Nonribosomal Peptide Synthetases: Quaternary Structure and Chemoenzymatic Synthesis of Macrocyclic Peptides

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem
Fachbereich Chemie
der Philipps-Universität Marburg
vorgelegt von

Stephan Axel Sieber

aus Marburg/Lahn

Marburg/Lahn 2004

Vom Fachbereich Chemie
der Philipps-Universität Marburg als Dissertation
am 10. Februar 2004 angenommen.

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

Tag der Disputation: 01. März 2004

Meinen Eltern ...

Summary

Nonribosomal peptide synthetases (NRPS) are multi-domain megaenzymes which are involved in the biosynthesis of structurally diverse natural peptides with broad biological activities. Recent NRPS research concentrates on protein reengineering of this machinery to alter product outcome. For these efforts, knowledge of the NRPS protein architecture is important which has been subject of speculations so far. In this work the monomeric structure of several NRPS systems was demonstrated by a set of techniques. Biophysical methods such as analytical ultracentrifugation, gel filtration and chemical cross-linking gave clear evidence for a monomeric architecture. In turn biochemical methods like mutant *in trans* complementation and a two affinity tag system revealed no indication for a dimeric state. All together, the results clearly show that in contrast to polyketide and fatty acid synthases, NRPSs work as monomers.

The second and predominant part of this work is dedicated to the chemoenzymatic engineering of NRPS products. A key determent of nonribosomal peptide product activity is the cyclic structure of many compounds. In this work the principles of enzymatic peptide cyclization were first investigated on the example of the lactone forming surfactin thioesterase (Srf TE). To evaluate recognition elements in the lipoheptapeptide chain by Srf TE, alterations to the fatty acyl group, the heptapeptide, and the thioester leaving group were made. Analysis of substrate tolerance with diaminopropionic acid and ε-aminohexanoic acid spacer substrates revealed an enzyme specific recognition of C- and N-terminal residues of the peptide sequence, leaving space for side chain replacements. Utilizing the recently solved crystal structure, the active site nucleophile, Ser80, was changed to Cys, and the other components of the catalytic triad, Asp107 and His207, were substituted by Ala, with the resulting mutants lacking detectable activity. Two cationic side chains in the active site, Lys111 and Arg120 that potentially coordinate the substrate, were altered to Ala, causing increased substrate hydrolysis, as did a Pro26 to Gly mutant, mimicking the behavior of lipases. To broaden the catalytic tool set, also fengycin (lipodecapeptide lactone), mycosubtilin (lipoheptapeptide lactam) and syringomycin (lipononapeptide lactone) TEs, which employ different cyclization strategies were cloned and overexpressed during the course of these studies. Surprisingly, neither cyclization nor hydrolysis activity with conventional peptidyl-SNAC substrates was observed. For a better substrate presentation, a new strategy was developed that took advantage of the promiscuity of the 4′-phosphopantetheine (ppan) transferase Sfp, which allowed to load synthetic peptidyl-CoA substrates onto the didomainal fengycin PCP-TE enzyme. The precise natural interaction of the TE with the lipodecapeptidyl fengycin substrate attached to the carrier domain PCP revealed regioselective cyclization activity with various substrate analogues. However, a limitation of this approach was the single turnover reaction. In order to overcome this limitation peptidyl thiophenol substrates were synthesized to force multiple turnover via transesterification with the free ppan thiol. Surprisingly, the new thiophenol substrates did already show activity with the TE domain alone which indicated a rapid direct acyclation of the TE active site serine. With this new approach it was possible to characterize previously SNAC-inactive fengycin, syringomycin and mycosubstilin TEs. Moreover, a protein evolution method was employed to develop “custom made” catalytic cyclization tools. Initial results of a mass screen of about 1700 mutant enzymes, derived from tyrocidine TE (252 residues), revealed 6 new enzyme variants which displayed altered activities. Due to stop codons within the mutant enzyme sequences (located in residues 141-178) these enzymes lacked 74 to 111 residues (30 – 44 %). Biochemical data with various substrate analogues indicated that the N-terminal part of Tyc TE is sufficient for catalysis.

In conclusion, a tool set of several efficient enzymes capable of stereo- and regioselective cyclization of various linear peptides is now available without the need of side chain protecting groups, overcoming limitations of chemical synthesis.

Zusammenfassung

Nichtribosomale Peptidsynthetasen (NRPS) sind multifunktionelle Enzymkomplexe, die für die Biosynthese einer Reihe strukturell diverser Peptide mit außerordentlichen biologischen Eigenschaften verantwortlich sind. In dieser Arbeit konnte mit einer Reihe von Techniken die monomere Struktur dieser Multienzymkomplexe gezeigt werden. Biophysikalische Methoden, wie z. B. analytische Ultrazentrifugation, Gel-Filtration und chemisches „cross-linking" gaben den Beweis für eine monomere Organisation. Zusätzlich haben biochemische Methoden, wie z. B. das „*in trans* Complementations" System und das Zwei-Affinitätstag System keinen Hinweis auf eine dimere Organisation ergeben. Zusammengenommen zeigen die Ergebnisse im Gegensatz zu den verwandten dimeren Polyketid- und Fettsäuresynthetasen eine monomere Architektur für NRPS Enzyme.

Im zweiten Teil der Arbeit wurden die Grundzüge der enzymatischen Peptidzyklisierung am Beispiel der lactonbildenden Surfactin Thioesterase (Srf TE) mit Hilfe artifizieller SNAC-Thioester Substrate untersucht. Um Erkennungssequenzen in der Lipoheptapeptidkette durch Srf TE herauszufinden, wurden systematische Veränderungen in der Fettsäuregruppe, dem Peptidrückgrat und in der Abgangsgruppe (SNAC) vorgenommen. Die Analyse der Substrattoleranz des Peptidrückgrats durch Austausch einzelner Aminosäurereste mit Diaminopropionsäure und ε-Aminohexansäure zeigte eine spezifische Erkennung C- und N-terminaler Reste der Peptidsequenz. Reste in der Mitte der Peptidsequenz konnten unter Erhalt der Zyklisierungsaktivität ausgetauscht werden. Mit Hilfe der Kristallstruktur von Srf TE konnten gezielte Mutationen in der aktiven Tasche des Enzyms vorgenommen werden. Das vermutliche Nukleophil Ser80 wurde gegen Cys, und alle anderen Reste der katalytischen Triade, Asp107 und His207, wurden durch Ala ersetzt. Alle Mutanten zeigten keine Aktivität. Zwei positiv geladene Seitengruppen, Lys111 und Arg120, die vermutlich das Substrat für die Zyklisierung koordinieren, wurden durch Ala ersetzt und zeigten verstärkt hydrolysiertes Produkt. Eine Mutation von Pro26 zu Gly in der Oxyanionentasche wurde basierend auf einem Vergleich mit Lipasen eingeführt und erhöhte drastisch die Substrathydrolyse. Um die Anwendbarkeit von Thioesterasen in diversen Zyklisierungsreaktionen zu vergrößern, wurden die Fengycin TE (Lipodekapeptid Lacton), Mycosubtilin TE (Lipoheptapeptid Lactam) und Syringomycin TE (Lipononapeptid Lacton) kloniert und überexpremiert. Überraschenderweise konnte weder eine Zyklisierungs- noch Hydrolyseaktivität mit herkömmlichen SNAC-Substraten festgestellt werden. In einem neuen Ansatz wurde die Flexibilität der Phosphopantethein (ppan) Transferase Sfp ausgenutzt, die eine direkte Beladung des Transportenzyms PCP im Zweidomänensystem PCP-TE mit synthetischen peptidyl-CoA ermöglicht. Durch kovalente Bindung des Peptids an das Transportenzym PCP konnte eine Substraterkennung erzielt und regioselektive Zyklisierungsaktivität mit dem Fengycin PCP-TE System für viele Substrate beobachtet werden. Ein Nachteil des Verfahrens stellt die einfache „turnover" Reaktion mit diesem System dar. Daher wurde in einem neuen Ansatz versucht durch eine Transthioesterifikationsreaktion mit Peptidyl-Thiophenolaten die freie Thiol-Gruppe des PCP-ppan mit neuem Peptidsubstrat zu beladen. Überraschenderweise zeigten die Thiophenol Substrate bereits Zyklisierungsaktivität mit der TE allein ohne die freie ppan Thiol Gruppe, was die direkte Acylierung des aktiven Zentrums der TE demonstriert. Mit dieser neuen Thiophenol-Methode war es dann auch möglich Aktivität für andere zuvor inaktive TEs zu erhalten, z. B. Mycosubtilin und Syringomycin TEs. Des weiteren wurde versucht durch Protein-Evolution maßgeschneiderte Zyklisierungsenzyme zu erhalten. Vorläufige Ergebnisse eines Massenscreens von 1700 zufällig mutierten Enzymen, die sich von der Tyrocidin TE (252 Reste) ableiten, zeigten 6 Mutanten mit veränderten Aktivitäten. Wegen Stop-Codons inmitten der Enzymsequenz (Reste 141-178) fehlten allen Enzymen etwa 30 – 40 % ihrer ursprünglichen Sequenz. Biochemische Untersuchungen mit verschiedenen Substratanaloga zeigten, dass der N-terminale Teil der TE ausreichend ist für eine katalytische Zyklisierungsaktivität, was nun rationale Veränderungen im C-terminalen Bereich ermöglicht.

The majority of the work presented here has been published:

Grünewald, J.*, Sieber, S. A.*, Marahiel, M. A. " Chemo- and Regioselective Peptide Cyclization triggered by the N-terminal Fatty Acid Chain Length: The Recombinant Cyclase of the Calcium-dependent Antibiotic from Streptomyces coelicolor" *Biochemistry.* 2004, in press.

Sieber, S. A., Tao, J., Walsh, C. T., Marahiel, M. A. "Peptidyl thiophenols as substrates for nonribosomal peptide cyclases ". *Angewandte Chemie.* 2004, 116, 499-504.

Sieber, S. A., Marahiel, M. A. "Learning from Nature´s drug factories: Nonribosomal Synthesis of cyclic peptides" *J. Bacteriol*. 2003, 185, 7036-7043.

Clugston, S. L., Sieber, S. A., Marahiel, M. A., Walsh, C. T. "Chirality of peptide-bond forming condensation domains in nonribosomal peptide synthetases: The C_5 domain of tyrocidine synthetase is a $^{D}C_{L}$ catalyst" *Biochemistry.* 2003, 42, 12095-12104.

Sieber, S. A., Walsh, C. T., Marahiel, M. A. "Loading peptidyl-coemzyme A onto peptidyl carrier proteins – a novel approach in characterizing macrocyclization by thioesterase domains" *J. Am. Chem. Soc.* 2003, 125, 10862-10866.

Sieber, S. A., Marahiel, M. A. "Antibiotika vom molekularen Fließband: Modulare Peptidsynthetasen als Biokatalysatoren" *Biospektrum*, Sonderausgabe, 9. Jahrgang.

Tseng, C. C.*, Bruner, S. D., Kholi, R. M., Marahiel, M. A., Walsh, C. T., Sieber, S. A.* "Characterization of the surfactin C-terminal thioesterase domain as a cyclic depsipeptide synthase" *Biochemistry.* 2002, 41, 13350-13359.

Sieber, S. A.*, Linne, U.*, Hillson, N. J., Roche, E., Walsh, C. T., Marahiel, M. A. "Evidence for a monomeric structure of nonribosomal peptide synthetases" *Chem. & Biol.*, 2002, 9, 997-1008.

*these authors contributed equally to this work

Table of Contents

1 Abbreviations 1
2 Introduction 4
2.1 Natural Products 4
2.2 The NRPS Natural Product Factory 7
2.3 Activation by the Adenylation Domain 9
2.4 Intermediates Transport by the Peptidyl Carrier Protein 10
2.4.1 Misacylation and Regeneration 12
2.5 Peptide Elongation by the Condensation Domain 12
2.6 Editing Domains 14
2.6.1 Epimerization 14
2.6.2 Methylation 16
2.7 Peptide Release 16
2.7.1 Cyclization Strategies: Nature´s Way 17
2.7.2 Chemoenzymatic Cyclization 20
2.7.3 Structural and Mechanistic Aspects of NRPS Peptide Cyclases 24
2.8 Quaternary Architecture 26
2.9 Task 29
3 Materials 30
3.1 Chemicals, Enzymes and General Materials 30
3.2 Equipment 31
3.3 Vector Systems 32
3.3.1 pQE60-Vector 32
3.3.2 pASK-IBA 2 and pASK-IBA 4 33
3.3.3 pET37b 33
3.3.4 pBAD202/D-TOPO 33
3.4 Microorganisms 34
3.4.1 E. coli XL1-Blue 34
3.4.2 E. coli BL21(DE3) 34
3.4.3 E. coli M15 34
3.5 Media 34
4 Methods 35
4.1 Methods of Molecular Biology 35
4.1.1 Construction of Plasmids 35
4.1.2 Site Directed Mutagenesis 36

4.1.3 Random Mutagenesis 37
4.1.4 DNA Sequencing 38
4.2 Protein Techniques 38
4.2.1 Gene Expression 39
4.2.1.1 Expression with the pQE- and pET-Vector Systems 39
4.2.1.2 Expression with the IBA-Vector Systems 39
4.2.1.3 Expression with the pBAD/TOPO-Vector System 39
4.2.2 Protein Purification 40
4.2.2.1 Preparation of Cell Extracts 40
4.2.2.2 Purification of His_6-tagged Proteins 40
4.2.2.3 Purification of Strep-tagged Proteins 40
4.2.3 Concentration of Protein Solutions 41
4.3 Biochemical Methods 41
4.3.1 Cyclization Assays with SNAC and Thiophenol Substrates 41
4.3.2 Loading Peptidyl-CoA onto PCP 42
4.3.3 PCP-TE Cyclization Assay 43
4.3.4 Protein Evolution 43
4.4 Analytical Methods 43
4.4.1 Analytical Ultracentrifugation 43
4.4.2 MALDI-MS 44
4.4.3 HPLC-MS 44
4.5 Solid Phase Peptide Synthesis (SPPS) 47
4.5.1 SPPS: Initiation, Elongation and Termination 48
4.5.1.1 Initiation: Coupling of the First Amino Acid to the Resin 48
4.5.1.2 Elongation 48
4.5.1.3 Termination 50
4.6 Chemical Synthesis 50
4.6.1 Synthesis of SNAC and Thiophenol Thioesters 50
4.6.2 Synthesis of Peptidyl-CoA 51
4.6.3 Synthesis of Cyclosporin A SNAC 51
5 Results 53
5.1 Protein Architecture: The Quaternary Structure of NRPS 53
5.1.1 Gel Filtration, Chemical Crosslinking and Equilibrium Analytical Ultracentrifugation 54
5.1.2 Dissociation and Reassociation of Subunits of a Putative Dimer 58

5.1.2.1 Two Affinity Tag System 59
5.1.2.2 Complementation 60
5.2 Natural Product Architecture: Chemoenzymatic Peptide Cyclization 63
5.2.1 Biochemical Characterization of Surfactin TE 63
5.2.1.1 Incubation of SrfTE with the Native SNAC Substrate 63
5.2.1.2 Dap Scan of the Heptapeptide Sequence 64
5.2.1.3 Spacer Scan 66
5.2.1.4 Reverse Reaction: Hydrolysis of the Macrolactone 67
5.2.1.5 Alterations to Leaving Group, Nucleophile, and Length of Substrate 68
5.2.1.6 Mutations of Catalytically Important Residues of Surfactin TE 70
5.2.1.7 Cocrystal Structure of Srf TE with a Dipeptidyl Boronate Inhibitor 73
5.2.2 Loading Peptidyl-CoA onto PCPs: A Novel Approach to Characterize Macrocyclization by TEs 74
5.2.2.1 Overexpression of the Proteins and Synthesis of Peptidyl-CoAs 75
5.2.2.2 Loading Peptidyl-CoAs onto PCP 76
5.2.2.3 Autonomous Cyclization Activity of FenPCP-TE 77
5.2.2.4 Regioselectivity of Cyclization 78
5.2.3 Peptidyl-Thiophenols as New Substrates for Peptide Cyclases 80
5.2.3.1 Cloning and Overexpression of the Proteins 80
5.2.3.2 Activity Based Enzyme Acylation with Thiophenol Substrates 81
5.2.3.3 Tuning the Reactivity of the Thiophenol Leaving Group 82
5.2.3.4 Characterization of New Peptide Cyclases 84
5.2.4 Directed Protein Evolution to Provide “Custom Made” Peptide Cyclases 85
5.2.4.1 Cloning and Overexpression of the Cyclosporin C-Domain 86
5.2.4.2 Synthesis of Cyclosporin SNAC 87
5.2.4.3 Cyclization of Cyclosporin A SNAC by TycTE? 88
5.2.4.4 Directed Protein Evolution of Tyc TE 89
5.2.4.5 Error Prone PCR and Periplasmatic Expression 92
5.2.4.6 High Throughput Mass Analysis 93
5.2.4.7 Biochemical Characterization of Hits 94
6 Discussion 98
6.1 Quaternary Architecture of NRPS 98
6.2 Enzymatic Peptide Cyclization 103
6.2.1 Characterization of Surfactin Thioesterase 103
6.2.1.1 Structure and Mechanism 103

6.2.1.2 Substrate Tolerance 104
6.2.1.3 Substrate Recognition - Cocrystallization 107
6.2.1.4 Comparison with Tyrocidine Cyclase 107
6.2.2 Characterization of Fengycin Thioesterase by Natural Substrate Presentation 108
6.2.3 Novel Thiophenol Based Leaving Groups 110
6.2.4 Protein Evolution 112
6.2.5 General Utility of Peptide Cyclases: Chemical vs. Enzymatic Cyclization 117
7 Literature 120

Inhaltsverzeichnis

1 Abkürzungen 1
2 Einleitung 4
2.1 Naturprodukte 4
2.2 Die NRPS Naturprodukt Fabrik 7
2.3 Aktivierung durch die Adenylierungsdomäne 9
2.4 Intermediat Transport durch das Peptidyl Carrier Protein 10
2.4.1 Falschbeladung und Regeneration 12
2.5 Peptid Elongation durch die Kondensationsdomäne 12
2.6 Editierende Domänen 14
2.6.1 Epimerisierung 14
2.6.2 Methylierung 16
2.7 Peptid Freisetzung 16
2.7.1 Zyklisierungs Strategien: Der natürliche Weg 17
2.7.2 Chemoenzymatische Zyklisierung 20
2.7.3 Strukturelle und mechanistische Aspekte von NRPS Peptid Zyklasen 24
2.8 Quartärstruktur 26
2.9 Aufgabenstellung 29
3 Material 30
3.1 Chemikalien, Enzyme und Materialen 30
3.2 Ausstattung 31
3.3 Vektor Systeme 32
3.3.1 pQE60-Vektor 32
3.3.2 pASK-IBA 2 and pASK-IBA 4 33
3.3.3 pET37b 33
3.3.4 pBAD202/D-TOPO 33
3.4 Microorganismen 34
3.4.1 E. coli XL1-Blue 34
3.4.2 E. coli BL21(DE3) 34
3.4.3 E. coli M15 34
3.5 Medien 34
4 Methoden 35
4.1 Methoden der Molekularbiology 35
4.1.1 Konstruktion von Plasmiden 35
4.1.2 Gerichtete Mutagenese 36

4.1.3 Zufällige Mutagenese 37
4.1.4 DNA Sequenzierung 38
4.2 Protein Techniken 38
4.2.1 Gen Expression 39
4.2.1.1 Expression mit dem pQE- und pET-Vektor Systemen 39
4.2.1.2 Expression mit dem IBA-Vektor System 39
4.2.1.3 Expression mit dem pBAD/TOPO-Vektor System 39
4.2.2 Protein Reinigung 40
4.2.2.1 Präparation von Zellextrakten 40
4.2.2.2 Reinigung von His_6-getaggten Proteinen 40
4.2.2.3 Reinigung von Strep-getaggten Proteinen 40
4.2.3 Konzentration der Proteinlösungen 41
4.3 Biochemische Methoden 41
4.3.1 Zyklisierungs Assays mit SNAC und Thiophenol Substraten 41
4.3.2 Beladung von Peptidyl-CoA auf PCP 42
4.3.3 PCP-TE Zyklisierungs Assay 43
4.3.4 Protein Evolution 43
4.4 Analytische Methoden 43
4.4.1 Analytische Ultrazentrifugation 43
4.4.2 MALDI-MS 44
4.4.3 HPLC-MS 44
4.5 Festphasenpeptidsynthese (SPPS) 47
4.5.1 SPPS: Initiation, Elongation und Termination 48
4.5.1.1 Initiation: Kupplung der ersten Aminosäure an das Harz 48
4.5.1.2 Elongation 48
4.5.1.3 Termination 50
4.6 Chemische Synthese 50
4.6.1 Synthese von SNAC und Thiophenol Thioestern 50
4.6.2 Synthese von Peptidyl-CoA 51
4.6.3 Synthese von Cyclosporin A SNAC 51
5 Ergebnisse 53
5.1 Protein Architektur: Die Quartärstruktur von NRPS 53
5.1.1 Gel-Filtration, Chemisches Crosslinking and Analytische Ultrazentrifugation 54

5.1.2 Dissoziation und Reassoziation von Untereinheiten in einem möglichen Dimer 58
5.1.2.1 Zwei-Affinitätstag System 59
5.1.2.2 Komplementation 60
5.2 Naturprodukt Architektur: Chemoenzymatische Peptidzyklisierung 63
5.2.1 Biochemische Charakterisierung der Surfactin TE 63
5.2.1.1 Inkubation von SrfTE mit dem nativen SNAC Substrat 63
5.2.1.2 Dap Scan in der Heptapeptidsequenz 64
5.2.1.3 Spacer Scan 66
5.2.1.4 Rückreaktion: Hydrolyse des Makrolactons 67
5.2.1.5 Veränderungen in der Abgangsgruppe, im Nukleophil und in der Länge des Substrates 68
5.2.1.6 Mutationen in katalytisch wichtigen Resten der Surfactin TE 70
5.2.1.7 Kocrystalstruktur der Srf TE mit einem Dipeptidyl Boronat Inhibitor 73
5.2.2 Beladung von Peptidyl-CoA auf PCPs: Ein neuer Ansatz zur Charakterisierung der Makrozyklisierung durch TEs 74
5.2.2.1 Überexpression der Proteine und Synthese der Peptidyl-CoAs 75
5.2.2.2 Beladung der Peptidyl-CoAs auf PCP 76
5.2.2.3 Autonome Zyklisierungsaktivität von FenPCP-TE 77
5.2.2.4 Regioselective Zyklisierung 78
5.2.3 Peptidyl-Thiophenole als neue Substrate für Peptid Zyklasen 80
5.2.3.1 Klonierung und Überexpression der Proteine 80
5.2.3.2 Aktivitätsbasierende Acylierung mit Thiophenol Substraten 81
5.2.3.3 Justierung der Reaktivität der Thiophenol Abgangsgruppe 82
5.2.3.4 Charakterisierung neuer Peptid Zyklasen 84
5.2.4 Gerichtete Proteinevolution zur Herstellung “maßgeschneiderter” Peptid Zyclasen 85
5.2.4.1 Klonierung und Überexpression der Cyclosporin C-Domäne 86
5.2.4.2 Synthese von Cyclosporin SNAC 87
5.2.4.3 Zyklisierung von Cyclosporin A SNAC durch TycTE? 88
5.2.4.4 Gerichtete Proteinevolution der Tyc TE 89
5.2.4.5 „Error Prone” PCR und periplasmatische Expression 92
5.2.4.6 Hochdurchsatz Massenanalyse 93
5.2.4.7 Biochemische Charakterisierung der Treffer 94

6 Diskussion 98
6.1 Quartäre Architektur von NRPS 98
6.2 Enzymatische Peptidzyklisierung 103
6.2.1 Charakterisierung der Surfactin Thioesterase 103
6.2.1.1 Struktur und Mechanismus 103
6.2.1.2 Substrat Toleranz 104
6.2.1.3 Substrat Erkennung - Kocrystallization 107
6.2.1.4 Vergleich mit der Tyrocidine Zyklase 107
6.2.2 Charakterisation der Fengycin Thioesterase durch nätürliche Substrat Präsentation 108
6.2.3 Neue Thiophenol basierende Abgangsgruppen 110
6.2.4 Protein Evolution 112
6.2.5 Anwendung von Peptid Zyklasen: Chemische vs. enzymatische Zyklisierung 117
7 Literatur 120

1 Abbreviations

ACP	acyl carrier protein
ACV	δ-(α-aminoadipyl)-cysteinyl-D-valine
A-domain	adenylation domain
Amp	ampicillin
ATP	adenosine-5'-triphosphate
BMH	bismaleimidohexane
bp	base pairs
C-domain	condensation domain
CLP	cyclosporin linear precursor
CoA	coenzyme A
Cy-domain	heterocyclization domain
Da	dalton
DBP	dibromopropanone
DCC	dicyclohexylcarbodiimide
DCM	dichlormethane
DEBS	desoxyerythronolid-B-synthase
DHB	dihydroxybenzoyl
DIPEA	diisopropylethylamine
DKP	D-Phe-L-Pro diketopiperazine
DMF	N,N-dimethylformamide
dNTP	2'-desoxynucleosid-5'-triphosphate
E-domain	epimerisation domain
EDTA	ethylene diamino tetra acedic acid
epPCR	error prone polymerase chain reaction
FAS	fatty acid synthase
Fen PCP-TE	PCP-TE didomain from the fengycin NRPS
FLP	fengycin linear precursor
HBTU	2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexaflurophosphate
HOBt	1-hydroxybenzotriazole
HPLC	high performance liquid chromatography
IMAC	immobilized metal ion affinity chromatography
IPTG	isopropyl-β-D-thiogalactoside
Kan	kanamycin
LC/MS	liquid chromatography/mass spectrometry
MALDI-TOF	matrix assisted laser desorption ionization-time of flight
MCS	multiple cloning site
MIC	minimal inhibitory concentration
min	minutes
MLP	mycosubtilin linear precursor

mRNA	messenger RNA
Myc PCP-TE	PCP-TE didomain from the mycosubtilin NRPS
n.d.	not detected
N-Mt-domain	*N*-methylation domain
NRPS	nonribosomal peptide synthetase
NTA	nitrilotriacetate
OD	optical density
ONAC	N-acetylethanolamine
Ox-domain	oxidation domain
PAGE	polyacrylamide gelelectrophoresis
PCP	peptidyl carrier protein or thiolation domain
PCR	polymerase chain reaction
PKS	polyketide synthase
ppan	4'-phosphopantetheine
PP_i	inorganic pyrophosphate
PyBOP	benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexaflourophosphate
RBS	ribosomal binding site
R-domain	reductase domain
rpm	rounds per minute
RT	room temperature
SDS	sodiumdodecylsulfate
SLP	surfactin linear precursor
SNAC	N-acetylcysteamine
SPPS	solid phase peptide synthesis
Srf TE	TE domain from the tyrocidine NRPS
SyLP	syringomycin linear precursor
Syr TE	TE domain from the syringomycin NRPS
T-domain	thiolation domain PCP
TE-domain	thioesterase domain
TFA	triflouroacedic acid
TFE	triflourethanol
THF	terahydrofuran
TLP	tyrocidine linear precursor
Tyc TE	TE domain from the tyrocidine NRPS
v/v	volume per volume
w/v	weight per volume
wt	wild type

Tab. 1.1 Amino acids: Abbreviations and molecular weights

Amino acid	3-Letter code	1-Letter code	MW [g/mol]
Alanine	Ala	A	89
Arginine	Arg	R	174
Asparagine	Asn	N	132
Aspartate	Asp	D	133
Cysteine	Cys	C	121
Glutamate	Glu	E	147
Glutamine	Gln	Q	146
Glycine	Gly	G	75
Histidine	His	H	155
Isoleucine	Ile	I	131
Leucine	Leu	L	131
Lysine	Lys	K	146
Methionine	Met	M	149
Ornithine	Orn	O	132
Phenylalanine	Phe	F	165
Proline	Pro	P	115
Serine	Ser	S	105
Threonine	Thr	T	119
Tryptophane	Trp	W	204
Tyrosine	Tyr	Y	181
Valine	Val	V	117

2 Introduction

Research on bioactive natural products was strongly inspired by A. Flemming´s exceptional discovery of the antibiotic activity of the peptidic molecule penicillin from the fungal host organism *Penicillium notatum*[1]. Since then microorganisms attracted considerable attention as a new source for pharmaceutical agents and screening of microbial extracts revealed a very large number of new compounds with antimicrobial, antiviral, immunosuppressive, and antitumor activities. These secondary metabolites have been optimised for their dedicated function during generations of evolution and now represent promising scaffolds for the development of new drug leads.

Within these substances small peptide molecules represent a large subclass of bioactive natural products, which comprise unique structural features like heterocyclic elements, D-amino acids, and glycosylated as well as N-methylated residues. Moreover, contrary to ribosomal protein synthesis, small peptide products contain not only the common 20 amino acids, but hundreds of different building blocks, suggesting a nonribosomal origin of biosynthesis. In the 1970s Lipmann et al. reported a nucleic acid independent synthesis of the peptide antibiotics gramicidin S and tyrocidine A from *Bacillus* sp. through the use of large enzyme complexes similar to fatty acid synthases[2]. In the following years more and more peptidic natural products were shown to be assembled by such large enzymes, referred to as nonribosomal peptide synthetases (NRPS). Significant progress has been made in the past decades toward understanding the principles of bioactive peptide synthesis in microorganisms and was the subject of recent extensive reviewing[3-6]. While research first focused on the principal molecular architecture of the assembled complex molecules and on the identification of their biosynthetic gene clusters, more recent biochemical work revealed high resolution structures of some of the catalytic core enzymes. Moreover, recent chemoenzymatic approaches were developed to reprogram natural peptide sequences by the combined action of chemical peptide synthesis and subsequent enzyme catalysis. This introduction chapter aims to give a brief overview on the nonribosomal peptide synthesis machinery with the major emphasis on recent progress made on chemoenzymatic synthesis of cyclic peptides.

2.1 Natural Products

Natural peptide products synthesized by NRPSs can be grouped according to their biological activities. A major class comprises antibiotic and antifungal agents, including the peptides tyrocidine, bacitracin, surfactin, pristinamycin, vancomycin and fengycin[3, 7-15]. Their biological functionality is strictly associated with their chemical structure, which constrains

the peptide sequence in its biologically active conformation and ensures specific interaction with a dedicated molecular target. This structural rigidity is achieved either by cyclization or oxidative cross-links, which contribute to stability. Moreover, the great diversity of chemical modifications like fatty acid chains, D-amino acids, aromatic and aliphatic side chains, glycosylated amino acids and heterocyclic elements add much to these specific interactions (Fig. 2.1).

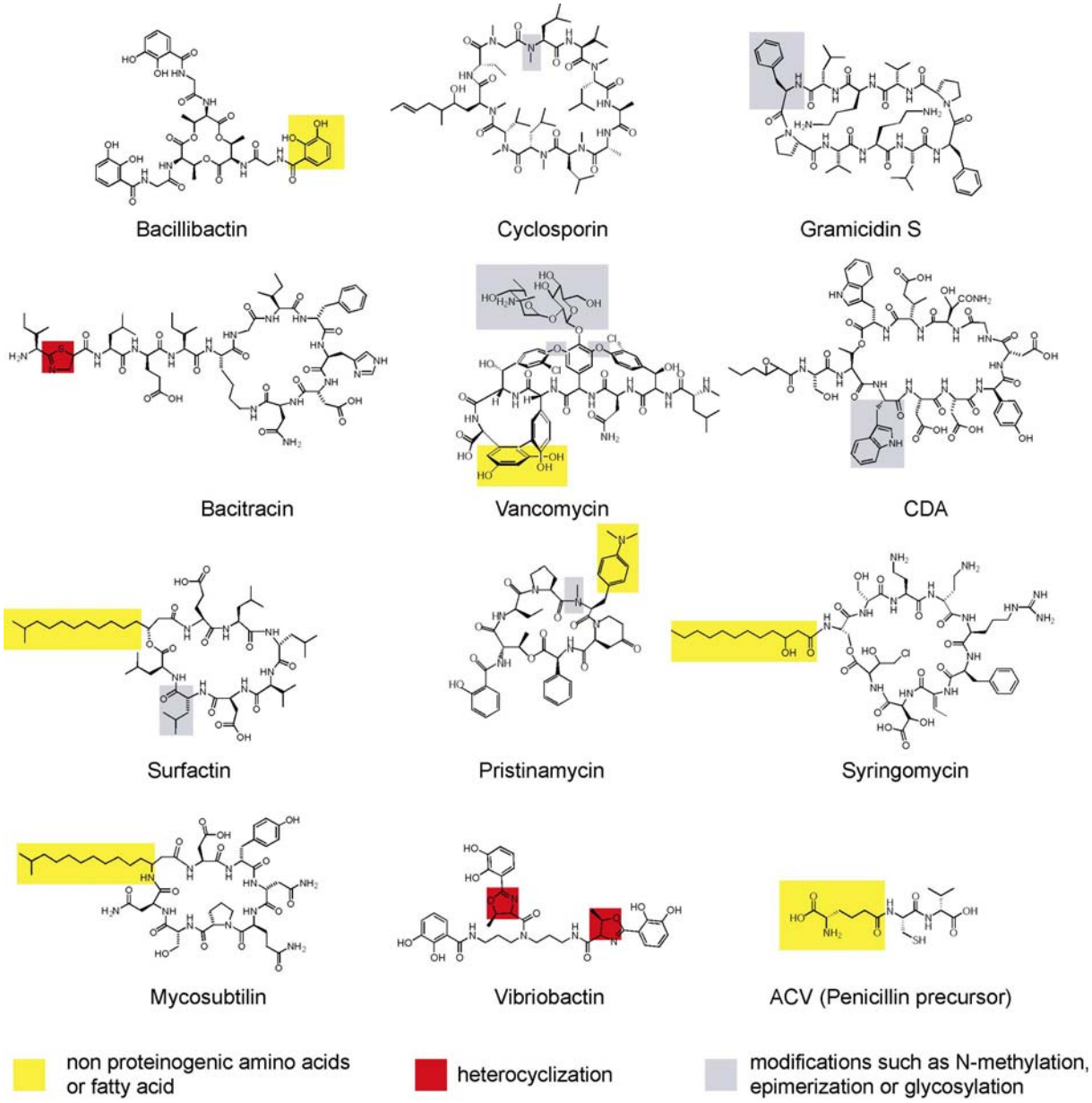


Figure 2.1 Natural peptide products. Characteristic structural features are highlighted on selected examples.

For many natural peptide products the relation between structural features and biological activities has been investigated. Some of these like bacitracin from *B. licheniformis* and gramicidin S form *B. brevis* act also non-specifically as membrane-inserting amphiphatic

species[16-19]. In addition to the macrocyclic ring, bacitracin contains a small thiazoline ring, which complements its antibiotic activity with a specific cation-dependent complexation of the phosphate moiety of the C_{55} lipid phosphate. This complexation leads to the inhibition of the lipid cycle[3, 20, 21]. Amphiphilic lipopetides like the surface tension reducing agent surfactin and the antifungal mycosubtilin, both produced by B. *subtilis*, are also thought to penetrate and disrupt the cell membrane, a process where the lipo-chain seems to play a key role[22-24]. In addition to the exceptional surfactant power provided by its amphiphilic sequence, surfactin has been reported to exhibit hemolytic, antivirial, antibacterial, and antitumor properties[10]. Also in case of the cytotoxic molecule syringomycin from *Pseudomonas syringae*, which exhibits toxicity against plant tissues, the amphiphatic nature of the polar peptide head and hydrophobic fatty acid tail allows insertation into the plant plasma membrane and formation of transmembrane pores permitting ions to flow freely across the membrane[25, 26]. Moreover, as seen for the *Streptomyces* lipopetidelactones CDA and daptomycin also metal ions like Ca^{2+} trigger antibiotic activity by stacking molecules[27-30]. These complexes may exert their antibacterial activity through membrane seeking, surface-active behaviour. Another close relationship between peptide structure and function is observed for the glycopeptide antibiotic vancomycin produced by *Streptomyces orientalis*. Vancomycin is a linear heptapeptide, whose backbone is constrained by oxidative cross-linking. This unique structure sequesters substrate peptidoglycan-D-Ala-D-Ala termini units with five hydrogen bonds and shuts down transpeptidation reaction[12]. A different cellular target is attacked by the antibiotic pristinamycin from *Streptomyces pristinaespiralis*, which blocks polypeptide translation by binding the 50S subunit of bacterial ribosomes at 23S rRNA sites. Investigations revealed interaction with the ribosom via the 3-hydroxy picolinic acid residue of pristinamycin emphasizing the importance of non-proteinogenic residues for antibiotic activity[31].

Cyclosporin produced by *Tolypocladium niveum* exhibits immunosuppressive and toxic properties due to the formation of a specific complex with cyclophilin which inhibits the protein phosphatase calcineurin, responsible for T-cell activation[32-34]. Cyclosporin is highly lipophilic and seven of its eleven amino acids are N-methylated. This high degree of methylation protects the peptide from proteolytic digest but also complicates chemical synthesis due to low coupling yields and side reactions[35]. In iron deficient environment, some bacteria such as *E. coli, B. subtilis* and *Vibrio Cholerae* synthesize and secrete iron-chelating molecules known as siderophores that scavenge Fe^{3+} with picomolar affinity, important for the host survival[36, 37]. Three catechol ligands derived from 2,3-

dihydroxybenzoyl (DHB) building blocks in bacillibactin, enterobactin and vibriobactin, complex iron by forming an intramolecular octahedral.

Many nonribosomal peptide products presented here show distinct chemical modifications, important to specifically interact and inhibit certain cellular functions, which are essential for survival. The high toxicity of the peptide products could therefore become a problem for the producer organism unless strategies for its own protection and immunity have coevolved with antibiotic biosynthesis. This immunity is achieved by several strategies including efflux pumps, temporary product inactivation and modifications of the target in the producer strain[3]. The latter strategy is used by vancomycin producing streptomycetes by changing the D-Ala-D-Ala terminus of the peptidoglycan pentapeptide precursor to a D-Ala-D-lactate terminus, which reduces binding affinity to vancomycin 1000-fold[12].

Due to their exceptional pharmacological activities many compounds like cyclosporin and vancomycin have been synthesized chemically[38, 39]. Regio- and stereoselective reactions require the use of protecting groups as well as chiral catalysts. Moreover, macrocyclization and coupling of N-methylated peptide bonds is difficult to achieve in satisfying yields, indicating an advantage of natural vs. synthetic strategies. Structural peculiarities of these complex peptide products suggested early on a nucleic acid independent biosynthesis facilitated by multiple catalytic domains expressed as a single multidomain protein. The diverse chemical reactions mediated by distinct enzymatic units will be the focus of the following section.

2.2 The NRPS Natural Product Factory

Although structurally diverse, most of the biologically produced peptides share a common mode of synthesis, the multienzyme thiotemplate mechanism[2, 6, 40]. According to this model, peptide bond formation takes place on large multienzyme complexes, which represent at the same time template and biosynthetic machinery. Sequencing of genes encoding NRPSs of bacterial and fungal origin provided insights into the molecular architecture and revealed a modular organization[6]. A module is a distinct section of the multienzyme that is responsible of the incorporation of one dedicated amino acid into the final product[3, 6, 41]. It is further subdivided into a catalytically independent set of domains responsible for substrate recognition, activation, binding, modification, elongation and release. Domains can be identified at the protein level by characteristic highly conserved sequence motivs. So far 10 different domains are known, catalysing independent chemical reactions. As an example to follow the principles, figure 2.2 shows a prototype NRPS assembly line of the cyclic lipoheptapetide surfactin[42].

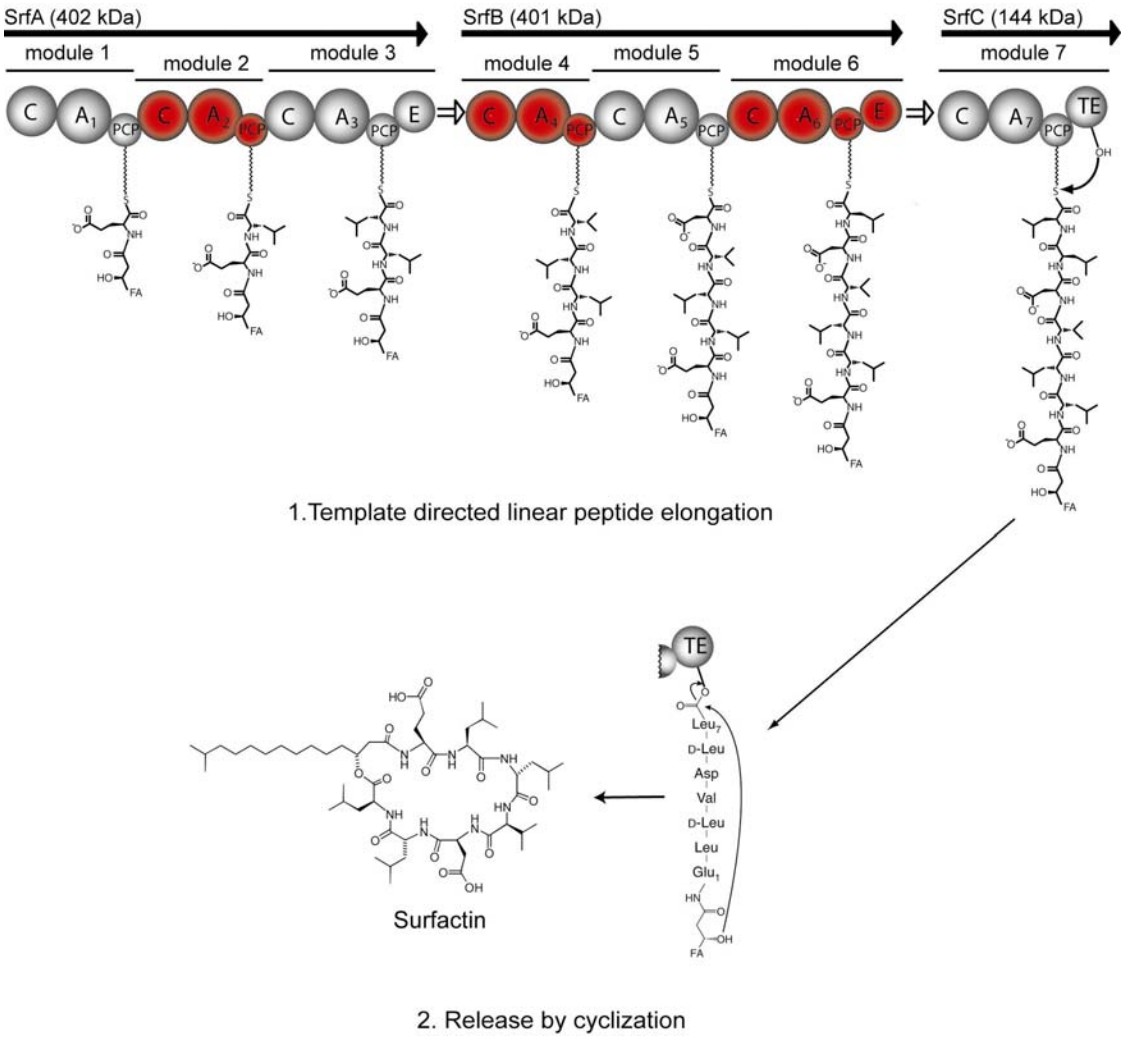


Figure 2.2. Surfactin assembly line. The multienzyme complex consists of 7 modules (grey and red) which are specific for the incorporation of 7 amino acids. 24 domains of 5 different types (C, A, PCP, E, and TE) are responsible for the catalysis of 24 chemical reactions. 23 reactions are required for peptide elongation while the last domain is unique and required for peptide release by cyclization.

The carboxy group of amino acid building blocks is first activated by ATP hydrolysis to the corresponding aminoacyl-adenylate. This reactive intermediate is transferred onto the free thiol-group of an enzyme bound 4´-phosphopantetheinyl cofactor (ppan), establishing a covalent linkage between enzyme and substrate. At this stage the substrate can undergo modifications such as epimerisation or N-methylation. Peptide assembly of the final product occurs then by a series of peptide bond formation steps (elongation) between the downstream building block with its free amine and the carboxy-acitvated thioester of the upstream substrate resulting in an N to C terminal elongation. The ppan cofactor facilitates the ordered transfer of thioester substrates between catalytically active units with all intermediates covalently tethered to the multienzyme until the product is released by the action of the C-

terminal TE-domain (termination). This principle minimizes side reactions as well as diffusion times, maximizing the catalytic efficiency. Type I polyketide synthases (PKS) and fatty acid synthases (FAS) also display a multienzymatic organization and catalyse a repeated reaction cycle of decarboxylative condensation of smaller acyl groups[43, 44].

2.3 Activation by the Adenylation Domain

Each nonribosomal peptide synthesis initiates by specific recognition and activation of the dedicated amino acid from a pool of substrates by the ca. 550 amino acid comprising adenylation domain (A-domain). For example, each of the seven amino acids found in the heptapeptide surfactin is respectively selected for incorporation into the growing peptide chain by one of the seven A-domains of the surfactin synthetase (Fig. 2.2). Substrate activation is achieved in a two-step chemical reaction. First, after binding of the cognate amino acid, the enzyme catalyses activation as an aminoacyl adenylate intermediate at the expense of Mg^{2+}-ATP and release of PP_i (Fig. 2.3). Second, the amino acid-O-AMP oxoester is converted into a thioester by a nucleophilic attack of the free thiol-ppan cofactor of the adjacent PCP domain, which will be discussed below (Fig. 2.4). This mechanism resembles amino acid activation catalysed by aminoacyl-tRNA-synthetases although these enzyme families share neither sequence nor structural relations[45]. Many A-domains can be heterologously expressed in *E. coli* and their activity as well as substrate specificity can be assayed *in vitro* by an equilibrium ATP-PP_i exchange assay with radiolabeled PP_i[46].

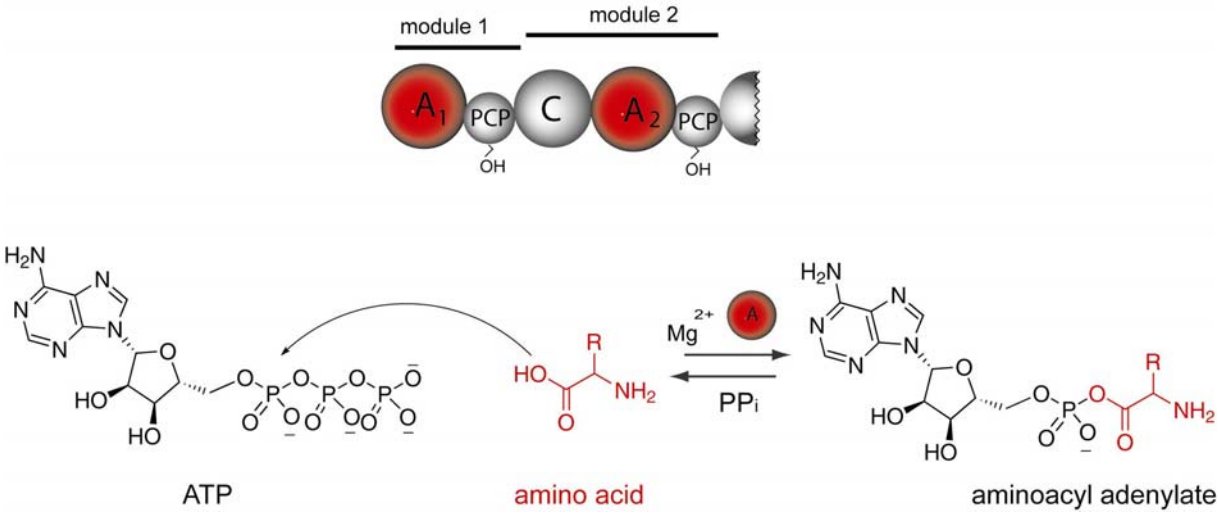


Figure 2.3. Amino acid activation. The adenylation domain activates a dedicated amino acid as aminoacyl adenylate at the expense of ATP.

2.4 Intermediates Transport by the Peptidyl Carrier Protein

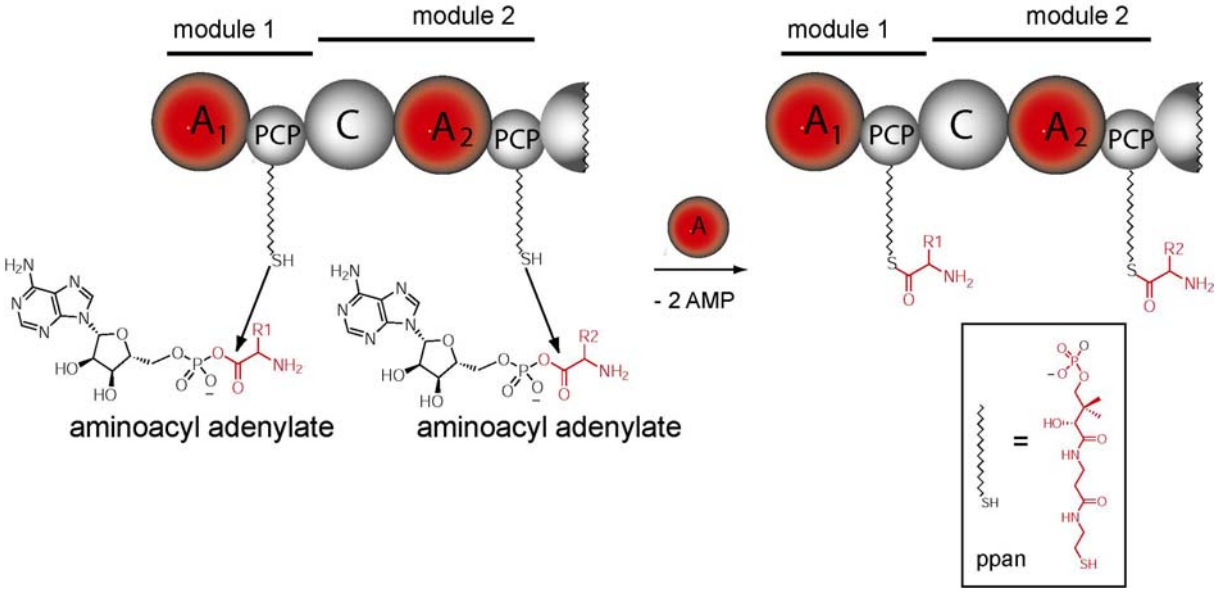


Figure 2.4. Thiolation of the activated amino acids.

The only NRPS domain without autonomous catalytic activity is the peptidyl carrier protein (PCP), also referred to as thiolation domain (T). The protein comprises ca. 100 amino acids and is located downstream of the A-domain. Within the NRPS assembly line PCP is responsible for transportation of substrates and elongation intermediates to the catalytic centres. As mentioned in the previous chapter, the A-domain catalyses in the second half-reaction the transfer of the activated aminoacyl-adenylate substrate onto the terminal cysteamine thiol group of the ppan cofactor bound to PCP (Fig. 2.4). In nonribosomal peptide synthesis the combination of A-domain and PCP is defined as initiation module, since both domains are required to activate and covalently tether the first building block for subsequent peptide synthesis. The activity of recombinant A-domains with adjacent holo-PCPs can be assayed *in vitro* by an aminoacylation assay with ATP and radioactive amino acids[46]. In contrast to ribosomal protein synthesis with tRNA bound oxo-ester intermediates, nonribosomal peptide synthetases use more reactive PCP-thioesters. This difference in reactivity is due to the lower mesomeric stabilization of the thioester, which forms less stable p_π-p_π double bonds and emphasizes that initial activation energy provided by the A-domain is preserved here for subsequent catalytic reactions such as condensation, hydrolysis and cyclization (see below)[47].

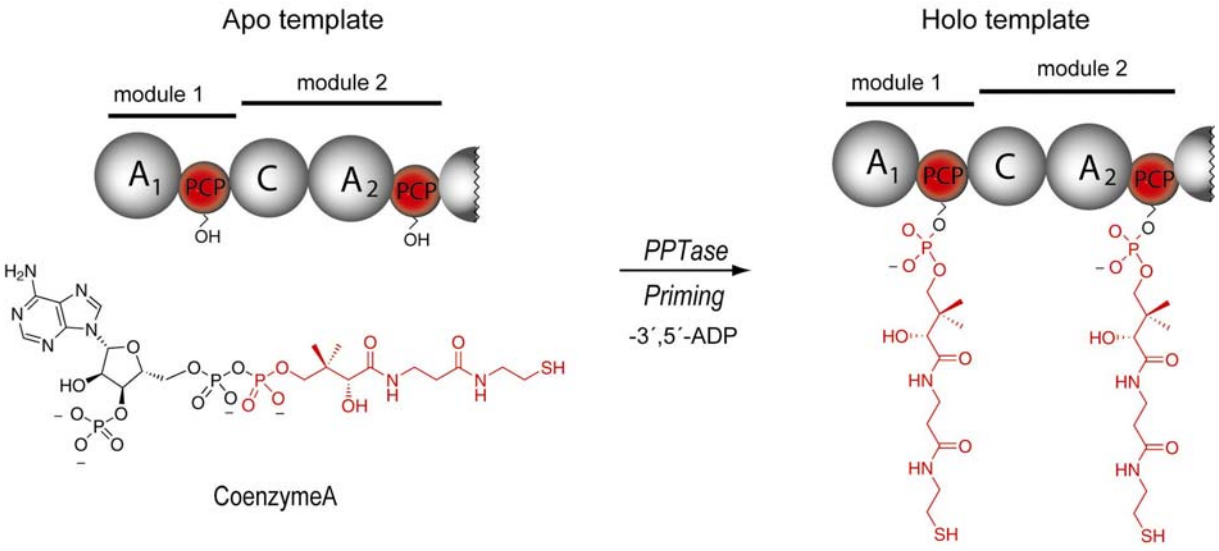


Figure 2.5. Priming of the enzymatic template. The phosphopantheteine moiety of coenzyme A is covalently attached onto the invariant serine residue of PCP by dedicated phosphopantheteine transferases.

The 20 Å long prosthetic ppan moiety of coenzyme A is covalently tethered to the side chain of a strictly conserved PCP serine residue and serves as a crane for building block delivery[48-51]. Transfer of ppan onto apo PCP is catalysed by NRPS specialized 4´-phosphopantetheinyl transferases such as Sfp and Gsp from *B. subtilis* and *B. brevis*, respectively[52-54] (Fig. 2.5). The conversion of inactive *apo*-PCP into its active ppan-PCP *holo*-form was monitored *in vitro* with recombinant Sfp and PCPs from the surfactin synthetase. These studies revealed very low selectivity of Sfp for the carrier proteins[55]. Sfp was shown to efficiently phosphopantetheinylate not only apo-PCPs from various NRPS systems but also acyl carrier proteins from fatty acid and polyketide synthases[52, 56]. In addition, also aminoacyl- and acetyl-CoAs were tolerated[57, 58]. This relaxed specificity of Sfp was useful for preparative applications which will be presented in this work[59, 60]. Insights into how Sfp mediates binding and protein recognition were provided by a crystal structure in complex with its substrate CoA[61] (Fig 2.6). The structure of the 224 amino acid comprising Sfp monomer shows a pseudo 2-fold symmetry, which divides the protein into two similar folds of almost identical size. The CoA substrate is bound in a bent conformation within a pocked formed by the two Sfp halves. The 3´-phospho-5´-ADP moiety of CoA is well defined in the electron density map and is coordinated by several Sfp residues and Mg^{2+}, while the main part of the ppan arm displays no interactions with Sfp and points out into bulk solvent.

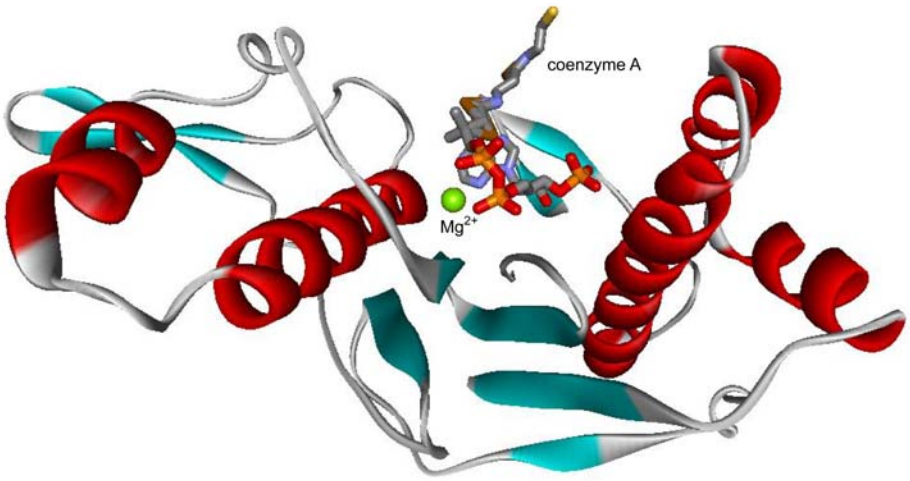


Figure 2.6. Crystal structure of Sfp. Cocrystal with bound coenzyme A and Mg^{2+}.

2.4.1 Misacylation and Regeneration

The promiscuity of Sfp causes also undesired misacylation of PCPs within NRPS assembly lines since not only CoA but also acyl-CoAs can serve as cofactor donors. In order to regenerate these misprimed NRPS templates (acetyl-ppan), nature has developed specific enzymes, which catalyse hydrolysis of the undesired acyl group. These so called thioesterase II domains (TEII) were shown to be specific for acyl-PCPs in *in vitro* assays, while there was no hydrolysis observed for acyl-ACPs which are essential in primary fatty acid metabolism. Comparison of the catalytic properties of TEII mediated aminoacyl- or peptidyl-PCP hydrolysis vs. acetyl-PCP showed a strong preference for the latter substrate which indicates that this proofreading enzyme is important for NRPS activity by deblocking misacetylated PCPs which was also confirmed by TEII knockout studies[62].

2.5 Peptide Elongation by the Condensation Domain

After activation and covalent binding of the first substrate amino acid by the A-PCP initiation module, peptide synthesis proceeds by stepwise condensation with amino acid building blocks bound to PCPs of several downstream elongation modules $(C\text{-}A\text{-}PCP)_n$. Peptide bond formation is mediated by a ca. 450 amino acids comprising condensation domain (C-domain). The C-domain catalyses the nucleophilic attack of the downstream PCP bound acceptor amino acid with its free α-amino group on the activated thioester of the upsteam PCP-bound donor amino acid or peptide[57] (Fig. 2.7). In surfactin synthetase, peptide bond formation occurs e.g. between modules 1 and 2 by the nucleophilic attack of the α-amino group of

leucine on the thioester-activated carboxy group of glutamate to give a dipeptide which is translocated to module 2 (Fig. 2.2).

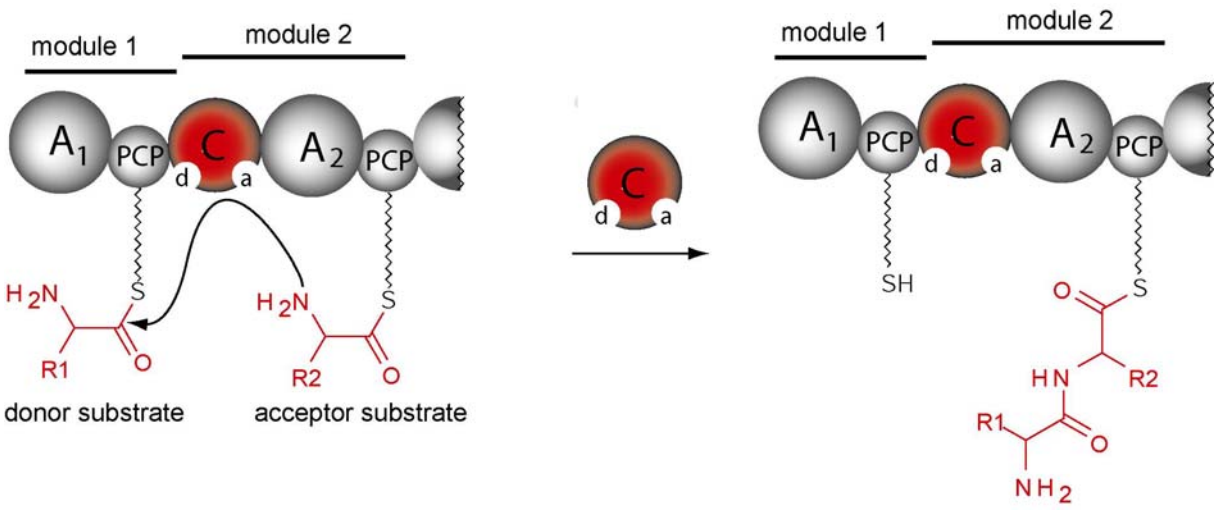


Figure 2.7. Peptide bond formation by the C-domain. Putative donor (d) and aceptor (a) sites are marked by circles.

Biochemical characterization of different C-domains revealed insights into their substrate specificities. Probing substrate specificity in the natural synthetase is difficult to achieve because the upstream donor and downstream acceptor substrates are defined by restrictive A-domains. In order to directly evaluate C-domain specificity, various aminoacyl-CoA substrates were synthesized and attached to the apo-PCPs of a minimal, bimodular NRPS system composed of module 1 from *B. brevis* gramicidine synthetase and module 2 from tyrocidine synthetase under catalysis of Sfp. Mischarging of PCPs from both modules revealed in condensation assays that the C-domain of tyrocidine module 2 seems to possess an acceptor position for the downstream PCP bound nucleophile, that discriminates against the non cognate D-enantiomer, as well as differences in the side chain. By contrast, low side chain selectivity was observed for the donor position of the upstream PCP bound electrophile. Interestingly, a preference for cognate D-enantiomers were observed[57]. This was confirmed by further investigations with the C-domain of tyrocidine elongation module 5 which revealed that the donor position exclusively selects a tetrapeptide with the cognate D-configuration of the C-terminal residue for condensation reactions[60]. This shows that besides A-domains and TEII-domains, C-domains represent a selectivity filter in nonribosomal peptide synthesis. Selection of the correct downstream nucleophile by the acceptor position prevents the formation of product mixtures and facilitates peptide synthesis in a directed manner. Sequence alignments of several C-domains revealed a highly conserved HHXXXDG motif that is also found in acyltransferases such as chloramphenicol acetyltransferase (CAT), NRPS

epimerisation and heterocyclization domains[63, 64]. Mutations in residues His147 in the conserved motif of the C-domain of tyrocidine module 2 and His138 in the motif of enterobactin synthetase module F revealed no activity in condensation assays[65, 66] which emphasizes the active role of the conserved histidine in catalysis.

2.6 Editing Domains

While the amino acid is covalently tethered onto the PCP several editing domains can carry out further modifications to increase the diversity of the final product. Structural features like D-amino acids and methylated peptide bonds increase the stability of the peptide against proteolytic digest and also contribute to unique conformations important for biological activity.

2.6.1 Epimerization

Almost every nonribosomally synthesized peptide contains D-configurated amino acids to various extends. NRPSs realize two different ways for their incorporation. The most common way is the epimerisation of L-amino acids by integrated 450 amino acids comprising epimerisation domains (E)[67] which racemize the C_α-carbon of the PCP tethered aminoacyl substrate to a D/L equilibrium[68] (Fig. 2.8). Racemization can either occur from L to D or the opposite way. Rapid quench kinetics revealed that this equilibrium is achieved within seconds[69]. Specific incorporation of only the D-amino acid into the growing peptide chain is ensured by the enantioselective donor site of the downstream condensation domain[57]. This principle is also used in the surfactin synthetase in modules 3 and 6 to incorporate two times D-Leu in the final product. The combination of D- and L-amino acids contributes to the unique conformation of surfactin that is important for its biological activity[10]. A second strategy of D-amino acid incorporation is often observed in fungal systems[32]. The A-domain of cyclosporin synthetase for example, exclusively incorporates D-Ala at position one which is provided by an external racemase[70].

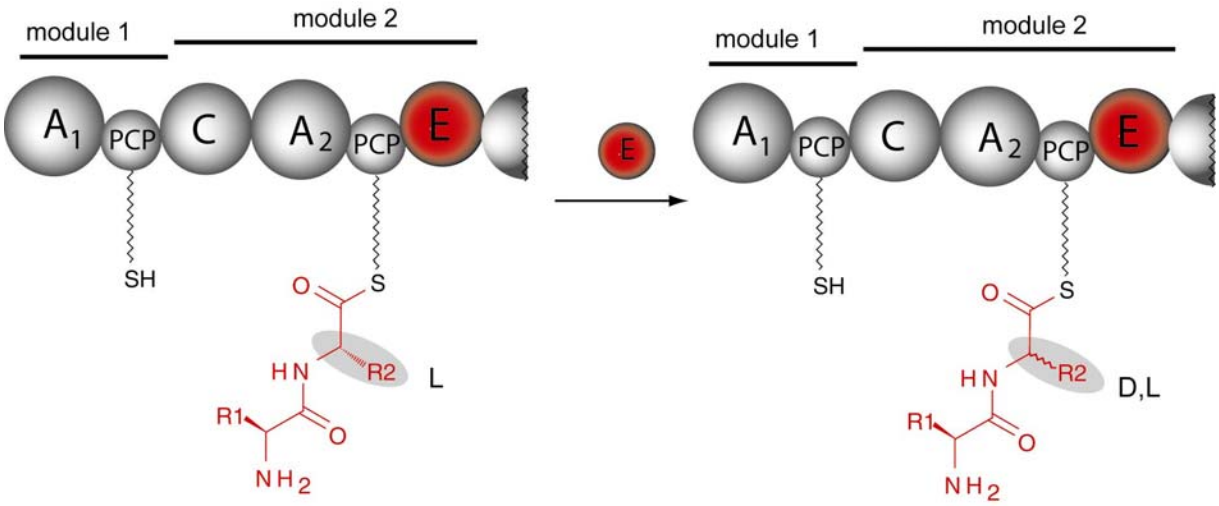


Figure 2.8. Epimerisation of the C-terminal amino acid by the E-domain. The stereocenter of the shaded residue is epimerised into a D,L equilibrium.

Biochemical characterization of E-domain substrate specificity revealed that also noncognate amino acids were racemized but with lower efficiency[71]. Further studies showed that artificial E domain constructs without preceding C-domain (as it would be observed in the native initiation module) could epimerise aminoacyl-PCP while in contrast the identical constructs with a preceding cognate C-domain (as in an elongation module) did not show epimerisation activity for the bound aminoacyl-S-ppan substrate. This led to the conclusion that C-domains bind tightly aminoacyl-PCP in the acceptor site until condensation occurs. The resulting peptidyl-PCP has a lower binding affinity to the acceptor site and is then transferred to the subsequent E-domain or next C-domain[60, 72]. These investigations contributed to the understanding of timing and directionality of nonribosomal peptide synthesis (Fig. 2.9).

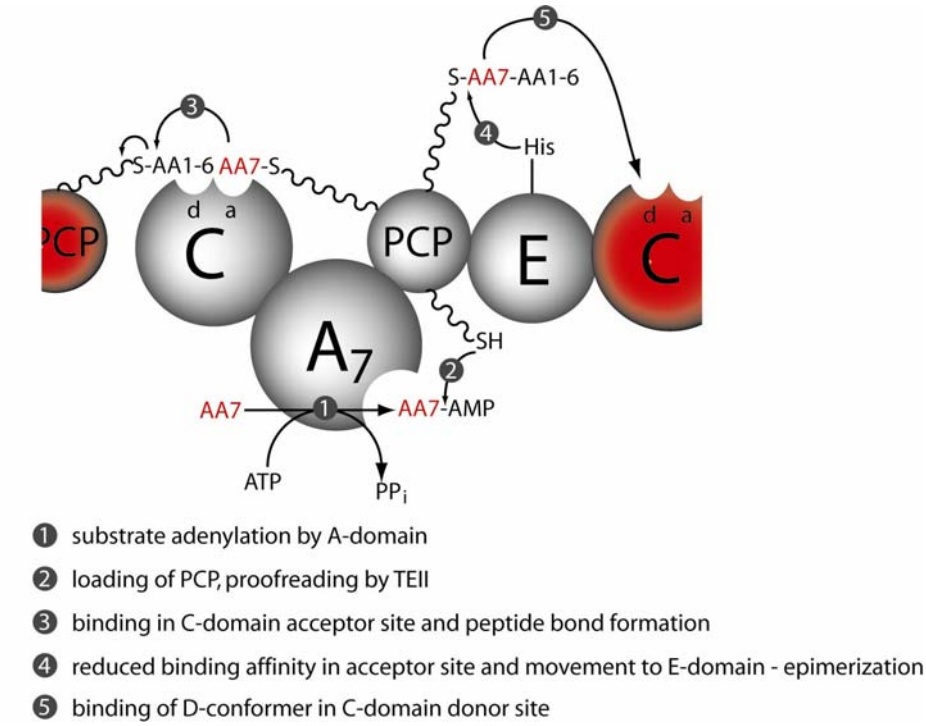


Figure 2.9. Directionality of peptide elongation.

2.6.2 Methylation

Some nonribosomal peptides such as cyclosporin[73], enniatin[74], actinomycin[75], and pristinamycin[11, 76] have N-methylated peptide bonds. This modification is introduced by a ca. 420 amino acid comprising N-methylation domain (N-Mt) which is inserted into the accompanying A-domain. The N-Mt-domain catalyses the transfer of the S-methyl group of S-adenosylmethionine (SAM) to the α-amino group of the thioesterified amino acid releasing S-adenosylhomocysteine as a reaction by-product. In comparison to other NRPS domains discussed previously less is known about N-Mt-domains.

2.7 Peptide Release

All catalytic domains discussed so far are repeating units of the enzymatic template and contribute to the synthesis of a linear peptide molecule tethered to the multienzyme. In order to reactivate the multienzyme machinery for the next synthesis cycle the mature peptide has to be cleaved once it reaches the end of the assembly line. This reaction is usually accomplished by a 280 amino acid comprising thioesterase domain (TE-domain) fused to the C-terminal module, also referred to as termination module. The peptide can either be released by hydrolysis as a linear acid or by an intramolecular reaction with an internal nucleophile to give a cyclic peptide (Fig. 2.10). Hydrolytic release is observed e.g. for vancomycin whose

peptide backbone is constrained by further postsynthetic oxidative cross linking reactions[12]. Alternatively, peptide release can also occur under reduction of the carboxy-group catalysed by the NADPH dependent reduction domain (R) to give linear aldehydes or alcohols such as in the yeast biosynthetic pathway for the essential amino acid lysine[77] and in the biosynthesis of linear gramicidin A in *B. brevis*[78] or by head to tail condensation mediated by the C-domain as seen in cyclosporin synthetase.

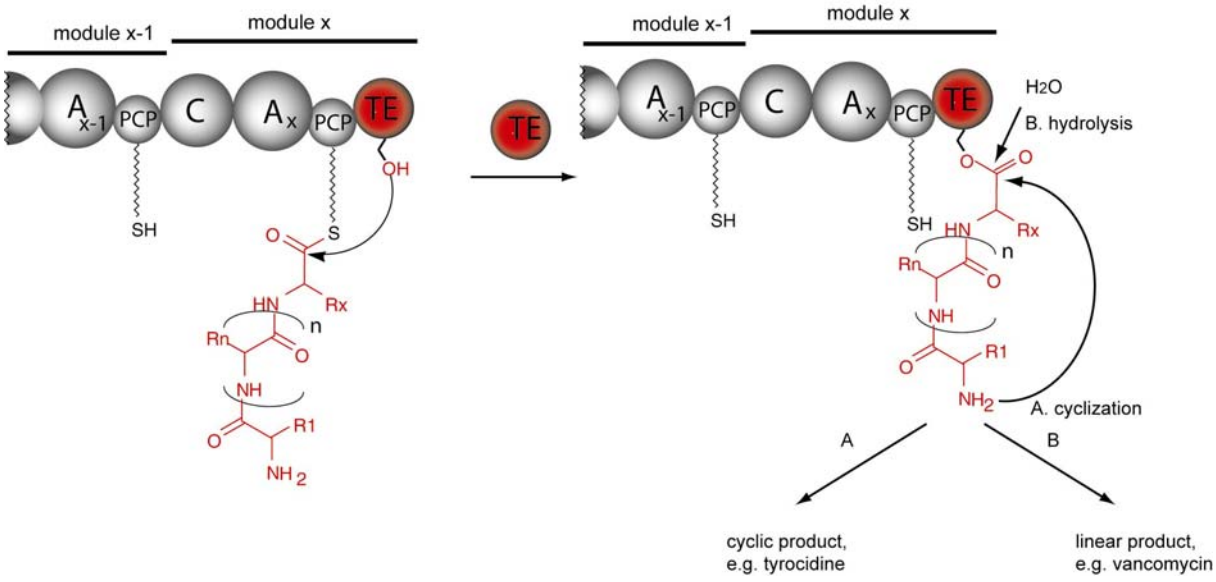


Figure 2.10. Release of the linear peptide chain by the TE-domain. Release can either be carried out by (A) cyclization or (B) hydrolysis.

In case of surfactin the peptide backbone is constrained by an intramolecular nucleophilic attack of a β-hydroxy group of the fatty acid moiety to give a branched chain lipodepsipeptide[79] (Fig. 2.2). Many other cyclization strategies are known so far giving rise to a large and diverse set of cyclic or cyclic branched molecules with distinct biological activities. Thioesterase domains catalyzing a cyclization reaction are also referred to as peptide cyclases and the main emphasis of this chapter is attributed to them.

2.7.1 Cyclization Strategies: Nature´s Way

All known NRPS macrocyclization strategies lead to cyclic or cyclic branched chain peptides (Fig. 2.11). In macrolactones the branch point can either be a hydroxylated amino acid side chain or a hydroxylated fatty acid moiety. In case of the surfactin peptide cyclase the ring closure is enzymatically catalysed between a N-terminal β-hydroxyl fatty acid and the C-terminal peptide end[79]. Cyclization was only observed when the (R)-β-hydroxy fatty acid was used while the (S)-enantiomer only showed enzymatic hydrolysis indicating

stereoselective recognition[80]. In contrast, syringomycin (*P. syringae*) and fengycin (*B. subtilis*) lipopeptide cyclases accept serine and tyrosine side chains of the peptide sequence as nucleophiles for cyclization, discriminating the N-terminal β-hydroxyl group of the attached fatty acid moiety[15, 26, 81](Fig. 2.11). Moreover, these peptide cyclases display a very high regioselectivity by selecting only one specific residue of the substrate from a large source of nucleophiles for cyclization.

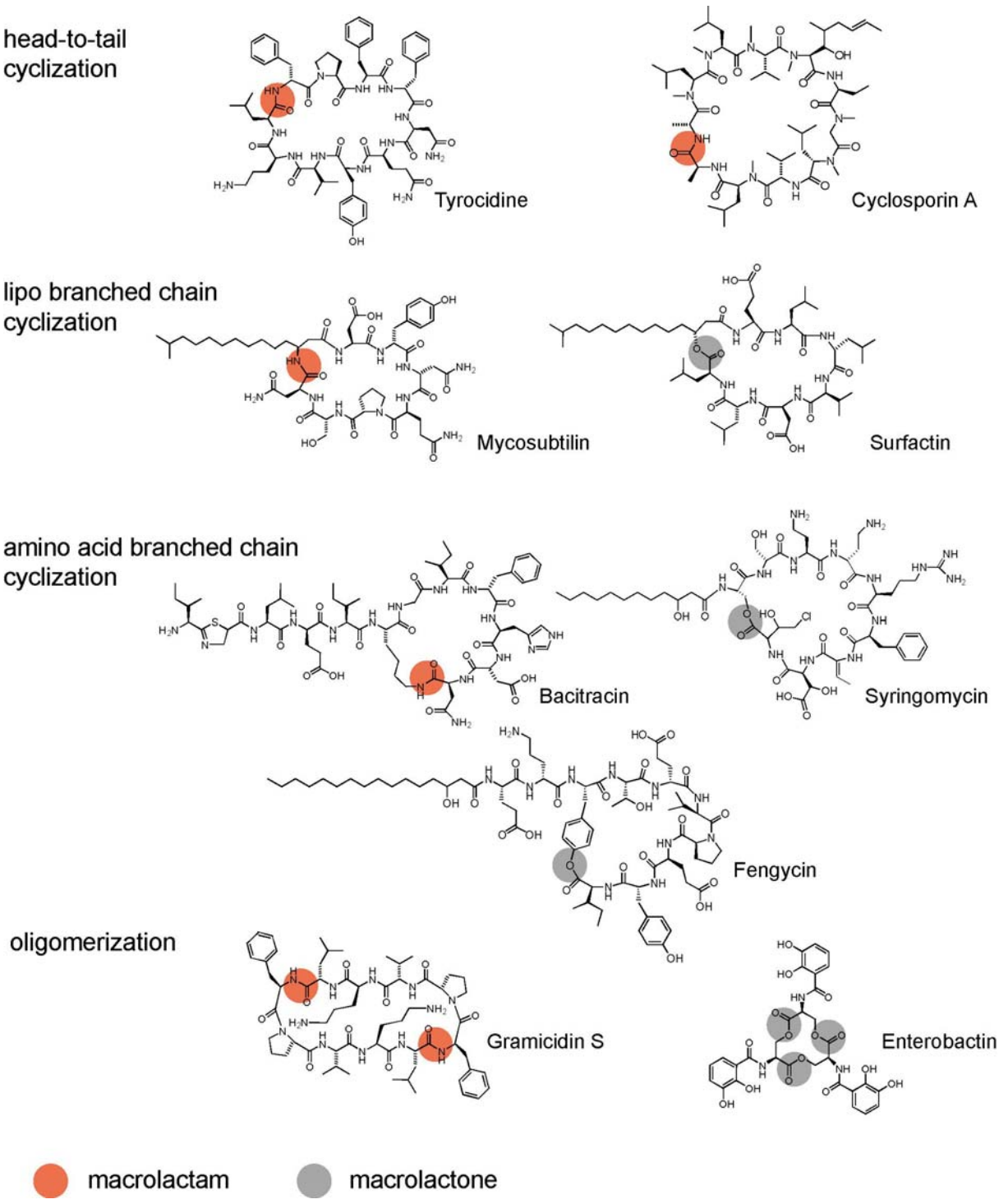


Figure 2.11. Macrocyclization strategies.

In addition to macrolactonization, natural product diversity is also increased by various enzymatic macrolactamization strategies. Basic head-to-tail peptide macrolactamization is observed in the antibiotic tyrocidine from *Bacillus brevis* and the potent immunosuppressive drug cyclosporin A from *tolypocladium niveum*. In cyclosporin A the final peptide bond is formed by a putative condensation domain[82] instead of a peptide cyclase, emphasizing that nature developed two different enzyme species capable of catalysing product release by cyclization. Recently, a new type of head to tail macrocyclization was reported for nostocyclopeptide from the terrestrial cyanobacterium *Nostoc* sp. The C-terminal residue of the linear peptide is reduced by the action of a R-domain to give an aldehyde which is then intramolecularly captured by the α-amino group of the N-terminal amino acid residue by an unknown mechanism to give a cyclic imine[83].

Besides head-to-tail cyclized lactams also branched chain lactams are observed. The peptide cyclase of the antifungal lipopetide mycosubtilin from *B. subtilis* forms regioselectively an amide bond between a N-terminal β-amino fatty acid and the peptide C-terminal end (Fig. 2.11). Similar to surfactin also in mycosubtilin a fatty acid chain is involved in the cyclization process providing an amine as a nucleophile[23]. While the precursors in both cases seem to be β-keto fatty acid residues derived from fatty acid synthases, nature processes these ketones in two different ways. They are either reduced to a hydroxyl-group as observed in surfactin or they are reductively aminated catalysed by amino transferases as observed in mycosubtilin. In both cases nature uses a common precursor motive which is subsequently diversified by the application of different synthetic strategies to increase the product outcome. Similar to fengycin and syringomycin, enzymatic amide bond formation can also occur between an ornithine side chain and the C-terminal peptide end as seen in the antibiotic bacitracin from *B. licheniformis*[17]. The scope of cyclic lactams and lactones can be further broadened by oligomerization of peptide monomers. This additional strategy enables the *B. brevis* gramicidin S peptide cyclase to cyclodimerize two linear pentapeptides by catalysing two subsequent peptide bond formation steps to form the cyclic antibiotic dilactam gramicidin S[80] (Fig. 2.11). Also cyclotrimerization is observed for the siderophore forming peptide cyclases of enterobactin and bacillibactin synthetases from *E. coli* and *B. subtilis*, respectively[84, 85]. Three units of 2,3-dihydroxybenzoyl-serine are fused together by three subsequent ester bond formation steps between the serine hydroxyl-group of one molecule and the C-terminal end of another molecule to give the cyclic trilactone enterobactin. This trilactone displays iron chelating activity which is closely related to its structure. Three catechol ligands provide electron donors required for the coordination of iron once more

emphasizing the close relationship between cyclic structural organization and biological activity. By contrast to enterobactin, it is postulated that the cyclic trimer enniatin is cyclized by a PCP-C didomain replacing the TE-domain and allowing the successive build up of oligomers on the PCP before C-domain mediated cyclization[86, 87].

NRPS peptide cyclases can generate diverse cyclic peptide molecules ranging in size from very small as in pristinamycin with seven residues to very large as in syringopeptin with 22 residues[76, 88]. At the large end of the scale there is also another source of macrocyclic molecules observed in nature referred to as naturally occurring circular proteins[89]. These proteins are of bacterial origin and have a well folded three dimensional structure. In contrast to NRPS they are produced by the translation of genes. Cyclization occurs posttranslationally only in a head-to-tail fashion to produce a seamless circle of peptide bonds. Unlike NRPSs not much is known about the cyclization mechanism of the linear precursors.

Cyclization strategies reported here emphasize that nature has developed a large enzymatic tool set which allows to introduce diversity into linear peptide sequences by a variety of different cyclization steps. The selection from different nucleophiles, epimers and positions in the peptide sequence makes peptide cyclases unique enzymatic tools, with very specific intrinsic stereo- and regioselective recognition elements. Moreover, to understand the catalysis of one, two or three subsequent condensation steps towards cyclization requires further studies on structural and mechanistic aspects of these enzymes.

2.7.2 Chemoenzymatic Cyclization

The great pharmacological potential of many cyclic peptides emphasizes their role in drug discovery, as they show specific interaction with cellular receptors and high resistance to proteolytic enzymes. They are therefore most promising scaffolds for phamacophores. So far, synthetic chemistry faces several difficulties in the production of cyclic compounds providing sufficiently good yields and regioselectivity. Synthetic macrocyclization is difficult to achieve since steric repulsion and misalignment of ring residues as well as the use of protecting groups to ensure proper regiochemistry decreases yields and makes chemical synthetic operations expensive and rather difficult. Since nature developed stereo- and regioselective peptide cyclization enzymes, researchers aimed to combine chemical linear peptide synthesis with enzymatically catalysed cyclization. This approach allows easy synthesis of linear peptide sequences by established solid phase peptide chemistry followed by selective and efficient enzymatic cyclization without the use of protecting groups and the formation of undesired by-products.

Similarities and differences between nonribosomal peptide synthesis and chemical solid phase peptide synthesis are highlighted in Figure 2.12. Both methods use a similar way of precursor activation and tethering on solid support, which facilitates a rapid and ordered synthesis of desired products in high yields. In chemical synthesis chain elongation proceeds from the C- to the N-terminal end while it is the opposite direction in enzymatic synthesis. One advantage of chemical synthesis is the huge diversity which can be incorporated into the linear peptide chain in order to create a variaty of substrate analogues which is desired in biochemical studies of excised TE domains.

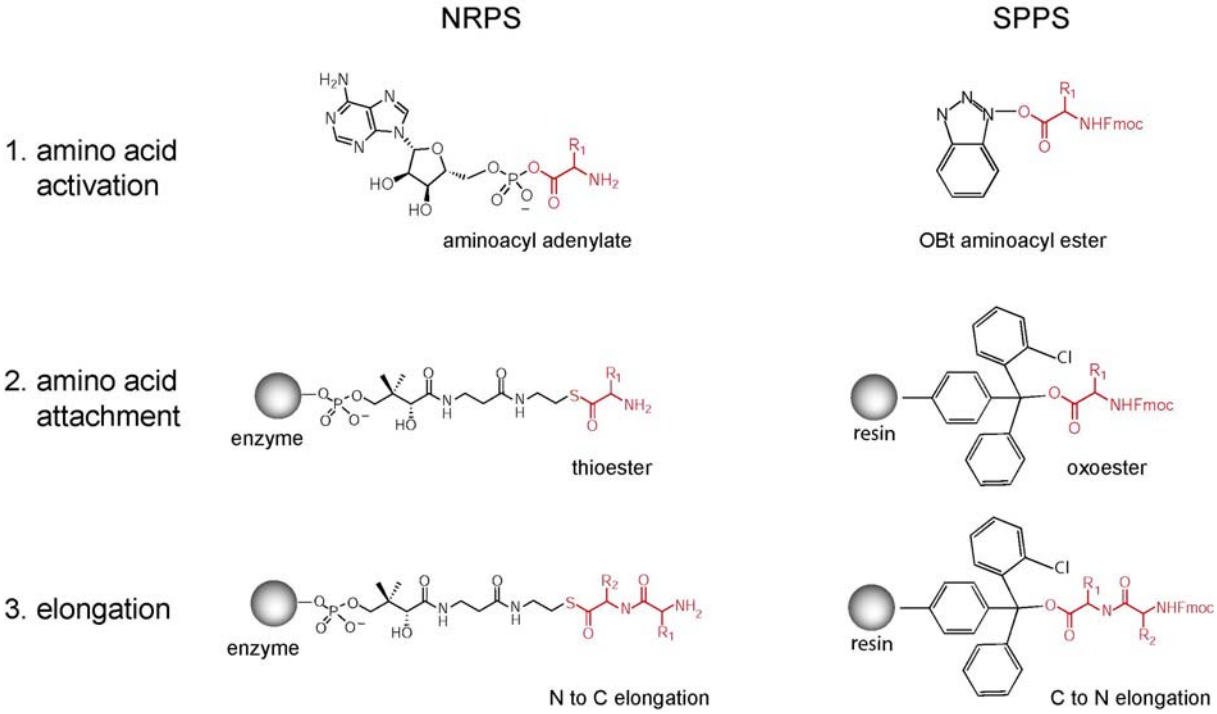


Figure 2.12. Comparison between nonribosomal peptide synthesis (NRPS) and solid phase peptide synthesis (SPPS). Both methods use active ester species for amino acid activation and tether these activated species on solid support as thioester (NRPS) or oxoester (SPPS) intermediates. While enzymatic peptide synthesis depends on the enzymatic specificity of the template, chemical synthesis allows flexibility for various amino acid substitutions, required for chemoenzymatic investigations.

In order to achieve chemoenzymology a translation between the language of chemistry and the language of biology has to be established by chemically mimicking the biological pathway as close as possible (Fig. 2.13 A and B). This was first achieved by a cooperation of the Walsh and Marahiel groups who cloned and overexpressed excised peptide cyclases (28 kDa) from the tyrocidine and surfactin synthetase[80, 90]. In order to prove the activity of the isolated enzyme vs. the natural enzyme which is embedded in a multienzyme complex, a short mimic of the natural cofactor phosphopantetheine, N-acetylcysteamine (SNAC), was attached

to the C-terminal end of chemically synthesized linear surfactin and tyrocidine peptides. The SNAC represents a link between natural and artificial systems compatible with both. Incubation of tyrocidine substrate mimic and excised tyrocidine peptide cyclase revealed activity with an observed cyclization to hydrolysis ratio of 6:1 (surfactin cyclization will be presented in the results section). The turnover of 59 min^{-1} indicated a very rapid conversion of the linear into the cyclic compound which makes this enzyme as an *in vitro* biocatalyst useful[90]. Follow up studies with various peptidyl-SNAC substrates with varied length, stereochemistry and amino acid composition revealed that the tyrocidine peptide cyclase only recognizes C- and N-terminal residues of the substrate by identity and stereochemistry, leaving space for making longer and shorter substrates as well as for replacements of single residues within the peptide backbone[80]. Alteration of backbone parts by replacement of three amino acid blocks with a flexible spacer or by replacement of amide bonds with ester bonds suggested that during the cyclization reaction the peptide substrate is preorganized by intramolecular hydrogen bonds analogous to those in the product tyrocidine. A minimal recognition model was postulated[91]. An important feature of cyclic peptides is the β-sheet content, which is high for molecules with (4n + 2) residues[80, 91]. A peptide with a high β-sheet content such as the decapeptide tyrocidine (n = 2) facilitates cyclization through substrate preorganization by backbone-to-backbone hydrogen bonds. This intrinsic property of tyrocidine facilitates easy cyclization which was also reported to occur un-catalysed but less efficiently[92]. A large number of changes in the peptide thioester substrate were incorporated for the synthesis of a cyclic integrin inhibitor retaining only the elements in the minimal recognition model for tyrocidine cyclization. Although 7 of 10 cognate residues were replaced by others with 2 also differing in stereochemistry, tyrocidine cyclase was still capable of cyclization but with reduced yield[93]. This method was extended by the synthesis of linear novel hybrid polyketide-tyrocidine substrates, which were cyclized in good yields[94]. Substrate tolerance of the excised tyrocidine cyclase allowed furthermore the synthesis of diverse tyrocidine variants in which position 4 (D-Phe) was replaced by one of 96 natural and unnatural amino acids. This library of tyrocidine product analogues was subsequently screened for improved or altered bioactivity. In contrast to the natural antibiotic tyrocidine A which does not discriminate between bacterial and eukaryotic cell membranes, the screen revealed that a substitution of D-Phe at position 4 by a positively charged D-amino acid led to a 30-fold increase in the selective recognition of bacterial membranes[95].

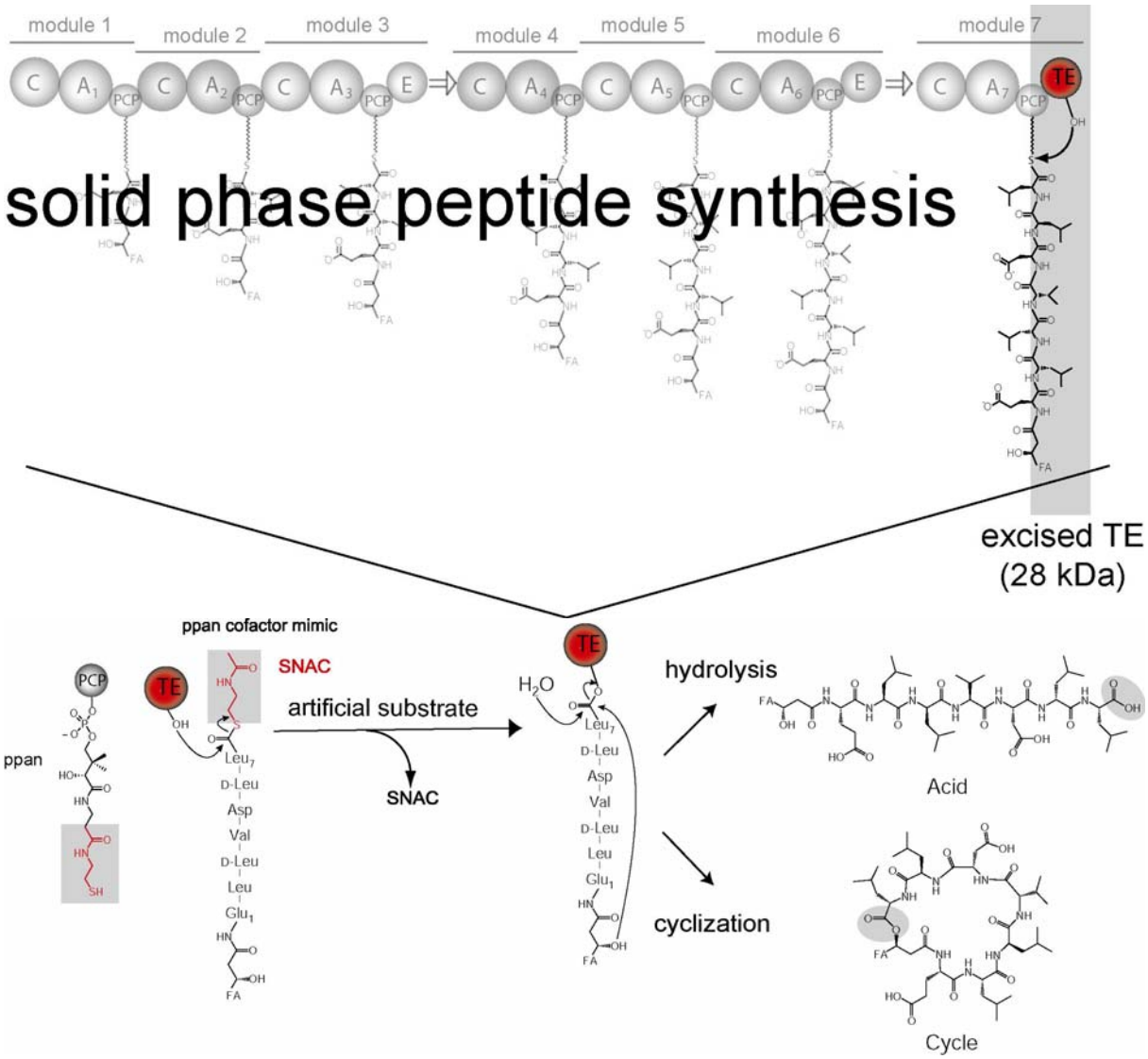


Figure 2.13. Chemoenzymatic cyclization. The NRPS multienzyme machinery required for peptide elongation is replaced by solid phase peptide synthesis and the TE domain is excised as a free standing enzyme. Recognition of the artificial substrate by the enzyme is ensured by the ppan cofactor mimic SNAC.

To expand the set of cyclization catalysts also peptide cyclases from other NRPS systems like mycosubtilin, fengycin from *B. subtilis* and syringomycin from *P. syringae* were recently cloned and overexpressed. Contrary to previous characterized cyclases, no activity was observed for those of mycosubtilin, fengycin and syringomycin cyclases with synthetically made peptidyl-SNAC substrates, which indicates a limitation in the chemoenzymatic potential of the latter cyclases. The inability to recognize or bind the SNAC substrates in the active site of the excised peptide cyclase could be affected by the way of presenting the short SNAC mimic to the enzyme and is subject of investigations presented in this work.

2.7.3 Structural and Mechanistic Aspects of NRPS Peptide Cyclases

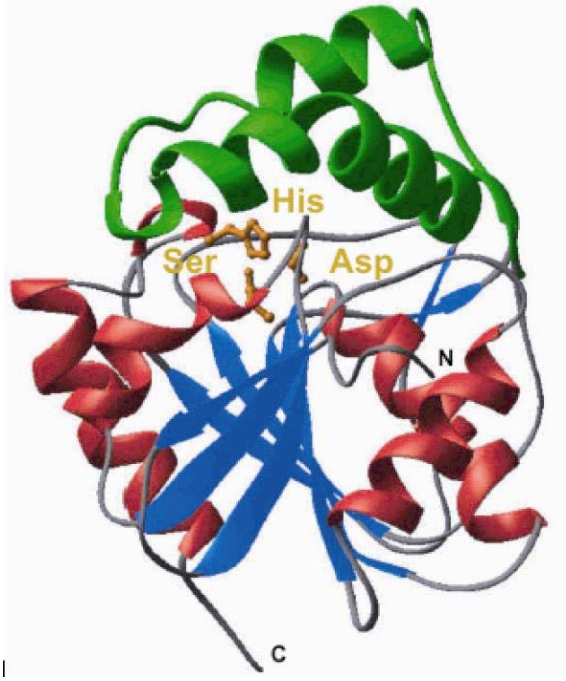


Figure 2.14. Crystal structure of the Srf TE domain. The Lid region is indicated in green.

The three dimensional organization of enzyme residues encodes all information required to understand principles and general features of macrocyclization. Regiospecific selection of only one nucleophile for cyclization as well as the exclusion of water to prevent undesired hydrolysis are features which are embedded in the structural fold. Moreover the question how thioesterases from different NRPS systems catalyse termination in one case by cyclization and in the other case by hydrolysis needs to be elucidated. Crystallographic data of the excised surfactin peptide cyclase (TE) showed that this enzyme is a member of the α,β-hydrolase family[96] (Fig. 2.14). The structural similarity to serine hydrolases suggested that an active site catalytic triad is responsible for the macrocyclization activity. This is in agreement with a structure model of the surfactin cyclase which suggests that the PCP-ppan bound peptidyl chain is directed via a cleft into the active site of the peptide cyclase and transferred onto an invariant serine residue (Ser80), which is activated by histidine (His207) and aspartate (Asp107)[96] (Fig. 2.15). The peptidyl chain of this acyl enzyme intermediate is accommodated in a predominantly hydrophobic binding pocket with two cationic residues (Lys111 and Arg120) predicted to direct cyclization through specific interactions with the substrate. In the deacylation step of the reaction, the β-hydroxyl group of the fatty acid moiety is activated by the same histidine and aspartate to facilitate an intramolecular nucleophilic attack on the acyl enzyme ester bond to release the final lactone product. Some thioesterase domains from other NRPS systems e.g. vancomycin as well as structurally related lipases only hydrolyse but do not cyclize their products. It has remained unclear what features direct

the acyl-O-TE intermediates between cyclization and hydrolysis. A sequence alignment between the surfactin peptide cyclase and other members of the α,β-hydrolase enzyme family, that catalyse only hydrolysis, revealed a glycine residue, involved in the formation of the oxyanion hole to be conserved among lipases. In the surfactin peptide cyclase a rigid proline residue is at this position, which might be important for cyclization (Fig. 2.15).

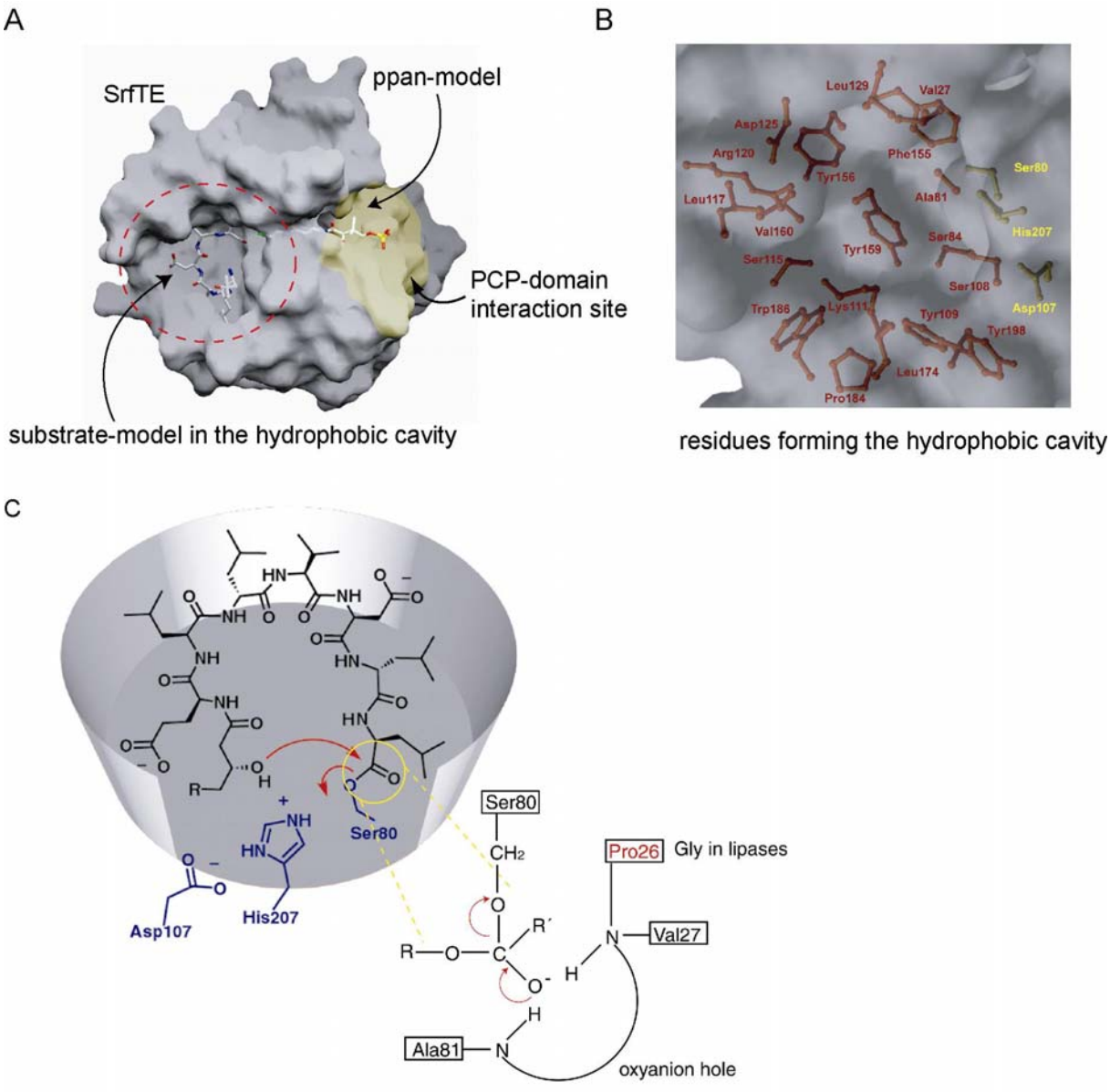


Figure 2.15. Catalytically important residues of Srf TE. (A) The peptide substrate and the ppan arm are modelled into a hydrophobic binding pocket and into a substrate channel, respectively. (B) Magnification of the binding pocked shows the catalytic triad (Ser80, His207 and Asp107) and several hydrophobic residues with the exception of Lys111 and Arg120. (C) Proposed oxyanion hole with Pro26 as a rigid gatekeeper.

In the surfactin assembly line the template modular sequence is collinear with the product peptide sequence which is also observed for many other NRPS systems. In these linear type A

NRPS assembly lines, TE domains catalyse only one reaction step which is either cyclization or hydrolysis of the linear precursor[5]. However, in iterative NRPS type B systems the TE domains have an additional function which allows the system to repeat the collinear synthesis once or twice. In this case the TE has to count the monomers stalled at the end of the assembly line and initiates release by cyclic dimer or trimer formation once the desired number is achieved. Less is known about mechanism and structure of iterative cyclases. Detailed mass spectrometric analysis was carried out for the last module of the enterobactin assembly line (EntF) containing a C-terminal peptide cyclase which catalyses cyclotrimerization of three 2,3-dihydroxybenzoyl serine (DHB-Ser) units to give the cyclic trilactone enterobactin[97]. In order to localize acyl-enzyme intermediates an active site histidine to alanine mutant enzyme was used with very slow substrate turnover. With this approach it was possible to provide evidence for a covalent acyl-O-TE intermediate and demonstrate that the peptide cyclase is involved in two reactions: acyl chain growth and cyclization. In the first step of acyl chain growth DHB-Ser is transferred to the active site of the peptide cyclase by a nucleophilic attack of active site serine onto the acyl-thioester of the upstream holo PCP-domain. The second step requires catalytic generation of a DHB-Ser alkoxide which in turn allows the nucleophilic attack onto another PCP bound DHB-Ser thioester to form a dimeric ester. The elongation step is repeated a third time before the final cyclic trilactone is released by an intramolecular nucleophilic attack of serine onto the acyl-O-TE ester bond. This mechanistic analysis suggests that the enterobactin cyclase serves as a waiting room while the phosphopantetheinylated PCP-domain is reacylated, which requires a stable ester bond and the exclusion of any water from the active site.
Insights into the mechanism of cyclization and hydrolysis as well as catalytic capacity of the surfactin cyclase are presented in this work.

2.8 Quaternary Architecture

Fatty acid synthases (FAS) and polyketide synthases (PKS) use an equivalent assembly line strategy with one multi-domain module for each monomeric acyl-CoA incorporated and elongated. The acyl chains in each module are tethered as thioesters to phosphopantetheinylated carrier protein domains. This is the same logic of a cascade of elongating acyl-S-pant-carrier protein intermediates utilized by the NRPS enzymatic machinery [44].
Many enzymes catalyzing sequential metabolic reactions aggregate by non-covalent linking of identical or non-identical subunits to form multi-enzyme complexes. The fatty acid synthases (FAS) of eucaryotes are examples of an enzyme complex composed of two

identical subunits (Fig. 2.16)[98]. Various investigations have unequivocally revealed the dimeric organization of multi-modular polyketide synthases [99, 100]. The oligomeric state of NRPS enzymatic machinery was unknown. The existence of NRPS-PKS hybrid enzymes in nature is well known [101-103] and it would be useful to know if NRPS modules that interface with PKS dimeric modules are themselves monomeric or oligomeric. Knowledge of the quaternary structure of NRPSs, defined as the spatial geometry of two identical neighbouring subunits, is not only important for understanding reactions occurring on or between enzymes, but may also be useful for engineering new hybrid enzymes by swapping modules or domains.

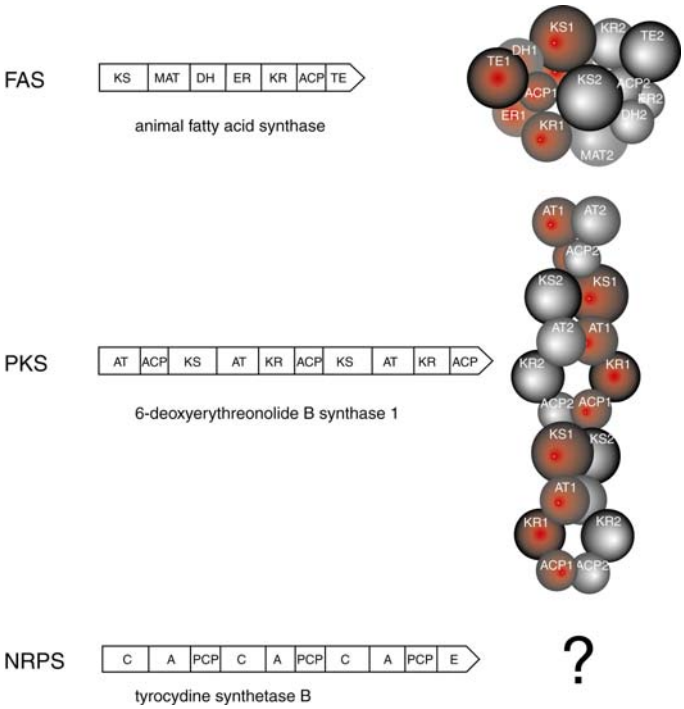


Figure 2.16. Quaternary structure of multienzyme complexes.

Although structures are now known for individual domains [96, 104-106], no high resolution structure is available for a multidomain NRPS or PKS (Fig. 2.17). Therefore, PKS and FAS enzymes were previously characterized by gel filtration, ultracentrifugation and chemical cross-linking experiments to ascertain oligomerization state of functional enzymes [100, 107, 108]. The topology of FAS and PKS enzymes has been additionally investigated by an *in vitro* complementation strategy. Site-directed mutagenesis was applied to produce genetically modified enzymes specifically inactivated in one domain. On inactivation of the ketosynthase (KS, carbon-carbon bond forming activity) or acyl carrier protein (ACP, homologous to NRPS PCP domains) all homodimeric enzymes lost their ability for product synthesis. An exchange of subunits from two differently mutated homodimeric enzymes generated a heterodimer carrying both (KS and ACP) mutations in opposite subunits. This heterodimer retains catalytic activity by cooperation of the remaining active KS and ACP domains across

both subunits. The functional interaction between KS and ACP from opposite subunits in a dimer has also been supported by cross-linking. This postulated mechanism has been shown for both PKS and FAS through product formation catalyzed by a heterodimer [98, 99, 109, 110]. Recently it was shown by Smith and coworkers that a FAS heterodimer, which consists of a wild type subunit and a subunit with mutations in all functional domains, still retains activity. This indicates that the quaternary organization of the dimer is required for structural integrity since monomers are not active, but in the context of the dimer structure, a monomer is self sufficient for catalytic activity.[111].

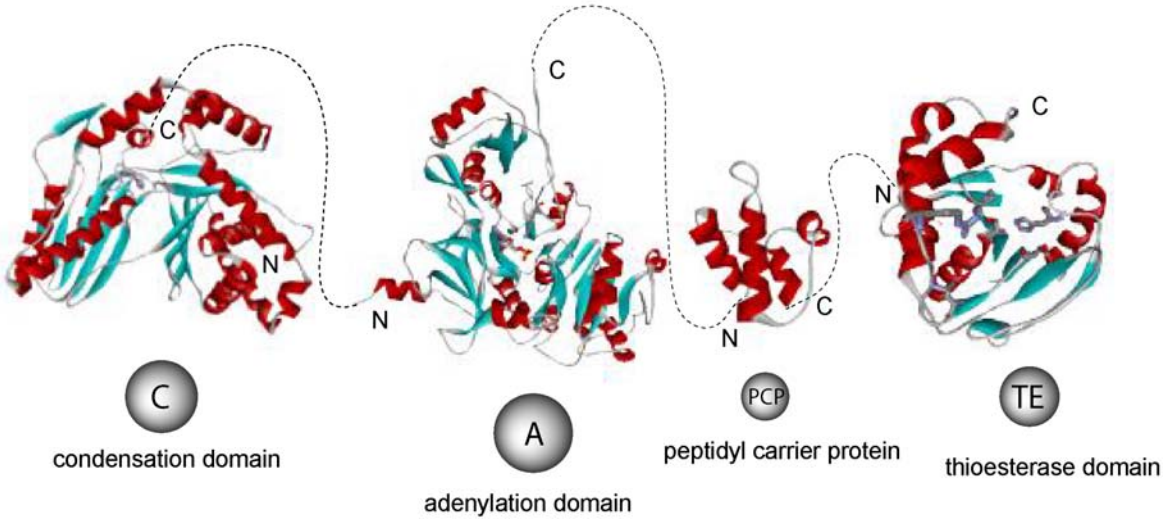


Figure 2.17. High resolution crystal structures of NRPS core domains assembled as a putative termination module. Shown are the structures of the vibriobactin synthetase C-domain VibH, the phenylalanine-activating A domain of gramicidin S synthetase PheA, the PCP domain of tyrocidine synthetase module C3 and the TE domain of the surfactin synthetase.

A double-helical structure model for modular PKS was proposed [100]. According to this model, the enzyme subunits are orientated head-to-head and folded in an interwound helical manner. KS, ACP and AT domains are building the core which is necessary for the observed interaction between KS and ACP domains from different strands (Fig. 2.16). Optional domains like the KR domain are accommodated in outside loops. The model for FAS enzymes considers also functional interactions within each of the two subunits which have been observed by cross-linking studies [112] [113].

Despite many analogies in the organization and reaction mechanism, in contrast to FAS and PKS little is known about the quaternary structure of NRPS. Mainly gel filtration experiments are mentioned in literature, e.g. the molecular weight of the enzymes EntB, EntE and EntF of the enterobactin synthetase were determined to be monomeric for EntE and EntF and trimeric

for EntB [84]. Other NRPS enzymes like the δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase as well as the pipecolate-incorporating enzyme for the biosynthesis of rapamycin also showed a monomeric behaviour on gel filtration [114, 115]. One objective of the research presented here was therefore to explore the quaternary structure of NRPS.

2.9 Task

One aim of this work was to elucidate the quaternary architecture of NRPSs with similar strategies to those which have been successfully applied to discover the structure of PKS and FAS. In addition to the work which has been started before, analytical ultracentrifugation should provide reliable information about protein organization under native conditions. Investigations should be carried out with several representative NRPS systems different in size and assembly line organization.

The main task of this work was to investigate the mechanism of peptide cyclization catalysed by thioesterase domains in order to evaluate their utility as catalytic cyclization tools. Based on a recent crystal structure of surfactin TE the following questions were addressed:

- Which residues are involved in catalysis?
- How does the enzyme distinguish between hydrolysis and cyclization?
- How tolerant is the enzyme for different substrates and what residues are recognized?
- Is it possible to engineer TE-domain architecture and alter its properties by *in vitro* protein evolution?

Moreover TE-domains from other NRPS systems should be explored and characterized to broaden a putative tool set of cyclization catalysts.

3 Materials

3.1 Chemicals, Enzymes and General Materials

Chemicals not listed were purchased as standard compounds from other manufacturers in p.a.-quality.

Table 3.1. Materials.

Manufacturer (Location)	Product
Amersham Pharmacia (Freiburg)	Source 15Q ion exchange column, Sephadex 200 16/60
Amersham/Buchler (Braunschweig)	various restriction endonucleases, ampicillin, IPTG, kanamycin, yeast extract, coomassie brilliant blue G and R250, HiTrap™ desalting columns
Bachem	N_α-Fmoc protected amino acids, N_α-Boc protected amino acids
Biomol Feinchemikalien (Darmstadt)	CHAPS
Böhringer Mannheim (Mannheim)	Expand™ Long Template PCR Kit
Difco (Detroit, USA)	Yeast extract
Eurogentech (Seraing, Belgien)	agarose, electro poration cuvettes
Fluka (Neu Ulm)	SDS, TEMED, fatty acid building blocks, DMF
Gibco BRL (Karlsruhe)	10 kDa protein molmarker, acrylamide
IBA (Göttingen)	Strep-Tactin sepharose column, Strep-tag HRP detection kit
Millipore (Bedford, USA)	dialysis membrane 0,025 µm
ICN Biomedicals GmbH (Eschwege)	DBP (dibromopropanone) cross-linker
MWG Biotech (Ebersberg), Qiagen (Hilden)	oligonucleotides
New England Biolabs (Schwalbach)	desoxyribonucleotides (dATP, dTTP, dGTP, dCTP), restriction endonucleases
Novabiochem (Darmstadt)	N_α-Fmoc protected amino acids, 2-chlorotritylchloride resin, HBTU, HOBT, PyBOP
Oxoid (Wesel)	agar Nr.1, trypton
Pierce (Rockford, USA)	BMH (bismaleimidohexane) cross-linker
Qiagen (Hilden)	QIAquick-spin PCR purification kit, Ni^{2+}-NTA-agarose, QIAexpress vector kit ATG, QIAEXII extraction kit
Qiagen-Operon (Köln)	oligonucleotides
Roth (Karlsruhe)	EtBr, β-mercaptoethanol, acrylamide for SDS-PAGE, piperidine
Schleicher & Schüll (Dassel)	Whatman-3MM paper
Serva (Heidelberg)	APS
Sigma (Deisenhofen)	EDTA, coenzyme A, N-acetylcysteamine, thiophenol and derivatives
Stratagene (Heidelberg)	QuikChange™ Site-Directed Mutagenesis Kit,

	PfuTurbo DNA polymerase, GeneMorph^{TM} PCR Mutagenesis Kit
Vivascience AG (Hannover)	Vivaspin 20 concentrators 10000 MWCO, 30000MWCO
Whatmann (Darmstadt)	96-deep well plates

3.2 Equipment

Table 3.2. Equipment

Device	Manufacturer
Autoclave	Tuttnauer 5075 ELV
Bidestilled water supply	Seral Seralpur Pro90CN
DNA-sequenze analyzer	Perkin-Elmer/ABI, ABI Prism 310 Genetic Analyzer
DNA-gel dokumentation	Cybertech CS1, thermoprinter Mitsubishi Video Copy processor
Elektroporation-pulse control	Bio-Rad Gene Pulser II
FPLC-System	Pharmacia FPLC-Biotechnology FPLC-System 250: Gradient-programmer GP-250 Pump P-500 Uvicord optical device UV-1 (l = 280 nm) Uvicord control element UV-1 2-channel printer REC-102 Injection valve V-7 3-way-valve PSV-100 Fraction collector FRAC-100
French Press	SLM Aminco; French-Presssure Cell-Version 5.1 20k Rapid-fill cell (40 mL)
HPLC-System	Agilent series 1100 HPLC-System with DAD-detection, vacuum degaser, quaternary pump, autosampler and HP-chemstation software columns: Macherey & Nagel Nucleosil 250/3, pore diameter 120 Å, particle size 3 µm; Nucleodur, 250/3, pore diameter 100 Å, particle size 3 µm
MALDI-TOF	Per Septive Biosystems Voyager-DE RP BioSpectrometry, Bruker FLEX III
Peptide synthesis	Millipore 9050 Plus
MS-MS analysis	Applied Biosystems, API Qstar Pulsar I
Shaker	New Brunswick Scientific Series 25 Incubator Shaker, New Brunswick Scientific Innova 4300 Incubator Shaker
Photometer	Pharmacia Biotech Ultraspec 3000
Speed-Vac	Savant Speed Vac Concentrator, Uniequip Univapo 150
Thermo-Cycler	Perkin-Elmer Thermal Cycler 480, Perkin Elmer Gene Amp PCR System 2400, Perkin Elmer Gene Amp PCR System 9700
Centrifugation	Heraeus Biofuge pico, Sorvall RC26 plus, rotors SS34 and SLA3000, Sorvall RC 5B Plus

3.3 Vector Systems

3.3.1 pQE60-Vector

The pQE60 vector system was used for cloning and overexpression of SrfTE active site mutants and PCP. The vector allows purification of recombinant proteins by Ni-NTA chromatography by fusing a His_6-tag to the C-terminal end of the overexpressed protein. This His_6 codon sequence is attached to the 3´ end of the multiple cloning site (MCS). The pQE60-vector carries two *lac*-operators in the promotor region. In the presence of a *lac*-repressor the gene can not be transcribed. After induction with IPTG, repression is abolished and gene transcription occurs. This system therefore allows a precise start of protein overexpression.

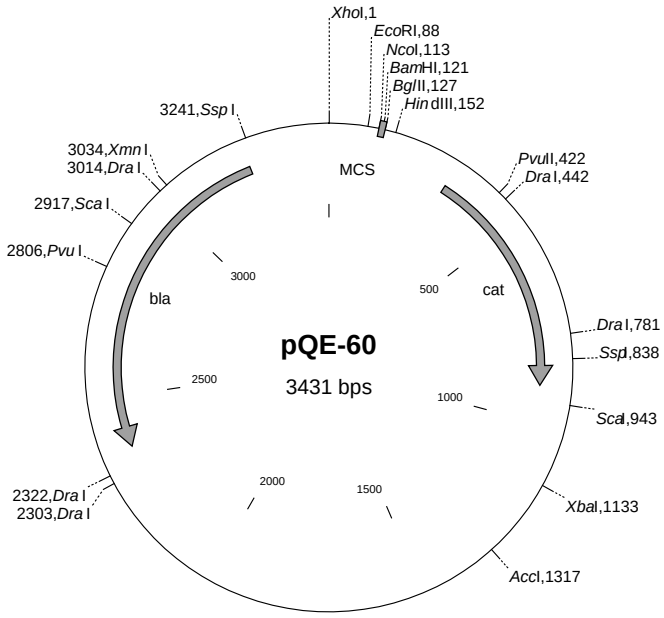


Fig. 3.1 Physical map of the pQE60-vector.

The vector contains the following components:

- replication origin from *E. coli* (ColE1)
- synthetic ribosomal binding site RBSII
- T5-promotor from *E. coli*-phages
- two *lac*-operator sequences for expression control by the *lac*-repressor
- MCS with recognition sequences for: *Nco*I (including the ATG-start codon), *Bam*HI and *Bgl*II
- Stop codons in all three reading frames
- codon sequence at the end of the MCS, which codes 6 histidines (C-terminal His_6-tag)
- two transcription terminators:
  t_0 of an λ-phage

T1 of the *rnnB*-operon from *E. coli*
- β-lactamase-gene *bla* for ampicillin resistance up to a final concentration of 100µg/mL

3.3.2 pASK-IBA 2 and pASK-IBA 4

These vector systems were used for protein evolution of TycTE. Both vectors allow purification of recombinant proteins by fusing a strep-tag sequence (WSHPQFEK) to the C-terminal (pASK-IBA 2) or N-terminal (pASK-IBA 4) end of the overexpressed protein. This strep tag sequence has a high affinity to genetically modified streptavidin (Strep-Tactin), which is used for affinity purification of the overexpressed proteins. Both vectors also contain N-terminal ompA signal peptide, which ensures that premature proteins are segregated into the periplasm of the bacterial cell. The signal peptide is cleaved by an *E. coli* signal peptidase while the protein passes trough the cell membrane.

pASK-IBA 2 and pASK-IBA 4 vectors are tightly regulated by the stringent *tet*-promotor. The *tet*-promotor is induced by low concentrations of anhydrotetracyclin which are below the level of antibiotic activities. The constitutive expression of the vector coded t*et*-repressor gene ensures high repression by the promotor in absence of inductor. There are no further cellular regulation mechanisms of the *tet*-promotor as seen in the lac-promotor.

The plasmid also contains the following components:

- origin of replication from pUC plasmids
- ampicilin-resistance
- multiple cloning site

3.3.3 pET37b

The pET37b (novagen) vector system was used for cloning and overexpression of fen PCP-TE, syr TE and myc PCP-TE as well as the fen PCP-TE active site mutant. The vector allows Ni-NTA chromatography purification of recombinant proteins by fusion of a His_8-tag to the C-terminal end of the overexpressed protein. In pET vectors, target genes are cloned under control of strong bacteriophage T7 transcription and translation signals, and expression is induced by providing a source of IPTG inducable T7 RNA polymerase in the host cell. *E. coli* BL21 cells were used for expression (chapter 3.4.2).

3.3.4 pBAD202/D-TOPO

The pBAD202/D-TOPO vector (Invitrogen) was used for the expression of the cyclosporin C-domain gene. The vector is regulated by the araBAD-promotor and is induced by low concentrations of arabinose. The His-patch thioredoxin leader increases the efficiency of translation and produces more soluble protein. The vector allows purification of recombinant

proteins by Ni-NTA chromatography by fusion of a His_6-tag to the C-terminal end of the overexpressed protein.

The plasmid also contains the following components:

- origin of replication from pUC plasmids
- rrnB transcription terminator
- codon sequence coding a V5 epitope
- CAP-(cAMP binding protein) binding site for transcription enhancement by binding of the CAP-cAMP-complex
- Kanamycin resistance

3.4 Microorganisms

3.4.1 *E. coli* XL1-Blue

This strain was used for cloning and sequencing purposes. The genotype is as follows: *rec*A1, *end*A1, *gyr*A96, *thi*-1, *hsd*R17, *sup*E44, *rel*A1, *lac*, F'(*pro*AB$^+$, *lac*I^q, *lac*ZDM15, Tn*10*(Tet^r))

3.4.2 *E. coli* BL21(DE3)

The *E. coli* strain BL21(DE3) with the genotype F^- *ompT* $r_B^- m_B^-$ is known as a bacterial host for the expression of plasmid DNA. It is characterized by a deficiency of Ion (protease) and by the lack of OmpT (protease of the outer membrane). It further contains an IPTG-induceble T7 polymerase gene, which is inserted in the chromosome after *lacZ* and the promotor *lac*uV5 on a λ-prophage. This is essential for the IPTG induction of genes under T7-promotor control.

3.4.3 *E. coli* M15

This strain lacks the T7-polymerase. The genotype is as follows: *nals*, *strs*, *rifs*, lac, *ara*, *gal*, *mtl*, *F*$^-$.

3.5 Media

E. coli strains were grown in LB-media. To prevent phage infections 20 mM Na^+-citrate were added.

16	g/L	bactotrypton
10	g/L	yeast-extract
5	g/L	NaCl
20	mM	Na-citrate

Culure plates: 1.2 % (w/v) of agar no.1 was added to the LB-media and heated at 121°C and 1,5 bar for 30 min. Antibiotics were added after cooling down to 55 °C in the following standard concentrations: 100 µg/mL ampicillin, 50 µg/ml kanamycin.

4 Methods

4.1 Methods of Molecular Biology

4.1.1 Construction of Plasmids

Amplification of all DNA gene fragments was carried out by polymerase chain reaction (PCR) with Pfu polymerase (Stratagene) according to the manufacturers protocol with synthetic oligonucleotides (MWG, Ebersberg or Qiagen-Operon, Köln). Purification of PCR-fragments was carried out by the "QIAquick-spin PCR purification kit" according to the manufacturers protocol (Qiagen, Hilden). All constructs were analyzed by restriction digests and DNA-sequencing.

Construction of pET37b[fenPCP-TE] – The fengycin gene fragment encoding Fen PCP-TE[14] was amplified from chromosomal DNA with the following oligonucleotides: 5´-AAA AAA AAC ATA TGC GTC AGG ACC TCA CAC CGC-3´ and 5´-AAA AAA CTC GAG ATG CTT ATT TGG CAG CAC TTT TTG-3´. The PCR product of fenPCP-TE was cloned into the NdeI/XhoI site of pET37b (Novagen). The plasmid was directly used to produce protein with a C-terminal octahistidine tag.

Construction of pQE60[srfPCP] - The surfactin gene fragment encoding Srf PCP (ATCC 21332) was amplified from chromosomal DNA with the following oligonucleotides: 5´-AAA CCA TGG AAT GGA TTG GAC CGC GGA AC-3´ and 5´-AAA GGA TCC GTT TTT CAA ATA CGC TGA AAT GC-3´. The PCR product of srf PCP was cloned into the NcoI/BamHI site of pQE60 (Qiagen). The plasmids were directly used to produce protein with a C-terminal hexahistidine tag.

Construction of pET37b[mycPCP-TE] - The mycosubtilin (ATCC 6633) gene fragment encoding Myc PCP-TE was amplified from chromosomal DNA with the following oligonucleotides: 5´-AAA AAC ATA TGG CTG CTC CCC GAA CGT TGA TT-3´ and 5`-AAA AAC TCG AGT GTT CGA TCA GAT TTT GTG CGA CTC-3´. The PCR product of myc PCP-TE was cloned into the *NdeI/XhoI* site of a pET37b vector (Novagene). *Escherichia coli* TOP 10 (Novagen) was used for preparation of recombinant plasmids. The plasmid was directly used to produce proteins with a C-terminal octahistidine tag.

Construction of pET37b[syrTE] - The syringomycin (DSMZ B-301D) gene fragment encoding Syr TE was amplified from chromosomal DNA with the following oligonucleotides: 5′-AAA AAC ATA TGG GCC AGG CAC GCC CC-3`and 5′-AAA AAC TCG AGC TCA GCT GGC GCG GTT ATC-3′. The PCR product of syr TE was cloned into the *NdeI/XhoI* site of a pET37b vector (Novagene). The plasmid was directly used to produce protein with a C-terminal octahistidine tag.

Construction of pASK-IBA2[tycTE] – The tycTE gene fragment was amplified from a pQE60 vector containing tycTE[90] with the following oligonucleotides: 5′-AAA AAG AAT TCG CAT AAG CGC TTT GAG AGC AG-3′ and 5′-AAA AAC TGC AGT TTC AGG ATG AAC AGT TCT TG-3′. The PCR product of tyc TE was cloned into the *EcoRI/PstI* site of a pASK-IBA2 vector (IBA). The plasmid was directly used to produce protein with a C-terminal strep-tag.

Construction of pASK-IBA4[tycTE]* – The tycTE gene fragment was amplified from the pASK-IBA2 [tycTE*] (chapter 4.1.2) with the following oligonucleotides: 5′-AAA AAG AAT TCC ATA AGC GCT TTG AGA GCA G-3′ and 5′-AAA AAC TGC AGT TTC AGG ATG AAC AGT TCT TG-3′. The PCR product of tyc TE was cloned into the *EcoRI/PstI* site of a pASK-IBA4 vector (IBA). The plasmid was directly used to produce protein with a N-terminal strep-tag.

Construction of pBAD202/D-TOPO[cssC]- The cssC gene fragment was amplified from chromosomal DNA (ATCC 34921) with the following oligonucleotides: 5′-CAC CCA ACA GCA GGG GTT CTC GGG-3′ and 5`-TGT TCT GCT CAA AGC GAG CAT GG-3′. The PCR product of cssC containing the sequence 5′-CACC at the 5′-end was incubated with a pBAD202/D-TOPO vector (Invitrogen) according to the manufacturers protocol. The plasmid was directly used to produce protein with a C-terminal hexahistidine tag.

4.1.2 Site Directed Mutagenesis

Construction of Srf TE active site mutants - The mutants were constructed by the "QuickChange site directed mutagenesis kit" (Stratagene) according to the manufacturers protocol. Constructs were obtained by PCR amplification of the Srf TE-containing pQE60 plasmid[80] with the following oligonucleotides (modified sequences underlined): **P26G**, 5′-ATT TTC GCA TTT CCG GGG GTC TTG GGC TAT GGC CT-3′ and 5′-AGG CCA TAG CCC AAG ACC CCC GGA AAT GCG AAA AT-3′; **S80C**, 5′-CAT TGT TTG GAT ATT

GCG CGG GAT GCA GCC TGG CG-3′ and 5′-CGC CAG GCT GCA TCC CGC GCA ATA TCC AAA CAA TG-3′; **D107A**, 5′-GCG GAT CAT CAT GGT CGC TTC CTA TAA AAA ACA AGG TGT C-3′ and 5′-GAC ACC TTG TTT TTT ATA GGA AGC GAC CAT GAT GAT CCG C-3′; **K111A**, 5′-GGT CGA TTC CTA TAA AGC ACA AGG TGT CAG TGA TCT GG-3′ and 5′-CCA GAT GAC TGA CAC CTT GTG CTT GTG CTT TAT AGG AAT CGA CC-3′; **R120A**, 5′-TGT CAG TGA TCT GGA CGG AGC CAC GGT TGA AAG TGA-3′ and 5′-TCA CTT TCA ACC GTG GCT CCG TCC AGA TCA CTG ACA-3′; **H207A**, 5′-AAG AGG CTT CGG AAC AGC CGC AGA AAT GCT GCA GG-3′ and 5′-CCT GCA TTT CTG CGG CTG TTC CGA AGC CTC TT-3′. *E. coli* XL1 Blue (Stratagene) was used for preparation of recombinant plasmids. Mutagenesis products were confirmed by sequencing with an ABI Prism 310 (Applied Biosystems).

Construction of pET37b[fenPCP-TE]* - The gene fragment encoding the Fen PCP-TE site directed mutant was constructed using the Quick-Change site directed mutagenesis kit (Stratagene) according to the manufacturers protocol. The construct was obtained by PCR amplification of Fen PCP-TE containing pET37b plasmid with the following oligonucleotides (modified sequences underlined): 5′-CTA TTA GGC TAC GCC GCG GGC GGA ACT T-3′ and 5′-TCC GCC CGC GGC GTA GCC TAA TAG AAC GT-3′. *E. coli* TOP10 (Novagen, USA) was used for preparation of recombinant plasmids. All constructs were confirmed by sequencing.

Construction of pASK-IBA2[tycTE]* – Two silent point mutations were introduced into the TycTE gene fragment by standard PCR amplification of the pASK-IBA2[tycTE] vector with the following oligonucleotides: 5′-CGA AAG AAC TCG AGG AGC GGG-3′ and 5′-CGT CTG CTT CGA GCT CGG ATT-5′. These silent point mutations give rise to two unique restriction sites within the lid region of Tyc TE (*XhoI/SacI*). The final PCR product was cloned into the *EcoRI/PstI* site of a pASK-IBA2 vector (IBA). The plasmid was directly used to produce protein with a C-terminal strep-tag.

4.1.3 Random Mutagenesis

Random mutagenesis was carried out with the "GeneMorph™ PCR Mutagenesis Kit" (Stratagene) according to the manufacturers protocol. In this kit an error rate is achieved by a genetically modified "Mutazyme DNA polymerase". A low, medium or high mutation frequency is produced by adjusting the initial target DNA amounts in the amplification reactions. For the same PCR yield, targets amplified from low amounts of target DNA

undergo more duplications than targets amplified from high concentrations of DNA. The more times a target is replicated, the more errors occur. Mutazyme DNA polymerase makes few insertion and deletion mutations (1.1 %) which create frame shifts.

In this study a mutation rate of 7-16/kb was selected by amplifying 10 ng of DNA template. This was repeated 3 times with 10 ng template at each amplification step. The PCR product yield must be between 500 ng and 10 µg in a 50 µl reaction in order to obtain the predicted mutation frequencies. To ensure sufficient product yield, sample PCR reactions are electrophoresed adjacent to a DNA standard. The mutation rate obtained by DNA sequencing was about 3-4 per 273 bp corresponding to 1-2 amino acid mutations. The mutated gene fragments of the tyc TE lid region were digested with *XhoI/SacI* and cloned into the pASK-IBA2[tycTE*] vector containing unique restriction sites in the lid region.

Table 4.1. PCR protocol for random mutagenesis

Cycles	Temperature [°C]	Time [min]
1	95	1:00
30	95	1:00
	50	1:00
	72	1:00
1	72	10:00

4.1.4 DNA Sequencing

Sequence analysis of double stranded DNA was carried out by the method of chain termination (147 RF) with the "ABI Prism dRhodamine terminator cycle sequencing ready reaction kit (ABI, Foster City (USA))" according to the manufacturers protocol. Sequencing reactions were performed with 100 ng of DNA per kbp, 5 pmol primer and 3-4 µl sequence mix. After 25-30 rounds of PCR and subsequent purification the sequence analysis was carried out on an "ABI Prism 310 Genetic Analyzer (Applied Biosystems, Weiterstadt)".

4.2 Protein Techniques

Standard methods used in protein chemistry like SDS-polyacrylamide-gelelectrophoresis (SDS-PAGE) and coomassie-staining have been described elsewhere[46, 116].

4.2.1 Gene Expression

Heterologous expression of TE and PCP-TE proteins was predominantly carried out in the pQE60 and pET37b vector systems. For protein evolution of Tyc TE also pASK-IBA and for cloning of the cyclosporin A C-domain pBAD202/D-TOPO were used. For this purpose *E. coli* BL21 (DE3) and M15 were transformed with the expression plasmids by electroporation, with the exception of tyc TE for protein evolution where *E. coli* XL1 Blue was transformed with expression plasmids.

4.2.1.1 Expression with the pQE- and pET-Vector Systems

5 ml overnight culture of the corresponding expression strain in LB-media was inoculated in 500 ml of LB-media. The culture was incubated at 30-37 °C and 220 rpm until an A_{600} of 0.5-0.7 was observed. The expression was induced by adding 1 mM IPTG. The culture was further incubated at 30 °C, 220 rpm for 2-3 h. The cells were subsequently harvested by centrifugation (5000 upm, 4 °C, 10 min) and the resulting pellet resuspended in Hepes-A buffer (50 mM Hepes, 300 mM NaCl (pH 8.0)). The cell suspension was either stored at –20 °C or immediately used for protein preparation.

4.2.1.2 Expression with the IBA-Vector Systems

3 ml overnight culture of the corresponding expression strain in LB-media was inoculated in 300 ml of LB-media (Amp^{100}). The culture was incubated at 30-37 °C and 220 rpm until an A_{550} of 0.5-0.7 was observed. The expression was induced by adding 3 µL of an anhydroteracycline solution in DMF (2 mg/ml) and the culture was further incubated at 30 °C, 220 rpm for 2-3 h. The cells were subsequently harvested by centrifugation (5000 upm, 4 °C, 10 min) and the resulting pellet was resuspended in 3 ml buffer P (100 mM Tris/HCl, 1 mM ethylendiaminteraactetate (EDTA), 250 mM sucrose (pH 8.0)). The cell suspension was either stored at –20 °C or immediately used for protein preparation.

4.2.1.3 Expression with the pBAD/TOPO-Vector System

5 ml overnight culture of the corresponding expression strain in LB-media was inoculated in 500 ml of LB-media (Kan^{50}). The culture was incubated at 22 °C and 220 rpm until an A_{600} of 0.5-0.7 was observed. The expression was induced by adding 0.01 % - 0.0005 % arabinose. The culture was further incubated at 22 °C, 220 rpm for 3 h. The cells were subsequently harvested by centrifugation (5000 upm, 4 °C, 10 min) and the resulting pellet resuspended in Hepes-A buffer (50 mM Hepes, 300 mM NaCl (pH 8.0)). The cell suspension was either stored at –20 °C or immediately used for protein preparation.

4.2.2 Protein Purification

Protein purification was carried out using Ni-NTA affinity chromatography, Strep-Tactin affinity chromatography or periplasmatic extraction for high throughput protein evolution.

4.2.2.1 Preparation of Cell Extracts

Preparation of cell extracts from cytoplasmatic expressions was carried out by using a pre-cooled 40 K French Press cell (SLM Aminco, Urbana, Illinois (USA)). Three cycles of compression and decompression were performed with each cell extract. Disrupted cell material was separated from the cell extract by centrifugation (17500 upm, 4 °C, 30 min). The supernatant was subsequently used for affinity chromatography.

Preparation of cell extracts from periplasmatic expression was carried out by 30 – 60 min incubation on ice with buffer P. Buffer P contains EDTA which destabilises the cell wall by complexing Ca^{2+} and Mg^{2+} ions. Subsequent centrifugation (17500 upm, 4 °C, 10 min) disrupts the outer membrane and releases proteins in solution. For high throughput assays no further purification was required.

4.2.2.2 Purification of His_6-tagged Proteins

His_6-tagged proteins can be easily purified by Ni^{2+}-NTA affinity chromatography. In standard purifications cell extracts from 500 ml expressions were loaded onto a Ni-NTA superflow (Qiagen, Hilden) column (HR 10/2, Amersham Biosciences, Freiburg). The column was loaded with a flow rate of 0.8 ml/min in 3 % buffer Hepes B (50 mM Hepes, 300 mM NaCl, 250 mM imidazole (pH 8.0)) on a "fast performance liquid chromatography" (FPLC) system (Amersham Biosciences, Freiburg). Protein binding was monitored at A_{220}. When the A_{220} was at baseline again, a 30 min linear gradient up to 50 % Hepes B followed by a 10 min gradient to 100 % Hepes B with a flow rate of 1 ml/min was applied. Fractions of 2 ml were collected. Proteins were identified by SDS-PAGE. All proteins were dialyzed with standard assay buffer (25 mM Hepes, 50 mM NaCl (pH 7.0)) by using "HiTrap Desalting" columns (Amersham Biosciences, Freiburg).

4.2.2.3 Purification of Strep-tagged Proteins

Strep-tagged proteins can be easily purified by Strep-Tactin affinity columns. 3 ml of periplasmatic extract were loaded onto a Step-Tactin gravity flow column (IBA, Göttingen). The column was washed 5 times with 1 ml buffer W and eluted with 6 times 0.5 ml buffer E (100 mM Tris/HCl, 1 mM EDTA, 2.5 mM Desthiobiotin (pH 8.0)). The column was regenerated with 5 ml buffer R (100 mM Tris/HCl, 1 mM EDTA, 1 mM 4´-hydroxo-

benzolazo-2-benzoic acid (HABA) (pH 8.0)) followed by equilibration of the column with 4 ml buffer W. Purity of proteins was evaluated by SDS-PAGE and desalting was carried out as described in 4.2.2.2.

4.2.3 Concentration of Protein Solutions

Concentrations of the purified proteins were determined photometrically at 280 nm using calculated extinction coefficients (table 4.2). Theoretical extinction coefficients were calculated by ProtParam (Expasy, Geneva (Switzerland)).

Table 4.2. Theoretical extinction coefficients

Protein	Theoretical extinction coefficient [mg/ml]
PCP	1.28
SrfTE and all mutants	0.87
SyrTE	0.83
FenPCP-TE	1.09
MycPCP-TE	0.80
TycTE	1.34
JH1	1.15
10B6	1.10
9E2	0.86

4.3 Biochemical Methods

4.3.1 Cyclization Assays with SNAC and Thiophenol Substrates

Reactions were carried out in standard assay buffer in a total volume of 50 µL. Substrate concentration was 100 - 250 µM for standard reactions and varied for kinetic investigations. For the cycle opening reactions, 250 mM of SNAC, ONAC, or N-acetylethylenediamine (NNAC) were added. Reactions were initiated by addition of enzyme to give final concentrations of 5 µM for Fen PCP-TE, 1 and 2.5 µM for Srf TE, 30 µM for Myc PCP-TE and 15 µM for Syr TE. Reactions were quenched at various time points by addition of 35 µL of 4 % TFA/H_2O, and the products were analysed by HPLC-MS on a reversed phase C_{18} Nucleodur column (Macherey and Nagel, 250/3, pore diameter 100 Å, particle size:3 µm)

with the following gradients: Fen PCP-TE 0-35 min, 30-60 % Acetonitrile/0.1 % TFA in water/0.1 % TFA; 0.4 mL/min, 40°C, Srf TE 0-40 min, 20-55 % Acetonitrile/0.1 % TFA in water/0.1 % TFA; 0.4 mL/min, 40°C, Myc PCP-TE 0-40 min, 5-50 % Acetonitrile/0.1 % TFA in water/0.1 % TFA; 0.4 mL/min, 40°C and Syr TE 0-40 min, 10-50 % Acetonitrile/0.1 % TFA in water/0.1 % TFA; 0.4 mL/min, 40°C. Identities of the products were confirmed by ESI-MS, and cyclic by-products were additionally verified by MS-MS analysis to confirm connection regiospecificity (API Qstar Pulsar i, Applied Biosystems).

Concentrations of peptidyl-thioesters were calculated using experimentally determined extinction coefficients at 220 nm and assumed to be the same for cyclized and hydrolysed products. Kinetic characterization of the cyclization and hydrolysis reactions were carried out by determining initial rates at 5 – 7 substrate concentrations using two time points at each concentration within the linear region of the enzymes determined by time courses.

The preparative scale synthesis of cyclic **SLP-wt** (see appendix) was carried out under standard assay conditions in a total volume of 10 mL (200-fold increase). After 2.5 hr, the reaction was quenched with 500 µL TFA and centrifuged at 2000 rpm for 5 min. The supernatant was lyophylized overnight and redissolved in 2 mL 1:1 H_2O/Acetonitrile. The cycle was isolated by preparative HPLC with a reversed-phase (C_{18}) column (Vydac). Cyclic **SLP-wt** elutes at 24 min (10 mL/min; 0-35 min, 20-100% B; A = H_2O, 0.1% TFA; B = acetonitrile).

4.3.2 Loading Peptidyl-CoA onto PCP

Loading reactions of apo PCPs were carried out in standard assay buffer in a total volume of 100 µL at room temperature. For the kinetic characterization of the PCP loading reaction the concentration of PCP was kept constant at 50 µM (K_M = 1.3-1.8 µM)[55]. Initial rates were determined at various peptidyl-CoA concentrations ranging from 5 µM to 100 µM using two time points (7 min and 14 min) at each concentration. A time course showed that modifications were linear up to 30 min. Reactions were initiated by addition of Sfp to give a final concentration of 2 µM and quenched by the addition of 30 µL of 4 % TFA/H_2O. The products were analyzed by analytical HPLC-MS (Agilent 1100) with a reversed phase C_{18} Nucleosil (Macherey and Nagel, 250/3, pore diameter:120 Å, particel size:3 µm) column. The identities of the products were confirmed by HPLC-MS and MALDI-TOF MS. Different gradients were applied according to resolution of the substrates (0-26 min, 55 to 70 % Acetonitrile/0.1 % TFA in water/0.1 % TFA, or 0-26 min, 40-57 % Acetonitrile/0.1 % TFA in water/0.1 % TFA; 0.9 mL/min, 45°C)[117]. The concentration of apo-PCP was calculated

using experimentally determined extinction coefficients at 220 nm which were assumed to be the same for the peptidyl holo-PCPs.

4.3.3 PCP-TE Cyclization Assay

PCP-TE cyclization assays were carried out in 100 µl standard assay buffer at room temperature. The peptidyl-CoA concentration ranged from 20 µM to 60 µM and the PCP-TE concentration was 60 µM. Reactions were initiated by addition of Sfp to give a final concentration of 5 µM to ensure a fast modification reaction. Reactions were quenched at various time points by addition of 40 µL 4 % TFA/H_2O and were analyzed by HPLC-MS on a reversed phase C_{18} Nucleodur column (Macherey and Nagel, 250/3, pore diameter:100 Å, particle size:3 µm) with the following gradient: 0-35 min, 30-60 % acetonitrile/0.1 % TFA in water/0.1 % TFA; 0.4 mL/min, 40°C. Identities of the products were confirmed by ESI-MS and MALDI-TOF MS, and the cyclic products were additionally verified by MS-MS analysis (API Qstar Pulsar i, Applied Biosystems). Cyclic product hydrolysis assays were carried out in 100 µL standard assay buffer under the same conditions as mentioned above with 30 µM PCP-TE and 50 µM fengycin cycle.

4.3.4 Protein Evolution

After expression of TycTe mutant variants in 96 deep well plates the cells were pelleted by centrifugation (4000 upm, 4 °C, 15 min) and resuspended in 50 µl buffer P. The cell fragments were pelleted again by centrifugation (4000 upm, 4 °C, 15 min) after 1 h incubation. 8 µl of 1 mM **TLP-Ala** (see appendix) was added to the supernatant and incubated for 1 h at room temperature. Reactions were stopped by adding 35 µl 4 % TFA/water. The products were analyzed by analytical HPLC-MS (Agilent 1100, API Qstar Pulsar I, Applied Biosystems) with a reversed phase C_8 Nucleosil (Macherey and Nagel, 30/2, pore diameter:120 Å, particle size:3 µm) column with the following method: 5 min, 40 % acetonitrile/0.1 % TFA in water/0.1 % TFA; 0.3 mL/min, 40°C.

4.4 Analytical Methods

4.4.1 Analytical Ultracentrifugation

Sedimentation equilibrium experiments were performed with a Beckman Optima XL-A ultracentrifuge. Equilibrium and Monte Carlo analyses were performed with UltraScan version 5.0 [118]. Hydrodynamic corrections for buffer conditions were made according to data published by Laue *et al.* [119] and as implemented in UltraScan. The partial specific volumes of $TycB_{2-3}$ and EntF were estimated according to the method of Cohn and Edsall

[120] and as implemented in UltraScan. $TycB_{2-3}$ samples were analyzed in a buffer containing 25mM Hepes pH 7 and 50mM NaCl. EntF samples were analyzed in a buffer containing 25mM TRIS pH 8, 10mM $MgCl_2$, 5mM DTT and 10% glycerol. Experiments were performed at 4°C at speeds ranging between 5,000 and 12,500 rpm for $TycB_{2-3}$ and between 8,000 and 17,000 for EntF. Samples were spun in a 6-channel 12mm external fill equilibrium centerpiece in an An-60 Ti rotor. 280nm scans were collected at equilibrium in radial step mode with 0.001cm steps and 50 point averaging. Loading concentrations were 0.39 and 0.57 O.D. for $TycB_{2-3}$ and 0.40 O.D. for EntF. Data exceeding 0.9 O.D. was excluded from the fits. Data fitting was performed with a monomer-dimer equilibrium model for $TycB_{2-3}$ and with a one-component (monomer only) model for EntF, as implemented in UltraScan.

4.4.2 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. Samples were prepared by mixing 1 µl of peptide sample with 1 µl DHB-matrix solution (Agilent Technologies). 1 µl of this solution was pipetted onto a metallic target probe and dried under air. The cocrystallized samples were investigated with a "Bruker FLEX III" (Bruker Daltronics, Leipzig).

4.4.3 HPLC-MS

High Performance Liquid Chromatography (HPLC) was used to characterize substrate and product by retention time on a chromatography column and by mass. Reversed phase chromatography is based on hydrophobic interactions with the unpolar stationary phase (C_{18} or C_8 coated silica gel). Elution is mediated by the unpolar organic solvent acetonitrile which concurs with the adsorbed analytical compound for binding positions. The retention time of the analytical compound is monitored by UV-detection. An electrospray-ionization mass detector allows the mass analysis of liquid compounds at atmospheric pressure. Ionization of the analytical compound was achieved by adding 0.1 % TFA. Experiments were carried out on a "Agilent 1100" system (Agilent, Waldbronn). Experimental details are listed in section 4.3.

All substrates and products were confirmed by HPLC-MS or MALDI-MS:

Table 4.3. Characterization of surfactin SNAC substrates and products by mass. N/A = not applicable; N/D = no product detected.

compound	species	observed mass (calculated mass) (Da)		
		Substrate	cyclized product	hydrolyzed product
SLP-wt	$[M-H]^-$	999.7 (999.5)	880.7 (880.5)	898.6 (898.5)
SLP-1	$[M+H]^+$	959.4 (958.6)	N/A	857.4 (857.5)
SLP-2	$[M+H]^+$	973.9 (974.5)	855.5 (855.5)	873.5 (873.5)
SLP-3	$[M+H]^+$	974.8 (974.5)	854.9 (855.5)	873.0 (873.5)
SLP-4	$[M+H]^+$	989.1 (988.5)	869.6 (869.5)	887.6 (887.5)
SLP-5	$[M+H]^+$	972.3 (972.6)	853.0 (853.5)	871.0 (871.5)
SLP-6	$[M+H]^+$	974.4 (974.5)	N/D	N/D
SLP-7	$[M+H]^+$	974.1 (974.5)	N/D	N/D
SLP-1/2	$[M+H]^+$	871.9 (872.5)	752.8 (753.5)	771.0 (771.5)
SLP-3/4	$[M+H]^+$	901.8 (902.5)	783.3 (783.4)	800.8 (801.5)
SLP-5/6	$[M+H]^+$	885.9 (886.5)	767.0 (767.5)	785.0 (785.5)
SLP-wt cycle	$[M-H]^-$	880.7 (880.5)	N/A	898.1 (898.5)
SLP-ONAC	$[M-H]^-$	983.8 (983.5)	880.6 (880.5)	898.6 (898.5)
SLP-εOH	$[M+H]^+$	1002.6 (1002.5)	883.5 (883.5)	901.5 (901.5)

Table 4.4. Characterization of peptidyl-coenzyme A substrates by mass.

compound	species	ionization method	observed mass (calculated mass) (Da)
SLP	$[M+H]^+$	ESI	1620.4 (1620.6)
MLP	$[M+H]^+$	ESI	1613.5 (1613.5)
FLP	$[M+H]^+$	ESI	2018.5 (2018.7)
FLP(Phe)	$[M+H]^+$	ESI	2003.6 (2003.7)
FLP(2)	$[M-H]^-$	MALDI	2016.6 (2016.7)
FLP(4)	$[M-H]^-$	MALDI	2016.6 (2016.7)

Table 4.5. Characterization of peptidyl-PCP and cyclic fengycin products by mass. N/D = no product detected.

compound	species	ionization method	observed mass (calculated mass) (Da)	
apo-PCP	$[M+H]^+$	MALDI	9876.0 (9877.3)	
SLP-PCP	$[M+H]^+$	MALDI	11069.0 (11068.9)	
MLP-PCP	$[M+H]^+$	MALDI	11063.0 (11061.7)	
FLP-PCP	$[M+H]^+$	MALDI	11467.0 (11466.9)	
			cyclized product	hydrolyzed product
FLP	$[M+H]^+$	ESI	1251.2 (1251.6)	1269.4 (1269.6)
FLP (Phe)	$[M+H]^+$	ESI	N/D	1253.4 (1253.6)
FLP (2)	$[M+H]^+$	ESI	1251.2 (1251.6)	1269.4 (1269.6)
FLP (4)	$[M+H]^+$	ESI	N/D	1269.4 (1269.6)

Table 4.6. Characterization of thiophenol substrates by mass.

compound	species	ionization method	observed mass (calculated mass) (Da)
FLP-tp	$[M+H]^+$	ESI	1361.4 (1361.6)
SLP-tp	$[M+H]^+$	ESI	965.4 (965.5)
MLP-tp	$[M+H]^+$	ESI	1055.5 (1055.5)
SyLP-tp	$[M+H]^+$	ESI	1175.6 (1175.5)
SLP-tc	$[M+H]^+$	ESI	979.4 (979.5)
SLP-mb	$[M+H]^+$	ESI	995.4 (995.5)

Table 4.7. Characterization of thiophenol products by mass.

compound	species	ionization method	observed mass (calculated mass) (Da)		Cy/Hy ratio
			cyclized product	hydrolyzed product	
FLP-tp	$[M+H]^+$	ESI	1251.4 (1251.6)	1269.4 (1269.6)	0.85
SLP-tp	$[M+H]^+$	ESI	855.4 (855.5)	873.4 (873.5)	0.78
MLP-tp	$[M+H]^+$	ESI	945.4 (945.4)	963.4 (963.5)	1.13
SyLP-tp	$[M+H]^+$	ESI	1065.3 (1065.5)	1083.6 (1083.5)	0.65
SLP-tc	$[M+H]^+$	ESI	855.4 (855.5)	873.4 (873.5)	0.73
SLP-mb	$[M+H]^+$	ESI	855.4 (855.5)	873.4 (873.5)	0.64

Table 4.8 Characterization of tyrocidine substrates and products by mass.

compound	species	observed mass (calculated mass) (Da)		
		Substrate	cyclized product	hydrolyzed product
TLP	$[M+H]^+$	1389.6 (1389.7)	1270.5 (1270.6)	1288.6 (1288.7)
TLP-Ala	$[M+H]^+$	1313.6 (1313.6)	1194.5 (1194.6)	1212.5 (1212.6)
allyl-TLP	$[M+H]^+$	1339.6 (1339.7)	1220.6 (1220.6)	1238.7 (1238.6)
Ala-Ala-TLP	$[M+H]^+$	1304.4 (1304.6)	1185.4 (1185.6)	1203.6 (1203.6)
Ala-Leu-TLP	$[M+H]^+$	1346.4 (1346.6)	1227.4 (1227.6)	1245.6 (1245.6)

4.5 Solid Phase Peptide Synthesis (SPPS)

Synthesis of peptide chains on an insoluble support has benefits like easy separation of intermediate peptides by filtration and washing with savings in time and labour. However, there are also limitations in this approach: By-products from incomplete reactions, side reactions, or impure reagents will accumulate on the resin during chain assembly and contaminate the final product. In this work Fmoc-based SPPS was used in a batch process[121]. The resin used was cross-linked polystyrene with an attached 2-chlorotritylchloride linker (Novabiochem, Darmstadt). Amino acid building blocks were side chain protected to prevent side reactions (Bachem, Novabiochem). Protecting groups included: tert-butyl (tBu), trityl (trt) and Boc.

4.5.1 SPPS: Initiation, Elongation and Termination

4.5.1.1 Initiation: Coupling of the First Amino Acid to the Resin

The first step in the process of solid phase peptide synthesis is the loading of the resin linker 2-chlorotritylchloride with the C-terminal amino acid. This step is very important because the extend of this reaction will determine the yield of the final product and also sites on the resin not reacted in this initial process can potentially be acylated in subsequent cycles leading to truncated products. For this reason the resin was incubated in dry DCM with 2 equivalents of Fmoc-protected amino acid and 8 equivalents of the weak base DIPEA (N,N-diisopropylethylamine) which ensures complete deprotonation of the carboxy-group. The carboxy-group carries out a nucleophilic attack onto the labile carbon-chlorine bond of the 2-chlorotrityl linker, attaching the first amino acid to the resin. After 2 h incubation the solvent was removed by filtration and the resin washed with DCM (Fig. 4.1).

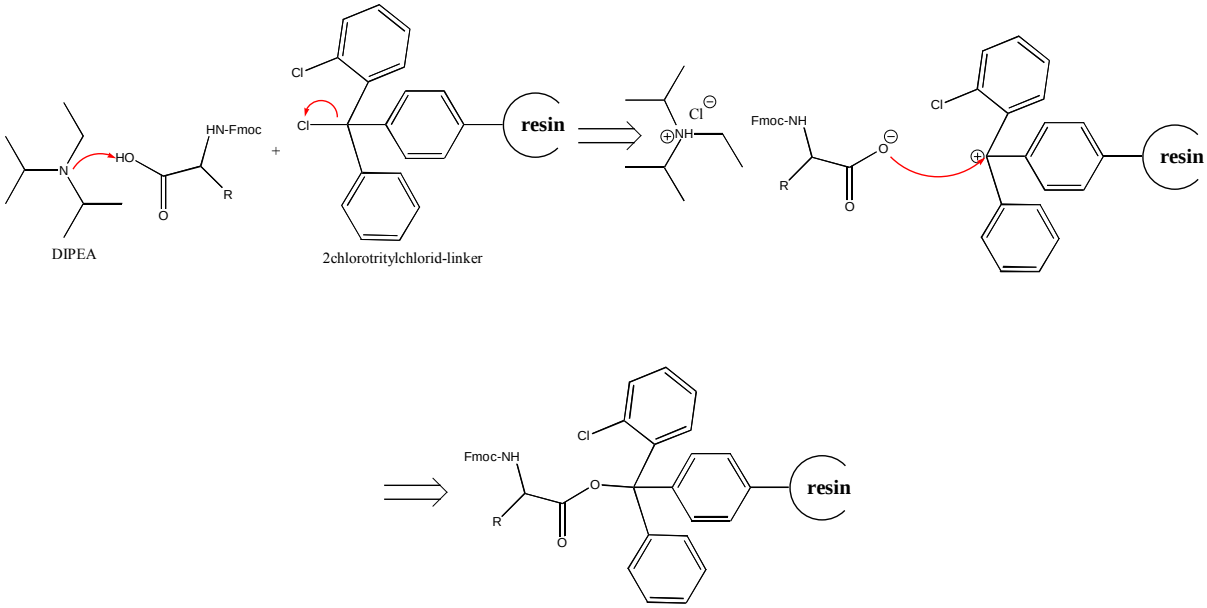


Figure 4.1. Coupling of first amino acid onto solid support.

4.5.1.2 Elongation

Elongation of the peptide chain requires deprotection of the N-terminal Fmoc-protecting group of the resin bound amino acid. The removal is usually achieved by treatment with 20 % piperidine in DMF for 5 – 10 min. The key step is initial deprotonation of the flourene ring to generate the aromatic cyclopentadiene-type intermediate. This rapidly eliminates to form dibenzofulvene, which is scavenged by piperidine (Fig. 4.2).

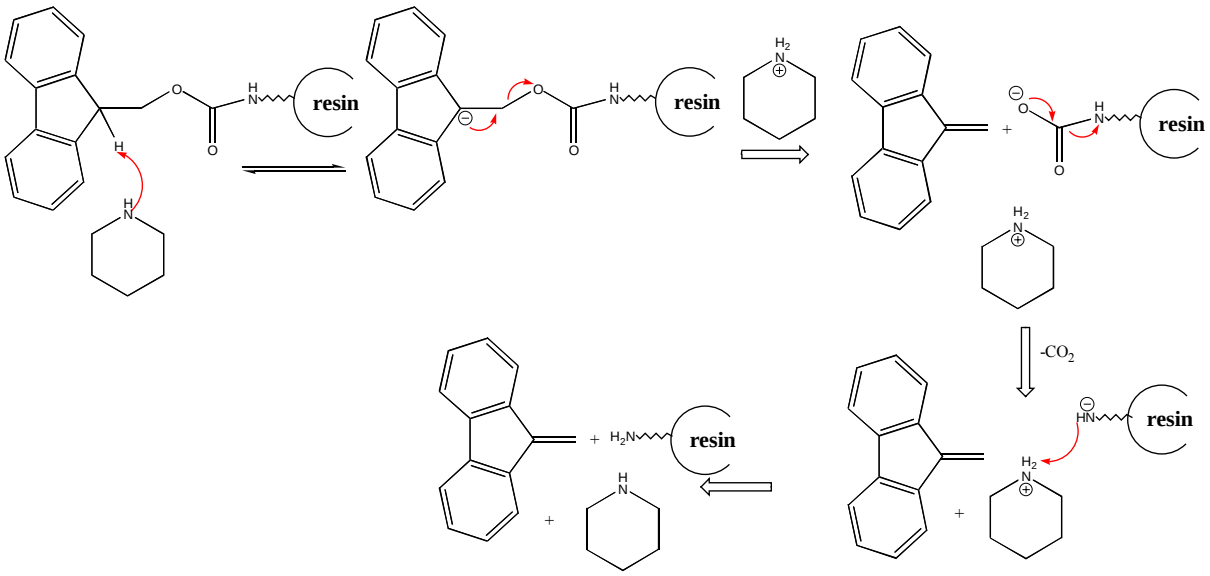


Figure 4.2. Deprotection of Fmoc protected amino acid.

Chemical methods to form peptide bonds require activation of the carboxy-group of the Fmoc-protected amino acid. In this work only *in situ* coupling methods were applied by using the coupling reagent HBTU with its additive HOBt in DMF. The carboxy-group is deprotonated by a 10 fold excess of DIPEA and attacks in the second step the electrophilic carbenium ion of HBTU. A very reactive N,N,N,N-tetramethylurea intermediate is formed which is attacked by the nucleophile HOBt leading to a reactive benzotriazole ester. The free amine of the resin bound amino acid attacks this intermediate releasing benzotriazol and giving rise to an elongated peptide chain (Fig. 4.3). Quantitative reactions are observed when a high excess of amino acids (3 times) is used.

An analytical method to monitor the qualitative success of each coupling step can be carried out with the Kaiser-test[121]. The reaction of ninhydrin with primary amines leads to intense blue color. No color is observed with Fmoc-protected amines, which allows to distinguish between success of deprotection and coupling.

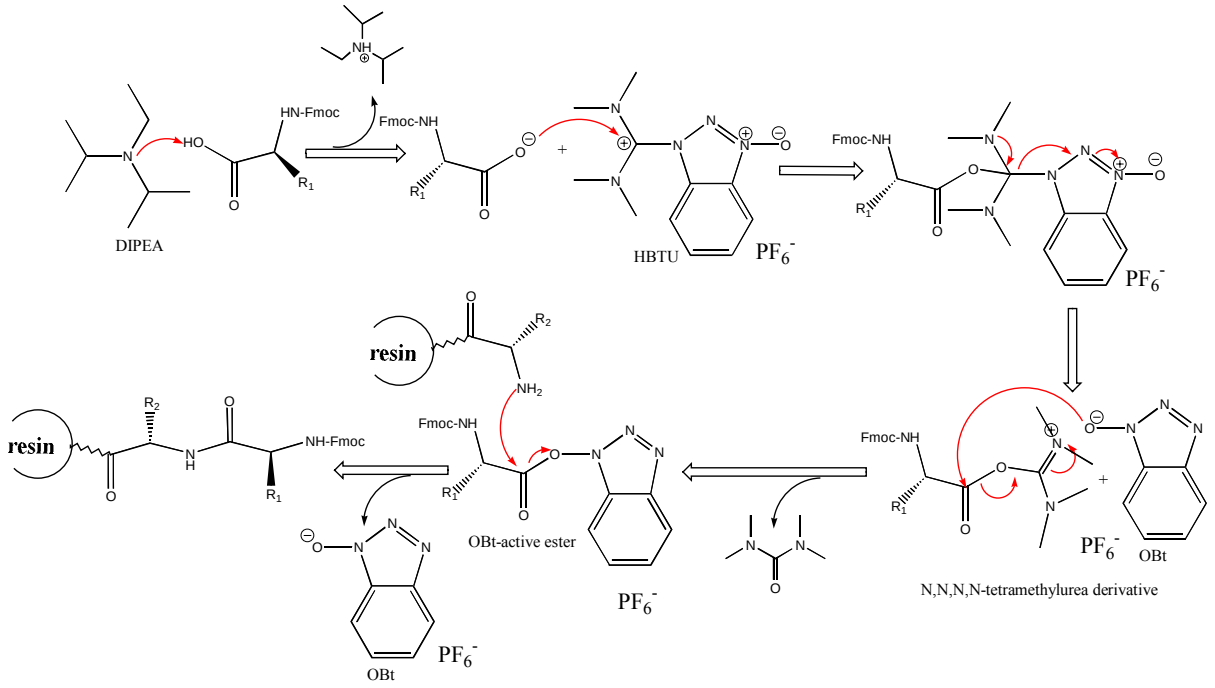


Figure 4.3. Amino acid activation and peptide elongation.

4.5.1.3 Termination

Cleavage of the mature peptide off the resin was carried out with a mixture of acetic acid/trifluoroethanol/DCM (1:2:7). The acid labile 2-chlorotritylchloride-linker releases the peptide while all other acid labile protecting groups are stable under these conditions. The peptide is washed in hexane.

4.6 Chemical Synthesis

4.6.1 Synthesis of SNAC and Thiophenol Thioesters

Activation of the peptide C-terminus was carried out *in situ* by using 2 eq of dicyclohexylcarbodiimide (DCC) and the additive HOBt in THF. K_2CO_3 was also added to ensure deprotonation of the carboxy-group which forms O-acylisourea with the C=N-double bond of DCC. This intermediate was labile and transformed into a benzotriazol-ester by nucleophilic attack of HOBt. The benzotriazol-ester leaving group of this active ester was finally replaced by SNAC or thiophenol, which were added in a 10 times excess. Therefore the reaction was stirred for 3 h at room temperature.

Acid labile side chain protecting groups were cleaved under strong acidic conditions. In this work a mixture of TFA/H_2O/triisopropylsilane 95:2.5:2.5 (v/v) was used. The peptide was incubated for 3 h and precipitated in cold ether. The deprotected peptides were purified by

preparative high performance liquid chromatography (HPLC) on a Äkta Purifier (Amersham Biosciences) HPLC-System with a reversed phase C18 Nucleodur (Macherey and Nagel) column. The identities of peptide thioesters were verified by high performance liquid chromatography-mass spectrometry (HPLC-MS) and matrix-assisted laser desorption ionization-time-of-flight (MALDI-TOF) MS.

4.6.2 Synthesis of Peptidyl-CoA

The general procedure for the synthesis of peptidyl-CoA substrates was based on a synthesis described previously[57]. To 1 eq of protected peptide 1 – 2 eq Coenzyme A, 1.5 eq PyBOP and 4 eq potassium carbonate were added and dissolved in a 1:1 THF/water mixture. The mixture was agitated for 2 h at rt and the solvent was removed. Side chain deprotection and purification were carried out as described in section 4.6.1. To confirm the thioester linkage all product peaks, identified by their mass, were additionally incubated with 0.1% KOH to hydrolyse the thioester bond which was again verified by LC-MS.

4.6.3 Synthesis of Cyclosporin A SNAC

Acetylation of Cyclosporine A - 1 g of cyclosporin A extracted from tablets (Novartis) was dissolved in 40 ml acetanhydride/pyridine 1:1 and 1.47 mg DMAP was added. The mixture was agitated for 2 h at 45°C. The solvent was removed by rotary evaporation, 40 ml chloroform was added and washed 3 times with water. After purification on silica gel ($CHCl_3$/CH_3OH, 98:2), 0.61g acetylcyclosporin was isolated and identified by NMR and MALDI based on literature data[122].

Synthesis of Acetylcyclosporin 7-thioamidate – 1 g of acetylcyclosporin was dissolved in 20 ml xylene and heated to 30 °C internal temperature. 0.12 g Lawesson reagent was added stepwise. After 30 min the reaction was cooled to room temperature and evaporated. The product was isolated by flash chromatography in diethylether/water 4:1 and identified by NMR[122].

Synthesis of Acetylcyclosporin 7-(benzyl thioamidate) – A solution of 251.7 mg acetylcyclosporin 7-thioamidate in 10 ml DCM, 0.3 ml DBU and 0.2 ml benzylbromide was kept at room temperature for 2 h. Tert-butyl methyl ether was added and washed 3 times with water. After flash chromatography with saturated diethylether the product was identified by NMR[122].

Synthesis of Acetylcyclosporin carboxylic acid S-benzylester – 46.5 mg Acetylcyclosporin 7-(benzyl thioamidate) in 1 ml acetonitrile was added within 2 min to a mixture of 0.25 ml 6 N HCl in 1 ml acetonitrile. The mixture was stirred for 20 min and the reaction was monitored

by TLC. The pH was adjusted to 11 and the product extracted with tert-butyl methyl ether. The product was confirmed by MALDI.

Subsequent conversion of the benzyl-thioester into the SNAC thioester was carried out according to standard procedures for SNAC synthesis.

5 Results

The exceptional biological and pharmacological activities of cyclic natural peptides produced by NRPS make them interesting targets for drug discovery efforts. In order to identify therapeutics with improved or altered activities, structural diversity has to be introduced into a selected molecular scaffold. The generation of such a library can be achieved by two different approaches:

1. Domain and module exchange in the multienzymatic machinery.
2. Chemical synthesis of linear peptides and enzymatic cyclization.

Protein engineering requires knowledge of the topology and architecture of NRPS multienzyme complexes in order to reprogram the enzymatic synthesis machinery by module or domain swapping. Data about the quaternary organization of NRPSs will be presented in the first part of the results section.

An alternative route to increase the diversity of cyclic peptide products can be achieved by a chemoenzymatic approach. In this approach the whole enzymatic machinery required for peptide elongation is replaced by chemical SPPS, which allows the easy and fast synthesis of diverse linear peptide analogues. Subsequently regio- and stereoselective peptide cyclization is catalysed by unique NRPS peptide cyclases. In the second part of the results section, mechanism, tolerance and activity of various peptide cyclases will be presented.

5.1 Protein Architecture: The Quaternary Structure of NRPS

The objective of this research effort was to explore the quaternary structure of NRPS with strategies similar to those, that have been successfully applied to discover the structure of PKS and FAS (chapter 2.8). Investigations were carried out with several representative NRPS systems different in size and assembly line organization: GrsA (domains A-PCP-E) [123], $TycB_1$ (domains C-A-PCP) [124], the dimodular enzyme $TycB_{2-3}$ (domains A-PCP.C-A-PCP-E) [71] and EntF (domains C-A-PCP-TE) [84, 125] (Fig. 5.1).

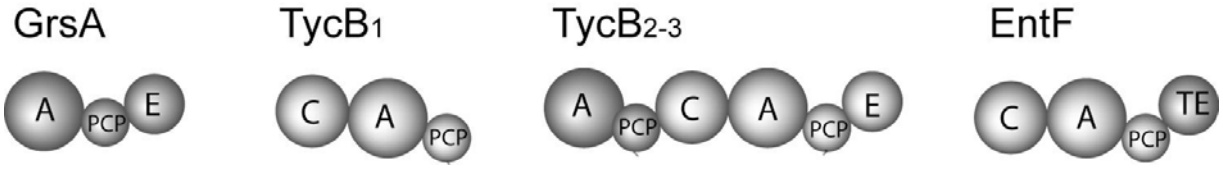


Figure 5.1. NRPS enzymes investigated. GrsA (127 kDa), $TycB_1$ (119 kDa), $TycB_{2-3}$ (237 kDa) and EntF (143 kDa).

The first part of the strategy focused on the chemical and biophysical properties of the enzymes. Gel filtration, ultracentrifugation and chemical cross-linking are useful methods to gain general information about size and molecular weight of a protein. The second part investigated the functional complementation between subunits carrying inactivated domains, which would additionally give a biochemical proof of a dimeric structure. In order to complete the biochemical part of the strategy, enzymes with two different affinity tags were constructed. Only a putative heterodimer with two different tags should be able to bind to two successive affinity columns. Parts of these investigations were already described in a previous diploma thesis and will therefore only be summarized here[126].

5.1.1 Gel Filtration, Chemical Crosslinking and Equilibrium Analytical Ultracentrifugation

First attempts to elucidate the quaternary structure of NRPS modules were carried out utilizing gel filtration experiments (Tab. 5.1). Each NRPS enzyme, GrsA (127 kDa), $TycB_1$ (119 kDa) and $TycB_{2-3}$ (237 kDa), eluted with an apparent monomeric weight. No protein elution at a corresponding dimeric or higher mass was observed. In contrast, gel filtration of the mixed NRPS-PKS hybrid enzyme HMWP1 (350 kDa) carried out under the same conditions revealed significant formation of a monomer (424 kDa) and an additional elution peak of a protein with the approximate weight of a dimer (721 kDa).

Table 5.1. Results of size exclusion chromatography. The table shows the theoretical (mt) and obtained (mo) molecular weights of all four applied proteins gained under the conditions as described in materials and methods. The quotient m_0/m_t is a measure how much the obtained weight exceeds the theoretical monomeric weight of the proteins (a value of 1 is obtained for monomers and a value of 2 for dimers).

enzyme	theoretical mass m_t (kDa)	obtained mass m_o (kDa)	quotient m_o/m_t
GrsA	127	167	1,31
ProCAT	119	162	1,36
$TycB_{2-3}$	237	364	1,53
HMWP1	350	721; 424	2.06; 1.21

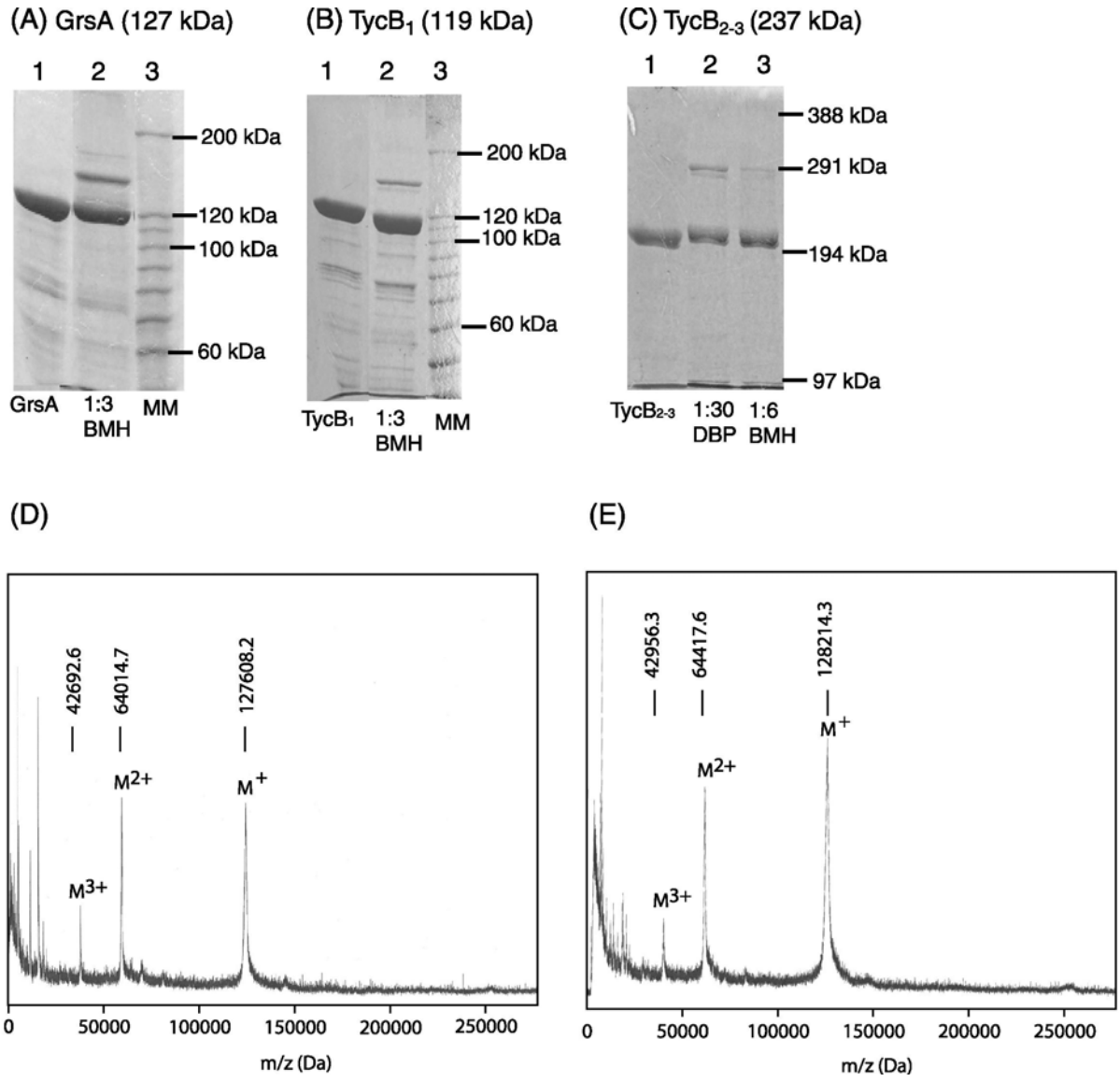


Figure 5.2. Results of cross-linking experiments. (A) 7.5 % polyacrylamide gel of GrsA (line 1) and GrsA after incubation with a 3-fold excess of BMH (line 2). A new protein species after incubation with BMH can be obtained corresponding to an electrophoretic mobility of ~140-160 kDa. (B) 7.5 % polyacrylamide gel of $TycB_1$ (line 1) and $TycB_1$ after incubation with a 3-fold excess of BMH. Also in case of $TycB_1$ a similar shift in electrophoretic mobility from ~120 before to ~140-160 kDa after incubation with cross-linker can be observed. (C) 5 % polyacrylamide gel of $TycB_{2-3}$ (line 1), $TycB_{2-3}$ after incubation with a 30-fold excess of DBP (line 2) and after incubation with a 6-fold excess of BMH (line 3). DBP as well as BMH lead to the formation of a new protein species corresponding to an electrophoretic mobility of ~290 kDa. (D) MALDI-TOF mass spectra of GrsA without BMH treatment. The molecule ion peak was obtained at 127608 Da. (E) MALDI-TOF mass spectra of GrsA after treatment with BMH. The molecule ion peak is shifted towards a higher mass (128214 Da) in comparison with the spectra of the untreated sample. Peaks corresponding to the masses of a dimer were not observed.

Bifunctional cross-linking reagents such as 1,3-dibromopropanone (DBP) and bismaleimidohexane (BMH) have been shown to cross-link PKS modules as homodimers [100, 127]. DBP and BMH covalently connect thiol groups within a distance of 5 Å and 16 Å, respectively. The same cross-linking experiments were applied to these NRPS modules. The molecular weights of any cross-linked proteins were investigated using SDS-PAGE and MALDI-TOF mass spectrometry. To avoid non-specific linkages the cross-linkers were used only up to a 3-10 fold molar excess according to the manufacturers protocol (ICN, Pierce). After incubation with BMH, the electrophoretic mobilities of GrsA and $TycB_1$ proteins decreased marginally, resulting in a shift to higher molecular weights (from ~120 kDa to ~140-160 kDa) on SDS gels (Fig. 5.2), whereas in case of DBP no change of migration was observed (data not shown). On the basis of these results there is no indication for the existence of a dimer since the electrophoretic mobilities of the new bands (~140-160 kDa) are in between the monomeric (~120 kDa) and dimeric weight (~240 kDa) of GrsA and $TycB_1$. Similar results were obtained for the larger dimodular enzyme $TycB_{2-3}$ (237 kDa) which revealed one new protein band after reaction with both BMH and DBP, corresponding to an electrophoretic mobility of a protein of ~290 kDa (Fig. 5.2). The determination of the electrophoretic mobility of the new protein band was carried out in direct comparison with a parallel cross-linking of phosphorylase b (97-582 kDa)[127]. Again, for the larger dimodular enzyme the electrophoretic mobility of the new protein band does not represent a dimer (474 kDa). Analogous experiments were carried out with dimethyl suberimidate (DMS), a cross-linker which connects primary amines in a distance of about 11 Å. None of these cross-linking reactions revealed altered protein patterns on SDS gels (data not shown). The cross-linking results were confirmed by application of MALDI-TOF mass spectrometry. Direct comparison of cross-linker treated and untreated samples revealed neither peaks for GrsA and $TycB_1$ corresponding to the singly charged molecular ions (~250 kDa) nor to odd multiply charged species (M^{3+} ~85 kDa and M^{5+} ~50 kDa) of the dimers (MALDI-TOF spectra of GrsA are shown in figure 2D and 2E). The differences of the molecular masses of treated and untreated samples, observed were “only” ~1500 Da for $TycB_1$ and ~600 Da for GrsA larger than the monomer molecular weights, indicated the reaction of 5 and 2 molecules of BMH (276 Da), respectively. $TycB_{2-3}$ was too large for MALDI-TOF investigations.

Analytical ultracentrifugation was applied to determine the native molecular weight for the dimodular enzyme fragment $TycB_{2-3}$ and the monomodular enzyme EntF.

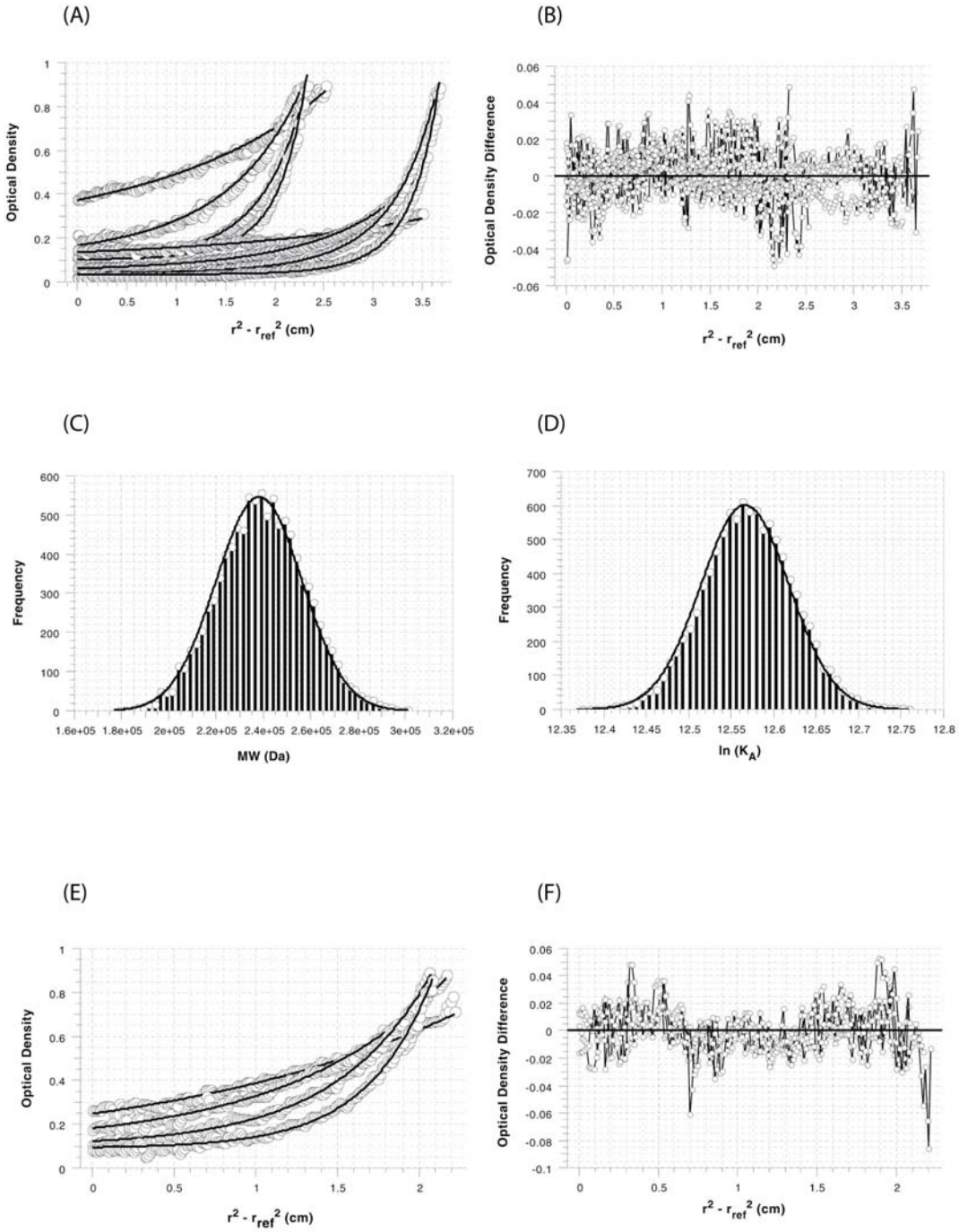


Figure 5.3. Sedimentation equilibrium analysis of $TycB_{2-3}$ and EntF. (A) Overlay of 8 wavelength scans and (B) residuals from the global fit of the data of $TycB_{2-3}$ as described in Materials and methods. The Monte Carlo distribution of the molecular weight (C) and the association constant (D) (ln Ka) are shown to establish an experimental M_r = 238.07 ± 18.21 kDa (the protein sequence predicted molecular weight is 237.69 kDa) for the monomer and a dissociation constant of K_d = 3.489 ± 0.186 µM for the $TycB_{2-3}$ enzyme. (E) Overlay of 4 wavelength scans and (F) residuals from the global fit of the data of EntF as described in Materials and methods. The Monte Carlo distribution of the molecular weight (not shown) establishes an experimental M_r = 137.17 ± 8.17 kDa for the monomer which is in agreement with the protein sequence predicted molecular weight (142.97 kDa).

To establish the association properties of $TycB_{2-3}$, sedimentation equilibrium experiments were conducted at several protein concentrations (Fig. 5.3). For the global equilibrium analysis, 8 scans (Fig. 5.3A) of speeds ranging from 5,000 to 12,500 r.p.m. and 0.39 and 0.57 O.D. loading concentrations were fit to a monomer-dimer model. Monomer molecular weight and the association constant were floated global parameters and were forced to be the same for all the included data sets. The system could be well described with this monomer-dimer model, which resulted in random residuals (Fig. 5.3B) and a monomer molecular weight of 238.07 ± 18.21 kDa (Fig. 5.3C), which is in agreement with the protein sequence predicted molecular weight (237.69 kDa). The dissociation constant of 3.489 ± 0.186 μM suggests the presence of dimer only at unphysiologically high concentrations, which were required for ultracentrifugation experiments (Fig. 5.3D).

Sedimentation equilibrium analysis of EntF was conducted at a single protein concentration (Fig. 5.3E and 5.3F). For the global equilibrium analysis, 4 scans (Fig. 5.3E) of speeds ranging from 8,000 to 17,000 r.p.m. and a 0.40 O.D. loading concentration were fit to a one-component (monomer only) model. Molecular weight was a floated global parameter and was forced to be the same for all the included data sets. The system could be well described with this monomer only model, which resulted in random residuals (Fig. 5.3F) and a molecular weight of 137.17 ± 8.17 kDa, which is in agreement with the protein sequence predicted molecular weight (142.97 kDa). The data fitting did not improve with a monomer-dimer model for EntF, and the experimentally fit dissociation constant (not shown) was many orders of magnitude larger than the loading concentration. Therefore, neither gel filtration, analytical ultracentrifugation, nor cross-linking supported a strong dimeric interaction of the investigated tyrocidin synthetase and enterobactin synthetase components.

5.1.2 Dissociation and Reassociation of Subunits of a Putative Dimer

In this study, two different biochemical methods were applied to address the quarternary structure of NRPS enzymes. Similar experiments were carried out previously to verify the dimeric structure of FAS and PKS [99, 109, 110, 113, 128]. Both methods were designed to detect heterodimeric enzymes, and give negative results for monomeric enzymes. The critical step, which is necessary for both methods is the dissociation and reassociation (scrambling) of putative homodimers to form heterodimers with different affinity tags or different domain-inactivating point mutations.

Two different approaches to perform scrambling were utilized. Anion exchange chromatography was applied in the same way as described previously for PKS-systems [127]. Dissociation is induced by a strong interaction of putative dimer subunits with the matrix

material and reassociation can occur by increasing the salt concentration resulting in the elution of the proteins. The aminoacylation activity of the tested enzymes did not drop after application on the anion exchange column. The aminoacylation activities before and after application to the anion exchange column of GrsA-ATE and of $TycB_{2-3}$-AT.CATE with [^{14}C]-phenylalanine increased from 20% to 25% and from 60% to 65%, respectively, that of $TycB_1$-CAT with [^{14}C]-proline decreased negligibly from 15% to 13%, with all differences being in the margin of error (5%) of the method. A second approach to perform dissociation and reassociation of putative dimers was carried out with the zwitterionic detergent CHAPS. One advantage of CHAPS is its non-denaturating character, which typically only directs the inhibition of protein-protein interactions leaving the tertiary structure intact. The attachment of CHAPS molecules onto the unpolar protein residues leads to the dissociation of putative dimers and to the formation of monomers covered in micelles [129]. Subsequent dialysis induces the destruction of micelles and allows the formation of putative heterodimers. To demonstrate the favoured properties of CHAPS some preliminary experiments were implemented. To allow micelle formation CHAPS was used at a concentration of 15 mM, which is well above the critical micellar concentration. The aminoacylation activities of GrsA and $TycB_1$ with their dedicated [^{14}C]-marked amino acids did not change during and after incubation with 15 mM CHAPS indicating a functioning tertiary structure. The aminoacylation activity of GrsA with [^{14}C]-phenylalanine changed only negligibly from 15% to 17% after addition of 15 mM CHAPS and remained at 12% after dialysis. Aminoacylation activity of $TycB_1$-CAT with [^{14}C]-labeled-proline as substrate resulted in a minor decrease from 21% to 16% after addition of 15 mM CHAPS and remained at 20% after dialysis.

The interaction of GrsA-ATE with $TycB_1$-CAT *in trans* leads to the formation of a linear dipeptide, which is cleaved by a non-catalyzed cyclization reaction resulting in the release of D-Phe-L-Pro-diketopiperazine (DKP-assay) [123]. This assay was applied to demonstrate that in the presence of 15 mM CHAPS the D-Phe-L-Pro-DKP (1,3-diketopiperazine) formation decreased by greater than 80 % in comparison with the untreated sample. The *in trans* reaction between GrsA and $TycB_1$ is conducted via protein-protein interactions and was effectively inhibited by CHAPS. These two experiments showed that CHAPS interrupts protein-protein interactions for Grs and Tyc proteins and therefore has the necessary prerequisite to perform scrambling.

5.1.2.1 Two Affinity Tag System

The specific and reversible binding of proteins with two different affinity tags to corresponding affinity columns was used to test for heterodimer formation. In this experiment

two different affinity tags were utilized. Affinity chromatography on immobilized Ni-ions allows the rapid purification of proteins with a fused hexahistidine-(His_6-)tag. The strep-tag is a small peptide with a high affinity to genetically modified streptavidin. The dimeric character of a protein can be demonstrated by scrambling of two identical homodimers, which only differ from their affinity tag. The resulting heterodimer consists of two subunits with two different affinity tags. This heterodimeric protein has the unique ability to bind with its hexahistidine tag first on a Ni-NTA column and subsequently with its strep-tag on a streptavidin column or the other way round. The detection of a protein after the second column would be a proof for a dimer. Similar experiments with two affinity tags were carried out with dimeric FAS previously [128].

In this work both scrambling methods were applied to conduct putative his_6-tag/strep-tag-heterodimer formation of GrsA and $TycB_1$, respectively. The scrambled protein samples were loaded on both affinity columns successively under conditions optimized to prevent non-specific binding, and all elution fractions were investigated with specific hexahistidine-tag and strep-tag antibodies. However, in all cases it was not possible to detect any protein in the elution fraction of the second column, (data not shown) indicating that no heterodimers existed in solution.

5.1.2.2 Complementation

Heterodimers of fatty acid synthases and polyketide synthases carrying mutations in distinct domains are able to balance their loss of activity by mutual complementation of the remaining active domains [99, 109, 110, 113]. The success of the mutual complementation was visualized by the ability to form product, which was a direct proof of a dimeric behaviour. In this work the same complementation strategy for NRPS was used and the cooperation of PCP- and C-domains from $TycB_{2\text{-}3}$ after scrambling was investigated (Fig. 5.4). Both scrambling methods were applied to conduct the formation of the putative heterodimers $TycB_{2\text{-}3(PCP)}$(S560A)/$TycB_{2\text{-}3(C\text{-}domain)}$(H741R) and $TycB_{2\text{-}3(PCP)}$(S1593A)/$TycB_{2\text{-}3(C\text{-}domain)}$(H741R). The phenotypes of the single mutants are the following: S560A, inactivation of the first PCP by mutation of the cofactor binding invariant serine residue resulting in the loss of the ability to covalently bind substrates to module one; H741R, inactivation of the C-domain resulting in the breakdown of condensation activity; S1593A inactivation of the second PCP by mutation of the cofactor binding invariant serine residue resulting in the loss of the ability to covalently bind substrates to module two.

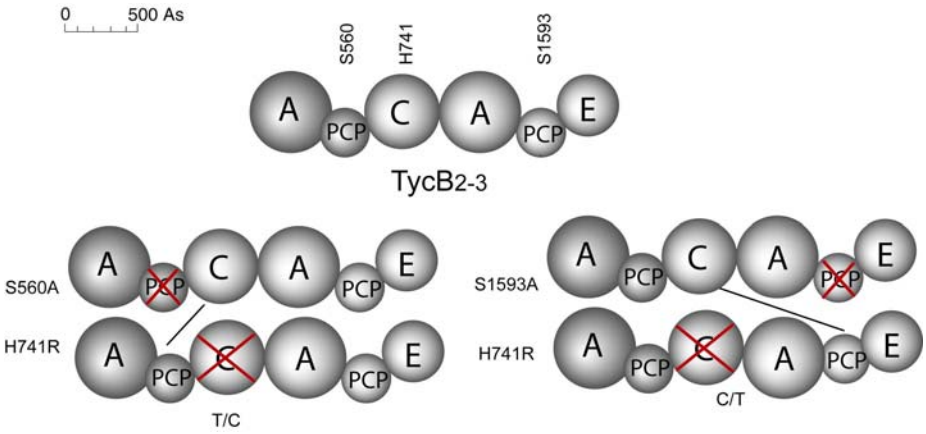


Figure 5.4. Complementation systems. For the complementation assays three different mutants of $TycB_{2-3}$ were used with mutations in the first PCP (S560A, resulting in the loss of the ability to covalently bind substrates to module one), the C-domain (H741R, resulting in the breakdown of condensation activity of the enzyme) and in the second PCP (S1593A, resulting in the loss of the ability to covalently bind substrates to module two) leading to two complementation systems. The putative heterodimeric systems would either allow the interaction of the upstream T-domain with the C-domain or the interaction of the C-domain with the downstream T-domain (indicated by red lines). In case of real heterodimers, product formation should be observed although all three mutated enzymes show no activity by themselves.

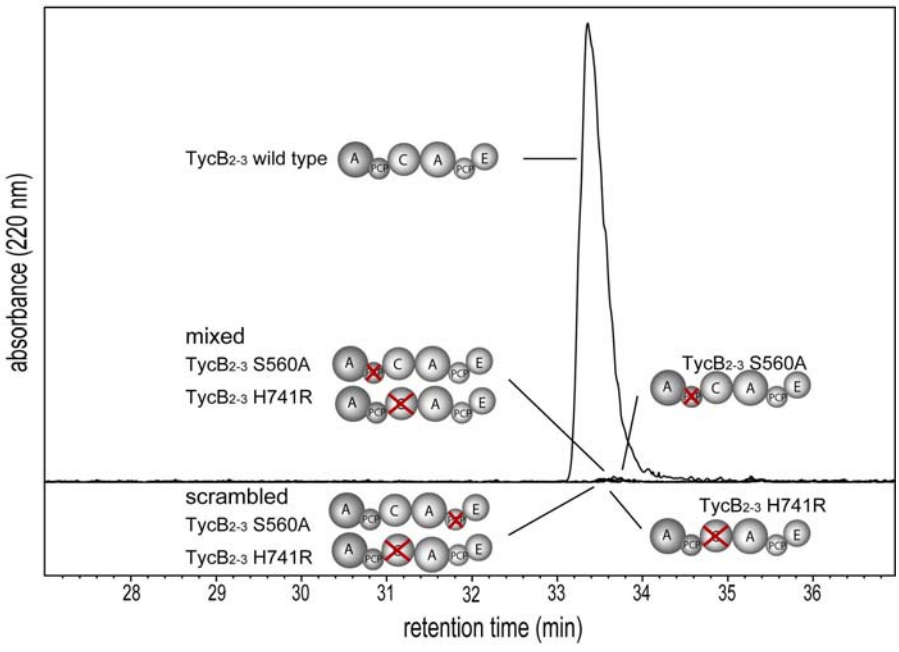


Figure 5.5. HPLC trace of the tripeptide product formation assay on the example of the S560A/H741R $TycB_{2-3}$ complementation system. Only the wild type enzyme is able to produce the tripeptide. The mutated enzymes S560A, H741R did not show any detectable activity alone. Also the mixture of S560A/H741R as well as a mixture of S1593A/H741R with and without Source Q scrambling did exhibit any activity above background. The same results were gained for the other complementation system and scrambling method.

Subsequently their ability for product formation was investigated via a tripeptide assay described previously [71]. Control reactions under the same conditions were carried out using the wild-type enzyme without mutations, the mutants alone, and a mixture of both mutants without the process of scrambling. As shown in figure 5.5, no product formation was observed in all experiments except in the case of the wild-type enzyme, indicating that there was no functional complementation between the enzymes inactivated in distinct domains.

5.2 Natural Product Architecture: Chemoenzymatic Peptide Cyclization

A common structural feature of nonribosomally produced peptides is the constrained structure which ensures precise functionality important for the interaction with the dedicated cellular receptor. In many cases rigidity is achieved by cyclization strategies of the peptide backbone leading to diverse macrolactams and macrolactones. Because peptide cyclization can be a difficult task in chemical synthesis there has been a significant interest in exploring enzymatic cyclization mechanisms to develop new synthesis routes for the production of modified natural cyclic peptides. Nature achieves the synthesis of linear peptide precursors by using the NRPS machinery and subsequent cyclization is catalysed by associated peptide cyclase domains (TE).

In order to use these enzymes for the cyclization of synthetic peptides, the autonomous activity of Tyc TE, Srf TE and Grs TE had been shown before[80, 90, 91]. The results presented here follow up on these initial findings and further investigate enzyme tolerated substrate variations on the example of Srf TE in order to evaluate its utility as a valuable cyclization tool for organic synthesis. This section then concentrates on investigations of the underlying mechanism of cyclization by mutational analysis of catalytically important residues of Srf TE based on its crystal structure. In the course of these studies peptide cyclases from various other systems were investigated. Their lack of activity with conventional SNAC substrates indicated an unsufficient translation between the language of biology and the language of chemistry by the small natural substrate surrogates. New methods to retain activity with all these enzymes will be presented here in order to generate a catalytic tool set of enzymes capable of cyclizing various precursors. Another approach to catalyse cyclization of various precursors is the generation of custom made peptide cyclases by protein evolution. The last part of this section focuses on initial results of such a protein evolution approach to engineer new peptide cyclases with altered substrate specificities.

5.2.1 Biochemical Characterization of Surfactin TE

All peptide substrates were characterized by mass spectrometry (table 4.3) and mutations confirmed by DNA sequencing. The work was carried out in cooperation with Claire Tseng.

5.2.1.1 Incubation of SrfTE with the Native SNAC Substrate

It was previously reported that Srf TE catalyzed the stereospecific cyclization of a linear surfactin precursor peptidyl-SNAC (N-acetylcysteamine) thioester that was a mimic of the acylheptapeptide sequence of authentic surfactin[80] and SNAC a mimic of the natural

cofactor ppan. To improve water solubility, the sequence of this substrate differed from that of the wild type substrate in that D-ornithine replaced D-Leu_3 and (*R*)-3-hydroxybutyric acid replaced the β-hydroxy C_{13}-C_{15} fatty acid (**SLP-Orn**). Reactions of **SLP-Orn** with Srf TE, in the presence of 1% glycerol as an enzyme stabilizer, produced a 1:1:1 ratio of the products of cyclization, hydrolysis, and glycerolysis, with similar k_{cat} (15-22 min^{-1}) and K_M (4.0-5.6 mM) values for all three products[80].

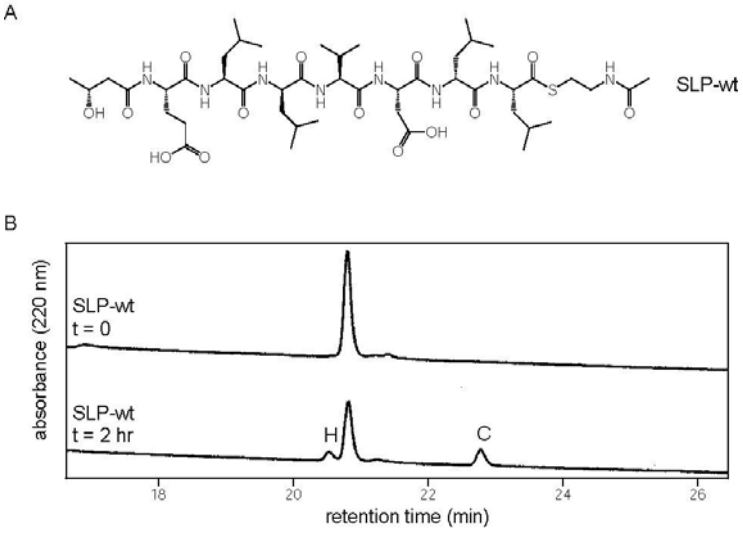


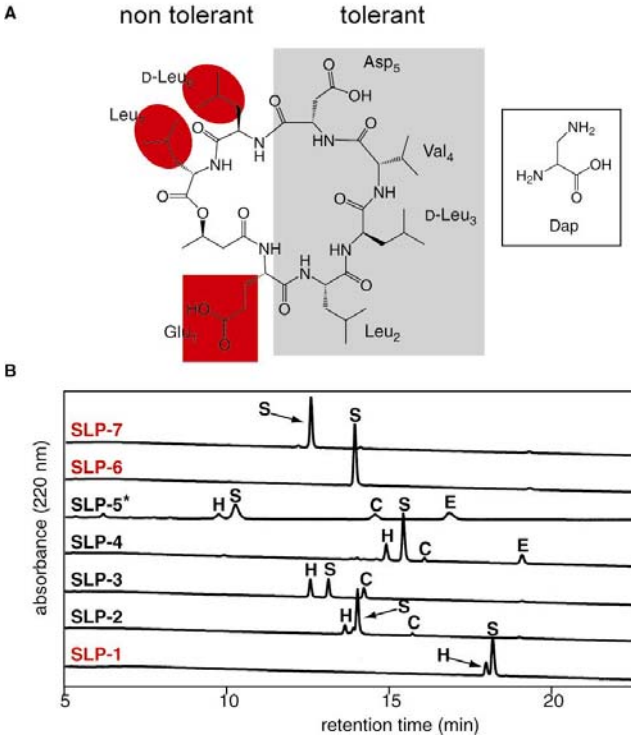
Figure 5.6. Cyclization and hydrolysis of a peptide-SNAC thioester substrate by Srf TE. (A) Structure of SLP-wt. (B) HPLC traces of 2 h reaction initially containing 2.5 μM Srf TE, 250 μM SLP-wt, and 25 mM MOPS, at pH 7.0 and 24 °C. The peaks corresponding to the cyclic (C) and hydrolyzed (H) product are labeled.

As a baseline to study a substrate more similar to the wild type substrate, a peptidyl-SNAC was synthesized that included the native surfactin heptapeptide sequence, but retained the β-hydroxy-C_4 fatty acid substitution of **SLP-Orn** (**SLP-wt**, Fig. 5.6A). **SLP-wt** was soluble in water to 2 mM, and in DMSO to 50 mM. Reactions of **SLP-wt** (250 μM, 0.5% DMSO, 25 mM MOPS, pH 7.0) with Srf TE (2 μM) in the absence of glycerol produced a 2.5:1 ratio of the cyclic:hydrolyzed products, with a k_{cat}/K_M of 0.86 $mM^{-1}min^{-1}$ for cyclization and 0.35 $mM^{-1}min^{-1}$ for hydrolysis (Fig. 5.6B). The k_{cat} and K_M values could not be deconvoluted, because K_M concentrations of substrate could not be reached due to the low solubility of **SLP-wt** in water. These solubility limitations are expected to be exacerbated with the C_{13}-C_{15} acyl chains in the wild type heptapeptide sequence of surfactin. Nonetheless, product analysis of **SLP-wt** showed that Srf TE could direct about 70% of the flux of β-hydroxy-acyl-heptapeptidyl-*O*-enzyme intermediate to the macrolactone ring of the natural product.

5.2.1.2 Dap Scan of the Heptapeptide Sequence

To evaluate enzymatic recognition elements in the heptapeptidyl substrate chain, a systematic alteration of each of the side chains of the substrate was considered. It was anticipated that an alanine scan would not improve the hydrophobic constraints on solubility; instead, a 2,3-diaminopropionic acid (Dap) residue with an ionizable 3-amino group was introduced at each

of the seven residue positions of **SLP-wt**, with the stereochemistry of each residue maintained (Fig. 5.7A). The resulting seven peptides, **SLP-1** to **SLP-7**, were assessed as substrates for Srf TE by determining the kinetic parameters and cyclization:hydrolysis ratio of Srf TE on each substrate, with the results summarized in Fig. 5.7. The solubilities of the Dap-containing substrates in water were much improved (25-50 mM) over **SLP-wt**, except when the Dap replaced an acidic residue, as in **SLP-1** and **SLP-5**, which displayed solubility properties similar to **SLP-wt**. However, separate k_{cat} and K_M values could still not be measured for those substrates with improved solubility, because severe substrate inhibition was observed at concentrations greater than 3 mM. This inhibition could be related to the critical micellar concentration (CMC) of these substrates[130], which were measured to be in the 0.6-2.0 mM range (data not shown).



C

	cyclization k_{cat}/K_M (mM^{-1} min^{-1})	hydrolysis k_{cat}/K_M (mM^{-1} min^{-1})	C/H ratio
SLP-wt	0.86	0.35	2.5
SLP-1	N/D	4.1	N/D
SLP-2	0.30	1.2	0.25
SLP-3	2.9	5.4	0.54
SLP-4	0.46	1.5	0.31
SLP-5	2.7	2.4	1.2
SLP-6	N/D	N/D	N/D
SLP-7	N/D	N/D	N/D

Figure 5.7. Results of Dap scan. (A) Cyclic product of SLP-wt. The residues for which replacement by Dap results in neither cyclization nor hydrolysis are highlighted by red ovals. The residue for which replacement results in only hydrolysis is highlighted by a red square. The grey squared residues can be both cyclized and hydrolyzed when replaced by Dap. (B) HPLC traces of 1 hr reactions initially containing 2 µM Srf TE, 500 µM indicated substrate, and 25 mM MOPS, at pH 7.0 and 24°C. The peaks corresponding to the cyclic (C) and hydrolyzed (H) products, substrate (S) and enzyme (E) are labeled. The trace corresponding to SLP-5 (labeled with asterisk) was run on a slower gradient than the other traces in order to achieve good peak separation. The identities of the substrates and products were verified by MALDI-TOF MS. (C) C/H ratios and k_{cat}/K_M values of cyclization and hydrolysis for each substrate (standard deviation ±10%). N/D = activity not detected.

Nonetheless, the Dap scan allowed to determine which residues were tolerant to a significant change of both size and charge. The residues at the N- (Glu_1) and C-terminus (D-Leu_6 and Leu_7) of the substrate were not tolerant for the Dap substitution, resulting in either no product formation or only hydrolysis (Fig. 5.7). On the other hand, the residues in the middle of the substrate (Leu_2, D-Leu_3, Val_4, and Asp_5) were tolerant to Dap substitution to varying degrees, with each forming measurable amounts of both cyclized and hydrolyzed products. Unexpectedly, the k_{cat}/K_M of Srf TE for the cyclization of **SLP-3** was almost four times greater than that for **SLP-wt**, although the k_{cat}/K_M for hydrolysis increased even more, with a cyclization:hydrolysis ratio of 0.54:1 (Fig. 5.7C). Qualitatively, this increase appeared to be due mainly to an increase in k_{cat}. **SLP-3** was therefore used as scaffold for further experiments.

5.2.1.3 Spacer Scan

A residue scan orthogonal to the Dap scan was also performed, in which two adjacent residues at a time, from Glu_1 to D-Leu_6, were replaced by the flexible spacer ε-aminohexanoic acid (ε-Ahx), in order to address which residue side chains are necessary for productive substrate positioning in the enzyme active site (Fig. 5.8). Kinetic parameters and cyclization:hydrolysis ratios of Srf TE on each substrate were determined, and found that, unlike in the Dap scan, the removal of residue side chains towards the N- (**SLP-1/2**) and C-terminus (**SLP-5/6**) of the substrate did not dramatically affect the cyclization:hydrolysis ratios of the resulting substrates. In contrast, the removal of residue side chains in the middle of the substrate (**SLP-3/4**) resulted in significantly decreased cyclization by Srf TE, with a cyclization:hydrolysis ratio of 0.057:1 (Fig. 5.8). These divergent results highlight the different substrate recognition properties probed in the two scans.

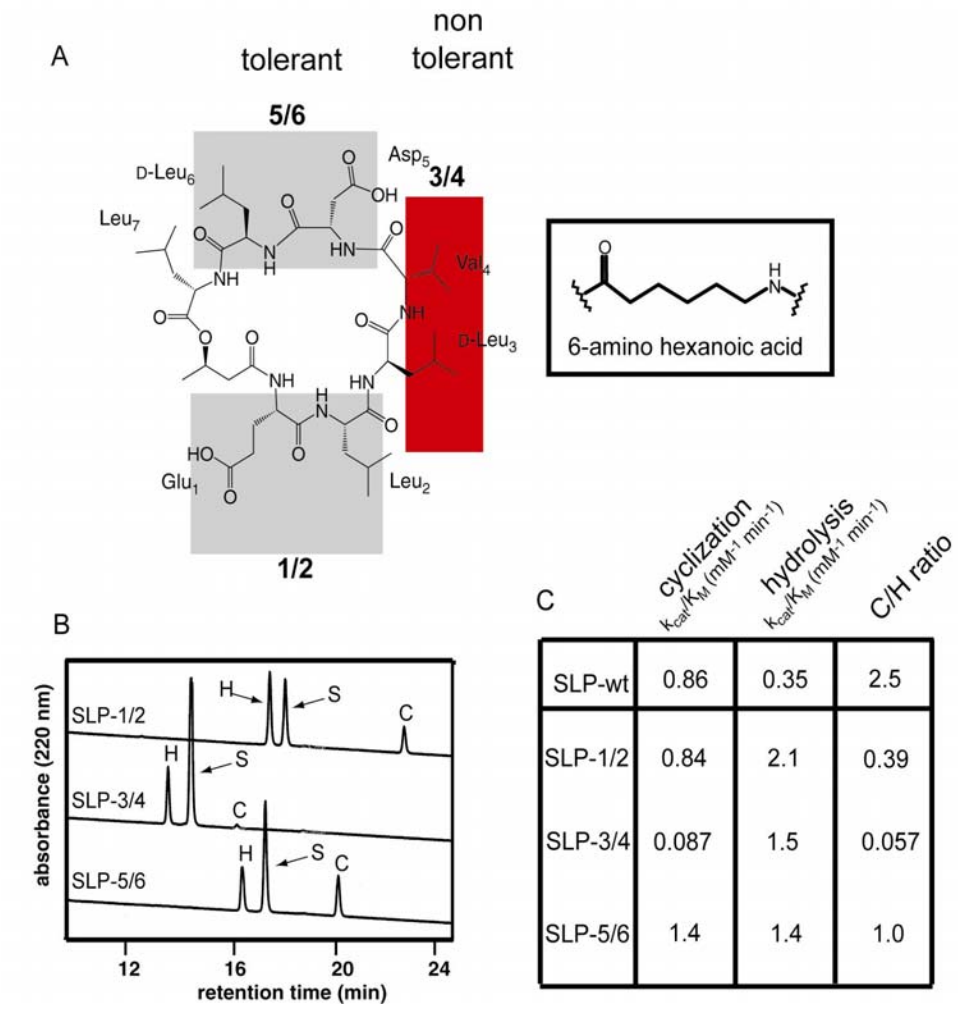


Figure 5.8. Results of spacer scan. (A) Cyclic product of SLP-wt. The residues for which replacement by ε-Ahx results in cyclization and hydrolysis are highlighted by grey squares. The residues for which replacement results in only hydrolysis and little cyclization are highlighted by a red square. (B) HPLC traces of 2 hr reactions initially containing 2 μM Srf TE, 500 μM indicated substrate, and 25 mM MOPS, at pH 7.0 and 24°C. The peaks corresponding to the cyclic (C) and hydrolyzed (H) products and substrate (S) are labeled. The identities of the substrates and products were verified by LC-MS. (C) C/H ratios and k_{cat}/K_M values of cyclization and hydrolysis for each substrate (standard deviation ±10%). FA = (R)-3-hydroxybutyric acid.

5.2.1.4 Reverse Reaction: Hydrolysis of the Macrolactone

The cyclic macrolactone product of **SLP-wt** was also a substrate for Srf TE with the enzyme reopening the macrolactone at the ester bond and catalyzing the formation of the linear acid through water attack (Fig. 5.9). A time course shows enzyme mediated conversion of cyclic substrate into linear hydrolysed product. In addition, the acyl-enzyme intermediate could also be captured through nucleophilic attack by ONAC (N-acetylethanolamine), SNAC, or NNAC (N-acetylethylenediamine), with the resulting formation of linear esters, thioesters, or amides of the substrate, respectively (Fig. 5.9C). As expected, the presence of exogenously added

macrolactone inhibited the cyclization of SNAC substrates in a concentration-dependent manner (data not shown). The k_{cat}/K_M for the hydrolysis of the macrolactone was 0.4 $mM^{-1}min^{-1}$.

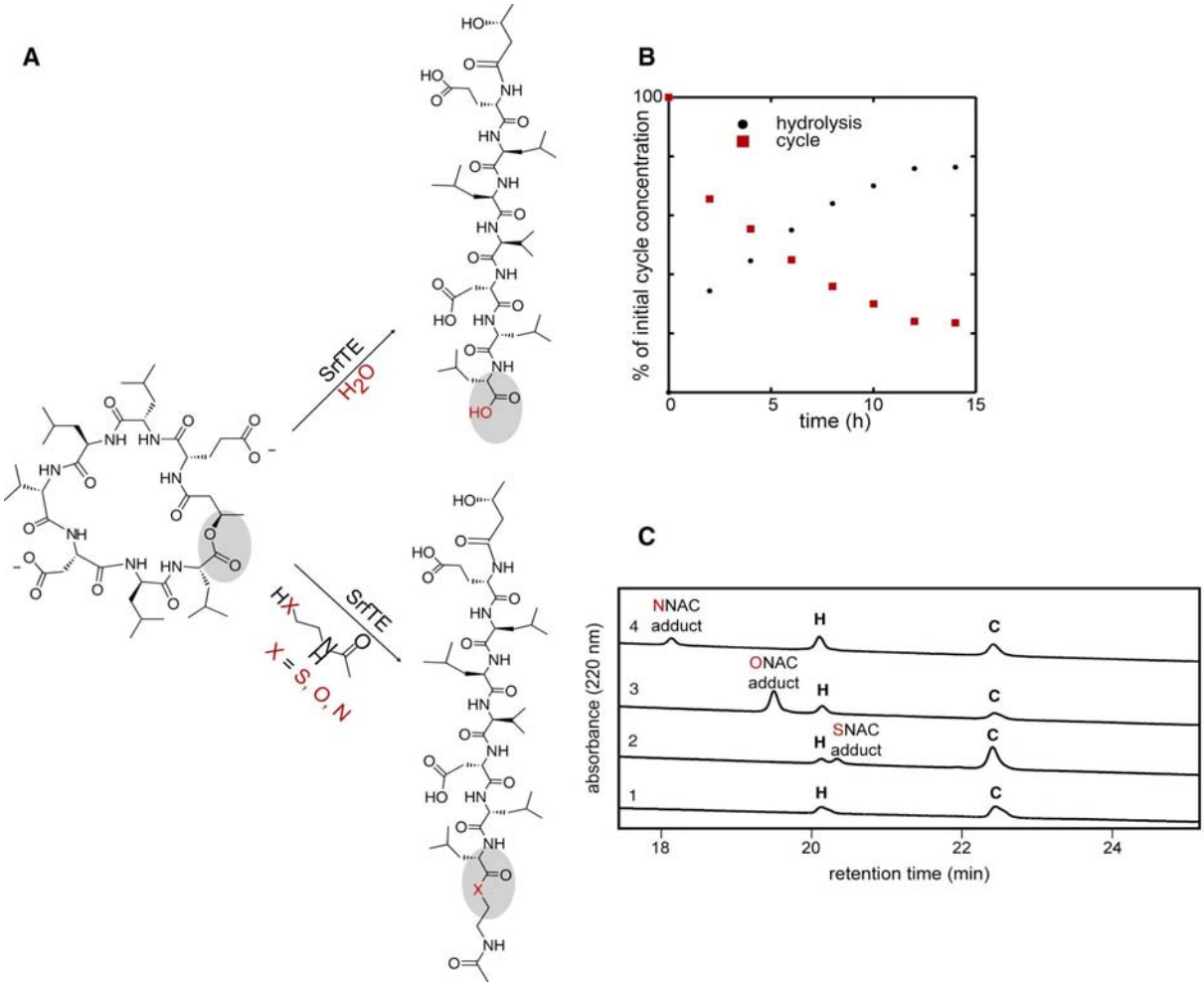


Figure 5.9. Enzymatic opening of macrolactone by Srf TE with different nucleophiles. (A) Mechanism of macrolactone opening. Srf TE can catalyze the attack of water or competing nucleophiles such as SNAC, ONAC, or NNAC to open the cyclic ester bond. (B) 14 hr time course of hydrolysis of the cyclic product of SLP-wt by Srf TE, with time points taken every 2 hr. (C) HPLC traces of 2 hr reactions initially containing 2 µM Srf TE, 250 µM cyclic product of SLP-wt, 25 mM MOPS, at pH 7.0 and 24 °C. Each reaction also contained 250 mM of the following nucleophile: 1) none, 2) SNAC, 3) ONAC, 4) NNAC. The peaks corresponding to the cyclic substrate (C), the hydrolyzed product (H), and the product resulting from attack by the added nucleophile are labeled. The structures of the substrate and products were verified by MALDI-TOF MS.

5.2.1.5 Alterations to Leaving Group, Nucleophile, and Length of Substrate

The ability of Srf TE to reopen the macrolactone product suggested that it might cyclize linear peptides with an ester leaving group. The thioester of **SLP-wt** was therefore changed to an ester through the replacement of the SNAC with ONAC. Srf TE was able to both cyclize and

hydrolyze the resulting substrate (**SLP-ONAC**), with a cyclization:hydrolysis ratio of 0.6:1 (Fig. 5.10), although the reaction occurred 5-10 times more slowly (data not shown). These results showed that Srf TE is able to tolerate a less active hydroxyl leaving group in addition to the native thiol leaving group. Assays with other thioester leaving groups will be presented below.

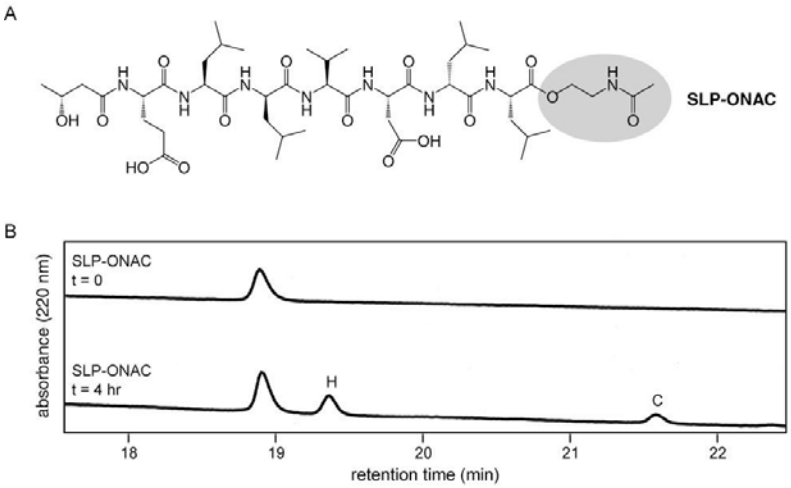


Figure 5.10. Cyclization and hydrolysis of peptide-ONAC substrate by Srf TE domain. (A) Structure of SLP-ONAC. (B) HPLC traces of 4 hr reaction initially containing 2 μM Srf TE, 250 μM SLP-ONAC, and 25 mM MOPS, at pH 7.0 and 24 °C. The peaks corresponding to the cyclic (C) and hydrolyzed (H) products are labeled. The structures of the substrate and products were verified by MALDI-TOF MS.

Variations were also made to the nucleophile-containing fatty acid of **SLP-3**, as summarized in Figure 5.11. The heptapeptide of **SLP-3** was used instead of that of **SLP-wt**, because its greater solubility in water allowed for easier synthesis and purification of the multiple alternate substrates. A substrate in which a secondary β-hydroxyl group was substituted for the secondary β-amino group of the fatty acid (**SLP-NH_2**) was not cyclized by Srf TE, although it was hydrolyzed (data not shown). Similar results were obtained on **(*R*)-** and **(*S*)-SLP-αOH**, which have a fatty acid containing a secondary α-hydroxyl group rather than a β-hydroxyl group, and also on a substrate in which the fatty acid was replaced by a serine (**SLP-Ser**), which contains a primary β-hydroxyl group. On the other hand, a substrate with a fatty acid containing a primary ε-hydroxyl group (**SLP-εOH**) was accepted as a substrate for both cyclization and hydrolysis, with a cyclization:hydrolysis ratio of 1.5:1 (Fig. 5.11). These initial data indicate altered macrolactone ring size is possible. To further test this possibility, a nonapeptide version of **SLP-wt** was synthesized, with a Leu inserted between Leu_2 and D-Leu_3 and another between Val_4 and Asp_5. However, similar to **SLP-NH_2**, this substrate was hydrolyzed but not cyclized by Srf TE (data not shown). In addition, structurally-unrelated peptidyl-SNAC substrates, including those with the peptide sequences of mycosubtilin (heptapeptide), fengycin (decapeptide), and tyrocidine (decapeptide), also showed substantial hydrolysis but no cyclization with Srf TE, indicating the enzyme's broad tolerance for hydrolysis but not cyclization of these substrates (data not shown).

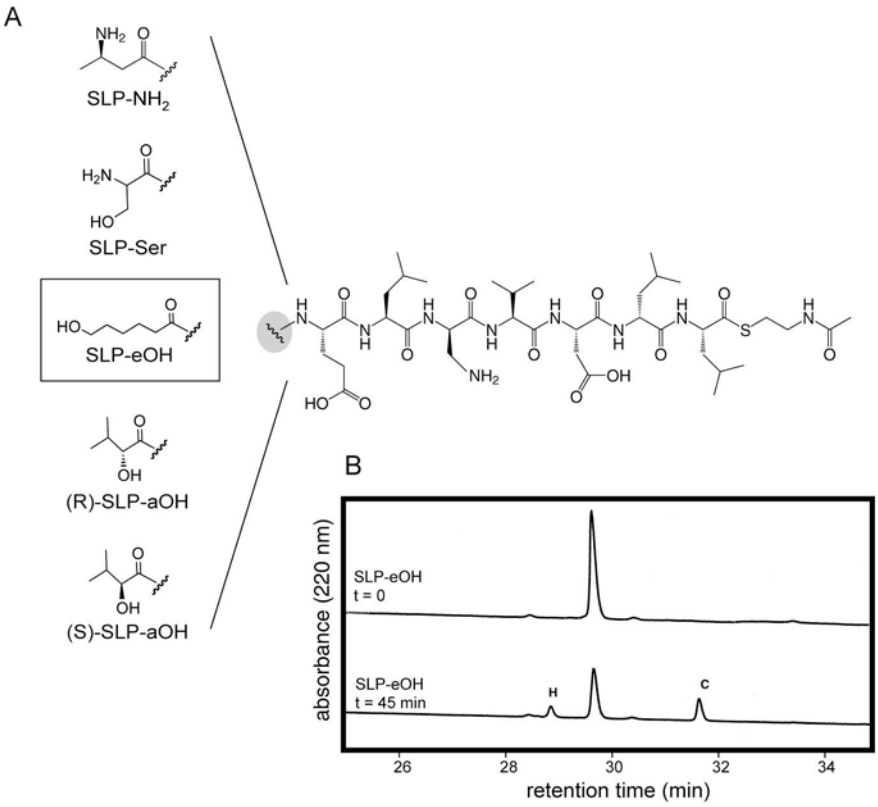


Figure 5.11. Variations to the nucleophilic attacking group of SLP-3. (A) The (R)-3-hydroxybutyric acid group was replaced by those shown. The altered substrate that could be cyclized by Srf TE, SLP-eOH, is boxed. (B) HPLC traces of 2 hr reaction initially containing 2.5 µM SrfTe, 250 µM SLP-εOH, and 25 mM MOPS at pH 7.0 and 24 °C. The peaks corresponding to the cyclic (C) and hydrolyzed (H) products are labeled. The structures of the substrate and products were verified by MALDI-TOF MS. Additionally the cyclic product was characterized by ESI-ion trap mass spectrometry.

5.2.1.6 Mutations of Catalytically Important Residues of Surfactin TE

Six point mutations were individually introduced into Srf TE based on analysis of the crystal structure[96]. All Srf TE mutants were heterologously expressed in *E. coli* and subsequently purified by Ni-NTA affinity chromatography. The yields of all proteins ranged from 2-5 mg per one liter of cell culture and their purity was evaluated by SDS-PAGE analysis (Fig. 5.12). The presumed catalytic serine, Ser80, was changed to cysteine, and the other two residues of the proposed catalytic triad, Asp107 and His207, were each changed to alanine as confirmed by DNA sequencing.

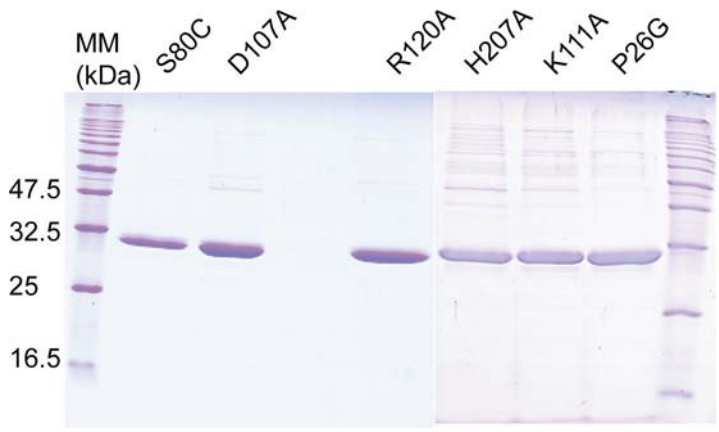
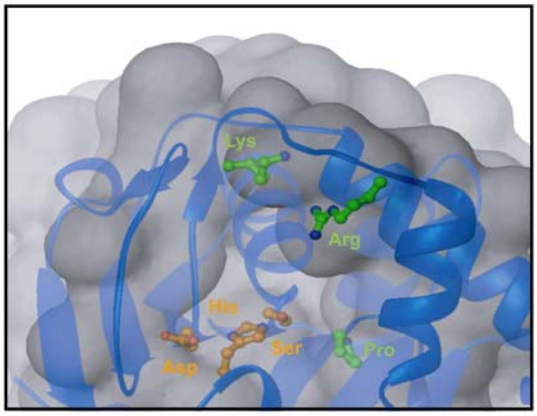


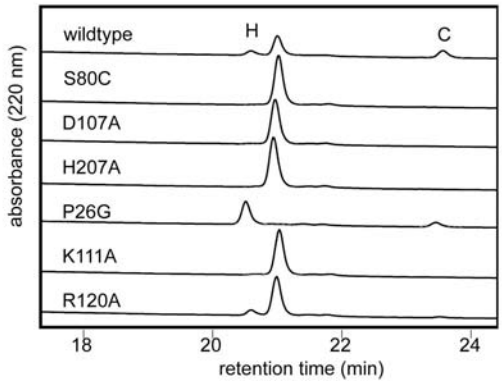
Figure 5.12. SDS-PAGE of purified mutant Srf TE proteins. The calculated mass (28 kDa for all mutants) corresponds to the observed mass. MM = molecular marker.

The catalytic triad mutants were inactive with **SLP-wt** as substrate, indicating that all three residues are critical for enzymatic activity and they can indeed be ascribed to a catalytic triad (slight hydrolysis activity was observed with the D107A mutant in overnight incubations) (Fig. 5.13). Two prominently placed basic residues in the largely hydrophobic active site of Srf TE, Lys111 and Arg120, were each changed to alanine. The K111A mutant was inactive on **SLP-wt** (slight hydrolysis activity was observed in overnight incubations), while the R120A mutant catalyzed the formation of decreased levels of the hydrolyzed, but not cyclized, product (Fig. 5.13B). In comparison with the wild type Srf TE, the R120A mutant showed a 11-fold decrease in k_{cat}/K_M for cyclization (0.078 $mM^{-1}min^{-1}$), and a comparable k_{cat}/K_M for hydrolysis (0.40 $mM^{-1}min^{-1}$) (Fig. 5.13C).

A



B



C

	k_{cat}/K_M for SLP-wt (mM^{-1} min^{-1})		SLP-wt C/H ratio	k_{cat}/K_M for SLP-3 (mM^{-1} min^{-1})		SLP-3 C/H ratio
	cyclization	hydrolysis		cyclization	hydrolysis	
wildtype	0.86	0.35	2.5	2.9	5.4	0.54
P26G	3.7	12	0.2	0.30	11	0.04
S80C	N/D	N/D	N/D	N/D	N/D	N/D
D107A	N/D	N/D	N/D	N/D	N/D	N/D
K111A	N/D	N/D	N/D	N/D	0.61	N/D
R120A	0.078	0.40	0.18	2.1	10	0.17
H207A	N/D	N/D	N/D	N/D	N/D	N/D

Figure 5.13. Characterization of Srf TE mutants. (A) Binding pocket of Srf TE with sites of mutations highlighted. (B) HPLC traces of 2 hr reactions initially containing 2 μM indicated enzyme, 250 μM SLP-wt, and 25 mM MOPS, at pH 7.0 and 24 °C. The peaks corresponding to the cyclic (C) and hydrolyzed (H) products, and substrate (S) are labeled. (C) C/H ratios and k_{cat}/K_M values of cyclization and hydrolysis for each mutant with SLP-wt and SLP-3 (standard deviation ±10%). N/D = activity not detected.

A P26G mutant was made based on the sequence alignment of Srf TE with other TE domains and lipases (data not shown), which are also members of the α,β-hydrolase enzyme family that catalyze only hydrolysis of ester/lactone substrates[131]. A proline conserved among TE domains, Pro26 in Srf TE, is instead conserved as a glycine in lipases, and the position of this residue near those responsible for the formation of the oxyanion hole in the active site (Val27 and Ala81) indicates that it may have an effect on product-determining catalysis of cyclization vs. hydrolysis by the enzyme (Fig. 5.13A). Indeed, reactions of **SLP-wt** with the P26G mutant produced two observations of note. One was a 12-fold change in product ratio to favor hydrolysis, down to a cyclization:hydrolysis ratio of 0.2:1 (Fig. 5.13C). The other was a net increase in catalytic flux to 15.7 $mM^{-1}min^{-1}$, up from 1.2 $mM^{-1}min^{-1}$ in wildtype Srf TE; of

this 13-fold increase in overall flux, 80% of **SLP-wt** underwent hydrolysis. P26G was also able to reopen the macrolactone with a k_{cat}/K_M of 0.21 $mM^{-1}min^{-1}$.

5.2.1.7 Cocrystal Structure of Srf TE with a Dipeptidyl Boronate Inhibitor

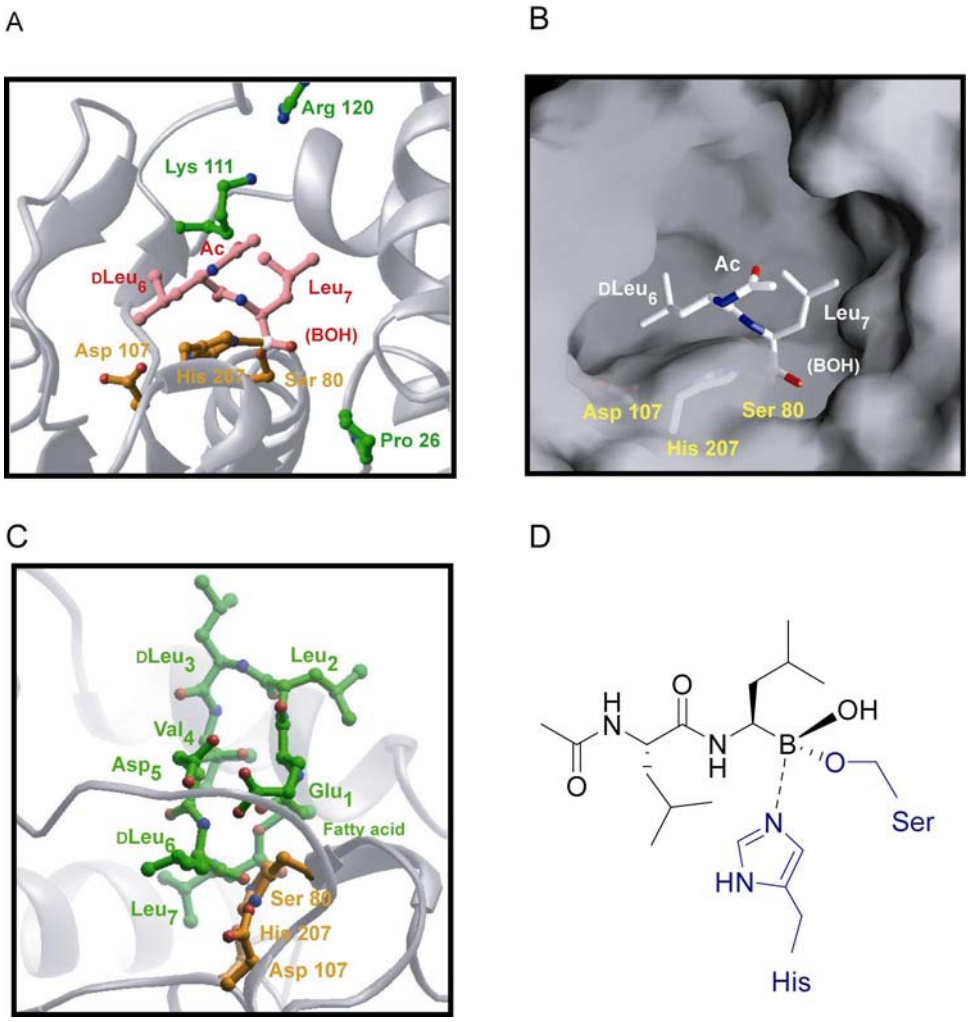


Figure 5.14. Crystal structure of Srf TE bound to a dipeptide boronic acid analogue. (A) Structure of N-acetyl-D-Leu-Leu-$B(OH)_2$ complexed to Srf TE. The dipeptide analogue is shown in pink, the catalytic triad in yellow, and additional active site chains in green. (B) Surface repesentation of the active site in the same orientation as in part A. (C) Model of the cyclizing conformation of the surfactin linear peptide. Superposition of the NMR structure of surfactin onto the structure of N-acetyl-D-Leu-Leu-B(OH) in the active site. (C) Structure of N-acetyl-D-Leu-Leu-B(OH) coordinated by His and Ser.

In order to further characterize the binding mode of surfactin linear peptides, Steve Bruner examined X-ray crystal structures of Srf TE bound to designed inhibitor molecules. A variety of linear peptides containing the surfactin heptapeptide sequence and a C-terminal boronic acid in place of the carboxyl group were synthesized[79]. Boronic acid-based inhibitors are widely used as potent inhibitors of serine proteases, and function by alkylating the catalytic

serine, forming a tetrahedral boronate that closely mimics the tetrahedral intermediate in the reaction pathway. Peptides were soaked into crystals of Srf TE, and the structures solved by molecular replacement using the structure of Srf TE[96]. Peptide electron density was only clearly present for D-Leu_6, Leu_7, and the boronic acid group, irrespective of the length of the peptide used in the soaking experiment. Figure 5.14 shows the structure of the enzyme bound to the dipeptide analog *N*-acetyl-D-Leu-Leu-$B(OH)_2$ (K_i = 50 μM), and is representative of the observed electron density with full-length heptapeptide boronic acid analogs. The structure shows a tetrahedral boron atom bound both to the hydroxyl side chain of Ser80 and the N3 of His270, a bidentate binding mode sometimes observed in serine protease/boronic acid complexes[132, 133]. The C-terminal Leu (corresponding to Leu_7) is bound specifically in a pocket adjacent to the triad, forming contacts with residues Tyr156, Tyr159, and Leu129. The D-Leu residue (corresponding to D-Leu_6) is also well ordered in the density maps, bound in a hydrophobic pocket consisting of Tyr109, Leu187, and Phe 181. The remaining residues of the substrate peptide are not well ordered in the maps.

5.2.2 Loading Peptidyl-CoA onto PCPs: A Novel Approach to Characterize Macrocyclization by TEs

To expand the set of cyclization catalysts also genes encoding peptide cyclases from other NRPS systems like mycosubtilin, and fengycin from *B. subtilis* and syringomycin from *P. syringae* were cloned and overexpressed. Contrary to surfactin and tyrocidine cyclases, no activity was observed for those of mycosubtilin, fengycin and syringomycin cyclases with synthetically made peptidyl-SNAC substrates, indicating a limitation in the chemoenzymatic potential of the latter cyclases. The inability to recognize or bind the SNAC substrates in the active site of the excised peptide cyclase could be affected by the way of presenting the short SNAC mimic to the enzyme. To overcome this limitation the cyclase domain (TE) can be excised with the preceeding cofactor binding PCP-domain as a PCP-TE didomain. Recombinant apo PCP-TE cyclases can then be loaded *in vitro* with chemically synthesized peptidyl-CoA using the 4′-phosphopantetheine (ppan) transferase Sfp. The resulting peptidyl-ppan-PCP-TE holo cyclase carries the covalent cofactor bound substrate in a way that mimics the natural substrate presentation in the NRPS assembly line (Fig. 5.15). A prerequisite of this method is the promiscuity of the *B. subtilis* ppan transferase Sfp for loading chemically synthesized peptidyl-CoA substrates onto apo PCP which was investigated first.

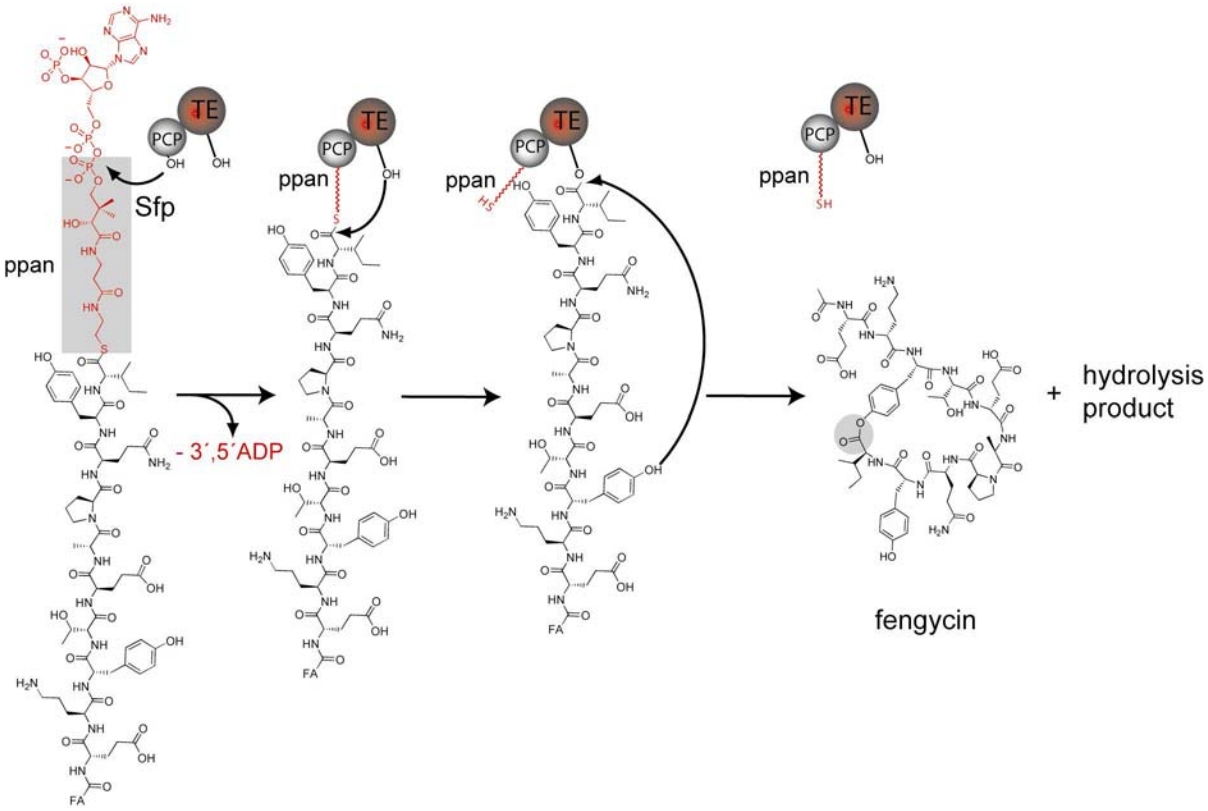


Figure 5.15. Natural substrate interaction. Synthetic fengycin-CoA is loaded onto apo Fen PCP-TE by Sfp to give the holo enzyme. The substrate is transferred onto the TE and subsequently released by cyclization and hydrolysis.

5.2.2.1 Overexpression of the Proteins and Synthesis of Peptidyl-CoAs

To investigate loading of peptidyl-CoA substrates with Sfp onto apo proteins, gene fragments encoding a PCP domain from module 7 of the surfactin synthetase in its apo form as well as an apo PCP-TE didomain from the fengycin synthetase and a TE active site Ser to Ala mutant of this didomain were cloned and overexpressed in *E. coli*. The proteins were purified by Ni-NTA affinity chromatography and the purity was evaluated by SDS-PAGE analysis (Fig. 5.16).

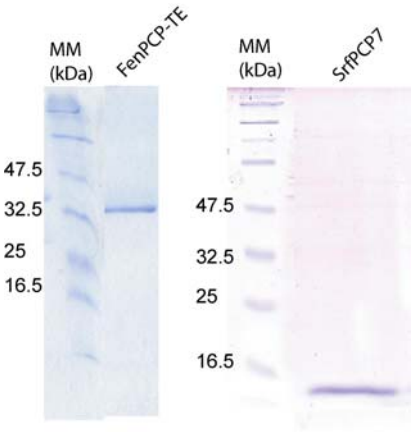


Figure 5.16. SDS-PAGE of purified Fen PCP-TE and Srf PCP7. The calculated mass of wt Fen PCP-TE (37 kDa) and Srf PCP7 (10 kDa) correspond to the observed mass. The Ser to Ala Fen PCP-TE mutant was obtained in the same yield and purity. MM = molecular marker.

Three different N-acyl-peptidyl-CoA substrates derived from surfactin (**SLP**), fengycin (**FLP**) and mycosubtilin (**MLP**) were synthesized. The structures of all peptidyl-CoA substrates are shown in figure 5.17. In comparison with all three wild type sequences the long fatty acid moiety was replaced with shorter acyl-chain mimics to increase solubility. In the surfactin sequence Asp at position 5 was replaced with 2,3-diaminopropionic acid as described before[79] and in the mycosubtilin sequence Asn at position 7 was replaced by Ala for synthetic reasons. The subsequent coupling with CoA yielded 50 – 60 % product[57]. The peptidyl-CoAs were characterized by mass analysis as shown in Table 4.4.

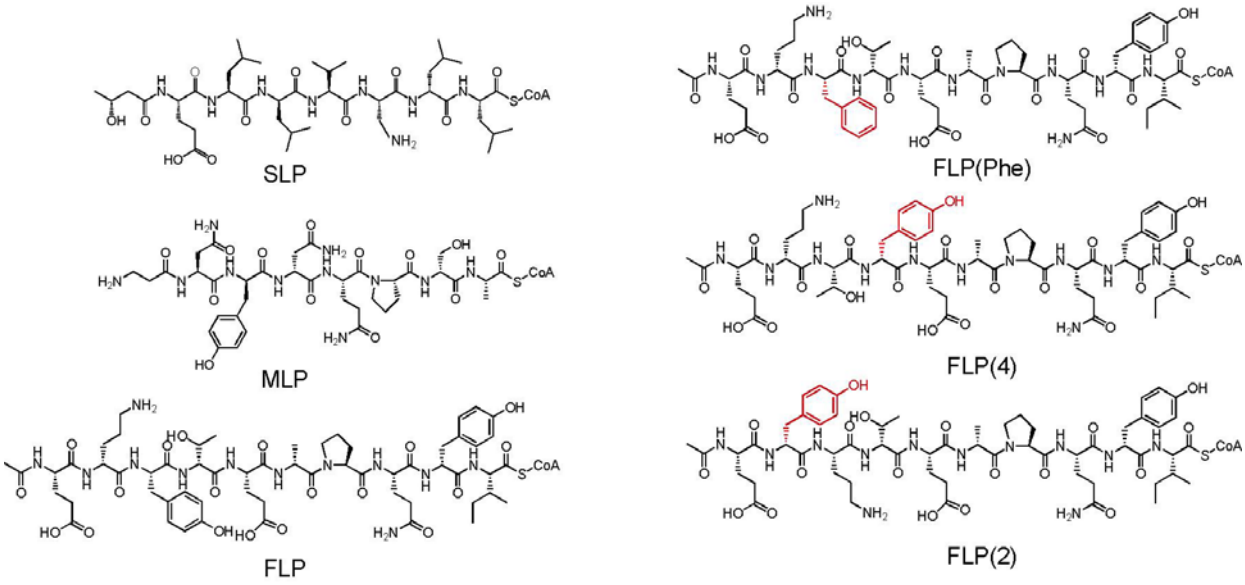


Figure 5.17. Structures of all CoA substrates. Residues important for regioselectivity experiments are highlighted in red. For amino acid sequences, please see also the appendix section.

5.2.2.2 Loading Peptidyl-CoAs onto PCP

Incubation of apo PCP with the three peptidyl-CoA substrates and Sfp revealed a substantial conversion into peptidyl holo PCPs (Fig. 5.18). The K_M values of the loading reactions were higher for **SLP-CoA** (17.2 µM), **FLP-CoA** (18.1 µM) and **MLP-CoA** (11.3 µM) than that of CoASH alone (0.7 µM) described previously[55]. Also a strong reduction of the k_{cat} values for the loading of the three peptidyl-CoAs was observed (**SLP-CoA**: 0.92 min^{-1}, **FLP-CoA**: 0.94 min^{-1}, **MLP-CoA**: 0.69 min^{-1}) (Fig. 5.19) when compared with CoASH alone (102 min^{-1})[55]. Sfp is promiscuous enough to tolerate dramatically larger derivatives of CoA with a reduced kinetic efficiency but preparatively useful. It is now possible to convert an apo PCP into peptidyl holo PCP and study its interaction with TE domains *in cis* or *in trans*. Products were identified by MS (Table 4.5).

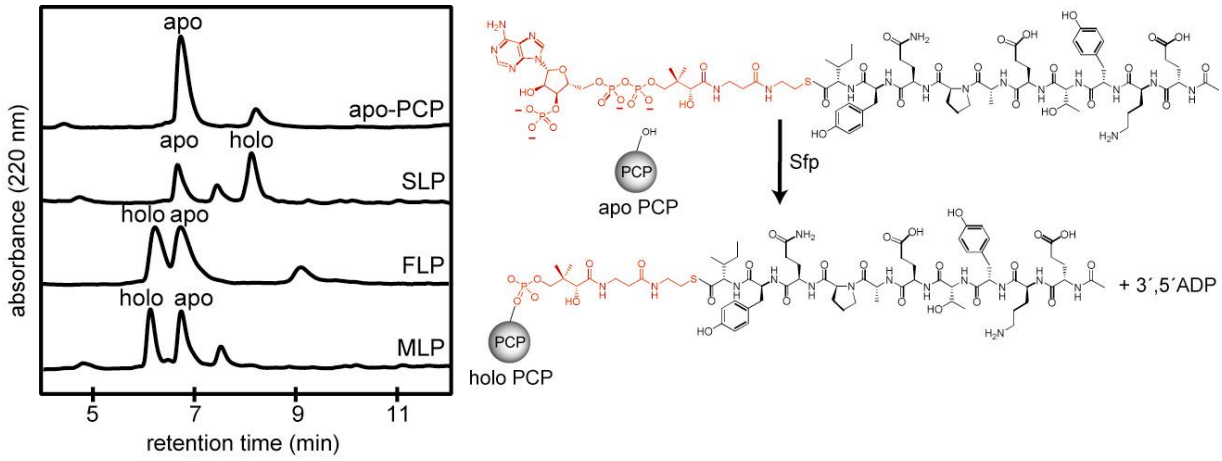


Figure 5.18. Loading peptidyl-CoA on PCP and interaction of peptidyl-PCP with the TE domain. HPLC traces of 14 min Sfp (2µM) loading reactions of apo PCP (50 µM) with surfactin-CoA (SLP), fengycin-CoA (FLP) and mycosubtilin-CoA (MLP) in 25 mM Hepes and 50 mM NaCl, pH 7.0. The peaks corresponding to the apo PCP (apo) and peptidyl-PCP (holo) are labeled. Sfp catalyses the nucleophilic attack of the invariant serine residue of apo PCP onto the phosphodiester bond of fengycin-CoA (FLP) to give peptidyl-holo PCP.

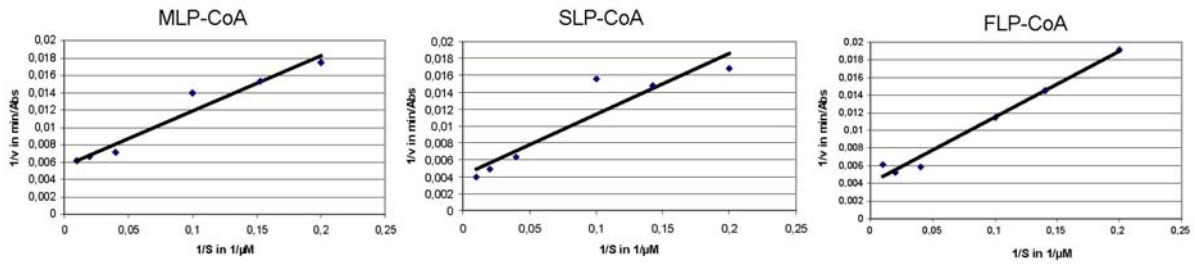


Figure 5.19. Lineweaver-Burk plots of MLP-CoA, SLP-CoA and FLP-CoA.

5.2.2.3 Autonomous Cyclization Activity of FenPCP-TE

Fengycin is a potent antifungal branched cyclic lipopetide which inhibits filamentous fungi but does not act on yeast and bacteria[15]. In order to understand principles of cyclization, the gene fragment coding for fengycin TE was cloned and overexpressed as an apo PCP-TE didomain. The PCP-TE didomain should first ensure a proper folding of the TE and secondly allow the loading of **FLP-CoA** substrate onto apo PCP to mimic the natural substrate for the TE (Fig. 5.15). Pure PCP-TE was first incubated with a linear fengycin SNAC substrate for several hours but did not show any cyclization or hydrolysis (data not shown). In contrast, using the new **FLP-CoA** substrate, PCP-TE and Sfp, cyclization and hydrolysis were observed while a PCP-TE mutant enzyme with a Ser to Ala mutation in the TE active site did not show any activity under the same conditions (Fig. 5.20, Table 4.5). Also a control reaction

with substrate and apo PCP-TE but without Sfp treatment did not reveal activity indicating that the covalent loading of the lipopeptidyl substrate onto the PCP is required for both cyclization and hydrolysis reactions. The observed cyclization to hydrolysis ratio was 2:1 validating that the TE in this PCP-TE didomain retains cyclase activity. A time course showed maximum cyclization after approximately one hour. With increasing time the amount of cyclic product decreased while hydrolysed product increased. Incubation of isolated cyclic peptide with PCP-TE showed progressive hydrolysis to linear lipopeptide. Thus the fengycin TE can like Srf TE also run backwards to a linear peptidyl-O-TE that can be captured by water.

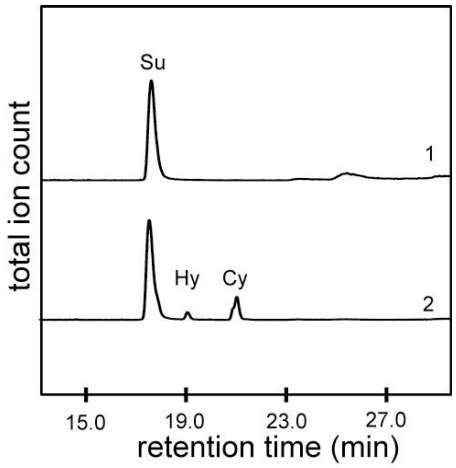


Figure 5.20. Results of the cyclization reaction with Fen PCP-TE. Peaks corresponding to the substrate (Su), hydrolysed product (Hy) and cyclic product (Cy) are labeled. LC-MS traces (TIC) of 30 min reactions containing 60 µM fengycin PCP-TE, 60 µM FLP-CoA and 5 µM Sfp in 25 mM Hepes and 50 mM NaCl, pH 7.0. Trace 1 shows the reaction with FLP-CoA and a mutated enzyme (Ser to Ala in the active site of the TE). Trace 2 shows the same reaction with the wild type enzyme.

5.2.2.4 Regioselectivity of Cyclization

The identity of the cyclic **FLP** peptidolactone was confirmed by MS-MS sequencing (Table 5.2). Masses of 2 characteristic cyclic fragments ((a)fragmentation of the peptide bond at the proposed cyclization position and (b) fragmentation of the adjacent NH-CH bond) were calculated for all possible nucleophiles and compared with the observed fragments.

Table 5.2. (a) Fragmentation of the peptide bond at the proposed cyclization position and (b) fragmentation of the adjacent NH-CH bond. *The observed mass belongs to a peptide fragment which carries Orn but is cyclized via Tyr. If cyclization would occur via Orn, no cyclic fragment would have been obtained with the apparent mass of Tyr cyclization. N/D = not detected.

Compound	Species	observed mass of fragments (calculated mass) (Da)		
		cyclization via Orn	Cyclization via Tyr	cyclization via Thr
FLP cycle (a)	$[M+H]^+$	1080.538* (1080.537)	966.459 (966.457)	N/D (803.394)
(b)		N/D (1063.510)	949.443 (949.431)	N/D (786.367)
FLP(2) (a)	$[M+H]^+$	N/D (917.473)	1080.542 (1080.537)	N/D (803.394)
cycle (b)		N/D (900.447)	1063.502 (1063.510)	N/D (786.367)

The results show that cyclization occured regioselectively via nucleophilic attack of Tyr at position 3, discriminating two other nucleophiles at position 2 (Orn) and 4 (Thr). This high selectivity raised the question if the TE specifically recognizes position 3 or if the identity of the amino acid residue (Tyr) is important. To answer this question three new fengycin-CoA substrate analogs were synthesized (Fig 5.17). The first carried a Phe residue instead of Tyr at position 3 (**FLP-Phe**). The enzymatic reaction with PCP-TE showed only hydrolysis but no cyclization as judged by LC-MS, indicating that none of the other two nucleophiles could replace tyrosine. In additional reactions substrates with an exchange of positions Tyr 3 and Orn 2 (**FLP-2**) as well as Tyr 3 and Thr 4 (**FLP-4**) were used. The relocation of Tyr at position 4 revealed only hydrolysis but Tyr relocation at position 2 showed hydrolysis and cyclization (Fig. 5.21, Table 4.5). The identity of Tyr 2 as the nucleophile for this cyclization was again confirmed by MS-MS sequencing (Table 5.2). These results show that fengycin TE specifically recognizes Tyr as the nucleophile in position 3 and 2 and an increase of the peptidolactone ring size by one residue (from 29 to 32 atoms) is possible.

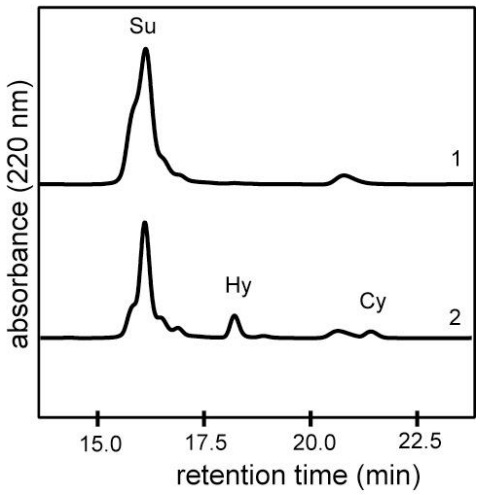


Figure 5.21. Cyclization of FLP-2. Peaks corresponding to the substrate (Su), hydrolysed product (Hy) and cyclic product (Cy) are labeled. LC-MS traces of 30 min reactions containing fengycin PCP-TE and a fengycin-CoA derivative (FLP-2) with Tyr at position 2 and Orn at postion 3 without Sfp (Trace 1) and in presence of Sfp (Trace 2).

5.2.3 Peptidyl-Thiophenols as New Substrates for Peptide Cyclases

The previous PCP-TE approach is limited to single turnover reactions, since after product release the cofactor ppan remains attached to the PCP-TE didomain cyclase, which blocks further Sfp catalyzed transfer of additional peptidyl-CoAs onto ppan-PCP. In order to force multiple rather than single turnover cycles a new strategy to expand the utility of this method was developed. This strategy is based on a thioester exchange reaction between the free ppan-PCP thiol and a soluble thiophenol-peptide substrate which should ensure chemical reloading of substrate onto the ppan-PCP-TE didomain[134, 135].

5.2.3.1 Cloning and Overexpression of the Proteins

The mycosubtilin (ATCC 6633) and syringomycin (DSMZ B-301D) gene fragments encoding Myc PCP-TE and Syr TE were amplified from chromosomal DNA and cloned into a pET37b expression plasmid. The proteins were heterologously expressed in *E. coli* and subsequently purified by Ni-NTA affinity chromatography. The yields of the proteins ranged from 2-5 mg per one liter of cell culture and their purity was evaluated by SDS-PAGE analysis (Fig. 5.22).

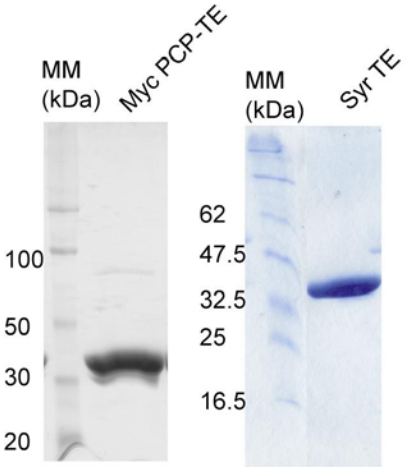


Figure 5.22. SDS-PAGE of purified Myc PCP-TE and Syr TE. The calculated mass of Myc PCP-TE (37 kDa) and Syr TE (32 kDa) correspond to the observed mass. MM = molecular marker.

5.2.3.2 Activity Based Enzyme Acylation with Thiophenol Substrates

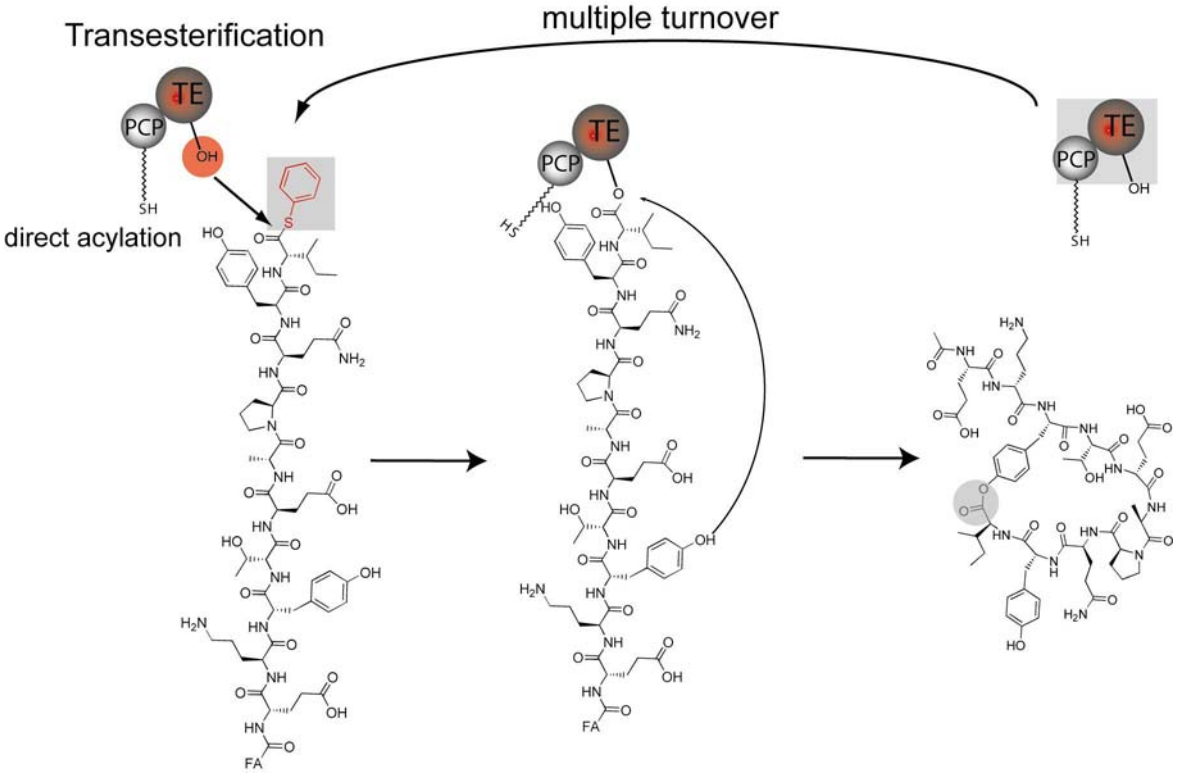


Figure 5.23. Activity based enzyme acyclation. The active site serine of the fengycin peptide cyclase (TE) is selectively acylated by a reactive peptidyl-thiophenol substrate. The acyl-enzyme intermediate is then captured by an intramolecular nucleophilic attack of tyrosine$_3$ to give the cyclic peptidolactone fengycin.

The initial aim was to achieve a thioester exchange reaction between the ppan free thiol group of fengycin bidomainal fragment ppan-PCP-TE[59] and synthetic fengycin-thiophenol (**FLP-tp**). Surprisingly, control reactions of **FLP-tp** with apo PCP-TE, lacking the free thiol group, revealed hydrolysis and cyclization activity, which was not observed with SNAC substrates (Fig. 5.24 and Fig. 5.25A). A control reaction with a Ser to Ala mutant in the fengycin cyclase active site abolished activity. Identities of all substrates (Table 4.6) and products (Table 4.7) were verified by HPLC-MS. In turn, no increased product turnover was observed with holo PCP-TE, indicating that the thiophenol substrate directly acylates the TE active site serine efficiently (Fig. 5.23). Kinetic data for the cyclization of **FLP-tp** (k_{cat} = 0.33 min^{-1} and K_M = 461 µM, Cy/Hy ratio = 0.85) show saturation kinetics and suggest sufficient catalytic efficiency for preparative scale reactions.

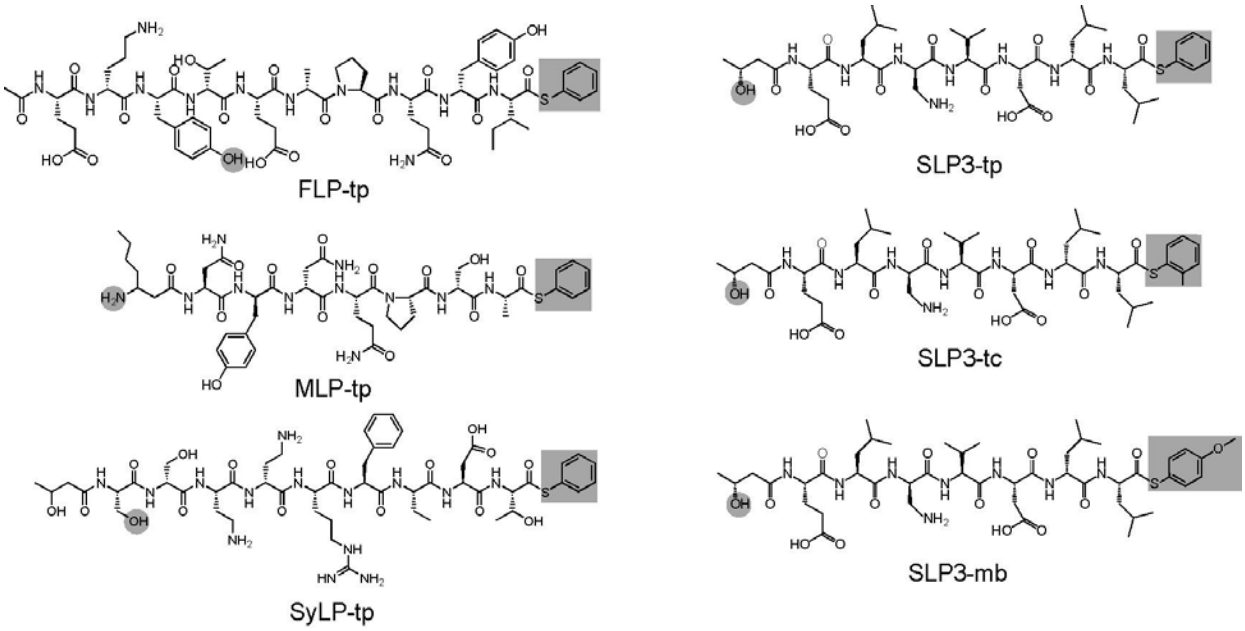


Figure 5.24. Structures of peptidyl-aryl thioester substrates. Leaving groups are indicated by squares and nucleophiles for enzymatic intramolecular cyclization are indicated by circles. For amino acid sequences, please see also the appendix section.

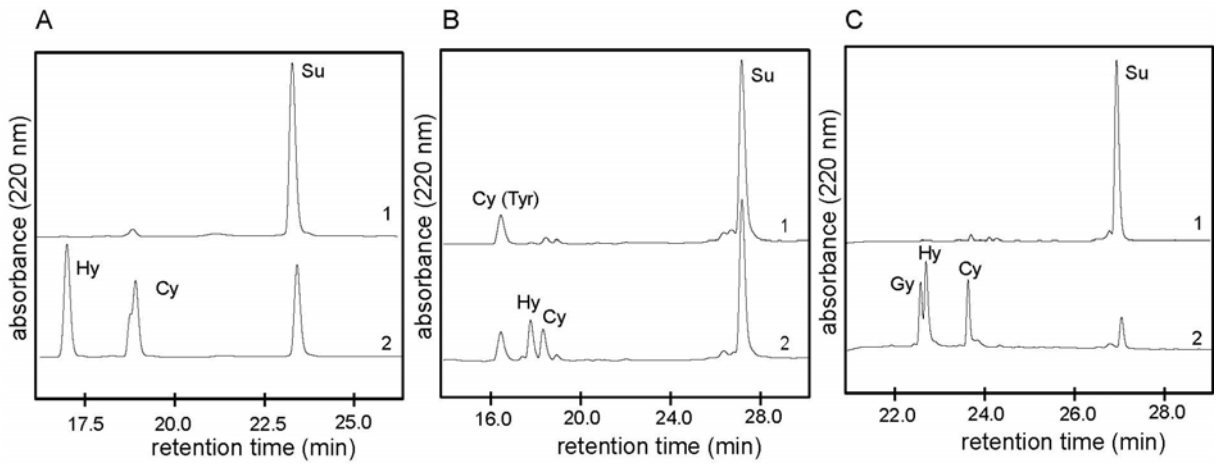


Figure 5.25. HPLC traces of peptide cyclases (SNAC inactive) incubated with peptidyl-thiophenol substrates. (A). Fengycin cyclase incubated with FLP-tp for 1.5 h (Trace 2). Trace 1 shows incubation of substrate with an active site Ser to Ala mutated cyclase. (B). Mycosubtilin cyclase incubated with MLP-tp for 5 min (Trace 2). Trace one shows substrate incubation without enzyme. Uncatalyzed cyclization via tyrosine$_3$ [Cy(Tyr)] is observed. (C) Syringomycin cyclase incubated with SyLP-tp for 30 min (Trace 2). Trace 1 shows substrate incubation without enzyme. Su = substrate, Hy = hydrolysed product, Cy = cyclized product, Gy = glycerol adduct.

5.2.3.3 Tuning the Reactivity of the Thiophenol Leaving Group

To evaluate the general utility of thiophenol leaving groups for other peptide cyclases, surfactin cyclase (Srf TE), which was previously characterized with SNAC substrates[79],

was incubated with surfactin-thiophenol (**SLP3-tp**) (Fig. 5.24). HPLC analysis revealed enzyme catalysed cyclization and hydrolysis activity with a K_M of 126 µM and a k_{cat} of 5.66 min^{-1} and 11.33 min^{-1} for cyclization and hydrolysis, respectively (Fig. 5.26A). Comparison of the catalytic cyclization efficiency of **SLP3-tp** (k_{cat}/K_M 44.9 mM^{-1} min^{-1}) with **SLP3** (2.9 mM^{-1} min^{-1})[79] revealed a 15 fold increased activity for the thiophenol substrate, with a Cy/Hy ratio of 0.78 for thiophenol and 0.54 for SNAC, consistent with a common acyl-enzyme intermediate for both substrates whose breakdown may now have become rate-determining with **SLP3-tp**. However, the better peptide activation by thiophenol seems to facilitate non-enzymatic catalysed cyclization as observed in the surfactin control reaction without enzyme (Fig. 5.26A). MS-MS sequencing of the cyclic species obtained in the absence of Srf TE revealed uncatalyzed cyclization via a diaminopropionic acid (Dap) residue at position three which was introduced due to solubility reasons and is not present in the natural sequence (Table 5.3). Such non-enzymatic peptide cyclization is an intrinsic property of some peptide sequences as has been reported before in the tyrocidine series[92]. Note that the Srf TE directs cyclization to the β-OH of the fatty acyl chain rather than the Dap amide formation seen nonenzymatically.

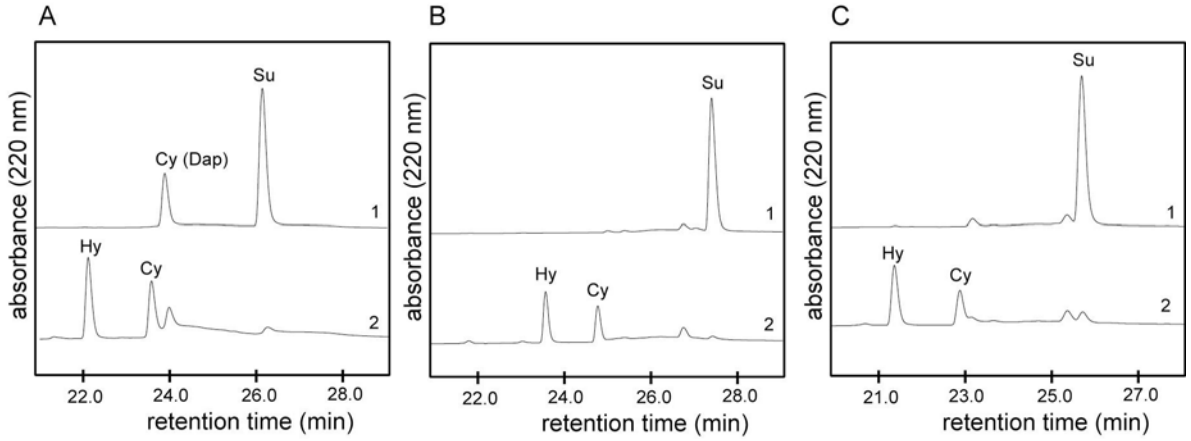


Figure 5.26. HPLC traces of surfactin cyclase (SNAC active) incubated with aromatic thioester substrates. (A). Incubation of surfactin cyclase with SLP-tp for 10 min (Trace 2). Trace one shows substrate incubation without enzyme. Uncatalyzed cyclization via diaminopropionic $acid_3$ [Cy(Dap)] is observed. (B) Incubation of surfactin cyclase with SLP-tc for 10 min (Trace 2). Trace one shows substrate incubation without enzyme. (C) Incubation of surfactin cyclase with SLP-mb for 10 min (Trace 2). Trace one shows substrate incubation without enzyme. Su = substrate, Hy = hydrolysed product, Cy = cyclized product.

Table 5.3. (a)Fragmentation of the peptide bond at the proposed cyclization position and (b) fragmentation of the adjacent NH-CH bond. N/D = not detected.

Compound	species	observed mass of fragments (calculated mass) (Da)	
		Cyclization via Dap 3	
SLP3 cycle (a)	$[M+H]^+$	527.316 (527.319)	
(b)		N/D (510.293)	
		Cyclization via Tyr 2	cyclization via Ser 6
MLP cycle (a)	$[M+H]^+$	704.278 (704.300)	N/D (202.083)
(b)		687.254 (687.274)	N/D (185.056)

To suppress the limitation of unspecific nonenzymatic cyclization of thiophenol-peptides via side chains, the reactivity of the leaving group was decreased by changing ligands. Substituents with a positive inductive or mesomeric effect such as methyl and methoxy groups were introduced into the aromatic thiophenol-ring at ortho, meta and para positions. All six new aryl thiol leaving groups were coupled to the free carboxylic acid of the surfactin peptide and subsequently incubated with the surfactin cyclase. Replacement of thiophenol with o-thiocresol (**SLP3-tc**) and p-methoxybenzenethiol (**SLP3-mb**) gave best results for suppression of unspecific background cyclization, while enzyme catalysed substrate cyclization and hydrolysis remained constant (Table 4.8, Fig. 5.24, Fig. 5.26B and C). The attachment of ligands onto the aromatic ring can therefore adjust the desired activity for any peptidyl thioester substrate. Ortho and para positions seem to be preferred, probably due to their impact in destabilization of the thiophenolate anion.

5.2.3.4 Characterization of New Peptide Cyclases

The observed catalytic activity of peptidyl thiophenol substrates with fengycin and surfactin peptide cyclases encouraged to investigate other peptide cyclases with this new technology. Therefore mycosubtilin PCP-TE and syringomycin TE were investigated because of their interesting structures and bioactivities[23, 136]. Both enzymes failed to show any activity with peptidyl-SNAC substrates (data not shown). Therefore thiophenol was attached to peptide sequences derived from mycosubtilin (**MLP-tp**) and syringomycin (**SyLP-tp**) (Fig. 5.24). In comparison to the syringomycin wild type sequence[81] the peptide analogue used here differs for synthetic reasons: 4-Cl-threonine at position 9 was replaced by threonine,

dehydrobutyrine by aminobutyric acid, β-hydroxy aspartate by aspartate and the long fatty acid moiety was replaced with a shorter β-hydroxy fatty acyl chain for solubility reasons. In contrast to syringomycin where cyclization occurs via a serine side chain, the fatty acid chain in the mycosubtilin sequence seems to be important since macrolactamization occurs via the β-amino fatty acid moiety (Fig. 5.24). The natural fatty acid chain was replaced by a C_7-β-aminoacyl chain, provided by chemical synthesis[137]. Incubation of mycosubtilin peptide cyclase with **MLP-tp** revealed enzyme catalysed hydrolysis and cyclization (Fig. 5.25B). Cyclization occured with a k_{cat} of 2.16 min^{-1} and a K_M of 3991 μM and a Cy/Hy ratio of 0.85. The high K_M value could be due to missing enzymatic recognition elements of the artificial cofactor thiophenol. Similar to the surfactin derived autocatalytic peptidyl-thiophenol self-cyclization of **MLP-tp** was also observed in a control reaction without enzyme. The identity of this new species was investigated by MS-MS sequencing and identified to be cyclized via tyrosine at position two (Table 5.3), again a regiospecificity distinct from that imposed during enzymatic cyclization. Incubation of syringomycin peptide cyclase with **SyLP-tp** revealed both cyclized and hydrolysed products (Fig. 5.25C). Also a glycerol adduct was observed due to the fact that 5 % glycerol were necessary to stabilize syringomycin peptide cyclase. Saturation behaviour was observed and kcat and KM of cyclization determined (k_{cat} = 0.32 min^{-1} and K_M = 28 μM, Cy/Hy ratio = 0.65). Previous data shows that in vivo cyclization of syringomycin occurs via L-serine at position one[81]. Interestingly, an adjacent D-serine residue at position two is not used, raising the question if a change in stereochemistry at position one and two could alter product outcome. Therefore two new thiophenol substrates were made with D-serine at position one and L-serine at position two as well as L-serine at position one and two. Incubation of these substrates with syringomycin cyclase revealed only hydrolysis in both cases emphasizing the importance of Ser1,2 stereochemistry for enzymatic cyclization (data not shown).

5.2.4 Directed Protein Evolution to Provide "Custom Made" Peptide Cyclases

Peptide cyclases presented here catalyse the chemo-, regio- and stereoselective cyclization of linear peptide chains of their dedicated substrates. The relaxed enzymatic specificity in defined parts of the peptide backbone as seen for surfactin TE and tyrocidine TE[90] make these enzymes valuable tools for the cyclization of related compounds.

However, if cyclization of a structurally unrelated linear compound is desired only synthetic chemical cyclization techniques with all their disadvantages can be applied so far. Therefore, in order to provide "custom made" catalytic tools for enzymatic peptide cyclization of nearly

any desired compound, *in vitro* protein evolution is the method of choice. By introducing random mutations into the protein sequence, mutants with altered activities should be generated and those with improved activity would be identified by a screening method. Subsequent rounds of random mutagenesis with an identified hit could further improve catalytic activity.

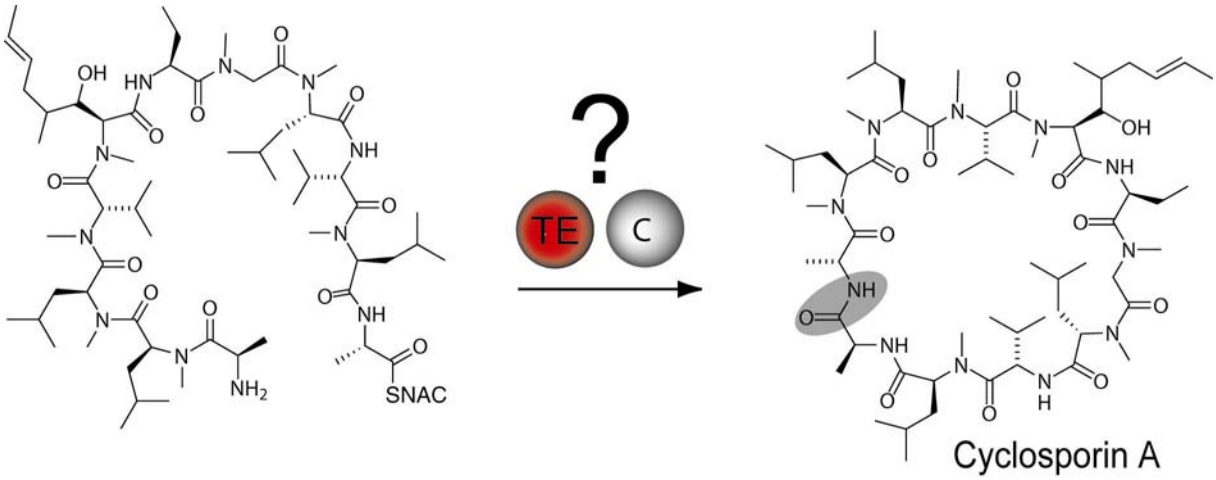


Figure 5.27. Task: Cyclization of cyclosporin SNAC either by cognate C-domain or engineered TE domain.

The objective of the project presented in this chapter was to provide an enzymatic catalyst for the cyclization of linear cyclosporin, which would allow to create libraries of new cyclosporin variants starting from linear compounds. Two different approaches have been tried (Fig. 5.27):

1. Cloning and overexpression of the putative natural cyclase, the cyclosporin C-domain.
2. Protein evolution of Tyc TE.

5.2.4.1 Cloning and Overexpression of the Cyclosporin C-Domain

The cyclosporin synthetase gene fragment encoding the C-terminal C-domain was amplified from chromosomal DNA and cloned into a pBAD202/D-TOPO expression plasmid. This plasmid facilitated C-terminal fusion of a His_6-tag and N-terminal fusion of a thioredoxin domain onto the recombinant protein. Thioredoxin is a 12 kDa protein which helps to increase the solubility of the recombinant protein. Induction of the araBAD promotor was facilitated by adding arabinose in concentrations ranging from 0.01% to 0.0005%. Cultures were grown at 22°C. Figure 5.28 shows a SDS-PAGE of a cyclosporin C-domain test-expression. Induction with 0.01% arabinose clearly reveals the formation of a new protein band with an observed weight of about 75 kDa (calculated: 74 kDa). Uninduced growth overnight does not lead to any protein production of the same size. After Ni-NTA purification only very little

protein with the expected size was eluted together with impurities. Mass spectrometry of this protein fraction failed so far so that His_6-tag antibodies will help to further clarify the identity of this species. The insoluble fraction shows a big protein band of the expected C-domain size which indicates that almost all of the desired protein was insoluble and therefore not suitable for cyclization studies.

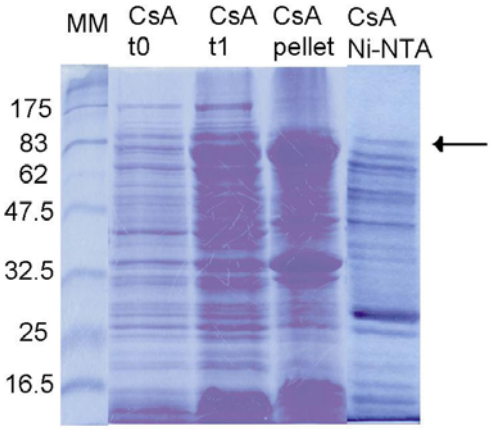


Figure 5.28. SDS-PAGE of cyclosporin C-domain test-expression. In comparison with the uninduced sample (t0), the induced (t1) sample as well as the pellet show a new protein band with the size of the C-domain (74 kDa) indicated by an arrow. The fraction after elution from a Ni-NTA column (100 mM imidazole) does not show specific binding of an individual protein. MM = molecular marker.

5.2.4.2 Synthesis of Cyclosporin SNAC

The chemical synthesis of linear cyclosporin starting from amino acid building blocks is problematic due to the formation of many N-methylated peptide bonds[35]. Therefore a different strategy starting with a ring opening of the cyclic compound was selected here (Fig. 5.29). The ring opening reaction of cyclosporin was carried out according to a previous protocol[122].

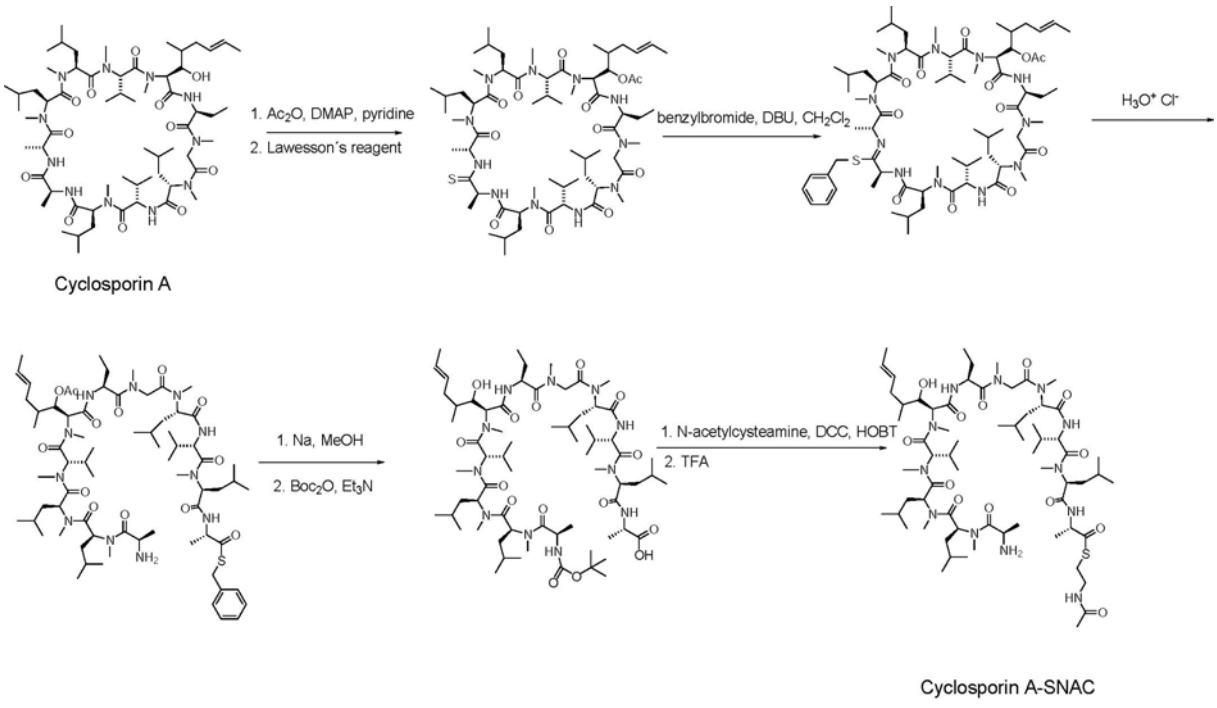


Figure 5.29. Synthesis of linear cyclosporin A-SNAC.

In the first step of synthesis cylosporin A was converted into cyclosporin A acetate by adding acetanhydride to protect the free hydroxyl group (60 % yield). The acetate was treated with Lawesson´s reagent at 130 °C for 30 min. A mixture of only three products was observed due to steric repulsion of the folded peptide. Acetyl cyclosporin 4,7-bis(thioamidate), acetyl cyclosporin 7-thioamidate and acetyl cyclosporin 4-thioamidate were isolated and used without purification for the next reaction step. The thioamidate mixture was incubated with benzylbromide for 2 h at room temperature. Extensive column chromatography on silica gel yielded 20% of the desired acetylcyclosporin A 7-(benzyl thioamidate) as identified by NMR and MALDI-MS. Subsequent hydrolysis of the benzyl thioamidate formed 98 % acetylcyclosporin A carboxylic acid S-benzylester. Cleavage of the acetyl-group, N-terminal Boc protection as well as C-terminal SNAC synthesis were carried out according to standard protocols. Pure cyclosporin A SNAC (**CLP**) was obtained after reversed phase HPLC purification with an observed mass ($M+H^+$) of 1321.81 Da (theoretical 1321.80). The overall yield was about 10 %.

5.2.4.3 Cyclization of Cyclosporin A SNAC by TycTE?

Alternatively to the C-domain derived from the cyclosporin biosynthesis cluster, Tyc TE represents a promising enzyme for the cyclization of **CLP** since the structures of tyrocidine A and cylosporin A feature striking similarities. Both have similar ring sizes and it is well known, that Tyc TE also tolerates larger rings for cyclization[90]. Moreover it is known, that Tyc TE is very selective for the stereochemistry at the cyclization position which is in both cases the same (D,L). A major difference of both compounds is the identity of the residue at position one which is D-Ala in case of cyclosporin A and D-Phe in case of tyrocidine. Previous studies revealed that Tyc TE is not tolerant for a change of D-Phe into D-Ala[90] (Fig. 5.30).

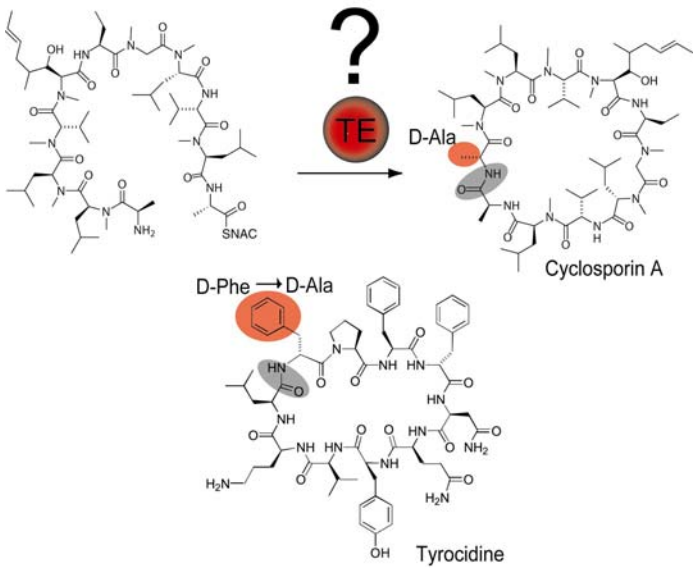


Figure 5.30. Can Tyc TE cyclize linear cyclosporin SNAC? Comparison of cyclosporin and tyrocidine peptides. Residues at position 1 are indicated by red circles.

Incubation of Tyc TE with **CLP** revealed complete hydrolysis after 30 min but no cyclization (Fig. 5.31). A control reaction without enzyme revealed no hydrolysis indicating that the substrate was stable under assay conditions. The identity of the hydrolysed species was confirmed by mass spectrometry ($M+H^+$) (1220.96 Da observed; 1220,80 theoretical).

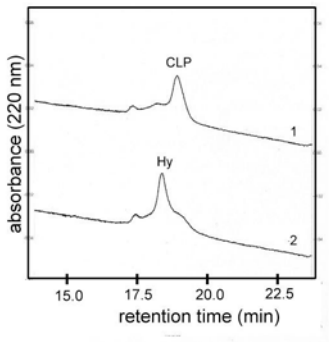


Figure 5.31. HPLC traces of tyrocidine cyclase incubated with CLP. Peaks corresponding to the substrate (CLP) and hydrolysed product (Hy) are labeled. Trace one shows 1 hr substrate incubation without enzyme and trace two with 5 µM tycTE.

5.2.4.4 Directed Protein Evolution of Tyc TE

The previous result indicated that Tyc TE can recognize **CLP** to form an acyl-enzyme intermediate which is released only by hydrolysis probably due to D-Ala at position one. The initial aim was therefore to identify a new species of Tyc TE by directed protein evolution which is capable of cyclizing a tyrocidine A SNAC with D-Ala (**TLP-Ala**) instead of the D-Phe at position one. A new species could then be further refined by protein evolution to finally cyclize the desired substrate **CLP**.

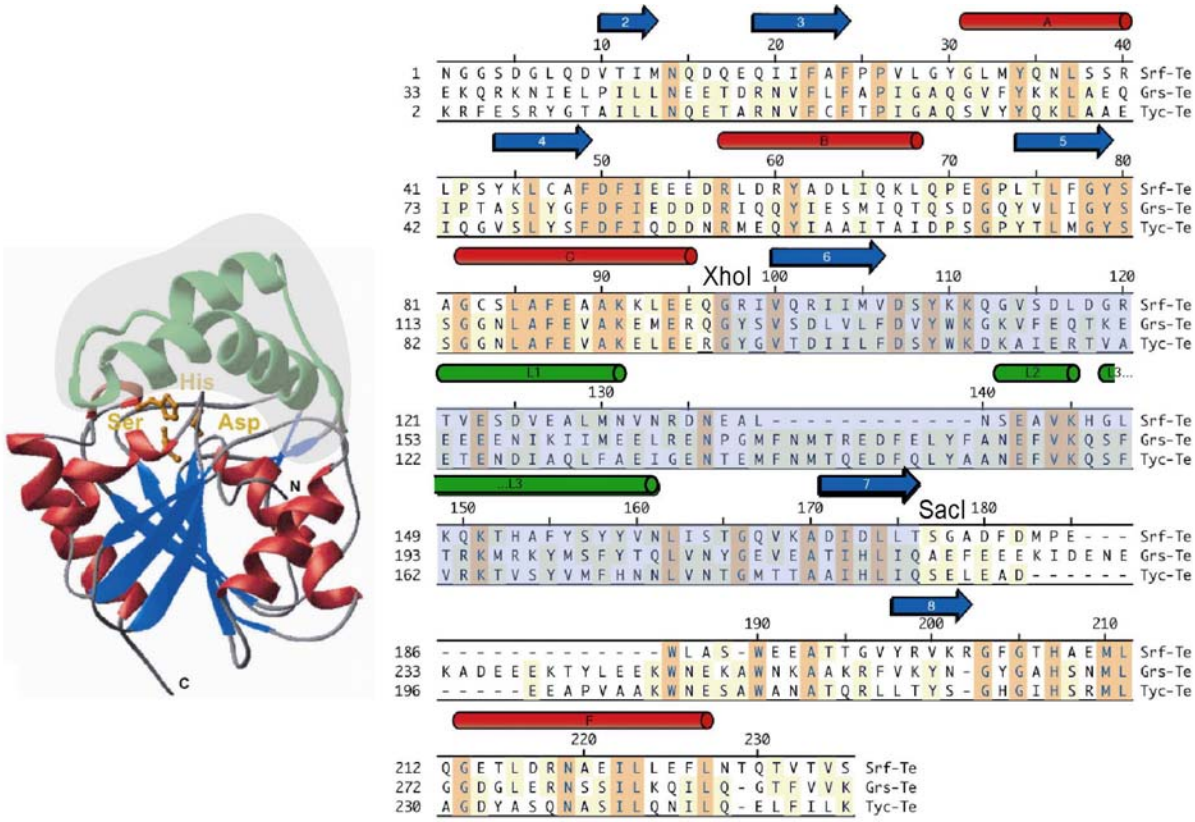


Figure 5.32. Selected sequence for random mutagenesis. The Lid region of the Srf TE crystal structure (grey circle) is highlighted by grey shade in the sequence alignment of Tyc TE, Grs TE and Srf TE.

In the approach selected here, random mutations were introduced by error prone PCR only into a defined part of the Tyc TE which is thought to be important for substrate recognition. Figure 5.32 shows a sequence alignment between Srf TE, Grs TE and Tyc TE together with the crystal structure of Srf TE. By comparison of the structure with the alignment one can see that the least conserved region of all three sequences is located in the lid region of the enzyme indicating that this region is likely to comprise the highest flexibility for substrate recognition. This observation is further supported by studies of Reetz and co-workers[138] who found mutations improving the stereoselectivity of a structurally related lipases to be located in this lid region. Previous investigations on the lid regions of hepatic and lipoprotein lipases supported their relevance for substrate interaction[139]. Therefore in order to minimize the size of an initial library, this 93 amino acid comprising region of Tyc TE was selected for random mutagenesis (Fig. 5.32).

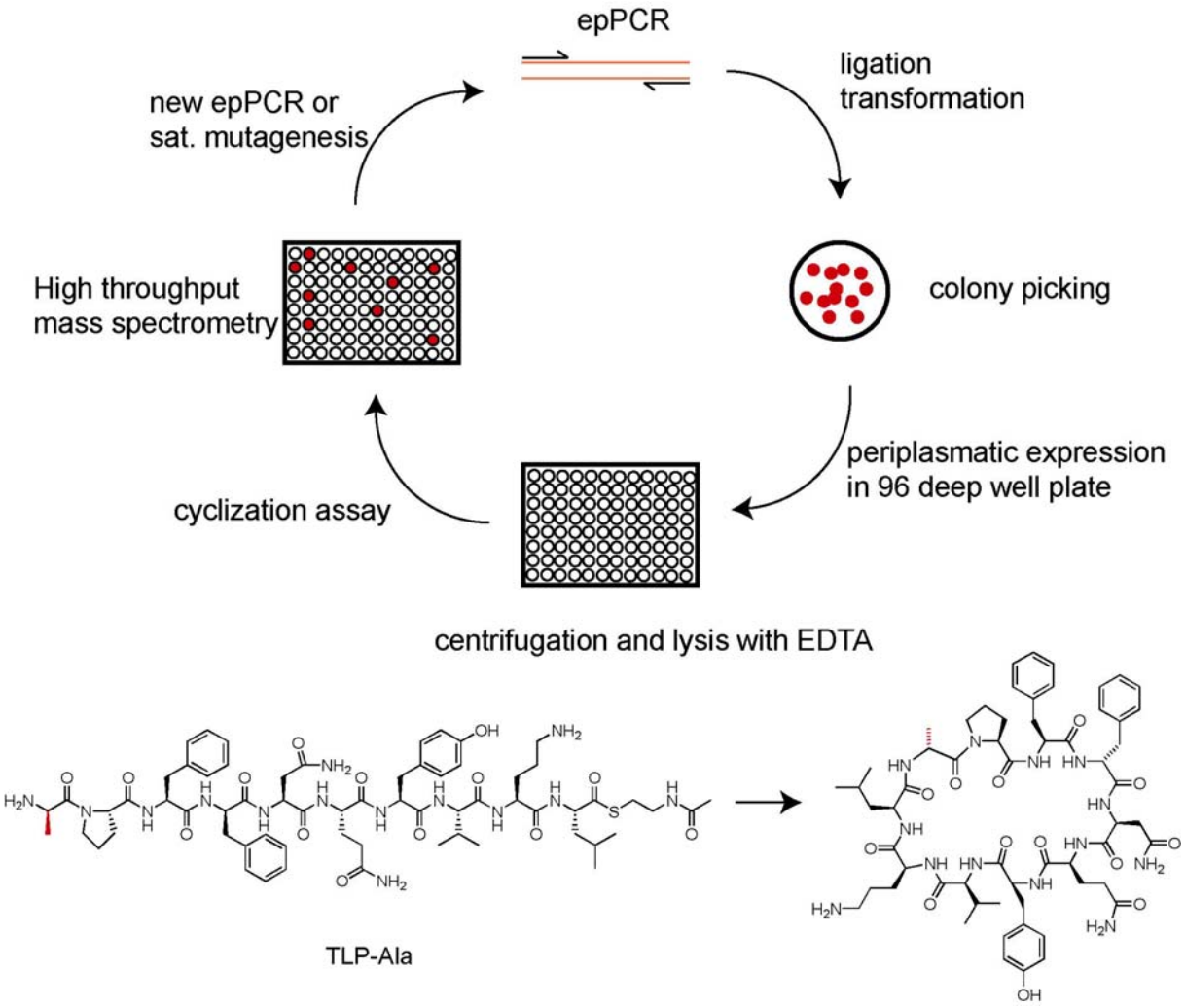


Figure 5.33. High troughput screen methodology. Any species which would cyclize linear TLP-Ala could be identified by a quantitative mass screen.

The strategy of the directed protein evolution approach is summarized in figure 5.33. Random mutations of 2-3 per 273 bp are introduced into the DNA fragment coding the lid region. These fragments are then ligated into a pASK IBA2 vector which allows periplasmatic expression of the protein. After transformation of cells the proteins are expressed in 96 well plates. The cells are harvested by centrifugation and EDTA treatment which permealize the cell wall and release the periplasmatic proteins into solution. Subsequent incubation with **TLP-Ala** would then reveal active species identified by high throughput mass spectrometry. An alternative screen could be based on the antibiotic activity of cyclic **TLP-Ala,** which inhibits the growth of *B. subtilis*. The MIC of cyclic **TLP-Ala** was experimentally determined to be 1.5 µM. Cyclic **TLP-Ala** was provided by a chemical cyclization method under high dilution conditions (data not shown).

5.2.4.5 Error Prone PCR and Periplasmatic Expression

Tyc TE was cloned into a pASK IBA2 vector with a C-terminal strep-tag and a N-terminal OmpA-sequence for periplasmatic expression. Two new unique restriction sites (XhoI and SacI) were incorporated by silent mutations before and after the 273 bp lid region and confirmed by DNA sequencing. Error prone PCR of this DNA region was performed with the GeneMorph mutagenesis kit (Invitrogen). A mutation rate of 2-3 on DNA level was selected leading to 1-2 mutations on protein level on average which were confirmed on randomly picked colonies by DNA sequencing (Fig. 5.34). According to the formula $N = 19^M X!/(X-M)! M!$[140] the number of colonies (N) which have to be screened in order to obtain a single amino acid exchanges at all 93 positions (X) by a mutation rate of 1 (M) is 1767. This number of colonies is reasonable to screen by the method described here. Error prone PCR helps to identify hot spot regions of the enzyme amino acid sequence. Identified hits will be further refined by saturation mutagenesis which introduces all other 19 amino acids at the identified position and can therefore help to select the best amino acid for this position.

EERGYGVTDIILFDSYWKDKAIERTVAETENDIAQLFAEIGENTEMFNMTQEDFQLYAANEFVKQSFVRKTVSYVMFHNNLVNTGMTTAAIHLIQSEL wt
EERGYGVTDIILFDSYWKDKAIERTVAETENDIAQLFAEIGENTKMFNMTQEDFQLYAANEFVKQSFVRKTVSYVMFHNNLVNTGMTTAAIHLIQSEL
EERGYGVTDIILFDXYWKDKAIERTVAETENDIAQLFAEIGENTEMFNMTQEDFQLYAANEFVKPSFVRKTVSY V I FHNNLVNTGMTTAAIHLIQSEL
EERGYGVTDIILFDSYWKDKAIERTVAETENDIVQLFAEIGENTEMFNMTQEDFQLYAANEFVKQSFVHKTVSYVK FHNNLVNTGMTTAAIHLIQSEL

Figure 5.34. Control sequencing at selected mutation rate. 1-2 mutations on protein level are observed.

Activity of Tyc TE in the periplasmatic cell extract was confirmed by incubation with **TLP** which revealed cyclization and hydrolysis activity. In a negative control **TLP** was stable with an extract of cells, which carried an expression vector lacking the tyc TE insert (data not shown). Tyc TE was also purified by strep-tag affinity chromatography (Fig. 5.35).

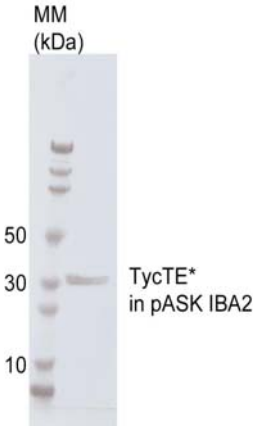


Figure 5.35. SDS-PAGE of Tyc TE* with new restriction sites in the lid region and strep tag. The expected mass (30 kDa) corresponds to the observed mass. MM = molecular marker.

5.2.4.6 High Throughput Mass Analysis

About 1700 mutant colonies of Tyc TE were overexpressed in 96 well plates and screened for cyclization activity by high throughput mass spectrometry together with Uwe Linne. Hits were manually identified by integration over the mass signals of the cyclic (1194.5 Da) and hydrolytic (1212.5 Da) peaks. Figure 5.36 shows mass signal and integration for a hit compared with a background signal. The cut-off for a hit identification was a 4-fold signal intensity increase of cyclization over background. The screen of about 1700 colonies revealed 6 hits in total (10B6, 10C12, 10E5, 6C10, JH1, and 9E2).

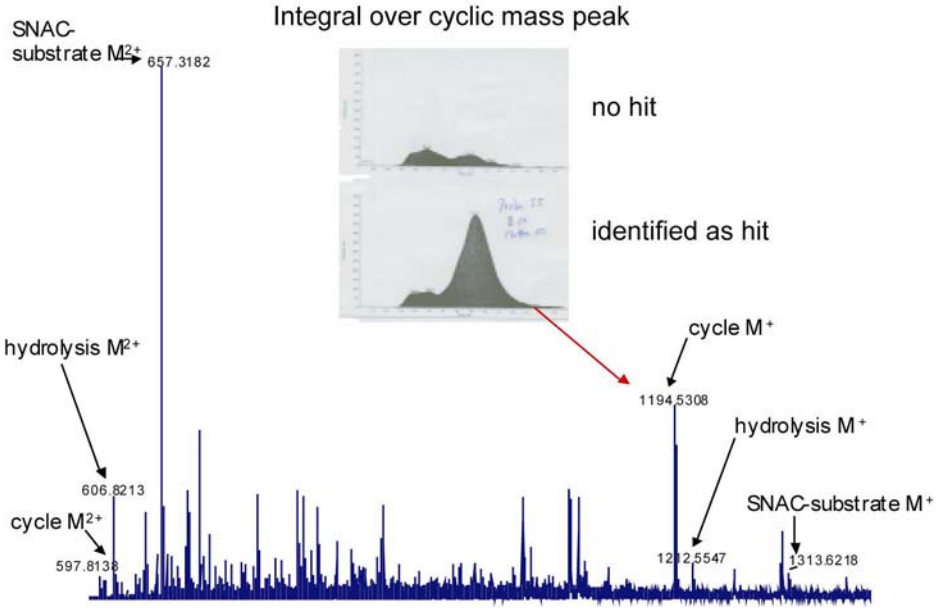


Figure 5.36. Mass spectra of an identified hit. All relevant mass signals are labled. Integrals over the cyclic mass peaks are shown for a background sample (no hit) and an identified hit (identified as hit).

5.2.4.7 Biochemical Characterization of Hits

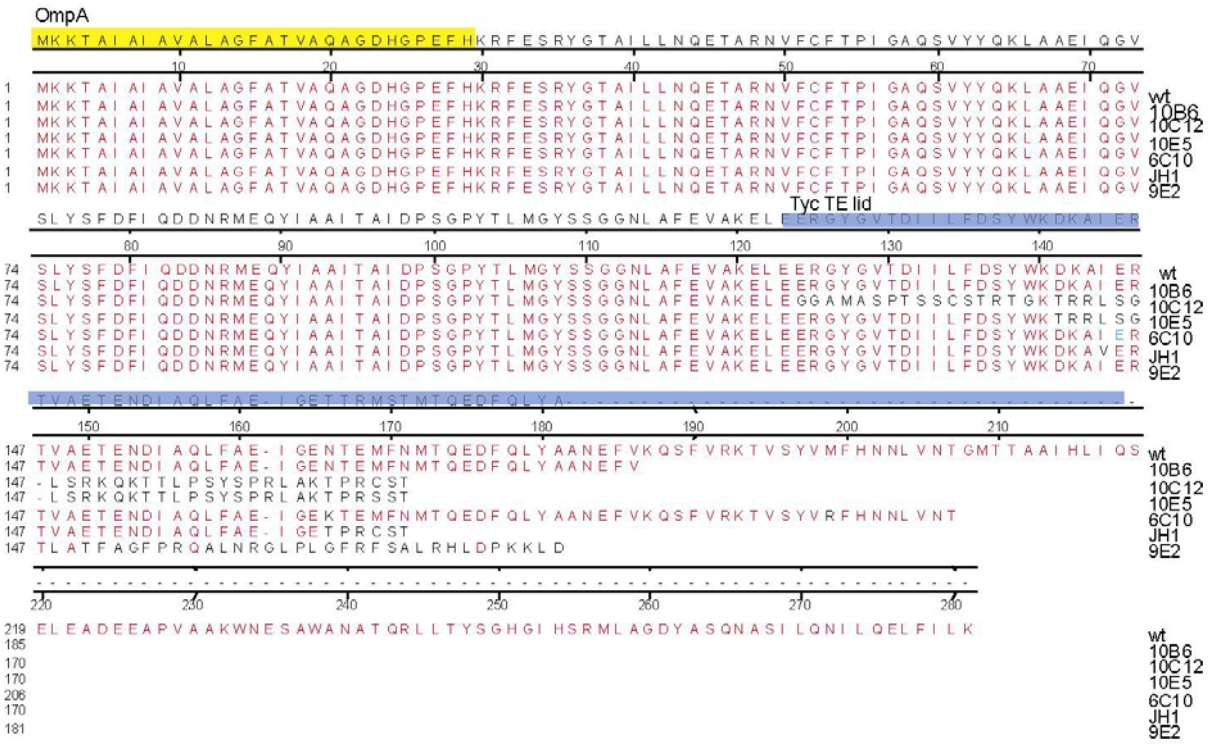


Figure 5.37. DNA sequnence alignment of the six selected hits with the wild type enzyme. Consensus sequences are indicated in red color and the lid region is highlighted by blue shade. The OmpA signal peptide is indicated in yellow shade.

All hits were first characterized by DNA sequencing and an alignment of the protein sequences in comparison with the wild type enzyme is shown in figure 5.37. Interestingly all 6 mutants carry stop codons in the lid region leading to shorter enzyme fragments. In 5 of the mutants these stop codons were derived from frame shifts due to a single nucleotide deletion. Stop mutations occurred in a region of residues 141 – 178 of the TE domain leading to enzyme mutants lacking 74 – 111 residues (30 – 40 %). According to the manufacture´s protocol (Invitrogen) these deletion mutants occur with a frequency of 1.1%. Since the shorter enzymes did not carry a C-terminal strep-tag anymore, all mutant inserts and the wild type insert were recloned into a pASK IBA4 vector, which carries a N-terminal strep-tag in order to perform biochemical characterization of purified proteins. A testexpression revealed induction of all proteins according to their predicted size except the 10C12 mutant, which did not show protein production (Fig. 5.38). Subsequent purification of wild type Tyc TE (29.1 kDa) and three selected mutants (9E2, 19.8 kDa; 10B6, 20.6 kDa; JH1,18.7 kDa) allowed for initial biochemical investigations of these shorter enzymes. The SDS-PAGE of the 4 enzymes shows double bands, which are probably due to incomplete splicing of the 2 kDa OmpA signal sequence (Fig. 5.38).

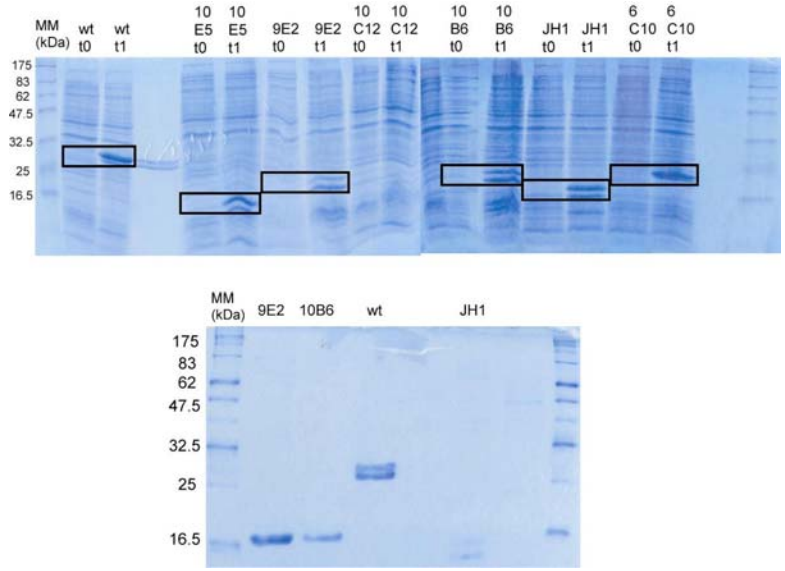


Figure 5.38. SDS-PAGE of testexpression and purification with selected Tyc TE variants. The testexpression shows protein fragments (boxed) which are smaller in comparison with the wild type enzyme. The second gel shows proteins which were purified on a strep tag affinity column. Observed masses of purified proteins fit well with theoretical masses: 9E2, 19.8 kDa; 10B6, 20.6 kDa; JH1,18.7 kDa, wt, 29.1 kDa (including the strep-tag and OmpA sequence). MM = molecular marker.

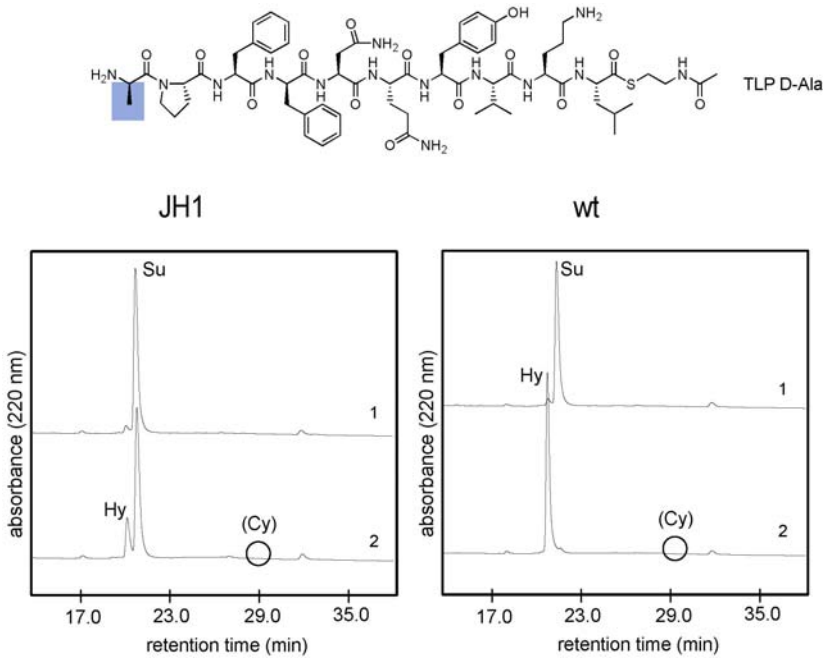


Figure 5.39. Incubation of TLP-Ala with mutant fragment JH1 and wild type Tyc TE (wt). HPLC traces of 1 hr incubation with 5 µM enzyme and 100 µM substrate in 25 mM Hepes and 50 mM NaCl, pH 7. Hy = hydrolysis product, Su = substrate, and (Cy) = trace amounts of cycle. The D-Ala residue of the substrate is indicated by a blue box in the peptide structure.

Surprisingly, none of the mutants showed increased cyclization activity of **TLP-Ala** in comparison with wild type Tyc TE. Traces of cyclic product were observed in the same range as produced by wild type Tyc TE (Fig. 5.39, Table 4.8). By changing the standard assay buffer to buffer P (containing 250 mM sucrose), which was used for the periplasmatic screen, a small increase of cyclic product was observed for the mutants as well as for the wild type. Hydrolysis of substrate was also detected for the mutants indicating that they were still able to form acyl-enzyme intermediates although lacking substantial parts of the wild type protein sequence including the catalytic triad His. In order to further characterize the activity of these short enzymes, substrate incubations with **allyl-TLP** and **TLP** were carried out. In comparison with **TLP**, **allyl-TLP** carries an allyl-glycine instead of phenylalanine at position 3 (allyl-TLP was used in a different study to investigate post-cyclization modifications on the functionalised allyl group). Interestingly, the assays revealed cyclization and hydrolysis activities for all mutants in varying amounts (Fig. 5.40, Table 4.8, Table 5.4).

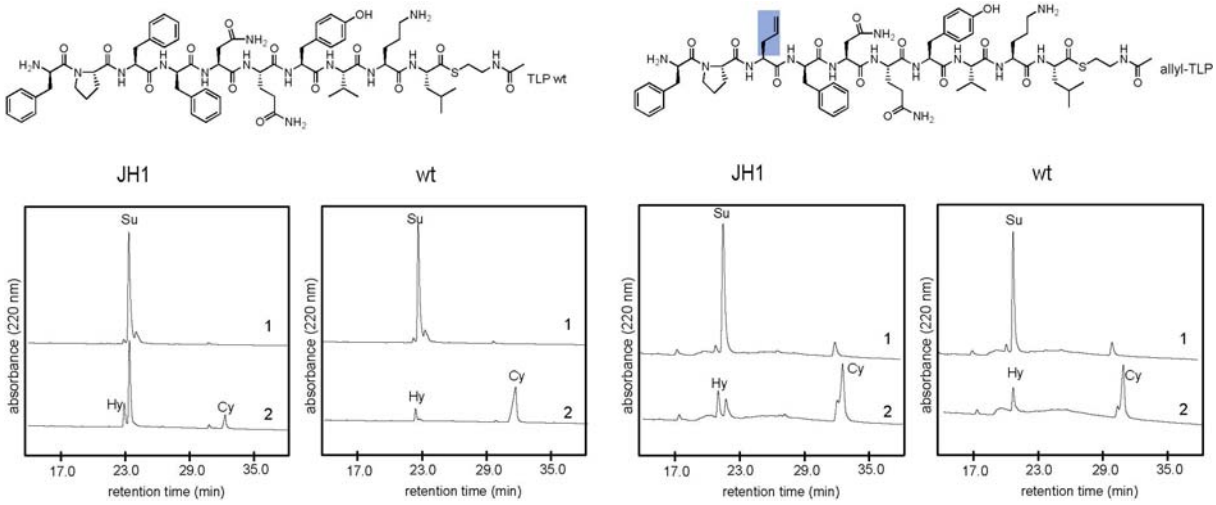


Figure 5.40. Incubation of TLP-wt and allyl-TLP with mutant fragment JH1 and wild type Tyc TE (wt). HPLC traces of 1 hr incubation with 5 µM enzyme and 100 µM substrate in 25 mM Hepes and 50 mM NaCl, pH 7. Hy = hydrolysis product, Su = substrate, and Cy = cyclic product. The allyl-Gly residue of allyl-TLP is indicated by a blue box in the peptide structure.

The best cyclization/hydrolysis ratios for **allyl-TLP** and **TLP** were observed for JH1 with 3.5 and 0.8 respectively (Table 5.4).

Table 5.4. Cyclization to hydrolysis ratios for stop fragments with TLP-wt and allyl-TLP.

TLP-wt		**allyl-TLP**	
	Cy/Hy		Cy/Hy
wt	7 (6 [90])	wt	3,2
10B6	0,24	10B6	0,5
JH1	0,8	JH1	3,5
9E2	0,13	9E2	0,03

Kinetic data for the cyclization of **TLP** (k_{cat} = 0.06 min^{-1} and K_M = 25.2 µM) catalysed by JH1 shows saturation kinetics. Compared to wild type Tyc TE with strep tag (k_{cat} = 55 min^{-1} and K_M = 29.5 µM) the catalytic efficiency is drastically reduced. Differences in protein-tagging, purification strategy, as well as buffer system may account for the slightly increased K_M value in comparison with his-tagged Tyc TE[90].

In order to evaluate the influence of reduced protein size on catalytic activity, two new substrates (**Ala-Leu-TLP** and **Ala-Ala-TLP**) lacking C-terminal recognition elements (Orn 9 in case of Ala-Leu-TLP and Orn 9, Leu 10 in case of Ala-Ala-TLP) were synthesized (Fig. 5.41). In contrast to wild type Tyc TE, which was able to cyclize both substrates, none of the other mutants revealed cyclization but only hydrolysis activity. Figure 5.41 shows the incubation of the wild type enzyme and JH1 with the substrate variants (Table 4.8).

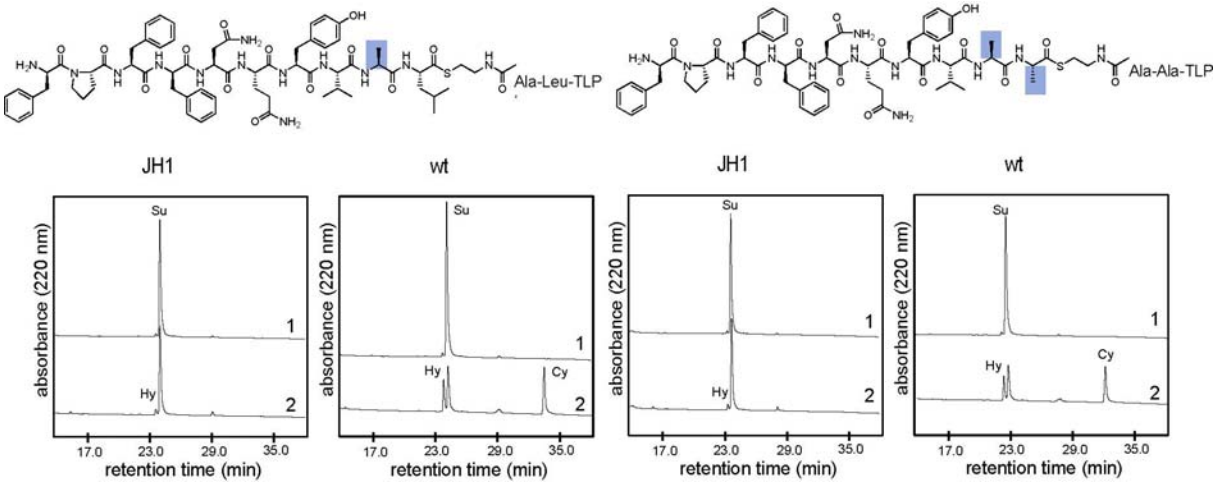


Figure 5.41. Incubation of Ala-Leu TLP and Ala-Ala-TLP with mutant fragment JH1 and wild type Tyc TE (wt). HPLC traces of 1 hr incubation with 5 µM enzyme and 100 µM substrate in 25 mM Hepes and 50 mM NaCl, pH 7. Hy = hydrolysis product and Su = substrate. The Ala residue of Ala-Leu-TLP and both Ala residues of Ala-Ala-TLP are indicated by blue boxes in the peptide structure.

6 Discussion

One of the current challenges in NRPS research is to reengineer natural products in order to increase or alter their biological activities. Research of the last years revealed that NRPS as well as NRPS-PKS hybrids can produce biologically active compounds with an outstanding potential for new drug discovery. A prominent example is the mixed NRPS-PKS product epothilone which is a promising candidate to combat cancer [141-143]. One goal is to improve and alter natural product based drugs by rational protein engineering of these enzymes (e.g. by module and domain swapping). Therefore understanding of any oligomeric structure of these enzymes might be crucial and was subject of investigations in the first part of the research presented here. Recently a second strategy has evolved which uses a combination of chemistry and enzymology to create libraries of novel compounds. Solid phase peptide chemistry is a well established method to provide linear peptides in sufficiently good yields. The critical part is cyclization which requires much effort in chemical synthesis[38] but can be overcome by enzymatic means. Therefore the chemoenzymatic approach presented in the second part utilizes chemistry and biology and concentrates on the strength of the individual discipline. Both approaches also aimed to contribute to the global understanding of NRPS natural product assembly in general, or for an individual domain, which would help for further biological or chemical engineering.

6.1 Quaternary Architecture of NRPS

The study addressed the question of whether NRPS need a higher-order quaternary structure to function or if they work as monomers. Several strategies, which were already successfully applied to FAS and PKS systems previously and which elucidated several independent aspects of protein-protein interactions were used.

Results gained from gel chromatography showed a monomeric behaviour of each of the NRPS enzymes. The experimental masses were slightly higher than the theoretical masses which is likely due to the structure of the proteins differing from the ideal globular shape [144]. The results obtained from gel filtration therefore indicate a monomeric structure of NRPS which is predominantly consistent with the literature [84, 96, 106, 114, 145]. A control experiment with NRPS-PKS hybrid HMWP1 subunit of yersiniabactin synthetase revealed dimers and monomers during gel filtration indicating an association. This shows that the method is robust enough to detect dimers and that the PKS/NRPS hybrid enzyme HMWP1 seems to be dimeric.

Analytical equilibrium ultracentrifugation is a reliable method to determine the molecular weight and detect oligomerization of solutes in the native state. Unlike gel filtration this method is independent of calibration or shape assumptions and is therefore a good supplement to the gel filtration results.

Ultracentrifugation experiments with the PKS enzyme 6-deoxyerythronolide B synthetase 1 (DEBS1) have substantially contributed to the finding of an overall dimeric structure for PKS [100]. Analytical ultracentrifugation experiments carried out to obtain the quaternary structure of NRPS are very rare in the literature. In a sedimentation velocity analysis experiment with cyclosporin synthetase (1.4 MDa) a sedimentation coefficient of 26.3 S was observed, which led to a reasonable structural interpretation only in case of a monomer with an oblate overall shape [146]. Here analytical equilibrium ultracentrifugation was applied to $TycB_{2-3}$ and EntF, peptide synthetases involved in the production of a cyclic decapeptide and a siderophore, one from a gram-positive *B. brevis*, one from a gram-negative *E. coli*. Results for EntF clearly indicate a monomeric structure and no aggregation at even high concentrations which is in agreement with results from gel filtration carried out previously [84]. The dissociation constant of $TycB_{2-3}$ (3.5 µM) calculated from the ultracentrifugation data indicates oligomerization at higher non-physiological concentrations. Assays described here were carried out at concentrations of about 0.17 µM which would correspond to 92 % monomer. Unlike dimeric PKS, these results from ultracentrifugation suggest a predominantly monomeric structure of the two investigated NRPS modules at assay and physiological concentrations.

The cross-linkers BMH and DBP differ in the length of their spacers and therefore in specificity. All three enzyme modules (GrsA, $TycB_1$ and $TycB_{2-3}$) reacted with BMH at low molar excess and one new protein species with an electrophoretic mobility in between monomeric and dimeric molecular weights appeared on the SDS gel. Witowski and coworkers already observed a similar phenomenon with cross-linked FAS [113]. They demonstrated that the new band belongs to a monomer which is intramolecular cross-linked, where altered protein shape explains the changed electrophoretic mobility of the internally cross-linked product. In consideration of the quantity of thiol groups available in the proteins (GrsA 8, $TycB_1$ 11, $TycB_{2-3}$ 13) multiple reactions with the cross-linker seem to be reasonable and support the assumption of an intramolecularly linking. This finding is supported by the results of MALDI-TOF mass spectrometry. The M^+ peaks of GrsA and $TycB_1$ were shifted to higher molecular masses indicating the reaction of 2 molecules of BMH with GrsA and 5 molecules with $TycB_1$. Signals belonging to a dimer were not observed. In contrast to FAS

and PKS no reaction with DBP occurred with GrsA and $TycB_1$, which could be explained by the short spacer of this cross-linker. In the case of FAS, DBP was able to connect the active site cysteine of the KS-domain with the closely located thiol group of the phosphopantetheine cofactor of the ACP inter- and intramolecularly [100, 113]. Such an active site cysteine does not exist in the C-domain of NRPS [65] which is consistent with lack of successful DBP cross-linking in $TycB_1$ while GrsA does not have a C-domain. However, the intramolecular reaction between DBP and the dimodular enzyme $TycB_{2-3}$ indicates that the flexibility of two modules may allow the approach of thiol groups into close proximity. All results gained from cross-linking did not show any dimer formation but demonstrated an intramolecular linking within a monomer.

A biochemical approach to test for oligomerization utilized a two affinity tag system which was adopted from investigations with FAS [128]. Neither GrsA nor $TycB_1$ demonstrated the ability to build heterodimers containing both the strep-tag and hexahistidine-tag, and thereby the capacity to bind on both affinity columns successively. The specificity of the columns and of the applied antibodies were investigated and confirmed by separate experiments. Provided that the two scrambling methods were sufficient to allow dissociation and reassociation of putative dimers, the results clearly demonstrate a monomeric character of the enzymes. Complementation experiments between C- and PCP-domains were also adapted from previous investigations of an analogous KS and ACP complementation in FAS and PKS enzymes. The dimodular NRPS system $TycB_{2-3}$ used for these studies had the advantage to provide C-PCP interactions *in cis* compared to *in trans* in a monomodular system. None of the complementation experiments revealed any product formation, indicating that no functioning heterodimers were formed, presuming a sufficient scrambling process. Based on the results from ultracentrifugation, a maximum population of 4% heterodimers would have been possible. Product formation catalyzed by them could have been detected within the limit of the HPLC assay.

The critical step to conduct the biochemical experiments is the scrambling of homodimeric subunits to form heterodimers. Positive results would have proved unequivocally the dimeric nature of the investigated enzymes. In contrast the negative results obtained may have two explanations: the enzymes are monomeric or the scrambling procedure was not sufficient. The zwitterionic detergent CHAPS associates with nonpolar protein residues and therefore disrupts hydrophobic interactions. The tertiary structure, at least of the A-domain, is not influenced by this process which was demonstrated by an unchanged aminoacylation activity of the enzymes (GrsA, $TycB_1$ and $TycB_{2-3}$) in presence of CHAPS and after dialysis.

Furthermore it was possible to show that CHAPS is in position to interrupt functional protein-protein interactions *in trans* between GrsA and $TycB_1$. Whether CHAPS would also be able to interrupt protein-protein interactions in a putative NRPS dimer is speculative. The application of source Q anion exchange chromatography was adapted from PKS experiments [127]. The principle of this method is the dissociation of proteins induced by the strong interaction with the matrix and the subsequent reassociation by an increasing salt gradient (David Cane, 2000, personal communication). Scrambling of dimeric PKS and FAS worked very well with anion exchange chromatography.

Do NRPS have an oligomeric quaternary structure?

The related organization and chain elongation logic of PKS and NRPS as well as the existence of naturally occurring PKS-NRPS hybrids led to the assumption of a similar quaternary structure of both enzymes.

Enzymes of physiological oligomeric significance other than NRPS vary greatly in degree and strength of interaction. Examples for strongly interacting enzymes are pyruvate dehydrogenase (PDH) systems with 24-mer or 60-mer cores [147]. The complex oligomeric organization of these systems is important for the channeling of intermediates, and stiff linkers between domains of each subunit allow motion without excessive degrees of freedom [147]. In contrast to these strong interactions, FAS dimers are able to dissociate in monomers at low temperatures whereas an equivalent behaviour has never been observed for related PKS. The observed tendency of NRPS to dimerize at higher concentrations can therefore be an evolutionary relic and does not appear to have functional significance.

Crystallographic data from the PKS thioesterase domain shows a hydrophobic leucine-rich dimer interface and a 2-fold symmetry axis of the dimer [148]. Such an interface has not been observed for any NRPS domain based on crystallographic data [96, 104, 105] indicating that one necessary prerequisite to form dimers does not exist for NRPS.

In the case of FAS and PKS the functional dimeric interaction takes place between the ketosynthase domains and the ACPs. Covalent intermediates can be isolated where the substrates are bound to a conserved cysteine residue of the ketosynthases and it is mechanistically necessary that this cysteine interacts with substrates bound to the ppan-cofactors of ACPs of the second polypeptide chain of the dimer [99]. For C-domains of NRPS systems, recently a reaction mechanism was suggested by the results of a mutational analysis that does not involve covalent acylenzyme intermediates [65]. There may be no such mechanistic need for NRPS to function as dimers. In a recent publication of Smith and

coworkers it was shown that there is contrary to previous believe also no mechanistic need for FAS to work as dimers[111]. Full inactivation of one polypeptide chain in a dimer did not abolish activity indicating that one functional subunit in a heterodimer is sufficient for product assembly.

In conclusion, although there is no direct experimental proof for the efficacy of the two methods for scrambling subunits between possible NRPS oligomers, the results from gel chromatography, analytical equilibrium ultracentrifugation and cross-linking demonstrate no indication of NRPS dimers for these NRPS subunits. A monomeric structure is also supported by recent studies of cyclosporin synthetase with transmission electron microscopy which demonstrates that nearly all 11 modules of one synthetase are not interacting with a second synthetase in a dimer [149]. It is also important to note that in a recent publication by Hillson et al. the dimeric structure of the unusual NRPS system VibF was demonstrated by analytical ultracentrifugation and complementation studies[150]. In contrast to cold labile FAS the VibF dimers did dissociate at elevated temperatures which suggestes a different pattern in dimer breakdown and re-formation. Crystallographic data of the accompanying free standing C-domain VibH revealed a monomeric structure which indicates two oligomeric states in the vibriobactin system. This is similar to the proposed difference in aggregation of PKS and NRPS systems with regard to the existence of hybrid enzymes. A possible structure for these hybrid enzymes could be a dimeric PKS core with monomeric NRPS loops. This structure is similar to the postulated PKS structure where optional domains are separated from the dimeric core in monomeric loops [44].

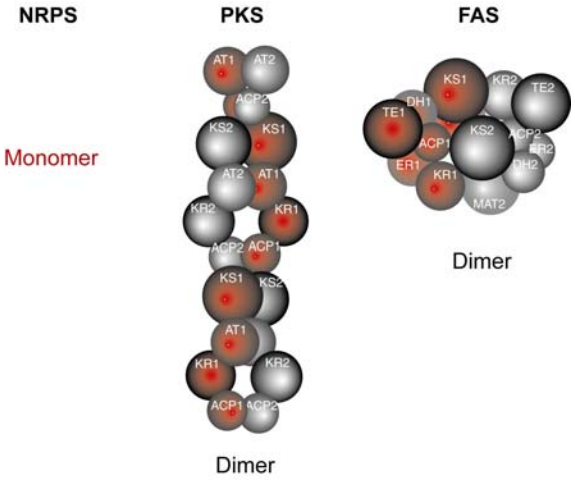


Figure 6.1. Quaternary architecture of multienzyme complexes. In contrast to PKS and FAS, NRPSs investigated here are monomeric.

6.2 Enzymatic Peptide Cyclization

6.2.1 Characterization of Surfactin Thioesterase

Previous studies with the tyrocidine thioesterase domain revealed its activity with synthetic SNAC substrate analogues and allowed to characterise enzymatic substrate tolerance[80, 90]. In contrast to Tyc TE, which catalyses head to tail cyclization, only little was known about branched chain cyclizing enzymes like the TE domains of surfactin, fengycin and syringomycin synthetases. Srf TE catalyzes the cyclization of linear β-hydroxy-acyl-heptapeptidyl thioesters to produce cyclic macrolactones through the capture of the Leu_7 carbonyl by the fatty acyl β-hydroxyl nucleophile (Fig. 2.2). This is representative of the terminal step in the biosynthesis of other cyclic peptidolactones, including globomycin[151] and tripropeptin[152], and analogous peptidolactams, such as the iturins[153]. The related nonribosomal peptidolactones in the fengycin and syringomycin classes of antibiotics are similarly macrocyclized at the end of the NRPS assembly line by TE domains that use as the cyclization nucleophile a β-hydroxyl group of an amino acid side chain instead of a fatty acid. To evaluate the catalytic capacity of these cyclic depsipeptide synthases, active site mutants of Srf TE were characterized for cyclization and thioester hydrolysis activity. An assessment of the contributions of the fatty acyl group, heptapeptide backbone, and SNAC leaving group to substrate recognition for both cyclization and hydrolysis was carried out.

6.2.1.1 Structure and Mechanism

The superfamily of serine esterases generally uses a catalytic triad of active site residues, with serine as nucleophile, histidine as acid/base catalyst, and aspartate to orient the histidine and serine residues optimally[154]. Mutational analysis on the chicken FAS TE domain[155] and enterobactin synthetase TE domain[97] validated the roles of the active site serine and histidine. In the enterobactin system, the Ser to Cys mutant showed a 1000-fold decrease in activity compared with the wild type enzyme, and the His to Ala mutant showed a 10,000-fold decrease[97]. The X-ray structure of the Srf TE domain indicated that Ser80, His207, and Asp107 were architecturally disposed to be such a catalytic triad[96]. Mutagenesis results reported here verify that the Ser80 to Cys and His207 to Ala mutants were without detectable activity in both cyclization and hydrolysis assays, with the Asp107 to Ala mutant also showing no cyclization activity and only a trace amount of hydrolysis activity. The lack of catalytic activity with the Ser80 to Cys mutant could be explained by the steric crowding in the active site due to the large sulfur atom which was also reported for a mutant version of subtilisin[156]. In acyl-*O*-Ser enzyme intermediate formation and breakdown, anionic

tetrahedral adducts are predicted, and electrostatic stabilization via oxyanion hole interaction is a general tenet of the superfamily catalytic mechanism[154]. For Srf TE, the backbone amides of Ala81 and Val27 are thought to serve this function (Fig. 5.13A)[96]. This work focused on Pro26 just adjacent to Val27, since it is a conserved residue in NRPS TE domains, but is instead conserved as a glycine in the lipase subfamily of serine esterases, which carry out acyl transfer exclusively to water as the nucleophilic cosubstrate. In fact, the Pro26 to Gly Srf TE mutant had a 13-fold increase in k_{cat}/K_M, with a 12-fold change in product partition ratio away from intramolecular cyclization in favor of intermolecular hydrolysis. The change from a rigid proline to a flexible glycine may increase the conformational freedom in this region of the active site, and create more access for water to capture the acyl-heptapeptidyl-*O*-Ser enzyme intermediate. In turn, it may be that mutation of lipase active sites to introduce the proline may enhance intramolecular cyclization outcomes. The active site pocket adjacent to the catalytic triad has the attributes of a hydrophobic bowl, likely a good match for the generally hydrophobic nature of the acyl-heptapeptidyl chain to be cyclized to the biosurfactant product[96]. There are two prominent cationic side chains in the bowl, Lys111 and Arg120; to evaluate whether removal of these side chains altered catalytic efficiency, the Lys111 to Ala and Arg120 to Ala mutants were prepared. The Lys111 to Ala mutant was inactive for cyclization of substrate thioesters and displayed only a low level of hydrolysis activity, while the Arg120 to Ala mutant retained a higher fraction of activity with product ratio now skewed towards hydrolysis. A prediction for such charged active site residues to be involved in the recognition of peptide backbone carbonyl groups in the substrate was made for the PKS DEBS TE domain[148].

6.2.1.2 Substrate Tolerance

One of the goals of this study was to evaluate features in substrates that allow the excised Srf TE to carry out the chemo-, regio-, and stereospecific macrocyclization to the lariat type of peptidolactone, as prelude to evaluation of its catalytic potential for biogenesis of novel macrocyclic products. A significant limitation in an initial study was the poor aqueous solubility of the fatty acyl-heptapeptidyl-SNACs[80]. As a result, the acyl chain from the natural C_{13}-C_{15} length was shortened down to C_4. In this work, β-hydroxybutyric acid was used for the ease of synthesis, purification, and assay, but it will be necessary to explore how longer fatty acyl chains may affect catalytic efficiency. Recent data suggests that CDA peptide cyclase catalyses cyclization of a CDA analogue with a natural hexanoic acid more efficiently with only trace amounts of hydrolysis. By contrast CDA analogues with shorter fatty acids show little regioselectivity and increased hydrolysis[157]. This indicates the

existence of a hydrophobic fatty acid binding pocket in the enzyme which will be discussed below. Even with the C_4 acyl chain, the surfactin substrates showed CMCs in the 0.6-2.0 mM range, creating problems in kinetic assays. Solubility was increased by systematically introducing a Dap residue at each of the seven positions of the heptapeptide sequence of surfactin. The substrate with Dap at the third position (**SLP-3**) was more tractable than the substrate with the wild type heptapeptide sequence (**SLP-wt**), but was still not optimal for deconvolution of k_{cat} and K_M values.

Nonetheless, HPLC assays could be performed on all of the substrates and relative rates of both cyclization and hydrolysis determined (Fig. 5.7C). Both the intermolecular hydrolysis and the intramolecular cyclization reactions were enzyme-mediated and are presumed to reflect competing fates for deacylation of the acyl-heptapeptidyl-*O*-TE enzyme intermediate (Fig. 2.2). Some structure/activity data for TE action were acquired on the nucleophile (X) in the acyl moiety, determinants in the heptapeptide chain (Y), and leaving group (Z), as highlighted in figure 6.2.

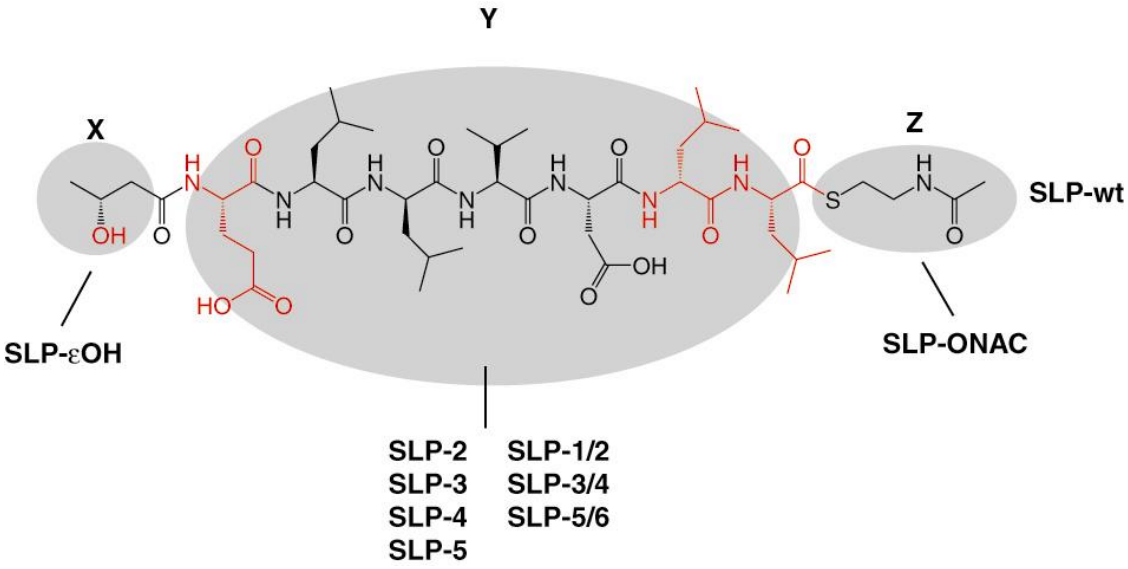


Figure 6.2. Summary of alternate substrates tolerated by SrfTE. In this study, different parts (X, Y, and Z) of a peptidyl-SNAC substrate (SLP-wt) were altered. X: Nucleophilic attacking group of fatty acid, Y: heptapeptide sequence, Z: leaving group. The substrates listed could be cyclized by Srf TE.

For the C_4 acyl chain, when X = OH, the peptidolactone could be formed by macrocyclization, while when X = NH_2, only hydrolysis occurred. It is possible this outcome may be altered for longer acyl chains, since the β-NH_2 of a $C_{14\text{-}17}$ acyl chain is the cyclizing nucleophile in the closely related iturin peptidolactam antibiotic[153]. When the attacking hydroxyl of the acyl chain was moved from the β- to α-position, only hydrolysis occurred,

suggesting misalignment or steric barriers to intramolecular lactonization. However, the straight ε-hydroxycaproyl chain was cyclized to generate a peptidolactone with three more carbon atoms, suggesting further exploration of ω-hydroxy-acylpeptidyl substrates may be useful. In terms of alteration of the heptapeptide backbone, lengthening to the nonamer gave hydrolysis but no cyclization, as though a cyclizing conformer in the active site was disfavored. Previous *in vivo* experiments by Mootz et al. who deleted the leucine activating module 2 of the surfactin synthetase showed, that a shorter substrate was tolerated by Srf TE[158].

The Dap scan, in which each side chain was replaced with one that differed in both size and charge, indicated tolerance of Dap in the middle of the peptidyl chain, but not at the N- and C-terminus, in some analogy with tolerance of the tyrocidine synthetase TE domain to treat internal residues in a cyclizing chain as substitutable cargo (Fig. 5.7)[80, 91]. With regards to the NMR structure of surfactin it is important to note that the two acidic residues Glu 1 and Asp 5 form a polar domain by pointing in an opposite direction as the other hydrophobic residues[159]. This "claw" configuration of acidic residues gives important clues for the cation binding and transporting ability of surfactin. Replacement of Asp 5 by Dap or other residues could therfore likely alter the bioactivity of surfactin analogues.

On the other hand, the results of the spacer scan, which replaced pairs of adjacent residues with a minimal linker containing no peptide bonds or side chains, indicated that, while the residues toward the N- and C-terminus of the substrate are not tolerant to significant side chain changes, the presence of sidechains at either end is not necessary for cyclization. In contrast, the removal of the residues in the middle of the substrate had a marked effect on cyclization (Fig. 5.8). In combination with the findings from the Dap scan, this suggests that the peptide backbone of the middle residues is an important feature for substrate placement and orientation for cyclization.

The leaving group, Z, is a ppan chain covalently tethered to the PCP of the seventh module of surfactin synthetase in the native environment of the surfactin synthetase assembly line (Fig. 2.2). This macromolecular portion of the substrate can be replaced by SNAC in case of the surfactin and tyrocidine cyclases. Changing from a thiol to hydroxyl leaving group still retained sufficient thermodynamic activation in the substrate for Srf TE-mediated cyclization. In the cognate Tyc TE case, such oxoester chemistry was used to make peptide libraries on beads to allow screening of libraries that can be enzymatically cyclized[95]. In addition, Srf TE domain also catalyzes the reverse reaction, the peptidolactone ring opening, where the

leaving group is the β-hydroxyl of the acyl chain. The influence of other leaving groups on Srf TE mediated cyclization will be discussed below.

6.2.1.3 Substrate Recognition - Cocrystallization

While the expected acylheptapeptidyl-O-TE intermediate has yet to be isolated, e.g., in rapid quench studies, the cocrystallization of the D-Leu-L-Leu-boronate with the boron bridging Ser80 and His207 is consistent with the mechanistic prediction for the catalytic pathway (Fig. 5.14A). In agreement with biochemical data, only the peptide side chains of D-Leu_6 and L-Leu_7 show specific binding in the active site pocket, each bound tightly in a well-defined hydrophobic pocket. These two amino acids fill up most of the active site, with the remainder of the peptide chain jutting out of the pocket (Fig. 5.14). Starting from this orientation of residues 6 and 7 of the heptapeptidyl chain and the NMR structure of the cyclic surfactin product[159], one can model the full heptapeptidyl chain in a product-like conformer as prelude to subsequent structure-activity-engineering mutagenesis on Srf TE and, by extension, other macrocyclizing NRPS TE domains. This model sets the surfactin linear peptide in a twisted β-sheet structure pointing away from the active-site pocket into solution (Fig. 5.14C). Consistent with the substrate tolerance of Srf TE, the enzyme makes specific contacts only with the two termini of the linear peptide. The proximity of the fatty acid of the natural surfactin linear lipopeptide to the nucleophilic hydroxy-group may suggest that it is a necessary determinant for cyclization. A pronounced hydrophobic pocket is present in the active site close to the predicted fatty acid position and may assist in increasing cyclization to hydrolysis product distribution.

6.2.1.4 Comparison with Tyrocidine Cyclase

Further comparison of Srf TE to Tyc TE and other macrolactonizing and macrolactamizing TE domains, excised from both NRPS and PKS assembly lines, is warranted to sort out the determinants of intramolecular cyclization vs. intermolecular hydrolysis for the acyl-*O*-TE intermediates. For example, Srf TE makes a branched chain peptidolactone, utilizing a D-hydroxyl stereocenter in the intramolecular nucleophile. Tyc TE also uses a D-stereocenter in the attacking nucleophile, but it is an amine nucleophile and the product is a head-to-tail cyclized macrolactam. Minimal substrate recognition by Tyc TE appears to consist only of a N-terminal D-phenylalanine and the potential of the substrate to preorganize into an antiparallel β-sheet, with multiple residue substitutions and length variations is well tolerated[90, 91, 93, 95]. In contrast, the analysis of Srf TE indicates that it is optimized for its naturally dedicated substrate, with catalysis of cyclization only retained by conservative

changes to the substrate sequence. This might be also explained by the fewer β-sheets displayed in surfactin, which limit substrate preorganization in comparison to tyrocidine. Deciphering the productive folding of acyclic conformers of the acyl chain in the binding pockets of acyl-*O*-TE intermediates will be crucial in future studies to understand cyclization outcomes.

6.2.2 Characterization of Fengycin Thioesterase by Natural Substrate Presentation

In order to expand the scope of macrocyclization catalysts, thioesterases form other NRPS systems were cloned and overexpressed. Surprisingly, no activity was observed for fengycin TE, mycosubtilin TE and syringomycin TE upon incubation with SNAC substrates, which indicated a limitation in the chemoenzymatic potential of these cyclases. There were two possible reasons for the observed inactivity: First, it could be that the enzymes are misfolded and unable to reconstitute their activity after heterologous expression. Alternatively, substrate presentation by the short leaving group SNAC might be insufficient because of problems to find the right way into the enzyme´s active site by diffusion. In order to rule out the second possibility, a new strategy was employed which allowed Sfp catalysed loading of peptidyl-CoAs onto apo PCPs mimicking the natural substrate presentation as close as possible. Remarkably Sfp is promiscous enough to tolerate peptidyl-CoA substrates instead of CoA and acetyl-CoA. This can be explained by the crystal structure which shows specific interactions with the adenine base but leaving the ppan arm pointing out of the enzyme (Fig. 2.6). This ensures enough space and freedom for the attached peptide chain which is accomodated in the solvent. Loading fengycin-CoA onto the fengycin PCP-TE didomain revealed indeed cyclization and hydrolysis activity which was not observed with SNACs before. The result demonstrates that the PCP tethered ppan arm is necessary to direct the substrate into the enzyme active site and ensures appropriate alignment for the nucleophilic attack of the active site serine.

Contrary to surfactin and tyrocidine, the fengycin peptide sequence carries two competing nucleophiles (Orn2 and Thr4) directly next to Tyr3, which is the naturally dedicated nucleophile. Since the hydroxy-group of tyrosine is less nucleophilic in comparison with Orn and Thr[47], it raises the question if one of these competing nucleophiles can complement Tyr if it is exchanged to Phe. Such an exchange in the peptide substrate did only reveal hydrolysis which shows that Tyr is specifically recognized and activated by the TE. Relocation of Tyr at position 2 and 4 did only reveal little cyclization outcome in case of Tyr2 which indicates that an increase of the peptidolactone ring size is possible by one residue. Fengycin is the first cyclic branched chain peptidolactone where regioselectivity of enzymatic macrocyclization

has begun to be explored. By contrast, recent studies on CDA cyclase revealed relaxed regioselectivity between Thr2 and Ser1[157]. This regioselectivity was dramatically increased by using longer N-terminal fatty acyl chains, which correlates to specific binding in a putative hydrophobic pocket in the CDA cyclase and induces correct alignment for the nucleophilic attack of Thr2. Such a pocket was also reported for surfactin TE. In case of fengycin TE, regiochemistry seems to be triggered by specific recognition of the nucleophile Tyr rather than fatty acyl chain length, which was acetate in the substrates investigated.

Fengycin TE is one example for a successful application of the new peptidyl-CoA loading method. Recently this approach was also used to characterize C- and R-domains in order to bypass NRPS specificity by directly loading peptidyl-CoA on any apo PCP within an assembly line[60, 78]. However, a limitation of the method is the single turnover (Fig. 6.3). After product release the cofactor ppan remains attached to the PCP-TE didomain cyclase, which blocks further Sfp catalyzed transfer of additional peptidyl-CoAs onto ppan-PCP.

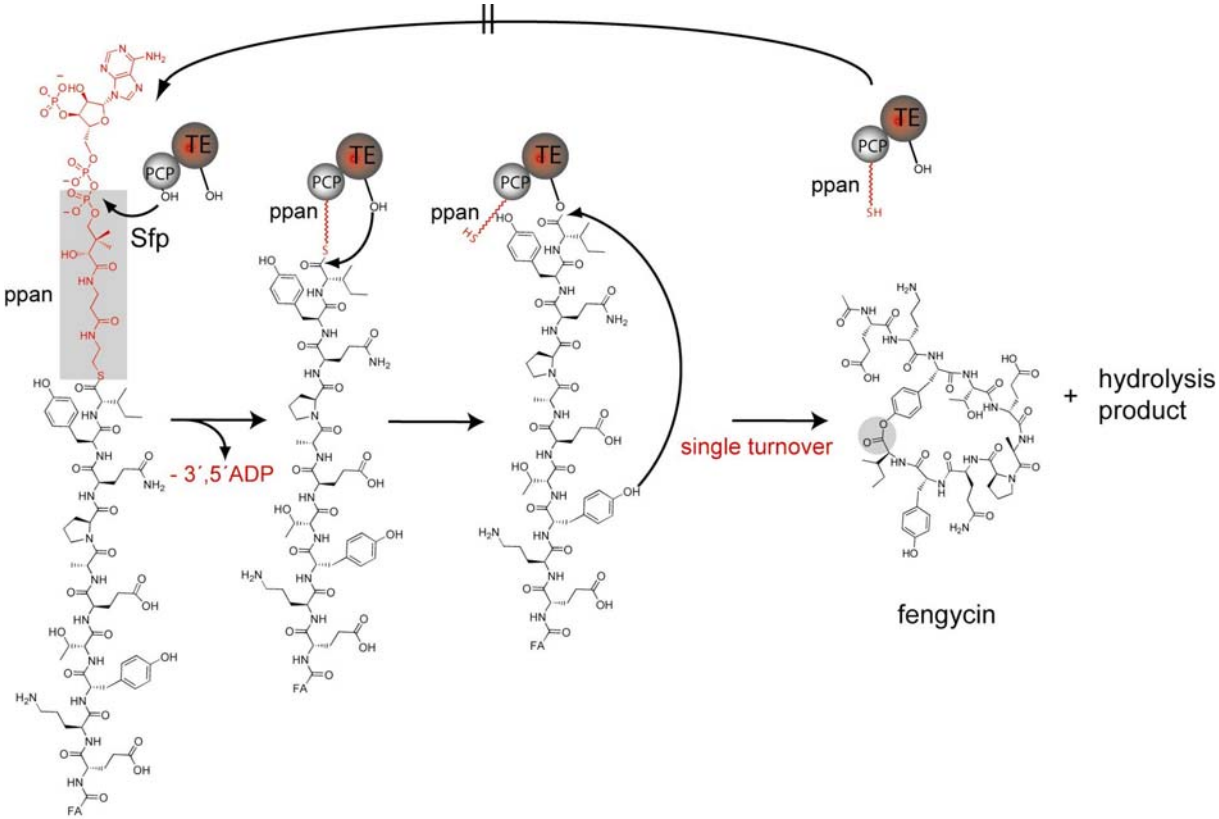


Figure 6.3. Limitation of natural substrate interaction approach: Single turnover. Once the peptide is released, the ppan moiety stays covalently attached to the PCP which blocks the enzyme for subsequent reloading of peptidyl-CoA.

6.2.3 Novel Thiophenol Based Leaving Groups

In order to force multiple rather than single turnover cycles a new strategy was developed to expand the utility of peptidyl-CoA loading. It is known that amino acid thioesters undergo transthioesterification reactions when exposed to thiol-containing compounds[135]. The strategy was initially based on a thioester exchange reaction between the free ppan-PCP thiol and a soluble thioester-peptide substrate which should ensure both chemical reloading of substrate onto the ppan-PCP-TE didomain[134, 135] and natural substrate presentation by the ppan arm. Inspired by expressed protein ligation[134], thiophenol was selected as thioester compound due to its good leaving group properties. In the course of studies it became obvious that soluble fengycin thiophenol did directly acylate the cyclase active site serine rather than the free ppan thiol. It indicates that the serine hydroxyl-group is a much better nucleophile for the attack onto the thiophenol-peptide thioester in comparison with the thiol of the ppan. This is just the opposite of what is expected from chemical predictions since thiols are better nucleophiles than hydroxy-groups because of their better polarizability[47]. If serine is embedded in a functional catalytic triad as observed for peptide cyclases its hydroxyl group proton is in conjugation with the neighbouring histidine which makes it much more nucleophilic compared to serine in solution[160]. The rapid direct acylation of the active site serine therefore makes sense and further confirms the autonomous activity of the excised enzyme. Moreover, this result shows that in contrast to previous believe no natural cofactor recognition elements as displayed in SNAC or CoA substrates are necessary for enzyme acylation but only the potency of the leaving group. After acyclation of the enzyme, thiophenol can delocalize the negative charge across the aromatic ring system while there is no such stabilization for SNAC or CoA substrates (Fig. 6.4).

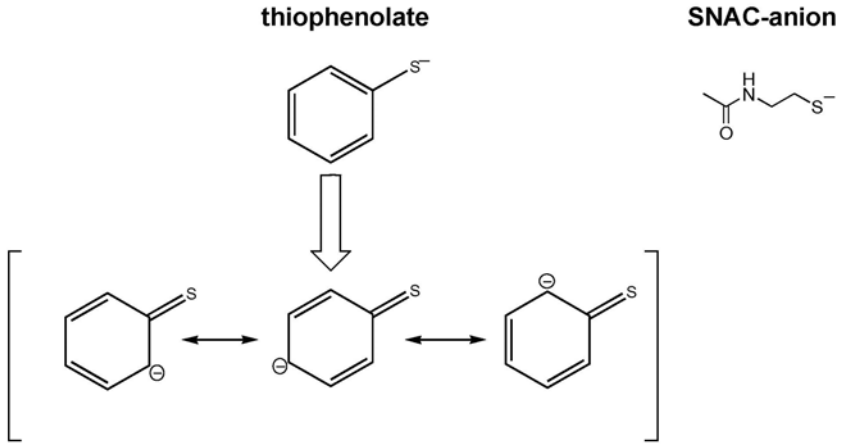


Figure 6.4. Mesomeric stabilization by electron delocalisation of the thiophenolate anion.

The results further suggest that nature has developed peptide cyclases with different catalytic activities. While tyrocidine TE and surfactin TE showed activity directly with SNAC substrates, mycosubtilin TE, fengycin TE and syringomycin TE appeared to be completely inactive. A 15 fold increase in catalytic activity was observed for Srf TE when the SNAC leaving group was exchanged by thiophenol. Recently, a similar increase of activity was also reported for comparison of CDA-SNAC with CDA-thiophenol for CDA cyclase[157]. The thiophenol leaving group increases the velocity of the acylation step. While the first step of acylation depends on substrate presentation, as seen in the peptidyl-CoA experiments, deacylation is an intrinsic property of the enzyme, visualized by comparable cyclization to hydrolysis ratios for **SLP3** and **SLP3-tp** which is consistent with a common acyl-enzyme intermediate for both substrates. By contrast, comparison of *in cis* experiments with PCP-TE bound **FLP-CoA** and *in trans* experiments with **FLP-tp**, the cyclization:hydrolysis ratio decreased from 2:1 to 0.85:1, respectively. This change in ratio could indicate that the ppan arm in the enzyme entrance excludes water from the active site and therefore reduces hydrolysis outcome. As expected, undesired hydrolytic by-products are observed in thiophenol based *in vitro* systems, since substrate presentation and enzyme environment are unnatural. The ratio of cyclization to hydrolysis could be increased by providing natural unmodified thiophenol substrates, that fit precisely into the enzyme active site. In contrast, not much is known about the hydrolysis rate of natural NRPS systems but it is likely that also the multienzyme complexes produce hydrolysed by-products to a certain extend. In general, selective enzyme acylation can now be achieved for many peptide cyclases by a variety of leaving groups (Fig. 6.5).

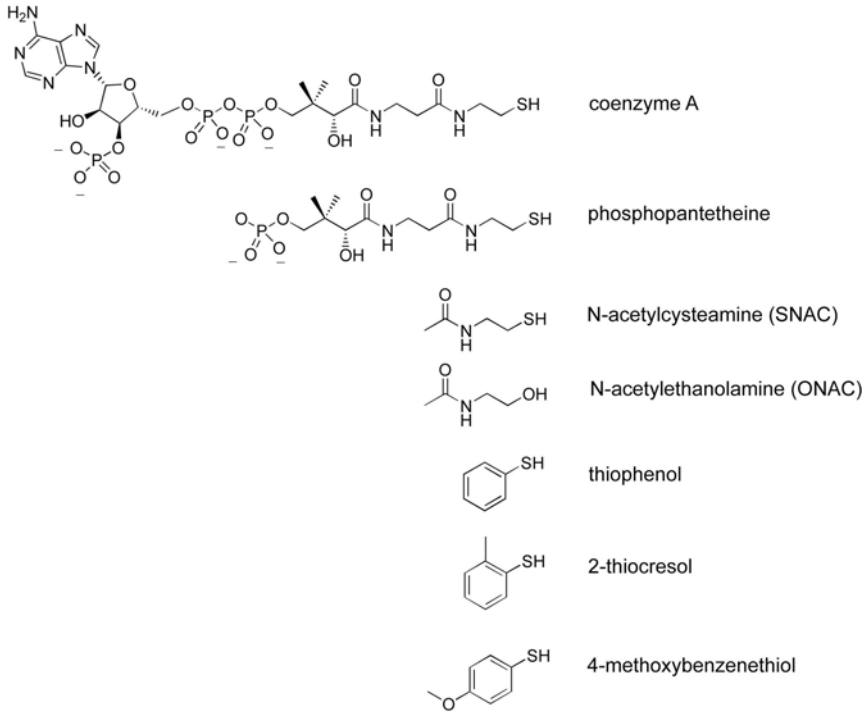


Figure 6.5. Leaving groups used in this study.

6.2.4 Protein Evolution

The previous studies on peptide cyclases revealed that cyclization depends on both, substrate preorganization and specific interaction with the enzymatic binding pocket. Within the cyclases investigated so far, tyrocidine TE exhibits the greatest tolerance for substrate variations in length, side chains, nucleophiles and leaving groups. This high tolerance makes Tyc TE an interesting candidate for cyclizing other substrates than tyrocidine analogues, which has been shown for NRPS-PKS hybrids as well as an intergrin inhibitor variant[93, 94]. Drastic changes, as observed for the integrin inhibitor substrate with 7 new residues in the decapeptide sequence, decrease the cyclization:hydrolysis ratio by reduction of productive substrate preorganization. Only two residues were identified to be specifically recognized by Tyc TE, D-Phe1 and Orn9 while changes in the rest of the peptide seemed to be tolerated[90]. Results of this study showed that Orn at position 9 could be exchanged by alanine under reconstitution of cyclization activity which shows that Tyc TE in contrast to Srf TE tolerates changes at the C-terminal end. Major limitations in cyclization ability of Tyc TE will therefore predominantly arise from changes at the N-terminal end and any alterations in the peptide sequence which lead to unproductive substrate preorganization or repulsion. Since problems in expression of soluble cyclosporin C-domain could not be solved, Tyc TE was a promising alternative possible candidate to facilitate cyclization of linear cyclosporin A. This

application would be very interesting in order to create libraries of cyclosporin A variants with improved or altered activities. Incubation with linear cyclosporin SNAC (**CLP**), which carries D-Ala instead of D-Phe at position 1, indicated only productive enzyme acylation but no release via cyclization. Therefore directed protein evolution was applied to change enzyme specificity of D-Phe1 towards D-Ala1. In order to minimize the number of mutant peptide products to screen by mass spectrometry, only the lid region was selected for random mutagenesis for several reasons. First, sequence alignments show very little conservation among TE domains in this region which could reflect individual optimisation for different substrates[96]. Second, Reetz and coworkers demonstrated that many mutations which increased the stereoselectivity in a related lipase were located in this lid-region[138]. The most evident support that the lid contributes to substrate binding was reported by Dugi et al., who showed that the lids of hepatic lipase and lipoprotein lipase are critical for the interaction with the lipid substrates[139]. The Srf TE crystal structure revealed that the lid surrounds the peptide binding pocket[96]. It can exist in an open and a closed conformation and in analogy to lipases it was suggested that the closed conformation represents the ground state without bound substrate. Together, these findings support the initial idea to concentrate random mutagenesis on one part of the enzyme which seems to be involved in substrate recognition.

The mass screen of about 1700 mutants revealed several hits, which did correlate with putative new enzyme species with increased **TLP-Ala** cyclization activity. DNA sequence analysis of 6 mutants showed that all of them carried stop codons within the lid region which led to shorter enzyme fragments. Most of the stop codons were derived from deletion mutants with corresponding frame shifts which occur during epPCR with a frequency of 1.1 % (Invitrogen). Since biochemical characterization of these purified stop fragments with **TLP** and **allyl-TLP** still revealed enzyme catalysed cyclization and hydrolysis activity, one can conclude that the lid region and subsequent parts of the enzyme are not essential for catalytic activity. This is specially intriguing since 30 – 44 % of the enzymes were missing in active stop fragments. Kinetic data of mutant JH1 revealed a drastic reduction in k_{cat} (1000 fold) but only very little deviation in K_M in comparison with wild type Tyc TE which suggests the relevance of the C-terminal enzyme part for catalytic efficiency. This is also supported by a drop in the cyclization:hydrolysis ratio from 6 to 0.8. Sequence alignments of several TEs show conserved regions for the N-terminal end up to the lid region. The lid region itself and the C-terminal part reveal high flexibility and almost no conserved regions which might explain why the short fragment mutants have been selected in this screen since they represent the shortest possible enzyme versions for productive cyclization (Fig. 6.6).

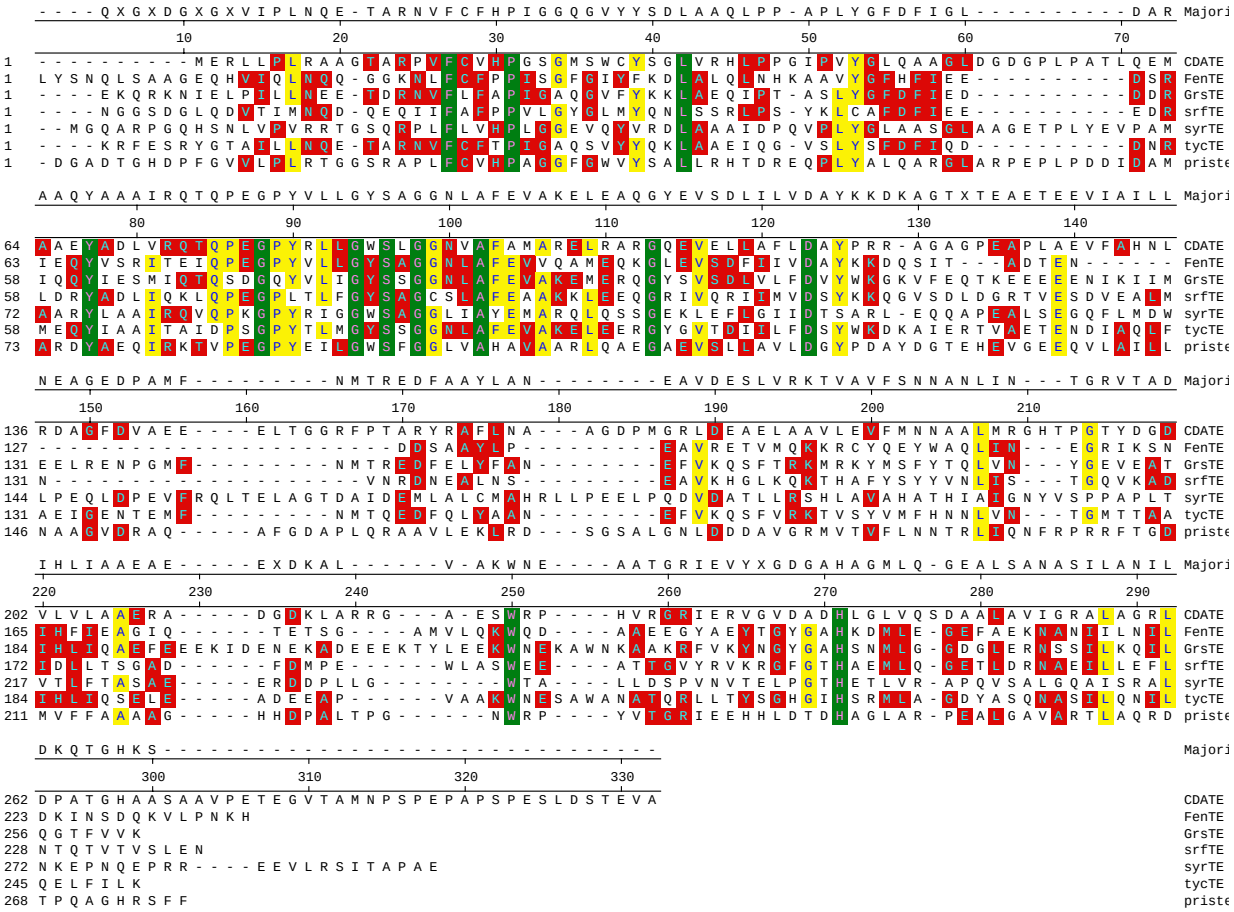


Figure 6.6. Sequence alignment of several TE domains. The C-terminal part of the enzymes show little homology, while the N-terminal part shows conserved regions. Also differences in length of the lid region can be observed.

It is also interesting to note that all mutant fragments lack the putative triad member His213 which may suggest that this residue is not essential for Tyc TE catalysis. Mutational exchange of the catalytic triad His to Ala in the serine protease subtilisin revealed drastic drop in activity but demonstrated still 1000 fold catalytic rate enhancement in comparison with the uncatalysed reaction[156]. This was explained by stabilization of the negative tetrahedral intermediate by the oxyanion hole similar to catalytic antibodies.

Previous research on hepatic lipase (HL) and lipoprotein lipase (LPL) suggested a crucial role of the lid region in substrate recognition for these TE related enzymes[139]. More detailed investigations revealed later that their substrate specificity can be altered. Characterization of chimeric enzymes in which LPL and HL lids have been exchanged demonstrated that the LPL core enzyme with the HL lid had an increased activity for the HL dedicated phospholipide substrates while HL with the LPL lid showed activity on triglyceride substrates which were only processed by LPL before[161]. This method could be applied to peptide cyclases and

would allow for directed protein engineering of Tyc TE by introduction of several lids or C-terminal fragments from other TE domains.

Since there is no crystal structure of Tyc TE available so far, the fragment JH1 was modeled by sequence homology into the crystal structures of Srf TE and Fen TE[162] (Fig. 6.7). Both structural models indicate a very rigid and stable N-terminal fold for the putative JH1 fragment which would explain why this part is sufficient for enzyme activity.

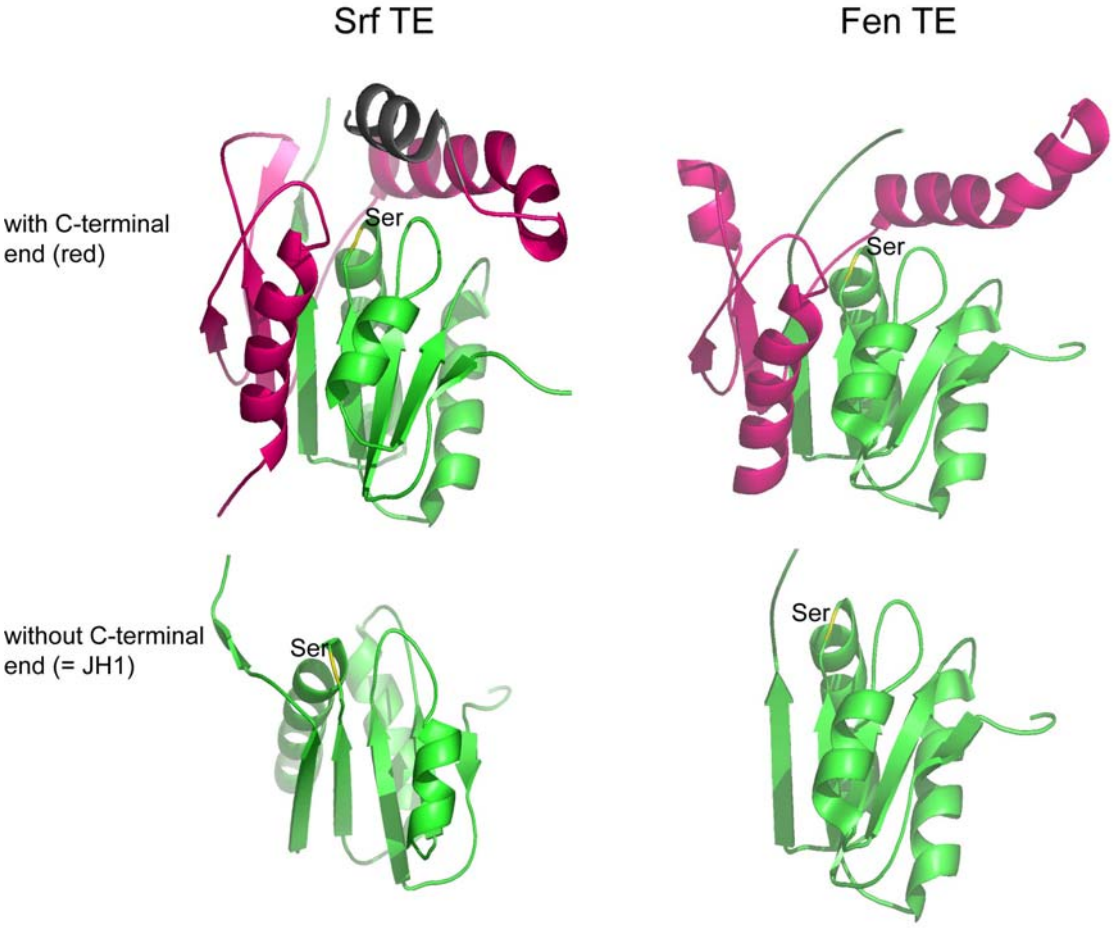


Figure 6.7. Structures of Srf TE and Fen TE to visualize JH1 fragment fold. Based on Srf TE and Fen TE[162] crystal structures, the short fragment was modeled into these structure based on sequence alignments. The green structure represents JH1 and the red part is the lid plus C-terminal regions which are missing in JH1.

However, although the selected hits demonstrated increased cyclization of **TLP-Ala** during the screen, no such behaviour could be detected with purified enzymes under standard assay conditions which raises the question why they were selected. All selected mutants revealed only 4 times the intensity of background cyclization during the screen which represents the lower detection limit. One possible answer could therefore be increased expression rates and better solubility of the shorter fragments in comparison with wild type Tyc TE, which is supported by the rigid and stable fold of the small fragment. This would lead to higher

enzyme concentrations during the screen and therefore explain the slightly increased substrate to product conversion. In order to obtain real hits with improved cyclization catalysis, a higher mutation rate in combination with other mutagenesis methods, such as DNA shuffling, cassette mutagenesis, could be applied[163].

Tyc TE and mutant fragment JH1 were both capable of cyclizing **TLP** and **allyl-TLP**, but surprisingly did not show activity for the N-terminal substrate variant **TLP-Ala**. Besides these similarities, a major difference was observed for the C-terminally modified substrates **Ala-Leu-TLP** and **Ala-Ala-TLP**. In contrast to Tyc TE, cyclization activity of the fragment JH1 was completely abolished for these substrates while trace amounts of hydrolysis were still observed. Substrate preorganization is limited in these two new tyrocidine variants since Orn9 is replaced by Ala which inhibits productive intramolecular H-bonds and complicates cyclization. This shows that the short C-terminal enzyme fragment is most likely involved in substrate recognition and coordination in the active site, especially if Orn9 is missing. Acylation and deacylation activity itself is probably contributed predominantly by the N-terminal region. This result would therefore encourage the exchange of C-terminal parts form other cyclases to explore new chimeric enzymes with putative altered substrate specificities as previously carried out with related lipases[161] (Fig. 6.8).

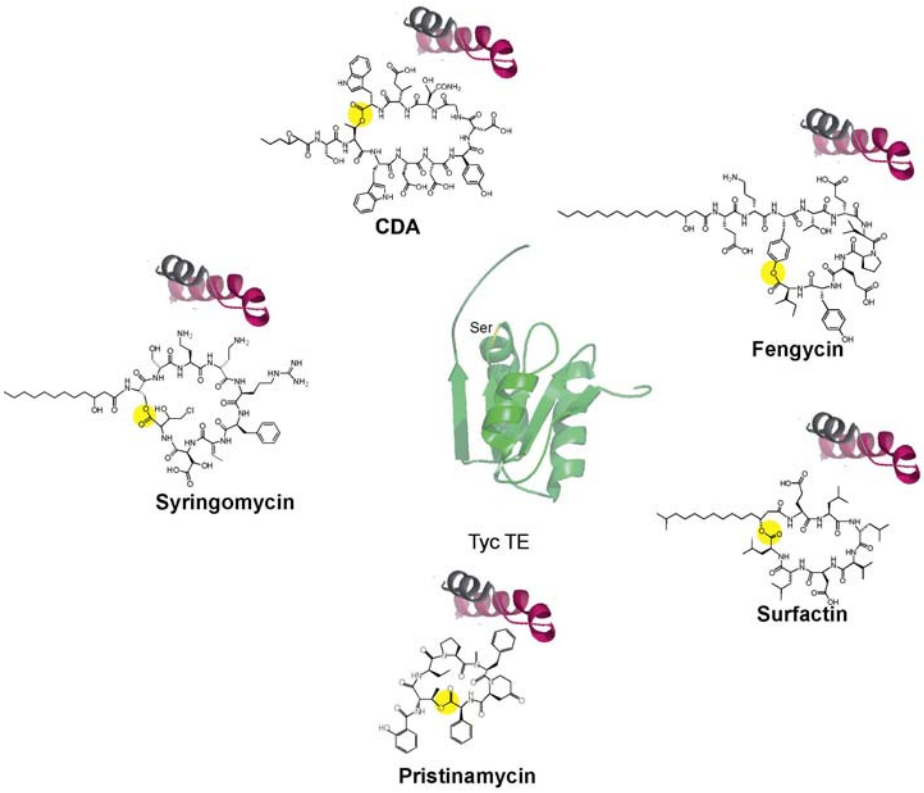


Figure 6.8. Engineering of hybrid TE domains by lid exchange. The lid of Tyc TE will be replaced by several lids from other TE domains.

6.2.5 General Utility of Peptide Cyclases: Chemical vs. Enzymatic Cyclization

In order to evaluate utility and potential of enzyme catalysed cyclization reactions, a comparison with established chemical methods is necessary. In principal, the chemical formation of macrocyclic rings is difficult, because of energetically disfavored ecliptic and transannular interactions[47]. If heteroatoms are present, as for example in macrocyclic natural products, the disfavored interactions are minimized[164] but ring closure proceeds still more slowly compared to normal peptide bond formations. In solution, only a few conformers have the right geometry to intramolecularly attack the C-terminal carboxy group. The entropic costs of populating these few productive conformations by several C-C-bond rotations are high and therefore disfavoured. To ensure regioselective cyclization, competing undesired nucleophiles such as hydroxy- or amino-groups have to remain protected, while the desired nucleophile needs to be deprotected which requires orthogonal protecting group strategies. Side reactions such as intermolecular peptide bond formation and subsequent cyclo-oligomerization may predominate since peptide bonds are usually trans-configured and favor a higher population of linear precursors. To minimize intermolecular reactions, high dilution conditions are applied with 10^{-4}-10^{-5} M solutions which make larger scale reactions difficult. Alternatively, cyclization of a peptide can also occur while it is still attached to the resin. This pseudo-dilution is a kinetic phenomenon and favors intramolecular reactions[121]. By contrast, turn inducing elements such as D-amino acids, proline, glycine, or N-alkyl amino acids also populate cis peptide bonds and are therefore easier to cyclize in solution. This emphazises that preorganization of the linear precursor by H-bonding is an important precondition for ring closure which was also observed for some peptide cyclases[91]. Another problem of chemical cyclization is the activation of the C-terminal carboxy-group without amino acid racemization. Coupling reagents like BOP and TBTU permit rapid cyclization but suffer from C-terminal racemization[121]. Less reactive coupling reagents minimize racemization but prolong reaction times. Better results were achieved with HOAt and DPPA[165-168]. Typically, with the above mentioned chemical cyclization methods 30-40 % yield of cyclic peptide products can be obtained[35, 164-166]. Reaction times range from several hours to days.

By contrast, enzymatically catalysed cyclization reactions do not require protecting groups and high dilution conditions due to enzymatic specificity. In literature, predominantly head to tail cyclizing enzymatic methods are reported. Cyclization of linear peptide esters was first reported for the subtilisin mutant subtiligase[169]. Subtiligase cyclizes peptide esters longer than 12 residues with yields of 30 - 88 % in a regioselective head to tail fashion. Hydrolysis

and dimerization are observed by-products. Head to tail cyclization without by-products was reported for an intramolecular version of an expressed protein ligation and split- intein approach which allow the generation of backbone cyclized peptides *in vitro* and *in vivo*[170]. In the work presented here, several enzymes such as surfactin TE, mycosubtilin TE, fengycin TE and syringomycin TE were shown to regiospecifically catalyse branched chain circularization between one dedicated nucleophile and the activated C-terminal peptide residue by discrimination of other unprotected competing nucleophiles. No oligomerization and C-terminal racemization were observed. The advantage of enzymatic vs. chemical cyclization illustrates the example of tyrocidine A synthesis. While chemical on resin cyclization reveals only 30 % yield, enzymatic cyclization yields 85 % product[90, 171]. However, the yield of enzymatically catalysed ring formation suffers from linear hydrolytic by-products due to competing nucleophilic attacks of water molecules. Typical yields of TE mediated cyclization range from 40-91 % and observed reaction times are several minutes to hours (Table 6.1).

Table 6.1. Cyclization yields and catalytic efficiency of several TE domains.

Enzyme	**Substrate**	**% Yield**	**k_{cat}/K_M ($min^{-1}mM^{-1}$)**
Tyc TE[90]	TLP-SNAC	85	19700
Srf TE	SLP3-SNAC	35	2.9
	SLP3-tp	44	44.9
Fen PCP-TE	FLP-CoA	67	N.A.
	FLP-tp	46	0.71
Myc PCP-TE	MLP-tp	46	0.54
Syr TE	SyLP-tp	40	11.4
CDA TE[157]	CDA-tp	91	29.8
SnbDE PCP-TE[172]	PLP-SNAC	67	0.77

Since these recombinant peptide cyclases are usually embedded in a hydrophobic multienzyme complex, their unprotected modification as a single TE or PCP-TE may enhance the exposure of the active site to water. Moreover, substrate analogues used in *in vitro* studies lack for the ease of synthesis some structural features like long chain fatty acyl chains which are important for a perfect fit into the enzyme binding pocket as seen for CDA[157]. This is also supported by *in vivo* studies with a genetically engineered surfactin synthetase lacking

module 2[158]. While it was not possible to detect hydrolysed surfactin product in the supernatant of the wild type producer cells, hydrolysis was observed for the shorter surfactin variant produced by the engineered strain. This indicates that hydrolysis can also occur *in vivo*, where the TE domain is embedded in a functional multienzyme complex. Therefore the presentation of the dedicated substrate seems to be a key determant to minimize hydrolysis by TE domains. Moreover, artificial thioester leaving groups like SNAC and thiophenol are short mimics of the natural cofactor ppan-PCP and may therefore be more suceptible for the entry of water into the active site rather than blocking it. The nucleophilic side reaction by water can further be eliminated by enzyme catalysis in organic solvents[173]. The high enzymatic selectivity for cognate substrates can also be a disadvantage when the cyclization of substrate analogues is desired. Changes of residues specially at the C- and N-terminal ends of the peptide sequence decrease or completely abolish cyclization yields and are subject of protein evolution. With the current set of active peptide cyclases described here, diverse cyclization reactions such as head to tail and branched chain lactamization and lactonization of various substrates can already be performed.

In comparison to organic synthesis the chemoenzymatic approach presented here minimizes time and side reactions and maximizes purity and yield of cyclic peptide production. Biochemical prerequisites such as enzyme characterization and leaving group technology have been solved which will now allow for applications such as formation of various cyclic peptide libraries[95]. In future, further exploration of structurally diverse peptide cyclases will help to generate a tool set of cyclization catalysts which in combination with protein evolution are likely to broaden the applications and increase the utility of enzymatic peptide cyclization similar to the use of lipases in stereochemistry today.

7 Literature

1. Flemming, A., *On the antibacterial action of cultures of a penicillinium, with special reference to their use in the isolation of B. influenzae.* Brit. J. Exp. Path., 1929. **10**: p. 226-236.
2. Lipmann, F., et al., *Polypeptide synthesis on protein templates: the enzymatic synthesis of gramicidin S and tyrocidine.* Adv. Enzymol. Relat. Areas Mol. Biol., 1971. **35**: p. 1-34.
3. Walsh, C.T., *Antibiotics: Actions, Origins, Resistance*. 2003: ASM Press, Washington DC. 335.
4. Schwarzer, D., Finking, R., and Marahiel, M. A., *Nonribosomal peptides: from genes to products.* Nat. Prod. Rep., 2003. **20**(3): p. 275-287.
5. Mootz, H.D., D. Schwarzer, and M.A. Marahiel, *Ways of assembling complex natural products on modular nonribosomal peptide synthetases.* Chembiochem, 2002. **3**(6): p. 490-504.
6. Marahiel, M.A., T. Stachelhaus, and H.D. Mootz, *Modular peptide synthetases involved in non-ribosomal peptide synthesis.* Chem. Rev., 1997. **97**(7): p. 2651-2673.
7. Zasloff, M., *Antimicrobial peptides of multicellular organisms.* Nature, 2002. **415**: p. 389-395.
8. Mootz, H.D. and M.A. Marahiel, *The tyrocidine biosynthesis operon of Bacillus brevis: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains.* J Bacteriol, 1997. **179**(21): p. 6843-50.
9. Simlot, M.M., P. Pfaender, and D. Specht, *Antibiotics producing enzymes of Bacillus licheniformis.* Hoppe Seylers Z Physiol Chem, 1972. **353**(5): p. 759.
10. Peypoux, F., J.M. Bonmatin, and J. Wallach, *Recent trends in the biochemistry of surfactin.* Appl Microbiol Biotechnol, 1999. **51**(5): p. 553-63.
11. de Crecy-Lagard, V., et al., *Pristinamycin I biosynthesis in Streptomyces pristinaespiralis: molecular characterization of the first two structural peptide synthetase genes.* J Bacteriol, 1997. **179**(3): p. 705-13.
12. Hubbard, B.K. and C.T. Walsh, *Vancomycin Assembly: Nature's Way.* Angew Chem Int Ed Engl, 2003. **42**(7): p. 730-65.
13. Steller, S., et al., *Structural and functional organization of the fengycin synthetase multienzyme system from Bacillus subtilis b213 and A1/3 [published erratum appears in Chem Biol 1999 May;6(5):R156].* Chem Biol, 1999. **6**(1): p. 31-41.
14. Steller, S. and J. Vater, *Purification of the fengycin synthetase multienzyme system from Bacillus subtilis b213.* J Chromatogr B Biomed Sci Appl, 2000. **737**(1-2): p. 267-75.
15. Vanittanakom, N., et al., *Fengycin--a novel antifungal lipopeptide antibiotic produced by Bacillus subtilis F-29-3.* J Antibiot (Tokyo), 1986. **39**(7): p. 888-901.
16. Rieber, M.T., T. Imaeda, and I.M. Cesari, *Bacitracin action on membranes of Mycobacteria.* J. Gen. Microbiol., 1969. **55**: p. 155-159.
17. Konz, D., et al., *The bacitracin biosynthesis operon of Bacillus licheniformis ATCC 10716: molecular characterization of three multi-modular peptide synthetases.* Chem Biol, 1997. **4**(12): p. 927-37.
18. Hori, K., et al., *Molecular cloning and nucleotide sequence of the gramicidin S synthetase 1 gene.* J Biochem (Tokyo), 1989. **106**(4): p. 639-45.
19. Katsu, T., H. Kobayashi, and Y. Fujita, *Mode of action of gramicidin S on Escherichia coli membrane.* Biochim Biophys Acta, 1986. **860**(3): p. 608-19.

20. Stone, K.J. and J.L. Strominger, *Mechanism of action of bacitracin: complexation with metal ions and C55-isoprenyl pyrophosphate.* Proc. Natl. Acad. Sci. USA, 1971. **68**: p. 3223-3227.
21. Storm, D.R. and J.L. Strominger, *Complex formation between bacitracin peptides and isoprenyl pyrophosphates.* J. Biol. Chem., 1973. **248**(11): p. 3940-3945.
22. Maget-Dana, R., et al., *Surfactin/iturin A interactions may explain the synergistic effect of surfactin on the biological properties of iturin A.* Biochimie, 1992. **74**(12): p. 1047-51.
23. Duitman, E.H., et al., *The mycosubtilin synthetase of Bacillus subtilis ATCC6633: a multifunctional hybrid between a peptide synthetase, an amino transferase, and a fatty acid synthase.* Proc Natl Acad Sci U S A, 1999. **96**(23): p. 13294-9.
24. Maget-Dana, R. and F. Peypoux, *Iturins, a special class of pore-forming lipopeptides: biological and physicochemical properties.* Toxicology, 1994. **87**(1-3): p. 151-174.
25. Hutchison, M.L. and D.C. Gross, *Lipopeptide phytotoxins produced by Pseudomonas syringae pv. syringae: comparison of the biosurfactant and ion channel-forming activities of syringopeptin and syringomycin.* Mol Plant Microbe Interact, 1997. **10**(3): p. 347-54.
26. Scholz-Schroeder, B.K., Hutchison, M. L., Grgurina, I., and Gross, D. C., *The Contribution of Syringopeptin and Syringomycin to Virolence of Pseudomonas syringae pv. syringae strain B301D on the Basis of sypA and syrB1 Biosynthesis Mutant Analysis.* MPMI, 2001. **14**(3): p. 336-348.
27. Hojati, Z., Milne, C., Harvey, B., Gordon, L., Borg, M., Flett, F., Wilkinson, B., Sidebottom, P. J., Rudd, B. A. M., Hayes, M. A., Smith, C. P., Micklefield, J., *Structure, Biosynthetic Origin, and Engineered Biosynthesis of Calcium-Dependent Antibiotics from Streptomyces coelicolor.* Chem. Biol., 2002. **9**: p. 1175-1187.
28. Alborn, W.E., Jr., N.E. Allen, and D.A. Preston, *Daptomycin disrupts membrane potential in growing Staphylococcus aureus.* Antimicrob Agents Chemother, 1991. **35**(11): p. 2282-7.
29. McHenney, M.A., et al., *Molecular cloning and physical mapping of the daptomycin gene cluster from Streptomyces roseosporus.* J Bacteriol, 1998. **180**(1): p. 143-51.
30. Canepari, P. and M. Boaretti, *Lipoteichoic acid as a target for antimicrobial action.* Microb Drug Resist, 1996. **2**(1): p. 85-9.
31. Di Giambattista, M., Engelborghs, Y., Nyssen, E., Clays, K., and Cocito, C., *Interaction between virginiamycin S and ribosomes is partly provided by a salt bridge with a Mg2+ ion.* Biochemistry, 1991. **30**: p. 7277-7282.
32. Weber, G., et al., *The peptide synthetase catalyzing cyclosporine production in Tolypocladium niveum is encoded by a giant 45.8-kilobase open reading frame.* Curr. Genet., 1994. **26**(2): p. 120-125.
33. Belshaw, P.J., Schreiber, S. L., *Cell-Specific Calcineurin Inhibition by a Modified Cyclosporin.* J. Am. Chem. Soc., 1997. **119**: p. 1805-1806.
34. Liu, J., Farmer, J. D. Jr., Lane, W. S., Friedman, J., Weissman, I.,and Schreiber, S. L., *Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes.* Cell, 1991. **66**: p. 807-815.
35. Thern, B., Rudolph, J., and Jung, G., *Total Synthesis of the Nematicidal Cyclodecapeptide Omphalotin A by Using Racemization-Free Triphosgene-Mediated Couplings in the Solid Phase.* Angew. Chem. Int. Ed., 2002. **41**(13): p. 2307-2309.
36. Quadri, L.E., *Assembly of aryl-capped siderophores by modular peptide synthetases and polyketide synthases.* Mol Microbiol, 2000. **37**: p. 1-12.
37. Crosa, J.H.a.W., C. T., *Genetics and assembly line enzymology of siderophore biosynthesis in bacteria.* Microbiol. Mol. Biol. Rev., 2002. **66**: p. 223-249.

38. Sewald, N., *Efficient, Racemization-Free Peptide Coupling of N-Alkyl Amino Acids by Using Amono Acid Chlorides Generated In Situ - Total Synthesis of the Cyclopeptides Cyclosporin O and Omphalotin A.* Angew. Chem. Int. Ed., 2002. **41**(24): p. 4661-4663.
39. Nicolaou, K.C., Boddy, C. N. C., Brase, S., Winssinger, N., *Chemistry, Biology and Medical Applications of Glycopeptide Antibiotics.* Angew. Chem. Int. Ed., 1999. **38**: p. 2096-2152.
40. Lipmann, F., *Bacterial production of antibiotic polypeptides by thiol-linked synthesis on protein templates.* Adv. Microb. Physiol., 1980. **21**: p. 227-66.
41. von Döhren, H., et al., *Multifunctional Peptide Synthetases.* Chem. Rev., 1997. **97**: p. 2675-2705.
42. Cosmina, P., et al., *Sequence and analysis of the genetic locus responsible for surfactin synthesis in Bacillus subtilis.* Mol Microbiol, 1993. **8**(5): p. 821-31.
43. Schwarzer, D. and M.A. Marahiel, *Multimodular biocatalysts for natural product assembly.* Naturwissenschaften, 2001. **88**(3): p. 93-101.
44. Cane, D.E. and C.T. Walsh, *The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases.* Chem Biol, 1999. **6**(12): p. R319-25.
45. Stryer, L., *Biochemistry*. 4 ed. 1996, Heidelberg-Berlin-Oxford: Spektrum Akademischer Verlag GmbH.
46. Linne, U., and Marahiel, M. A., *Reaction catalysed by mature and recombinant nonribosomal peptide synthetases.* Methods in Enzymology, 2004. **in press**.
47. Brückner, R., *Reaktionsmechanismen*. 1996, Heidelberg: Spektrum Verlag.
48. Stein, T., et al., *Detection of 4'-phosphopantetheine at the thioester binding site for L-valine of gramicidinS synthetase 2.* FEBS Lett, 1994. **340**(1-2): p. 39-44.
49. Stein, T., et al., *The multiple carrier model of nonribosomal peptide biosynthesis at modular multienzymatic templates.* J Biol Chem, 1996. **271**(26): p. 15428-35.
50. Stachelhaus, T., A. Hüser, and M.A. Marahiel, *Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases.* Chem. Biol., 1996. **3**(11): p. 913-921.
51. Schlumbohm, W., et al., *An active serine is involved in covalent substrate amino acid binding at each reaction center of gramicidin S synthetase.* J Biol Chem, 1991. **266**(34): p. 23135-41.
52. Lambalot, R.H., et al., *A new enzyme superfamily - the phosphopantetheinyl transferases.* Chem Biol, 1996. **3**(11): p. 923-36.
53. Walsh, C.T., et al., *Post-translational modification of polyketide and nonribosomal peptide synthetases.* Curr. Opin. Chem. Biol., 1997. **1**: p. 309-315.
54. Gocht, M. and M.A. Marahiel, *Analysis of core sequences in the D-Phe activating domain of the multifunctional peptide synthetase TycA by site-directed mutagenesis.* J Bacteriol, 1994. **176**(9): p. 2654-62.
55. Quadri, L.E., et al., *Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases.* Biochemistry, 1998. **37**(6): p. 1585-95.
56. Kealey, J.T., Liu, L., Santi, D. V., Betlach, M. C., and Barr, P. J., *Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts.* Proc. Natl. Acad. Sci., 1998. **95**: p. 505-509.
57. Belshaw, P.J., C.T. Walsh, and T. Stachelhaus, *Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis.* Science, 1999. **284**(5413): p. 486-9.
58. Chen, H., et al., *Epothilone biosynthesis: assembly of the methylthiazolylcarboxy starter unit on the EpoB subunit.* Chem Biol, 2001. **8**(9): p. 899-912.

59. Sieber, S.A., Walsh, C. T., and Marahiel, M. A., *Loading Peptidyl-Coenzyme A onto Peptidyl Carrier Proteins: A Novel Approach in Characterizing Macrocyclization by Thioesterase Domains.* J. Am. Chem. Soc., 2003. **125**: p. 10862-10866.
60. Clugston, S.L., Sieber, S. A., Marahiel, M. A., and Walsh, C. T., *Chirality of Ppeptide Bond-Forming Condensation Domains in Nonribosomal Peptide Synthases: The C5 Domain of Tyrocidine Synthetase is a DCL Catalyst.* Biochemistry, 2003. **42**: p. 12095-12104.
61. Reuter, K., et al., *Crystal structure of the surfactin synthetase-activating enzyme Sfp: a prototype of the 4´-phosphopantetheinyl transferase superfamily.* Embo J., 1999. **18**: p. 6823-6831.
62. Schneider, A. and M.A. Marahiel, *Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in non-ribosomal peptide biosynthesis in Bacillus subtilis.* Arch Microbiol, 1998. **169**(5): p. 404-10.
63. Keating, T.A., et al., *The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains.* Nat Struct Biol, 2002. **9**(7): p. 522-6.
64. De Crecy-Lagard, V., P. Marliere, and W. Saurin, *Multienzymatic non ribosomal peptide biosynthesis: identification of the functional domains catalysing peptide elongation and epimerisation.* C R Acad Sci III, 1995. **318**(9): p. 927-36.
65. Bergendahl, V., U. Linne, and M.A. Marahiel, *Mutational analysis of the C-domain in nonribosomal peptide synthesis.* Eur J Biochem, 2002. **269**(2): p. 620-9.
66. Roche, E.D., and Walsh, C. T., *Dissection of the EntF Condensation Domain Boundary and Active Site Residues in Nonribosomal Peptide Synthesis.* Biochemistry, 2002. **42**: p. 1334-1344.
67. Pfeifer, E., et al., *Characterization of tyrocidine synthetase 1 (TY1): requirement of posttranslational modification for peptide biosynthesis.* Biochemistry, 1995. **34**(22): p. 7450-9.
68. Stachelhaus, T. and C.T. Walsh, *Mutational analysis of the epimerization domain in the initiation module PheATE of gramicidin S synthetase.* Biochemistry, 2000. **39**(19): p. 5775-87.
69. Luo, L. and C.T. Walsh, *Kinetic analysis of three activated phenylalanyl intermediates generated by the initiation module PheATE of gramicidin S synthetase.* Biochemistry, 2001. **40**(18): p. 5329-37.
70. Hoffmann, K., et al., *Purification and characterization of eucaryotic alanine racemase acting as key enzyme in cyclosporin biosynthesis.* J Biol Chem, 1994. **269**(17): p. 12710-4.
71. Linne, U. and M.A. Marahiel, *Control of directionality in nonribosomal peptide synthesis: role of the condensation domain in preventing misinitiation and timing of epimerization.* Biochemistry, 2000. **39**(34): p. 10439-47.
72. Linne, U., S. Doekel, and M.A. Marahiel, *Portability of epimerization domain and role of peptidyl carrier protein on epimerization activity in nonribosomal peptide synthetases.* Biochemistry, 2001. **40**(51): p. 15824-34.
73. Velkov, T., and Lawen, A., *Nonribosomal peptide synthetases as technological platforms for the synthesis of highly modified peptide bioeffectors - cyclosporin synthetase as a complex example.* Biotechnol. Annu. Rev., 2003. **9**: p. 151-197.
74. Billich, A., et al., *Monoclonal antibodies to the multienzyme enniatin synthetase. Production and use in structural studies.* Biol Chem Hoppe Seyler, 1987. **368**(5): p. 521-9.
75. Schauwecker, F., et al., *Construction and in vitro analysis of a new bi-modular polypeptide synthetase for synthesis of N-methylated acyl peptides.* Chem Biol, 2000. **7**(4): p. 287-97.

76. de Crecy-Lagard, V., et al., *Streptogramin B biosynthesis in Streptomyces pristinaespiralis and Streptomyces virginiae: molecular characterization of the last structural peptide synthetase gene.* Antimicrob Agents Chemother, 1997. **41**(9): p. 1904-9.
77. Ehmann, D.E., A.M. Gehring, and C.T. Walsh, *Lysine biosynthesis in Saccharomyces cerevisiae: mechanism of alpha-aminoadipate reductase (Lys2) involves posttranslational phosphopantetheinylation by Lys5.* Biochemistry, 1999. **38**(19): p. 6171-7.
78. Kessler, N., Schuhmann, H., Morneweg, S., Linne, U., and Marahiel, M. A., *The linear pentadecapeptide gramicidin is assembled by four multimodulare nonribosomal peptide synthetases that comprise 16 modules with 56 catalytic domains.* J Biol Chem, 2004. **in press**.
79. Tseng, C.C., et al., *Characterization of the surfactin synthetase C-terminal thioesterase domain as a cyclic depsipeptide synthase.* Biochemistry, 2002. **41**(45): p. 13350-9.
80. Kohli, R.M., et al., *Generality of peptide cyclization catalyzed by isolated thioesterase domains of nonribosomal peptide synthetases.* Biochemistry, 2001. **40**(24): p. 7099-108.
81. Segre, A., Bachmann, R. C., Ballio, A., Bossa, F., Grgurina, I., Iacobellis, N. S., Marino, G., Pucci, P., Simmaco, M., and Takemoto, J. Y., *The structure of syringomycins A1, E and G.* FEBS Lett, 1989. **255**: p. 27-31.
82. Weber, G. and E. Leitner, *Disruption of the cyclosporin synthetase gene of Tolypocladium niveum.* Curr Genet, 1994. **26**(5-6): p. 461-7.
83. Becker, J.E., R.E. Moore, and B.S. Moore, *Cloning, sequencing, and biochemical characterization of the nostocyclopeptide biosynthetic gene cluster: molecular basis for imine macrocyclization.* Gene, 2004. **325**: p. 35-42.
84. Gehring, A.M., I. Mori, and C.T. Walsh, *Reconstitution and characterization of the Escherichia coli enterobactin synthetase from EntB, EntE, and EntF.* Biochemistry, 1998. **37**(8): p. 2648-59.
85. May, J.J., T.M. Wendrich, and M.A. Marahiel, *The dhb operon of Bacillus subtilis encodes the biosynthetic template for the catecholic siderophore 2,3-dihydroxybenzoate-glycine-threonine trimeric ester bacillibactin.* J Biol Chem, 2001. **276**(10): p. 7209-17.
86. Pieper, R., et al., *Arrangement of catalytic sites in the multifunctional enzyme enniatin synthetase.* Eur J Biochem, 1995. **230**(1): p. 119-26.
87. Haese, A., et al., *Molecular characterization of the enniatin synthetase gene encoding a multifunctional enzyme catalysing N-methyldepsipeptide formation in Fusarium scirpi.* Mol Microbiol, 1993. **7**(6): p. 905-14.
88. Scholz-Schroeder, B.K., Soule, J. D., and Gross, D. C., *The sypA, sypS, and sypC synthetase genes encode twenty-two modules involved in the non-ribosomal peptide synthesis of syringopeptin by Pseudomonas syringae pv. syringae strain B301D.* Mol. Plant-Microbe Interact., 2003. **16**: p. 271-280.
89. Craik, D.J., Daly, N. L., Saska, I., Trabi, M., and Rosengren, K. J., *Structures of naturally occuring circular proteins from bacteria.* J. Bacteriol., 2003. **185**: p. 4011-4021.
90. Trauger, J., et al., *Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase.* Nature, 2000. **407**(6801): p. 215-8.
91. Trauger, J.W., R.M. Kohli, and C.T. Walsh, *Cyclization of backbone-substituted peptides catalyzed by the thioesterase domain from the tyrocidine nonribosomal peptide synthetase.* Biochemistry, 2001. **40**(24): p. 7092-8.

92. Bu, X., Wu, X., Xie G. and Guo, Z., *Synthesis of tyrocidine A and its analogues by spontaneous cyclization in aqueous solution.* Organic Letters, 2002. **4**(17): p. 2893-2895.
93. Kohli, R.M., Takagi, J., and Walsh, C. T., *The thioesterase domain from a nonribosomal peptide synthetase as a cyclization catalyst for integrin binding peptides.* PNAS, 2002. **99**(3): p. 1247-1252.
94. Kohli, R.M., Burke, M. D., Tao, J., and Walsh, C. T., *Chemoenzymatic Route to Hybrid Peptide/Polyketide-like Molecules.* J. Am. Chem. Soc., 2003. **125**: p. 7160-7161.
95. Kohli, R.M., C.T. Walsh, and M.D. Burkart, *Biomimetic synthesis and optimization of cyclic peptide antibiotics.* Nature, 2002. **418**(6898): p. 658-61.
96. Bruner, S.D., et al., *Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE.* Structure (Camb), 2002. **10**(3): p. 301-10.
97. Shaw-Reid, C.A., et al., *Assembly line enzymology by multimodular nonribosomal peptide synthetases: the thioesterase domain of E. coli EntF catalyzes both elongation and cyclolactonization.* Chem Biol, 1999. **6**(6): p. 385-400.
98. Witkowski, A., A. Joshi, and S. Smith, *Fatty acid synthase: in vitro complementation of inactive mutants.* Biochemistry, 1996. **35**(32): p. 10569-75.
99. Kao, C.M., et al., *Evidence for two catalytically independent clusters of active sites in a functional modular polyketide synthase.* Biochemistry, 1996. **35**(38): p. 12363-8.
100. Staunton, J., et al., *Evidence for a double-helical structure for modular polyketide synthases.* Nat Struct Biol, 1996. **3**(2): p. 188-92.
101. Silakowski, B., et al., *New lessons for combinatorial biosynthesis from myxobacteria. The myxothiazol biosynthetic gene cluster of Stigmatella aurantiaca DW4/3-1.* J Biol Chem, 1999. **274**(52): p. 37391-9.
102. Quadri, L.E., et al., *Identification of a Mycobacterium tuberculosis gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin.* Chem Biol, 1998. **5**(11): p. 631-45.
103. Pelludat, C., et al., *The yersiniabactin biosynthetic gene cluster of Yersinia enterocolitica: organization and siderophore-dependent regulation.* J. Bacteriol., 1998. **180**: p. 538-546.
104. Weber, T. and M.A. Marahiel, *Exploring the domain structure of modular nonribosomal peptide synthetases.* Structure (Camb), 2001. **9**(1): p. R3-9.
105. Conti, E., et al., *Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S.* Embo J, 1997. **16**(14): p. 4174-83.
106. Weber, T., et al., *Solution structure of PCP, a prototype for the peptidyl carrier domains of modular peptide synthetases.* Structure Fold Des, 2000. **8**(4): p. 407-18.
107. Stoops, J.K. and S.J. Wakil, *Animal fatty acid synthetase. A novel arrangement of the beta-ketoacyl synthetase sites comprising domains of the two subunits.* J Biol Chem, 1981. **256**(10): p. 5128-33.
108. Stoops, J.K. and S.J. Wakil, *Animal fatty acid synthetase. Identification of the residues comprising the novel arrangement of the beta-ketoacyl synthetase site and their role in its cold inactivation.* J Biol Chem, 1982. **257**(6): p. 3230-5.
109. Joshi, A.K., A. Witkowski, and S. Smith, *Mapping of functional interactions between domains of the animal fatty acid synthase by mutant complementation in vitro.* Biochemistry, 1997. **36**(8): p. 2316-22.
110. Gokhale, R.S., et al., *Functional orientation of the acyltransferase domain in a module of the erythromycin polyketide synthase.* Biochemistry, 1998. **37**(8): p. 2524-8.
111. Joshi, A.K., Vangipuram, S. R., Witkowski, A., and Smith, S., *Engineering of an Active Animal Fatty Acid Synthase Dimer with Only One Competent Subunit.* Chem. Biol., 2003. **10**: p. 169-173.

112. Rangan, V.S., A.K. Joshi, and S. Smith, *Fatty acid synthase dimers containing catalytically active beta- ketoacyl synthase or malonyl/acetyltransferase domains in only one subunit can support fatty acid synthesis at the acyl carrier protein domains of both subunits.* J Biol Chem, 1998. **273**(52): p. 34949-53.
113. Witkowski, A., et al., *Dibromopropanone cross-linking of the phosphopantetheine and active- site cysteine thiols of the animal fatty acid synthase can occur both inter- and intrasubunit. Reevaluation of the side-by-side, antiparallel subunit model.* J Biol Chem, 1999. **274**(17): p. 11557-63.
114. Schwecke, T., et al., *Enzymatic characterisation of the multifunctional enzyme delta-(L-alpha- aminoadipyl)-L-cysteinyl-D-valine synthetase from Streptomyces clavuligerus.* Eur J Biochem, 1992. **205**(2): p. 687-94.
115. Konig, A., et al., *The pipecolate-incorporating enzyme for the biosynthesis of the immunosuppressant rapamycin--nucleotide sequence analysis, disruption and heterologous expression of rapP from Streptomyces hygroscopicus.* Eur J Biochem, 1997. **247**(2): p. 526-34.
116. Sambrook, J., E.F. Fritsch, and T. Maniatis, *Molecular Cloning: A Laboratory Manual*. 1989, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
117. Schwarzer, D., et al., *Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases.* Proc Natl Acad Sci U S A, 2002. **99**(22): p. 14083-8.
118. Lambert, L.J., et al., *Flipping a genetic switch by subunit exchange.* Embo J, 2001. **20**(24): p. 7149-59.
119. Laue, T.M., et al., *Computer-aided interpretation of analytical sedimentation data for proteins*, in *Analytical ultracentrifugation in biochemistry and polymer science*, S.E. Harding, A.J. Rowe, and J.C. Horton, Editors. 1992, Royal Society of Chemistry: Cambridge, UK. p. 90-125.
120. Cohn, E.J. and J.T. Edsall, *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions*. 1943, New York: Reinhold.
121. Chan, W.C., and White P. D., *Fmoc Solid Phase Peptide Synthesis*, ed. B.D. Hames. 2000, Oxford: Oxford University Press.
122. Eberle, M.K., Jutzi-Eme, A. M., and Nuninger, F., *Cyclosporin A: Regioselective Ring Opening and Fragmentation Reactions via Thioamides. A Route to Semisynthetic Cyclosporins.* J. Org. Chem., 1994. **59**: p. 7249-7258.
123. Stachelhaus, T., et al., *Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain.* J Biol Chem, 1998. **273**(35): p. 22773-81.
124. Mootz, H.D., D. Schwarzer, and M.A. Marahiel, *Construction of hybrid peptide synthetases by module and domain fusions.* Proc Natl Acad Sci U S A, 2000. **97**(11): p. 5848-53.
125. Rusnak, F., et al., *Biosynthesis of the Escherichia coli siderophore enterobactin: sequence of the entF gene, expression and purification of EntF, and analysis of covalent phosphopantetheine.* Biochemistry, 1991. **30**(11): p. 2916-27.
126. Sieber, S.A., *Untersuchungen zur Quartärstruktur von nichtribosomalen Peptidsynthetasen*, in *Department of Chemistry*. 2001, Philipps-Universität: Marburg. p. 121.
127. Pieper, R., et al., *Purification and characterization of bimodular and trimodular derivatives of the erythromycin polyketide synthase.* Biochemistry, 1997. **36**(7): p. 1846-51.
128. Joshi, A.K., V.S. Rangan, and S. Smith, *Differential affinity labeling of the two subunits of the homodimeric animal fatty acid synthase allows isolation of heterodimers consisting of subunits that have been independently modified.* J Biol Chem, 1998. **273**(9): p. 4937-43.

129. Dingley, A.J., et al., *Measuring protein self-association using pulsed-field-gradient NMR spectroscopy: application to myosin light chain 2.* J Biomol NMR, 1995. **6**(3): p. 321-8.
130. Osman, M., Ishigami, Y., Ishikawa, K., Ishizuka, Y., and Holmsen, H., *Dynamic Transition of α-Helix to β-Sheet Structure in Linear Surfactin Correlating to Critical Micelle Concentration.* Biotechnol. Lett., 1994. **16**: p. 913-918.
131. Wong, H., and Schotz, M. C., *The lipase gene family.* J. Lipid Res., 2002. **43**: p. 993-999.
132. Stoll, V.S., Eger, B. T., Hynes, R. C., Martichonok, V., Jones, J. B., and Pai, E. F., *Differences in binding modes of enantiomers of 1-acetamido boronic acid based protease inhibitors: crystal structures of gamma-chymotrypsin and subtilisin Carlsberg complexes.* Biochemistry, 1998. **37**: p. 451-462.
133. Takahashi, L.H., Radhakrishnan, R., Rosenfield, R. E., Jr., and Meyer, E. F., Jr., *Crystallographic analysis of the inhibition of porcine pancreatic elastase by a peptidyl boronic acid: structure of a reaction intermediate.* Biochemistry, 1989. **28**: p. 7610-7617.
134. Muir, T.W., Sondhi, D. and Cole, P. A., *Expressed protein ligation: A general method for protein engineering.* Proc. Natl. Acad. Sci., 1998. **95**: p. 6705-6710.
135. Dawson, P.E., Churchill, M. J., Ghadiri, M. R. and Kent, S. B. H., *Modulation of Reactivity in Native Chemical Ligation through the Use of Thiol Additives.* J. Am. Chem. Soc., 1996. **119**(19): p. 4325-4329.
136. Guenzi, E., et al., *Characterization of the syringomycin synthetase gene cluster. A link between prokaryotic and eukaryotic peptide synthetases.* J Biol Chem, 1998. **273**(49): p. 32857-63.
137. Sieber, S.A., Tao, J., Walsh, C. T., and Marahiel, M. A., *Peptidyl Thiophenols as Substrates for Nonribosomal Peptide Cyclases.* Angew. Chem., 2003. **116**: p. 449-504.
138. Liebeton, K., Zonta, A., Schimossek, K., Nardini, M., Lang, D., Dijkstra, B. W., Reetz, M. T., Jaeger, K. E., *Directed evolution of an enantioselective lipase.* Chem. Biol., 2000. **7**(9): p. 709-718.
139. Dugi, K.A., Dichek, H. L., Talley, G. D., Brewer, H. B., Jr., and Sanatamarina-Fojo, S., *Human lipoprotein lipase: the loop covering the catalytic site is essential for interaction with lipid substrates.* J Biol Chem, 1992. **267**: p. 25086-25091.
140. Jaeger, K.E., Eggert, T., Eipper, A., and Reetz, M. T., *Directed evolution and the creation of enantioselective biocatalysts.* Appl. Microbiol. Biotechnol., 2001. **55**: p. 519-530.
141. Bollag, D.M., et al., *.Epothilones, a new class of microtubule-stabilizing agents with a taxol-like mechanism of action.* Cancer Res., 1995. **55**: p. 2325-2333.
142. Molnar, I., et al., *The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from Sorangium cellulosum So ce90.* Chem Biol, 2000. **7**(2): p. 97-109.
143. Tang, L., et al., *Cloning and heterologous expression of the epothilone gene cluster.* Science, 2000. **287**(5453): p. 640-2.
144. Lottspeich, F. and H. Zorbas, *Bioanalytik*. 1998, Heidelberg, Berlin: Spektrum Akademischer Verlag.
145. Konig, A., et al., *The pipecolate-incorporating enzyme for the biosynthesis of the immunosuppressant rapamycin--nucleotide sequence analysis, disruption and heterologous expression of rapP from Streptomyces hygroscopicus.* Eur J Biochem, 1997. **247**(2): p. 526-34.

146. Schmidt, B., et al., *Cyclosporin synthetase is a 1.4 MDa multienzyme polypeptide. Re-evaluation of the molecular mass of various peptide synthetases.* FEBS Lett, 1992. **307**(3): p. 355-60.
147. Perham, R.N., *Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions.* Annu Rev Biochem, 2000. **69**: p. 961-1004.
148. Tsai, S.C., et al., *Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: versatility from a unique substrate channel.* Proc Natl Acad Sci U S A, 2001. **98**(26): p. 14808-13.
149. Hoppert, M., C. Gentzsch, and K. Schorgendorfer, *Structure and localization of cyclosporin synthetase, the key enzyme of cyclosporin biosynthesis in Tolypocladium inflatum.* Arch Microbiol, 2001. **176**(4): p. 285-93.
150. Hillson, N.J., and Walsh, C. T., *Dimeric Structure of the six-Domain VibF Subunit of Vibriobactin Synthetase: Mutant Domain Activity Regain and Ultracentrifugation Studies.* Biochemistry, 2003. **42**: p. 766-775.
151. Kiho, T., Nakayama, M., Yasuda, K., Miyakoshi, S., Inukai, M., and Kogen, H., *Synthesis and antimicrobial activity of novel globomycin analogues.* Bioorg Med Chem Lett., 2003. **13**(14): p. 2315-2318.
152. Hashizume, H., Igarashi, M., Hattori, S., Hori, M., Hamada, M., and Takeuchi, T., *Tripropeptins, novel antimicrobial agents produced by Lysobacter sp. I. Taxonomy, isolation and biological activities.* J. Antibiot. (Tokyo), 2001. **54**(12): p. 1054-1059.
153. Tsuge, K., Akiyama, T., Shoda, M., *Cloning, sequencing, and characterization of the iturin A operon.* J. Bacteriol., 2001. **183**(21): p. 6265-6273.
154. Kraut, J., *Serine proteases: structure and mechanism of catalysis.* Annu. Rev. Biochem., 1977. **46**: p. 331-358.
155. Pazirandeh, M., S.S. Chirala, and S.J. Wakil, *Site-directed mutagenesis studies on the recombinant thioesterase domain of chicken fatty acid synthase expressed in Escherichia coli.* J. Biol. Chem., 1991. **266**(31): p. 20946-20952.
156. Carter, P., and Wells, J. A., *Dissecting the catalytic triad of a serine protease.* Nature, 1988. **332**: p. 564-568.
157. Grünewald, J., Sieber, S. A., and Marahiel, M. A., *Chemo- and Regioselective Peptide Cyclization triggered by the N-terminal Fatty Acid Chain Length: The Recombinant Cyclase of the Calcium-dependent Antibiotic from Streptomyces coelicolor.* Biochemistry, 2004. **in press**.
158. Mootz, H.D., et al., *Decreasing the ring size of a cyclic nonribosomal Peptide antibiotic by in-frame module deletion in the biosynthetic genes.* J Am Chem Soc, 2002. **124**(37): p. 10980-1.
159. Bonmatin, J.M., et al., *Solution three-dimensional structure of surfactin: a cyclic lipopeptide studied by 1H-NMR, distance geometry, and molecular dynamics.* Biopolymers, 1994. **34**(7): p. 975-86.
160. Silverman, R.B., *The Organic Chemistry of Enzyme-Catalyzed Reactions*. 2000, San Diego, San Francisco, New York, Boston, London, Sydney, Tokyo: Academic Press.
161. Dugi, K.A., Dichek, H. L., and Santamarina-Fojo, S., *Human hepatic and lipoprotein lipase: The loop covering the catalytic site mediates lipase substrate specificity.* J Biol Chem, 1995. **270**(43): p. 25396-25401.
162. Samel, S., *Kristallisation der Fengycin Thioesterase*, in *Biochemie*. 2004, Philipps Universität: Marburg.
163. Reetz, M.T., Wilensek, S., Zha, D., and Jaeger, K. E., *Directed Evolution of an Enantioselective Enzyme through Combinatorial Multiple-Cassette Mutagenesis.* Angew. Chem. Int. Ed., 2001. **40**(19): p. 3589-3591.

164. Sewald, N., and Jakubke, *Peptides: Chemistry and Biology*. 2002, Weinheim: Wiley-VCH.
165. Zhu, J., and Ma, D., *Total synthesis of Microsclerodermin E.* Angew. Chem., 2003. **43**: p. 5506-5509.
166. Davies, J.S., *The Cyclization of Peptides and Depsipeptides.* J. Peptide Sci., 2003. **9**: p. 471-501.
167. Schmidt, R., and Neubert, K., *Cyclization studies with tetra- and pentapeptide sequences corresponding to beta-casomorphins.* Int. J. Peptide Protein Res., 1991. **37**: p. 502-507.
168. Ehrlich, A., Heyne, H.-U., Winter, R., Beyermann, M., Haber, H., Carpino, L. A., and Bienert, M., *Cyclization of all L-pentapeptides by means of 1-hydroxy-7-azabenzotriazole-derived uronium and phosphonium reagents.* J. Org. Chem., 1996. **63**: p. 8831-8838.
169. Jackson, D.Y., Burnier, J. P., and Wells, J. A., *Enzymatic cyclization of linear peptide esters using subtiligase.* J Am Chem Soc, 1995. **117**: p. 819-820.
170. Muir, T.W., *Semisynthesis of proteins by expressed protein ligation.* Annu Rev Biochem, 2003. **72**: p. 249-289.
171. Nishino, N., Xu, M., Mihara, H., Fujimoto, T., Ueno, Y., and Kumagai, H., Tetrahedron Lett, 1992. **33**: p. 1479.
172. Mahlert, C., *Untersuchungen zur enzymatischen Zyklisierung von Pristinamycin*, in *Chemistry/Biochemistry*. 2004, Philipps-University: Marburg.
173. Klibanov, A.M., *Improving enzymes by using them in organic solvents.* Nature, 2001. **409**: p. 241-246.

Name	Peptide sequence
SLP wt	FA-Glu-Leu-D-Leu-Val-Asp-D-Leu-Leu
SLP-1	FA-Dap-Leu-D-Leu-Val-Asp-D-Leu-Leu
SLP-2	FA-Glu-Dap-D-Leu-Val-Asp-D-Leu-Leu
SLP-3	FA-Glu-Leu-D-Dap-Val-Asp-D-Leu-Leu
SLP-4	FA-Glu-Leu-D-Leu-Dap-Asp-D-Leu-Leu
SLP-5	FA-Glu-Leu-D-Leu-Val-Dap-D-Leu-Leu
SLP-6	FA-Glu-Leu-D-Leu-Val-Asp-D-Dap-Leu
SLP-7	FA-Glu-Leu-D-Leu-Val-Asp-D-Leu-Dap
SLP-1/2	FA-ε-Ahx -D-Leu-Val-Asp-D-Leu-Leu
SLP-3/4	FA-Glu-Leu-ε-Ahx -Asp-D-Leu-Leu
SLP-5/6	FA-Glu-Leu-D-Leu-Val-ε-Ahx -Leu
MLP	FA-Asn-D-Tyr-D-Asn-Gln-Pro-D-Ser-Ala
FLP	FA-Glu-D-Orn-Tyr-D-Thr-Glu-D-Ala-Pro-Glu-D-Tyr-Ile
FLP(Phe)	FA-Glu-D-Orn-Phe-D-Thr-Glu-D-Ala-Pro-Glu-D-Tyr-Ile
FLP(2)	FA-Glu-D-Tyr-Orn -D-Thr-Glu-D-Ala-Pro-Glu-D-Tyr-Ile
FLP(4)	FA-Glu-D-Orn-Thr-D-Tyr-Glu-D-Ala-Pro-Glu-D-Tyr-Ile
SyLP	FA-Ser-D-Ser-Dab-D-Dab-Arg-Phe-Abu-Asp-Thr
TLP	D-Phe-Pro-Phe-D-Phe-Asn-Gln-Tyr-Val-Orn-Leu
allyl-TLP	D-Phe-Pro-allyl-Gly-D-Phe-Asn-Gln-Tyr-Val-Orn-Leu
Ala-Leu-TLP	D-Phe-Pro-Phe-D-Phe-Asn-Gln-Tyr-Val-Ala-Leu
Ala-Ala-tlp	D-Phe-Pro-Phe-D-Phe-Asn-Gln-Tyr-Val-Ala-Ala
TLP-DAla	D-Ala-Pro-Phe-D-Phe-Asn-Gln-Tyr-Val-Orn-Leu

Acknowledgements

I would like to thank Prof. Marahiel for his excellent scientific guidance and generous support during my diploma and doctoral thesis in his lab. During the course of my studies he provided not only important scientific advice and education but also encouraged interesting and helpful discussions. I would like to thank him for his scientific dedication which made work in his lab very exciting. He always supported new ideas and helped to develop them further which influenced my way of scientific thinking. I am in particular grateful for the generous opportunity to spend one year of my doctoral research in the lab of Prof. Christopher Walsh at Harvard Medical School in Boston. I would like to thank Chris for the very interesting time in his lab. In this time I learned many scientific techniques which were not only essential for my thesis but will also determine my future research. I am grateful for his excellent scientific advise and support and most important for the inspiration and spirit in his lab. I would like to thank him for being in my thesis committee and for providing an expert opinion.
I would like to thank my both supervisors for their excellent cooperation and the many opportunities to exchange scientific expertise. I would be very grateful for their support in my future.

I would like to thank Prof. Schrader and Prof. Klebe for being in my thesis committee.

I would also like to thank all members of the Marahiel lab and junior labs for their generous help and the nice atmosphere inside and outside the lab. I would like to thank Dr. Uwe Linne for fruitful collaborations and help with the structure and evolution projects, Jan Grünewald for his motivation and commitment on TE projects, Verena Senn and Antje Schäfer for great work on the evolution project, Christina Dauth for excellent support of the C-domain project and Björn Wagner, Christoph Mahlert, Georg Schönafinger and Daniel Garbe for interesting research discussions. Special thanks also to Dr. Jürgen May for helping me many times with computer problems. I would like to thank Prof. Lars O. Essen for providing crystal structure models of JH1.
I would like to thank people from the Walsh lab for their kind help and support. They made my start in Boston easy and helped me to feel comfortable in a new environment. I would like to thank Claire Tseng for excellent cooperation with the Srf TE project, Natahn Hillson for help and cooperation with the ultracentrifugation, Steve Bruner and Rahul Kohli for advice and help with many scientific problems, and Junhua Tao for collaboration with chemical synthesis. I would also like to thank these people for having a nice time outside the lab.

I would like to thank Dr. Uwe Linne, Dr. Henning Mootz, Nathan Hillson and Jan Grünewald for proof reading the manuscript of the thesis.

I would like to acknowledge the “Studienstiftung des deutschen Volkes” for financial support during my thesis and specially I would like to thank my “Vertrauensdozent” Prof. Melsheimer for many interesting discussions. I would also like to thank the Graduiertenkolleg “Proteinfunktion auf atomarer Ebene” for general support and the many opportunities to present my research to an international audience.

The biggest support came from my parents. They were my first mentors and without all their help in many difficult situations I would have not been able to accomplish this thesis. For their dedication to me, I wish to dedicate this thesis to them.