to an almost complete inactivation of the resulting variant to the point where the activity could not be quantified (<0.005 s^{-1}), confirming the essential role of D24, K133, and C166. Residue exchanges N72A, Y136A, and Y136F yielded significant decreases of k_{cat} (<10% wild-type) and Y29F resulted in a moderate reduction in k_{cat} (approximately 20% wild-type). Notably, E80A and S92A lead to a more than 1.5-fold increase in k_{cat}. Residue exchanges W74F, E80A, E80D, and S92A increased the K_M value more than 10-fold in relation to the wild-type, while Y136A produced an approximately 7-fold increase in K_M. Introduction of W74A, S92T, or S92Y leads to K_M values too high to be determined.

Table 4.2: Catalytic parameters of RutB variants for ureidoacrylate
Steady-state enzyme kinetics for the hydrolysis of ureidoacrylate (UAc) at 25°C were measured in triplicate (see subsubsection 2.4.3.2). Fitting of the initial conversion rates plotted against substrate concentrations with the Michaelis-Menten equation yielded the catalytic constant (k_{cat}) and the Michaelis constant (K_M), from which the catalytic efficiency (k_{cat}/K_M) could be calculated. In variants with radically increased K_M values, a linear fit had to be applied which revealed only the catalytic efficiency. The standard error is given for all values.

Protein	k_{cat} [s^{-1}]	K_M [μM]	k_{cat}/K_M [$M^{-1}s^{-1}$]
WT	0.86±0.017	52±4.6	1.7e4±1.8e3
Y29F	0.17±3.3e-3	25±2.7	6.6e3±8.4e2
Y35F	0.62±0.017	1.7e2±16	3.7e3±4.5e2
N72A	0.086±4.7e-3	94±19	9.2e2±2.3e2
W74A	-	-	3.7±0.31
W74F	0.76±0.093	2.3e3±4.5e2	3.3e2±1.1e2
E80A	1.5±0.12	1.1e3±1.9e2	1.4e3±3.6e2
E80D	1.0±0.099	8.0e2±1.4e2	1.3e3±3.5e2
S92A	1.4±0.095	5.1e3±7.7e2	2.8e2±61
S92T	-	-	3.0±0.044
S92Y	-	-	0.42±6.4e-3
Q105A	1.2±0.025	84±6.6	1.4e4±1.4e3
Y136A	0.034±3.1e-3	3.5e2±1.1e2	99±40
Y136F	0.031±2.3e-3	20±8.4	1.5e3±7.4e2
D24A	<0.005	-	-
D24N	<0.005	-	-
K133A	<0.005	-	-

Protein	k_{cat} [s^{-1}]	K_M [μM]	k_{cat}/K_M [$M^{-1}s^{-1}$]
C166A	<0.005	-	-
C166S	<0.005	-	-
C166T	<0.005	-	-
C166K	<0.005	-	-

4.3 Discussion

4.3.1 RutB Crystal Structure

In this work, the crystal structure of the ureidoacrylate amidohydrolase RutB from *E. coli* was solved for the first time. RutB exhibits a typical isochorismatase-like hydrolase (IHL) fold with a Rossman-like core structure consisting of a six strand parallel β-sheet surrounded by five α-helices (see subsection 4.2.4). The enzyme features a catalytic triad, consisting of D24, K133, and C166 (see Table 4.2), which is also found in other IHL enzymes [126].
Comparison of the crystal structure of RutB with other IHL superfamily enzymes revealed general similarities and characteristic features of RutB. Relevant structures for comparison are isochorismatase from *Pseudomonas aeruginosa* (PhzD; PDB: 1NF8; [112]), isochorismatase from *E. coli* (EntB; PDB: 2FQ1; [111]), a putative isochorismatase hydrolase from *Oleispira antarctica* (OaIHL; PDB: 3LQY; [126]), isochorismatase from *Vibrio cholerae* (VibB; PDB: 3TB4; [166]), biuret hydrolase from *Rhizobium leguminasorum* (BiuH; PDB: 6AZN; [140]), and triuret hydrolase from *Herbaspirillum sp. BH-1* (TrtA; PDB: 6XJE;[139]).
In some IHL enzymes (PhzD, EntB, VibB), the N-terminus is a long, mostly disordered region, reaching around the respective enzyme's core structure and partially closing off the active. This is not the case in RutB, where the N-terminus forms a β-hairpin structure, that is part of the main dimerization interface. The β-hairpin structure is shared with TrtA, while the N-terminus of BiuH adopts a similar conformation but lacks β-strands. Based on visual inspection, the N-termini of TrtA and BiuH may similarly be involved in dimerization.
RutB possesses a short loop with two α-helical elements (α-1 and α-2) which contain a number of aromatic and hydrophobic residues (Y29, Y35, L36, F41) exposed to the active site cavity. This structural element has been noted in other IHL enzymes and a correlation between the identities of the exposed residues and substrate specificity has been investigated in some cases. In BiuH, exchange of F41 for alanine, leucine, tyrosine, or tryptophan reduced the enzyme's catalytic efficiency for biuret to below 1.5% of the wild-type value [140]. Mutation of the exposed residues of this loop element was conducted in an unsuccessful attempt to shift the activity of TrtA towards biuret hydrolysis [139]. In another part of this work, full combinatorial randomization of Y29, Y35, and F41 failed to establish an isochorismatase activity on the RutB scaffold (see subsec-

tion 3.3.2). Involvement of Y29 and Y35 in the reaction mechanism of RutB is discussed below (see subsection 4.3.2).
The RutB structure features a relatively long extension near its active site cavity, encompassing a total of 46 residues (72-117). This portion of the protein consists of short α-helical segments interspersed with connecting loops and contains the residues N72, W74, and S92 which were shown to be essential for RutB activity in this work (see Table 4.2). Similar features found in other members of the IHL superfamily are often significantly shorter at around 30 residues in PhzD, EntB, and VibB and as little as little as 19 residues in OaIHL. BiuH and TrtA possess extensions similar in length to that of RutB (43 and 42 residues, respectively) but with different topologies. Taking into consideration further structures of IHL enzymes (PDB: 2A67; 4H17, 3MCW, 3OQP; all from unpublished results), shorter versions of this extension typically correlate with the presence of an additional β-hairpin structure (residues 152-159 in OaIHL) as noted before [126]. Apart from its apparent role in the formation of the active site of RutB, this long extension might fulfill the role of an interface add-on [167], possibly facilitating a specific complex with the aminoacrylate peracid reductase (RutC), which is the next downstream enzyme in the Rut pathway. While no such complex has been shown to exist yet, it might indeed be beneficial for the organism to tightly control the release of the RutB product aminoacrylate peracid, which has been described to be toxic [52]. The hypothetical RutBC complex might aid in alleviating this danger and accelerate the degradation of aminoacrylate peracid and the similarly toxic downstream product malonic semialdehyde. The position of the extension as well as the direction of the active site cleft in the oligomeric state of RutB (see Figure 4.5) would both allow for the formation of such a complex.
By extension of the asymmetric unit, the oligomeric state of RutB was found to be a dimer-of-dimers (see Figure 4.5), with the two dimerization interfaces drastically differing in area and the number of residues involved (see subsection 4.2.4). The mode of dimerization across the main dimerization interface is common in IHL enzymes, with PhzD, EntB, VibB, and TrtA exhibiting similar dimeric complex structures. Notably, the tetrameric assembly of RutB is similarly exhibited by BiuH.
In an attempt to generate *holo* structures of RutB, wild-type RutB was co-crystallized with the substrate analogue ureidopropionate and catalytically inactive variants (D24N, C166S) were further co-crystallized with the RutB substrate ureidoacrylate. The resulting structures revealed that neither the mutations nor the presence of ligands cause any structural or conformational change (see subsection 4.2.4). In particular, no conformational change from an open to a closed state, as described for TrtA [139], could be observed.

4.3.2 Reaction Mechanism

The crystal structure determination and in-depth mutational studies conducted in this work allow for the formulation of a detailed reaction mechanism of RutB. Two classes of IHL enzymes have previously been distinguished [126]. The catalysis of one class, containing an aspartate-lysine catalytic dyad, likely proceeds *via* simple acid-base catalysis [112], while the other class of enzymes, containing an aspartate-lysine-cysteine catalytic triad, initiates their respective reactions by a nucleophilic attack of the catalytic cysteine [140] [139]. RutB belongs to the second class of enzymes, with the corresponding catalytic residues being D24, K133, and C166. Any mutation of these residues investigated in this work lead to almost complete loss of enzyme activity (see Table 4.2), confirming their relevance for the catalysis of RutB.
A structural homology search with DALI [34], reveals the recently discovered enzymes biuret hydrolase (BiuH; PDB: 6AZN; [140]) and triuret hydrolase (TrtA; PDB: 6XJE; [139]) as the closest structural homologues of RutB, with all-atom RMSD values of 5.8 Å and 4.1 Å, respectively. Pairwise sequence identities of RutB to BiuH and TrtA are 28% and 24%, respectively. Structural similarities of the three enzymes are discussed above (see subsection 4.3.1). Like RutB, these two enzymes feature an aspartate-lysine-cysteine catalytic triad and catalyze hydrolysis reactions at ureido moieties of their respective substrates. Accordingly, the detailed reaction mechanisms described for these enzymes [140] [139] are reasonable blueprints for the mechanism of RutB.
Co-crystallization of wild-type RutB with the substrate analogue ureidopropionate (UPro) yielded a protein structure with excess electron density within the active site. In some monomers of the asymmetric unit, this density can plausibly be fitted with UPro (see Figure 4.6), revealing a coordination of the compound's carboxy group to S92 of RutB, which places the ureido moiety of UPro towards the catalytic triad, although in an apparently catalytically unproductive conformation. In the crystal structure of the catalytically inactive variant RutB C166S, co-crystallized with ureidoacrylate (UAc), electron density was observed in the active site that can plausibly be explained by two acetate molecules, which are coordinated to S92 and S166, respectively. Both complex structures display coordination of a carboxy group to S92 of RutB, which may similarly coordinate the carboxy group of UAc in the RutB reaction.
In accordance with this assumption, introduction of the residue exchange S92A into RutB leads to a drastic increase in K_M (approximately 100-fold). Moreover, RutB S92T and S92Y are virtually inactive, indicating the importance of a precisely positioned and sterically unhindered hydroxyl group for catalysis. The residue W74 apparently plays a significant role in the mechanism of RutB, with the W74A residue exchange leading to almost complete inactivation of the enzyme. The W74F exchange leads to a notable increase in K_M (approximately 45-fold) with little effects on the k_{cat}, implying an involvement of the tryptophan in substrate binding. The residue exchange N72A produces an approximately 10-

fold reduction in k_{cat} implying a modest involvement in the catalytic mechanism. Based on the location of N72, an activation of the ureido moiety of ureidoacrylate appears plausible. Both Y136A and Y136F produce a significant reduction in k_{cat} (>25-fold decrease), indicating a relevance of the hydroxyl group of Y136 in the catalytic mechanism, plausibly through activation of the substrate's ureido moiety. A minor K_M effect (7-fold increase), observed only in Y136A, may point to an additional role of the aromatic ring in substrate binding. Despite being conceived as negative controls, residue exchanges E80A and E80D lead to a significant increase in K_M and a slight increase in k_{cat}. This effect is likely caused by a conformational shift of the long α-helical loop, in which position 80 is located, that indirectly causes a rearrangement of S92. This theory appears plausible, as the introduction of S92A similarly leads to increased K_M and k_{cat} values.

Based on the results discussed above, a plausible reaction mechanism for the ureidoacrylate amidohydrolase activity of RutB can be proposed (see Figure 4.10). The ureidoacrylate substrate is bound in the active site by S92 and may further be guided by W74 towards the catalytic triad. N72 and Y136 likely activate the ureido moiety and position the substrate in a catalytically productive conformation. The catalytic cysteine residue, C166, is deprotonated by D24, which may in turn be deprotonated by K133, as proposed for BiuH [140]. C166 then initiates the hydrolysis reaction through a nucleophilic attack to the carbonyl carbon of the ureido group of the substrate, creating a covalent tetrahedral intermediate, which releases aminoacrylate after proton donation from D24. The remaining carbamoyl group bound to C166 is then hydrolytically removed by nucleophilic attack of a water molecule, which may be activated through deprotonation by D24, liberating carbamate as the second product and restoring the original state of the enzyme. This reaction mechanism is also applicable to the other RutB substrates ureidoacrylate peracid and (Z)-2-methylureidoacrylate.

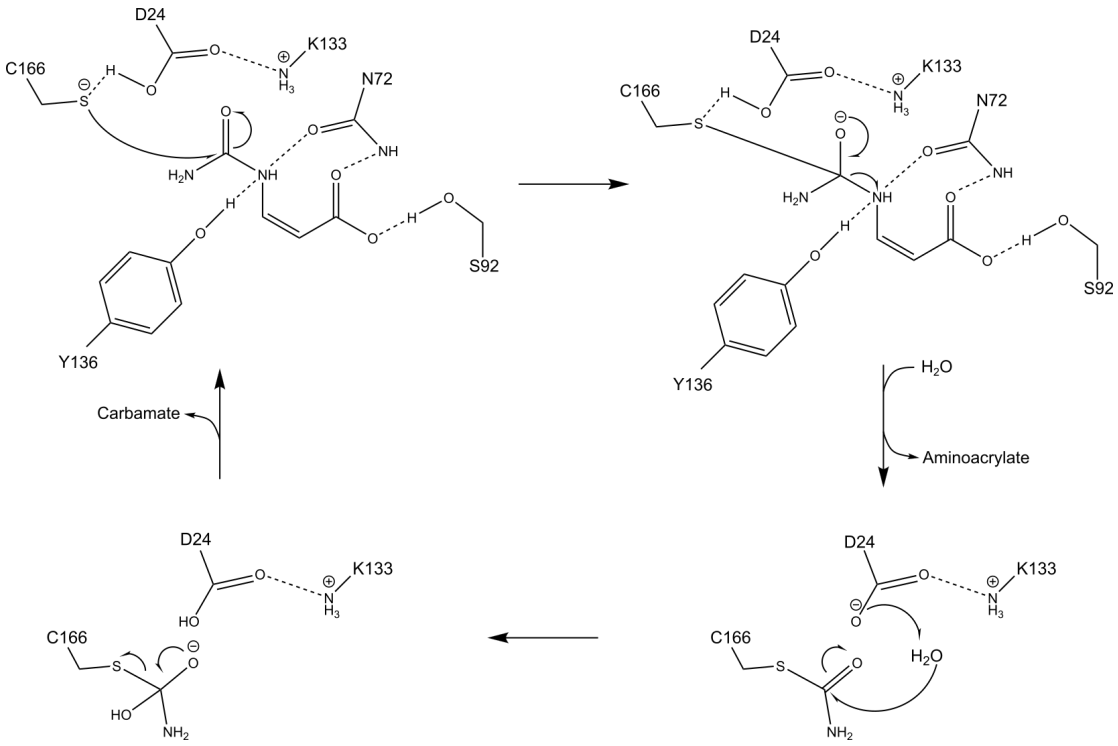


Figure 4.10: Proposed reaction mechanism of the ureidoacrylate amidohydrolase RutB
The RutB substrate ureidoacrylate (UAc) is bound in the enzyme's active site by S92 (top left). W74 (not shown) is relevant for substrate binding in a way that is currently not understood. N72 and Y136 precisely position UAc and possibly activate the ureido moiety for nucleophilic attack of the catalytic cysteine residue (C166). C166 is deprotonated by D24 which is in turn deprotonated by K133 and the cysteine attacks the carbonyl carbon of the ureido moiety, forming a covalent tetrahedral intermediate (top right). The first reaction product, aminoacrylate, is released from this intermediate with D24 serving as a proton donor, leaving a carbamoyl group bound to C166 (bottom right). This covalent intermediate is hydrolyzed by a water molecule activated through deprotonation by D24, which leads to release of carbamate as a second product (bottom left) and reversion of the enzyme to the starting state.

4.4 Outlook

Owing to its relatively late discovery in 2006 [154], many aspects of the Rut pathway and its enzymes remain to be explored. The results presented in this work particularly reveal new directions for the further research of the ureidoacrylate amidohydrolase RutB from *E. coli*.
The only currently known natural substrates of RutB are (Z)-3-ureidoacrylate and its peracid [52] [110] with (Z)-2-methylureidoacrylate and its peracid be-

ing other very likely substrates due to the facts, that *E. coli* utilizes thymine through the Rut pathway and a thymine oxygenase activity has been confirmed in RutA [48]. Exploring the substrate range of RutB might in turn reveal previously unknown substrates of the Rut pathway or possibly hint at the enzyme's evolutionary origin. RutB substrate analogs are easily accessible by chemical synthesis [168].
According to the KEGG database [169] the overwhelming majority of complete Rut pathways is found in the *Enterobacterales* group with notable exceptions being *Pseudomonadales*, *Agrobacterium*, and *Rhizobium*. This limited phylogenetic spread of the Rut pathway points towards its relatively recent emergence. The evolutionary history of the Rut enzymes may thus be explored by evolutionary reconstruction, possibly aided by ancestral sequence reconstruction [170].
Considering the functional similarity of the initial two steps of the Rut pathway to the degradation of cyanuric acid as part of the Atz pathway [171] [172] [173], it might well be possible to evolve the Rut pathway enzymes towards degradation of anthropogenic heterocyclic compounds like s-triazines, which might open up entirely new avenues for the field of bioremediation techniques.
With the determination of the crystal structure of RutB achieved in this work, the only remaining Rut enzymes without a known structure are RutE and RutF. Furthermore, the pyrimidine transport protein RutG remains unresolved. Solution of these protein structures might lead to a deeper understanding of the enzymes of the Rut pathway and might particularly inform research into possible heteromeric complexes of the Rut enzymes. Complex formation between RutB and certain downstream enzymes (RutC, RutD, RutE) appears plausible, as the pathway intermediates have been noted to be toxic [52]. A heteromeric complex might increase processivity and prevent release of the toxic intermediates into the cytoplasm. A long, α-helical loop in RutB (residues 73-116) may be involved in complex formation and could convey specificity for the interaction partner by acting as an interface add-on [167].
RutB [159] and other IHL enzymes [162] [163] have previously been employed for the enantiopure synthesis of the pharmaceutically relevant substance Vince lactam (2-azabicyclo[2.2.1]hept-5-en-3-one). The structural and functional understanding of RutB generated in this work may potentially aid in engineering RutB for increased selectivity, activity, and stability.

5 Experimental Reconstruction of the Evolution of the Recently Emerged Hydroxyatrazine Ethylaminohydrolase AtzB

5.1 Introduction

5.1.1 Environmental Effects of s-Triazine Herbicides

The advent of modern industrial chemistry has led to pervasive changes of the global ecosystem due to a multitude of anthropogenic substances being introduced into the environment at a large scale [174] [175]. Many of these compounds, including herbicides, insecticides, and synthetic antibiotics do not occur in Nature and are thus referred to as *xenobiotic*. Exposure of various organisms to these novel compounds has driven a number of specific evolutionary adaptations. Different resistance mechanisms to pesticides have evolved in target species, giving rise to herbicide resistant plants [176], insecticide resistant insects [177][175], and antibiotic resistant bacteria [178] [179]. Interestingly, the introduction of pesticides into the environment also induced evolutionary adaptations in non-target species. Several microbial strains have been isolated that utilize anthropogenic substances as nutrient sources through newly evolved enzymes and catabolic pathways [180] [181] [182] [183] [175] [184].
An important class of anthropogenic substances is derived from 1,3,5-triazine (s-triazine; see Figure 5.1). Within this substance group there are a number of herbicidal compounds which kill susceptible plants by coordinating to the quinone binding protein in photosystem II, thereby inhibiting photosynthetic electron transfer [185] [186]. The most common of these s-triazine herbicides are atrazine, ametryn, and simazine. Another industrially relevant s-triazine compound is melamine, which is used as basic material for the production of resins, foams, and plastic products [187].

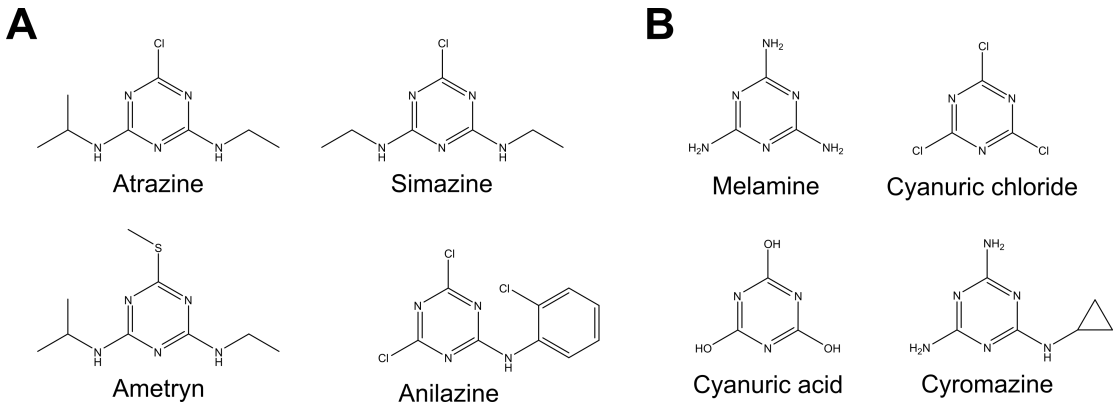


Figure 5.1: Anthropogenic s-triazine compounds
(A) The herbicides atrazine, simazine, ametryn, and anilazine all share an s-triazine core structure. (B) Melamine and cyanuric chloride are industrially relevant precursors of plastics and substituted s-triazines, respectively. Cyanuric acid is used as an additive in the chlorination of pool water and cyromazine is an s-triazine insecticide.

Since the 1950s, atrazine (2-chloro-4-ethylamino-6-isopropylamino-1,3,5-triazine) has been one of the most widely used s-triazine herbicides for the selective control of grassy and broadleaf weeds in corn, sorghum, sugarcane, and pineapple [188] [189] [190]. In the late twentieth century, atrazine and related herbicides became increasingly abundant and highly concentrated in soil and ground water due to their widespread application in agriculture and their low biodegradability [191] [192] [193] [194] [195]. Toxicological studies indicated that atrazine is hazardous to human health [196]. Nevertheless, s-triazine herbicides are still used today in many countries, with their consumption in the United States alone amounting to 30000 tons per year between 1992 and 2017 [197].
The widespread release of atrazine and structurally similar s-triazine herbicides into the environment constituted a novel evolutionary factor for plant life as well as for soil microbial communities. The immediate herbicidal properties generated a strong selection pressure for susceptible plant life to develop resistance mechanisms, the first of which has been described in 1970 [198], which had spontaneously emerged after merely a decade of periodic application of s-triazine herbicides to the area studied. The mode of resistance was later determined to be due to a single S→G amino acid exchange in protein D1 of plant photosystem II [199], decreasing affinity of the protein for s-triazine herbicides [200]. Notably, some photosynthetic bacteria with photosystems similar to those in plants are also susceptible to some s-triazine herbicides and have developed similar resistance mutations under laboratory conditions [201]. Other mechanisms of herbicide resistance in plants were later identified [202]. Notably, some plants may also acquire atrazine resistance by entering a symbiotic relationship with atrazine degrading microorganisms [203].

The introduction of anthropogenic s-triazine herbicides into the environment further constituted a new potential nutrient source for soil bacteria, leading to the emergence of s-triazine degrading enzymes within only a few decades [204]. Various bacterial species possessing these metabolic capabilities have since been isolated all over the globe [205] [183] [204]. Full mineralization of atrazine and other s-triazines is achieved by certain strains of *Agrobacterium* [206], *Arthrobacter* [207], *Chelatobacter* [208], *Pseudoaminobacter* [209], *Pseudomonas* [210] [211] [212] [186], and *Ralstonia* [213] [214]. Notably, genes for the degradation of atrazine have also been found in soils, that have not previously been treated with herbicides [215] [216].
Comparison of multiple s-triazine degrading species indicates strong variations in the presence of relevant genes [208] [217], with many organisms only producing single enzymes or an incomplete degradation pathway. Partial degradation of s-triazine herbicides is found in species of *Arthrobacter* [208], *Chelatobacter* [208], *Clavibacter* [218], *Ochrobactrum* [219], *Pseudomonas* [218] [219], *Rhizobium* [220], *Stenotrophomonas* [208], and *Streptomyces* [221]. Some of these species have been shown to contribute to the complete mineralization of atrazine within microbial consortia [218] [219].
The main evolutionary advantage of the degradation of s-triazine herbicides is the use of the breakdown products as carbon and nitrogen sources. Many of the isolates previously described can grow on a number of s-triazine substrates as the only sources of carbon and nitrogen [207] [209] or as the only source of nitrogen [208]. Depending on the metabolic background, the alkylamine side chains liberated from the s-triazine ring can be used as sources of N and C [207] and the heterocyclic core structure can be utilized as N source [222]. Interestingly, in some bacteria only specific alkylamines can be metabolized [209].

5.1.2 The Atz Pathway

The Atz pathway is the most abundant and best researched pathway for the degradation of atrazine and other s-triazine herbicides. It was first identified in *Pseudomonas sp.* strain ADP, which was isolated from an atrazine spillage site in Minnesota [211]. The Atz pathway is comprised of eight proteins (AtzABCDEFGH) and can be divided into an *upper pathway* (see Figure 5.2A) and a *lower pathway* (see Figure 5.2B).

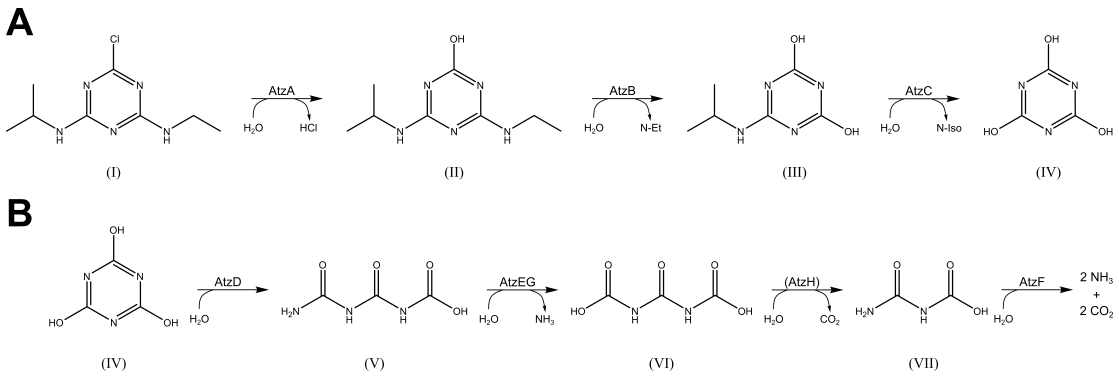


Figure 5.2: The Atz pathway
(A) The herbicide atrazine (I) is converted by atrazine chlorohydrolase (AtzA) into hydroxyatrazine (II), which serves as substrate for hydroxyatrazine ethylaminohydrolase (AtzB). AtzB releases ethylamine (N-Et) from the s-triazine ring, yielding N-isopropylammelide (III), which is in turn hydrolyzed by N-isopropylammelide isopropylaminohydrolase (AtzC), giving isopropylamine (N-Iso) and cyanuric acid (IV) as products. The sequential actions of AtzA, AtzB, and AtzC is considered the upper pathway. (B) Cyanuric acid hydrolase (AtzD) converts (IV) into 1-carboxybiuret (V), which is further hydrolyzed by 1-carboxybiuret hydrolase (AtzE) in complex with the small protein AtzG [223]. The resulting 1,3-dicarboxyurea (VI) is decarboxylated through an unknown mechanism, possibly involving AtzH [224] producing allophanate (VII), which is hydrolyzed by allophanate hydrolase (AtzF) to give two molecules each of NH_3 and CO_2. The reactions leading to the mineralization of cyanuric acid are considered the lower portion of the Atz pathway.

In the first step of the upper pathway, atrazine is hydrolytically dechlorinated by atrazine chlorohydrolase (AtzA), yielding hydroxyatrazine [225] [226] [227]. AtzA is a homo-hexamer, with one essential Fe(II) metal ion per monomer [226] [228]. The second metabolic step requires hydroxyatrazine ethylaminohydrolase (AtzB) for the hydrolytic deamination of hydroxyatrazine giving N-isopropylammelide and ethylamine as products. AtzB was reported to be a Zn(II) metalloenzyme, which preferentially populates a homo-dimeric oligomerization state [229] [230]. In the third reaction, N-isopropylammelide isopropylaminohydrolase (AtzC) converts N-isopropylammelide into cyanuric acid by the hydrolytic removal of the second alkylamine group [231]. AtzC was found to be a homotetramer with Zn(II) in its active site [232] [233].
In the first step of the lower pathway, cyanuric acid is further metabolized via a ring-opening reaction catalyzed by cyanuric acid hydrolase (AtzD) [171] [234] [235]. 1-carboxybiuret hydrolase (AtzE), which requires the small protein AtzG as complex partner [223], and allophanate hydrolase (AtzF) catalyze the final two hydrolytic steps to give a total of three moles of both carbon dioxide and ammonia as the final products of cyanuric acid mineralization [236] [237] [238]

[239] [240]. Notably AtzH may play an additional role in the decarboxylation of the AtzE product, 1,3-dicarboxyurea [224].
In *Pseudomonas sp.* strain ADP the genes encoding the enzymes required for atrazine catabolism are located on a single, self-transmissible plasmid (pADP-1; [241]. The *lower pathway* genes (*atzDEFGH*) are arranged in an operonic structure, which has been shown to be controlled by nitrogen availability and cyanuric acid concentration [242] [243]. While cyanuric acid is a major industrial chemical, it already existed on prebiotic earth in low quantities [244] and thus is not exclusively anthropogenic. Hence, it is assumed that AtzDEFGH represent an ancient route for cyanuric acid degradation that became more prevalent due to industrial release of cyanuric acid and s-triazine precursors into the environment [245].
The genes of the *upper pathway* enzymes (*atzABC*), which are expressed constitutively and without any known regulation, are spread over a wide area of the pADP-1 plasmid with long interjacent sequences. Moreover, those sequences contain inverted repeats that point to the recent assembly of the plasmid by way of transposase activity [231] [246]. Together, these findings indicate that AtzA, AtzB, and AtzC have only recently evolved in response to the presence of atrazine in the environment [181] [247]. These proteins therefore represent one of the rare cases that allow for studying the rapid emergence of new enzymatic activities and the assembly of novel catabolic pathways [248].
An investigation into the origin of one of these enzymes has previously been conducted by Noor and co-workers, who studied in detail the evolutionary relationship between AtzA and the melamine deaminase TriA [249]. TriA differs in only nine amino acids from AtzA, yet both enzymes feature distinct activities and little promiscuity, implying a very recent divergence. Stepwise exchange of all differing positions established an evolutionary trajectory, along which natural evolution might have proceeded [249].
The recent emergence of AtzA, AtzB, and AtzC suggests that these enzymes have not evolved *de novo* but by gene duplication and diversification from hitherto unknown progenitor enzymes [250] [251] [252] [204]. The Innovation-Amplification-Divergence (IAD) model proposed by Bergthorsson and coworkers [9] [253] (see section 1.1) may be suitable to explain this process. The model suggests neofunctionalization in a progenitor gene, resulting in a new, physiologically relevant promiscuous side reaction. Following gene duplication, the two gene copies can independently undergo specialization towards one of the two functions. The possibility that new enzymatic functions may develop in an existing protein without impairing its original function (as proposed in the IAD model) is supported by the occurrence of promiscuous activities, which are low-level catalytic activities towards other substrates than the physiological one [254].
Previous work in our group [21] has identified a weak promiscuous guanine deaminase side activity in AtzB and established a rudimentary evolutionary trajectory towards an increase of this activity. These findings imply a guanine

deaminase progenitor of AtzB which possessed a hydroxyatrazine ethylaminohydrolase side activity and rapidly evolved into the extant AtzB enzyme after the introduction of atrazine into the environment. On the basis of these findings, it was attempted to further increase the promiscuous guanine deaminase activity of AtzB and, *vice versa*, to establish AtzB activity on the scaffold of certain homologous guanine deaminases.

5.2 Results

5.2.1 Bioinformatics

Models of AtzB (Accession: NP_862481.1) and relevant variants were generated by Dr. Julian Nazet utilizing AlphaFold [30] (see subsection 2.5.1). Homologues of AtzB were identified by BLAST (see subsubsection 2.2.1.1). Multiple sequence alignments (MSAs) of all AtzB sequences and the 40 closest sequence homologues (see Table 8.4) were generated and visualized as described (see subsubsection 2.2.1.2).

5.2.2 Mutational Planning

Residues conveying hydroxyatrazine ethylaminohydrolase activity can be presumed to be recent acquisitions of AtzB. Accordingly, divergent residues between AtzB and its closest homologues are likely to be causal for the respective enzymatic activity and introduction of the correlating residue exchanges into AtzB could potentially further increase the enzyme's guanine deaminase activity, as was previously achieved in our group [21]. Sequence logos of (a) all known AtzB sequences and (b) the 40 closest sequence homologues were aligned and compared (see Figure 5.3).

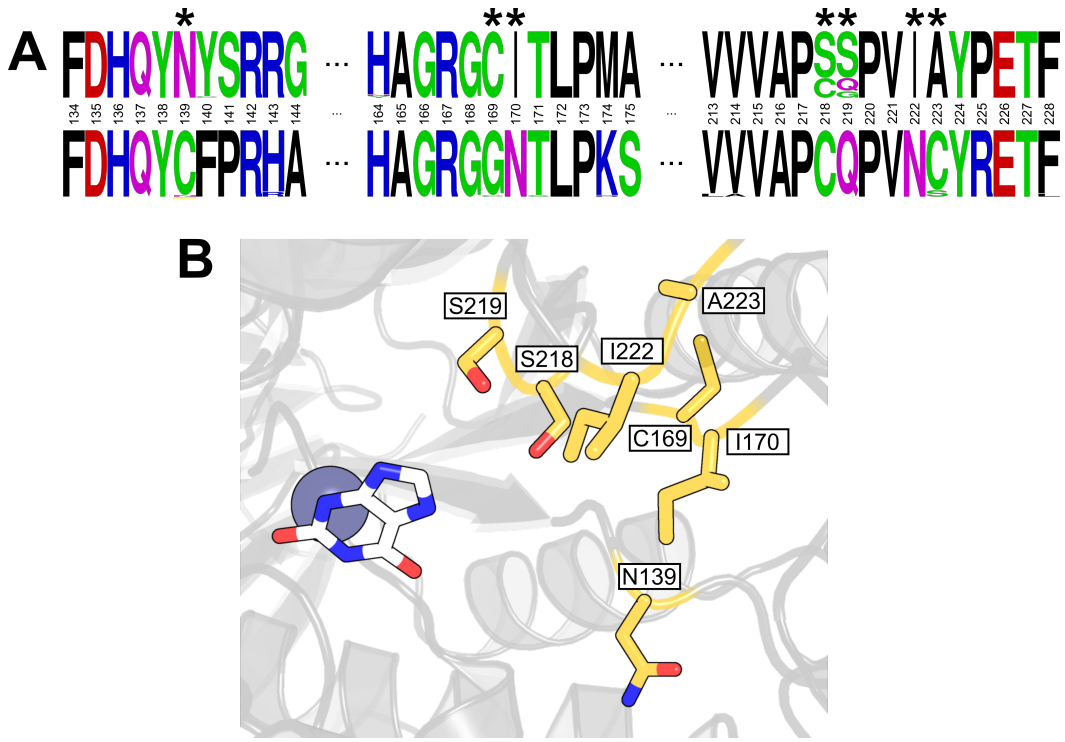


Figure 5.3: Sequence logos of AtzB and homologues in relation to the AtzB active site
(A) Relevant regions of the aligned sequence logos of all AtzB sequences (top) and the 40 closest sequence homologues (bottom). Residue numbers are relative to AtzB from *Pseudomonas sp.* ADP. Residue exchanges selected (N139C, C169G, I170N, S218C, S219Q, I222N, A223S) are indicated by asterisks. (B) Location of all residues to be exchanged (yellow sticks) within the structural model of AtzB. Xanthine (white sticks), representing a substrate analogue, was introduced into the AtzB structure by superposition of yeast guanine deaminase (PDB: 6OHA).

The logic behind the identification of promising residue exchanges was the requirement for close spatial proximity to the presumed active site and strongly conserved alternative residue identities within both the AtzB sequences and the closest AtzB homologues (see Figure 5.3). Spatial parameters were inferred from structural models of AtzB (see subsection 5.2.1). Residue exchanges identified in this way were N139C, C169G, I170N, S218C, S219Q, I222N, and A223S. A223S was chosen instead of A223C, implied by the sequence logos, as serine is present at the equivalent position in the two closest AtzB homologues. All of these residue exchanges were introduced into AtzB in a stepwise manner, with the best guanine deaminase variant of each iteration becoming the basis of the next one (see subsection 5.2.6).
The two closest homologues of AtzB, two amidohydrolases of unknown function from *Haliea sp. SAOS-164* (Accession: WP_135441588.1; AtzB_Hom_Hal) and *Pleomorphomonas oryzae* (Accession: WP_026793093.1; AtzB_Hom_Pleo)

were selected for further characterization. Substrate screening, described in subsection 5.2.5, revealed them to be guanine deaminases with a small hydroxyatrazine hydrolase side activity. It was reasoned that these two enzymes could potentially be converted into efficient hydroxyatrazine hydrolases, prompting the introduction of the inverse residue exchanges found to be relevant for the functional conversion of AtzB (see subsection 5.2.6).

5.2.3 Protein Purification

Required expression plasmids based on pUR22 (see Table 2.2) were generated as described (see subsubsection 2.2.3.3). Almost all enzyme variants were produced in the *E. coli* BW25113 Δ*guaD* strain to avoid contamination with *E. coli* guanine deaminase (GuaD). AtzB_Hom_Hal C209S and AtzB_Hom_Hal Q210S couldn't be produced in this strain but could be purified from BL21 Gold (DE3). Despite multiple attempts, AtzB I170N could not be generated. All other variants were purified to homogeneity by immobilized metal ion affinity chromatography (see subsubsection 2.2.4.2) and preparative size exclusion chromatography (see subsubsection 2.2.4.3). Yields for variants of AtzB ranged from 2.4 mg of protein per liter of culture medium to 12 mg/l. Variants of AtzB_Hom_Pleo gave yields of around 3.0 mg/l while AtzB_Hom_Hal could only be produced in small quantities with yields ranging from 0.1 mg/l to 0.4 mg/l. The purification of wild-type AtzB (see Figure 5.4) exemplifies the overall success and applicability of the purification protocol. The respective purification steps of further relevant enzyme variants are shown in subsection 8.4.2.

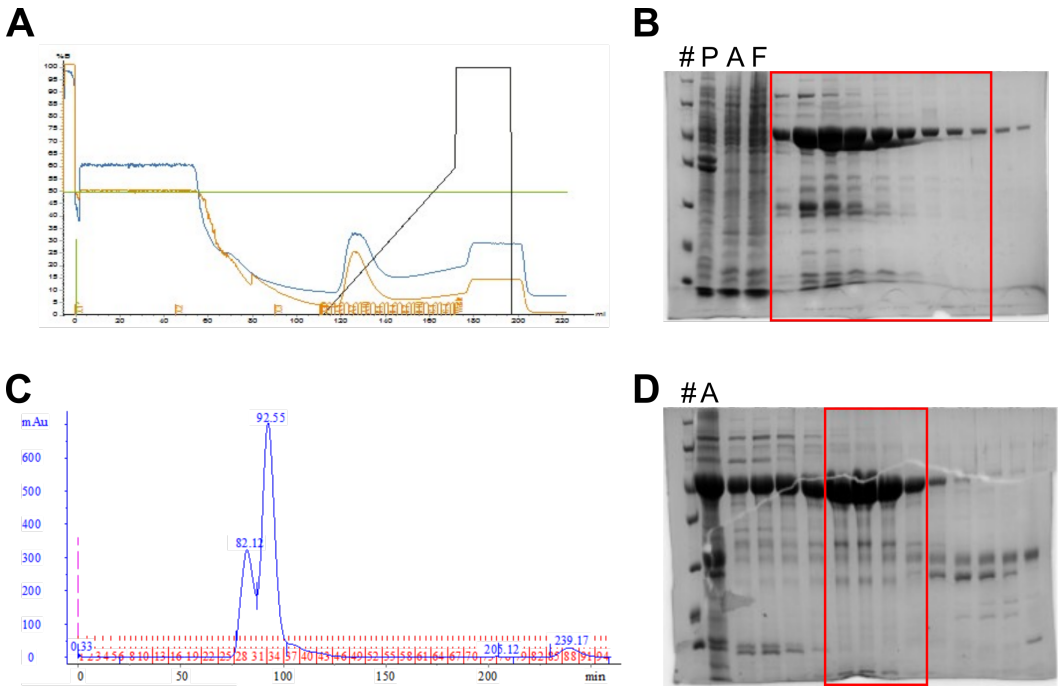


Figure 5.4: Purification of wild-type AtzB
Purification of wild-type AtzB proceeded by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (black). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 53.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) S75 elution profile of AtzB. (D) SDS-PAGE (13.5% acrylamide) of S75 elution fractions: 5 μl each of protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

5.2.4 Circular Dichroism

Circular dichroism (CD) spectra of relevant variants of AtzB were measured as described (see subsubsection 2.2.5.3) to determine effects of the residue exchanges introduced on the secondary structural content of the protein. Generally the spectra are similar, pointing to the overall fold being conserved (see Figure 5.5), even though all variants tested displayed a slightly reduced secondary structure content. This effect is most pronounced in variants S218C and S218C S219Q.

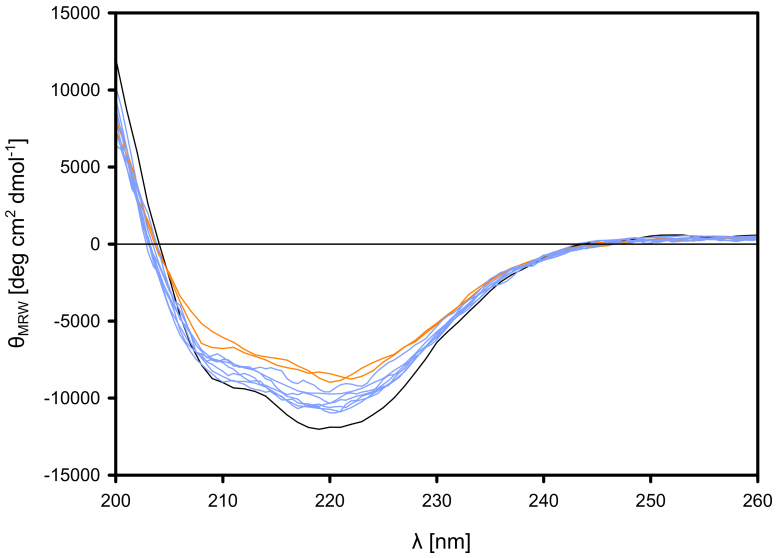


Figure 5.5: Superposition of CD spectra of AtzB variants
Most AtzB variants (blue) show only a slight reduction in secondary structural content compared to the wild type (black). These changes are more pronounced in AtzB S218C and S218C S219Q (orange).

Thermal denaturation curves were measured as described (see subsubsection 2.2.5.4) for select enzyme variants (see subsection 8.4.3) to assay protein stability and the effect of the residue exchanges introduced. Variants of AtzB_Hom_Hal and AtzB_Hom_Pleo could only partially be unfolded at temperatures of up to 90°C and no melting points (T_M) could be determined. T_M values of AtzB variants revealed only minor effects of the residue exchanges introduced (see Table 5.1).

Table 5.1: Melting points of enzyme variants of AtzB
Thermal denaturation curves of select AtzB variants were measured as described (see subsubsection 2.2.5.4) and fitted with a sigmoidal function to determine the respective melting points (T_M; see subsection 8.4.3). Values from signals at λ = 208 nm and λ = 220 nm, correlate to α-helix and β-sheet content, respectively. The residue exchanges introduced do not exert significant effects on protein stability.

Protein	T_M (208 nm) [°C]	T_M (220 nm) [°C]
WT	60.8	63.4
I222N	62.6	64.5
S218C	65.8	66.7
S219Q	61.9	64.6

Protein	T_M (208 nm) [°C]	T_M (220 nm) [°C]
S218C S219Q	62.2	63.0
S218C I170N	57.8	60.6
S218C I222N	65.2	67.4
S218C S219Q I170N	58.6	60.9
S218C S219Q I222N	59.0	60.6
S218C S219Q I170N I222N	61.3	61.5

5.2.5 Substrate Screening of AtzB Homologues

The closest known sequence homologues of AtzB, AtzB_Hom_Hal and AtzB_Hom-Pleo, have not previously been investigated and are annotated simply as amidohydrolases. To identify the enzymatic activities of AtzB_Hom_Hal and AtzB_Hom-Pleo, an HPLC-based substrate screening was performed as described (see subsubsection 2.5.4.1). Both enzymes displayed complete conversion of guanine as well as partial degradation of hydroxyatrazine (see Figure 5.6. No activity of either enzyme for atrazine, cytosine, adenine, isoguanine, 8-oxoguanine, or melamine was observed, while isocytosine, ammeline, and ammelide may be converted to a minimal degree.

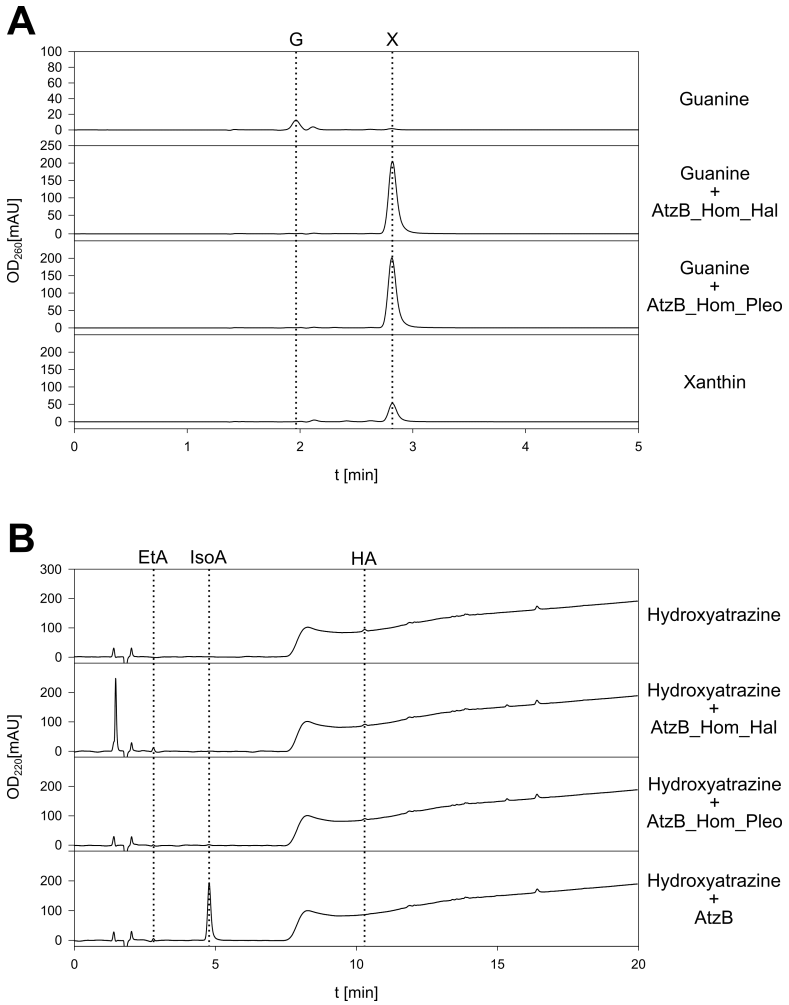


Figure 5.6: Turnover of guanine and hydroxyatrazine by AtzB_Hom_Hal and AtzB_Hom_Pleo
AtzB_Hom_Hal and AtzB_Hom_Pleo were incubated with various substrates as described (see subsubsection 2.5.4.1) and the reaction mixtures were analyzed by HPLC (see subsubsection 2.2.5.5) alongside applicable reference substances. (A) Both enzymes displayed complete conversion of guanine (G) into xanthine (X). (B) Hydroxyatrazine (HA) is partially hydrolyzed to N-ethylammelide (EtA) by AtzB_Hom_Hal and AtzB_Hom_Pleo. As the latter enzyme additionally displays a hydrolytic activity on EtA (not shown) cyanuric acid may be produced which is not photometrically detectable.

5.2.6 Steady State Enzyme Kinetics

The hydroxyatrazine ethylaminohydrolase AtzB was previously shown to possess a small promiscuous guanine deaminase activity [21], which was the foundation of the evolutionary investigations conducted in this work. Furthermore, recent experiments conducted by Lukas Drexler established a promiscuous N^2,N^2-di-

methylguanine dimethylaminohydrolase activity of AtzB. Reactions known to be catalyzed by AtzB are shown in Figure 5.7.

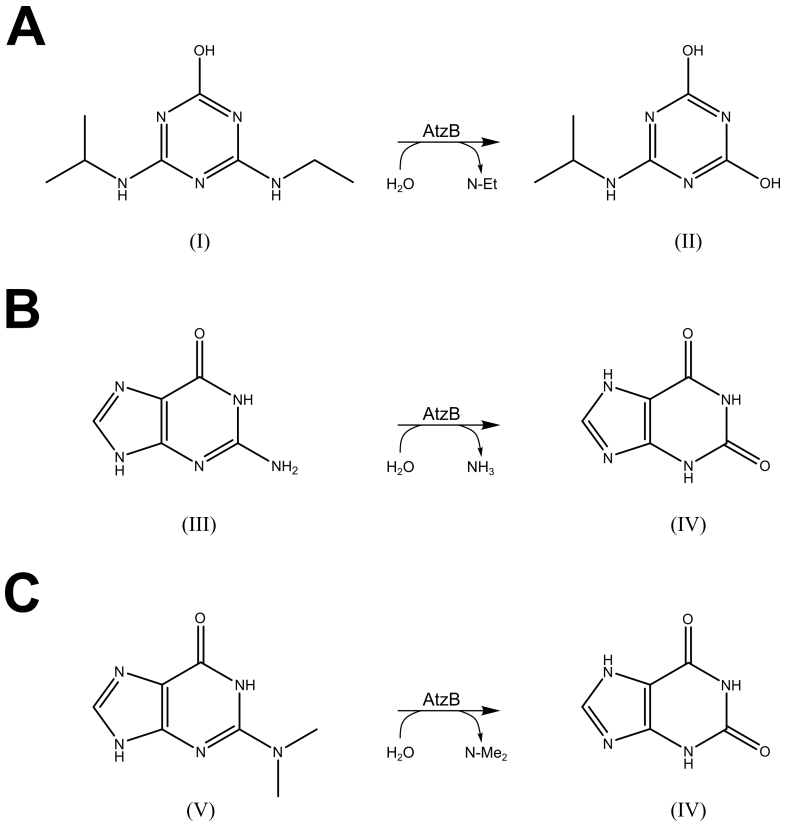


Figure 5.7: Reactions catalyzed by AtzB
(A) AtzB is mainly a hydroxyatrazine ethylaminohydrolase [230], converting hydroxyatrazine (I) into N-isopropylammelide (II) and ethylamine (N-Et). (B) A promiscuous guanine deaminase activity of AtzB has previously been described [21]. Guanine (III) is deaminated by AtzB to give ammonia and xanthine (IV). (C) Recent work conducted by Lukas Drexler has identified N^2,N^2-dimethylguanine dimethylaminohydrolase activity as a further promiscuous activity of AtzB. Dimethylamine (N-Me_2) is liberated from N^2,N^2-dimethylguanine (V) resulting in (IV).

Steady-state enzyme kinetics of all variants of AtzB and its homologues with hydroxyatrazine and guanine were measured in triplicate as described (see subsubsection 2.5.4.2). Initial slopes were plotted against substrate concentrations and the resulting data points were fitted with the Michaelis-Menten equation, yielding the catalytic parameters k_{cat} and K_M from which the catalytic efficiency (k_{cat}/K_M) was calculated. In variants with exceedingly high K_M values, where only the initial slope of the Michaelis-Menten curve could be observed, a linear fit was applied, which solely revealed the catalytic efficiency k_{cat}/K_M. Notably, the enzymatic assay for hydroxyatrazine employed here does not distin-

guish between hydroxyatrazine ethylaminohydrolase and hydroxyatrazine isopropylaminohydrolase activities. Accordingly, the enzymes will be denoted as hydroxyatrazine hydrolases.
Measurements were initially conducted in UV-transparent plates (UV-STAR® 96-well; Greiner Bio-One, Frickenhausen) to screen for variants of AtzB with increased guanine deaminase activity (see Figure 8.95). Kinetic measurements of these relevant variants were then confirmed by cuvette measurements (see subsection 8.4.4). Kinetic values from these experiments are listed in Table 5.2. Despite considering additional residue exchanges (see subsection 5.2.2), AtzB S218C S219Q, identified in previous work in our group [21], was confirmed as the best double mutant based on k_{cat}/K_M values. Introduction of I170N in this context further increased guanine deaminase activity approximately 1.5-fold. The effect of the I170N single mutant could not be assayed as this variant could not be produced. AtzB S218C S219Q I170N could be further improved by introduction of I222N, leading to an approximately 60-fold increase in guanine deaminase activity. Introduction of the remaining residue exchanges failed to improve activity for guanine, indicating AtzB S218C S219Q I170N I222N as the current endpoint of the evolutionary trajectory. Notably, this trajectory from wild-type AtzB to AtzB S218C S219Q I170N I222N proceeds over enzyme variants with continuously decreased hydroxyatrazine hydrolase activity, despite this trait not being selected for. All permutations of these four relevant residue exchanges were then assayed to fully understand this limited sequence space (see Figure 5.11).
The AtzB homologues AtzB_Hom_Hal and AtzB_Hom_Pleo were confirmed as modest guanine deaminases with k_{cat}/K_M values of 1.1e3 M^{-1} s^{-1} and 2.2e3 M^{-1} s^{-1}, respectively. Introduction of the inverse residue exchanges that increased guanine deaminase activity in AtzB into AtzB_Hom_Pleo yielded a quadruple variant (AtzB_Hom_Pleo N165I C213S Q214S N217I) with greatly increased hydroxyatrazine hydrolase activity and nonexistent guanine deaminase activity. Similarly, introduction of the inverse equivalent of S218C S219Q into AtzB_Hom-Hal yielded the double variant AtzB_Hom_Hal C209S Q210S that exhibited increased activity for hydroxyatrazine and lacked activity for guanine.
Recent work conducted by Lukas Drexler has revealed, that AtzB, AtzB_Hom_Hal, AtzB_Hom_Pleo, and their respective variants prefer the rare natural substance N^2,N^2-dimethylguanine over guanine as a substrate. While complete confirmation is still pending, guanine deaminase and N^2,N^2-dimethylguanine dimethylaminohydrolase activity seem to be correlated in the enzyme variants investigated in this work. N^2,N^2-dimethylguanine dimethylaminohydrolase activity may thus represent the true natural activity of AtzB_Hom_Hal, AtzB_Hom_Pleo, and the progenitor of AtzB.

Table 5.2: Catalytic parameters of variants of AtzB, AtzB_Hom_Hal, and AtzB_Hom_Pleo

Steady-state enzyme kinetics for the hydrolysis of guanine (Gua), hydroxyatrazine (HA), and N^2,N^2-dimethylguanine (DMG) were measured in cuvettes (see subsubsection 2.5.4.2). Fitting of the initial conversion rates plotted against substrate concentrations with the Michaelis-Menten equation yielded the catalytic constant (k_{cat}) and the Michaelis constant (K_M), from which the catalytic efficiency (k_{cat}/K_M) could be calculated. In variants with radically increased K_M values, a linear fit had to be applied which revealed only the catalytic efficiency. The standard error is given for measurements conducted in triplicate. Some values could be determined (ND). The lower limit for the detection of enzyme activity was approximately 2 $M^{-1}s^{-1}$.
The k_{cat}/K_M value of AtzB I170N S219Q I222N for hydroxyatrazine is derived from single plate reader measurements. Measurements of AtzB_Hom_Pleo N165I C213S Q214S N217I and DMG kinetics were conducted as single measurements by Lukas Drexler.

Protein	k_{cat} [s^{-1}]	K_M [μM]	k_{cat}/K_M [$M^{-1}s^{-1}$]
Gua			
AtzB	ND	ND	<2.0
S218C	ND	ND	9.8±0.23
S219Q	ND	ND	<2.0
I222N	ND	ND	<2.0
S218C I170N	ND	ND	13±0.66
S218C S219Q	ND	ND	43±1.3
S218C I222N	ND	ND	2.9±0.10
I170N S219Q	ND	ND	<2.0
I170N I222N	ND	ND	7.0±0.34
S219Q I222N	4.4e-3±3.6e-4	93±20	47±14
S218C S219Q I170N	ND	ND	61±2.3
S218C S219Q I222N	ND	ND	<2.0
I170N S218C I222N	ND	ND	2.5e2±8.3
I170N S219Q I222N	ND	ND	45±1.3
S218C S219Q I170N I222N	0.21±8.0e-3	53±6.6	3.9e3±6.4e2
AtzB_Hom_Hal	0.062±2.1e-3	56±6.1	1.1e3±1.6e2
C209S	0.028±0.010	7.4e2±3.5e2	38±32
Q210S	0.094±0.012	6.5e2±1.1e2	1.4e2±44

Protein	k_{cat} [s^{-1}]	K_M [μM]	k_{cat}/K_M [$M^{-1}s^{-1}$]
C209S Q210S	ND	ND	<2.0
AtzB_Hom_Pleo	0.15±0.0027	67±3.6	2.2e3±1.6e2
C213S	8.5e-3±1.2e-3	5.3e2±1.1e2	16±5.6
Q214S	0.017±2.7e-3	5.6e2±1.3e2	30±12
C213S Q214S	0.076±1.8e-3	22±1.4	3.4e3±2.8e2
C213S Q214S N165I N217I	ND	ND	<2.0
HA			
AtzB	5.3±0.66	21±7.8	2.4e5±1.2e5
S218C	2.1±0.13	41±5.3	5.2e4±1.0e4
S219Q	1.7±0.13	44±6.9	3.8e4±9.1e3
I222N	0.53±0.018	35±3.9	1.0e4±1.6e3
S218C I170N	1.9±0.20	1.3e2±19	1.5e4±3.7e3
S218C S219Q	0.60±0.081	34±11	1.8e4±8.1e3
S218C I222N	0.43±0.035	60±8.7	7.1e3±1.6e3
I170N S219Q	0.086±4.1e-3	65±5.4	1.3e3±1.7e2
I170N I222N	0.062±0.022	9.3±1.3	6.7e3±1.2e3
S219Q I222N	0.030±1.7e-3	48±5.2	6.3e2±1.0e2
S218C S219Q I170N	0.13±0.012	27±6.4	4.8e3±1.6e3
S218C S219Q I222N	0.33±0.021	45±5.7	7.4e3±1.4e3
I170N S218C I222N	0.078±2.3e-3	18±1.5	4.4e3±5.2e2
I170N S219Q I222N	ND	ND	11
S218C S219Q I170N I222N	7.9e-3±2.6e-4	27±2.2	2.9e2±33
AtzB_Hom_Hal	ND	ND	<2.0
C209S	ND	ND	<2.0
Q210S	7.7e-3±4.0e-4	16±2.6	4.8e2±1.0e2
C209S Q210S	4.4e-3±2.0e-4	9.9±1.4	4.4e2±81
AtzB_Hom_Pleo	4.9e-3±6.7e-4	1.2e2±23	42±14
C213S	1.0e-3±9.3e-5	15±4.3	69±27
Q214S	0.020±1.1e-3	28±3.9	7.0e2±1.4e2
C213S Q214S	0.076±1.8e-3	22±1.4	3.4e3±2.8e2
N165I C213S Q214S N217I	7.5	25	3.0e5
DMG			
AtzB	4.3	4.6e3	9.4e2

Protein	k_{cat} [s^{-1}]	K_M [μM]	k_{cat}/K_M [$M^{-1}s^{-1}$]
S218C S219Q I170N I222N	12	61	2.0e5
AtzB_Hom_Pleo	14	24	5.7e5

5.2.7 Synthesis of N-Ethylammelide and N-Isopropylammelide

While N-isopropylammelide is the main product of the AtzB-catalyzed hydroxyatrazine ethylaminohydrolase reaction, N-ethylammelide appears as a minor side product [230] (see Figure 5.8). It was reasoned that variants of AtzB and its homologues might have different product ratios due to alternate orientations of the hydroxyatrazine substrate.

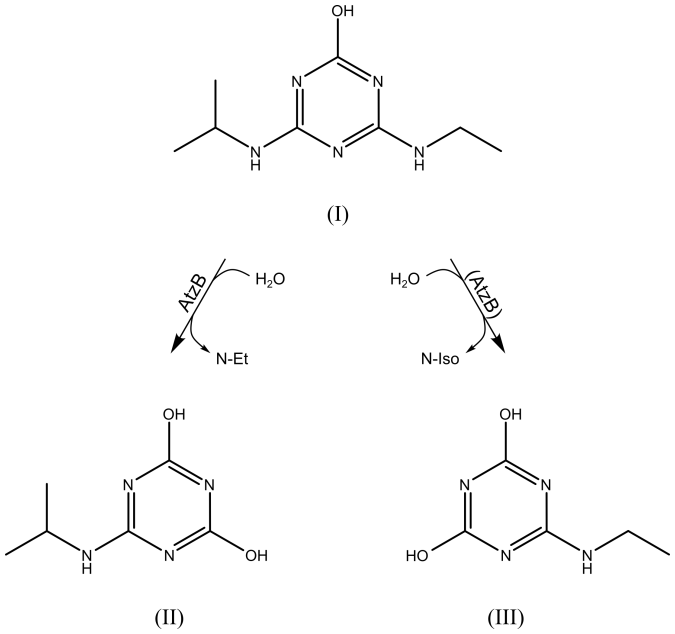


Figure 5.8: Possible reaction products of hydroxyatrazine hydrolysis
The hydrolysis of hydroxyatrazine (I) can result in two different sets of products. Hydrolytic removal of the ethylamine (N-Et) group results in N-isopropylammelide (II), while liberation of the isopropylamine (N-Iso) group produces N-ethylammelide (III). Wild-type AtzB preferentially produces (II) and N-Et [230].

N-ethylammelide and N-isopropylammelide were required for identification of the respective substance in HPLC-based assays (see subsubsection 2.5.4.1) and were thus synthesized. 6-Chloro-1,3,5-triazine-2,4(1H,3H)-dione (CT) was reacted with either ethylamine or isopropylamine (see subsubsection 2.5.2.1).

HPLC analysis of the reaction mixtures (see subsubsection 2.2.5.5) revealed the appearance of new peaks (see Figure 5.9) which were confirmed by mass spectrometry (see Figure 8.107) as N-ethylammelide and N-isopropylammelide, respectively. These peaks further co-eluted with the reaction products of AtzB (see subsection 5.2.8). Impurities are present in the reaction mixture but do not impair the intended purpose of comparative identification of the products of the AtzB reaction and thus no further purification steps were applied.

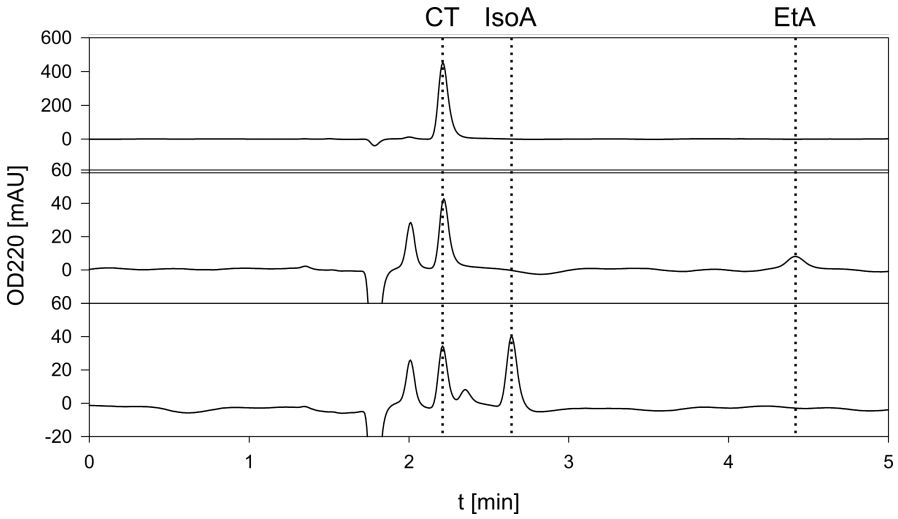


Figure 5.9: HPLC analysis of chemically synthesized N-ethylammelide and N-isopropylammelide
N-ethylammelide (EtA) and N-isopropylammelide (IsoA) were synthesized by reacting 6-Chloro-1,3,5-triazine-2,4(1H,3H)-dione (CT; top) with ethylamine and isopropylamine, respectively (see subsubsection 2.5.2.1). Reaction mixtures of EtA (middle) and IsoA (bottom) were diluted to a theoretical product concentration of 500 μM and analyzed *via* HPLC (see subsubsection 2.2.5.5). Resulting peaks were identified by mass spectrometry and comparison to known substances, indicating that EtA and IsoA have indeed been produced. In both reactions significant amounts of CT remain unreacted and an unknown side product arises which elutes at 2 min.

5.2.8 Product Analysis of Hydroxyatrazine Hydrolases

The variants of AtzB, AtzB_Hom_Hal, and AtzB_Hom_Pleo making up the main trajectories between hydroxyatrazine ethylaminohydrolase activity and guanine deaminase activity were incubated with hydroxyatrazine as described (see subsubsection 2.5.4.1) and the reaction mixtures were analyzed by HPLC (see subsubsection 2.2.5.5). As N-ethylammelide and N-isopropylammelide can not currently be quantified, the resulting peak areas were considered for the estimation of relative product ratios (see Figure 5.10).

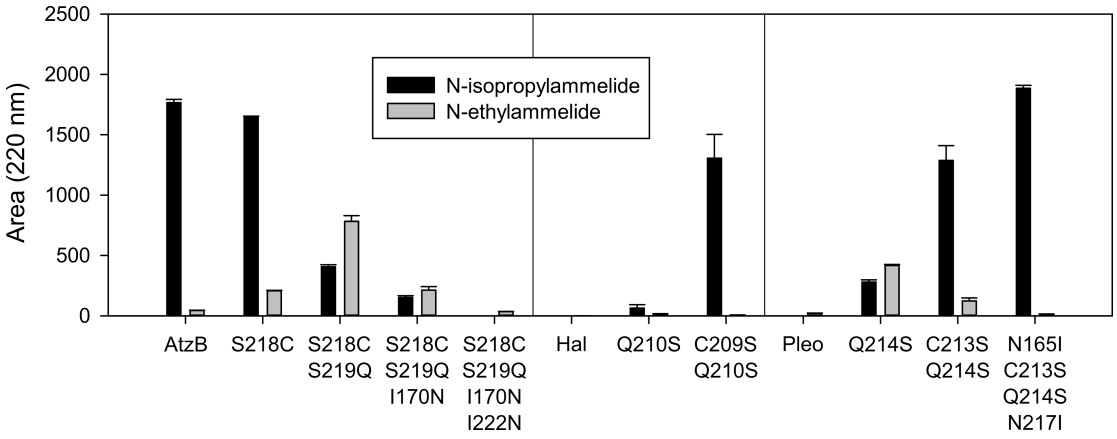


Figure 5.10: Products of hydroxyatrazine hydrolysis of selected protein variants
The enzyme variants leading from wild-type AtzB towards the highest guanine deaminase activity as well as those leading from wild-type AtzB_Hom_Hal (Hal) and AtzB_Hom_Pleo (Pleo) towards the best respective hydroxyatrazine hydrolase variant were investigated for their product composition resulting from conversion of hydroxyatrazine. 2 μM of each enzyme variant was incubated with 500 μM of hydroxyatrazine in 50 mM potassium phosphate buffer (pH 7.0) for 19 h at room temperature. Reaction mixtures were ultrafiltrated to remove protein and analyzed by HPLC (see subsubsection 2.2.5.5). Identities of resulting peaks were determined by comparison to chemically synthesized control substances (see subsection 5.2.7), which were in turn confirmed by mass spectrometry. Peak areas were calculated and plotted to estimate the resulting relative product ratio of each variant. Experiments were conducted in triplicate and the standard error is indicated in the graph.
In all three enzymes, the presence of a serine dyad at positions 218/219, 209/210, and 213/214, respectively, is decisive for a selective hydroxyatrazine ethylaminohydrolase activity.

Wild-type AtzB almost exclusively produces N-isopropylammelide with only a minimal N-ethylammelide peak. This specificity is mostly maintained in AtzB S218C but completely lost in further AtzB variants along the main trajectory. Interestingly, the inverse trend is observed along the respective trajectories from AtzB_Hom_Hal and AtzB_Hom_Pleo towards variants with higher hydroxyatrazine hydrolase activity. High specificity for ethylaminohydrolase activity is established by introduction of C209S Q210S and C213S Q214S into AtzB_Hom_Hal and AtzB_Hom_Pleo, respectively. The specificity is further increased by additional introduction of N165I and N217I into AtzB_Hom_Pleo.

5.3 Discussion

5.3.1 Overcoming Data Limitations

The central challenge of this project was the limited availability of sequence data. Currently, there are only ten known full-length (>300 amino acids) unique amino acid sequences of AtzB enzymes, all of which feature >97% global identity. A multiple sequence alignment (MSA) of the available AtzB sequences revealed, that most of the existing sequence diversity is located at the N-terminus, making a relevance for the catalytic process of the core domain (residues 68-481) unlikely. Considering only the sequence of this core domain, there are only two positions, in which relevant deviations occur. Two AtzB sequences (Accession: AAY40323.2 and BAD69556.1) feature the residue exchanges S218C and S219Q (relative to AtzB from *Pseudomonas sp.* ADP) and one sequence (Accession: WP_192321895.1) carries the mutation S219G. Another AtzB variant from *Leucobacter triazinivorans* (Accession: WP_130110726.1) carries a number of unique deviations, including a three amino acid deletion, but their relevance and validity is unclear. This limited sequence diversity prohibited most bioinformatic analysis methods like correlation analysis or ancestral sequence reconstruction.
The closest AtzB homologue (AtzB_Hom_Hal) is only 64% identical to AtzB from *Pseudomonas sp.* ADP. Considering the generally low rate of nucleotide mutations (approximately 10^{-9} per generation at each position [255] [256]) it is highly unlikely that the last common ancestor of AtzB and AtzB_Hom_Hal existed at the time of introduction of atrazine into the environment, i.e. when a selection pressure for the establishment of AtzB activity first occurred. Accordingly, there may be as of yet undiscovered homologues of AtzB that have diverged only in the last decades and which may be representative of the true AtzB progenitor. The degree of similarity to be expected between AtzB and its true progenitor may be inferred by comparison of atrazine chlorohydrolase (AtzA) and melamine deaminase (TriA) [249] which very recently diverged towards the degradation of two different anthropogenic substances and differ in only 9 out of 474 amino acids. Notably, there are no insertions in AtzB relative to its closest homologues, despite insertions being repeatedly named as a driving force in enzyme diversification [257] [258].
These limitations notwithstanding, the rational comparison of AtzB sequences and their closest homologues (see subsection 5.2.2) was ultimately sufficient to identify residue exchanges that significantly shifted the enzyme activity of AtzB and its two closest homologues (AtzB_Hom_Hal and AtzB_Hom_Pleo) from hydroxyatrazine ethylaminohydrolase to guanine deaminase activity and *vice versa*. This is likely due to the fact that AtzB has not yet significantly diverged from its true progenitor, the active site of which is conserved in the known homologues.

5.3.2 Towards an AtzB Progenitor

In an effort to reconstruct the evolutionary history of the hydroxyatrazine ethylaminohydrolase AtzB, previous work in our group [21] has identified a guanine deaminase side-activity in AtzB and established a set of mutations (S218C S219Q) that increase this activity approximately 10-fold. The present work expands this reverse evolutionary trajectory by step-wise introduction of an expanded set of residue exchanges, which are implied by comparison of AtzB and its closest homologues (see subsection 5.2.2). Kinetic measurements of the resulting activities for hydroxyatrazine and guanine (see subsection 5.2.6) established an evolutionary trajectory consisting of residue exchanges S218C, S219Q, I170N, and I222N (see Figure 5.11). Each successive step within this trajectory is based on the best guanine deaminase variant of the previous iteration, leading to continuously increased guanine deaminase activities, which interestingly correlate with reductions in hydroxyatrazine hydrolase activity, despite this latter trait not being considered in the establishment of the trajectory. The endpoint of the trajectory, AtzB S218C S219Q I170N I222N (AtzB CQNN), cannot further be improved by addition of any of the residue exchanges considered in this work (see Figure 8.95).

AtzB CQNN shares certain traits with the closest AtzB homologues, AtzB_Hom-Hal and AtzB_Hom_Pleo, which were identified as modest guanine deaminases in this work (see subsection 5.2.5). AtzB CQNN exhibits a k_{cat}/K_M value of 3.9e3 $M^{-1}s^{-1}$ for guanine, which is higher than those of AtzB_Hom_Hal and AtzB_Hom_Pleo (1.1e3 $M^{-1}s^{-1}$ and 2.2e3 $M^{-1}s^{-1}$, respectively). The K_M values of AtzB CQNN, AtzB_Hom_Hal, and AtzB_Hom_Pleo for guanine are nearly identical at 53 μM, 56 μM, and 67 μM, respectively. AtzB CQNN further displays a minor hydroxyatrazine hydrolase activity with a k_{cat}/K_M of 292 $M^{-1}s^{-1}$. This activity was also identified in AtzB_Hom_Hal and AtzB_Hom_Pleo although it is significantly weaker in these enzymes (see Figure 5.6 and Table 5.2).

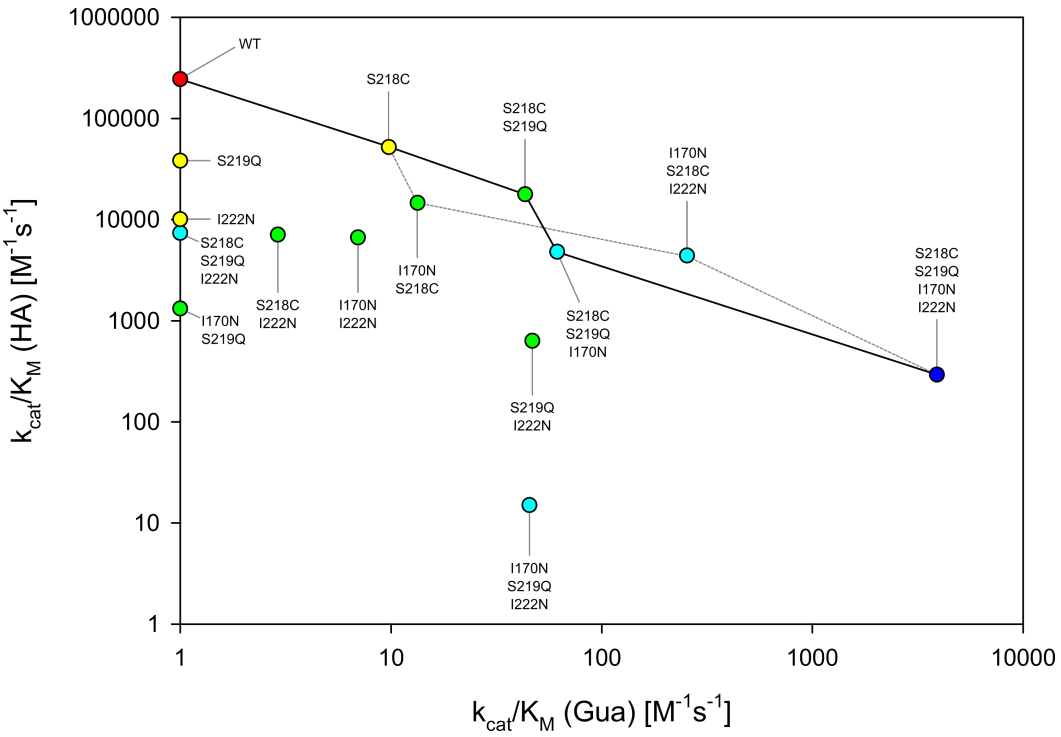


Figure 5.11: Sequence space between AtzB and AtzB CQNN
Catalytic efficiency (k_{cat}/K_M) for hydroxyatrazine (HA) is plotted against k_{cat}/K_M for guanine (Gua) for all AtzB variants contained in the sequence space between wild-type AtzB (WT) and AtzB S218C S219Q I170N I222N (CQNN). Values shown are the average of triplicate measurements in cuvettes (see subsection 5.2.6). AtzB I170N could not be produced. Activities too low to be determined are represented as k_{cat}/K_M = 1 $M^{-1}s^{-1}$. Colors indicate the number of mutations relative to the wild type (0 = red; 1 = yellow; 2 = green; 3 = light blue; 4 = dark blue).
Wild-type AtzB, which is predominantly a hydroxyatrazine ethylaminohydrolase, was converted into a modest guanine deaminase by step-wise introduction of S218C, S219Q, I170N, and I222N (black line). The trajectory was designed for maximum increase in guanine deaminase activity with each step. Notably, hydroxyatrazine hydrolase activity continuously decreased but is never completely abolished along this path. AtzB CQNN represents the current endpoint of this reverse evolutionary trajectory and constitutes a pseudo-progenitor of AtzB due to its physiologically relevant guanine deaminase activity, promiscuous hydroxyatrazine side-activity, and strong evolvability of the latter activity.
The hypothetical evolution of AtzB CQNN into wild-type AtzB can be assumed to have proceeded over bi-functional intermediates, maintaining the physiologically relevant guanine deaminase activity (see Figure 1.1). This forward evolutionary trajectory is likely identical to the previously established reverse one and would consist of the step-wise introduction of N222I, N170I, Q219S, and C218S (relative to AtzB CQNN). An alternative trajectory, with similar activity trade-offs of its intermediates might have proceeded along the successive acquisition of Q219S, N222I, N170I, and C218S (dashed, gray lines).

While it is unclear, if AtzB CQNN ever existed in Nature, its traits make it a plausible progenitor of AtzB in the Innovation-Amplification-Divergence model [9] (see Figure 1.1). Its guanine deaminase activity could potentially have been physiologically relevant before introduction of atrazine into the environment, similar to the presumable roles of AtzB_Hom_Hal and AtzB_Hom_Pleo. The promiscuous activity for hydroxyatrazine might then have constituted an early selection advantage after this substance became available to soil organisms. At this stage, gene duplication and accumulation of mutations increasing hydroxyatrazine hydrolase activity would have been evolutionarily favored, leading to the rapid evolution of AtzB CQNN into the extant AtzB (see Figure 5.11), which is accompanied by an approximately 800-fold increase in catalytic efficiency for hydroxyatrazine. Notably, naturally occurring AtzB variants have been identified (Accession: AAY40323.2 and BAD69556.1) that carry the two amino acid exchanges S218C and S219Q (relative to AtzB from *Pseudomonas sp.* ADP). This could imply that these two exchanges were the most recent to be acquired, partially supporting the hypothesis of AtzB CQNN as a progenitor of AtzB.
The limited number of residue exchanges required for functional conversion of AtzB into a predominant guanine deaminase allowed for the full analysis of the intermediate sequence space. Two viable routes exist connecting the hypothetical progenitor AtzB CQNN to wild-type AtzB across variants with gradually increased hydroxyatrazine hydrolase activity, accompanied by decreases in guanine deaminase activity (see Figure 5.11). Evolution might have proceeded either by the reverse trajectory that first established AtzB CQNN (stepwise introduction of N222I, N170I, Q219S, and C218S relative to AtzB CQNN) or, alternatively, through the successive acquisition of Q219S, N222I, N170I, and C218S. The process of weak negative trade-off between hydroxyatrazine hydrolase and guanine deaminase, observed along either of the two trajectories, has been described by Khersonsky & Tawfik [10] as a principle in the evolution of novel enzyme functions. Weak negative trade-offs have similarly been observed in the interconversion of AtzA and TriA [249] and in the reconstruction of the evolutionary history of phosphotriesterases [259].
Two of the three types of residue exchanges determined to induce functional conversion in AtzB (S ↔ C; I ↔ N) are favored by the genetic code insofar as a single base exchange may be sufficient for them to arise (AGY/TCY ↔ TGY; ATY ↔ AAY). Only the exchange of serine for glutamine (S ↔ Q) requires at least two base exchanges (AGY/TCN ↔ CAR).

5.3.3 Evolutionary Potential of AtzB Homologues

The two closest sequence homologues of AtzB (AtzB_Hom_Hal and AtzB_Hom-Pleo) were investigated in this work for the first time. They were initially shown to be modest guanine deaminases with k_{cat}/K_M values of 1.1e3 $M^{-1}s^{-1}$ and 2.2e3 $M^{-1}s^{-1}$, respectively. Both enzymes exhibit a promiscuous hydroxyatrazine hy-

drolase activity, with a k_{cat}/K_M of 42 $M^{-1}s^{-1}$ in the case of AtzB_Hom_Pleo, while the activity of AtzB_Hom_Hal is too small to be quantified (see Figure 5.6 and Table 5.2). Since functional conversion of AtzB into a guanine deaminase was achieved with only four residue exchanges (S218C S219Q I170N I222N), it seemed plausible that AtzB activity could similarly be enhanced in the AtzB homologues. While the respective sequence space has not yet been fully investigated, the introduction of the inverse of some of the residue exchanges that lead to functional conversion in AtzB have produced variants of the homologues with increased hydroxyatrazine hydrolase activity (see Figure 5.12).
Stepwise introduction of Q214S and C213S into AtzB_Hom_Pleo led to a continuous increase in hydroxyatrazine activity, coupled to a continuous decrease in guanine deaminase activity. The equivalent residue exchanges have a different effect in AtzB_Hom_Hal, where Q210S leads to an increased activity for hydroxyatrazine that was not further increased in the C209S Q210S variant. However, the guanine deaminase activity is continuously decreased by step-wise introduction of the two amino acid exchanges to the point of being completely abolished in AtzB_Hom_Hal C209S Q210S. Amazingly, AtzB_Hom_Pleo N165I C213S Q214S N217I exhibited a k_{cat}/K_M of 3.0e5 $M^{-1}s^{-1}$ for hydroxyatrazine, surpassing even wild-type AtzB, while simultaneously exhibiting no measurable guanine deaminase activity.
Investigating the product composition of the conversion of hydroxyatrazine by the trajectory variants of AtzB and its homologues, allowed for the differentiation between hydroxyatrazine ethylaminohydrolase and hydroxyatrazine isopropylaminohydrolase activity (see Figure 5.10). Interestingly, enzyme variants with higher activity for hydroxyatrazine generally also display a preference for hydrolysis of the ethylamine group, as is the case in wild-type AtzB, which is highly selective for this reaction [230]. In the trajectory from AtzB towards AtzB CQNN this selectivity is lost after introduction of S218C and S219Q. In the incomplete trajectories from the AtzB homologues towards increased hydroxyatrazine hydrolase activity, the introduction of the two inverse exchanges (C209S Q210S and C213S Q214S) is sufficient to introduce strong selectivity for ethylamine hydrolysis, indicating that the identity of these two residues is responsible for determination of substrate orientation.

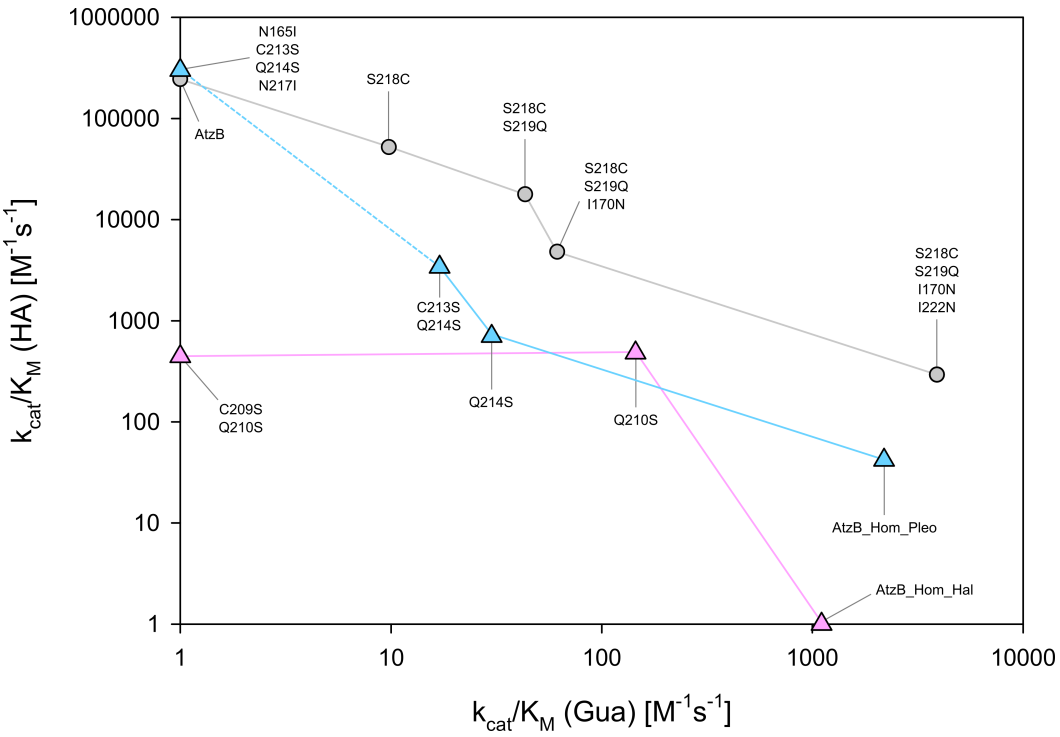


Figure 5.12: Preliminary evolutionary trajectories of AtzB homologues
Catalytic efficiency (k_{cat}/K_M) for hydroxyatrazine (HA) is plotted against k_{cat}/K_M for guanine (Gua) for all relevant variants of AtzB (gray circles), AtzB_Hom_Hal (pink triangles), and AtzB_Hom_Pleo (blue triangles). Activities too low to be determined are represented as $k_{cat}/K_M = 1\ M^{-1}s^{-1}$. Evolutionary trajectories of the different enzymes are indicated by lines of the respective color. The trajectory of AtzB shown here is identical to the one indicated by black lines in Figure 5.11.
Introduction of Q210S and Q214S (inverse to S219Q from AtzB) into AtzB_Hom_Hal and AtzB_Hom_Pleo, respectively, increases the innate hydroxyatrazine hydrolase activity found in each of the two enzymes. The hydroxyatrazine activity of AtzB_Hom_Hal cannot be further increased by introduction of C209S (inverse to S218C from AtzB) but the equivalent exchange in AtzB_Hom_Pleo (C213S) further increases this activity in the Q214S background. The quadruple variant AtzB_Hom_Pleo N165I C213S Q214S N217I even exceeds the catalytic efficiency for hydroxyatrazine of wild-type AtzB.

Despite the relatively low sequence identity of AtzB to AtzB_Hom_Hal and AtzB-Hom_Pleo (64% and 56% global identity, respectively), increased levels of hydroxyatrazine ethylaminohydrolase activity can easily be established in the background of the homologues. Furthermore, even though the substrate hydroxyatrazine is near-symmetrical, the mutations introduced have specifically conveyed hydroxyatrazine ethylaminohydrolase activity, implying that ethylamine and isopropylamine groups are distinguished by the enzyme variants. This result

shows that there exist potential progenitors for alternative AtzB-like enzymes in Nature. Theoretically, any enzyme with a sufficiently similar active site to AtzB_Hom_Hal or AtzB_Hom_Pleo could undergo or could have undergone the same evolutionary process described in this work (see Figure 5.12). There exists precedence for the functionally convergent evolution of atrazine degrading enzymes in that two distinct enzymes (AtzA [226] and TrzN [260]) can catalyze the dechlorination of a number of s-triazine herbicides as the first step of the Atz pathway [217]. However, to the best of the author's knowledge, no such isofunctional enzyme to AtzB has yet been identified. Homologues of AtzB are typically found within operonic structures, often between a glutathione S-transferase and an ABC-type transporter e.g. in the genome of *Pleomorphomonas oryzae* (Accession: NZ_AUHB01000022.1), pointing to a specific physiological function. A single example exists, in which a gene, coding for an AtzB homolog (Accession: QLF71979.1; 53.3% global identity), is located next to a transposase, possibly pointing to an involvement in anthropogen degradation.

5.3.4 Structure-Function Relationships of the Functional Conversion of AtzB and its Homologues

Near total functional conversion of AtzB (see subsection 5.3.2) and AtzB_Hom-Pleo (see subsection 5.3.3) has been achieved by an exchange of an equivalent set of four residues (S218C S219Q I170N I222N and N165I C213S Q214S N217I, respectively). While no crystal structure is available for any of these enzymes, a basic understanding of the specific function and effect of each residue exchange can be gained by inspection of structural models (see subsection 5.2.1). Comparison of models of wild-type AtzB and AtzB CQNN generated with AlphaFold [30] (see Figure 5.13) reveals a set of plausible interactions of the relevant residues with the respectively preferred substrate. The two substrates considered, hydroxyatrazine and guanine, share similar core structures in the s-triazine ring and the pyrimidine portion of the purine ring, respectively. Accordingly, the increase in hydroxyatrazine ethylaminohydrolase and guanine deaminase activities observed in this work is likely mainly determined by a shift in binding affinity for the isopropylamine group and the imidazole portion of the purine ring, respectively.

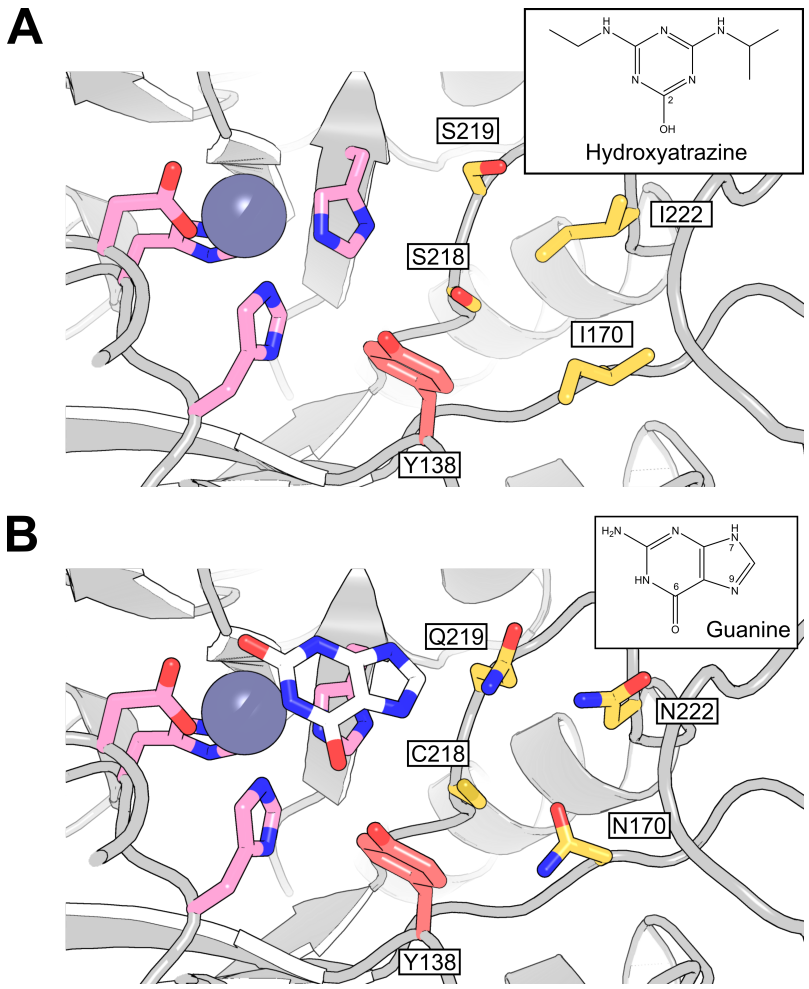


Figure 5.13: Comparison of the active sites of wild-type AtzB and AtzB CQNN

Both wild-type AtzB (A) and AtzB CQNN (B) contain a metal binding site (H74, H76, H245, D331; pink) and a catalytic histidine (H280; not shown) [261] [230]. A zinc ion (gray) is superimposed from the crystal structure of yeast guanine deaminase (PDB: 6OHA; [262]). Y138 (red) likely coordinates O2 of hydroxyatrazine and O6 of guanine, respectively. Residues relevant for substrate specificity are shown in yellow. Structural models were generated with AlphaFold [30] (see subsection 5.2.1).

(A) The location of hydroxyatrazine in AtzB is unknown but may be similar to the position of xanthine in (B). I170 and I222 likely interact with the isopropyl moiety of hydroxyatrazine, while S218 and S219 possibly form H-bonds to the nitrogen atoms of the s-triazine ring or the isopropylamine group. This binding mode would place the ethylamine group towards the catalytic zinc, which is in accordance with the preference of AtzB for ethylaminohydrolysis [229].

(B) In the structural model of AtzB CQNN, xanthine (white sticks) is introduced by superposition with yeast guanine deaminase (PDB: 6OHA; [262]) and represents the likely guanine binding conformation. Guanine binding is likely based on H-bonds to its N7 and N9 atoms, particularly by involvement of Q219.

Enzymes belonging to subtype III of the amidohydrolase superfamily possess a highly conserved divalent metal binding site [261] consisting of three histidines and one aspartate (H74, H76, H245, and D331 in AtzB). A catalytic histidine (H280 in AtzB) further serves to activate a metal bound water molecule, which then initiates the hydrolytic reaction. The observed preference of AtzB for ethylaminohydrolysis [230] requires the ethylamine group of hydroxyatrazine to be positioned towards the metal center. Accordingly, the isopropylamine group of the substrate may interact with the two hydrophobic residues I170 and I222, while S218 and S219 may coordinate the ring nitrogens N3 and N5. However, it was found that the presence of a serine dyad at equivalent positions of 218 and 219 of AtzB was already sufficient for establishing specificity for hydroxyatrazine ethylaminohydrolase activity in the AtzB homologues (see Figure 5.10), implying a further interaction between these residues and the isopropylamine group. Inversely, specificity for ethylaminohydrolysis is lost in AtzB S218C S219Q, despite the presence of the two isoleucines at positions 170 and 222, indicating that these residues alone are not sufficient to cause a specific orientation of the substrate. Y138, which was not considered for mutations as it is equally present in the AtzB homologues, could further coordinate O2 of hydroxyatrazine. In AtzB CQNN, which is predominantly a guanine deaminase, N170, Q219 and N222 may be suited to direct guanine into a catalytically active conformation by interaction with the nitrogens of the imidazole portion of the molecule. Y138 likely forms an H-bond to O6 of guanine, similar to its presumed role in wild-type AtzB. The same substrate binding mechanisms described above likely also apply to the AtzB homologues and their respective variants.
Lately it was revealed, that the enzymes investigated in this work are significantly more active on N^2,N^2-dimethylguanine than they are on guanine (see Table 5.2). AtzB_Hom_Pleo displays a catalytic efficiency of 5.7e5 $M^{-1}s^{-1}$ for N^2,N^2-dimethylguanine compared to only 2.2e3 $M^{-1}s^{-1}$ for guanine and the same holds true for wild-type AtzB (9.4e2 $M^{-1}s^{-1}$ versus <2 $M^{-1}s^{-1}$). Interestingly, activity for N^2,N^2-dimethylguanine is significantly increased in AtzB CQNN (2.0e5 $M^{-1}s^{-1}$) just as was observed for guanine (3.9e3 $M^{-1}s^{-1}$). Accordingly, the previously undescribed activity N^2,N^2-dimethylguanine dimethylaminohydrolase is likely the main activity of the AtzB homologues and the real AtzB progenitor, although the biological relevance of this function is still unclear. A utilization pathway specific for N^2,N^2-dimethylguanine is conceivable, as this substance does occur in Nature in certain contexts [263] [264] [265] [266] [267]. Notably, N^2,N^2-dimethylguanine dimethylaminohydrolase activity does not occur in guanine deaminase from *E. coli* (Lukas Drexler, personal communications).
Mechanistically, this specific activity may have been the basis of the emergence of hydroxyatrazine hydrolase activity, as the presumable affinity for a dimethylamine group near the catalytic metal implies a tolerance of the alkylamine groups of hydroxyatrazine, thus generating the initial promiscuous activity for hydroxyatrazine, that was observed in AtzB CQNN, AtzB_Hom_Hal and AtzB-Hom_Pleo. N^2,N^2-Dimethylguanine dimethylaminohydrolase activity requires

orientation of the dimethylamine group towards the catalytic metal center, placing it away from the residue exchanges that establish the evolutionary trajectories identified in this work (see Figure 5.12). Accordingly, these residues are likely not involved in the discrimination between guanine and N^2,N^2-dimethylguanine and the same substrate binding considerations discussed above for guanine (see Figure 5.13) similarly apply to N^2,N^2-dimethylguanine .

5.4 Outlook

Investigating the evolutionary history of newly emerged enzymes acting on anthropogenic substances is a valuable opportunity to understand enzyme evolution and evolvability. The evolutionarily short time scales since the introduction of the relevant substances have only allowed for minor mutations of the respective progenitors to occur, in many cases allowing for a full reconstruction of the recent evolutionary history of an enzyme. The Atz pathway for the degradation of the s-triazine herbicide atrazine offers tremendous further potential in this regard.

N^2,N^2-Dimethylguanine dimethylaminohydrolase activity, detected in AtzB and AtzB_Hom_Pleo is a new enzymatic activity not previously described and thus merits further investigation. Both AtzB_Hom_Hal and further AtzB homologues may also exhibit this activity and expanded substrate screening may identify even further tolerated substrates. Analysis of the substrate specificity, its structural basis, and genomic context of a set of AtzB homologues may serve to establish a previously unrecognized metabolic pathway for the utilization of specific modified nucleobases. Such modifications are common in Nature [263] and may represent a source of nutrients for specialized microorganisms. N^2,N^2-Dimethylguanine is known to occur in archaea [267] [265] [266] and eukaryota [268] [269] [270].

It may further be interesting to investigate how this novel enzyme activity plays into the evolutionary history of AtzB. As has been hypothesized (see subsection 5.3.4), the tolerance of a dimethylamine group in the active site may have favored the promiscuous activity for hydroxyatrazine, which requires binding of an ethylamine or isopropylamine group in a similar position. It remains to be seen how N^2,N^2-dimethylguanine dimethylaminohydrolase activity changes along the evolutionary trajectories established in this work, although it is likely correlated to guanine deaminase activity as the dimethylamine group can be expected to point away from the residue exchanges constituting the trajectories. Taking into account the rather broad substrate spectrum of the AtzB enzyme, which can liberate a number of alkylamines from s-triazine compounds [230], testing the AtzB variants of the main trajectory with various N^2-modified guanines might give additional insights into the gradual changes of evolving substrate specificity. N^2-alkylated guanines are accessible by a three-step synthesis, starting from 2-amino-6-chloropurine [271].

While AlphaFold [30] has achieved high accuracy of *de novo* protein structure prediction and models of AtzB created by this method have been instrumental in rationalizing the mutant effects studied in this work (see Figure 5.13), crystal structure determination of AtzB and a number of its homologues may nevertheless prove instrumental for the tasks described above. Initial crystallization screens have successfully been conducted for wild-type AtzB (not shown) and refinement of the crystallization conditions followed by structure determination would further close the final knowledge gap regarding the enzymes of the Atz pathway, as crystal structures of AtzA [272], AtzC [233], AtzD [235], AtzE (in complex with AtzG) [223], AtzF [233], and AtzH [224] have previously been identified. Co-crystallization or soaking may provide complex structures of AtzB variants with either of the two substrates hydroxyatrazine and N^2,N^2-dimethylguanine, which might give detailed insight into each single state of the evolutionary process.

Introduction of four residue exchanges (I170N, S218C, S219Q, and I222N) into wild-type AtzB was sufficient to change a catalytic efficiency of 2.4e5 $M^{-1}s^{-1}$ for hydroxyatrazine into a 2.0e5 $M^{-1}s^{-1}$ catalytic efficiency for N^2,N^2-dimethylguanine. While this evolutionary trajectory can be considered mostly complete, the sequence logos used for the identification of presumably substrate defining residues in this work (see Figure 5.3) imply further residue exchanges at a number of second-shell positions close to the ones relevant for functional conversion (M174K and P225R). Additionally, the distant proline exchanges S141P, P146K, and P205A are implied by comparison of the two MSAs. These residue exchanges may be suited to further increase activity for guanine and N^2,N^2-dimethylguanine.

Similar to AtzB, the enzymes AtzA and AtzC can be assumed to have evolved in the last couple of decades. Accordingly, the general process for the reconstruction of recent progenitors applied to AtzB in this work could similarly be repeated for AtzA and AtzC. The same applies to the enzyme TrzN [273], which is isofunctional to AtzA. Other enzymes for the degradation of anthropogenic substances are potential targets for evolutionary reconstruction. Such capacities for degradation have been identified for melamine [187], a number of synthetic chelating agents [274], particularly EDTA [275], haloalkanes [276], chlorophenols [277], and phosphotriesters [278].

A significant impediment in this project was the limited sequence diversity of the known AtzB sequences and the lack of close homologues, a problem which has also been noted for recently evolved phosphotriesterases [259]. Identifying further variants and close homologues of AtzB could be achieved by targeted sequencing of soil samples. Protocols for the purification of DNA from soil have previously been presented [279]. Targeting close homologues of *atzB* with a set of degenerate primers has the potential to identify a diverse set of genes which have only recently diverged from the *atzB* gene. Such primers should not target 5'-terminal sequences of *atzB*, as these locations are generally variable between the different *atzB* genes identified.

It has been noted, that the mode of metal binding and the presence of different metal ion cofactors has an important effect on enzyme activity and substrate specificity and may particularly guide the evolution of novel enzyme functions [280] [281] [258]. Identifying similar effects along the evolutionary trajectories of AtzB and its homologues was outside the scope of this work but may be an important subject of further investigations. While, the residues identified in this work to be relevant to the substrate specificity of AtzB are distal to the metal ion binding site, they may still have an effect on specific metal affinity owing to long-range interactions [280]. Furthermore, as of yet unknown residue exchanges may have accompanied the evolution of extant AtzB by tuning metal affinity into a desired direction. Identifying the metal content of the AtzB variants along the evolutionary trajectory defined in this work, as well as metal exchange studies could help deepen our understanding of the molecular processes accompanying the emergence of hydroxyatrazine ethylaminohydrolase activity.
The high thermal stability observed for the AtzB homologues subsection 5.2.4 together with their evolutionary potential for hydroxyatrazine ethylaminohydrolase activity make them interesting targets for use in bioremediation. Directed evolution could further increase the hydroxyatrazine activity of AtzB_Hom_Pleo N165I C213S Q214S N217I and the resulting construct could be applied to contaminated soils or waters as has previously been done with Atz pathway enzymes [190].

6 Summary

Enzymes are highly specialized and efficient biocatalysts, absolutely essential for all cellular life on earth. Like species as a whole, enzymes are subject to gradual changes, based on spontaneous mutations in their respective genes, which can bring about new structural and functional features in the resulting novel enzyme. Where these features are beneficial to the organism's survival, they present a selection advantage, leading to the retention and amplification of the new enzyme variant. The emergence of new catalytic functions as the result of enzyme evolution is a particularly interesting and complex process the understanding of which can provide valuable new impulses for the field of enzyme design.

6.1 Evolution of Secondary Metabolic Enzymes

Isochorismatase (EntB) and 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase (EntA) are part of the secondary metabolic enterobactin pathway of *Escherichia coli*, contributing to iron uptake of the organism. As it is generally accepted that secondary metabolic enzymes evolved from primary metabolic ones, it was attempted to establish EntB and EntA activity in their closest respective primary metabolic homologues, ureidoacrylate amidohydrolase (RutB) and 3-oxoacyl-[acyl-carrier-protein] reductase (FabG), in an effort to understand the underlying evolutionary processes.
Due to a rational approach previously remaining unsuccessful, a number of strategies was applied to achieve functional conversion. Bioinformatic tools, specifically aimed at functional conversion or diversification, were employed to suggest sets of residue exchanges for RutB and FabG, that might alter the respective substrate specificity. The involvement of the target enzymes in the production of the siderophore enterobactin allowed the establishment of selective growth conditions, based on limiting the bio-availability of iron in growth media by addition of chelating agents. Strains with genomic deletions of *entB-IC* (coding for the isochorismatase domain of EntB) and *entA* did not grow under these conditions but could be rescued by a plasmid-borne copy of the respective gene, establishing a viable system for gene library selection. Fully-randomized gene libraries of *rutB* and *fabG* were created alongside a focused gene library of *rutB*. Transformation into the applicable deletion strain and subjection to the selective growth conditions revealed single gene variants which lead to colony growth on agar plates or overperformance in liquid medium.

However, all efforts eventually proved unsuccessful as none of the isolated or computationally predicted variants of RutB and FabG displayed EntB and EntA activity, respectively. One notable effect observed for several FabG variants was an improved growth of the *entA* deletion strain under iron limiting conditions, which may be due to an unknown, newly acquired function.

6.2 Characterization of Ureidoacrylate Amidohydrolase

The ureidoacrylate amidohydrolase RutB from *E. coli* is part of the Rut pathway for pyrimidine utilization. It was considered as a potential progenitor of the isochorismatase EntB and further was one of only two enzymes of the Rut pathway for which a crystal structure had not been reported, thus meriting detailed investigation. A further motivating factor is the potential relevance of RutB for the pharmaceutical industry due to its (+)-γ-lactamase side activity.
In this work, a protocol for the synthesis of the RutB substrate ureidoacrylate was developed, yielding milligram quantities of the compound. Wild-type RutB was produced in sufficient quantities to conduct crystallization screens and subsequent refinement of crystallization conditions. Crystals of selenomethionine-labeled RutB were produced, allowing for the initial solution of a crystal structure of wild-type RutB at a resolution of 1.9 Å. RutB was further co-crystallized with the substrate analogue ureidopropionate, revealing the mode of substrate binding. Additionally, the inactive RutB variants D24N and C166S were co-crystallized with the substrate ureidoacrylate and structures were solved, revealing the presence of acetate within the active site of RutB C166S.
Various residue exchanges of D24, K133, and C166 produced inactive variants, identifying these three residues as the catalytic triad of RutB. Further residue exchanges were selected based on visual inspection of the active site of RutB and their various effects on the catalytic parameters of ureidoacrylate amidohydrolysis, as determined by a guanine deaminase-coupled assay, allowed for the formulation of a detailed reaction mechanism of RutB.
The detailed insight into the mechanism of RutB, that was gained in this work, could be the foundation of an investigation into the evolutionary history of RutB. A future project for the directed evolution of the (+)-γ-lactamase activity of RutB, based on the data presented in this work, could also be conceived.

6.3 Evolution of Hydroxyatrazine Ethylaminohydrolase

Non-natural substances introduced into the environment through human action, including pesticides, fertilizer, and chemical waste, represent novel evolutionary

factors to natural life. These so called anthropogenic substances have lead to a number of specific adaptations of target and non-target organisms, among them the emergence of novel enzyme functions. A number of microbial species employ the Atz pathway for the utilization of the herbicide atrazine and a number of similar s-triazine compounds, mainly as a nitrogen source. Certain enzymes of the Atz pathway can be assumed to have emerged only after the widespread introduction of atrazine into the environment, starting in the 1950s.

AtzB, the second enzyme in this pathway, is a hydroxyatrazine ethylaminohydrolase, which was previously shown to exhibit a minor guanine deaminase side activity [21]. In this work, the evolutionary history of AtzB was reconstructed by introduction of a set of residue exchanges inferred by comparison of the closest known homologues. Step-wise introduction of these residue exchanges into AtzB produced continuously improved guanine deaminase variants with each iteration. The sequential introduction of S218C, S219Q, I170N, and I222N established an evolutionary trajectory, along which catalytic efficiency (k_{cat}/K_M) for guanine increased from <2 $M^{-1}s^{-1}$ to 3.9e3 $M^{-1}s^{-1}$, while the catalytic efficiency for hydroxyatrazine dropped from 2.4e5 $M^{-1}s^{-1}$ to 2.9e2 $M^{-1}s^{-1}$. This quadruple variant (AtzB CQNN) represents a pseudo-progenitor of AtzB, due to its physiologically relevant guanine deaminase activity, promiscuous hydroxyatrazine hydrolase activity, and its direct evolvability into wild-type AtzB.

The closest known sequence homologues of AtzB, AtzB_Hom_Hal and AtzB_Hom-Pleo, could be shown to be modest guanine deaminases with k_{cat}/K_M values of 1.1e3 $M^{-1}s^{-1}$ and 2.2e3 $M^{-1}s^{-1}$, respectively. Both enzymes exhibited a minor side activity for hydroxyatrazine. Introduction of some of the residue exchanges constituting the evolutionary trajectory established for AtzB produced variants with increased hydroxyatrazine ethylaminohydrolase activities. AtzB_Hom_Hal C209S Q210S possesses a catalytic efficiency of 4.4e2 $M^{-1}s^{-1}$ and AtzB_Hom_Pleo N165I C213S Q214S N217I even surpasses wild-type AtzB with 3.0e5 $M^{-1}s^{-1}$. The AtzB homologues are thus viable evolutionary progenitors of alternative AtzB enzymes.

HPLC-based analysis of the products of hydroxyatrazine hydrolysis revealed that the presence of a serine dyad (S218 S219 in AtzB) is sufficient and required for the specific substrate binding orientation resulting in hydroxyatrazine ethylaminohydrolase activity.

Recent work has further identified a previously undescribed N^2,N^2-dimethylguanine dimethylaminohydrolase activity in AtzB and AtzB_Hom_Pleo. This function may represent the actual enzymatic activity of the AtzB progenitor.

7 Zusammenfassung

Enzyme sind hochspezifische und effiziente Biokatalysatoren, die für jede Form von Leben auf der Erde unerlässlich sind. Ähnlich den Individuen einer Spezies, verändern sich auch Enzyme im Laufe der Evolution. Spontane Mutationen in einem entsprechenden Gen können die strukturellen und funktionellen Eigenschaften eines Enzyms verändern. Sind diese Änderungen für den Organismus vorteilhaft, stellt die neue Enzymvariante einen Selektionsvorteil dar, der schließlich zu ihrer Verbreitung führt.
Die Entstehung neuer Enzymfunktionen als Resultat dieser Enzymevolution ist ein faszinierender und komplexer Vorgang, dessen Erforschung wertvolle neue Impulse für das Feld des Enzymdesigns geben kann.

7.1 Evolution Sekundärmetabolischer Enzyme

Die Enzyme Isochorismatase (EntB) und 2,3-Dihydro-2,3-dihydroxybenzoat-Dehydrogenase (EntA) sind Teil des sekundärmetabolischen Enterobactin-Stoffwechselweges, der in *Escherichia coli* einen Mechanismus zur Eisenaufnahme darstellt. Es ist generell akzeptiert, dass sekundärmetabolische Enzyme aus primärmetabolischen Vorläufern entstanden sind, weshalb in dieser Arbeit versucht wurde, die Aktivitäten von EntB und EntA in ihren nächsten primärmetabolischen Homologen zu etablieren. Diese sind Ureidoacrylat-Amidohydrolase (RutB) und 3-Oxoacyl-[Acyl-Carrier-Protein]-Reduktase (FabG). Diese Funktionsumwandlungen sollten Hinweise auf die der Entstehung von EntB und EntA zugrundeliegenden evolutionären Prozesse geben.
Nachdem ein rationaler Ansatz der funktionellen Umwandlung zuvor erfolglos geblieben war, wurden in dieser Arbeit eine Reihe von Strategien angewandt. Zwei bioinformatische Methoden, die speziell für funktionelle Umwandlung beziehungsweise Diversifizierung entwickelt worden waren, ergaben eine Reihe von Vorschlägen für Residuenaustausche in RutB und FabG, die die Substratspezifität dieser Enzyme verändern könnten. Aufgrund der Beteiligung der Zielenzyme an der Synthese des Siderophors Enterobactin, konnte ein spezifisches Selektionssystem entwickelt werden, indem durch Zugabe geeigneter Chelatoren zum Wachstumsmedium Eisenmangelbedingungen erzeugt wurden. Stämme mit deletiertem *entB-IC* (Bereich des Gens, der für die Isochorismatasedomäne von EntB kodiert) beziehungsweise *entA* Gen konnten unter diesen Bedingungen nicht wachsen, jedoch durch eine plasmidkodierte Kopie des jeweils deletierten Gens gerettet werden. Vollständig randomisierte Genbanken

der Gene *rutB* und *fabG* wurden erzeugt und in den jeweiligen Deletionsstamm transformiert. Unter selektiven Wachstumsbedingungen traten einzelne Genvarianten hervor, die das Wachstum der Deletionsstämme unter Eisenmangelbedingungen wiederherstellten.
Funktionelle Untersuchungen zeigten jedoch, dass keine der bioinformatisch vorgeschlagenen oder aus den Genbanken isolierten Varianten von RutB und FabG die Aktivitäten von EntB beziehungsweise EntA besaßen. Einige der FabG Varianten zeigten jedoch reproduzierbar eine Verbesserung der Wachstumsgeschwindigkeit des *entA* Deletionsstamms unter Eisenmangelbedingungen, was auf eine unbekannte, neue Funktion zurückzuführen sein könnte.

7.2 Charakterisierung des Enzyms Ureidoacrylat-Amidohydrolase

Das Enzym Ureidoacrylat-Amidohydrolase (RutB) aus *E. coli* ist Teil des Rut-Stoffwechselweges zum Abbau von Pyrimidinen. Da es einerseits als ein Vorläufer der Isochorismatase (EntB) angesehen wurde und außerdem eines von nur zwei Enzymen seines Stoffwechselweges war, von denen keine Kristallstruktur bekannt war, wurde eine detaillierte Untersuchung dieses Enzyms unternommen. Ferner existiert eine mögliche Relevanz von RutB für die pharmazeutische Industrie aufgrund seiner (+)-γ-Lactamase Aktivität.
Die Etablierung eines Protokolls für die Synthese von Ureidoacrylat, dem Substrat von RutB, ermöglichte die Produktion dieser Substanz im Milligramm-Maßstab. Wildtypisches RutB konnte in ausreichender Menge hergestellt werden, um Kristallisationsbedingungen für dieses Enzym zu etablieren und zu optimieren. Mit Selenomethionin markiertes RutB wurde kristallisiert und zur initialen Strukturaufklärung verwendet, wodurch erstmalig eine Struktur von RutB, mit einer Auflösung von 1.9 Å, gelöst werden konnte. Weiterhin wurde RutB mit dem Substratanalogon Ureidopropionat kokristallisiert, wodurch der Modus der Substratbindung aufgeklärt werden konnte. Die Kokristallisation der katalytisch inaktiven RutB Varianten D24N und C166S mit dem Substrat Ureidoacrylat und nachfolgende Strukturaufklärung zeigte das Vorhandensein zweier Moleküle Azetat im aktiven Zentrum von RutB C166S.
Zur Bestimmung des Umsatzes von Ureidoacrylat wurde ein an Glutamat-Dehydrogenase gekoppelter Enzymassay etabliert. Verschiedene Austausche der Residuen D24, K133 und C166 führten zu nahezu vollständig inaktiven Enzymvarianten, was diese drei Reste als die katalytische Triade von RutB bestätigte. Weitere Residuenaustausche in der Substratbindetasche von RutB hatten spezifische Effekte auf den k_{cat} und den K_M Wert des Enzyms, durch die der Reaktionsmechanismus von RutB aufgeklärt werden konnte.
Die detaillierte strukturelle und mechanistische Untersuchung von RutB, die in dieser Arbeit präsentiert wird, kann als Grundlage für die Erforschung der evo-

lutionären Vergangenheit von RutB dienen. Ein weiteres Projekt könnte auf die Verstärkung der (+)-γ-Lactamase Aktivität von RutB abzielen.

7.3 Evolution des Enzyms Hydroxyatrazin-Ethylaminohydrolase

Die Freisetzung synthetischer Substanzen in Form von Pestiziden, Düngern und Chemieabfällen hat die Natur vor neue Herausforderungen gestellt. Verschiedene Spezies haben auf die Anwesenheit dieser anthropogenen Substanzen mit spezifischen Anpassungen reagiert, darunter auch die Evolvierung neuer Enzymfunktionen. Einige Mikroorganismen nutzen den Atz-Stoffwechselweg um das Herbizid Atrazin und ähnliche s-Triazin Herbizide abbauen und als Stickstoffquelle nutzen zu können. Es ist anzunehmen, dass die Enzyme dieses Stoffwechselweges sich erst in jüngster Zeit, nach Beginn der großflächigen Ausbringung von Atrazin ab den 1950er Jahren, evolviert haben.
AtzB, das zweite Enzym dieses Stoffwechselweges, ist eine Hydroxyatrazin-Ethylaminohydrolase, in der jüngst eine promiskuitive Guanin-Desaminaseaktivität nachgewiesen werden konnte [21]. In dieser Arbeit wurde die evolutionäre Vergangenheit von AtzB rekonstruiert, indem verschiedene Residuenaustausche, basierend auf den Sequenzen der nächsten Homologen von AtzB, untersucht wurden. Die schrittweise Einführung dieser Austausche in AtzB führte zu Varianten mit sukzessive verbesserten Guanin-Desaminaseaktivitäten. Die aufeinanderfolgende Einführung von S218C, S219Q, I170N und I222N stellt eine reverse evolutionäre Trajektorie dar, entlang derer die katalytische Effizienz (k_{cat}/K_M) für Guanin von <2 $M^{-1}s^{-1}$ auf 3.9e3 $M^{-1}s^{-1}$ steigt, während die katalytische Effizienz für Hydroxyatrazin von 2.4e5 $M^{-1}s^{-1}$ auf 2.9e2 $M^{-1}s^{-1}$ fällt. Die resultierende Vierfachvariante (AtzB CQNN) stellt aufgrund ihrer potentiell physiologisch relevanten Guanin-Desaminaseaktivität, ihrer promiskuitiven Hydroxyatrazin-Hydrolaseaktivität und ihres evolutionären Potentials in Richtung des wildtypischen AtzB, einen Pseudovorläufer von AtzB dar.
Die nächsten bekannten Sequenzhomologen von AtzB (AtzB_Hom_Hal und AtzB_Hom_Pleo) wurden in dieser Arbeit als moderat aktive Guanin-Desaminasen mit k_{cat}/K_M Werten von 1.1e3 $M^{-1}s^{-1}$ beziehungsweise 2.2e3 $M^{-1}s^{-1}$ identifiziert. In beiden Enzymen zeigte sich zudem eine geringe Hydroxyatrazin-Hydrolaseaktivität, die durch Einführung inverser Residuenaustausche der oben beschriebenen evolutionären Trajektorie jeweils verbessert werden konnte. AtzB-Hom_Hal C209S Q210S zeigte eine katalytische Effizienz von 4.4e2 $M^{-1}s^{-1}$ und AtzB_Hom_Pleo N165I C213S Q214S N217I übertraf mit 3.0e5 $M^{-1}s^{-1}$ sogar das wildtypische AtzB Enzym. Die beiden Homologen von AtzB sind somit plausible Ausgangspunkte für die Evolution alternativer AtzB Enzyme.
HPLC-basierte Analysen der Produktzusammensetzung der Hydrolyse von Hydroxyatrazin durch verschiedene Enzymvarianten zeigte, dass das Vorhanden-

sein einer Serin-Dyade (S218 S219 in AtzB) notwendig und hinreichend für die spezifische Orientierung des Substrates ist, die zur Hydroxyatrazin-Ethylaminohydrolyse führt.
Weiterhin konnte jüngst gezeigt werden, dass AtzB und AtzB_Hom_Pleo eine bisher unbekannte N^2,N^2-Dimethylguanin-Dimethylaminohydrolaseaktivität besitzen. Diese ist möglicherweise die eigentliche enzymatische Aktivität des Vorläufers von AtzB.

8 Supplement

8.1 Materials

Table 8.1: Nucleotide primers used in this work
This table lists the entirety of all nucleotide primers used in this work. *Hal* and *Pleo* denote *atzB*_Hom_Hal and *atzB*_Hom_Pleo, respectively

Name	Target	Sequence
EcEntA_fw_NdeI	*entA*	CAGGGCATATGGATTTCAGCGGTAAAAATGTCTG
EcEntA_rw_XhoI	*entA*	CCCTGCTCGAGTGCCCCCAGCGTTG
EcFabG_fw_NdeI	*fabG*	CAGGGCATATGAATTTTGAAGGAAAAATCGCACTG
EcFabG_rw_XhoI	*fabG*	CCCTGCTCGAGGACCATGTACATCCCGC
EcRutB_fw_NdeI	*rutB*	CAGGGCATATGACGACCTTAACCGCTC
EcRutB_rw_XhoI	*rutB*	CCCTGCTCGAGAGCGATATGAGCAAAGGAC
pKD_EntA_Del_5'	pKD3	CCGACCATCGACGCCTGGTGGAAGCTACTCTCCCGC GAGGTGAAATAATGGTGTAGGCTGGAGCTGCTTC
pKD_EntA_Del_3'	pKD3	GGCTCAACGCTGGGGGCATAAGCATGATCTGGAAAC GCCATTTAACGCTCCATATGAATATCCTCCTTA
EntA_InterPrime_5'	*entA*	GCGGTAAAAATGTCTGGGTAACC
EntA_InterPrime_3'	*entA*	CCATCGACCACAATATCCTGTAGGG
pKD_EntA_Del_3'_Corr	pKD3	GAGCGTTAAATGGCGTTTCCAGATCATGCTTATGCCC CCAGCGTTGAGCCCATATGAATATCCTCCTTAG
FabG_pET21a_BsaI_fw	*fabG*	ATTAAAGAGGAGAAATTAAGCGGTCTCACATGAATT TTGAAGGAAAAATCGC
FabG_His6_BsaI_rw	*fabG*	TTTTTTGGTCTCTCTTAGTGGTGGTGGTGGTGGTGC TCGAG
FabG_pExp_BsaI_fw	*fabG*	ATTAAAGAGGAGAAATTAAGCGGTCTCACATG
pKD_EntA_Del_5'PAGE	pKD3	CCGACCATCGACGCCTGGTGGAAGCTACTCTCCCGC GAGGTGAAATAATGGTGTAGGCTGGAGCTGCTTC
pKD_EntA_Del_3'PAGE	pKD3	GAGCGTTAAATGGCGTTTCCAGATCATGCTTATGCCC CCAGCGTTGAGCCCATATGAATATCCTCCTTAG
FabG_ad_pUR28_fo	*fabG*	AAAAAAGGTCTCAGATGAATTTTGAAGGAAAAATCG
FabG_ad_pUR28_re	*fabG*	TTTTTTGGTCTCTGTCAGACCATGTACATCC
EntB_Region_Seq	*entB*	GGTGATGACTGAAGAATTACTGCC
EntA_fw_BsaI_pUR21	*entA*	AAAAAAGGTCTCACATGGATTTCAGCGGTAAAAATGT

Name	Target	Sequence
		CTGGG
EntA_rv_BsaI_pUR21	*entA*	TTTTTTGGTCTCTCTTATGCCCCCAGCGTTGAG
EntB_fw_BsaI_pUR21	*entB*	AAAAAAGGTCTCACATGGCTATTCCAAAATTACAGGC
EntB_rv_BsaI_pUR21	*entB*	TTTTTTGGTCTCTCTTATTTCACCTCGCGGGAG
EcEntB_5'Region_Seq	*entB*	GGTGCGTCGTTTCCTGCGTGAACAGGG
EcRutB_pET21a_fw_BsaI	*rutB*	ATTAAAGAGGAGAAATTAAGCGGTCTCACATGACGAC CTTAACCGCTCGACC
EcEntB_pET21a_fw_BsaI	*rutB*	ATTAAAGAGGAGAAATTAAGCGGTCTCACATGGCTAT TCCAAAATTACAGGC
EcRutB_Deletion_5'	pKD3	ATGCAATGCCGCGCCCATCTACCTGCGCTGACTCAGG AGGTGGCATGATGGTGTAGGCTGGAGCTGCTTC
EcRutB_Deletion_3'	pKD3	AATTACGGATTTTGGCATCGTTAAACTCCTTAAGCGA TATGAGCAAAGGACATATGAATATCCTCCTTAG
RutB_InterPrime_5'	*rutB*	CCTTAACCGCTCGACCGGAAGCC
RutB_InterPrime_3'	*rutB*	GCAAAGGACGTGGGAGAAAGCGC
CyRI	pUR22	TCACGAGGCCCTTTCGTCTT
CyPstI	pUR22	TCGCCAAGCTAGCTTGGATTCT
RutB_D24N_1_2	*rutB*	AATATGCAAAACGCTTATGCC
RutB_D24A_1_2	*rutB*	GCTATGCAAAACGCTTATGCC
RutB_D24A_2_2	*rutB*	CACCACGATCAGCGCAC
RutB_N72A_1_2	*rutB*	GCTGGCTGGGATGAACAGTATGTCG
RutB_N72A_2_2	*rutB*	TTGAAACCAGATGATCAGCATCCCTGC
RutB_K133A_1_2	*rutB*	GCGCCGCGCTACAGCGG
RutB_K133A_2_2	*rutB*	CGGCAGCACAATATCGCCAGG
RutB_Y136A_1_2	*rutB*	GCCAGCGGTTTCTTCAATACGC
RutB_Y136A_2_2	*rutB*	GCGCGGCTTCGGCAG
RutB_C166A_1_2	*rutB*	GCCGTCGAATCGACGCTAC
RutB_C166A_2_2	*rutB*	GACGTTGGTAGCGATGCCG
RutB_C166T_1_2	*rutB*	ACCGTCGAATCGACGCTAC
RutB_C166S_1_2	*rutB*	TCCGTCGAATCGACGCTAC
RutB_C166K_1_2	*rutB*	AAAGTCGAATCGACGCTAC
RutB_OuterPrime_5'	*rutB*	AACATCAATATGGGGACGTTAGTCG
RutB_OuterPrime_3'	*rutB*	GTCATCGGCAAACAGCACG
RutB_pUR_pExp_fw	*rutB*	AAAAAAGGTCTCACATGACGACCTTAACCGCTCG
RutB_pUR_pExp_rv	*rutB*	TTTTTTGGTCTCTCTTAAGCGATATGAGCAAAGGACG
RutB_pUR22_rv	*rutB*	TTTTTTGGTCTCTCGAGAGCGATATGAGCAAAGGACG
RutB_C166A_1_2_NEU	*rutB*	GCCGTCGAATCGACGCTACGCG
RutB_C166S_1_2_NEU	*rutB*	AGCGTCGAATCGACGCTAC
RutB_C166A_1_2_NEO	*rutB*	GCCGTCGAATCGACG
RutB_C166S_1_2_NEO	*rutB*	TCCGTCGAATCGACGC
RutB_C166X_2_2_NEO	*rutB*	GACGTTGGTAGCGATGC
RutB_E80A_1_2	*rutB*	GCAGCTGGCGGAC

Name	Target	Sequence
RutB_E80D_1_2	*rutB*	GATGCTGGCGGACCC
RutB_E80X_2_2	*rutB*	GACATACTGTTCATCCCAGCC
RutB_W74F_1_2	*rutB*	TTCGATGAACAGTATGTCGAG
RutB_W74A_1_2	*rutB*	GCGGATGAACAGTATGTCG
RutB_W74X_2_2	*rutB*	GCCATTTTGAAACCAGATGATC
RutB_Q105A_1_2	*rutB*	GCCGGGAAATTGCTGG
RutB_Q105N_1_2	*rutB*	AACGGGAAATTGCTGGC
RutB_Q105X_2_2	*rutB*	CAGCTGCGGCTGC
RutB_C166X_fw	*rutB*	GTCGAATCGACGCTACGCGACG
RutB_C166A_rv	*rutB*	CGCGACGTTGGTAGCGATGCCG
FabG_pUR22_rv	*fabG*	TTTTTTGGTCTCTCGAGGACCATGTACATCCC
RutB_Y29F_fw	*rutB*	TTTGCCACGCCAGGCGGCTAC
RutB_Y29X_rv	*rutB*	AGCGTTTTGCATATCCACCACGATC
RutB_Y35F_fw	*rutB*	TTCTTAGATCTCGCCGGGTTTGATGTCTC
RutB_Y35X_rv	*rutB*	GCCGCCTGGCGTGGC
RutB_S92A_fw	*rutB*	GCGAACGCCCTGAAAACCATGCG
RutB_S92X_rv	*rutB*	TTTATGAAAATTCGGTGAGCCGGGTCC
T7_prom	pET21a	TAATACGACTCACTATAGGG
T7_term	pET21a	GCTAGTTATTGCTCAGCGG
FlexLoop-Rand1	*rutB*	NNNGCCGCCTGGCGTGGCNNNAGCGTTTTGCATATCC ACCACG
FlexLoop-Rand2	*rutB*	TTAGATCTCGCCGGGNNNGATGTCTCAACCACTCGCC CGGTCATTGC
EcEntB_Deletion_5'	pKD3	ACGTCAGTGGCTGGCGTCACGCGCATCAGCCTGAAGG AGAGAACACGATGGTGTAGGCTGGAGCTGCTTC
EcEntB_Deletion_ArCP	pKD3	AGAACGTCCGGCCACATATTTCAGCGACATCAAATGC CTCCTACGGCTGACATATGAATATCCTCCTTAG
EcEntB_ArCP_pUR_fw	*entB*	AAAAAAGGTCTCACATGTCGCTGAAATATGTGGC
EcEntB_ArCP_pUR_rv	*entB*	TTTTTTGGTCTCTCTTATTTCACCTCGCGGGAGAGTAGC
Priming_Site_1	pKD3	GTGTAGGCTGGAGCTGCTTC
Priming_Site_2	pKD3	CATATGAATATCCTCCTTAG
Priming_Site_2_outward	pKD3	CTAAGGAGGATATTCATATG
EcEntB_IC_pUR_rv	*entB*	TTTTTTGGTCTCTCTTACAAATGCTCGTCACGGCTG
Admetos_ORF2_fw	*fabG*	AAAAAAGGTCTCACATGTTGAATGTGACCGACCCGGC
Fec_Deletion_5'	pKD3	CTCATATTAATATGACTACGTGATAATTAACTTTTGA TGCACTCCGCATGGTGTAGGCTGGAGCTGCTTC
Fec_Deletion_3'	pKD3	CGTGGTTTGGTTCTTACGGCCTGTGCAATCTACCTCA TTAGGCACATCGGCATATGAATATCCTCCTTAG
Outside_FecOperon	*E. coli*	TGGTCAGAAACTGGATTAGC
RutB_S92T_fw	*rutB*	ACCAACGCCCTGAAAACCATGCG
RutB_S92Y_fw	*rutB*	TATAACGCCCTGAAAACCATGCG
EcEntB_IC-long_rv	*rutB*	TTTTTTGGTCTCTCTTACAGTAATTCTTCAGTCATCA

Name	Target	Sequence
		CC
RutB_Y136F_fw	*rutB*	TTTAGCGGTTTCTTCAATACGC
Fec_Deletion_3'_langePS	*rutB*	CGTGGTTTGGTTCTTACGGCCTGTGCAATCTACCTCA
		TTAGGCACATCGGGCCATGGTCCATATGAATATCCTCC
fec_InterPrime_5'	*fec*	CCTGAACGTTATCGCGC
fec_InterPrime_3'	*fec*	GCGCCGGTTTTGGCG
RutB_Y136I_fw	*rutB*	ATTAGCGGTTTCTTCAATACGC
RutB_Y136K_fw	*rutB*	AAAAGCGGTTTCTTCAATACGC
RutB_Y136T_fw	*rutB*	ACCAGCGGTTTCTTCAATACGC
RutB_V165X_fw	*rutB*	TGCGTCGAATCGACGCTACGCG
RutB_V165I_rv	*rutB*	AATGTTGGTAGCGATGCCGGTG
RutB_S92S_fw	*rutB*	AGCAACGCCCTGAAAACCATGCG
RutB_W74Y_fw	*rutB*	TATGATGAACAGTATGTCG
RutB_W74X_in_MIMD_rv	*rutB*	GCCCATTTGAAACCAGATGATCAGC
RutB_N72H_fw	*rutB*	CATGGCTGGGATGAACAGTATGTCG
RutB_N72Y_fw	*rutB*	TATGGCTGGGATGAACAGTATGTCG
RutB_V43L_fw	*rutB*	CTGTCAACCACTCGCCCGGTC
RutB_V43V_fw	*rutB*	GTGTCAACCACTCGCCCGGTC
RutB_V43X_in_MIMD_rv	*rutB*	ATCAAACCCGGCGAGATCTAACATGCC
RutB_F41I_in_MIMD_fw	*rutB*	ATTGATATTTCAACCACTCGCC
RutB_F41Y_in_MIMD_fw	*rutB*	TATGATATTTCAACCACTCGCC
RutB_F41X_in_MIMD_rv	*rutB*	CCCGGCGAGATCTAACATGCCG
RutB_Y35Y_fw	*rutB*	TATTTAGATCTCGCCGGGTTTG
RutB_Y35F_fw_kurz	*rutB*	TTCTTAGATCTCGCCGGGTTTG
RutB_Y35A_fw	*rutB*	GCGTTAGATCTCGCCGGGTTTG
FabG_M188L_L192F_fw	*fabG*	CTGACACGTGCGTTTAGCGATGACCAGC
FabG_M188L_L192M_fw	*fabG*	CTGACACGTGCGATGAGCGATGACCAGC
FabG_L192M_fw	*fabG*	ATGACACGTGCGATGAGCGATGACCAGC
FabG_M188L_T189S_fw	*fabG*	CTGAGCCGTGCGCTGAGCGATGACCAGC
FabG_T189S_L192M_fw	*fabG*	ATGAGCCGTGCGATGAGCGATGACCAGC
FabG_M188L_fw	*fabG*	CTGACACGTGCGCTGAGCGATGACCAGC
FabG_L192F_fw	*fabG*	ATGACACGTGCGTTTAGCGATGACCAGC
FabG_MTL_fw	*fabG*	ATGACACGTGCGCTGAGCGATGACCAGC
FabG_T186S_rv	*fabG*	GTCGCTTTCAATAAAGCCCGGAGC
FabG_F183W_rv	*fabG*	GTCCGTTTCAATCCAGCCCGGAGC
FabG_F183W_T186S_rv	*fabG*	GTCGCTTTCAATCCAGCCCGGAGC
FabG_F183Y_rv	*fabG*	GTCCGTTTCAATATAGCCCGGAGC
FabG_F183_T186_rv	*fabG*	GTCCGTTTCAATAAAGCCCGGAGC
FabG_V140_fw	*fabG*	GTGGTTGGTACCATGGGAAATGGCG
FabG_S138_rv	*fabG*	AGAACCGATAGTGATAATACGACC
FabG_V140L_fw	*fabG*	GTGCTGGGTACCATGGGAAATGGCG
FabG_S138A_rv	*fabG*	CGCACCGATAGTGATAATACGACC

Name	Target	Sequence
FabG_D92_fw	*fabG*	CGTGATAACCTGTTAATGCG
FabG_D92E_fw	*fabG*	CGTGAAAACCTGTTAATGCG
FabG_T90_rv	*fabG*	AGTGATACCGGCATTATTGACC
FabG_T90V_rv	*fabG*	CACGATACCGGCATTATTGACC
FabG_T90S_rv	*fabG*	GCTGATACCGGCATTATTGACC
RutB_F41I_V43V_fw	*rutB*	ATTGATGTGTCAACCACTCGCC
RutB_F41Y_V43L_fw	*rutB*	TATGATCTGTCAACCACTCGCC
FabG_211Q_fw	*fabG*	CAGCCGCAGGAAATCGCCAACG
FabG_211X_rv	*fabG*	GCCGAGGCGACCCGCAG
EcAtzB_pET21a_fw	*atzB*	AAAAAAGGTCTCACATGACCACCACACTGTATACCGG
EcAtzB_pET21a_rv	*atzB*	TTTTTTGGTCTCTCGAGACACGGGGTAACACCTGTAC C
RutB_N72H_W74Y_fw	*rutB*	CATGGCTATGATGAACAGTATGTCG
RutB_S92A_var_rv	*rutB*	GGGCGTTCGCTTTATGAAAATTCGGTGAGCCGGG
RutB_S92S_var_rv	*rutB*	GGGCGTTCGATTTATGAAAATTCGGTGAGCCGGG
RutB_S92X_var_fw	*rutB*	TGAAAACCATGCGTAAGCAGCCGCAGCTGCAGGGG
EcGuaD_fw_pET21a_BsaI	*guaD*	AAAAAAGGTCTCACATGATGTCAGGAGAACAC
EcGuaD_rv_pET21a_BsaI	*guaD*	TTTTTTGGTCTCTCGAGGTTGCGTTCGTACAC
RutB_N88X_rv	*rutB*	GAGCCGGGTCCGCCAGC
EcRutB_Q26H_fw	*rutB*	CATAACGCTTATGCCACGCC
EcRutB_Q26_rv	*rutB*	CATATCCACCACGATCAGCG
Priming Site_1_pKD13	pKD13	ATTCCGGGGATCCGTCGACC
Priming Site_2_pKD13	pKD13	TGTAGGCTGGAGCTGCTTCG
PrimSite_2_Outward	pKD13	CGAAGCAGCTCCAGCCTACA
Prim Site_guaD_del	pKD13	GAAGTACAGTTGGCTGGAGTGCACC
AtzB_N139C_fw	*atzB*	TGCTATAGCCGTCGTGGTGG
AtzB_N139C_rv	*atzB*	ATACTGGTGATCAAATGCGGTGG
AtzB_I170N_fw	*atzB*	AACACCCTGCCGATGGCAG
AtzB_I170N_rv	*atzB*	ACAACCACGACCTGCATGAAAACG
AtzB_S218C_fw	*atzB*	TGCAGTCCGGTTATTGCCTATCCGG
AtzB_S219Q_fw	*atzB*	AGCCAGCCGGTTATTGCCTATCCGG
AtzB_S218C_S219Q_fw	*atzB*	TGCCAGCCGGTTATTGCCTATCCGG
AtzB_S218C-S219Q_rv	*atzB*	CGGTGCAACAACAACACGCTGC
AtzB_I222N_fw	*atzB*	AACGCCTATCCGGAAACCTTTGTTGAAAGC
AtzB_A223S_fw	*atzB*	ATTAGCTATCCGGAAACCTTTGTTGAAAGCGC
AtzB_I222N_A223S_rv	*atzB*	AACCGGACTGCTCGGTGCAAC
AtzB_D331_fw	*atzB*	GCGGGTCATGCAAGCAATGATAGC
AtzB_D331_rev	*atzB*	AACGCCAAAACCAACACGAACACCG
AtzB_E248A_fw	*atzB*	GCGGGTGAAACACCGGCAATGGTTGC
AtzB_E248A_rev	*atzB*	ACCTAAATGGGTATGCAGGCTAACACC
AtzB_SS_rev_mix	*atzB*	CGGTGCTACGACCACACGCTGCATTGCAAACGG
AtzB_SSrv_lang	*atzB*	CGGTGCAACAACAACACGCTGCATTGCAAACGGAC

Name	Target	Sequence
AtzB_CQ_I222N_A223S_rv	*atzB*	AACCGGCTGGCACGGTGCAAC
AtzB_C_I222N_A223S_rv	*atzB*	AACCGGACTGCACGGTGCAAC
AtzB_I222N_A223S_fw	*atzB*	AACAGCTATCCGGAAACCTTTGTTGAAAGC
AtzB_C169G_fw	*atzB*	GGCATTACCCTGCCGATGGC
AtzB_C169G_rv	*atzB*	ACCACGACCTGCATGAAAACGC
AtzB_A175S_fw	*atzB*	AGCGAAGGTAGCACCATTCCG
AtzB_A175S_rv	*atzB*	CATCGGCAGGGTAATACAACCACG
AtzB_A183S_fw	*atzB*	AGCATGCGTGAAAGCACCGATACC
AtzB_A183S_rv	*atzB*	ATCCGGAATGGTGCTACCTTCTGCC
AtzB_V255E_fw	*atzB*	GAAGCACGTTTTGGTGAACGTAGCC
AtzB_V255E_rv	*atzB*	CATTGCCGGTGTTTCACCTTCACC
AtzB_A183S_fw_var	*atzB*	CCGGATAGCATGCGTGAAAGCACCGATACC
AtzB_A183S_rv_var	*atzB*	AATGGTGCTACCTTCTGCCATCGGCAGG
Hal_fw_pET21a_BsaI	*Hal*	AAAAAAGGTCTCACATGAGCACCGTTCTGTTTCG
Hal_rv_pET21a_BsaI	*Hal*	TTTTTTGGTCTCTCGAGGCCCGGTGTACGCAGC
Pleo_fw_pET21a_BsaI	*Pleo*	AAAAAAGGTCTCACATGGGCAACTATCTGCTG
Pleo_rv_pET21a_BsaI	*Pleo*	TTTTTTGGTCTCTCGAGAAAATCAATGGTTGCCAGGG
Hal_S166A_fw	*Hal*	GCCGATGGCGGCAGGTAGCACCATTCCGG
Hal_S166S_fw	*Hal*	GCCGATGAGCGCAGGTAGCACCATTCCGG
Hal_N161I_rv	*Hal*	AGGGTAATACAACCACGACCAACATGCAGACGGG
Hal_N161N_rv	*Hal*	AGGGTATTACAACCACGACCAACATGCAGACGGG
Hal_N213I_S214S_fw	*Hal*	GTTATTAGCCTGCCGGAAACCTTTCCGGAAGCC
Hal_N213N_S214A_fw	*Hal*	GTTAATGCGCTGCCGGAAACCTTTCCGGAAGCC
Hal_N213I_S214A_fw	*Hal*	GTTATTGCGCTGCCGGAAACCTTTCCGGAAGCC
Hal_N213N_S214S_fw	*Hal*	GTTAATAGCCTGCCGGAAACCTTTCCGGAAGCC
Hal_C209S_Q210Q_rv	*Hal*	CGGCTGGCTCGGTGCAACAACAACCTGTGCC
Hal_C209C_Q210S_rv	*Hal*	CGGGCTACACGGTGCAACAACAACCTGTGCC
Hal_C209S_Q210S_rv	*Hal*	CGGGCTGCTCGGTGCAACAACAACCTGTGCC
Hal_C209C_Q210Q_rv	*Hal*	CGGCTGACACGGTGCAACAACAACCTGTGCC
Pleo_S170A_fw	*Pleo*	CCGAAAGCGGAAGGTAGCACCATTCCGG
Pleo_S170S_fw	*Pleo*	CCGAAAAGCGAAGGTAGCACCATTCCGG
Pleo_N165I_rv	*Pleo*	CAGGGTAATACCACCACGACCGGCATGATAACG
Pleo_N165N_rv	*Pleo*	CAGGGTATTACCACCACGACCGGCATGATAACG
Pleo_N217I_S218S_fw	*Pleo*	GTTATTAGCTATCGTGAAACCTTTGTTGAAAGCG
Pleo_N217N_S218A_fw	*Pleo*	GTTAATGCGTATCGTGAAACCTTTGTTGAAAGCG
Pleo_N217I_S218A_fw	*Pleo*	GTTATTGCGTATCGTGAAACCTTTGTTGAAAGCG
Pleo_N217N_S218S_fw	*Pleo*	GTTAATAGCTATCGTGAAACCTTTGTTGAAAGCG
Pleo_C213S_Q214Q_rv	*Pleo*	CGGCTGGCTCGGGCTAATAACAACC
Pleo_C213C_Q214S_rv	*Pleo*	CGGGCTACACGGGCTAATAACAACC
Pleo_C213S_Q214S_rv	*Pleo*	CGGGCTGCTCGGGCTAATAACAACC
Pleo_C213C_Q214Q_rv	*Pleo*	CGGCTGACACGGGCTAATAACAACC
AtzB_C169G_I170N_fw	*atzB*	GGCAACACCCTGCCGATGGC

Name	Target	Sequence
AtzB_QNS_fw	*atzB*	AGCCAGCCGGTTAACAGCTATCCGG
Hal_rv_pUR23_BsaI	*Hal*	TTTTTTGGTCTCTCTTAGCCCGGTGTACGCAGC
Pleo_rv_pUR23_BsaI	*Pleo*	TTTTTTGGTCTCTCTTAAAAATCAATGGTTGCCAGGG
AtzB_Q_I222N_A223S_rv	*atzB*	AACCGGCTGGCTCGGTGCAAC

Table 8.2: Synthetic genes used in this work
This table lists the entirety of all synthetic genes used in this work.

Name	Sequence
rutB FuncLib MIMD	ATGACGACCTTAACCGCTCGACCGGAAGCCATTACCTTCGATCCGCAGCA AAGTGCGCTGATCGTGGTGGATATGCAAAACGCTTATGCCACGCCAGGCG GCATGTTAGATCTCGCCGGGTTTGATATTTCAACCACTCGCCCGGTCATT GCCAACATTCAAACCGCCGTGACCGCAGCGCGAGCGGCAGGGATGCTGAT CATCTGGTTTCAAATGGGCTGGGATGAACAGTATGTCGAGGCTGGCGGAC CCGGCTCACCGAATTTTCATAAAGATAACGCCCTGAAAACCATGCGTAAG CAGCCGCAGCTGCAGGGGAAATTGCTGGCGAAAGGCTCCTGGGATTATCA ACTGGTGGATGAACTGGTGCCGCAGCCTGGCGATATTGTGCTGCCGAAGC CGCGCTACAGCGGTTTCTTCAATACGCCGCTGGACAGCATTTTGCGCAGC CGCGGAATACGCCATCTGGTTTTCACCGGCATCGCTACCAACGTCTGCGT CGAATCGACGCTACGCGACGGCTTTTTTCTGGAGTATTTCGGCGTGGTGC TTGAAGACGCAACTCACCAGGCGGGGCCGAAATTTGCGCAGAAAGCCGCG TTGTTCAATATCGAAACCTTTTTTGGCTGGGTCAGCGACGTCGAAACATT CTGCGACGCGCTTTCTCCCACGTCCTTTGCTCATATCGCTTAA
*fabG*2*entA*16	ATGAATTTTGAAGGAAAAATCGCACTGGTAACCGGTGCAAGCCGCGGAAT TGGCCGCGCAGTGGCTGAAACGCTCGCAGCCCGTGGCGCGAAAGTTATTG GCACTGCGACCAGTGAAAATGGCGCTCAGGCGATCAGTGATTATTTAGGT GCCAACGGCAAAGGTCTGATGTTGGATGTGACCGACCCGGCATCTATCGA ATCTGTTCTGGAAAAAATTCGCGCAGAATTTGGTGAAGTGGATATCCTGG TCAATGCGGCCGGTATCCTGCGTCTGAACCTGTTAATGCGAATGAAAGAT GAAGAGTGGAACGATACCTTTGAAGTGAACGTGTCATCTGTTTTCCGTCT GTCAAAAGCGGTAATGCGCGCTATGATGAAAAAGCGTCATGGTCGTATTA TCACTGTGGGTTCTAACGTTGGTACCATGCCGAATGGCGGTATGGCCGCG TACGCTGCGGCGAAAGCGGGCTTGATCGGCTTCAGTAAATCACTGGCGCG CGAAGTTGCGTCACGCGGTATTACTGTAAACGTTGTTGCTCCGGGCAGCA CCGAAACGGACATGACACGTGCGCTGAGCGATGACCAGCGTGCGGGTATC CTGGCGCAGGTTCCTGCGGGTCGCCTCGGCGGCGCACAGGAAATCGCCAA CGCGGTTGCATTCCTGGCATCCGACGAAGCAGCTTACATCACGGGTGAAA CTTTGCATGTGAACGGCGGGATGTACATGGTCTGA

Name	Sequence
*rutB*2*entB*15	ATGACGACCTTAACCGCTCGACCGGAAGCCATTACCTTCGATCCGCAGCA AAGTGCGCTGATCGTGCATGATATGCAAAACTATTTTGCCACGCCAGGCG GCTACTTAGATCTCGCCGGGTTTGATGTCTCAACCACTCGCCCGGTCATT GCCAACATTCAAACCGCCGTGACCGCAGCGCGAGCGGCAGGGATGCTGAT CATCTGGACCCAAAATGCGCAGGATGAACAGTATGTCGAGGCTGGCGGAC CCGGCTCACCGAATTTTCATAAATCGAACGCCCTGAAAACCATGCGTAAG CAGCCGCAGCTGCAGGGGAAAGGCCTGGCGAAAGGCTCCTGGGATTATCA ACTGGTGGATGAACTGGTGCCGCAGCCTGGCGATATTGTGCTGCCGAAGC CGCGCTACAGCGGTTTCTTCAATACGCCGCTGGACAGCATTTTGCGCAGC CGCGGAATACGCCATCTGGTTTTCACCGGCATCTATACCCATATTGGCGT CGAATCGACGCTACGCGACGGCTTTTTTCTGGAGTATTTCGGCGTGGTGC TTGAAGACGCAACTGCGGATTTTAGCCCGAAATTTGCGCAGAAAGCCGCG TTGTTCAATATCGAAACCTTTTTTGGCTGGGTCAGCGACGTCGAAACATT CTGCGACGCGCTTTCTCCCACGTCCTTTGCTCATATCGCT
atzB	ATGACCACCACACTGTATACCGGTTTTCATCAGCTGGTTACCGGTGATGT TGCAGGCACCGTTCTGAATGGTGTTGATATTCTGGTTCGTGATGGTGAAA TTATTGGTCTGGGTCCTGATCTGCCTCGTACACTGGCACCGATTGGTGTT GGTCAAGAACAGGGTGTTGAAGTTGTTAATTGTCGTGGTCTGACCGCATA TCCGGGTCTGATTAATACCCATCATCATTTTTTTCAGGCCTTTGTGCGTA ATCTGGCACCGCTGGATTGGACCCAGCTGGATGTTCTGGCATGGCTGCGT AAAATCTATCCGGTTTTTGCACTGGTTGATGAGGATTGCATTTATCATAG CACCGTTGTTAGCATGGCCGAACTGATTAAACATGGTTGTACCACCGCAT TTGATCACCAGTATAACTATAGCCGTCGTGGTGGTCCGTTTCTGGTTGAT CGTCAGTTTGATGCAGCAAATCTGCTGGGTCTGCGTTTTCATGCAGGTCG TGGTTGTATTACCCTGCCGATGGCAGAAGGTAGCACCATTCCGGATGCAA TGCGTGAAAGCACCGATACCTTTCTGGCAGATTGTGAACGTCTGGTTAGC CGCTTTCATGATCCGCGTCCGTTTGCAATGCAGCGTGTTGTTGTTGCACC GAGCAGTCCGGTTATTGCCTATCCGGAAACCTTTGTTGAAAGCGCACGTC TGGCACGTCATCTGGGTGTTAGCCTGCATACCCATTTAGGTGAAGGTGAA ACACCGGCAATGGTTGCACGTTTTGGTGAACGTAGCCTGGATTGGTGTGA AAATCGTGGTTTTGTTGGTCCGGATGTTTGGCTGGCACATGGTTGGGAAT TTACCGCAGCAGATATTGCCCGTCTGGCAGCAACCGGCACCGGTGTTGCA CATTGTCCGGCACCTGTTTTTCTGGTGGGTGCAGAAGTTACCGATATTCC TGCAATGGCAGCAGCCGGTGTTCGTGTTGGTTTTGGCGTTGATGGTCATG CAAGCAATGATAGCAGCAATCTGGCAGAATGTATTCGTCTGGCATACCTG CTGCAGTGTCTGAAAGCAAGCGAACGTCAGCATCCGGTTCCGGCACCGTA TGATTTTCTGCGTATGGCAACCCAAGGTGGTGCAGATTGTCTGAATCGTC CGGATCTGGGTGCACTGGCAGTTGGTCGTGCAGCCGATTTTTTTGCAGTG GATCTGAATCGCATTGAATATATTGGTGCAAATCATGATCCTCGTAGCCT GCCTGCAAAAGTTGGTTTTAGCGGTCCGGTTGATATGACCGTTATTAATG GTAAAGTTGTGTGGCGCAATGGTGAATTTCCTGGTCTGGATGAAATGGAA CTGGCACGTGCAGCAGATGGTGTTTTTCGTCGTGTTATTTATGGTGATCC GCTGGTTGCAGCACTGCGTCGTGGTACAGGTGTTACCCCGTGT

Name	Sequence
*atzB*_Hom_Hal	ATGAGCACCGTTCTGTTTCGTAATTTTCGTCAGCTGGTTTGTGCGGGTGC ACCGGGTAGCGTTCTGCGTGATGTTGATCTGTGTGCACGTGATGGTATGA TTACCGCAATTGGTCCGCAGCTGCCGCTGACCGATGTTGATGAAGTTGTT GATTGTGGTGGTCTGACCGCATATCCTGGTCTGGTTAATACCCATCATCA TTTTTTTCAGGCCCTGGTTCGTAATCTGCCAGGTCTGGATTGGACCACAC TGAGCCTGCTGGAATGGCTGGATACCATTTATCCGATTTTTGCACGTCTG GATGAGGATTGTATTTATCATGCAAGCCTGATTAGCCTGGCCGATCTGCT GAAACATGGTTGTACCACCGCATTTGATCACCAGTATAACTTTAATAGCA ATATGGGTAGCCGTGTTGTGGATCGTCAGTTTGAAGCAGCAGCACTGCTG GGTGCCCGTCTGCATGTTGGTCGTGGTTGTAATACCCTGCCGATGAGCGC AGGTAGCACCATTCCGGATGCAATGCTGGAAACCACCGATGCATTTCTGG CCGATTGTGAACGTCTGATTGGTGCATTTCATAATCCGGCACCGGGTGCA ATGGCACAGGTTGTTGTTGCACCGTGTCAGCCGGTTAATAGCCTGCCGGA AACCTTTCCGGAAGCCGCAGCGCTGGCACGTCGTCATGGTGTTCGTCTGC ATACCCATCTGAGCGAAGGTGAAAATGCAGCCATGCTGGATCGTTTTGGT ATGCGTAGCCTGGATTGGTGTGAAAGCGTTGGTTTTGTTGGTCCGGATGT TTGGTTTGCACATGGTTGGGAATTTACCCCTCCGGAAATTGCGCGTCTGG CAGCAACCGGTACAGGTGTTGCACATTGTCCGGCTCCGGTTTTTCTGGTT GGTGCAGAAGTTACCGATCTGCCTGCAATGGTTGCAGCAGATATGACCGT TGGTATGGGTGTTGATGGTCAGGCAAGCAATGATAGCAGCAATCTGGCAG AATGTATGCGTCTGGCCTACCTGCTGCAGTGTCTGAATGCACGCCATAAT CCGCTGCCTGCACCGCCTCCGGAACGTTATCTGCATATGGCAACCGCAGG CGGTGCCGCATGTCTGGGTCGTACCGATATTGGTGAACTGGCAGTTGGTA AAGCAGCAGATTTTTTCTGTGCAGATCTGAATGGCCTGGATTATGCCGGT GCAGATAGCGATCCGCTGAGTCTGCCTGCCAAAGTGGGTTTTGCAGGTCC GGCAGCAATGACCGTGGTTCATGGTCGTGTTGTTTGGCGTGATGGTGAAT TTCCGGGTTTAGATGAAACCCAGCTGCGTAGCGCAGCAGATGCCCTGCTG CGTGAAAAACTGGATGGTCATCTGGCACCGCTGCGTACACCGGGT

Name	Sequence
*atzB*_Hom_Pleo	ATGGGCAACTATCTGCTGAAAAATTGTGCAGCAGTTATGGTTGATGATGG TGCAGGTCTGAATGCACGTCGTAATGTTGATATTCTGACCGATGGTCCGG CAATTAAAGCAATTGAACCGCATCTGGCAGAAACACCGCAGAGCGTTGGT GCCGAAGTTATTGATGCAAGCGGTTGGTTTGTTTATCCTGGTCTGGTTAA TACCCACCACCATTTTTTTCAGACCTTTGTTCGTAATCGTGCAGAACTGG ATTGGACCAAACTGAGCGTTCTGGAATGGCTGGATCGTATTTATCCGATT TTTAGCCAGCTGACCGAAGATTGCTTTTATCATAGCAGCCTGACCGCAAT GGCAGAACTGATTAAACATGGTTGTACCACCGCACTGGATCATCAGTATT GTTTTCCGCGTCATGCAGGTAAATATCTGGTTGATCGTCAGTTTGAAGCA GCAGAACGTCTGGGTATTCGTTATCATGCCGGTCGTGGTGGTAATACCCT GCCGAAAAGCGAAGGTAGCACCATTCCGGATGCAATGCTGGAAACCACCG ATGAATTTCTGGCAGATTGTGAACGTCTGATTGATCGCTATCATGATGCA AGTCCGTTTAGCCTGCGTCAGGTTGTTATTAGCCCGTGTCAGCCGGTTAA TAGCTATCGTGAAACCTTTGTTGAAAGCGTTGCACTGGCACGTGATAAAG GTGTTTTTCTGCATACCCATGTTGGTGAAGGTGAAAGTCCGGTTATGGAA GCACGTCATGGTAAACGTACCGTTGATTATCTGGAAGAAATGGGTTTTGC AGGTCCGGATGTTTTTTATGCACATTGTTGGGAACTGACCCATACCGAAC TGGCAAAACTGGCAGCAAGCGGCACCGGTGTTAGCCATTGTCCGGAACCG GTTTATCTGGTGGGTGCAGAAGTTACCGATATTCCAGCCATGGCAGCACT GGGTGTTCGTGTTGGTCTGGGTTGTGATGGTAGCGCAAGCAATGATAATA GTAATCTGATGCATTGCATTCATAGCGCCTATATGCTGCAGTGTCTGGTT GCAAGCAGCCGTAGCCATCCTGTTCCGGCACCGGCAGAATTTCTGCGTTT TGCAACCACCGGTAGCGCCAGCCTGCTGGGTCGTGCAGATATTGGTCGTC TGGCACCTGGTATGGCAGCAGACCTGTTTGCAATTGATACCCGTCGTATG GATTATGTTGGCACCCGTCATGATCCGCTGAGCCTGCCTGCAAAATTAGG TATTGGTATGGCCACCGATCTGACCATGATTAATGGTCGTATTGTTTGGG CCAATGGTGAATTTCCGGGTATTGATGAAGCAGAAATGGCAGCCGAAGCA GAAGCAACCCTGGCAACCATTGATTTT

8.2 Evolutionary Origin of EntB and EntA

8.2.1 FuncLib

Table 8.3: FuncLib variants selected for analysis
These enzyme variants were selected from a list of FuncLib predictions to be analyzed for possible side activities based on their high stability score.

Enzyme	Residue Exchanges
RutB	Y35M F41Y V43I N72M S92A
RutB	Y35M F41Y V43L N72M S92D

Enzyme	Residue Exchanges
RutB	Y35M F41I V43I N72M Y136T
RutB	Y35M F41I N72M S92D Y136K
RutB	Y35M V43I N72M W74Y S92D
RutB	Y35A V43I N72M S92D Y136K
RutB	Y35A V43L N72M S92D Y136T
RutB	Y35M V43I N72M S92A Y136I
RutB	Y35M V43I N72M S92D Y136F
RutB	Y35M V43L N72M S92D Y136I
RutB	Y35A V43I N72M S92D V165I
RutB	Y35M V43I N72Y S92D V165I
RutB	Y35M V43I N72M Y136T V165I
RutB	Y35M N72M W74Y S92D Y136F
RutB	Y35F N72M W74Y S92D Y136T
RutB	V43I N72H W74Y S92D V165I
RutB	V43I N72M S92D Y136K V165I
FabG	T90V D92E S138A V140L F183Y
FabG	T90V D92E S138A F183Y M188L
FabG	T90V D92E V140L
FabG	T90V D92E M188L T189S
FabG	T90S S138A V140L
FabG	T90V S138A F183W L192M
FabG	T90V S138A M188L
FabG	T90S S138A M188L L192F
FabG	T90V S138A T189S L192M
FabG	T90V V140L M188L L192M
FabG	D92E S138A V140L L192M
FabG	D92E V140L F183W
FabG	D92E V140L M188L
FabG	S138A V140L F183W T186S M188L
FabG	S138A V140L M188L
FabG	S138A F183Y M188L L192F
FabG	S138A F183Y L192M
FabG	V140L T186S M188L L192M
FabG	F183W T186S L192F
FabG	T186S M188L L192M V140L

8.2.2 Protein Purification

The bands of the protein standard (LMW, Thermo Fisher Scientific) used in the following SDS-PAGEs have the following molecular weights: 116.0, 66.2, 45.0,

35.0, 25.0, 18.4, 14.4 kDa. Preparative size exclusion chromatography runs were followed at 280 nm.

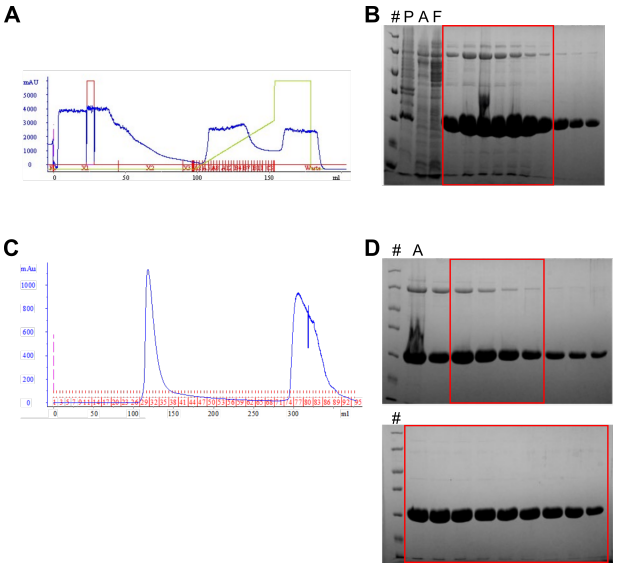


Figure 8.1: Purification of RutB Q26H
RutB Q26H was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB Q26H appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB Q26H. (D) Top: SDS-PAGE (13.5% acrylamide) of the first SEC peak: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Bottom: SDS-PAGE (13.5% acrylamide) of the second SEC peak: 5 μl protein standard (#), 10 μl of elution fractions. Pure fractions (red rectangles) were collected in two separate pools, concentrated, and frozen in liquid nitrogen.

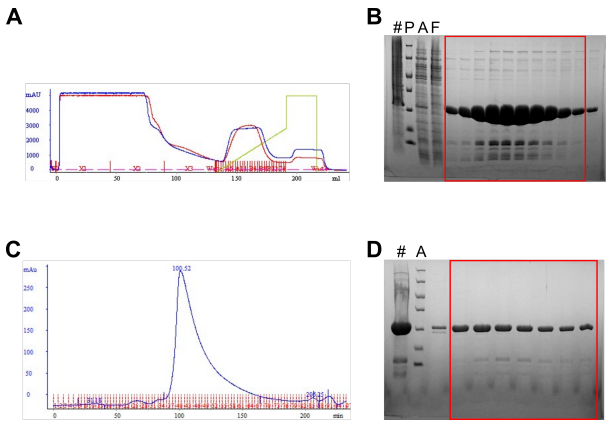


Figure 8.2: Purification of RutB2EntB15
RutB2EntB15 was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB2EntB15 appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 25.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB2EntB15. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

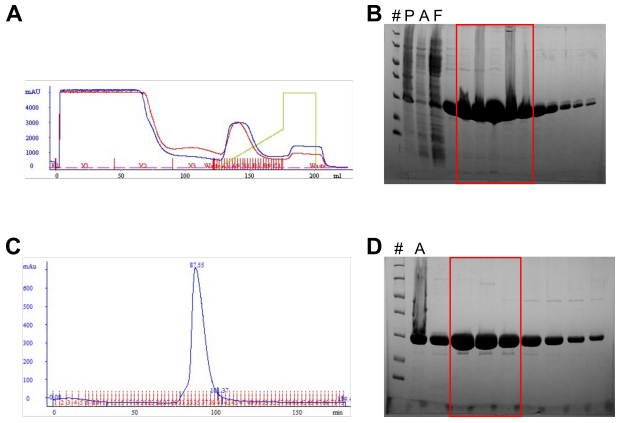


Figure 8.3: Purification of FabG2EntA16
FabG2EntA16 was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. FabG2EntA16 appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.4 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of FabG2EntA16. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

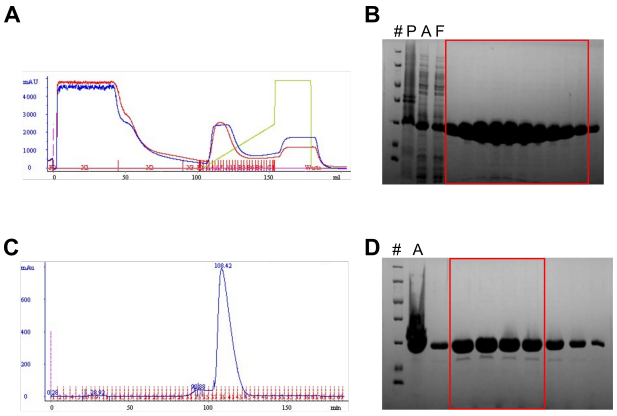


Figure 8.4: Purification of FabG S138A V140L F183W T186S M188L
FabG S138A V140L F183W T186S M188L was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. FabG S138A V140L F183W T186S M188L appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.4 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of FabG S138A V140L F183W T186S M188L. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

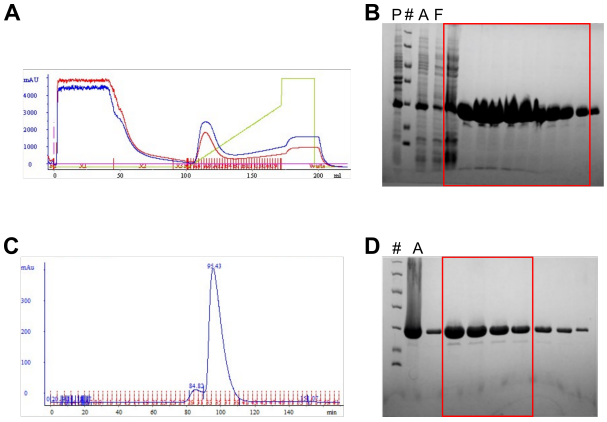


Figure 8.5: Purification of FabG T90S S138A V140L
FabG T90S S138A V140L was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl pellet (P), protein standard (#), crude extract (A), flowthrough (F), and 10 μl of elution fractions. FabG T90S S138A V140L appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.4 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of FabG T90S S138A V140L. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

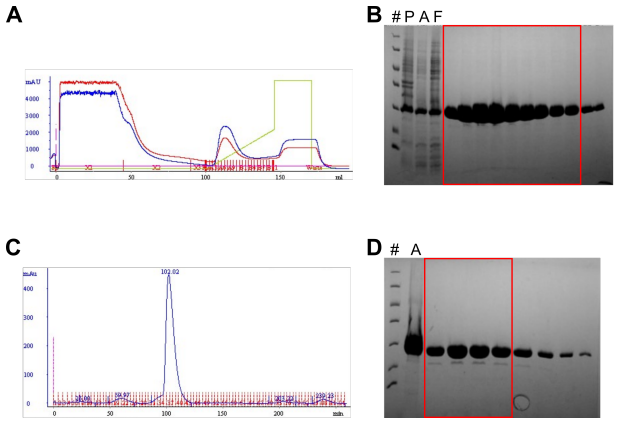


Figure 8.6: Purification of FabG T90V D92E S138A F183Y M188L
FabG T90V D92E S138A F183Y M188L was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. FabG T90V D92E S138A F183Y M188L appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.4 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of FabG T90V D92E S138A F183Y M188L. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

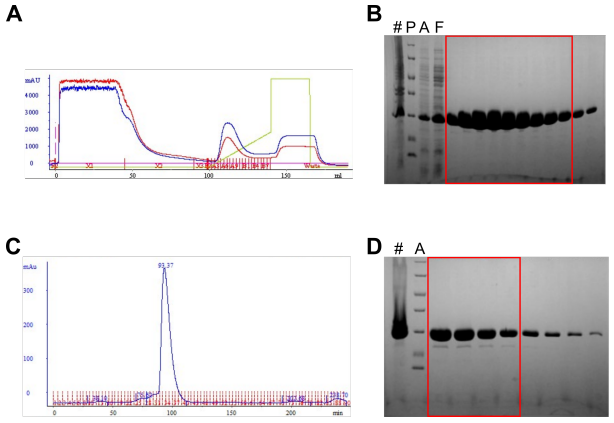


Figure 8.7: Purification of FabG T90V D92E S138A V140L F183Y
FabG T90V D92E S138A V140L F183Y was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. FabG T90V D92E S138A V140L F183Y appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.4 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of FabG T90V D92E S138A V140L F183Y. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

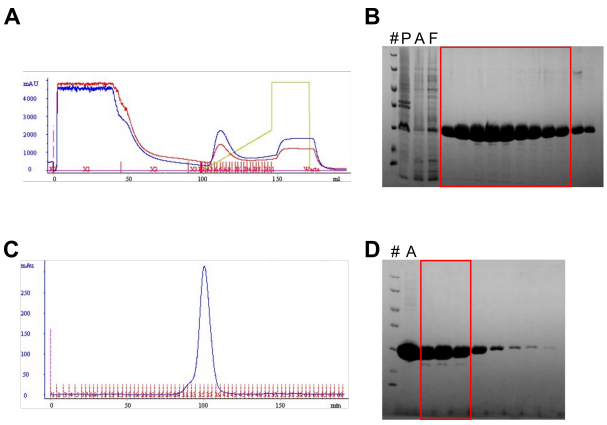


Figure 8.8: Purification of FabG T90V S138A M188L
FabG T90V S138A M188L, was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. FabG T90V S138A M188L appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.4 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of FabG T90V S138A M188L. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

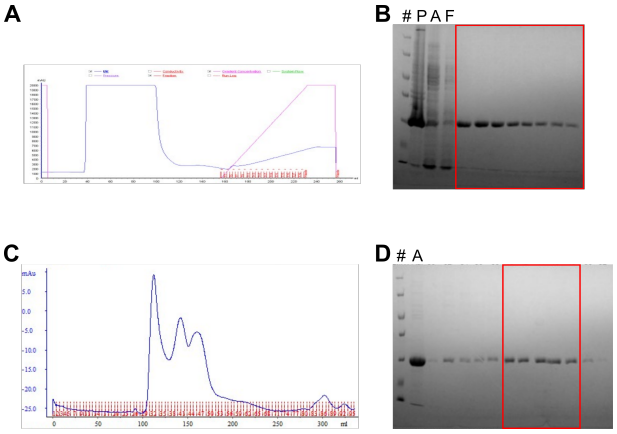


Figure 8.9: Purification of FabG A1
FabG A1 (see Table 3.3), was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), volume% of 1 M imidazole (pink). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. FabG A1 appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.6 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of FabG A1. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

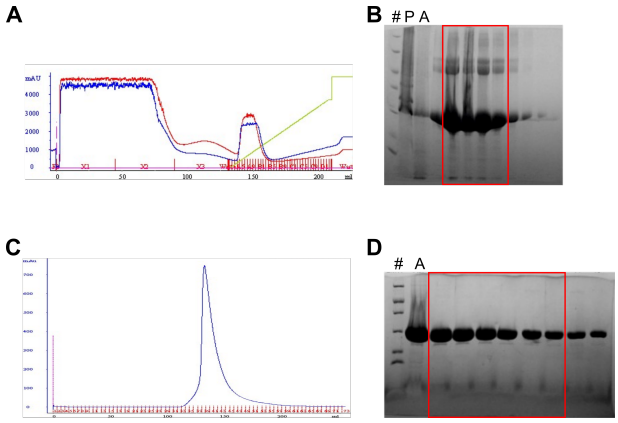


Figure 8.10: Purification of FabG A2
FabG A2 (see Table 3.3), was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), and 10 μl of elution fractions. FabG A2 appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.4 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of FabG A2. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

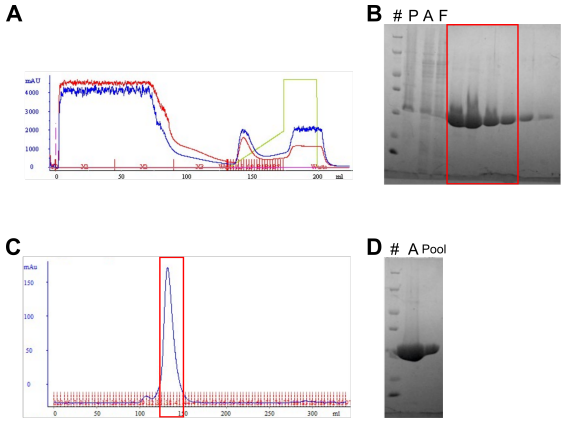


Figure 8.11: Purification of FabG A5
FabG A5 (see Table 3.3), was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. FabG A5 appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.4 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of FabG A5. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen. (D) SDS-PAGE (13.5% acrylamide) of SEC purification: 5 μl protein standard (#), IMAC pool (A), and SEC pool (Pool).

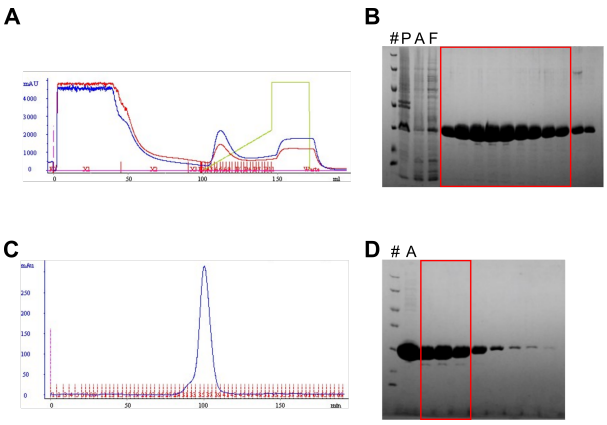


Figure 8.12: Purification of FabG A6
FabG A6 (see Table 3.3), was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), volume% of 1 M imidazole (pink). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. FabG A6 appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.4 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of FabG A6. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

8.3 Characterization of the Ureidoacrylate Amidohydrolase RutB

8.3.1 Protein Purification

The bands of the protein standard (LMW, Thermo Fisher Scientific) used in the following SDS-PAGEs have the following molecular weights: 116.0, 66.2, 45.0, 35.0, 25.0, 18.4, 14.4 kDa. Preparative size exclusion chromatography runs were followed at 280 nm.

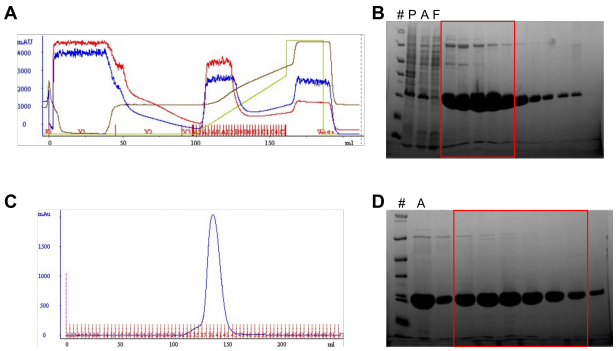


Figure 8.13: Purification of RutB
RutB was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

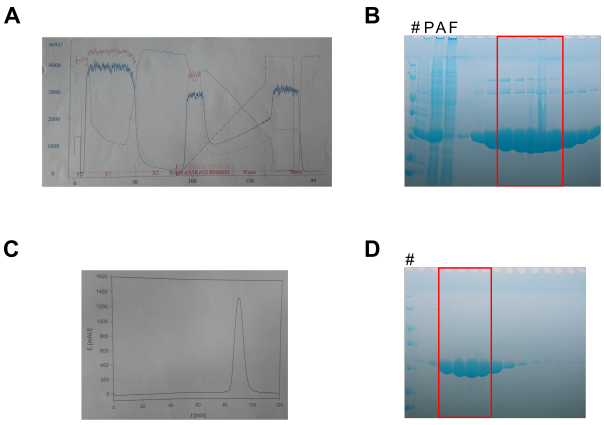


Figure 8.14: Purification of SeMet RutB
Selenomethionine-labeled RutB (SeMet RutB) was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (dashed green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. SeMet RutB appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of SeMet RutB. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

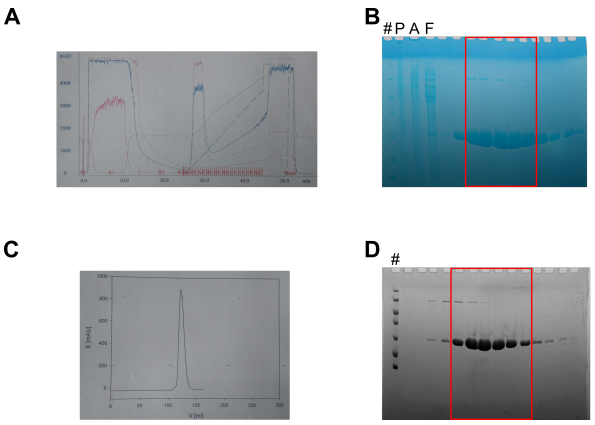


Figure 8.15: Purification of RutB D24A
RutB D24A was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (dashed green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB D24A appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB D24A. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

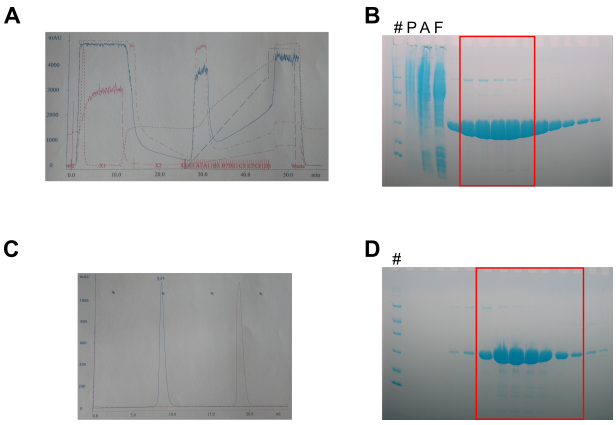


Figure 8.16: Purification of RutB D24N
RutB D24N was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (dashed green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB D24N appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB D24N. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

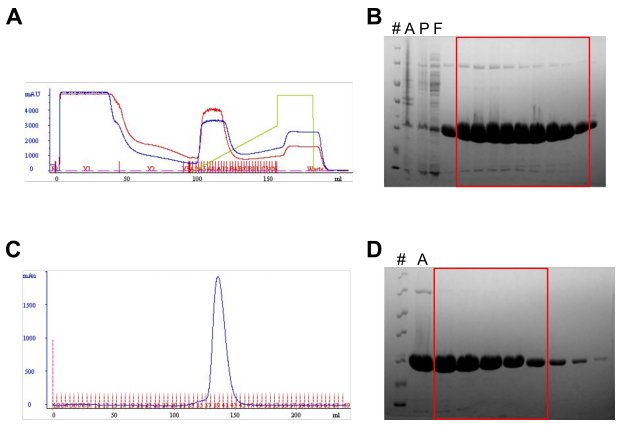


Figure 8.17: Purification of RutB Y29F
RutB Y29F was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), crude extract (A), pellet (P), flowthrough (F), and 10 μl of elution fractions. RutB Y29F appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB Y29F. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

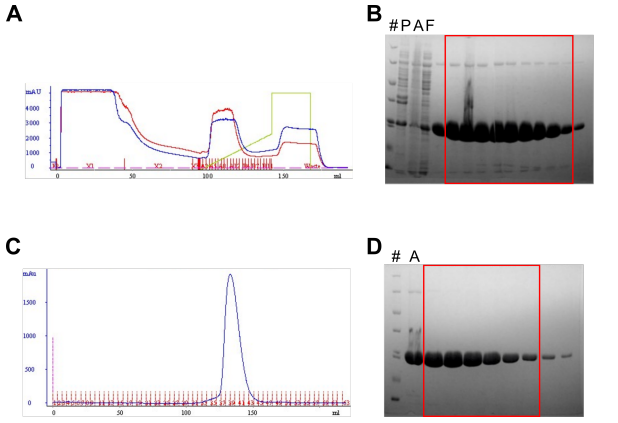


Figure 8.18: Purification of RutB Y35F
RutB Y35F was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB Y35F appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB Y35F. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

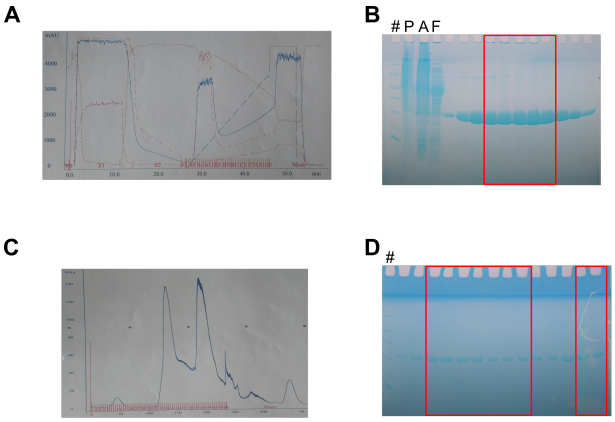


Figure 8.19: Purification of RutB N72A
RutB N72A was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (dashed green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB N72A appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB N72A. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

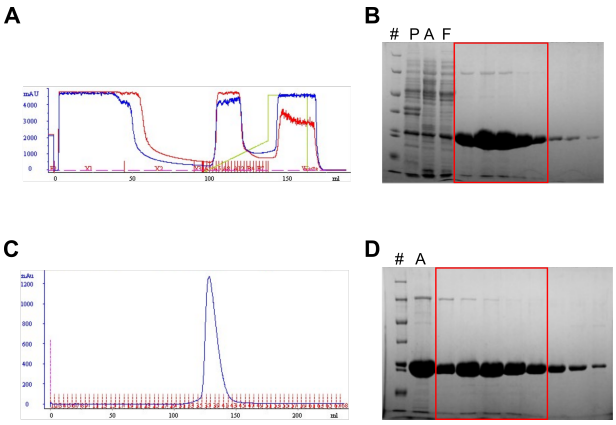


Figure 8.20: Purification of RutB W74A
RutB W74A was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB W74A appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB W74A. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

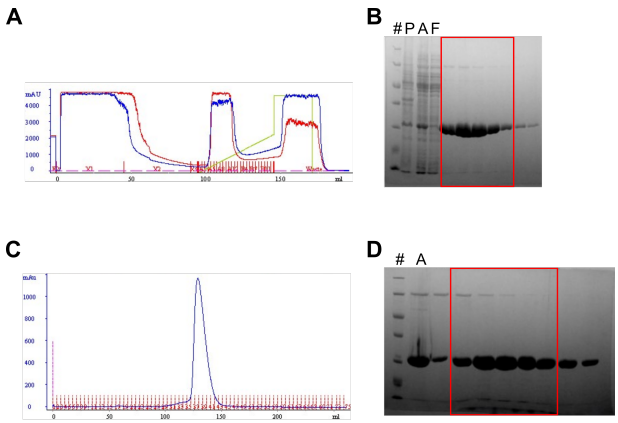


Figure 8.21: Purification of RutB W74F
RutB W74F was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB W74F appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB W74F. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

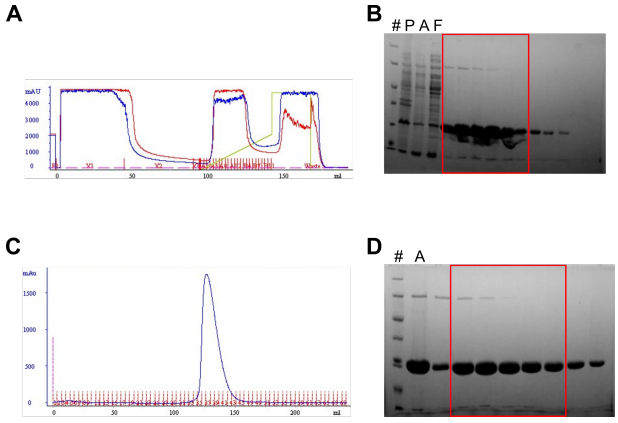


Figure 8.22: Purification of RutB E80A
RutB E80A was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB E80A appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB E80A. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

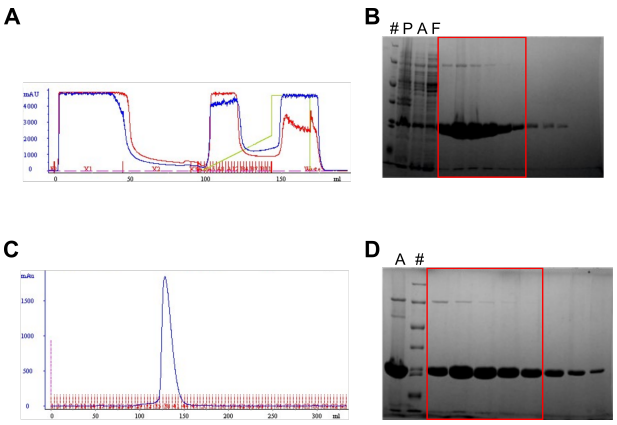


Figure 8.23: Purification of RutB E80D
RutB E80D was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB E80D appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB E80D. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl IMAC pool (A) and protein standard (#), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

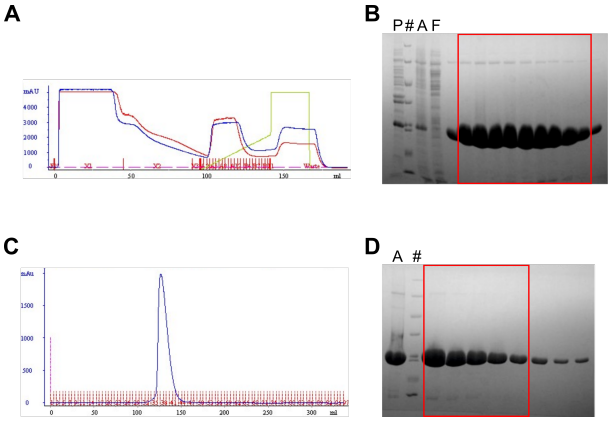


Figure 8.24: Purification of RutB S92A
RutB S92A was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl pellet (P), protein standard (#), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB S92A appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB S92A. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl IMAC pool (A) and protein standard (#), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

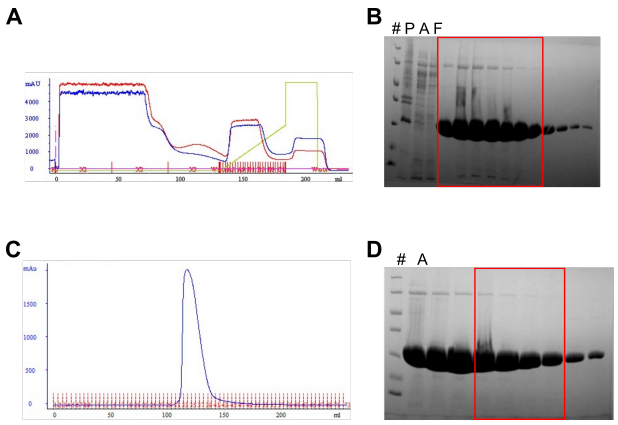


Figure 8.25: Purification of RutB S92T
RutB S92T was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB S92T appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB S92T. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

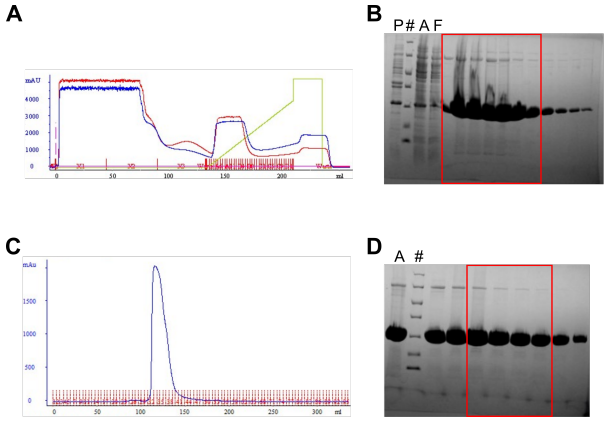


Figure 8.26: Purification of RutB S92Y
RutB S92Y was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl pellet (P), protein standard (#), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB S92Y appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB S92Y. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl IMAC pool (A) and protein standard (#), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

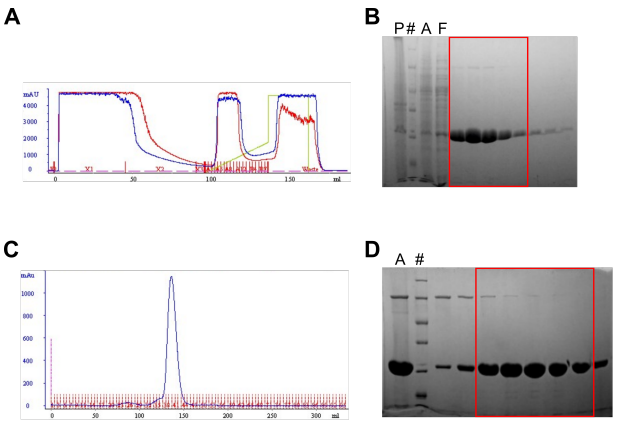


Figure 8.27: Purification of RutB Q105A
RutB Q105A was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl pellet (P), protein standard (#), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB Q105A appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB Q105A. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl IMAC pool (A) and protein standard (#), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

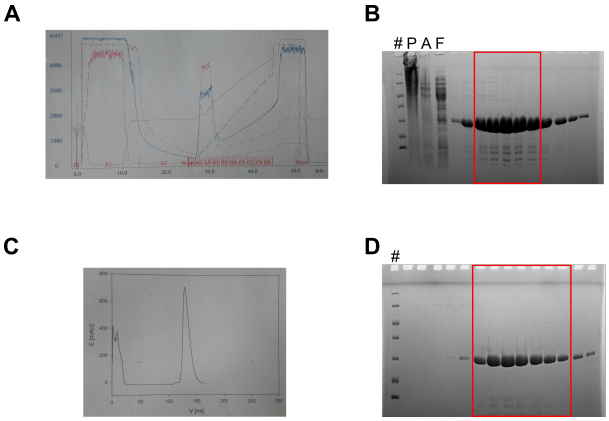


Figure 8.28: Purification of RutB K133A
RutB K133A was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (dashed green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB K133A appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB K133A. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

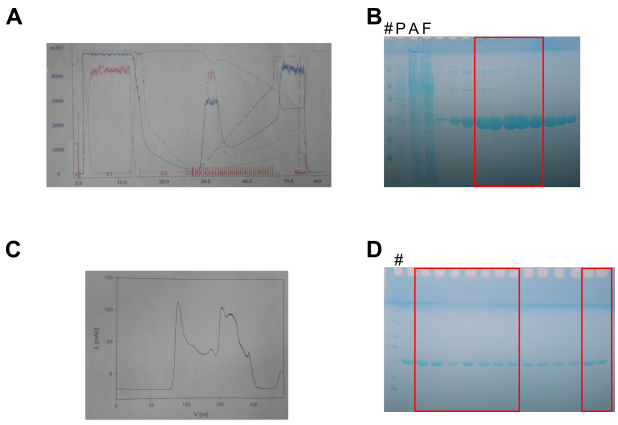


Figure 8.29: Purification of RutB Y136A
RutB Y136A was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (dashed green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB Y136A appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB Y136A. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#), 10 μl of elution fractions. Pure fractions (red rectangles) were pooled, concentrated, and frozen in liquid nitrogen.

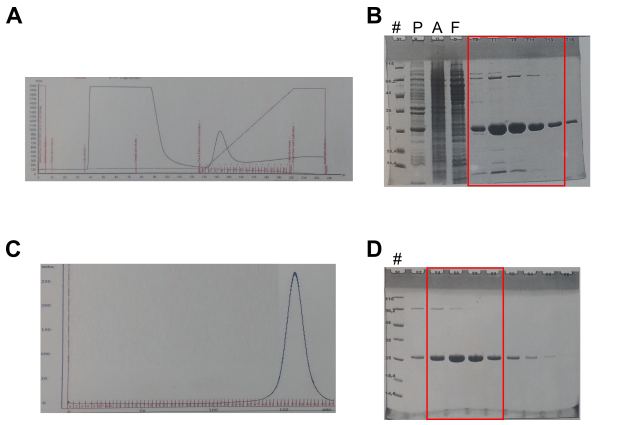


Figure 8.30: Purification of RutB Y136F
RutB Y136F was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB Y136F appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB Y136F. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

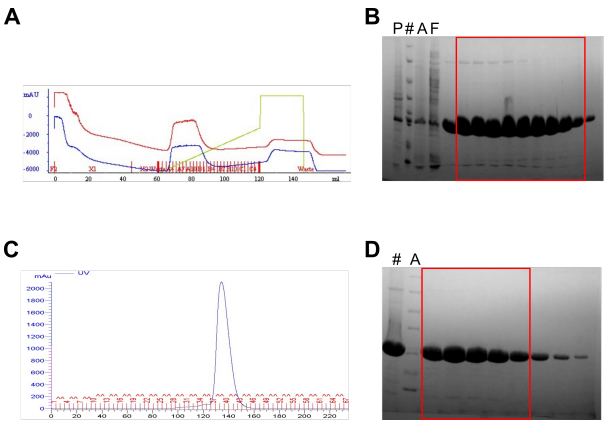


Figure 8.31: Purification of RutB C166A
RutB C166A was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl pellet (P), protein standard (#), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB C166A appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB C166A. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

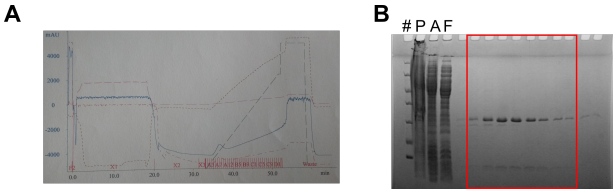


Figure 8.32: Purification of RutB C166K
RutB C166K was purified by immobilized metal ion affinity chromatography (IMAC) and dialysis. (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (dashed green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB C166K appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were dialyzed and frozen in liquid nitrogen.

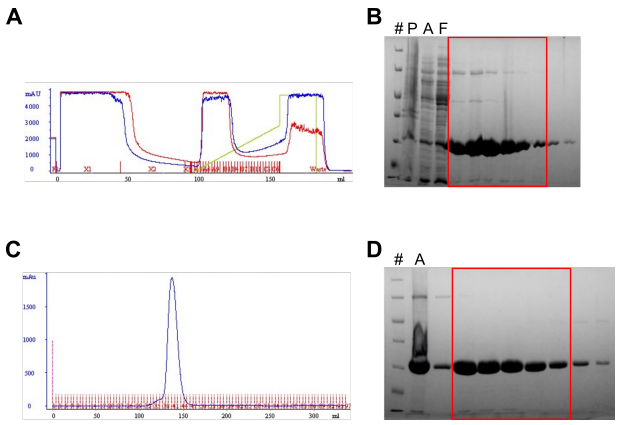


Figure 8.33: Purification of RutB C166S
RutB C166S was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB C166S appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB C166S. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

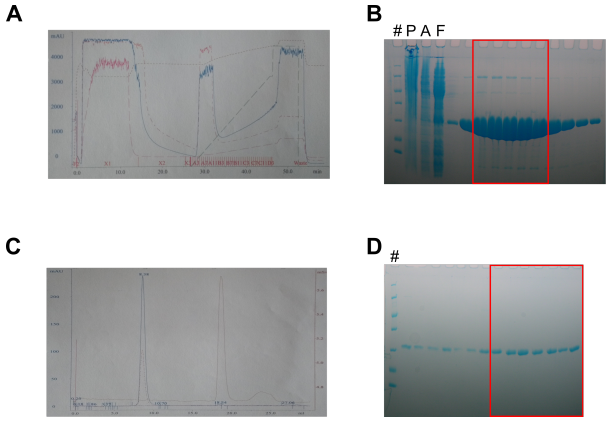


Figure 8.34: Purification of RutB C166T
RutB C166T was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (dashed green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. RutB C166T appears at a molecular weight (MW) of approximately 25 kDa (theoretical MW = 26.3 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of RutB C166T. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

8.3.2 Circular Dichroism Spectra

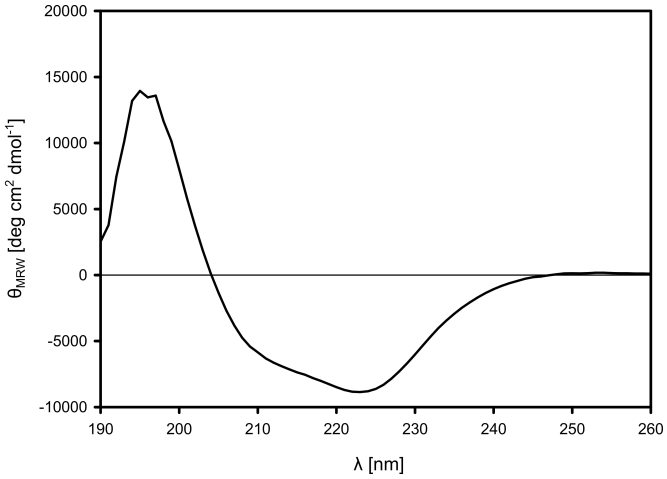


Figure 8.35: Far-UV CD spectrum of wild-type RutB
The far-UV CD spectrum of wild-type RutB reveals the overall structural integrity of the protein.

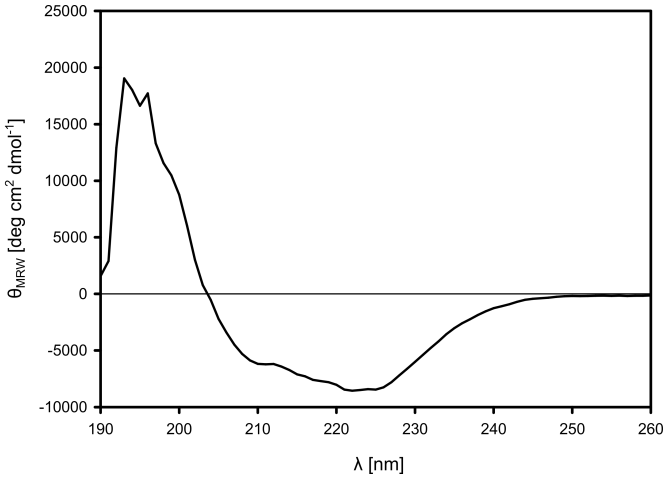


Figure 8.36: Far-UV CD spectrum of SeMet RutB
The far-UV CD spectrum of selenomethionine-labeled RutB (SeMet RutB) reveals the overall structural integrity of the protein. The CD signal is virtually identical to wild-type RutB.

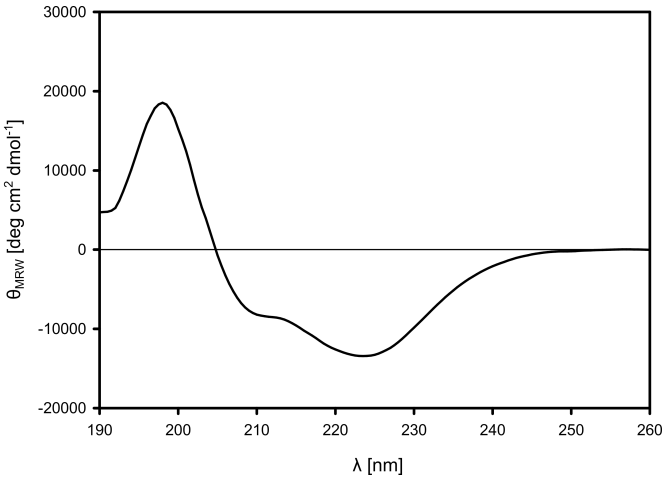


Figure 8.37: Far-UV CD spectrum of RutB D24A
The far-UV CD spectrum of RutB D24A reveals the overall structural integrity of this protein variant. However it display an increased α-helical content compared to the wild-type.

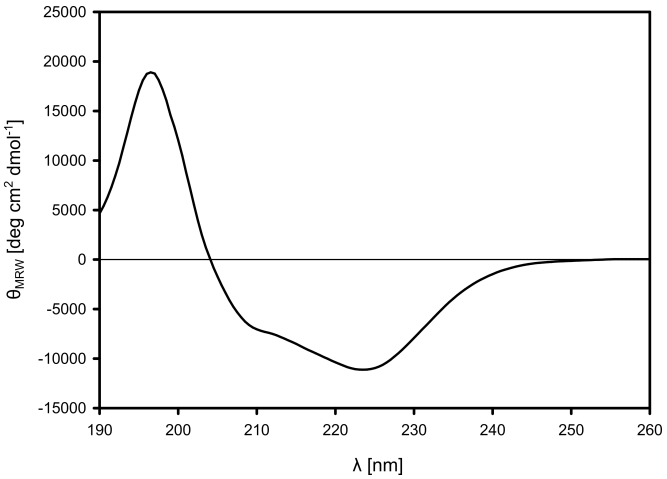


Figure 8.38: Far-UV CD spectrum of RutB D24N
The far-UV CD spectrum of RutB D24N reveals the overall structural integrity of this protein variant. However it display an increased β-sheet content compared to the wild-type.

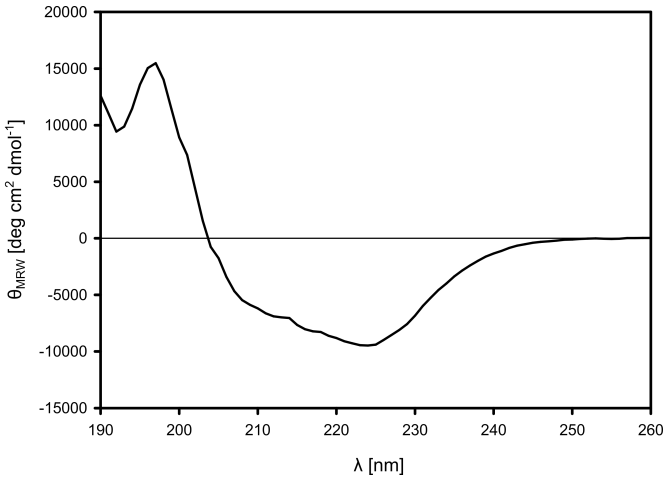


Figure 8.39: Far-UV CD spectrum of RutB Y29F
The far-UV CD spectrum of RutB Y29F reveals the overall structural integrity of this protein variant.

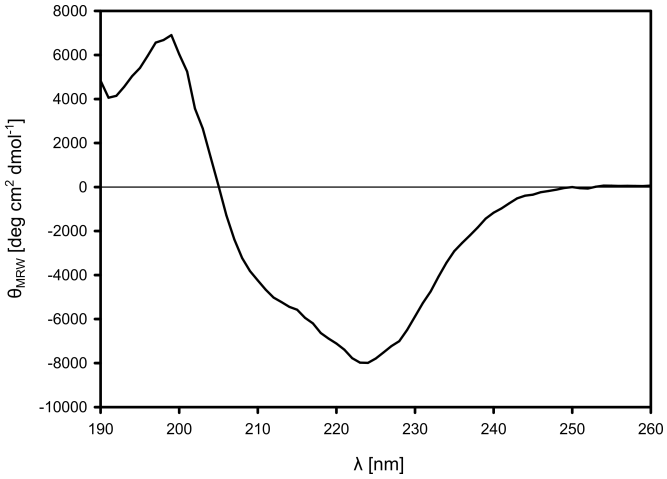


Figure 8.40: Far-UV CD spectrum of RutB Y35F
The far-UV CD spectrum of RutB Y35F reveals a partial loss of secondary structural content in this protein variant.

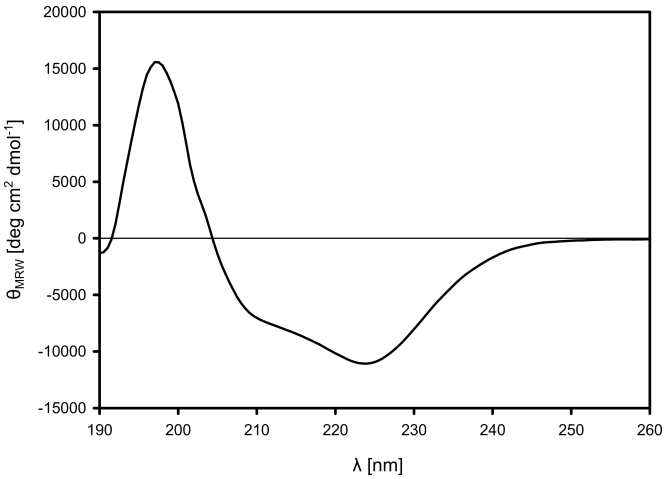


Figure 8.41: Far-UV CD spectrum of RutB N72A
The far-UV CD spectrum of RutB N72A reveals the overall structural integrity of this protein variant. However it display an increased β-sheet content compared to the wild-type.

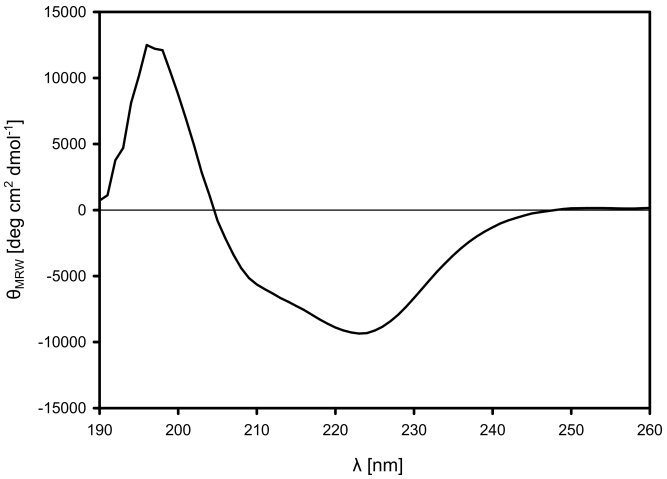


Figure 8.42: Far-UV CD spectrum of RutB W74A
The far-UV CD spectrum of RutB W74A reveals the overall structural integrity of this protein variant.

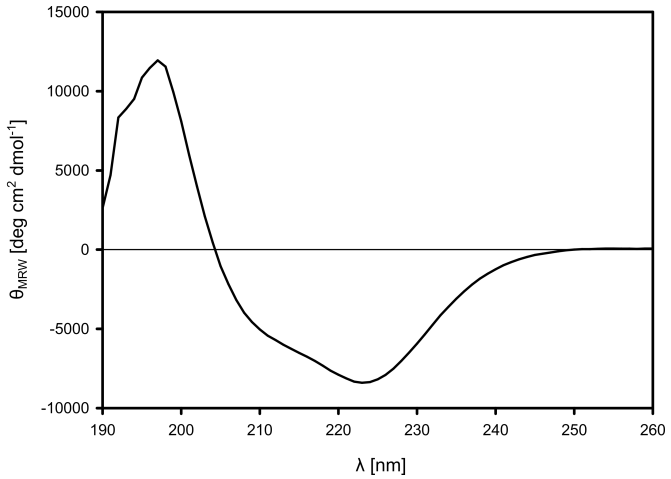


Figure 8.43: Far-UV CD spectrum of RutB W74F
The far-UV CD spectrum of RutB W74F reveals the overall structural integrity of this protein variant.

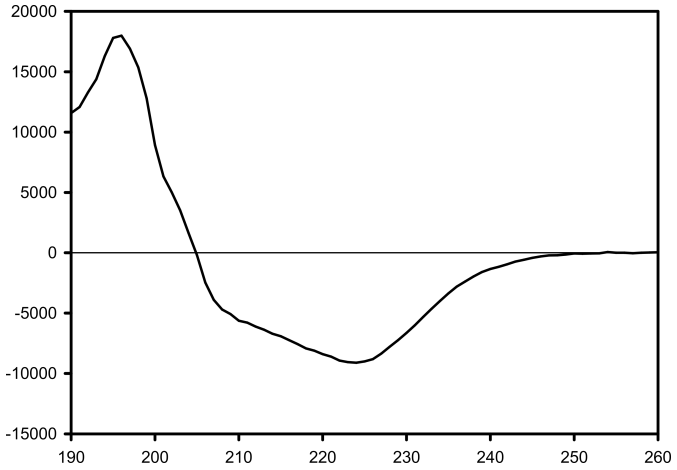


Figure 8.44: Far-UV CD spectrum of RutB E80A
The far-UV CD spectrum of RutB E80A reveals the overall structural integrity of this protein variant.

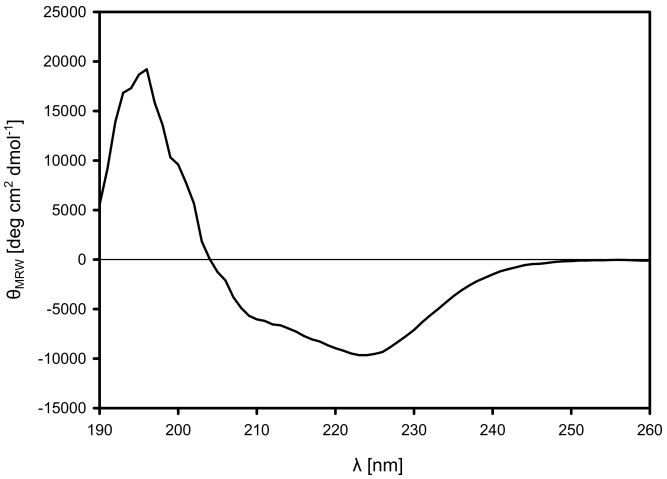


Figure 8.45: Far-UV CD spectrum of RutB E80D
The far-UV CD spectrum of RutB E80D reveals the overall structural integrity of this protein variant.

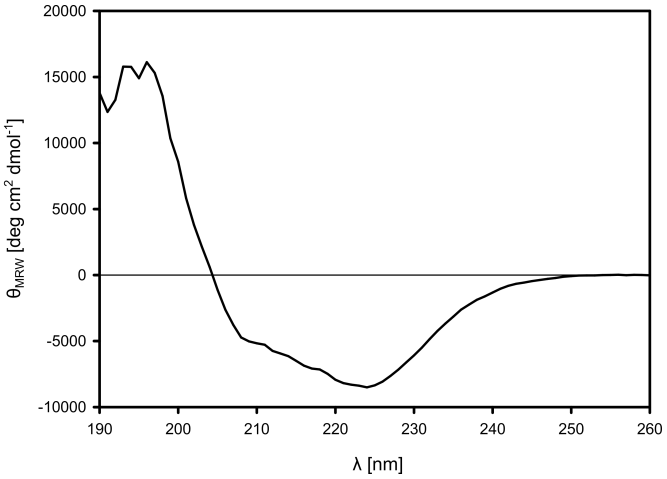


Figure 8.46: Far-UV CD spectrum of RutB S92A
The far-UV CD spectrum of RutB S92A reveals the overall structural integrity of this protein variant.

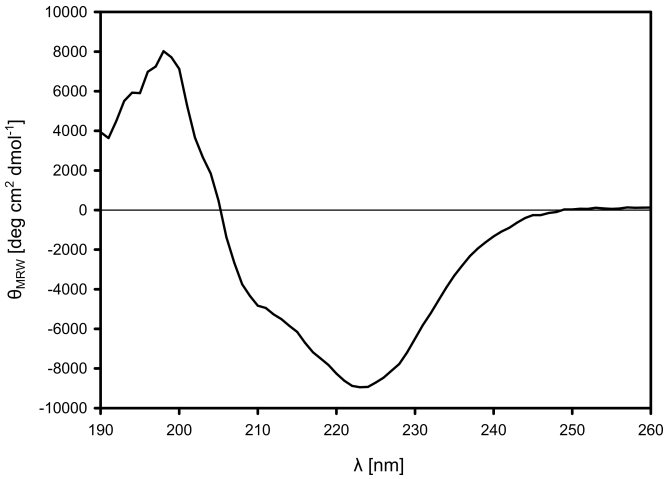


Figure 8.47: Far-UV CD spectrum of RutB S92T
The far-UV CD spectrum of RutB S92T reveals the overall structural integrity of this protein variant.

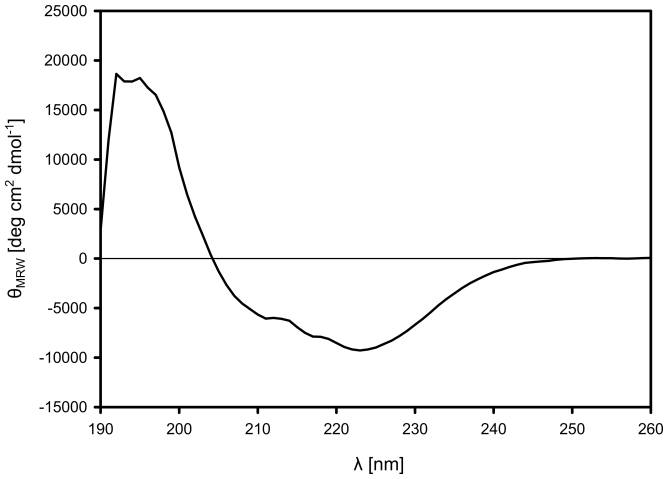


Figure 8.48: Far-UV CD spectrum of RutB S92Y
The far-UV CD spectrum of RutB S92Y reveals the overall structural integrity of this protein variant.

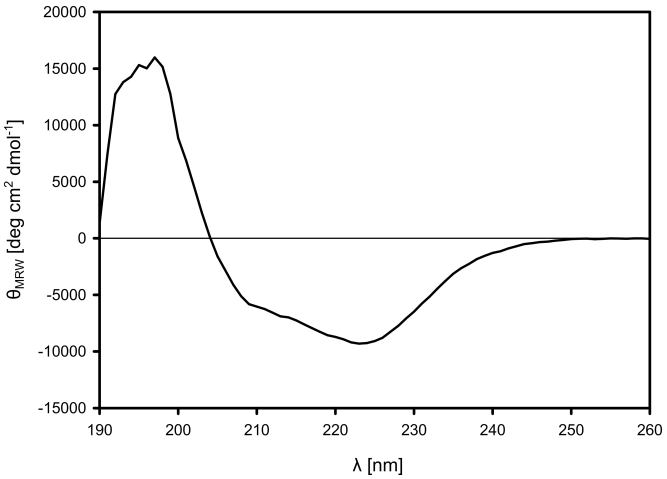


Figure 8.49: Far-UV CD spectrum of RutB Q105A
The far-UV CD spectrum of RutB Q105A reveals the overall structural integrity of this protein variant.

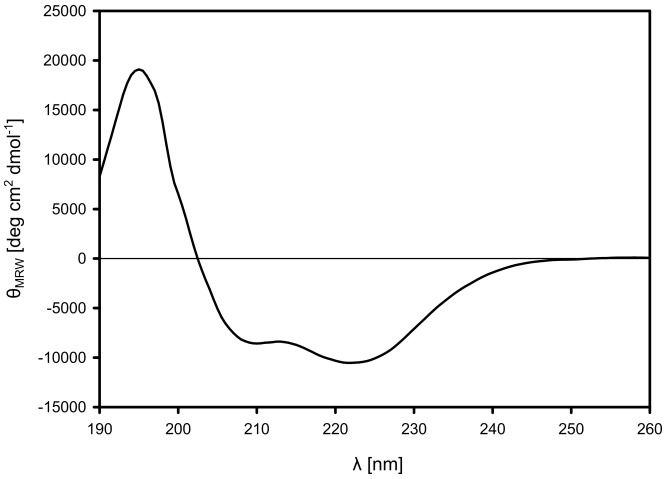


Figure 8.50: Far-UV CD spectrum of RutB K133A
The far-UV CD spectrum of RutB K133A reveals the overall structural integrity of this protein variant. However it display an increased α-helical content compared to the wild-type.

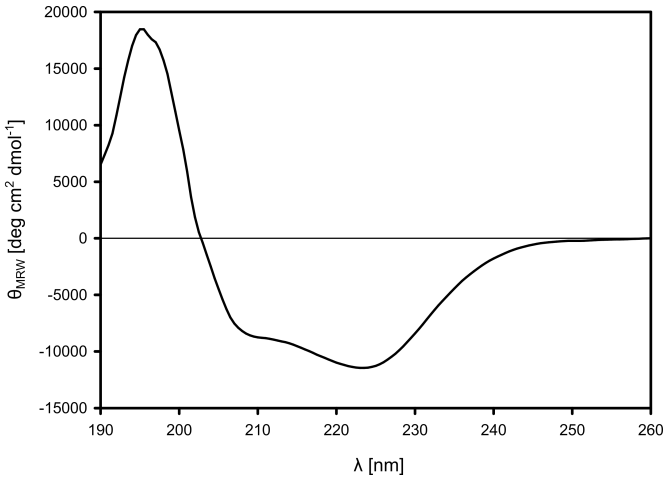


Figure 8.51: Far-UV CD spectrum of RutB Y136A
The far-UV CD spectrum of RutB Y136A reveals the overall structural integrity of this protein variant. However it display an increased α-helical content compared to the wild-type.

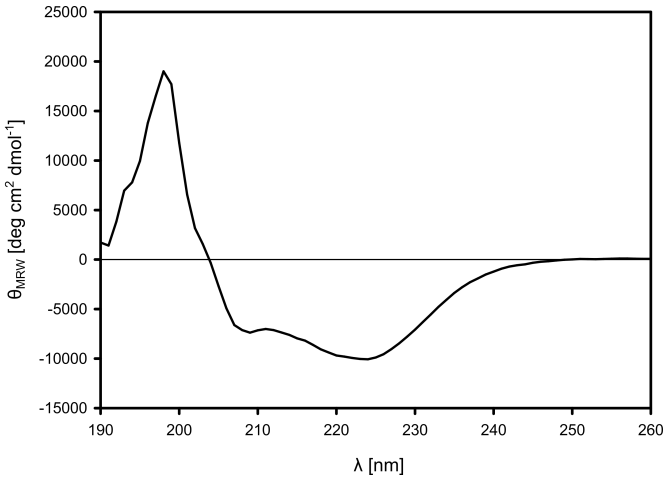


Figure 8.52: Far-UV CD spectrum of RutB Y136F
The far-UV CD spectrum of RutB Y136F reveals the overall structural integrity of this protein variant. However it display an increased α-helical content compared to the wild-type.

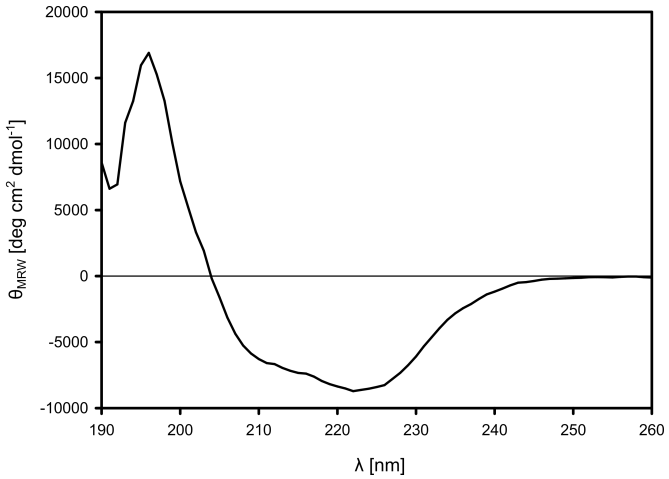


Figure 8.53: Far-UV CD spectrum of RutB C166A
The far-UV CD spectrum of RutB C166A reveals the overall structural integrity of this protein variant.

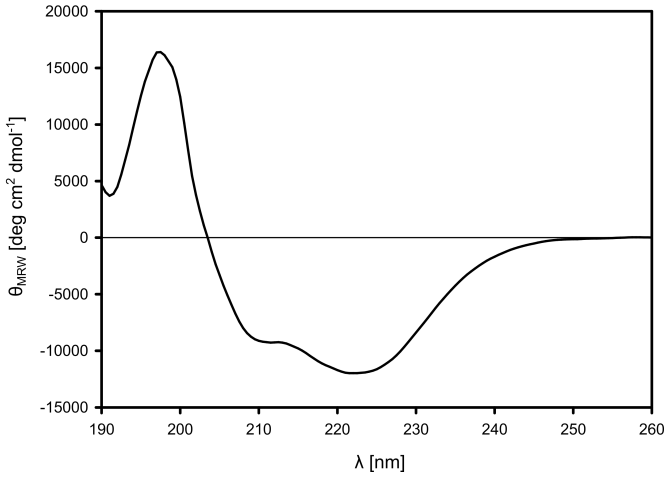


Figure 8.54: Far-UV CD spectrum of RutB C166K
The far-UV CD spectrum of RutB C166K reveals the overall structural integrity of this protein variant. However it display an increased α-helical content compared to the wild-type.

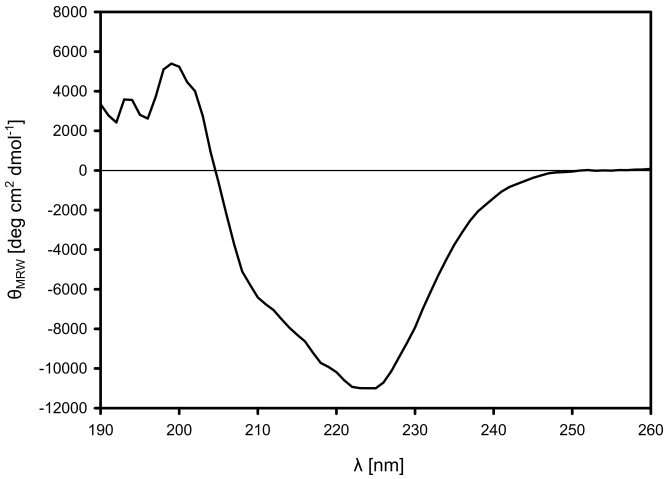


Figure 8.55: Far-UV CD spectrum of RutB C166S
The far-UV CD spectrum of RutB C166S reveals the overall structural integrity of this protein variant. However it display an increased β-sheet content compared to the wild-type.

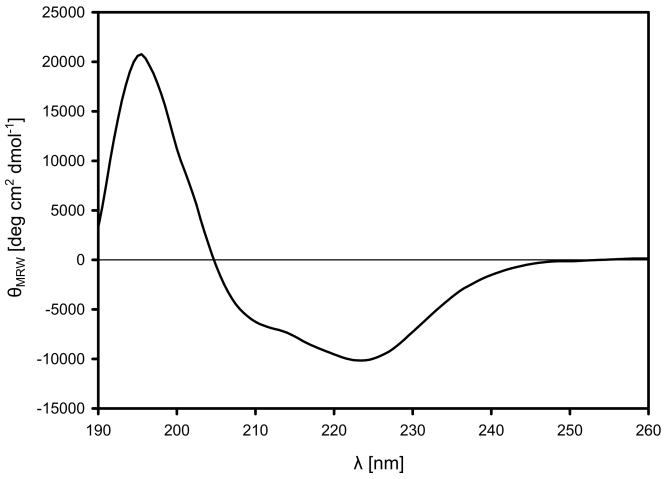


Figure 8.56: Far-UV CD spectrum of RutB C166T
The far-UV CD spectrum of RutB C166T reveals the overall structural integrity of this protein variant.

8.3.3 Protein Crystallization

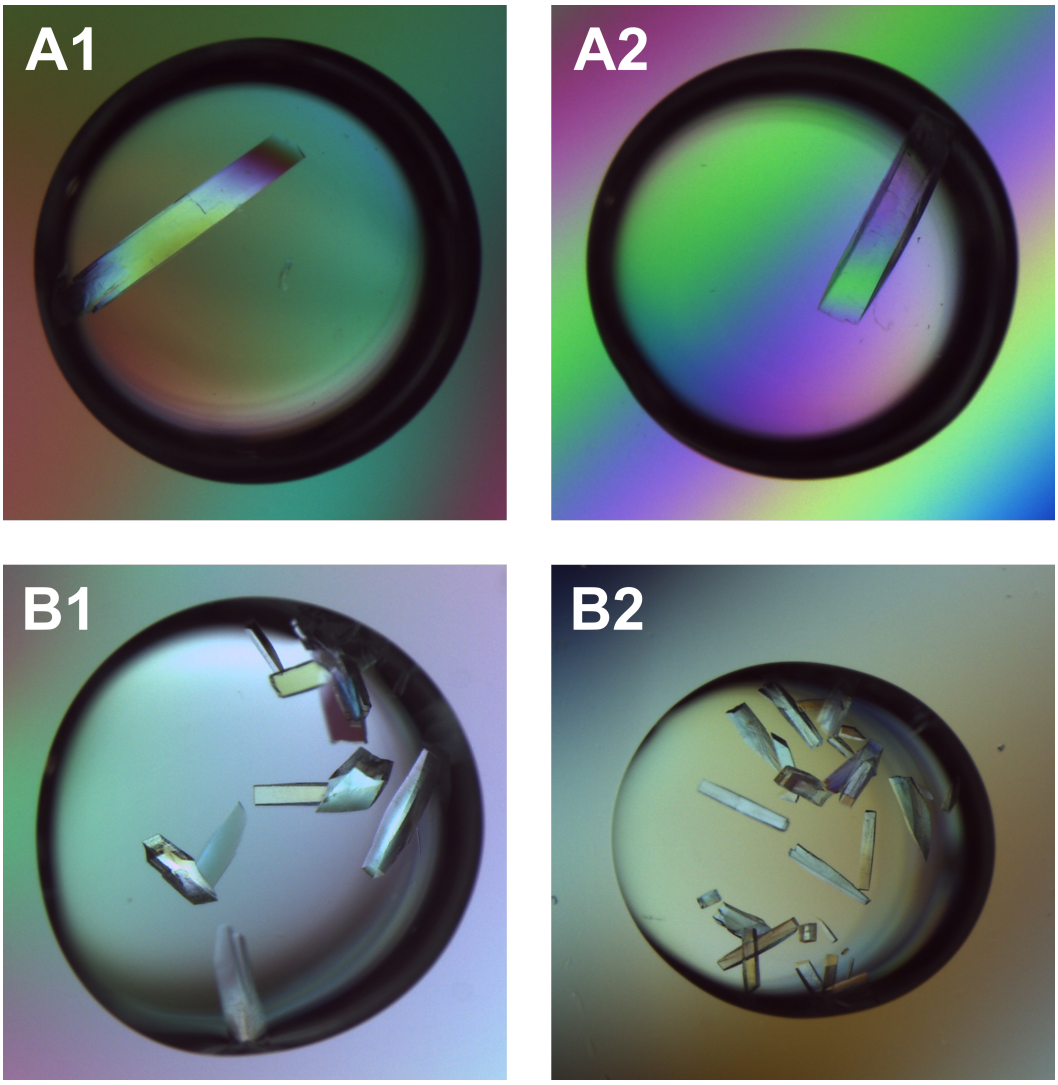


Figure 8.57: Protein crystals of RutB and RutB C166S
Protein crystals of wild-type RutB (A1; A2) grew in 2 μl drops containing equal volumes of protein solution (40 mg/ml) and crystallization solution (0.9 M NaAc, 0.1 M imidazole, pH 6.00). Protein crystals of RutB C166S grew in 2 μl drops containing equal volumes of protein solution (40 mg/ml) and crystallization solution (B1: 0.9 M NaAc, 0.1 M imidazole, pH 6.25; B2: 1.2 M NaAc, 0.1 M imidazole, pH 6.25). Optimal conditions varied slightly for variants of RutB (not shown).

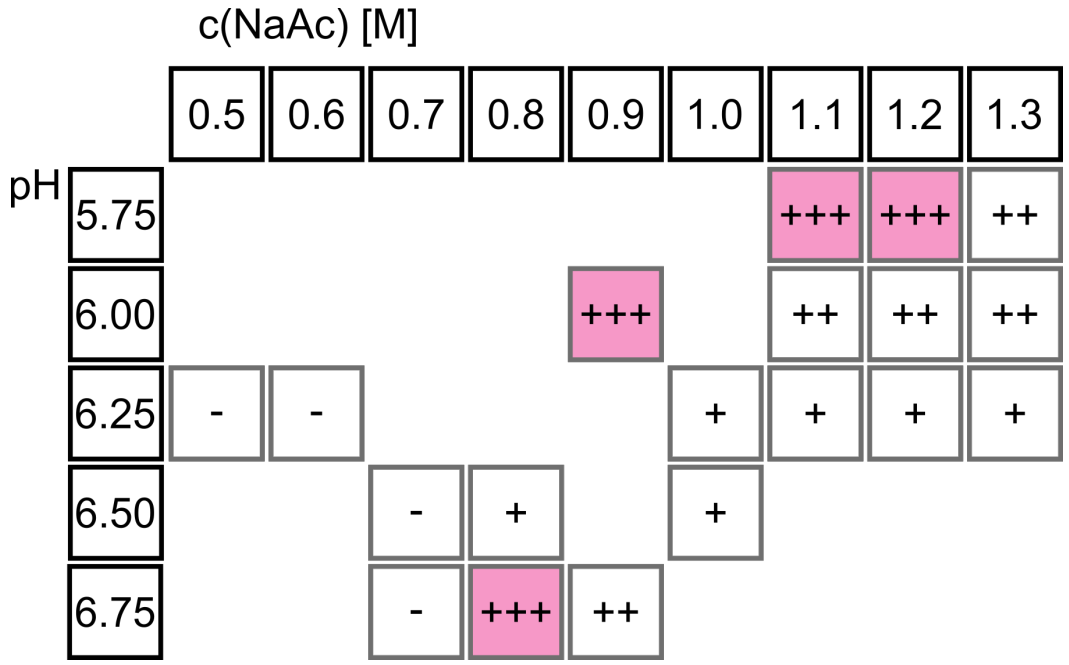


Figure 8.58: Optimization of crystallization conditions for wild-type RutB
Crystallization conditions for RutB were optimized by varying the pH and sodium acetate (NaAc) concentration based on the successful screening condition MD1-30 B5 (1 M NaAc; 0.1 M imidazole; pH 6.5; see subsection 4.2.4). The quality of the protein crystals arising from each condition was graded by visual inspection with "-" indicating no crystal formation. Ideal conditions are indicated in red.

8.3.4 Steady-State Kinetics of RutB Variants

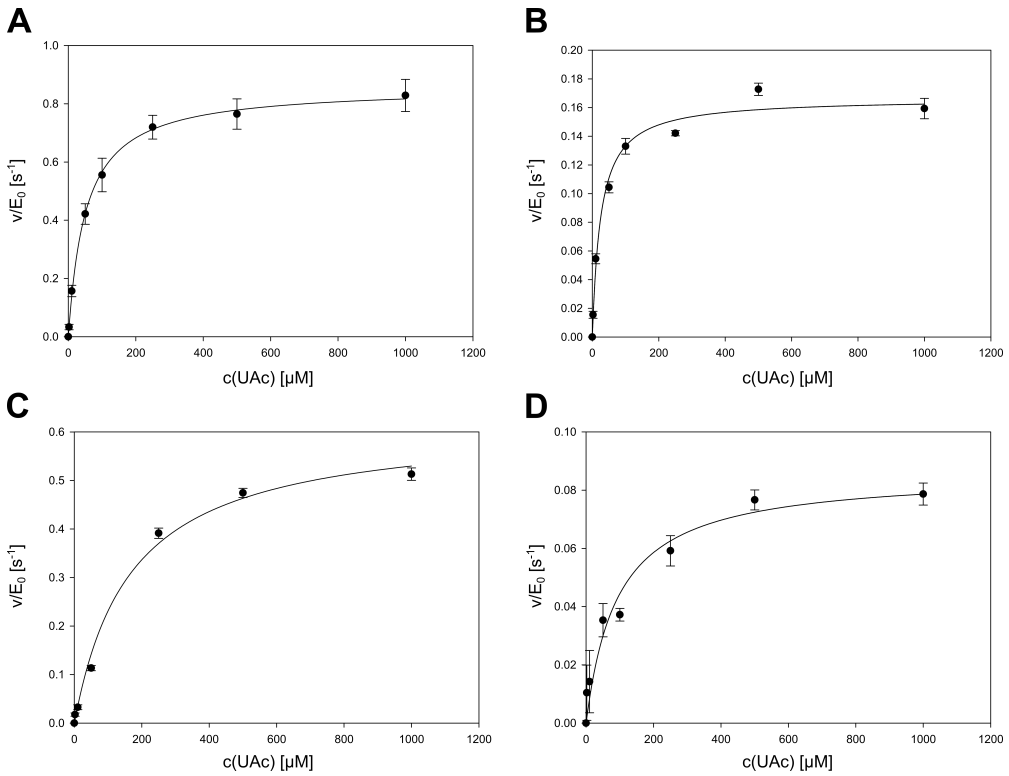


Figure 8.59: Enzyme kinetics of wild-type RutB and variants Y29F, Y35F, and N72A
Wild-type RutB (A) and variants Y29F (B), Y35F (C), and N72A (D) were reacted with ureidoacrylate as described (see subsubsection 2.4.3.2). Resulting kinetic parameters are given in Table 4.2.

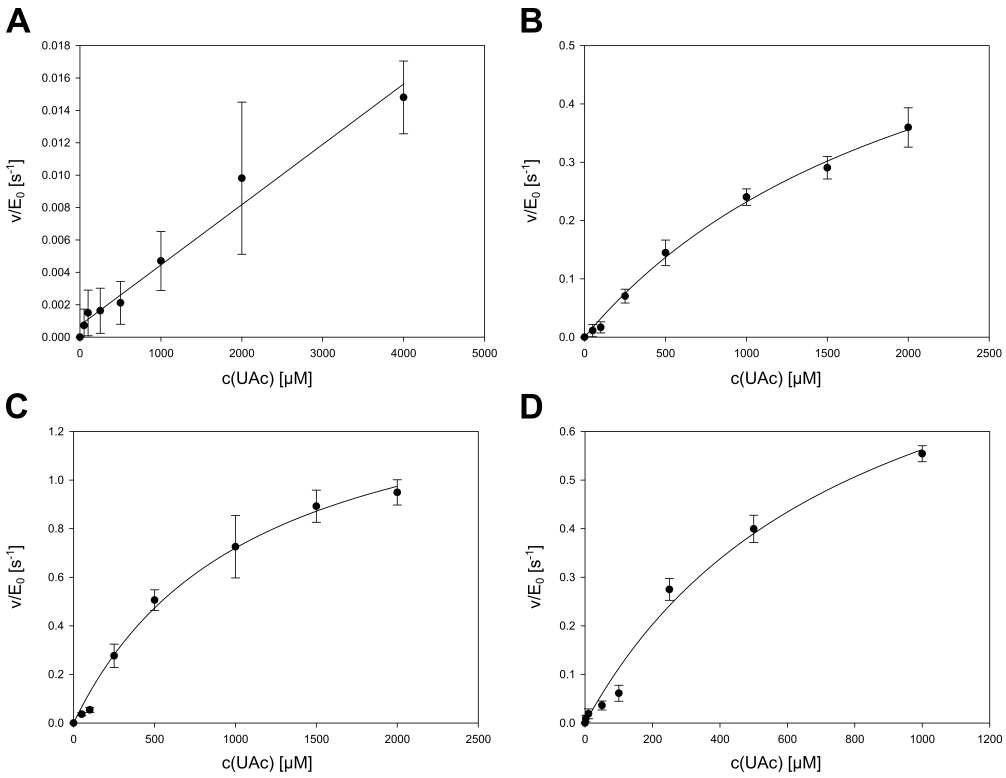


Figure 8.60: Enzyme kinetics of RutB variants W74A, W74F, E80A, and E80D

RutB variants W74A (A), W74F (B), E80A (C), and E80D (D) were reacted with ureidoacrylate as described (see subsubsection 2.4.3.2). Resulting kinetic parameters are given in Table 4.2.

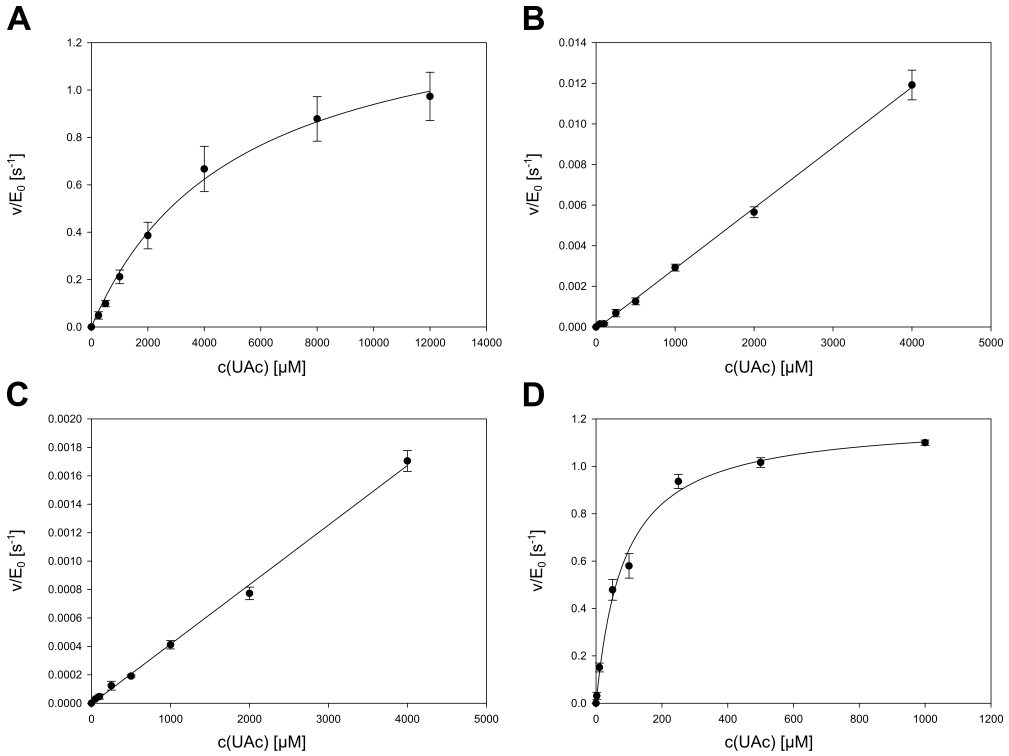


Figure 8.61: Enzyme kinetics of RutB variants S92A, S92T, S92Y, and Q105A
RutB variants S92A (A), S92T (B), S92Y (C), and Q105A (D) were reacted with ureidoacrylate as described (see subsubsection 2.4.3.2). Resulting kinetic parameters are given in Table 4.2.

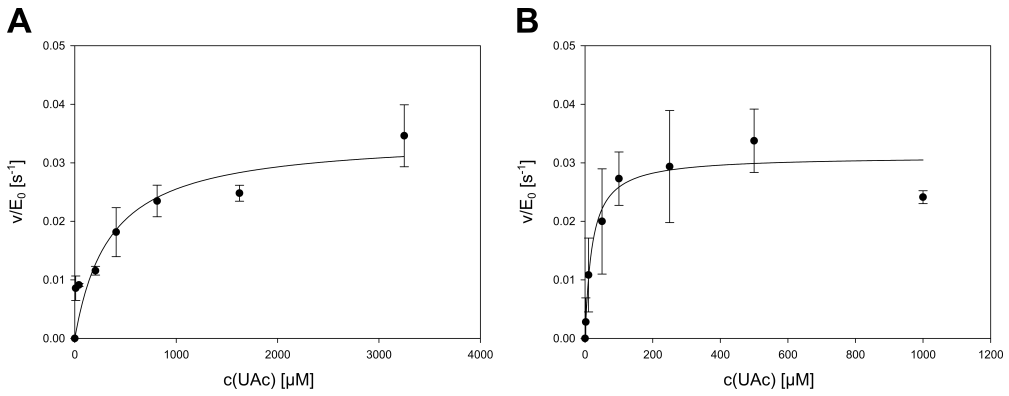


Figure 8.62: Enzyme kinetics of RutB variants Y136A and Y136F
RutB variants Y136A (A) and Y136F (B) were reacted with ureidoacrylate as described (see subsubsection 2.4.3.2). Resulting kinetic parameters are given in Table 4.2.

8.4 Evolution of Hydroxyatrazine Ethylaminohydrolase AtzB

8.4.1 Bioinformatics

Table 8.4: Homologues of AtzB
The 40 closest sequence homologues of AtzB that are not annotated as hydroxyatrazine hydrolases were identified with BLAST [32] and aligned into a multiple sequence alignment that served as a reference for the selection of active site mutations in AtzB (see subsection 5.2.2).

Accession	Annotation	Species
WP_135441588.1	amidohydrolase family protein	*Haliea sp. SAOS-164*
WP_127421824.1	amidohydrolase family protein, partial	*Mesorhizobium sp. M4A.F.Ca.ET.022.05.2.1*
TIL78870.1	amidohydrolase, partial	*Mesorhizobium sp.*
TIX15364.1	amidohydrolase, partial	*Mesorhizobium sp.*
TIU68665.1	amidohydrolase, partial	*Mesorhizobium sp.*
TJW68030.1	amidohydrolase, partial	*Mesorhizobium sp.*
WP_092584509.1	amidohydrolase	*Rhizobium loessense*
TIV83471.1	amidohydrolase, partial	*Mesorhizobium sp.*
WP_183923408.1	amidohydrolase family protein	*Rhizobium mongolense*
RWD29693.1	amidohydrolase, partial	*Mesorhizobium sp.*
WP_074071449.1	MULTISPECIES: amidohydrolase	*Rhizobium*
RVD70582.1	amidohydrolase	*Mesorhizobium sp. M4A.F.Ca.ET.029.04.2.1*
RWD10288.1	amidohydrolase	*Mesorhizobium sp.*
WP_126056427.1	MULTISPECIES: amidohydrolase	unclassified *Mesorhizobium*
RUX00053.1	hypothetical protein EOA35_19915, partial	*Mesorhizobium sp. M8A.F.Ca.ET.023.01.1.1*
WP_138396983.1	amidohydrolase	*Rhizobium sp. MHM7A*
WP_064709432.1	amidohydrolase	*Rhizobium bangladeshense*
WP_183756471.1	MULTISPECIES: amidohydrolase family protein	unclassified *Rhizobium*
WP_097521850.1	amidohydrolase	*Sinorhizobium sp. BJ1*

Accession	Annotation	Species
RUW55174.1	hypothetical protein EOA36_07870	*Mesorhizobium sp. M8A.F.Ca.ET.021.01.1.1*
WP_097598987.1	amidohydrolase	*Rhizobium sp. L9*
WP_092749835.1	amidohydrolase	*Rhizobium aethiopicum*
WP_026614989.1	MULTISPECIES: amidohydrolase	*Sinorhizobium/Ensifer* group
WP_183864506.1	amidohydrolase family protein	*Rhizobium sp. BK399*
WP_023775217.1	amidohydrolase	*Mesorhizobium sp. LNHC229A00*
WP_085035411.1	amidohydrolase	*Ensifer aridi*
WP_064713025.1	amidohydrolase	*Rhizobium bangladeshense*
WP_008532560.1	amidohydrolase	*Rhizobium sp. Pop5*
WP_184457550.1	amidohydrolase family protein	*Rhizobium aethiopicum*
WP_037129375.1	amidohydrolase	*Rhizobium sp. CF394*
WP_180897215.1	MULTISPECIES: amidohydrolase family protein	unclassified *Martelella*
RVC56219.1	amidohydrolase, partial	*Mesorhizobium sp. M4B.F.Ca.ET.088.02.2.1*
WP_127392470.1	MULTISPECIES: amidohydrolase	unclassified *Mesorhizobium*
WP_064243158.1	amidohydrolase	*Ensifer glycinis*
WP_183699729.1	amidohydrolase family protein	*Rhizobium sp. BK049*
TIW44722.1	amidohydrolase, partial	*Mesorhizobium sp.*
WP_202294944.1	amidohydrolase family protein	*Mesorhizobium sp. 131-2-1*
TIW20014.1	amidohydrolase	*Mesorhizobium sp.*
TIS64324.1	amidohydrolase	*Mesorhizobium sp.*
WP_016557374.1	amidohydrolase	*Rhizobium grahamii*

8.4.2 Protein Purification

The bands of the protein standard (LMW, Thermo Fisher Scientific) used in the following SDS-PAGEs have the following molecular weights: 116.0, 66.2, 45.0, 35.0, 25.0, 18.4, 14.4 kDa. Preparative size exclusion chromatography runs were followed at 280 nm.

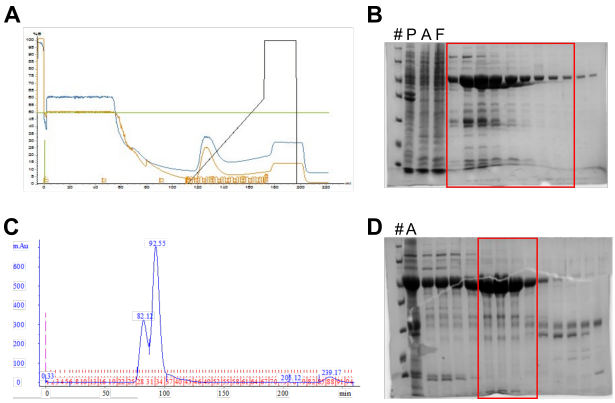


Figure 8.63: Purification of AtzB
AtzB was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (black). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 53.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of AtzB. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

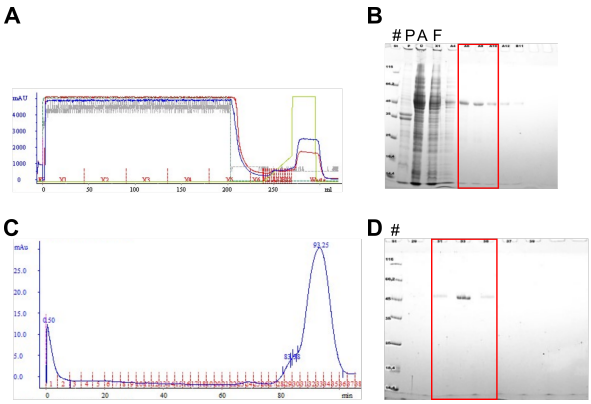


Figure 8.64: Purification of AtzB S218C
AtzB S218C was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB S218C appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 53.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of AtzB S218C. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

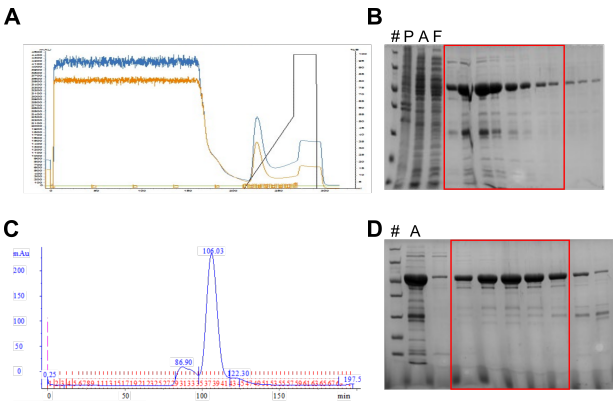


Figure 8.65: Purification of AtzB S219Q
AtzB S219Q was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (black). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB S219Q appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 53.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of AtzB S219Q. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

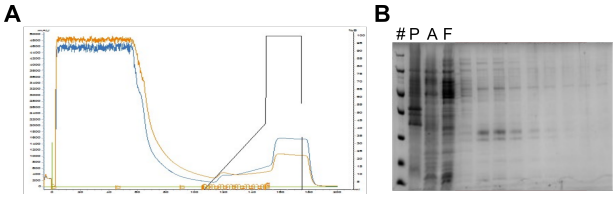


Figure 8.66: Purification attempt of AtzB I170N
An attempt to purify AtzB I170N by immobilized metal ion affinity chromatography (IMAC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (black). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB I170N could not be identified.

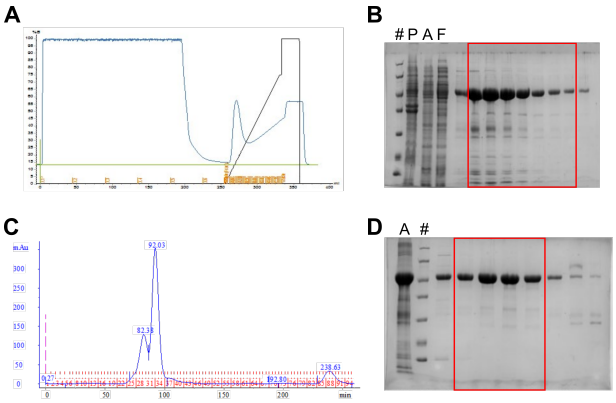


Figure 8.67: Purification of AtzB S218C S219Q
AtzB S218C S219Q was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), volume% of 1 M imidazole (black). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB S218C S219Q appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 53.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of AtzB S218C S219Q. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl IMAC pool (A) and protein standard (#), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

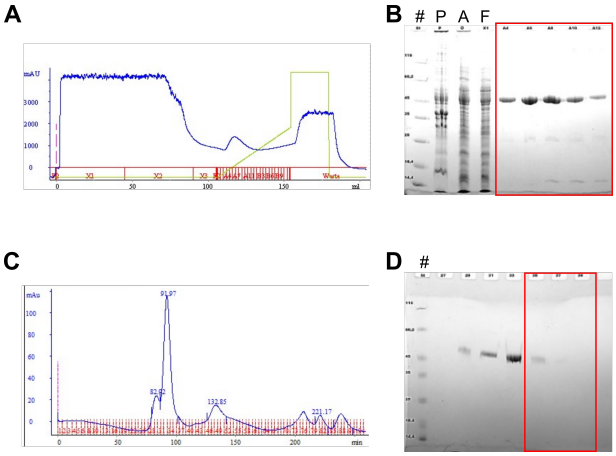


Figure 8.68: Purification of AtzB S218C I170N
AtzB S218C I170N was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB S218C I170N appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 53.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of AtzB S218C I170N. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

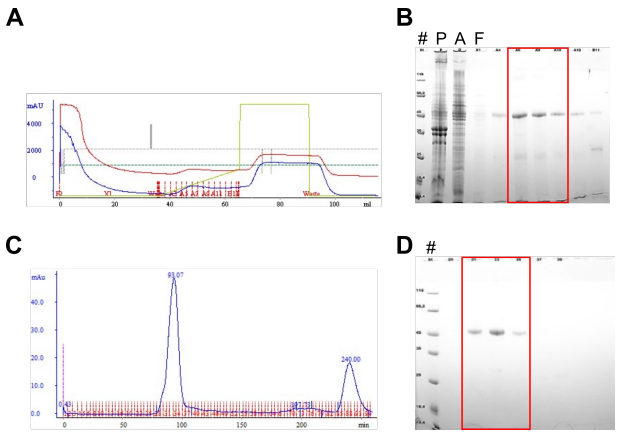


Figure 8.69: Purification of AtzB S218C I222N
AtzB S218C I222N was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB S218C I222N appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 53.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of AtzB S218C I222N. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

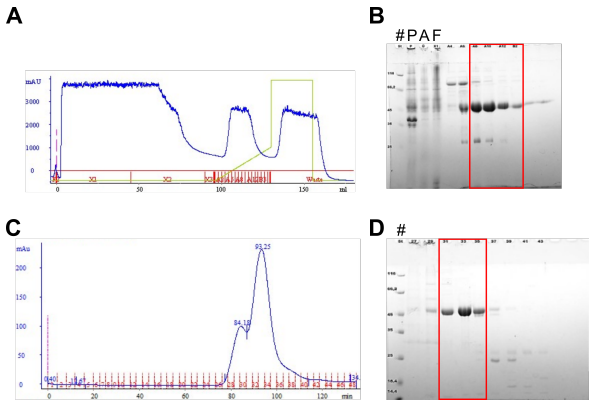


Figure 8.70: Purification of AtzB S218C S219Q I170N
AtzB S218C S219Q I170N was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB S218C S219Q I170N appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 53.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of AtzB S218C S219Q I170N. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

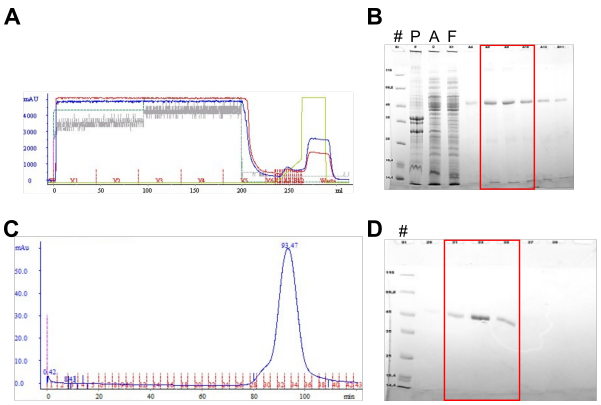


Figure 8.71: Purification of AtzB S218C S219Q I222N
AtzB S218C S219Q I222N was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (green). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB S218C S219Q I222N appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 53.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of AtzB S218C S219Q I222N. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

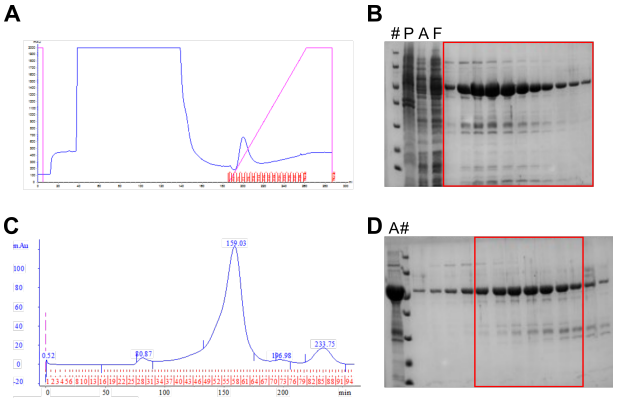


Figure 8.72: Purification of AtzB S218C S219Q I170N I222N
AtzB S218C S219Q I170N I222N was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), volume% of 1 M imidazole (pink). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB S218C S219Q I170N I222N appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 53.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of AtzB S218C S219Q I170N I222N. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl IMAC pool (A) and protein standard (#), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

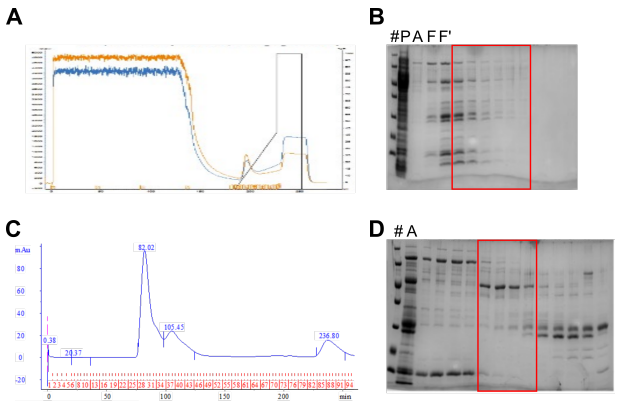


Figure 8.73: Purification of AtzB_Hom_Hal
AtzB_Hom_Hal was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (black). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB_Hom_Hal appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 50.7 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of AtzB_Hom_Hal. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

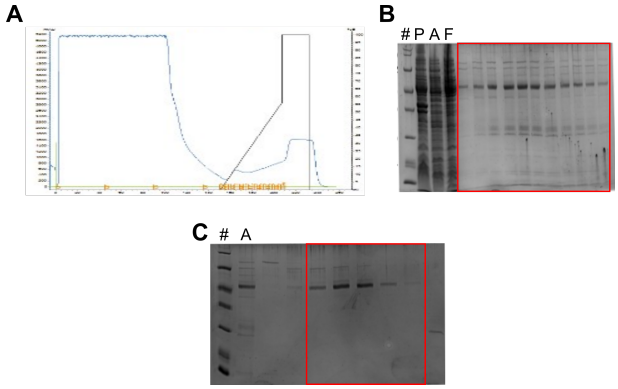


Figure 8.74: Purification of AtzB_Hom_Hal C209S Q210S
AtzB_Hom_Hal C209S Q210S was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), volume% of 1 M imidazole (black). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB_Hom_Hal C209S Q210S appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 50.7 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

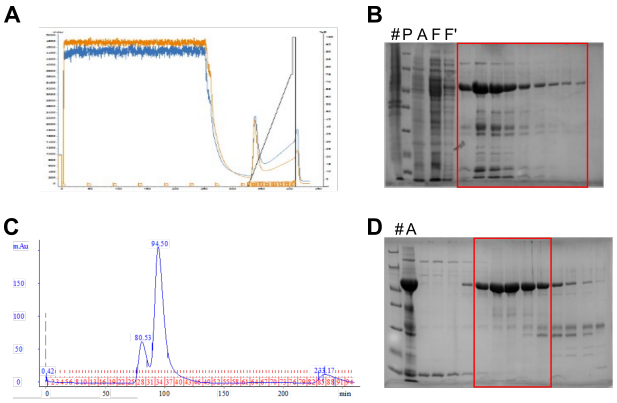


Figure 8.75: Purification of AtzB_Hom_Pleo
AtzB_Hom_Pleo was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (black). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F, F'), and 10 μl of elution fractions. AtzB_Hom_Pleo appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 51.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of AtzB_Hom_Pleo. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

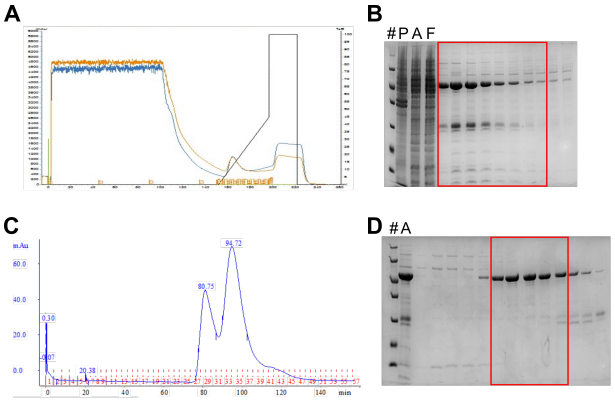


Figure 8.76: Purification of AtzB_Hom_Pleo C213S
AtzB_Hom_Pleo C213S was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (black). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB_Hom_Pleo C213S appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 51.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of AtzB_Hom_Pleo C213S. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl protein standard (#) and IMAC pool (A), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

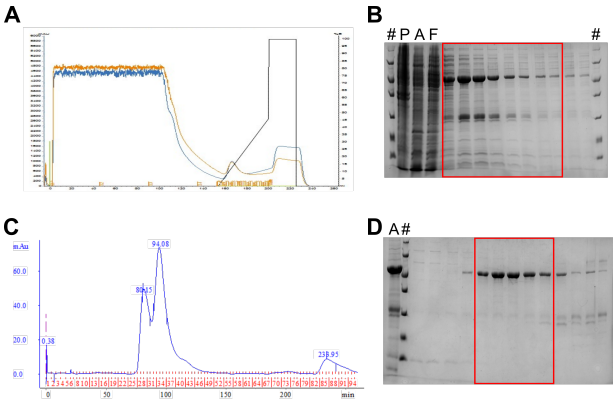


Figure 8.77: Purification of AtzB_Hom_Pleo Q214S
AtzB_Hom_Pleo Q214S was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), OD_{260} (red), volume% of 1 M imidazole (black). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl protein standard (#), pellet (P), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB_Hom_Pleo Q214S appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 51.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SEC elution profile of AtzB_Hom_Pleo Q214S. (D) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl IMAC pool (A) and protein standard (#), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

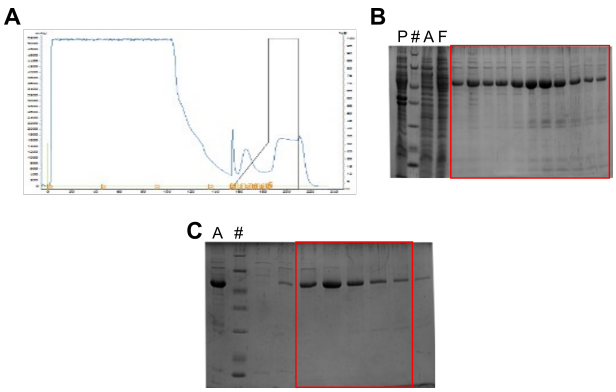


Figure 8.78: Purification of AtzB_Hom_Pleo C213S Q214S
AtzB_Hom_Pleo C213S Q214S was purified by immobilized metal ion affinity chromatography (IMAC) and preparative size exclusion chromatography (SEC). (A) Chromatogram of IMAC: OD_{280} (blue), volume% of 1 M imidazole (black). (B) SDS-PAGE analysis (13.5% acrylamide) of IMAC elution fractions: 5 μl pellet (P), protein standard (#), crude extract (A), flowthrough (F), and 10 μl of elution fractions. AtzB_Hom_Pleo C213S Q214S appears at a molecular weight (MW) of approximately 45 kDa (theoretical MW = 51.2 kDa). Fractions indicated by a red rectangle were further purified by preparative SEC. (C) SDS-PAGE (13.5% acrylamide) of SEC elution fractions: 5 μl IMAC pool (A) and protein standard (#), 10 μl of elution fractions. Pure fractions (red rectangle) were pooled, concentrated, and frozen in liquid nitrogen.

8.4.3 Circular Dichroism and Thermal Stability Analysis

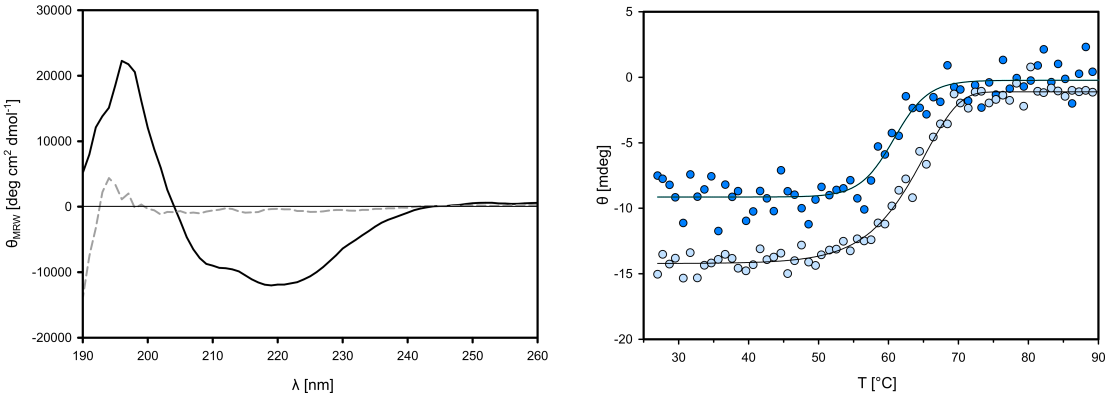


Figure 8.79: Far-UV CD spectrum and thermal denaturation curve of wild-type AtzB
The far-UV CD spectrum of wild-type AtzB before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) were fitted with sigmoidal functions (black line), yielding T_M values of 60.8°C and 63.4°C, respectively.

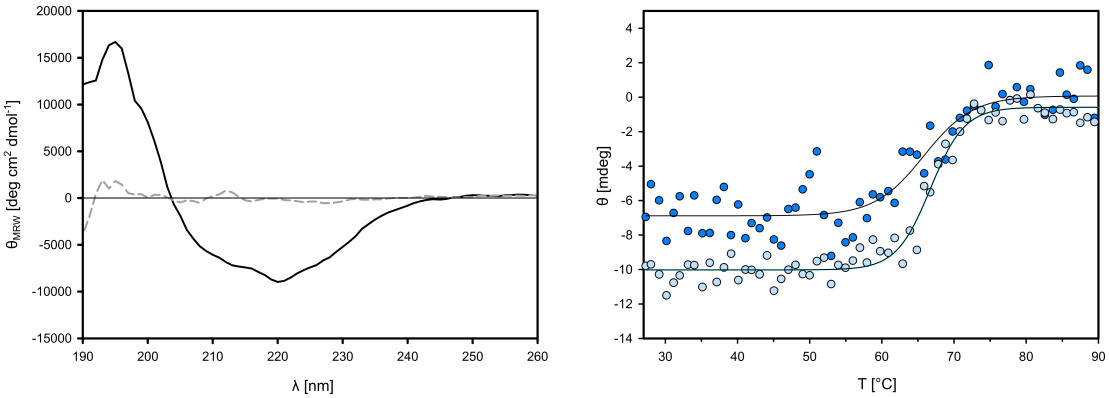


Figure 8.80: Far-UV CD spectrum and thermal denaturation curve of AtzB S218C
The far-UV CD spectrum of AtzB S218C before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) were fitted with sigmoidal functions (black line), yielding T_M values of 65.8°C and 66.7°C, respectively.

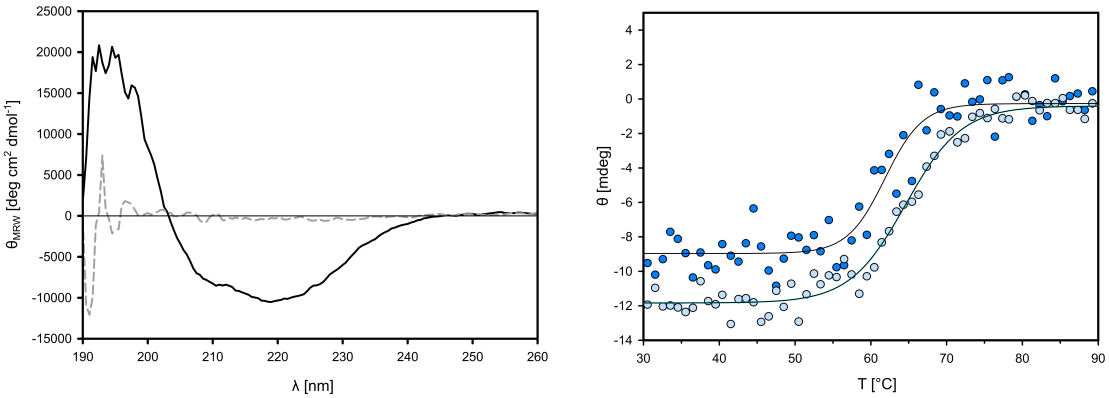


Figure 8.81: Far-UV CD spectrum and thermal denaturation curve of AtzB S219Q
The far-UV CD spectrum of AtzB S219Q before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) were fitted with sigmoidal functions (black line), yielding T_M values of 61.9°C and 64.6°C, respectively.

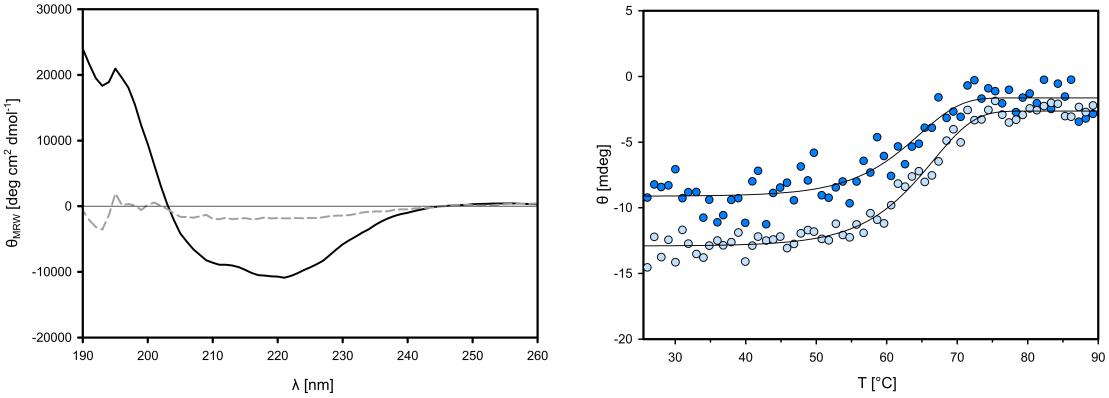


Figure 8.82: Far-UV CD spectrum and thermal denaturation curve of AtzB I222N
The far-UV CD spectrum of AtzB I222N before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) were fitted with sigmoidal functions (black line), yielding T_M values of 62.6°C and 64.5°C, respectively.

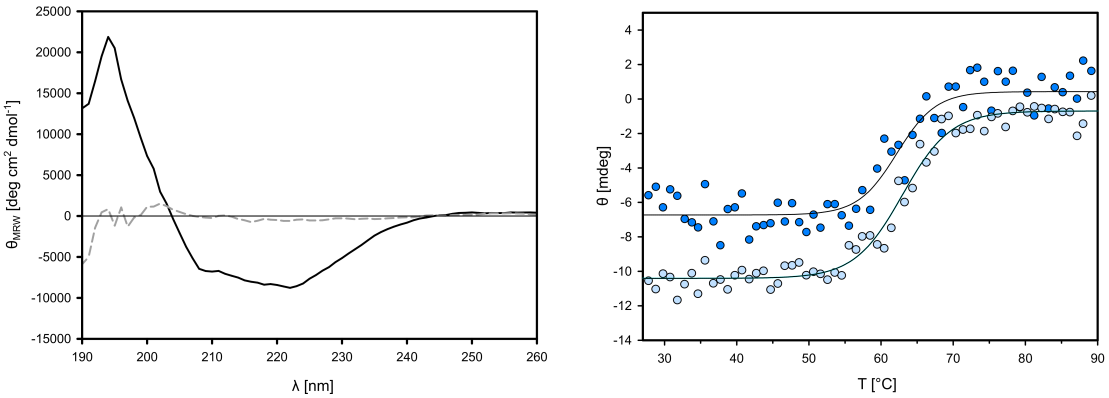


Figure 8.83: Far-UV CD spectrum and thermal denaturation curve of AtzB S218C S219Q
The far-UV CD spectrum of AtzB S218C S219Q before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) were fitted with sigmoidal functions (black line), yielding T_M values of 62.2°C and 63.0°C, respectively.

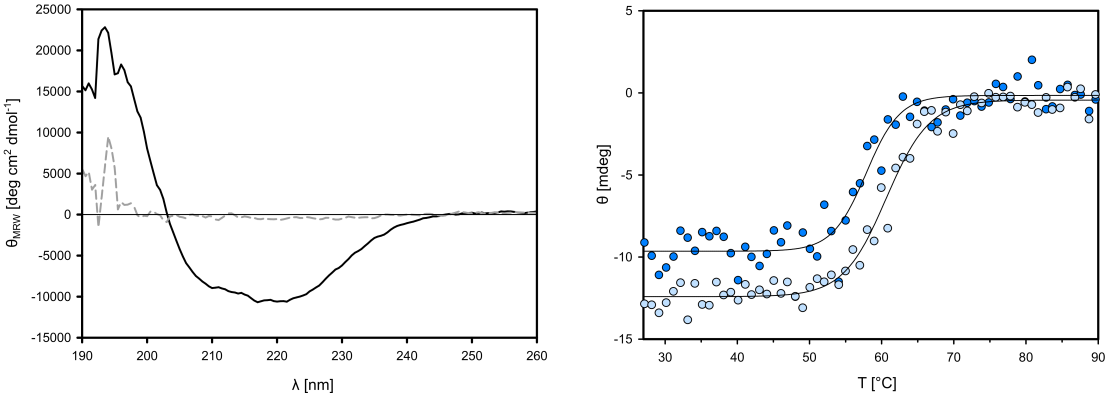


Figure 8.84: Far-UV CD spectrum and thermal denaturation curve of AtzB S218C I170N
The far-UV CD spectrum of AtzB S218C I170N before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) were fitted with sigmoidal functions (black line), yielding T_M values of 57.8°C and 60.6°C, respectively.

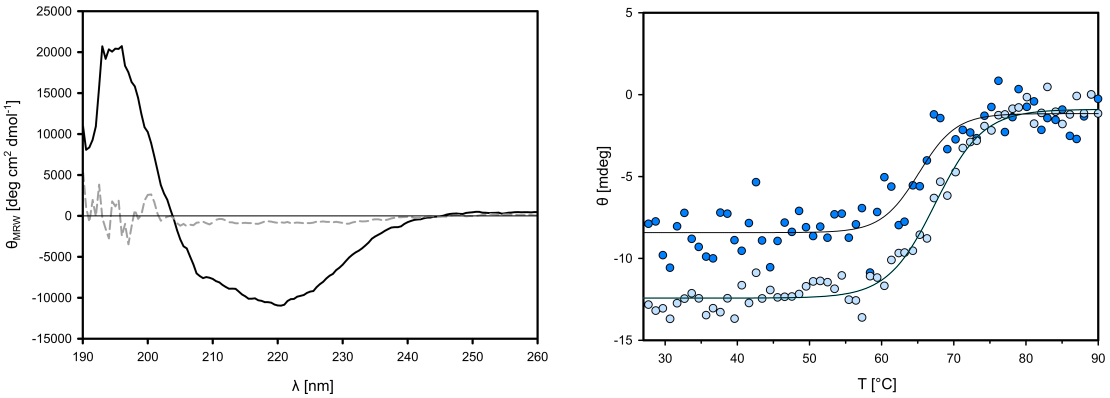


Figure 8.85: Far-UV CD spectrum and thermal denaturation curve of AtzB S218C I222N
The far-UV CD spectrum of AtzB S218C I222N before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) were fitted with sigmoidal functions (black line), yielding T_M values of 65.1°C and 67.4°C, respectively.

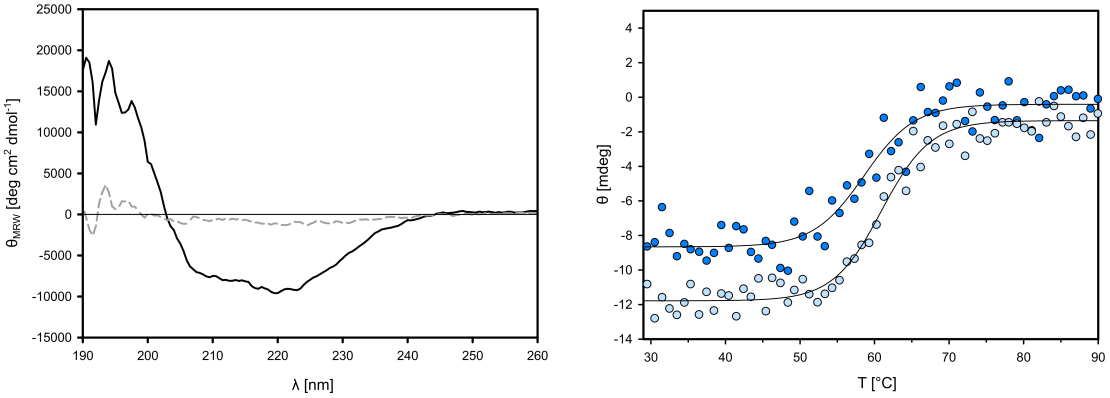


Figure 8.86: Far-UV CD spectrum and thermal denaturation curve of AtzB S218C S219Q I170N
The far-UV CD spectrum of AtzB S218C S219Q I170N before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) were fitted with sigmoidal functions (black line), yielding T_M values of 58.6°C and 60.9°C, respectively.

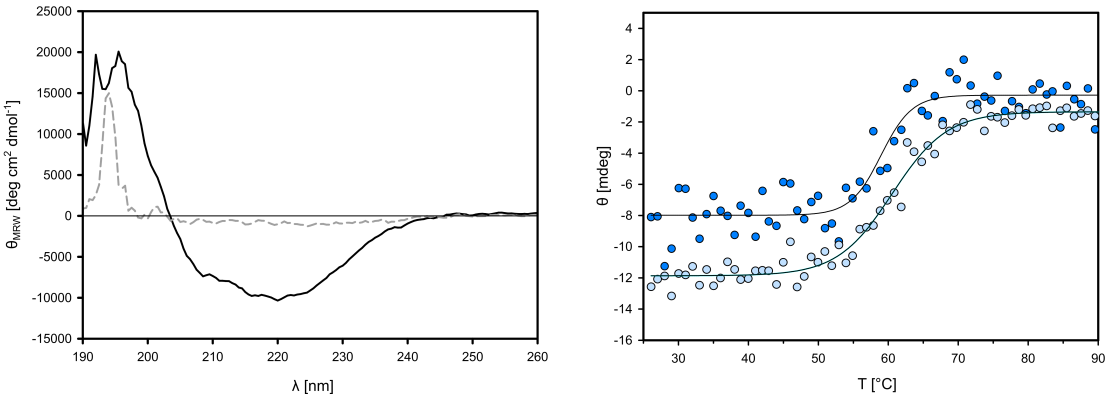


Figure 8.87: Far-UV CD spectrum and thermal denaturation curve of AtzB S218C S219Q I222N
The far-UV CD spectrum of AtzB S218C S219Q I222N before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) were fitted with sigmoidal functions (black line), yielding T_M values of 59.0°C and 60.6°C, respectively.

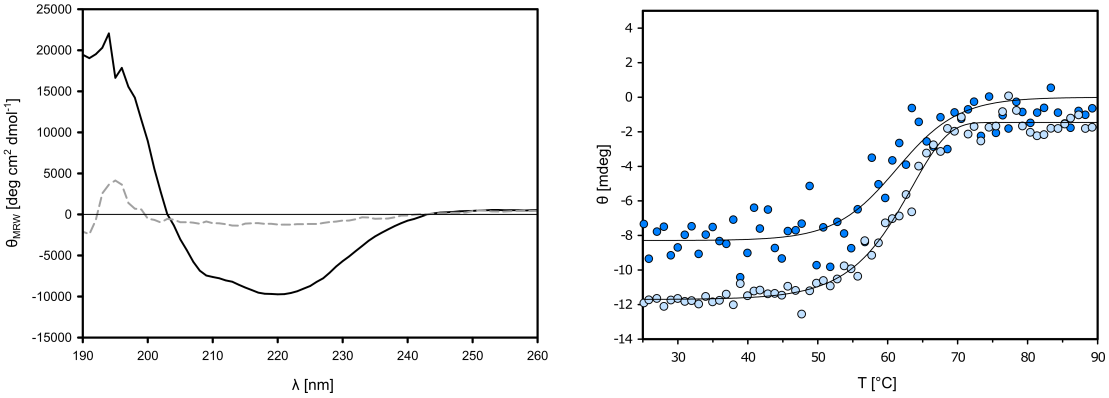


Figure 8.88: Far-UV CD spectrum and thermal denaturation curve of AtzB S218C S219Q I170N I222N
The far-UV CD spectrum of AtzB S218C S219Q I170N I222N before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) were fitted with sigmoidal functions (black line), yielding T_M values of 61.3°C and 61.5°C, respectively.

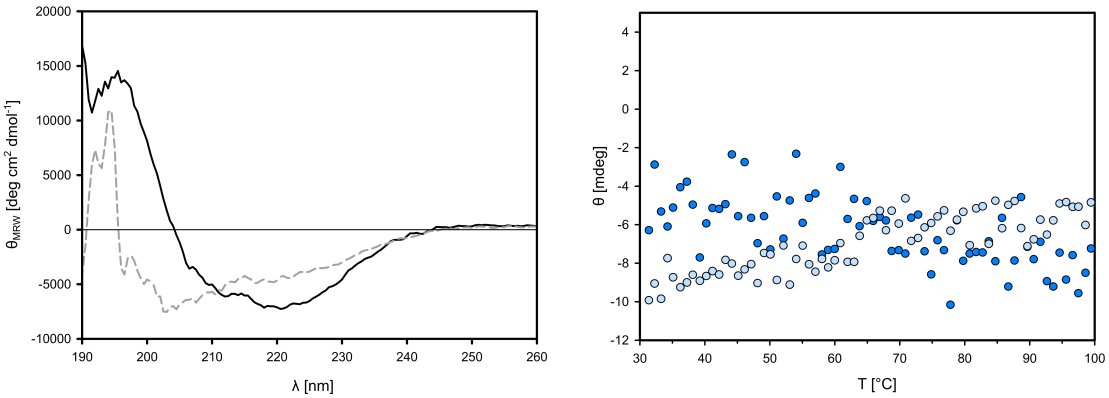


Figure 8.89: Far-UV CD spectrum and thermal denaturation curve of AtzB_Hom_Hal C209S

The far-UV CD spectrum of AtzB_Hom_Hal C209S before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) could not be fitted and no T_M values could be determined

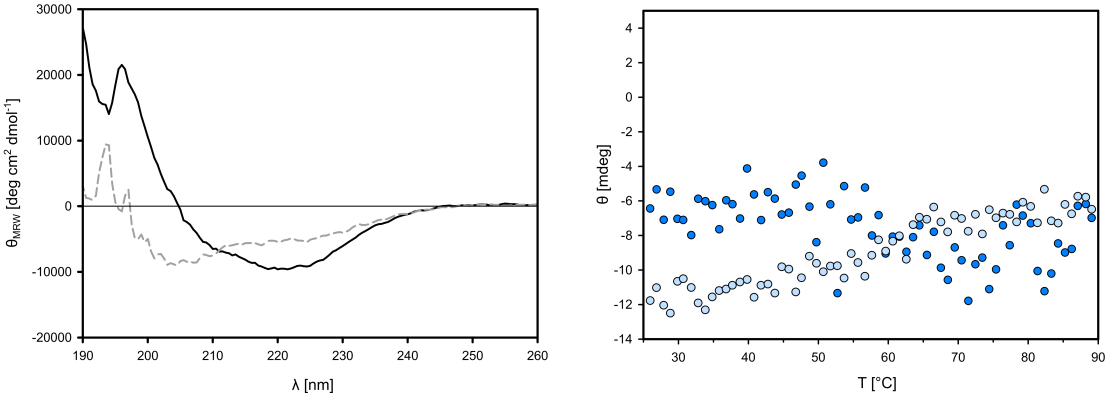


Figure 8.90: Far-UV CD spectrum and thermal denaturation curve of AtzB_Hom_Hal Q210S

The far-UV CD spectrum of AtzB_Hom_Hal Q210S before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) could not be fitted and no T_M values could be determined

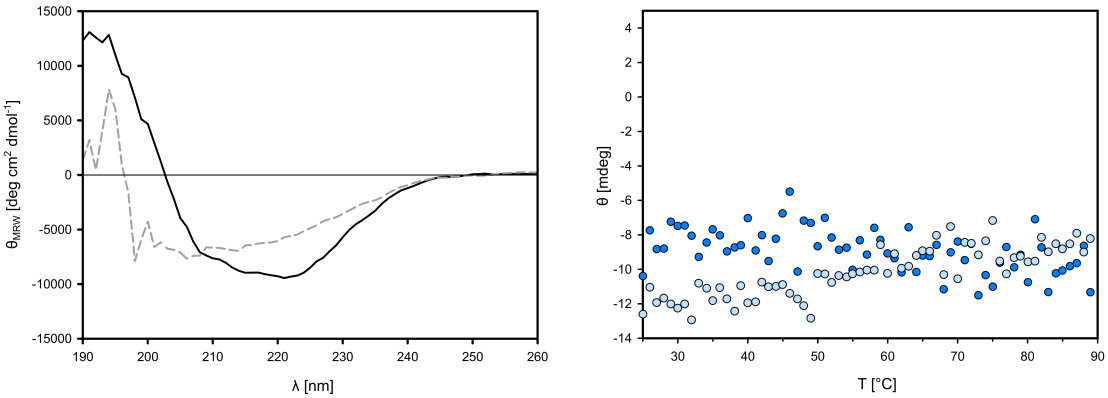


Figure 8.91: Far-UV CD spectrum and thermal denaturation curve of wild-type AtzB_Hom_Pleo

The far-UV CD spectrum of wild-type AtzB_Hom_Pleo before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) could not be fitted and no T_M values could be determined.

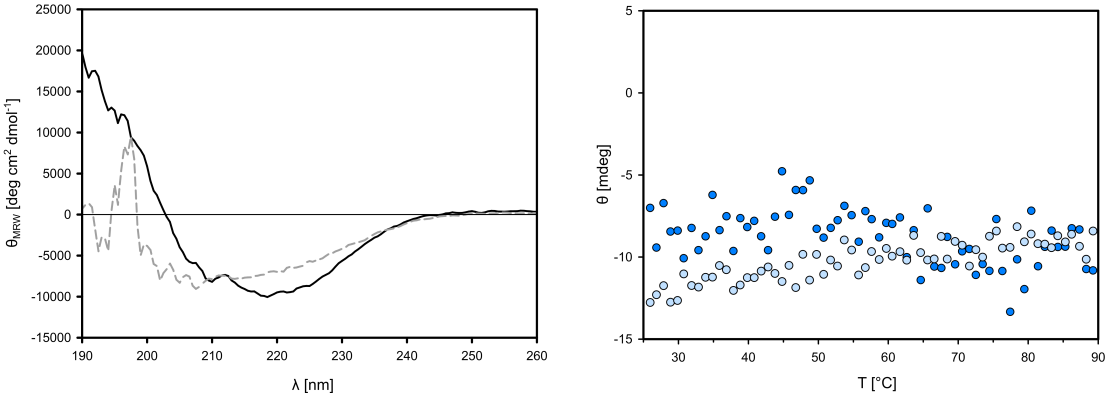


Figure 8.92: Far-UV CD spectrum and thermal denaturation curve of AtzB_Hom_Pleo C213S

The far-UV CD spectrum of AtzB_Hom_Pleo C213S before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) could not be fitted and no T_M values could be determined

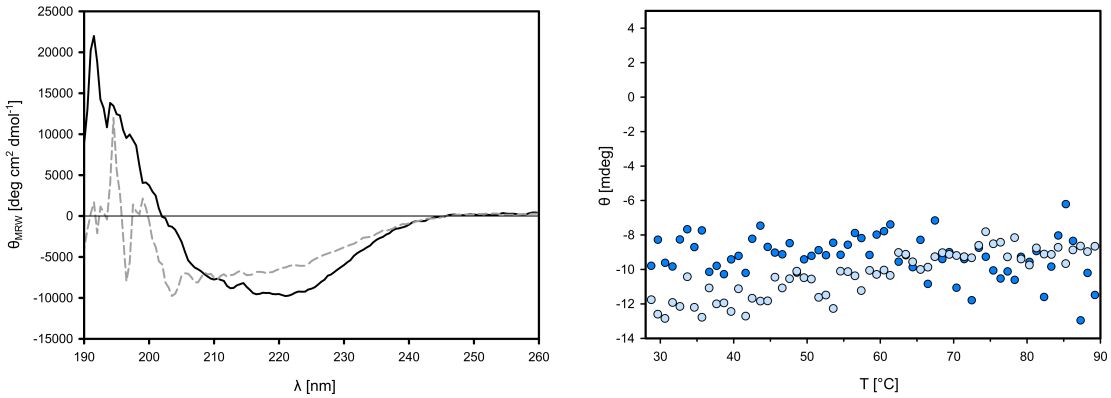


Figure 8.93: Far-UV CD spectrum and thermal denaturation curve of AtzB_Hom_Pleo Q214S
The far-UV CD spectrum of AtzB_Hom_Pleo Q214S before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) could not be fitted and no T_M values could be determined

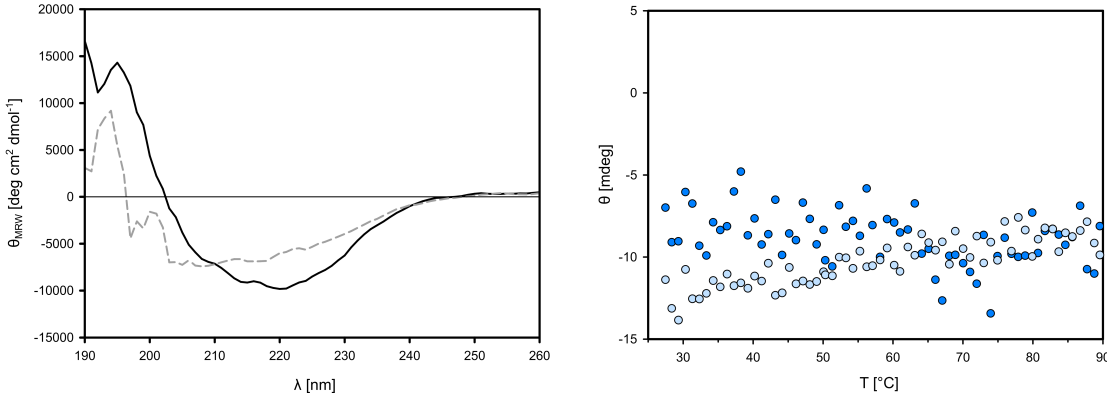


Figure 8.94: Far-UV CD spectrum and thermal denaturation curve of AtzB_Hom_Pleo C213S Q214S
The far-UV CD spectrum of AtzB_Hom_Pleo C213S Q214S before (black) and after (gray) thermal denaturation is shown on the left. The CD signals at 208 nm (dark blue) and 220 nm (light blue) could not be fitted and no T_M values could be determined

8.4.4 Steady-State Kinetics of AtzB Variants

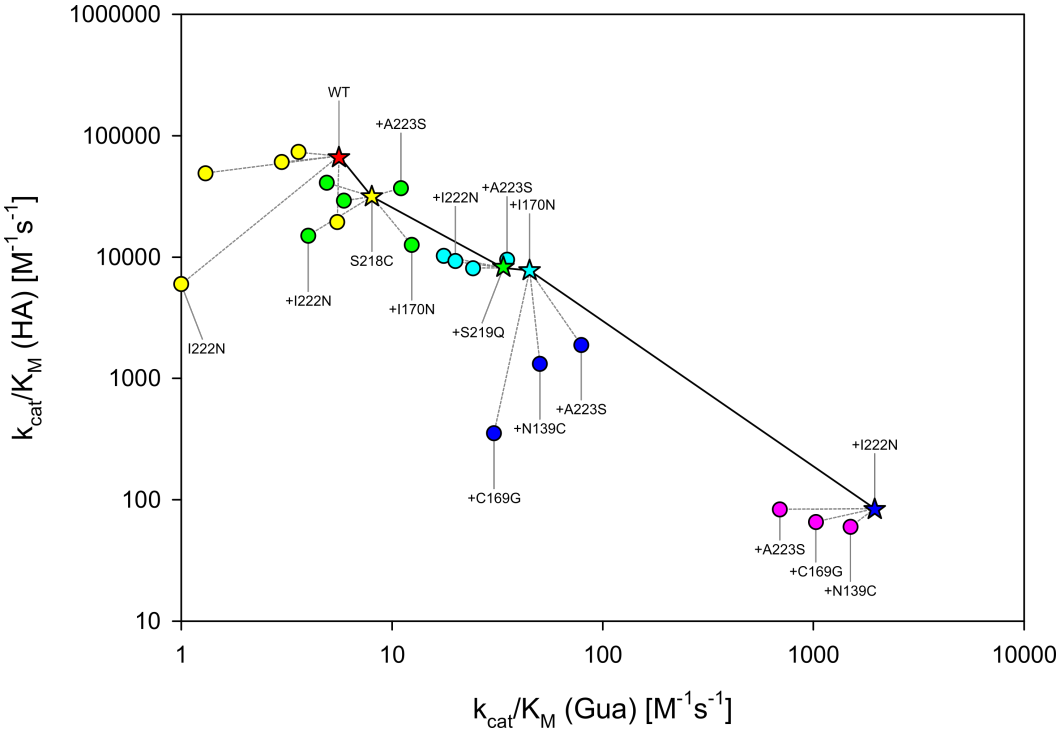


Figure 8.95: Screening of hydroxyatrazine deaminase activity and guanine deaminase activity of AtzB variants
Catalytic efficiency (k_{cat}/K_M) for hydroxyatrazine (HA) is plotted against k_{cat}/K_M for guanine (Gua) for all AtzB variants analyzed in this work. Values were determined by plate reader measurements. Most active guanine deaminase variants of each round of mutation are shown as stars. Other AtzB variants are indicated as circles, color-coded by the number of residue exchanges (0 = red; 1 = yellow; 2 = green; 3 = light blue; 4 = dark blue; 5 = purple). Variants separated by a single residue exchange are connected by dashed gray lines, while the main trajectory is indicated by solid black lines. Residue exchanges leading to relevant changes in activity are indicated.
Stepwise introduction of S218C, S219Q, and I170N leads to modest increases in Gua deaminase activity and significant reduction in HA hydrolase activity. The final addition of I222N causes a drastic increase in Gua deaminase activity and a strong drop-off in activity for HA. Introduction of N139C, C169G, or A223S into this quadruple variant produces reductions in Gua activity despite N139C and A223S slightly improving this activity in the S218C S219Q I170N background. A223S is further beneficial in the S218C and S218C-S219Q backgrounds. Notably, the residue exchange I222N is detrimental to guanine deaminase activity in all contexts except for S218C-S219Q-I170N. The activities of all permutations of the four functionally relevant residue exchanges (S218C, S219Q, I170N, and I222N) were confirmed by cuvette measurements (see Table 5.2).

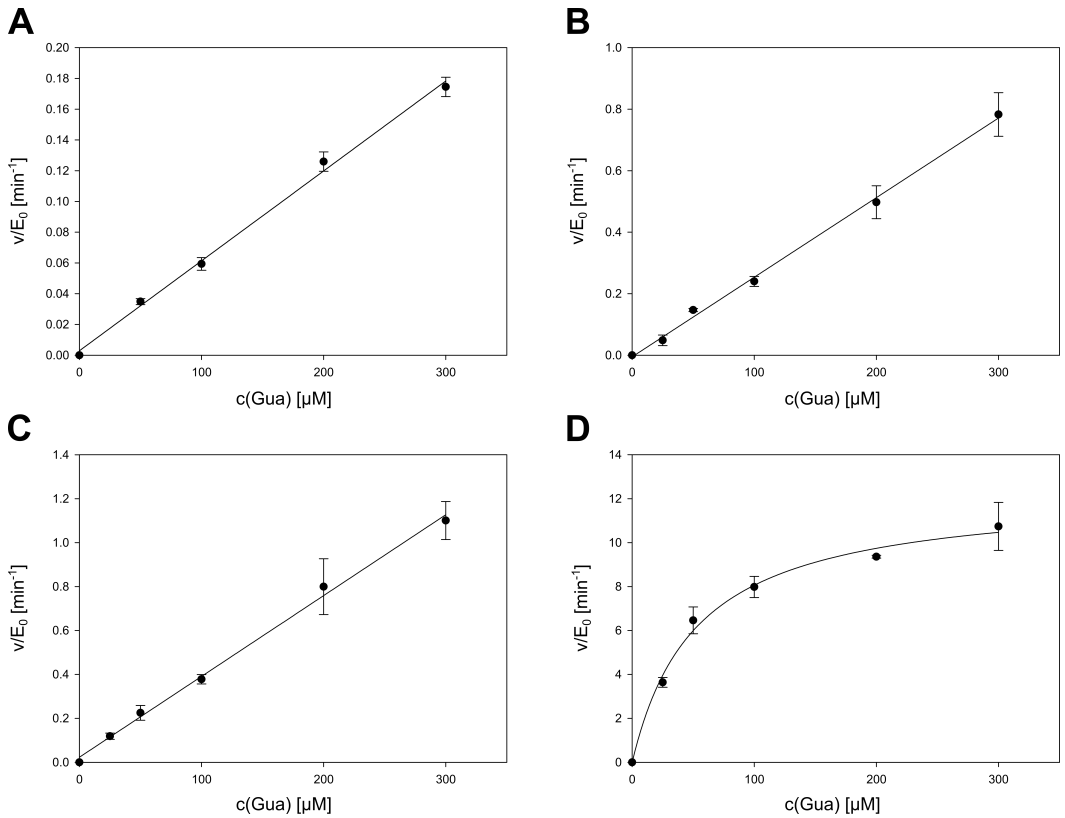


Figure 8.96: Steady-state kinetics for guanine of the variants of the AtzB trajectory

AtzB S218C (A), S218C S219Q (B), S218C S219Q I170N (C), and S218C S219Q I170N I222N (D) were reacted with guanine as described (see paragraph 2.5.4.2.3). Resulting kinetic parameters are given in Table 5.2.

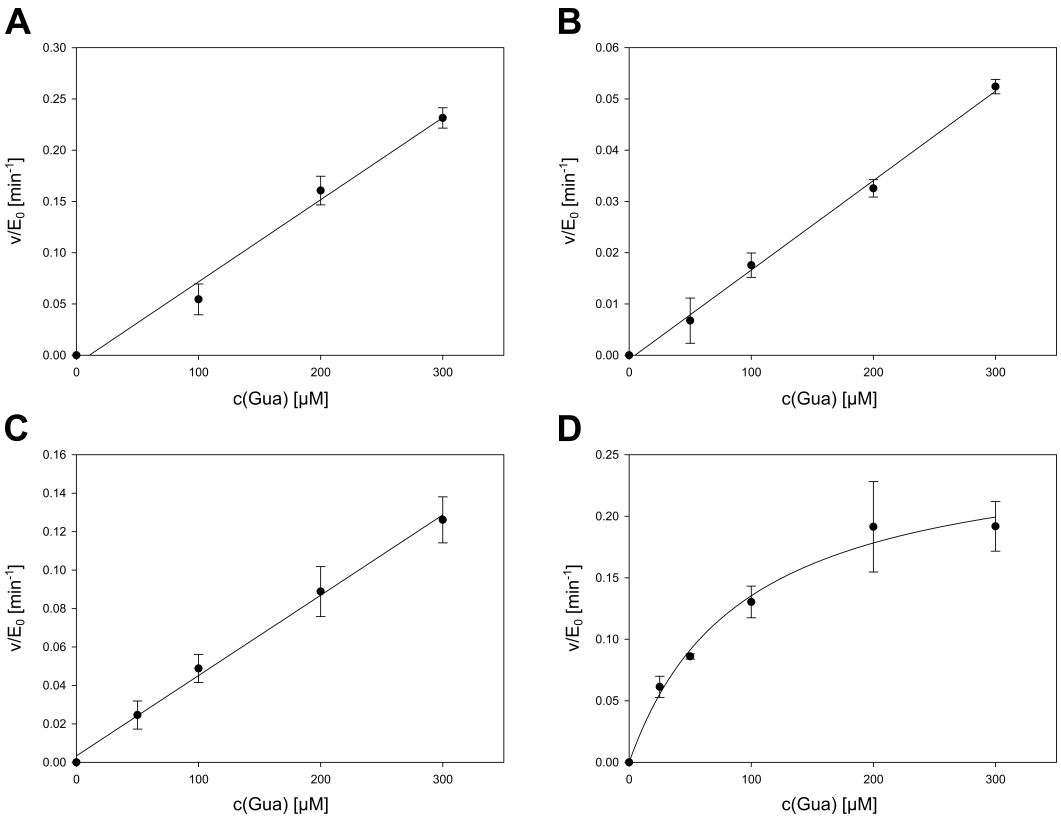


Figure 8.97: Steady-state kinetics for guanine of the remaining double-mutants of AtzB
AtzB S218C I170N (A), S218C I222N (B), I170N I222N (C), and S219Q I222N (D) were reacted with guanine as described (see paragraph 2.5.4.2.3). Resulting kinetic parameters are given in Table 5.2. AtzB I170N S219Q did not possess a measurable activity (not shown).

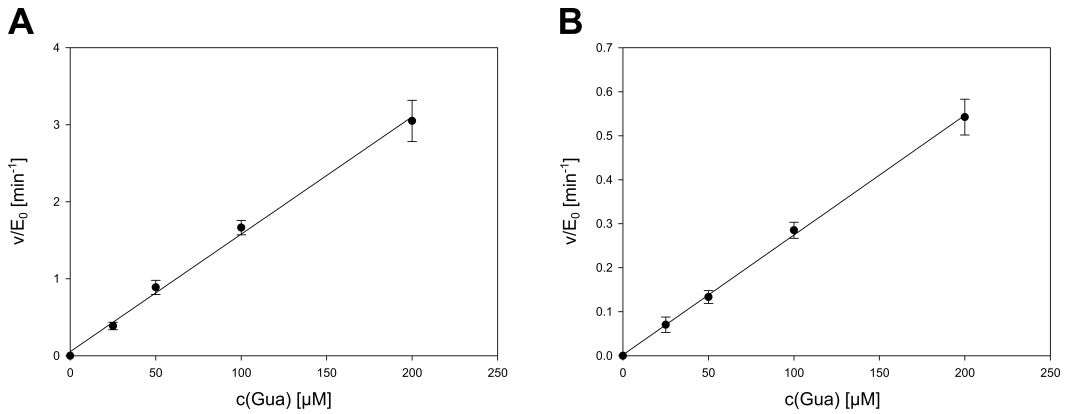


Figure 8.98: Steady-state kinetics for guanine of the remaining triple-mutants of AtzB

AtzB I170N S218C I222N (A) and I170N S219Q I222N (B) were reacted with guanine as described (see paragraph 2.5.4.2.3). Resulting kinetic parameters are given in Table 5.2. AtzB S218C S219Q I222N did not possess a measurable activity (not shown).

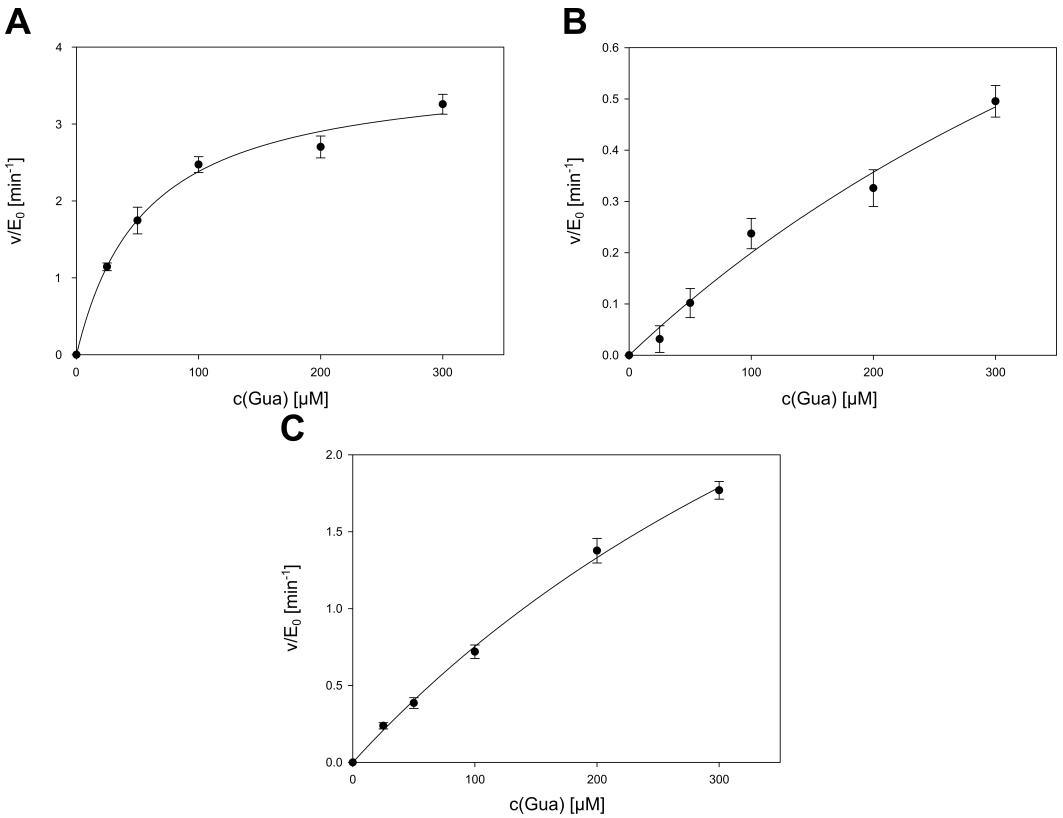


Figure 8.99: Steady-state kinetics for guanine of the variants of AtzB_Hom_Hal

Wild-type AtzB_Hom_Hal (A), AtzB_Hom_Hal C209S (B), and AtzB_Hom_Hal Q210S (C) were reacted with guanine as described (see paragraph 2.5.4.2.3). Resulting kinetic parameters are given in Table 5.2. AtzB_Hom_Hal C209S Q210S did not possess a measurable activity (not shown).

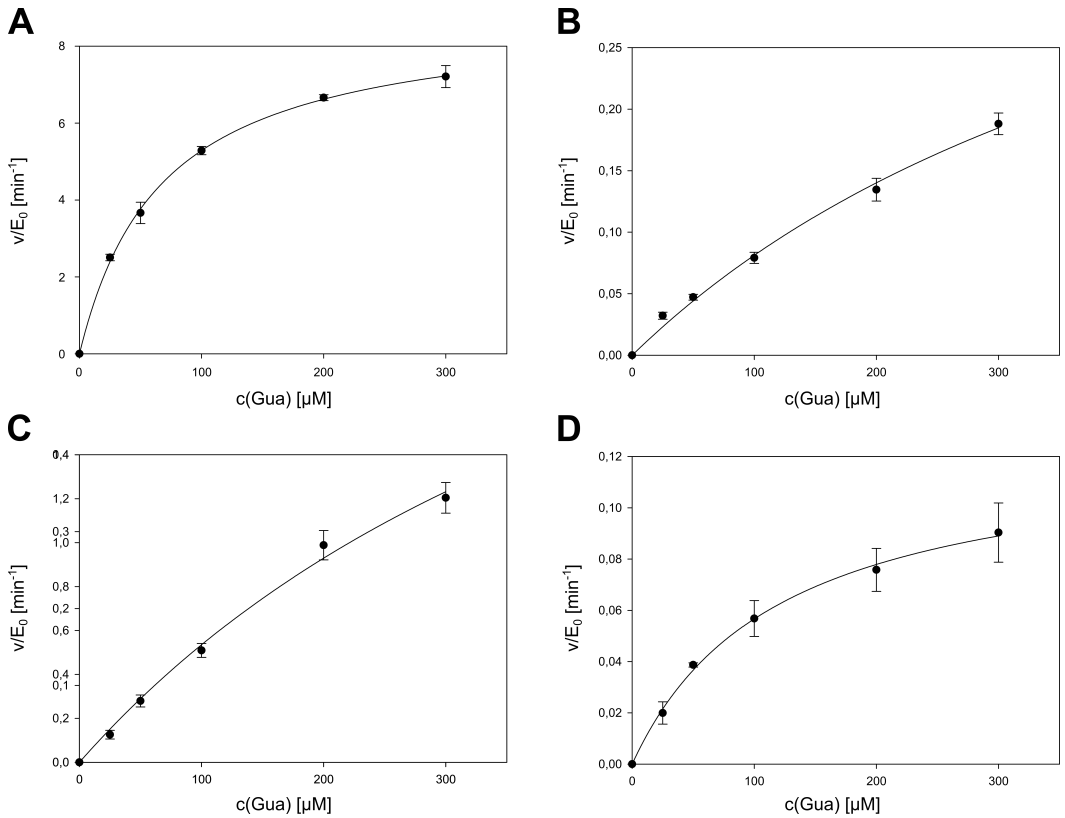


Figure 8.100: Steady-state kinetics for guanine of the variants of AtzB_Hom_Pleo

Wild-type AtzB_Hom_Pleo (A), AtzB_Hom_Pleo C213S (B), AtzB_Hom_Pleo Q214S (C), and AtzB_Hom_Pleo S213S Q214S (D) were reacted with guanine as described (see paragraph 2.5.4.2.3). Resulting kinetic parameters are given in Table 5.2.

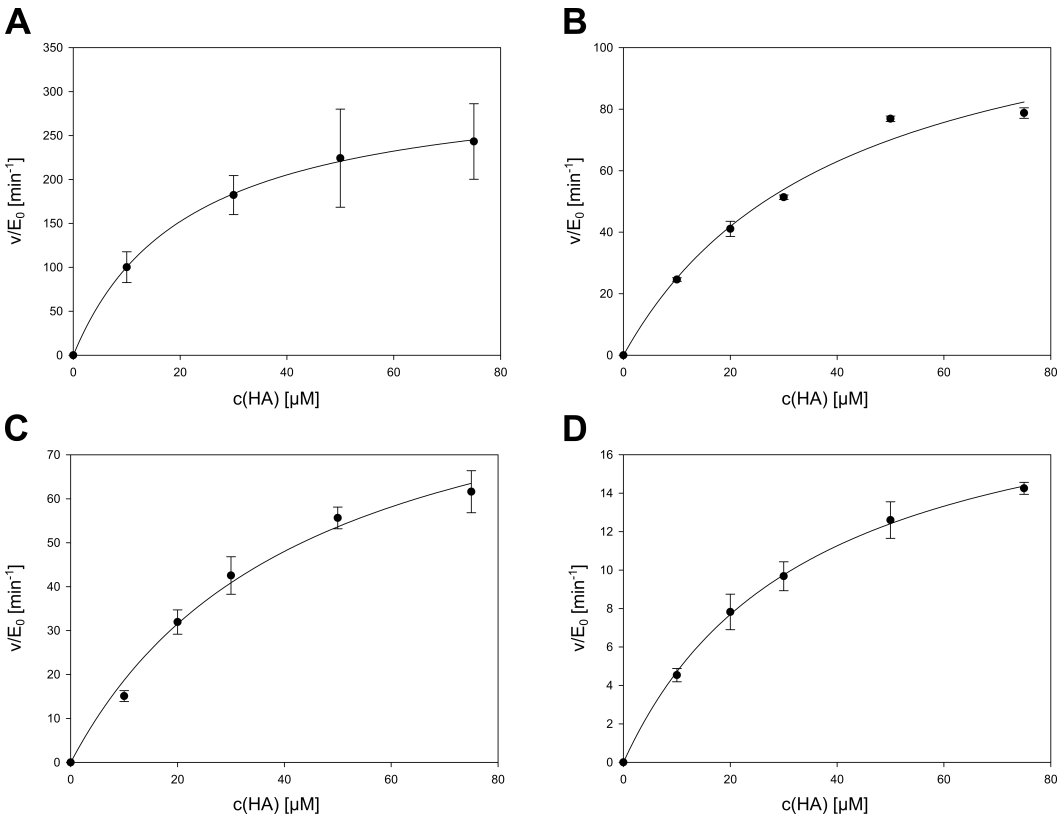


Figure 8.101: Steady-state kinetics for hydroxyatrazine of wild-type AtzB and variants S218C, S219Q, and I222N
Wild-type AtzB (A), and variants S218C (B), S219Q (C), and I222N (D) were reacted with hydroxyatrazine as described (see paragraph 2.5.4.2.1). Resulting kinetic parameters are given in Table 5.2.

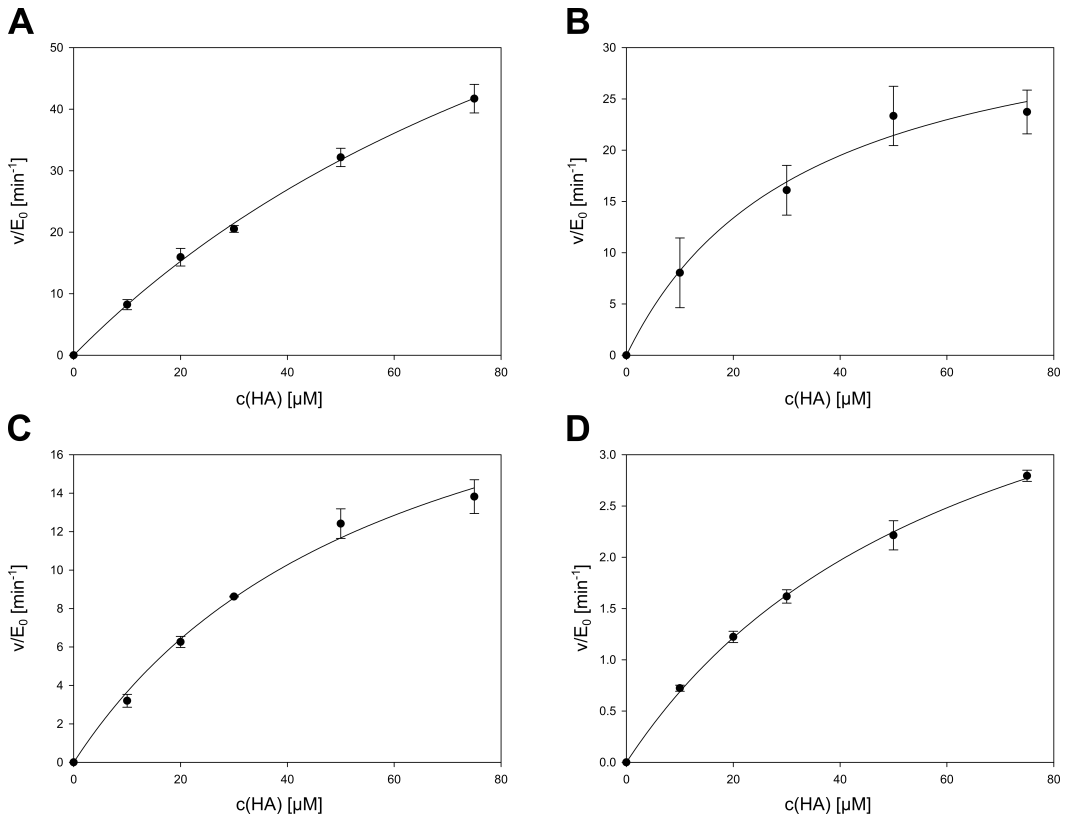


Figure 8.102: Steady-state kinetics for hydroxyatrazine of AtzB variants I170N S218C, S218C S219Q, S218C I222N, and I170N S219Q
AtzB variants I170N S218C (A), S218C S219Q (B), S218C I222N (C), and I170N S219Q (D) were reacted with hydroxyatrazine as described (see paragraph 2.5.4.2.1). Resulting kinetic parameters are given in Table 5.2.

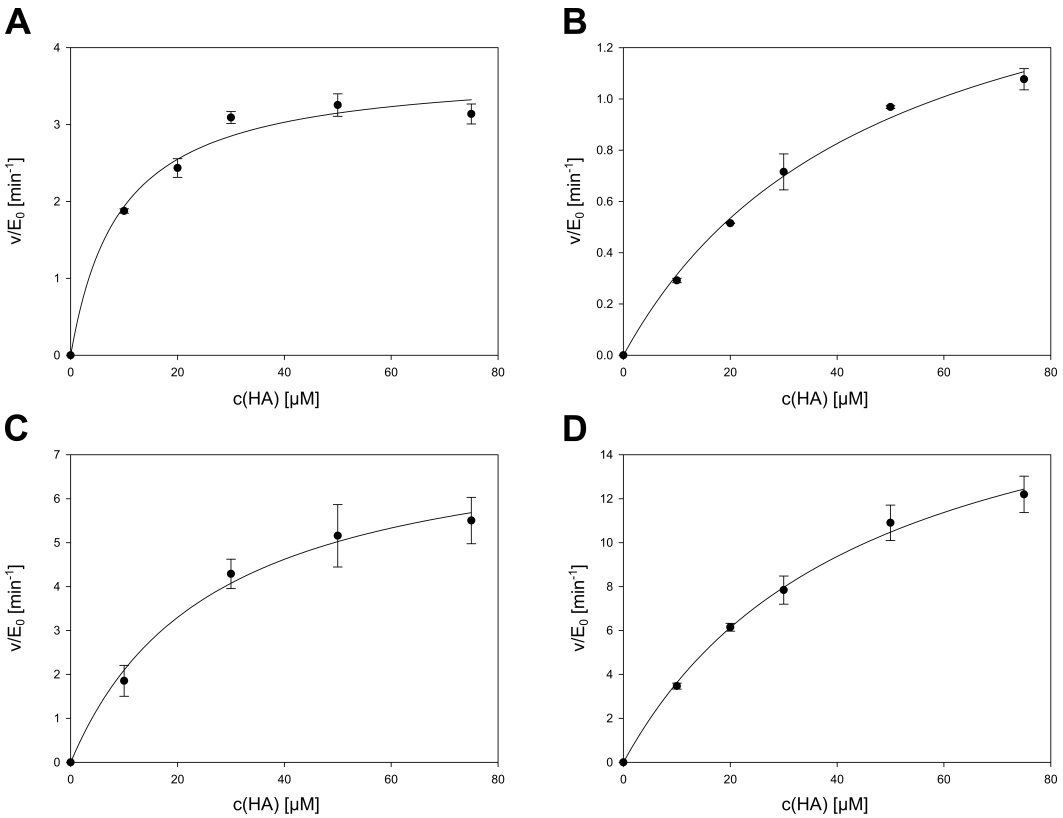


Figure 8.103: Steady-state kinetics for hydroxyatrazine of AtzB variants I170N I222N, S219Q I222N, S218C S219Q I170N, and S218C S219Q I222N
AtzB variants I170N I222N (A), S219Q I222N (B), S218C S219Q I170N (C), and S218C S219Q I222N (D) were reacted with hydroxyatrazine as described (see paragraph 2.5.4.2.1). Resulting kinetic parameters are given in Table 5.2.

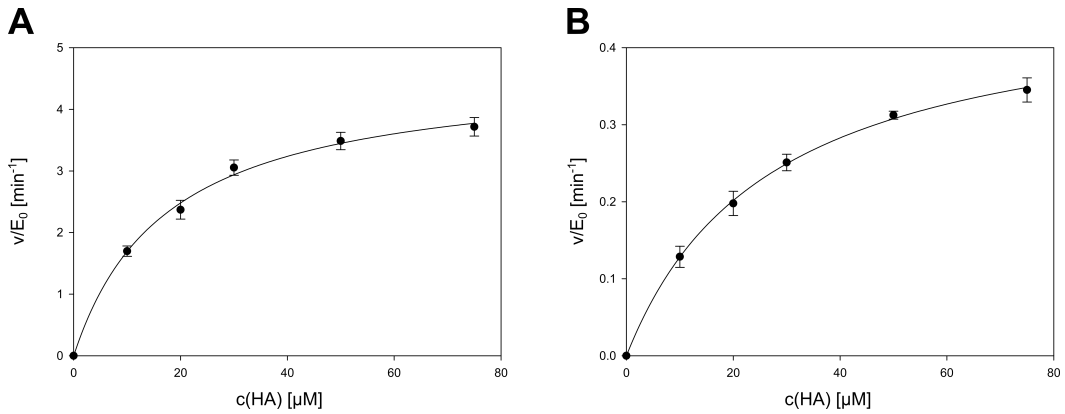


Figure 8.104: Steady-state kinetics for hydroxyatrazine of wild-type AtzB variants I170N S218C I222N and S218C S219Q I170N I222N
AtzB variants I170N S218C I222N (A) and S218C S219Q I170N I222N (B) were reacted with hydroxyatrazine as described (see paragraph 2.5.4.2.1). Resulting kinetic parameters are given in Table 5.2.

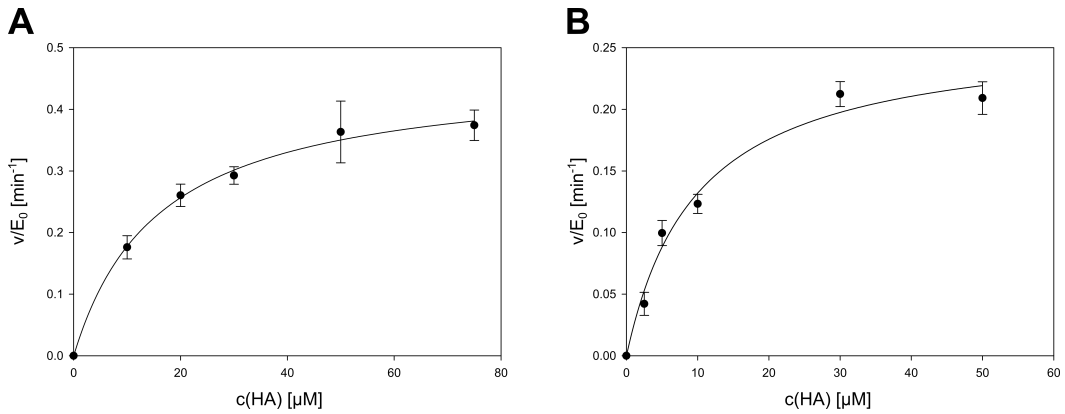


Figure 8.105: Steady-state kinetics for hydroxyatrazine of AtzB_Hom_Hal variants Q210S and C209S Q210S
AtzB_Hom_Hal Q210S (A) and AtzB_Hom_Hal C209S Q210S (B) were reacted with hydroxyatrazine as described (see paragraph 2.5.4.2.1). Resulting kinetic parameters are given in Table 5.2. Wild-type AtzB_Hom_Hal and AtzB_Hom_Hal C209S did not possess a measurable activity (not shown).

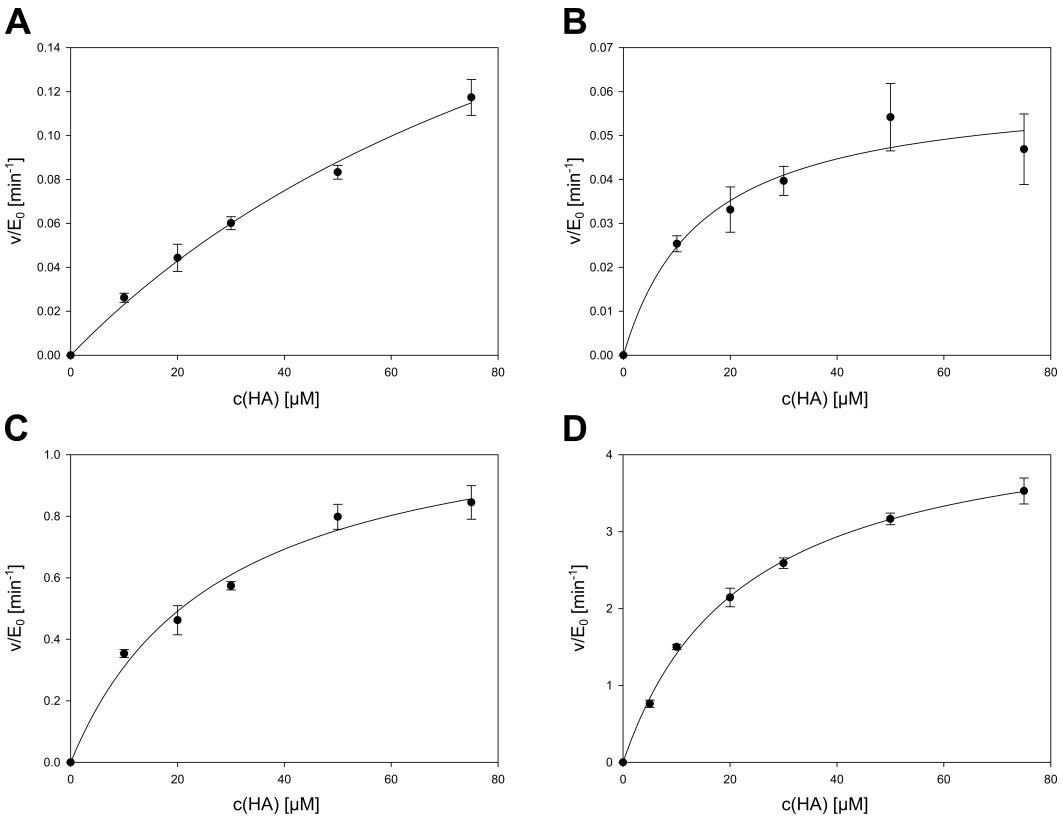


Figure 8.106: Steady-state kinetics for hydroxyatrazine of wild-type AtzB_Hom_Pleo and variants C213S, Q214S, and C213S Q214S
Wild-type AtzB_Hom_Pleo (A), and variants C213S (B), Q214S (C), and C213S Q214S (D) were reacted with hydroxyatrazine as described (see paragraph 2.5.4.2.1). Resulting kinetic parameters are given in Table 5.2.

8.4.5 Synthesis of N-Ethylammelide and N-Isopropylammelide

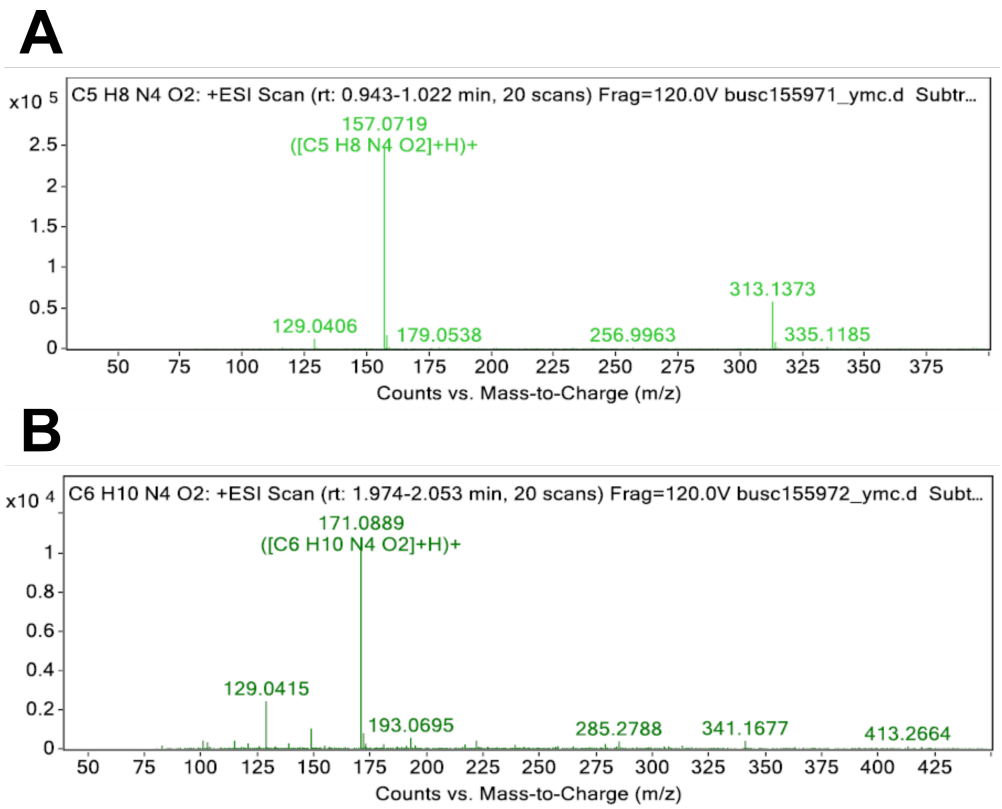


Figure 8.107: Mass spectrometry (MS) analysis of chemically synthesized N-ethylammelide and N-isopropylammelide
The reaction mixtures of N-ethylammelide and N-isopropylammelide were analyzed by HPLC-ESI-MS runs (see subsubsection 2.2.5.6). (A) The presumed HPLC peak of N-ethylammelide produced a primary MS product with a mass-to-charge ratio of 157.07 m/z, likely representing the protonated form of N-ethylammelide (156.14 Da). (B) The HPLC peak of N-isopropylammelide produced a main MS peak with 171.09 m/z, indicative of the protonated form of N-isopropylammelide (170.17 Da).

Bibliography

[1] Jing Cai, Ruoping Zhao, Huifeng Jiang, and Wen Wang. De novo origination of a new protein-coding gene in Saccharomyces cerevisiae. *Genetics*, 179(1):487–496, 2008.

[2] David G. Knowles and Aoife McLysaght. Recent de novo origin of human protein-coding genes. *Genome Research*, 19(10):1752–1759, 2009.

[3] Barth F. Smets and Tamar Barkay. Horizontal gene transfer: perspectives at a crossroads of scientific disciplines. *Nature Reviews Microbiology*, 3(9):675–678, 2005.

[4] Christian Roth, Shruti Rastogi, Lars Arvestad, Katharina Dittmar, Sara Light, Diana Ekman, and David A. Liberles. Evolution after gene duplication: models, mechanisms, sequences, systems, and organisms. *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, 308(1):58–73, 2007.

[5] In-Geol Choi and Sung-Hou Kim. Evolution of protein structural classes and protein sequence families. *Proceedings of the National Academy of Sciences*, 103(38):14056–14061, 2006.

[6] Norman H. Horowitz. On the evolution of biochemical syntheses. *Proceedings of the National Academy of Sciences of the United States of America*, 31(6):153–157, 1945.

[7] Susumu Ohno. *Evolution by gene duplication*. Springer, New York, 1970.

[8] Roy A. Jensen. Enzyme recruitment in evolution of new function. *Annual Review of Microbiology*, 30:409–425, 1976.

[9] Ulfar Bergthorsson, Dan I. Andersson, and John R. Roth. Ohno's dilemma: Evolution of new genes under continuous selection. *Proceedings of the National Academy of Sciences*, 104(43):17004–17009, 2007.

[10] Olga Khersonsky and Dan S. Tawfik. Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annual Review of Biochemistry*, 79:471–505, 2010.

[11] Shelley D. Copley. Moonlighting is mainstream: paradigm adjustment required. *BioEssays*, 34(7):578–588, 2012.

[12] Shelley D Copley. Shining a light on enzyme promiscuity. *Current Opinion in Structural Biology*, 47:167–175, 2017.

[13] Vaishali Katju and Ulfar Bergthorsson. Copy-number changes in evolution: rates, fitness effects and adaptive significance. *Frontiers in Genetics*, 4:273, 2013.

[14] Andrew B. Reams and John R. Roth. Mechanisms of gene duplication and amplification. *Cold Spring Harbor Perspectives in Biology*, 7(2):a016592, 2015.

[15] Dan I. Andersson, Jon Jerlström-Hultqvist, and Joakim Näsvall. Evolution of new functions *de novo* and from preexisting genes. *Cold Spring Harbor Perspectives in Biology*, 7(6):a017996, 2015.

[16] Linus Sandegren and Dan I. Andersson. Bacterial gene amplification: implications for the evolution of antibiotic resistance. *Nature Reviews Microbiology*, 7(8):578–588, 2009.

[17] Michael Lynch and Alan Force. The probability of duplicate gene preservation by subfunctionalization. *Genetics*, 154(1):459–473, 2000.

[18] John A. Gerlt and Patricia C. Babbitt. Divergent evolution of enzymatic function: mechanistically diverse superfamilies and functionally distinct suprafamilies. *Annual Review of Biochemistry*, 70:209–246, 2001.

[19] Nicholas Furnham, Natalie L. Dawson, Syed A. Rahman, Janet M. Thornton, and Christine A. Orengo. Large-Scale Analysis Exploring Evolution of Catalytic Machineries and Mechanisms in Enzyme Superfamilies. *Journal of Molecular Biology*, 428(2 Pt A):253–267, 2016.

[20] Markus Busch. Attempts to establish a secondary metabolic enzyme activity on the protein scaffold of a primary metabolic enzyme. *Master's Thesis, University of Regensburg*, 2017.

[21] Caroline Hiefinger. Biochemical analysis of the evolutionary origin of enzymes from striazine metabolism. *Master's Thesis, University of Regensburg*, 2020.

[22] Corinna Grisostomi, Peter Kast, Rosalino Pulido, Judy Huynh, and Donald Hilvert. Efficient *in vivo* synthesis and rapid purification of chorismic acid using an engineered *Escherichia coli* strain. *Bioorganic Chemistry*, 25(5):297–305, 1997.

[23] Florian Hubrich, Michael Müller, and Jennifer N. Andexer. *In vitro* production and purification of isochorismate using a two-enzyme cascade. *Journal of Biotechnology*, 191:93–98, 2014.

[24] Olga Khersonsky, Rosalie Lipsh, Ziv Avizemer, Yacov Ashani, Moshe Goldsmith, Haim Leader, Orly Dym, Shelly Rogotner, Devin L. Trudeau, Jaime Prilusky, Pep Amengual-Rigo, Victor Guallar, Dan S. Tawfik, and Sarel J. Fleishman. Automated design of efficient and functionally diverse enzyme repertoires. *Molecular Cell*, 72(1):178–186.e5, 2018.

[25] Trevor A. Addington, Robert W. Mertz, Justin B. Siegel, James M. Thompson, Andrew J. Fisher, Vladimir Filkov, Nicholas Fleischman, Alisa Suen, Chensong Zhang, and Michael D. Toney. JANUS: prediction and ranking of mutations required for functional interconversion of enzymes. *Journal of Molecular Biology*, 425(8):1378–1389, 2013.

[26] Bettina Rohweder, Florian Semmelmann, Christiane Endres, and Reinhard Sterner. Standardized cloning vectors for protein production and generation of large gene libraries in Escherichia coli. *BioTechniques*, 64(1):24–26, 2018.

[27] Kirill A. Datsenko and Barry L. Wanner. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America*, 97(12):6640–6645, 2000.

[28] Daiguan Yu, Hilary M. Ellis, E-Chiang Lee, Nancy A. Jenkins, Neal G. Copeland, and Donald L. Court. An efficient recombination system for chromosome engineering in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 97(11):5978–5983, 2000.

[29] Tomoya Baba, Takeshi Ara, Miki Hasegawa, Yuki Takai, Yoshiko Okumura, Miki Baba, Kirill A. Datsenko, Masaru Tomita, Barry L. Wanner, and Hirotada Mori. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*, 2:2006.0008, 2006.

[30] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, Martin Steinegger, Michalina Pacholska, Tamas Berghammer, Sebastian Bodenstein, David Silver, Oriol Vinyals, Andrew W. Senior, Koray Kavukcuoglu, Pushmeet Kohli, and Demis Hassabis. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, 2021.

[31] Paul Shannon, Andrew Markiel, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11):2498–2504, 2003.

[32] Stephen F. Altschul, Warren Gish, Webb Miller, Eugene W. Myers, and David J. Lipman. Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410, 1990.

[33] Wei Tian, Chang Chen, Xue Lei, Jieling Zhao, and Jie Liang. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Research*, 46(W1):W363–W367, 2018.

[34] Liisa Holm. Using Dali for protein structure comparison. *Methods in Molecular Biology*, 2112:29–42, 2020.

[35] Fabian Sievers and Desmond G. Higgins. The Clustal Omega Multiple Alignment Package. *Methods in Molecular Biology*, 2231:3–16, 2021.

[36] John A. Gerlt, Jason T. Bouvier, Daniel B. Davidson, Heidi J. Imker, Boris Sadkhin, David R. Slater, and Katie L. Whalen. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochimica Et Biophysica Acta*, 1854(8):1019–1037, 2015.

[37] Elisabeth Gasteiger, Alexandre Gattiker, Christine Hoogland, Ivan Ivanyi, Ron D. Appel, and Amos Bairoch. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, 31(13):3784–3788, 2003.

[38] Evgeny Krissinel and Kim Henrick. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology*, 372(3):774–797, 2007.

[39] Gavin E. Crooks, Gary Hon, John-Marc Chandonia, and Steven E. Brenner. WebLogo: a sequence logo generator. *Genome Research*, 14(6):1188–1190, 2004.

[40] Kary Mullis, Fred Faloona, Stephen Scharf, Randall Saiki, Glenn Horn, and Henry Erlich. Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology*, 51 Pt 1:263–273, 1986.

[41] Wenyan Wang and Bruce A Malcolm. Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange site-directed mutagenesis. *BioTechniques*, 26(4):680–682, 1999.

[42] Anna Gąciarz, Narendar Kumar Khatri, M. Lourdes Velez-Suberbie, Mirva J. Saaranen, Yuko Uchida, Eli Keshavarz-Moore, and Lloyd W. Ruddock. Efficient soluble expression of disulfide bonded proteins in the cytoplasm of *Escherichia coli* in fed-batch fermentations on chemically defined minimal media. *Microbial Cell Factories*, 16(1):108, 2017.

[43] Norma J. Greenfield. Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols*, 1(6):2876–2890, 2006.

[44] Lia Addadi, Eileen K. Jaffe, and Jeremy R. Knowles. Secondary tritium isotope effects as probes of the enzymic and nonenzymic conversion of chorismate to prephenate. *Biochemistry*, 22(19):4494–4501, 1983.

[45] Antje Chang, Lisa Jeske, Sandra Ulbrich, Julia Hofmann, Julia Koblitz, Ida Schomburg, Meina Neumann-Schaal, Dieter Jahn, and Dietmar Schomburg. BRENDA, the ELIXIR core data resource in 2021: new developments and updates. *Nucleic Acids Research*, 49(D1):D498–D508, 2021.

[46] I. G. Young, L. Langman, R. K. J. Luke, and F. Gibson. Biosynthesis of the iron-transport compound enterochelin: mutants of *Escherichia coli* unable to synthesize 2,3-dihydroxybenzoate. *Journal of Bacteriology*, 106(1):51–57, 1971.

[47] Kathleen M. Meneely, Jesse A. Sundlov, Andrew M. Gulick, Graham R. Moran, and Audrey L. Lamb. An open and shut case: the interaction of magnesium with MST enzymes. *Journal of the American Chemical Society*, 138(29):9277–9293, 2016.

[48] Tathagata Mukherjee, Yang Zhang, Sameh Abdelwahed, Steven E. Ealick, and Tadhg P. Begley. Catalysis of a flavoenzyme-mediated amide hydrolysis. *Journal of the American Chemical Society*, 132(16):5550–5551, 2010.

[49] Stephen S. Washburne, W. R. Peterson, and Dennis A. Berman. Reaction of trimethylsilyl azide with anhydrides and imides. Uracil synthesis via nitrogen insertion. *The Journal of Organic Chemistry*, 37(11):1738–1742, 1972.

[50] Gretchen M. Rehberg and Brian M. Glass. Improved synthesis of 2h-1,3(3h)-oxazine-2,6-dione (oxauracil). *Organic Preparations and Procedures International*, 27(6):651–652, 1995.

[51] Jiří Farkaš, Jaroslav Hapala, Olga Jindrová, and Jan Škoda. Reaction of 2,3-dihydro-1,3-6H-oxazine-2,6-dione with aliphatic amines and amino acids. *Collection of Czechoslovak Chemical Communications*, 47(11):2932–2945, 1982.

[52] Kwang-Seo Kim, Jeffrey G. Pelton, William B. Inwood, Ulla Andersen, Sydney Kustu, and David E. Wemmer. The Rut pathway for pyrimidine degradation: novel chemistry and toxicity problems. *Journal of Bacteriology*, 192(16):4089–4102, 2010.

[53] Romelia Salomon-Ferrer, David A. Case, and Ross C. Walker. An overview of the Amber biomolecular simulation package. *WIREs Computational Molecular Science*, 3(2):198–210, 2013.

[54] Albrecht Kossel. Über die chemische Zusammensetzung der Zelle. *Archiv für Physiologie*, VI:181–186, 1891.

[55] Renato Fani and Marco Fondi. Origin and evolution of metabolic pathways. *Physics of Life Reviews*, 6(1):23–52, 2009.

[56] Madeline C. Weiss, Filipa L. Sousa, Natalia Mrnjavac, Sinje Neukirchen, Mayo Roettger, Shijulal Nelson-Sathi, and William F. Martin. The physiology and habitat of the last universal common ancestor. *Nature Microbiology*, 1(9):1–8, 2016.

[57] Madeline C. Weiss, Martina Preiner, Joana C. Xavier, Verena Zimorski, and William F. Martin. The last universal common ancestor between ancient Earth chemistry and the onset of genetics. *PLoS Genetics*, 14(8):e1007518, 2018.

[58] Walaa Kamel Mousa and Manish N. Raizada. The diversity of anti-microbial secondary metabolites produced by fungal endophytes: an interdisciplinary perspective. *Frontiers in Microbiology*, 4, 2013.

[59] V. Betina. Biological effects of the antibiotic brefeldin A (decumbin, cyanein, ascotoxin, synergisidin): a retrospective. *Folia Microbiologica*, 37(1):3–11, 1992.

[60] Chandrabhan Seniya, Harshal Mishra, Ajay Yadav, Nitin Sagar, Babita Chaturvedi, Kuldeep Uchadia, and Gulshan Wadhwa. Antiviral potential of 4-hydroxypanduratin A, secondary metabolite of fingerroot, *Boesenbergia pandurata* (Schult.), towards Japanese encephalitis virus NS2B/NS3 protease. *Bioinformation*, 9(1):54–60, 2013.

[61] Tünde Pusztahelyi, Imre J. Holb, and István Pócsi. Secondary metabolites in fungus-plant interactions. *Frontiers in Plant Science*, 6, 2015.

[62] Peter Proksch. Pflanzliche Sekundärstoffe als chemischer Fraßschutz gegen herbivore Insekten. *Pharmazie in unserer Zeit*, 20(5):217–224, 1991.

[63] John W. Daly, H. Martin Garraffo, Thomas F. Spande, Michael W. Decker, James P. Sullivan, and Michael Williams. Alkaloids from frog skin: the discovery of epibatidine and the potential for developing novel non-opioid analgesics. *Natural Product Reports*, 17(2):131–135, 2000.

[64] Martin Eisendle, Harald Oberegger, Ivo Zadra, and Hubertus Haas. The siderophore system is essential for viability of *Aspergillus nidulans*: functional analysis of two genes encoding l-ornithine N 5-monooxygenase (*sidA*) and a non-ribosomal peptide synthetase (*sidC*). *Molecular Microbiology*, 49(2):359–375, 2003.

[65] Ramakrishnan Balasubramanian and Amy C. Rosenzweig. Copper methanobactin: a molecule whose time has come. *Current Opinion in Chemical Biology*, 12(2):245–249, 2008.

[66] Simon Irmer, Nora Podzun, Dorothee Langel, Franziska Heidemann, Elisabeth Kaltenegger, Brigitte Schemmerling, Christoph-Martin Geilfus, Christian Zörb, and Dietrich Ober. New aspect of plant–rhizobia interaction: alkaloid biosynthesis in *Crotalaria* depends on nodulation. *Proceedings of the National Academy of Sciences of the United States of America*, 112(13):4164–4169, 2015.

[67] Jinichiro Koga, Masaru Shimura, Kiyomi Oshima, Noriko Ogawa, Toyozo Yamauchi, and Nagahiro Ogasawara. Phytocassanes A, B, C and D, novel diterpene phytoalexins from rice, *Oryza sativa* L. *Tetrahedron*, 51(29):7907–7918, 1995.

[68] Ralf Horbach, Aura Rocio Navarro-Quesada, Wolfgang Knogge, and Holger B. Deising. When and how to kill a plant cell: infection strategies of plant pathogenic fungi. *Journal of Plant Physiology*, 168(1):51–62, 2011.

[69] Wolfgang Wohlleben, Yvonne Mast, Evi Stegmann, and Nadine Ziemert. Antibiotic drug discovery. *Microbial Biotechnology*, 9(5):541–548, 2016.

[70] Paul R. Shipley, Caitlin C. A. Donnelly, Cuong H. Le, Ashley D. Bernauer, and Andis Klegeris. Antitumor activity of asukamycin, a secondary metabolite from the actinomycete bacterium *Streptomyces nodosus* subspecies *asukaensis*. *International Journal of Molecular Medicine*, 24(5):711–715, 2009.

[71] Supriya Baikar and Nutan Malpathak. Secondary metabolites as DNA topoisomerase inhibitors: A new era towards designing of anticancer drugs. *Pharmacognosy Reviews*, 4(7):12–26, 2010.

[72] Edmund Stone. An account of the success of the bark of the willow in the cure of agues. In a letter to the right honourable George Earl of Macclesfield, President of R. S. from the Rev. Mr. Edmund Stone, of Chipping-Norton in Oxfordshire. *Philosophical Transactions*, 53:195–200, 1763.

[73] Diarmuid Jeffreys. *Aspirin: The Remarkable Story of a Wonder Drug*. Bloomsbury Publishing, 2008.

[74] Michael A. Fischbach, Christopher T. Walsh, and Jon Clardy. The evolution of gene collectives: How natural selection drives chemical innovation. *Proceedings of the National Academy of Sciences of the United States of America*, 105(12):4601–4608, 2008.

[75] Gaurav D. Moghe and Robert L. Last. Something old, something new: conserved enzymes and the evolution of novelty in plant specialized metabolism. *Plant Physiology*, 169(3):1512–1523, 2015.

[76] Arren Bar-Even, Elad Noor, Yonatan Savir, Wolfram Liebermeister, Dan Davidi, Dan S. Tawfik, and Ron Milo. The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry*, 50(21):4402–4410, 2011.

[77] Richard D. Firn and Clive G. Jones. The evolution of secondary metabolism - a unifying model. *Molecular Microbiology*, 37(5):989–994, 2000.

[78] Michael A. Fischbach and Christopher T. Walsh. Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. *Chemical Reviews*, 106(8):3468–3496, 2006.

[79] T. Cavalier-Smith. Origins of secondary metabolism. *Ciba Foundation Symposium*, 171:64–80; discussion 80–87, 1992.

[80] Yuriko Carrington, Jia Guo, Cuong H. Le, Alexander Fillo, Junsu Kwon, Lan T. Tran, and Jürgen Ehlting. Evolution of a secondary metabolic pathway from primary metabolism: shikimate and quinate biosynthesis in plants. *The Plant Journal*, 95(5):823–833, 2018.

[81] Meimei Xu, P. Ross Wilderman, and Reuben J. Peters. Following evolution's lead to a single residue switch for diterpene synthase product outcome. *Proceedings of the National Academy of Sciences of the United States of America*, 104(18):7397–7401, 2007.

[82] Maximilian G. Plach, Patrick Löffler, Rainer Merkl, and Reinhard Sterner. Conversion of anthranilate synthase into isochorismate synthase: implications for the evolution of chorismate-utilizing enzymes. *Angewandte Chemie (International Ed. in English)*, 54(38):11270–11274, 2015.

[83] Simon C. Andrews, Andrea K. Robinson, and Francisco Rodríguez-Quiñones. Bacterial iron homeostasis. *FEMS microbiology reviews*, 27(2-3):215–237, 2003.

[84] Kenneth N. Raymond and Carl J. Carrano. Coordination chemistry and microbial iron transport. *Accounts of Chemical Research*, 12(5):183–190, 1979.

[85] Volkmar Braun and Helmut Killmann. Bacterial solutions to the iron-supply problem. *Trends in Biochemical Sciences*, 24(3):104–109, 1999.

[86] Robert C. Hider and Xiaole Kong. Chemistry and biology of siderophores. *Natural Product Reports*, 27(5):637–657, 2010.

[87] Eugene D. Weinberg. Iron and infection. *Microbiological Reviews*, 42(1):45–66, 1978.

[88] Ketil Thorstensen and Inge Romslo. The role of transferrin in the mechanism of cellular iron uptake. *Biochemical Journal*, 271(1):1–9, 1990.

[89] Ulrich E. Schaible and Stefan H. E. Kaufmann. Iron and microbial infection. *Nature Reviews Microbiology*, 2(12):946–953, 2004.

[90] Sandro Silva-Gomes, Sílvia Vale-Costa, Rui Appelberg, and Maria S. Gomes. Iron in intracellular infection: to provide or to deprive? *Frontiers in Cellular and Infection Microbiology*, 3, 2013.

[91] Joy R. Paterson, Marikka S. Beecroft, Raminder S. Mulla, Deenah Osman, Nancy L. Reeder, Justin A. Caserta, Tessa R. Young, Charles A. Pettigrew, Gareth E. Davies, J. A. Gareth Williams, and Gary J. Sharples. Insights into the antibacterial mechanism of action of chelating agents by selective deprivation of iron, manganese, and zinc. *Applied and Environmental Microbiology*, 88(2):e0164121, 2022.

[92] Hubertus Haas, Martin Eisendle, and B. Gillian Turgeon. Siderophores in fungal physiology and virulence. *Annual Review of Phytopathology*, 46:149–187, 2008.

[93] Volker Römheld and Horst Marschner. Evidence for a Specific Uptake System for Iron Phytosiderophores in Roots of Grasses 1. *Plant Physiology*, 80(1):175–180, 1986.

[94] Stephan M. Kraemer, D. E. Crowley, and Ruben Kretzschmar. Geochemical aspects of phytosiderophore-promoted iron acquisition by plants. In *Advances in Agronomy*, volume 91, pages 1–46. Academic Press, 2006.

[95] I. G. O'Brien and F. Gibson. The structure of enterochelin and related 2,3-dihydroxy-N-benzoyne conjugates from *Eschericha coli*. *Biochimica et Biophysica Acta - General Subjects*, 215(2):393–402, 1970.

[96] J. R. Pollack and J. B. Neilands. Enterobactin, an iron transport compound from Salmonella typhimurium. *Biochemical and Biophysical Research Communications*, 38(5):989–992, 1970.

[97] Hans-Peter Fiedler, Philipp Krastel, Johannes Müller, Klaus Gebhardt, and Axel Zeeck. Enterobactin: the characteristic catecholate siderophore of Enterobacteriaceae is produced by Streptomyces species.(1). *FEMS microbiology letters*, 196(2):147–151, 2001.

[98] Filip Dosselaere and Jozef Vanderleyden. A metabolic node in action: chorismate-utilizing enzymes in microorganisms. *Critical Reviews in Microbiology*, 27(2):75–131, 2001.

[99] Tobias Sahr, Stéphane Ravanel, Gilles Basset, Brian P. Nichols, Andrew D. Hanson, and Fabrice Rébeillé. Folate synthesis in plants: purification, kinetic properties, and inhibition of aminodeoxychorismate synthase. *Biochemical Journal*, 396(Pt 1):157–162, 2006.

[100] Ignacio Lopez-Goñi, Ignacio Moriyon, and JB Neilands. Identification of 2,3-dihydroxybenzoic acid as a *Brucella abortus* siderophore. *Infection and Immunity*, 60(11):4496–4503, 1992.

[101] Kenneth N. Raymond, Emily A. Dertz, and Sanggoo S. Kim. Enterobactin: an archetype for microbial iron transport. *Proceedings of the National Academy of Sciences of the United States of America*, 100(7):3584–3588, 2003.

[102] A. M. Gehring, K. A. Bradley, and C. T. Walsh. Enterobactin biosynthesis in *Escherichia coli*: isochorismate lyase (EntB) is a bifunctional enzyme that is phosphopantetheinylated by EntD and then acylated by EntE using ATP and 2,3-dihydroxybenzoate. *Biochemistry*, 36(28):8495–8503, 1997.

[103] C. A. Shaw-Reid, N. L. Kelleher, H. C. Losey, A. M. Gehring, C. Berg, and C. T. Walsh. Assembly line enzymology by multimodular nonribosomal peptide synthetases: the thioesterase domain of E. coli EntF catalyzes both elongation and cyclolactonization. *Chemistry & Biology*, 6(6):385–400, 1999.

[104] David E. Ehmann, Cathryn A. Shaw-Reid, Heather C. Losey, and Christopher T. Walsh. The EntF and EntE adenylation domains of *Escherichia coli* enterobactin synthetase: Sequestration and selectivity in acyl-AMP transfers to thiolation domain cosubstrates. *Proceedings of the National Academy of Sciences*, 97(6):2509–2514, 2000.

[105] Carl J. Carrano and Kenneth N. Raymond. Ferric ion sequestering agents. 2. Kinetics and mechanism of iron removal from transferrin by enterobactin and synthetic tricatechols. *Journal of the American Chemical Society*, 101(18):5401–5404, 1979.

[106] S. K. Buchanan, B. S. Smith, L. Venkatramani, D. Xia, L. Esser, M. Palnitkar, R. Chakraborty, D. van der Helm, and J. Deisenhofer. Crystal structure of the outer membrane active transporter FepA from Escherichia coli. *Nature Structural Biology*, 6(1):56–63, 1999.

[107] C. Sprencel, Z. Cao, Z. Qi, D. C. Scott, M. A. Montague, N. Ivanoff, J. Xu, K. M. Raymond, S. M. Newton, and P. E. Klebba. Binding of ferric enterobactin by the *Escherichia coli* periplasmic protein FepB. *Journal of Bacteriology*, 182(19):5359–5364, 2000.

[108] K. T. Greenwood and R. K. Luke. Enzymatic hydrolysis of enterochelin and its iron complex in *Escherichia coli* K-12. Properties of enterochelin esterase. *Biochimica Et Biophysica Acta*, 525(1):209–218, 1978.

[109] Klaus Hantke. Dihydroxybenzolyserine—a siderophore for E. coli. *FEMS Microbiology Letters*, 67(1-2):5–8, 1990.

[110] Rebecca E. Parales and John L. Ingraham. The surprising Rut pathway: an unexpected way to derive nitrogen from pyrimidines. *Journal of Bacteriology*, 192(16):4086–4088, 2010.

[111] Eric J. Drake, David A. Nicolai, and Andrew M. Gulick. Structure of the EntB multidomain nonribosomal peptide synthetase and functional analysis of its interaction with the EntE adenylation domain. *Chemistry & Biology*, 13(4):409–419, 2006.

[112] James F. Parsons, Kelly Calabrese, Edward Eisenstein, and Jane E. Ladner. Structure and mechanism of *Pseudomonas aeruginosa* PhzD, an isochorismatase from the phenazine biosynthetic pathway. *Biochemistry*, 42(19):5684–5693, 2003.

[113] Charles O. Rock and Suzanne Jackowski. Forty years of bacterial fatty acid synthesis. *Biochemical and Biophysical Research Communications*, 292(5):1155–1166, 2002.

[114] R. E. Toomey and S. J. Wakil. Studies on the mechanism of fatty acid synthesis. XV. Preparation and general properties of beta-ketoacyl acyl carrier protein reductase from Escherichia coli. *Biochimica Et Biophysica Acta*, 116(2):189–197, 1966.

[115] A. C. Price, Y. M. Zhang, C. O. Rock, and S. W. White. Structure of beta-ketoacyl-[acyl carrier protein] reductase from *Escherichia coli*: negative cooperativity and its structural basis. *Biochemistry*, 40(43):12772–12781, 2001.

[116] Allen C. Price, Yong-Mei Zhang, Charles O. Rock, and Stephen W. White. Cofactor-induced conformational rearrangements establish a catalytically competent active site and a proton relay conduit in FabG. *Structure*, 12(3):417–428, 2004.

[117] Cyprian D. Cukier, Anthony G. Hope, Ayssar A. Elamin, Lucile Moynie, Robert Schnell, Susanne Schach, Holger Kneuper, Mahavir Singh, James H. Naismith, Ylva Lindqvist, David W. Gray, and Gunter Schneider. Discovery of an allosteric inhibitor binding site in 3-oxo-acyl-ACP reductase from *Pseudomonas aeruginosa*. *ACS chemical biology*, 8(11):2518–2527, 2013.

[118] Christopher T. Nomura, Kazunori Taguchi, Zhihua Gan, Kazuhiro Kuwabara, Tomoyo Tanaka, Kazuma Takase, and Yoshiharu Doi. Expression of 3-ketoacyl-acyl carrier protein reductase (*fabG*) genes enhances production of polyhydroxyalkanoate copolymer from glucose in recombinant *Escherichia coli* JM109. *Applied and Environmental Microbiology*, 71(8):4297–4306, 2005.

[119] Ying-Hui Sun, Qing Cheng, Wei-Xi Tian, and Xiao-Dong Wu. A substitutive substrate for measurements of beta-ketoacyl reductases in two fatty acid synthase systems. *Journal of Biochemical and Biophysical Methods*, 70(6):850–856, 2008.

[120] Yong-Mei Zhang, Bainan Wu, Jie Zheng, and Charles O. Rock. Key residues responsible for acyl carrier protein and beta-ketoacyl-acyl carrier protein reductase (FabG) interaction. *The Journal of Biological Chemistry*, 278(52):52935–52943, 2003.

[121] M. G. Rossmann, D. Moras, and K. W. Olsen. Chemical and biological evolution of nucleotide-binding protein. *Nature*, 250(463):194–199, 1974.

[122] Zhe Hu, Jincheng Ma, Yicai Chen, Wenhua Tong, Lei Zhu, Haihong Wang, and John E. Cronan. *Escherichia coli* FabG 3-ketoacyl-ACP reductase proteins lacking the assigned catalytic triad residues are active enzymes. *Journal of Biological Chemistry*, 296:100365, 2021.

[123] Jesse A. Sundlov, Julie A. Garringer, Jill M. Carney, Albert S. Reger, Eric J. Drake, William L. Duax, and Andrew M. Gulick. Determination of the crystal structure of EntA, a 2,3-dihydro-2,3-dihydroxybenzoic acid dehydrogenase from *Escherichia coli*. *Acta Crystallographica. Section D, Biological Crystallography*, 62(Pt 7):734–740, 2006.

[124] Israel Hanukoglu. Proteopedia: Rossmann fold: A beta-alpha-beta fold at dinucleotide binding sites. *Biochemistry and Molecular Biology Education*, 43(3):206–209, 2015.

[125] Jörg Claren, Christoph Malisi, Birte Höcker, and Reinhard Sterner. Establishing wild-type levels of catalytic activity on natural and artificial $(\beta\alpha)$8-barrel protein scaffolds. *Proceedings of the National Academy of Sciences*, 106(10):3704–3709, 2009.

[126] Anna M. Goral, Karolina L. Tkaczuk, Maksymilian Chruszcz, Olga Kagan, Alexei Savchenko, and Wladek Minor. Crystal structure of a putative isochorismatase hydrolase from *Oleispira antarctica*. *Journal of Structural and Functional Genomics*, 13(1):27–36, 2012.

[127] Rebecca L. Scholz and E. Peter Greenberg. Sociality in *Escherichia coli*: enterochelin is a private good at low cell density and can be shared at high cell density. *Journal of Bacteriology*, 197(13):2122–2128, 2015.

[128] Claudia Andreini, Ivano Bertini, Gabriele Cavallaro, Gemma L. Holliday, and Janet M. Thornton. Metal ions in biological catalysis: from enzyme databases to general principles. *Journal of Biological Inorganic Chemistry*, 13(8):1205–1218, 2008.

[129] Kevin J. Waldron, Julian C. Rutherford, Dianne Ford, and Nigel J. Robinson. Metalloproteins and metal sensing. *Nature*, 460(7257):823–830, 2009.

[130] Volkmar Braun and Michael Braun. Iron transport and signaling in Escherichia coli. *FEBS Letters*, 529(1):78–85, 2002.

[131] V. de Lorenzo, A. Bindereif, B. H. Paw, and J. B. Neilands. Aerobactin biosynthesis and transport genes of plasmid ColV-K30 in Escherichia coli K-12. *Journal of Bacteriology*, 165(2):570–578, 1986.

[132] Erin C. Garcia, Ariel R. Brumbaugh, and Harry L. T. Mobley. Redundancy and specificity of *Escherichia coli* iron acquisition systems during urinary tract infection. *Infection and Immunity*, 79(3):1225–1235, 2011.

[133] Zhe Zhou, Jonathan R. Lai, and Christopher T. Walsh. Directed evolution of aryl carrier proteins in the enterobactin synthetase. *Proceedings of the National Academy of Sciences of the United States of America*, 104(28):11621–11626, 2007.

[134] Katherine M. Digianantonio, Maria Korolev, and Michael H. Hecht. A non-natural protein rescues cells deleted for a key enzyme in central metabolism. *ACS synthetic biology*, 6(4):694–700, 2017.

[135] Bettina Rohweder, Gerhard Lehmann, Norbert Eichner, Tino Polen, Chitra Rajendran, Fabian Ruperti, Mona Linde, Thomas Treiber, Oona Jung, Katja Dettmer, Gunter Meister, Michael Bott, Wolfram Gronwald, and Reinhard Sterner. Library selection with a randomized repertoire of $(\beta\alpha)$8-barrel enzymes results in unexpected induction of gene expression. *Biochemistry*, 58(41):4207–4217, 2019.

[136] Sammy Pontrelli, Riley C. B. Fricke, Shao Thing Teoh, Walter A. Laviña, Sastia Prama Putri, Sorel Fitz-Gibbon, Matthew Chung, Matteo Pellegrini, Eiichiro Fukusaki, and James C. Liao. Metabolic repair through emergence of new pathways in *Escherichia coli*. *Nature Chemical Biology*, 14(11):1005–1009, 2018.

[137] B. Schwyn and J. B. Neilands. Universal chemical assay for the detection and determination of siderophores. *Analytical Biochemistry*, 160(1):47–56, 1987.

[138] Stephanie D. Himpsl and Harry L. T. Mobley. Siderophore detection using chrome azurol S and cross-feeding assays. *Methods in Molecular Biology*, 2021:97–108, 2019.

[139] Lambros J. Tassoulas, Mikael H. Elias, and Lawrence P. Wackett. Discovery of an ultraspecific triuret hydrolase (TrtA) establishes the triuret biodegradation pathway. *Journal of Biological Chemistry*, 296:100055, 2021.

[140] Lygie Esquirol, Thomas S. Peat, Matthew Wilding, Del Lucent, Nigel G. French, Carol J. Hartley, Janet Newman, and Colin Scott. Structural and biochemical characterization of the biuret hydrolase (BiuH) from the cyanuric acid catabolism pathway of *Rhizobium leguminasorum* bv. *viciae* 3841. *PLOS ONE*, 13(2):e0192736, 2018.

[141] C. Wasternack. Degradation of pyrimidines and pyrimidine analogs–pathways and mutual influences. *Pharmacology & Therapeutics*, 8(3):629–651, 1980.

[142] Rita Zrenner, Heike Riegler, Cathleen R. Marquard, Peter R. Lange, Claudia Geserick, Caren E. Bartosz, Celine T. Chen, and Robert D. Slocum. A functional analysis of the pyrimidine catabolic pathway in *Arabidopsis*. *The New Phytologist*, 183(1):117–132, 2009.

[143] L. Leon Campbell. Reductive degradation of pyrimidines I. the isolation and characterization of a uracil fermenting bacterium, *Clostridium uracilicum* nov. spec. *Journal of Bacteriology*, 73(2):220–224, 1957.

[144] T. P. West. Pyrimidine base catabolism in *Pseudomonas putida* biotype B. *Antonie Van Leeuwenhoek*, 80(2):163–167, 2001.

[145] Klaus D. Schnackerz and Doreen Dobritzsch. Amidohydrolases of the reductive pyrimidine catabolic pathway: Purification, characterization, structure, reaction mechanisms and enzyme deficiency. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1784(3):431–444, 2008.

[146] Di Zhu, Yifeng Wei, Jinyu Yin, Dazhi Liu, Ee Lui Ang, Huimin Zhao, and Yan Zhang. A pathway for degradation of uracil to acetyl coenzyme A in *Bacillus megaterium*. *Applied and Environmental Microbiology*, 86(7):e02837–19, 2020.

[147] Jinyu Yin, Yifeng Wei, Dazhi Liu, Yiling Hu, Qiang Lu, Ee Lui Ang, Huimin Zhao, and Yan Zhang. An extended bacterial reductive pyrimidine degradation pathway that enables nitrogen release from β-alanine. *The Journal of Biological Chemistry*, 294(43):15662–15671, 2019.

[148] G D Vogels and C Van der Drift. Degradation of purines and pyrimidines by microorganisms. *Bacteriological Reviews*, 40(2):403–468, 1976.

[149] C. L. Soong, J. Ogawa, and S. Shimizu. Novel amidohydrolytic reactions in oxidative pyrimidine metabolism: analysis of the barbiturase reaction and discovery of a novel enzyme, ureidomalonase. *Biochemical and Biophysical Research Communications*, 286(1):222–226, 2001.

[150] Chee-Leong Soong, Jun Ogawa, Eiji Sakuradani, and Sakayu Shimizu. Barbiturase, a novel zinc-containing amidohydrolase involved in oxidative pyrimidine metabolism. *The Journal of Biological Chemistry*, 277(9):7051–7058, 2002.

[151] S. Simaga and E. Kos. Properties and regulation of pyrimidine catabolism in *Escherichia coli*. *The International Journal of Biochemistry*, 13(5):615–619, 1981.

[152] T. P. West. Isolation and characterization of an *Escherichia coli* B mutant strain defective in uracil catabolism. *Canadian Journal of Microbiology*, 44(11):1106–1109, 1998.

[153] Ryota Hidese, Hisaaki Mihara, Tatsuo Kurihara, and Nobuyoshi Esaki. *Escherichia coli* dihydropyrimidine dehydrogenase is a novel NAD-dependent heterotetramer essential for the production of 5,6-dihydrouracil. *Journal of Bacteriology*, 193(4):989–993, 2011.

[154] Kevin D. Loh, Prasad Gyaneshwar, Eirene Markenscoff Papadimitriou, Rebecca Fong, Kwang-Seo Kim, Rebecca Parales, Zhongrui Zhou, William Inwood, and Sydney Kustu. A previously undescribed pathway for pyrimidine catabolism. *Proceedings of the National Academy of Sciences of the United States of America*, 103(13):5114–5119, 2006.

[155] Arne Matthews, Raspudin Saleem-Batcha, Jacob N. Sanders, Frederick Stull, K. N. Houk, and Robin Teufel. Aminoperoxide adducts expand the catalytic repertoire of flavin monooxygenases. *Nature Chemical Biology*, 16(5):556–563, 2020.

[156] Aleksandra Alicja Knapik, Janusz Jurand Petkowski, Zbyszek Otwinowski, Marcin Tadeusz Cymborowski, David Robert Cooper, Maksymilian Chruszcz, Wanda Małgorzata Krajewska, and Wladek Minor. Structure of *Escherichia coli* RutC, a member of the YjgF family and putative aminoacrylate peracid reductase of the rut operon. *Acta Crystallographica. Section F, Structural Biology and Crystallization Communications*, 68(Pt 11):1294–1299, 2012.

[157] Aleksandra A. Knapik, Janusz J. Petkowski, Zbyszek Otwinowski, Marcin T. Cymborowski, David R. Cooper, Karolina A. Majorek, Maksymilian Chruszcz, Wanda M. Krajewska, and Wladek Minor. A multi-faceted analysis of RutD reveals a novel family of α/β hydrolases. *Proteins*, 80(10):2359–2368, 2012.

[158] Phu Nguyen Le Minh, Sergio de Cima, Indra Bervoets, Dominique Maes, Vicente Rubio, and Daniel Charlier. Ligand binding specificity of RutR, a member of the TetR family of transcription regulators in *Escherichia coli*. *FEBS Open Bio*, 5:76–84, 2015.

[159] Jianjun Wang, Junge Zhu, and Sheng Wu. Immobilization on macroporous resin makes *E. coli* RutB a robust catalyst for production of (-) Vince lactam. *Applied Microbiology and Biotechnology*, 99(11):4691–4700, 2015.

[160] Rohit Singh and Robert Vince. 2-Azabicyclo[2.2.1]hept-5-en-3-one: chemical profile of a versatile synthetic building block and its impact on the development of therapeutics. *Chemical Reviews*, 112(8):4642–4686, 2012.

[161] Shaozhou Zhu and Guojun Zheng. Dynamic kinetic resolution of Vince lactam catalyzed by γ-lactamases: a mini-review. *Journal of Industrial Microbiology & Biotechnology*, 45(12):1017–1031, 2018.

[162] Jianjun Wang, Yaxin Zhu, Guogang Zhao, Junge Zhu, and Sheng Wu. Characterization of a recombinant (+)-γ-lactamase from *Microbacterium hydrocarbonoxydans* which provides evidence that two enantiocomplementary γ-lactamases are in the strain. *Applied Microbiology and Biotechnology*, 99(7):3069–3080, 2015.

[163] Shuaihua Gao, Rong Huang, Shaozhou Zhu, Hongxia Li, and Guojun Zheng. Identification and characterization of a novel (+)-γ-lactamase from *Microbacterium hydrocarbonoxydans*. *Applied Microbiology and Biotechnology*, 100(22):9543–9553, 2016.

[164] Shuaihua Gao, Shaozhou Zhu, Rong Huang, Hongxia Li, Hao Wang, and Guojun Zheng. Engineering the enantioselectivity and thermostability of a (+)-γ-lactamase from *Microbacterium hydrocarbonoxydans* for kinetic resolution of Vince lactam (2-azabicyclo[2.2.1]hept-5-en-3-one). *Applied and Environmental Microbiology*, 84(1):e01780–17, 2018.

[165] V. P. Saxena and D. B. Wetlaufer. A New Basis for Interpreting the Circular Dichroic Spectra of Proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 68(5):969–972, 1971.

[166] Shiheng Liu, Conggang Zhang, Ning Li, Bei Niu, Mengyuan Liu, Xiuhua Liu, Tiandi Wei, Deyu Zhu, Yan Huang, Sujuan Xu, and Lichuan Gu. Structural insight into the ISC domain of VibB from *Vibrio cholerae* at atomic resolution: a snapshot just before the enzymatic reaction. *Acta Crystallographica. Section D, Biological Crystallography*, 68(Pt 10):1329–1338, 2012.

[167] Maximilian G. Plach, Florian Semmelmann, Florian Busch, Markus Busch, Leonhard Heizinger, Vicki H. Wysocki, Rainer Merkl, and Reinhard Sterner. Evolutionary diversification of protein-protein interactions by interface add-ons. *Proceedings of the National Academy of Sciences of the United States of America*, 114(40):E8333–E8342, 2017.

[168] James D. Warren, John H. MacMillan, and Stephen S. Washburne. Synthesis of substituted 2H-1,3-oxazine-2,6-diones by reaction of trimethylsilyl azide with maleic anhydrides. *The Journal of Organic Chemistry*, 40(6):743–746, 1975.

[169] M. Kanehisa and S. Goto. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28(1):27–30, 2000.

[170] Rainer Merkl and Reinhard Sterner. Ancestral protein reconstruction: techniques and applications. *Biological Chemistry*, 397(1):1–21, 2016.

[171] Isaac Fruchey, Nir Shapir, Michael J. Sadowsky, and Lawrence P. Wackett. On the origins of cyanuric acid hydrolase: purification, substrates, and prevalence of AtzD from *Pseudomonas sp.* strain ADP. *Applied and Environmental Microbiology*, 69(6):3653–3657, 2003.

[172] Jennifer L. Seffernick, Jasmine S. Erickson, Stephan M. Cameron, Seunghee Cho, Anthony G. Dodge, Jack E. Richman, Michael J. Sadowsky, and Lawrence P. Wackett. Defining sequence space and reaction products within the cyanuric acid hydrolase (AtzD)/barbiturase protein family. *Journal of Bacteriology*, 194(17):4579–4588, 2012.

[173] Thomas S. Peat, Sahil Balotra, Matthew Wilding, Carol J. Hartley, Janet Newman, and Colin Scott. High-resolution X-ray structures of two functionally distinct members of the cyclic amide hydrolase family of toblerone fold enzymes. *Applied and Environmental Microbiology*, 83(9):e03365–16, 2017.

[174] S. D. Copley. Evolution of a metabolic pathway for degradation of a toxic xenobiotic: the patchwork approach. *Trends in Biochemical Sciences*, 25(6):261–265, 2000.

[175] Robyn J. Russell, Colin Scott, Colin J. Jackson, Rinku Pandey, Gunjan Pandey, Matthew C. Taylor, Christopher W. Coppin, Jian-Wei Liu, and John G. Oakeshott. The evolution of new enzyme function: lessons from xenobiotic metabolizing bacteria versus insecticide-resistant insects. *Evolutionary Applications*, 4(2):225–248, 2011.

[176] Todd A. Gaines, Stephen O. Duke, Sarah Morran, Carlos A. G. Rigon, Patrick J. Tranel, Anita Küpper, and Franck E. Dayan. Mechanisms of evolved herbicide resistance. *Journal of Biological Chemistry*, 295(30):10307–10330, 2020.

[177] M. E. Whalon, D. Mota-Sanchez, and R. M. Hollingworth. Analysis of global pesticide resistance in arthropods. In M. E. Whalon, D. Mota-Sanchez, and R. M. Hollingworth, editors, *Global pesticide resistance in arthropods*, pages 5–31. CABI, Wallingford, 2008.

[178] Dan I Andersson. Persistence of antibiotic resistant bacteria. *Current Opinion in Microbiology*, 6(5):452–456, 2003.

[179] Sojib Bin Zaman, Muhammed Awlad Hussain, Rachel Nye, Varshil Mehta, Kazi Taib Mamun, and Naznin Hossain. A review on antibiotic resistance: alarm bells are ringing. *Cureus*, 9(6):e1403, 2017.

[180] R. Elaine Dick and John P. Quinn. Control of glyphosate uptake and metabolism in *Pseudomonas sp.* 4ASW. *FEMS Microbiology Letters*, 134(2-3):177–182, 1995.

[181] J. L. Seffernick and L. P. Wackett. Rapid evolution of bacterial catabolic enzymes: a case study with atrazine chlorohydrolase. *Biochemistry*, 40(43):12747–12753, 2001.

[182] Luísa Barreiros, Balbina Nogales, Célia M. Manaia, António C. Silva Ferreira, Dietmar H. Pieper, Maria A. Reis, and Olga C. Nunes. A novel pathway for mineralization of the thiocarbamate herbicide molinate by a defined bacterial mixed culture. *Environmental Microbiology*, 5(10):944–953, 2003.

[183] L. Jason Krutz, Dale L Shaner, Mark A Weaver, Richard MT Webb, Robert M Zablotowicz, Krishna N Reddy, Yanbo Huang, and Steven J Thomson. Agronomic and environmental implications of enhanced s-triazine degradation. *Pest Management Science*, 66(5):461–481, 2010.

[184] Simranjeet Singh, Vijay Kumar, Jatinder Pal Kaur Gill, Shivika Datta, Satyender Singh, Vaishali Dhaka, Dhriti Kapoor, Abdul Basit Wani, Daljeet Singh Dhanjal, Manoj Kumar, S. L. Harikumar, and Joginder Singh. Herbicide glyphosate: toxicity and microbial degradation. *International Journal of Environmental Research and Public Health*, 17(20):7519, 2020.

[185] John E. Mullet and Charles J. Arntzen. Identification of a 32–34-kilodalton polypeptide as a herbicide receptor protein in photosystem II. *Biochimica et Biophysica Acta - Bioenergetics*, 635(2):236–248, 1981.

[186] L. P. Wackett, M. J. Sadowsky, B. Martinez, and N. Shapir. Biodegradation of atrazine and related s-triazine compounds: from enzymes to field studies. *Applied Microbiology and Biotechnology*, 58(1):39–45, 2002.

[187] Anthony G. Dodge, Lawrence P. Wackett, and Michael J. Sadowsky. Plasmid localization and organization of melamine degradation genes in *Rhodococcus sp.* strain Mel. *Applied and Environmental Microbiology*, 78(5):1397–1403, 2012.

[188] D. A. Belluck, S. L. Benjamin, and T. Dawson. Groundwater contamination by atrazine and its metabolites: risk assessment, policy, and legal implications. *ACS Symposium series - American Chemical Society (USA)*, 1991.

[189] Marc Ribaudo and Aziz Bouzaher. Atrazine: environmental characteristics and economics of management. URL: http://www.ers.usda.gov/publications/pub-details/?pubid=40594.

[190] Xuxiao Fan and Fuqiang Song. Bioremediation of atrazine: recent advances and promises. *Journal of Soils and Sediments*, 14(10):1727–1737, 2014.

[191] L. L. McCormick and A. E. Hiltbold. Microbiological decomposition of atrazine and diuron in soil. *Weeds*, 14(1):77–82, 1966.

[192] S. R. Obien and R. E. Green. Degradation of atrazine in four hawaiian soils. *Weed Science*, 17(4):509–514, 1969.

[193] H. M. LeBaron. Ways and means to influence the activity and the persistence of triazine herbicides in soils. *Residue Reviews*, 32:311–353, 1970.

[194] T. J. Sheets. Persistence of triazine herbicides in soils. *Residue Reviews*, 32:287–310, 1970.

[195] Sajid Noor, Frédérique Changey, John G. Oakeshott, Colin Scott, and Fabrice Martin-Laurent. Ongoing functional evolution of the bacterial atrazine chlorohydrolase AtzA. *Biodegradation*, 25(1):21–30, 2014.

[196] Lubow Jowa and Robert Howd. Should atrazine and related chlorotriazines be considered carcinogenic for human health risk assessment? *Journal of Environmental Science and Health. Part C, Environmental Carcinogenesis & Ecotoxicology Reviews*, 29(2):91–144, 2011.

[197] USDS. Pesticide National Synthesis Project. URL: https://water.usgs.gov/nawqa/pnsp/usage/maps/show_map.php?year=2016&map=ATRAZINE&hilo=L. 2016.

[198] G. F. Ryan. Resistance of common groundsel to simazine and atrazine. *Weed Science*, 18(5):614–616, 1970.

[199] Joseph Hirschberg, Anthony Bleecker, David J. Kyle, Lee McIntosh, and Charles J. Arntzen. The molecular basis of triazine-herbicide resistance in higher-plant chloroplasts. *Zeitschrift für Naturforschung C*, 39(5):412–420, 1984.

[200] Klaus Pfister, Steven R. Radosevich, and Charles J. Arntzen. Modification of herbicide binding to photosystem II in two biotypes of *Senecio vulgaris* L. *Plant Physiology*, 64(6):995–999, 1979.

[201] R. R. Stein, A. L. Castellvi, J. P. Bogacz, and C. A. Wraight. Herbicide-quinone competition in the acceptor complex of photosynthetic reaction centers from *Rhodopseudomonas sphaeroides*: a bacterial model for PS-II-herbicide activity in plants. *Journal of Cellular Biochemistry*, 24(3):243–259, 1984.

[202] W. Oettmeier. Herbicide resistance and supersensitivity in photosystem II. *Cellular and molecular life sciences: CMLS*, 55(10):1255–1277, 1999.

[203] Dmitry P. Bazhanov, Kai Yang, Hongmei Li, Chengyun Li, Jishun Li, Xiangfeng Chen, and Hetong Yang. Colonization of plant roots and enhanced atrazine degradation by a strain of *Arthrobacter ureafaciens*. *Applied Microbiology and Biotechnology*, 101(17):6809–6820, 2017.

[204] Nikolina Udiković-Kolić, Colin Scott, and Fabrice Martin-Laurent. Evolution of atrazine-degrading capabilities in the environment. *Applied Microbiology and Biotechnology*, 96(5):1175–1189, 2012.

[205] Mervyn L. de Souza, Jennifer Seffernick, Betsy Martinez, Michael J. Sadowsky, and Lawrence P. Wackett. The atrazine catabolism genes *atzABC* are widespread and highly conserved. *Journal of Bacteriology*, 180(7):1951–1954, 1998.

[206] J. K. Struthers, K. Jayachandran, and T. B. Moorman. Biodegradation of atrazine by *Agrobacterium radiobacter* J14a and use of this strain in bioremediation of contaminated soil. *Applied and Environmental Microbiology*, 64(9):3368–3375, 1998.

[207] Lisa C. Strong, Charlotte Rosendahl, Gilbert Johnson, Michael J. Sadowsky, and Lawrence P. Wackett. *Arthrobacter aurescens* TC1 metabolizes diverse s-triazine ring compounds. *Applied and Environmental Microbiology*, 68(12):5973–5980, 2002.

[208] S. Rousseaux, A. Hartmann, and G. Soulas. Isolation and characterisation of new Gram-negative and Gram-positive atrazine degrading bacteria from different French soils. *FEMS microbiology ecology*, 36(2-3):211–222, 2001.

[209] Edward Topp, Walter M. Mulbry, Hong Zhu, Sarah M. Nour, and Diane Cuppels. Characterization of s-triazine herbicide metabolism by a *Nocardioides sp.* isolated from agricultural soils. *Applied and Environmental Microbiology*, 66(8):3134–3141, 2000.

[210] C. Yanze-Kontchou and N. Gschwind. Mineralization of the herbicide atrazine as a carbon source by a *Pseudomonas strain*. *Applied and Environmental Microbiology*, 60(12):4297–4302, 1994.

[211] R. T. Mandelbaum, D. L. Allan, and L. P. Wackett. Isolation and characterization of a *Pseudomonas sp.* that mineralizes the s-triazine herbicide atrazine. *Applied and Environmental Microbiology*, 61(4):1451–1457, 1995.

[212] M. Wenk, T. Baumgartner, J. Dobovsek, T. Fuchs, J. Kucsera, J. Zopfi, and G. Stucki. Rapid atrazine mineralisation in soil slurry and moist soil by inoculation of an atrazine-degrading *Pseudomonas sp.* strain. *Applied Microbiology and Biotechnology*, 49(5):624–630, 1998.

[213] M Radosevich, S J Traina, Y L Hao, and O H Tuovinen. Degradation and mineralization of atrazine by a soil bacterial isolate. *Applied and Environmental Microbiology*, 61(1):297–302, 1995.

[214] David M Stamper, Mark Radosevich, Kevin B Hallberg, Samuel J Traina, and Olli H Tuovinen. *Ralstonia basilensis* M91-3, a denitrifying soil bacterium capable of using s-triazines as nitrogen sources. *Canadian Journal of Microbiology*, 48(12):1089–1098, 2002.

[215] Ana Flavia Tonelli Fernandes, Michelle Barbosa Partata da Silva, Vinicius Vicente Martins, Carlos Eduardo Saraiva Miranda, and Eliana Guedes Stehling. Isolation and characterization of a *Pseudomonas aeruginosa* from a virgin Brazilian Amazon region with potential to degrade atrazine. *Environmental Science and Pollution Research International*, 21(24):13974–13978, 2014.

[216] R. L. Yale, M. Sapp, C. J. Sinclair, and J. W. B. Moir. Microbial changes linked to the accelerated degradation of the herbicide atrazine in a range of temperate soils. *Environmental Science and Pollution Research International*, 24(8):7359–7374, 2017.

[217] Dmitry P. Bazhanov, Chengyun Li, Hongmei Li, Jishun Li, Xinjian Zhang, Xiangfeng Chen, and Hetong Yang. Occurrence, diversity and community structure of culturable atrazine degraders in industrial and agricultural soils exposed to the herbicide in Shandong Province, P.R. China. *BMC microbiology*, 16(1):265, 2016.

[218] M. L. de Souza, D. Newcombe, S. Alvey, D. E. Crowley, A. Hay, M. J. Sadowsky, and L. P. Wackett. Molecular basis of a bacterial consortium: interspecies catabolism of atrazine. *Applied and Environmental Microbiology*, 64(1):178–184, 1998.

[219] Nikolina Udiković Kolić, Dubravka Hršak, Ana Begonja Kolar, Ines Petrić, Sanja Stipičevic, Guy Soulas, and Fabrice Martin-Laurent. Combined metabolic activity within an atrazine-mineralizing community enriched from agrochemical factory soil. *International Biodeterioration & Biodegradation*, 60(4):299–307, 2007.

[220] C. Bouquard, J. Ouazzani, J. Prome, Y. Michel-Briand, and P. Plesiat. Dechlorination of atrazine by a *Rhizobium sp.* isolate. *Applied and Environmental Microbiology*, 63(3):862–866, 1997.

[221] D. R. Shelton, S. Khader, J. S. Karns, and B. M. Pogell. Metabolism of twelve herbicides by *Streptomyces*. *Biodegradation*, 7(2):129–136, 1996.

[222] Fernando Govantes, Odil Porrúa, Vicente García-González, and Eduardo Santero. Atrazine biodegradation in the lab and in the field: enzymatic activities and gene regulation. *Microbial Biotechnology*, 2(2):178–185, 2009.

[223] Lygie Esquirol, Thomas S. Peat, Matthew Wilding, Jian-Wei Liu, Nigel G. French, Carol J. Hartley, Hideki Onagi, Thomas Nebl, Christopher J. Easton, Janet Newman, and Colin Scott. An unexpected vestigial protein complex reveals the evolutionary origins of an s-triazine catabolic enzyme. *The Journal of Biological Chemistry*, 293(20):7880–7891, 2018.

[224] Lygie Esquirol, Thomas S. Peat, Matthew Wilding, Carol J. Hartley, Janet Newman, and Colin Scott. A novel decarboxylating amidohydrolase involved in avoiding metabolic dead ends during cyanuric acid catabolism in *Pseudomonas sp.* strain ADP. *PLOS ONE*, 13(11):e0206949, 2018.

[225] M. L. de Souza, L. P. Wackett, K. L. Boundy-Mills, R. T. Mandelbaum, and M. J. Sadowsky. Cloning, characterization, and expression of a gene region from *Pseudomonas sp.* strain ADP involved in the dechlorination of atrazine. *Applied and Environmental Microbiology*, 61(9):3373–3378, 1995.

[226] M L de Souza, M J Sadowsky, and L P Wackett. Atrazine chlorohydrolase from *Pseudomonas sp.* strain ADP: gene sequence, enzyme purification, and protein characterization. *Journal of Bacteriology*, 178(16):4894–4900, 1996.

[227] Colin Scott, Colin J. Jackson, Chris W. Coppin, Roslyn G. Mourant, Margaret E. Hilton, Tara D. Sutherland, Robyn J. Russell, and John G. Oakeshott. Catalytic improvement and evolution of atrazine chlorohydrolase. *Applied and Environmental Microbiology*, 75(7):2184–2191, 2009.

[228] Jennifer L. Seffernick, Hugh McTavish, Jeffrey P. Osborne, Mervyn L. de Souza, Michael J. Sadowsky, and Lawrence P. Wackett. Atrazine chlorohydrolase from *Pseudomonas sp.* strain ADP is a metalloenzyme. *Biochemistry*, 41(48):14430–14437, 2002.

[229] K L Boundy-Mills, M L de Souza, R T Mandelbaum, L P Wackett, and M J Sadowsky. The atzB gene of *Pseudomonas sp.* strain ADP encodes the second enzyme of a novel atrazine degradation pathway. *Applied and Environmental Microbiology*, 63(3):916–923, 1997.

[230] Jennifer L. Seffernick, Asma Aleem, Jeffrey P. Osborne, Gilbert Johnson, Michael J. Sadowsky, and Lawrence P. Wackett. Hydroxyatrazine N-ethylaminohydrolase (AtzB): an amidohydrolase superfamily enzyme catalyzing deamination and dechlorination. *Journal of Bacteriology*, 189(19):6989–6997, 2007.

[231] M. J. Sadowsky, Z. Tong, M. de Souza, and L. P. Wackett. AtzC is a new member of the amidohydrolase protein superfamily and is homologous to other atrazine-metabolizing enzymes. *Journal of Bacteriology*, 180(1):152–158, 1998.

[232] Nir Shapir, Jeffrey P. Osborne, Gilbert Johnson, Michael J. Sadowsky, and Lawrence P. Wackett. Purification, substrate range, and metal center of AtzC: the N-isopropylammelide aminohydrolase involved in bacterial atrazine metabolism. *Journal of Bacteriology*, 184(19):5376–5384, 2002.

[233] Sahil Balotra, Andrew C. Warden, Janet Newman, Lyndall J. Briggs, Colin Scott, and Thomas S. Peat. X-ray structure and mutagenesis studies of the N-isopropylammelide isopropylaminohydrolase, AtzC. *PLOS ONE*, 10(9):e0137700, 2015.

[234] Seunghee Cho, Ke Shi, Lawrence P. Wackett, and Hideki Aihara. Crystallization and preliminary X-ray diffraction studies of cyanuric acid hydrolase from *Azorhizobium caulinodans*. *Acta Crystallographica. Section F, Structural Biology and Crystallization Communications*, 69(Pt 8):880–883, 2013.

[235] Thomas S. Peat, Sahil Balotra, Matthew Wilding, Nigel G. French, Lyndall J. Briggs, Santosh Panjikar, Nathan Cowieson, Janet Newman, and Colin Scott. Cyanuric acid hydrolase: evolutionary innovation by structural concatenation. *Molecular Microbiology*, 88(6):1149–1163, 2013.

[236] Gang Cheng, Nir Shapir, Michael J. Sadowsky, and Lawrence P. Wackett. Allophanate hydrolase, not urease, functions in bacterial cyanuric acid metabolism. *Applied and Environmental Microbiology*, 71(8):4437–4445, 2005.

[237] Nir Shapir, Michael J. Sadowsky, and Lawrence P. Wackett. Purification and characterization of allophanate hydrolase (AtzF) from *Pseudomonas sp.* strain ADP. *Journal of Bacteriology*, 187(11):3731–3738, 2005.

[238] Nir Shapir, Gang Cheng, Michael J. Sadowsky, and Lawrence P. Wackett. Purification and characterization of TrzF: biuret hydrolysis by allophanate hydrolase supports growth. *Applied and Environmental Microbiology*, 72(4):2491–2495, 2006.

[239] Stephan M. Cameron, Katharina Durchschein, Jack E. Richman, Michael J. Sadowsky, and Lawrence P. Wackett. A new family of biuret hydrolases involved in s-triazine ring metabolism. *ACS Catalysis*, 2011(1):1075–1082, 2011.

[240] Sahil Balotra, Janet Newman, Nigel G. French, Lyndall J. Briggs, Thomas S. Peat, and Colin Scott. Crystallization and preliminary X-ray diffraction analysis of the amidase domain of allophanate hydrolase from *Pseudomonas sp.* strain ADP. *Acta Crystallographica. Section F, Structural Biology Communications*, 70(Pt 3):310–315, 2014.

[241] M. L. de Souza, L. P. Wackett, and M. J. Sadowsky. The *atzABC* genes encoding atrazine catabolism are located on a self-transmissible plasmid in *Pseudomonas sp.* strain ADP. *Applied and Environmental Microbiology*, 64(6):2323–2326, 1998.

[242] Vicente García-González, Fernando Govantes, Odil Porrúa, and Eduardo Santero. Regulation of the *Pseudomonas sp.* strain ADP cyanuric acid degradation operon. *Journal of Bacteriology*, 187(1):155–167, 2005.

[243] Fernando Govantes, Vicente García-González, Odil Porrúa, Ana Isabel Platero, Alicia Jiménez-Fernández, and Eduardo Santero. Regulation of the atrazine-degradative genes in *Pseudomonas sp.* strain ADP. *FEMS Microbiology Letters*, 310(1):1–8, 2010.

[244] Klaus Huthmacher and Dieter Most. Cyanuric Acid and Cyanuric Chloride. In *Ullmann's Encyclopedia of Industrial Chemistry*. John Wiley & Sons, Ltd, 2000.

[245] Jennifer L. Seffernick and Lawrence P. Wackett. Ancient evolution and recent evolution converge for the biodegradation of cyanuric acid and related triazines. *Applied and Environmental Microbiology*, 2016.

[246] B. Martinez, J. Tomkins, L. P. Wackett, R. Wing, and M. J. Sadowsky. Complete nucleotide sequence and organization of the atrazine catabolic plasmid pADP-1 from *Pseudomonas sp.* strain ADP. *Journal of Bacteriology*, 183(19):5684–5697, 2001.

[247] Lawrence P Wackett. Questioning our perceptions about evolution of biodegradative enzymes. *Current Opinion in Microbiology*, 12(3):244–251, 2009.

[248] Shelley D. Copley. Setting the stage for evolution of a new enzyme. *Current Opinion in Structural Biology*, 69:41–49, 2021.

[249] Sajid Noor, Matthew C. Taylor, Robyn J. Russell, Lars S. Jermiin, Colin J. Jackson, John G. Oakeshott, and Colin Scott. Intramolecular epistasis and the evolution of a new enzymatic function. *PLoS ONE*, 7(6):e39822, 2012.

[250] M. Kolár, K. Urbánek, and T. Látal. Antibiotic selective pressure and development of bacterial resistance. *International Journal of Antimicrobial Agents*, 17(5):357–363, 2001.

[251] Amir Aharoni, Leonid Gaidukov, Olga Khersonsky, Stephen McQ Gould, Cintia Roodveldt, and Dan S. Tawfik. The 'evolvability' of promiscuous protein functions. *Nature Genetics*, 37(1):73–76, 2005.

[252] Misha Soskine and Dan S. Tawfik. Mutational effects and the evolution of new protein functions. *Nature Reviews Genetics*, 11(8):572–582, 2010.

[253] Shelley D. Copley. Evolution of new enzymes by gene duplication and divergence. *The FEBS Journal*, 287(7):1262–1283, 2020.

[254] Dan S. Tawfik. Enzyme promiscuity and evolution in light of cellular metabolism. *The FEBS journal*, 287(7):1260–1261, 2020.

[255] W J Mackay, S Han, and L D Samson. DNA alkylation repair limits spontaneous base substitution mutations in *Escherichia coli*. *Journal of Bacteriology*, 176(11):3224–3230, 1994.

[256] Richard Ellis Hudson, Ulfar Bergthorsson, John R. Roth, and Howard Ochman. Effect of chromosome location on bacterial mutation rates. *Molecular Biology and Evolution*, 19(1):85–92, 2002.

[257] Eyal Akiva, Janine N. Copp, Nobuhiko Tokuriki, and Patricia C. Babbitt. Evolutionary and molecular foundations of multiple contemporary functions of the nitroreductase superfamily. *Proceedings of the National Academy of Sciences*, 114(45):E9549–E9558, 2017.

[258] Pablo Perez-Garcia, Stefanie Kobus, Christoph G. W. Gertzen, Astrid Hoeppner, Nicholas Holzscheck, Christoph Heinrich Strunk, Harald Huber, Karl-Erich Jaeger, Holger Gohlke, Filip Kovacic, Sander H. J. Smits, Wolfgang R. Streit, and Jennifer Chow. A promiscuous ancestral enzyme´s structure unveils protein variable regions of the highly diverse metallo-β-lactamase family. *Communications Biology*, 4(1):132, 2021.

[259] Livnat Afriat-Jurnou, Colin J. Jackson, and Dan S. Tawfik. Reconstructing a missing link in the evolution of a recently diverged phosphotriesterase by active-site loop remodeling. *Biochemistry*, 51(31):6047–6055, 2012.

[260] Nir Shapir, Charlotte Pedersen, Omer Gil, Lisa Strong, Jennifer Seffernick, Michael J. Sadowsky, and Lawrence P. Wackett. TrzN from *Arthrobacter aurescens* TC1 is a zinc amidohydrolase. *Journal of Bacteriology*, 188(16):5859–5864, 2006.

[261] Clara M. Seibert and Frank M. Raushel. Structural and catalytic diversity within the amidohydrolase superfamily. *Biochemistry*, 44(17):6383–6391, 2005.

[262] Roger Shek, Tylene Hilaire, Jasper Sim, and Jarrod B. French. Structural determinants for substrate selectivity in guanine deaminase enzymes of the amidohydrolase superfamily. *Biochemistry*, 58(30):3280–3292, 2019.

[263] Patrick A. Limbach, Pamela F. Crain, and James A. McCloskey. Summary: the modified nucleosides of RNA. *Nucleic Acids Research*, 22(12):2183–2196, 1994.

[264] Hiroshi Takeda, Hiroyuki Hori, and Yaeta Endo. Identification of *Aquifex aeolicus* tRNA (m2(2G26) methyltransferase gene. *Nucleic Acids Research*, (2):229–230, 2002.

[265] Jean Armengaud, Jaunius Urbonavicius, Bernard Fernandez, Guylaine Chaussinand, Janusz M. Bujnicki, and Henri Grosjean. N2-methylation of guanosine at position 10 in tRNA is catalyzed by a THUMP domain-containing, S-adenosylmethionine-dependent methyltransferase, conserved in Archaea and Eukaryota. *The Journal of Biological Chemistry*, 279(35):37142–37152, 2004.

[266] Jaunius Urbonavicius, Jean Armengaud, and Henri Grosjean. Identity elements required for enzymatic formation of N2,N2-dimethylguanosine from N2-monomethylated derivative and its possible role in avoiding alternative conformations in archaeal tRNA. *Journal of Molecular Biology*, 357(2):387–399, 2006.

[267] Akira Hirata, Takeo Suzuki, Tomoko Nagano, Daishiro Fujii, Mizuki Okamoto, Manaka Sora, Todd M. Lowe, Tamotsu Kanai, Haruyuki Atomi, Tsutomu Suzuki, and Hiroyuki Hori. Distinct modified nucleosides in tRNA-Trp from the hyperthermophilic archaeon *Thermococcus kodakarensis* and requirement of tRNA m2G10/m2 2G10 methyltransferase (archaeal Trm11) for survival at high temperatures. *Journal of Bacteriology*, 201(21):e00448–19, 2019.

[268] Suresh K. Purushothaman, Janusz M. Bujnicki, Henri Grosjean, and Bruno Lapeyre. Trm11p and Trm112p are both required for the formation of 2-methylguanosine at position 10 in yeast tRNA. *Molecular and Cellular Biology*, 25(11):4359–4370, 2005.

[269] Kazuki Okada, Yuki Muneyoshi, Yaeta Endo, and Hiroyuki Hori. Production of yeast (m2G10) methyltransferase (Trm11 and Trm112 complex) in a wheat germ cell-free translation system. *Nucleic Acids Symposium Series (2004)*, (53):303–304, 2009.

[270] Marie-Hélène Mazauric, Léon Dirick, Suresh K. Purushothaman, Glenn R. Björk, and Bruno Lapeyre. Trm112p is a 15-kDa zinc finger protein essential for the activity of two tRNA and one protein methyltransferases in yeast. *The Journal of Biological Chemistry*, 285(24):18505–18515, 2010.

[271] H. T. Zhu, L. Y. Qin, T. Liu, and Y. Luo. A Convenient Synthesis of N2-Alkylated Guanines. *Russian Journal of Organic Chemistry*, 55(6):874–878, 2019.

[272] T. S. Peat, J. Newman, S. Balotra, D. Lucent, A. C. Warden, and C. Scott. The structure of the hexameric atrazine chlorohydrolase AtzA. *Acta Crystallographica Section D: Biological Crystallography*, 71(Pt 3):710–720, 2015.

[273] Jennifer L. Seffernick, Erik Reynolds, Alexander A. Fedorov, Elena Fedorov, Steven C. Almo, Michael J. Sadowsky, and Lawrence P. Wackett. X-ray structure and mutational analysis of the atrazine chlorohydrolase TrzN. *The Journal of Biological Chemistry*, 285(40):30606–30614, 2010.

[274] Margarete Bucheli-Witschel and Thomas Egli. Environmental fate and microbial degradation of aminopolycarboxylic acids. *FEMS Microbiology Reviews*, 25(1):69–106, 2001.

[275] Se-Young Jun, Kevin M. Lewis, Buhyun Youn, Luying Xun, and ChulHee Kang. Structural and biochemical characterization of EDTA monooxygenase and its physical interaction with a partner flavin reductase. *Molecular Microbiology*, 100(6):989–1003, 2016.

[276] Andrea Jesenská, Marta Monincová, Tána Koudeláková, Khomaini Hasan, Radka Chaloupková, Zbynek Prokop, Arie Geerlof, and Jirí Damborsky. Biochemical characterization of haloalkane dehalogenases DrbA and DmbC, representatives of a novel subfamily. *Applied and Environmental Microbiology*, 75(15):5157–5160, 2009.

[277] Michelle R. Gisi and Luying Xun. Characterization of chlorophenol 4-monooxygenase (TftD) and NADH:flavin adenine dinucleotide oxidoreductase (TftC) of *Burkholderia cepacia* AC1100. *Journal of Bacteriology*, 185(9):2786–2792, 2003.

[278] K. A. Brown. Phosphotriesterases of *Flavobacterium sp*. *Soil Biology and Biochemistry*, 12(2):105–112, 1980.

[279] C Yeates, MR Gillings, AD Davison, N Altavilla, and DA Veal. Methods for microbial DNA extraction from soil for PCR amplification. *Biological Procedures Online*, 1:40–47, 1998.

[280] Elena Sugrue, Carol J. Hartley, Colin Scott, and Colin J. Jackson. The evolution of new catalytic mechanisms for xenobiotic hydrolysis in bacterial metalloenzymes. *Australian Journal of Chemistry*, 69(12):1383–1395, 2016.

[281] Dave W. Anderson, Florian Baier, Gloria Yang, and Nobuhiko Tokuriki. The adaptive landscape of a metallo-enzyme is shaped by environment-dependent epistasis. *Nature Communications*, 12(1):3867, 2021.

Acknowledgements

I want to thank Prof. Dr. Reinhard Sterner for supporting me as well as my thesis for many years and for having more patience with me than I had with myself. This work would not have been possible without his vast experience and expert guidance. Likewise, I want to thank Prof. Dr. Rainer Merkl for years of personal and professional support. Furthermore, I would like to thank Prof. Dr. Thomas Dresselhaus for mentoring this thesis and Prof. Dr. Frank Raushel for agreeing to act as assessor. My gratidute further goes out to Dr. Chitra Rajendran, whose assistance was indispensible for the structural elucidation of RutB.
I would like to thank all colleagues from the Sterner group, particularly Thomas Klein and Caroline Hiefinger, for the conversations, discussions, and the pleasant work environment. Special thanks go out to Lukas Drexler for continuing the AtzB project and assisting me with the final experiments. A warm thanks to our technical assistants Christiane Endres, Sonja Fuchs, Sabine Laberer, and Jeanette Ueckert, whose tireless practical support has made my work much easier.
Furthermore, this work would not have been possible without the support of my friends and family.