Unusual Building Blocks and Domain Organization of Non-Ribosomal Peptide Synthetases

---

Ungewöhnliche Synthesebausteine und Domänenorganisation von Nicht-Ribosomalen Peptidsynthetasen

Dissertation

zur

Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem

Fachbereich Chemie

der Philipps-Universität Marburg

vorgelegt von

Matthias Strieker

aus Lingen an der Ems

Marburg an der Lahn, 2009

Vom Fachbereich Chemie
der Philipps-Universität Marburg als Dissertation
am 01.09.2009 angenommen.

Erstgutachter : Prof. Dr. M. A. Marahiel
(Philipps-Universität Marburg)
Zweitgutachter : Prof. Dr. C. T. Walsh
(Harvard Medical School, Boston, MA, Vereinigte Staaten von Amerika)

Tag der Disputation: 25.09.2009

Dedicated to my parents

in loving memory

Abstract

The diverse class of non-ribosomal peptides consists of manifold pharmacologically important natural products. They are clinically used in antibiotic, antiviral and antitumor therapy, furthermore some are known immunosuppresants. The biological activity is based on their structural diversity, as they contain various non-proteinogenic building blocks and amino acids of which many are β-modified. It was shown that the latter are important for biological activity, but little is known about their biosynthetic origin. In particular, these building blocks are key determinants of the class of acidic lipopeptide antibiotics and kutznerides, which are in the focus of this thesis.

To determine the mechanism underlying the biosynthetic origin of the synthetically challenging β-hydroxylated asparagine (hAsn) moieties, found in the acidic lipopeptides CDA and A54145, the corresponding recombinant non-heme iron (II)/α-ketoglutarate dependent hydroxylases AsnO and LptL have been examined *in vitro*. Direct hydroxylation of the free amino acid was observed in both cases, clearly indicating a precursor synthesis pathway. The crystal structure of one of the two hydroxylases (AsnO) was determined at high resolution and revealed a substrate induced fit mechanism of the enzyme. Upon addition of asparagine, a lid-like region seals the active site and shields it from sterically demanding substrates, which explains the observed specificity for free asparagine. Furthermore, the AsnO structure could be seen as an archetype enzyme for non-heme iron hydroxylases acting on free amino acids. It was possible to predict amino acid binding residues for homologous enzymes by 3D modeling.

In order to fully understand the mechanisms of β-hydroxylated building blocks synthesis, the hydroxylases KtzO and KtzP, predicted to be responsible for the generation of the two 3-hydroxyglutamic acid isomers found in the mixture of antifungal and antimicrobial kutznerides, were produced recombinantly and analyzed *in vitro*. Notably, they were found to work *in trans* to the assembly line on PCP-tethered glutamic acid rather than on the free amino acid. Unexpectedly, as the two isomers are found in approximately equal amounts in mature kutznerides, KtzO was shown to stereospecifically generate *threo*-hydroxyglutamate, while KtzP catalyzed the formation of the *erythro* isomer by co-elution HPLC experiments with synthetic dabsylated standards. A powerful method that employs non-hydrolyzable coenzyme A analogs was developed, which allowed the determination of the kinetic parameters of enzymes working on PCP-bound substrates for the first time. Furthermore, a hitherto unknown mechanism of NRPS assembly line restoration was observed. The corresponding adenylation (A) domain for glutamic acid activation in the kutzneride NRPS was found to be corrupted. Herein, it is shown that this lack of a functional A domain is compensated *in trans* by a stand-alone A domain. These findings elucidated the mechanism for the *in trans* compensation and the stereospecific hydroxyglutamate generation in detail and may guide the usage of *in trans* hydroxylation/compensation enzymes in biocombinatorial engineering approaches.

In the third part of this work, the acquired knowledge about the mechanisms underlying enzymatic β-hydroxylation of amino acids was exploited for the synthesis of the pharmaceutically relevant β-hydroxyaspartate. Primarily, this was facilitated by the structure elucidation of AsnO in which the substrate binding residues were identified. By site directed mutagenesis, an AsnO variant was generated, which notably did not hydroxylate the original substrate asparagine, instead it was found to stereospecifically catalyze the formation of L-*threo*-hydroxyaspartic acid, even in commercially interesting amounts. Therefore, the AsnO variant is an excellent example for the application of basic research in order to generate pharmacologically relevant non-proteinogenic amino acids.

Zusammenfassung

Die mannigfaltige Klasse der nicht-ribosomalen Peptide beinhaltet viele pharmakolisch relevante Wirkstoffe, die in der Klinik in der Antibiotika-, Krebs- und Immunsuppressionstherapie Anwendung finden. Die biologische Aktivität dieser Naturstoffe geht zurück auf deren strukturelle Vielfalt, da sie neben den proteinogenen Aminosäuren noch viele nicht-proteinoge aufweisen, von denen eine Vielzahl an der β-Position modifiziert sind. Es wurde gezeigt, dass diese Modifikationen zu der beobachteten Bioaktivität beitragen, aber die biosynthetischen Ursprünge sind noch weitestgehend unerforscht. Diese Bausteine sind Schlüsselmerkmale in den Naturstoffklassen der aziden Lipopeptidantibiotka und der Kutzneride, auf die sich diese Arbeit fokussiert.

Um den Mechanismus des biosynthetischen Ursprungs des synthetisch anspruchsvollen β-Hydroxyasparagin-(hAsn)-Restes, der Teil der aziden Lipopeptide CDA und A54145 ist, aufzuklären, wurden die entsprechenden nicht-Häm Eisen(II)/α-Ketoglutarat-abhängigen Hydroxylasen, AsnO und LptL, rekombinant hergestellt und *in vitro* untersucht. Für beide Enzyme konnte gezeigt werden, dass sie die freie Aminosäure regio- und stereospezifisch hydroxylieren, was eindeutig auf einen Vorläufermolekülsynthesemechanismus hinweist. Die Kristallstruktur von AsnO konnte gelöst werden und offenbarte einen unerwarteten substratinduzierten Mechanismus. In diesem schließt sich nach Substratzugabe ein Deckel über der aktiven Tasche und schirmt sterisch anspruchsvolle Substrate ab, was weiter die beobachtete Substratspezifität für freies Asparagin erklärt. Außerdem kann die AsnO-Struktur als Prototyp für Aminosäure-modifizierende nicht-Häm Eisen Hydroxylasen dienen. Durch Vergleich der AsnO-Strukur mit 3D Modellen von homologen Enzymen war es möglich, die substratbindenden Reste in den letzteren erfolgreich vorherzusagen.

Um einen umfangreichen Einblick in die Mechanismen der β-Hydroxybausteinsynthese zu erhalten, wurden die Hydroxylasen KtzO und KtzP, welche voraussichtlich die Synthese der 3-Hydroxyglutamat-(hGlu)-Isomere, die in den Antimykotika der Kutzneride gefunden wurden, katalysieren, rekombinant produziert und *in vitro* untersucht. Es wurde beobachtet, dass beide Hydroxylasen *in trans* zum NRPS Fließband an PCP-gebundenem Glutamat arbeiten. In den Kutzneriden sind die beiden hGlu Isomere in gleichen Mengen vorhanden, daher war die Beobachtung, dass KtzO stereospezifisch *threo*-hGlu herstellt, wohingegen KtzP die Synthese des *erythro*-Isomers katalysierte, unerwartet. Desweiteren wurde eine wirkungsvolle Methode entwickelt, in der nicht-hydrolisierbare Coenzym A Derivate zum ersten Mal eine direkte Bestimmung der kinetischen Eigenschaften von Enzymen, die an PCP-gebundenen Substraten arbeiten, möglich machten. Außerdem wurde ein bisher unbekannter *in trans* Mechanismus einer NRPS-Fließband Wiederherstellung beobachtet. In dieser Arbeit wird gezeigt, dass die einzelne Adenylierungs-(A)-domäne KtzN die verkürzte und daher nicht-funktionierende A-Domäne des Kutzneridfließbandes *in trans* kompensiert. Zusammengefasst erläutern diese Ergebnisse den Mechanismus, auf dem die *in trans* Wiederherstellung des Fließbandes und stereospezifische Hydroxyglutamatsynthese basieren.

Im dritten Teil dieser Arbeit wurde das gesammelte Wissen über die enzymatische β-Hydroxylierung von Aminosäuren angewendet, um eine Methode zur Herstellung des pharmakologisch relevanten β-Hydroxyaspartats zu entwickeln. In erster Sicht wurde dies möglich durch die Strukturaufklärung von AsnO, in der die Substrat-bindenden Reste des Enzyms identifiziert wurden. Durch gerichtete Mutagenese wurde eine AsnO Variante hergestellt, die anstelle des ursprünglichen Substrats Asparagin stereospezifisch Aspartat zu L-*threo*-Hydroxyaspartat umwandelt, sogar in kommerziell interessanten Mengen. Daher ist die geschaffene AsnO-Variante ein exzellentes Beispiel für die Anwendung von Grundlagenforschung für die Synthese von nicht-proteinogen Naturstoffbausteinen.

The majority of the work presented herein has been published:

Strieker, M.; Nolan, E. M.; Walsh, C. T.; Marahiel, M. A. Stereospecific Synthesis of *threo*- and *erythro*-β-Hydroxyglutamic Acid During Kutzneride Biosynthesis. *J Am Chem Soc* **2009**, *131*, 13523-13530.

Strieker, M.; Marahiel, M. A. The structural diversity of acidic lipopeptide antibiotics. *ChemBioChem* **2009**, *10*, 607-616.

Strieker, M.; Essen, L. O.; Walsh, C. T.; Marahiel, M. A. Non-heme hydroxylase engineering for simple enzymatic synthesis of L-*threo*-hydroxyaspartic acid. *ChemBioChem* **2008**, *9*, 374-376.

Fujimori, D. G.; Hrvatin, S.; Neumann, C. S.; Strieker, M.; Marahiel, M. A.; Walsh, C. T. Cloning and characterization of the biosynthetic gene cluster for kutznerides. *Proc Natl Acad Sci U S A* **2007**, *104*, 16498-16503.

Strieker, M.; Kopp, F.; Mahlert, C.; Essen, L. O.; Marahiel, M. A. Mechanistic and structural basis of stereospecific C_β-hydroxylation in calcium-dependent antibiotic, a daptomycin-type lipopeptide. *ACS Chem Biol* **2007**, *2*, 187-196.

Furthermore the enzymatic synthesis of L-threo-hydroxyaspartic acid has been patented:

Strieker, M.; Essen, L. O.; Marahiel, M. A. Protein for chemoenzymatic production of L-*threo*-hydroxyaspartate. Publication No**.:** WO/2008/125080; International Application No.: PCT/DE2008/000580 (07.04.2008). National Application No.: DE 10 2007 017 861.3 (13.04.2007).

Table of Contents

Inhaltsverzeichnis

1. Abbreviations

1.1 General Abbreviations

Abbreviation	Full name
A domain	adenylation domain
A_{core}	*N*-terminal core domain of A domains
ACP	acyl carrier protein
AcpS	acyl carrier protein synthase
ACS	acetyl-CoA synthase
αKG	α-ketoglutarate
AMP	adenosine monophosphate
AsnO	asparagine oxygenase
A_{sub}	*C*-terminal subdomain of A domains
ATCC	American Type Culture Collection
ATP	adenosine-5'-triphosphate
BLAST	Basic Local Alignment Search Tool
Boc	*tert*-butoxycarbonyl
C domain	condensation domain
CAS	clavaminate synthase
CDA	calcium-dependent antibiotic
CLEC	chiral ligand exchange chromatography
C-MT	carbon methylation domain
CNS	central nervous system
CoA	coenzyme A
CSA	(±)-camphor-10-sulphoric acid
Cy domain	heterocyclization domain
dabsyl	4-dimethylamino-azobenzene-4-sulphonyl chloride
DHB	2,5-dihydroxybenzoic acid or gentisic acid
DIPEA	*N*,*N*-diisopropylethylamine or Hünig's base
DMF	dimethyl formamide
DNA	deoxyribonucleic acid
DPCK	dephospho coenzyme A kinase
DSMZ	Deutsche Sammlung für Mikroorganismen und Zellkulturen
DTT	dithiothreitol
E domain	epimerization domain
EAA	excitatory amino acid
EAAT	excitatory amino acid transporter
EDA	ethylenediamine
EDC	*N*-(3-dimethylaminopropyl)-*N'*-ethylene carbodiimide hydrochloride
EDTA	ethylenediaminetetraacetic acid
Ent	enterobactin
ESI	electron-spray ionization
EXAFS	extended x-ray absorption fine structure
F domain	*N*-formylation domain
FAD	flavin adenine dinucleotide
Ferene S	3-(2-pyridyl)-5,6-bis(2-(5-furyl sulfonic acid))-1,2,4,-triazine
fH4F	formyltetrahydrofolate
FIH	factor inhibiting HIF
FMOC	fluorenylmethyloxycarbonyl
FPLC	fast protein liquid chromatography

FT-ICR	Fourier transform ion cyclotron resonance
GlmT	glutamate-3-methyl-transferase
Gtf	glycosyltransferase
HasP	hydroxyasparagine phosphotransferase
HBTU	O-Benzotriazole-*N*,*N*,*N*',*N*'-tetramethyl-uronium-hexafluoro-phosphate
HEPES	4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid
HIF	hypoxia-inducible factor
HmaS	4-hydroxymandelate synthase
Hmo	hydroxymandelate oxidase
HOAt	1-hydroxy-7-azabenzotriazol
HOBt	hydroxybenzotriazole
HpgT	4-hydroxyphenylglycine transaminase
HPLC	high performance liquid chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	kilo Dalton
Ktz	kutzneride
lac	lactose
mal	maltose
MBP	maltose-binding protein
MDa	mega Dalton
MeCN	acetonitrile
MIC	minimal inhibitory concentration
MIO	4-methylideneimidazole-5-one
mRNA	messenger ribonucleic acid
MRSA	methicillin-resistant *Staphylococcus aureus*
MS	mass spectrometry
MT domain	methylation domain
NDP	nucleoside diphosphate
NH_2-CoA	amino coenzyme A
N-MT	nitrogen methylation domain
NRP	non-ribosomal peptide
NRPS	non-ribosomal peptide synthetase
NRRL	Northern Regional Research Laboratory strain collection
NTA	nitrilotriacetic acid
O-MT	oxygen methylation domain
ORF	open reading frame
P4H	prolyl-4-hydroxylase
PADA	*p*-anisaldehyde dimethylacetate
PanK	pantotheine kinase
PCP domain	peptidyl-carrier protein domain
PCR	polymerase chain reaction
PDB	protein data bank
PK	polyketide
PKS	polyketide synthase
ppan	4'-phosphopantetheine
PPAT	phosphopantetheine adenosine transferase
PP_i	inorganic pyrophosphate
PPTase	4'-phosphopantetheine transferase
QTOF	quadrupole time-of-flight
R domain	reductase domain
r.m.s.d.	root mean square deviation

RNA	ribonucleic acid
RP	reversed-phase
RT	room temperature
SAM	*S*-adenosylmethionine
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
sp.	species
Srf	surfactin
Syr	syringomycin E
TauD	taurine dioxygenase
TE domain	thioesterase domain
TEII	type II thioesterase
TES	triethyl silane
TFA	trifluoroacetic acid
TIPS	triisopropyl silane
TLC	thin layer chromatography
TRIS	tris(hydroxymethyl)aminomethane
tRNA	transfer ribonucleic acid
Tyc	tyrocidine
Vio	viomycin
VRE	vancomycin-resistant *enterococcus*

1.2 Non-Proteinogenic Amino Acids

Abbreviation	Full name
Abu	aminobutyrate
alloCMA	(1*S,*2*R*)-allocoronamic acid
ADH	2-amino-3,5-dimethylhex-4-enoic acid
Dab	diaminobutyrate
Dap	diaminopropionate
ΔTrp	*Z*-2'-3'-dehydro-tryptophan
hAsn	3-hydroxyasparagine
hGlu	3-hydroxyglutamic acid
Hpg	4-hydroxyphenylglycine
KIV	keto-isovaleric acid
Kyn	kynurenine
L-THA	L-*threo*-3-hydroxyaspartic acid
MeAsp	L-*threo*-3-methylaspartic acid
MecPG	2-(1-methylcyclopropyl)glycine
MeGlu	3-methylglutamic acid
MeOAsp	3-methoxyaspartic acid
Orn	ornithine
PIC	(2*S*,3a*R*,8a*S*)-6,7-dichloro-3a-hydroxyhexa-hydropyrrolo[2,3b]indole-2-carboxylic acid
PIP	pipecolinic acid
Sar	sarcosine

2. Introduction

The complex structures of microbial natural products have fascinated chemists for decades. As the tools of chemistry and biochemistry were sharpened, huge advances in understanding natural product biosynthesis were made, but there were still barriers to a satisfactory understanding.[1] Many of these microbial natural products are secondary metabolites derived from non-ribosomal peptide or polyketide synthesis. During evolution these metabolites were optimized for chemical defense or communication and they show a broad range of biological activity.[2] Many non-ribosomal peptides (NRPs) and polyketides (PKs) are used as therapeutical agents to fight infectious diseases (erythromycin,[3] penicillin,[4] daptomycin[5]), cancer (epothilone[6]) or to suppress the immune response after organ transplantations (cyclosporine[7]). These molecules are biosynthesized by the consecutive condensation of amino acids or acyl-CoAs, which is achieved by large multi-domain enzymes, non-ribosomal peptide synthetases (NRPSs)[8,9] and polyketide synthases (PKSs)[10,11] that follow a parallel biosynthetic mechanism.

From the chemical point of view, NRPs are especially intriguing as they contain besides the 20 proteinogenic amino acids several hundred other building blocks. 504 different monomers, which include non-proteinogenic amino acids, fatty acids and α-hydroxy acids, have been identified to date for NRPs.[12] The non-proteinogenic building blocks contribute to structural versatility[13] of NRPs and are likely to be responsible for proper biological activity, as they e.g. prevent them to get proteolytically cleaved in the target strain. An important class of these unusual building blocks are β-hydroxylated amino acids, which are generated by non-heme iron(II) and α-ketoglutarate (αKG) dependent oxygenases[14] either in a precursor pathway or by *in trans* modification while the amino acid is tethered to the modular assembly line.[15,16]

This introduction gives an overview of NRPS biosynthetic principles, which follow and differ from the co-linearity principle[17] of NRP synthesis. Furthermore, it focuses on the biosynthesis of NRPs precursor biosynthesis and on *in cis* and *in trans* modification reactions. Additionally, an overview of non-heme iron oxygenases and their catalysis mechanism is given. Special attention is assigned to the medically important acidic lipopeptide antibiotics as well as to the unusual biosynthetic cluster organization and the structurally intriguing kutznerides.[18,19]

2.1 Non-Ribosomal Peptide Synthetases

Non-ribosomal peptide synthetases are large multimodular enzyme complexes that catalyze the synthesis of biologically active peptides,[20] in contrast to ribosomal peptide synthesis, in an mRNA-template free manner. NRPS encoding gene clusters are wide-spread in nature, they can be mainly found in fungi and microbes.[21] The largest NRPS known from the antibiotic producing fungi *Trichoderma virens* is the peptaibol synthetase, which encodes for a 20,925 residue containing protein. This synthetase has a weight of approximately 2.3 MDa, which is almost the size of the 70S ribosome of prokaryotes (2.52 MDa),[22] but on a single polypeptide chain. Gene clusters, which encode for secondary metabolite producing enzymes, make up to five or more % of the whole genome of the producer.[23] The answer, why organisms spend this much energy in producing non essential metabolites, is that they gain an enormous advantage over others i.e. by producing toxins or cytostatic, antimycotic, antibiotic or antiviral agents. The best examined NRPS producing organisms are soil actinobacteria, which have proven to be a rich source of potent therapeutic NRPs.[24] As bacteria are easier cultivatable and faster growing than fungi or plants, the major research on the synthesis of non-ribosomal peptides has focused on these producers during the last decades.

The results of the biochemical and biophysical examination of these producer strains and their corresponding NRPs producing enzymes are presented in this chapter. First, the general features of NRPSs are given, then the structure and architecture of single domains and of an entire termination module of the assembly line are presented, followed by non-essential domains acting *in cis* and finally modifying enzymes, which act prior to, during or after the peptide assembly are detailed. Details on the heterologous expression and manipulation of NRPS gene clusters can be found elsewhere.[25-28]

2.1.1 Structures and Mechanisms of Essential Domains

2.1.1.1 General Synthesis Mechanism

Many reviews explaining the basic chemistry of NRPS can be found in contemporary literature,[13,29,30] thus this introduction will just briefly summarize the biochemistry and will focus on domain structures instead. The unique structures of many hundreds of non-ribosomal peptides share a common mode of synthesis, which is the co-linear modularity principle of NRPSs.[17] Some recently discovered NRPS gene clusters implicate

that a different synthesis mechanism is possible,[31,32] but for most natural products the co-linearity principle is applicable.

Generally, NRPSs are comprised of an array of distinct modular sections, each of which is responsible for the incorporation of one defined monomer into the final peptide product.[33] The identity and order of a module in an assembly line specifies first the sequence of monomer units activated and incorporated, second the chemistry that occurs at each way station in the assembly line, and third the length and functionality of the product released from the distal end of the assembly line.[29] The modules can be further divided into catalytic domains. Three domains are ubiquitous in non-ribosomal peptide synthesis and essential for peptide elongation (Figure 2.1).

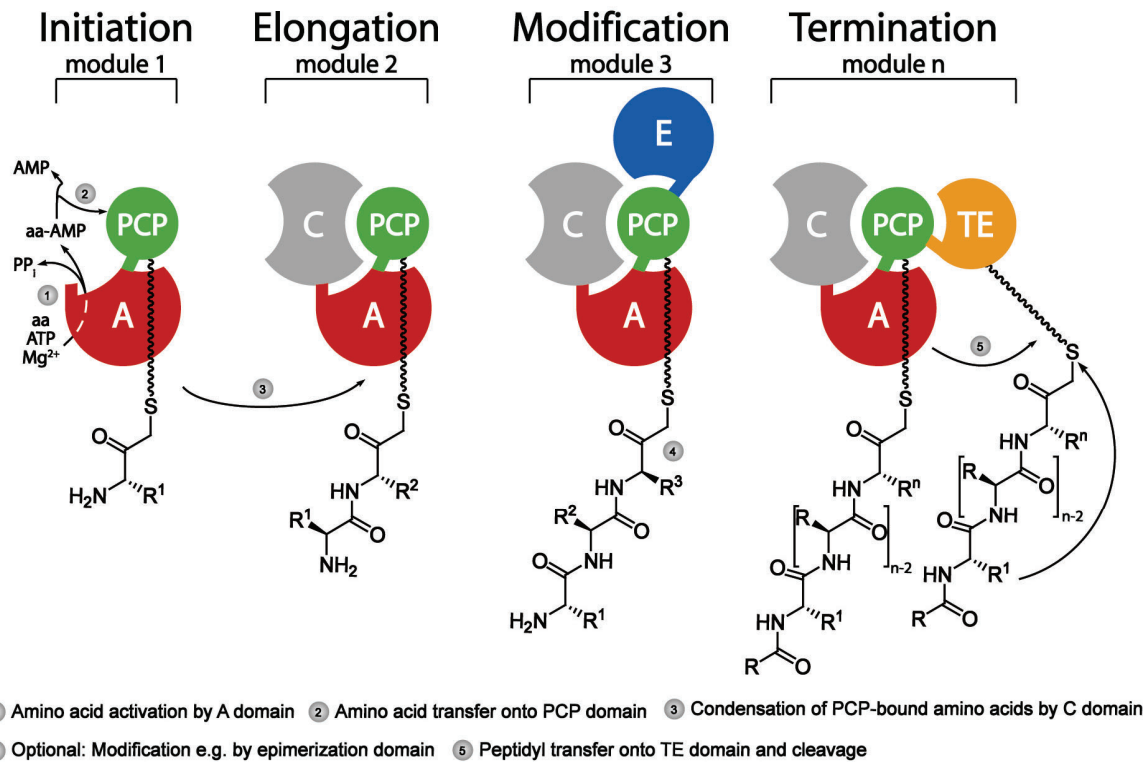


Figure 2.1: Mechanism of non-ribosomal peptide synthesis. The scheme details the sequential reactions catalyzed by individual NRPS domains during initiation, elongation, modification and termination.

The domains are responsible for the activation of the amino acid [adenylation (A) domains], the propagation of the growing peptide chain [peptidyl carrier protein (PCP) domains] and the condensation of the amino acids [condensation (C) domains]. A fourth essential NRPS catalytic unit associated with product release is the thioesterase (TE) domain (Figure 2.1). In contrast to the other, repetitively occurring NRPS domains, TE domains are only present once within an NRPS assembly line in a termination module

with the domain order C-A-PCP-TE.[34] An overview of domain structures in a termination module can be found in Figure 2.2. The structures of the stand-alone VibH[35] C domain, the DhbE[36] A domain, the TycC3[37] PCP domain and the Fen TE[38] domain are depicted.

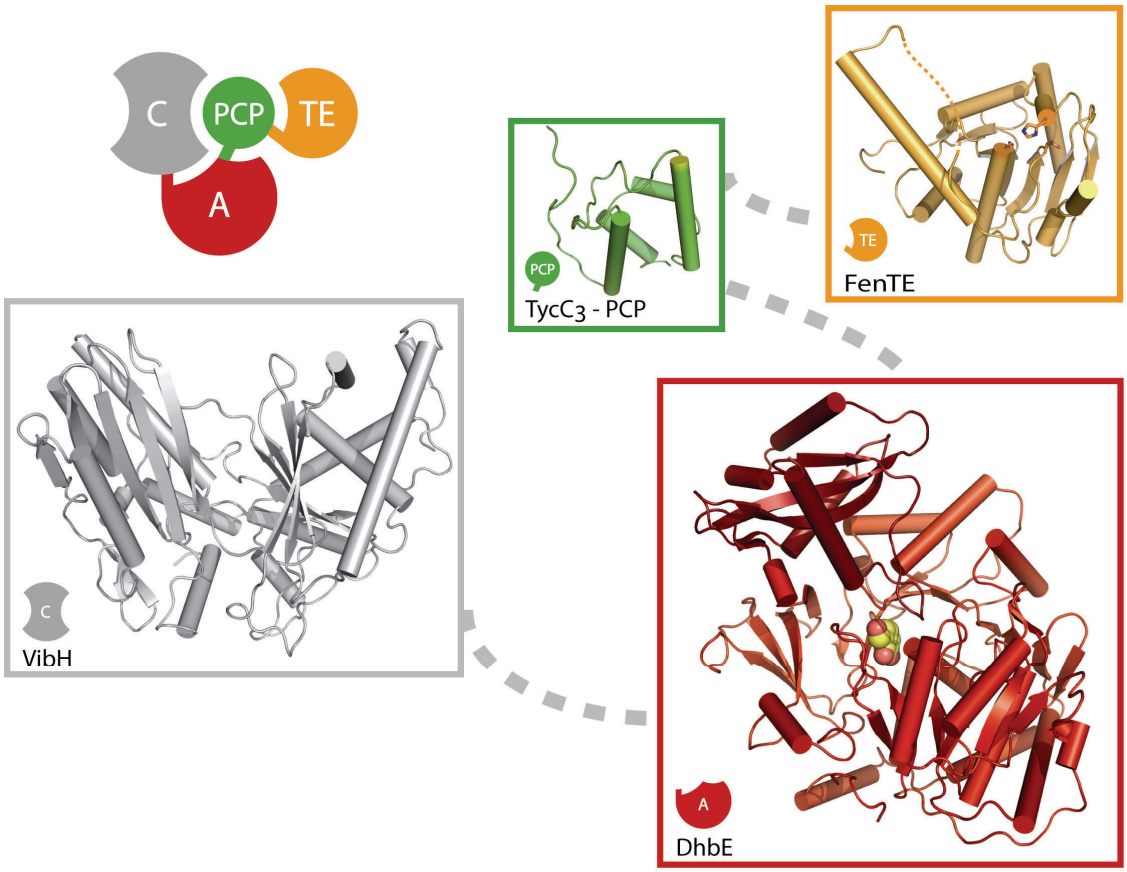


Figure 2.2: Overview of single domain structures. The structures of the C domain VibH, the A domain DhbE, the PCP TycC3 and the TE from the fengycin system are shown. The single domains are connected by dashed lines, illustrating the natural state.

2.1.1.2 Adenylation Domain

The adenylation domains are highly selective enzymes, which act as gatekeepers for incoming monomeric building blocks. Each adenylation domain is approximately 550 amino acids in size and comprises a large *N*-terminal core domain (A_{core}, ~ 450 aa's) and a small *C*-terminal subdomain (A_{sub}, ~ 100 aa's), which are connected by a small hinge (5-10 residues). The core and subdomain organization of A domains was observed in all three currently known A domain structures. These are the 2,3-dihydroxy benzoate (DhbE,[36] Figure 2.2), the L-phenylalanine (PheA,[39] Figure 2.3), and the D-alanine (DltA,[40] Figure 2.3) activating adenylation domains. Furthermore, the described fold is highly conserved even in other adenylate generating enzymes, like acetyl-CoA synthase (ACS)[41]

isolated from *Salmonella enterica* or firefly luciferase (Figure 2.3) from *Photinus pyralis*.[42]

Notably, the small subdomain A_{sub} undergoes a significant structural reorganization during the catalytic cycle of amino acid adenylation of altogether 140° (Figure 2.3). In the open state the substrate amino acid, cofactor Mg^{2+} and ATP can be bound. After substrate binding, the subdomain bends towards the core domain and a closed conformation is achieved in which the amino acid gets adenylated. After pyrophosphate release, A_{sub} approaches the active site further and shields the AMP-amino acids from the bulk solvent.[43] The latter conformation allows the transfer of the aminoacyl moiety onto the ppan-arm of the downstream PCP domain (Figure 2.3).

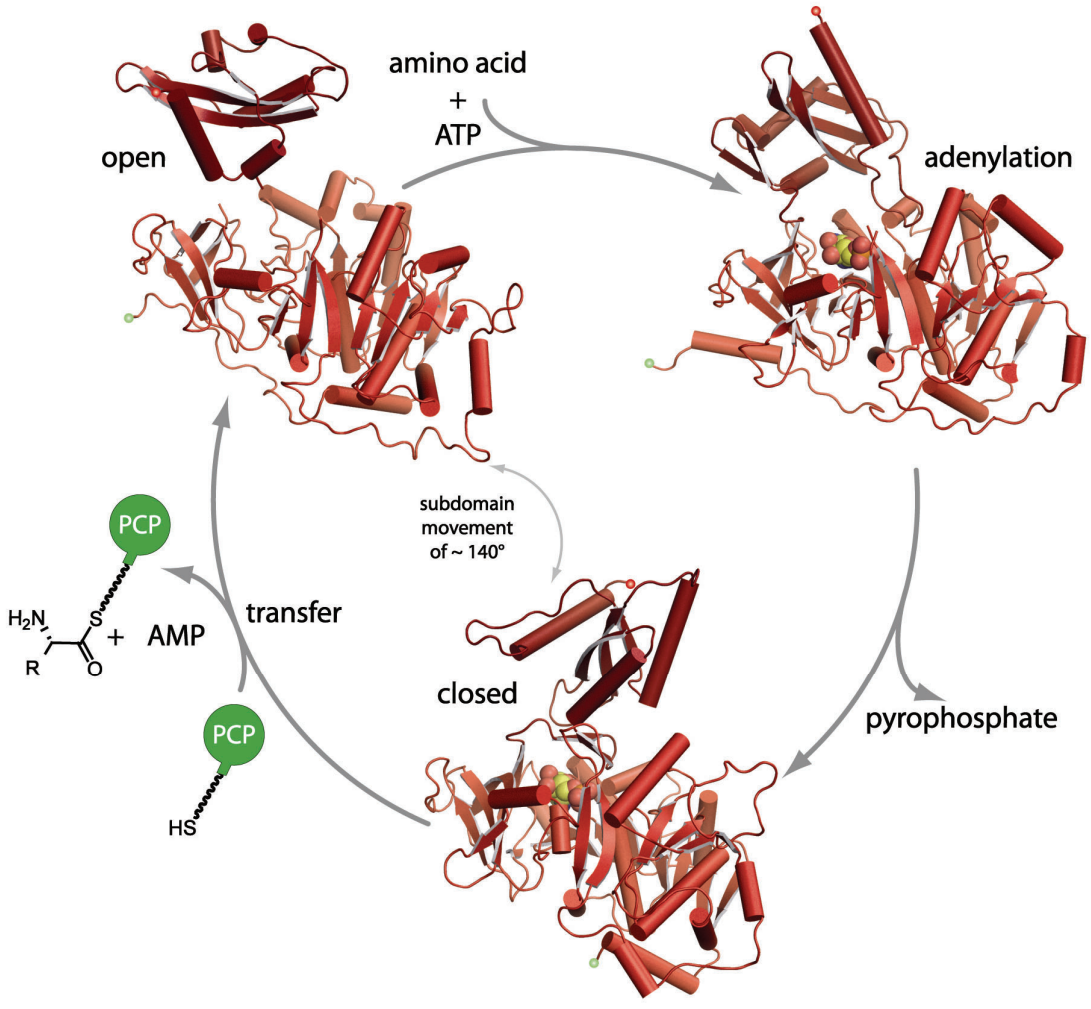


Figure 2.3: Adenylation domain catalysis cycle. The open form of the A domain (as observed for luciferase) is able to bind the amino acid and ATP. The adenylation intermediate as found for PheA then generates the acyl-AMP and pyrophosphate is cleaved off. The acyl-AMP intermediate is protected from bulk solvent in a closed formation (observed in the DltA structure), which facilitates transfer onto the PCP domain. The subdomain A_{sub} is shown in dark red, the core domain A_{core} in light red, the *N*-terminus has a green, the *C*-terminus a red dot.

Pioneering work from the late 1990s on a large number of adenylation domains, in which biochemical, structural and phylogenic studies were combined, revealed the so called specificity-conferring code of A domains.[44] This code is comprised of ten amino acids, which are by sequence alignment with PheA,[39] responsible for substrate binding within the active site of A domains. For α-amino acids two out of the ten residues are invariant, the α-carboxyl coordinating lysine and the α-amino coordinating aspartic acid residues. Subsequent research and bioinformatics allowed the prediction of the substrate amino acids directly from the primary sequence.[45] Nowadays, a couple of online tools for A domain specificity prediction are available[46,47] and might facilitate the isolation of predicted NRPS products in genome mining approaches.[48,49] In contrast to aminoacyl-tRNA-synthetases,[50] A domains lack a proof-reading mechanism, combined with relaxed substrate specificity, this often results in the synthesis of NRPs with different amino acid composition but produced by the same NRPS.

2.1.1.3 Peptidyl-Carrier Protein or Thiolation Domain

The peptidyl-carrier protein or thiolation domain is only 80 amino acids in size, but as it is responsible for the transportation, propagation and presentation of the aminoacyl or peptidyl substrates of the growing NRP chain, it is from paramount importance for a functional NRPS assembly line. The simple four helix bundle protein[37,51] is posttranslationally primed with a 4'-phosphopantetheine (ppan) cofactor, which is attached to a highly conserved serine residue (GGXS-motif) by cognate ppan-transferases (PPTases, e.g. Sfp, see below).[52-54] To this ppan-arm, all substrates and peptidyl intermediates are attached as thioesters during non-ribosomal synthesis. The mobile arm delivers these acyl/peptidyl-*S*-intermediates to the catalytic sites of the adjacent NRPS domains, for peptide bond formation, modification or product release.[29,55]

In NMR studies on the TycC3-PCP domain, a carrier protein dissected from the 7th module of the tyrocidine assembly line, in its apo- and holo-forms different chemical shifts suggest the existence of three different and slowly converting conformations: the apo (A), the A/H and the holo (H) state. Furthermore, in the apo-PCP, the A and A/H conformations coexist and in the case of holo-PCP the presence of H and A/H states were identified (Figure 2.4).[51] Thus, the common state for apo and holo is the A/H state.

The PCP state can be frozen in the A-state by mutation of the ppan-attachment site serine and in the H-state by acylation of the ppan-arm. The most compact PCP conformation resembling a classical four-helix bundle is the common A/H form. The apo-form (A) is less compact and has a large loop region between helix I (Figure 2.4, cyan) and helix II (Figure 2.4 yellow), whereas these helices change their relative orientation in the H-state.

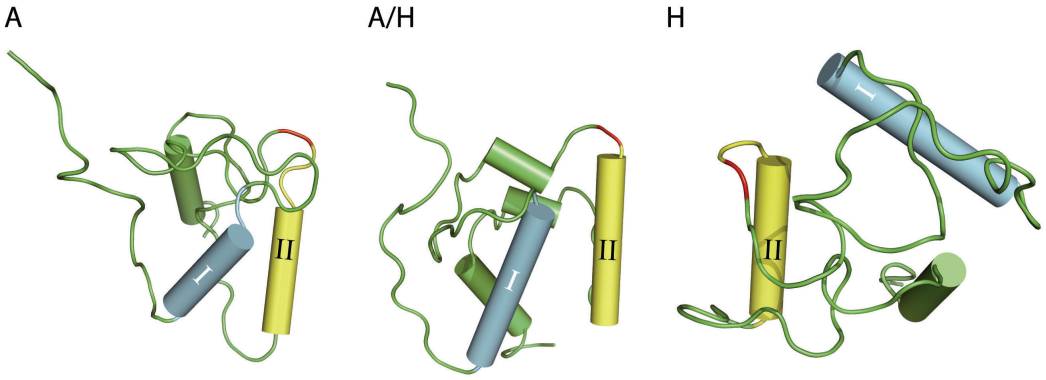


Figure 2.4: The TycC3-PCP domain in different states. The apo (A) state is slowly converted into the A/H state. The latter is also the common conformation of the holo (H) state. Between the apo and holo state, helix I (cyan) moves about 100°. The ppan-cofactor attachment site serine is shown in red. By superpositioning of all three states onto helix II (yellow), the movement and structural rearrangement is illustrated.

Additionally in the NMR study, the interactions of the amide protons and nitrogens of the cofactor in the H and A/H-state have been monitored. It was found that the cofactor relocates its position on the PCP surface of more than 100°. A consequence of this is that the terminal SH-group of the ppan-arm approximately is able to move 16 Å. With these findings, the first evidence for the long proposed swinging mode of the ppan prosthetic group during non-ribosomal peptide synthesis was given.[51]

2.1.1.4 Condensation Domain

Condensation domains, which are approximately 450 residues in size, carry out the peptide bond formation within an elongation module (Figure 2.1). Usually the number of C domains is in agreement with the number of peptide bonds found in the mature NRPS products (co-linearity principle). C domains contain an acceptor and a donor site, which harbor aminoacyl-*S*-PCP and peptidyl-*S*-PCP substrates, respectively.[56,57] The peptide bond formation is thought to be initiated by nucleophilic attack of the α-amino group of the acceptor substrate onto the thioester group of the donor substrate. During this reaction, the amide bond is formed and the entire peptide is transferred from upstream

donor site to downstream acceptor site. After peptide elongation, the extended chain serves as a donor substrate in the downstream condensation step. Although the reasons are not fully elucidated yet, it was found that all condensation reactions are strictly unidirectional towards the termination module.[57,58] Substrate specificity of the C domain were tested by PCP misloading studies, in which various synthetic aminoacyl-CoA substrates were attached to apo-PCPs of minimal dimodular NRPSs using the promiscuous ppan-transferase Sfp[53] (see below). Mischarged carrier molecules were then tested in condensation assays, which revealed strict substrate specificity at the acceptor site and a relaxed at the donor site. Neither D-amino acids instead of L-configured nor different aminoacyl side chain length were accepted at the acceptor site, whereas the donor site accepted different side chains.[57,58]

The biochemical studies were complemented by crystal structures of the stand-alone C domain VibH[35] (part of the vibriobactin biosynthesis machinery)[59,60] and of a PCP-C didomain derived from modules 5 and 6 of the tyrocidine synthetase C (Figure 2.5).[61]

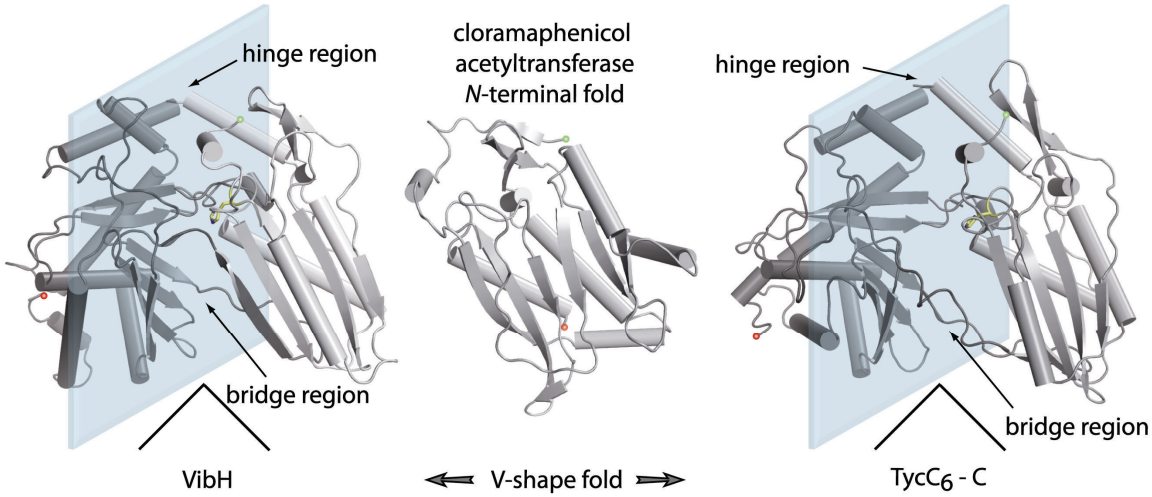


Figure 2.5: Overall structures of the stand-alone C domain VibH (vibrobaction biosynthesis) and $TycC_6$ C domain (tyrocidine biosynthesis). Both enzymes show a V-shaped fold. The N-terminus of each C domain show high homology to the chloamphenicol acetyltransferase (CAT).[62] A small hinge region connects the C-terminal domain from CAT-like half of the V. Furthermore, a bridging loop is observed, connecting the two halves in the middle. The catalytic histidine is shown in yellow. The *N*- and *C*-terminus of each enzyme is illustrated by a green and red dot, respectively.

Instead of a classical deep cavity for substrate binding, a canyon-like structure is found in the overall fold. The V-shaped fold allows the two up- and downstream PCP domains to position the condensation substrates at each opening (acceptor and donor sites). The open V-shape allows the acceptor and donor substrates to orient in a way that they

could reach the highly conserved catalytic histidine (Figure 2.5, shown in yellow) of the HHxxxDG motif. The condensation domains are found to consist of two similar big subdomains of which the *N*-terminal one shares high sequence homology to the well-known chloramphenicol acetyltransferase fold.[62] The two equally big subdomains are arranged in a V-shaped fashion, connected only by a small hinge region at the bottom of the V and by a bridging loop in the middle (Figure 2.5), which results in a canyon-like active site structure with the catalytically active His in the middle. Some mutational analysis suggested that the second histidine (His 224 in PCP-C) of the HHxxxD motive may act as a catalytic base by deprotonation of the α-ammonium group of the acceptor substrate to restore its nucleophilicity to attack the thioester of the donor substrate. More recent mutational studies and pK_a value calculations of active site residues suggested that peptide bond formation may depend on electrostatic interactions rather than on general acid/base catalysis.[61] However, the exact mechanism has not been determined yet.

2.1.1.5 Thioesterase Domain

The thioesterase domain is the fourth essential subunit of NRPSs. TE domains are located in the termination module of the assembly line (Figure 2.1), as they catalyze the product release either by hydrolysis or by cyclization. In some cases TE domains are not part of the assembly line and are found as single enzymes downstream of the cluster. They contain approximately 230 to 270 residues and often catalyze a two-step product release mechanism: First, formation of an acyl-O-TE intermediate at the active site serine, second, cleavage by intramolecular nucleophilic attack of water or of an internal nucleophile resulting in hydrolysis or in the generation of a cyclic or branched-cyclic peptide.[63,64] Due to the latter activity TEs are also referred to as cyclases. Instead of cyclization or hydrolysis, reduction of the thioester can be carried out for product release by a C-terminal NADPH-dependent reductase domain resulting in imine macrocycles.[65]

The crystal structures of two TE domains were elucidated in the last years, namely the surfactin (SrfTE)[66] and the fengycin (FenTE)[38] thioesterase (Figure 2.6). TEs show a conserved α/β-hydrolase fold similar to serine esterases and lipases.

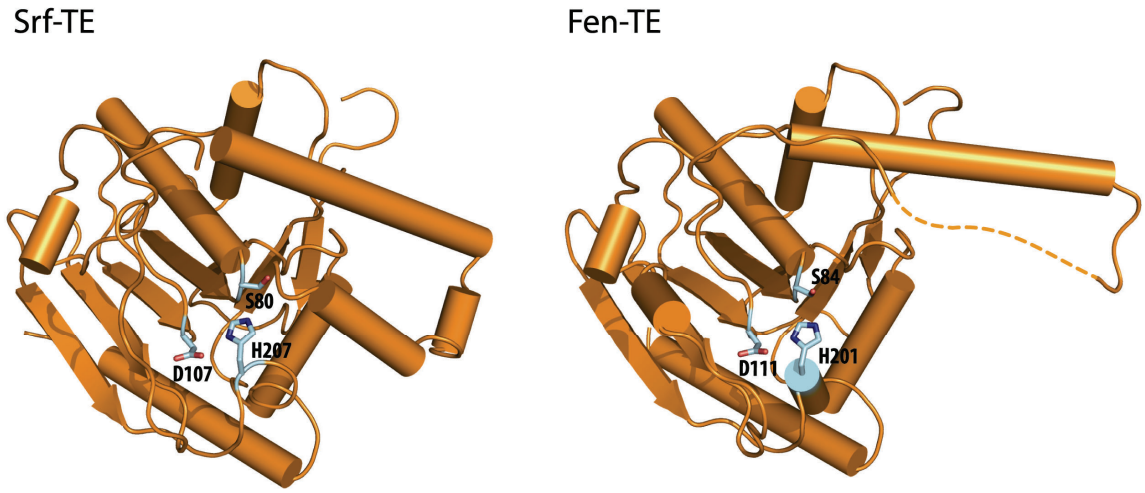


Figure 2.6: Crystal structures of SrfTE and FenTE. The TE domains show a conserved α/β-hydrolase fold. The catalytic triad Ser-His-Asp is highlighted in light blue and is shielded from the bulk solvent by an α-helical lid region.

The catalytic triad (Ser-His-Asp) is found in a canyon-like structure, which likely accommodates the linear peptide, where it is shielded from bulk solvent by an α-helical lid region to catalyze cyclization rather than hydrolysis. An open conformation would allow substrate entry, whereas the closed lid would resemble the cyclization case.[66] In an alternative model, hydrolysis is prevented by the substrate itself, where it seals the active site serine from water, in a so called edge-on binding mechanism.[38] The exact mechanism is still unknown, although NMR studies of a PCP-TE didomain from the enterobactin system were carried out recently. In that study the active site serine of the PCP was mutated to alanine, thus no conclusions could be made for the exact mechanism of substrate binding in the TE domain.[67]

TE domains are catalytically independent working subunits of an NRPS assembly line as exemplified by using excised TEs and chemically synthesized peptidyl-*S*-*N*-acetylcysteamines or peptidyl-*S*-thiophenols to evaluate substrate specificity. The results of these studies were that only the substitution of the nucleophile or the electrophile or changes in the substrate length were not tolerated by the excised TEs.[63,68]

2.1.2 SrfA-C – Crystal Structure of a Termination Module

Very recently, the crystal structure of the four domain containing termination module was determined.[69] The module contains 1274 residues and the four essential domains in the order C-A-PCP-TE. The individual fold of the domains within the SrfA-C structure show the same overall fold as previously found in the dissected domains (Figures 2.2 and

2.7 A). It was necessary to mutate the ppan binding serine within the PCP domain to obtain a uniform *apo* population of protein and to reduce the mobility of domains during crystallization.[69]

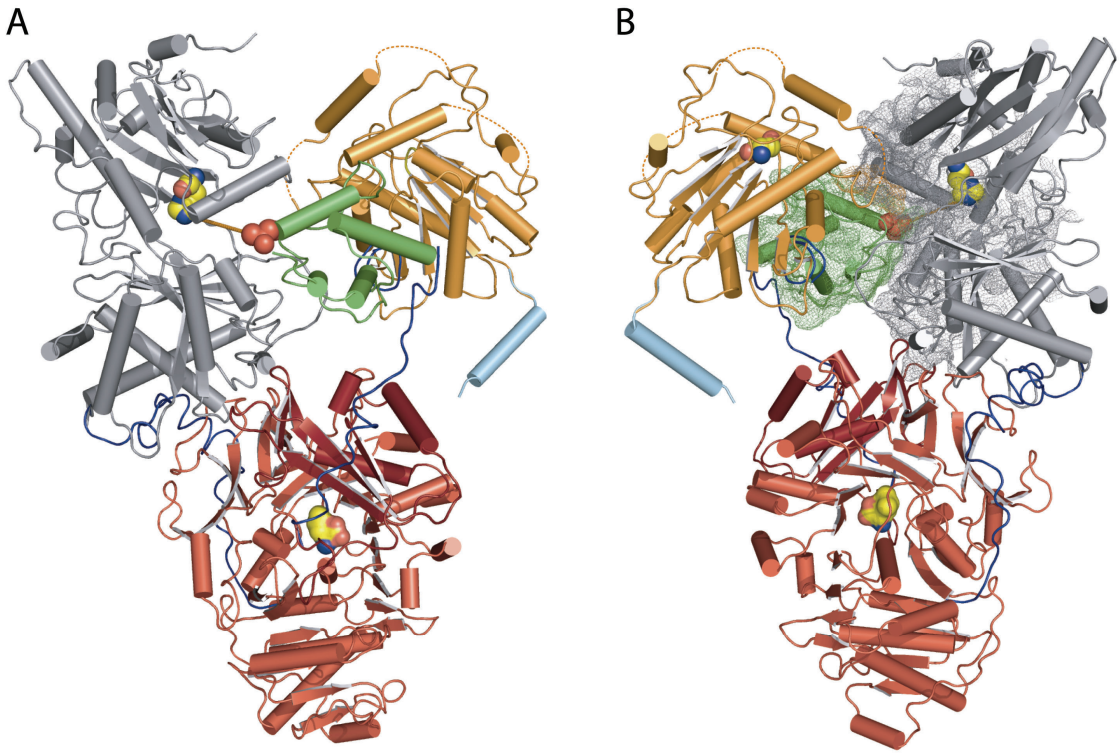


Figure 2.7: Crystal Structure of the surfactin termination module SrfA-C. A) Overall fold of the module. The C domain is shown in grey, the A domain in red, the C-terminal A_{sub} domain in dark red, the PCP domain in green, the thioesterase in orange and the myc-helix in cyan. The serine to alanine mutation of the ppan-attachment site of the PCP is highlighted in red spheres. The active site histidine of the C domain is also depicted in spheres. Interdomain linkers are colored in blue. Dashed orange lines in the TE domain show unordered regions. B) Back view of the module structure. The 20 Å mesh around the ppan-attachment site of the PCP shows the reachable region by the ppan-arm. The active sites of the TE and the A domain substrate Leu (shown in spheres) cannot be reached. Domains and linkers are colored accordingly to A).

The structural core of the SrfA-C module is a compact rectangular catalytic platform built by the C-domain and the major part of the A-domain with both active sites arrayed on the same site of the platform. The strong interaction between the A_{core} and the C domain explains why these two domains have to be shuffled together in NRPS reengineering approaches, which was observed in biocombinatorial engineering trials for the daptomycin assembly line.[25,70] In this structure the adenylation domain is positioned in a conformation similar to the adenylate-forming conformation. The *C*-terminal subdomain is rotated away from the *N*-terminal domain by approximately 40° compared to the PheA structure, resulting in a more open active site. The PCP domain and the A_{sub} domain are tethered together by a flexible 15-residue linker and located on

top of the catalytic platform. They have little interaction with platform and thus can move to adopt the required conformation during the catalytic cycle as the active sites of the A and the C domains are separated by approximately 60 Å. In the observed state for the PCP domain, it is impossible to cover this distance, as the ppan-arm is only 20 Å in length. The regions, which could be reached in the observed SrfA-C structure by the ppan cofactor is covered with a mesh in Figure 2.7 B.

The PCP domain of the SrfA-C protein is positioned to interact productively with the upstream condensation domain (Figure 2.7). The domain is likely to be positioned in the acceptor site of the C domain hence the catalytic center of the PCP domain is only 16 Å from the active site histidine of the C-domain (Figure 2.7 A). The PCP active site serine is not positioned where it could donate the pantetheine arm to either the adenylation domain or the thioesterase domain. The conformational mechanism to direct the substrate to the downstream thioesterase domain remains to be determined.

A finding which might assist in the *in trans* association of engineered NRPS, is the discovery of a "COM-hand", a motif found at the N-terminus of the SrfA-C module and which is able to bind a C-terminal α-helix of an upstream NRPS module. In the SrfA-C structure this is resembled by the myc-epitope-His_6-affinity tag fused to the *C*-terminus of the module, which forms an α-helix, which perfectly fits into the *N*-terminal hand-shaped motif of the SrfA-C C domain.[69]

2.1.3 Additional NRPS Domains and Related Enzymes

In addition to the presented essential domains, auxiliary domains are known and are responsible for the *in cis* modifications such as epimerization (Figure 2.1), cyclization and others.[15] Furthermore, an NRPS repair mechanism by a type II thioesterase,[71-73] which regenerates misprimed PCP domains, is presented as well as the priming of PCPs by ppan-transferases.[52,53]

2.1.3.1 In Cis Acting Modification Domains

NRPS modules can comprise auxiliary domains to catalyze substrate modifications, such as epimerization at C_α atoms (E), methylation (MT), *N*-formylation (F), or heterocyclization (Cy) of the peptidic backbone, utilizing serine and threonine side chains.[13,15]

Methylation. Often methylated amino acids are found in NRPS products. These are introduced by methyltransferases (MT). They catalyze the transfer of a methyl group from *S*-adenosylmethionine (SAM) to the carbon, nitrogen or oxygen atoms at various positions on the backbone of non-ribosomal peptides, therefore have been classified as C-MT, N-MT and O-MT, respectively, depending upon their site of methylation. These enzymatic domains in general have a bidomain structure, where the first subdomain contains the binding site for methyl group donor, while the second subdomain harbors the binding site for acceptor substrate. Various studies have suggested that the size of N-MT domain is typically 450 amino acids, while C-MT and O-MT are generally 300 amino acids long. A set of 3 conserved sequence motifs has been identified in most MTs.[74-76] In the following, some selected instances of the huge class of MTs are given. An excellent example is the cyclosporin synthetase, where 7 of the 11 amino acids are *N*-methylated.[77] The typical domain order in these seven modules of CssA synthetase is C-A(MT)-PCP, supporting the concept that each aminoacyl-*S*-PCP thioester intermediate in these modules can then be *N*-methylated via the cosubstrate *S*-adenosyl methionine. *N*-methylation prior to condensation was also observed in a domain-swapped actinomycin synthetase, indicating that an A-MT pair can retain its activities when moved to a new location in the synthetase assembly line.[78] *O*-Methylation is observed for A domains activating serine or threonine residues as in the cases of kutznerides[18,19] and perthamide.[79]

Formylation. *N*-formylated residues are found in linear gramicidin produced by *Bacillus brevis* ATCC 8185 and in the anabaenopeptilides from *Anabaena* strain 90.[80,81] The *N*-formylated Val, found in the linear gramicidin is generated by the F domain of the first module of the corresponding peptide synthetase LgrA. The formylation domain transfers the formyl group of formyltetrahydrofolate (fH4F) onto the amino acid using both cofactors N10- and N5-fH4F, respectively. Interestingly, the necessity of the formylated starter unit formyl-valine for the initiation of the gramicidin biosynthesis was observed. Without formyl group donor, no condensation product of valine with the subsequent building block glycine was detected.[81]

Epimerization. One of the hallmarks of non-ribosomal peptide natural products is the presence of D-amino acids. For example, four of the seven amino acids of the aglycone of vancomycin[82] and chloroeremomycin[83] have the D-configuration. In general, the A

domains of NRPS assembly lines select the readily available L-amino acids for activation, and epimerization (E) domains epimerize them once they are installed as aminoacyl- or peptidyl-*S*-PCP acyl enzyme intermediates. Exceptions to this rule are found in cyclosporin and in the *Helminthosporium carbonum* toxin assembly, where the corresponding monomer of the first and third module, respectively, is provided in a racemase catalyzed precursor mechanism.[84,85] Condensation domains with dual epimerization and condensation activity have also been discovered.[86] The E domains embedded in assembly lines occur precisely in the modules responsible for incorporating D-monomers into the growing chain. E domains in initiation modules rapidly produces an equilibrium mixture of the PCP-*S*-L,D-monomer, therefore the downstream C domain must be D-specific in its peptide-bond forming step.[87] For E domains embedded in elongating modules it was shown that efficient epimerization occurs not at the L-aminoacyl-*S*-PCP stage before condensation but rather at the peptidyl-*S*-PCP before transfer to the next elongation module of the assembly line. In both initiation module (aminoacyl-*S*-PCP) and elongation modules (peptidyl-*S*-PCP), the adjacent thioester linkage provides the thermodynamic accessibility and kinetic stability to a C-carbanion required for epimerization. It is then expected that all C domains immediately downstream of E domains will be D-specific for peptidyl donor chirality.[57]

Cyclization. Heterocyclizations of cysteine and serine or threonine residues to thiazoline and oxazoline five-membered ring heterocycles, respectively, are found in several NRPS products. These heterocycles alter the residue connectivity patterns and provide both metal-chelating and intercalating groups to the modified peptides. The responsible enzymes, the cyclization domains, are variants of C domains. Cy domains first catalyze peptide bond condensation and then carry out cyclization of the thiol sidechain of cysteine or the hydroxyl sidechain of serine or threonine onto the just-formed peptide bond to form thiohemiaminal or hemiaminal intermediates that are then dehydrated to yield the C=N bond in the thiazoline and oxazoline rings.[88] In yersiniabactin, bleomycin, and pyochelin biosynthesis, two adjacent Cy-A-PCP modules activate cysteines, and cyclize them into tandem 4,2-linked bis-heterocycles. In the bleomycin case, the oxidized bisheterocycle is the DNA-intercalating moiety of this clinically used antitumor drug. In the vibriobactin synthetase assembly line two threonyl-*S*-PCPs are heterocyclized, not in a 4,2-tandem linkage but instead are transferred to N1 and N5 of norspermidine.[15]

2.1.3.2 Repair Mechanism – Thioesterase II

In addition to the *normal* thioesterases found in termination modules of NRPS assembly lines, a second type of thioesterases (TEII) is found. The latter is a crucial repair enzyme for the regeneration of functional ppan cofactors of holo-T domains of NRPS and PKS systems.[72,73,89] Mispriming of ppan cofactors by acetyl- and short-chain acyl-residues interrupts the biosynthetic system. This repair reaction is very important, because approximately 80% of CoA, the precursor of the ppan cofactor, is acetylated in bacteria. Recent NMR solution structure elucidation of the surfactin TEII (Srf-TEII) in complex with a PCP domain from the tyrocidine biosynthesis system (TycC3-PCP) provided insights into substrate selectivity and mode of interaction of TEII and PCP in this repair mechanism.[71]

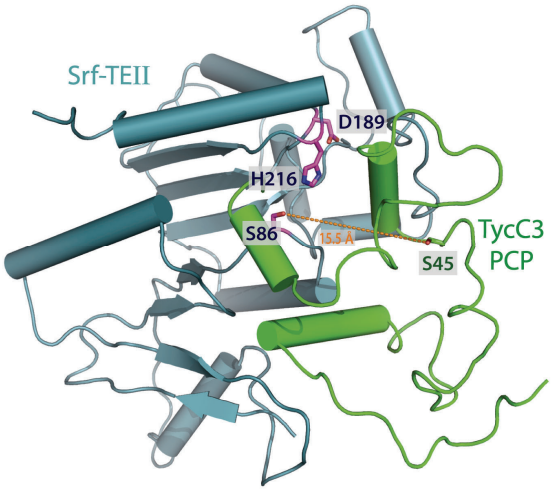


Figure 2.8: The type II thioesterase of the surfactin biosynthesis cluster. The catalytic triad S86-H216-D189 is more accessible than in normal thioesterases. The active site comprises a shallow groove, as resembled by small distance between the active sites of the holo-PCP and the TEII.

SrfTEII shows the typical α/β-hydrolase fold found for thioesterases (Figure 2.6, 2.8). Comparison with the structure of the *cis*-acting SrfTEI shows a similar overall architecture, but with the notable difference that the active site residues in the SrfTEII structure (S86, D189,H216) are just partially covered by a short loop (Gln 138–Ala 144) and are more accessible compared to Srf-TE, resulting in a shallow but easily accessible active site. This allows unspecific interactions of the TEII with small acyl-misprimed PCP, which explained the observed promiscuity of TEII's e.g. in the surfactin cluster[90] and why peptidyl substrates are not cleaved off.

2.1.3.3 Ppan-Transferases

All peptidyl- or acyl-carrier protein domain (PCP or ACP) containing multimodular enzymes such as PKS and NRPS require activation by a phosphopantetheinyl transferase.[52] Carrier proteins are activated by ppan-transferases (PPTase) by the addition of a 340 Da 4'-phosphopantetheinyl arm, which is covalently bound from the coenzyme A (CoA) substrate. PPTases are also essential for the activation of carrier proteins in fatty acid and heterocyst synthesis.[91-93] While all organisms require a PPTase for primary metabolism, some also encode additional PPTases that are capable of activating carrier proteins from secondary metabolite pathways.[94,95] PPTases can be classified into two main groups: the acyl carrier protein synthase (AcpS)-type PPTases, which are capable of activating ACPs from PKSs and fatty acid synthesis,[52,96] and the Sfp-type PPTases, which are named for the broad range activity PPTase from surfactin synthesis in *Bacillus subtilis* and typically activate NRPS/PKS multienzymes.[95,97] As Sfp was shown to transfer the ppan-arm to a wide set of PCPs,[98] it is important for the manipulation and artificial loading of PCP-domains, and thus this enzyme is introduced here more detailed.

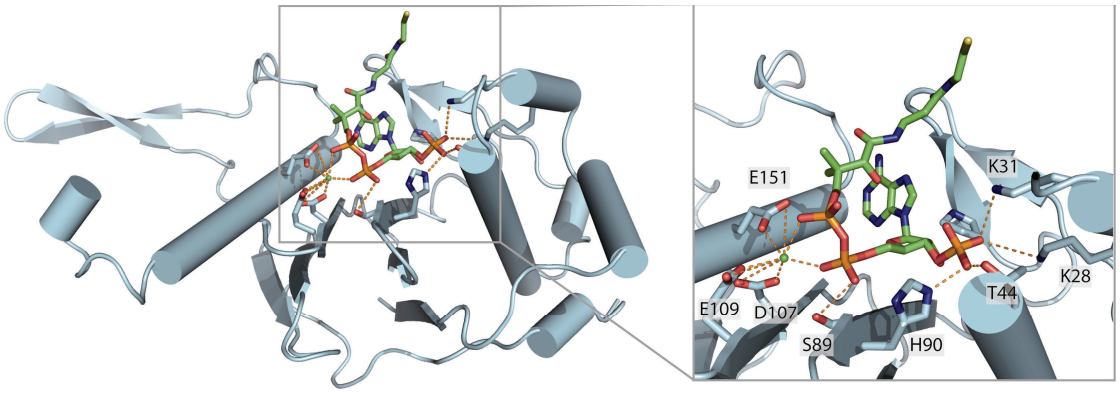


Figure 2.9: Overall fold and CoA binding site of Sfp. The CoA is shown in green, a detailed view on the CoA binding site reveals that CoA is mainly bound via its phosphate groups and that the magnesium cofactor is needed. The thiol end of CoA is not bound by Sfp at all, explaining the observed promiscuity of the PPTase.

The structure of Sfp (Figure 2.9) was elucidated ten years ago and revealed a novel α/β-fold.[53,99] The active site (detailed in the magnification in Figure 2.9) harbors a magnesium ion which is complexed by the pyrophosphate of CoA and the three acidic side chains of E151, E109 and D107. This complex resembles the main interaction between Sfp and CoA as well as the necessity of the presence of Mg^{2+}. In addition, the 3'-phosphate is bound tightly to the protein by the side chains of T44, K28, K31 and by the ε-nitrogen of the imidazole group of H90. The most important finding of the

described work was that the pantetheinyl part of the CoA substrate is not bound at all, which allowed the usage of Sfp to transfer non-cognate aminoacyl or peptidyl-ppan residues onto PCP domains for artificial loading of PCP domains. The described method was used in manifold assays, such as intentional PCP mispriming, examination of condensation domains, evaluation of on-line tailoring enzyme activity and many more.[98,100] The fact that the thiol group of CoA is also not bound by Sfp allowed the transfer of aminoacyl CoA analogs.[101] A finding which was exploited in this study by the usage of amino CoA as described in chapter 5.2.3.

2.2 Tailoring and Modification of NRPS Building Blocks

Beside the *in cis* acting modification domains, a vast number of so called *in trans* acting tailoring enzymes are associated with non-ribosomal peptide synthesis.[16] These enzymes contribute to a large extend to the remarkable structural diversity found in NRPs either in precursor synthesis of non-proteinogenic building blocks or *in trans* modification while the amino acid is bound to the corresponding carrier protein domain.[15] Often these modifications are carried out at unreactive aliphatic carbon atoms, which makes this class of enzymes interesting from a chemical point of view.[13] An incomplete compendium of possible carbon functionalized NRPs is given in Figure 2.10. Furthermore, post-assembly decoration with sugar or prenyl[102] moieties is known.[103]

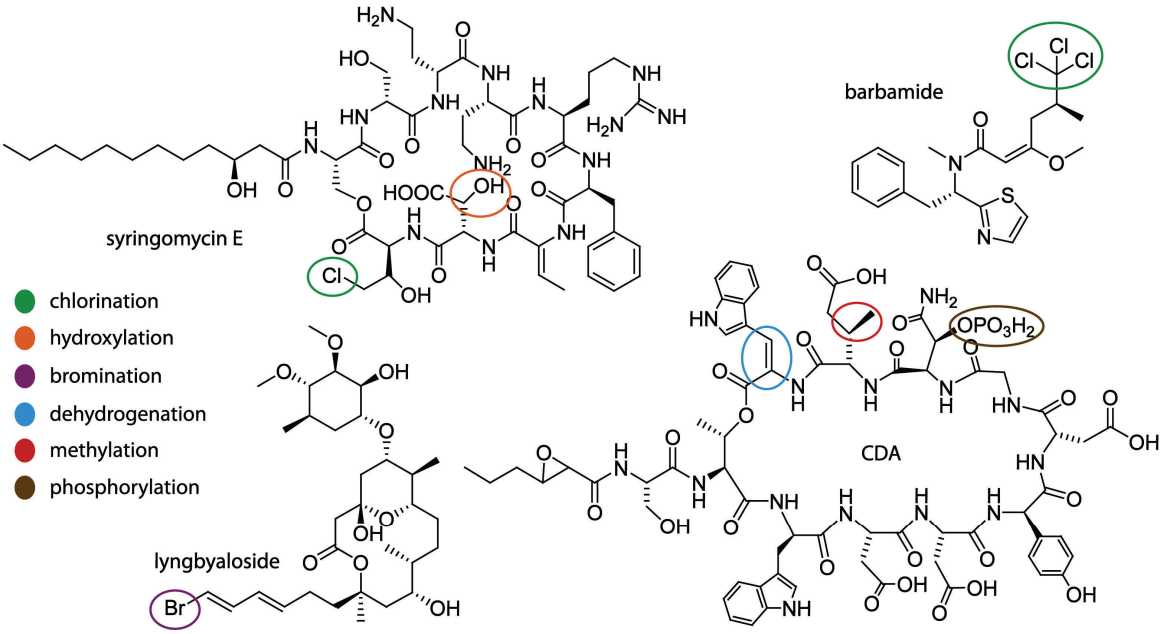


Figure 2.10: Digest of natural products containing carbon functionalized building blocks. Modifications range from methylation over hydroxylation, dehydrogenation and halogenation to phosphorylation. The modification reactions of the CDAs are described in chapter 2.4.

2.2.1 Precursor Synthesis of Non-Proteinogenic Building Blocks

In addition to α-amino acids, α-keto and α-hydroxy acids derived from amino acid metabolism are utilized as precursors in non-ribosomal peptide biosynthesis.* For example, cyclodepsipeptides exist that contain alternating ester and amide bonds which arise through the condensation of α-amino and α-hydroxy acids. Typically, in fungal depsipeptides such as enniatin B[104] (Figure 2.11) and the octadepsipeptides PF1022A,[105] α-keto acids are reduced to α-hydroxy acids prior to activation by the NRPS A-domain.

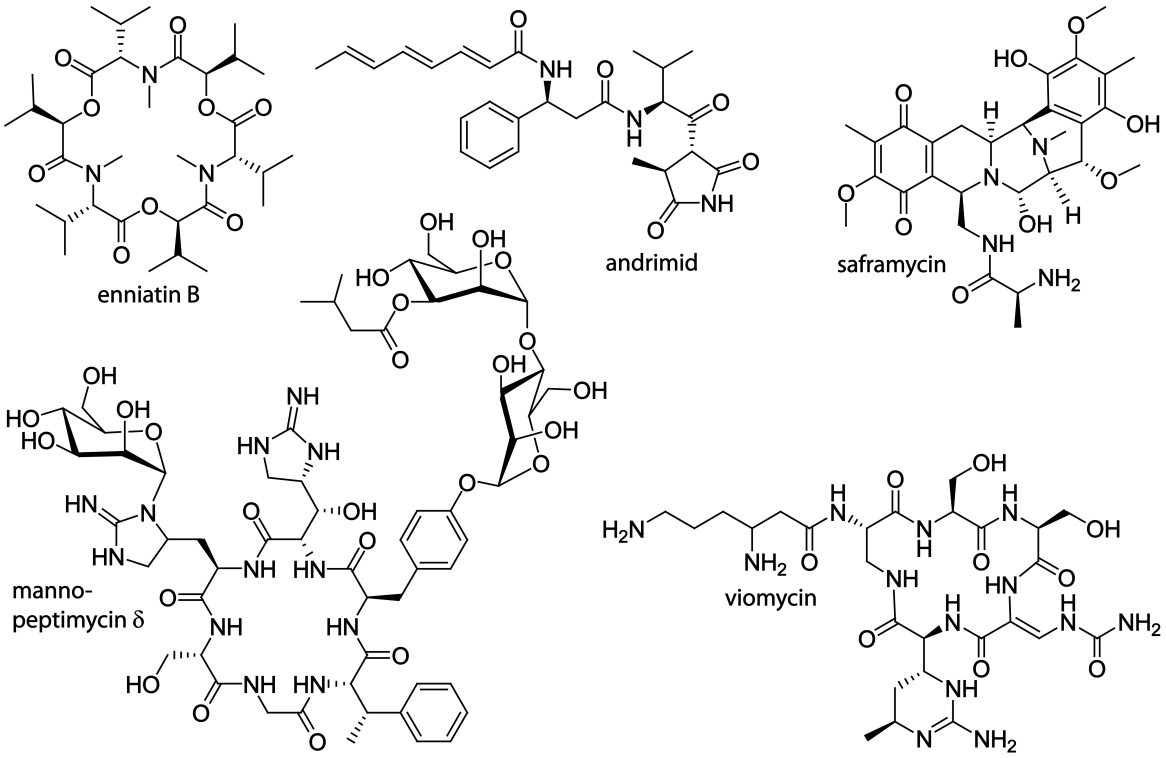


Figure 2.11: Selection of natural products containing non-proteinogenic precursors. The site of the unusual building block in each molecule is detailed in the text.

Furthermore, some peptides possess both β- and α-amino acids. For example, 4-methylideneimidazole-5-one (MIO)-dependent aminomutases (CmdF and AdmH) are responsible for generating the β-Tyr and β-Phe precursors of the hybrid non-ribosomal peptide-polyketides chondramide[106] and andrimid,[107] respectively (Fig. 2.11). MIO participates in covalent catalysis, enabling the elimination followed by conjugate addition of ammonia onto an α,β-unsaturated acid intermediate.[108] Alternatively, for example, a lysine 2,3-aminomutase (VioP) is implicated in the biosynthesis of the β-lysine precursor of viomycin (Figure 2.11).[109] The lysine 2,3-aminomutases belong to the

* The precursor synthesis during lipopeptide assembly is described there (chapter 2.4).

radical *S*-adenosyl methionine (SAM) superfamily utilizing pyridoxal-5'-phosphate, [4Fe-4S]$^{2+}$ and SAM cofactors to transfer the amino group via a radical mechanism.[110]

β-hydroxylation of amino acids is also frequently encountered in the biosynthesis of non-ribosomal peptides. In the biosynthesis of viomycin and mannopeptimycins (Figure 2.11) the free amino acids L-Arg and L-enduracididine are hydroxylated by non-heme iron(II)/αKG-dependent oxygenases VioC[111,112] and MppO.[113] The subsequent processing of β-hydroxylated amino acids by further oxidation, glycosylation, macrolactonization, methylation, or phosphorylation further serves to increase the structural diversity of non-ribosomal peptides (Figure 2.11).[16]

Methylation is another general theme in the structural diversification of non-ribosomal peptides and precursors. For example, a SAM-dependent catechol 4'-*O*-methyltransferase (SafC) delivers a 4'-*O*-methyl-L-dopa precursor in saframycin (Figure 2.11) biosynthesis.[114] There are also a number of *C*-methyl transferases, including MppJ and CymG, which are suggested to catalyze methylation reactions in the biosynthesis of the β-methylphenylalanine and 2-amino-3,5-dimethylhex-4-enoic acid (ADH), precursors of the mannopeptimycins[115] (Figure 2.11) and cyclomarins,[102] respectively.

2.2.2 Modification Enzymes Acting on PCP-Tethered Substrates

Beside precursor synthesis, modification enzymes were also found to work on PCP-tethered substrates. In the biosynthesis of the bacterial cyclododecadepsipeptides cereulide (Figure 2.12) and valinomycin (Figure 2.12), α-keto acids are activated by the A domain of the NRPS and an α-ketoreductase (α-KR) domain within the NRPS assembly line is responsible for chiral reduction to give peptidyl carrier protein (PCP)-tethered α-hydroxy acids.[116] In addition, there is a family of heme protein hydroxylases (e.g., ORF20, NovI, and NikQ) that catalyze the hydroxylation of PCP-tethered Tyr and His residues during the biosynthesis of chloroeremomycin (Figure 2.12), novobiocin[117] (Figure 2.12) and nikkomycin[118] (Figure 2.12).

A number of non-ribosomal peptides contain α,β-unsaturated amino acids. Enamine residues are most likely generated during peptide assembly due to their instable nature. Dehydrothreonine and *N*-methyldehydroalanine residues found in syringomycin[119] (Figure 2.12) and microcystin[120] (Figure 2.12) are predicted to arise by dehydration of threonine and *N*-methylalanine residues of PCP-peptidyl intermediates.

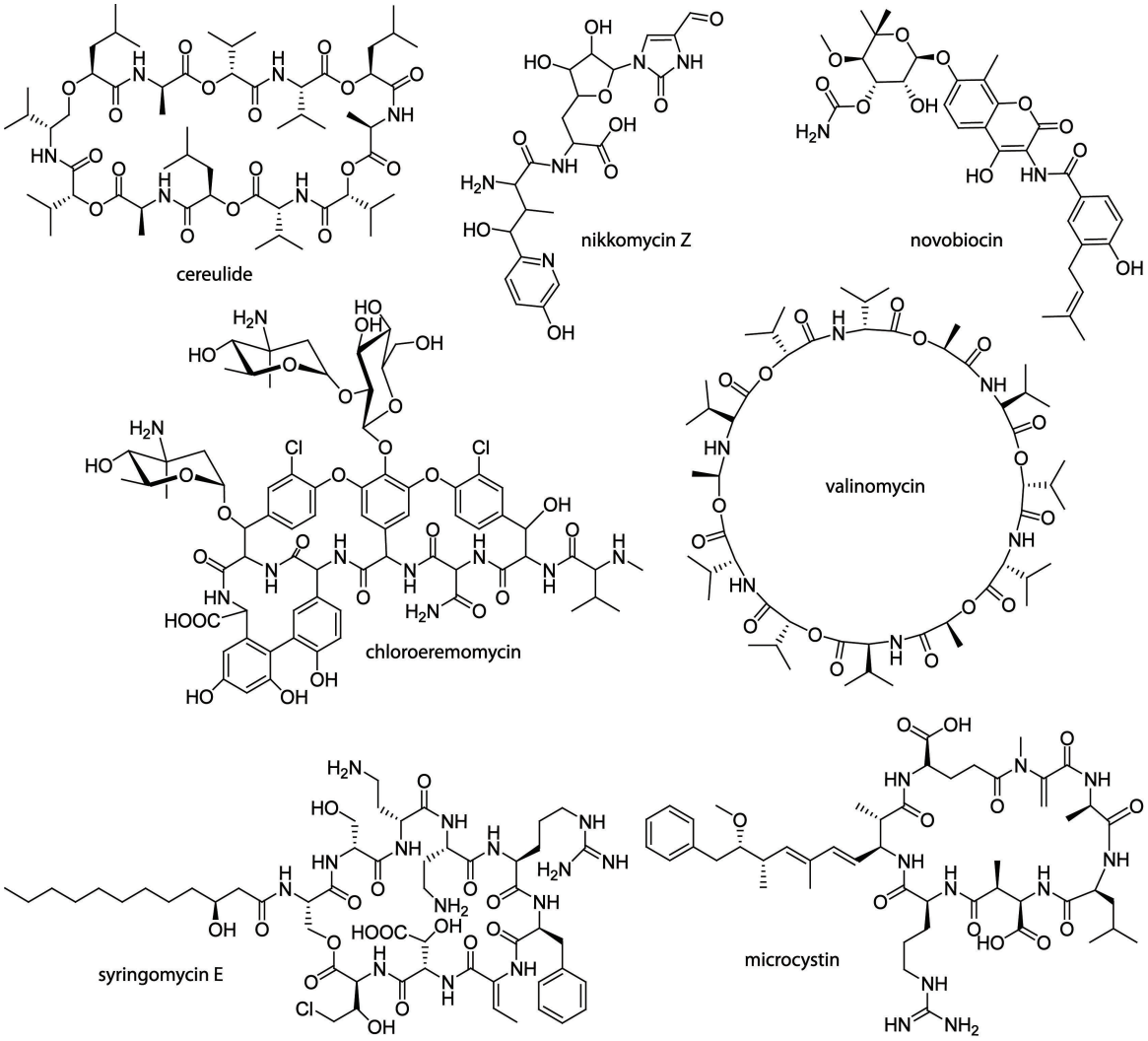


Figure 2.12: Selection of natural products comprising non-proteinogenic building blocks resulting from modification enzymes working on PCP-tethered substrates.

In the biosynthesis of the β-ureidodehydroalanine residue of viomycin (Fig. 2.11), a putative FAD-dependent oxidase/dehydrogenas (VioJ) is predicted to catalyze α,β-desaturation of a 2,3-diaminopropionyl-*S*-PCP intermediate prior to carbamylation.[109]

Of the many halogenated amino acid residues found in non-ribosomal peptides, chlorinated derivatives are most prevalent.[121,122] In the case of electronrich substrates such as aromatic groups, chlorine-carbon bond formation is typically catalyzed by flavin-dependent halogenases.[121,122] For example, chlorination of β-hydroxytyrosine in vancomycin (Figure 2.13) and other aromatic residues in related antibiotics is catalyzed by flavin-dependent enzymes.[123,124] Chlorination of inactivated carbon centers requires more powerful non-heme iron(II)/αKG halogenases[†] such as SyrB2, which catalyzes the

[†] For a comprehensive review of this enzyme class refer to chapter 2.3.2.

γ-chlorination of Thr-*S*-PCP intermediate during syringomycin E biosynthesis (Figure 2.12).[125] Additionally, in the syringomycin E (Figure 2.12) biosynthesis the non-heme iron (II)/αKG-dependent oxygenase SyrP catalyzes the hydroxylation of PCP-tethered aspartic acid to yield carrier bound *threo*-3-hydroxyaspartic acid.[126]

2.2.3 Post-Assembly Tailoring

Post-assembly modifications of the peptide antibiotics mainly contain of decoration with sugar-molecules. To date 66 different NDP-activated deoxysugars have been identified to be transferred onto natural product aglycones.[103] Many of the natural products are regio- and stereospecifically modified and often these glycosylated patterns are important for biological activity. Dehydratases, epimerases, and amino- and methyltransferases are responsible for the biosynthesis of these NDP-deoxysugars which derive mostly from D-glucose.[103]

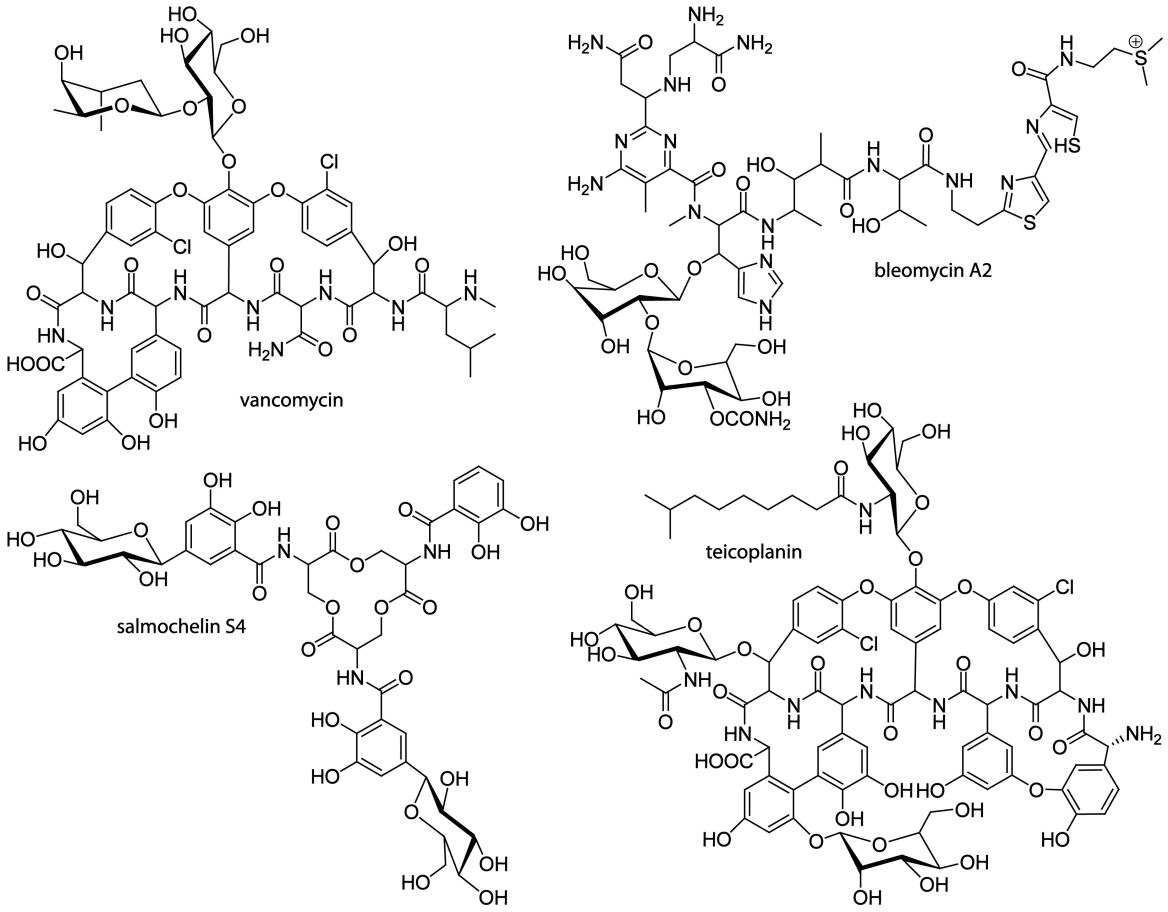


Figure 2.13: Structures of the glycopeptide antibiotics vancomycin, bleomycin A2, teicoplanin and the glycosylated enterobactin-like siderophore salmochelin S4.

Glycosyltransferases (Gtfs) are responsible for the attachment of these carbohydrates (mono-, di- or oligo-saccharides). Some 20 Gtfs from actinomycete-derived natural and biologically active agents have been characterized until now. The most prominent acceptor NRPS aglycones for these Gtfs are vancomycin[82] (Figure 2.13), teicoplanin[127] (Figure 2.13), chloroeremomycin[128] (Figure 2.12), bleomycin[129,130] (Figure 2.13) and novobiocic acid[130,131] (Figure 2.12). An overview of the antibiotics, producer strains, glycosyltransferases, substrates and acceptors can be found in Table 2.1.

Compound	Producer	Gtfs	Details
vancomycin	*Amycolatopsis orientalis* ATCC 19795	GtfD	transfers L-vancosamine to D-glucose
		GtfE	transfers D-glucose to heptapeptide aglycone
teicoplanin	*Actinomyces teichomyceticus*	tGtfA	transfers (*N*-acetyl)-glucosamine to heptapeptide aglycone
		tGtfB	transfers glucosamine to heptapeptide aglycone
		ORF3	transfers mannose to heptapeptide aglycone
chloroeremomycin	*Amycolatopsis orientalis* A82846	GtfA	transfers L-epivancosamine to heptapeptide aglycone
		GtfB	transfers D-glucose to heptapeptide aglycone
		GtfC	transfers L-epivancosamine to D-glucose
bleomycin	*Streptomyces verticillus* ATCC 15003	BlmE	transfers L-gulose to bleomycin-aglycone
		BlmE	transfers 3-*O*-carbamoyl-D-mannose to L-gulose
novobiocin	*Streptomyces spheroides* NCIMB 11891	NovM	transfers L-noviose to novobiocic acid
salmochelin S4	*Salmonella typhimurium* LT2 and other	IroB	transfers D-glucose via *C*-glycosylation of the catecholic arms of enterobactin

Table 2.1: The producer strains, Gtfs and their reaction details of the antibiotics vancomycin, teicoplanin, chloroeremomycin, bleomycin and novobiocin are listed as well of the siderophore salmochelin S4.

The glycosyl moieties were shown to be the basis of or to at least enhance the biological activity of the antibiotics.

The glucosylated enterobactin-like siderophore salmochelin (Figure 2.13) is also listed in Table 2.1 due to interesting chemical and evolutionary features.[132] Siderophores (*Greek: "iron carrier"*) are small, high-affinity iron chelating compounds secreted by microorganisms such as bacteria, fungi and grasses. Siderophores are amongst the greatest binders to Fe^{3+} known. Often they are found to be the pathogenic factors of infecting microorganisms by *stealing* essential iron from the host.[133] The human body evolved an efficient mechanism to sequester enterobactin by production of a lipocalin, called siderocalin. This protein has been shown to bind ferric-catecholates, one of which is enterobactin and hence prevents the iron from being delivered to the bacteria. It thus

can be seen as part of an innate immune response for protection against infection.[134] Interestingly, in some *Salmonella* and *Klebsiella* strains enterobactin was found to be modified by the glucosyltransferase IroB[135] to evade sequestration by siderocalin.[136] This is due to the fact that the binding pocket of the ferric-catecholate binding protein does not appear to be large enough to bind glucose attached to enterobactin.[132] Interestingly, IroB catalyzes the formation of a *C*-glycosidic bond, rarely found in natural products. Another example for *C*-glycosylation is the Trojan horse antibiotic microcin E492.[137]

Beside glycosylation, another post-assembly tailoring mechanism is prenylation. Prenylated non-ribosomal peptides are found e.g. in the cyclomarins and cyclomarazines produced by the marine bacterium *Salinispora arenicola* CNS-205.[102] Another example is the observed geranyl group transfer during lyngbyatoxin A synthesis, which is produced by the cyanobacterium *Lyngbya majuscula*, also known as fireweed.[138]

2.3 Non-Heme Iron Enzymes

The iron-containing enzymes that are involved in dioxygen activation can be divided into two groups based on the active site structures, i.e. heme and non-heme containing enzymes. The heme enzymes have been studied extensively and are well understood, with cytochrome P450 as the prototypical example.[139]

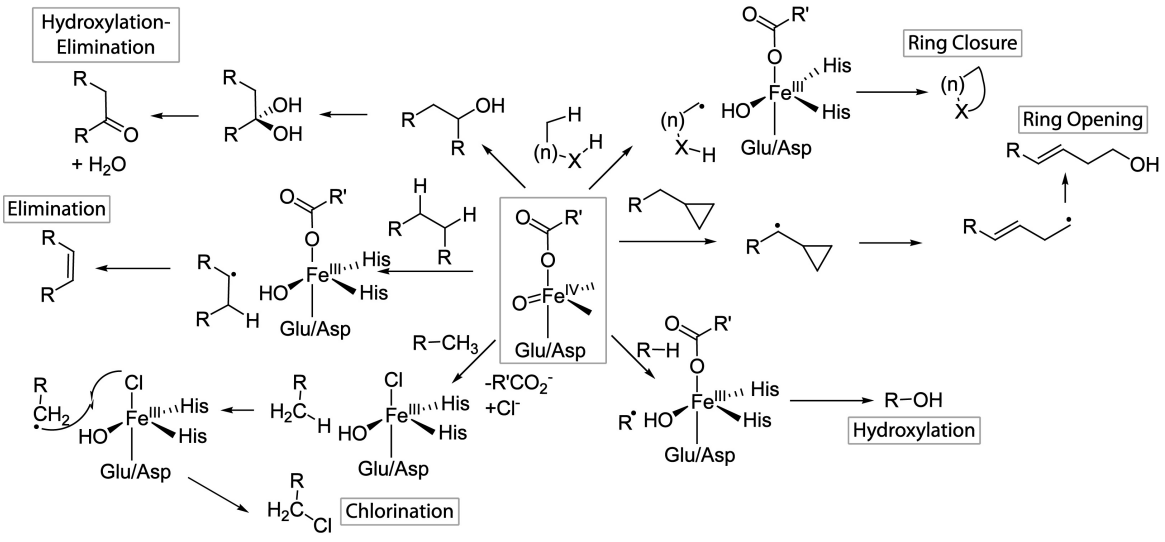


Figure 2.14: Overview of reactions catalyzed by non-heme iron(II)/aKG-dependent oxygenases.

On the other hand, mononuclear non-heme iron oxygenases have received the most attention recently, primarily because of the recent availability of crystal structures of many different enzymes and the stunningly diverse oxidative transformations that these enzymes catalyze.[140] This is briefly mentioned in chapter 2.2 and illustrated by a

comprehensive overview of the catalyzed reactions in Figure 2.14. The remarkable scope of oxidative transformations seems to be even broader than that associated with oxidative heme enzymes. Not only are many of these oxidative transformations of key biological importance, many of these selective oxidations are also unprecedented in synthetic organic chemistry. Beside the most important reactions, hydroxylation and halogenation, these enzymes were found to catalyze ring closure and opening, eliminations, which is in some case preceded by hydroxylation (Figure 2.14). As a result, these enzymes are involved in processes as diverse as the synthesis of antibiotics, DNA repair, oxygen sensing, and transcription regulation. The subfamily constitutes probably the most versatile group of oxidizing biological catalysts identified to date.[14,141,142] As most of the work described in this study deals with the β-hydroxylation of unreactive carbon atoms by AsnO, LptL, KtzO and KtzP and the AsnO D241N variant, non-heme iron(II)/αKG-dependent hydroxylases are detailed below. Due to the structural and likely evolutionary relationship between non-heme iron hydroxylases and halogenases the latter is also introduced in the following.

2.3.1 Iron(II)- and αKG-Dependent Hydroxylases

Facial triad or iron binding motif. Recent crystal structure availability of mononuclear non-heme iron oxygenases established the so-called 2-His-1-carboxylate facial triad as a new common structural motif for the activation of dioxygen.[143] The largest subfamily of non-heme iron enzymes, the hydroxylases couple the oxidative transformation of substrates to the oxidative decarboxylation of the cofactor α-ketoglutarate αKG to carbon dioxide and succinate.[14] Many of these often regio- and stereospecific reactions are at the moment beyond the scope of synthetic organic chemists and are therefore of special chemical interest.[141] A recent comprehensive sampling of the reactivities of the currently known αKG-dependent enzymes is available.[144] Therefore here only some general features and typical examples for hydroxylation reactions are addressed. Two types of hydroxylation reactions have recently received increased attention out of medical interest. Damage of RNA and DNA by nucleotide alkylation, for instance, results in lesions that are both cytotoxic and mutagenic. Several αKG-dependent oxygenases have been shown to repair these alkylated DNA and RNA bases. *Escherichia coli* AlkB and

the homologous human enzyme ABH3 fix the lesion by hydroxylating the alkyl group, after which spontaneous deformylation yields the unmodified base (Figure 2.15 A).[145]

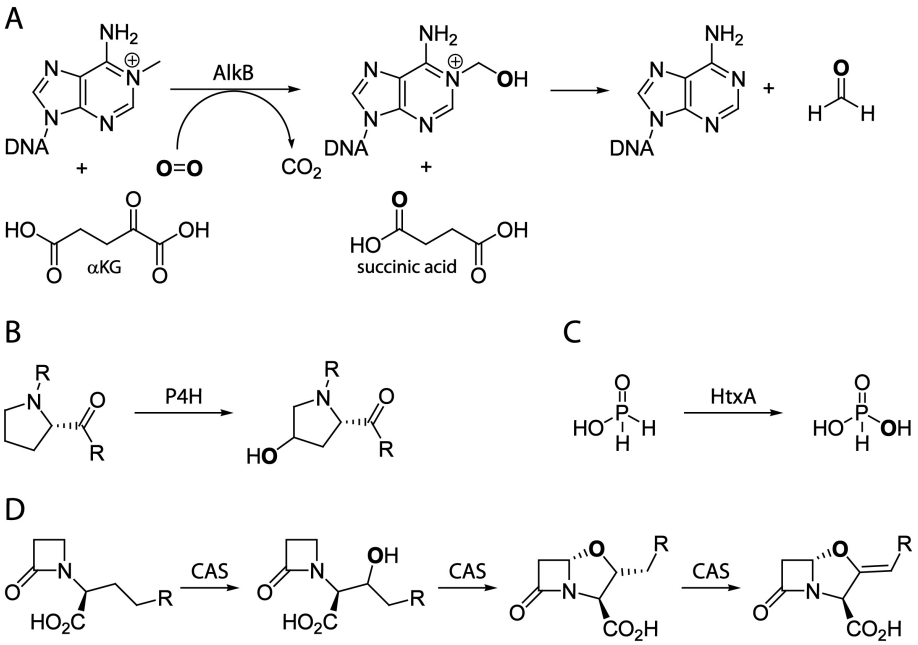


Figure 2.15: Selection of hydroxylation catalyzed by non-heme iron(II)/αKG-dependent enzymes. The incorporated oxygen atoms are highlighted in bold letters. One oxygen atom is transferred to succinate (see reaction A) and one to the substrate. The reactions catalyzed are: A) DNA dealkylation by AlkB; B) Proline hydroxylation at position 4; C) Oxidation of hypophosphite to phosphite by HtxA; D) The clavaminate synthase (CAS) catalyzes three steps in the synthesis of the β-lactamase inactivator clavulanic acid. First hydroxylation yields proclavaminic acid, the second conversion dihydroclavaminic acid and in the last step clavaminic acid is generated.

The hydroxylation of specific residues of protein side chains and more specifically in oxygen sensing in the cell has also attracted recent interest. Hypoxia-inducible factor (HIF) is responsible for mediating the mammalian response to low oxygen tension (hypoxia). αKG HIF hydroxylases have been implicated in this hypoxic response and are therefore interesting targets for the development of new therapies for the different diseases associated with this system.[146] Other functions of the αKG-dependent oxygenases include the biosynthesis of antibiotics and plant products, lipid metabolism and biodegradation.[14] Some selected examples are shown in Figure 2.15.

Crystallographic data on many different αKG-dependent oxygenases have been reported over the last 10 years and the more than 50 available structures now provide a wealth of structural information (for a recent review see: Clifton *et al.*[141]). The ternary TauD-Fe(II)-αKG-substrate complex, for instance, shows a five-coordinated metal center with the

αKG co-substrate bound in a bidentate way.[147] The highly conserved jelly-roll fold, which consists of two parallel β-sheets, typical for this enzyme class, is also found (Figure 2.16).

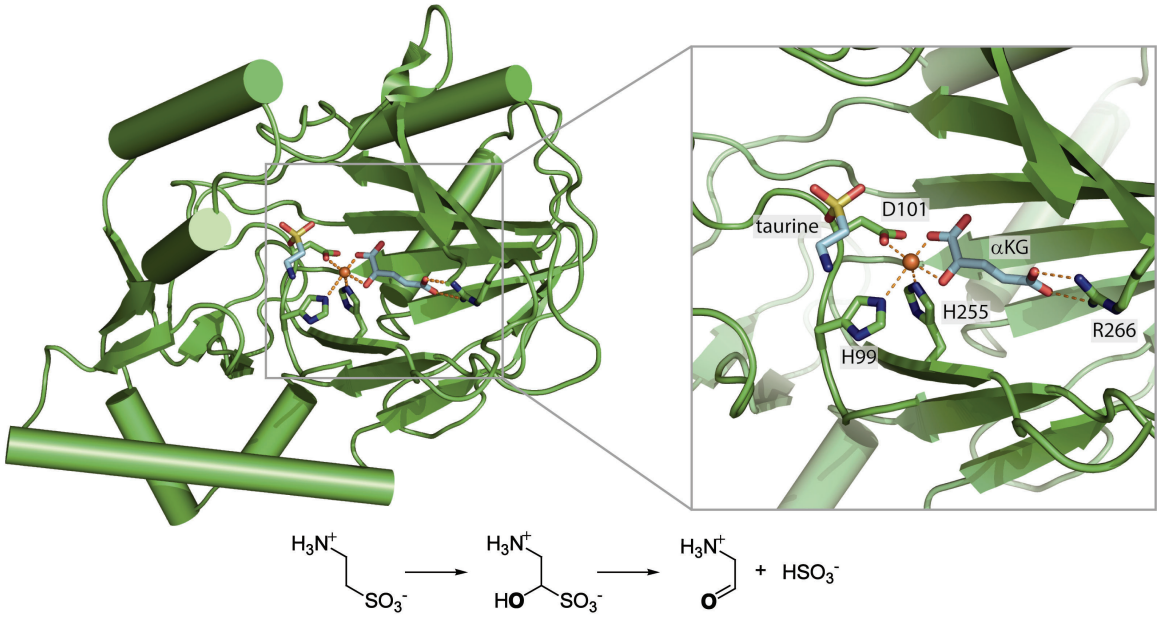


Figure 2.16: Overall fold of one monomer of the ternary TauD. The active site of the ternary TauD-Fe(II)-αKG-taurine complex is magnified (pdb: 1OS7). The bottom scheme show the reaction catalyzed by TauD.

The co-substrate is further held in place by an additional interaction of the C5-carboxylate with a conserved Arg residue. The primary substrate is not bound directly to the metal center, but is found close to the open coordination site, which is believed to be the site of dioxygen binding. TauD catalyzes the hydroxylation of taurine (2-aminoethane-1-sulfonic acid), which leads eventually to sulfite elimination from the product (Figure 2.16, bottom scheme).[148]

Mechanism. The large body of crystallographic data together with spectroscopic studies on especially TauD[147] and CAS (clavaminate synthase)[149] has led to the proposal of a common, conserved mechanism for the aKG-dependent oxygenases [Figure 2.17, steps numbered (1-8)]. In this general mechanism, an ordered sequential binding of first αKG, then substrate and finally dioxygen is assumed. Structural and spectroscopic studies show an octahedral iron(II) metal center in the holo enzymes with the facial triad and three water molecules completing the coordination sphere (1).[144] Subsequently, the α-ketoglutarate binds in a bidentate fashion by displacing two water molecules and the metal center remains six-coordinated (2).

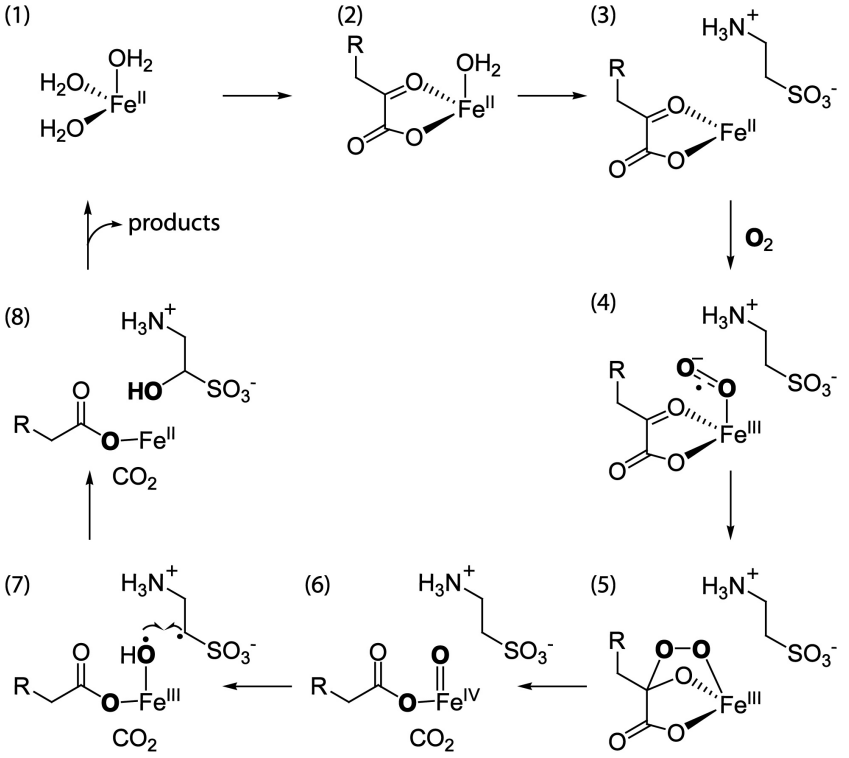


Figure 2.17: Consensus mechanism for αKG-dependent hydroxylases, exemplified for TauD (adapted from Bollinger *et al.*[147]). For clarity reason the facial triad has been removed.

The bidentate binding of the α-keto acid gives rise to a very characteristic set of metal to ligand charge transfer transitions in the visible region (around 500 nm).[150] The native substrate then binds in the proximity of (but not directly to) the metal ion and upon substrate-binding the remaining water ligand is lost and, in almost all cases, the coordination around the ferrous ion changes to five-coordinate square-pyramidal (3). This change in coordination number has been crystallographically established in, among others, the structures of TauD, DAOCS, and Alkylsulfatase (AtsK), which were crystallized in the presence of αKG and their respective native substrates.[141] The displacement of the last water molecule is thought to be essential for catalysis and poises the metal for reaction with dioxygen. The greatly enhanced reactivity of the metal center towards dioxygen after substrate binding effectively favors the generation of the (potentially damaging) reactive intermediates only in the presence of the target substrate and in this way protects the enzyme from inactivation by self-hydroxylation reactions. Such selfhydroxylation of nearby amino acid residues in the absence of substrate has in fact been observed for a number of non-heme iron enzymes, including some αKG-dependent enzymes.[151] It has been proposed that this self-modification serves to protect the

enzyme from chemical damage in the absence of substrate, but the physiological function and significance is currently not completely clear and subject of discussion. In the presence of substrate, the binding of dioxygen results in the formation of an adduct with significant Fe(III)-superoxide radical anion character (4).[152] Nucleophilic attack of the carbonyl carbon atom by this species would then result in a cyclic ferryl bridged-peroxo species (5). In the next step, decarboxylation and O-O bond cleavage yield succinate, carbon dioxide and a high-valent Fe(IV)=O species (6), which is the intermediate responsible for the substrate oxidation leading to hydroxylation or a related two-electron oxidation.[147] This ferryl intermediate was first observed experimentally for TauD by rapid freeze-quench Mößbauer spectroscopy,[153] and its presencewas confirmed by extended x-ray absorption fine structure (EXAFS),[154] and resonance Raman studies.[155] These spectroscopic studies revealed species with a high-spin (S = 2) configuration,[153] a characteristic isotope-sensitive iron-oxo vibration in the 800 cm^{-1} region,[155] and a short Fe-O interaction of 1.62 Å.[154] Most importantly, these studies provided the first direct evidence of the involvement of an Fe(IV)-intermediate in reactions catalyzed by mononuclear non-heme iron enzymes. From this point on the chemistry of each enzyme diverges and different subsequent steps result in the rich variety of oxidative transformations that this subgroup catalyzes. In the hydroxylase subfamily of non-heme iron αKG-dependent enzymes, the common Fe(IV)-intermediate (see below) abstracts a hydrogen atom (7), followed by recombination of the coordinated hydroxyl radical with the substrate radical yieldind the hydroxylated product (8). The observation of a large substrate deuterium kinetic isotope effect on the decay of the ferryl species identified it as the hydrogen-abstracting intermediate.[153] A C-H cleaving high-spin Fe(IV) complex with strikingly similar spectroscopic and kinetic characteristics has also been detected in prolyl-4-hydroxylase (P4H), which is only distantly related to TauD.[156] These results support the view of a conserved mechanism within the αKG-dependent hydroxylases. The hydroxylases involved in the acidic lipopeptide antibiotic synthesis are discussed in chapter 2.4 and in the results as well as in the discussion section.

2.3.2 Iron(II)- and αKG-Dependent Halogenases

The recently discovered group of non-heme iron halogenases posses a unique reactivity.[122] Rather than hydroxylation, these enzymes catalyze the halogenation of aliphatic C-H bonds and require Fe(II), αKG, dioxygen and chloride for activity. Pioneering work of Vaillancourt *et al.* on the origin of the chlorinated threonine residue found in the phytotoxin syringomycin E (Figure 2.12) resulted in the discovery of this intriguing enzyme class. It was shown that SyrB2 catalyzes the chlorination of the threonine methyl group (Figure 2.18 A).[125]

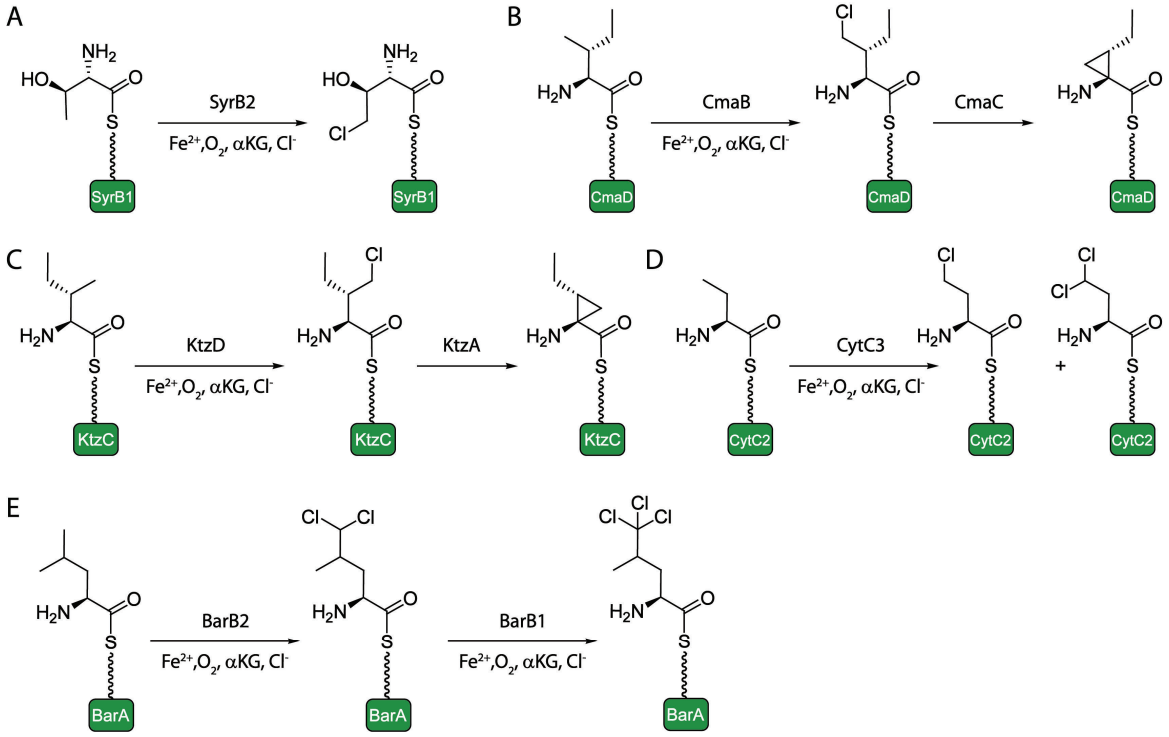


Figure 2.18: Overview of reactions catalyzed by non-heme iron(II)/αKG-dependent halogenases. A) Threonine halogenation in the syringomycin E biosynthesis; Cryptic chlorination pathways in the biosynthesis of the non-proteinogenic amino acids B) coronamic acid and C) (1*S*,2*R*)-allocoronamic acid; D) CytC3 catalyzes the mono and dichlorination of α-aminobutyric acid; E) Coupled triple chloride transfer onto PCP-bound leucine by tandem action of BarB2 and BarB1 during barbamide biosynthesis.

Additionally, the non-heme halogenase CmaB[157] was examined and found to be responsible for the γ-halogenation of L-allo-isoleucine as part of a cryptic synthesis strategy for cylcopropyl ring formation during phytotoxin precursor coronamic acid synthesis (Figure 2.18 B). Another example for cyclopropyl ring formation is the halogenase KtzD. Initially the gene cassette KtzABCD was predicted to catalyze the production of the non-proteinogenic amino acid 2-(1-methylcyclopropyl)glycine (MecPG)[158] a component of all kutzneride molecules (see chapter 2.5) reported to

date.[18,19] Instead it was reported that KtzABCD function to produce an alternative cyclopropyl-containing amino acid, (1*S*,2*R*)-allocoronamic acid (alloCMA), tethered to the PCP KtzC (Figure 2.18 C) by KtzD-mediated hydroxylation of the γ-CH_2 of L-Ile.[159] Bioinformatic analysis has provided leads to additional members of this interesting new group.[122] CytC3, for example, chlorinates L-Abu to γ-chloroaminobutyrate or, if allowed to proceed through multiple turnovers, even to γ,γ-dichloroaminobutyrate (Figure 2.18 D).[160] The free dichloroamino acid is a known antibacterial natural product called armentomycin. Multiple chlorinations were also observed for the SyrB2 homologues BarB1 and BarB2, which catalyze the tandem trichlorination of L-Leu as part of the biosynthesis of the marine natural product barbamide (Figure 2.10 and Figure 2.18 E).[161] The crystal structures of the Syr2B-Fe(II)-αKG (Figure 2.19 A)[162] and the CytC3 Fe(II)-αKG (Figure 2.19B)[163] complexes were solved very recently. As a surprising result, it presented the first αKG-dependent non-heme iron enzyme not coordinating the iron by the 2-His-1-carboxylate facial triad. Instead, the metal is coordinated by αKG, a chloride anion and only two endogenous ligands. i.e. two histidines (Figure 2.19). An alanine replaces the aspartate of the facial triad.

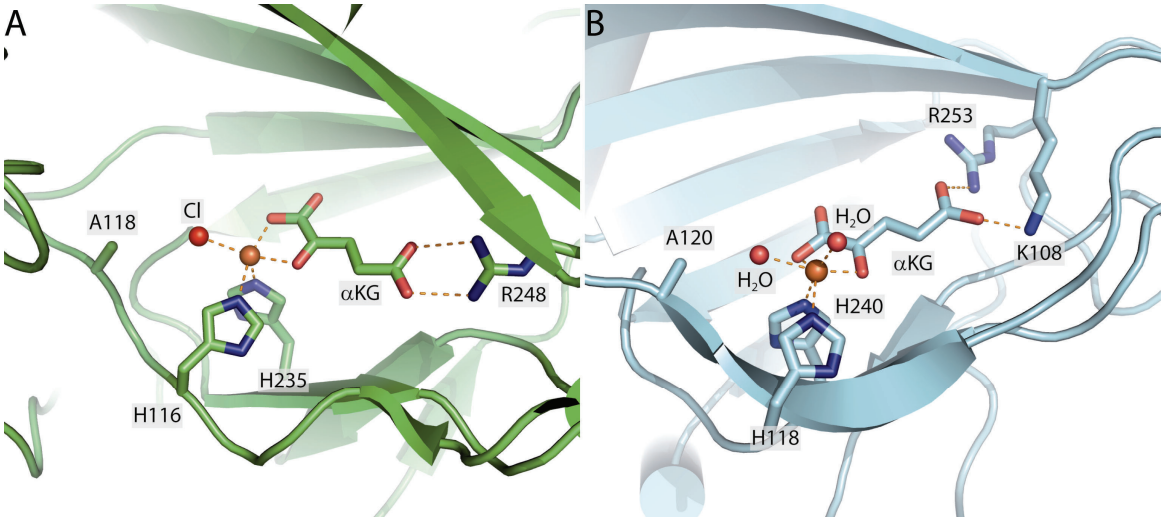


Figure 2.19: Active sites of the SyrB2-Fe(II)-αKG (A) and CytC3-Fe(II)-αKG (B) complexes. Instead of the facial triad (HXD/E…H) observed in hydroxylases only two ligands coordinate the iron cofactor and the acidic residue was found to be exchanged by alanine, which provides enough space for chloride binding.

After the generation of the pivotal ferryl-oxo species and the hydroxylradical, the mechanism of halogenation is thought to be similar to the common hydroxylation mechanism (Figure 2.17) with chloride abstraction instead of hydroxyl abstraction by the carbon radical in the final step (see also Figure 2.14). Mößbauer and absorption

spectroscopic studies on the chlorination reaction by CytC3 identified two rapidly interconverting high-spin Fe(IV) intermediates and a marked substrate deuterium kinetic isotope effect on the decay of these ferryl species implicated them as the hydrogen atom-abstracting complex.[164] The presence of two rather than one Fe(IV) intermediate contrasts with the observations made for TauD and P4H. The authors speculate that the two intermediates might be two rapidly-converting conformers of a Cl-Fe(IV)-oxo complex.[165] A spectroscopic study of CytC3 in the analogous reaction with bromide provided further evidence for direct coordination of a halogen to the high-valent iron and the assignment of these high-spin iron(IV) intermediates as a Br-Fe(IV)-oxo species.[166] Freeze-quench Fe K-edge X-ray absorption spectroscopy revealed short Fe-O and Fe-Br interactions of 1.62 Å and 2.43 Å, respectively. These results further emphasize the similarities between the αKG-dependent halogenases and hydroxylases.

2.4 Acidic Lipopeptides

There has been a steady rise in the prevalence of multi-drug resistant, especially Gram-positive, and other pathogens. Furthermore, concerns were expressed about the clinical effectiveness of glycopeptides in treating infections due to *Staphylococcus aureus*.[167] Promising novel antibiotics to fight the latter are acidic lipopeptides and their efficiency is resembled by the success story of daptomycin. This antibiotic (Cubicin, Cubist Pharmaceuticals) was the first natural antibiotic in a new structural class, which has entered clinical use since the 1960s.[168] Daptomycin, produced by *Streptomyces roseosporus*, was discovered in the early 1980s.[5] The United States FDA approved daptomycin for the treatment of serious skin and soft tissue infections in 2003 and for methicillin-resistant *S. aureus* (MRSA) infections of the bloodstream (bacteremia) in 2006.[169] Besides daptomycin, the best known representative of the acidic lipopeptide antibiotics, there are a number of other ten-membered cyclic lipopeptides. These natural products are also secondary metabolites produced by soil bacteria and include the depsipeptides, calcium-dependent antibiotic (CDA) produced by the model streptomycete *Streptomyces coelicolor* A(3)2[170] and A54145 produced by *Streptomyces fradiae.*[171] Other cyclic lipopeptide antibiotics produced by actinomycetes include friulimicin[172] produced by *Actinoplanes friuliensis* and the amphomycins,[173] laspartomycins[174] and glycinocins[175] (Table 2.2).

Compound	Producer
A-1437	*Actinoplanes* sp. (=*Actinoplanes friuliensis* DSM 7358)
A21978C/daptomycin	*Streptomyces roseosporus* NRRL11379
A54145	*Streptomyces fradiae* NRRL18158
	Streptomyces refuineus spp*. thermotolerans*
amphomycin	*Streptomyces canus* ATCC 12237
aspartocin	*Streptomyces griseus* var. *spiralis* ATCC13733
	Streptomyces violaceus var. *aspartocinius* ATCC 13734
CDA	*Streptomyces coelicolor* A3(2)
	Streptomyces violaceoruber Kutner 673
	Streptomyces lividans
friulimicin	*Actinoplanes friuliensis* DSM 7358
glycinocin	*Actinomycete* AW 998
laspartomycin	*Streptomyces viridochromogens* ATCC 29814
parvuline	*Streptomyces parvulus* var. *parvuli* NRRL 5740
tsushimycin	*Streptomyces griseoflavus* ATCC 21139

Table 2.2: Overview of important acidic lipopeptide antibiotics and the producing strains.[176]

Within these lipopeptides the positioning of D-amino acids and achiral amino acids is conserved, as is the Asp-X-Asp-Gly motif in the macrocycle. Some of the amino acids are heavily modified during the biosynthesis. Tailoring enzymes, which include oxygenases, oxidases, methyltransferases and others, are responsible for these modifications and thus provide a large number of non-proteinogenic amino acids and a fascinating structural diversity.[15,28,177] The fascinating pool of unusual amino and fatty acid building blocks of the acidic lipopeptide antibiotics can be used for the manipulation of their biosynthesis and contributes to their biological activity. In this summary about acidic lipopeptides, the structural diversity of acidic lipopeptide antibiotics is detailed and the proposed modes of action are presented. The recent advances in the field of the modifying enzymes are described.

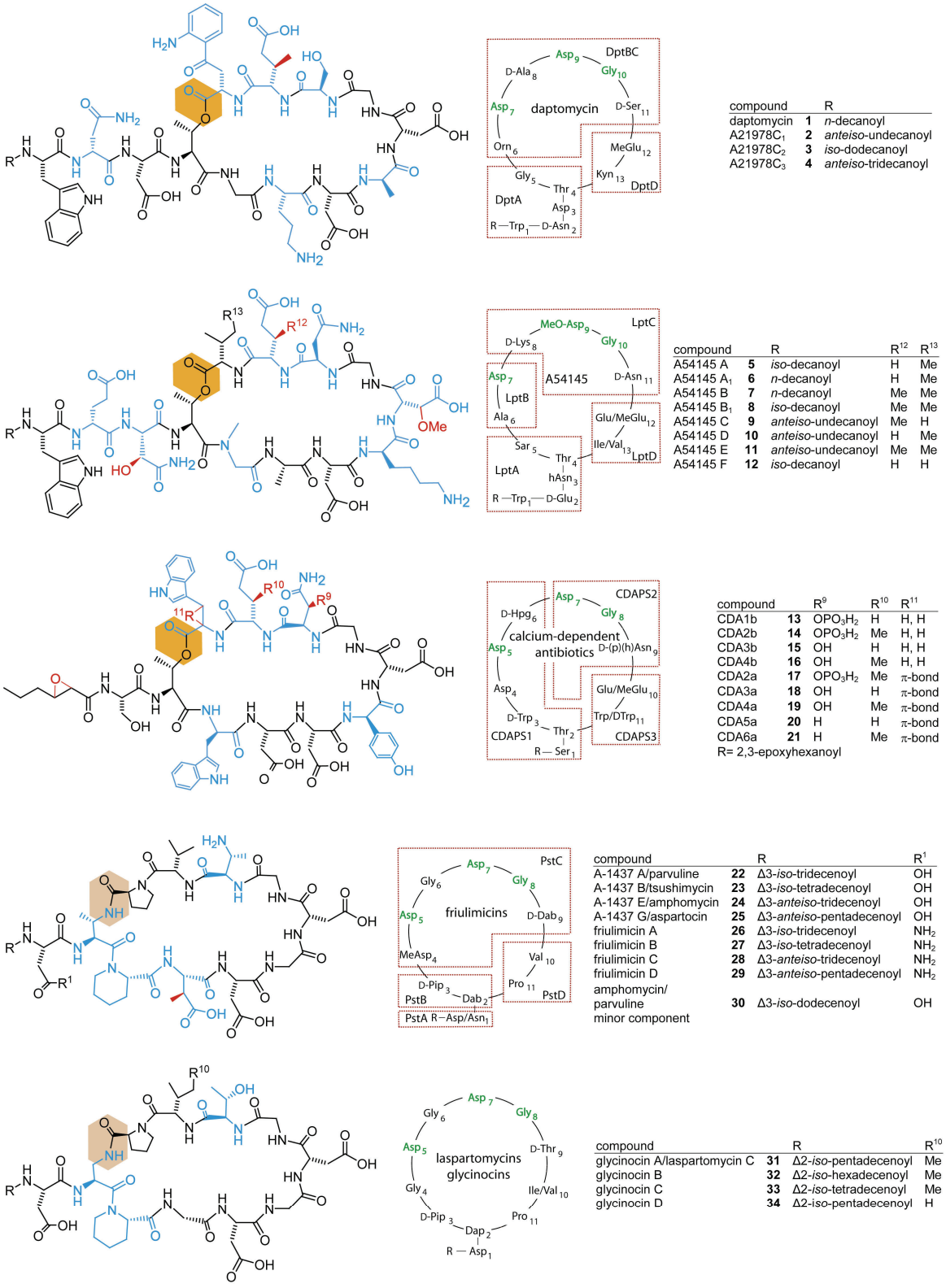
2.4.1 Structures

In addition to their cyclic nature, a key structural feature of lipopeptide antibiotics is the long chain fatty acid, which is invariably attached to the macrocyclic peptide core. All members of this class of antibiotics are produced by non-ribosomal peptide synthetases with variation of the fatty acid tail. Straight and branched chain fatty acids that significantly differ in the degree of saturation and the oxidation state are frequently found and contribute to the high structural diversity of this class of compounds (Figure 2.20). In particular, the lipid portion impacts the biological properties of these molecules; antimicrobial behavior and toxicity are dramatically affected by the nature of

the incorporated fatty acid group.[178] The length of the fatty acid chain is variable and ranges from 6 to 16 carbon units. CDA has the shortest fatty acid, an expoxidized hexanoic acid, which in contrast to other lipopeptides with their complex fatty acid mixtures, is invariant. The longest fatty acid tail containing 16 carbon atoms is found in glycinocin B. An overview of the acidic lipopeptides can be found in Figure 2.20, where all variants are listed.

A21978C/daptomycin. The three major components of A21978C (Figure 2.20; **2**-**4**) originally found to be produced by *Streptomyces roseosporus*, were determined by deacetylase-mediated removal of the lipid tail, followed by classical degradation of the amino acid sequence paired with MS analysis. The peptide core was found to consist of 13 amino acids and that ten of these residues create a macrolactone by formation of an ester bond between Thr_4 and the C-terminal kynurenine (Kyn_{13}).[5] Besides Kyn, other non-proteinogenic and D-amino acids are present in A21978C, namely D-Asn_2,[179] ornithine (Orn_6), D-Ser_{11} and 3-methylglutamic acid ($MeGlu_{12}$). The most famous representative of the A21978C complex, daptomycin (Figure 2.20; **1**), originally was found to be part of a minor component ($A21978C_0$) but showed superior biological acitivity compared with other members of the complex.[178] Thus, the fermentation process was optimized to yield high-quantity daptomycin production by feeding decanoic acid.[180] Although the complex was known since the early 1980s, the biosynthetic gene cluster was not described before 2005. It encodes for three NRPSs DptA, DptBC and DptD. The catalytic domains are predicted to couple five, six and two amino acids, respectively, with three epimerization domains present, consistent with the three D-amino acids found in the antibiotic.[179]

A54145. *Streptomyces fradiae* produces the lipopeptide A54145, a complex antibiotic mixture active against Gram-positive bacteria.[181] The identities of the acyl side chains were established as *iso*-decanoyl (Figure 2.20; **5**, **8**, **12**), *n*-decanoyl (Figure 1, **6**-**7**), and *anteiso*-decanoyl (Figure 2.20; **9**-**11**).[182] Seven of the thirteen amino acids are non-proteinogenic, including three D-configured peptide building blocks. The six amino acids



compound		R
daptomycin	1	*n*-decanoyl
A21978C$_1$	2	*anteiso*-undecanoyl
A21978C$_2$	3	*iso*-dodecanoyl
A21978C$_3$	4	*anteiso*-tridecanoyl

compound		R	R^{12}	R^{13}
A54145 A	5	*iso*-decanoyl	H	Me
A54145 A$_1$	6	*n*-decanoyl	H	Me
A54145 B	7	*n*-decanoyl	Me	Me
A54145 B$_1$	8	*iso*-decanoyl	Me	Me
A54145 C	9	*anteiso*-undecanoyl	Me	H
A54145 D	10	*anteiso*-undecanoyl	H	Me
A54145 E	11	*anteiso*-undecanoyl	Me	Me
A54145 F	12	*iso*-decanoyl	H	H

compound		R^9	R^{10}	R^{11}
CDA1b	13	OPO_3H_2	H	H, H
CDA2b	14	OPO_3H_2	Me	H, H
CDA3b	15	OH	H	H, H
CDA4b	16	OH	Me	H, H
CDA2a	17	OPO_3H_2	Me	π-bond
CDA3a	18	OH	H	π-bond
CDA4a	19	OH	Me	π-bond
CDA5a	20	H	H	π-bond
CDA6a	21	H	Me	π-bond

R= 2,3-epoxyhexanoyl

compound		R	R^1
A-1437 A/parvuline	22	Δ3-*iso*-tridecenoyl	OH
A-1437 B/tsushimycin	23	Δ3-*iso*-tetradecenoyl	OH
A-1437 E/amphomycin	24	Δ3-*anteiso*-tridecenoyl	OH
A-1437 G/aspartocin	25	Δ3-*anteiso*-pentadecenoyl	OH
friulimicin A	26	Δ3-*iso*-tridecenoyl	NH_2
friulimicin B	27	Δ3-*iso*-tetradecenoyl	NH_2
friulimicin C	28	Δ3-*anteiso*-tridecenoyl	NH_2
friulimicin D	29	Δ3-*anteiso*-pentadecenoyl	NH_2
amphomycin/ parvuline minor component	30	Δ3-*iso*-dodecenoyl	OH

compound		R	R^{10}
glycinocin A/laspartomycin C	31	Δ2-*iso*-pentadecenoyl	Me
glycinocin B	32	Δ2-*iso*-hexadecenoyl	Me
glycinocin C	33	Δ2-*iso*-tetradecenoyl	Me
glycinocin D	34	Δ2-*iso*-pentadecenoyl	H

Figure 2.20: Comprehensive overview of the chemical structures of the acidic lipopeptide antibiotics. The cyclic depsipeptides A21978C/daptomycin, A54145, CDA and the cyclic peptides friulimicins, laspartomycins and glycinocins are shown. The macrolactonization site is shaded in orange, the macrolactamization in rose-color. D- and unusual amino acids are highlighted in blue, sites where tailoring enzymes act in red. Next to the structure the peptide-cores are represented schematically. The position of the conserved Asp-X-Asp-Gly-motif is accentuated in green. The red dashed-line boxes resemble the NRPSs responsible for the incorporation of the framed amino acid(s) (if known). Next to the schemes the observed variants of the lipopeptide antibiotics are listed.

D-Glu_2, 3-hydroxy-L-asparagine ($hAsn_3$), sarcosine (Sar_5), D-Lys_8, 3-methoxyaspartic acid ($MeOAsp_9$), and D-Asn_{11} are found in all A54145s, whereas $MeGlu_{12}$ is only observed in **7**-**9** and **11**. The higher complexity of A54145 compared with A21978C is also due to the variation at position 13, where either Ile (Figure 2.20; **5**-**8**, **10**, **11**) or Val (Figure 2.20; **9**, **12**) is incorporated. The assumption that D-amino acids are present at position 2, 8 and 11 was made after the biosynthetic gene cluster was characterized and three epimerization domains were found in the corresponding modules.[183] Four NRPSs are also encoded by the biosynthetic gene cluster (LptA-D) responsible for the incorporation of five, two, four and two amino acids into the growing peptide chain, respectively.

Calcium-dependent antibiotics (CDA). In 1983, Hopwood *et al.*[170] isolated a substance from *Streptomyces coelicolor* A3(2) fermentations that inhibits the growth of Gram-positive bacteria in the presence of Ca^{2+} ions. Because of this activity, the substance was named calcium-dependent antibiotic (CDA). Originally four (CDA1b, CDA2b, CDA3b and CDA4b) of the nine currently known CDAs were structurally characterized (Figure 2.20; **13**-**16**).[184] Macrolactonization of the CDAs occurs through the side chain of Thr_2 and the *C*-terminal amino acid Trp_{11} to yield the ten-membered ring. The exocyclic tail only consists of one additional amino acid and the fatty acid acyl chain is invariant and exclusively found to be 2,3-epoxy-hexanoyl. The sites of variations were found to be in the amino acid core. For the first four characterized CDAs these are position 9 and 10. Asn_9 is either hydroxylated (Figure 2.20; **15**-**16**) or in a subsequent reaction phosphorylated (Figure 2.20; **13**-**14**) at the β-position. The tenth residue, Glu_{10}, is methylated in two cases (Figure 2.20; **14**, **16**). The C-terminal Trp was found to be a third site of amino acid modification by Hojati *et al.*[185] and lead to the discovery of three new CDA variations that each contain *Z*-2'-3'-dehydro-Trp (ΔTrp) (Figure 2.20, **17**-**19**). In the same work, the organization of the biosynthetic gene cluster was determined and revealed three NRPSs (CDAPS1-3) which are responsible for the incorporation of six, three and two amino acids, respectively. Recently, two further CDA derivatives have been isolated from genetically manipulated *S. coelicolor* in which the *asnO* and *hasP* genes were knocked out.[186] AsnO[187] is an iron(II)/α-ketoglutarate (αKG) oxygenase, which hydroxylates L-Asn at the β-position (see chapter 5.1 and 6.1) and *hasP* has been annotated as a putative phosphotransferase.[185] Accordingly, CDA variants produced by

this strain lack the phosphohydroxy group at Asn-9 (Figure 2.20; **20**-**21**). CDA is also produced in some closely related *Streptomyces* species (Table 2.2).[176,188]

Cyclic lipopeptides. The following antibiotics are also consistent of a ten-membered ring, but are true peptides, i.e. macrocycles are formed by macrolactamization. This family of undecapeptide antibiotics can be further divided into two sub-groups, the ampho-mycins/friulimicins and the laspartomycins/glycinocins. All representatives of a subclass are very similar. They differ only at one position in the peptide core and in the branching of the fatty acids.

Amphomycins/friulimicins. Since the discovery of amphomycin,[173] similar or identical antibiotics have been published under different names; this has caused significant confusion in this substance class. In 2000, Wink *et al.*[172] reported the structural elucidation of eight bioactive lipopeptides isolated from *Actinoplanes friuliensis*. That work and an extensive review by Baltz *et al.*[176] clarified the former uncertainties in the literature. These lipopeptides can be divided in the friulimicins and the A-1437 compounds. A-1437s (Figure 2.20; **22**-**25**) were found to be linked via the *N*-terminal diaminobutyrate (Dab) to an exocyclic aspartate residue with different fatty acid substituents. The study of Wink *et al.*[172] revealed that the different A-1437s are identical to previously described[176] lipopeptide antibiotics (see Figure 2.20 for details). The friulimicins A-D (Figure 2.20; **26-29**) are closely related to the A-1437s. They differ only in the exocyclic amino acid, which is Asn instead of Asp. The peptide component of all of the above described lipopeptides is featured by unusual amino acids, such as L-threo and D-erythreo-2,3-Dab, D-pipecolinic acid (D-Pip) as well as L-*threo*-3-methyl-Asp (MeAsp) at positions 2, 9, 3 and 4 in the peptide core. Later studies revealed that the minor friulimicin components contain Asp_4 instead of $MeAsp_4$.[189] The gene clusters for the amphomycins/friulimicins were unidentified until recently. Initially, only a fragment of the friulimicins gene cluster was known,[189] but the entire gene cluster has been described recently.[190] It shows an unusual organization; the first peptide synthetase encoding gene, *pstA*, is divided from the next synthetase, *pstB*, by nine genes. The gene products PstC and PstD complete the assembly line. PstA has one, PstB two, PstC six and PstD two adenylation domains responsible for incorporation of the corresponding number of amino acids into the peptide core. The presence of two epimerization domains is consistent with the observed D-Dab_9 and D-Pip_3 residues.

Laspartomycins/glycinocins. Laspartomycin was originally isolated and characterized as a lipopeptide antibiotic related to amphomycin in 1968,[174] but its molecular weight and structure remained unknown. Novel purification methods lead to the determination of the structure of the major component, laspartomycin C (Figure 2.20, **31**).[191] The structure differs from that of the amphomycins/friulimicins and all related antibiotics. The amino acid branching into the side chain is diaminopropionate (Dap_2) rather than Dab. In addition, the fatty acid chain is 2,3-unsaturated. Furthermore, the second D-amino acid is D-Thr_9 instead of D-Dab. Gly replaces the $MeAsp_4$ residue and Ile substitutes the Val residue at position 10. Two additional peptides were isolated during laspartomycin C purification, but these species were not structurally characterized.[191]
Glycinocin A, a compound of the glycinocins family (Figure 2.20; **31**-**34**) was shown to have the identical structure as laspartomycin C (Figure 2.20; **31**).[175] The glycinocin B **32** and C **33** have the same peptide core as **31** and differ only in the length of the fatty acid tail. Glycinocin D **34** on the other hand has the same lipid part as **31** but instead of having Ile at position 11, Val is found. Biosynthetic gene clusters neither for laspartomycin nor glycinocin production are known.

2.4.2 Mode of Action

The mode of action of cyclic lipopeptide antibiotics is still poorly understood. It is known that calcium ions are essential for the antimicrobial activity of daptomycin, CDA and A54145.[176] For daptomycin other divalent ions were tested for antibiotic activity but had at least a 32-fold higher minimal inhibitory concentration (MIC) than Ca^{2+}.[192] Current research is largely focused on the mode of action of daptomycin. Three different NMR structures of apo daptomycin have been published.[193-195] In very recent studies on daptomycin's mode of action by Ho *et al.*,[192] the structure of daptomycin in the presence of Mg^{2+} was solved. The authors proposed a new mode of action for bacterial cell death by daptomycin. The divalent cation binding actually facilitates micelle formation by neutralizing the negative charge of the acidic side chains. In this scenario, the micelles are vehicles to deliver high local concentrations of daptomycin to the bacterial cell membrane. Micelle dissociation once the lipid bilayer is reached would allow the daptomycin lipid tail to insert into the membrane. Inside the membrane oligomerization of daptomycin might occur again for large pore generation (Figure 2.21).

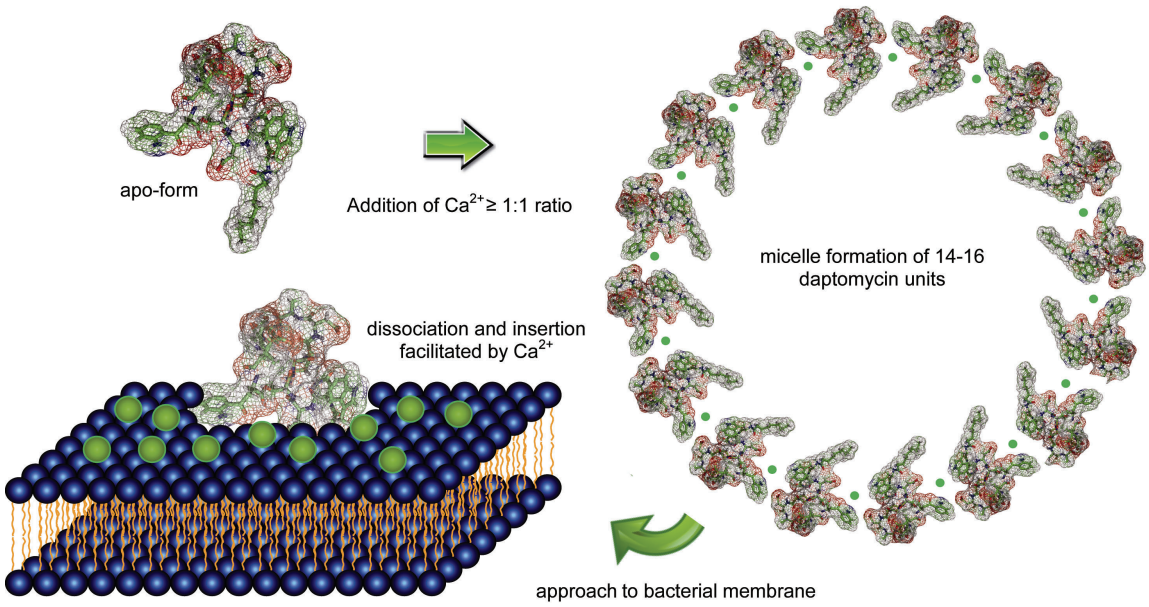


Figure 2.21: Proposed mode of action of daptomycin. The not highly amphiphilic apo form of daptomycin forms micelles of 14-16-mers upon addition of Ca^{2+}. The minimum amount for micelle formation is a 1:1 molar ration. This process might facilitate rapid approach to bacterial membranes. Once the lipid bilayer is reached, the complex dissociates and daptomycin inserts into the membrane. In the final stages it is proposed that daptomycin perforates the membrane by depolarization, which subsequently leads to efflux of essential potassium ions and bacterial cell death (not shown).

The insertion should lead to pores formation, which will lead to potassium efflux and subsequent bacterial cell death.[196] However, the remaining questions if daptomycin will build a multimer in the membrane and if oligomerization is crucial for bactericidal activity as well as if potassium efflux is the cause of cell death or other membrane-associated processes, e.g. cell wall synthesis, division, energetics, etc., are unanswered.

Micelle formation has also been proposed for tshushimycin for which the crystal structure with bound Ca^{2+} at 1.0 Å resolution was solved recently.[197] This work implicates a common mode of action of the tsushimycin/A-1437 B (Figure 2.20; **23**) and the acidic lipodepsipeptides (Figure 2.20; **1**-**21**). The calcium was found to be bound between Asp_1 and Asp_5 of the tshushimycin peptide core.

Although friulimicin B (Figure 2.20; **27**) shows high structural similarities with tshushimycin, only one residue of the peptide core differs (Asp_1 instead of Asn_1), it has a very different mode of action. Studies showed that **27**, in the presence of calcium, formed a complex with a lipid carrier. Friulimicin B appears to block two essential pathways in the cell envelope and the friulimicin lipid carrier complex prevents the formation of a functional cell wall. In contrast to daptomycin, no potassium leakage was observed. This unique mode of action is most likely due to the missing negatively

charged side chain of Asp_1 and offers an explanation for the lack of cross resistance between daptomycin and friulimicin B.[198] Amphomycin (A-1437 E, **24**) was found to target the peptidoglycan biosynthesis and cause the accumulation of UDP-MurNAc-pentapeptide.[199] Furthermore, amphomycin inhibits the formation of mannosylphosphoryldolichol, glucosylphosphoryldolichol and *N*-acetylglucosaminylphosphoryl dolichol in eukaryotic cells, by complex formation with the dolichol monophosphate in the presence of Ca^{2+}.[200]

2.4.3 Tailoring

Structural changes, resulting from tailoring reactions, improve the biological activity of the secondary metabolites and provide the producer with an evolutionary advantage by, for instance, making the antibiotic more stable to protease cleavage by the target strains. Altogether, twelve non-proteinogenic amino acids are present in the acidic lipopeptides (Figure 2.20) and contribute thereby to the observed structural diversity. Six of these (Hpg, Kyn, Dab, Dap, Pip and Orn) have been extensively discussed,[176] therefore this introduction focuses on more recent studies addressing methylation, dehydrogenation and epoxidation mainly in CDA.

Methylation. One common feature of daptomycin, A54145 and CDA is the presence of a β-methylated Glu residue at the same relative ring position. Whereas CDA and A54145 lipopeptides are produced as mixtures of compounds containing MeGlu or Glu,[181,183] fermentation of *Streptomyces roseosporus* results exclusively in MeGlu containing peptides.[201] The MeGlu containing acidic lipopeptides exhibit a higher bioactivity than the Glu-containing analogs.[181,201-204]

Following determination of the CDA biosynthetic gene cluster,[185] a gene, *sco3215*, was predicted to encode a *S*-adenosyl methionine (SAM)-dependent glutamate-3-methyl-transferase (GlmT). Genes homologous to *glmT* are found in the daptomycin (*dptI*)[201] and A54145 (*lptI*)[183] pathway. It was observed previously that fermentation of deletion mutants of *S. coelicolor* (Δ*glmT*) and *S. roseosporus* (Δ*dptGHIJ*) led to production of CDA and daptomycin analogs containing exclusively Glu instead of MeGlu. Complementation of the Δ*dptGHIJ* mutant by *dptI* or *glmT* and complementation of Δ*glmT* with synthetic MeGlu restored the biosynthesis of the MeGlu-containing compounds.[201,204] These results showed that GlmT, DptI and LptI are methyltransferases involved in the

biosynthesis of MeGlu residues in CDA, daptomycin and A54145. By characterization of recombinant GlmT, DptI and LptI, Mahlert *et al*.[205] determined the substrate specificity of the methyltransferases. Combined with the *in vivo* results, complete mechanistic details for the biosynthesis of MeGlu and its incorporation into the acidic lipopeptides were described (Figure 2.22 A). It was shown that GlmT, DptI and LptI are indeed all SAM-dependent methyltransferases. But instead of catalyzing the methylation of Glu directly, all enzymes act exclusively on αKG leading to 3-methyl-2-oxoglutarate. Together with a branched chain aminotransferase (IlvE) from the primary metabolism of *S. coelicolor*, the mechanism was established: The coupling of the methylation and transamination reaction leads to the synthesis of MeGlu starting from αKG. Subsequent activation and incorporation of MeGlu by the adenylation domain of the tenth module of the CDA and the twelfth module of the daptomycin/A54145 NRPSs will lead to the corresponding acidic lipopeptide antibiotics.[205]

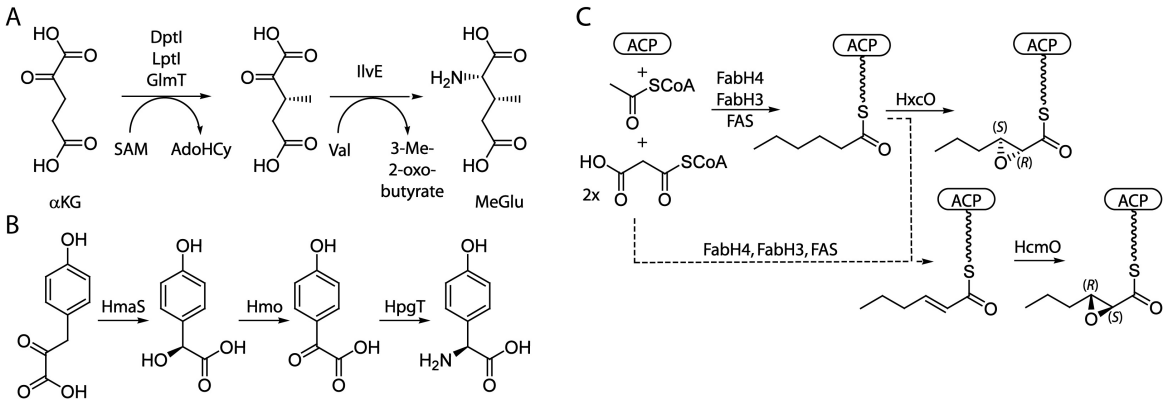


Figure 2.22: Selection of observed tailoring steps during the biosynthesis of acidic lipopeptides (mainly CDA). A) Mechanism of MeGlu precursor synthesis. The SAM-dependent methyltransferases DptI, LptI and GlmT transfer a methyl group to αKG leading to 3-methyl-2-oxoglutarate. Subsequent amino group transfer from valine to the methylated glutarate catalyzed by branched chain aminotransferases yields MeGlu. B) Hydroxyphenylglycine biosynthesis in the CDA biosynthesis of *Streptomyces coelicolor*. C) Proposed mechanism of the fatty acid tailoring events leading to 2,3-epoxyhexanoyl, the fatty acid tail of CDAs.

The methylation of an aspartate residue observed at position four of the amphomycins/friulimicins **22**-**30** (Figure 2.20) was not fully examined biochemically, but there is genetic evidence that this methyl group results from mutase-catalyzed rearrangement of Glu into MeAsp.[189] This notion was validated by *in vitro* activity of the gene products predicted to show the described mutase activity in the friulimicin pathway, GlmA and GlmB, which were expressed heterologously in *S. lividans*. Disruption of these genes in *A. friuliensis* caused a 60% reduction in the yield of MeAsp-

amphomycins. If this is due to a leaky recombinant phenotype or to production by an alternative pathway is unclear.[189]

Hydroxylation. The observed hydroxylated asparagines are found at position 9 in the peptide core of some CDAs (Figure 2.20, **15**-**16**, **18**-**19** or subsequent phosphorylated **12**, **14**, **17**) and at position 3 of A54145 (Figure 2.20, **5**-**12**) and their origin is discussed in chapters 5.1 and 6.1 Additionally, a methoxylated aspartate is found at position 9 in A54145. For the latter modification, a mechanism was proposed by sequence analysis of the genes from the A54145 biosynthesis cluster.[183] The gene *lptK* was annotated as an *O*-methyltransferase because it fits high sequence homology to known enzymes and *lptJ* was annotated as a regulatory protein with sequence homology to the putative regulatory enzyme SyrP, which is a gene product of the syringomycin biosynthetic gene cluster in *Pseudomonas syringae*.[206] The SyrP sequence was recently re-examined and showed high homology to iron(II)/αKG-dependent hydroxylases. Biochemical characterization of SyrP revealed hydroxylation activity on aspartic acid bound to the PCP of the eighth module of syringomycin NRPS.[126] Therefore, a possible MeOAsp precursor pathway is that LptJ hydroxylates Asp at the β-position, while tethered to the PCP, followed by *O*-methylation via LptK catalysis.

In the broadest sense the synthesis of the non-proteinogenic amino acid 4-hydroxyphenylglycine (Hpg, Figure 2.22 B) could be seen as hydroxylation reaction, in which Hpg is incorporated into CDA (Figure 2.20; **13**-**21**). The Hpg synthesis requires HmaS, a 4-hydroxymandelate synthase, to convert 4-hydroxyphenylpyruvate to 4-hydroxymandelate. Two additional enzymes, the 4-hydroxymandelate oxidase (Hmo) and the 4-Hpg transaminase (HpgT), complete Hpg biosynthesis (Figure 2.22 B).[185]

Dehydrogenation. The a-series of the calcium-antibiotics (Figure 2.20; **17**-**21**) exhibit a C-terminal *Z*-2'*S*-3'*S*-dehydrotryptophan (ΔTrp) residue.[185] Little is known about the nature of the enzymes involved in the dehydrogenation of amino acid residues during NRPS biosynthesis. Nevertheless, related enzyme activities have been characterized, namely the tryptophan side chain oxidase from *Pseudomonas* and the L-tryptophan-2'-3'-oxidase from *Chromobacterium violaceum*.[16] No gene was identified within and outside the CDA biosynthetic cluster whose product shares sequence homology to any of these enzymes. Amir-Heidari *et al.*[207] indirectly supported the stereochemical course of the Trp dehydrogenation. By feeding a Trp auxotroph *S. coelicolor* strain with

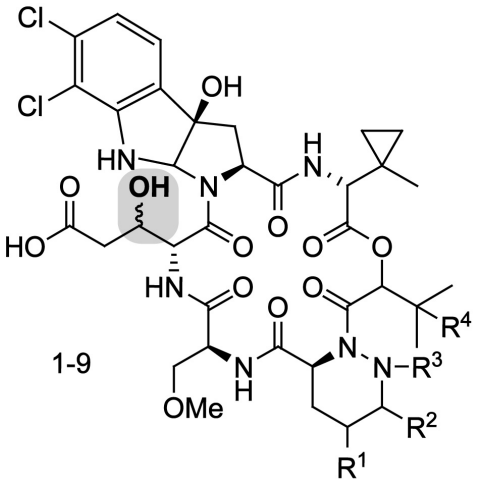
synthetic C3'-deuterated tryptophans, it was possible to follow the fate of the C3' deuterium atoms during CDA biosynthesis using mass spectroscopy. The possibility that hydroxylation is followed by dehydration was ruled out, because no hydroxylated Trp-containing intermediates were observed. It is suggested that the dehydrogenation of Trp during CDA biosynthesis takes place directly as in the case of L-tryptophan-2'-3'-oxidase from *Chromobacterium violaceum*. This oxidase was shown to catalyze the dehydrogenation of *N*-Boc-L-Trp to the ΔTrp derivative abstracting the C3'-*proS*-proton with *syn* stereochemistry,[208] which was also observed in the feeding experiments. The unidentified enzyme responsible for ΔTrp formation during CDA biosynthesis should therefore follow the logic of the L-tryptophan-2'-3'-oxidase.

Fatty acid tailoring. The mechanism of the incorporation of 2,3-epoxyhexanoic acid as an invariant fatty acid tail of the CDAs (Figure 2.20; **13**-**21**) was determined recently by Kopp *et al.*[209] It was found that the tailoring of unique 2,3-epoxyhexanoyl moiety of CDA is acyl carrier protein (ACP) mediated. It was shown that a new type of a FAD-dependent oxidase (HxcO) with intrinsic enoyl-ACP epoxidase activity and a second enoyl-ACP epoxidase (HcmO) are responsible for the generation of the epoxidized fatty acid tail. It was postulated that hexanyol-ACP is produced by FabH4 and FabH3, putative β-ketoacyl-ACP synthases encoded within the *fab* operon, together with enzymes from primary metabolism.[185] Subsequently, fatty acid tailoring by HxcO occurs on the ACP-bound hexanoyl moiety and results in the (2*R*,3*S*)-2,3-epoxyhexanoyl product. Alternatively, the hex-2-enoyl-ACP, a side product of HxcO or derived from primary metabolism, can be epoxidized directly by HcmO during fatty acid synthesis on the ACP (Figure 2.22 C). Based on the finding that two epoxidases exist in the CDA cluster, it is suggested that HxcO has gained its epoxidation activity during evolution. HcmO and HxcO produce the two possible 2,3-epoxyhexanoyl enantiomers, but it remains unclear if both enantiomers are transferred from the ACP to the first module of the CDA peptide synthetase.[209]

2.5 Kutznerides

2.5.1 Structures

The recently discovered antifungal and antimicrobial kutznerides,[18,19] hexadepsipeptides produced by *Kutzneria sp.* 744, are chemically interesting natural products and another excellent example for the observed structural diversity in NRPs. Kutznerides, secondary metabolites produced non-ribosomally by Kutzneria sp. 744, are cyclic hexadepsipeptides and bear a lot of potential to expand and investigate mechanisms underlying the structural diversity found in non-ribosomal natural products. Structural elucidation of kutznerides revealed nine related compounds composed of five unusual non-proteinogenic amino acids and one hydroxy acid (Figure 2.23).



	3- **OH**-Glu	R^1	R^2	R^3	R^4
Kutzneride 1	(*S*)	H	H	H	Me
Kutzneride 2	(*S*)	Cl	H	H	Me
Kutzneride 3	(*R*)	H	H	H	Me
Kutzneride 4	(*R*)	H	π-bond		Me
Kutzneride 5	(*R*)	H	H	H	H
Kutzneride 6	(*R*)	OH	π-bond		Me
Kutzneride 7	(*S*)	H	H	H	H
Kutzneride 8	(*R*)	Cl	H	H	Me
Kutzneride 9	(*S*)	H	p-bond		Me

Figure 2.23: Chemical structures of kutznerides 1-9. The examined hGlu residue is highlighted in grey. All known kutzneride species where found to contain D-β-hydroxylated glutamic acid either in the *erythro* (1, 2, 7, 9) or *threo* (3-6, 8) form.

All kutznerides contain 2-(1-methylcyclopropyl)-glycine (MecPG) connected to the α-hydroxyl moiety of either (*S*)-2-hydroxy-3-methylbutyric or (*S*)-2-hydroxy-3,3-dimethylbutyric acid. The hydroxy acid residue is followed by a piperazic acid moiety, found in four distinct forms: as piperazic acid in kutznerides 1, 3, 5, and 7; as dehydropiperazate in kutznerides 4 and 9; as γ-chloro-piperazate in kutznerides 2 and 8; and as γ-hydroxy-dehydropiperazate in kutzneride 6. Furthermore, kutznerides contain *O*-methyl-L-serine and either the *threo* or *erythro* isomer of 3-hydroxy-D-glutamate (highlighted in grey, Figure 2.23). The origin of the latter will be described in this work. Finally, an unusual tricyclic dihalogenated (2*S*,3a*R*,8a*S*)-6,7-dichloro-3a-hydroxyhexa-hydropyrrolo[2,3-b]indole-2-carboxylic acid (PIC) is conserved in all structurally characterized kutznerides.

2.5.2 Discovery and Organization of the Biosynthetic Gene Cluster

Isolation of the Gene Cluster. The gene cluster encoding the kutzneride biosynthesis enzymes was identified using degenerate primer amplification. Due to the highly chlorinated peptide scaffold of the kutznerides, degenerated primers were designed for both flavin and non-heme iron halogenases. Amplicons obtained by PCR using *Kutzneria sp.* 744 genomic DNA as a template were sequenced and identified two distinct flavin-dependent halogenases and one Fe(II)/αKG-dependent halogenase. From the obtained sequences, specific primers were designed and used to screen a cosmid library of genomic DNA. Two cosmids with sequence overlap, one containing genes for both flavin halogenases and the other with the non-heme iron halogenase were identified and fully sequenced.[158]

Analysis of Kutzneride Gene Cluster. The biosynthetic gene cluster of kutznerides was found to span 56 kb of genomic DNA and to consist of 29 ORFs, 17 of which can be assigned roles in kutzneride biosynthesis. The kutzneride cluster encodes three multidomain NRPS proteins, KtzE, KtzG, and KtzH that provide the six required NRPS modules for kutzneride assembly. It was proposed that the gene cassette KtzABCD would function in the assembly of MecPG, but was shown to catalyze the formation of (1*S,*2*R*)-allocoronamic acid (alloCMA) and therefore is likely not to be part of the kutzneride cluster.[159] The location of the MecPG synthesizing gene cassette thus remains to be discovered. Furthermore, the stand-alone A domain KtzN and the α/β-hydrolase-fold enzyme KtzF were suggested to be part of the gene cluster. After this exclusion seven tailoring enzymes remain in the cluster of which six are likely to catalyze oxidative transformations, including the flavoprotein monooxygenase KtzI, the cytochrome P450 monooxygenase KtzM, the mononuclear non-heme iron dioxygenases KtzO and KtzP (see results and discussion), and the flavin-dependent halogenases KtzQ and KtzR, which were shown to catalyze the formation of 6,7-dichlorotryptophan as part of the tricyclic PIC.[210] KtzS is predicted to be a dedicated flavin reductase. The kutznerides gene cluster harbors a SAM-dependent methyltransferase KtzL and several genes postulated to be involved in regulating the production of and resistance to kutznerides.[158]

Proposed Biosynthesis. The organization of the kutzneride NRPS assembly line is as unusual as the structure (Figure 2.24).

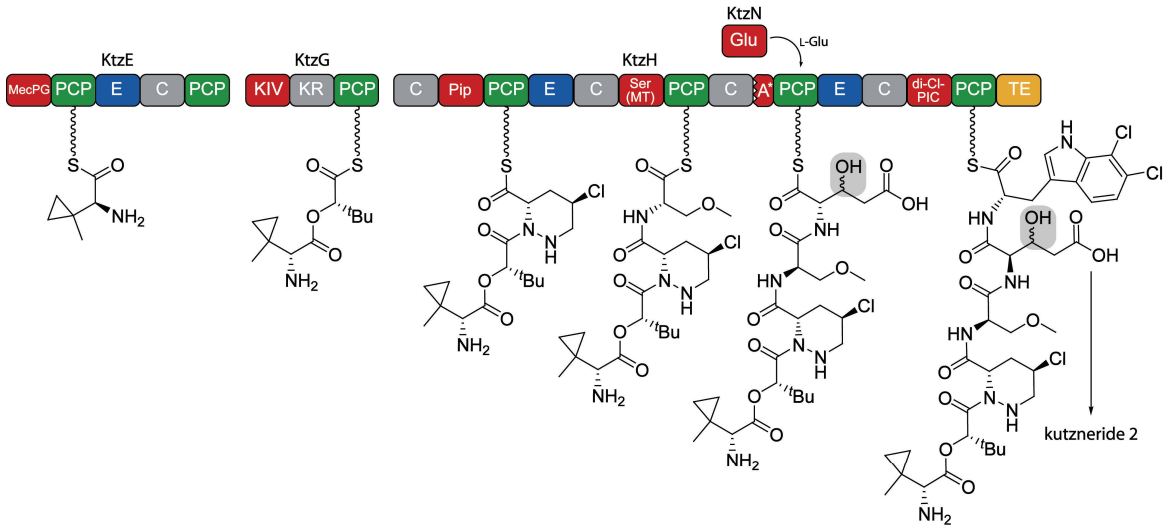


Figure 2.24: Proposed biosynthesis of kutznerides, exemplified for kutzneride 2. Adenylation (A) domains are shown in red, peptidyl carrier protein (PCP) domains in green, condensation (C) domains in grey and epimerization (E) domains in blue. The third A domain of KtzH (A*) is only 88-aa in size and non-functional. The thioesterase (yellow) is proposed to cleave the mature kutzneride from the assembly line by cyclization. The site of glutamic acid hydroxylation is highlighted in grey.

Kutznerides are thought to be assembled by three non-ribosomal peptide synthetases, KtzE, KtzG and KtzH. KtzE is predicted to activate MecPG and catalyze its condensation with the hydroxy acid, which is activated by the adenylation (A) domain of KtzG. KtzG was shown to adenylate 2-keto-isovaleric acid (KIV) and it is postulated that after *in situ* reduction of the keto function by the ketoreductase domain (KR), a methyltransferase forms the *t*-butyl group found in many kutznerides. Whether the methylated hydroxy acid is transferred onto the second peptidyl carrier protein (PCP) domain of KtzE prior to condensation is unclear. KtzH contains the remaining four modules needed for kutzneride assembly. The first module is predicted to activate Pip and the second to activate Ser. The A domain of module two of KtzH is also thought to be responsible for *O*-methylation because it contains a methyltransferase. The third module of KtzH deviates from standard NRPS logic. The A domain found in module three, designated as A*, is only approximately 20% of the size of typical A domains[39] and thus likely to be non-functional. The lack of a functional amino acid activating domain in the third module may be compensated by the stand-alone A domain KtzN, which was predicted to activate Glu. Lastly, the fourth module of KtzH is predicted to incorporate the differently substituted PICs prior to cyclization by the thioesterase (TE) domain and peptide release. The presence of three epimerization (E) domains in the NRPSs agrees with the occurrence of three D-amino acids.[158]

2.6 Tasks of This Study

One aim of this study was to elucidate the biosynthetic origin of the 3-hydroxy-asparagine (hAsn) moieties of the acidic lipopeptide antibiotics CDA and A54145. In addition to the validation of the involvement of the putative non-heme iron(II)/αKG-dependent hydroxylases AsnO (CDA) and LptL (A54145), the biochemical characterization of these should give clues about the enzymes' substrate and stereo-specificity. Furthermore, it should give rise to the mechanistic aspects of hAsn biosynthesis in the two systems. In addition, the preliminary crystal structure of AsnO should be refined and discussed. Based on these results, the potency to predict the reactivity of AsnO-like enzymes and the nature of AsnO as archetype for free amino acid modifying hydroxylases should be examined.

In the context of β-hydroxylated amino acid generating enzymes, like AsnO and LptL, the role of the putative non-heme iron oxygenases KtzO and KtzP, predicted to hydroxylate PCP-bound glutamic acid to yield 3-hydroxyglutamic acid (hGlu) in the biosynthesis of the antifungal and antimicrobial kutzneride, should be assigned. By the biochemical characterization of these enzymes, it should be possible to explain the observance of two distinct hGlu isomers, i.e. *threo* and *erythro*. Furthermore, the potency of non-hydrolyzable PCP-substrates for the kinetic parameter determination of *in trans* to the NRPS assembly line acting enzymes should be determined. In addition the characterization of the, likely *in trans* NRPS restoring, adenylation enzyme KtzN should give a complete mechanistic pathway for hGlu incorporation into kutznerides and finally should give insights into NRPS assembly line mechanism, which differ from the classical co-linearity principle.

After a solid foundation regarding the mechanisms underlying β-hydroxylated amino acid generation either in precursor pathway (AsnO and LptL) or in thiotemplated fashion (KtzO/KtzP) was gained, the acquired knowledge should be exploited in hydroxylase engineering efforts. The possibility of substrate selectivity manipulation by rational design of AsnO should be investigated and the potency of one AsnO variant (AsnO D241N) to synthesize the pharmaceutical relevant neurotransmitter inhibitor L-*threo*-hydroxyaspartic acid even on a preparative scale should be resolved.

3. Materials

3.1 Chemicals, Enzymes, Consumables

All chemicals not listed below were purchased from Sigma Aldrich (Schnelldorf, Germany or St. Louis, MO, USA) in p.a. quality.

Table 3.1: Chemicals, enzymes, consumables, general materials.

Manufacturer (site)	Product
AppliChem (Darmstadt, Germany)	ampicillin, kanamycin, potassium phosphate, potassium hydrogen phosphate, media components
Bio-Rad (Hercules, CA, USA)	Econo-Pac 10DG desalting columns Bio-Gel P-6 gel, gels for PAGE
Corning Incorporated (Corning, NY, USA)	consumables (pipette tips, Petri dishes, inoculators, cell scraper)
Dana Farber Cancer Institute (Boston, MA, USA)	DNA dideoxy sequencing
EMD Biosciences (Bad Soden, Germany)	N^{α}-protected amino acids
Eppendorf (Hamburg, Germany)	1.5 and 2.0 mL reaction tubes
Eurofins MWG Operon (Cologne, Germany)	oligonucleotides
Eurogentech (Seraing, Belgium)	agarose, electroporation cuvettes
Fermentas (St. Leon-Rot, Germany)	protein markers
GATC (Konstanz, Germany)	DNA dideoxy sequencing
GE Healthcare (Freiburg, Germany)	Hi-Trap desalting columns size exclusion chromatography columns
Integrated DNA Technologies (Coralville, IA, USA)	oligonucleotides
Invitrogen	*E. coli* strains (BL21, TOP10, BL21 Star)
Macharey und Nagel (Düren, Germany)	standard HPLC columns
Millipore (Billerica, MA, USA)	Ultracel-YM10 and Ultracel-YM30 protein concentration discs
Millipore (Schwalbach, Germany)	Amicon Ultra-15 concentrators
New England Biolabs (Ipswitch, MA, USA)	Phusion DNA polymerase, restriction endonucleases, DNA ladders, protein markers
Phenomenex (Torrance, CA, USA)	chiral and fusion HPLC columns
Qiagen (Hilden, Germany)	QIAprep Miniprep kit, QIAquick gel extraction and PCR purification kit, Ni-NTA-resin
Roth (Karlsruhe, Germany)	ethidium bromide, acrylamide solution
Sarstedt (Nümbrecht, Germany)	pipette tips, Falcon tubes (15 and 50 mL)
Stratagene (La Jolla, CA, USA)	Quick Change II site directed mutagenesis kit, including *E. coli* XL-10 Gold
Thermo Scientific (Waltham, MA, USA)	Hypercarb HPLC columns
Tocris Bioscience (Ellisville, MO, USA)	L-threo-hydroxyaspartic acid, L-erythro-hydroxyaspartic acid
VWR (West Chester, PA, USA)	electroporation cuvettes, 15 and 50 mL Falcon tubes, disposable culture tubes, sterile filters, inoculators

3.2 Equipment

Table 3.2: Equipment used in this study with manufacturers.

Device	Manufacturer/trade name
Anaerobic chamber (glove box)	M. Braun LABstar workstation; Coy Laboratory Basic vinyl glove box
Autoclave	Tuttnauer 5075 ELV; HMS in-house
Centrifuges	Heraeus Biofuge pico; Sorvall RC 5B Plus (SS-34, SLC-300 rotors); Sorvall RC 5B Plus (SLC-4000 rotor); Eppendorf 5415D; Eppendorf 5810R; VWR Galaxy microcentrifuge 16
Clean bench	Antair BSK
Documentation system	Cybertech CS1; Mitsubishi thermo printer video copy
Electrophoresis	Agarose gel chamber (manufactured in-house Philipps-University Marburg); Bio-Rad Mini-Protean III-PAGE chamber
Electroporation	Bio-Rad gene pulser electroporator
Fast Protein Liquid Chromatography (FPLC)	Pharmacia FPLC-System 250: Gradient programmer GP-250, pump P-500, Uvicord optical device UV-1 (λ= 280 nm), Uvicord control element UV-1, 2-channel flat bed printer, injection valve V-7, 3-way-magnet valve PSV-100, fraction collector Frac-100
French-Press	SLM Aminco french pressure cell version 5.1, 20k Rapid-fill Cell (35 mL)
HPLC-System	Agilent series 1100 HPLC-system with DAD-, FLD- and ESI-QUAD-MS-detection, vacuum degasser, quaternary pump, auto sampler, HP-ChemStation software; Beckman-Coulter Gold System with UV detector 168
Incubators	Eppendorf Thermomixer Comfort, New Brunswick Scientific Series 25 Incubator, New Brunswick Scientific Innova 4300 Incubator
MALDI-TOF	Per Septive Biosystems Voyager-DE RP BioSpectrometry coupled with Bruker Flex III
NMR	Varian 600 Unity Plus; Bruker Avance 300
Photometers	PEQLab Nanodrop ND-1000; Pharmacia Ultraspec 3000
Q-TOF MS	Applied Biosystems API QStar Pulsar i
Pipettes	Eppendorf Research series; Abimed Labmate
Speed-Vac	Uniequip Univapo 150
Thermocycler	Eppendorf Mastercycler Personal
Vortexer	Scientific Industries Vortex Genie 2

3.3 Plasmid Vectors

3.3.1 pQTEV

The pQTEV vector (GenBank accession number: AY243506) is a derivatized pQE30 vector (Qiagen) allowing expression of the cloned gene of interest as a N-terminal His_7-fusion (Figure 3.1). Additionally, the vector encodes for a tobacco etch virus (TEV) protease recognition site between the overproduced protein and the His_7-tag to allow cleavage after purification. The pQTEV vector carries two *lac*-operators in the promoter region and thus induction of the overexpression of the gene of interest is achieved by IPTG addition.

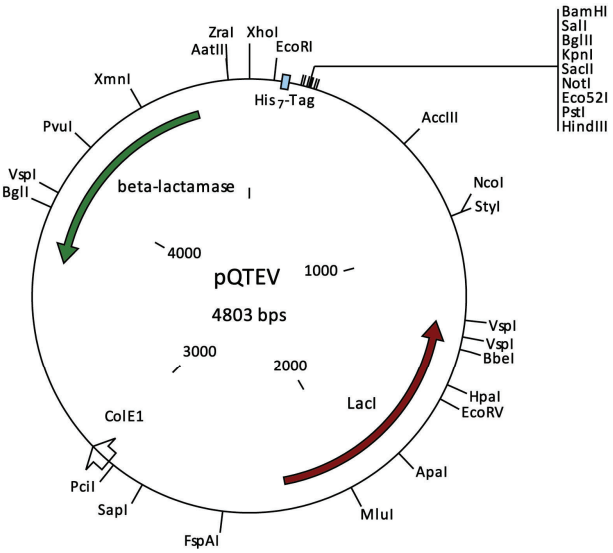


Figure 3.1: Physical map of the pQTEV vector.

3.3.2 pET28a(+)

The pET28a(+) vector system (Novagen) is an expression vector and was used to produce proteins recombinantly (Figure 3.2). The vector allows Ni^{2+}-NTA affinity chromatography purification by introduction of a His_6-tag fusion to the C- or N-terminus of the overproduced protein. The transcription of the cloned genes is dependent on the T7 RNA polymerase and can be induced by IPTG addition.

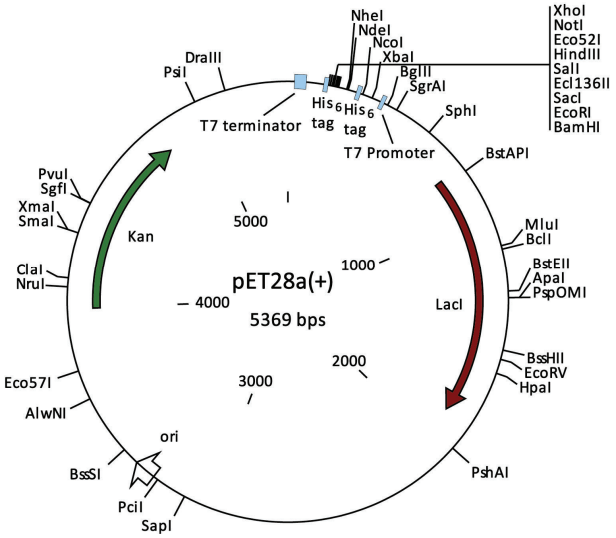


Figure 3.2: Physical map of pET28a(+).

3.3.3 pMAL-c2X

In the pMAL-c2X vector the gene of interest is inserted downstream from the *malE* gene, which encodes maltose-binding protein (MBP). This results in the expression of an MBP-fusion protein (Figure 3.3). The vector contains the strong P_{tac} promoter and the translation initiation signals of MBP to express the fusion protein upon addition of IPTG. The fusion protein can be purified by an amylose column, which specifically binds MBP. The pMAL-c2X vector includes a sequence coding for the Factor Xa protease recognition site adjacent to the multiple cloning site. This allows the protein of interest to be cleaved from MBP after purification.

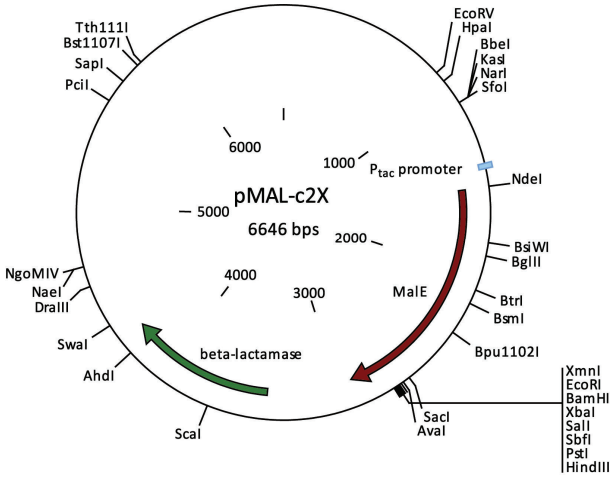


Figure 3.3: Physical map of pET28a(+).

3.4 Oligonucleotides

The oligonucleotides which were used as PCR primers to amplify the targeted genes or for site directed mutagenesis are listed in Table 3.3.

Table 3.3: PCR primers used in this study (restriction/mutagenesis sites are underlined).

Oligoname	Sequence (5'→ 3')	Restriction Enzyme	Vector	Targeted Gene
5'-asnO	AAAAAAGGATCCGCTGCGAATGCCGCGGG	BamHI	pQTEV	*asnO*
3'-asnO	AAAAAAGCGGCCGCTCAGGCGGGCTGCGGG	NotI		
5'-cdaPCP$_9$	AAAAAAGGATCCGGCCGGGCCCCGGTCACGG	BamHI	pQTEV	*PCP$_9$* of *cdaPSII*
3'-cdaPCP$_9$	AAAAAACTGCAGCTGCGCCTCGGCGAGTTCGA	PstI		
5'-lptL	AAAAAAGGATCCATGGAACCCGAGAACACCTT	BamHI	pET28a(+)	*lptL*
3'-lptL	AAAAAAAAGCTTTCAGCTCAGTACACGCGAAT	HindIII		
5'-lpt-PCP$_3$	AAAAAAGGATCCTACGGCACCGCCTCCGTCG	BamHI	pET28a(+)	*pcp$_3$* of lptA
3'-lpt-PCP$_3$	AAAAAAAAGCTTTCACTGTGCGGCGATCTGCGG	HindIII		
5'-ktzN	AAAAAATCTAGAATGGAAACCGGGACCGG	XbaI	pMAL-c2X	*ktzN*
3'-ktzN	AAAAAACTGCAGGGACCGGCGTGCGAGGGCG	PstI		
5'-ktzO	AAAAAAGGATCCATGACGAACTCGACGGACGA	BamHI	pQTEV	*ktzO*
3'-ktzO	AAAAAAGCGGCCGCTCAGCGGTCGCGGTGGACGG	NotI		
5'-ktzP	AAAAAAGGATCCGTGACTGTGGAACCCAGGCG	BamHI	pQTEV	*ktzP*
3'-ktzP	AAAAAAGCGGCCGCTCAGCGGGAATCGAGAGTCG	NotI		
5'-A*PCP$_3$	AAAAAAGGATCCGACCTCGCGCTGGTCGAGGC	BamHI	pQTEV	*a*pcp$_3$* of *ktzH*
3'-A*PCP$_3$	AAAAAAGCGGCCGCTCAGTCGGCCACCCGTGCGGCCA	NotI		
5'-PCP$_3$	AAAAAAGGATCCATCAGGGCCCCGCGGACGGA	BamHI	pQTEV	*pcp3* of ktzH
3'-A*PCP$_3$	AAAAAAGCGGCCGCTCAGTCGGCCACCCGTGCGGCCA	NotI		
5'-D241N	CCCCGACCTGCGGGTGAACCTGGCGGCCACCGAGC	muta-genesis	pQTEV	*asnO* D241N
3'-D241N	GCTCGGTGGCCGCCAGGTTCACCCGCAGGTCGGGG			

3. 5 Microorganisms

3.5.1 *E. coli* TOP10

E. coli TOP 10 (Invitrogen) was routinely used for standard cloning and sequencing purposes. The genotype is: F-mcrA. (mrr-hsdRMS-mcrBC) 80lacZ.M15.lacX74 deoR recA1 araD139. (ara-leu)7697 galU galK rpsL (StrR) endA1 nupG.

3.5.2 *E. coli* XL-10 Gold

The cloning strain *E. coli* XL-10 Gold was used for site directed mutagenesis, as the Hte phenotype results in high transformation efficiencies for large plasmid inserts. The genotype is: ndA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Hte Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 tet^R F'[proAB $lacI^qZ\Delta M15$ Tn10(Tet^R Amy Cm^R)].

3.5.3 *E. coli* BL21(DE3)

E. coli BL21(DE3) (Invitrogen), with the following genotype F^- ompT gal dcm lon $hsdS_B(r_B^- m_B^-)$ λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]), was used as expression strain for plasmid vectors. Facilitated by the knock-out of Ion and OmpT proteases, the overproduced recombinant proteins have an extended half-life within the bacterial host, compared to other expression strains. Furthermore, the strain contains IPTG-inducible T7 RNA polymerase, which is inserted in the chromosome after *lacZ* and the promoter *lac*uV5 on a λ-prophage. This is mandatory for the IPTG induction of genes on expression plasmids under T7 promoter control.

3.5.4 *E. coli* BL21 Star (DE3)

Alternatively, *E. coli* BL21 Star (DE3) (Invitrogen) was used as expression strain. In addition to the genotype mentioned above (3.5.3), the BL21 Star strain contains a mutation in the gene encoding RNase E (*rne*131), which is one of the major sources of mRNA degradation in conventional BL21 strains. Therefore, BL21 Star cells improve the stability of the gene of interest mRNA transcripts and increase protein expression yield from T7 promoter-based vectors.

3.5.4 Streptomyces

Streptomyces coelicolor A3(2) (DSM 40783) and *Streptomyces fradiae* (NRRL 18158) were cultivated for chromosomal DNA retrieval.

3.6 Media

The following media were used for the cultivation of the microorganisms stated above. For cultivation on agar plates Agar No 1 was added to a concentration of 1.5% (w/v) prior to autoclaving (121°C, 1.5 bar, 20 min). Selection antibiotics were added to the autoclaved medium (temperature below 55°C) in the following concentrations: kanamycin 50 µg/mL, ampicillin 100 µg/mL.

3.6.1 LB Medium

E. coli strains were grown in LB medium with a pH adjusted to 7.0.

LB medium (1 L)
10 g tryptone
5 g yeast extract
5 g sodium chloride

3.6.2 Medium 65

Streptomyces species were grown in medium 65 with a pH adjusted to 7.2.

Medium 65 (1 L)
4 g glucose
4 g yeast extract
2 g calcium carbonate
10 g malt extract

4. Methods

4.1 Molecular Biology Techniques

4.1.1 Cultivation of *Streptomyces* Species and Genomic DNA Isolation

Liquid media (5 mL each) were inoculated with freeze-dried spores of the corresponding *Streptomyces* species, as supplied by the microbial strain collection (DSMZ or NRRL). The cultures were grown for five days (28°C, 250 rpm) and harvested by centrifugation (13,000 rpm, 3 min). Pellets were washed with 1 mL double distilled water, suspended in 500 µL lysis buffer (100 mM TRIS pH 8.0, 50 mM EDTA, 1% (w/v) SDS) and acid washed glass beads were added to final volume of 1.25 mL. The reaction tube was vortexed for two minutes. The liquid phase was recovered by perforating the bottom of the reaction tube, followed by brief centrifugation into another tube. 275 µL of 7 M ammonium acetate (pH 7.0) solution were added and incubated for 5 min at 65°C, followed by 5 min incubation on ice. Chloroform (500 µL) was added, briefly shaken and centrifuged (13,000 rpm, 2 min). The aqueous phase was transferred into a new tube and genomic DNA was precipitated by addition of isopropyl alcohol (1 mL) and incubation at RT for 5 min, followed by centrifugation (13,000 rpm, 5 min). The DNA pellet was washed twice (200 µL 70% aq. EtOH), dried for 5 min at RT and dissolved in 200 µL dd. H_2O.

4.1.2 Plasmid DNA Preparation

4 mL LB medium were inoculated with a single colony and grown over night. 2 mL of the culture were transferred into a reaction tube and centrifuged (13,000 rpm, 1 min). The supernatant was discarded and the resulting pellet was suspended in 200 µL of buffer P1 (50 mM TRIS pH 8.0, 10 mM EDTA, RNase A 100 µg/mL). 200 µL of buffer P2 (0.2 M aq. NaOH, 1% (w/v) SDS) were added and the tube was shaken gently. After incubation at RT for 5 min, 200 µL of buffer P3 (2.55 M aq. KOAc pH 4.8) were added, again, the tube was shaken gently and spun at RT (13,000 rpm, 5 min). 590 µL of the supernatant were transferred into a new tube and 900 µL isopropyl alcohol were added, followed by centrifugation (13,000 rpm, 10 min, 4°C). The supernatant was discarded and the pellet was washed twice (200 µL 70% aq. EtOH), dried at 37°C for 10 min and dissolved in 40 µL dd. H_2O.

4.1.3 Construction of Expression Plasmids

Amplification of the targeted genes or gene fragments were carried out by polymerase chain reaction (PCR) using the Fusion polymerase (New England Biolabs) according to the manufacturer's manual for GC-rich templates. Purification of the PCR products was achieved by the usage of the QIAquick PCR purification kit (Qiagen). Digestion with restriction endonucleases and ligation was carried out according to the protocol provided by the manufacturer (New England Biolabs). Plasmids were isolated as described above (4.1.2) or by usage of the QIAprep Miniprep kit (Qiagen). All constructed expression plasmids were analyzed and validated by restriction mapping and dideoxy sequencing (GATC or DNA sequencing facility Dana-Farber Cancer Institute). Transformation of either E. coli BL21 (DE3) with plasmids containing *Streptomyces* genes or of E. coli BL21 Star (DE3) with plasmids containing *Kutzneria* genes was carried out following the manufacter's protocol (Invitrogen).

Construction of pQTEV(*asnO*). The *asnO* gene was amplified from genomic DNA of *S. coelicolor* A3(2) by PCR using the oligonucleotides listed in Table 3.3. The resulting amplicon was digested with the corresponding enzymes (Table 3.3) and cloned into appropriate restriction sites of pQTEV.

Construction of pQTEV(*cdaPCP*$_9$). The *cdaPCP*$_9$ gene fragment coding for the ninth PCP domain of the CDA peptide synthetase (*cdaPSII*) was amplified from genomic DNA of *S. coelicolor* A3(2) by PCR using the oligonucleotides listed in Table 3.3. The resulting amplicon was digested with the corresponding enzymes (Table 3.3) and cloned into appropriate restriction sites of pQTEV.

Construction of pET28a(+)(*lptL*). The *lptL* gene was amplified from genomic DNA of *S. fradiae* by PCR using the oligonucleotides listed in Table 3.3. The resulting amplicon was digested with the corresponding enzymes (Table 3.3) and cloned into appropriate restriction sites of pET28a(+).

Construction of pET28a(+)(*lpt-pcp*$_3$). The *lpt-pcp*$_3$ gene fragment coding for the third PCP domain of the A54145 synthetase (*lptA*) was amplified from genomic DNA of *S. fradiae* by PCR using the oligonucleotides listed in Table 3.3. The resulting amplicon was

digested with the corresponding enzymes (Table 3.3) and cloned into appropriate restriction sites of pET28a(+).

Construction of pMAL-c2X(*ktzN*). The *ktzN* gene was amplified from fosmid DNA containing the kutzneride biosynthesis cluster by PCR using the oligonucleotides listed in Table 3.3. The resulting amplicon was digested with the corresponding enzymes (Table 3.3) and cloned into appropriate restriction sites of pMAL-c2X.

Construction of pQTEV(*ktzO*). The *ktzO* gene was amplified from fosmid DNA containing the kutzneride biosynthesis cluster by PCR using the oligonucleotides listed in Table 3.3. The resulting amplicon was digested with the corresponding enzymes (Table 3.3) and cloned into appropriate restriction sites of pQTEV.

Construction of pQTEV(*ktzP*). The *ktzP* gene was amplified from fosmid DNA containing the kutzneride biosynthesis cluster by PCR using the oligonucleotides listed in Table 3.3. The resulting amplicon was digested with the corresponding enzymes (Table 3.3) and cloned into appropriate restriction sites of pQTEV.

Construction of pQTEV(*a*PCP*$_3$). The *a*PCP*$_3$ gene fragment coding for the truncated A domain and third PCP domain of *ktzH* was amplified from fosmid DNA containing the kutzneride biosynthesis cluster by PCR using the oligonucleotides listed in Table 3.3. The resulting amplicon was digested with the corresponding enzymes (Table 3.3) and cloned into appropriate restriction sites of pQTEV.

Construction of pQTEV(*PCP*$_3$). The *PCP*$_3$ gene fragment coding for the third PCP domain of *ktzH* was amplified from fosmid DNA containing the kutzneride biosynthesis cluster by PCR using the oligonucleotides listed in Table 3.3. The resulting amplicon was digested with the corresponding enzymes (Table 3.3) and cloned into appropriate restriction sites of pQTEV.

4.1.4 Site Directed Mutagenesis

The mutagenesis was carried out using the *QuickChange II Kit* (Stratagene) in accordance to the manufacturer's manual. The previously described pQTEV(*asnO*) expression vector served as the template. A point mutation was introduced in the aspartic acid codon (GAC, bases 721-723, D241 in the gene product) by PCR of the entire vector using the oligonucleotides listed in Table 3.3. The exchange of the 721st base guanine to adenine

resulted in an asparagine codon (AAC, bases 721-723) and therefore the gene product is the AsnO D241N variant. The identity of the constructed mutation-carrying plasmid pQTEV(*asnOD241N*) was confirmed by DNA dideoxy sequencing (GATC).

4.2 Biochemical Techniques

Standard protein analysis techniques, e.g. SDS-PAGE and gel staining, were carried out as described elsewhere.[211]

4.2.1 Gene Expression

4.2.1.1 Expression of pQTEV Vectors

Expressions were carried out by inoculating 500 mL LB medium supplemented with ampicillin (100 µg/mL) in a 2 L culture flask with 5 mL of an overnight culture of the corresponding production strain. Overproduction of the protein of interest was carried out on a 5 L scale. Initially, cells were grown at 34°C to an optical density of 0.6 (λ= 600 nm), induced with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and again grown at 28°C for 4 h. The cells were harvested by centrifugation (7,000 rpm, 4°C, 15 min), suspended in 15 mL buffer (50 mM HEPES, 250 mM NaCl, pH 8.0) and stored at -20°C.

4.2.1.2 Expression of pET28a(+) Vectors

Expressions were carried out by inoculating 500 mL LB medium supplemented with kanamycin (50 µg/mL) in a 2 L culture flask with 5 mL of an overnight culture of the corresponding production strain. Overproduction of the protein of interest was carried out on a 5 L scale. Initially, cells were grown at 37°C to an optical density of 0.5 (λ= 600 nm), induced with 0.1 mM IPTG, and again grown at 30°C for 4 h. The cells were harvested by centrifugation (7,000 rpm, 4°C, 15 min), suspended in 15 mL buffer (50 mM HEPES, 250 mM NaCl, pH 8.0) and stored at -20°C.

4.2.1.3 Expression of pMAL-c2X Vectors

Expressions were carried out by inoculating 2 L LB medium supplemented with ampicillin (100 µg/L) and glucose (2 g/L) in a 5 L culture flask with 25 mL of an overnight culture of the corresponding production strain. Overproduction of the protein of interest was carried out on a 12 L scale. Initially, cells were incubated at 35°C, then 25°C, and finally 15°C until an optical density of 0.5 (λ= 600 nm) was reached. Protein expression was

induced with 0.3 mM IPTG and cultures were incubated for an additional 15 h at 15°C. The cells were harvested by centrifugation (7,000 rpm, 4°C, 15 min), suspended in 35 mL buffer (20 mM TRIS, 200 mM NaCl, 1mM EDTA, pH 7.4) and stored at -20°C.

4.2.2 Protein Purification

4.2.2.1 His-tagged Proteins

For purification of the His-tagged proteins, cell suspensions were thawed and disrupted either by using an EmulsiFlex-C5 High Pressure Homogenizer (Avestin) or a French Pressure Cell (SLM Aminco). After centrifugation (17,000 rpm, 4°C, 30 min) the supernatant was carefully removed and the recombinant protein was purified by Ni-NTA (Qiagen) affinity chromatography using a FPLC system (Amersham Pharmacia Biotech). The Ni column was equilibrated in 50 mM HEPES, 250 mM NaCl pH 8.0 buffer, the supernatant was injected with a flow rate of 0.7 mL/min and an increasing gradient of imidazole was employed (7.5 mM to 250 mM in 30 min). Fractions containing the recombinant protein were identified by SDS-PAGE analysis, combined, and subjected to buffer exchange into 25 mM HEPES, 50 mM NaCl, pH 7.0 using HiTrap Desalting Columns (GE Health Care). The proteins were aliquoted, flash-frozen in liquid nitrogen and stored at -80°C until usage.

4.2.2.2 MBP-tagged Proteins

For purification of the MBP-tagged proteins, cell suspensions were thawed and disrupted by using an EmulsiFlex-C5 High Pressure Homogenizer (Avestin). After centrifugation (17,000 rpm, 4°C, 30 min) the supernatant was carefully removed and diluted to 200 mL with column buffer (20 mM TRIS, 200 mM NaCl, 1mM EDTA, pH 7.4). The MBP-fusion protein was purified by loading the diluted cell lysate onto an amylose gravity flow column (2.5 x 10 cm, 15 mL amylose resin, equilibrated with 8 column volumes of column buffer) with a flow rate of approximately 1 mL/min. The column was washed with 12 column volumes of column buffer. Fractional elution of the MBP-fusion protein was carried out with column buffer containing maltose (10 mM). Fractions containing the recombinant protein were identified by SDS-PAGE analysis, combined, and subjected to buffer exchange into 25 mM HEPES, 50 mM NaCl, 2 mM $MgCl_2$, 10%

glycerol, pH 8.0 using HiTrap Desalting Columns (GE Health Care). The proteins were aliquoted, flash-frozen in liquid nitrogen and stored at -80°C until usage.

4.2.3 Protein Concentration Determination

The protein concentrations were determined spectrophotometrically using calculated extinction coefficients at 280 nm. The coefficients were determined using *Protean* software (DNAStar).

Recombinant Protein	Calculated Extinction Coefficient (mg/mL)
AsnO	0.95
CDA-PCP_9	4.53
LptL	1.64
Lpt-PCP_3	7.57
KtzN	0.90
KtzO	1.00
KtzP	0.66
A*PCP3	4.36
PCP3	2.95

Table 4.1: Calculated extinction coefficients (λ= 280 nm).

4.3 Analytical Methods

4.3.1 MALDI-MS

Matrix assisted laser desorption/ionization mass spectrometry (MALDI-MS) is an analytical method to determine the molecular mass of peptides and proteins in high vacuum. In the ion source of the mass spectrometer a stream of gaseous ions is generated by distinct laser pulses. The matrix sensitive for the applied laser wavelength is excited by the laser pulse and the excitation energy is gently transferred to the analyte. Samples were prepared by mixing 0.5 µL of the sample with 0.5 µL of DHB-matrix solution (Agilent Technologies), the mixture was transferred onto the metallic probe target and dried at RT. Co-crystals of matrix and analyte were measured on a Bruker Flex III system.

4.3.2 HPLC-MS

High performance liquid chromatography (HPLC) was used to characterize substrates and products by retention time on a chromatography column and by mass spectrometric analysis. Reversed-phase (RP) chromatography is based on hydrophobic interactions with the non-polar stationary phase (carbon, C_{18} or C_4 coated silica gel). Elution is

mediated by non-polar acetonitrile, which competes with the adsorbed analytical compounds for non-covalent binding positions. For the chiral separation of optical isomers, chiral ligand exchange chromatography (CLEC) was used instead of RP-HPLC. In CLEC complexes of transition metals (here: Cu(II)-ions) and enantiomeric molecules are formed. Chiral separation is a result of the differences between the free energies of the intermediate diasteromeric complexes formed between the chiral stationary phase (D-penicillamine) and the analyte. Samples molecules must have two polar functional groups with the correct spacing, which can simultaneously act as a ligand for the divalent transition metal ions. For this reason, CLEC was used for the separation of L-*threo* and L-*erythro*-hydroxyaspartic acid. In each case the retention time of the analytical compound is monitored by UV-vis or fluorescence detection. An electrospray ionization mass detector allows the mass analysis of liquid compounds at atmospheric pressure. Ionization of the analytical compound was achieved by addition of 0.1% TFA or 20 mM nonafluoropentanoic acid as ion pairing reagents. In general, experiments were carried out on an Agilent 1100 system or on a Beckman Coulter Gold system. For the protein mass fingerprinting from tryptic digestions and for the measurements of whole proteins quadrupole time-of-flight (QTOF) detection was used. For the acquisition of high accuracy MS data Fourier transform ion cyclotron resonance-MS (FT-ICR) was applied. The high resolution technique FTICR-MS allows composition elucidation of molecules based on their accurate mass. Experiments were carried out on a Finnigan-LTQ-FT-mass spectrometer (Thermo Electron).

4.3.3 Peptide Mass Fingerprinting

The identities of the produced recombinant proteins were verified by peptide mass fingerprinting. Gel bands of correct size were excised after SDS-PAGE analysis. Single bands were transferred into separate 1.5 mL reaction tubes, wash solution (200 µL, 200 mM NH_4HCO_3, 50% (v/v) MeCN) was added and incubated at 37°C for 30 min. Solvents were removed using a speed-vac manifold and the bands were vacuum dried (37°C, 30 min). *In-band* tryptic digestion was carried out by addition of trypsin solution (20 µL, 0.02 µg/µL trypsin, 10% NH_4HCO_3, 10% (v/v) MeCN, pH 8.1). Excess trypsin solution was removed after initial incubation at 37°C for 45 min, followed by additional incubation for 16h. Peptide fragments were eluted with diffusion solution (25 µL, 1% (v/v) TFA, 10%

(v/v) MeCN, pH 8.1), while sonicated (45 min, RT). Sample analysis was carried out by nano spray-HPLC-QTOF-MS. Proteins were identified by comparison of the peptide mass fingerprints with the MASCOT database.

4.4 Chemical Syntheses

4.4.1 Synthesis of Amino-Coenzyme A

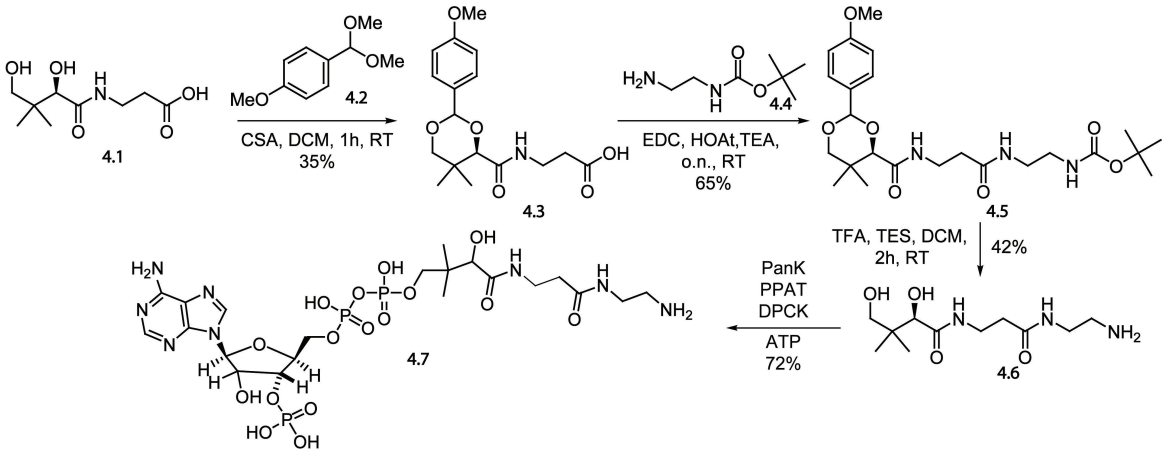


Figure 4.1: Synthesis of amino coenzyme A.

Step 1. Protection of pantothenic acid.

Chemicals: pantothenic acid hemicalcium salt, *p*-anisaldehylde dimethylacetate (PADA), (±)-camphor-10-sulphoric acid (CSA).

compound	MW(g/mol)	n/mmol	equivalents	amount
Pantothenic acid	238.27	64.1	1.0	15.3 g
PADA	182.22 (ρ = 1.07)	64.1	1.0	10.9 mL
CSA	232.29	6.4	0.1	1.5 g

Pantothenic acid hemi calcium salt was dissolved in 200 mL of double distilled water and the pH value was adjusted to 3 by addition of four molar aqueous hydrochloric acid, and the pantothenic acid was extracted with ethyl acetate (3 x 300 mL). The solvent was removed under reduced pressure and the and campher sulfonic acid were added. The reaction mixture was stirred overnight at room temperature. The solvent was removed using a rotary evaporator and the remaining slurry was subject to flash column purification on silica gel using hexane:ethyl acetate 6:1 as the starting eluent and increasing ethyl acetate concentration to 100%. Product-containing fractions were pooled (TLC: hexane:ethyl acetate 3:2, R_f= 0.04), and the solvent removed under

reduced pressure, which yielded the protected pantothenic acid **4.3** as white solid (7.6 g, 22.4 mmol, 35%).

^{1}H-NMR ($CDCl_3$, 300 MHz): δ= 7.40 (d, *J*= 8.69 Hz, 2H); 6.90 (d, *J*= 8.80 Hz, 2H); 5.45 (s, 1H); 4.09 (s, 1H); 3.79 (s, 3H); 3.67 (q, *J*= 11.53 Hz, 2H); 3.52 (dq, *J*= 13.86, 8.22 Hz, 2H); 2.58 (t, *J* = 6.24 Hz, 2H); 1.09 (s, 3H), 1.08 (s, 3H).

^{13}C-NMR ($CDCl_3$, 100 MHz): δ= 176, 170, 160, 130, 127, 114, 101, 84, 78, 55, 34, 33.5, 33, 22, 19.

ESI-MS: $[M+H]^+$= 338.4 m/z; $[M+Na]^+$= 360.4 m/z; $[M+K]^+$= 376.4 m/z; theoretical mass: $[M+H]^+$= 338.2 m/z.

Step 2. Coupling of compound **4.3** with Boc-protected ethylenediamine.

Chemicals: *N*-(3-dimethylaminopropyl)-*N*'-ethylene carbodiimide hydrochloride (EDC), 1-hydroxy-7-azabenzotriazol (HOAt), *N*-Boc-ethylenediamine (*N*-Boc-EDA), triethylamine.

compound	MW(g/mol)	n/mmol	equivalents	amount
4.3	337.37	14.4	1.0	4.9 g
EDC	181.71	28.8	2.0	5.2 g
HOAt	136.11	28.8	2.0	3.9 g
N-Boc-EDA	160.21 (ρ= 1.01)	14.4	1.0	2.3 mL
triethylamine	101.19 (ρ= 0.73)	43.2	3.0	6.0 mL

Protected pantothenic acid S3, EDC, HOAt and ethylendiamine **4.4** were dissolved in 15 mL of dimethyl formamide (DMF). Triethylamine was added dropwise and the reaction mixture was stirred vigorously overnight at room temperature. The DMF was removed by vacuum distillation (5 mbar, 40°C) and 30 mL of ethyl acetate and 20 mL of saturated sodium hydrogencarbonate solution were added to the crude product. To isolate the product, the resulting layers were separated and the organic layer was washed with saturated sodium hydrogencarbonate (2 x 30 mL) and with saturated sodium chloride solution (2 x 30 mL) prior to solvent removal under reduced pressure. Purification by flash chromatography as described for Step 1 (TLC: hexane:ethyl acetate 3:2, $R_{f,\,S5}$= 0) afforded the fully protected aminopantetheine **4.5** as white solid after solvent removal (4.5 g, 9.4 mmol, 65%).

^{1}H-NMR ($CDCl_3$, 300 MHz): δ= 7.43 (d, *J*= 8.8 Hz, 2H); 6.91 (d, *J*= 8.7 Hz, 2H); 5.46 (s, 1H); 4.07 (s, 1H); 3.81 (s, 3H); 3.68 (q, *J*= 11.4 Hz, 2H); 3.55 (m, 1H); 3.25 (m, 2H); 2.94 (d, *J*= 12.1 Hz, 2H); 2.47 (td, *J*= 6.0 Hz, *J*= 12.4 Hz, 2H); 1.43 (s, 9H); 1.09 (s, 6H).

^{13}C-NMR (acetone-d_6, 100 MHz): δ= 169, 161, 150, 132, 129, 114, 102, 85, 84, 79, 56, 40, 37, 36, 35.5, 34, 22, 20.

ESI-MS: $[M+H]^+$= 480.3 m/z; $[M+Na]^+$= 502.4 m/z; theoretical mass: 480.3 m/z.

Step 3. Deprotection of aminopantetheine **4.5.**

Chemicals: trifluoroacetic acid (TFA), triethylsilane (TES).

compound	MW(g/mol)	n/mmol	equivalents	amount
4.5	479.57	9.38	1	4.50 g
TFA	114.02 (ρ= 1.48)	93.80	10	7.23 mL
TES	116.28 (ρ= 0.73)	46.90	5	7.47 mL

The protected compound **4.5** was dissolved in 100 mL of dichloromethane, and trifluoroacetic acid and triethylsilane were added. The reaction mixture was stirred at room temperature for four hours. The solvent and the excess reagents were removed under reduced pressure and the remaining slurry was dissolved in 70 mL 0.1% (v/v) aq. TFA solution. The water-insoluble components were removed by centrifugation (4000 rpm, 10 min). The supernatant was separated into 14 x 5 mL aliquots and each subjected to preparative HPLC purification (Macherey and Nagel, Nucleodur C-18 ec, 250 x 21 mm, particle size 5 µM, pore size 110 Å) the flow rate was 10 mL/min, solvent A was 0.1% (v/v) TFA/dd. H_2O and solvent B was 0.1% (v/v) TFA/acetonitrile) using gradient elution (0-60% B over 30 min). The fractions containing the aminopantetheine (t_R= 13.1 min) were pooled and freeze-dried. Aminopantetheine **4.6** was isolated (1.03 g, 3.94 mmol, 42%) as highly viscous oil and was stored in 30 mg aliquots at -80°C.

^{1}H-NMR (D_2O, 300 MHz): δ= 3.94 (s, 1H); 3.49-3.43 (m, 5H); 3.35 (d, *J*= 11.1 Hz, 1H); 3.10 (t, *J*= 5.9 Hz, 2H); 2.48 (t, *J*= 6.5 Hz, 2H); 0.88 (s, 3H); 0.85 (s, 3H).

^{13}C-NMR (D_2O, 100 MHz): δ= 175, 174.5, 76, 69, 39, 38.5, 37, 35.5, 21, 19.

ESI-MS: $[M+H]^+$= 262.1 m/z; $[M+Na]^+$= 284.2 m/z; theoretical mass: 262.2 m/z.

Step 4. Enzymatic synthesis of amino-coenzyme A.

Chemicals: adenosine triphosphate (ATP), magnesium chloride ($MgCl_2$), potassium chloride (KCl), tris-(hydroxymethyl)aminomethane (TRIS)

compound	MW(g/mol)	n/µmol	equivalents	amount
4.6 (20 mM)	261.43	20	1	1000 µL
PanK (161 µM)	36360	0.050	0.00250	310 µL
PPAT (457 µM)	17836	0.075	0.00375	146 µL
DPCK (277 µM)	22621	0.100	0.00500	240 µL
ATP (0.1 M, pH 7)	551.14	200	10	2000 µL
$MgCl_2$ (3M)	95.21	400	20	133 µL
KCl (3M)	74.55	400	20	133 µL
TRIS/HCl (1M, pH 9)	121.14	1000	50	1000 µL

The genes coding for PanK, PPAT, DPCK were provided by Eric R. Strieter (University of Wisconsin). All compounds were dissolved in dd. H_2O as stated in the table above and the pH was adjusted to the values shown in parenthesis, when applicable. The corresponding amounts of the stock solutions were combined in a 15 mL Falcon tube and incubated on a nutating mixer for 1.5 hours. The reaction mixture was injected onto an equilibrated Sep Pak C-18 SPE column (previously washed with 10 mL acetonitrile followed by 10 mL dd H_2O). The loaded column was washed with 5 mL dd. H_2O and the reaction product was eluted with 6 mL of H_2O:acetonitrile (1:1) solution containing 0.1% (v/v) TFA. The solution containing product was flash-frozen in liquid nitrogen and freeze-dried. The resulting solids were dissolved in 2 mL of an aq. 0.1% (v/v) TFA solution and subjected to preparative HPLC purification (Macherey and Nagel, Nucleodur C-18 ec, 250 x 21 mm, particle size 5 µM, pore size 110 Å; the flow rate was 18 mL/min, solvent A was 0.1% (v/v) dd. H_2O and solvent B was 0.1% (v/v) TFA acetonitrile) using gradient elution (0-10% B over 10 min). Freeze-drying of the product containing fractions (t_R= 15.8 min) yielded the amino-Coenzyme A as white solid (10.8 mg, 14.4 µmol, 72%), which was stored at -20°C until further usage.

^{1}H NMR (D_2O, 300 MHz): δ= 8.69 (s, 1H); 8.45 (s, 1H); 6.24 (d, *J* = 5.5Hz, 1H); 4.89-4.87 (m, 2H); 4.62 (br s, 1H); 4.29-4.27 (m, 2H); 4.04 (s, 1H); 3.84 (q, *J* = 5.0 Hz, 1H); 3.66 (q, *J* = 5.0 Hz, 1H); 3.53-3.48 (m, 4H); 3.14 (t, *J* = 6.0 Hz, 2H); 2.51 (t, *J* = 6.5 Hz, 2H); 0.95 (s, 3H); 0.89 (s, 3H).

^{13}C NMR (D_2O, 100 MHz) δ= 175.4, 175.0, 150.1, 148.7, 144.9, 142.7, 118.8, 87.7, 83.8, 74.7, 74.3, 74.2, 71.9, 65.3, 39.4, 38.6, 37.0, 35.8, 35.6, 20.8, 19.1.

HR-ESI-MS: $[M+H]^+$= 751.1617 m/z; theoretical mass: 751.1613 m/z.

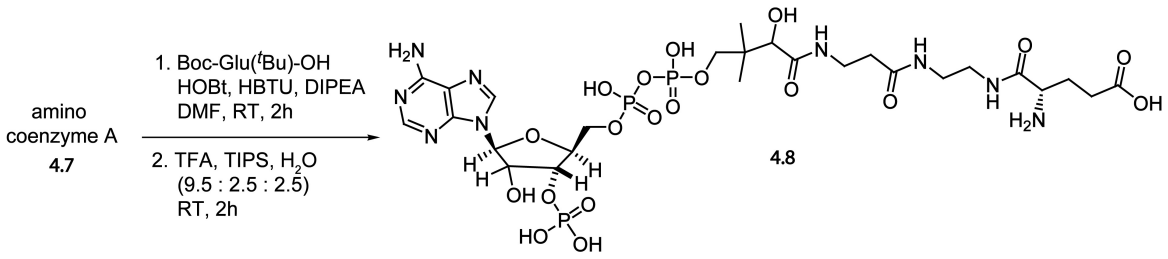
4.4.2 Synthesis of Aminoacyl-CoA

Aminoacyl-CoA thioesters were synthesized by adding 1.5 eq. of CoA, 1.5 eq. of PyBOP, and 10 eq. of DIPEA to 1 eq. of the corresponding N^{α}-Boc and side chain protected amino acid in DMF/H_2O (4:1). The reaction mixture was stirred at room temperature for 2 h, followed by *in situ* cleavage of the protection groups by addition of TFA/TIPS/H2O (9.5:0.25:0.25). The deprotected aminoacyl-CoA compounds were precipitated in ice-cold diethyl ether and subject to preparative HPLC purification (Macherey and Nagel, Nucleodur C-18 ec, 250 x 21 mm, particle size 5 µM, pore size 110 Å, the flow rate was 18 mL/min, solvent A was 0.1% (v/v) dd. H_2O and solvent B was 0.1% (v/v) TFA acetonitrile) using gradient elution (5-30% B over 30 min). Fractions containing the desired CoA-*S*-aa were identified by MALDI-MS, pooled, and freeze-dried. The identities of the synthesized CoA-*S*-aa were verified by high resolution MS. The retention times, observed and calculated masses as well as the yields are summarized in the following table:

Compound	Retention Time [min]	Observed Mass $[M+H]^+$	Calculated Mass $[M+H]^+$	Yield
CoA-*S*-Asn	13.2	882.1647	882.1654	29%
CoA-S-D-Asn	12.8	882.1666	882.1654	22%
CoA-*S*-Asp	12.4	883.1510	883.1494	30%
CoA-*S*-D-Asp	12.0	883.1479	883.1494	24%
CoA-*S*-Gln	16.7	896.1809	896.1816	28%
CoA-*S*-Glu	14.1	897.1664	897.1651	25%
CoA-*S*-D-Glu	14.0	897.1659	897.1651	22%
CoA-*S*-Leu	18.2	881.2071	881.2065	46%
CoA-*S*-Phe	22.2	915.1911	915.1909	51%

Table 4.2. Summary of synthesized aminoacyl-CoA compounds.

4.4.3 Synthesis of CoA-H*N*-Glu



compound	µmol	eq	amount
amino coenzyme A **4.7** (750.15 g/mol)	4.0	1.0	3.0 mg
Boc-Glu(O^{t}Bu)-OH (303.4 g/mol)	6.0	1.5	1.8 mg
HBTU (379.30 g/mol)	12.0	3.0	4.6 mg
HOBt (with 20% water, anhyd. MW 135.13 g/mol)	12.0	3.0	2.0 mg
DIPEA (129.25 g/mol, ρ= 0.76 g/cm^3)	40.0	10.0	6.8 µL

The N^{α}-protected and side chain protected glutamic acid, HBTU and HOBt were dissolved in 300 µL of DMF. DIPEA was added and the mixture was heavily agitated to dissolve all solids, and amino coenzyme A was then added. The mixture was incubated for 1.5 hours at 25°C in an Eppendorf Thermomixer at 700 rpm to allow amide bond formation, followed by *in situ* deprotection by addition of 1.5 mL of the deprotection mixture (TFA/H_2O/TIPS; 95%:2.5%:2.5%). The deprotection reaction was agitated for 2 hours at 25°C (Thermomixer, 700 rpm). The mixture was then transferred into a 50 mL Falcon tube containing 30 mL of cold (-20°C) diethyl ether and incubated at -20°C for 2 hours. A precipitate formed and was pelletized by centrifugation (4°C, 4,000 rpm, 20 min). The pellet was dissolved in 5% aqueous acetonitrile solution and subject to preparative HPLC purification (Macherey and Nagel, Nucleodur C-18 ec, 250 x 21 mm, particle size 5 µM, pore size 110 Å, the flow rate was 18 mL/min, solvent A was 0.1% (v/v) dd. H_2O and solvent B was 0.1% (v/v) TFA acetonitrile) using gradient elution (5-30% B over 30 min). Fractions containing the CoA-H*N*-Glu were identified by MALDI-MS (t_R= 13.7 min), combined, freeze-dried and CoA-H*N*-Glu was yielded as white solid (2.64 mg, 3.0 µmol, 75%).

HR-ESI-MS CoA-H*N*-Glu: $[M+H]^+$= 880.2045 m/z; theoretical mass: 880.2039 m/z.

HR-ESI-MS CoA-H*N*-D-Glu: $[M+H]^+$= 880.2051 m/z; theoretical mass: 880.2039 m/z.

4.5 Biochemical Assays

4.5.1 Analysis of Iron Cofactor Content

The iron cofactor content of the recombinant hydroxylases was determined by using the chromogen Ferene S as previously published.[212] To the corresponding hydroxylase (10 nmol in 800 µL dd. H_20) concentrated hydrochloric acid (100 µL) was added and incubated in an Eppendorf Thermomixer (10 min, 400 rpm, 25°C). Proteins were precipitated by addition of 80% trichloroacetic acid followed by centrifugation (13,000 rpm, 10 min, RT). The supernatant was transferred into an acid washed clean glass tube and 200 µL of a 45% (w/v) sodium acetate solution was added. 1.8 mL of the Ferene S solution (0.75 mM Ferene S, 10 mM L-ascorbic acid, 45% (w/v) sodium acetate) was admixed and 1 mL of the mixture was measured photometrically at λ= 593 nm. The obtained absorbance values were compared with the recorded calibration curve, which was recorded analogously with distinct iron concentration. $FeCl_3$ was used as iron source and the amount was varied between 4 and 32 nmol.

4.5.2 *In Vitro* 4'-Phosphopantetheinylation of PCP

The recombinant PCPs (200 µM) were artificially loaded with either coenzyme A (CoA), amino-coenzyme A (NH_2-CoA) or aminoacyl-CoAs by using the promiscuous *B. subtilis* phosphopantetheinyl (ppan) transferase Sfp (2 µM).[52,54] The loading reaction was incubated at 28°C for 15 min in buffered aqueous solution (pH 7, 25 mM HEPES, 50 mM NaCl, 1 mM $MgCl_2$). To ensure that complete loading of the PCP was achieved, the loading assays were stopped by addition of formic acid (final concentration of 4% (v/v)) and the reaction mixture was analyzed directly by HPLC-QTOF-MS.

4.5.3 Hydroxylation Assays

4.5.3.1 Free Amino Acid Substrates

In a typical hydroxylation assay with free amino acids, the recombinant hydroxylase (5 µM) was incubated for different time spans with the free amino acid (2 mM), the ferrous iron cofactor (0.5 mM $(NH_4)_2Fe(SO_4)_2$) and co-substrate αKG (2 mM) in 100 µL 50 mM HEPES buffer (pH 7.5). Controls were carried out either in the absence the recombinant hydroxylase or without co-substrate αKG. The reaction was either stopped by adding nonafluoropentanoic acid (4% (v/v) final concentration) or by adding FMOC-

chloride solution (see 4.5.5.1). The reactions were either analyzed for possible hydroxylation by reversed-phase HPLC-MS analysis on a Hypercarb column (Thermo Electron Corporation, pore diameter of 250 Å, particle size of 5 µm, 100% carbon) or by fluorescence detection of the FMOC derivatized amino acid (see 4.5.5.1). Mobile phases were: A, 20 mM aqueous nonafluoropentanoic acid; B, acetonitrile. The following gradient was applied: 0-10 % B in 12 min, with a flow rate of 0.2 mL min^{-1} at 17.5°C.

Kinetics were carried out accordingly, the substrate concentration was varied between 50 µM and 2 mM. The reaction was stopped in the linear range (after 5 and 30 min) and analyzed as described above. Integration of the mass or fluorescence signals, respectively, of the starting material and product gave relative conversion values from which starting velocities were calculated.

4.5.3.2 PCP-Bound Amino Acid Substrates

The PCP was loaded with the corresponding CoA-*S*-amino acid or CoA-H*N*-amino acid as described in 4.5.2. Then, the loaded PCP domain, the hydroxylase, αKG and iron(II) were added and the assay was carried out as described in 4.5.3.1. The reactions were stopped by addition of formic acid (4% (v/v) final concentration) and analyzed by reversed phase HPLC-MS using a QTOF-MS QStar Pulsar i (Applied Biosystems) coupled to a HP 1100 HPLC (Agilent) equipped with a C-4 Nucleosil guard column (Macherey and Nagel, 10 x 3 mm, pore diameter of 300 Å, particle size 5 µm) with the following conditions: solvent A (water/0.45% formic acid), solvent B (acetonitrile/0.45% formic acid), flow rate 0.2 mL/min, temperature 45 °C with a gradient of 10-95% solvent B in 10 min, the gradient was then held for 7 min.

For kinetic measurements with A*PCP_3-H*N*-Glu and non-cognate CDA-T9-H*N*-Glu, the hydroxylation assays were set-up as described above and stopped after 30 min, which was determined to be the optimal reaction time for being in the linear conversion range. Substrate concentrations were varied between 10 µM and 450 µM. Kinetic parameters were determined using the calculated starting velocity and the Enzyme Kinetic Module for Sigma Plot 8.0 (SPSS).

4.5.3.3 Coupled Amino Acid Transfer and Hydroxylation Assays

The A domain KtzN (2 µM) was incubated with ATP (1 mM), L-glutamic acid (5 mM), DTT (0.5 mM), $MgCl_2$ (1 mM) holo-A*PCP_3 (100 µM) or holo PCP_3 (200 µM). Holo-PCPs were

generated by pre-incubation of the PCPs with CoA (1 mM) and Sfp (2 μM) at 28°C for 15 min (complete conversion validated by QTOF-MS). The coupled assays were also conducted in the presence of KtzO or KtzP (5 μM), αKG (2 mM) and ferrous iron (0.5 mM). All reactions were incubated in buffered aqueous solution (25 mM HEPES, 50 mM NaCl, pH 7), stopped after two hours by addition of formic acid (final concentration of 4% v/v) and then analyzed by HPLC-QTOF-MS as described in 4.5.3.2.

4.5.4 Cleavage of PCP-Bound Amino Acids

The hydroxylation assay was carried out as described in 4.5.3.2 except that the final reaction volume was 200 μL. The proteins were precipitated by addition of 1 mL 10% TCA, and the resulting mixture centrifuged (45 min, 13,000 rpm, 4°C). The supernatant was discarded and the pellet was washed with 200 μL diethyl ether/absolute ethanol (3:1) twice and with 200 μL diethyl ether once. The washed pellet was dried for 5 min at 37°C and 100 μL of 0.1 M aqueous potassium hydroxide solution was added to cleave the thioester bond and release the hydroxylated amino acid from the PCP domain. The pellet was resuspended in the KOH(aq) solution and the resulting mixture was incubated at 70°C for 20 min at 1400 rpm in a Thermomixer (Eppendorf). After thioester cleavage, methanol (1 mL) was added, and the mixture was incubated over night at -20°C. Precipitated protein debris was separated from the supernatant by centrifugation (13,000 rpm, 30 min, 4°C). The supernatant was transferred into a new tube and the solvents were removed under reduced pressure by using a SpeedVac-manifold (13,000 rpm, 2 mbar, 30°C). The pellets were stored at -20°C or derivatized immediately for HPLC analysis (see below).

4.5.5 Amino Acid Derivatization

As most amino acids neither show specific UV absorption nor fluorescence, except for these which are aromatic, they had to be derivatized with a hydrophobic reagent to allow RP-HPLC and UV or fluorescence detection.

4.5.5.1 Derivatization with FMOC-Cl

Based on a previously published protocol,[213] FMOC-Cl derivatization was carried out by adding 200 μL of a 3 mM FMOC-Cl solution in acetone to 50 μL of the hydroxylation assay (4.5.3.1) or to amino acids solutions. Prior to FMOC-Cl addition, the assay/solution

was diluted to 250 µL using a 50 mM HEPES buffer (pH 7.5). The reaction mixture was incubated for 2 min in an Eppendorf Thermomixer (750 rpm, 25°C). The reaction was stopped by pipetting 50 µL of a heptylamine solution (3 mL heptylamine in 15 mL acetonitrile, pH adjusted to 7.5 with 175 mL 0.1 M HCl) to the derivatization mixture. The heptylamine quenched excess FMOC-Cl and the resulting FMOC-heptylamine does not interfere with HPLC-fluorescence analysis of FMOC-amino acids due to its strong hydrophobicity. The FMOC-N^{α}-derivatized amino acids were analyzed by HPLC separation and fluorescence detection using an Agilent 1100 series HPLC equipped with a Synergi Fusion-RP column (Phenomenex, C18 polar ec, pore diameter of 80 Å, particle size of 3 µm) with the following gradient: 0-20% eluent B (100% acetonitrile) and eluent A (78% 0.1 M aqueous sodium acetate solution, 22% acetonitrile) in 17 min at 0.3 mL/min and 17.5°C.

4.5.5.2 Derivatization with Dabsyl-Cl

The pellets of from the cleavage reactions (4.5.4) or the corresponding amino acid were dissolved in buffered aqueous solution (100 µL, 50 mM TRIS, pH 9) and a saturated solution of dabsyl chloride in acetone (100 µL) was added.[205,214] The mixture was heated at 70°C for 15 min at 1400 rpm in a Thermomixer and subsequently diluted with absolute ethanol (1 mL). The resulting mixture was centrifuged at 8,000 rpm for 5 min, and 100 µL of the supernatant were analyzed by reversed-phase HPLC-QTOF-MS. Gradient elution was carried out using a Synergi Fusion-RP column (Phenomenex, polar embedded C-18 column, 250 x 3 mm, pore diameter of 80 Å, particle size 4 µm) and the following conditions: solvent A was water/0.45% formic acid; solvent B was acetonitrile/0.45% formic acid. The flow rate was 0.3 mL/min with a gradient from 50% B to 75% B over 20 min and then increasing to 95% B over the next min. The gradient was then held for further 5 min.

4.5.6 ATP/PP_i Exchange Assay

ATP-PP_i exchange was assayed in 500 µL of reaction buffer (50 mM TRIS, 40 mM KCl, 10 mM $MgCl_2$, pH 7.5) containing 0.1 mM ATP, 0.5mM tetrasodium PPi (0.72 µCi of [^{32}P]PP_i), 0.1 mM DTT, and 10 mM substrate. Reactions were initiated by adding purified protein to a final concentration of 2 µM. After incubating at 37°C for 10 min, reactions were quenched by the addition of charcoal suspension (0.1 M tetrasodium PP_i, 0.35 M

$HClO_4$, 16 g/L charcoal). Free [^{32}P]PP_i was removed by centrifugation of the sample followed by washing twice with wash solution (0.1 M tetrasodium PP_i, 0.35 M $HClO_4$). Controls were carried out either without enzyme or without ATP or PP_i. Charcoal-bound radioactivity was measured on a Beckman LS 6500 scintillation counter.

4.5.7 Stereospecificity Determination and Product Isolation by Chiral HPLC

In order to determine the exact stereochemistry of AsnO D241N assay products and to synthesize L-*threo*-hydroxyaspartic acid (L-THA) chiral HPLC was used. 10 µm AsnO D241N was incubated with L-aspartic acid (18.9 mg, 142 µmol), the co-substrate αKG (26 mg, 18 mmol) and ammonium-iron(II) sulfate hexahydrate (0.39 mg, 100 µmol) as source of the ferrous iron cofactor. All solids were dissolved in 9 mL of 25 mM HEPES buffer, pH 7.0. The reaction was started by addition of 1 mL enzyme solution (100 µM, same buffer) and was allowed to stir for 14 h at 16°C. L-THA was isolated by HPLC separation on a Chirex 3126 column (Phenomenex, D-penicillamine, 150 x 4.6 mm) with an isocratic mobile phase of 1 mm $CuSO_4$ in 95% H_2O and 5% isopropyl alcohol on a Beckman Coulter Gold System with UV detector 168 (Fullerton, USA). The reaction product was compared with *threo* and *erythro*-hydroxyglutamic acid standards and L-THA containing fractions were identified. The fractions were pooled and loaded twice on a gravity column containing NTA-Agarose resin to remove the copper ions. The through-flow was freeze-dried to yield enantiomerically pure L-THA (14.4 mg, 97 µmol, 68%).

4.6 Crystallography and Structure Elucidation

4.6.1 Crystallization, Soaking and Data Collection

Crystals of AsnO from *S. coelicolor* were grown at 18°C by the sitting-drop vapor-diffusion method. Native crystals were obtained in 2.5 M sodium acetate, 0.1 M HEPES, pH 4.6 and in 4 M sodium formate at a protein concentration of 8.3 mg/mL in 25 mM HEPES, 50 mM NaCl, pH 7.0. The product-complex was achieved by soaking AsnO crystals with degassed 10 mM $(NH_4)_2Fe(SO_4)_2$, 20 mM α-ketoglutarate and 50 mM L-Asn-solution for 60 minutes under protection gas (97% nitrogen, 3% hydrogen) in an anaerobic chamber. Soaking of the AsnO crystals with αKG and Fe^{2+}, under the conditions described above, led to an AsnO·Fe^{2+} complex. Cryoprotection was achieved by washing the crystals in the mother liquor that contained 30% (v/v) glycerol. Intensity

data from flash-frozen crystals were collected at beamline X13 at EMBL, Hamburg, Germany, at 1.45 Å resolution for the apo-form of AsnO and 1.66 Å for the AsnO·Fe(II)·succinate·hAsn complex. A 1.92 Å dataset for the AsnO·Fe(II) complex was recorded on an inhouse CuK_α rotating anode. Diffraction data were integrated and scaled with the programs MOSFLM and SCALA of the CCP4 program suite.[215]

4.6.2 Structure Solution and Refinement

The crystal structure of AsnO was solved by molecular replacement using the program MOLREP[216] and chain A of the clavaminate synthase (PDB ID code 1DRY) as a search model. Structure refinement was carried out using REFMAC5[217] and COOT.[218] The final model of AsnO contains residues S10-P330, but misses a disordered region spanning L218-E223. Anomalous difference electron density maps indicated that no iron atom was bound in the active site. Instead, the octahedrally coordinated ion observed at the iron binding site was assigned as a sodium ion originating from the crystallization buffer. Anomalous difference electron density maps indicated for Fe^{2+}/αKG soaked AsnO crystals the binding of the iron atom to the active site. For data statistics refer to Table 5.2. PyMol was used to prepare all figures displaying structural information also in the introduction and discussion of non-AsnO-related enzymes.[219]

5. Results

5.1 β-Hydroxylated Acidic Lipopeptide Antibiotics Precursors

Acidic lipopeptide antibiotics are a new class of potent antibiotics. Members are amongst others the clinically approved daptomycin, the calcium-dependent antibiotics (CDA) and A54145. They display activity against a wide range of Gram-positive bacteria including multi-drug resistant bacteria, such as methicillin-resistant *Staphylococcus aureus* and vancomycin-resistant *Enterococci*. The potency of acidic lipopeptides is an intrinsic feature of their chemical structure. The non-ribosomally synthesized peptide cores consist of eleven to thirteen amino acids, which are rigidified by the formation of a ten-membered ring. Within these antibiotics a remarkable structural diversity is generated by the incorporation of modified non-proteinogenic amino acids.

In this study, the origin of one of these non-proteinogenic amino acids, namely β-hydroxyasparagine (hAsn), found in the peptide backbone of CDA and A54145 was elucidated. In the case of the calcium-dependent antibiotics, the hydroxylase AsnO, responsible for hAsn formation, was characterized biochemically and crystallized.[‡] The thorough investigation of AsnO's biochemical properties led to the elucidation of a precursor pathway, in which hAsn is generated. Furthermore, the 3D structure determination of AsnO allowed mechanism prediction for other non-heme iron hydroxylases involved antibiotic biosynthesis by utilizing bioinformatic tools.

To prove the predictions based on AsnO's structure for the reactivity of other hydroxylases, it was shown that the non-heme iron and αKG dependent hydroxylase LptL also catalyzes the formation of hAsn in a precursor pathway during A54145 biosynthesis. Biochemical characterization of LptL was conducted including kinetic parameter and substrate specificity determination. Together, the elucidation of these precursor pathways might facilitate the generation of novel building blocks for NRPS, potentially resulting in natural products with improved or altered bioactivity.

[‡] The work described in chapter 5.1.3, in particular the biochemical characterization of the enzyme AsnO was to some extent published previously in the diploma thesis: M. Strieker. Ungewöhnliche Hydroxylierung und Phosphorylierung eines aliphatischen Kohlenstoffatoms während der Biosynthese von D-3-Phosphohydroxyasparagin. Marburg **2006**. Some parts presented in the following paragraphs might therefore be redundant to that diploma thesis. This was necessary for clarity reasons and to provide complete characterization of the studied enzyme. The x-ray crystal structure elucidation was conducted by Prof. Dr. Lars-Oliver Essen (Philipps-University Marburg).

5.1.1 Recombinant Expression of *asnO, cda-pcp$_9$*, *lptL* and *lpt-pcp$_3$*

AsnO and *cda-pcp$_9$* were cloned from *S. coelicolor lptL* and *lpt-pcp$_3$* from *S. fradiae*, respectively and expressed in *E. coli* BL21 (DE3) as described in the Methods section. The recombinant proteins were purified by Ni-NTA affinity chromatography and concentrated proteins were analyzed by SDS-PAGE and visualized using Coomassie dye (Figure 5.1). The final protein yields per liter culture after purification and concentration were 7 mg for AsnO, 50 mg for CDA-PCP$_9$, 12 mg for LptL and 8 mg for Lpt-PCP$_3$.

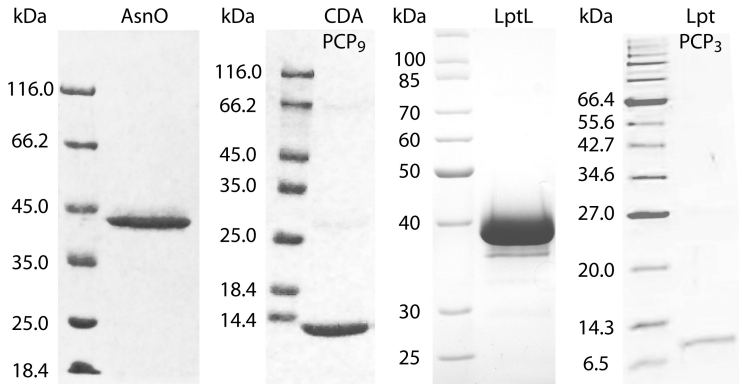


Figure 5.1: SDS-PAGE analysis of purified recombinant AsnO (38.9 kDa), CDA-PCP$_9$ (12.1 kDa), LptL (37.1 kDa) and Lpt-PCP$_3$ (10.6 kDa). Protein markers were from left to right Fermentas Molecular Weight Marker, Fermentas PageRuler and NEB Broad Range Protein Marker.

5.1.2 Iron Content of AsnO and LptL

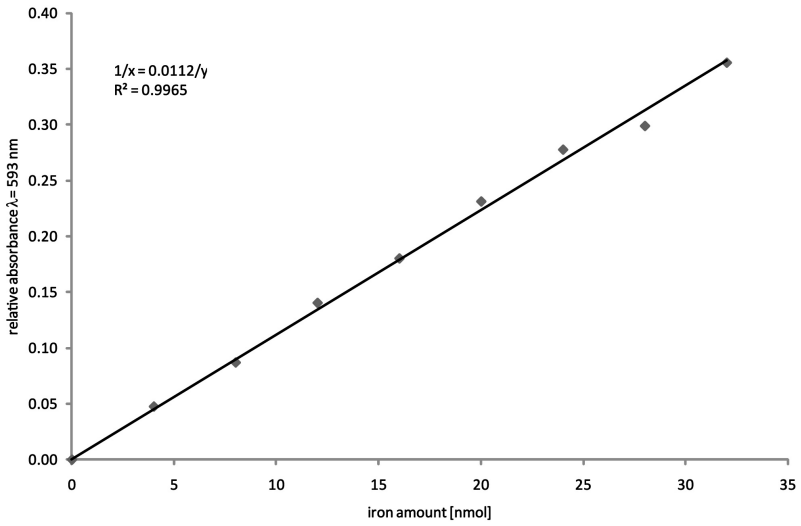


Figure 5.2: Calibration curve for the Ferene S iron content determination. Iron amounts were varied between 4 and 32 nmol and measured versus Ferene S solution without added iron.

The iron content of the non-heme iron(II) and αKG dependent dioxygenases AsnO and LptL was determined by using the Ferene S [3-(2-pyridyl)-5,6-bis(2-(5-furyl sulfonic

acid))-1,2,4,-triazine] dye.[212] The determination is based on the tight complex formation between iron ions and the iron chelating Ferene S dye, which has an absorption maximum at 593 nm. A standard curve for calibration of the Ferene S-iron complex was obtained by using stock iron solutions and measuring the absorption at 593 nm versus a blank run (Figure 5.2). For AsnO the A_{593nm}-value was 0.011 and for LptL 0.008. From this value the iron content was calculated using the equation of the linear regression of the calibration curve. This led to a calculated iron content of 0.98 nmol for AsnO and 0.71 nmol for LptL, as 10 nmol of each enzyme were used, these values correspond to 9.8% and 7.1% iron content. The low iron content is either due to oxidation or the iron was exchanged with other divalent ions during the purification steps. Therefore in all subsequent assays ferrous iron in form of ammonium ferrous iron sulfate was added to grant a constant supply of ferrous iron cofactor.

5.1.3 Characterization and Structure Elucidation of AsnO

5.1.3.1 Hydroxylation Activity of AsnO

Asparagine Hydroxylation. To evaluate the biological activity of AsnO, the recombinant enzyme was incubated with the amino acids L-Asn and D-Asn. To facilitate identification and detection of the modified amino acid, L-Asn was derivatized with 9-fluorenylmethyl chloroformate (FMOC-Cl) to yield highly fluorescent FMOC derivatives. The hydroxylation and derivatization was carried out as described in Methods section. In the presence of the ferrous iron cofactor and the co-substrate αKG, hydroxylation activity was observed with the L-enantiomer of asparagine. The reaction was monitored by fluorescence detection by exciting the FMOC residue at 265 nm and measure the emission at 315 nm. It was found that the incubation of L-asparagine with AsnO (Figure 5.3, red trace) and subsequent derivatization with FMOC-Cl gave FMOC-hAsn. Under assay conditions the conversion was nearly quantitative after 30 min. In the control reaction (Figure 5.3, blue trace), without enzyme added, only substrate-derived FMOC-Asn was detected. In addition to HPLC fluorescence detection, MS analysis was carried out to confirm the identity of the starting material Asn and the product hAsn. For peak 1 from the HPLC-separated enzymatic reaction workup the mass of 371.2 m/z (theoretical mass: 371.1 m/z) was observed, which corresponds to L-hAsn-derived fluorenylmethyl.

Peak 2 on the other hand showed a mass of 355.2 m/z (theoretical mass: 355.1 m/z) L-Asn FMOC-adduct.

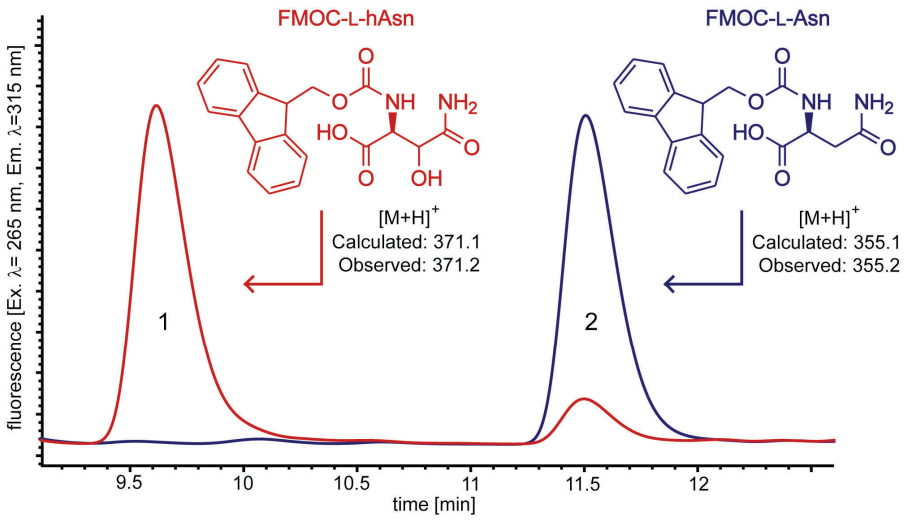


Figure 5.3: β-carbon hydroxylation of L-Asn by AsnO. HPLC traces and mass analyses of FMOC-derivatized hAsn (peak 1) and L-Asn (peak 2). The blue and red traces represent the fluorescence signals before and after L-Asn incubation with AsnO under assay conditions.

Kinetic Parameter Determination. For the determination of the kinetic parameters of Asn hydroxylation by AsnO, the inability of the enzyme to hydroxylate FMOC-Asn and the fact that the FMOC-derivatization is completed within a few seconds was exploited.

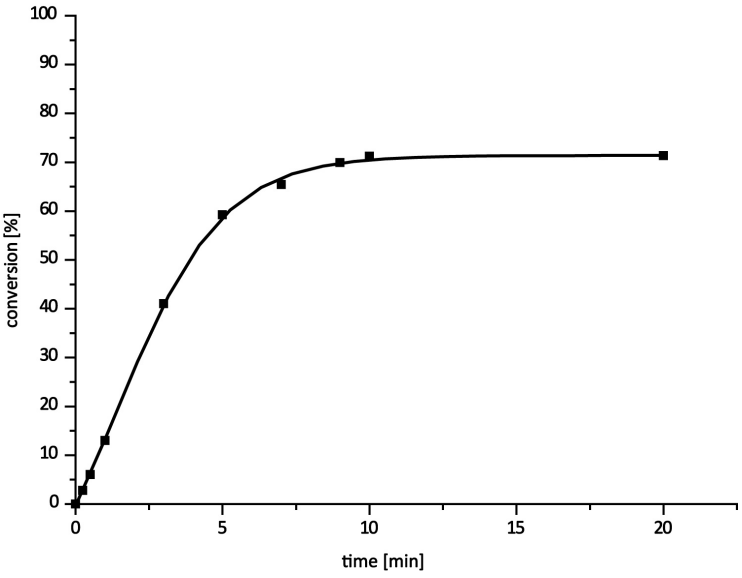


Figure 5.4: Time dependent conversion of Asn to hAsn.

To determine the linear range of time dependent conversion assays with 0.5 µM AsnO and 250 µM asparagine were stopped at different time points (Figure 5.4). The time dependent conversion revealed that the hydroxylation is in the linear range between zero and approximately four minutes. Thus, the time span between one and three minutes was used to determine the starting velocities for different substrate concentrations. Therefore, the enzyme assays were stopped by addition of

FMOC-Cl and the kinetic parameters determined using Michaelis-Menten and Lineweaver-Burk equations (Figure 5.5).

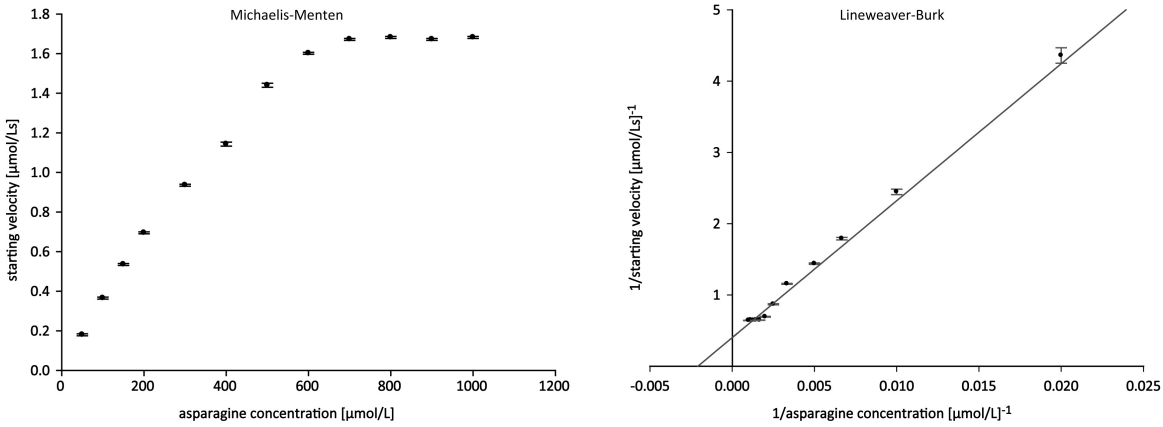


Figure 5.5: Michaelis-Menten and Lineweaver-Burk plots for kinetic parameter determination of AsnO.

The kinetics for AsnO mediated L-Asn hydroxylation were calculated to an apparent K_M of 0.479 ± 0.067 mM and k_{cat}= 298.8 ± 19.2 min^{-1}, leading to a catalytic efficiency of k_{cat}/K_M= 0.62 ± 0.15 $min^{-1}µM^{-1}$.

5.1.3.2 Substrate Specificity

For the evaluation of the substrate specificity of AsnO, the hydroxylase was incubated with a broad range of different substrates including the cognate PCP domain of module nine (CDA-PCP_9) loaded with L- and D-Asn. The latter was generated by incubation of the PCP with synthesized coenzyme A asparagine thioesters (CoA-*S*-Asn) and the promiscuous *B. subtilis* ppan transferase Sfp (see Methods for details). The PCP-tethered asparagine was tested in order to validate that AsnO is involved in precursor synthesis exclusively. Amino acids representing the different class of amino acids as well as the D-isomer of Asn were tested for AsnO-mediated hydroxylation. Furthermore, the FMOC-Asn and CoA-S-Asn as PCP mimic were tested as putative substrates. An overview of substrates tested for AsnO-mediated hydroxylation is given in Table 5.1. Neither the L- nor the D-isomer of PCP tethered asparagine was hydroxylated by AsnO. Furthermore, the FMOC-Asn as well as the CoA-*S*-Asn were not substrates for enzyme-mediated hydroxylation neither were the other tested amino acid, thus AsnO exclusively catalyses the oxygenation of the β-position of L-asparagine.

Tested Substrate	Mass [Da]	Calculated Mass Hydroxylation Product [Da]	Observed Mass [Da]	Hydroxylation
CDA-PCP_9-*S*-L-Asn	12600	12616	12600	no
CDA-PCP_9-*S*-D-Asn	12600	12616	12600	no
aspartic acid	134.0	150.0	133.0	no
glutamine	147.1	163.1	147.1	no
glutamic acid	148.1	164.1	148.1	no
valine	118.1	134.1	118.1	no
isoleucine	132.2	148.2	132.2	no
lysine	147.2	163.2	147.2	no
tyrosine	182.2	198.2	182.2	no
FMOC-Asn	355.1	371.1	355.1	no
CoA-*S*-Asn	881.2	897.2	881.2	no
D-asparagine	133.1	149.1	131.1	no
L-asparagine	133.1	149.1	149.1	yes

Table 5.1: Substrate specificity evaluation of AsnO. Different amino acids including the D-enantiomer of Asn and PCP bound Asn isomers were tested for hydroxylation by AsnO. All masses correspond to the $[M+H]^+$ as detected in positive mode MS, except for the PCP masses which are neutral mass reconstructions.

5.1.3.3 Crystal Structure of AsnO

Overall Structure Description. Trigonal crystals of AsnO were grown within two days and found to comprise one AsnO molecule per asymmetric unit (see Methods). The crystal structure[§] of AsnO was solved at 1.45 Å resolution by molecular replacement, using the related structure of CAS[220] as a search model. Like other members of the CSL-superfamily, the refined structure of AsnO contains a core of nine β-strands (A-I), eight of which are folded into a jelly roll topology, building a major and a minor sheet with five (B, G, D, I, C) and three β-strands (F, E, H), respectively (Figure 5.6 A). The jelly roll fold is sandwiched between two, largely α-helical subdomains. Two flexible loop regions border the active site (residue V135-P165 and A205-T249), and near the *C*-terminus a 3_{10}-helix (residues L312-M318) is found. In the active site region of the native AsnO, the

[§] As mentioned in the introduction to this paragraph, the initial structure elucidation was carried out prior to this study. Refinements and additional structure elucidation ($AsnO{\cdot}Fe^{2+}$ complex) as well as more profound interpretations and analyses of the structures were conducted in this study in collaboration with Prof. Dr. Essen (Philipps-University Marburg). As the crystal structures give crucial insights into this hydroxylase family and as they are the basis of most of the other research presented in this study, a complete overview of the x-ray structure elucidation is given in this paragraph, although some parts might be redundant to M. Strieker, Ungewöhnliche Hydroxylierung und Phosphorylierung eines aliphatischen Kohlenstoffatoms während der Biosynthese von D-3-Phosphohydroxyasparagin. Marburg **2006**.

electron density reveals an ion that is octahedrally coordinated to the residues H155, E157, and H287, in addition to three water molecules.

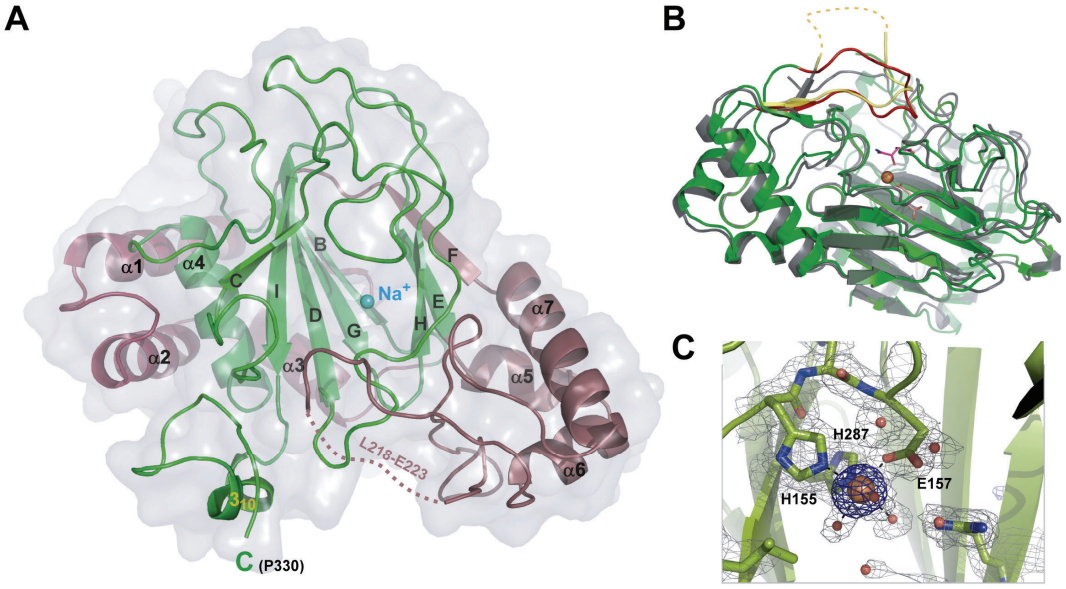


Figure 5.6: Crystal structure of AsnO. A) Overall structure of the apo-form of AsnO. B) Comparison of the structure of the AsnO·Fe(II) complex (green) with clavaminate synthase (PDB accession code 1DRY) (grey). A dashed line indicates the disordered region (residues G208-A214) of CAS. (c) SIGMAA-weighted $2F_{obs}$-F_{calc} electron density showing the binding of an iron ion to the active site of AsnO (contouring level 1σ≡ 0.11 $e^-/Å^3$). Anomalous difference electron density for the iron atom is shown in blue (contouring level 5 σ≡ 0.06 $e^-/Å^3$, λ_{Xray}=1.54179 Å).

Due to a lack of sufficient electron density and anomalous signal, this ion was assigned as a sodium ion rather than an Fe^{2+} ion. In further studies the crystal structure of AsnO with the Fe^{2+}-cofactor bound was solved at 1.92 Å resolution by soaking AsnO crystals under anaerobic conditions with Fe^{2+} and αKG. This AsnO·Fe^{2+} complex is almost identical to the apo-form, except for the central ion. No additional electron density for αKG was observed. Comparison of CAS^{220} and AsnO (Figure 5.6 B) shows the high structural similarity of these two enzymes with an overall r.m.s.d. of 1.3 Å for 257 C_α-positions. The residues H155, E157 and H287 form the conserved HXD/E…H iron binding motif (Figure 5.6 C) that is located at the mouth of the minor β-sheet. The same metal coordinating triad is conserved in almost all characterized non-heme iron dependent oxygenases, including those with folds different to αKG-dependent oxygenases.

Active and Substrate Binding Site of AsnO. Crystals soaked with the substrate L-Asn and Fe^{2+}-cofactor/αKG-co-substrate led to the structure of a product complex comprising iron either in the Fe^{2+} or Fe^{3+}-state, (2*S*,3*S*)-3-hydroxyasparagine and succinate, instead

of the expected starting material complex. After removing the crystals from the anaerobic chamber, crystals were exposed to atmospheric oxygen for approximately 30 seconds prior to flash freezing, which caused rapid formation of the product complex. The hAsn is intimately bound within the active site, and the C3-hydroxy group can be clearly assigned by electron density at 1.66 Å resolution, to correspond to the pro*S*-hydrogen of the L-Asn substrate.

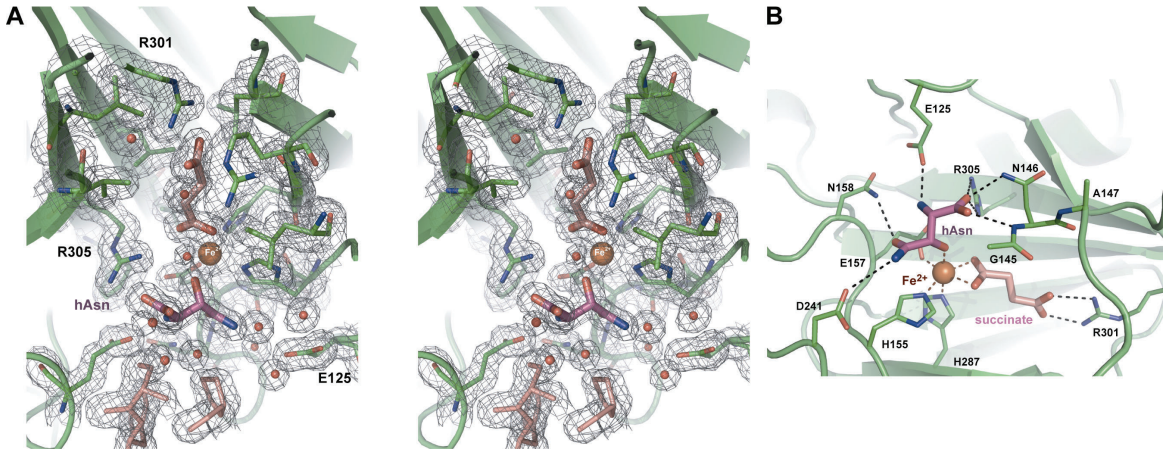


Figure 5.7: Active site in the AsnO·Fe(II)·succinate·hAsn complex. A) A Stereo diagram shows the $2F_{obs}$-F_{calc} electron density for the active site (contouring level 1.0 $\sigma \equiv 0.24$ $e^-/Å^3$). The coordination of the reaction products (2*S*,3*S*)-3-hydroxyasparagine and succinate to the catalytic iron atom allows unambiguous assignment of the stereochemistry of the hydroxylated asparagine. B) Product-enzyme interactions within the active site. The succinate molecule forms a salt bridge to the conserved arginine R301. The reaction product (2*S*,3*S*)-3-hydroxyasparagine is coordinated to the active site iron via its β-hydroxy group and forms seven additional interactions with the protein scaffold.

Its coordination to the catalytic iron (Figure 5.7 A) suggests that the starting material asparagine is similarly bound to the active site with its pro*S*-hydrogen located next to the reactive oxygen of the ferryl species. Thus, this hydrogen atom is abstracted during hydroxylation with retention of configuration[221] to give 2*S*,3*S*-hAsn. Furthermore, hAsn and, as anticipated, the L-Asn substrate is bound by seven amino acids of the peptide backbone (Figure 5.7 B). The α-carboxy group of hAsn forms a salt bridge with R305 and hydrogen bonds to N146 and the peptide group preceding this residue. The α-amino group forms another salt bridge to E125 and is also bound by one oxygen atom of the carboxyl group of the iron binding E157 side chain. The carboxamide group of the sidechain of hAsn participates in hydrogen bonds with N158, D241 and via a water molecule with the peptide group of Q144 (Figure 5.7 B).

Induced Fit Mechanism. The active site of AsnO is canopied by a flexible lid region (residues F208-E223) that is partly disordered in the apo-form (Figure 5.6 A, residues

L218-E223), but becomes ordered upon complexation of iron in the AsnO·Fe^{2+} complex (Figure 5.6 B, green), albeit with high overall B-factors.

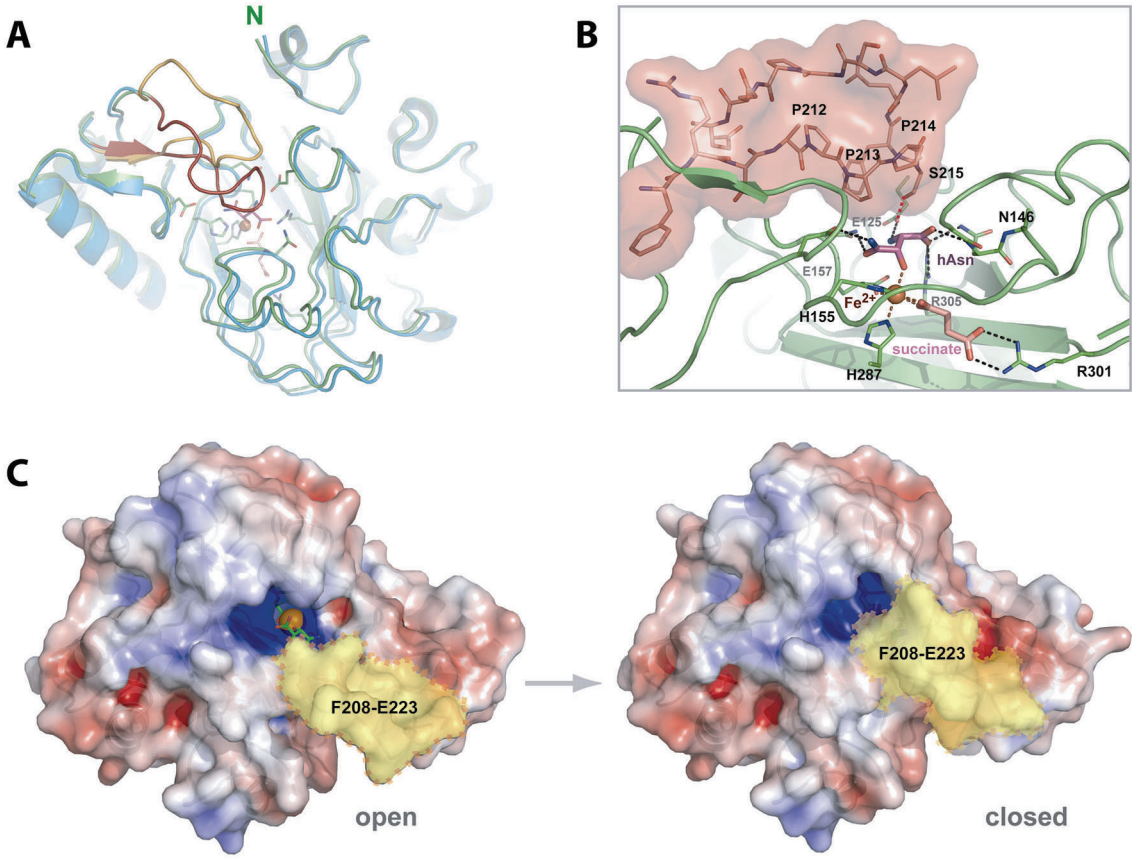


Figure 5.8: Induced fit upon substrate binding. A) Structural comparison between the product-bound AsnO·Fe(II)·succinate·hAsn complex (green) and the AsnO·Fe(II) complex (cyan). The lid region (F208-E223) is marked in red and orange. B) Sealing of the active site by the lid region. The residue E125 that binds to the α-amino group of the substrate L-Asn is favorably positioned by forming a hydrogen bond with the side chain of S215, projecting from the lid region. C) Electrostatic representation of the open and closed state of AsnO as calculated by APBS (blue: + 5 kT/e, red: -5kT/e; ion concentration 0.1 M). In the open state, the lid region (accentuated in yellow) keeps the active site accessible for the binding of substrate and co-substrates. Concomitant with substrate binding, the lid mostly occludes the active site (right side) from bulk solvent. For clarity, the reaction products are shown as green sticks around the catalytic iron cofactor (orange).

The product complex shows a product-induced structural change for this lid region (Figure 5.8 A**)** that causes almost complete shielding of the products from bulk solvent access. The three abreast proline residues (P212-214) build a hydrophobic wedge (Figure 5.8 B) that seals the active site. The interaction between the carboxyl group of the E125 side chain and the α-amino group of bound asparagine might be crucial to orient E125 in such a way that it establishes a hydrogen bond with the hydroxy group of the side chain of S215 (Figure 5.8 B). Thus, the lid closes concomitantly with L-Asn substrate binding. In the AsnO·Fe^{2+} structure, without αKG or Asn bound, the lid-like region F208-E223 is

open and leaves the enzyme center accessible for binding of substrate and co-substrate. With the bound products, hAsn and succinate, the lid bends towards the product binding site, burying it completely inside the enzyme (Figure 5.8 C**)**.

Summary of Data Statistics for AsnO Structures. The structure of AsnO was elucidated as apo-form, in a complex with iron and in complex with hAsn and succinate. The corresponding coordinate files have been deposited in the PDB database and are accessible under the codes 2OG5, 2OG6 and 2OG7. The data statistics for each dataset can be found in Table 5.2.

Crystallographic parameters	AsnO (apo)	AsnO· Fe^{2+}	AsnO·hAsn·αKG· Fe^{2+}
Data collection			
beamline	X13, EMBL	CuK_{α}, FR591	X13, EMBL
wavelength, Å	0.8080	1.5418	0.8080
detector	MARCCD 165	MAR 345	MARCCD 165
space group	$P3_121$	$P3_121$	$P3_121$
unit cell dimensions	*a*=*b*=91.6 Å	*a*=*b*=90.7 Å	*a*=*b*=91.3 Å
	c=90.9 Å	*c*=89.8 Å	*c*=90.9 Å
resolution, Å	23-1.45	20-1.90	23-1.65
structure amplitudes |F|	300230	260083	191425
unique reflections	78000	33610	51815
Wilson B-factor, $Å^2$	21.0	30.6	24.9
I/σ(I)	12.6 (0.9)	38.3 (2.1)	15.1 (1.7)
R_{merge}*	0.060 (0.992)	0.070 (0.416)	0.052 (0.539)
completeness, %	99.7 (99.7)	98.6 (78.8)	99.4 (98.4)
Refinement			
Used resolution	23-1.45 (1.48-1.45)	20-1.92 (1.97-1.92)	23-1.66 (1.70-1.66)
reflections (F>0)	75342	32018	50194
R_{Cryst}, %	17.1 (29.4)	16.9 (23.2)	16.8 (26.0)
R_{Free}, %	18.6 (29.4)	19.7 (25.2)	18.4 (24.2)
no. of atoms	2793	2721	2725
B-factor, $Å^2$	18.5	34.0	23.1
r.m.s.d. from ideal			
bonds, Å	0.007	0.007	0.008
angles, °	1.175	1.193	1.178

Table 5.2: Statistics on data collection and refinements of AsnO structures. The values in parenthesis indicate the highest resolution shell.

* $R_{merge} = \Sigma\left(\left|I-\langle I\rangle\right|\right)/\Sigma I$; $R_{Cryst}/R_{Free} = \Sigma\left(\left||F_0|-|F_c|\right|\right)/\Sigma|F_0|$.

5.1.4 Characterization of LptL

5.1.4.1 Hydroxylation Activity of LptL

Asparagine Hydroxylation. Based on the crystal structure of AsnO and its comparison with homology models of LptL and other putative hydroxylases (see Discussion), LptL was predicted to be responsible for (2*S*,3*S*)-3-hydroxyasparagine precursor synthesis during the assembly of the acidic lipoeptide A54145. In order to prove this prediction, the non-heme iron dioxygenase LptL (5 µM) was incubated with L-Asn (250 µM), the iron(II) cofactor and the αKG co-substrate (see Methods) for 10 min. Initial analysis was carried out by LC-MS with EIC for Asn (Figure 5.9 A) and hAsn (Figure 5.9 B).

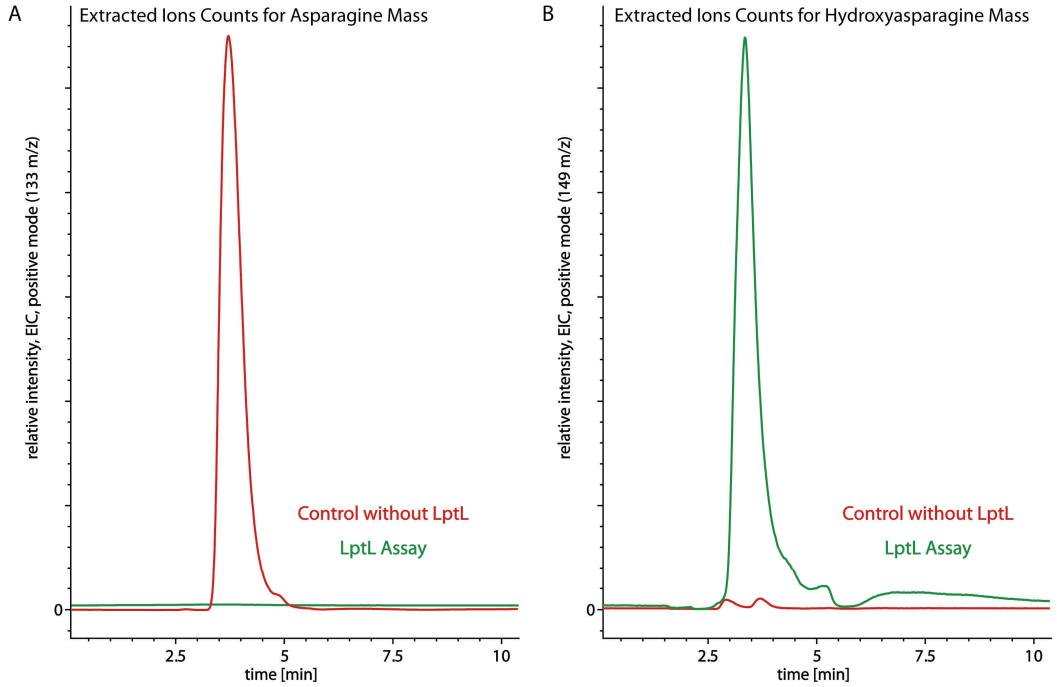


Figure 5.9: EIC analysis of the LptL assay. A) EIC for Asn. The red trace shows the control reaction without added LptL, the green trace the LptL assay, where no Asn was detected. B) EIC for hAsn. The green trace of the LptL assay shows complete conversion of Asn to hAsn after the assay. No hAsn was detected in the control reaction (red trace).

Hydroxylated asparagine and Asn were found to co-elute in standard HPLC-MS analysis, thus the ions of the starting material and product were extracted (EIC). After incubation with LptL no Asn was detected in the EIC spectrum (Figure 5.9 A, green trace), whereas a prominent peak representing the mass of hAsn was detected (Figure 5.9 B, green trace). This demonstrates that under the applied condition LptL catalyzes the quantitative conversion of Asn to hAsn. In the control reaction without LptL added, only Asn was detected (Figure 5.9, red traces), which excluded enzyme-independent hydroxylation.

Kinetic Parameter Determination. As the starting material Asn co-elutes with the product hAsn in standard HPLC-MS analysis, the same approach for kinetic measurements as in the case of AsnO was applied, namely FMOC-derivatization. Prior to substrate concentration variation, the time dependency of LptL-mediated hAsn generation was examined to determine the linear conversion range. To appoint the linear range, assays with LptL (0.5 µM and 5 µM) and 250 µM asparagine were stopped at different time points (Figure 5.10). The time dependent conversion revealed that the hydroxylation with a 0.5 µM LptL concentration was in the linear range between zero and approximately two and a half minutes, whereas a 5 µM LptL concentration lead to almost complete conversion after 1 min. Therefore the time span between zero and 1.5 minutes was used to determine the starting velocities for different substrate concentrations. The enzyme assays were stopped by addition of FMOC-Cl and the kinetic parameters were determined (Figure 5.11).

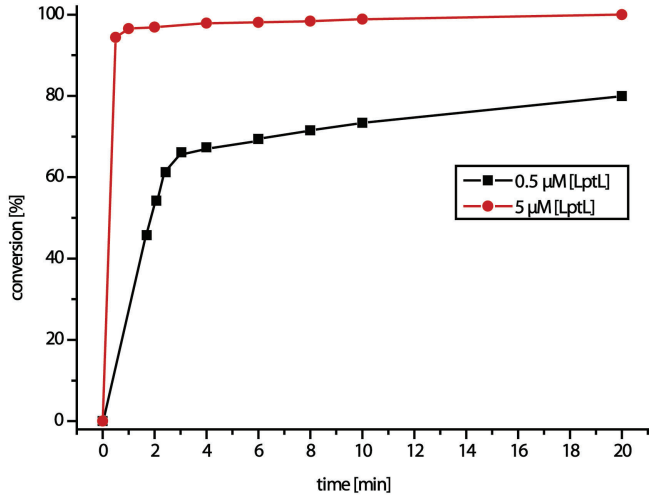


Figure 5.10: Time dependent Asn hydroxylation by LptL.

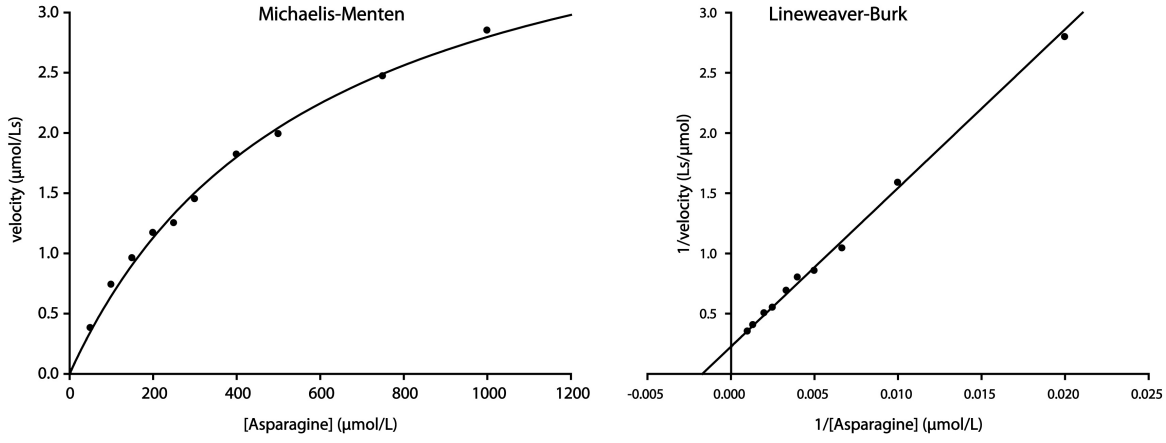


Figure 5.11: Michaelis-Menten and Lineweaver-Burk plots of LptL-mediated Asn hydroxylation.

The kinetics for LptL catalyzed L-Asn hydroxylation were calculated to an apparent K_M of 0.584 ± 0.049 mM and k_{cat}= 531 ± 24 min^{-1}, leading to a catalytic efficiency of k_{cat}/K_M= 0.91 ± 0.11 $min^{-1}µM^{-1}$.

5.1.4.2 Substrate Specificity

For the evaluation of the substrate specificity of LptL, the hydroxylase was incubated with a broad range of different substrates including the cognate PCP domain of module three of the A54145 NRPS (Lpt-PCP_3) loaded with L- and D-Asn. The latter was generated by incubation of the PCP with synthesized coenzyme A asparagine thioesters (CoA-*S*-Asn) and the promiscuous *B. subtilis* ppan transferase Sfp (see Methods for details). The PCP-tethered asparagine was tested in order to validate that LptL is involved in precursor synthesis only and not in on-line modification additionally. Amino acids representing the different classes of amino acids as well as the D-isomer of Asn were tested for LptL-mediated hydroxylation. Furthermore, FMOC-Asn and CoA-*S*-Asn as PCP mimic were tested as putative substrates. An overview of substrates tested for LptL catalyzed hydroxylation is given in Table 5.3.

Tested Substrate	Mass [Da]	Calculated Mass Hydroxylation Product [Da]	Observed Mass [Da]	Hydroxylation
arginine	175.1	191.1	175.1	no
aspartic acid	134.0	150.0	134.0	no
L-asparagine	133.1	149.1	149.1	yes
D-asparagine	133.1	149.1	133.1	no
glutamic acid	148.1	164.1	148.1	no
glutamine	147.1	163.1	147.1	no
isoleucine	132.1	148.2	132.2	no
lysine	147.1	163.2	147.2	no
methionine	150.1	166.1	150.1	no
phenylalanine	166.1	182.1	166.1	no
tyrosine	182.1	198.2	182.2	no
valine	118.1	134.1	118.1	no
FMOC-Asn	355.1	371.1	355.1	no
CoA-*S*-Asn	881.2	897.2	881.2	no
Lpt-PCP_3-*S*-L-Asn	11060	11076	11060	no
Lpt-PCP_3-*S*-D-Asn	11060	11076	11060	no

Table 5.3: Substrate specificity evaluation of LptL. Different amino acids including the D-enantiomer of Asn and PCP bound Asn isomers were tested for hydroxylation by LptL. All masses correspond to the $[M+H]^+$ as detected in positive mode MS, except for the PCP masses which are neutral mass reconstructions.

Neither the L- nor the D-isomer of PCP tethered asparagine was hydroxylated by LptL. Furthermore the FMOC-Asn as well as the CoA-*S*-Asn were not substrates for enzyme-mediated hydroxylation neither were the other tested amino acid, thus LptL exclusively catalyses the oxygenation of the β-position of L-asparagine. Therefore, LptL and AsnO are responsible for the synthesis of the hAsn building blocks found in CDA and A54145.

5.2 *In Trans* Acting Hydroxylases and Regeneration Domains during NRP Assembly

Another excellent example for the observed structural diversity observed in non-ribosomally produced bioactive agents, is the class of antifungal and antimicrobial kutznerides. The kutzneride hexadepsipeptide backbone is comprised of one α-hydroxy acid and five non-proteinogenic amino acids. One of the latter is D-3-hydroxyglutamic acid, which is found in the *threo* and *erythro* isomers in mature kutznerides. In this study the origin of these chemically intriguing residues was determined. The non-heme iron dioxygenases KtzO and KtzP were recombinantly expressed, characterized biochemically *in vitro*, and found to hydroxylate the β-position of glutamic acid. In contrast to the acidic lipopeptide hydroxylases, KtzO and KtzP were found to stereospecifically generate the *threo* and *erythro* isomers of 3-hydroxy-L-glutamic acid bound to the peptidyl carrier protein of the third module of the non-ribosomal peptide synthetase KtzH prior to epimerization by the module's E domain.

Interestingly, the third module of KtzH has a truncated adenylation domain and is unable to activate and incorporate glutamic acid. This study furthermore shows that the lack of a functional adenylation domain in the third KtzH module is compensated *in trans* by the stand-alone adenylation domain KtzN, which activates and transfers glutamic acid onto the carrier of KtzH in the presence of the truncated adenylation domain and either KtzO or KtzP. KtzN was identified and characterized as the first *in trans* NRPS assembly line regeneration domain.

Furthermore, a method that employs non-hydrolyzable coenzyme A analogs was developed and used to determine the kinetic parameters for KtzO- and KtzP-catalyzed hydroxylation of glutamic acid bound to the carrier protein. This is favorable compared with the native thioester bond which gets readily hydrolyzed. These insights may guide the use of KtzO/KtzP and KtzN or other *in trans* modification/restoration tools in biocombinatorial engineering approaches.

Together with the results from the precursor generating enzymes AsnO and LptL, the results concerning the *in trans* acting enzymes KtzO and KtzP are valuable additions towards the complete understanding of the mechanisms underlying β-hydroxylated amino acid biosynthesis and how they get incorporated into non-ribosomal peptides.

5.2.1 Recombinant Expression of *ktzN*, *ktzO*, *ktzP*, *a*pcp$_3$* and *pcp$_3$*

All constructs were cloned from fosmid DNA containing the kutzneride biosynthesis cluster and were expressed in *E. coli* BL21 Star (DE3) as described in the Methods section. The recombinant proteins were purified by amylose (KtzN) or Ni-NTA (all other) affinity chromatography and concentrated proteins were analyzed by SDS-PAGE and visualized using Coomassie dye (Figure 5.12). The final protein yields per liter culture after purification and concentration were 0.1 mg for KtzN, 3.1 mg for KtzO, 2.2 mg for KtzP, 5.7 mg for A*PCP$_3$ and 0.2 mg for PCP$_3$.

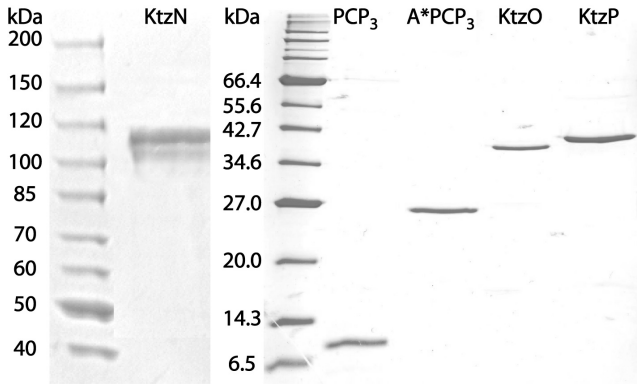


Figure 5.12: SDS-PAGE analysis of purified recombinant KtzN (110.1 kDa), PCP$_3$ (11.3 kDa), A*PCP$_3$ (22.3 kDa), KtzO (39.5 kDa) and KtzP (40.1 kDa). Protein markers were from left to right Fermentas PageRuler and NEB Broad Range Protein Marker.

5.2.2 Iron Content of KtzO and KtzP

The iron content of the non-heme iron(II) and αKG dependent dioxygenases KtzO and KtzP was determined analogously to AsnO and LptL by using the Ferene S [3-(2-pyridyl)-5,6-bis(2-(5-furyl sulfonic acid))-1,2,4,-triazine] dye.[212] The same standard curve for calibration of the Ferene S-iron complex was used (Figure 5.2). For KtzO the A_{593nm}-value was 0.009 and for KtzP 0.013, corresponding to 0.80 nmol for KtzO and 1.16 nmol for KtzP. As 10 nmol of each enzyme were used, these values correspond to 8% and 11.6% iron content. The low iron content is either due to oxidation or the iron was exchanged with other divalent ions during the purification steps. Therefore in all subsequent assays ferrous iron in form of ammonium ferrous iron sulfate was added to grant a constant supply of ferrous iron cofactor.

5.2.3 Characterization of KtzO and KtzP

5.2.3.1 Hydroxylation Activity of KtzO and KtzP

Hydroxylation of Carrier Protein-Bound Glutamic Acid. Sequence alignments of KtzO and KtzP revealed a putative ferrous iron binding motif HXD/E…H and similarity to the prototype non-heme iron(II) and αKG-dependent taurine dioxygenase TauD[222] and other oxygenases (Table 5.4).

Enzyme	Res.	First His motif	Res.	Second His and Arg
KtzO	98	PAQHAIQMHNESSYTNTWP	294	DNMLMSHGRRPYRGARQIRV
KtzP	95	PVDQRIMLHNENSYTKSFP	291	DNLLAAHGRDPYRGDRMILV
SyrP	123	PEREMILYHNESSHLESWP	319	DNMLAAHARDPYEEPRLIVV
TauD	92	PPD-NDNWHTDVTFIETPP	249	DNRVTQHYANADYLPQRRIM
		* *.: : : *		** : * . : :

Table 5.4: ClustalW-Alignment of KtzO, KtzP with the known iron(II) and αKG dependent hydroxylases SyrP and TauD.

KtzO and KtzP also share high sequence similarity with SyrP, a hydroxylase acting on the β-position of aspartic acid bound to the phosphopantetheinyl (ppan) co-factor of the eighth peptidyl carrier protein of the syringomycin E megasynthetase (SyrE-T8-*S*-Asp).[126] Therefore, KtzO and KtzP were predicted to hydroxylate glutamic acid at the β-position during kutzneride biosynthesis and to act on glutamic acid tethered to the third PCP of KtzH. One can speculate that either one enzyme hydroxylates PCP-bound glutamic acid and generates both isomers of L-3-hydroxyglutamic acid (*threo*- and *erythro*-hGlu), and that the other is not involved in hGlu biosynthesis, or that both work on PCP-tethered Glu. In the latter case, it is possible that the enzymes act stereospecifically to generate one hGlu isomer each. To determine the biochemical origin of the hGlu residues found in mature kutznerides, KtzO or KtzP was incubated with the glutamic acid ppan thioesters of the truncated A* domain-PCP (A*PCP$_3$-*S*-Glu), as A* might play a role in hydroxylase-PCP interaction. Preincubation of A*PCP$_3$ with synthetic CoA-*S*-Glu and the promiscuous ppan-transferase Sfp[52,95] generated the loaded carrier A*PCP$_3$-*S*-Glu, which was subsequently incubated with αKG, iron(II) and either KtzO or KtzP. QTOF-MS revealed that the apo-form of A*PCP$_3$ (Figure 5.13 A) with an apparent mass of 22,343 Da was loaded with ppan-*S*-Glu to generate A*PCP$_3$-*S*-Glu (Figure 5.13 B), indicated by the mass increase of 470 Da. Incubations with KtzO (Figure 5.13 C) and KtzP (Figure 5.13 D) resulted in the formation of hydroxylated A*PCP$_3$-bound L-glutamic acid (A*PCP$_3$-*S*-

hGlu), shown by the mass difference of 16 Da. Hydrolysis of A*PCP_3-*S*-Glu occurred at neutral pH (Figure 5.13 B-D, mass: 22,684 Da) after 2 hours. The control reaction lacking αKG only showed the mass of A*PCP_3-S-Glu and holo-A*PCP_3 (Figure 5.13 B), providing no evidence for hydroxylase activity.

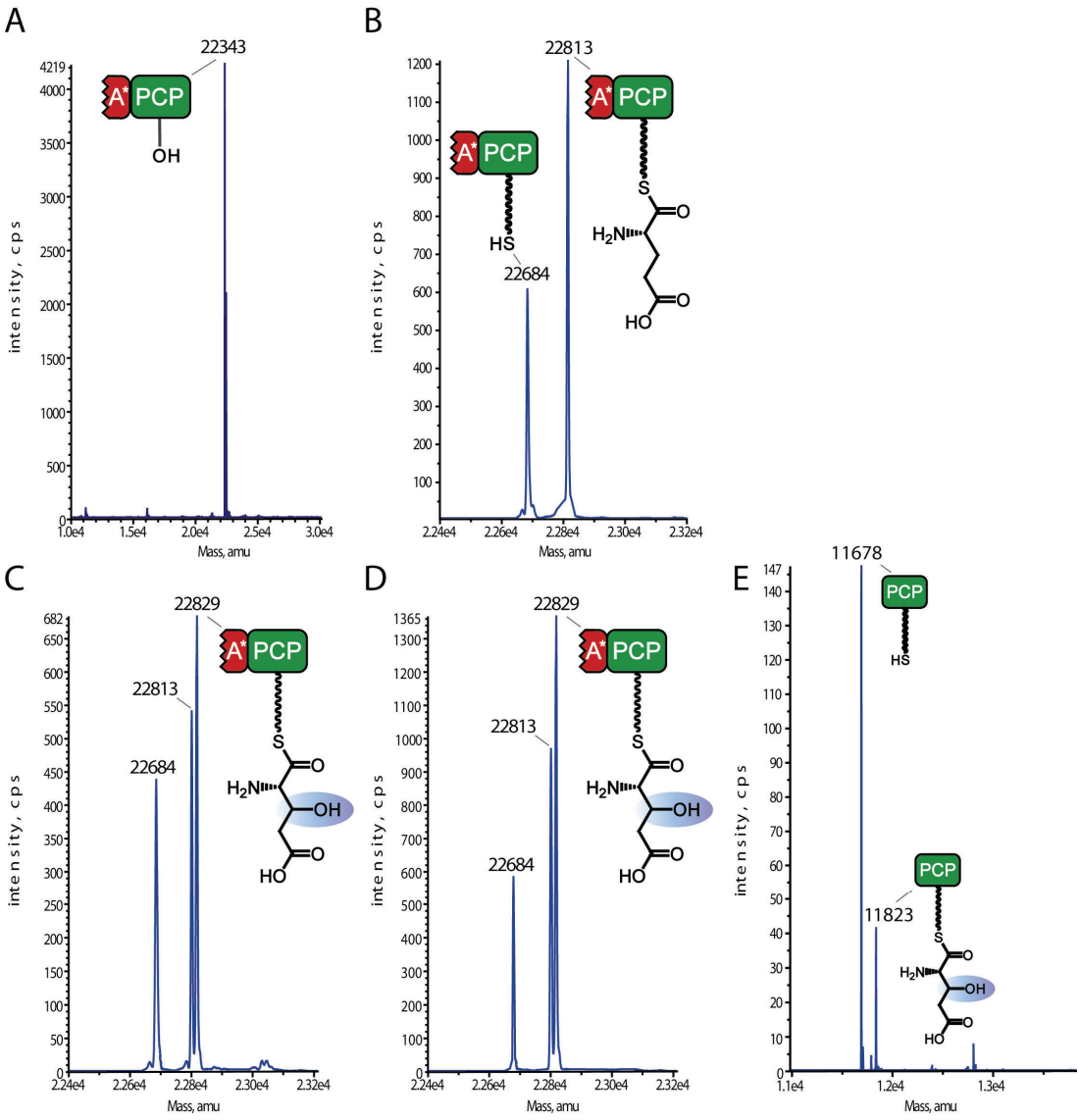


Figure 5.13: QTOF-MS analysis of hydroxylation activity of KtzO and KtzP on PCP-tethered glutamic acid. A) apo-A*PCP_3 with an apparent mass of 22343 Da. B) control reaction: A*PCP_3 loaded with synthetic Coenzyme A glutamic acid by Sfp and incubated for 2h. The mass shift of 470 Da shows that ppan-*S*-Glu was transferred on the PCP domain. C) A*PCP_3 after incubation with αKG, $(NH_4)_2Fe(SO_4)_2$ and KtzO, preceded by incubation with Sfp and CoA-*S*-Glu. The mass difference of 16 compared with B) indicates hydroxylation of the PCP-bound glutamic acid. D) Same as C, but after incubation with KtzP. Hydrolysis is observed for B and C and D, and the mass of 22684 Da corresponds to the holo-A*PCP_3. E) Typical MS analysis of PCP_3 loaded with CoA-*S*-Glu by Sfp and incubated with KtzP for 60 min at 28°C. The holo-PCP_3 peak with 11,678 Da was observed as hydrolysis of the PCP_3-*S*-Glu. The peak at 11,823 Da has a mass shift of 145 Da compared to the holo-peak, which corresponds to the mass of the hydroxylated glutamic acid moiety bound as thioester. MS-analysis for KtzO assays were similar (data not shown).

Reactions conducted with the single PCP_3 (PCP_3-*S*-Glu), also revealed a mass shift of 16 Da, indicating that hydroxylation occurred (Figure 5.13 E). These findings demonstrate that both KtzO and KtzP are indeed non-heme iron(II)- and αKG-dependent oxygenases that catalyze the hydroxylation of PCP-tethered glutamic acid *in vitro*. The truncated A domain observed in module three of KtzH is not required for hydroxylation of PCP-bound substrate.

β-Hydroxylation and Stereospecificity. In order to confirm the Glu-β-hydroxylation and to ascertain the stereospecificity of KtzO and KtzP, the *threo* and the *erythro* isomers of β-OH-Glu were synthesized in collaboration with the group of Prof. Dr. C. T. Walsh based on a previously published protocol (Figure 5.14).[18]

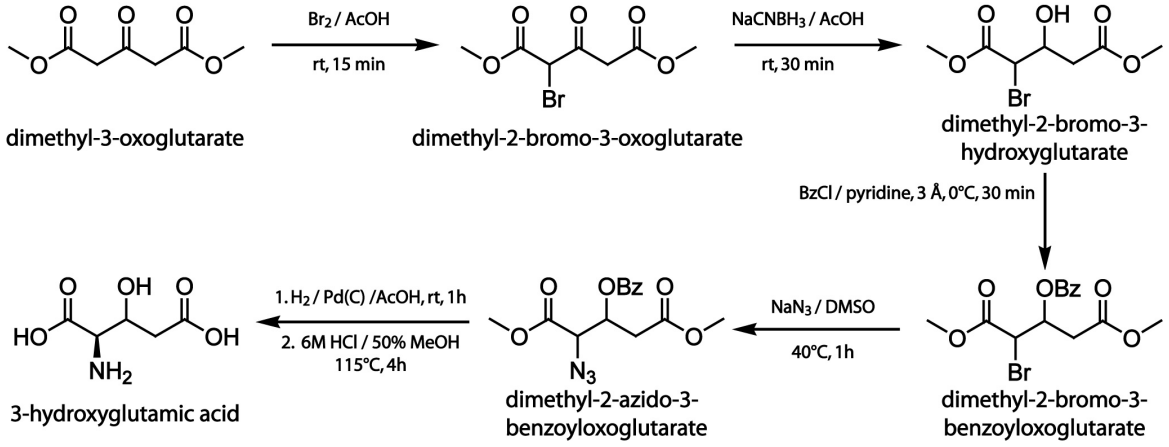


Figure 5.14: Synthesis of L-3-hydroxyglutamic acid as racemic mixture.[**]

The enantiomerically pure *threo*-L-hydroxyglutamic acid and the *erythro*-L-glutamic acid were dabsylated to facilitate separation and identification by reversed phase HPLC-MS as described in the Methods section. The hydroxylation assays were carried out as described for PCP-bound glutamic acid followed by subsequent alkaline hydrolysis of the PCP-*S*-hGlu thioesters. Co-elution analyses employing synthetic and dabsylated hGlu standards revealed that KtzO and KtzP are stereospecific hydroxylases acting on the β-position of Glu. KtzO generates *threo*-L-hydroxyglutamic acid and KtzP catalyzes the formation of *erythro*-hGlu (Figure 5.15). This system is the first example for a NRPS where products with two isomers of a C_β modified amino acid are generated by two

[**] The synthesis of the hydroxyglutamic standards was carried out by Dr. Elizabeth M. Nolan as part of the collaboration with Prof. Dr. C. T. Walsh (Harvard Medical School, Boston, MA, USA). The synthesis route is based on a previously published protocol by Broberg *et al.*[18]

modifying enzymes that catalyze the very same reaction but with opposite stereospecificity.

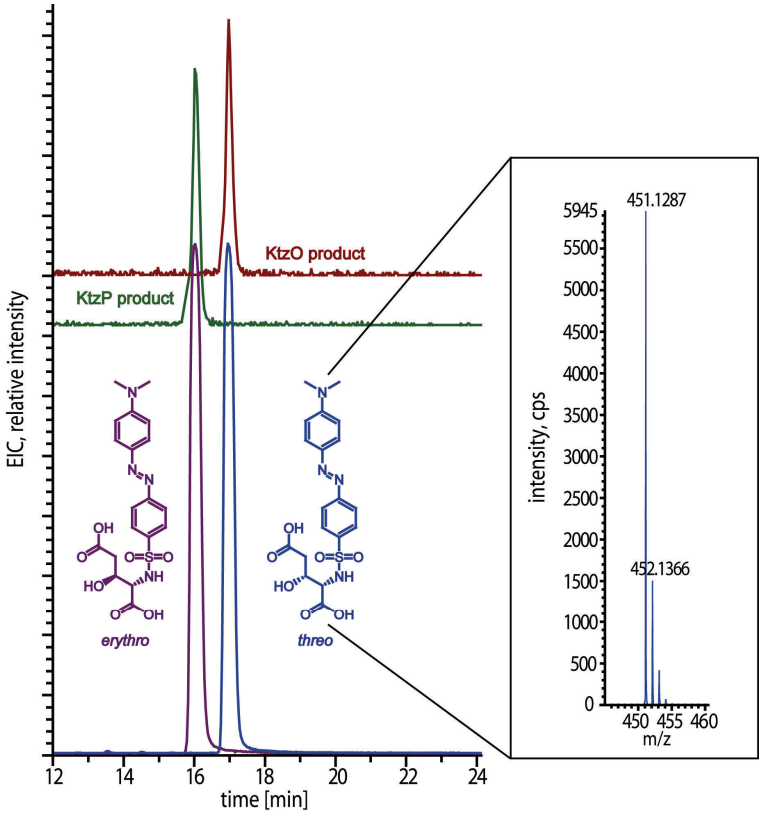


Figure 5.15. Stereospecificity of KtzO and KtzP. KtzO and KtzP products were cleaved off the PCP, derivatized with dabsyl-Cl, analyzed via HPLC-MS and compared with synthetic standards of the dabsylated L-β-hydroxyglutamic acids. The KtzP product co-elutes with the Dab-L-*erythro*-hGlu and the KtzO product with the *threo* form. High resolution mass spectrometric analyses of the peaks (observed $[M+H]^+$= 451.1287 m/z) are in accordance with the theoretical masses (451.1282 m/z).

Kinetic Parameter Determination Using CoA Analogs. The unstable nature of the thioester bond between glutamic acid and PCP, as observed in the assays stated above (Figure 5.13 BCD), makes it difficult to collect accurate biochemical data for determining the kinetic parameters for KtzO- and KtzP-catalyzed reactions. Although different strategies have been developed to circumvent thioester hydrolysis, e.g. amide ligation,[209] basic cleavage or thioesterase type II mediated cleavage,[125] a convincing method for direct kinetic parameter determination of carrier protein ppan-thioesters has not been established to date. The reported methods involve thioester cleavage, and thus the amount of substrate hydrolyzed prior to cleavage is impossible to quantify.

The method introduced herein, inspired by work of the Bruner group[101] where phosphopantetheinyl thioester mimics, including amino-coenzyme A (NH_2-CoA), were used to manipulate the geometry of carrier domains in multidomain NRPS assemblies

involves replacement of the labile thioester linkage with a hydrolytically stable amide bond. Amino-coenzyme A was used as a synthetic precursor to prepare a non-hydrolyzable amino coenzyme A coupled to glutamic acid (CoA-H*N*-Glu, Figure 5.16) as detailed in the Methods section.

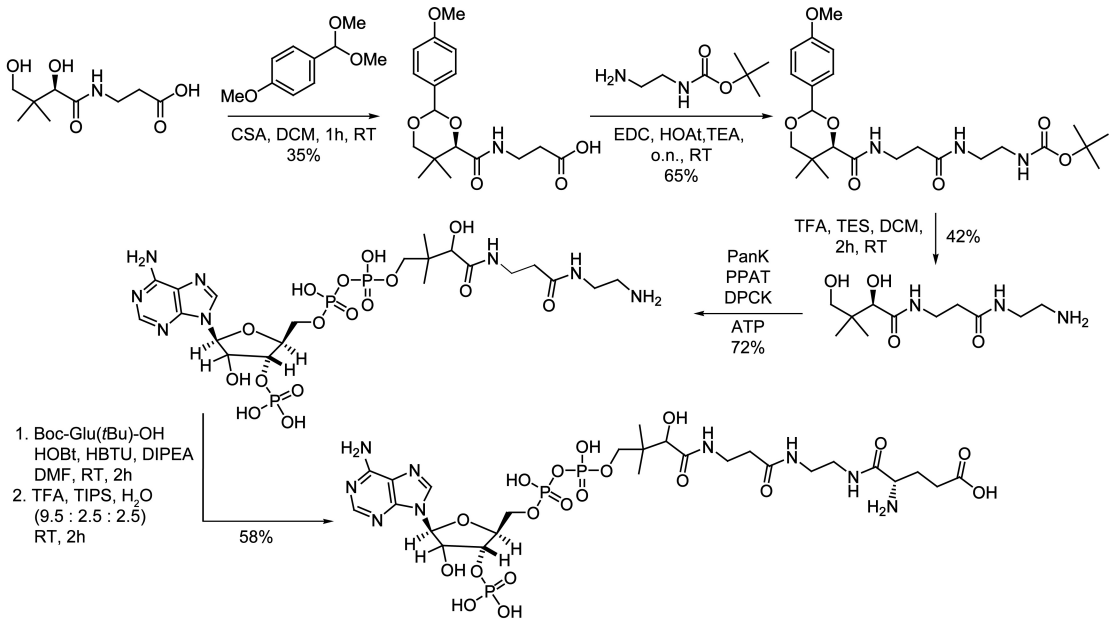


Figure 5.16: Synthesis of amino coenzyme A and its coupling to glutamic acid.

Instead of starting the synthesis with trityl-diaminoethane resin, which was used in previous studies, free pantothenic acid was used. The diol was protected and the pantothenic acid was coupled to Boc-diaminoethane to yield fully protected aminopantetheine. After deprotection aminopantetheine was incubated with the three coenzyme A biosynthesis enzymes[††] PanK, PPAT, DPCK,[223] which were recombinantly expressed in *E. coli*, and ATP to yield the Coenzyme A analog, in which the thiol group is replaced by an amine group. The amino-CoA was then coupled to the Boc-protected glutamic acid using standard amide coupling reagents. Upon acidic *in situ* deprotection, CoA-H*N*-Glu was obtained with an overall yield of 4% (Figure 5.16).

A*PCP$_3$ was loaded with CoA-H*N*-Glu by Sfp to yield the A*PCP$_3$-H*N*-Glu and MS analysis revealed complete conversion of apo-A*PCP$_3$ to A*PCP$_3$-H*N*-Glu (Figure 5.17 A). After incubation with either KtzO or KtzP, αKG and ferrous iron, a mass shift of 16 Da was

[††] The three coenzyme A biosynthesis enzymes were expressed from plasmids provided by Dr. Eric R. Strieter (Harvard Medical School) during the course of the ongoing collaboration with Prof. Dr. C. T. Walsh. Expression and purification was carried out by the technician Gabriele Schimpff-Weiland (Philipps-University Marburg). Details on the expression and purification can be found in *Liu et al.*[101]

observed (KtzO, Figure 5.17 B, KtzP, Figure 5.17 C), demonstrating that KtzO and KtzP catalyzed the hydroxylation of A*PCP$_3$-H*N*-Glu.

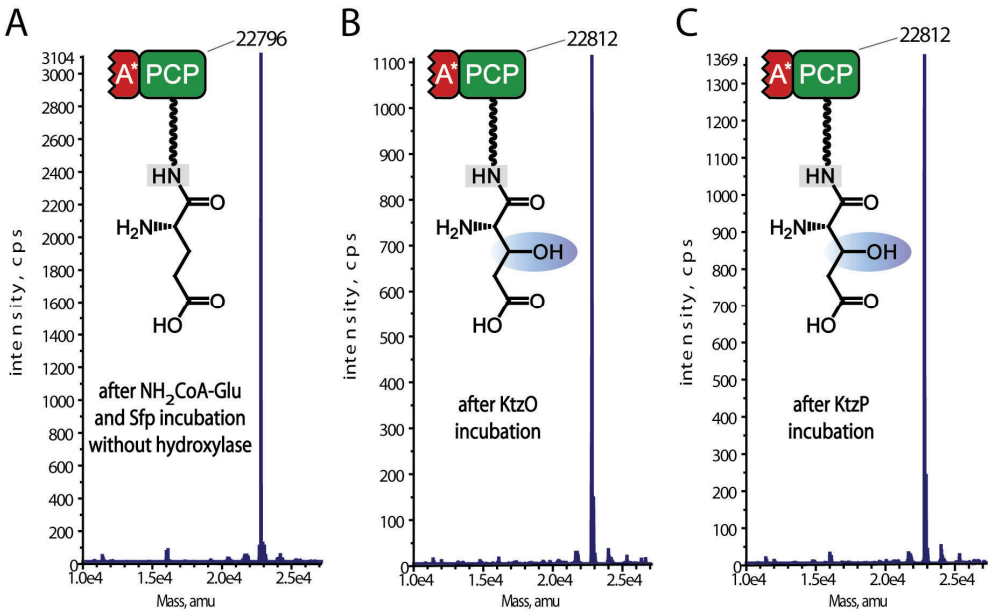


Figure 5.17: QTOF-MS analysis of CoA-H*N*-glutamic acid loading on A*PCP$_3$ and hydroxylation assays. A) A*PCP$_3$ incubated with Sfp and CoA-H*N*-Glu. The mass shift of 453 Da compared to apo-A*PCP$_3$ (Figure 5.13 A, 22343 Da) is in agreement with the calculated ppan-H*N*-Glu addition. B, C) Incubation of ppan-H*N*-Glu loaded A*PCP$_3$ with KtzO (B) or KtzP (C) yielded in a mass shift of 16 Da, suggesting hydroxylation.

The additional oxygen introduced by the hydroxylation reaction is of no consequence for the ionization properties of the examined protein. Thus, the mass spectrometer counts for the loaded A*PCP$_3$-H*N*-Glu and its hydroxylated form were used to determine the percentage of the hydroxylated protein species relative to the non-hydroxylated one. By variation of the A*PCP$_3$-H*N*-Glu concentration (10 to 450 µM), it was possible to measure the starting velocities and therefore to determine Michaelis-Menten kinetic parameters (Figure 5.18).

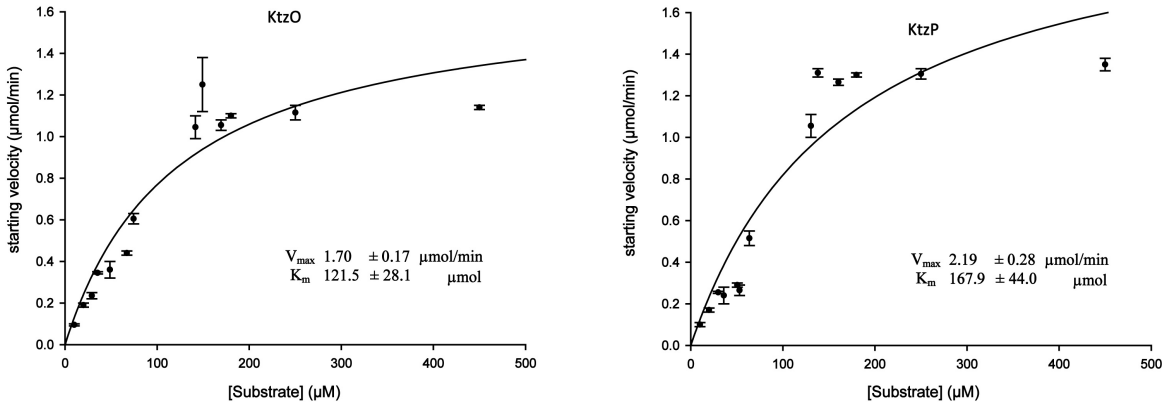


Figure 5.18: Michaelis-Menten plots of KtzO and KtzP-mediated hydroxylation of A*PCP-H*N*-Glu.

For KtzO catalyzed hydroxylation the kinetic parameters were: K_M 121.5 ± 28.1 mM and k_{cat}= 0.34 ± 0.03 min^{-1}, leading to a catalytic efficiency of k_{cat}/K_M= $2.8 \cdot 10^{-3}$ min^{-1} μM^{-1}; and for KtzP they were: K_M 168.9 ± 44.0 mM and k_{cat}= 0.44 ± 0.05 min^{-1}, leading to a catalytic efficiency of k_{cat}/K_M= $2.6 \cdot 10^{-3}$ $min^{-1} \mu M^{-1}$. A*PCP_3 was chosen for kinetic measurements over the single PCP as its production level and its ionization properties are much better. However, the incubation of the single PCP with KtzO or KtzP yielded the same conversion percentage (data not shown), indicating that that similar kinetic properties should be expected.

5.2.3.2 Substrate specificity

To further characterize KtzO and KtzP, their substrate specificities were evaluated. Free amino acids (L-Glu, D-Glu), different PCP-bound amino acids, especially aspartic acid because of its observed activity in KtzN adenylation assays (Figure 5.21), CoA-*S*-Glu as PCP mimic as well as a non-cognate PCP domain from the CDA biosynthetic cluster (Figure 5.1) were considered and the results are summarized in Table 5.5.

Substrate	$[M+H]^+$	$[M+H]^+$ hydroxylated	$[M+H]^+$ observed	Hydroxylation
L-glutamic acid	148.1	164.1	148.1	no
D-glutamic acid	148.1	164.1	148.1	no
CoA-*S*-L-Glu	897.2	913.2	897.2	no
CoA-*S*-D-Glu	897.2	913.2	897.2	no
KtzHA*T3-*S*-L-Glu	22813	22829	22829	**yes**
KtzHA*T3-*S*-D-Glu	22813	22829	22813	no
KtzHT3-*S*-L-Glu	11807	11823	11823	**yes**
KtzHT3-*S*-D-Glu	11807	11823	11807	no
KtzHA*T3-H*N*-L-Glu	22796	22812	22812	**yes**
KtzHT3-H*N*-L-Glu	11790	11806	11806	**yes**
CDA-PCP_9-*S*-L-Glu	12616	12632	12632	**yes**
CDA-PCP_9-H*N*-L-Glu	12599	12615	12615	**yes**
KtzHA*T3-*S*-L-Leu	22797	22813	22797	no
KtzHA*T3-*S*-L-Phe	22831	22847	22831	no
KtzHA*T3-*S*-L-Asp	22799	22815	22799	no
KtzHA*T3-*S*-L-Asn	22798	22814	22798	no
KtzHA*T3-*S*-L-Gln	22812	22828	22812	no

Table 5.5: Substrate specificity of KtzO and KtzP.

Interestingly, KtzO and KtzP are both able to hydroxylate glutamic acid bound to a non-cognate PCP, in this case the ninth PCP domain of the CDA NRPS assembly line (CDA-PCP_9) (Figure 5.19).

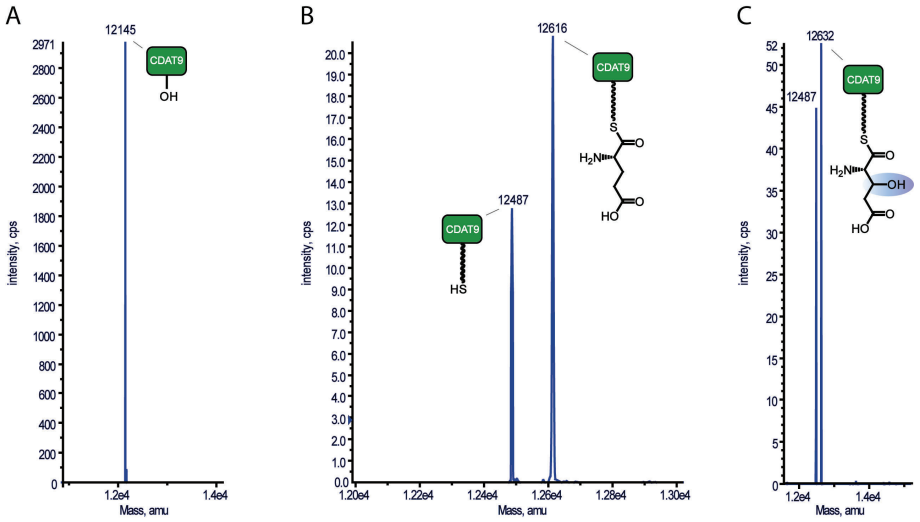


Figure 5.19: Hydroxylation assay with the non-cognate PCP of the ninth module of the CDA synthetase. A) MS analysis of apo-CDA-PCP$_9$; B) apo-CDA-PCP$_9$ loaded with CoA-*S*-Glu by Sfp. Hydrolysis occurred as observed by the holo-peak of CDA-PCP$_9$ (12,487 Da). C) CDA-PCP$_9$-*S*-Glu after incubation with KtzO, the recorded mass shift of 16 Da indicates hydroxylation. KtzP spectra showed similar patterns (data not shown). As hydrolysis occurred, the kinetic measurements were carried out using CoA-H*N*-Glu.

The kinetic parameters for this reaction were determined for both hydroxylases (Figure 5.20) in the same way as described for the cognate PCP domain, i.e. by the usage of the non-hydrolyzable CoA analog coupled to glutamic acid.

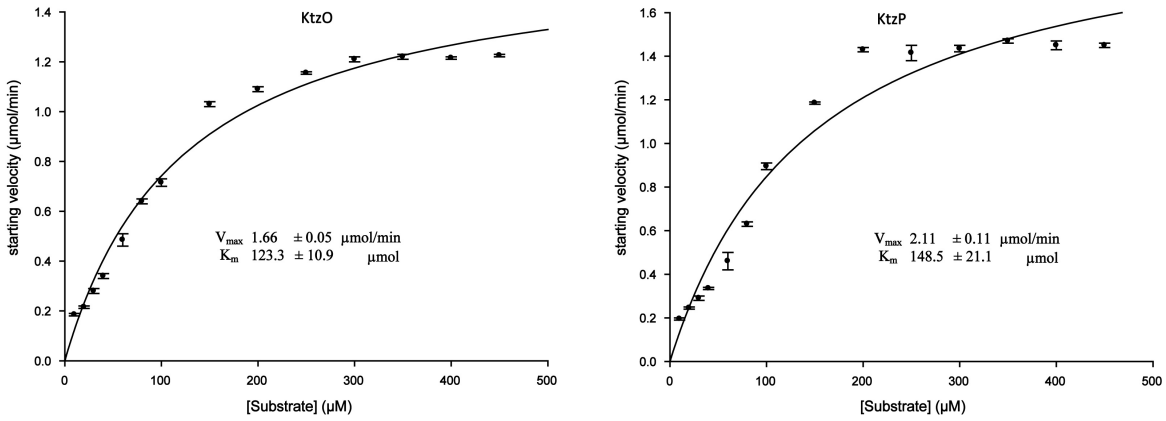


Figure 5.20: Michelis-Menten-Plot of the starting velocities vs. the substrate concentrations for CDA-PCP$_9$-H*N*-Glu hydroxylation by KtzO and KtzP. The values are in the same range as for the cognate PCP.

It was found that KtzO and KtzP were similarly efficient in hydroxylating CDA-PCP$_9$-H*N*-Glu compared with A*PCP$_3$-H*N*-Glu *in vitro*. The kinetic parameters for CDA-PCP$_9$-H*N*-Glu hydroxylation by KtzO were K_M 123.3 ± 10.9 mM and k_{cat}= 0.33 ± 0.01 min^{-1}, leading to a catalytic efficiency of k_{cat}/K_M= 2.7·10^{-3}min^{-1}µM^{-1} and for KtzP they were K_M 148.5 ± 21.1 mM and k_{cat}= 0.42 ± 0.02 min^{-1}, leading to a catalytic efficiency of k_{cat}/K_M= 2.8·10^{-3} min^{-1}µM^{-1}.

5.2.4 Adenylation Activity of KtzN

Adenylation activity of KtzN was tested, because it was postulated that this enzyme loads the third PCP domain of KtzH with glutamic or 3-hydroxyglutamic acid *in trans* and thus substitutes for the activity of incomplete A* domain in module three of KtzH.[158] Analysis of the primary sequence within substrate-binding pockets of adenylation domains allows for the prediction of A domain selectivities. KtzN was predicted to be specific for glutamic or aspartic acid in agreement with the primary structure of kutznerides. To obtain experimental evidence for the proposed specificity of KtzN, the ability of the recombinant MBP-fusion protein to reversibly adenylate various amino acids was tested by using the well established ATP/PPi exchange assay[224] (see Methods).

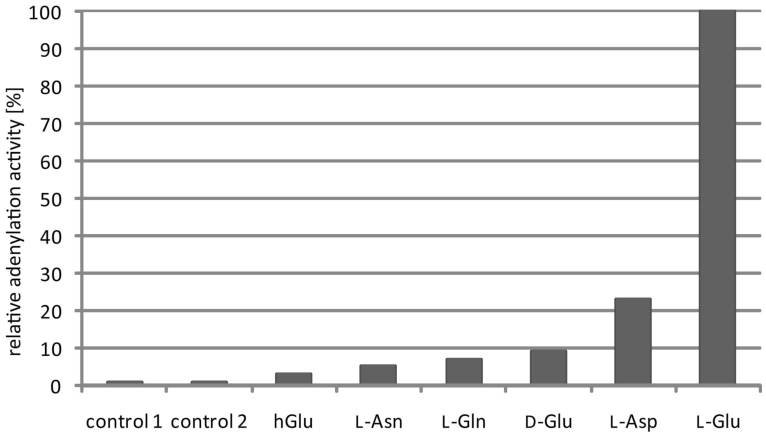


Figure 5.21: Adenylation activity of KtzN. Control 1 is without ATP, control 2 is without enzyme, hGlu, racemic mixture of 3-hydroxyglutamic acid.

KtzN showed a preference towards reversible formation of L-Glu-AMP with modest activation of L-Asp (Figure 5.21). Incubation without ATP or enzyme as well as incubation with hydroxyglutamic acid showed no significant formation of radioactive ATP.

5.2.5 *In Trans* Activity of KtzN

Another intriguing factor regarding tailoring of hGlu and the kutzneride assembly line is the mechanism of activation and transfer of the glutamic acid onto the third PCP domain of KtzH. It was postulated that the stand-alone adenylation domain KtzN (Figure 1) acts as an *in trans* substitute domain for the truncated A* domain in the third module of KtzH by activating glutamic acid as an amino acyl adenylate, and transferring it to the third module.[158] KtzN was shown to activate glutamic acid (see 5.2.3.3), but the transfer event and factors that may mediated this reaction, such as protein-protein interaction, the

role of the hydroxylases or A*, are unknown. To evaluate what protein-protein interactions are required for Glu transfer by KtzN and to gain more mechanistic insight into this unusual NRPS recovery strategy, Glu transfer was monitored in an assay containing recombinant KtzN (Glu A domain), ATP, L-Glu, holo-A*PCP$_3$ and in the absence or presence of one hydroxylase (KtzO or KtzP), αKG and ferrous iron as described in the Methods section.

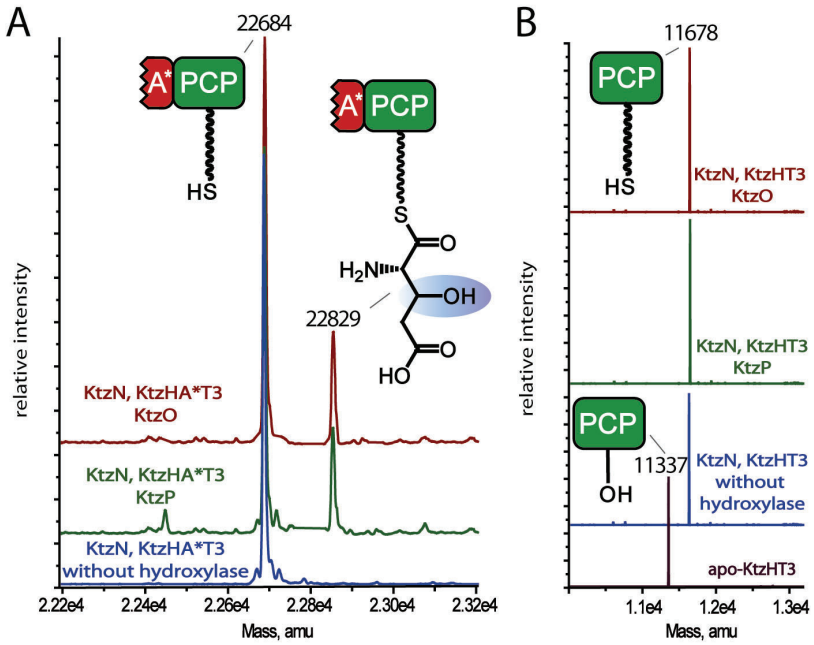


Figure 5.22: MS analysis of coupled glutamic acid transfer assays. A) Incubation of KtzN with holo A*PCP$_3$, KtzO (red trace) or KtzP (green trace) and their co-substrates yielded the transfer of glutamic acid and subsequent hydroxylation by KtzO and KtzP. No transfer of glutamic acid by KtzN was observed without hydroxylase added (blue trace). B) Role of A*: The KtzN coupled assay was carried out as in A, but the single PCP was used instead of A*PCP of the third module of KtzH. In all cases, only holo-PCP was detected after the coupled assays.

MS-Analysis of the assay of KtzN with holo-A*PCP$_3$ in the absence of hydroxylase revealed that no glutamic acid was transferred (Figure 5.22, blue trace). Upon addition of either KtzO (Figure 5.22 A, dark red trace) or KtzP (Figure 5.22 A, green trace), co-substrate and cofactor, glutamic acid was transferred onto A*PCP$_3$ and subsequently hydroxylated. The reaction was either not complete or hydrolysis of the PCP-*S*-hGlu thioester occurred prior to analysis, indicated by the prominent holo-A*PCP$_3$ peak (Figure 5.22 A). To evaluate the role of the truncated adenylation domain A*, the assay was carried out using the single PCP$_3$ domain instead of A*PCP$_3$. MS analyses showed that no glutamic acid was transferred regardless of the presence or absence of any hydroxylase (Figure 5.22 B).

5.3 Rational Manipulation of β-Hydroxylases for the Synthesis of Pharmaceutical Relevant Agents

L-THA (L-*threo* or (2*S*,3*S*)-3-hydroxyaspartic acid), the β-hydroxylated form of L-aspartic acid, is of current medicinal interest because it inhibits the function of excitatory amino acid (EAA) transporters as a L-glutamic acid mimic.[225] As the key player in primary neurotransmission in the mammalian central nervous system (CNS) the latter participates in diverse and complex neuronal communication by activating a broad assortment of the EAA receptors.[226] With its potential to over-activate these receptors L-glutamate can contribute to CNS damage in acute injuries or chronic diseases.[227] Thus, regulation of extracellular L-glutamate concentration, carried out by the EAA transporters, is crucial. A readily available source of L-THA could help to further investigate complex signaling processes.

The synthesis of L-THA is a challenging task for the organic chemists, therefore a direct stereospecific enzymatic route is more favorable. The development of the enzymatic L-THA synthesis is described in the following paragraphs. The acquired knowledge on the origin and mechanisms underlying β-hydroxylated amino acid biosynthesis facilitated rational manipulation of the substrate specificity of AsnO.

AsnO, which was shown to stereospecifically synthesize (2*S*,3*S*)-3-hydroxyasparagine (see chapter 5.1.3), binds the carboxamide function of the asparagine substrate via the side chain of the aspartic acid residue 241 (D241) as shown in the crystal structure (Figure 5.7 B). In this study, the residue 241 was exchanged by site-directed mutagenesis to asparagine in order to generate a binding site for the side chain of the aspartic acid substrate. The D241N AsnO variant was able to directly hydroxylate aspartic acid at the aliphatic β-position. Furthermore the manipulated enzyme was biochemically characterized and it became evident, that a highly specific hydroxylase was generated. The AsnO variant exclusively accepted aspartic acid as a substrate and generated L-THA in enantiomerically pure form, even on a semi-preparative scale.

5.3.1 Recombinant Expression of *asnO D241N*

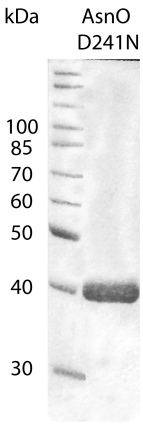


Figure 5.23: SDS-PAGE analysis of purified recombinant AsnO D241N (38.9 kDa). Protein marker was Fermentas PageRuler.

AsnO D241N was constructed by site-directed mutagenesis with the pQTEV(*asnO*) expression vector as template. The 721st *asnO* base guanine was mutated to adenine. This resulted in an asparagine codon (AAC, bases 721-723) which replaced the originial GAC aspartic acid codon. Therefore the gene product is the AsnO D241N variant, which was expressed in *E. coli* BL21 (DE3) as described in the Methods section. The recombinant protein was purified by Ni-NTA affinity chromatography. The concentrated enzyme was analyzed by SDS-PAGE and visualized using Coomassie dye (Figure 5.23). The final protein yield per liter culture after purification and concentration was 6 mg.

5.3.2 Iron Content Determination of AsnO

The iron content of the non-heme iron(II) and αKG dependent dioxygenases AsnO D241N was determined analogously to 5.1.2 and 5.2.2 using the Ferene S dye.[212] The same standard curve for calibration of the Ferene S-iron complex was used (Figure 5.2). The A_{593nm}-value for AsnO D241N was 0.007 corresponding to 0.62 nmol. As 10 nmol of the enzyme were used, this value corresponds to 6.2% iron content. Again, the low iron content is either due to oxidation or the iron was exchanged with other divalent ions during the purification steps. Therefore in all subsequent assays ferrous iron in form of ammonium ferrous iron sulfate was added to grant a constant supply of ferrous iron cofactor.

5.3.3 Characterization of AsnO D241N

5.3.3.1 Hydroxylation Activity of AsnO D241N

Aspartic Acid Hydroxylation. To evaluate the activity of AsnO D241N on an analytical scale, the purified enzyme was incubated with aspartic acid, $(NH_4)_2Fe(SO_4)_2$, as source of the ferrous iron cofactor, and αKG co-substrate at different temperatures. The reaction was monitored and the identity of the reaction product was proven by HPLC-MS analysis

by scanning for the masses of aspartic acid ($[M+H]^+$= 134.05 Da) and its hydroxylated form L-THA ($[M+H]^+$= 150.04 Da), respectively.

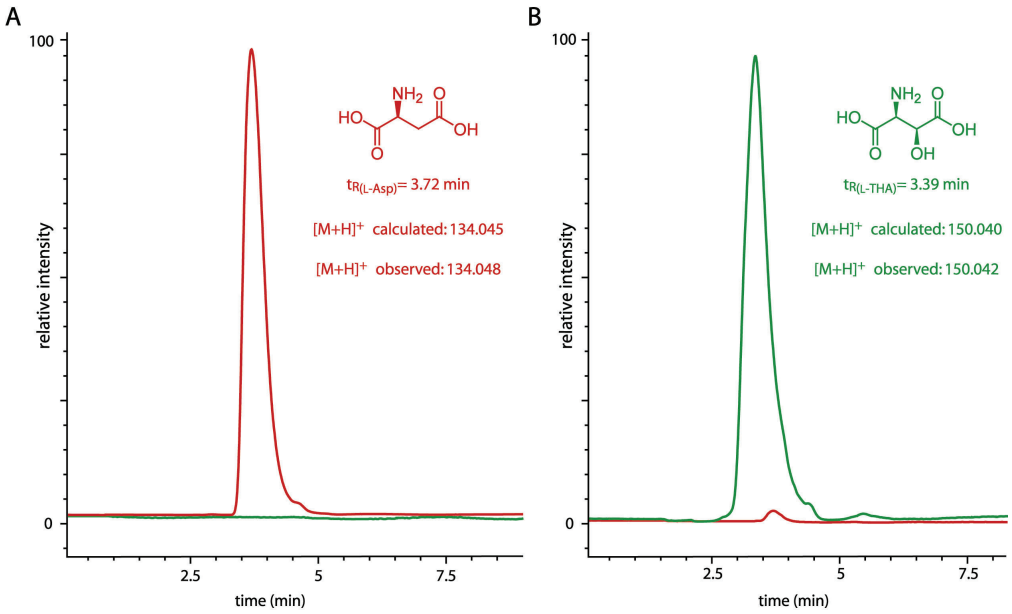


Figure 5.24: HPLC-MS activity analysis of AsnO D241N. A) Control without enzyme added. Only the mass of L-Asp (green trace) is observed, while L-THA (red trace) is not detected. B) Enzyme Assay. Almost all of the starting material (red trace) was consumed by the enzyme to generate L-THA (green trace).

The HPLC-MS chromatogram (Figure 5.24 B) of the assay after incubation of Asp with AsnO D241N revealed nearly quantitative conversion to L-THA at 16°C. In the control reaction without enzyme added (Figure 5.24 A), only the substrate L-Asp was observed. Complete conversion was only observed at 16°C; higher temperatures yielded in 80, 67, 55 and 32% conversion for 20, 25, 30 and 37°C (data not shown).

Kinetic Parameter Determination. As the L-THA standard was commercially available, it was possible to determine if the additional hydroxyl group of L-THA influences the ionization properties compared with Asp. MS analysis in the positive mode revealed that both compounds show same ionization patterns. Therefore, kinetic parameters were determined directly by MS analyses using extracted ions counts instead of FMOC or dabsyl derivatization. To determine the kinetic parameters of the variant, assays with 5 µM enzyme were quenched at appropriate time points for different substrate concentrations (50 µM to 2 mM), by addition of perfluoropentanoic acid, which also acts as an ion pairing reagent for HPLC-MS analysis. The reaction follows Michaelis-Menten kinetics (Figure 5.25).

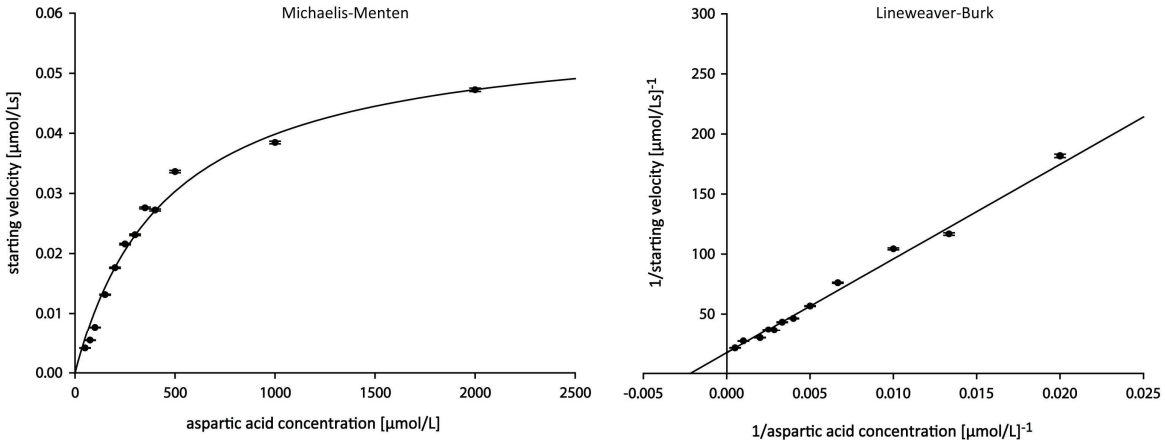


Figure 5.25: Michaelis-Menten and Lineweaver-Burk plots of AsnO D241N-mediated Asp hydroxylation.

The Lineweaver-Burk plot facilitated kinetic parameter calculation and revealed an apparent K_M of 0.46 ± 0.03 mM for the AsnO D241N variant. The k_{cat} was determined to 1.0 ± 0.1 min^{-1}, which afforded a catalytic efficiency of k_{cat}/K_M= 2.2 ± 0.4 min^{-1} mM^{-1}.

5.3.3.2 Substrate Specificity

The substrate specificity of the AsnO variant was evaluated by incubation of the enzyme with different amino acids overnight at 16°C (Table 5.6).

Amino Acid	Substrate $[M+H]^+$	Calculated $[M+H]^+$ Product	Mass Observed	Hydroxylation
L-Asp	134.0	150.0	150.0	yes
D-Asp	134.0	150.0	134.0	no
L-Asn	133.1	149.1	133.0	no
L-Glu	148.1	164.1	148.2	no
L-Gln	147.1	163.1	147.2	no
L-Ile	132.1	148.1	132.2	no
L-Phe	166.1	182.1	166.1	no
L-Trp	205.1	221.1	205.3	no
L-Val	118.1	134.1	118.0	no

Table 5.6: Evaluation of the substrate specificity of AsnO D241N.

LC-MS analyses of these assays showed that the engineered protein exclusively accepted L-Asp as its cognate substrate. Neither the AsnO wild-type substrate L-Asn nor L-Asp-related amino acids, like L-Glu and or the D-isomer of Asp, or representatives of other amino acid classes were accepted as substrates for protein-mediated hydroxylation.

5.3.4 Preparative Enzymatic Synthesis of L-THA

To evaluate the ability of the AsnO D241N variant to be used as an alternative to classical organic L-THA synthesis, the enzymatic reaction was carried out on a preparative scale (10 µM enzyme, 14.2 mM L-Asp, 18 mM αKG in 10 mL buffer) at 16°C for 14 hours. The reaction was analyzed by HPLC, equipped with a chiral column to distinguish between possible *threo*- and *erythreo*-products and absorbance was measured at 220 nm and compared with authentic L-3-hydroxyaspartic acid standards (Tocris Bioscience, Ellisville, USA). The overlaid HPLC traces (Figure 5.26 A) of L-*erythro*-hydroxyaspartic acid (Figure 5.26 A, orange), L-aspartic acid (Figure 5.26 A, green), L-*threo*-hydroxyaspartic acid (Figure 5.26 A, red) and the enzyme assay (Figure 5.26 A, black) revealed that the only reaction product co-elutes with the L-THA standard. L-*erythro*-hydroxyaspartic acid was not observed. Integration of the starting material and product peaks evidenced 70% conversion of L-Asp to L-THA.

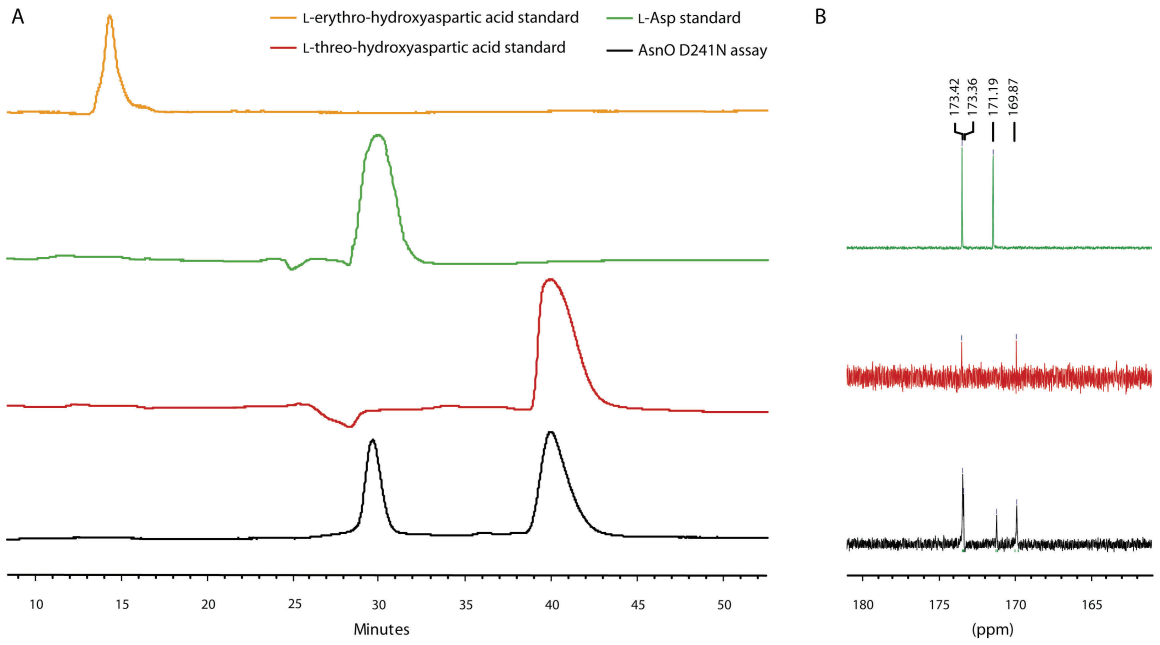


Figure 5.26. Analysis of preparative assay via chrial HPLC. A) The HPLC chromatogram shows the standards of L-*erythro*-hydroxyaspartic acid (orange), L-aspartic acid 2 (green), and L-THA 1 (red). The main product peak of the assay (black) co-elutes with L-THA and its integral revealed 70% conversion at 16°C. B) ^{13}C-NMR spectra of 1,4-^{13}C-labeled aspartic acid (green), unlabeled L-THA standard (red) and reproduction of the assay with 1,4-^{13}C-labeled 2 (black). The signal at 169.97 ppm proves that L-THA is generated and the ^{13}C-integral confirmed 70% conversion.

The identity and conversion from Asp to L-THA in the reaction mixture was validated by employing 1,4-^{13}C-labeled Asp (Cambridge Isotopes, Andover, USA) in the assay for NMR analysis. The labeled aspartic acid was used to observe specific ^{13}C-NMR signals for Asp and L-THA as well as to suppress interfering carbon signals from αKG co-substrate and succinate, its decomposition product. The NMR spectrum of the reaction mixture (Figure 5.26 B, black) shows that the two signals (δ = 173.36, 171.19 ppm) of L-Asp (Figure 5.26 B, green) are shifted towards the L-THA signals (Figure 5.26 B, red, δ = 173.42, 169.87 ppm). Integration of the ^{13}C signals revealed 70% conversion, in agreement with product quantification from the HPLC-assays. In this one-step enzymatic reaction, enantiomerically pure L-THA was isolated in 68% yield (14.8 mg, 0.1 mmol) following preparative HPLC and freeze-drying of the L-THA-containing fractions.

6. Discussion

6.1 β-Hydroxylated Acidic Lipopeptide Antibiotics Precursors

Natural products from traditional microbial sources have played an important role as valuable antibiotics and other drugs but interest in them waned in the 1990s when major pharmaceutical companies decided that their discovery was no longer cost-effective and instead concentrated on synthetic chemistry as a source of novel compounds.[23] With the development of molecular genetic methods to isolate and manipulate the complex microbial enzymes that produce natural products, unexpected chemistry has been revealed and interest in the compounds has again flowered.[13] An excellent example is the success story of the clinically approved daptomycin, which is produced by *S. roseosposrus* and is used for the treatment of skin infections and bacteremia caused by multi-drug resistant bacteria, including MRSA and VRE.[176] Daptomycin belongs to a new class of potent antibiotics. Members are beside the aforementioned daptomycin,[178] the complex A54145,[182] produced by *S. fradiae*, the calcium-dependent antibiotics (CDA),[170] produced by *S. coelicolor* and other actinomycete-produced antibiotics. The acidic lipopeptides contain various non-proteinogenic amino acids, which are often highly functionalized by the action of modifying or so called tailoring enzymes.[15,16] Little was known about the origin of the non-proteinogenic amino acids as well as the mechanisms underlying their generation when this study was initialized.[176] Furthermore the mode of action of these tailoring enzymes was not clear, as they could either functionalize an amino acid prior to, during or after the non-ribosomal assembly of the antibiotic.[185]

This study revealed the origin of the hAsn residues found in CDA and A54145. By biochemical characterization of AsnO and LptL, it was shown that these enzymes are non-heme iron(II)- and αKG- dependent dioxygenases, which are directly involved in 3-OH-asparagine precursor synthesis. It was possible to exclude an *in trans* mechanism by incubation of the tailoring enzyme with Asn bound to the cognate carrier protein. Furthermore, kinetic parameter and substrate specificity determination showed that AsnO and LptL are highly specific and fast hAsn precursor biocatalysts.

Remarkably, the crystal structure of AsnO facilitated the determination of the exact stereochemistry of the hAsn residue found in CDA. Additionally, it allowed prediction of

the substrate specificity of uncharacterized hydroxylases from other microbial sources, by comparison of the AsnO structure with homology models of putative oxygenases. In fact, the motivation to characterize LptL, was the result of these predictions as a proof of principle. Notably, it turned out to be correct.

In summary, the characterization of AsnO and LptL elucidated the mechanism of hAsn incorporation during CDA and A54145 assembly and together with 3D structural information from AsnO, it might facilitate the generation of novel, β-hydroxylated NRPS building blocks.

6.1.1 Biochemical Characterization of AsnO and LptL

Sequence Analyses. On the basis of sequence homology to the known iron(II) and αKG dependent oxygenases MppO,[113] VioC[111] and CAS[220] (Figure 6.1), the *asnO* and *lptL* gene were identified within the biosynthetic gene cluster of CDA and A54145 respectively as putative hydroxylases involved in hAsn formation.

```
AsnO ------------MAANAAGPASRYDVTLDQSDAELVEEIAWKLATQATGRPDDAEWVEA  47
LptL ------------MEPENT-------FTLSEAERDDVAALAQELTRARPGLVDEREWLDR  40
MppO ----------------------MLTLHLQDDDVAAIDAVADELS-RRYDSVESTEFQAE  36
VioC MTESPTTHHGAAPPDSVATPVRPWSEFRLTPAEAAAAAALAARCA-QRYDETDGPEFLLD  59
CAS  --------------------------MTSVDCTAYGPELRALAARLPRTPRADLYAFLDA  34
                                 .                 *          :   :

AsnO ARNAWHAWPATLRRDLAGFRRDSGPDGAIVLRGLPVDSMGLPPTPRVN--GSVQREASLG  105
LptL CRTLSCHLPARLQDRLRAFRHDPGPTGRLLLRNLPAADS-VPATPREP--DSVERRATLS  97
MppO SRLYADELPRRVRRALHEYRSTEKS-GILVVTGLPVDDSALGATPADRRHKPVPSTSLRQ  95
VioC APVIAHELPRRLRTFMARARLDAWP-HALVVRGNPVDDAALGSTPVHWRTARTPG-SRPL  117
CAS  AHTAAASLPGALATALDTFNAEGSEDGHLLLRGLPVEADADLPTTPSSTPAPEDRSLLTM  94
     .        *   :    :    .          ::: . *.       .*.

AsnO AAVLLMTACGLGDPGAFLPEKNGALVQDVVPVPGMEEFQGNAGS-TLLTFHNENAFHEHR  164
LptL ASVLCALSMELGDVIAYRNEKQGALVQNVVPVPGREGQQSNAGS-VPLEMHTENAFHPHR  156
MppO DIAFYLIANLLGDPIGWATQQDGFIMHDVYPVQGFEHEQIGWGSEETLTWHTEDAFHPLR  155
VioC SFLLMLYAGLLGDVFGWATQQDGRVVTDVLPIKGGEHTLVSSSSRQELGWHTEDAFSPYR  177
CAS  EAMLGLVGRRLGLHTGYRELRSGTVYHDVYPSPG-AHHLSSETSETLLEFHTEMAYHRLQ  153
        :    .   **    .:    :.* :  :* *  *         .   *    *   *.* *:    :

AsnO PDFVMLLCLRADPTGRAGLRTACVRRVLPLLSDSTVDALWAPEFRTAPPPSFQLSGPEE-  223
LptL PDYVGLFCVRSDHDRAAGLRVASVRAVMDHLDAGTREMLRQPLFTTEPPPSFGRPDSGT-  215
MppO TDYLGLMCLRNP----DGVETTACDIADVEIDDETRETLSQERFRILPDDAHRIHGKAPG  211
VioC ADYVGLLSLRNP----DGVATTLAGVPLDDLDERTLDVLFQERFLIRPDDSHLQVNNSTA  233
CAS  PNYVMLACSRADHERTAATLVASVRKALPLLDERTRARLLDRRMPCCVDVAFRGGVDDPG  213
     .:::  *  .  *         .   .:          :.   *    *    :       :.

AsnO ------------------------APAPVLLGDRSDPDLRVD--LAATEPVTERAAEALR  257
LptL ------------------------KPHAVLTGDAEDPDIRVD--FHATHTSDPWGRQAME  249
MppO DESARESALRERSRQRVASALESPDPVAVLFGDRDDPYLRIDPHYMQGVQG-ETEQRALE  270
VioC QQGRVEFEG-------IAQAADRPEPVAILTGHRAAPHLRVDGDFSAPAEGDEEAAAALG  286
CAS  ----------------------AIAQVKPLYGDADDPFLGYDRELLAPEDP--ADKEAVA  249
                             * *.   *  :  *                  *:

AsnO ELQAHFDATAVTHRLLPGELAIVDNRVTVHGRTEFTPRYDGTDRWLQRTFVLTDLRRSRA  317
LptL ALAEAVRTVSEELVLEPADLVYVDNRVALHGRTAFVPRYDGQDRWLQRAFVHLDHRRSRA  309
MppO TIGAAIDDAMSGVVLSPGDIVFIDNYRVVHGRKPFRARFDGTDRWLRRLNIARDLRKSRE  330
VioC TLRKLIDASLYELVLDQGDVAFIDNRRAVHGRRAFQPRYDGRDRWLKRINITRDLHRSRK  346
CAS  ALSKALDEVTEAVYLEPGDLLIVDNFRTTHARTPFSPRWDGKDRWLHRVYIRTDRNGQLS  309
      :   .        *   .::   :**   . *.*   *  .*:** ***:*   :   *  .  .
```

Figure 6.1: ClustalW alignment of AsnO, LptL with known non-heme iron oxygenases. The conserved iron binding HXE…H motif (shaded in red) as well as two Arg were found 14/18 residues after the second His, thought to responsible for αKG-binding (shaded in green).

AsnO and LptL contain the conserved iron binding motif HXE…H and the conserved arginine residues thought to be responsible for co-substrate αKG binding.[14] They perfectly align with sequence homology of 40% or more with the known free amino acid precursor generating enzymes MppO and VioC. Furthermore they are similar to CAS (clavaminate synthase from *S. clavuligerus*), thus they are likely to belong to the CAS-like non-heme iron and αKG dependent oxygenases subclass rather than to the *E. coli* taurine catabolism enzyme TauD-like subclass of this enzyme family.[228]

Catalysis. When an excess of the ferrous iron co-factor was added (see iron content determination 5.1.2), the two enzymes were indeed capable to functionalize the unreactive β-position of asparagine by hydroxylation (Figures 5.3 and 5.9). The following mechanism for this reaction can be proposed (Figure 6.2):

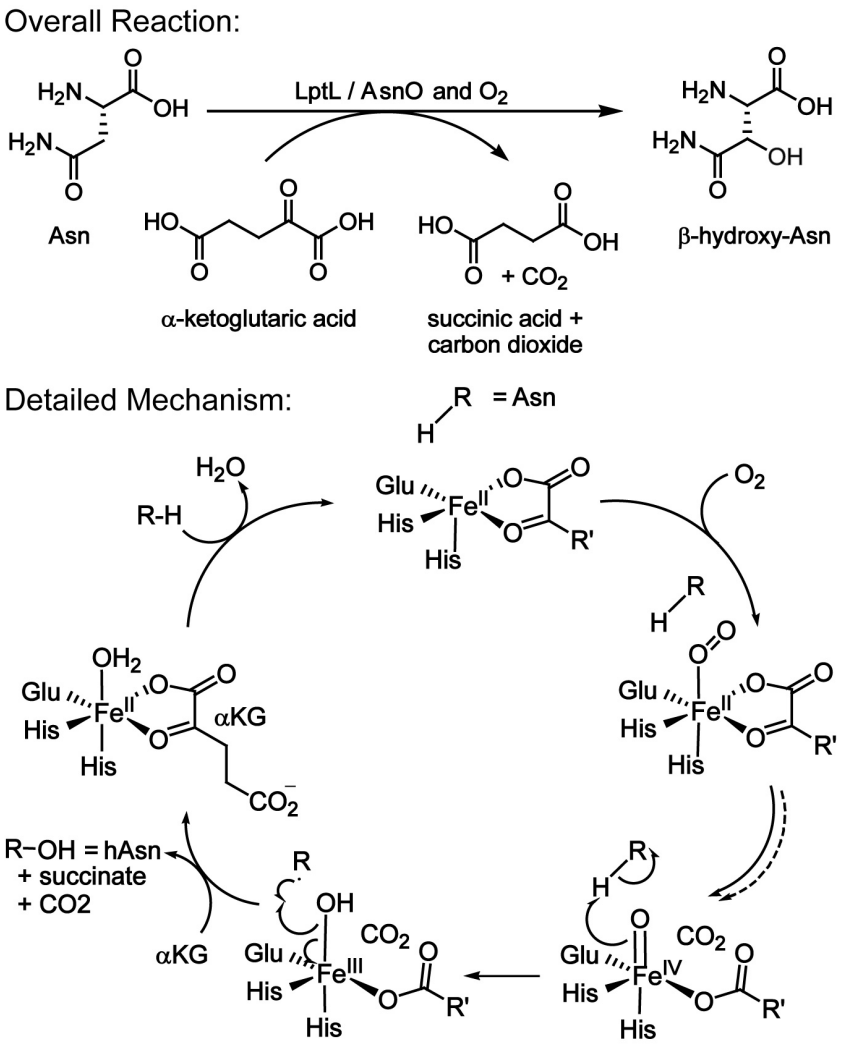


Figure 6.2: Proposed reaction mechanism for AsnO and LptL catalyzed hAsn formation.

Like other members of the iron(II)/αKG-dependent oxygenases,[147] AsnO and LptL decarboxylate α-ketoglutarate by utilizing dioxygen to generate carbon dioxide, succinate and a reactive oxidizing species that mediates substrate oxidation (Figure 6). One of the oxygen atoms from O_2 is incorporated into succinate and the other into the alcohol of the product. It is proposed that after substrate binding iron(II) and oxygen react to give a Fe(III)-superoxo species. The anion subsequently attacks the 2-ketogroup of αKG, which is activated by the Lewis acidity of the iron. This reaction leads to a persuccinate species, and after its collapse results in a potent Fe(IV)=O species, which in turn abstracts a hydrogen radical from the substrate asparagine β-CH_2 group. This results in a close proximity of the Fe(III)-OH and the substrate radical in the active site. Radical recombination involves $OH^{\cdot}$ transfer to the substrate radical, yielding the observed alcohol product and regenerating the starting iron(II)-oxidation state (Figure 6.2). The key to the catalytic strategy is the generation of the high-valent oxoiron (Fe(IV)=O) species that removes a hydrogen atom from even the most unactivated carbon sites of a bound substrate. [156]

Kinetics. It was possible to determine the kinetic parameters for AsnO and LptL-mediated hydroxylation by FMOC-Cl quenching and derivatization of the Asn starting material and the hAsn product. The results together with parameters obtained for different small molecule modifying oxygenases are presented in Table 6.1:

Enzyme	Substrate	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($min^{-1}\mu M^{-1}$)	Source
AsnO	asparagine	479	299	0.62	this study
LptL	asparagine	584	531	0.91	this study
VioC	arginine	3400	2611	0.77	Helmetag[111]
CAS	(2*S*,3*S*)-pro-clavaminate	190	45.9	0.24	Salowe[229]
PtlH	1-deoxypentalenic acid	570	252	0.44	You[230]
RdpA	2-(4-chloro-2-methyl-phenoxy)-propionic acid	380	252	0.66	Müller[231]

Table 6.1: Kinetic parameters of AsnO, LptL and other small molecule hydroxylating enzymes.

The general trend of this enzyme class is that the higher the affinity (low K_M) to the substrate is, the slower the enzyme is. In the case of the clavaminate synthase, resembling the smallest K_M value with 190 μM only 46 molecules are converted within a minute, whereas in the case of VioC, the arginine β-hydroxylase involved in viomycin synthesis, with the K_M being in the 3 mM range, over 2600 substrate molecules are

converted. All enzymes have been characterized *in vitro*, therefore these kinetic parameters have to be taken cautious as they might not represent the *in vivo* situation. However, a general trend is present, i.e. that these small molecule oxygenases are relatively fast compared to NRPS enzymes (k_{cat} approximately 1/min) and that they are all similarly efficient. With a k_{cat}/K_M value of 0.91 $min^{-1}\mu M^{-1}$ LptL was found to be the most efficient known small molecules modifying oxygenase known so far *in vitro*. By being that efficient and fast, the NRPS related enzymes AsnO, LptL and VioC[111] as well as the PKS related enzyme PtlH[230] grant a constant supply of the corresponding substrate for the molecular assembly line. Although these oxygenases are relatively fast compared with the NRPS, they are still much slower than enzymes involved in the primary metabolism. Therefore the oxygenases are unlikely to interfere with the primary metabolism by snatching essential amino acids or precursors.

Substrate Specificity and Incorporation Mechanism. AsnO and LptL were both found to have a stringent substrate specificity. They only catalyze the conversion of Asn to hAsn, and neither other free amino acids nor the corresponding D-isomer of Asn are accepted as substrates (Tables 5.1 and 5.3). They both exclusively modify the free amino acids, whereas carrier bound asparagine was not hydroxylated. Therefore, they are both fast and efficient precursor generating enzymes. This finding was somewhat unexpected as AsnO and LptL were predicted to work on the PCP-tethered Asn.[185]

Based on these results, for AsnO the following mechanism of hAsn generation, activation and incorporation into the growing CDA peptide chain could be postulated:

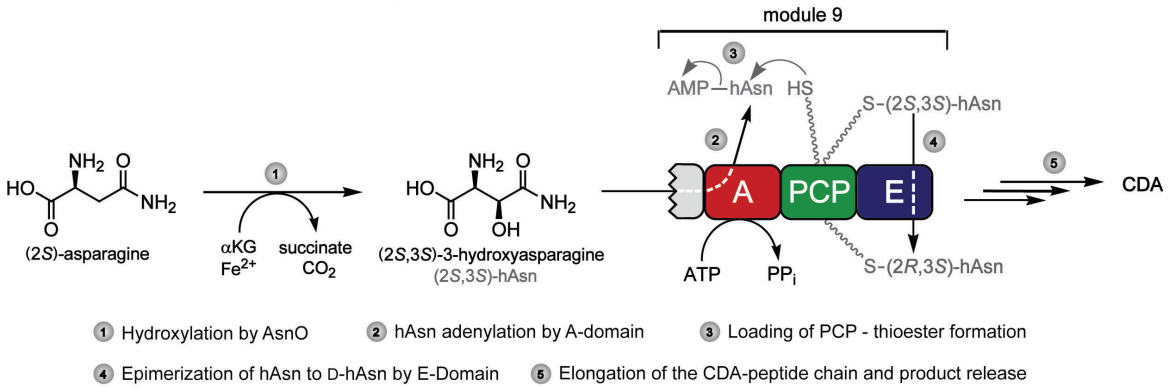


Figure 6.3: Reaction sequence scheme for hAsn generation, activation, epimerization and incorporation during the biosynthesis of cyclic CDA peptides bearing hydroxylated asparagines at position 9.

After activation of the amino acid as hAsn-AMP, a thioester with the sulfhydryl-group of the PCP-cofactor 4'phosphopantheteine could be formed. The PCP domain then would transfer hAsn to the E domain, where it is converted from the 2*S*,3*S* to the 2*R*,3*S* configuration. After that, the latter one would be accepted by the neighboring condensation domains, which catalyze the peptide bond formations. Subsequent chain elongation and product release would yield naturally occurring CDAs with hydroxylated D-Asn-moieties at position nine of the peptide backbone (Figure 6.3). Interestingly, in mature CDA only D-hAsn, which is generated by the epimerization (E) domain of the ninth module (Figure 6.3), is found. Although the recombination of the cognate adenylation domain turned out to be difficult, one may postulate that the A domain of the ninth module preferably accepts hAsn but is also capable for L-Asn activation with reduced efficiency. The latter is due to the fact, that derivatives of CDA with Asn instead of hAsn at the canonical position were isolated from Δ*asnO*-deletion mutant strains with decreased CDA production.[186]

For the origin of the hAsn residue found at position three of the A54145 backbone, the following mechanism is reasonable:

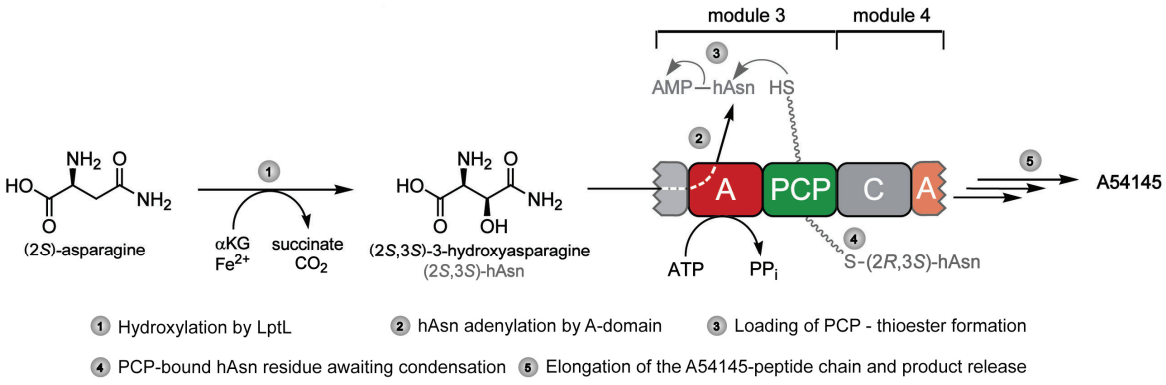


Figure 6.4: hAsn generation, activation and incorporation scheme during A54145 biosynthesis.

LptL directly hydroxylates the chemically unreactive β-position of Asn. The adenylation domain of the third module of the first A54145 NRPS then should activate the hAsn as AMP, which subsequently could be attacked by the cognate PCP domain. The PCP-tethered hAsn is presented afterwards to the third condensation domain, and the formed tripeptide bound to the carrier protein should be further elongated downstream. After elongation is completed, the tridecapeptide is transferred to the TE domain, and mature A54145s with hAsn at position 9 are cleaved off by cyclization.

A Domain Analysis. Again, the recombinant production in *E. coli* of the *Streptomyces'* adenylation domain (module three of the A54145 NRPS), only yielded insoluble protein. The alignment of the primary sequence with PheA,[39] the adenylation domain used for the determination of the specificity-conferring code of adenylation domains in non-ribosomal peptide synthetases,[44] revealed a similar binding pocket composition compared with the ninth A domain of the CDA cluster (Table 6.2). The latter is likely to activate the hAsn as amino acyl AMP, which is discussed above.

		Residue numbers, referring to the corresponding position in PheA									
Amino acid	A domain	235	236	239	278	299	301	322	330	331	517
Phe	PheA	D	A	W	T	I	A	A	V	**C**	K
Asn	BlmX1	D	L	T	K	V	G	E	V	**G**	K
Asn	BlmVI2	D	L	T	K	V	G	E	V	**G**	K
Asn	LptC4	D	L	T	K	V	G	D	V	**S**	K
hAsn	LptA3	D	L	T	K	V	G	D	V	**N**	K
hAsn	CDA9	D	L	T	K	V	G	E	V	**N**	K

Table 6.2: Specificity residues for Asn and hAsn adenylation domains. The Blm A domains are from the *S. verticillus* bleomycin cluster and LptC4 is from the 11th A domain of the A54145 cluster. Position 331, located above the β-position of the amino acid within the binding pocket is highlighted.

Generally, as highlighted in bold letters in Table 6.2, usual Asn-activating A domains contain either a small side chain at position 331, as in the case of LptC4 or no side chain at all is found, as observed in the bleomycin[232] Asn A domains. The presumably hAsn activating A domains from the A54145 and CDA clusters contain an asparagine residue at this position, which could stabilize the additional oxygen at the β-position. This hypothesis is supported by the PheA structure, the prototype A domain involved in gramicidin S biosynthesis.[44] The structure elucidation of the Phe A domain of GrsA allowed the identification of the residues, which provide the substrate amino acid binding pocket, and led to the formulation of the specificity-conferring code of adenylation domains. The PheA structure (Figure 6.5 A) revealed e.g. that all amino acid A domains known hitherto contain aspartic acid at position 235 (relative to PheA sequence) and lysine at 517. These are responsible for the binding of the α-amino and α-carboxy function.

The structure of PheA (Figure 6.5 A) was manipulated *in silico* and the substrate Phe was replaced by Asn. Furthermore, the cysteine residue 331 was mutated to asparagine (C331N, Figure 6.5 B), which is the case in the likely hAsn activating A domains CDA9 and LptA3 (see Table 6.2).

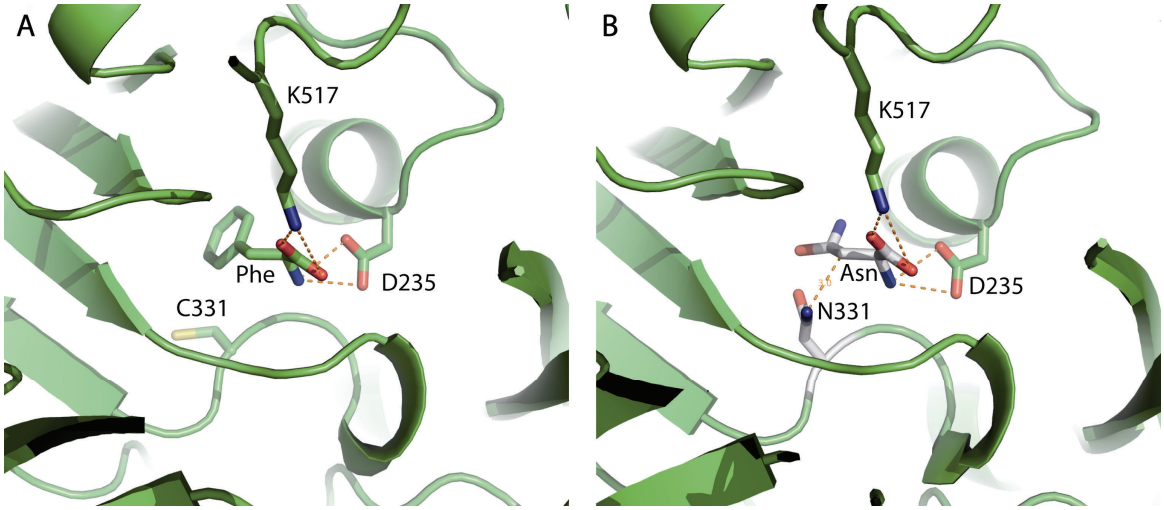


Figure 6.5: *In silico* mutation of the PheA active site. A) Active site of PheA. B) Active site where the substrate is Asn instead of Phe and the C331 is mutated to N331. In the C331N mutant model the distance between the NH_2-group of the carboxamide side chain of residue N331 and the β-position of the substrate Asn is only 3 Å, thus allowing tight interaction if the Asn β-position would be hydroxylated. The mutated residues are shown in grey.

In the model (Figure 6.5 B), the NH_2-group of the side chain of N331 is well positioned for β-hydroxy coordination. Normal Asn activating A domains usually have a glycine or serine at this canonical position. Therefore, the above depicted model supports the hypothesis that hAsn rather than Asn activation is likely to occur for LptA3 and CDA9 A domains, which further supports the postulated hAsn flow schemes (Figures 6.3, 6.4).

6.1.2 Crystal Structure of AsnO

General. The observed iron binding motif of the residues H155, E157 and H287, which projects from the rigid β-barrel core of the iron(II)/αKG-dependent oxygenases, fulfills a role similar to that of the porphyrin ring in cytochrome P450 enzymes, which is a stable platform to ligate the iron and the subsequently formed reactive intermediates.[139] The presence of a jelly roll topology[141] in AsnO like those in CAS, IPNS, and DAOCS, in which active site residues are located on analogous strands of the core domain,[14] implies a divergent evolutionary relationship between AsnO and other members of the extended family of αKG/iron(II)-requiring and related enzymes, e.g. the halogenase SyrB2.[162]

Substrate Binding. As found by a BLAST search with the entire AsnO sequence as the query, AsnO shares high sequence homology with several putative as well as biochemically characterized non-heme iron(II)/αKG oxygenases (Figure 6.1), which participate in the biosynthesis of NRPS products. On the basis of the AsnO crystal

structure of the product bound state, five proteins exhibiting the greatest similarity were selected to generate 3D-homology models with Modeller 7 (not shown).[233] These are the putative proteins SCO2693 (57% identity),[234] produced by *S. coelicolor*, LptL (45%) from *Streptomyces roseosporus*, TobO[235] (37%) from *Streptomyces sp.* DSM 40477 and the biochemically examined enzymes VioC[112] (33%) as well as MppO[113] (31%) produced by *Streptomyces vinaceus* and *Streptomyces hygroscopicus*, respectively. Strikingly, most of the α-carboxy and α-amino group binding residues are conserved in the 3D-homology models for all five proteins (Figure 6.6).

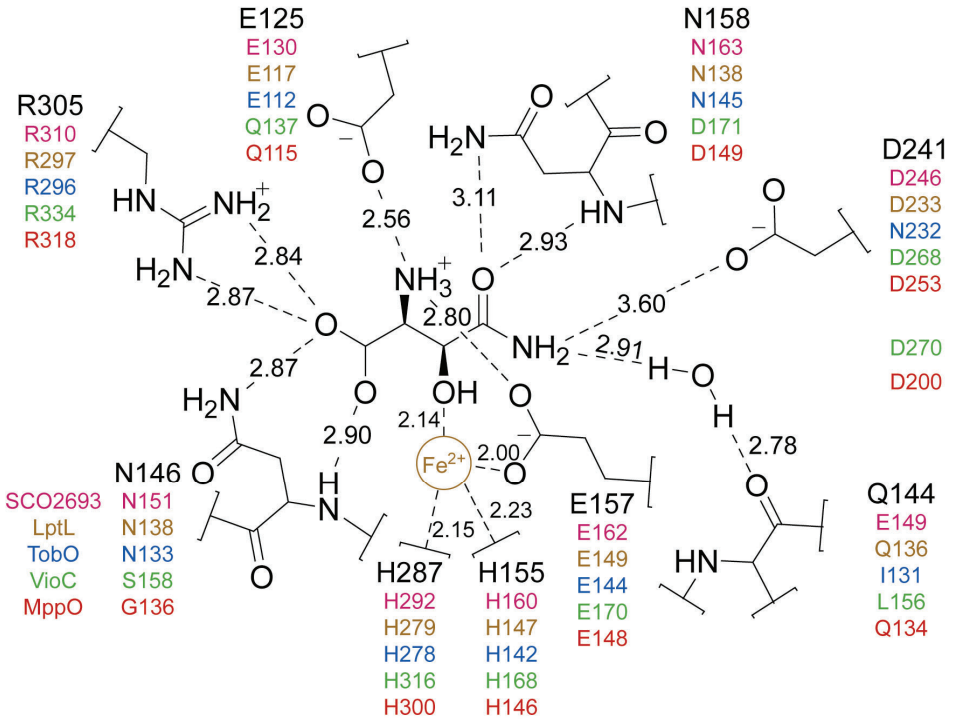


Figure 6.6: Schematic representation of hAsn-AsnO interactions and homologous enzymes. The (2*S*,3*S*)-3-hydroxyasparagine interactions with the protein scaffold are represented. 3D-Homology models based on the AsnO product complex revealed the substrate binding residues of SCO2693 (magenta), LptL (brown), TobO (blue), VioC (green) and MppO (red). These are stated below the corresponding AsnO residues. Additional putative substrate binding residues of VioC and MppO can be found on the right side.

In the case of SCO2693 and LptL (Figure 6.6, residues in magenta and brown), all residues are conserved, except Q144, which is a glutamate in SCO2693. As this residue is only indirectly involved in substrate binding via its peptide group, one can predict that both enzymes will also use L-Asn as a substrate for the generation of (2*S*,3*S*)-3-hydroxyasparagine. For LptL, this finding perfectly agrees with the prediction that it is involved in the hAsn precursor synthesis for the lipopeptide antibiotic A54145, which bears a hAsn-moiety at position three of the peptide backbone.[183]

Indeed, this prediction/hypothesis was the motivation to clone, express and biochemically characterize LptL. As shown above, the substrate specificity prediction was correct. LptL was found to exclusively hydroxylate Asn at the β-position as predicted (chapter 5.1.4).

TobO, which is part of the aminoglycoside tobramycin biosynthesis cluster, may encode for a protein involved in the synthesis of hydroxylated precursors or inhibitors for this antibiotic. At current knowledge, nothing is known about a possible substrate. Compared with AsnO, TobO (Figure 6.6, residues in brown) notably differs only in the acidic D241 residue, which is changed to N232. Thus, we predict that TobO catalyzes the oxygenation of substrates with an acidic side chain, such as Asp or Glu. This prediction was indirectly proven by the generation of the AsnO D241N variant (see 5.3, 6.3).

It was shown that VioC selectively catalyzes the formation of (2*S*,3*S*)-3-hydroxyarginine with L-Arg as a substrate during viomycin biosynthesis.[112,236] The model (Figure 6.6, residues in green) revealed that Q137, which corresponds to α-amino group binding E125 in AsnO, rather binds the α-carboxy group and the side chain of D171, which is positioned analogously to the carbonyl-group binding N158 residue of AsnO, the α-amino group. The D241 residue is conserved (D268) and could establish a salt bridge with the guanidinium group of the L-Arg substrate. Furthermore an additional acidic residue (D270) in close proximity to the substrate may stabilize the basic arginine side chain. Interestingly after the AsnO crystal structure was published in 2007, the crystal structure of VioC[111] was solved in 2009 and the predicted residues were indeed involved in substrate binding except for D171, therefore the predictions made on the AsnO structure were proven again to be correct to a great extent. Even the different binding mode, leading to *erythro* products of VioC rather than *threo* as in the case of AsnO, was predicted correctly. This exemplifies the prototype character of the AsnO structure for amino acid hydroxylating enzymes.

Finally, MppO,[113] catalyzes the hydroxylation of the arginine derived unnatural amino acid enduracididine at the β-position to generate (2*S*,3*S*)-3-hydroxyenduracididine building blocks for the biosynthesis of the potent glycopeptide antibiotics mannopeptimycins.[115] The MppO model (figure 6.6, residues in red) showed high similarity to VioC and the α-carboxy and α-amino group should be bound identically. The

side chains of D253 (corresponding to D241 in AsnO) and of the additional acidic residue D200 should bind the cyclic guanidinium group and thereby stabilize enduracididine in the substrate binding pocket. Furthermore, the second reaction product, succinate, is bound by AsnO via a salt bridge to the guanidinium group of the conserved residue R301 that is similarly formed in other structures of α-KG/Iron(II)-dependent oxygenases complexed to α-ketoglutarate. This arginine is usually found 14-22 residues after the conserved second His-residue of the iron binding motif. In all 3D-homology models, this arginine is highly conserved, located exactly 14 residues after the second iron binding histidine.[14,141]

The high number of specific interactions between the protein scaffold of AsnO and L-Asn rationalizes why the enzyme exclusively catalyzes the hydroxylation of L-Asn and why the product has to be 2*S*,3*S*-configured. The 3D-homology models based on the product complex of AsnO share this high quantity of specific interactions with their substrates and explain the observed 2*S*,3*S*-stereoselectivity for the hydroxylation product in the case of MppO. Furthermore, AsnO has archetype character for putative non-heme iron(II)/αKG dependent oxygenases that are involved in the generation of β-hydroxylated amino acid building blocks for NRPS products as exemplified by 3D-homology models of SCO2693, LptL and VioC.

Substrate Induced Fit. At current knowledge, the observed substrate induced-fit mechanism, explaining the high substrate specificity of AsnO, was not observed before in the clavaminate synthase like superfamily of iron(II)/αKG-dependent oxygenases. A similar lid closure was observed recently in the case of VioC.[111] Sterically demanding peptidyl-Asn substrates could either not bind to the active site or trigger the lid closure. Thus, it could be concluded that these compounds are incapable of accessing the active site and only the free amino acid will be hydroxylated. That is quite the contrary to factor inhibiting HIF-1 (FIH-1),[237] which hydroxylates a conserved asparaginyl residue in the master switch of cellular hypoxia responses, hypoxia-inducible factor 1 (HIF-1), under normoxia, which suppresses transcriptional activity of HIF-1.[238] In FIH-1, the substrate binding site is a large groove (15 Å wide and 40 Å long), providing enough space for the binding of a large α-helical region of HIF-1.[237]

6.2 *In Trans* Acting Hydroxylases and Regeneration Domains during NRP assembly

Extraordinary examples of structural diversity and the occurrence of non-proteinogenic amino acids are the antifungal and antimicrobial kutznerides 1-9 (Figure 2.23).[18,19] All kutznerides contain one α-hydroxy acid and five non-proteinogenic amino acids, one of the latter is either the *threo* or *erythro* isomer of 3-hydroxy-D-glutamic acid (D-hGlu). Evaluation of KtzO and KtzP showed that there are two on-line hydroxylating enzymes in kutzneride biosynthesis, which stereospecifically catalyze the formation of PCP-*S*-L-*threo*-hGlu (KtzO) and PCP-*S*-L-*erythro*-hGlu (KtzP) prior to epimerization by the E domain found in KtzH.

The transfer of the glutamic acid onto the third PCP domain of KtzH was investigated. The ability of the stand-alone A domain KtzN to reconstitute the non-functional A* domain of KtzH *in trans* was examied. Thus, this study represents the first characterization of an *in trans* acting adenylation domain, which simultaneously restores the ability of an NRPS assembly line to produce a secondary metabolite.

To determine the kinetic parameters of KtzO and KtzP catalyzed hydroxylations, the problem of the labile thioester bond of A*PCP_3-*S*-Glu was circumvented by using synthetic coenzyme A analogs coupled to glutamic acid where the labile thioester was replaced by a hydrolytically stable amide bond. This method is a simple and robust way to investigate carrier proteins, which have substrates bound via a prosthetic ppan arm. Finally, KtzO and KtzP exclusively hydroxylate PCP-bound glutamic acid, but do not require the cognate PCP of module three of KtzH for *in vitro* activity. If the NRPS on-line modification by non-cluster enzymes can be transferred to *in vivo* applications, the insights from this study might assist in future reengineering approaches of NRPS gene clusters in order to generate new secondary metabolites.

The results obtained for KtzO and KtzP complete the knowledge about β-hydroxylated building blocks found in natural products. Interestingly, in nature two different mechanisms evolved in non-ribosomal peptide synthesis for the generation of these building blocks, namely precursor synthesis and *in trans* modification, as exemplified by AsnO/LptL and KtzO/KtzP, respectively.

6.2.1 Biochemical Characterization of KtzO and KtzP

Hydroxylation Activity and Stereospecificity. As observed in the iron content determination for AsnO and LptL, the hydroxylases KtzO and KtzP either lose their ferrous iron co-factor during purification or it gets exchanged by other divalent ions. Thus, for the activity of KtzO and KtzP, the addition of iron(II), was necessary for hydroxylation activity. KtzO and KtzP were both able, when incubated with the co-substrate αKG, and cofactor Fe^{2+} and substrate, to selectively hydroxylate the β-position of PCP bound glutamic acid *in vitro*. This finding was somewhat expected as both share high sequence homology with the known PCP-*S*-Asp β-hydroxylase SyrP.[126] SyrP and KtzO/KtzP are the only known enzymes able to exclusively hydroxylate PCP bound amino acids. Unfortunately, for none of these enzymes 3D structural information is available. But models of these[233] enzymes based on 2.4 Å resolution crystal structure of the protein from gene locus At3g21360, a putative iron(II)/αKG-dependent enzyme from *Arabidopsis thaliana*,[239] suggested a thiotemplated mechanism (Figure 6.7).

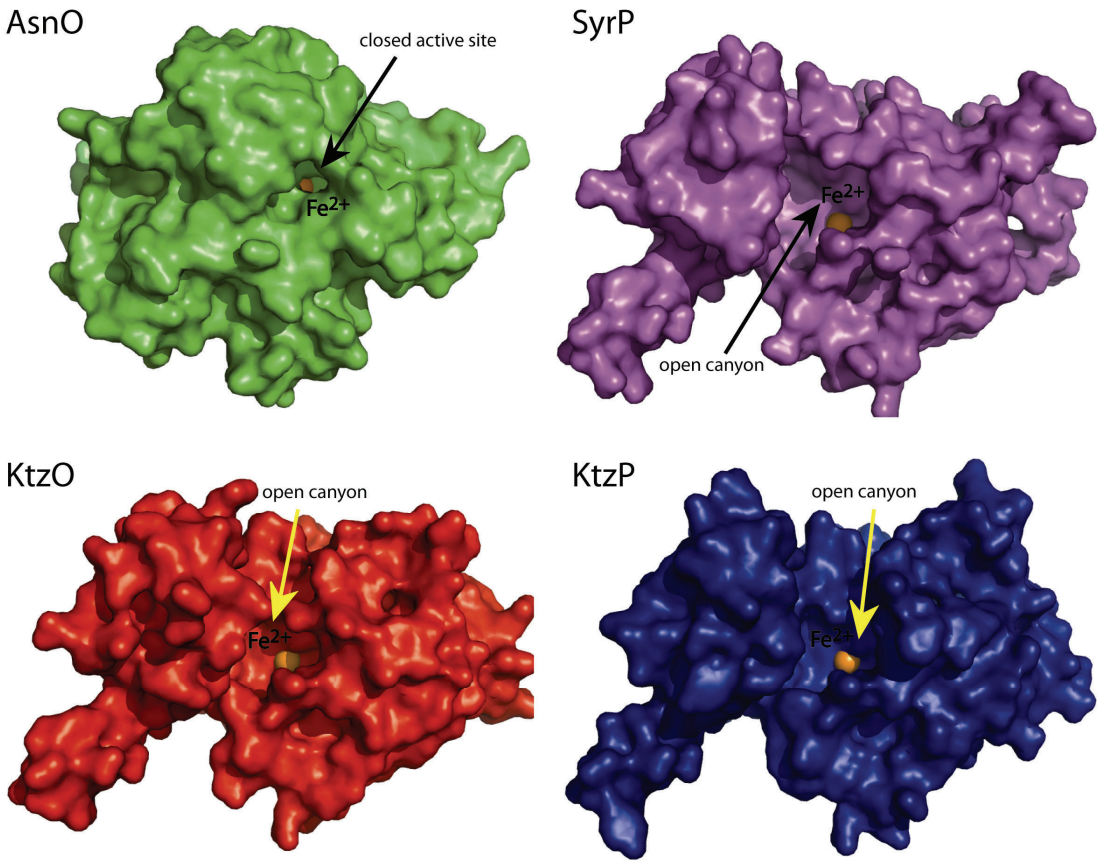


Figure 6.7: Comparison of the AsnO structure with structure models of SyrP, KtzO and KtzP. The open canyon towards the ferrous iron cofactor of SyrP, KtzO and KtzP suggest PCP-tethered substrates. AsnO on the contrary has a closed active site, burying the free amino acid substrate completely.

The active site of AsnO is completely buried within the enzyme structure (Figure 6.7 A), whereas in the SyrP, KtzO and KtzP 3D models (Figure 6.7 B-D) the active site is open and a large canyon provides enough space for the amino acid substrate bound to the bulky ppan-arm of the corresponding PCP domain. Although these are just models, they give broad hints that KtzO and KtzP would accept PCP-tethered amino acids, as it was shown for SyrP. The observed reactivity of the kutzneride hydroxylases met these expectations.

On the other hand, the finding that both enzymes were found to be stereospecific was unexpected. One would expect from the kutzneride structures, where the *threo* and *erythro* 3-hydroxyglutamic acids are found in similar amounts,[18,19] that it would be more economical for the kutzneride NRPS to have only one enzyme for Glu hydroxylation. This enzyme could catalyze β-hydroxylation of PCP-*S*-L-Glu, providing a racemic mixture in a kind of epimerization mechanism. Instead, evaluation of KtzO and KtzP shows there are two on-line hydroxylating enzymes in kutzneride biosynthesis, which stereospecifically catalyze the formation of PCP-*S*-L-*threo*-hGlu (KtzO) and PCP-*S*-L-*erythro*-hGlu (KtzP). This system is the first example for a NRPS system where products with two isomers of a C_β modified amino acid are generated by two modifying enzymes that catalyze the very same reaction but with opposite stereospecificity.

Substrate Specificity and Kinetics. The free amino acids were not accepted as substrates by KtzO or KtzP and therefore a precursor synthesis pathway can be excluded (see Table 5.5). KtzO and KtzP were found to be highly specific for hydroxylation of PCP-bound L-glutamic acid; PCP-*S*-D-Glu was not a substrate. The D-hGlu moieties as observed in mature kutznerides are thus likely to be generated exclusively by the action of the epimerization domain, which is present in the third module of KtzH.[158]

An intriguing finding was that KtzO and KtzP exclusively hydroxylate PCP-bound glutamic acid, but do not require the cognate PCP of module three of KtzH for *in vitro* activity, as shown by the hydroxylation of CDA-PCP_9-*S*-L-Glu (Figure 5.19). These findings make *ktzO* and *ktzP* valuable targets for gene transfer into NRPS gene clusters, which produce natural products with a glutamic acid moiety. A strong overproduction of KtzO and KtzP within these clusters could yield in NRPS products containing hGlu moieties, which might improve the bioactivity of these natural products. Given that these *in vitro* results could be transferred to *in vivo* applications, the insights from this study might assist in future reengineering approaches of NRPS gene clusters.

The kinetic parameter determination was carried out using non-hydrolyzable CoA analogs for cognate and non-cognate PCP domains. One characteristic of these hydroxylases is the high affinity to their substrate, which is shown by the relatively low K_M values (KtzO: 121.5 µM, KtzP: 168.9 µM) for the cognate-PCP compared with free amino acids modifying enzymes (AsnO: 479 µM, VioC[111]: 3400 µM) (Table 6.3).

Substrate	Enzyme	K_M [µmol]	k_{cat} [min^{-1}]	k_{cat}/K_M [$min^{-1}µmol^{-1}$]
A*PCP_3-H*N*-ppan-Glu	KtzO	121.5 ± 28.1	0.34 ± 0.03	$2.8 \cdot 10^{-3}$
A*PCP_3-H*N*-ppan-Glu	KtzP	168.9 ± 44.0	0.44 ± 0.05	$2.6 \cdot 10^{-3}$
CDA-T9-H*N*-ppan-Glu	KtzO	123.3 ± 10.9	0.33 ± 0.01	$2.7 \cdot 10^{-3}$
CDA-T9-H*N*-ppan-Glu	KtzP	148.5 ± 21.1	0.42 ± 0.02	$2.8 \cdot 10^{-3}$
asparagine	AsnO	479 ± 6.7	298.8 ± 19.2	0.62
arginine	$VioC^{111}$	3400 ± 450	2611 ± 196	0.76

Table 6.3: Comparison of the kinetic parameters of KtzO and KtzP-mediated hydroxylation of PCP-bound glutamic acid with free amino acid modifying enzymes.

Other features are their low turn-over numbers (KtzO: 0.34 min^{-1}, KtzP: 0.44 min^{-1}) compared with free amino acid hydroxylases (AsnO: 299 min^{-1}, $VioC^{111}$: 2611 min^{-1}) and the low catalytic efficiencies (Table 6.3). However, the catalytic efficiencies of KtzO and KtzP are identical, which explains why both isomers of the hydroxylated glutamic are found at equal amounts in the mature kutznerides, and only differ in the grade of substitution of other residues (Figure 2.23). Although the kinetic parameters suggest that the kutzneride hydroxylases are relatively slow and inefficient enzymes compared to other members of this class, it is reasonable that they are not as fast as free amino acid modifying enzymes for several reasons: First, the hydroxylases have to interact with the NRPS KtzH, second, they have to recognize readily loaded PCP domains (see below) and third, the concentrations of NRPS modules *in vivo* are relatively low, thus there is no need for the hydroxylases to act fast. In the case of free amino acid modifying hydroxylases, the enzymes have to be fast as they are in a constant battle about their substrates with the primary metabolism, which is not the case for KtzO and KtzP.

For the non-cognate PCP domain from the CDA cluster the kinetic parameters for KtzO and KtzP were found to be in the same range as for the cognate PCP (Table 6.3). These results give another broad hint that *ktzO* and *ktzP*, if introduced into a NRPS gene cluster, of which the gene products catalyze the synthesis of glutamic acid-containing natural products, could yield hGlu containing natural products.

6.2.2 *In Trans* Regeneration of the Third Module of KtzH by KtzN

The third A domain, A*, of KtzH, which only has 20% the size of a regular A domain[17,21,34] is non-functional, thus KtzN was predicted to be the first example for specific NRPS assembly line *in trans* recovery, which most likely involves specific protein-protein interactions between the third module of KtzH and KtzN. Without such specific protein-protein interactions, it would be likely that KtzN would load all four PCPs of KtzH with glutamic acid. KtzN was shown to activate L-glutamic acid in ATP/PP_i exchange assays (5.2.4). Furthermore, it was demonstrated that hGlu was not a substrate for KtzN-mediated adenylation, hence excluding a precursor mechanism. The transfer of the adenylated glutamic acid onto the PCP of the third module of KtzH revealed what protein-protein interactions are required for Glu transfer by KtzN and gave more mechanistic insight into this unusual NRPS recovery strategy. These results of the KtzN assays with holo-A*PCP_3, or holo-PCP_3 alone respectively, as well as in the absence/presence of KtzO or KtzP demonstrated that two factors are critical for KtzN-mediated transfer of glutamic acid onto the third PCP domain of KtzH *in vitro:* the presence of one of the two hydroxylases and the presence of A*, the truncated A domain in KtzH. These requirements explain the selective transfer of Glu to the third module of KtzH by KtzN during the *in trans* compensation/regeneration mechanism. Although *in trans* loading, even multiple times, is not unusual for polyketide synthase systems,[31] KtzN is the first biochemically examined example for such a reaction in non-ribosomal peptide synthesis and thus the first *in trans* acting NRPS recovery domain.

6.2.2 Mechanism of the Origin of the hGlu Residues Found in Kutznerides

From this biochemical characterization of KtzO and KtzP, in addition to the studies of KtzN-mediated Glu transfer onto the third PCP domain of KtzH, a reaction mechanism for the origin of the *threo*- and *erythro*-D-3-hydroxyglutamic acid moieties found in mature kutznerides can be proposed (Figure 6.8). The stand-alone adenylation domain KtzN first activates L-glutamic acid as amino acyl adenylate, which then can be transferred site-specifically to the third PCP domain of KtzH. The specificity of this interaction is mediated by the presence of the truncated A domain (A*) as well as one hydroxylase, either KtzO or KtzP.

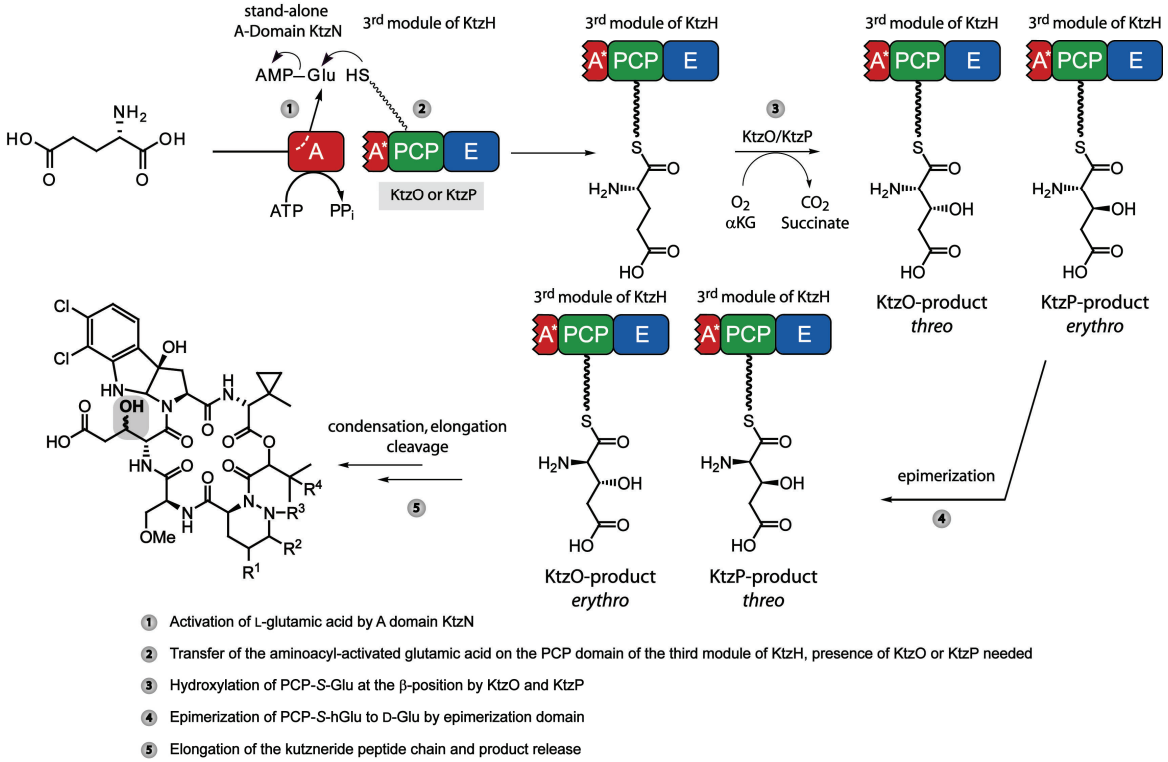


Figure 6.8: Proposed reaction mechanism for glutamic acid activation, incorporation and hydroxylation during kutzneride assembly.

Subsequently, the PCP-bound glutamic acid is hydroxylated by KtzO to yield L-*threo*-3-hydroxyglutamic acid or by KtzP to afford L-*erythro*-3-hydroxyglutamic acid. Epimerization of the two PCP-*S*-L-hGlu to D-amino acids catalyzed by the E domain, will also convert the relative positioning of the hydroxyl group such that the KtzO product becomes *erythro*-D-hGlu and the KtzP product *threo*-D-hGlu. Subsequently, condensation and peptide elongation occur, and release of the mature peptide is achieved by cyclization by the thioesterase domain.

6.3 Rational Manipulation of AsnO for the Synthesis of Pharmaceutical Relevant Agents

The knowledge acquired during the course of this study on the mechanism underlying the enzymatic synthesis of β-hydroxylated amino acids laid the foundations for the rational manipulation of AsnO. Nowadays rational protein design, based on 3D-structural information, has become a common tool in the manipulation of enzymes for biocatalysis.[240,241] The AsnO enzyme was manipulated based on its structure in complex with hAsn by site-directed mutagenesis of the oxygenase encoding gene. By a point mutation of a single base, resulting in an AsnO D241N variant on the protein level, the substrate specificity of this enzyme was completely changed and aspartic acid instead of asparagine is hydroxylated by the enzyme.

L-*threo*-hydroxyaspartic acid (L-THA), the β-hydroxylated form of L-aspartic acid, is of current medicinal interest because it inhibits L-asparagine synthetase, is a key-constituent of several proteins in the blood-clotting cascade, and inhibits the function of excitatory amino acid transporters as a L-glutamic acid mimic.[225,242] The latter function is of greater importance, as L-glutamate plays a key role as a primary neurotransmitter in the mammalian central nervous system and participates in diverse and complex neuronal communication by activating a broad assortment of the EAA receptors.[226] With its potential to over-activate these receptors L-glutamate can contribute to CNS damage in acute injuries or chronic diseases.[227] Thus, regulation of extracellular L-glutamate concentration, carried out by the EAA transporters, is crucial. As the organic synthesis of L-THA is complex and circuitous,[243,244] a fast enzymatic route is desired. Furthermore a readily available source of L-THA, which is achieved by the examined AsnO D241N variant, could help to further investigate these transporters and the complexity of L-glutamate-mediated signaling processes.

With its ability to synthesize L-THA even on a preparative scale the AsnO D241N variant is an excellent example in which basic research on non-proteinogenic amino acid building blocks of natural products paved the way for the efficient and enantiomerically pure synthesis of a pharmaceutical relevant agent.

6.3.1 Structural Basis for the Rational Manipulation

The AsnO structure and the identification of the substrate binding residues revealed that the residue 241 (Figure 6.9, labeled in red), which is an aspartic acid, binds the NH_2-group of the carboxamide group of the substrate L-asparagine (Figure 6.9 A, highlighted in orange).

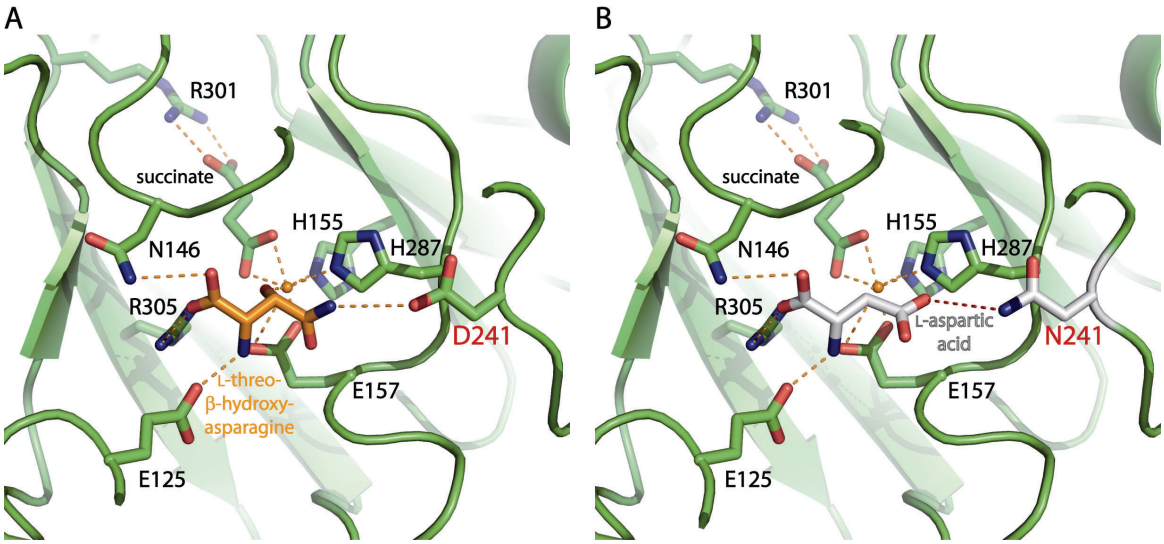


Figure 6.9. Substrate binding of AsnO and the D241N variant. A) The side chain of D241 binds the NH_2 of the carboxamide group of the asparagine substrate. B) Model of AsnO D241N. The mutation of D241 to N241 (shown in grey) alters the substrate specificity of the variant from Asn to Asp with retention of stereochemistry. The N241 could bind to the aspartic acid substrate (grey) with the side chain NH_2 to the side chain carboxyl group of Asp.

In the AsnO D241N variant, the carboxyl group of the side chain of residue 241 is replaced by a carboxamide group of which the NH_2 can bind to the substrate aspartic acid (Figure 6.9 B, highlighted in grey), as depicted in the AsnO D241N model (Figure 6.9 B). Retention of stereochemistry would also be expected, as the α-carboxy and α-amino coordinating residues (E125, N146, R305) are unaffected by the mutagenesis of D241 to N241 (Figure 6.9). Therefore, the product should be the medically interesting L-*threo*-hydroxyaspartic acid (L-THA), which indeed was observed (5.3.4 and below).

6.3.1 Biochemical Characterization of the AsnO Variant

Hydroxylation Activity and Substrate Specificity. The AsnO D241N variant was shown to hydroxylate aspartic acid, as expected by the exploration of the AsnO structure bound with product L-*threo*-β-hydroxyasparagine. Compared with the wild-type AsnO, the hydroxylation assays have to be carried out at lower temperatures, i.e. 16°C instead of 25°C. However, incubation at 16°C yielded quantitative conversion on an analytical scale.

The necessity for lower temperatures for the AsnO variant is likely due to reduced enzyme stability compared to the wild-type, putatively introduced by the AsnO manipulation. This is exemplified by the conversion rate decrease at higher temperatures [>98% (16°C) to 32% (37°C)].

The substrate specificity of the variant was evaluated by incubation of the enzyme with different amino acids overnight at 16°C (Table 5.6). LC-MS analyses of these assays indicated that the engineered protein is highly specific and exclusively accepts aspartic acid as its cognate substrate. Notably, neither the AsnO wild-type substrate L-Asn nor L-Asp related amino acids, such as L-Glu and or the D-configured isomer of Asp or other amino acids were accepted as substrates for protein-mediated hydroxylation. Thus, a highly specific Asp β-hydroxylase was generated.

Kinetic Parameters. The kinetic parameters have been determined to be: K_M= 0.46 ± 0.03 mM and k_{cat}= 1.0 ± 0.1 min^{-1}. Compared with the wild-type, the AsnO variant shows an almost identical K_M (0.48 ± 0.07 mm) for L-Asn, but has a 300-fold higher k_{cat} (298 ± 19 min^{-1}). The side chain swap therefore maintains the substrate affinity in the AsnO variant, but results in decelerated hydroxylation. One could anticipate that less obvious secondary interactions between the peptide backbone and the substrate are responsible for the accelerated reaction in the wild-type. In order to increase the conversion rate of the variant, the AsnO structure has to be closer examined and these secondary interactions have to be identified.

Preparative Scale. The AsnO D241N variant was capable to generate L-THA on a preparative scale in a simple one-pot enzymatic reaction and yielded 14.8 mg at the first try (68% yield). The potent, competitive, transportable EAAT1-4 inhibitor and non-transportable EAAT5 inhibitor[225] L-THA is commercially available exclusively at Tocris Bioscience, the biggest package size is 10 mg. Although not optimized yet, the synthesis presented in this study could compete with the current supplier even at the current state. In summary, a highly specific hydroxylase was designed that catalyzes the direct hydroxylation of the aliphatic β-position of L-Asp for the enantiomerically pure generation of the medicinally interesting L-THA even on a preparative scale.

Therefore, this enzyme is an excellent example for successful rational protein design based on basic research on hydroxylated amino acid building blocks and its application in the synthesis of a pharmaceutical relevant agent.

7. References

(1) Hopwood, D. A. Developments in the study of natural products biosynthesis. Preface. *Methods Enzymol* **2009**, *458*, xix-xxi.

(2) Li, J. W.; Vederas, J. C. Drug discovery and natural products: end of an era or an endless frontier? *Science* **2009**, *325*, 161-165.

(3) Leadlay, P. F.; Staunton, J.; Aparicio, J. F.; Bevitt, D. J.; Caffrey, P. et al. The erythromycin-producing polyketide synthase. *Biochem Soc Trans* **1993**, *21*, 218-222.

(4) Smith, D. J.; Burnham, M. K.; Edwards, J.; Earl, A. J.; Turner, G. Cloning and heterologous expression of the penicillin biosynthetic gene cluster from penicillum chrysogenum. *Biotechnology (N Y)* **1990**, *8*, 39-41.

(5) Debono, M.; Barnhart, M.; Carrell, C. B.; Hoffmann, J. A.; Occolowitz, J. L. et al. A21978C, a complex of new acidic peptide antibiotics: isolation, chemistry, and mass spectral structure elucidation. *J Antibiot (Tokyo)* **1987**, *40*, 761-777.

(6) Gerth, K.; Bedorf, N.; Hofle, G.; Irschik, H.; Reichenbach, H. Epothilons A and B: antifungal and cytotoxic compounds from Sorangium cellulosum (Myxobacteria). Production, physico-chemical and biological properties. *J Antibiot (Tokyo)* **1996**, *49*, 560-563.

(7) Ruegger, A.; Kuhn, M.; Lichti, H.; Loosli, H. R.; Huguenin, R. et al. [Cyclosporin A, a Peptide Metabolite from Trichoderma polysporum (Link ex Pers.) Rifai, with a remarkable immunosuppressive activity]. *Helv Chim Acta* **1976**, *59*, 1075-1092.

(8) Schwarzer, D.; Finking, R.; Marahiel, M. A. Nonribosomal peptides: from genes to products. *Nat Prod Rep* **2003**, *20*, 275-287.

(9) Walsh, C. T. Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* **2004**, *303*, 1805-1810.

(10) Weissman, K. J. Polyketide synthases: mechanisms and models. *Ernst Schering Res Found Workshop* **2005**, 43-78.

(11) Weissman, K. J. Introduction to polyketide biosynthesis. *Methods Enzymol* **2009**, *459*, 3-16.

(12) Caboche, S.; Pupin, M.; Leclere, V.; Fontaine, A.; Jacques, P. et al. NORINE: a database of nonribosomal peptides. *Nucleic Acids Res* **2008**, *36*, D326-331.

(13) Walsh, C. T. The chemical versatility of natural-product assembly lines. *Acc Chem Res* **2008**, *41*, 4-10.

(14) Hausinger, R. P. FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. *Crit Rev Biochem Mol Biol* **2004**, *39*, 21-68.

(15) Walsh, C. T.; Chen, H.; Keating, T. A.; Hubbard, B. K.; Losey, H. C. et al. Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr Opin Chem Biol* **2001**, *5*, 525-534.

(16) Wilkinson, B.; Micklefield, J. Chapter 14. Biosynthesis of nonribosomal peptide precursors. *Methods Enzymol* **2009**, *458*, 353-378.

(17) Mootz, H. D.; Schwarzer, D.; Marahiel, M. A. Ways of assembling complex natural products on modular nonribosomal peptide synthetases. *Chembiochem* **2002**, *3*, 490-504.

(18) Broberg, A.; Menkis, A.; Vasiliauskas, R. Kutznerides 1-4, depsipeptides from the actinomycete Kutzneria sp. 744 inhabiting mycorrhizal roots of Picea abies seedlings. *J Nat Prod* **2006**, *69*, 97-102.

(19) Pohanka, A.; Menkis, A.; Levenfors, J.; Broberg, A. Low-abundance kutznerides from Kutzneria sp. 744. *J Nat Prod* **2006**, *69*, 1776-1781.

(20) Schwarzer, D.; Marahiel, M. A. Multimodular biocatalysts for natural product assembly. *Naturwissenschaften* **2001**, *88*, 93-101.

(21) Finking, R.; Marahiel, M. A. Biosynthesis of nonribosomal peptides1. *Annu Rev Microbiol* **2004**, *58*, 453-488.

(22) Tsai, C. S. *Biomacromolecules: Intoduction to Structure, Function and Informatics*; 1st ed.; Wiley & Sons, : New York, 2006; 141.

(23) Baltz, R. H. Renaissance in antibacterial discovery from actinomycetes. *Curr Opin Pharmacol* **2008**, *8*, 557-563.

(24) Hopwood, D. A. *Streptomyces in Nature and Medicine - The Antibiotic Makers*; 1st ed.; Oxford University Press: New York, 2007.

(25) Baltz, R. H. Daptomycin: mechanisms of action and resistance, and biosynthetic engineering. *Curr Opin Chem Biol* **2009**, *13*, 144-151.

(26) Baltz, R. H.; Brian, P.; Miao, V.; Wrigley, S. K. Combinatorial biosynthesis of lipopeptide antibiotics in Streptomyces roseosporus. *J Ind Microbiol Biotechnol* **2006**, *33*, 66-74.

(27) Sattely, E. S.; Fischbach, M. A.; Walsh, C. T. Total biosynthesis: in vitro reconstitution of polyketide and nonribosomal peptide pathways. *Nat Prod Rep* **2008**, *25*, 757-793.

(28) Strieker, M.; Marahiel, M. A. The structural diversity of acidic lipopeptide antibiotics. *Chembiochem* **2009**, *10*, 607-616.

(29) Fischbach, M. A.; Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. *Chem Rev* **2006**, *106*, 3468-3496.

(30) Linne, U.; Marahiel, M. A. Reactions catalyzed by mature and recombinant nonribosomal peptide synthetases. *Methods Enzymol* **2004**, *388*, 293-315.

(31) Wenzel, S. C.; Muller, R. Formation of novel secondary metabolites by bacterial multimodular assembly lines: deviations from textbook biosynthetic logic. *Curr Opin Chem Biol* **2005**, *9*, 447-458.

(32) Wenzel, S. C.; Muller, R. Myxobacterial natural product assembly lines: fascinating examples of curious biochemistry. *Nat Prod Rep* **2007***, 24*, 1211-1224.
(33) Marahiel, M. A.; Stachelhaus, T.; Mootz, H. D. Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. *Chem Rev* **1997***, 97*, 2651-2674.
(34) Marahiel, M. A.; Essen, L. O. Chapter 13. Nonribosomal peptide synthetases mechanistic and structural aspects of essential domains. *Methods Enzymol* **2009***, 458*, 337-351.
(35) Keating, T. A.; Marshall, C. G.; Walsh, C. T.; Keating, A. E. The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains. *Nat Struct Biol* **2002***, 9*, 522-526.
(36) May, J. J.; Kessler, N.; Marahiel, M. A.; Stubbs, M. T. Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases. *Proc Natl Acad Sci U S A* **2002***, 99*, 12120-12125.
(37) Weber, T.; Baumgartner, R.; Renner, C.; Marahiel, M. A.; Holak, T. A. Solution structure of PCP, a prototype for the peptidyl carrier domains of modular peptide synthetases. *Structure* **2000***, 8*, 407-418.
(38) Samel, S. A.; Wagner, B.; Marahiel, M. A.; Essen, L. O. The thioesterase domain of the fengycin biosynthesis cluster: a structural base for the macrocyclization of a non-ribosomal lipopeptide. *J Mol Biol* **2006***, 359*, 876-889.
(39) Conti, E.; Stachelhaus, T.; Marahiel, M. A.; Brick, P. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *Embo J* **1997***, 16*, 4174-4183.
(40) Yonus, H.; Neumann, P.; Zimmermann, S.; May, J. J.; Marahiel, M. A. et al. Crystal structure of DltA. Implications for the reaction mechanism of non-ribosomal peptide synthetase adenylation domains. *J Biol Chem* **2008***, 283*, 32484-32491.
(41) Hegg, E. L. Unraveling the structure and mechanism of acetyl-coenzyme A synthase. *Acc Chem Res* **2004***, 37*, 775-783.
(42) Conti, E.; Lloyd, L. F.; Akins, J.; Franks, N. P.; Brick, P. Crystallization and preliminary diffraction studies of firefly luciferase from Photinus pyralis. *Acta Crystallogr D Biol Crystallogr* **1996***, 52*, 876-878.
(43) Gulick, A. M. Conformational Dynamics in the Acyl-CoA Synthetases, Adenylation Domains of Non-ribosomal Peptide Synthetases, and Firefly Luciferase. *ACS Chem Biol* **2009**.
(44) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem Biol* **1999***, 6*, 493-505.
(45) Rausch, C.; Weber, T.; Kohlbacher, O.; Wohlleben, W.; Huson, D. H. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Res* **2005***, 33*, 5799-5808.
(46) Bachmann, B. O.; Ravel, J. Chapter 8. Methods for in silico prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data. *Methods Enzymol* **2009***, 458*, 181-217.
(47) Weber, T.; Rausch, C.; Lopez, P.; Hoof, I.; Gaykova, V. et al. CLUSEAN: a computer-based framework for the automated analysis of bacterial secondary metabolite biosynthetic gene clusters. *J Biotechnol* **2009***, 140*, 13-17.
(48) Challis, G. L. Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology* **2008***, 154*, 1555-1569.
(49) Zerikly, M.; Challis, G. L. Strategies for the discovery of new natural products by genome mining. *Chembiochem* **2009***, 10*, 625-633.
(50) Francklyn, C. S. DNA polymerases and aminoacyl-tRNA synthetases: shared mechanisms for ensuring the fidelity of gene expression. *Biochemistry* **2008***, 47*, 11695-11703.
(51) Koglin, A.; Mofid, M. R.; Lohr, F.; Schafer, B.; Rogov, V. V. et al. Conformational switches modulate protein interactions in peptide antibiotic synthetases. *Science* **2006***, 312*, 273-276.
(52) Lambalot, R. H.; Gehring, A. M.; Flugel, R. S.; Zuber, P.; LaCelle, M. et al. A new enzyme superfamily - the phosphopantetheinyl transferases. *Chem Biol* **1996***, 3*, 923-936.
(53) Reuter, K.; Mofid, M. R.; Marahiel, M. A.; Ficner, R. Crystal structure of the surfactin synthetase-activating enzyme sfp: a prototype of the 4'-phosphopantetheinyl transferase superfamily. *Embo J* **1999***, 18*, 6823-6831.
(54) Mootz, H. D.; Finking, R.; Marahiel, M. A. 4'-phosphopantetheine transfer in primary and secondary metabolism of Bacillus subtilis. *J Biol Chem* **2001***, 276*, 37289-37298.
(55) Stachelhaus, T.; Huser, A.; Marahiel, M. A. Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases. *Chem Biol* **1996***, 3*, 913-921.
(56) Stachelhaus, T.; Mootz, H. D.; Bergendahl, V.; Marahiel, M. A. Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain. *J Biol Chem* **1998***, 273*, 22773-22781.
(57) Linne, U.; Marahiel, M. A. Control of directionality in nonribosomal peptide synthesis: role of the condensation domain in preventing misinitiation and timing of epimerization. *Biochemistry* **2000***, 39*, 10439-10447.
(58) Luo, L.; Kohli, R. M.; Onishi, M.; Linne, U.; Marahiel, M. A. et al. Timing of epimerization and condensation reactions in nonribosomal peptide assembly lines: kinetic analysis of phenylalanine activating elongation modules of tyrocidine synthetase B. *Biochemistry* **2002***, 41*, 9184-9196.
(59) Wyckoff, E. E.; Stoebner, J. A.; Reed, K. E.; Payne, S. M. Cloning of a Vibrio cholerae vibriobactin gene cluster: identification of genes required for early steps in siderophore biosynthesis. *J Bacteriol* **1997***, 179*, 7055-7062.
(60) Keating, T. A.; Marshall, C. G.; Walsh, C. T. Reconstitution and characterization of the Vibrio cholerae vibriobactin synthetase from VibB, VibE, VibF, and VibH. *Biochemistry* **2000***, 39*, 15522-15530.
(61) Samel, S. A.; Schoenafinger, G.; Knappe, T. A.; Marahiel, M. A.; Essen, L. O. Structural and functional insights into a peptide bond-forming bidomain from a nonribosomal peptide synthetase. *Structure* **2007***, 15*, 781-792.

(62) Leslie, A. G. Refined crystal structure of type III chloramphenicol acetyltransferase at 1.75 A resolution. *J Mol Biol* **1990**, *213*, 167-186.
(63) Kopp, F.; Marahiel, M. A. Where chemistry meets biology: the chemoenzymatic synthesis of nonribosomal peptides and polyketides. *Curr Opin Biotechnol* **2007**, *18*, 513-520.
(64) Kopp, F.; Marahiel, M. A. Macrocyclization strategies in polyketide and nonribosomal peptide biosynthesis. *Nat Prod Rep* **2007**, *24*, 735-749.
(65) Kopp, F.; Mahlert, C.; Grunewald, J.; Marahiel, M. A. Peptide macrocyclization: the reductase of the nostocyclopeptide synthetase triggers the self-assembly of a macrocyclic imine. *J Am Chem Soc* **2006**, *128*, 16478-16479.
(66) Bruner, S. D.; Weber, T.; Kohli, R. M.; Schwarzer, D.; Marahiel, M. A. et al. Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE. *Structure* **2002**, *10*, 301-310.
(67) Frueh, D. P.; Arthanari, H.; Koglin, A.; Vosburg, D. A.; Bennett, A. E. et al. Dynamic thiolation-thioesterase structure of a non-ribosomal peptide synthetase. *Nature* **2008**, *454*, 903-906.
(68) Grunewald, J.; Marahiel, M. A. Chemoenzymatic and template-directed synthesis of bioactive macrocyclic peptides. *Microbiol Mol Biol Rev* **2006**, *70*, 121-146.
(69) Tanovic, A.; Samel, S. A.; Essen, L. O.; Marahiel, M. A. Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science* **2008**, *321*, 659-663.
(70) Doekel, S.; Coeffet-Le Gal, M. F.; Gu, J. Q.; Chu, M.; Baltz, R. H. et al. Non-ribosomal peptide synthetase module fusions to produce derivatives of daptomycin in Streptomyces roseosporus. *Microbiology* **2008**, *154*, 2872-2880.
(71) Koglin, A.; Lohr, F.; Bernhard, F.; Rogov, V. V.; Frueh, D. P. et al. Structural basis for the selectivity of the external thioesterase of the surfactin synthetase. *Nature* **2008**, *454*, 907-911.
(72) Linne, U.; Schwarzer, D.; Schroeder, G. N.; Marahiel, M. A. Mutational analysis of a type II thioesterase associated with nonribosomal peptide synthesis. *Eur J Biochem* **2004**, *271*, 1536-1545.
(73) Schwarzer, D.; Mootz, H. D.; Linne, U.; Marahiel, M. A. Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. *Proc Natl Acad Sci U S A* **2002**, *99*, 14083-14088.
(74) Kagan, R. M.; Clarke, S. Widespread occurrence of three sequence motifs in diverse S-adenosylmethionine-dependent methyltransferases suggests a common structure for these enzymes. *Arch Biochem Biophys* **1994**, *310*, 417-427.
(75) Gaurav, K.; Gupta, N.; Sowdhamini, R. FASSM: enhanced function association in whole genome analysis using sequence and structural motifs. *In Silico Biol* **2005**, *5*, 425-438.
(76) Katz, J. E.; Dlakic, M.; Clarke, S. Automated identification of putative methyltransferases from genomic open reading frames. *Mol Cell Proteomics* **2003**, *2*, 525-540.
(77) Weber, G.; Schorgendorfer, K.; Schneider-Scherzer, E.; Leitner, E. The peptide synthetase catalyzing cyclosporine production in Tolypocladium niveum is encoded by a giant 45.8-kilobase open reading frame. *Curr Genet* **1994**, *26*, 120-125.
(78) Schauwecker, F.; Pfennig, F.; Grammel, N.; Keller, U. Construction and in vitro analysis of a new bi-modular polypeptide synthetase for synthesis of N-methylated acyl peptides. *Chem Biol* **2000**, *7*, 287-297.
(79) Gulavita, N. K.; Pomponi, S. A.; Wright, A. E.; Yarwood, D.; Sills, M. A. Isolation and Structure Elucidation of Perthamide-B, a Novel Peptide from the Sponge Theonella Sp. *Tetrahedron Lett* **1994**, *35*, 6815-6818.
(80) Rouhiainen, L.; Paulin, L.; Suomalainen, S.; Hyytiainen, H.; Buikema, W. et al. Genes encoding synthetases of cyclic depsipeptides, anabaenopeptilides, in Anabaena strain 90. *Mol Microb* **2000**, *37*, 156-167.
(81) Schoenafinger, G.; Schracke, N.; Linne, U.; Marahiel, M. A. Formylation domain: an essential modifying enzyme for the nonribosomal biosynthesis of linear gramicidin. *J Am Chem Soc* **2006**, *128*, 7406-7407.
(82) Hubbard, B. K.; Walsh, C. T. Vancomycin assembly: Nature's way. *Angew Chem Int Ed Engl* **2003**, *42*, 730-765.
(83) van Wageningen, A. M. A.; Kirkpatrick, P. N.; Williams, D. H.; Harris, B. R.; Kershaw, J. K. et al. Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. *Chem Biol* **1998**, *5*, 155-162.
(84) Hoffmann, K.; Schneider-Scherzer, E.; Kleinkauf, H.; Zocher, R. Purification and characterization of eucaryotic alanine racemase acting as key enzyme in cyclosporin biosynthesis. *J Biol Chem* **1994**, *269*, 12710-12714.
(85) Cheng, Y. Q.; Walton, J. D. A eukaryotic alanine racemase gene involved in cyclic peptide biosynthesis. *J Biol Chem* **2000**, *275*, 4906-4911.
(86) Balibar, C. J.; Vaillancourt, F. H.; Walsh, C. T. Generation of D amino acid residues in assembly of arthrofactin by dual condensation/epimerization domains. *Chem Biol* **2005**, *12*, 1189-1200.
(87) Stachelhaus, T.; Walsh, C. T. Mutational analysis of the epimerization domain in the initiation module PheATE of gramicidin S synthetase. *Biochemistry* **2000**, *39*, 5775-5787.
(88) Gehring, A. M.; Mori, I.; Perry, R. D.; Walsh, C. T. The nonribosomal peptide synthetase HMWP2 forms a thiazoline ring during biogenesis of yersiniabactin, an iron-chelating virulence factor of Yersinia pestis. *Biochemistry* **1998**, *37*, 11637-11650.
(89) Yeh, E.; Kohli, R. M.; Bruner, S. D.; Walsh, C. T. Type II thioesterase restores activity of a NRPS module stalled with an aminoacyl-S-enzyme that cannot be elongated. *Chembiochem* **2004**, *5*, 1290-1293.
(90) Schneider, A.; Marahiel, M. A. Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in non-ribosomal peptide biosynthesis in Bacillus subtilis. *Arch Microb* **1998**, *169*, 404-410.

(91) Black, T. A.; Wolk, C. P. Analysis of a Het- mutation in Anabaena sp. strain PCC 7120 implicates a secondary metabolite in the regulation of heterocyst spacing. *J Bacteriol* **1994***, 176*, 2282-2292.
(92) Hopwood, D. A. Molecular-Genetics of Polyketides and Its Comparison to Fatty-Acid Biosynthesis. *Annu Rev Genet* **1990***, 24*, 37-66.
(93) Lee, S. G.; Lipmann, F. Tyrocidine synthetase system. *Methods Enzymol* **1975***, 43*, 585-602.
(94) Copp, J. N.; Neilan, B. A. The phosphopantetheinyl transferase superfamily: phylogenetic analysis and functional implications in cyanobacteria. *Appl Environ Microbiol* **2006***, 72*, 2298-2305.
(95) Quadri, L. E.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P. et al. Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* **1998***, 37*, 1585-1595.
(96) McAllister, K. A.; Peery, R. B.; Zhao, G. S. Acyl carrier protein synthases from gram-negative, gram-positive, and atypical bacterial species: Biochemical and structural properties and physiological implications. *J Bacteriol* **2006***, 188*, 4737-4748.
(97) Mofid, M. R.; Finking, R.; Marahiel, M. A. Recognition of hybrid peptidyl carrier proteins/acyl carrier proteins in nonribosomal peptide synthetase modules by the 4'-phosphopantetheinyl transferases AcpS and Sfp. *J Biol Chem* **2002***, 277*, 17023-17031.
(98) Sunbul, M.; Zhang, K.; Yin, J. Chapter 10 using phosphopantetheinyl transferases for enzyme posttranslational activation, site specific protein labeling and identification of natural product biosynthetic gene clusters from bacterial genomes. *Methods Enzymol* **2009***, 458*, 255-275.
(99) Mofid, M. R.; Marahiel, M. A.; Ficner, R.; Reuter, K. Crystallization and preliminary crystallographic studies of Sfp: a phosphopantetheinyl transferase of modular peptide synthetases. *Acta Crystallogr D Biol Crystallogr* **1999***, 55*, 1098-1100.
(100) Koglin, A.; Walsh, C. T. Structural insights into nonribosomal peptide enzymatic assembly lines. *Nat Prod Rep* **2009***, 26*, 987-1000.
(101) Liu, Y.; Bruner, S. D. Rational manipulation of carrier-domain geometry in nonribosomal peptide synthetases. *Chembiochem* **2007***, 8*, 617-621.
(102) Schultz, A. W.; Oh, D. C.; Carney, J. R.; Williamson, R. T.; Udwary, D. W. et al. Biosynthesis and structures of cyclomarins and cyclomarazines, prenylated cyclic peptides of marine actinobacterial origin. *J Am Chem Soc* **2008***, 130*, 4507-4516.
(103) Harle, J.; Bechthold, A. Chapter 12. The power of glycosyltransferases to generate bioactive natural compounds. *Methods Enzymol* **2009***, 458*, 309-333.
(104) Feifel, S. C.; Schmiederer, T.; Hornbogen, T.; Berg, H.; Sussmuth, R. D. et al. In vitro synthesis of new enniatins: Probing the alpha-D-hydroxy carboxylic acid binding pocket of the multienzyme enniatin synthetase. *Chembiochem* **2007***, 8*, 1767-1770.
(105) Yanai, K.; Sumida, N.; Okakura, K.; Moriya, T.; Watanabe, M. et al. Para-position derivatives of fungal anthelmintic cyclodepsipeptides engineered with Streptomyces venezuelae antibiotic biosynthetic genes. *Nat Biotechnol* **2004***, 22*, 848-855.
(106) Rachid, S.; Krug, D.; Weissman, K. J.; Muller, R. Biosynthesis of (R)-beta-tyrosine and its incorporation into the highly cytotoxic chondramides produced by Chondromyces crocatus. *J Biol Chem* **2007***, 282*, 21810-21817.
(107) Fortin, P. D.; Walsh, C. T.; Magarvey, N. A. A transglutaminase homologue as a condensation catalyst in antibiotic assembly lines. *Nature* **2007***, 448*, 824-U811.
(108) Christianson, C. V.; Montavon, T. J.; Festin, G. M.; Cooke, H. A.; Shen, B. et al. The mechanism of MIO-based aminomutases in beta-amino acid biosynthesis. *J Am Chem Soc* **2007***, 129*, 15744-15745;
(109) Thomas, M. G.; Chan, Y. A.; Ozanick, S. G. Deciphering tuberactinomycin biosynthesis: Isolation, sequencing, and annotation of the viomycin biosynthetic gene cluster. *Antimicrob Agents Chemother* **2003***, 47*, 2823-2830.
(110) Lepore, B. W.; Ruzicka, F. J.; Frey, P. A.; Ringe, D. The x-ray crystal structure of lysine-2,3-aminomutase from Clostridium subterminale. *Proc Natl Acad Sci U S A* **2005***, 102*, 13819-13824.
(111) Helmetag, V.; Samel, S. A.; Thomas, M. G.; Marahiel, M. A.; Essen, L. O. Structural basis for the erythro-stereospecificity of the L-arginine oxygenase VioC in viomycin biosynthesis. *Febs J* **2009***, 276*, 3669-3682.
(112) Yin, X. H.; Zabriskie, T. M. VioC is a non-heme iron, alpha-ketoglutarate-dependent oxygenase that catalyzes the formation of 3S-hydroxy-L-arginine during viomycin biosynthesis. *Chembiochem* **2004***, 5*, 1274-1277.
(113) Haltli, B.; Tan, Y.; Magarvey, N. A.; Wagenaar, M.; Yin, X. H. et al. Investigating beta-hydroxyenduracididine formation in the biosynthesis of the mannopeptimycins. *Chem Biol* **2005***, 12*, 1163-1168.
(114) Nelson, J. T.; Lee, J.; Sims, J. W.; Schmidt, E. W. Characterization of SafC, a catechol 4-O-methyltransferase involved in saframycin biosynthesis. *Appl Environ Microb* **2007***, 73*, 3575-3580.
(115) Magarvey, N. A.; Haltli, B.; He, M.; Greenstein, M.; Hucul, J. A. Biosynthetic pathway for mannopeptimycins, lipoglycopeptide antibiotics active against drug-resistant gram-positive pathogens. *Antimicrob Agents Chemother* **2006***, 50*, 2167-2177.
(116) Magarvey, N. A.; Ehling-Schulz, M.; Walsh, C. T. Characterization of the cereulide NRPS alpha-hydroxy acid specifying modules: Activation of alpha-keto acids and chiral reduction on the assembly line. *J Am Chem Soc* **2006***, 128*, 10698-10699.

(117) Chen, H. W.; Thomas, M. G.; O'Connor, S. E.; Hubbard, B. K.; Burkart, M. D. et al. Aminoacyl-S-enzyme intermediates in beta-hydroxylations and alpha,beta-desaturations of amino acids in peptide antibiotics. *Biochemistry* **2001**, *40*, 11651-11659.
(118) Chen, H.; Hubbard, B. K.; O'Connor, S. E.; Walsh, C. T. Formation of beta-hydroxy histidine in the biosynthesis of nikkomycin antibiotics. *Chem Biol* **2002**, *9*, 103-112.
(119) Singh, G. M.; Vaillancourt, F. H.; Yin, J.; Walsh, C. T. Characterization of SyrC, an aminoacyltransferase shuttling threonyl and chlorothreonyl residues in the syringomycin biosynthetic assembly line. *Chem Biol* **2007**, *14*, 31-40.
(120) Tillett, D.; Dittmann, E.; Erhard, M.; von Dohren, H.; Borner, T. et al. Structural organization of microcystin biosynthesis in Microcystis aeruginosa PCC7806: an integrated peptide-polyketide synthetase system. *Chem Biol* **2000**, *7*, 753-764.
(121) Neumann, C. S.; Fujimori, D. G.; Walsh, C. T. Halogenation strategies in natural product biosynthesis. *Chem Biol* **2008**, *15*, 99-109.
(122) Vaillancourt, F. H.; Yeh, E.; Vosburg, D. A.; Garneau-Tsodikova, S.; Walsh, C. T. Nature's inventory of halogenation catalysts: Oxidative strategies predominate. *Chem Rev* **2006**, *106*, 3364-3378.
(123) Puk, O.; Bischoff, D.; Kittel, C.; Pelzer, S.; Weist, S. et al. Biosynthesis of chloro-beta-hydroxytyrosine, a nonproteinogenic amino acid of the peptidic backbone of glycopeptide antibiotics. *J Bacteriol* **2004**, *186*, 6093-6100.
(124) Stegmann, E.; Pelzer, S.; Bischoff, D.; Puk, O.; Stockert, S. et al. Genetic analysis of the balhimycin (vancomycin-type) oxygenase genes. *J Biotechnol* **2006**, *124*, 640-653.
(125) Vaillancourt, F. H.; Yin, J.; Walsh, C. T. SyrB2 in syringomycin E biosynthesis is a nonherne Fe-II alpha-ketoglutarate- and O-2-dependent halogenase. *Proc Natl Acad Sci U S A* **2005**, *102*, 10111-10116.
(126) Singh, G. M.; Fortin, P. D.; Koglin, A.; Walsh, C. T. beta-Hydroxylation of the aspartyl residue in the phytotoxin syringomycin E: characterization of two candidate hydroxylases AspH and SyrP in Pseudomonas syringae. *Biochemistry* **2008**, *47*, 11310-11320.
(127) Howard-Jones, A. R.; Kruger, R. G.; Lu, W.; Tao, J.; Leimkuhler, C. et al. Kinetic analysis of teicoplanin glycosyltransferases and acyltransferase reveal ordered tailoring of aglycone scaffold to reconstitute mature teicoplanin. *J Am Chem Soc* **2007**, *129*, 10082-10083.
(128) Mulichak, A. M.; Losey, H. C.; Lu, W.; Wawrzak, Z.; Walsh, C. T. et al. Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway. *Proc Natl Acad Sci U S A* **2003**, *100*, 9238-9243.
(129) Shen, B.; Du, L. C.; Sanchez, C.; Edwards, D. J.; Chen, M. et al. Cloning and characterization of the bleomycin biosynthetic gene cluster from Streptomyces verticillus ATCC15003. *J Nat Prod* **2002**, *65*, 422-431.
(130) Pacholec, M.; Tao, J. H.; Walsh, T. W. CouO and NovO: C-methyltransferases for tailoring the aminocoumarin scaffold in coumermycin and novobiocin antibiotic biosynthesis. *Biochemistry* **2005**, *44*, 14969-14976.
(131) Jakimowicz, P.; Freel Meyers, C. L.; Walsh, C. T.; Buttner, M. J.; Lawson, D. M. Crystallization and preliminary X-ray studies on the putative dTDP sugar epimerase NovW from the novobiocin biosynthetic cluster of Streptomyces spheroides. *Acta Crystallogr D Biol Crystallogr* **2003**, *59*, 1507-1509.
(132) Fischbach, M. A.; Lin, H.; Liu, D. R.; Walsh, C. T. How pathogenic bacteria evade mammalian sabotage in the battle for iron. *Nat Chem Biol* **2006**, *2*, 132-138.
(133) Miethke, M.; Marahiel, M. A. Siderophore-based iron acquisition and pathogen control. *Microbiol Mol Biol Rev* **2007**, *71*, 413-451.
(134) Goetz, D. H.; Holmes, M. A.; Borregaard, N.; Bluhm, M. E.; Raymond, K. N. et al. The neutrophil lipocalin NGAL is a bacteriostatic agent that interferes with siderophore-mediated iron acquisition. *Mol Cell* **2002**, *10*, 1033-1043.
(135) Fischbach, M. A.; Lin, H.; Liu, D. R.; Walsh, C. T. In vitro characterization of IroB, a pathogen-associated C-glycosyltransferase. *Proc Natl Acad Sci U S A* **2005**, *102*, 571-576.
(136) Raffatellu, M.; George, M. D.; Akiyama, Y.; Hornsby, M. J.; Nuccio, S. P. et al. Lipocalin-2 resistance confers an advantage to Salmonella enterica serotype Typhimurium for growth and survival in the inflamed intestine. *Cell Host Microbe* **2009**, *5*, 476-486.
(137) Nolan, E. M.; Fischbach, M. A.; Koglin, A.; Walsh, C. T. Biosynthetic tailoring of microcin E492m: post-translational modification affords an antibacterial siderophore-peptide conjugate. *J Am Chem Soc* **2007**, *129*, 14336-14347.
(138) Edwards, D. J.; Gerwick, W. H. Lyngbyatoxin biosynthesis: sequence of biosynthetic gene cluster and identification of a novel aromatic prenyltransferase. *J Am Chem Soc* **2004**, *126*, 11432-11433.
(139) Zhu, Y. Q.; Silverman, R. B. Revisiting heme mechanisms. a perspective on the mechanisms of nitric oxide synthase (NOS), heme oxygenase (HO), and cytochrome p450s (CYP450s). *Biochemistry* **2008**, *47*, 2231-2243.
(140) Abu-Omar, M. M.; Loaiza, A.; Hontzeas, N. Reaction mechanisms of mononuclear non-heme iron oxygenases. *Chem Rev* **2005**, *105*, 2227-2252.
(141) Clifton, I. J.; McDonough, M. A.; Ehrismann, D.; Kershaw, N. J.; Granatino, N. et al. Structural studies on 2-oxoglutarate oxygenases and related double-stranded beta-helix fold proteins. *J Inorg Biochem* **2006**, *100*, 644-669.
(142) Ozer, A.; Bruick, R. K. Non-heme dioxygenases: cellular sensors and regulators jelly rolled into one? *Nat Chem Biol* **2007**, *3*, 144-153.
(143) Que, L., Jr. One motif--many different reactions. *Nat Struct Biol* **2000**, *7*, 182-184.

(144) Purpero, V.; Moran, G. R. The diverse and pervasive chemistries of the alpha-keto acid dependent enzymes. *J Biol Inorg Chem* **2007**, *12*, 587-601.
(145) Yu, B.; Edstrom, W. C.; Benach, J.; Hamuro, Y.; Weber, P. C. et al. Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. *Nature* **2006**, *439*, 879-884.
(146) Schofield, C. J.; Ratcliffe, P. J. Oxygen sensing by HIF hydroxylases. *Nat Rev Mol Cell Biol* **2004**, *5*, 343-354.
(147) Bollinger, J. M.; Price, J. C.; Hoffart, L. M.; Barr, E. W.; Krebs, C. Mechanism of taurine: alpha-ketoglutarate dioxygenase (TauD) from Escherichia coli. *Eur J Inorg Chem* **2005**, 4245-4254.
(148) Price, J. C.; Barr, E. W.; Hoffart, L. M.; Krebs, C.; Bollinger, J. M. Kinetic dissection of the catalytic mechanism of taurine: alpha-ketoglutarate dioxygenase (TauD) from Escherichia coli. *Biochemistry* **2005**, *44*, 8138-8147.
(149) Zhou, J.; Gunsior, M.; Bachmann, B. O.; Townsend, C. A.; Solomon, E. I. Substrate binding to the alpha-ketoglutarate-dependent non-heme iron enzyme clavaminate synthase 2: Coupling mechanism of oxidative decarboxylation and hydroxylation. *J Am Chem Soc* **1998**, *120*, 13539-13540.
(150) Pavel, E. G.; Zhou, J.; Busby, R. W.; Gunsior, M.; Townsend, C. A. et al. Circular dichroism and magnetic circular dichroism spectroscopic studies of the non-heme ferrous active site in clavaminate synthase and its interaction with alpha-ketoglutarate cosubstrate. *J Am Chem Soc* **1998**, *120*, 743-753.
(151) Farquhar, E. R.; Koehntop, K. D.; Emerson, J. P.; Que, L., Jr. Post-translational self-hydroxylation: a probe for oxygen activation mechanisms in non-heme iron enzymes. *Biochem Biophys Res Commun* **2005**, *338*, 230-239.
(152) Schenk, G.; Pau, M. Y. M.; Solomon, E. I. Comparison between the geometric and electronic structures and reactivities of {FeNO}(7) and {FeO2}(8) complexes: A density functional theory study. *J Am Chem Soc* **2004**, *126*, 505-515.
(153) Price, J. C.; Barr, E. W.; Glass, T. E.; Krebs, C.; Bollinger, J. M. Evidence for hydrogen abstraction from C1 of taurine by the high-spin Fe(IV) intermediate detected during oxygen activation by taurine :alpha-ketoglutarate dioxygenase (TauD). *J Am Chem Soc* **2003**, *125*, 13008-13009.
(154) Riggs-Gelasco, P. J.; Price, J. C.; Guyer, R. B.; Brehm, J. H.; Barr, E. W. et al. EXAFS spectroscopic evidence for an Fe = O unit in the Fe(IV) intermediate observed during oxygen activation by taurine :alpha-ketoglutarate dioxygenase. *J Am Chem Soc* **2004**, *126*, 8108-8109.
(155) Proshlyakov, D. A.; Henshaw, T. F.; Monterosso, G. R.; Ryle, M. J.; Hausinger, R. P. Direct detection of oxygen intermediates in the non-heme Fe enzyme taurine/alpha-ketoglutarate dioxygenase. *J Am Chem Soc* **2004**, *126*, 1022-1023.
(156) Hoffart, L. M.; Barr, E. W.; Guyer, R. B.; Bollinger, J. M.; Krebs, C. Direct spectroscopic detection of a C-H-cleaving high-spin Fe(IV) complex in a prolyl-4-hydroxylase. *Proc Natl Acad Sci U S A* **2006**, *103*, 14738-14743.
(157) Vaillancourt, F. H.; Yeh, E.; Vosburg, D. A.; O'Connor, S. E.; Walsh, C. T. Cryptic chlorination by a non-haem iron enzyme during cyclopropyl amino acid biosynthesis. *Nature* **2005**, *436*, 1191-1194.
(158) Fujimori, D. G.; Hrvatin, S.; Neumann, C. S.; Strieker, M.; Marahiel, M. A. et al. Cloning and characterization of the biosynthetic gene cluster for kutznerides. *Proc Natl Acad Sci U S A* **2007**, *104*, 16498-16503.
(159) Neumann, C. S.; Walsh, C. T. Biosynthesis of (-)-(1S,2R)-allocoronamic acyl thioester by an Fe(II)-dependent halogenase and a cyclopropane-forming flavoprotein. *J Am Chem Soc* **2008**, *130*, 14022-14023.
(160) Ueki, M.; Galonic, D. P.; Vaillancourt, F. H.; Garneau-Tsodikova, S.; Yeh, E. et al. Enzymatic generation of the antimetabolite gamma,gamma-dichloroaminobutyrate by NRPS and mononuclear iron halogenase action in a streptomycete. *Chem Biol* **2006**, *13*, 1183-1191.
(161) Galonic, D. P.; Vaillancourt, F. H.; Walsh, C. T. Halogenation of unactivated carbon centers in natural product biosynthesis: Trichlorination of leucine during barbamide biosynthesis. *J Am Chem Soc* **2006**, *128*, 3900-3901.
(162) Blasiak, L. C.; Vaillancourt, F. H.; Walsh, C. T.; Drennan, C. L. Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis. *Nature* **2006**, *440*, 368-371.
(163) Wong, C.; Fujimori, D. G.; Walsh, C. T.; Drennan, C. L. Structural Analysis of an Open Active Site Conformation of Nonheme Iron Halogenase CytC3. *J Am Chem Soc* **2009**, *131*, 4872-4879.
(164) Galonic, D. P.; Barr, E. W.; Walsh, C. T.; Bollinger, J. M.; Krebs, C. Two interconverting Fe((IV)) intermediates in aliphatic chlorination by the halogenase CytC3. *Nat Chem Biol* **2007**, *3*, 113-116.
(165) Krebs, C.; Fujimori, D. G.; Walsh, C. T.; Bollinger, J. M. Non-heme Fe(IV)-oxo intermediates. *Acc Chem Res* **2007**, *40*, 484-492.
(166) Fujimori, D. G.; Barr, E. W.; Matthews, M. L.; Koch, G. M.; Yonce, J. R. et al. Spectroscopic evidence for a high-spin Br-Fe(IV)-Oxo intermediate in the alpha-ketoglutarate-dependent halogenase CytC3 from Streptomyces. *J Am Chem Soc* **2007**, *129*, 13408-+.
(167) Enoch, D. A.; Bygott, J. M.; Daly, M. L.; Karas, J. A. Daptomycin. *Journal of Infection* **2007**, *55*, 205-213.
(168) Raja, A.; LaBonte, J.; Lebbos, J.; Kirkpatrick, P. Daptomycin. *Nature Reviews Drug Discovery* **2003**, *2*, 943-944.
(169) Sauermann, R.; Rothenburger, M.; Graninger, W.; Joukhadar, C. Daptomycin: A review 4 years after first approval. *Pharmacology* **2008**, *81*, 79-91.
(170) Lakey, J. H.; Lea, E. J. A.; Rudd, B. A. M.; Wright, H. M.; Hopwood, D. A. A New Channel-Forming Antibiotic from Streptomyces-Coelicolor A3(2) Which Requires Calcium for Its Activity. *J Gen Microbiol* **1983**, *129*, 3565-3573.
(171) Boeck, L. D.; Papiska, H. R.; Wetzel, R. W.; Mynderse, J. S.; Fukuda, D. S. et al. A54145, a New Lipopeptide Antibiotic Complex - Discovery, Taxonomy, Fermentation and Hplc. *J Antibiot* **1990**, *43*, 587-593.

(172) Aretz, W.; Meiwes, J.; Seibert, G.; Vobis, G.; Wink, J. Friulimicins: Novel lipopeptide antibiotics with peptidoglycan synthesis inhibiting activity from Actinoplanes friuliensis sp nov I. Taxonomic studies of the producing microorganism and fermentation. *J Antibiot* **2000***, 53*, 807-815.
(173) Heinemann, B.; Kaplan, M. A.; Muir, R. D.; Hooper, I. R. Amphomycin, a New Antibiotic. *Antibiot Chemother* **1953***, 3*, 1239-1242.
(174) Naganawa, H.; Hamada, M.; Maeda, K.; Okami, Y.; Takeuchi, T. et al. Laspartomycin a New Anti-Staphylococcal Peptide. *J Antibiot* **1968***, 21*, 55-&.
(175) Kong, F. M.; Carter, G. T. Structure determination of glycinocins A to D, further evidence for the cyclic structure of the amphomycin antibiotics. *J Antibiot* **2003***, 56*, 557-564.
(176) Baltz, R. H.; Miao, V.; Wrigley, S. K. Natural products to drugs: daptomycin and related lipopeptide antibiotics. *Nat Prod Rep* **2005***, 22*, 717-741.
(177) Samel, S. A.; Marahiel, M. A.; Essen, L. O. How to tailor non-ribosomal peptide products - new clues about the structures and mechanisms of modifying enzymes. *Mol Biosyst* **2008***, 4*, 387-393.
(178) Debono, M.; Abbott, B. J.; Molloy, R. M.; Fukuda, D. S.; Hunt, A. H. et al. Enzymatic and Chemical Modifications of Lipopeptide Antibiotic A21978c - the Synthesis and Evaluation of Daptomycin (Ly146032). *J Antibiot* **1988***, 41*, 1093-1105.
(179) Miao, V.; Coeffet-Legal, M. F.; Brian, P.; Brost, R.; Penn, J. et al. Daptomycin biosynthesis in Streptomyces roseosporus: cloning and analysis of the gene cluster and revision of peptide stereochemistry. *Microbiology* **2005***, 151*, 1507-1523.
(180) Huber, F. M.; Pieper, R. L.; Tietz, A. J. The Formation of Daptomycin by Supplying Decanoic Acid to Streptomyces-Roseosporus Cultures Producing the Antibiotic Complex A21978c. *Journal of Biotechnology* **1988***, 7*, 283-292.
(181) Counter, F. T.; Allen, N. E.; Fukuda, D. S.; Hobbs, J. N.; Ott, J. et al. A54145 a New Lipopeptide Antibiotic Complex - Microbiological Evaluation. *J Antibiot* **1990***, 43*, 616-622.
(182) Fukuda, D. S.; Dubus, R. H.; Baker, P. J.; Berry, D. M.; Mynderse, J. S. A54145, a New Lipopeptide Antibiotic Complex - Isolation and Characterization. *J Antibiot* **1990***, 43*, 594-600.
(183) Miao, V.; Brost, R.; Chapple, J.; She, K.; Gal, M. F. et al. The lipopeptide antibiotic A54145 biosynthetic gene cluster from Streptomyces fradiae. *J Ind Microbiol Biotechnol* **2006***, 33*, 129-140.
(184) Kempter, C.; Kaiser, D.; Haag, S.; Nicholson, G.; Gnau, V. et al. CDA: Calcium-dependent peptide antibiotics from Streptomyces coelicolor A3(2) containing unusual residues. *Angew Chem Intl Ed Engl* **1997***, 36*, 498-501.
(185) Hojati, Z.; Milne, C.; Harvey, B.; Gordon, L.; Borg, M. et al. Structure, biosynthetic origin, and engineered biosynthesis of calcium-dependent antibiotics from Streptomyces coelicolor. *Chem Biol* **2002***, 9*, 1175-1187.
(186) Neary, J. M.; Powell, A.; Gordon, L.; Milne, C.; Flett, F. et al. An asparagine oxygenase (AsnO) and a 3-hydroxyasparaginyl phosphotransferase (HasP) are involved in the biosynthesis of calcium-dependent lipopeptide antibiotics. *Microbiology* **2007***, 153*, 768-776.
(187) Strieker, M.; Kopp, F.; Mahlert, C.; Essen, L. O.; Marahiel, M. A. Mechanistic and structural basis of stereospecific Cbeta-hydroxylation in calcium-dependent antibiotic, a daptomycin-type lipopeptide. *ACS Chem Biol* **2007***, 2*, 187-196.
(188) Penn, J.; Li, X.; Whiting, A.; Latif, M.; Gibson, T. et al. Heterologous production of daptomycin in Streptomyces lividans. *J Ind Microbiol Biotechnol* **2006***, 33*, 121-128.
(189) Heinzelmann, E.; Berger, S.; Puk, O.; Reichenstein, B.; Wohlleben, W. et al. A glutamate mutase is involved in the biosynthesis of the lipopeptide antibiotic friulimicin in Actinoplanes friuliensis. *Antimicrob Agents Chemother* **2003***, 47*, 447-457.
(190) Muller, C.; Nolden, S.; Gebhardt, P.; Heinzelmann, E.; Lange, C. et al. Sequencing and analysis of the biosynthetic gene cluster of the lipopeptide antibiotic friulimicin in Actinoplanes friuliensis. *Antimicrob Agents Chemother* **2007***, 51*, 1028-1037.
(191) Borders, D. B.; Leese, R. A.; Jarolmen, H.; Francis, N. D.; Fantini, A. A. et al. Laspartomycin, an acidic lipopeptide antibiotic with a unique peptide core. *Journal of Natural Products* **2007***, 70*, 443-446.
(192) Ho, S. W.; Jung, D.; Calhoun, J. R.; Lear, J. D.; Okon, M. et al. Effect of divalent cations on the structure of the antibiotic daptomycin. *European Biophysics Journal with Biophysics Letters* **2008***, 37*, 421-433.
(193) Jung, D.; Rozek, A.; Okon, M.; Hancock, R. E. W. Structural transitions as determinants of the action of the calcium-dependent antibiotic daptomycin. *Chemistry & Biology* **2004***, 11*, 949-957.
(194) Ball, L. J.; Goult, C. M.; Donarski, J. A.; Micklefield, J.; Ramesh, V. NMR structure determination and calcium binding effects of lipopeptide antibiotic daptomycin. *Organic & Biomolecular Chemistry* **2004***, 2*, 1872-1878.
(195) Rotondi, K. S.; Gierasch, L. M. A well-defined amphipathic conformation for the calcium-free cyclic lipopeptide antibiotic, daptomycin, in aqueous solution. *Biopolymers* **2005***, 80*, 374-385.
(196) Silverman, J. A.; Perlmutter, N. G.; Shapiro, H. M. Correlation of daptomycin bactericidal activity and membrane depolarization in Staphylococcus aureus. *Antimicrob Agents Chemother* **2003***, 47*, 2538-2544.
(197) Bunkoczi, G.; Vertesy, L.; Sheldrick, G. M. Structure of the lipopeptide antibiotic tsushimycin. *Acta Cryst D* **2005***, 61*, 1160-1164.
(198) Schneider, T.; Gries, K.; Josten, M.; Wiedemann, I.; Pelzer, S. et al. The Lipopeptide Antibiotic Friulimicin B Inhibits Cell Wall Biosynthesis through Complex Formation with Bactoprenol Phosphate. *Antimicrob Agents Chemother* **2009***, 53*, 1610-1618.

(199) Tanaka, H.; Oiwa, R.; Matsukura, S.; Inokoshi, J.; Omura, S. Studies on Bacterial-Cell Wall Inhibitors .10. Properties of Phospho-N-Acetylmuramoyl-Pentapeptide-Transferase in Peptidoglycan Synthesis of Bacillus-Megaterium and Its Inhibition by Amphomycin. *J Antibiot* **1982**, *35*, 1216-1221.
(200) Banerjee, D. K.; Scher, M. G.; Waechter, C. J. Amphomycin - Effect of the Lipopeptide Antibiotic on the Glycosylation and Extraction of Dolichyl Monophosphate in Calf Brain Membranes. *Biochemistry* **1981**, *20*, 1561-1568.
(201) Nguyen, K. T.; Kau, D.; Gu, J. Q.; Brian, P.; Wrigley, S. K. et al. A glutamic acid 3-methyltransferase encoded by an accessory gene locus important for daptomycin biosynthesis in Streptomyces roseosporus. *Mol Microbiol* **2006**, *61*, 1294-1307.
(202) Grunewald, J.; Sieber, S. A.; Mahlert, C.; Linne, U.; Marahiel, M. A. Synthesis and derivatization of daptomycin: a chemoenzymatic route to acidic lipopeptide antibiotics. *J Am Chem Soc* **2004**, *126*, 17025-17031.
(203) Kopp, F.; Grunewald, J.; Mahlert, C.; Marahiel, M. A. Chemoenzymatic design of acidic lipopeptide hybrids: New insights into the structure-activity relationship of daptomycin and A54145. *Biochemistry* **2006**, *45*, 10474-10481.
(204) Milne, C.; Powell, A.; Jim, J.; Al Nakeeb, M.; Smith, C. P. et al. Biosynthesis of the (2S,3R)-3-methyl glutamate residue of nonribosomal lipopeptides. *J Am Chem Soc* **2006**, *128*, 11250-11259.
(205) Mahlert, C.; Kopp, F.; Thirlway, J.; Micklefield, J.; Marahiel, M. A. Stereospecific enzymatic transformation of alpha-ketoglutarate to (2S,3R)-3-methyl glutamate during acidic lipopeptide biosynthesis. *J Am Chem Soc* **2007**, *129*, 12011-12018.
(206) Zhang, J. H.; Quigley, N. B.; Gross, D. C. Analysis of the syrP gene, which regulates syringomycin synthesis by Pseudomonas syringae pv syringae. *Appl Environ Microb* **1997**, *63*, 2771-2778.
(207) Amir-Heidari, B.; Thirlway, J.; Micklefield, J. Stereochemical course of tryptophan dehydrogenation during biosynthesis of the calcium-dependent lipopeptide antibiotics. *Org Lett* **2007**, *9*, 1513-1516.
(208) Amir-Heidari, B.; Micklefield, J. NMR confirmation that tryptophan dehydrogenation occurs with Syn stereochemistry during the biosynthesis of CDA in Streptomyces coelicolor. *J Org Chem* **2007**, *72*, 8950-8953.
(209) Kopp, F.; Linne, U.; Oberthur, M.; Marahiel, M. A. Harnessing the chemical activation inherent to carrier protein-bound thioesters for the characterization of lipopeptide fatty acid tailoring enzymes. *J Am Chem Soc* **2008**, *130*, 2656-2666.
(210) Heemstra, J. R., Jr.; Walsh, C. T. Tandem action of the O2- and FADH2-dependent halogenases KtzQ and KtzR produce 6,7-dichlorotryptophan for kutzneride assembly. *J Am Chem Soc* **2008**, *130*, 14024-14025.
(211) Sambrook, J.; Russel, D. W. *Molecular Cloning: A Laboratory Manual*; 3rd ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 2000.
(212) Haigler, B. E.; Gibson, D. T. Purification and Properties of Nadh-Ferredoxinnap Reductase, a Component of Naphthalene Dioxygenase from Pseudomonas Sp Strain Ncib-9816. *J Bacteriol* **1990**, *172*, 457-464.
(213) Kirschbaum, J.; Luckas, B.; Beinert, W. D. Precolumn Derivatization of Biogenic-Amines and Amino-Acids with 9-Fluorenylmethyl Chloroformate and Heptylamine. *J Chromatogr A* **1994**, *661*, 193-199.
(214) Vendrell, J.; Aviles, F. X. Complete Amino-Acid-Analysis of Proteins by Dabsyl Derivatization and Reversed-Phase Liquid-Chromatography. *J Chromatogr* **1986**, *358*, 401-413.
(215) CCP4 The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D* **1994**, *50*, 760-763.
(216) Vagin, A.; Teplyakov, A. MOLREP: an Automated Program for Molecular Replacement. *J Appl Cryst* **1997**, *30*, 1022-1025.
(217) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr D* **1997**, *D53*, 240-255.
(218) Emsley, P.; Cowtan, K. COOT: model-building tools for molecular graphics. *Acta Crystallogr D* **2004**, *60*, 2126-2132.
(219) DeLano, W. L. The PyMOL molecular graphics system; DeLano Scientific LLC, http://www.pymol.org: San Carlos, CA, USA, 2002.
(220) Zhang, Z.; Ren, J.; Stammers, D. K.; Baldwin, J. E.; Harlos, K. et al. Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase. *Nat Struct Biol* **2000**, *7*, 127-133.
(221) Baldwin, J. E.; Field, R. A.; Lawrence, C. C.; Merritt, K. D.; Schofield, C. J. Proline 4-Hydroxylase: Stereochemical Course of the Reaction. *Tetrahedron Lett* **1993**, *34*, 7489-7492.
(222) Eichhorn, E.; van der Ploeg, J. R.; Kertesz, M. A.; Leisinger, T. Characterization of alpha-ketoglutarate-dependent taurine dioxygenase from Escherichia coli. *J Biol Chem* **1997**, *272*, 23031-23036.
(223) Meier, J. L.; Mercer, A. C.; Rivera, H., Jr.; Burkart, M. D. Synthesis and evaluation of bioorthogonal pantetheine analogues for in vivo protein modification. *J Am Chem Soc* **2006**, *128*, 12174-12184.
(224) Berg, P.; Bergmann, F. H.; Ofengand, E. J.; Dieckmann, M. Enzymic Synthesis of Amino Acyl Derivatives of Ribonucleic Acid. *J Biol Chem* **1961**, *236*, 1726-1734.
(225) Bridges, R. J.; Esslinger, C. S. The excitatory amino acid transporters: pharmacological insights on substrate and inhibitor specificity of the EAAT subtypes. *Pharmacol Ther* **2005**, *107*, 271-285.
(226) Schoepp, D. D. Unveiling the functions of presynaptic metabotropic glutamate receptors in the central nervous system. *J Pharmacol Exp Ther* **2001**, *299*, 12-20.
(227) Olney, J. W. Excitotoxicity, apoptosis and neuropsychiatric disorders. *Curr Opin Pharm* **2003**, *3*, 101-109.
(228) Elkins, J. M.; Ryle, M. J.; Clifton, I. J.; Dunning Hotopp, J. C.; Lloyd, J. S. et al. X-ray crystal structure of Escherichia coli taurine/alpha-ketoglutarate dioxygenase complexed to ferrous iron and substrates. *Biochemistry* **2002**, *41*, 5185-5192.

(229) Salowe, S. P.; Marsh, E. N.; Townsend, C. A. Purification and characterization of clavaminate synthase from Streptomyces clavuligerus: an unusual oxidative enzyme in natural product biosynthesis. *Biochemistry* **1990***, 29*, 6499-6508.
(230) You, Z.; Omura, S.; Ikeda, H.; Cane, D. E.; Jogl, G. Crystal structure of the non-heme iron dioxygenase PtlH in pentalenolactone biosynthesis. *J Biol Chem* **2007***, 282*, 36552-36560.
(231) Muller, T. A.; Zavodszky, M. I.; Feig, M.; Kuhn, L. A.; Hausinger, R. P. Structural basis for the enantiospecificities of R- and S-specific phenoxypropionate/alpha-ketoglutarate dioxygenases. *Prot Sci* **2006***, 15*, 1356-1368.
(232) Du, L. C.; Sanchez, C.; Chen, M.; Edwards, D. J.; Shen, B. The biosynthetic gene cluster for the antitumor drug bleomycin from Streptomyces verticillus ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase. *Chem Biol* **2000***, 7*, 623-642.
(233) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **1993***, 234*, 779-815.
(234) Bentley, S. D.; Chater, K. F.; Cerdeno-Tarraga, A. M.; Challis, G. L.; Thomson, N. R. et al. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature* **2002***, 417*, 141-147.
(235) Piepersberg, W. Directed Deposit to GenBank (AJ810851.1). **2004**.
(236) Ju, J.; Ozanick, S. G.; Shen, B.; Thomas, M. G. Conversion of (2S)-arginine to (2S,3R)-capreomycidine by VioC and VioD from the viomycin biosynthetic pathway of Streptomyces sp. strain ATCC11861. *Chembiochem* **2004***, 5*, 1281-1285.
(237) Lee, C.; Kim, S. J.; Jeong, D. G.; Lee, S. M.; Ryu, S. E. Structure of human FIH-1 reveals a unique active site pocket and interaction sites for HIF-1 and von Hippel-Lindau. *J Biol Chem* **2003***, 278*, 7558-7563.
(238) Semenza, G. L. Regulation of mammalian O2 homeostasis by hypoxia-inducible factor 1. *Annu Rev Cell Dev Biol* **1999***, 15*, 551-578.
(239) Bitto, E.; Bingman, C. A.; Allard, S. T. M.; Wesenberg, G. E.; Aceti, D. J. et al. The structure at 2.4 angstrom resolution of the protein from gene locus At3g21360, a putative Fe-II/2-oxoglutaratedependent enzyme from Arabidopsis thaliana. *Acta Cryst F* **2005***, 61*, 469-472.
(240) Bornscheuer, U. T.; Pohl, M. Improved biocatalysts by directed evolution and rational protein design. *Curr Opin Chem Biol* **2001***, 5*, 137-143.
(241) Hult, K.; Berglund, P. Engineered enzymes for improved organic synthesis. *Curr Opin Biotechnol* **2003***, 14*, 395-400.
(242) Esslinger, C. S.; Agarwal, S.; Gerdes, J.; Wilson, P. A.; Davis, E. S. et al. The substituted aspartate analogue L-beta-threo-benzyl-aspartate preferentially inhibits the neuronal excitatory amino acid transporter EAAT3. *Neuropharmacology* **2005***, 49*, 850-861.
(243) Benoiton, L.; Birnbaum, S. M.; Winitz, M.; Greenstein, J. P. The Enzymic Resolution of Beta-Methylaspartic Acid with Acylase-Ii. *Arch Biochem Biophys* **1959***, 81*, 434-438.
(244) Hafez, A. M.; Dudding, T.; Wagerle, T. R.; Shah, M. H.; Taggi, A. E. et al. A multistage, one-pot procedure mediated by a single catalyst: A new approach to the catalytic asymmetric synthesis of beta-amino acids. *J Org Chem* **2003***, 68*, 5819-5825.

Acknowledgment

First of all, I like to thank Prof. M. A. Marahiel for being an excellent supervisor, his scientific guidance and constant generous support during my Ph. D. thesis. I gratefully acknowledge his open-minded character and his support of new ideas. Paired with his enthusiasm about science, he generated an atmosphere of inspiration, which was of great benefit of my work. Most of all, I am grateful for the opportunity to spend a most inspiring year at Harvard Medical School and to participate in numerous conference and summer schools. Especially, the NRPS/PKS meetings in Hirschegg (Austria) and Providence, RI as well as the summer school on “Microbial Secondary Metabolites: Genomes, Signals and Communities” under administration of Sir D. Hopwood and Prof. J. Davies have been outstanding experiences.

I am much indebted to Prof C. T. Walsh for welcoming me in his laboratory and for the constant support during my one year stay. With his decades of experiences and his never ebbing wellspring of wisdom, discussions with him have always been very inspiring and fruitful. I am grateful for the experience to work in such a professional environment and for his generous support. Furthermore, I like to thank him for his scientific supervision and his openness towards new thoughts. The one-year stay in Boston was a priceless experience, which I always will be grateful for.

I thank Prof. L.-O. Essen for an excellent collaboration on the AsnO structure and variant. Furthermore, I like to acknowledge him and Prof. A. Geyer for being in my thesis committee.

I gratefully acknowledge Prof. Elizabeth Nolan (MIT) for a fruitful collaboration and for being a good friend during my time in Boston. For the latter, I also like to thank Dr. Jay A. Read, Prof. Darren Hansen (Rutgers), Dr. Alexander Koglin, Prof. Eric Strieter (Wisconsin), Dr. Gita Singh and Prof. Danica Fujimori (UCSF). Furthermore, I am indebted to all members of the Walsh group for the warm welcome, the nice working atmosphere and inspiring discussions.

I acknowledge the whole Marahiel group for the nice working atmosphere and the great time in the lab. In particular, I would like to thank my current and former colleagues of Lab 4707, Dr. Christoph Mahlert, Dr. Florian Kopp and Verena Helmetag, Florian Ritterbusch and Femke Kraas, for many helpful discussions and the oustanding co-operation. I really enjoyed being part of the Marahiel group no matter if at work or in the various leisure time activities.

Furthermore, I like to thank Dr. Uwe Linne for his expert opinion in bioanalytical problems and his helpful advises. I gratefully acknowledge the technical support of Antje Schäfer and her trainees Karola Kisselbach and Anke Botthoff.

I thank Verena Helmetag and Lars Robbel for proof-reading this thesis.

Special thanks go to my friends and my family, without them I would never have made it. I am much indebted to my mother, who supported me her whole lifetime, which I will never forget.

Most of all, I like to thank my wife Svenja for being a delightful partner, for the wonderful past experiences we shared and for those which will come.

Erklärung

Ich versichere, dass ich meine Dissertation „*Unusual Building Blocks and Domain Organization of Non-Ribosomal Peptide Synthetases*“ selbständig, ohne erlaubte Hilfe angefertigt und mich dabei keiner als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, 14.08.2009 _____

(Ort, Datum) (Unterschrift)