

**Structural and Functional Investigation of the  
Selectivity in Interactions of Peptidyl Carrier Proteins  
from *Bacillus***

Dissertation  
zur Erlangung des Doktorgrades  
der Naturwissenschaften

vorgelegt beim Fachbereich  
Biochemie, Chemie und Pharmazie (FB 14)  
der Johann Wolfgang Goethe-Universität  
in Frankfurt am Main

von  
Peter Tufar  
aus Marburg/Lahn

Frankfurt am Main, 2014  
(D 30)

Vom Fachbereich Biochemie, Chemie und Pharmazie (FB 14) der  
Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Thomas Prisner

Gutachter: Prof. Dr. Volker Dötsch

Prof. Dr. Mohamed A. Marahiel (Philipps-Universität Marburg)

Datum der Disputation: 09. April 2014

## Eidesstattliche Versicherung

Ich erkläre hiermit an Eides Statt, dass ich die vorgelegte Dissertation über "*Structural and Functional Investigation of the Selectivity in Interactions of Peptidyl Carrier Proteins from Bacillus*" selbständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe, insbesondere, dass alle Entlehnungen aus anderen Schriften mit Angabe der betreffenden Schrift gekennzeichnet sind.

Ich versichere, nicht die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen zu haben.

Frankfurt am Main, den 08. Januar 2014

(Peter Tufar)

*“– und es war alles, alles gut!”*  
(Joseph Freiherr von Eichendorff, 1788-1857)

## Table of Contents

Abbreviations.....	V
Summary.....	X
Zusammenfassung.....	XIII
1. Introduction.....	1
1.1 Structure and Function of Nonribosomal Peptide Synthetases.....	1
1.1.1 The Peptidyl Carrier Protein.....	3
1.1.2 The Adenylation Domain.....	6
1.1.3 The Condensation Domain.....	9
1.1.4 Thioesterases.....	12
1.1.5 Tailoring Domains and Enzymes.....	14
1.2 Non-PCP Carrier Proteins.....	16
1.3 Phosphopantetheine Transferases.....	17
1.4 Inter-domain Interactions in NRPSs.....	20
1.4.1 Domain/Domain Interactions.....	20
1.4.2 Communication-mediating Domains.....	24
2. Materials.....	26
2.1 Equipment.....	26
2.2 Chemicals and Reagents.....	28
2.3 Primers.....	32
2.3.1 Primers for PCR.....	32
2.3.2 Primers for Site Directed Mutagenesis.....	32
2.4 Common Media and Buffers.....	33
2.4.1 Media and Solutions for Cultivation of Microorganisms.....	33
2.4.2 Buffers for Protein Purification.....	36
2.4.3 Buffers for DNA Preparation.....	38
2.4.4 Buffers for Polyacrylamide Gel Electrophoresis.....	38
2.5 Enzymes.....	40
2.6 Bacterial strains.....	40
2.7 Plasmids.....	41

---

2.8 Nucleic Acids Purification Kits.....	42
2.9 Crystallization Kits.....	42
2.10 Software and Online-Tools.....	42
3. Methods.....	44
3.1 DNA Techniques.....	44
3.1.1 Polymerase Chain Reaction.....	44
3.1.2 Analytical and Preparative Agarose Gel Electrophoresis.....	44
3.1.3 PCR Purification.....	44
3.1.4 Restriction Digestion.....	45
3.1.5 Dephosphorylation of DNA.....	45
3.1.6 Site Directed Mutagenesis.....	45
3.1.7 DNA Ligation.....	45
3.1.8 Transformation of Ligated DNA or DNA from Site Directed Mutagenesis.....	46
3.1.9 Transformation of Purified DNA.....	46
3.1.10 Preparation of Cryogenic Cultures.....	46
3.1.11 Plasmid DNA Preparation.....	47
3.1.12 Quantification of DNA.....	47
3.1.13 DNA Sequencing.....	47
3.2 Protein Expression and Purification.....	47
3.2.1 Expression of Proteins in Rich Media.....	47
3.2.2 Expression of Proteins in M9 Minimal Medium.....	48
3.2.3 Cell Free Expression of Proteins.....	48
3.2.4 Cell Lysis.....	49
3.2.5 Immobilized Metal Ion Affinity Purification.....	49
3.2.6 Reversed Immobilized Metal Ion Affinity Purification.....	49
3.2.7 Size Exclusion Chromatography.....	50
3.2.8 Dialysis.....	51
3.2.9 Proteolytic Cleavage with TEV Protease.....	51
3.2.10 Concentration of Proteins.....	52
3.2.11 Polyacrylamide Gel Electrophoresis.....	52
3.2.12 Determination of Protein Concentration.....	53
3.2.13 Protein Analysis by MALDI-TOF.....	53

3.3 Synthesis of CoA Derivatives.....	53
3.3.1 Synthesis of Dephospho Amino-CoA.....	53
3.3.2 Coupling of Peptides with Amino-CoA.....	55
3.3.3 Coupling of Peptides and Amino Acids with CoA.....	56
3.3.4 Modification of PCP Domains with Ppan Derivatives.....	56
3.4 Biochemical and Biophysical Methods.....	57
3.4.1 NMR Titration experiments.....	57
3.4.2 NMR Measurements for Structure Determination.....	58
3.4.3 NMR Assignment and Structure Calculation.....	58
3.4.4 Crystallization Trials.....	59
3.4.5 X-ray Diffraction and Crystal Structure Determination.....	60
3.4.6 Phosphopantetheinylation Assay.....	60
3.4.7 Circular Dichroism Spectroscopy.....	60
3.4.8 Isothermal Titration Calorimetry.....	61
3.4.9 Analysis of PCPs by Liquid Chromatography-Mass Spectrometry.....	61
4. Results.....	63
4.1 Recognition of PCP-bound Peptides.....	63
4.1.1 Protein Expression and Purification.....	63
4.1.2 Stability of Acylated PCPs and Activity of TycC4_C.....	66
4.1.3 Comparison of Amide- and Thioester-bound Peptidyl-PCP.....	67
4.1.4 Structure of Peptidyl(NH)-TycC3_PCP in Solution.....	68
4.1.5 Crystal Structure of holo-TycC5-6_PCP-C.....	72
4.2 Structural and Functional Basis for Phosphopantetheinylation of PCP.....	76
4.2.1 Crystal Structure of the PCP/Sfp Complex.....	76
4.2.2 NMR Investigation of the Active Site Mutant PCP in Solution.....	87
4.2.3 Production of Mutant TycC3_PCP and Sfp.....	94
4.2.4 Impact of Mutations on the Stability of the PCP/Sfp Complex.....	98
4.2.5 Velocity of the PCP Priming Reaction for Different Mutants.....	99
5. Discussion and Outlook.....	100
5.1 Interactions in the Peptide Bond Formation.....	100
5.1.1 Activity of TycC4_C.....	100
5.1.2 The Structure of peptidyl(NH)-TycC3_PCP.....	101

---

5.1.3 Structure of holo-TycC5-6_PCP-C.....	103
5.2 Phosphopantetheinylation of PCPs.....	104
5.2.1 Structural Flexibility of Sfp.....	104
5.2.2 Mechanism of the Phosphopantetheine Transfer by Sfp.....	105
5.2.3 Carrier Protein Recognition by Sfp.....	108
6. References.....	114
7. Appendix.....	136
7.1 Protein Sequences and Properties.....	136
7.2 Resonance Assignment of peptidyl-TycC3_PCP.....	138
7.3 Results of Side Projects.....	143
7.3.1 Cyclization Domain.....	143
7.3.2 In vivo Expression of Peptides.....	144
Curriculum Vitae.....	145
Acknowledgments.....	148



## Abbreviations

°	Degree
1D	One-dimensional
<sup>1</sup> H	Hydrogen-1
<sup>13</sup> C	Carbon-13
<sup>15</sup> N	Nitrogen-15
2xYT	Two times yeast extract and tryptone
μ	Micro
ρ	Density
Å	Angstrom
A	Absorption
AA	Amino acid
ACP	Acyl carrier protein (domain)
AcpS	<i>holo</i> -acyl carrier protein synthase
A domain	Adenylation domain
AMP	Adenosine monophosphate
APS	Ammonium persulfate
ArCP	Aryl carrier protein (domain)
ATP	Adenosine triphosphate
<i>B. brevis</i>	<i>Bacillus brevis</i>
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
Bac	Bacitracin A synthetase
cm	Centimeter
C	Celsius
CD	Circular dichroism
C domain	Condensation domain
CoA	Coenzyme A
COM domain	Communication-mediating domain
CS(P)	Chemical shift (perturbation)
CV	Column volume
Cy domain	Cyclization domain
Da	Dalton
ddH <sub>2</sub> O	Ultrapure water
DIPEA	<i>N,N</i> -diisopropylethylamine

---

DMF	<i>N,N</i> -dimethylformamide
DNA	Deoxyribonucleic acid
DPCCK	Dephosphocoenzyme A kinase
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
E domain	Epimerization domain
eq	Equivalents
EtOH	Ethanol
FA	Fatty acid
FAS	Fatty acid synthase
g	Gram
* <i>g</i>	Gravitational acceleration
Grs	Gramicidin S synthetase
h	Hour
HBTU	1H-benzotriazole- <i>N,N,N',N'</i> -tetramethyl-uronium hexafluorophosphate
HOBt	1-Hydroxybenzotriazole
HPLC	High performance liquid chromatography
HSQC	Hetero spin quantum correlation
Hz	Hertz
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
k	Kilo
K	Kelvin
$K_b$	Dissociation constant
l	Length
L	Liter
LB	Lysogeny broth
LC-MS	Liquid chromatography-mass spectrometry
m	Meter
M	Molar
MALDI-TOF	Matrix assisted laser desorption/ionization-time of flight
mAU	Milli absorption units
Mbt	Mycobactin synthetase
MeCN	Acetonitrile
MHz	Megahertz
min	Minute

mL	Milliliter
mM	Millimolar
mol	Mole
MS	Mass spectrometry
MWCO	Molecular weight cutoff
n	Amount of substance
N	Number of theoretical binding sites
NADPH	Nicotine adenine dinucleotide phosphate
nm	Nanometer
NMR	Nuclear magnetic resonance (spectroscopy)
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
NRP	Nonribosomal peptide
NRPS	Nonribosomal peptide synthetase
NS	Number of scans
OD <sub>600</sub>	Optical density at 600 nm
Orn	Ornithine
PAGE	Polyacrylamide gel electrophoresis
PanK	Pantothenate kinase
PCP	Peptidyl carrier protein (domain)
PCR	Polymerase chain reaction
PDB	Protein database (accession code)
PEG	Polyethylene glycol
pH	Negative decimal logarithm of the proton concentration
P <sub>i</sub>	Phosphate
PKS	Polyketide synthase
ppan	Phosphopantetheine
PPAT	Phosphopantetheine adenylyltransferase
ppm	parts per million
PPT	Phosphopantetheine transferase
PyBOP	Benzotriazole-1-yl-oxytrypyrrolidinophosphonium hexafluorophosphate
RMSD	Route-mean-square deviation
rpm	Revolutions per minute
RT	Room temperature
s	Second
SDS	Sodium dodecyl sulfate

---

SEC	Size exclusion chromatography
Sfp	PPT from the surfactin A gene cluster
SrfA	Surfactin A synthetase
T	Temperature
TCA	Trichloroacetic acid
TCEP	<i>Tris</i> (2-carboxyethyl)phosphine
TE I/II	Thioesterase type I/II
TEMED	<i>N,N,N',N'</i> -tetramethylethane-1,2-diamine
TEV	Tobacco etch virus
TFA	Trifluoroacetic acid
TIPS	Triisopropylsilane
T <sub>M</sub>	Melting temperature
TOCSY	Total correlation spectroscopy
tRNA	Transfer ribonucleic acid
TRICINE	<i>N</i> -(Tri(hydroxymethyl)methyl)glycine
Tyc	Tyrocidine A synthetase
UV	Ultraviolet
V	Volt
v/v	Volume per volume
Vib	Vibriobactin synthetase
W	Watt
w/v	Weight per volume
wt	Wild type

For amino acids the standard single and three letter codes were used.

Regarding nuclear magnetic resonance (NMR) spectroscopy, the atoms and groups of amino acids were named in accordance with the standard nomenclature of CYANA. An exception was made for diastereotopic atoms and groups. For diastereotopic atoms, the atom with the larger chemical shift and for diastereotopic groups, the group with the downfield carbon were assigned with the smaller index number.

Domains and multi-domain fragments of nonribosomal peptide synthetases (NRPS) were denominated by the name of the synthetase's subunit, the number(s) of the module(s) within the subunit and the name of the domain(s). For the borders of a modules, the definition was used that they start with the N-terminus of the condensation domain (C domain).

For example, TycC3\_PCP is the peptidyl carrier protein (PCP) of the third module of the Tyrocidine A synthetase's subunit C (TycC), and TycC5-6\_PCP-C is a bidomain construct of the PCP from the fifth and the C domain from the sixth module of TycC.

## Summary

Nonribosomal peptide synthetases (NRPSs), often found in bacteria and fungi, are a source for many bioactive secondary metabolites. Considering the growing number of pathogenic strains with resistance to known antibiotics, nonribosomal peptides (NRPs) with antimicrobial properties are of special interest. The high diversity of NRPs results from the numerous building blocks, which can be incorporated into an NRP. The number of building blocks exceeds the 20 proteinogenic amino acids by far, and further diversity is gained by modifications like epimerization, methylation, and the formation of macrocyclic structures.

An NRPS consists of several modules, each harboring the catalytic domains which are responsible for the activation and incorporation of a certain amino acid into the growing peptide chain. In addition to the catalytic domains, so-called peptidyl carrier protein (PCP) domains are present in these modules. PCP domains show no enzymatic activity, but they are responsible for the transport of the intermediates between the different catalytic centers. In most cases, only substrates bound to PCPs via the thiol group of a prosthetic phosphopantetheine (ppan) moiety are accepted by the catalytic domains.

Beside the search for naturally occurring peptides, the artificial redesign of NRPSs has recently become a promising approach to produce new NRPs. However, the swapping of domains or modules in order to produce peptides with an altered sequence often results in very low yields. Usually, the most likely explanation for the drop in productivity is an intrinsic selectivity in the domains that is not compatible with their rearrangement. Therefore, it is necessary to understand this selectivity, not only with regard to the protein/protein interactions, but also with respect to the recognition of the PCP-bound intermediates by the different domains.

In the presented work, a structure-based approach with subsequent biochemical analysis has been made to get further insights into interactions during the maturation of the NRP tyrocidine A from *Bacillus brevis*.

In the first part of this work, the peptide bond formation was examined. In this reaction, condensation domains (C domains) catalyze the transfer of a peptide bound to a PCP onto an amino acid bound to another PCP. For this investigation, the well-characterized TycC3\_PCP was chosen as the peptide-loaded PCP. Together with peptidyl-TycC3\_PCP, the C domain TycC4\_C and valyl-TycC4\_PCP, which are its natural interaction partners, were used. All three proteins could be produced as homogeneous samples. However, biochemical assays indicated that TycC4\_C was not catalytically active. Therefore, the recognition of the PCP-bound peptide by the C domain could not be investigated.

Instead, the interaction between PCP and peptide in peptidyl-TycC3\_PCP was examined. In a first step, the stability of the thioester bond, which links the peptide to the ppan-arm, was analyzed. This analysis revealed that the hydrolysis was slow enough to record two-dimensional nuclear magnetic resonance (NMR) spectra of a  $^{15}\text{N}$ -labeled sample. Yet, the thioester bond was not stable enough to record a set of multi-dimensional NMR spectra, which is required for structure determination. Comparison of NMR spectra indicated that peptidyl-TycC3\_PCP does not change its structure when the thioester bond is substituted by an amide bond. Thus, the structure of the peptide-loaded PCP was determined using the non-hydrolyzable peptide. In this structure, the PCP adopts the A/H state conformation, with the ppan-arm only in loose contact to helix 2 and 3 of the protein. Contacts between protein and peptide were not observed, indicating that the peptide is solvent-exposed and not selectively recognized by the PCP.

Another attempt to get structural information on the recognition of a PCP-bound peptide by a C domain was made using the bidomain TycC5-6\_PCP-C, which had previously been shown to be catalytically active. However, the bidomain could not be loaded with the non-hydrolyzable peptidyl-ppan, as already a significant fraction of the protein was loaded with ppan during the expression. On the other hand, quantitative loading with ppan *in vitro* was successful, and the bidomain in its *holo*-form was crystallized. Yet, the overall structure of *holo*-TycC5-6\_PCP-C was the same like the one previously solved for its *apo*-form, in which the two domains are not in the orientation required for their interaction.

In the second part of this work, the modification of PCPs with ppan was investigated. This reaction is catalyzed by phosphopantetheine transferases (PPTs), which use coenzyme A (CoA) as substrate. The crystal structure of Sfp, a group II PPT from *Bacillus subtilis*, in complex with CoA and the active site S45A mutant of TycC3\_PCP was solved. In this structure, two interaction sites between the two proteins are apparent: TycC3\_PCP(S45A) forms a hydrogen bond to the N-terminal domain of Sfp, and its second helix forms hydrophobic contacts to the C-terminal domain.

These interactions were further investigated in biochemical assays. Mutation of L46 (to alanine, asparagine, or aspartate) and M49 (to aspartate) in helix 2 of the PCP, harboring also the active site mutation, weakened the interaction of the two proteins to a great extent. All these mutants showed no binding to Sfp in isothermal titration calorimetry (ITC) experiments at 5°C and 25°C, whereas the binding of TycC3\_PCP(S45A) to Sfp had a dissociation constant ( $K_D$ ) below 1  $\mu$ M at both temperatures. Introduction of the same mutations in active TycC3\_PCP reduced the velocity of the ppan transfer reaction in a priming assay. In this context, the reduction of the velocity of the reaction was stronger when polar and charged residues were introduced, and the effect of mutation of L46 appeared to be more severe. On the other hand, disruption of the intermolecular hydrogen bond resulted in a faster ppan transfer, although it weakened the binding of the two proteins ( $K_D \sim 50 \mu$ M at 5°C).

In the structure of the complex, TycC3\_PCP(S45A) is in the A/H state conformation, which is not in accordance with a previous model. Nevertheless, determination of the structure of free TycC3\_PCP(S45A) in solution by NMR spectroscopy and NMR titration experiments confirmed that TycC3\_PCP(S45A) adopts this conformation in its free form as well as in complex with Sfp.



## Zusammenfassung

### ***Strukturelle und funktionale Untersuchung der Selektivität in Interaktionen von Peptidyl Carrier Proteinen aus Bacillus***

Nichtribosomale Peptid Synthetasen, die häufig in Bakterien und Pilzen zu finden sind, sind Quelle für eine Vielzahl von Sekundärmetaboliten mit biologischer Aktivität. Angesichts der in den letzten Jahrzehnten auftretenden Zunahme von Krankheitserregern, die gegen standardmäßig verwendete Antibiotika Resistenzen ausgebildet haben, sind besonders jene nichtribosomal gebildeten Peptide mit antibiotischer Wirkung von besonderem Interesse. Die hohe strukturelle und funktionale Diversität dieser Peptide ergibt sich aus der Vielzahl ihrer Bausteine. Die Anzahl dieser Bausteine umfasst weit mehr als die 20 proteinogenen Aminosäuren, und eine weitere Vielfalt wird durch Modifikationen wie Epimerisierungen, Methylierungen und die Bildung makrozyklischer Strukturen erreicht.

Die Synthetasen bestehen in der Regel aus mehreren Modulen, wobei jedes Modul die katalytischen Domänen für die Aktivierung und den Einbau einer bestimmten Aminosäure während der Bildung des Peptids enthält. Außerdem gibt es in jedem Modul eine *Peptidyl Carrier Protein* (PCP) Domäne, die selber nicht katalytisch aktiv ist, sondern die wachsende Peptidkette zu den verschiedenen katalytischen Zentren transportiert. In den meisten Fällen werden von den katalytischen Domänen nur solche Substrate akzeptiert, die an die Thiolgruppe eines kovalent mit der PCP Domäne verbundenen Phosphopantetheinrests (Ppan) gebunden sind.

Neben der Suche nach neuen, natürlich vorkommenden Peptiden ist die künstliche Veränderung von bereits vorhandenen Synthetasen ein immer häufiger verfolgter Ansatz für die Gewinnung neuer nichtribosomal gebildeter Peptide. Der Austausch einzelner Domänen oder ganzer Module, um Peptide mit einer veränderten Sequenz zu erhalten, geht jedoch häufig damit einher, dass die Ausbeute des veränderten Peptids nur sehr gering ist. In der Regel ist die naheliegendste Erklärung für diesen Verlust an Produktivität eine den Domänen innewohnende Selektivität, die mit dem

veränderten Aufbau der Synthetase nicht vereinbar ist. Für die erfolgreiche Veränderung der Synthetasen ist es daher von essentieller Bedeutung, diese Selektivitäten sowohl in Bezug auf Protein/Protein Wechselwirkungen zwischen den Domänen als auch in Hinsicht auf die Wechselwirkung der einzelnen Domänen mit den PCP-gebundenen Zwischenstufen der wachsenden Peptidkette zu verstehen.

In der vorliegenden Arbeit wurde zur Untersuchung solcher Selektivitäten ein Struktur-basierter Ansatz mit nachfolgender biochemischer Analyse gewählt. Als Modellsysteme dienten hierbei Fragmente aus der Tyrocidin A Synthetase aus *Bacillus brevis* und Sfp, die mit der Surfactin A Synthetase assoziierte Gruppe II Phosphopantethein Transferase (PPT) aus *Bacillus subtilis*.

Im ersten Teil der Arbeit wurde die Peptidbindungsbildung untersucht. In Nichtribosomalen Peptid Synthetasen wird diese Reaktion durch Kondensationsdomänen (C Domänen) bewirkt, welche die Verknüpfung eines Peptids, das an eine PCP Domäne gebunden ist, mit einer Aminosäure, die an einer anderen PCP Domäne hängt, katalysieren. Als Peptid-beladene PCP Domäne wurde TycC3\_PCP, die dritte PCP Domäne aus der Untereinheit C der Tyrocidin A Synthetase verwendet, deren natürliche Interaktionspartner in der Kondensationsreaktion die C Domäne TycC4\_C und die Valin-beladene TycC4\_PCP sind. Von allen drei Proteinen war es möglich, homogene Proben durch heterologe Expression in *Escherichia coli* und anschließender chromatographischer Aufreinigung herzustellen, und auch die quantitative Beladung der PCP Domänen mit den gewünschten Substraten *in vitro* war erfolgreich. Biochemische Untersuchungen zeigten jedoch, dass TycC4\_C im isolierten Zustand keine katalytische Aktivität aufwies, weswegen auf weitere Untersuchungen zur Erkennung des PCP-gebundenen Peptids durch die C Domäne vorerst verzichtet wurde.

Stattdessen wurde eine mögliche Interaktion zwischen der PCP Domäne und des an sie gebundenen Peptids untersucht. Tests zur spontanen Hydrolyse der Thioesterbindung, durch welche Peptide an die PCP Domäne gebunden sind, mittels gekoppelter Flüssigchromatographie/Massenspektrometrie ergaben für peptidyl-TycC3\_PCP, dass die Geschwindigkeit der Hydrolyse langsam genug war, um

Untersuchungen mittels zweidimensionaler Kernspinresonanzspektroskopie (NMR) vorzunehmen. Die Bindung war jedoch nicht stabil genug, um die zur Strukturbestimmung nötigen höherdimensionalen Spektren zu messen. Der Vergleich von Spektren ergab, dass sich die Struktur von peptidyl-TycC3\_PCP nicht ändert, wenn der Thioester durch eine Hydrolyse-stabile Amidbindung ersetzt wird. Durch die Verwendung des Hydrolyse-stabilen Peptids war es möglich, die Struktur von peptidyl-TycC3\_PCP mittels Lösungs-NMR zu bestimmen. In dieser Struktur nimmt die PCP Domäne die als *A/H state* bezeichnete Konformation ein, welche bisher bei ähnlichen Domänen zumeist beobachtet wurde. Zwischen PCP Domäne und Ppan-Arm wurden nur wenige, und zwischen PCP Domäne und Peptid keine Kern-Overhauser-Effekt Kontakte beobachtet. Dies weist darauf hin, dass das Peptid frei beweglich ist und in keinem engen Kontakt zur PCP Domäne steht, was auf eine selektive Erkennung hindeuten würde.

In einem weiteren Versuch, strukturelle Informationen über die Erkennung eines PCP-gebundenen Peptids durch eine C Domäne und die damit einhergehenden Wechselwirkungen zwischen den beiden Domänen zu erhalten, wurde die PCP-C Bidomäne TycC5-6\_PCP-C verwendet, für die bereits zuvor von Samel *et al.* eine Aktivität der C Domäne nachgewiesen worden war. Auch diese Bidomäne sollte Hydrolyse-stabil mit einem Peptid beladen und anschließend kristallisiert werden. Ein signifikanter Anteil des Proteins wurde jedoch bereits während der Expression mit Ppan beladen und somit in die *holo*-Form überführt. Da die Bidomäne nicht vollständig in der *apo*-Form vorlag, in der die PCP Domäne unmodifiziert ist, konnte die quantitative Beladung mit dem Peptid nicht durchgeführt werden. Eine quantitative Überführung in die *holo*-Form war hingegen *in vitro* möglich, so dass *holo*-TycC5-6\_PCP-C erfolgreich kristallisiert werden konnte. Die gelöste Kristallstruktur entsprach aber jener, die bereits zuvor von Samel *et al.* für *apo*-TycC5-6\_PCP-C gelöst worden war. In dieser Struktur haben die PCP und C Domäne eine Orientierung zueinander, die aufgrund des großen Abstands zwischen der Bindungsstelle für den Ppan-Rest und dem katalytischen Zentrum der C Domäne während der Kondensationsreaktion nicht vorliegen kann.

Im zweiten Teil der Arbeit wurde die Modifizierung von PCP Domänen mit Ppan untersucht. Diese Beladung, bei der die PCP Domäne von der inaktiven *apo*- in die aktive *holo*-Form überführt wird, wird von PPTs katalysiert, die in einer Magnesium-abhängigen Reaktion Coenzym A (CoA) als Quelle für den Ppan-Rest verwenden. Für die Untersuchungen wurden TycC3\_PCP und Sfp verwendet. Bei diesen beiden bereits gut charakterisierten Proteinen handelt es sich um keine natürlichen Interaktionspartner, da die sie in verschiedenen *Bacillus* Stämmen vorkommen. Aufgrund der großen Ähnlichkeit der Tyrocidin A und Surfactin A Synthetasen wurde dies jedoch als vernachlässigbar angesehen.

Zur Kristallisation des Komplexes von PCP Domäne und Sfp wurde das Serin, an dessen Seitenkette der Ppan-Rest angehängt wird, in TycC3\_PCP mittels ortsspezifischer Mutagenese durch Alanin ersetzt um eine Modifikation zu verhindern und den Komplex zu stabilisieren. Dadurch war es möglich, den PCP/Sfp Komplex in Anwesenheit von CoA zu kristallisieren und die Kristallstruktur mit einer Auflösung von 2.0 Å zu lösen.

In der gelösten Struktur bilden je ein TycC3\_PCP(S45A)- und Sfp-Molekül einen Komplex, wobei sich zwischen den beiden Proteinen ein CoA-Molekül und zwei Magnesiumionen befinden. Aufgrund der beobachteten Anordnung scheint es möglich, dass die Seitenkette des in PCP Domänen konservierten Serins entweder durch ein in Gruppe II PPTs konservierten Glutamats oder durch ein Wassermolekül, dessen Reaktivität sich aus seiner Koordinierung an das eine Magnesiumion ergeben könnte, deprotoniert wird. Der anschließende nukleophile Angriff der deprotonierten Seitenkette auf die 5'  $\beta$ -Phosphatgruppe von CoA könnte durch das zweite Magnesiumion begünstigt werden, welche an die 5'  $\alpha$ -Phosphatgruppe koordiniert ist.

Neben diesen Einblicken in den möglichen Reaktionsmechanismus sind in der Struktur des Komplexes zwei Bereiche, in denen die beiden Proteine in Kontakt stehen, zu erkennen: Zum einen bildet der Bereich vor der zweiten Helix der PCP Domäne eine intermolekulare Wasserstoffbrückenbindung zur N-terminalen Domäne von Sfp, und zum anderen gibt es zwischen Bereichen in der zweiten Helix der PCP Domäne und der C-terminalen Domäne von Sfp hydrophobe Wechselwirkungen. Diese Interaktionen wurden durch Mutationsstudien weiter

untersucht. Mutation der Aminosäuren, welche den hydrophoben Bereich in der zweiten Helix der PCP Domäne bilden, führte zu einer signifikanten Erniedrigung der Affinität zu Sfp in Isothermalen Titrationskalorimetrie-Messungen und zu einer Verlangsamung der Ppan-Transferreaktion. Mutationen, welche die intermolekulare Wasserstoffbrücke aufbrechen, führten ebenfalls zu einer erniedrigten Komplexstabilität, jedoch wurde die Modifizierungsreaktion durch diese Mutationen beschleunigt.

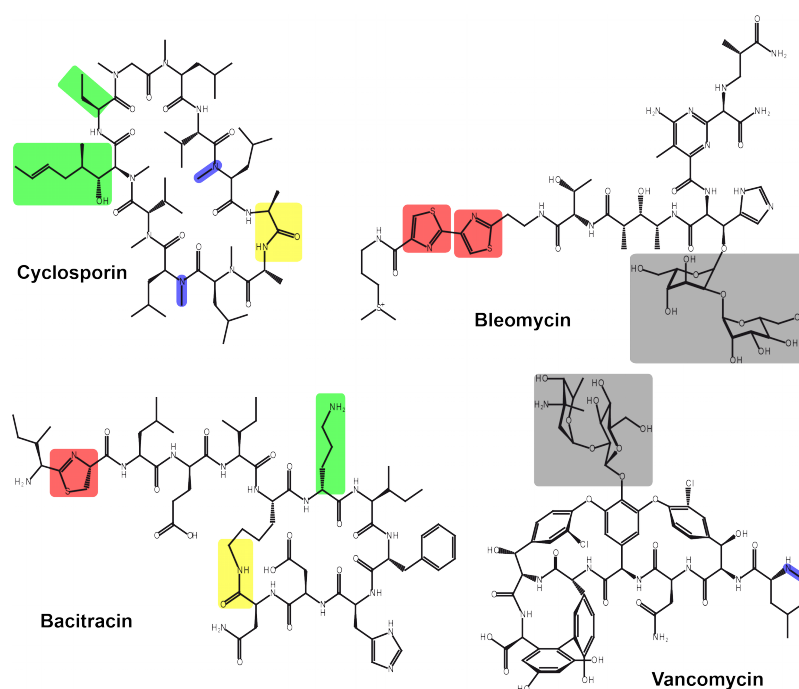
TycC3\_PCP(S45A) liegt in der gelösten Struktur des PCP/Sfp Komplexes in der *A/H state* Konformation vor, was vorherigen Untersuchungen von Koglin *et al.* widerspricht. Um auszuschließen, dass es sich bei der beobachteten Konformation um ein durch die Kristallisation hervorgerufenes Artefakt handelt, wurde die Struktur des ungebundenen TycC3\_PCP(S45A) mittels Lösungs-NMR bestimmt und die Komplexbildung mit Sfp in Lösung durch NMR-Titrationsmessungen untersucht. Da auch hier jeweils die *A/H state* Konformation beobachtet wurde, ist davon auszugehen, dass es sich bei der in der Kristallstruktur beobachteten Konformation der PCP Domäne nicht um ein Artefakt handelt.

Die Struktur des PCP/Sfp Komplexes hat eine große Ähnlichkeit zur von Bunkoczi *et al.* gelösten Kristallstruktur des Komplexes aus der humanen Gruppe II PPT und der *Acyl Carrier Protein* Domäne aus der humanen Fettsäure Synthase. Außerdem weist, in Bezug auf die an der Katalyse des Ppan-Transfers und die an den Wechselwirkungen zwischen PPT und *Carrier Protein* beteiligten Reste, eine Vielzahl anderer Proteine aus verschiedensten Organismen eine große Sequenzhomologie auf. Diese Beobachtungen deuten darauf hin, dass die in dieser Arbeit am Beispiel von TycC3\_PCP und Sfp untersuchten Wechselwirkungen zwischen *Carrier Protein* und Gruppe II Phosphopantethein Transferase in der Natur weit verbreitet sind.

# 1. Introduction

## 1.1 Structure and Function of Nonribosomal Peptide Synthetases

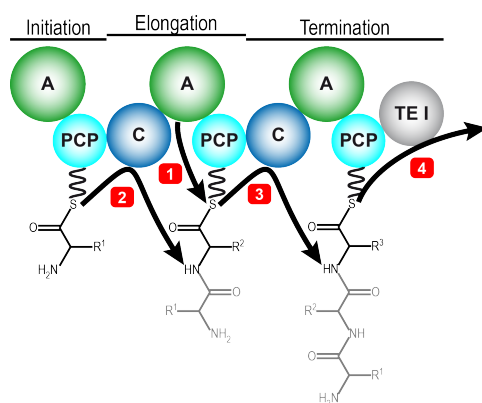
The formation of secondary metabolites by nonribosomal peptide synthetases (NRPSs) is often used in bacteria and fungi<sup>1-3</sup>. Among the various nonribosomal peptides (NRPs) are many with interesting features for medical application. Beside immunosuppressants like cyclosporin A from *Tolypocladium inflatum*<sup>4, 5</sup> and anti-tumor agents like bleomycin from *Streptomyces verticillus*<sup>6, 7</sup>, many NRPs show antibiotic properties like bacitracin A from *Bacillus licheniformis*<sup>8, 9</sup> or vancomycin from *Amycolatopsis orientalis*<sup>10, 11</sup> (Figure 1). The high diversity of NRPs results from their numerous building blocks and modifications. Incorporated building blocks are not only proteinogenic amino acids but comprise, amongst others, non-canonical amino acids, fatty acids and aryl acids. Common modifications are epimerization of amino acids, methylation, glycosylation, cyclization of side chains, and the formation of macrocyclic structures.



**Figure 1:** Selected NRPs with pharmacological relevance. Examples for non-proteinogenic amino acids (green), methylation (blue), glycosylation (gray), cyclization (red), and macrolactams (yellow) are highlighted.

Throughout the last decades, the widespread application of antibiotics has been accompanied by the rise of bacteria with resistance to common antibiotics not only in health care<sup>12, 13</sup> but also in food production<sup>14, 15</sup>. Prominent examples are methicillin-resistant *Staphylococcus aureus*<sup>16</sup> and *Enterococcus faecium* strains with resistance to vancomycin<sup>17</sup> and daptomycin<sup>18</sup>. Facing the rapidly growing number of pathogens with multi-drug resistance, the development of new antimicrobial agents is a subject of high importance. One strategy on this way is the search for microorganisms which produce new bioactive NRPs<sup>19, 20</sup> and their use in large-scale production. Although the cultivation of many natural NRP-producers is challenging, the production of NRPs by heterologous expression of their corresponding synthetases has been shown successfully in organisms suitable for an easy fermentation in an industrial scale like *Escherichia coli* (*E. coli*)<sup>21</sup> and yeast<sup>22, 23</sup>.

Another potential strategy to gain new antibiotics arises from the architecture of NRPSs. In general, an NRPS consists of modules with a repetitive domain arrangement, in which each module catalyzes the incorporation of one amino acid into the peptide, and the peptide sequence is coded in the succession of the modules<sup>24, 25</sup>. A basic module consists of a condensation (C) and an adenylation (A) domain as well as of a peptidyl carrier protein (PCP). The A domain activates an amino acid and transfers it onto the PCP domain. The PCP presents the amino acid to its preceding C domain, which catalyzes the reaction with a peptide tethered to the previous PCP. Afterwards, the PCP presents the elongated peptide to its subsequent C domain. In the first, so-called initiation module of an NRPS, the C domain is usually missing, and in the last (termination) module the full-length peptide is presented to a domain catalyzing its release (Figure 2).



**Figure 2**

Schematic representation of the processes during the maturation of an NRP. An amino acid is activated and loaded onto the PCP by the A domain (1). The PCP presents the amino acid to the upstream C domain, where the peptide bond is formed (2). Loaded with a peptide, the PCP interacts with the downstream C domain, and the growing peptide chain is again transferred to the next PCP (3). Finally, the full length peptide is set free from the endmost PCP by a releasing domain (e.g. a type I thioesterase; 4).

Depending on the overall arrangement of the modules in the synthetase and its relation to the produced NRP, different types of NRPSs are classified. On the one hand, type I NRPSs, where all modules involved in the formation of the NRP are present in one protein, are distinguished from type II, where the modules are distributed among several proteins that interact during the formation of the product. On the other hand, linear (type A), iterative (type B), and non-linear (type C) NRPSs are distinguished. In a linear NRPS, the succession and quantity of the building blocks in the NRP correspond to the order and number of the modules, in an iterative NRPS more than one peptide chain is formed in the same fashion like in a type A NRPS and fused to oligomers in the final NRP. In a non-linear NRPS, no relation between the arrangement of the modules to the building blocks in the final product is evident.

By manipulation of the selectivity and the order of domains or modules within an NRPS, it should be possible to reprogram it for the production of an altered product<sup>26</sup>.<sup>27</sup> Beyond that, a combination of fragments from different NRPSs could give access to an almost unlimited number and diversity of new peptides and would be straight forward, if suitable libraries of such fragments would be at hand. However, up to now such attempts have had varying success, and often newly designed NRPSs were hardly productive, if any product was obtained at all<sup>28-30</sup>. The reason for this is the limited knowledge of the selectivity in the domain/domain interactions and in the recognition of the intermediates by the domains.

### 1.1.1 The Peptidyl Carrier Protein

The fact that almost every intermediate of an NRP is bound to a PCP (sometimes also referred to as thiolation domain) emphasizes the central role of this domain. By the covalent binding, the local concentrations are increased, which is favorable for a faster reaction, and a higher selectivity over the vast pool of potential unbound substances in the cellular environment is achieved. Binding of the intermediates to the PCP is facilitated via a 4'-phosphopantetheine cofactor (ppan) that is covalently bound as a phosphoester to the side chain hydroxyl group of a serine residue invariant in every PCP<sup>31-33</sup>. When this serine was initially identified, it was introduced

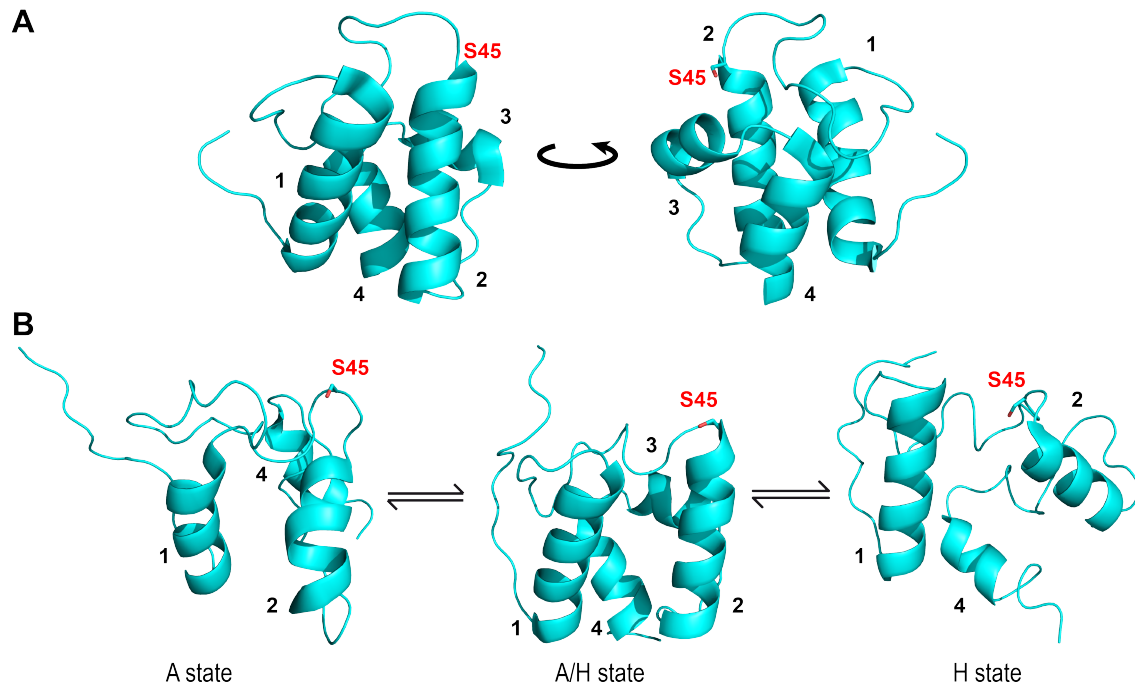


as being part of a conserved GG(D/L)S(I/L) motif, forming the recognition site for the modification<sup>32, 34</sup>, but with the growing number of known protein sequences of NRPSs it became evident that non of the residues flanking the serine is completely invariant, though only a few examples are known where the first glycine is not conserved. The aminoacyl and peptidyl intermediates are bound to the distal thiol group of the cofactor with the energy required for the peptide bond formation stored in the reactive thioester bond. Accordingly, *apo*-PCPs lacking the cofactor are not functional and need to be posttranslationally modified by phosphopantetheine transferases (PPTs) to become active *holo*-proteins<sup>35, 36</sup>.

Similar to carrier proteins from other types of megasynthases, PCPs consist of ~80-100 amino acids that fold into a four-helix bundle structure with the conserved serine at the N-terminus of the second helix. The first PCP of which the structure was solved is TycC3\_PCP, the third PCP domain of subunit C of the tyrocidine A synthetase. The structure (PDB 1DNY) was solved for the *apo*-form by nuclear magnetic resonance spectroscopy (NMR), but due to the high similarity in the [<sup>1</sup>H;<sup>15</sup>N]-heteronuclear single quantum coherence (HSQC) spectra of the *apo*- and *holo*-protein, it was suggested that both forms adopt the same structure<sup>37</sup>. In this structure, the helices 1 (residues 13-28), 2 (45-59), and 4 (73-83) align in a twisted parallel orientation, and the shorter helix 3 (64-71) lies almost orthogonal to them. Consisting of 16 residues, the loop between the first two helices is much longer than the connections of the other helices. It is positioned on top of one side of the helical bundle with contacts to the N-terminus of helix 4 and the loop preceding the first helix (Figure 3A).

Further structural studies on TycC3\_PCP revealed that the *apo*- as well as the *holo*-protein adopt dynamic equilibria between two conformers with one conformation common to the PCP with and without the cofactor<sup>38</sup> (Figure 3B). The common conformation, denominated A/H state (PDB 2GDW), resembles the known structure of the PCP with a slightly increased distance between the three parallel helices. The conformations specific for the *apo*- (A state) and the *holo*-PCP (H state) show a less ordered secondary structure and were not known so far<sup>38</sup>. In the A state (PDB 2GDY) the third helix is disintegrated, whereas helix 1 (residues 14-24), helix 2 (48-59), and helix 4 (73-83) are shortened. In the H state (PDB 2GDY) the third helix is dissolved,

too, but from the remaining helices only helix 2 and helix 4 are shortened significantly compared to the A/H state (Figure 3B).



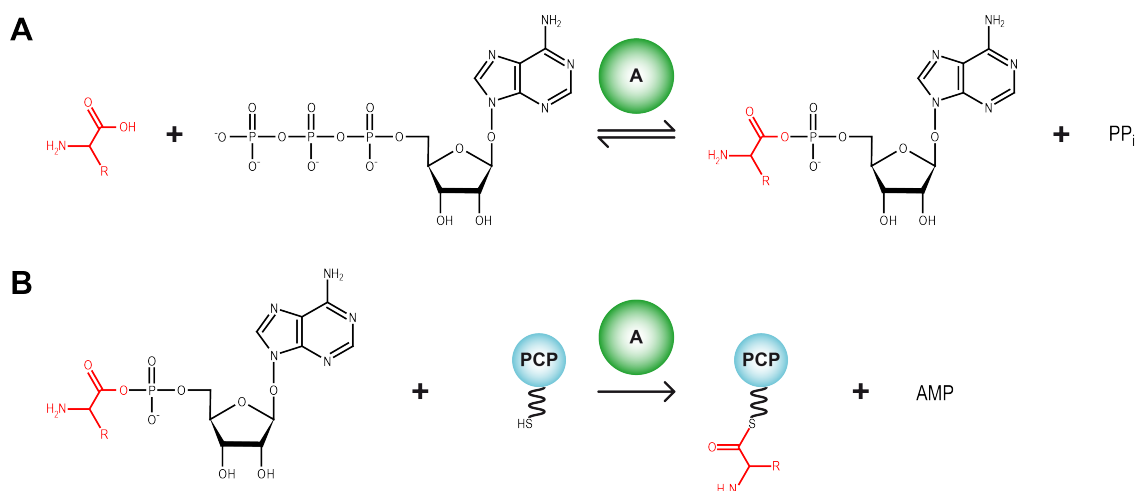
**Figure 3:** Structures of the TycC3\_PCP. In the first structure of the PCP, the typical fold as a four helix bundle was observed (A; PDB 1DNY). Refinement of the structure revealed that the previously observed conformation (A/H state; PDB 2GDW) is in a dynamic equilibrium. In the *apo*-PCP it coexists with the A state (PDB 2GDY) and in *holo*-PCP with the H state (PDB 2GDY) (B). The helices are numbered from the N- to the C-terminus, and the active site serine is shown as sticks and labeled.

Since helix 2 is shortened in the A state as well as in the H state, in these two conformations the active site serine S45 is not located at the N-terminus of the helix, but is part of the preceding loop. Furthermore, it was shown that the ppan cofactor undergoes a reorientation by  $\sim 100^\circ$  when the conformation of the *holo*-PCP changes.

NMR titration experiments with interaction partners of the PCP revealed that the different conformers are specifically selected for different interactions, and a role of the conformational switch and the mobility of the cofactor for the substrate shuttling by the PCP was proposed<sup>38, 39</sup>.

### 1.1.2 The Adenylation Domain

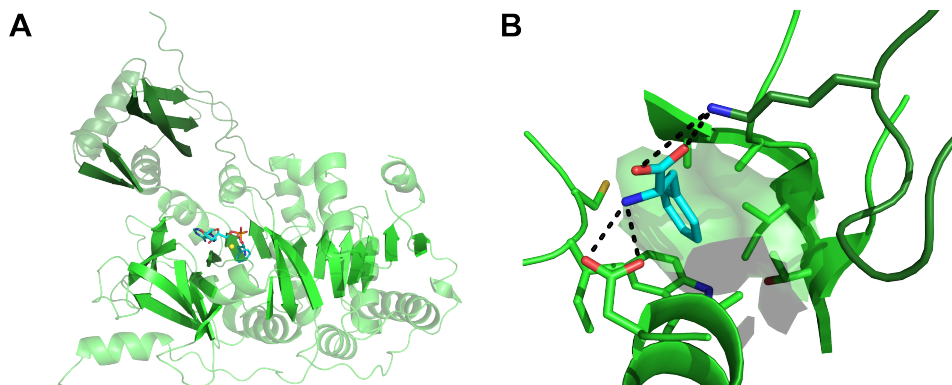
A domains catalyze the activation and transfer of the amino acid onto the thiol group of the ppan-arm of a *holo*-PCP in a two step reaction<sup>34, 40-42</sup>. In the first, reversible step of the reaction, the amino acid is activated under consumption of ATP, throughout which an amino acid adenylate is formed. By hydrolysis of pyrophosphate, which is a side product of the reaction, the reversibility of this reaction step can be circumvented. In the second step, the thiol group of the PCP's cofactor acts as a nucleophile and attacks the adenylated amino acid. Thus, the thioester bond between amino acid and ppan is formed, and AMP is released (Figure 4).



**Figure 4:** In the first half-reaction, the A domain activates an amino acid as an adenylate under consumption of ATP (A). The activated amino acid is subsequently transferred onto the thiol group of the ppan-arm of the PCP (B).

The A domain is the first NRPS domain whose three-dimensional structure was solved: In 1997 the crystal structure of GrsA\_A (due to its selectivity for phenylalanine often referred to as PheA) from the gramicidine S NRPS was determined<sup>43</sup> (PDB 1AMU). It revealed that the protein consists of a large N-terminal (~410 residues) and a smaller C-terminal subdomain (~100 residues). The structure of the N-terminal subdomain has three  $\beta$ -sheets with 8, 7, or 6 strands, respectively, as striking structural elements with several  $\alpha$ -helices located around and between them. Connected to the larger domain by a short linker, the so-called hinge-region, the C-terminal domain of the A domain consists of a two- and a three-stranded

$\beta$ -sheet as well as two helices and shows relatively few direct protein/protein contacts to the other subdomain (Figure 5A). Since one phenylalanine and one AMP molecule were co-crystallized with GrsA\_A, information about the cofactor binding could be derived from the structure, too. The two molecules are located in close proximity to each other at the interface of the two subdomains. Of special interest was the existence of a binding pocket for the amino acid (Figure 5B): The side chain of phenylalanine points into this binding pocket formed by side chains of residues from the N-terminal domain, whereas the amino and carboxyl group are in contact with both subdomains. Further analysis of the corresponding residues from other A domains revealed a correlation between the nature of these residues and the specificity of the A domain<sup>44, 45</sup>. This code, defined by 10 residues, was successfully applied in directed alteration of the specificity of A domains<sup>30</sup> and in the prediction of the specificity of newly found A domains<sup>46</sup>. However, recent studies have shown that the binding motif for amino acids with very large side chains comprises more than 10 residues<sup>47</sup>, and that the selectivity is influenced by other residues, too<sup>48</sup>. Furthermore, it was observed that the selectivity *in vitro* can even depend on the composition of the reaction buffer, especially with respect to the addition of crowding agents<sup>49</sup>.



**Figure 5:** The structure of GrsA\_A reveals that adenylation domains consist of two subdomains. The large N-terminal subdomain (green) is built by three  $\beta$ -sheets and several helices, whereas the smaller C-terminal portion (dark green) is formed by two sheets and three helices. At the interface of the two domains AMP, Phe (sticks), and a magnesium ion (yellow sphere) are co-crystallized (A. For a better clarity all structural elements but the  $\beta$ -sheets are shown semitransparent). A detailed view of the bound amino acids (cyan) shows that its side chain points into a hydrophobic pocket (surface), and the amino- and carboxyl-group form hydrogen bonds (dashed lines) to the protein (B).

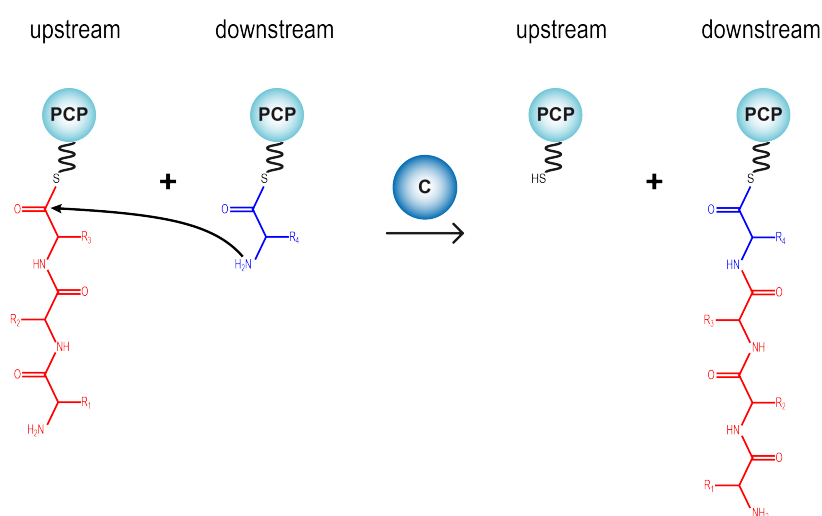
The function of the A domain is similar to the tRNA synthetase, which activates amino acids as adenylates, too, and transfers them onto a tRNA. Still, A domains

show no structural similarity to this class of synthetases, but resemble the structures of the firefly luciferase<sup>50</sup> and the acyl-CoA synthetase<sup>51</sup>, with which they share sequential homology, too<sup>52, 53</sup>. Comparison of known structures of these type of proteins showed that the C-terminal domain can adopt different orientations relative to the N-terminal domain. These different conformations were assigned to three different steps in the catalytic cycle of the A domain<sup>54, 55</sup>. The domain without any ligands bound is in the open conformation, in which the binding sites for the amino acid and ATP are not covered by the C-terminal subdomain. Upon binding of the two substrates, the A domain changes to the adenylation conformation. Now the C-terminal subdomain gets into contact with the substrates and the adenylation reaction occurs. When the adenylate is formed, the A domain again undergoes a conformational change, and pyrophosphate is released. In the thioester/transfer conformation the C-terminal subdomain covers the adenylate to shield it against hydrolysis, and a channel for the binding of the ppan moiety is formed at the interface of the two subdomains. The PCP-bound cofactor enters this channel, and the amino acid is transferred to its thiol group. After the thioester is formed, the A domain opens up again and all products are released so the enzyme can undergo the next catalytic cycle.

In some cases the insertion of a stretch of ~200 residues between the two  $\beta$ -sheets of the C-terminal subdomain of the A domain is observed<sup>56, 57</sup>. Analysis of the excised stretch revealed that it is an oxidase domain, whose function is independent of the adenylation domain and the adenylation reaction<sup>58</sup>. Then again, a type of protein is known which has been shown to influence the activity of certain A domains. MbtH-like proteins<sup>59</sup>, named after the protein from the mycobactin NRPS<sup>60</sup>, are ~70 amino acids in size. Their exact function is still not yet known, but they have been shown to bind to certain A domains with a high affinity<sup>61</sup> and modulate their activity<sup>61, 62</sup>, their selectivity<sup>61</sup>, and their stability<sup>63</sup>. Although it was shown that certain A domains bind MbtH-like proteins in a 1:1 ratio<sup>61</sup> it is unlikely that the activity of these A domains is regulated by the expression level of the MbtH-like proteins, as the transcription level of their genes was found to be similar<sup>64</sup>, and an example for a naturally occurring fusion of the two proteins is known<sup>65</sup>.

### 1.1.3 The Condensation Domain

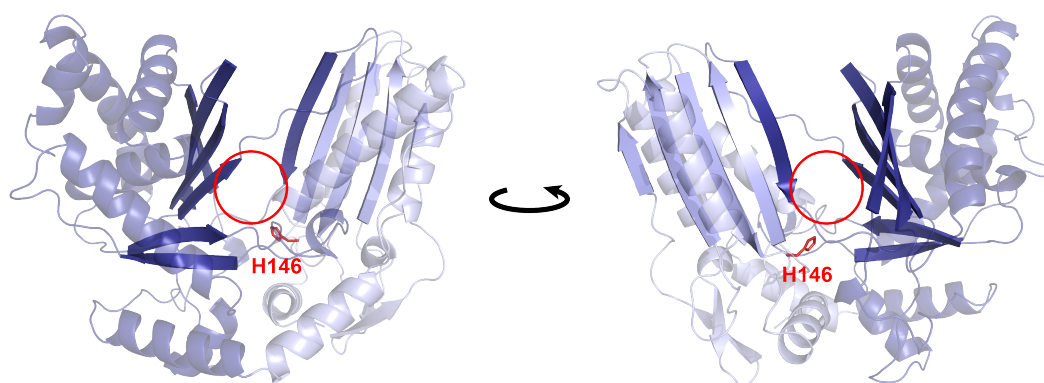
The central reaction in an NRPS, the formation of the peptide bond, is catalyzed by the C domain<sup>66</sup>, which comprises of ~440 residues. In this reaction two PCP domains interact with one C domain. The peptide-loaded PCP of the preceding module (“upstream”) acts as the peptide donor, and the aminoacylated PCP, usually separated by an A domain from the C-terminus of the C domain (“downstream”), as the peptide acceptor. There is no evidence that the peptide is bound to the C domain at any stage of the condensation reaction, but it is directly transferred to the amino group of amino acid tethered to the downstream PCP (Figure 6). The most prominent residue for the catalysis is the second histidine of a conserved HHxxxDG motif, which is thought to stabilize the intermediate of the nucleophilic attack of the amino group on the carbonyl of the thioester, linking the peptide to the upstream PCP<sup>66-69</sup>.



**Figure 6:** Condensation domains catalyze peptide bond formations. The peptide-loaded (red) upstream and the aminoacylated (blue) downstream PCP present their charged cofactors to the C domain, where the nucleophilic attack of the amino group on the thioester occurs (arrow).

The first detailed information about the structure of C domains was gained by resolving the crystal structure of VibH, a stand-alone C domain from the vibriobactin synthetase<sup>68</sup> (PDB 1L5A). The structural basis of the C domain is a subdomain similar to the chloramphenicol acetyltransferase<sup>70</sup>, which is, amongst other structural elements, characterized by a six-stranded  $\beta$ -sheet. In contrast to the chloramphenicol acetyltransferase, which forms an intermolecular trimer of such domains, the

C domain consists of two such domains within one peptide chain, and the  $\beta$ -sheet in the N-terminal (sub-)domain consists of only five strands, of which one is a part of the C-terminal domain with respect to the primary sequence. Furthermore, another part of the C-terminal domain points towards the N-terminal domain, too. The arrangement of the two domains is described as V-shaped with a groove in the space between them. Remarkably, the side chain of the histidine that was suspected to be the catalytic residue of the C domain is located approximately in the middle of this groove. The two ends of the groove were attributed to the donor and acceptor site, from where the peptide- and amino acid-loaded ppan-arm, respectively, enter the groove to undergo the peptide transfer reaction (Figure 7).



**Figure 7:** Structure of the C domain VibH. The N- (light blue with respect to the primary sequence) and the C-terminal domain (dark blue) are arranged in a V-shape. At the interface of the domains a groove with the catalytic histidine (red sticks) is formed. The two entry-sites of the groove are the donor (red circle in the left picture) and acceptor site (right). For a better clarity only the two large  $\beta$ -sheets are shown as massive and the rest of the protein as semi-transparent.

The description of the C domain as V-shaped suggests that the two subdomains are arranged more or less symmetrically, which is not the case: The two large  $\beta$ -sheets are orientated in different directions. With the growing number of structures of C domains<sup>68, 69 71, 72</sup> solved, it became evident that the arrangement of the two subdomains with respect to each other is not always the same. The transition between the different orientations was shown to be possible by molecular dynamics simulation and was suggested to be relevant for the catalysis of the peptide bond formation<sup>72</sup>. However, up to now the different orientations were not assigned to specific steps in the catalysis, and different conformations for the same C domain were not observed so far.

Studies on the substrate selectivity of C domains revealed a high selectivity at the acceptor position. The donor substrate is less strictly selected<sup>73, 74</sup>, although a high preference for peptides with either D- or L-configuration for the amino acid at the thioester forming position was observed<sup>69, 73, 75, 76</sup>. Furthermore, C domains are often located at the junction of polyketide synthase (PKS) to NRPS parts in PKS/NRPS hybrid systems, where they accept an acyl carrier protein (ACP) bound polyketide as the donor substrate<sup>7, 77</sup>.

Beside the catalysis of the peptide bond formation between PCP-bound amino acids and peptides, there are other functions known that certain C domains fulfill. As the activation energy of the reaction is stored in the reactive thioester bond of the donor substrate, the transfer can occur with an acceptor substrate which is not bound to a PCP, too<sup>68</sup>. Furthermore, C domains can catalyze the incorporation of a fatty acid (FA) into the NRP. This so-called lipoinitiation is catalyzed by a C domain located at the N-terminus of the first module, and the donor substrate is a fatty acid either bound to CoA<sup>74</sup> or a carrier protein<sup>78</sup>. Especially in fungal NRPSs, terminal C domains often catalyze the macrocyclization and release of the mature peptide<sup>79, 80</sup>. It has also been shown that an excised bacterial elongation C domain can catalyze macrocyclization of peptides *in vitro*<sup>69</sup>.

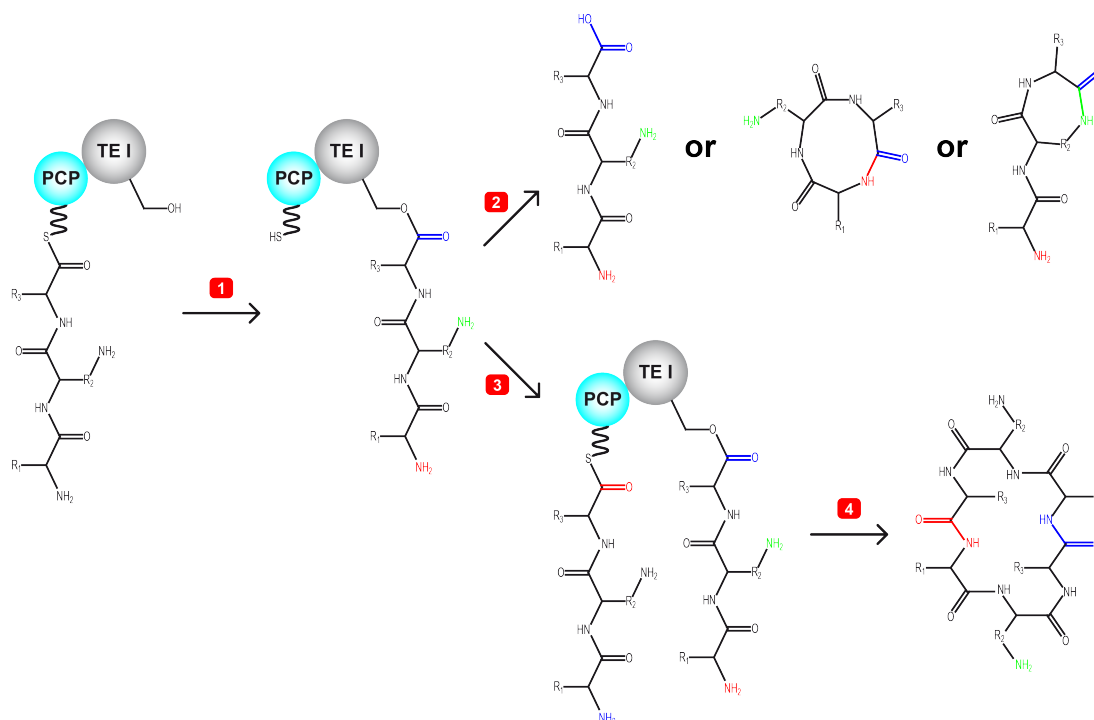
Variants of C domains are the cyclization and dual condensation/epimerization domains, which share a high sequence similarity with the C domain. Cyclization domains catalyze not only the peptide bond formation but a subsequent reaction between a side chain hydroxy- or thiol-group and the backbone carbonyl of the two condensed amino acids, resulting in the formation of an oxazoline- or thiazoline-ring<sup>9</sup>. The step-by-step character of this reaction could be shown by introduction of point mutations in a cyclization domain, which reduced its activity to the catalysis of the condensation reaction<sup>81</sup>. Dual condensation/epimerization domains catalyze the epimerization of the donor substrate prior to the condensation reaction, for which they only accept the D-enantiomer as substrate<sup>82</sup>. In contrast to these variants of the classic C domain, which all share a high sequence similarity, a non-related domain can substitute the C domain, too: Transglutaminase-like proteins have been shown to catalyze the peptide bond formation in some NRPSs<sup>83, 84</sup>.



## 1.1.4 Thioesterases

### 1.1.4.1 Type I Thioesterases

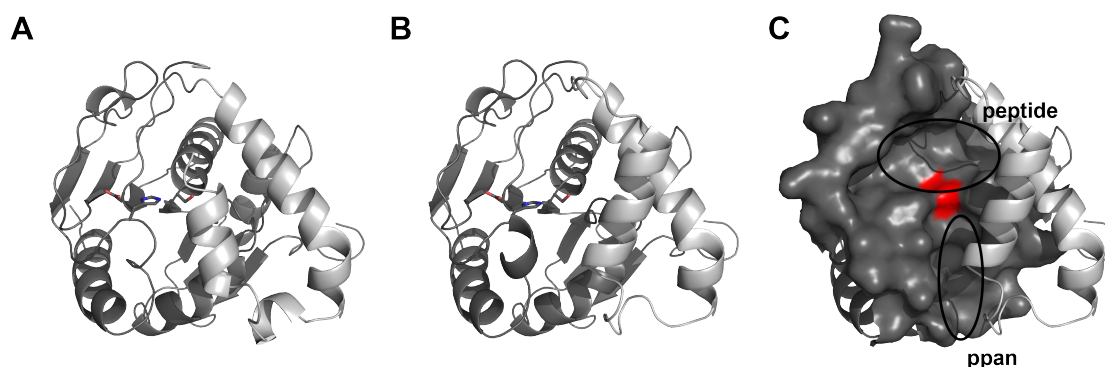
Thioesterase domains of type I (TE Is) are usually located at the C-terminus of an NRPS and catalyze the release of the mature peptide in *cis*, although examples are known where the TE I is located in another position within the NRPS<sup>85</sup> or is a separated protein<sup>46</sup>, albeit the uncommon arrangements of the in *trans* acting TEs are accompanied by a different domain size and probably a different structure, too. The mature peptide is hydrolyzed as a linear product, or a macrocycle is formed<sup>86</sup>. In the latter case, the ring formation can occur as macrolactam<sup>87</sup>, macrolactone<sup>88, 89</sup>, or macrothiolactone<sup>90</sup>, and head-to-tail<sup>87</sup> (cyclic), head-to-side chain<sup>88</sup> (cyclic branched), and oligomeric<sup>42, 90</sup> peptides are distinguished (Figure 8).



**Figure 8:** For its release the full length peptide is transferred from the last PCP to the catalytic serine of the TE I (1). If it is released as a monomer (2), the peptide can either be hydrolyzed as a linear product or cyclized by the N- and C-terminus or by a side chain and the C-terminus. If the peptide is oligomerized, the first monomer is stored at the TE I, until a second peptide is produced (3), then the two peptides are released as one large macrocycle (4). Important groups are colored to illustrate the different types of bond formation.

TE I s are members of the  $\alpha/\beta$  hydrolase superfamily, with the residues of the conserved catalytic Ser/His/Asp triad located in different parts of the domain which contains  $\sim 270$  residues. In the first step of the release, the side chain hydroxyl group of the catalytic serine, activated by the histidine, attacks the carbonyl carbon of the thioester, which links the peptide to the ppan-arm, in a nucleophilic fashion. Thus, the peptide is transferred onto the TE I, where it is bound as an oxoester. The release of the peptide happens by another nucleophilic attack. The nucleophile can be a water molecule (resulting in a linear product), a nucleophile within the peptide (cyclic product) or from a second, PCP-bound peptide (oligomerization).

Structure elucidation of SrfAC\_TE I, the TE I of the surfactin synthetase (PDB 1JMK), confirmed the expected  $\alpha/\beta$  hydrolase overall fold, although the N-terminal  $\beta$ -strand is missing. A special feature of TE I is the so-called lid-region, which is inserted between the fifth and sixth strand of the central  $\beta$ -sheet<sup>91</sup>. The core domain forms two binding pockets for the ppan-arm of the PCP and the peptide, with the catalytic serine and the so-called oxygen hole formed by backbone amides for stabilization of the intermediate of the nucleophilic attack in the middle. In the crystal structure of SrfAC\_TE I, an open and a closed conformation of the lid region were observed (Figure 9), whereas the corresponding regions were partly disordered in other crystal structures<sup>71, 92, 93</sup> and showed a high flexibility in NMR investigations<sup>94</sup>. The proposed function of the lid region is the binding and alignment of the peptidyl-ppan, as well as prevention of the hydrolysis of the TE I-bound peptide<sup>91</sup>.



**Figure 9:** Structure of SrfAC\_TE I in the closed (A) and open (B) conformation. The C-terminal part of the lid region (light gray) is differently oriented towards the catalytic triad (sticks) in the core domain (dark gray). Surface representation illustrates the binding sites for the ppan-arm and the peptide with the catalytic serine of the TE I (red) in the middle (C).

Analysis of the selectivity of TE I has demonstrated that they act even on peptides bound to synthetic molecules with high promiscuity to different peptides, although the tolerance towards alterations depends on their nature and the position in the peptide<sup>95-98</sup>.

#### 1.1.4.2 Type II Thioesterases

The necessity for a thioesterase of type II (TE II) arises from the high cellular content of acylated Co A derivatives<sup>99</sup> and the low selectivity of NRPS associated PPTs<sup>35</sup>. Due to this situation, the priming of an NRPS *in vivo* will mainly result in PCP domains carrying a blocked cofactor, so the whole synthetase is inactive<sup>100</sup>. The function of TE II is to remove the group bound to the thiol group of ppan in order to regain the activity of the NRPS<sup>101</sup>. For each NRPS, there is only one TE II that is a separate protein and acts in *trans* on all PCPs. It has been shown that TE IIs can act on aminoacylated PCPs, so they can remove amino acids falsely loaded by the A domain that would block the NRPS as well, but only small peptides are hydrolyzed<sup>101, 102</sup>. This biochemical data is in agreement with the observed intensity of binding of TE II to a PCP loaded with different non-hydrolyzable substrates<sup>39</sup>.

The structure of the TE II associated with the surfactin A NRPS (SrfAD) has been solved by NMR techniques<sup>39</sup> (PDB 2RON). It shows high similarity to TE I with respect to the secondary structure and has a catalytic Ser/His/Asp triad<sup>103</sup>, too, but the tertiary structure differs in some parts. The lid region shows flexibility as observed for TE I, but it is smaller and hardly covers the catalytic triad.

#### 1.1.5 Tailoring Domains and Enzymes

Considering the large number of modifications of NRPs and the involved enzymes only a short overview is given about tailoring domains that are part of an NRPS and individual tailoring proteins that are associated with NRPS.

Although other strategies for the incorporation of amino acids with D-configuration are known<sup>82, 104</sup>, the most common way is the utilization of an epimerization domain (E domain). E domains are integrated into the NRPS, where they are located at the

C-terminus of a module. They act on the substrate bound to the preceding PCP in the same module. In elongation modules, potentially the aminoacyl- and the peptidyl-PCP could be the substrate, but *in vitro* studies showed that E domains from elongation modules act on peptidyl-PCPs much faster than on aminoacyl-PCPs, and the opposite is the case when they are substituted by an E domain derived from an initiation module<sup>105</sup>. Furthermore, it has been shown that E domains have a moderate selectivity with respect to the amino acids that are epimerized<sup>106</sup>. The crystal structure of TycA\_E (2XHG) confirmed the structural similarity between E and C domains, which was proposed on the basis of their high sequence similarity.

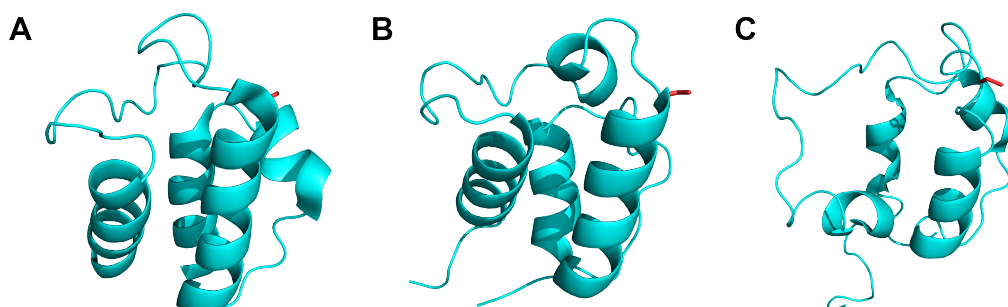
Flavine mononucleotide dependent oxidation domains are either inserted into the C-terminal subdomain of an A domain or following a PCP<sup>107, 108</sup>. In both cases they are part of a module with a cyclization domain, and their substrate is the PCP-bound heterocycle, which is oxidized to an aromatic<sup>108</sup>. The reduction of the heterocycle by an external reductase in a nicotine adenine dinucleotide phosphate (NADPH) dependent reaction is known, too<sup>109, 110</sup>. Integrated reduction domains with the same cofactor can be involved in peptide release as an aldehyde<sup>111, 112</sup>, whereas a different kind of oxidation is catalyzed by separate enzymes that form C-C and C-O bonds between side chains in a heme dependent reaction<sup>113, 114</sup>.

Beside these modifications, there is a wide range of atoms and groups that can be incorporated in NRPs. S-adenosyl methionine is the donor substrate for N-, C-, and O-methylation. N-methyltransferases are found as domains being incorporated<sup>115</sup> in or following to A domains<sup>5, 116</sup>, or as stand-alone enzymes<sup>11, 117</sup>. The methylation can occur on the amino acid<sup>118, 119</sup> prior to or on the peptide after the condensation reaction<sup>110</sup> as well as on the full-length peptide after release<sup>117, 120</sup>. C-methyltransferases are found as integrated domains<sup>121</sup> or separate enzymes<sup>122</sup>, too, and they act on PCP-bound intermediates<sup>121</sup> or on amino acid precursors prior to their activation by the A domain<sup>122</sup>. The latter phenomenon was also observed for O-methyltransferases<sup>123</sup>. Hydroxylation by stand-alone enzymes was observed, which either modify PCP-bound amino acids heme-dependently<sup>124</sup>, or act on precursors by utilization of NADPH<sup>125</sup>. Formylation with N<sup>10</sup>-formyl tetrahydrofolate as substrate happens at the same two points of the NRP formation with an external enzyme acting on the precursor<sup>125</sup> or an incorporated domain on the aminoacylated

PCP<sup>126</sup>. Acetylation, on the other hand, is exclusively carried out by separate proteins which modify precursors<sup>127</sup> or the full-length peptide<sup>128</sup> with acetyl-CoA as cofactor. Full-length peptides, so-called aglycons, are the only substrates for glycosyltransferases which transfer UDP-coupled sugar moieties<sup>129, 130</sup>. For halogenation, two types of stand-alone enzymes are known, that are acting on PCP-bound amino acids: Non-heme iron dependent halogenases use  $\alpha$ -ketoglutarate<sup>131, 132</sup>, whereas the second type uses flavine adenine dinucleotide<sup>133</sup> to activate the substrate for halogenation.

## 1.2 Non-PCP Carrier Proteins

Contrary to PCPs, which are only found in the production of secondary metabolites in bacteria and fungi, ACPs involved in the fatty acid synthase (FAS) of the primary metabolism are ubiquitous. Two types of FASs are distinguished. In type I which is found in some bacteria and all eukaryotes but plants, the ACP and the catalytic domains are incorporated in one or two multifunctional peptide chains, whereas in type II common in bacteria, mitochondria, and plants, the ACPs like all other involved enzymes are separate proteins. Another difference between PCPs and ACPs is the fact that there is only one ACP which carries all intermediates throughout the maturation of the FA. Amongst others, the structures of the ACP from the FAS of *Bacillus subtilis* (*B. subtilis*)<sup>134</sup> (PDB 1HY8), of the isolated ACP domain from the rat FAS (PDB 2PNG)<sup>135</sup>, and of the ACP domain within the whole yeast FAS<sup>136</sup> (PDB 3HMJ) were solved and illustrate the high structural similarities of ACPs throughout all kingdoms of life (Figure 10).



**Figure 10:** The structures of ACPs from the FAS of *B. subtilis* (A), yeast (B), and rat (C) share the same overall fold. The conserved serine side chain is shown as red sticks.

The four-helical fold and the conserved serine as attachment site for ppan resembles the situation in PCPs (A/H state), even though the two types of carrier proteins share only a small sequence similarity and differ in the electrostatic properties of their surfaces<sup>37</sup>. The structure of ACPs from FAS was found to be flexible, too, but this flexibility was mainly located in the loop preceding helix 2<sup>134</sup>. No other overall conformations were observed up to now. Further studies revealed that this loop region is an entry point for the FA: For several ACPs from type II FAS, it has been shown that the growing acyl chain is buried in the hydrophobic core between the helices of the ACP without altering the overall arrangement of the helices, and it was proposed that the linking thioester is shielded against hydrolysis in that way<sup>137-139</sup>.

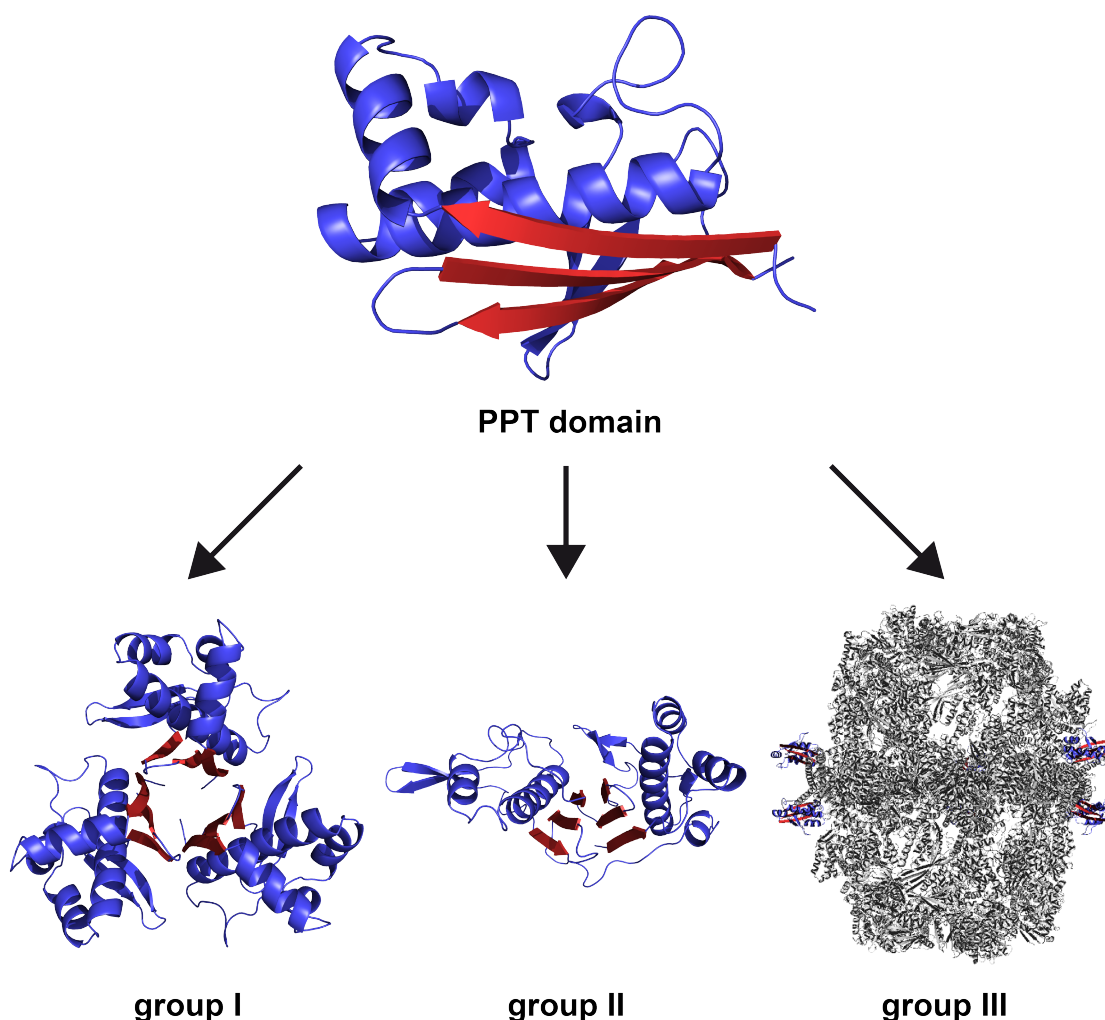
For PKSs, the same two types of assembly are known except for the difference that the arrangement in type II PKSs is more like the one in NRPSs with repetitive domains, and the intermediates are passed from one ACP to another. A feature that is found exclusively in PKSs are several ACP domains in direct succession<sup>140, 141</sup>, though mutational studies have shown that only one of these domains is essential<sup>142</sup>. Conformational dynamics observed in ACPs from PKS are rather small and comprise the rearrangement of side chains as a result of phosphopantetheinylation<sup>143</sup> or different orientations of the ppan-arm to the ACP<sup>144</sup>. Loading of an ACP from a type II PKS showed no interaction with its small intermediates that were natural binding partners, but the same ACP formed a binding pocket for Fas, which are no natural binding partners<sup>145</sup>. In another case, a binding pocket for large intermediates in such an ACP was observed<sup>146</sup>.

Beside the biosynthesis of NRPs, FAs, and polyketides other processes like the bacterial cell-wall formation<sup>147</sup>, the metabolism of lysine<sup>148, 149</sup>, and the deformylation of *N*<sup>10</sup>-formyl tetrahydrofolate<sup>150</sup> are known to also have carrier protein bound intermediates.

### 1.3 Phosphopantetheine Transferases

The utilization of carrier proteins is accompanied by the need for PPTs, which convert them to their active *holo*-form. Structurally, three different groups of PPTs are known which all share a common domain as their structural basis (Figure 11). This

domain consists of a central three-stranded  $\beta$ -sheet and a mainly helical part that is folded on one side of the  $\beta$ -sheet, whereas the other side of the  $\beta$ -sheet functions as an interface for oligomerization.



**Figure 11:** PPT domains consist of a three stranded  $\beta$ -sheet (red) and a mainly helical part (blue). They are the basic structural element of every phosphopantetheine transferase, in which they either form intermolecular trimers like in the *B. subtilis* AcpS (group I), intramolecular pseudo-dimers like in Sfp (group II), or are incorporated into a synthase as in the yeast FAS (group III).

*E. coli holo* acyl carrier protein synthase (AcpS), the PPT modifying the ACP from the FAS, has been known for a long time<sup>151</sup> and further characterization revealed that it is a ~14 kDa protein which was proposed to form dimers<sup>152</sup>. However, the crystal structure of the *B. subtilis* AcpS has shown that the single polypeptide forms one PPT domain, and the active enzyme is a trimer<sup>153</sup> (PDB 1F7T). Co-crystallization of CoA (PDB 1F7L) and *holo*-ACP (PDB 1F80), respectively, has shown that the

oligomerization interfaces form the binding sites for both substrates, so each trimeric enzyme has three binding sites<sup>153</sup>. A second group of PPTs has about twice the mass of an AcpS monomer and does not form oligomers<sup>35</sup>. Structure elucidation of Sfp, a group II PPT involved in the production of surfactin A in *B. subtilis*, has revealed that these PPTs consist of two PPT domains that form an intramolecular pseudo-dimer<sup>154</sup> (PDB 1QR0). Again, the binding site for CoA is at the domain/domain interface, so each protein has one binding site. The third group of PPTs are those PPTs, that are no distinct proteins, but are incorporated in a multi-domain type I FAS. In the first crystal structure of such a FAS, the PPT domains were flexible and showed no electron density<sup>155</sup>. Other structures revealed that the PPT domains are at the outside of these dome-shaped synthases<sup>136, 156</sup> (PDB 2PFF and 3HMJ), so they are separated from the ACP domains. The six PPT domains which are present in these structures are all separated from each other, too, but the FAS still shows catalytic activity with respect to ppan transfer<sup>136</sup>. However, the excised PPT domains form trimers like AcpS<sup>136</sup>, and mutational studies have shown that there is an interaction between them<sup>157</sup>. Thus, it has been proposed that the phosphopantetheinylation of the ACP occurs before the FAS is assembled, and the activity of the assembled FAS results from transient dimer formation by PPT domain pairs that was not observed in the known structures, but is sterically possible<sup>136</sup>.

The natural target for group I PPTs are ACPs from bacterial FASs, whereas type I FASs with integral (group III) PPT domains are only found in fungi and some bacteria. Thus, group II PPTs are the most widespread type of PPTs. In some bacteria, they are loading the ACPs of the primary metabolism<sup>158</sup>, they are modifying the carrier proteins of the secondary metabolism in bacteria and fungi<sup>148, 159</sup>, and they are the only group of PPTs found in animals<sup>160</sup>. Remarkably, a group II PPT from yeast can be substituted by its bacterial or its human homolog *in vivo*<sup>158, 160</sup>, indicating a high functional conservation among this group of PPTs.

The function of all PPTs requires the presence of M<sup>2+</sup> (metal) ions and it has been shown that such an ion is coordinated by the PPT and the pyrophosphate of CoA in PPT/CoA complexes<sup>136, 153, 154</sup>. Group I and II PPTs seem to have only a low selectivity for substitutions at the thiol group of CoA<sup>161</sup>, so AcpS and Sfp have been used for loading of carrier proteins with CoA derivatives<sup>73, 145, 162-164</sup>.



## 1.4 Inter-domain Interactions in NRPSs

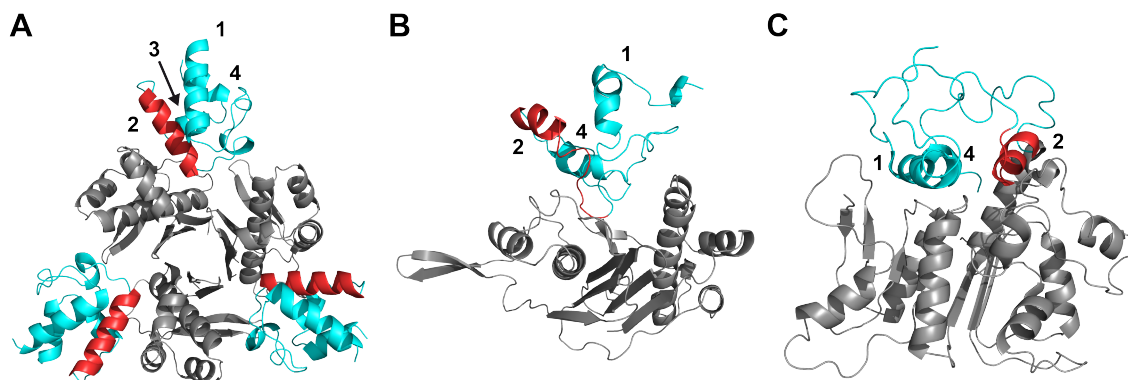
A crucial point in the function of an NRPS is the crosstalk between the domains. In the last years, great efforts have been made to get insight into the involved interactions and their selectivity<sup>25, 165-168</sup>. These studies mainly focused on the PCP as the central interaction partner for all other domains.

### 1.4.1 Domain/Domain Interactions

Following the logic of its catalytic cycle, a PPT is the first interaction partner of a PCP. It has been shown that Sfp can carry out the function of AcpS *in vitro* whereas the opposite is not the case<sup>35</sup>. From the crystal structure of an AcpS/ACP complex, it has become evident that the carrier protein binding is driven by a combination of hydrophobic interactions and polar contacts with a central role of helix 2 of the ACP<sup>153</sup> (Figure 12A). Comparison of ACPs and PCPs revealed that a residue 2 positions C-terminal to the active site serine with a positively charged side chain (lysine or arginine) is conserved in PCPs, whereas at the same position in ACPs, an aspartate is conserved that was involved in a salt bridge to AcpS in the crystal structure. Thus, by introduction of a K/D mutation at this position, a PCP could be converted into a substrate for AcpS; still, mutation of the complementary residue in AcpS was not sufficient to gain selectivity for a wild type PCP<sup>169</sup>. By NMR titration experiments it has been shown that the PCP interacts with Sfp selectively in its A state conformation: The proposed model (PDB 2GE1) gives no insight into the contribution of single side chains to the binding, but illustrates the necessity of the A state conformation to fit into the binding pocket<sup>38</sup> (Figure 12B). Mutational analysis by alanine scanning in a PCP revealed that the mutation of several residues flanking the active site serine disrupts the interaction with group II PPTs<sup>170, 171</sup>.

In case of mispriming, the acyl group tethered to the thiol group of ppan has to be removed from the cofactor of the PCP by a TE II. It has been shown that TE IIs are promiscuous for different PCPs, but do not interfere with the FAS by acting on acetyl- or malonyl-ACPs<sup>101</sup>. For the PCP/TE II complex only a NMR-based model with the PCP in the H state exists (PDB 2K2Q) which shows that helices 1, 2, and 4 are

involved in the binding, but lacks detailed insights into the protein/protein interactions at atomic resolution<sup>39</sup> (Figure 12C).



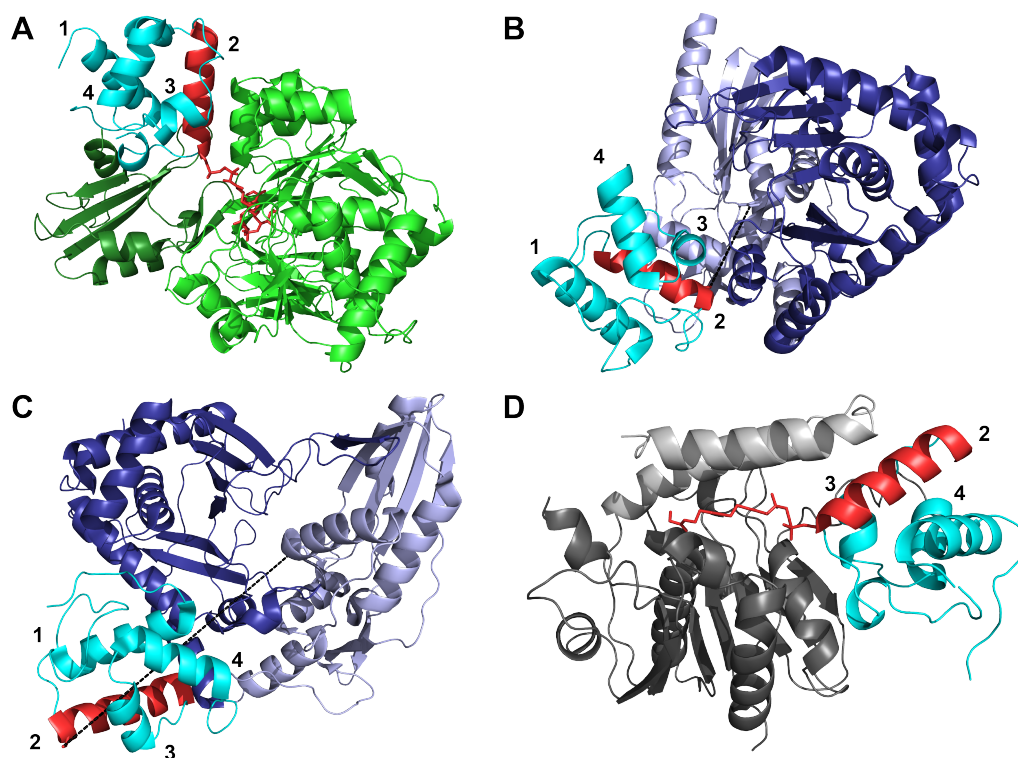
**Figure 12:** Known structures of carrier proteins in complex with interaction partners involved in the priming reaction. ACP (cyan) binds to group I PPTs (gray) via helix 2 and 3 and parts of the loop preceding helix 2 (A). For interaction with group II PPTs (gray), the PCP (cyan) binds in the A state via the N-terminal portion of helix 2, which is disordered in this conformation (B). In case of mispriming, the acylated PCP (cyan) undergoes a conformational change to the H state in which it binds to TE II (gray) via its three remaining helices. In the interest of easier comprehension, the residues forming helix 2 of the PCP in the A/H state are depicted in red, and all helices are numbered.

For the loading of the PCP by an A domain, structural information was gained from a crystal structure of an A-PCP bidomain by using an analogue of the adenylated amino acid, which binds to the active site of the A domain and selectively crosslinks to the ppan-arm of the PCP<sup>172</sup> (PDB 4DG9). In this structure, the PCP binds to the A domain, which is in the thioester/transfer orientation, via the N-terminal parts of helix 1 and 2 as well as the loop connecting them. In this orientation, helix 2 forms non-polar contacts to the N-terminal subdomain of the A domain, whereas helix 1 and the loop form a hydrogen bond network to the C-terminal subdomain. The crosslinked ppan-arm is not in contact with the PCP, but is located between the two subdomains of the A domain (Figure 13A). The presence of the cofactor seems to be necessary to stabilize the orientation of the PCP, as only the protein/protein interactions are not sufficient to lock the PCP in a distinct orientation in the crystal structure of the apo-form of the same A-PCP bidomain<sup>172</sup> (PDB 4DG8). Surprisingly, residues that were previously found to be involved in productive carrier protein/A domain interaction<sup>173</sup> were not in the region that shows direct contacts between the two proteins in the crystal structure.

In the crystal structure of SrfAC (PDB 2VSQ), the C domain and the downstream PCP, whose active site serine was mutated to alanine, were proposed to represent the acceptor site orientation<sup>71</sup>. In this situation, which is mainly stabilized by hydrophobic contacts, helix 2 of the PCP aligns parallel with the first helix of the C domain and helix 3 of the PCP with the tenth helix of the C domain. In this orientation, the mutated active site of the PCP and the catalytic histidine in the C domain are  $\sim 16$  Å apart. With a length of  $\sim 20$  Å, the ppan-arm could span this distance, so the observed orientation is likely to represent the acceptor site orientation (Figure 13B).

Up to now, only limited structural information for the interaction of the upstream PCP with the C domain is available. An *apo*-PCP-C bidomain has been crystallized (PDB 2JGP), but the observed orientation cannot represent the donor site complex, as the active site serine of the PCP and the catalytic histidine of the C domain are  $\sim 50$  Å apart<sup>69</sup> (Figure 13C). However, mutational analysis indicates a participation of helix 2 and 3 of the PCP in the interaction<sup>174</sup>.

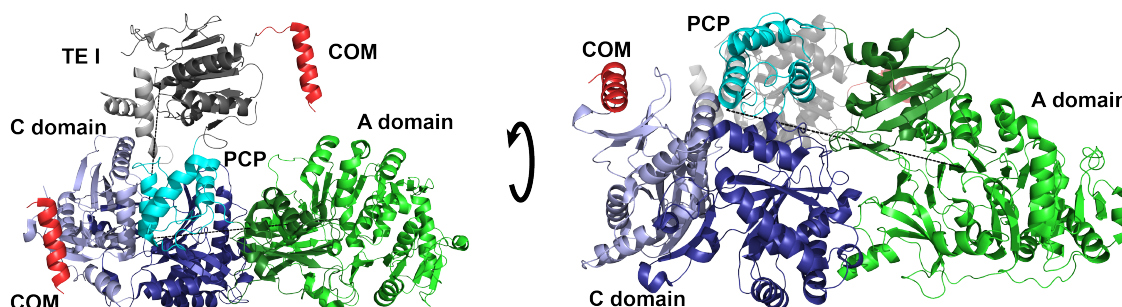
Two structures have been solved of PCP-TE I bidomains. The first structure was solved by liquid-state NMR with the active site serine of the PCP mutated to alanine<sup>94</sup> (PDB 2ROQ). In this structure, the PCP seems to be in the right orientation towards the TE I but is only in loose contact with it, so a second set of peaks in the NMR spectra were observed for some residues of the PCP. These signals were assigned to a different orientation of the PCP, in which it is not associated with the TE I<sup>94</sup>. After loading the same wild type bidomain with a CoA derivative, it was successfully crystallized<sup>93</sup>. In the crystal structure (PDB 3TEJ), the PCP is in closer contact with the TE I. The N-terminal part of helix 2, the loop preceding it, and helix 3 of the PCP are in contact with the core domain of TE I, whereas the middle part of helix 2 is in contact with the lid region. The ppan-arm, whose thiol group is substituted with  $\alpha$ -hydroxy-acetyl amide, lies in the region that was previously appointed to be its binding region<sup>91</sup>, and in doing so, it is pointing away from the PCP (Figure 13D). This structure is in agreement with the finding that the mutation of residues located in helix 3 of the PCP disrupts the productive interaction of the two domains<sup>171</sup>.



**Figure 13:** Structural insights into the maturation of NRPs. For the aminoacylation, the PCP (cyan with helix 2 in red; all helices numbered) binds to both subdomains of the A domain (light and dark green), so the ppan-arm can point to the active site, where it reacts with the adenylated amino acid (red sticks) (A). In the acceptor site orientation, the downstream PCP binds at one end of the catalytic groove to both subdomains of the C domain (light and dark blue). Although the ppan-arm is missing in this structure, it could span the distance between the active site of the PCP and the catalytic histidine of the C domain (dashed line) (B). For the donor site orientation, no structure is known up to now. In a known structure of a PCP-C bidomain, the PCP is located at the entry of the catalytic groove but its active site serine is pointing in the wrong direction, so the distance to the catalytic core of the C domain (dashed line) cannot be bridged by the cofactor (C). For its interaction with TE I, the PCP binds simultaneously to the core (dark gray) and the lid region (light gray), so the cofactor (red sticks) can point towards the catalytic triad of TE I (D).

Comparing these structures of PCPs bound to their interaction partners, it becomes evident that the same parts of a PCP are directly involved in several or all binding events. This finding underlines the need for large conformational changes of the synthetase during its catalytic cycle. Up to now, only one structure is known of an NRPS fragment which comprises of more domains than the PCP and one of its interaction partners: The aforementioned structure of the SrfAC termination module<sup>71</sup> consists of four domains (C-A-PCP-TE I), with the PCP oriented towards the acceptor site of the C domain. In this situation, the distances from the active site of the PCP to

the catalytic centers of the A domain and the TE I are  $\sim 56$  Å and  $\sim 44$  Å, respectively. Again, this emphasizes the need for large reorientations during the catalytic cycle (Figure 14).

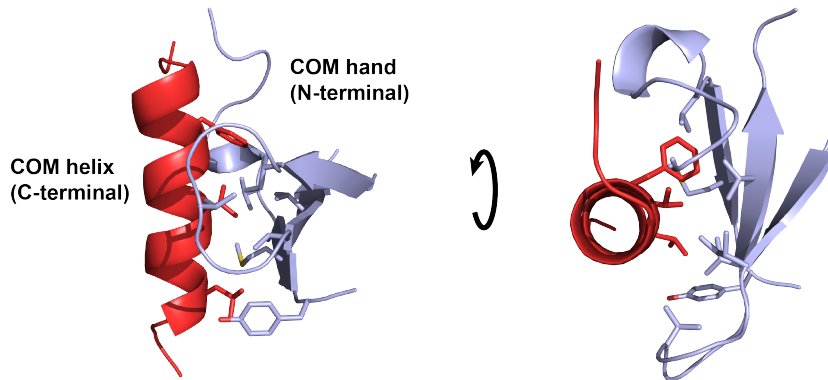


**Figure 14:** Overall structure of the SrfAC termination module. The PCP (cyan) is oriented towards the C domain (N-terminal subdomain light blue, C-terminal subdomain dark blue) so it is far apart from the catalytic centers of the A domain (N-terminal subdomain light green, C-terminal subdomain dark green) and the TE I (dark gray with lid region in light gray) as indicated by the dashed lines. The function of the COM domain (red) will be discussed in the next section.

In the structure of SrfAC, the C-terminal half of the C domain and the N-terminal subdomain of the A domain build a large interface with many directed interactions. Based on this observation, it was proposed that the C domain and the N-terminal subdomain of the A domain are in a rigid orientation to each other with the C-terminal subdomain of the A domain and the PCP moving dynamically around this “catalytic platform”<sup>71</sup>. This hypothesis is in agreement with the observed flexibility of these two (sub-)domains, although it might be modified by the recently proposed motion of the N-terminal subdomain of the C domain.

#### 1.4.2 Communication-mediating Domains

In NRPS systems, in which the assembly line is distributed among several enzymes, the correct intermolecular interaction between the different subunits has to be maintained. Interacting synthetases have short amino acid sequences at their termini, the communication-mediating domains (COM domains), which were found to be essential for this purpose *in vitro* and *in vivo*<sup>175</sup>. By crosslinking experiments, it has been shown that deletion of these regions disrupts any protein/protein contacts<sup>176</sup>, and the portability of COM domains to facilitate productive interactions of subunits from different NRPS systems has been demonstrated<sup>175</sup>.



**Figure 15:** The COM domain interaction observed in the crystal lattice of SrfAC. The C-terminus of the protein forms a helix (red), which binds intermolecularly to the hand motif formed by a three stranded  $\beta$ -sheet and its neighboring regions (light blue) in the N-terminal part of the C domain. The binding is mainly based on hydrophobic interactions (selected residues shown as sticks).

Up to now, structural information about the interaction of COM domains was only derived from an artificial system. The construct used for crystallization of SrfAC (PDB 2VSQ) has a C-terminal cloning tag that was incidentally found to mimic the COM domain of SrfAB. It forms a helix that interacts intermolecularly with a three-stranded  $\beta$ -sheet and its flanking loops from another SrfAC molecule in the crystal lattice (Figure 14+15). Contrary to previous assumptions, this so-called hand-motif is not only formed by the very N-terminus of SrfAC, but also by two  $\beta$ -strands which are part of the core of the N-terminal subdomain of the C domain<sup>71</sup>.

## 2. Materials

### 2.1 Equipment

Agar plate incubator	Memmert GmbH & Co. KG, Germany
Agilent 1100 Series HPLC Value System	Hewlett Packard GmbH, Germany
ÄKTAbasic FPLC system	GE Healthcare Europe GmbH, Germany
ÄKTAprime plus FPLC system	GE Healthcare Europe GmbH, Germany
Alpha 2-4 LSC freeze dryer	M. Christ Gefriertrockenanlagen GmbH, Germany
Amicon centrifugal filter units (3,500, 10,000 MWCO)	Millipore GmbH, Germany
Analytical balance CPA124S-OCE	Sartorius AG, Germany
Analytical balance PB3002 DeltaRange	Mettler-Toledo GmbH, Germany
Autoclave GE 446 EC-1	Getinge AB, Sweden
Avance NMR spectrometers (500, 600, 700, 800, 900, 950 MHz)	Bruker, Germany
BA-VC-300 H vacuum concentrator	Bachofer GmbH, Germany
Canon EOS 1000D	Canon Deutschland GmbH, Germany
CC 250/2 Nucleodur 100-3 C18ec column	Macherey-Nagel GmbH & Co. KG, Germany
CC 250/3 Nucleosil 120-3 C18 column	Macherey-Nagel GmbH & Co. KG, Germany
Centrifuge Centricon H-401	Kontron-Hermle, Germany
Centrifuge Megafuge 16R	Heraeus, Germany
Centrifuge rotor A 6.9	Kontron-Hermle, Germany
Centrifuge rotor F34-6-38	Eppendorf AG, Germany
Centrifuge rotor GS-3	Sorvall Instruments, Germany
Centrifuge rotor GSA	Sorvall Instruments, Germany
Centrifuge rotor SS-34	Sorvall Instruments, Germany
Centrifuge rotor TX-400	Heraeus, Germany
Centrifuge Sorvall Evolution RC	Sorvall Instruments, Germany
Centrifuge Sorvall RC-5B	Sorvall Instruments, Germany

## 2. MATERIALS

---

Centrifuge Sorvall RC-5C	Sorvall Instruments, Germany
Centriprep concentrators (YM-3, -10)	Millipore GmbH, Germany
Cryogenic Vials	Corning Inc, USA
CrystalMation crystallization robot	Rigaku Europe SE, Germany
Digital camera Powershot G3	Canon GmbH, Germany
FR-E+ diffractometer with	
Saturn 994+ CCD detector	Rigaku Europe SE, Germany
French Pressure Cell Press	SLM Instruments, USA
Heating block	VWR International GmbH, Germany
Hellmanex	Hellma GmbH & Co. KG, Germany
HiLoad 16/60 Superdex 200 PG	GE Healthcare Europe GmbH, Germany
HiLoad 16/60 Superdex 75 PG	GE Healthcare Europe GmbH, Germany
Incubation shaker Innova 4330	New Brunswick Scientif. GmbH, Germany
Incubation shaker Multitron	Infors-HT AG, Switzerland
JASCO J-810 CD-spectrometer	Jasco Germany GmbH, Germany
Kühl-Brutschrank 3001 incubator	Rubarth Apparate GmbH, Germany
Leica M205C	Leica Microsystems AG, Switzerland
Magnetic stirrer RTC basic	IKA-Werke GmbH & Co. KG, Germany
Membrane vacuum pump	VACUUBRAND GmbH & Co. KG, Germany
Milli-Q Academic ultrapure water system	Millipore GmbH, Germany
Mini-Protean Tetra Cell SDS-PAGE system	Bio-Rad Laboratories GmbH, Germany
Mosquito crystallization robot	TTP LabTech Inc, USA
NanoDrop 1000 UV-/Vis spectrophotometer	PEQLAB Biotechnology GmbH, Germany
NMR tube 528-PP-7	Wilmad-LabGlass Inc, USA
NMR tube BMS-3	Shigemi Inc, USA
PCR Cycler TPersonal	Biometra GmbH, Germany
PCR Gradient cycler peqSTAR	PEQLAB Biotechnology GmbH, Germany
PD 10 columns (1.0 and 2.5 mL)	GE Healthcare Europe GmbH, Germany
pH-meter PHM210	Radiometer-Analytical SAS, France
Pipettes Research and Research Plus	Eppendorf AG, Germany



Power supply Power PAC 3000	Bio-Rad Laboratories GmbH, Germany
Precision cell 110-QS	Hellma GmbH & Co. KG, Germany
PX3 beamline	Swiss Light Source, Switzerland
Refrigerated centrifuge 5810 R	Eppendorf AG, Germany
Refrigerated table top centrifuge Micro 22R	A. Hettich GmbH & Co. KG, Germany
Sepharose 6 Fast Flow	GE Healthcare Europe GmbH, Germany
Sep-Pak C-18 SPE column	Waters S.A.S., France
Slide-A-Lyzer dialysis cassettes (3,500, 10,000 MWCO)	Pierce / Thermo Scientific, Germany
Sonifier Labsonic U	B. Braun Biotech International, Germany
Spectra/Por dialysis membrane (1,000, 3,500, 10,000 MWCO)	Spectrum Europe B. V., Netherlands
Superdex 200 10/300 GL	GE Healthcare Europe GmbH, Germany
Syringe filter 0.22 µm polysulfone	Carl-Roth GmbH & Co. KG, Germany
Table top centrifuge 5415 D	Eppendorf AG, Germany
Table top centrifuge Biofuge 13	Heraeus, Germany
Universal indicator sticks (pH 6.0-7.7)	Carl-Roth GmbH & Co. KG, Germany
UV-/Vis-spectrometer Cary 3	Varian GmbH, Germany
Vortex MS2 minishaker	IKA-Werke GmbH & Co. KG, Germany
VP-ITC calorimeter	GE Healthcare Europe GmbH, Germany
Water bath TW12	JULABO Labortechnik GmbH, Germany

## 2.2 Chemicals and Reagents

If not indicated otherwise, all chemical and reagents were used in *pro analysi* (p. a.) quality.

<sup>13</sup> C-glucose	Cambridge Isotope Laboratories Inc, USA
<sup>15</sup> N-ammonium chloride ( <sup>15</sup> NH <sub>4</sub> Cl)	Cambridge Isotope Laboratories Inc, USA

1H-benzotriazole- <i>N,N,N',N'</i> -tetramethyl-uronium hexafluorophosphate (HBTU)	Merck Chemicals GmbH, Germany
1-Hydroxybenzotriazole (HOBt)	Merck Chemicals GmbH, Germany
2-Mercaptoethanol	Carl-Roth GmbH & Co. KG, Germany
4-(2-Hydroxyethyl)-1-piperazineethane-sulfonic acid (HEPES)	Carl-Roth GmbH & Co. KG, Germany
4,4-dimethyl-4-silapentane-1-sulfonic acid	Sigma-Aldrich Chemie GmbH, Germany
Acetic acid	Carl-Roth GmbH & Co. KG, Germany
Acetonitrile, HPLC grade (MeCN)	VWR International GmbH, Germany
Acrylamide solution (30% w/v)	Carl-Roth GmbH & Co. KG, Germany
Adenosine triphosphate (ATP)	Carl-Roth GmbH & Co. KG, Germany
Agar-agar	Carl-Roth GmbH & Co. KG, Germany
Ammonium chloride (NH <sub>4</sub> Cl)	Carl-Roth GmbH & Co. KG, Germany
Ammonium persulfate (APS)	Carl-Roth GmbH & Co. KG, Germany
Ampicillin sodium salt	Carl-Roth GmbH & Co. KG, Germany
Benzotriazol-1-yl-oxytripyrrolidino-phosphonium hexafluorophosphate (PyBOP)	Merck Chemicals GmbH, Germany
Boric acid (H <sub>3</sub> BO <sub>3</sub> )	Carl-Roth GmbH & Co. KG, Germany
Bromphenol blue sodium salt	Carl-Roth GmbH & Co. KG, Germany
Calcium chloride dihydrate (CaCl <sub>2</sub> *2H <sub>2</sub> O)	Carl-Roth GmbH & Co. KG, Germany
Choline chloride	Sigma-Aldrich Chemie GmbH, Germany
Cobalt(II) chloride hexahydrate (CoCl <sub>2</sub> *6H <sub>2</sub> O)	Sigma-Aldrich Chemie GmbH, Germany
Coenzyme A trilithium salt (CoA*Li <sub>3</sub> )	MP Biomedicals LLC, Germany
Coomassie brilliant blue	Sigma-Aldrich Chemie GmbH, Germany
Copper(II) chloride dihydrate (CuCl <sub>2</sub> *2H <sub>2</sub> O)	Sigma-Aldrich Chemie GmbH, Germany
Deuterium oxide (D <sub>2</sub> O)	Sigma-Aldrich Chemie GmbH, Germany
Diethyl ether	Carl-Roth GmbH & Co. KG, Germany
Disodium hydrogen phosphate dihydrate (Na <sub>2</sub> HPO <sub>4</sub> *2H <sub>2</sub> O)	Carl-Roth GmbH & Co. KG, Germany
Dithiothreitol (DTT)	Carl-Roth GmbH & Co. KG, Germany
Ethanol (EtOH)	Carl-Roth GmbH & Co. KG, Germany

Ethidium bromide	Carl-Roth GmbH & Co. KG, Germany
Ethylenediaminetetraacetic acid (EDTA)	Carl-Roth GmbH & Co. KG, Germany
Folic acid	Sigma-Aldrich Chemie GmbH, Germany
Glucose monohydrate	Carl-Roth GmbH & Co. KG, Germany
Glycerol	Carl-Roth GmbH & Co. KG, Germany
Hydrochloric acid, 32% (HCl)	Carl-Roth GmbH & Co. KG, Germany
Imidazole	Sigma-Aldrich Chemie GmbH, Germany
Iron(II) chloride tetrahydrate (FeCl <sub>2</sub> *4H <sub>2</sub> O)	Sigma-Aldrich Chemie GmbH, Germany
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Carl-Roth GmbH & Co. KG, Germany
Kanamycin sulfate	Carl-Roth GmbH & Co. KG, Germany
Magnesium chloride hexahydrate (MgCl <sub>2</sub> *6H <sub>2</sub> O)	Carl-Roth GmbH & Co. KG, Germany
Magnesium sulfate heptahydrate (MgSO <sub>4</sub> *7H <sub>2</sub> O)	Carl-Roth GmbH & Co. KG, Germany
Manganese(II)chloride tetrahydrate (MnCl <sub>2</sub> *4H <sub>2</sub> O)	Sigma-Aldrich Chemie GmbH, Germany
Myo-inositol	Sigma-Aldrich Chemie GmbH, Germany
Nicotinamide	Sigma-Aldrich Chemie GmbH, Germany
<i>N,N,N',N'</i> -tetramethylethane-1,2-diamine (TEMED)	Carl-Roth GmbH & Co. KG, Germany
<i>N,N</i> -Diisopropylethylamine (DIPEA)	Sigma-Aldrich Chemie GmbH, Germany
<i>N,N</i> -Dimethylformamide (DMF)	Sigma-Aldrich Chemie GmbH, Germany
<i>N</i> -(Tri(hydroxymethyl)methyl)glycine (TRICINE)	Carl-Roth GmbH & Co. KG, Germany
Pantothenic acid calcium salt	Sigma-Aldrich Chemie GmbH, Germany
Potassium chloride (KCl)	Carl-Roth GmbH & Co. KG, Germany
Potassium dihydrogen phosphate (KH <sub>2</sub> PO <sub>4</sub> )	Carl-Roth GmbH & Co. KG, Germany

Potassium hydroxide (KOH)	Carl-Roth GmbH & Co. KG, Germany
Pyridoxal hydrochloric	Sigma-Aldrich Chemie GmbH, Germany
Riboflavin	Carl-Roth GmbH & Co. KG, Germany
Sodium chloride (NaCl)	Carl-Roth GmbH & Co. KG, Germany
Sodium dihydrogen phosphate (NaH <sub>2</sub> PO <sub>4</sub> )	Carl-Roth GmbH & Co. KG, Germany
Sodium dodecyl sulfate (SDS)	Carl-Roth GmbH & Co. KG, Germany
Sodium hydroxide (NaOH)	Carl-Roth GmbH & Co. KG, Germany
Sodium molybdate dihydrate (Na <sub>2</sub> MoO <sub>4</sub> *2H <sub>2</sub> O)	Merck Chemicals GmbH, Germany
Thiamine hydrochloric	Sigma-Aldrich Chemie GmbH, Germany
Trichloroacetic acid (TCA)	Carl-Roth GmbH & Co. KG, Germany
Trifluoroacetic acid (TFA)	VWR International GmbH, Germany
Triisopropylsilane (TIPS)	Sigma-Aldrich Chemie GmbH, Germany
<i>Tris</i> (2-carboxyethyl)phosphine (TCEP)	Carl-Roth GmbH & Co. KG, Germany
<i>Tris</i> (hydroxymethyl)aminomethane (TRIS)	Carl-Roth GmbH & Co. KG, Germany
Tryptone	Carl-Roth GmbH & Co. KG, Germany
Urea	Carl-Roth GmbH & Co. KG, Germany
L-Valine, boc-protected	Merck Chemicals GmbH, Germany
Yeast extract	Carl-Roth GmbH & Co. KG, Germany
Zinc chloride dihydrate (ZnCl <sub>2</sub> *2H <sub>2</sub> O)	Merck Chemicals GmbH, Germany

Amino-pantetheine was synthesized in the group of Prof. M. A. Marahiel.

Protected peptides were obtained from in-house synthesis by the group of Prof. H. Schwalbe.

## 2.3 Primers

Oligo-nucleotide primers were purchased from BioSpring GmbH (Germany) and biomers.net GmbH (Germany).

### 2.3.1 Primers for PCR

In the oligonucleotide sequences, the restriction sites are depicted in bold and the annealing sequence in italic.  $T_M$  of the annealing sequences are given in parentheses.

tycC3_PCP(pBH4)-fw	ATAT <b>GGATCC</b> GCGCAATATGTCGCGCCG ( $T_M= 61^\circ\text{C}$ )
tycC3_PCP(pBH4)-rv	ATAT <b>CTCGAG</b> TTATTTTCCGCTCGTGGCGACATACTGG GCCAAC ( $70^\circ\text{C}$ )
tycC4_PCP-fw	CCCC <b>GGATCC</b> AGCGCCTTTGTCGCTGCG ( $61^\circ\text{C}$ )
tycC4_PCP-rv	CCCC <b>CTCGAG</b> CTACTCGAATGCGCTCTCGGC ( $61^\circ\text{C}$ )
tycC5-6_PCP-C-fw	ATAT <b>GGATCC</b> GAGTATGTAGCGCCGCGC ( $61^\circ\text{C}$ )
tycC5-6_PCP-C-rv	ATAT <b>CTCGAG</b> TTAAAGCATGTCGATCTCGCCC ( $59^\circ\text{C}$ )

### 2.3.2 Primers for Site Directed Mutagenesis

In the primer sequences, the mutated bases are depicted in bold and the  $T_M$  of the annealing sequences are given in parentheses.

sfp(Y36P)-fw	GCCGGAGATTT <b>CCT</b> CATAAAGAAGATGCTCACC ( $T_M= 71^\circ\text{C}$ )
sfp(Y36P)-rv	GGTGAGCATCTTCTTTATGAG <b>GG</b> AAATCTCCGGC ( $71^\circ\text{C}$ )
tycC3_PCP(G42A)-fw	CCAGATCG <b>CC</b> GGACATTCCTTGAAAGC ( $70^\circ\text{C}$ )
tycC3_PCP(G42A)-rv	GCTTTCAAGGAATGTCCG <b>G</b> CGATCTGG ( $70^\circ\text{C}$ )
tycC3_PCP(G42A,S45A)-fw	CCAGATCG <b>CC</b> GGACATGCCTTGAAAGC ( $71^\circ\text{C}$ )
tycC3_PCP(G42A,S45A)-rv	GCTTTCAAGGCATGTCCG <b>G</b> CGATCTGG ( $71^\circ\text{C}$ )
tycC3_PCP(G43A)-fw	CCAGATCGGCG <b>C</b> ACATTCCTTGAAAGC ( $70^\circ\text{C}$ )

tycC3_PCP(G43A)-rv	GCTTTCAAGGAATGT <b>G</b> CGCCGATCTGG (70°C)
tycC3_PCP(G43A,S45A)-fw	CCAGATCGGCG <b>C</b> ACATGCCTTGAAAGC (71°C)
tycC3_PCP(G43A,S45A)-rv	GCTTTCAAGGCATGT <b>G</b> CGCCGATCTGG (71°C)
tycC3_PCP(S45A)-fw	CGGCGGACAT <b>G</b> CCTTGAAAGC (63°C)
tycC3_PCP(S45A)-rv	GCTTTCAAGG <b>C</b> ATGTCCGCCG (63°C)
tycC3_PCP(L46A)-fw	CGGCGGACATTCC <b>G</b> CGAAAGCGATGG (71°C)
tycC3_PCP(L46A)-rv	CCATCGCTTT <b>C</b> GCGGAATGTCCGCCG (71°C)
tycC3_PCP(S45A,L46A)-fw	CGGCGGACATGCC <b>G</b> CGAAAGCGATGG (72°C)
tycC3_PCP(S45A,L46A)-rv	CCATCGCTTT <b>C</b> GCGGCATGTCCGCCG (72°C)
tycC3_PCP(L46D)-fw	CGGCGGACATTCC <b>G</b> ATAAAGCGATGG (68°C)
tycC3_PCP(L46D)-rv	CCATCGCTTT <b>A</b> T <b>C</b> GGAATGTCCGCCG (68°C)
tycC3_PCP(S45A,L46D)-fw	CGGCGGACATGCC <b>G</b> ATAAAGCGATGG (70°C)
tycC3_PCP(S45A,L46D)-rv	CCATCGCTTT <b>A</b> T <b>C</b> GGCATGTCCGCCG (70°C)
tycC3_PCP(L46N)-fw	CGGCGGACATTCC <b>A</b> ATAAAGCGATGG (68°C)
tycC3_PCP(L46N)-rv	CCATCGCTTT <b>A</b> TTGGAATGTCCGCCG (68°C)
tycC3_PCP(S45A,L46N)-fw	CGGCGGACATGCC <b>A</b> ATAAAGCGATGG (70°C)
tycC3_PCP(S45A,L46N)-rv	CCATCGCTTT <b>A</b> TTGGCATGTCCGCCG (70°C)
tycC3_PCP(M49D)-fw	CCTTGAAAGCG <b>G</b> ATGCTGTGCTGC (68°C)
tycC3_PCP(M49D)-rv	GCAGCGACAGC <b>A</b> T <b>C</b> CGCTTTCAAGG (68°C)
tycC3_PCP(S45A,M49D)-fw	GCCTTGAAAGCG <b>G</b> ATGCTGTGCTGC (70°C)
tycC3_PCP(S45A,M49D)-rv	GCAGCGACAGC <b>A</b> T <b>C</b> CGCTTTCAAGGC (70°C)

## 2.4 Common Media and Buffers

### 2.4.1 Media and Solutions for Cultivation of Microorganisms

All media were autoclaved for 30 min at 121°C prior to use. Stock solutions were stored at -20°C.

#### 2.4.1.1 Lysogeny Broth (LB) Medium

10 g/L tryptone

5 g/L yeast extract

10 g/L NaCl

were dissolved in desalted water.

#### 2.4.1.2 2xYT Medium

16 g/L tryptone  
10 g/L yeast extract  
5 g/L NaCl

were dissolved in desalted water.

#### 2.4.1.3 M9 Minimal Medium

8.5 g  $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$   
3 g  $\text{KH}_2\text{PO}_4$   
0.5 g NaCl

were dissolved in 970 mL ultrapure water (ddH<sub>2</sub>O). After the pH was adjusted to 7.25 with 10 M NaOH, the solution was autoclaved followed by supplementation with:

2 mL of a 1 M  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$  solution (autoclaved)  
1 mL vitamin mixture  
2 mL Solution Q  
4.4 g glucose monohydrate dissolved in 20 mL ddH<sub>2</sub>O (sterile-filtered)  
1 g  $\text{NH}_4\text{Cl}$  dissolved in 5 mL ddH<sub>2</sub>O (sterile-filtered)

For isotopic labeling  $^{15}\text{NH}_4\text{Cl}$  was used and, if indicated, glucose monohydrate was substituted with 2 g/L uniformly  $^{13}\text{C}$ -labeled glucose.

#### 2.4.1.4 SOC Medium

20 g/L tryptone  
5 g/L yeast extract  
0.5 g/L NaCl

were dissolved in ddH<sub>2</sub>O and the pH was adjusted to 7.0 with 10 M NaOH. After autoclaving sterile solutions of

40 mL KCl (1 M; autoclaved)  
10 mL  $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$  (1 M; autoclaved)  
20 mL glucose monohydrate (1 M; sterile-filtered)

were added.

### 2.4.1.5 LB Agar

The same solution like for LB medium with additional 15 g/L agar-agar was made.

### 2.4.1.6 Antibiotic Stocks

Solutions of 30 mg/mL kanamycin sulfate in ddH<sub>2</sub>O and 100 mg/mL ampicillin (sodium salt) in 50% (v/v) aqueous ethanol were used as stocks with 1000x concentration.

### 2.4.1.7 IPTG

A 1 M, sterile-filtered solution of IPTG was used as stock solution for the induction of protein expression with final concentrations ranging from 0.1 to 1.0 mM.

### 2.4.1.8 Vitamin Mixture

20 mg choline chloride

25 mg folic acid

25 mg pantothenic acid

25 mg nicotinamide

50 mg myo-inositol

25 mg pyridoxal hydrochloride

25 mg thiamin hydrochloride

2.5 mg riboflavin

were dissolved in 50 mL ddH<sub>2</sub>O. After sterile filtration, the solution was used as a 1000x stock.



#### 2.4.1.9 Solution Q

40 mg  $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$

605 mg  $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$

340 mg  $\text{ZnCl}_2$

4 mg  $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$

18 mg  $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$

64 mg  $\text{H}_3\text{BO}_3$

184 mg  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$

5000 mg  $\text{FeCl}_2 \cdot 4\text{H}_2\text{O}$

were suspended in one liter ddH<sub>2</sub>O. 5 mL hydrochloric acid (5 M) were added, and the suspension was autoclaved.

#### 2.4.2 Buffers for Protein Purification

The pH was adjusted at room temperature. All buffers for protein purification were sterile-filtered. In addition, buffers for size exclusion chromatography (SEC) were degassed.

##### 2.4.2.1 Buffers for Nickel Affinity Purification

5.96 g/L (25 mM) HEPES

14.61 g/L (250 mM) NaCl

1.70 g/L (25 mM) imidazole for buffer A

or 17.02 g/L (250 mM) imidazole for buffer B

were dissolved in ddH<sub>2</sub>O, and the pH was adjusted to 8.0 with 10 M NaOH.

##### 2.4.2.2 Buffers for Reversed Nickel Affinity Purification

5.96 g/L (25 mM) HEPES

2.92 g/L (50 mM) NaCl

1.70 g/L (25 mM) imidazole for buffer A

or 17.02 g/L (250 mM) imidazole for buffer B

were dissolved in ddH<sub>2</sub>O, and the pH was adjusted to 7.2 with 10 M NaOH.

### 2.4.2.3 Imidazole Stock Solution

The solution was prepared by dissolving 68.08 g (1 mol) in 175 mL ddH<sub>2</sub>O. After the pH was adjusted to 7.2 with hydrochloric acid and the volume was set to 250 mL, the solution was sterile-filtered.

### 2.4.2.4 Buffer for Cleavage with TEV Protease

3.03g/L (25 mM) TRIS

2.92 g/L (50 mM) NaCl

0.15 g/L (1 mM) DTT

0.15 g/L (0.5 mM) EDTA

were dissolved in ddH<sub>2</sub>O, and the pH of 7.2 was adjusted with hydrochloric acid.

### 2.4.2.5 Buffer for NMR

4.06 g/L (26 mM) NaH<sub>2</sub>PO<sub>4</sub>

4.27 g/L (24 mM) Na<sub>2</sub>HPO<sub>4</sub>\*2H<sub>2</sub>O

were dissolved in ddH<sub>2</sub>O. The pH was 6.8 without further adjustment.

### 2.4.2.6 Buffer for Crystallization

5.96 g/L (25 mM) HEPES

2.92 g/L (50 mM) NaCl

were dissolved in ddH<sub>2</sub>O, and the pH was adjusted with 10 M NaOH to 7.2.

### 2.4.2.7 Buffer for Phosphopantetheinylation

5.96 g/L (25 mM) HEPES

2.92 g/L (50 mM) NaCl

1.02 g/L (5 mM) MgCl<sub>2</sub>\*6H<sub>2</sub>O

were dissolved in ddH<sub>2</sub>O. The pH was adjusted to 7.2 with 10 M NaOH.

### 2.4.3 Buffers for DNA Preparation

#### 2.4.3.1 TRIS, Acetate and EDTA Buffer (TAE)

242.28 g/L (2 M) TRIS

5.7% (v/v) mL (1 M) acetic acid ( $\rho=1.05$  g/mL)

14.61 g/L (50 mM) EDTA

were dissolved in ddH<sub>2</sub>O and pH 8.0 adjusted with 10 M NaOH. For use, this stock was diluted 50x with ddH<sub>2</sub>O.

#### 2.4.3.2 DNA Sample Buffer

The buffer was made of 37.5 mg bromphenol blue (sodium salt) dissolved in 50 mL 45% (v/v) glycerol. The buffer was used with final dilutions of 4-6x.

#### 2.4.3.3 DNA Storage Buffer

0.61 g/L (5 mM) TRIS were dissolved in 900 mL ddH<sub>2</sub>O, and the pH was adjusted to 8.0 with hydrochloric acid. After the volume was set to 1 L, the solution was sterile-filtered.

### 2.4.4 Buffers for Polyacrylamide Gel Electrophoresis

The pH of the buffers for polyacrylamide gel electrophoresis (PAGE) was adjusted at room temperature.

#### 2.4.4.1 Gel Buffer for PAGE

363.42 g/L (3 M) TRIS

3.00 g/L (10.4 mM) SDS

were dissolved in ddH<sub>2</sub>O, and the pH was adjusted to 8.45 with hydrochloric acid.

### 2.4.4.2 Separation Gel for PAGE

For six gels

11 mL acrylamide solution (30% w/v)

10 mL gel buffer for PAGE

3 mL glycerol

6 mL ddH<sub>2</sub>O

were mixed. The polymerization was started by adding 150  $\mu$ L APS solution (10% w/v) and 15  $\mu$ L TEMED.

### 2.4.4.3 Stacking Gel for PAGE

For six gels

0.825 mL acrylamide solution (30% w/v)

3.175 mL ddH<sub>2</sub>O

2 mL gel buffer for PAGE

were mixed. The polymerization was started by adding 30  $\mu$ L APS solution (10% w/v) and 12  $\mu$ L TEMED.

### 2.4.4.4 Running Buffers for PAGE

As anode buffer, a 100 mM (12.11 g/L) TRIS solution with pH 8.9, adjusted with hydrochloric acid, was used. The cathode buffer contained 100 mM (12.11 g/L) TRIS, 100 mM (17.92 g/L) TRICINE and 0.1% (w/v) SDS. The pH of the cathode buffer was ~8.25 without further adjustment.

### 2.4.4.5 Sample Buffer for PAGE

The buffer was a 4x stock solution with

25% (v/v) 1 M TRIS (pH 8.0)

7.5% (w/v) SDS

25% (v/v) glycerol

12.5% (v/v) 2-mercaptoethanol

0.025% (w/v) bromphenol blue sodium salt.

#### 2.4.4.6 Buffers for Staining and Destaining

The buffer for staining of SDS-PAGE gels was a solution of 0.25 g coomassie brilliant blue G250 in 10% (v/v) aqueous acetic acid. For destaining a solution containing 10% (v/v) acetic acid and 40% (v/v) methanol was used.

## 2.5 Enzymes

Antarctic Phosphatase	New England Biolabs GmbH, Germany
<i>Bam</i> HI-HF	New England Biolabs GmbH, Germany
DPCCK (Dephosphocoenzyme A kinase)	Marahiel group (plasmid from Wright Lab <sup>177</sup> )
<i>Dp</i> nl	New England Biolabs GmbH, Germany
<i>Nco</i> I-HF	New England Biolabs GmbH, Germany
<i>Xho</i> I	New England Biolabs GmbH, Germany
PanK (Pantothenate kinase)	Marahiel group (plasmid from Wright Lab <sup>177</sup> )
PfuTurbo DNA polymerase	Agilent Technologies Deutschland GmbH, Germany
PPAT (Ppan adenylyltransferase)	Marahiel group (plasmid from Wright Lab <sup>177</sup> )
Sfp	self-produced
T4 DNA ligase	New England Biolabs GmbH, Germany
TEV protease	self-produced
Vent DNA polymerase	New England Biolabs GmbH, Germany

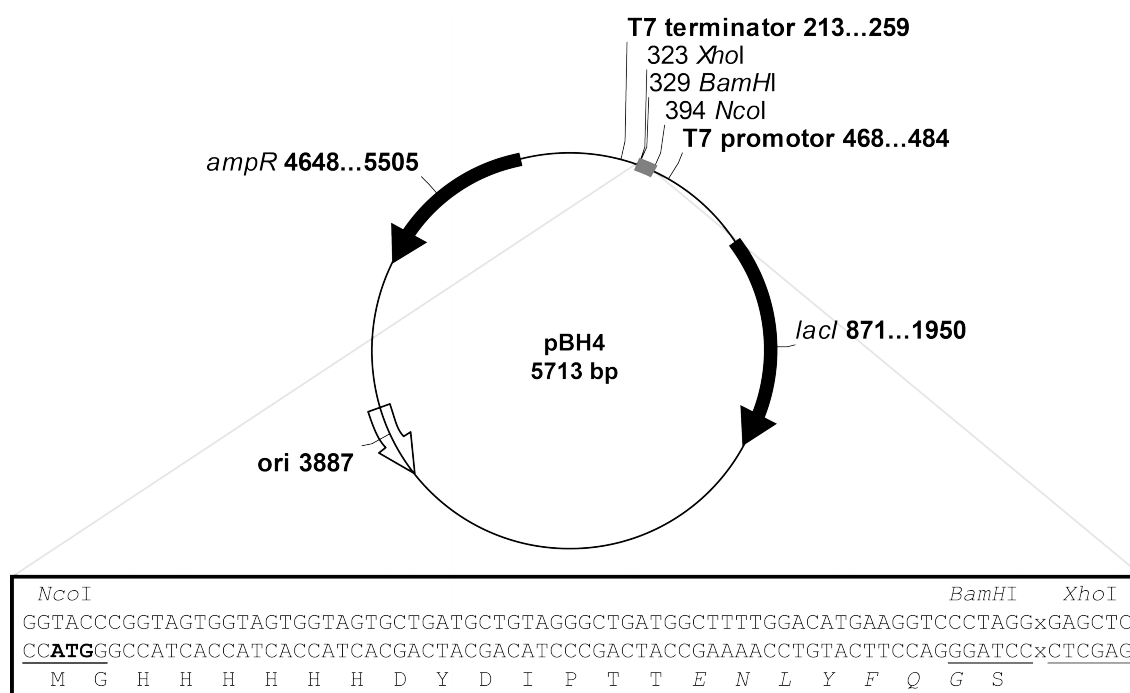
## 2.6 Bacterial strains

NEB 5-alpha Competent <i>E. coli</i>	New England Biolabs GmbH, Germany
T7 Express Competent <i>E. coli</i>	New England Biolabs GmbH, Germany
<i>E. coli</i> M15(pREP4)	Qiagen GmbH, Germany

## 2.7 Plasmids

pQE60	Qiagen GmbH, Germany
pQE70	Qiagen GmbH, Germany
pET28a	Novagen GmbH, Germany
pBH4 (variant of pET15b)	Gift from Lim Lab, UCSF

Detailed information about the purchasable plasmids is available at the distributors' homepages. pBH4 was derived from pET15b (Novagen GmbH, Germany) by inserting a sequence coding for a six- or tenfold his-tag and a recognition site for the TEV protease between the *NcoI* and *BamHI* restriction sites (Figure 16).



**Figure 16:** Vector map of pBH4(His6) with a detailed view of the sequence coding for the his-tag, the recognition site for TEV protease and the gene of interest. The start codon is depicted in bold, and the gene of interest (x) is located between the *BamHI* and *XhoI* restriction sites. In case of pBH4(His10), five of the his-tag coding CATCAC-repeats are present.

## 2.8 Nucleic Acids Purification Kits

JETsorb Gel Extraction	Genomed GmbH, Germany
NucleoBond PC 100 Midi	Macherey-Nagel GmbH & Co. KG, Germany
NucleoSpin Plasmid	Macherey-Nagel GmbH & Co. KG, Germany
QIAquick PCR Purification	Qiagen GmbH, Germany

## 2.9 Crystallization Kits

Clear Strategy Screen I+II	Molecular Dimensions Ltd., UK
Crystal Screen 1+2	Hampton Research Corp., USA
Crystal Screen Cryo 1+2	Hampton Research Corp., USA
Index	Hampton Research Corp., USA
JCSG Core Suite I-IV	Qiagen GmbH, Germany
JCSG+ Suite	Qiagen GmbH, Germany
MembFac	Hampton Research Corp., USA
Morpheus	Molecular Dimensions Ltd., UK
Natrix 1+2	Hampton Research Corp., USA
PEG/Ion 1+2	Hampton Research Corp., USA
PEGRx 1+2	Hampton Research Corp., USA
SaltRx 1+2	Hampton Research Corp., USA

## 2.10 Software and Online-Tools

Adobe Creative Suite 5	Adobe Systems, Germany
ApE 1.17	M. W. Davis, University of Utah, USA
BioDocAnalyze 2.0	Biometra GmbH, Germany
ChemStation B.01.03	GE Healthcare Europe GmbH, Germany
Chromas LITE 2.01	Technelysium Pty Ltd, Australia
Clustal Omega ( <a href="http://www.ebi.ac.uk/Tools/msa/clustalo/">http://www.ebi.ac.uk/Tools/msa/clustalo/</a> ) <sup>178</sup>	EMBL-EBI, UK
COOT 0.7 <sup>179</sup>	P. Emsley, University of York, UK

CrystalTrak software	Rigaku Europe SE, Germany
CYANA 3.96 <sup>180, 181</sup>	P. Güntert, Universität Frankfurt, Germany
HKL-2000 <sup>182</sup>	HKL Research Inc, USA
MOLREP 11.0 <sup>183</sup>	A. A. Vagin, University of York, UK
MOSFLM 7.0.3 <sup>184</sup>	A. G. W. Leslie, H. Powell, MRC Laboratory of Molecular Biology, UK
mosquito v3.5.0	TTP LabTech Inc, USA
MS Office Professional Plus 2010	Microsoft GmbH, Germany
Oligo Calculator v3.26 ( <a href="http://www.basic.northwestern.edu/biotools/oligocalc.html">http://www.basic.northwestern.edu/biotools/oligocalc.html</a> ) <sup>185</sup>	W. Kibbe, Northwestern University, USA
OPALp <sup>186</sup>	P. Güntert, Universität Frankfurt, Germany
Origin 6.1	OriginLab Corporation, USA
Protein Calculator v3.3 ( <a href="http://www.scripps.edu/~cdputnam/">http://www.scripps.edu/~cdputnam/</a> )	C. Putnam, Scripps Institute, USA
Pymol 1.3	protcalc.html
REFMAC 5.5 <sup>187</sup>	Schrodinger LLC, USA
Sparky 3.114	G. Murshudov, University of York, UK
Spectra Manager 2.08.02	T. D. Goddard and D. G. Kneller, UCSF, USA
TALOS+ <sup>188</sup>	Jasco Germany GmbH, Germany
Topspin (2.1, 3.0 and 3.1)	Y. Shen, National Institutes of Health, USA
UNICORN 5.11	Bruker BioSpin GmbH, Germany
VP Viewer 2000	GE Healthcare Europe GmbH, Germany
	GE Healthcare Europe GmbH, Germany



## **3. Methods**

### **3.1 DNA Techniques**

#### **3.1.1 Polymerase Chain Reaction**

For amplification of DNA fragments, polymerase chain reaction<sup>189, 190</sup> (PCR) with Vent DNA polymerase was used. Standard reactions were made in a 50  $\mu$ L scale with 25-100 ng template DNA and 2  $\mu$ M forward and reverse primer. The annealing temperature was set 1-5°C below the melting temperature ( $T_M$ ), as predicted with the Oligo Calculator web-tool<sup>185</sup>.

#### **3.1.2 Analytical and Preparative Agarose Gel Electrophoresis**

Analytical gels were cast from 30 mL molten 1% (w/v) agarose in TAE buffer supplemented with ethidium bromide. A sample of 3  $\mu$ L was mixed with 1  $\mu$ L DNA sample buffer and loaded on the gel. Separation was achieved by application of a constant voltage of 110 V for 30 min with TAE as running buffer. For documentation photos of the gels trans-illuminated with UV-light were taken.

Preparative gels were cast accordingly from 60 mL molten 1% (v/v) agarose in TAE buffer. 40  $\mu$ L sample were mixed with 8  $\mu$ L DNA sample buffer, and then the gels were run for 60 min at 120 V.

#### **3.1.3 PCR Purification**

The purity of PCR products was determined by analytical agarose gel electrophoresis.

Reactions containing only the desired DNA fragment were purified with the QIAquick PCR Purification Kit following the manufacturer's instructions. Reactions containing side products were purified by preparative agarose gel electrophoresis, and the DNA was extracted with the JETsorb Gel Extraction Kit according to the

manufacturer's protocol. In both cases, 5 mM TRIS (pH 8.0) was used as elution buffer and the DNA was stored at -20°C for further use.

#### **3.1.4 Restriction Digestion**

Purified PCR products and plasmids were digested with the appropriate restriction endonucleases for 3 h at 37°C. Buffer conditions and enzyme concentrations were chosen following the manufacturer's recommendations for a 60 µL reaction scale.

#### **3.1.5 Dephosphorylation of DNA**

Digested, purified plasmid DNA was dephosphorylated prior to ligation. The 5'-phosphate group was removed by incubation for 2 h at 37°C with antarctic phosphatase in the provided reaction buffer. After heat inactivation of the enzyme for 10 min at 70°C, the plasmid was stored at -20°C.

#### **3.1.6 Site Directed Mutagenesis**

Point mutations were introduced by site directed mutagenesis, using PfuTurbo DNA polymerase in a 50 µL scale with ~50 ng template DNA and both primers at final concentrations of 0.2 µM. Annealing temperatures were varied between +1°C and -6°C relative to  $T_M$  of the primers, as predicted by Oligo Calculator.

The reaction products were incubated with *DpnI* for 6-18 h at 37°C and could be stored at -20°C prior to transformation.

#### **3.1.7 DNA Ligation**

Ligations were performed in 20 µL reactions containing 2 µL T4 DNA ligase and the buffer according to the manufacturer's protocol. The vector DNA concentration ranged between 10-50 ng, and the insert was used in molar ratios of 1:3 or 1:5 with respect to the vector as approximated by (1) using the lengths  $l$  of vector and insert in base pairs (bp).

---

$$m(\text{insert}) = \text{ratio} \frac{m(\text{vector}) \cdot l(\text{insert})}{l(\text{insert})} \quad (1)$$

Ligations were carried out for 1 h at room temperature (RT).

### 3.1.8 Transformation of Ligated DNA or DNA from Site Directed Mutagenesis

10  $\mu\text{L}$  of a ligation or mutagenesis reaction were added to a 100  $\mu\text{L}$  aliquot of the appropriate chemically competent *E. coli* cells (M15(pREP4) for pQE vectors, DH5 $\alpha$  for others) thawed on ice. The cells were incubated on ice for 20 min, followed by a heat shock for 45 s at 42°C. After another incubation on ice for 3 min, 650  $\mu\text{L}$  SOC medium were added, and the cells were grown for 60 min at 37°C/180 rpm. Then the cells were pelleted by centrifugation for 2 min at 1000\*g, the pellet was resuspended in 100  $\mu\text{L}$  SOC and spread on an agar plate with the desired antibiotics and incubated over night (ON) at 37°C.

### 3.1.9 Transformation of Purified DNA

Transformation of purified DNA was performed like for ligation reactions, but with 1  $\mu\text{L}$  containing ~50 ng plasmid DNA. The first incubation step on ice was shortened to 5 min, and for vectors with ampicillin resistance, the cells were not grown in SOC, but directly spread on an agar plate.

### 3.1.10 Preparation of Cryogenic Cultures

For one culture, 5 mL LB medium supplemented with the appropriate antibiotics were inoculated with a single colony from an agar plate, and the cells were grown at 37°C/180 rpm. At an optical density measured at 600 nm ( $\text{OD}_{600}$ ) of ~0.5, 750  $\mu\text{L}$  of the culture were transferred into a sterile cryogenic vial, mixed with the same volume glycerol, shock-frozen in liquid nitrogen and stored at -80°C.

#### **3.1.11 Plasmid DNA Preparation**

5 mL LB medium supplemented with the appropriate antibiotics were inoculated either with a single colony from an agar plate, or a small portion of a cryogenic culture. Cells were grown ON at 37°C/180 rpm and pelleted by centrifugation for 1 min at 13000 rpm. Plasmid DNA was extracted from the cell pellet with the NucleoSpin Plasmid Kit according to the user's manual with 5 mM TRIS (pH 8.0) as elution buffer.

For larger quantities, 250 mL cultures were grown in the same way, and DNA was extracted with the NucleoBond PC 100 Midi kit with ddH<sub>2</sub>O for the final dissolving of the DNA.

#### **3.1.12 Quantification of DNA**

The DNA concentration was quantified by the absorption at 260 nm, measured at a NanoDrop 1000 UV/Vis spectrophotometer with the respective buffer as blank reference.

#### **3.1.13 DNA Sequencing**

Sequencing of plasmid DNA was performed by Seqlab – Sequence Laboratories Göttingen GmbH (Germany) in the extended Hotshot mode. Samples were prepared in 7 µL scale containing 200-600 ng DNA and 3 µM primer.

### **3.2 Protein Expression and Purification**

#### **3.2.1 Expression of Proteins in Rich Media**

*E. coli* M15(pREP4) cells were used for expression with the pQE vector system. For other constructs, *E. coli* NEB T7 cells were used.

Precultures were made with 250 mL 2xYT medium by inoculation either from a cryogenic culture and subsequent growth ON at 37°C/180 rpm or with all clones from a freshly prepared agar plate, followed by incubation for 1-3 h at 37°C/180 rpm.

For the expression culture, LB or 2xYT medium was inoculated with 10 mL preculture per liter medium and grown at 37°C and 180 rpm. When the OD<sub>600</sub> was ~0.4, the temperature was lowered to the desired expression temperature. 20 min later, the expression was induced by addition of IPTG with final concentrations of 0.1-1.0 mM. After an appropriate time, cells were harvested by centrifugation for 10 min at 6000 rpm at 4°C.

If the cell pellet was not lysed immediately, it was stored at -20°C for further use.

### **3.2.2 Expression of Proteins in M9 Minimal Medium**

LB medium was inoculated with a preculture, prepared like for the expression in rich medium and incubated at 37°C/180 rpm. When the OD<sub>600</sub> was 0.8-1.0, the cells were pelleted by centrifugation for 10 min at 4000 rpm at RT, and the cell pellet from 1 L was carefully resuspended in ~20 mL supplemented M9 medium. The resuspended cells were used to inoculate supplemented M9 medium to a theoretical OD<sub>600</sub> of 0.2. The cells were grown at 37°C/180 rpm to an OD<sub>600</sub> of ~0.4. At this point, the temperature was reduced to the expression temperature, and after another 30 min the expression was started by addition of IPTG. At the end of the expression, the cells were harvested by centrifugation for 10 min at 6000 rpm at 4°C, and the resulting pellet was either lysed instantly or stored at -20°C for further use.

### **3.2.3 Cell Free Expression of Proteins**

For the cell free expression, a T7 RNA polymerase based coupled transcription/translation system, developed in our lab<sup>191</sup>, was used. After a small-scale screening for optimal magnesium concentrations in the reaction mix, 2 mL preparative-scale expressions for each construct were made ON at 30°C. As no soluble proteins were obtained, only the protein pellets derived by centrifugation for

10 min at 4°C/13000 rpm were used furthermore. They were dissolved in 6 M urea, and the proteins were refolded by dialysis.

#### **3.2.4 Cell Lysis**

The cell pellet from 2 L expression was thawed on ice and resuspended into 25 mL buffer A for nickel affinity purification. The cells were lysed by application of 30 ultrasonic pulses, each with a duration of 0.6 s and an intensity of 200 W, followed by an incubation for 2 min on ice. This procedure was repeated five times, and the lysate was clarified by centrifugation for 30 min at 38000\*g at 4°C.

#### **3.2.5 Immobilized Metal Ion Affinity Purification**

For this purification, a column packed with 12 mL Sepharose 6 Fast Flow resin loaded with nickel ions was used. The column was connected to an ÄKTAbasic system running with Unicorn 5.11 and operated at 4°C. During equilibration with buffer A and the following purification, the absorption of the eluate at 280 nm was monitored.

The clarified lysate was loaded onto the column at a flow rate of 1 mL/min. Then the column was washed with 5 column volumes (CV) buffer A at a flow rate of 1 mL/min, followed by elution of the protein with a linear gradient to 60% buffer B over 8 CV with a flow rate of 3 mL/min. During the elution, 3 mL fractions were collected. The fractions were analyzed by SDS-PAGE, and the fractions containing the target protein were pooled for further use.

After usage, the column was washed with 3 CV buffer B and 5 CV ddH<sub>2</sub>O, followed by equilibration with 3 CV 20% (v/v) EtOH for storage.

#### **3.2.6 Reversed Immobilized Metal Ion Affinity Purification**

For purification, a column packed with 5 mL nickel ion loaded Sepharose 6 Fast Flow resin was used. The column was used at 4°C in the gravity flow mode. Prior to use, the column was washed with ddH<sub>2</sub>O and equilibrated with 3 CV buffer A. The

sample was mixed with an imidazole stock solution to a final concentration of 25 mM, centrifugated for 10 min at 18000\*g, and loaded onto the column. The first 2 mL of the flow through were discarded, and the rest was collected. After loading of the sample, the column was washed with 2.5 CV buffer A, and the flow through was collected.

Afterwards, the collected flow through was combined for further use, and the column was washed with 4 CV buffer B, 3 CV ddH<sub>2</sub>O and 3 CV 20% (v/v) EtOH for storage.

### 3.2.7 Size Exclusion Chromatography

To prevent precipitation during the runs, every sample was subjected to dialysis against the running buffer prior to size exclusion chromatography (SEC). Then it was centrifuged for 10 min at 18000\*g at 4°C before injection. For proteins with a theoretical molecular weight larger than 50 kDa a Superdex 200, for smaller proteins a Superdex 75 was used.

#### 3.2.7.1 Preparative Scale SEC

After a Superdex 16/60 column was connected to an ÄKTAprime system operated at 4°C, it was washed with 140 mL degassed ddH<sub>2</sub>O, followed by 140 mL buffer at a flow rate of 1.0 mL/min.

5 mL sample, containing more than 5 mg protein, were injected onto the column and separated at a flow rate of 1.0 mL/min with continuous monitoring of the UV-absorption at 280 nm. Between elution volumes of 35 and 140 mL, fractions of 3 mL were collected. After the purification, the column was washed with 140 mL degassed ddH<sub>2</sub>O and 140 mL degassed 20% (v/v) EtOH for storage. Fractions containing monomeric protein as approximated by their apparent molecular weight were pooled for further use.

#### 3.2.7.2 Semi-preparative Scale SEC

Semi-preparative SEC was used for protein amounts of less than 5 mg. For semi-preparative SEC, a Superdex 10/300 column was connected to an ÄKTAbasic system operated at 4°C. The column was equilibrated with 30 mL degassed ddH<sub>2</sub>O and 30 mL buffer at a flow rate of 0.5 mL/min, before 100 or 500 µL sample were injected and separated. During the run, the UV-absorption at 280 nm was monitored and 500 µL fractions were collected. After the SEC runs, the column was washed with 30 mL degassed ddH<sub>2</sub>O and 30 mL degassed 20% (v/v) EtOH for storage.

The monomeric protein fractions were pooled for further use.

#### 3.2.8 Dialysis

1-5 L of the particular buffer were prepared and cooled for more than 3 h at 4°C before the start of the dialysis. Samples with volumes larger than 4 mL were filled into a Spectra/Por dialysis membrane with an appropriate molecular weight cut-off (MWCO), whereas for smaller samples volumes Slide-A-Lyzer dialysis cassettes were used. Dialysis of CoA was possible with Slide-A-Lyzer dialysis cassettes with a MWCO of 3500 Da without any loss of CoA.

All dialyses were carried out at 4°C ON with slow stirring on a magnetic stirrer.

#### 3.2.9 Proteolytic Cleavage with TEV Protease

The protein samples were subjected to dialysis against buffer for cleavage with TEV protease for 30 min, before TEV protease was added in a molar ratio of 1:50, and the dialysis was continued ON.

For unfolded proteins in buffer with urea, the dialysis was extended to 6 h, before the protease was added.



### 3.2.10 Concentration of Proteins

The type of concentrator was chosen by the sample volume. For sample volumes larger than 50 mL, Centriprep concentrators were used, whereas Amicon concentrators were chosen for smaller volumes. For proteins with molecular masses larger than 20 kDa, a MWCO of 10000 Da and for smaller proteins a MWCO of 3000 or 3500 Da was chosen.

#### 3.2.10.1 Concentration with Centriprep Concentrators

The concentrators were equilibrated with the appropriate buffer by centrifugation for 20 min at 3000\*g at 4°C. Afterwards, the concentrator was filled with 13 mL sample and centrifuged at 3000\*g at 4°C. Every 20 min the centrifuge was stopped to discard the flow through and, if indicated, to add another portion of the sample.

#### 3.2.10.2 Concentration with Amicon Concentrators

For equilibration, 4 mL buffer were filled into the concentrator, which was spun for 20 min at 3500\*g at 4°C. Then the buffer was discarded and the concentrator was filled with 4 mL of the sample. The sample was concentrated by centrifugation for 15 min at 3500\*g at 4°C, followed by adding new sample and mixing by pipetting up and down. The centrifugation/mixing cycles were repeated until the desired volume or concentration was reached.

### 3.2.11 Polyacrylamide Gel Electrophoresis

TRICINE/SDS gels<sup>192</sup> were cast and run, using a Mini-Protean Tetra Cell SDS-PAGE system. After the separation and stacking gels were cast, the system was assembled and filled with the cathode and anode buffer. Samples were prepared by mixing with the SDS-PAGE sample buffer and heating them to 95°C for 10 min. After the samples were loaded onto the gels, the gels were run for 15 min at a constant voltage of 70 V, before the voltage was increased to 150 V an for additional 50 min.

The gels were stained and fixed by mild shaking for 2 h at RT in ~100 mL staining/fixing solution and subsequently destained by shaking in ~100 mL destaining solution. If indicated, the destaining solution was exchanged every 60 min to reduce the background. For documentation, the destained gels were scanned.

#### **3.2.12 Determination of Protein Concentration**

The protein concentration was determined with a NanoDrop 1000 spectrophotometer by measurement of the absorption at 280 nm, with the buffer as blank reference. The theoretical molar absorption was calculated with the Protein Calculator web-tool. For phosphopantetheinylated proteins, the absorption of the cofactor was neglected.

#### **3.2.13 Protein Analysis by MALDI-TOF**

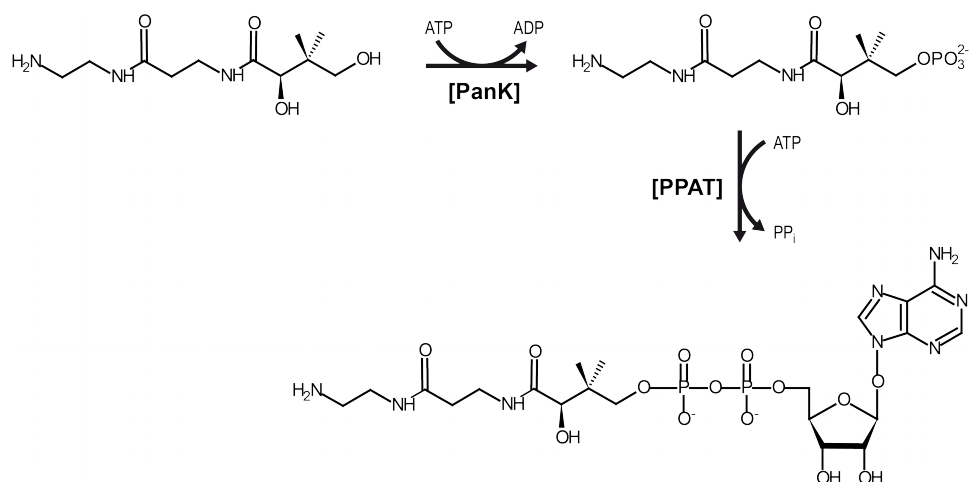
Protein solutions for analysis by matrix assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometry were given to Dr. Ute Bahr from the group of Prof. M. Karas, who performed sample preparation and measurement.

### **3.3 Synthesis of CoA Derivatives**

#### **3.3.1 Synthesis of Dephospho Amino-CoA**

Dephospho amino-CoA was synthesized in a chemo-enzymatic reaction, previously reported for the *in vitro* synthesis of CoA<sup>177</sup> and amino-CoA<sup>164, 193</sup>, omitting DPCK to obtain the 3'-dephosphorylated product (Figure 17).

Amino-pantetheine and the purified enzymes were prepared by members of the Marahiel group. The compounds listed in Table 1 were mixed in a 15 mL reaction vial, into which the enzymes, thawed on ice, were the last to be added.



**Figure 17:** Enzymatic synthesis of (3')-dephospho amino-CoA. Amino-pantetheine is phosphorylated by PanK. Subsequently, PPAT catalyzes the formation of dephospho amino-CoA.

The mixture was shaken at room temperature and 600 rpm. After 1.5 h it was applied onto a Sep-Pak C18 column, previously equilibrated with 10 mL MeCN and 10 mL ddH<sub>2</sub>O. The column was washed with 5 mL ddH<sub>2</sub>O and the product eluted with 6 mL 0.1% (v/v) TFA in MeCN/ddH<sub>2</sub>O (1:1), frozen in liquid nitrogen and lyophilized. The resulting white solid was dissolved in 2 mL 0.1% (v/v) TFA in ddH<sub>2</sub>O and further purified by HPLC.

**Table 1:** Reagents for synthesis of dephospho amino-CoA

Compound	c <sub>stock</sub> /mM	V/mL	eq
Amino-pantetheine (261.40 g/mol)	20	1.000	1.0
PanK (36360 g/mol)	0.297	0.168	0.0025
PPAT (17836 g/mol)	0.172	0.437	0.00375
ATP (551.14 g/mol)	100	2.000	10
MgCl <sub>2</sub> *6H <sub>2</sub> O (203.30 g/mol)	3000	0.133	20
KCl (74.56 g/mol)	3000	0.133	20
TRIS (pH 9.0 with hydrochloric acid)	1000	1.750	
ddH <sub>2</sub> O		to 6.500	

The purification was carried out with a CC 250/2 Nucleodur 100-3 C18ec column, connected to an Agilent 1100 Series HPLC Value system, equilibrated with 0.1% (v/v) TFA in ddH<sub>2</sub>O (buffer A; 0.1% (v/v) TFA in MeCN as buffer B) at RT. The product was eluted with a linear gradient to 30% buffer B over 30 min at a flow rate of 18 mL/min,

while fractions of 9 mL were collected, and the UV-absorption at 215 nm was monitored. The column was washed with 95% buffer B and buffer A, respectively, for storage.

Fractions showing UV-absorption were analyzed by ESI-MS for dephospho amino-CoA ( $m = 670.3$  Da). The fractions containing the product were pooled, shock frozen and lyophilized, resulting in a white powder.

### 3.3.2 Coupling of Peptides with Amino-CoA

The protected peptide, HBTU, and HOBt were dissolved in 300  $\mu$ L DMF. After adding of DIPEA, the mixture was heavily agitated and transferred to a vial with the solid dephospho amino-CoA (Table 2). The mixture was incubated for 1.5 h at 25°C at 700 rpm.

**Table 2:** Reagents for coupling of peptides with dephospho amino-CoA

Compound	n/ $\mu$ mol	amount	eq
Dephospho amino-CoA (670.30 g/mol)	10	6.70 mg	1.0
Tyc-Heptapeptide (protected; 1602.93 g/mol)	15	24.04 mg	1.5
HBTU (379.30 g/mol)	30	11.38 mg	3.0
HOBt (with 20% water; 135.13 g/mol)	30	5.07 mg	3.0
DIPEA (129.25 g/mol; $\rho = 0.76$ g/mL)	100	17.00 $\mu$ L	10

After completion of the reaction, peptidyl dephospho amino-CoA was deprotected immediately by adding 1.5 mL deprotection mixture (TFA/ddH<sub>2</sub>O/TIPS; 95:2.5:2.5). The deprotection reaction was carried out for 2 h at 25°C at 850 rpm. Afterwards, the reaction mixture was frozen in liquid nitrogen and dried by lyophilisation. The dry product was dissolved in 1 mL 15% (v/v) aqueous MeCN and applied to purification by preparative HPLC.

The purification was carried out with a CC 250/2 Nucleodur 100-3 C18ec column connected to an Agilent 1100 Series HPLC Value system with 0.1% (v/v) TFA in ddH<sub>2</sub>O as buffer A and 0.1% (v/v) TFA in MeCN as buffer B. The column was equilibrated with 10% buffer B at RT. After injection, a linear gradient to 60% buffer B over 30 min at a flow rate of 18 mL/min was used for purification, and 9 mL fractions

were collected. The column was washed with 95% buffer B and buffer A, respectively, for storage.

Fractions containing peptidyl dephospho amino-CoA ( $m = 1614.4$  Da) were identified by ESI-MS, pooled, and lyophilized for further use.

### 3.3.3 Coupling of Peptides and Amino Acids with CoA

The protected peptide, CoA and PyBOP were dissolved in 0.3 mL DMF/ddH<sub>2</sub>O (4:1), DIPEA was added (Table 3) and mixed for 2 h at 25°C and 850 rpm. Afterwards, the mixture was frozen in liquid nitrogen and lyophilized.

**Table 3:** Reagents for synthesis of peptidyl CoA

Compound	n/ $\mu$ mol	amount	eq
Coenzyme A*Li <sub>3</sub> (821.30 g/mol)	15	12.33 mg	1.5
Tyc-Heptapeptide (protected; 1602.93 g/mol)	10	16.03 mg	1.0
PyBOP (520.39 g/mol)	15	7.81 mg	1.5
DIPEA (129.25 g/mol; $\rho = 0.76$ g/mL)	100	17.00 $\mu$ L	10

The resulting solid was dissolved in 1.5 mL deprotection solution and agitated for 30 min at 25°C at 850 rpm, followed by shock-freezing and lyophilisation. The product was dissolved in 1 mL ddH<sub>2</sub>O and applied to HPLC purification like peptidyl dephospho amino-CoA.

Aminoacyl-CoA was synthesized and purified in the same way with protected amino acids, but for purification a linear gradient from 0-30% buffer B over 30 min was used.

### 3.3.4 Modification of PCP Domains with Ppan Derivatives

For the phosphopantetheinylation using CoA or non-hydrolyzable CoA-derivatives, a solution of 100  $\mu$ M apo-protein in the buffer for phosphopantetheinylation was prepared. The CoA-derivatives were dissolved in an appropriate volume of the buffer to get a theoretical concentration of 100 mM. Then they were added to the protein solution to a final concentration of 1.0 mM. Afterwards, a 1.3 mM Sfp solution was

added to a final concentration of 0.5  $\mu\text{M}$ . The mixture was incubated ON at 4°C under mild shaking and purified by reversed nickel affinity purification and subsequent SEC, or by salt exchange on a PD10 column.

For hydrolyzable derivatives, the reaction was carried out similarly with 1.0  $\mu\text{M}$  Sfp and incubation for 30 min at 37°C in 50 mM  $\text{NaP}_i$ /5 mM  $\text{MgCl}_2$  (pH 6.8). In these cases, only salt exchange was used for purification.

## 3.4 Biochemical and Biophysical Methods

### 3.4.1 NMR Titration experiments

The last purification step for all proteins used in a titration experiment was a SEC run with the same aliquot of NMR buffer to minimize artifacts due to different buffer compositions.

The first sample was prepared by mixing the labeled protein with 3 mM DSS in  $\text{D}_2\text{O}$  and NMR buffer to get 550  $\mu\text{L}$  sample with 100  $\mu\text{M}$  protein, 150  $\mu\text{M}$  DSS, and 5% (v/v)  $\text{D}_2\text{O}$ . The sample was centrifuged for 5 min at 4°C at 13000 rpm, filled into a NMR tube and measured on a Bruker Avance spectrometer. For signal referencing, an one-dimensional (1D)  $^1\text{H}$ -spectrum at the desired temperature was recorded, and the center of the signal from the methyl-groups of DSS was set to 0.000 ppm. Afterwards,  $^{15}\text{N}$ - and, if indicated,  $^{13}\text{C}_{\text{aromatic}}$ - as well as  $^{13}\text{C}_{\text{aliphatic}}$ -HSQC spectra with adequate number of scans (NS), increments in the indirect dimension, and spectral width were recorded and calibrated to the 1D-spectrum.

Samples for the further titration steps were prepared by mixing a portion of the sample from the previous step with the stock solutions of the labeled and unlabeled proteins and the DSS/ $\text{D}_2\text{O}$  solution. All samples contained 100  $\mu\text{M}$  labeled protein, 150  $\mu\text{M}$  DSS, and 5% (v/v)  $\text{D}_2\text{O}$  and the concentration of the unlabeled interaction partner was varied from 25 to 400  $\mu\text{M}$  (0.25-4.0 eq with respect to the labeled protein). The same types of spectra as before were recorded and only the NS was changed, if the signal intensity decreased over the titration steps.

### 3.4.2 NMR Measurements for Structure Determination

Uniformly [ $^{13}\text{C};^{15}\text{N}$ ]-labeled samples were used for the recording of NMR spectra for resonance assignment and structure determination. The purified protein in NMR buffer was concentrated to a volume of  $\sim 250\ \mu\text{L}$ , of which  $238\ \mu\text{L}$  were mixed with  $12\ \mu\text{L}$  3 mM DSS in  $\text{D}_2\text{O}$ , centrifuged for 5 min at  $4^\circ\text{C}$  at 13000 rpm and filled into a Shigemi NMR tube.

A set of three-dimensional spectra was recorded by Dr. Frank Löhr for assignment and structure calculation at the desired temperature at different Bruker Avance spectrometers with field strengths of 500-950 MHz. As not all spectra were consecutively recorded at the same spectrometer, for each session a 1D  $^1\text{H}$ -spectrum was recorded for referencing to DSS, as described for titration experiments, and a [ $^1\text{H};^{15}\text{N}$ ]-HSQC was recorded to confirm the integrity of the sample.

For assignment and determination of structural restraints of the unlabeled peptidyl ppan cofactor [ $^{13}\text{C};^{15}\text{N}$ ] double-filtered two- and three-dimensional NOESY<sup>194</sup> spectra were recorded.

### 3.4.3 NMR Assignment and Structure Calculation

With the Sparky NMR analysis software resonances from backbone amide protons and nitrogen atoms as well as CAs and CBs were manually assigned using a [ $^1\text{H};^{15}\text{N}$ ]-HSQC and a HNCACB<sup>195</sup> spectrum. Aliphatic side chain hydrogen and carbon resonances were assigned in a H(CCCO)NH- and (H)C(CCO)NH-TOCSY<sup>196</sup> in combination with a [ $^1\text{H};^{13}\text{C}_{\text{aliphatic}}$ ]-HSQC. For assignment of the aromatic resonances spectra of (HB)CB(CGCD)HD<sup>197</sup>, Phe- or Tyr-optimized (HB)CB(CGCC-TOCSY)H<sup>198</sup> and a [ $^1\text{H};^{13}\text{C}_{\text{aromatic}}$ ]-HSQC were used. All resonances that were not assigned at this point were assigned using the NOESY spectra in combination with the HSQCs.

NOE restraints for structure calculation were derived from the NOESY spectra with the automated, restricted peak picking function of Sparky, followed by manual revision of the picked peaks. The peak lists and a list of the assigned resonances were given to Donata Kirchner from the Güntert group, who performed the structure

calculation. The NOE restraints were assigned and calibrated with CYANA, and torsion angle restraints, deduced from the resonances, were calculated with TALOS+. The final structures were calculated with CYANA and refined using OPALp.

#### **3.4.4 Crystallization Trials**

The last purification step for all protein samples used for crystallization was a SEC run in the crystallization buffer. Afterwards, the proteins were concentrated and samples with different concentrations were prepared.

##### **3.4.4.1 Sitting Drop Method**

Samples were prepared in a 96-well format together with Barbara Rathmann, using the CrystalMation robot at the Max-Planck-Institute of Biophysics, or with Dr. Simin Rahighi, using a Mosquito robot at the Buchmann Institute for Molecular Life Science.

The protein solutions were mixed in a 1:1 ratio with the condition to drops of total volumes between 0.2 and 1.0  $\mu\text{L}$ . Plates with 100-200  $\mu\text{L}$  of the conditions in the reservoirs were incubated at different temperatures and checked frequently for crystals.

##### **3.4.4.2 Hanging Drop Method**

Trials for hanging drop were prepared manually for conditions showing crystal growth with the sitting drop method. For each drop, 1  $\mu\text{L}$  protein solution and 1  $\mu\text{L}$  of the condition were mixed, and 250  $\mu\text{L}$  of the condition were filled into the reservoir. Incubation at different temperatures was tested for optimal crystal growth.



### 3.4.5 X-ray Diffraction and Crystal Structure Determination

Crystals were tested for diffraction together with Dr. Simin Rahighi at the in-house beamline at the Max-Planck-Institute of Biophysics. Additionally, in case of the PCP/Sfp complex, a whole dataset was recorded.

Other datasets of crystals from promising conditions were recorded at the PX3 beamline at the Swiss Light Source. Datasets were processed with MOSFLM, the structures were solved by molecular replacement with MOLREP and adequate search models, and then they were refined, using COOT and REFMAC5. Recording and processing of the data was performed by Dr. Simin Rahighi.

### 3.4.6 Phosphopantetheinylation Assay

The final purification step for all proteins used in the phosphopantetheinylation assay was a SEC run with NMR buffer supplemented with DTT and  $MgCl_2$ , each to a final concentration of 5 mM.

PCPs were adjusted to a concentration of 105  $\mu M$ , and 95  $\mu L$  aliquots were made. The reaction was started by adding 5  $\mu L$  of a solution containing 10  $\mu M$  Sfp and 20 mM CoA. Then the reaction was carried out for several periods of time at 25°C with mild shaking. The reactions were quenched by adding 100  $\mu L$  of a 20% (w/v) TCA solution. The precipitates were pelleted by centrifugation for 30 min at 20000\*g at 4°C, washed with 100  $\mu L$  20% (w/v) TCA solution, centrifuged as before and washed with 100  $\mu L$  cold diethyl ether. After a final centrifugation step. the pellets were dried in a vacuum concentrator and stored at -20°C, before each pellet was dissolved in 50  $\mu L$  0.1% (v/v) aqueous TFA and applied to analysis by LC-MS.

### 3.4.7 Circular Dichroism Spectroscopy

Protein samples for circular dichroism (CD) spectroscopy were dialyzed against 10 mM  $NaP_i$  (pH 6.8; a 1:5 dilution of the NMR buffer). The concentration was adjusted to 25  $\mu M$ , and 300  $\mu L$  sample were filled into a clean cuvette with a path length of 0.1 cm.

On a JASCO J-810 spectrometer, CD-spectra were recorded at 5°C as the averages of three scans. Each scan was measured from 350-190 nm with a bandwidth of 1.0 nm and a speed of 100 nm/min. For melting curves, the temperature dependency of the ellipticity at 222 nm was measured. Starting at 4°C, the temperature was linearly increased to 95°C by 1°C/min. After each measurement, the cuvette was filled with a 2% (v/v) hellmanex solution and warmed to 50°C for 10 min. After repeated rinsing with ddH<sub>2</sub>O, the cuvette was dried for further use.

#### **3.4.8 Isothermal Titration Calorimetry**

After purification by SEC, using NMR buffer with an additional 5 mM MgCl<sub>2</sub>, all proteins and CoA used in one series of experiments were dialyzed against the same aliquot of buffer with the same composition used for the previous SEC. For measurement of protein/protein interactions, solutions of ~50 μM (titrand) or ~800 μM (titrant) protein with 3 mM CoA and 2 mM TCEP were prepared and degassed. For Protein/ligand interactions, solutions of ~50 μM protein and ~800 μM CoA, both with 2 mM TCEP, were used.

The titrand was filled into the cell, and the titrant into the syringe of a VP-ITC calorimeter (1.453 mL cell volume) tempered to 5°C. After the system's temperature was equilibrated, the titrant was injected in 27 steps of 10 μL each lasting 10 s with 5 min pauses for temperature equilibration in between. If the proteins were stable enough, the measurements were repeated at 25°C in the same manner. Data analysis was performed with Origin 6.1 using a titration of titrant to buffer as blank.

#### **3.4.9 Analysis of PCPs by Liquid Chromatography-Mass Spectrometry**

PCP domains were analyzed, using a CC250/3 Nucleosil 120-3 C18 column connected to an Agilent 1100 Series HPLC Value system with 0.1% (v/v) TFA in ddH<sub>2</sub>O as buffer A and 0.1% (v/v) TFA in MeCN as buffer B. The protein solutions were centrifuged for 5 min at 13000 rpm prior to injection. The operation temperature and gradient were optimized for each PCP individually. Generally, the temperature of the column was between 40 and 50°C and the initial concentrations were 35-45%

buffer B. For separation of the PCPs, linear gradients to concentrations of 50-60% buffer B over 27 min at a flow rate of 0.8 mL/min were used. In all cases the column was washed with 95% buffer B and equilibrated with the initial buffer concentration of the next run afterwards.

During separation the UV-absorption at 214 nm and the MS-signal in the m/z range of 400-1400 were monitored. The different forms of PCP were identified by deconvolution of the m/z signals and quantified by the peak areas of the UV-signals, analyzed with ChemStation.

## 4. Results

### 4.1 Recognition of PCP-bound Peptides

#### 4.1.1 Protein Expression and Purification

Expression of NRPS proteins from *Bacillus* and other bacteria in *E. coli* often yields high amounts of soluble protein. This might be related to the fact that *E. coli* itself is a natural producer of the NRP enterobactin<sup>199</sup> and therefore a very suitable expression system for NRPS proteins. Then again, problems with the expression of constructs containing PCP domains in *E. coli* have been reported<sup>71, 200 69</sup>. In these studies, the overexpressed proteins were partly phosphopantetheinylated *in vivo*, which was associated with the activity of EntD, the PPT from the enterobactin NRPS gene cluster<sup>201</sup>. Thus, it is not only necessary to find conditions for expression and purification that result in stable monomeric proteins, but to also ensure a homogeneous form of the PCP domain.

##### 4.1.1.1 Expression and Purification of TycC3\_PCP

In the presented work, the construct of *tycC3\_PCP* in pQE70, already used for several structural and functional studies<sup>37, 101 38, 39</sup>, was used. In contrast to these studies, the protein was expressed soluble and not purified under denaturing conditions. This was accomplished by expression in *E. coli* M15(pREP4) cells, induced by adding 0.25 mM IPTG and carried out for 6 h at 28°C. Monomeric protein with yields between 4 and 12 mg per liter expression (depending on the medium) could be gained by nickel affinity purification, followed by SEC on a Superdex 75 with the buffers for NMR or phosphopantetheinylation.

As analyzed by MALDI-TOF, protein from expression in LB, 2xYT, and unlabeled as well as isotopic labeled M9 medium was exclusively in the *apo*-form. Furthermore, in all cases the observed mass indicated a quantitative *in vivo* demethioninylation of the protein.

#### 4.1.1.2 Expression and Purification of TycC4\_C

For the expression of TycC4\_C, a construct from the Marahiel lab in pET28a with a N-terminal his-tag was used. The protein was expressed in *E. coli* BL21(DE3) cells, grown in 2xYT. The expression was started by adding IPTG to a final concentration of 0.5 mM, and the cells were further cultivated at 16°C ON.

The protein was purified by nickel affinity purification, and subsequent SEC. Monomeric protein (~10 mg/L) was obtained when the NMR buffer was used, whereas the protein formed soluble aggregates in the buffer for phosphopantetheinylation.

#### 4.1.1.3 Expression and Purification of TycC4\_PCP

*tycC4\_PCP* in pBH4 was cloned from genomic DNA of *B. brevis* strain ATCC 8185<sup>202</sup> with the primers *tycC4\_PCP*-fw and -rv.

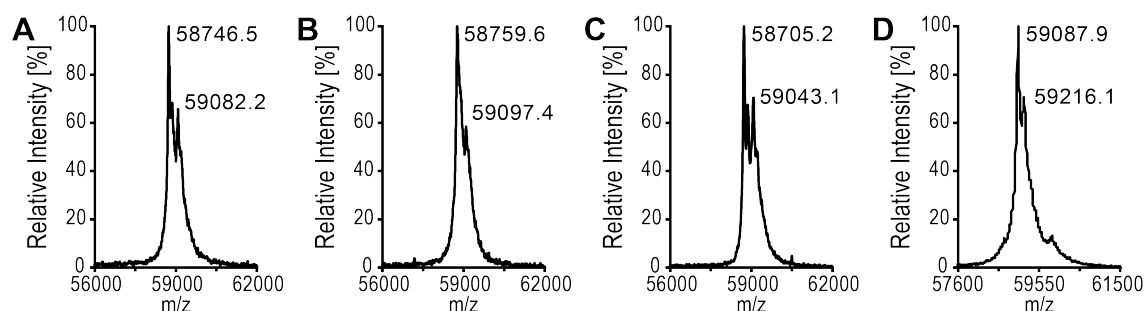
Expression and purification of TycC4\_PCP was performed according to the procedure for TycC4\_C. In contrast to the C domain, the obtained apo-PCP was monomeric in the NMR buffer and in the buffer for phosphopantetheinylation with yields of ~12 mg/L.

#### 4.1.1.4 Expression and Purification of TycC5-6\_PCP-C

The gene coding for TycC5-6\_PCP-C could be subcloned into the pBH4(His10) vector, using *tycC5-6\_PCP-C* in pQE60<sup>69</sup> as template and *tycC5-6\_PCP-C*-fw and -rv as primers for PCR.

The bidomain was expressed in *E. coli* BL21(DE3) cells, cultivated in 2xYT. After IPTG was added to a final concentration of 0.25 mM, the cells were grown ON at 16°C. The protein was purified by nickel affinity purification, followed by a cleavage with TEV protease, reverse nickel affinity purification and SEC with the phosphopantetheinylation or the crystallization buffer. In both buffers, the protein was a stable monomer with a yield of ~50 mg/L. Analysis of the purified protein revealed that a significant portion was already phosphopantetheinylated *in vivo*, as strong

signals of the *apo*- (58749 Da) as well as of the *holo*-bidomain (59090 Da) were observed. Neither variation of the duration of the expression (2-16 h), of the temperature (10-30°C), and of the cultivation medium (2xYT, LB, M9), nor supplementation of the media with FeCl<sub>3</sub> (0-50 µM) resulted in pure *apo*-protein (Figure 18A-C), though the yields differed under some conditions tremendously (for example ~3 mg/L from expression in 2xYT at 10°C for 2 h).



**Figure 18:** MALDI-TOF spectra of Tyc5-6\_PCP-C expressed in 2xYT, supplemented with 20 µM FeCl<sub>3</sub> for 4 h at 10°C (A), 16°C (B), and 28°C (C), respectively. In all cases a mixture of the *apo*- and *holo*-form was detected. Similar results were obtained, when other expression parameters were varied. *In vitro* modification with Sfp resulted in homogeneous *holo*-bidomain (D). In some cases the matrix-adduct (theoretical mass +136 Da) was detected, too.

Though no homogeneous *apo*-bidomain was obtained, the mixture could be converted to pure *holo*-Tyc5-6\_PCP-C by phosphopantetheinylation with Sfp *in vitro*, followed by reverse nickel affinity purification and SEC with the crystallization buffer (Figure 18D).

#### 4.1.1.5 Expression of Sfp

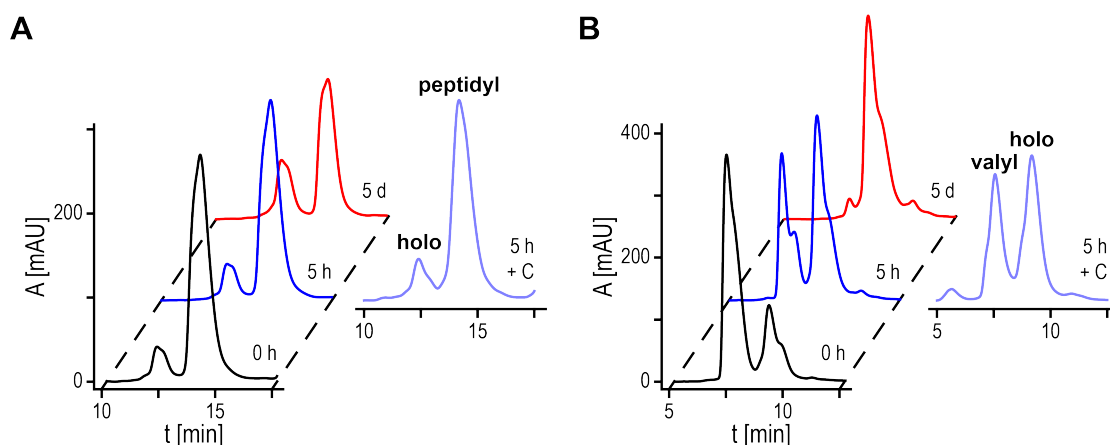
*E. coli* M15(pREP4) cells harboring *sfp* in pQE60<sup>203</sup> were used for expression in LB medium. The expression was started by adding 0.4 mM IPTG and carried out for 4 h at 30°C. After nickel affinity purification the eluted protein was dialyzed against 50 mM HEPES/250 mM NaCl/2 mM MgCl<sub>2</sub> (pH 8.0), concentrated to ~1.3 mM, and stored in aliquots at -80°C. The enzyme was still active after storage periods exceeding one year.

#### 4.1.2 Stability of Acylated PCPs and Activity of TycC4\_C

While a non-active fraction of interaction partners can often be deduced in a biochemical assay, it is disastrous for most structural investigations. In the context of NRPS proteins, 'non-active' is not limited to misfolded, aggregated, or not posttranslationally modified proteins, but includes proteins with a wrong or no attachment to the ppan's thiol group. The vulnerability to hydrolysis of the thioester linking the attachment to the ppan-arm of the PCP raises the question, if the PCP's form of interest, in case of investigation of its interaction with the C domain the peptidyl- and aminoacyl-form, is stable enough to endure the experiment.

To investigate the hydrolysis of the thioester bond in aqueous solution, L-valine and the native tyrocidine heptapeptide (DPhe-LPro-LPhe-DPhe-LAsn-LGln-LTyr) were coupled to CoA and loaded onto TycC4\_PCP and TycC3\_PCP, respectively. Synthesis and purification of the CoA derivatives were successful and showed no impurities as analyzed by ESI-MS. Valyl- as well as peptidyl-CoA could be used for quantitative loading to the PCPs with Sfp *in vitro*. Incubation of 50  $\mu$ M PCP solutions in the NMR buffer at 25°C for different time periods followed by quick-freezing. Following analysis by LC-MS with the column heated to 45°C, using a gradient from 45 to 54.5% MeCN in 0.1% aqueous TFA over 27 min, revealed the time dependency of the hydrolysis. Even before the incubation at 25°C, peptidyl-TycC3\_PCP contained 11.2% *holo*-PCP, whose content increased to 11.8% after 5 h and 29.7% after 5 days. Similar incubation in presence of 2 eq TycC4\_C showed approximately the same hydrolysis rate, as 11.9% *holo*-PCP were detected after 5 h (Figure 19A).

The thioester bond between valine and TycC4\_PCP was even more labile. Prior to incubation the samples contained already 27.4% *holo*-PCP. After 5 h 58.3% of the PCP was hydrolyzed, and after 5 days only *holo*-PCP was detectable (Figure 19B). Again, TycC4\_C had no significant influence on the hydrolysis (56.8% *holo*-TycC4\_PCP after 5 h).



**Figure 19:** LC-MS analysis of the non-catalyzed hydrolysis of the thioester bond in peptidyl-TycC3\_PCP (A) and valyl-TycC4\_PCP (B). In both cases the content of *holo*-PCP increases over time (black 0 h, blue 5 h and red 5 days at 25°C), though the hydrolysis is more rapid for the valyl-PCP. The presence of TycC4\_C does not influence the velocity of the hydrolysis (light blue).

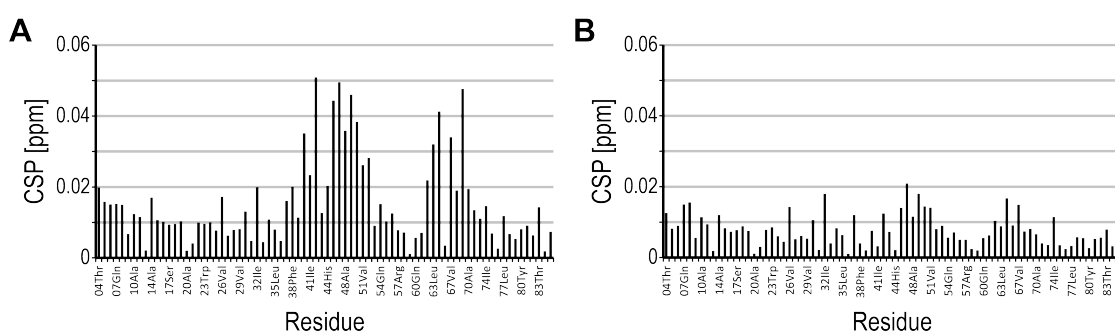
Since both proteins used for loading contained only *apo*-protein, as analyzed by the same LC-MS assay, the *holo*-PCP, detected prior to incubation at 25°C, must have been formed by hydrolysis during the *in vitro* phosphopantetheinylation, the subsequent purification, and the LC-MS run. Due to the fast hydrolysis of valyl-TycC4\_PCP, further structural studies on the aminoacyl-PCP were abandoned. Furthermore, TycC4\_C was assumed to not interact with peptidyl-TycC3\_PCP and valyl-TycC4\_PCP, as it had no influence on their hydrolysis. To further probe the activity of the C domain, a mixture of 25  $\mu$ M peptidyl-TycC3\_PCP, 25  $\mu$ M valyl-TycC4\_PCP, and 50  $\mu$ M TycC4\_C was incubated and analyzed as before. After 5 h incubation at 25°C, no octapeptidyl-TycC4\_PCP was detected, suggesting that no productive interaction between the three proteins occurred.

#### 4.1.3 Comparison of Amide- and Thioester-bound Peptidyl-PCP

Analysis of the stability of peptidyl-TycC3\_PCP against hydrolysis suggests that it is not stable enough for time-consuming multidimensional NMR experiments, as required for structure determination. Still, the stability was sufficient to record a [ $^1\text{H};^{15}\text{N}$ ]-HSQC of the uniformly  $^{15}\text{N}$ -labeled protein loaded with unlabeled peptidyl-ppan in NMR buffer at 295 K. Comparison with a HSQC of *holo*-TycC3\_PCP, recorded under the same conditions, showed significant chemical shift perturbations



(CSPs) clustering between Q40-A52 and E62-F69, whereas the HSQC of TycC3\_PCP loaded with the non-hydrolyzable peptidyl(NH)-ppan showed only minor differences (Figure 20). The differences in the spectra of *holo*- and peptidyl-TycC3\_PCP indicate a structural change of these two forms of the PCP. Although the structure determination of the peptidyl-PCP is not possible due to its hydrolytic instability, the PCP loaded with the non-hydrolyzable peptide seems to have the same structure, as its HSQC shows only small differences.



**Figure 20:** Analysis of the CSPs between the  $[^1\text{H};^{15}\text{N}]$ -HSQCs of TycC3\_PCP loaded with the hydrolyzable peptide and *holo*-TycC3\_PCP (A), or loaded with the non-hydrolyzable peptide, respectively (B). While the attachment of the peptide has an effect on the PCP, it seems to make no difference whether it is linked via an amide or a thioester bond.

#### 4.1.4 Structure of Peptidyl(NH)-TycC3\_PCP in Solution

The structure of peptidyl(NH)-TycC3\_PCP was solved to address the questions, which conformation it adopts, and if there is an interaction between the protein and the peptide, indicating a possible recognition of the peptide by the PCP.

##### 4.1.4.1 Sample Preparation and NMR Measurements

Dephospho amino-CoA was successfully synthesized and coupled with the tyrocidine heptapeptide. The product of this reaction could be used to load peptidyl(NH)-ppan onto  $[^{13}\text{C};^{15}\text{N}]$ -labeled *apo*-TycC3\_PCP *in vitro*, and a sample, containing 450  $\mu\text{M}$  of the modified protein, was used for structure determination. Recording of the spectra for resonance assignment and structure calculation was performed at 295 K.

#### 4.1.4.2 Resonance Assignment and Structure Calculation

For the core protein without the his-tag (P2-S85) the resonances of almost all groups with non-labile protons could be assigned, except for the backbone amide proton and nitrogen of V3 and S45 as well as the methanediyl group of I34. From the his-tag (H86-91), only the resonances of the aliphatic atoms of H86 and H87 as well as the amide signals of H87 could be assigned.

For structure calculation, torsion restraints were calculated with TALOS+, based on the resonance assignment. NOE restraints were derived from [ $^1\text{H};^{13}\text{C}_{\text{aliphatic}}$ ]-, [ $^1\text{H};^{13}\text{C}_{\text{aromatic}}$ ]-, and [ $^1\text{H};^{15}\text{N}$ ]-edited NOESY spectra by automated restricted peak picking, followed by manual revision in Sparky. The combined information was used for calculation of a bundle of 20 structures, not considering the cofactor (Table 4).

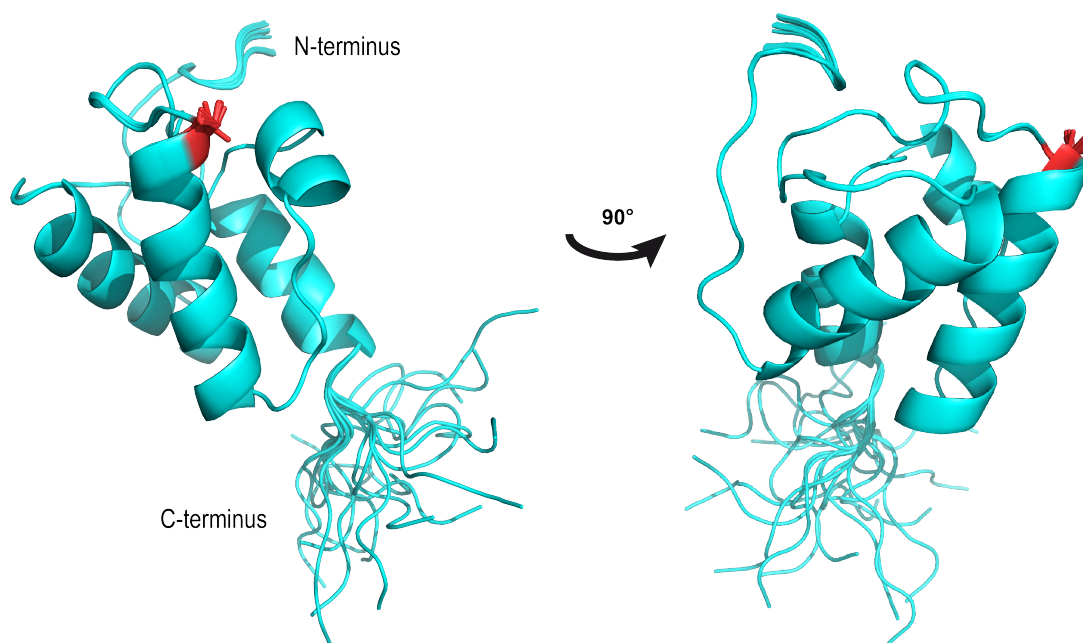
Afterwards a 2D filtered NOESY and comparison with a previous assignment of the ppan-arm<sup>204</sup> were used to assign the cofactor. Resonances of peptidyl(NH)-ppan were used for manual peak picking and assignment in a filtered [ $^1\text{H};^{13}\text{C}_{\text{aliphatic}}$ ]-edited NOESY spectrum to gain distance information about the position of the cofactor relative to the protein. During this process, it was difficult to distinguish NOEs between the cofactor and the PCP from those within the protein, as the latter could not be suppressed completely. The reason for this is the need of identical  $^1\text{J}(^1\text{H}-^{13}\text{C}_{\text{aliphatic}})$ -coupling constants throughout the whole protein, if all signals should be suppressed completely<sup>194</sup>. As these coupling constants vary from ~80-125 Hz, this cannot be accomplished in reality. This problem was solved by searching NOEs, which cluster in a region close to S45 in the preliminary structure. However, by doing so only a few NOEs from the pantoyl portion of the cofactor to the side chains of L46, M49, and L65 were identified. Lacking further distance information, another structure calculation considering peptidyl(NH)-ppan was not performed.

**Table 4:** NMR structure determination statistics for peptidyl(NH)-TycC3\_PCP

<b>NOE assignment<sup>a</sup></b>	
Total number of NOESY cross peaks	6107
Assigned cross peaks	5401 (=88% of total)
in <sup>13</sup> C <sub>aliphatic</sub> resolved NOESY	3812 (89%)
in <sup>13</sup> C <sub>aromatic</sub> resolved NOESY	287 (94%)
in <sup>15</sup> N resolved NOESY	1302 (85%)
<b>Conformational restraints</b>	
Total NOE distance restraints	3301
Short range $ i - j  \leq 1$	1294 (39%)
Medium range $1 <  i - j  < 5$	867 (26%)
Long range $ i - j  > 5$	1140 (35%)
Dihedral angle restraints ( $\psi/\phi$ )	140
<b>Structure statistics<sup>b</sup></b>	
Average CYANA target function ( $\text{\AA}^2$ )	2.78±0.24
AMBER energies (kcal/mol)	-3116.89±38.36
<b>Restraint violations<sup>c</sup></b>	
Max. distance restraint violation ( $\text{\AA}$ )	0.13
Violated distance restraints $> 0.2 \text{\AA}$	0
Max. dihedral angle restraint violation ( $^\circ$ )	3.42
Violated dihedral angles $> 5^\circ$	0
<b>Ramachandran plot</b>	
Residues in most favored regions	86.4%
Residues in additionally allowed regions	13.5%
Residues in generously allowed regions	0.1%
Residues in disallowed regions	0.0%
<b>RMSD (residues 2-91)</b>	
Average backbone RMSD ( $\text{\AA}$ )	1.52±0.38
Average heavy atom RMSD ( $\text{\AA}$ )	2.03±0.35
<sup>a</sup> using the automated NOE assignment and structure calculation functionalities of CYANA	
<sup>b</sup> after restrained energy minimization with OPALp	
<sup>c</sup> after energy minimization, calculated with CYANA	

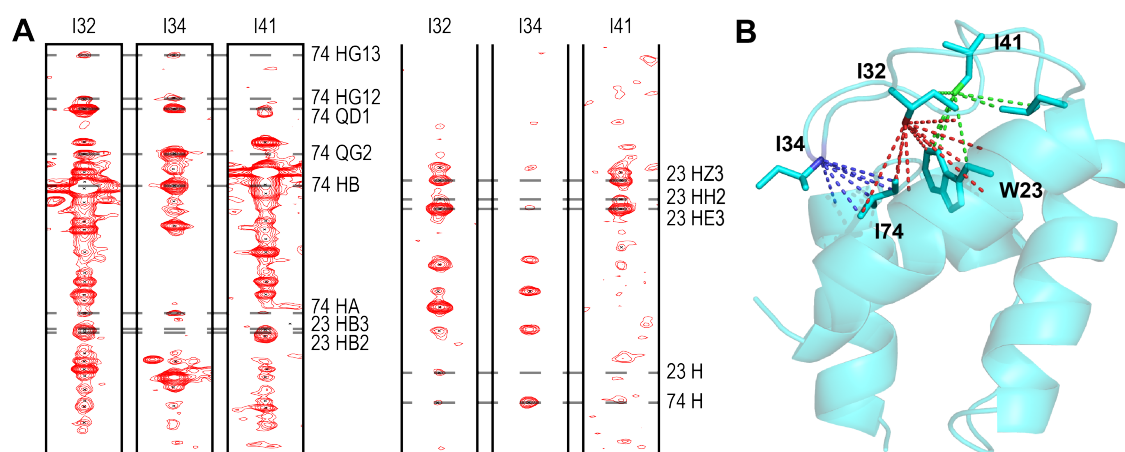
#### 4.1.4.3 The Preliminary Structure of Peptidyl(NH)-TycC3\_PCP

The structure solved (Figure 21) shows that the peptide-loaded PCP is in the normal carrier protein conformation, referred to as A/H state. Beside a small part of the N-terminus and the his-tag, the 20 calculated structures align very well.



**Figure 21:** Presentation of the 20 lowest energy structures of peptidyl(NH)-TycC3\_PCP. The cofactor attached to S45 (red) was not considered for the structure calculation.

Analysis of the NOESY spectra suggests that the N-terminus is less ordered than it appears in the structure bundle: P2-E5 exhibit no long-range NOEs, though from the structure one would expect to observe NOEs from T4 and E5 to the side chain of N37. Furthermore, the absence of the signal from V3 in the  $[^1\text{H};^{15}\text{N}]$ -HSQC is likely to be caused by line broadening due to orientational flexibility. However, as the rest of the loop, starting with A6, shows NOE contacts to the core domain, it seems as if this part of the loop preceding the first helix is presented correctly in the structure. The same is valid for the loop connecting helix 2 and 3. Being only  $\sim 11$  Å apart, the C-terminus of helix 2 and the N-terminus of helix 3 are bridged by 17 residues. However, the whole loop is ordered and tightly associated with the hydrophobic core of the PCP which can be retraced by a widespread network of long-range NOE signals of residues from the loop mainly to W23 as well as I74 located in helix 1 and 4, respectively (Figure 22).



**Figure 22:** Stripes from a  $^{13}\text{C}_{\text{aliphatic}}$ -NOESY of the peptide-loaded PCP, which correspond to the  $\gamma$ -methyl group of I32,  $\text{C}^{\alpha}/\text{H}^{\alpha}$  of I34, and the  $\delta$ -methyl group of I41 (A. The left part shows the aliphatic region and the right part the aromatic/amide region, with the resonances of W23 and I74 indicated as dashed lines.) Presentation of the long-range NOE contacts (dashed lines), derived from the three shown stripes, illustrates the close contact between the loop comprising I32 (red), I34 (blue), and I41 (green) to the core of the PCP.

Structural alignment of the helical core of the PCP (P11-T83) with known structures of PCP domains from *Bacillus* in the A/H state reveals high similarity. Compared to the first NMR structure of TycC3\_PCP solved by Weber *et al.*<sup>37</sup> (PDB 1DNY), it has an RMSD of 1.26 Å, whereas the structure later solved by Koglin *et al.*<sup>38</sup> (PDB 2GDW) with an RMSD of 2.45 Å is less similar, yet still very alike. The similarity to crystal structures of PCP domains that are part of larger constructs is remarkable: The RMSD is 1.27 Å to the structure of *apo*-TycC5-C\_PCP-C (PDB 2JGP) and 1.40 Å to the PCP from the termination module of the surfactin synthetase.

#### 4.1.5 Crystal Structure of *holo*-TycC5-6\_PCP-C

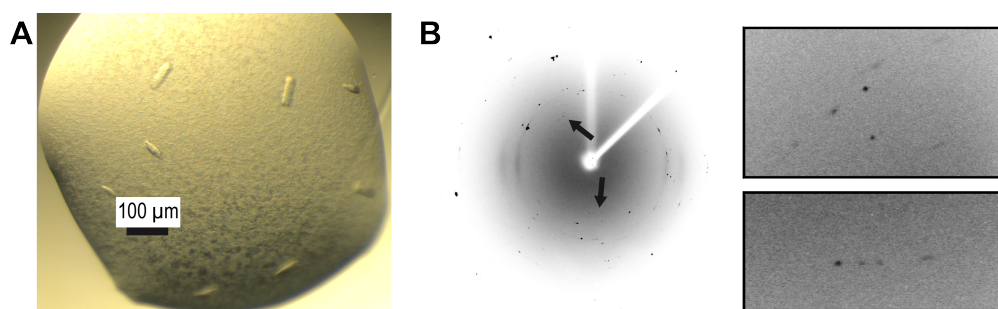
In the beginning of this project, it was intended to crystallize TycC5-6\_PCP-C loaded with the non-hydrolyzable native tyrocidine nonapeptide (dPhe-LPro-LPhe-dPhe-LAsn-LGln-LTyr-LVal-LOrn), using the amino-CoA approach. In the known structure of the *apo*-bidomain<sup>69</sup> (PDB 2JGP), the orientation of the PCP to the C domain cannot resemble the donor site conformation, as the distance between the active site serine of the PCP and the catalytic histidine of the C domain is far too large to be bridged by the ppan-arm. It was expected that the C domain has a higher

affinity for the peptide-loaded PCP, which causes a conformational change, so the PCP can adopt a position, in which the ppan-arm with the attached peptide can extend into the catalytic groove of the C domain. Yet, this hypothesis could not be tested, as the bidomain could not be expressed in a homogeneous *apo*-form, which is required for loading with peptidyl-ppan. Instead, it was tested if the C domain's affinity for the ppan-arm alone was high enough to induce the desired conformational change of the PCP.

#### 4.1.5.1 Crystallization Trials and Optimization

After successful quantitative loading of TycC5-6\_PCP-C with ppan, using Sfp *in vitro* (Figure 18D), the bidomain was again purified by reverse nickel affinity chromatography, followed by SEC in the crystallization buffer. After a sequence of five purification steps and two enzymatic reactions, 30 mg homogeneous, monomeric *holo*-bidomain per liter expression culture were finally yielded.

Solutions of the protein with concentrations of 16.4 and 8.0 mg/mL, respectively, were used for high-throughput crystallization trials with a CrystalMation robot at the Max Planck Institute of Biophysics, Frankfurt. In total, more than 1000 conditions were tested with the sitting drop method at a drop size of 200 nL and 100  $\mu$ L reservoir volume, each at 4 and 16°C. Crystal growth was observed only in three of the tested conditions. As two of the conditions were not reproducible, only one condition was further optimized. After 10 days at 16°C, crystals had started growing in a 1:1 mixture of the 16.4 mg/mL protein solution and condition 51 of the JCSG core III crystallization kit [1.6 M  $(\text{NH}_4)_2\text{SO}_4$ , 0.1 M MES, 10% (v/v) dioxane at pH 6.5].



**Figure 23:** Image of the crystals obtained from the initial condition screening (A). Their diffraction pattern (B) indicates a low quality of the crystals, though some reflections (arrows; magnification) hint at the presence of protein in the crystal.

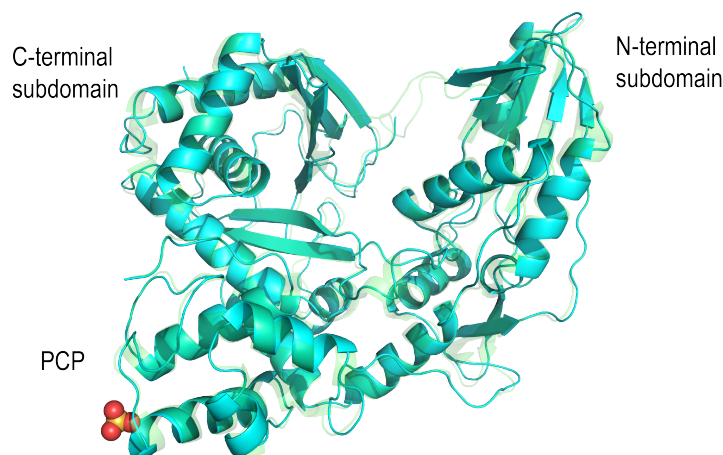
Remarkably, this condition was the same as reported for the initial crystallization of the *apo*-form of the bidomain<sup>69</sup>. Quality control of the crystals at the in-house beamline resulted in a diffraction pattern, indicating a protein crystal, but with a quality too low for recording of a whole dataset, as required for resolving the structure of the protein (Figure 23).

During the optimization of the initial condition, crystals were obtained from multiple combinations of MES, varying from 0.01-0.15 M, and dioxane, varying from 1.0-5.7% (v/v), in a pH range from 5.5 to 6.7. In all cases, the crystal growth was dependent on high concentrations of  $(\text{NH}_4)_2\text{SO}_4$ , as crystals only formed in presence of 1.6 and 1.8 M of this agent, whereas a reduction to 1.3 M already suspended their growth. On the other hand, the protein concentration had only a little influence on the crystal growth, as all tested concentrations (4.0-37.0 mg/mL) resulted in crystals. However, a slight correlation between a higher protein concentration and an increasing number of conditions, in which crystal growth was observed, was apparent.

Surprisingly, reproduction of any condition with the hanging drop method yielded in no crystals at all.

#### 4.1.5.2 Structure Determination of *holo*-TycC5-6\_PCP-C

Datasets for several crystals from different conditions obtained in the optimization process were collected at the Swiss Light Source, Villingen. Most crystals had a similar resolution to ~3.5 Å and showed the same diffraction pattern. In all of these cases, the structure (Figure 24) could be solved by molecular replacement, using the structure of the *apo*-bidomain as a search model.



**Figure 24:** Overlay of the structures of *holo*- (cyan) and *apo*-TycC5-6\_PCP-C (semi-transparent green). The two forms have the same overall structure, but the sulfate ion (spheres) close to the conserved serine of the PCP is not present in the *holo*-structure.

The overall arrangement and the crystal unit cell of the derived structure were virtually identical with the structure of the *apo*-form. The only difference was the lack of electron density, on the one hand for the two loops bridging the active site canyon between the two subdomains of the C domain, and on the other hand for the co-crystallized sulfate ion close to the active site serine of the PCP domain. The latter makes sense, as the sulfate ion in the *apo*-structure was in a position that is likely to be occupied by the ppan-arm in the *holo*-bidomain, though the absence of electron density for the cofactor indicates its disorder. As the determined structure did not resemble the wanted conformation and due to the low resolution of the crystals, a refinement of the structure seemed pointless and was not performed.



## 4.2 Structural and Functional Basis for Phosphopantetheinylation of PCP

Although a model for the complex of a PCP in the A-state conformation and a PPT had been proposed<sup>38</sup> (PDB 2GE1) it was intended to obtain a more detailed view on the interaction of PCP and its modifying enzyme on an atomic level. X-ray crystallography was judged as a more promising method for structure determination than NMR, as the reactive complex of PCP and PPT comprises not only the two proteins, but also CoA and magnesium as cofactors. Still, NMR studies were performed simultaneously to get a better insight into the dynamics of the interaction and to exclude artifacts from the crystallization. Afterwards, possible interaction sites postulated from the structural data were validated and correlated with a function by mutational analysis.

### 4.2.1 Crystal Structure of the PCP/Sfp Complex

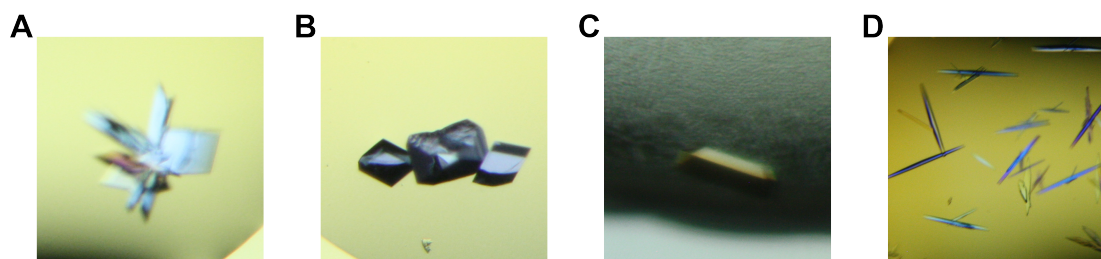
Sfp and TycC3\_PCP were chosen as proteins for the crystallization of the complex. Since it was already successfully crystallized, Sfp seemed to be a promising PPT for crystallization. TycC3\_PCP was chosen, as it exhibits properties convenient for crystallization like a high thermal stability and a low tendency towards aggregation or precipitation even at high concentrations. TycC3\_PCP is not a natural substrate of Sfp, as it is part of the tyrocidine NRPS, while Sfp is associated with the surfactin synthetase. Nevertheless, the choice of the two proteins is reasonable, as Sfp shares a high sequence identity of ~40% compared to the PPT from the tyrocidine cluster (GenBank entry BAH43769), and the similarity of PCPs from the two synthetases is even higher (e. g. ~44% identity for TycC3\_PCP and SrfAC\_PCP).

#### 4.2.1.1 Crystallization Trials and Structure Determination of the PCP/Sfp Complex

As the use of the wild type proteins in presence of CoA would have instantly led to the ppan transfer reaction, S45 of TycC3\_PCP was mutated to alanine by site

directed mutagenesis with the *tycC3\_PCP* gene in pQE70 and the primers *tycC3\_PCP*(S45A)-fw and -rv.

Both proteins were expressed in LB medium and purified by nickel affinity chromatography, as described in the previous section, followed by a SEC run in the crystallization buffer with an additional 5 mM DTT and MgCl<sub>2</sub>. The monomeric fractions of each protein were pooled separately and concentrated to ~1.6 mM. Samples for crystallization were prepared by mixing the two proteins to final concentrations of 750, 500, or 300 μM for each protein, whilst 1.5% (v/v) CoA (100 mM, pH 7.2) was added. The samples were used for crystallization trials with the sitting drop method. They were prepared on a Mosquito crystallization robot, using several purchasable kits mixed with the protein solution in a 1:1 ratio to a total drop size of 400 nL and 200 μL reservoir solution. Upon incubation at 20°C, crystal growth was observed within 1-5 days under several conditions, most of which contained a polyethylene glycol (PEG) species and sodium cacodylate (Figure 25).



**Figure 25:** Examples of crystals from the first condition screening for the PCP/Sfp complex grown in Natrix condition 17 (A) and 95 (B), Salt Rx condition 51 (C), and JCSG core III condition 55 (D). The latter were used for resolving the preliminary structure of the complex. (See table 5 for composition of the conditions)

These conditions were tested for reproducibility, and crystals were successfully obtained in 7 out of 11 tested conditions (Table 5). Crystals from the reproducible conditions were tested for diffraction on the FR-E+ in-house diffractometer at the Max Planck Institute of Biophysics, Frankfurt. Diffraction of the crystals grown in a 1:1 mixture of the 500 μM protein solution and condition 55 of JCSG core III [160 mM magnesium acetate; 80 mM sodium cacodylate (pH 6.5), 16% (w/v) PEG 8000 and 20% (v/v) glycerol] was even good enough to record a whole dataset, whereas crystals from other conditions showed minor diffraction or no diffraction at all.

**Table 5:** Conditions showing initial hits for the crystallization of the PCP/Sfp complex

Condition	Salts	Buffer	Precipitant	Additives	Reproducible
Natrix (17)	KCl, 200 mM CaCl <sub>2</sub> , 10 mM	Sodium cacodylate, 50 mM (pH 6.0)	PEG 4000, 10% (w/v)	-	YES
Natrix (23)	KCl, 200 mM MgCl <sub>2</sub> , 10 mM	Sodium cacodylate, 50 mM (pH 6.5)	PEG 4000, 10% (w/v)	-	NO
Natrix (26)	KCl, 200 mM Mg acetate, 100 mM	Sodium cacodylate, 50 mM (pH 6.5)	PEG 8000, 10% (w/v)	-	YES
Natrix (95)	CaCl <sub>2</sub> , 200 mM	HEPES, 50 mM (pH 7.5)	PEG 400, 28% (v/v)	Spermine, 2 mM	NO
JCSG+ (19)	-	Na acetate, 100 mM (pH 4.6)	PEG 4000, 8% (w/v)	-	NO
JCSG II (56)	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , 500 mM	TRIS, 100 mM (pH 7.0)	PEG 600, 30% (v/v)	Glycerol, 10% (v/v)	YES
JCSG III (55)	Mg acetate, 160 mM	Sodium cacodylate, 80 mM (pH 6.5)	PEG 8000, 16% (w/v)	Glycerol, 20% (v/v)	YES (for structure)
CrystalSuite (18)	Mg acetate, 200 mM	Sodium cacodylate, 100 mM (pH 6.5)	PEG 8000, 25.5% (w/v)	-	YES
CrystalCryo (15)	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , 170 mM	Sodium cacodylate, 85 mM (pH 6.5)	PEG 300, 20% (v/v)	Glycerol, 15% (v/v)	YES
PEG Rx (6)	-	BICINE, 100 mM (pH 8.5)	-	-	NO
Salt Rx (51)	(NH <sub>4</sub> ) <sub>2</sub> HPO <sub>4</sub> , 1.5 M	TRIS, 100 mM (pH 8.5)	-	-	YES

Structure determination was achieved with the recorded dataset by molecular replacement. A first attempt, using the known structures of Sfp in complex with CoA<sup>154</sup> (PDB 1QR0) and of TycC3\_PCP in the A state<sup>38</sup> (PDB 2GDY) as search models, was only partially successful, as the structure of Sfp fit into the electron density, but the PCP could not be fitted. In a second attempt, the model of TycC3\_PCP(S45A) solved simultaneously by NMR was used instead, whereupon the structure of the whole complex could be solved to 2.6 Å resolution (Table 6).

**Table 6:** Statistics for data collection and refinement of the structure of the PCP/Sfp complex

<b>Data collection</b>	<b>Preliminary structure</b>	<b>Final Structure</b>
X-ray source	Rigaku FR-E+	Swiss Light Source PX3
Space group	C2	C2
Cell dimensions		
a, b, c (Å)	160.78, 39.10, 53.37	160.58, 39.16, 53.26
$\alpha$ , $\beta$ , $\gamma$ (°)	90.00, 106.92, 90.00	90.00, 106.96, 90.00
Wavelength (Å)	1.50	1.00
Resolution (Å)	50.00-2.60 (2.64-2.60)	50.00-2.00 (2.05-2.00)
$R_{\text{merge}}$	0.10 (0.22)	0.26 (0.65)
$I / \sigma I$	15.4 (3.3)	6.7 (2.7)
Completeness (%)	94.2 (72.0)	99.4 (98.9)
Redundancy	3.2 (2.5)	6.2 (6.1)
<b>Refinement</b>		
Resolution (Å)		50.00-2.00
Number of reflections		20491
$R_{\text{work}} / R_{\text{free}}$		18.8/23.3
Number of atoms		
Protein		2424
Water		173
Ligand/Ion		61
B-factors		
Protein		29.1
Water		34.8
Ligand/Ion		25.1
RMSD		
Bond length (Å)		0.009
Bond angles (°)		1.486

The same 11 conditions that were used for testing the reproducibility of the crystallization with the sitting drop method were used for crystallization as hanging drops with 2  $\mu$ L drop size and 250  $\mu$ L condition in the reservoir. The grown crystals were tested for diffraction at the Swiss Light Source, Villigen. Again, the crystals from condition 55 of JCSG core III gave the best diffraction pattern, and a dataset was recorded. With the non-refined, preliminary structure of the complex as search model for molecular replacement, the structure of the complex could be solved from these data to a resolution of 2.0 Å and was further refined (Table 6). In doing so, parts of Sfp had to be remodeled, as their position had changed compared to the structure of Sfp and were missing in the preliminary structure.

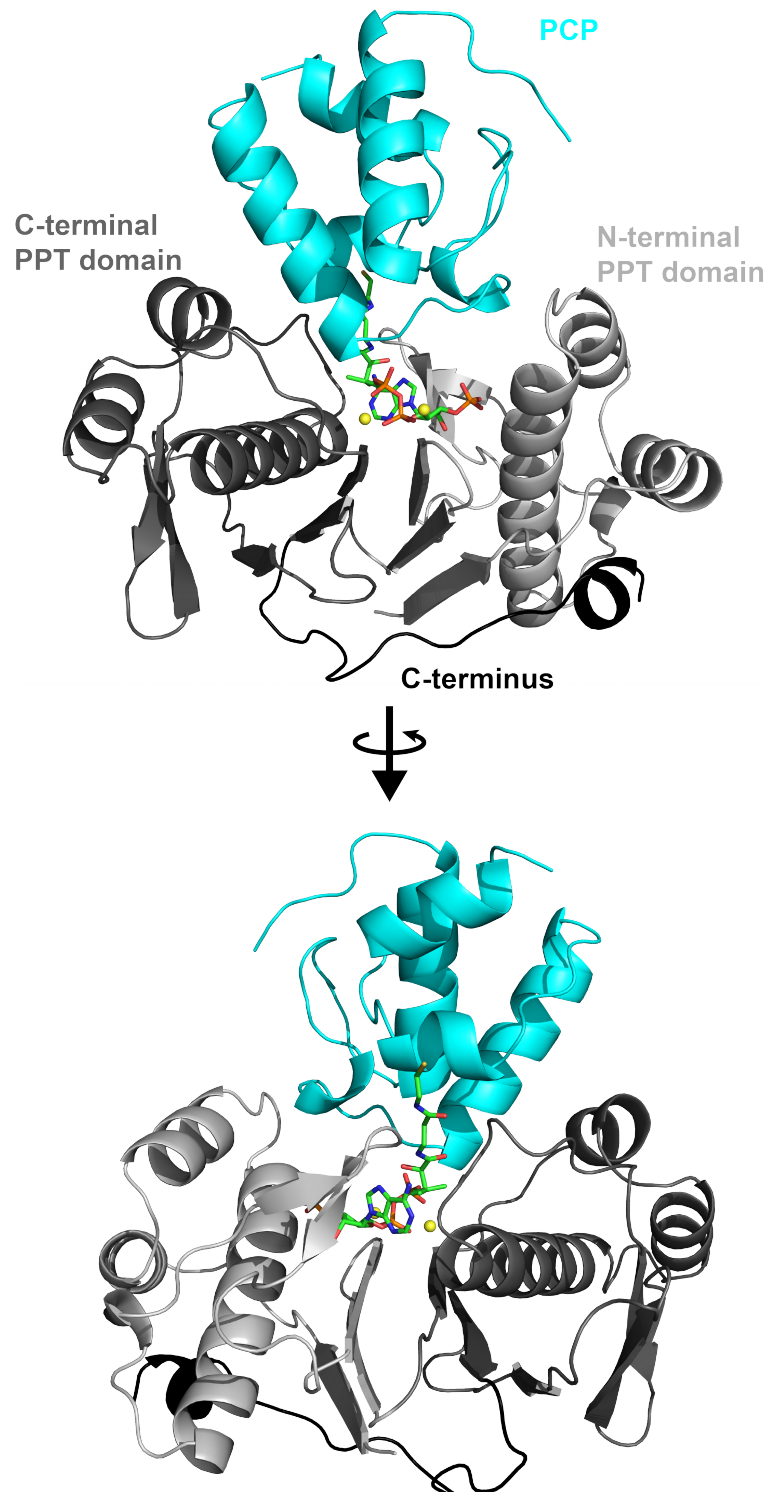
#### 4.2.1.2 The Three-dimensional Structure of the PCP/Sfp Complex

The structure, which was solved of the complex, comprises not only the two proteins, but also CoA, two magnesium ions, and five water molecules that seem to play an important role in the active site. Sfp is completely resolved from M1 to L224, and only its C-terminally tag is disordered, whereas to both termini of the PCP (P2-Q7 and S85-H91, respectively) no electron density could be assigned (Figure 26).

Sfp within the complex shows the typical structure of a group II PPT: It consists of two PPT domains (M1-D100 and I104-A206) with a similar fold (RMSD  $\sim$ 3.0 Å over 93 C $^{\alpha}$ s): Each domain consists of a central three-stranded antiparallel  $\beta$ -sheet ( $\beta$ 1/ $\beta$ 3/ $\beta$ 2), which acts as the interface for the pseudo-dimer alignment of the domains, and a stretch of  $\sim$ 80 residues inserted between  $\beta$ 1 and  $\beta$ 2, which is folded on one side of the  $\beta$  sheet. This stretch consists of two shorter  $\alpha$ -helices ( $\alpha$ 1 and  $\alpha$ 2) preceding a longer helix ( $\alpha$ 3), which is followed by a short, two-stranded antiparallel  $\beta$ -sheet ( $\beta$ 1'/ $\beta$ 2'). The C-terminus of Sfp (A207-L224) points towards the N-terminal domain, where it passes  $\beta$ 1 and ends as a short helix located in close proximity to  $\alpha$ 1 and  $\alpha$ 3.

A cavity is formed at one side of the interface. The edges of  $\beta$ 2 of the N-terminal and  $\beta$ 1 of the C-terminal PPT domain are located at the bottom of this cavity. One side is formed by  $\alpha$ 2 and  $\alpha$ 3 of the N-terminal and the opposite side by  $\alpha$ 1 and  $\alpha$ 3 of

the C-terminal PPT domain, respectively. Another side is blocked by  $\beta 1'/\beta 2'$  of the N-terminal domain, and the fourth side and the top of the cavity are open.



**Figure 26:** Overall structure of the complex of PCP (cyan) and Sfp (N-terminal domain: light gray; C-terminal domain: dark gray; C-terminus: black). Between the two proteins lies a coenzyme A molecule (sticks), complexing two magnesium ions (yellow spheres).

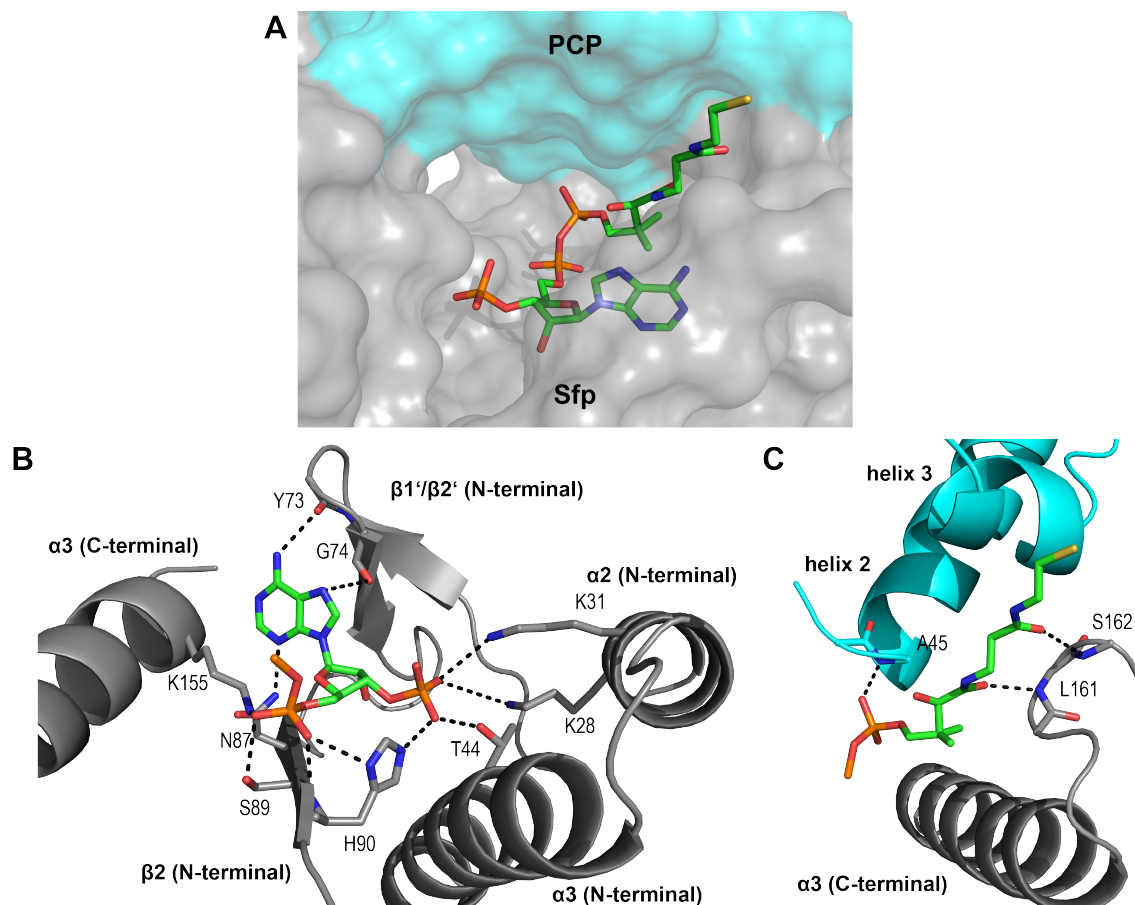
Compared to the structure of Sfp by Reuter *et al.*<sup>154</sup>, which was initially used for molecular replacement, the structure aligns well with an RMSD of 1.58 Å over 205 residues. The residues that do not align (P164-I180), form  $\beta 1'/\beta 2'$  in the C-terminal domain, which are relocated by  $\sim 90^\circ$ . This is accompanied by a movement of  $\alpha 1$  and  $\alpha 2$  towards  $\beta 1'/\beta 2'$  in this domain. Compared to the structure of the human PPT in complex with CoA<sup>205</sup> (PDB 2CG3), Sfp in complex with PCP shares a high similarity, too, as the RMSD for all residues is only 2.05 Å.

In contrast to the previous model of the PCP/Sfp complex<sup>38</sup> (PDB 2GE1), the PCP in the presented structure is not in the A state, but in the A/H state. Again, the structure is very similar to the known structures of TycC3\_PCP in this conformation with RMSDs of 2.44 Å compared to the structure by Koglin *et al.*, and of 1.34 Å to the structure of Weber *et al.*. However, it shares the highest similarity with the structure of peptide-loaded TycC3\_PCP, presented in this work (RMSD 0.62 Å). As seen before, the four helices of the PCP form a tight bundle, with the long loop between helix 1 and 2, as well as a part of the N-terminus associated with it.

CoA lies in the cavity of Sfp with its base in the *anti* conformation, oriented towards  $\beta 1'/\beta 2'$  of the N-terminal domain of Sfp and its sugar pointing to the bottom of the cavity. As the molecule kinks at the 5' pyrophosphate, which links the nucleoside to the pantetheine moiety, the latter is folded back and positioned on top of the sugar and the base. Above CoA lies the PCP, covering the top of the cavity (Figure 27A).

Binding of sugar, base, and the 3', as well as the 5'  $\alpha$ -phosphate group resembles the situation observed for the structure of Sfp in complex with CoA<sup>154</sup>, where mostly residues from the N-terminal PPT domain are involved: The backbone carbonyl groups of Y73 and G74 as well as the side chain amide of N87 form hydrogen bonds to the base. The 3' phosphate group is in contact with the side chains of K28, K31, T44, and H90, while the 5'  $\alpha$ -phosphate forms hydrogen bonds to the backbone amide as well as to the side chain of H90, to S89, and to K155 (Figure 27B). In contrast to the structure without PCP, the whole ppan portion of CoA shows electron density. The 5'  $\beta$ -phosphate group forms a hydrogen bond to the backbone amide of A45 from the PCP, and the pantetheine moiety is in contact with the third helix of the PCP as well as the loop following  $\alpha 3$  of Sfp's C-terminal PPT domain. The interaction with the helix is characterized by hydrophobic contacts to the side chains of F69 and

A70 of the PCP, whereas the two carbonyl groups of pantetheine form hydrogen bonds to the backbone amides of Sfp's L161 and S162 (Figure 27C).

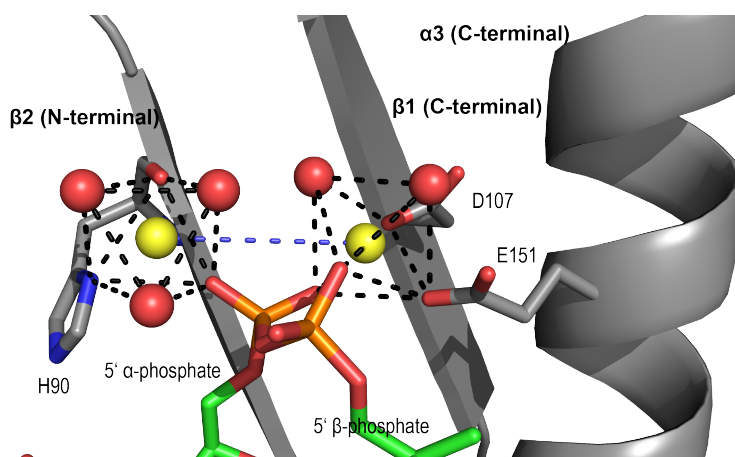


**Figure 27:** CoA lies in a cavity between PCP (cyan) and Sfp (gray) (A). Binding of CoA (sticks) to the proteins is facilitated by a widespread network of hydrogen bonds (dashed lines) and hydrophobic interactions. Mainly the N-terminal PPT domain of Sfp is involved in the binding of the adenosine 3',5' bisphosphate portion of CoA (B), whereas the ppan part is in contact with the C-terminal domain and PCP (C).

Beside the contacts with the two proteins, CoA is coordinated to two magnesium ions, which are present in the active site, too. The first magnesium ion was already observed in the previous structures of group II PPTs and was assigned as being involved in the catalysis of the ppan transfer<sup>154, 205</sup>. It is coordinated by the 5'  $\alpha$ - and  $\beta$ -phosphate groups of CoA, by the side chains of D107 and E151 in the C-terminal PPT domain of Sfp, and by two water molecules. The coordination of the magnesium ion by a second water molecule is a difference to the structure of Sfp in complex with CoA, where the coordination site of the second water molecule is occupied by the side chain of Sfp's E109. The presence of a second magnesium ion is a novelty that

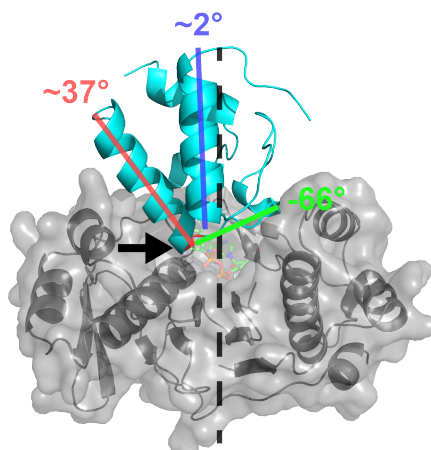


was not described in any of the known structures so far. It is oriented towards the N-terminal PPT domain of Sfp, where it is coordinated by the backbone and side chain of H90. The remaining four coordination sites are occupied by the 5'  $\alpha$ -phosphate group of CoA and three water molecules. Though the two magnesium ions do not share a ligand in their inner coordination shells, their distance of  $\sim 5.0$  Å is relatively small (Figure 28).



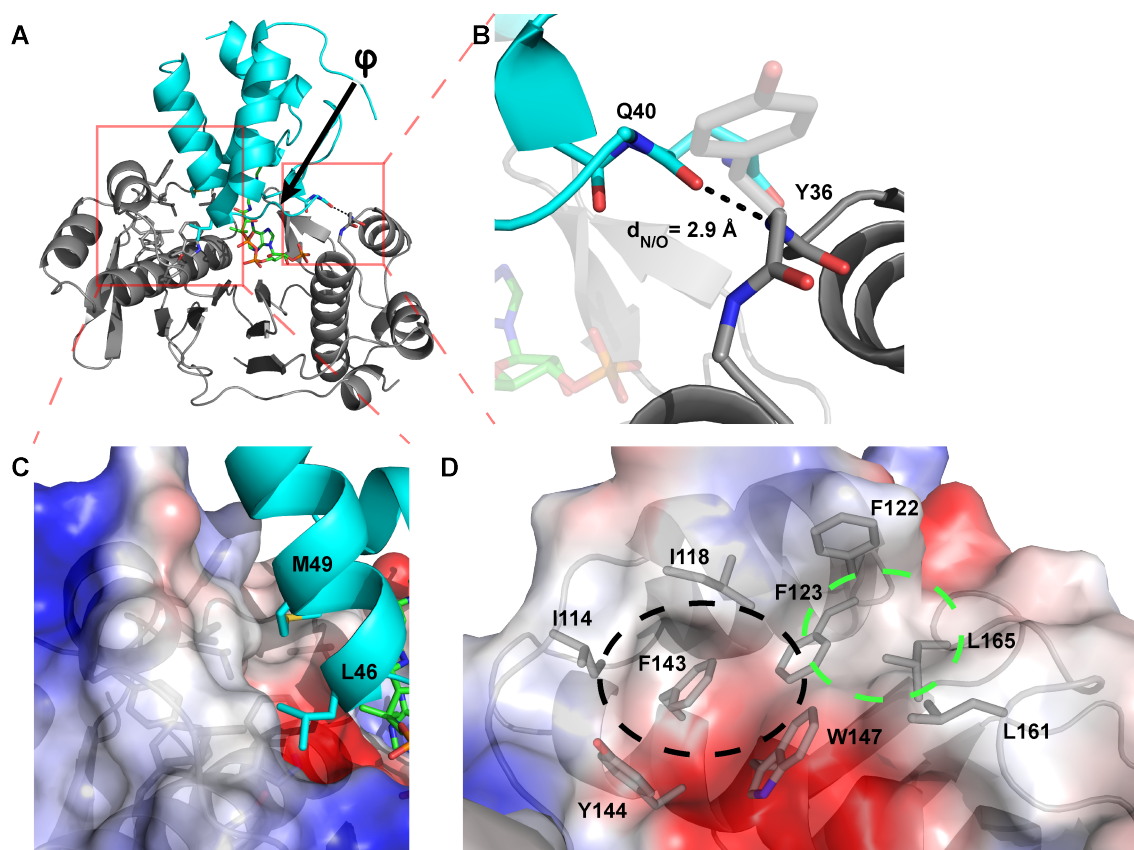
**Figure 28:** Two magnesium ions (yellow spheres) are located in the active site between CoA (sticks) and Sfp (gray). Both ions are  $\sim 5$  Å apart from each other (blue lines) and coordinated octahedrally (black lines) by CoA, residues from Sfp and water molecules (red spheres).

The structure gives insight not only into the binding of CoA and the magnesium ions in the complex, but also reveals the protein/protein interactions involved in the binding of the PCP by Sfp. A surface built by parts of the N-terminal half of helix 2 (A45-L46, M49-A50 and A52-A53 oriented towards  $\alpha 1$  and  $\alpha 3$  in the C-terminal PPT domain of Sfp), its preceding loop (F39-H44 oriented towards the loop connecting  $\alpha 2$  and  $\alpha 3$  of the N-terminal PPT domain), and helix 3 (L65-K66 and F69-A70 oriented towards  $\beta 1'/\beta 2'$  in the N-terminal PPT domain) points into the binding pocket of Sfp, where it lies on top of CoA. The helical axis of helix 1 of the PCP aligns almost parallel with the axis for the pseudo-C2 symmetry of Sfp. In this orientation, helix 2 deviates from the symmetry axis by  $\sim 37^\circ$ , so its N-terminus, where the invariant S45 would be located in the wild type protein, is the part of the PCP which is buried deepest in the binding pocket of Sfp (Figure 29).



**Figure 29:** Helix 1 of the PCP (blue line) aligns almost parallel with the symmetry axis of Sfp (black dashed line). Helix 2 (red) and the preceding loop (green) are oriented by  $\sim 37^\circ$  and  $-66^\circ$  towards this axis, so the N-terminus of helix 2 with the conserved serine (arrow) is pointing deepest into the cavity of Sfp.

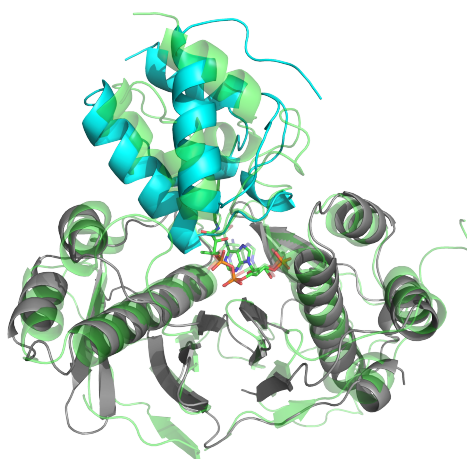
Since CoA lies between Sfp and PCP, only a few parts of the two proteins are in direct contact. The first of the two detectable protein/protein interaction sites is built by R34-Y36 from the loop between  $\alpha 2$  and  $\alpha 3$  in the N-terminal PPT domain and Q40/I41 from the loop preceding helix 2 of the PCP. In this relatively small area, the only directed intermolecular interaction that might contribute to the complex formation is a hydrogen bond between the backbones of the two proteins: The carbonyl group of Q40 points towards the amide nitrogen of Y36, from which it is  $\sim 2.9 \text{ \AA}$  away (Figure 30B). The second, larger protein/protein interface consists of L46 and M49 of the PCP as well as I114, I118, F122, F143, Y144, W147, L161, and L165 from Sfp. The residues of Sfp are located in  $\alpha 1$  (I118 and F122), the loops preceding it (I114),  $\alpha 3$  (F143, Y144 and W147), and the loop following it (L161 and L165), in the C-terminal PPT domain (Figure 30D). The side chains of the involved residues of Sfp point towards the binding pocket, of which they form one side. The side chains of L46 and M49, which are located at the same side of helix 2 of the PCP, are in contact with this surface (Figure 30B). As this interface is formed exclusively by non-polar side chains of both proteins, hydrophobic interactions are likely to be the driving force for this protein/protein contact. Q40, which forms the intermolecular hydrogen bond, and helix 2, which is involved in the hydrophobic interactions, are connected by a stretch of amino acids (I41-H44), which is arranged in a fashion in which G42 has a positive backbone torsion  $\phi$  of  $\sim 114^\circ$  (Figure 30A).



**Figure 30:** The two parts of PCP (cyan) that interact with Sfp (gray) are connected by a stretch with G42 (arrow), having a positive backbone torsion (A). One interaction site is formed by an intermolecular hydrogen bond (dashed line) between the backbones of the two proteins (B). Furthermore, the side chains of L46 and M49, located in helix 2 of the PCP, interact with a hydrophobic patch in the C-terminal domain of Sfp (shown in electrostatic surface representation) (C). A front view on the hydrophobic patch reveals two binding pockets for L46 (black circle) and M49 (green circle), respectively (D).

As already mentioned, Sfp exhibits a high similarity to the human PPT. From this protein in complex with a carrier protein, a crystal structure is known, too<sup>205</sup>. The structure of the human PPT in complex with the ACP domain of the human FAS resembles many aspects of the PCP/Sfp complex. The two PPTs from the complexes align very well with an RMSD of  $\sim 2.0$  Å over 221 residues. The position of CoA in the binding pocket is similar, too, though for parts of the ppan-arm in the human structure no electron density was observed. Due to the lower resolution, no water molecules and ions could be resolved in the active site of the human PPT, which is why it cannot be said for sure, if there are any magnesium ions present, and if they are coordinated similarly. Nevertheless, in the structure of the human PPT in complex with CoA, a magnesium ion at the previously known position is present, and its

coordination is similar to the situation in the active site of the PCP/Sfp complex. The protein/protein interaction in the complex of the human proteins is stabilized by an intermolecular hydrogen bond and by hydrophobic interactions in the same fashion as in the presented complex. Therefore, the orientation of helix 2 and its preceding loop of the PCP as well as of the ACP in relation to the PPTs are very similar, though the other parts of the carrier proteins do not align very well (Figure 31).



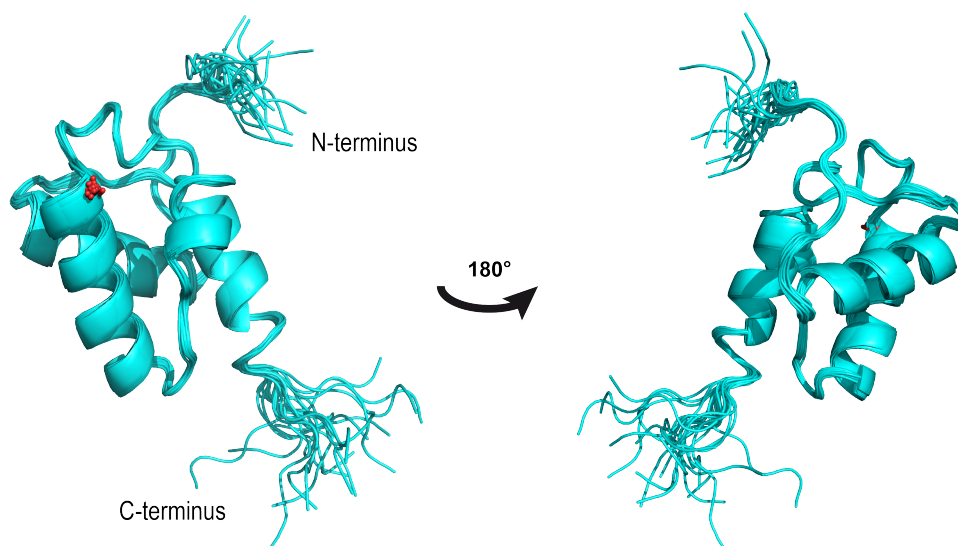
**Figure 31:** Overlay of the structures of the PCP/Sfp complex (cyan and gray) with the structure of the human ACP/PPT complex (semi-transparent green) illustrates the high similarity, though parts of the carrier proteins differ in their orientation.

#### 4.2.2 NMR Investigation of the Active Site Mutant PCP in Solution

In contrast to the crystal structure of the PCP/Sfp complex presented here, the study, in which it was proposed that the PCP in complex with Sfp is in the A state conformation, was based on NMR measurements in solution<sup>38</sup>. This raised the question, if the observed A/H state conformation of the PCP was an artifact of the crystallization process, since the A state is less well structured, which probably is accompanied by a reduced tendency to crystallize. To address this question, the structure of the active site mutant TycC3\_PCP(S45A) was solved by liquid state NMR, and titration experiments with Sfp were performed to analyze the PCP/Sfp complex in solution.

#### 4.2.2.1 Structure Determination of TycC3\_PCP(S45A)

Protein expression and purification, sample preparation, NMR measurements, resonance assignment, and structure calculation for TycC3\_PCP(S45A) were performed the same way like described for peptidyl(NH)-TycC3\_PCP, except that the sample used for NMR measurements had a protein concentration of 700  $\mu\text{M}$ , and the spectra were recorded at 298 K, omitting the  $[^1\text{H};^{13}\text{C}_{\text{aliphatic}}]$ -NOESY (Table 7). The completeness of the resonance assignment of TycC3\_PCP(S45A) was the same as for peptidyl(NH)-TycC3\_PCP. To ensure that the lack of NOE restraints from a  $[^1\text{H};^{13}\text{C}_{\text{aliphatic}}]$ -NOESY did not alter the quality of the structure calculation, the structure of peptidyl(NH)-TycC3\_PCP was calculated once again without including the information from the  $[^1\text{H};^{13}\text{C}_{\text{aliphatic}}]$ -NOESY. Since the structure of peptidyl(NH)-TycC3\_PCP looks virtually identical, no matter if it is calculated with or without the  $[^1\text{H};^{13}\text{C}_{\text{aliphatic}}]$ -NOESY, it was assumed that the structure calculation for TycC3\_PCP(S45A) was valid, although no restraints from a  $[^1\text{H};^{13}\text{C}_{\text{aliphatic}}]$ -NOESY were included.



**Figure 32:** Bundle of the 20 lowest-energy solution structures of TycC3\_PCP(S45A). The side chain of A45 is shown as red sticks.

The calculated structure (Figure 32) is very similar to the PCP in the crystal structure of the PCP/Sfp complex with an RMSD of  $\sim 0.7$  Å. This contradicts the previously made statement that the active site mutant PCP is exclusively in the A state<sup>38</sup>. This statement was made for the same construct of TycC3\_PCP(S45A)

under the same buffer conditions at the same temperature. Yet, it was not made based on a structure calculation, but on a comparison of the peak pattern in  $[^1\text{H};^{15}\text{N}]$ -HSQC spectra of *apo*- and mutant TycC3\_PCP.

**Table 7:** NMR structure determination statistics for TycC3\_PCP(S45A)

<b>NOE assignment<sup>a</sup></b>	
Total number of NOESY cross peaks	1693
Assigned cross peaks	1584 (=94% of total)
in $^{13}\text{C}_{\text{aromatic}}$ resolved NOESY	302 (95%)
in $^{15}\text{N}$ resolved NOESY	1282 (93%)
<b>Conformational restraints</b>	
Total NOE distance restraints	1419
Short range $ i - j  \leq 1$	661 (47%)
Medium range $1 <  i - j  < 5$	457 (32%)
Long range $ i - j  > 5$	301 (21%)
Dihedral angle restraints ( $\psi/\phi$ )	140
<b>Structure statistics<sup>b</sup></b>	
Average CYANA target function ( $\text{\AA}^2$ )	1.19±0.18
AMBER energies (kcal/mol)	-3192±59
<b>Restraint violations<sup>c</sup></b>	
Max. distance restraint violation ( $\text{\AA}$ )	0.13
Violated distance restraints $> 0.2 \text{\AA}$	0
Max. dihedral angle restraint violation ( $^\circ$ )	3.63
Violated dihedral angles $> 5^\circ$	0
<b>Ramachandran plot</b>	
Residues in most favored regions	89.8%
Residues in additionally allowed regions	10.1%
Residues in generously allowed regions	0.1%
Residues in disallowed regions	0.0%
<b>RMSD (residues 2-91)</b>	
Average backbone RMSD ( $\text{\AA}$ )	0.27±0.03
Average heavy atom RMSD ( $\text{\AA}$ )	0.67±0.04

<sup>a</sup> using the automated NOE assignment and structure calculation functionalities of CYANA

<sup>b</sup> after restrained energy minimization with OPALp

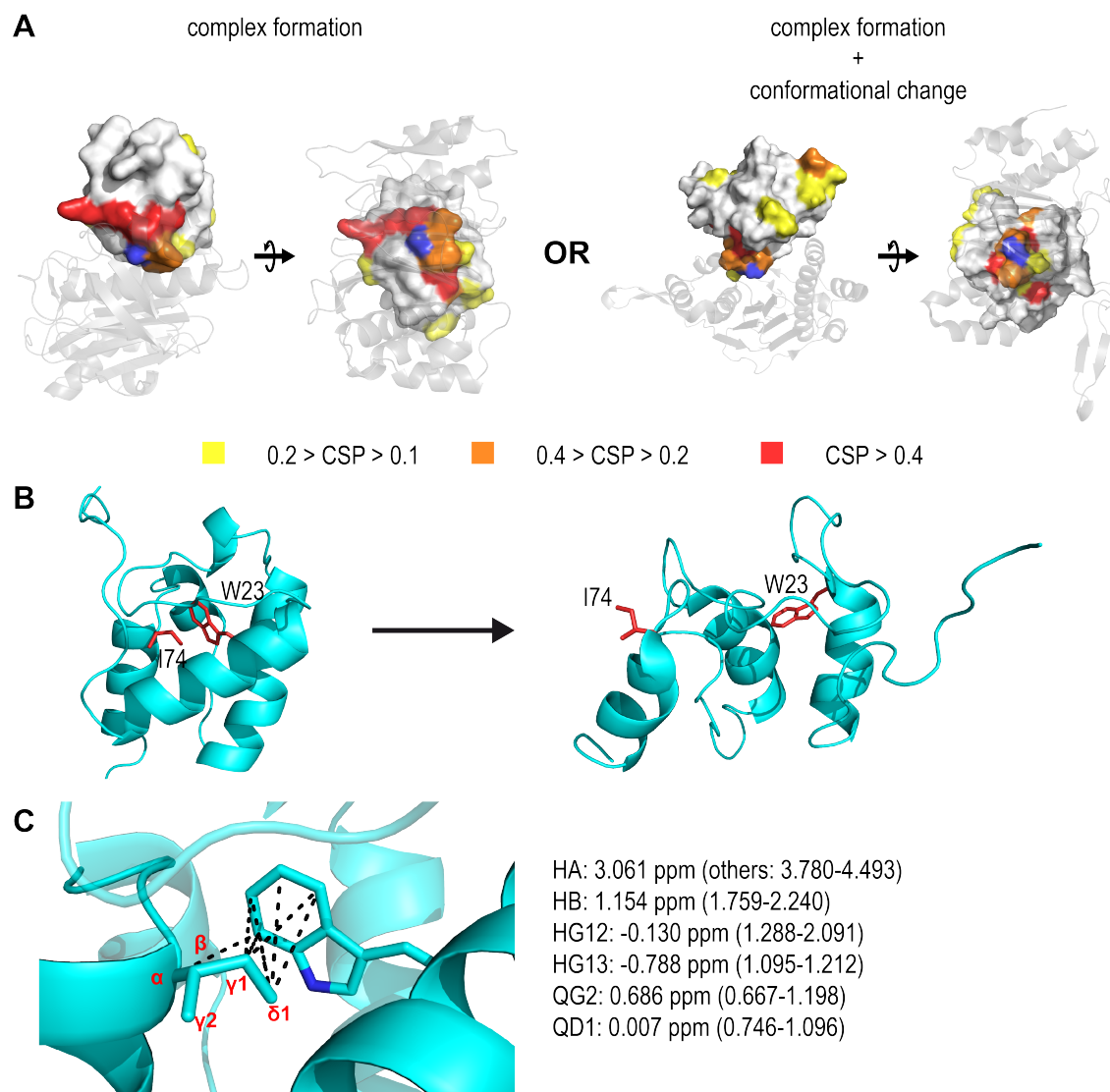
<sup>c</sup> after energy minimization, calculated with CYANA

#### 4.2.2.2 NMR Titration Experiments of TycC3\_PCP(S45A) with Sfp

After it had been shown that TycC3\_PCP(S45A) is in the A/H state conformation in solution, too, the question, if the PCP changes its conformation upon binding to Sfp, was raised. To answer this, titration experiments with uniformly [ $^{13}\text{C};^{15}\text{N}$ ]-labeled TycC3\_PCP(S45A) and unlabeled Sfp were performed. The proteins were expressed as before and purified by nickel affinity purification and subsequent SEC with the buffer for NMR with additional DTT and  $\text{MgCl}_2$ , each with a concentration of 5 mM. After purification, CoA was added to both proteins to a final concentration of 1 mM. As references [ $^1\text{H};^{15}\text{N}$ ]-, [ $^1\text{H};^{13}\text{C}_{\text{aliphatic}}$ ]-, and [ $^1\text{H};^{13}\text{C}_{\text{aromatic}}$ ]-HSQC spectra of a 100  $\mu\text{M}$  PCP sample were recorded. Afterwards, Sfp was added stepwise, and for each titration step the same set of spectra was recorded.

In all three types of spectra of the PCP alone, only one set of peaks was detectable. In the course of the titration experiment, additional peaks were rising and the intensity of some of the original peaks decreased and peaks vanished in presence of 4 eq Sfp. These circumstances indicate a complex formation in the slow exchange time regime: The life-times of the two states (free and complexed PCP) are long enough to be resolved as two separate peaks. CSPs of signals from the same residue in the [ $^1\text{H};^{15}\text{N}$ ]-HSQC spectra were mapped on the presented crystal structure of TycC3\_PCP(S45A) in complex with Sfp and on the model of TycC3\_PCP in the A state conformation in complex with Sfp, respectively. In both complexes, the residues showing the strongest CSPs cluster around the active site and other parts of the PCP which are in close proximity to Sfp, making it impossible to distinguish between the two cases by this mapping approach (Figure 33A). Instead, conformation specific side chain signals were searched for in the other spectra. The most prominent case are the signals of W23 and I74 (Figure 33B). Being more than 50 residues apart in the primary sequence of the PCP, their side chains are in close proximity in the structure solved for the mutant PCP in the A/H state. This contact can be retraced directly by several NOEs in the  $^{13}\text{C}_{\text{aromatic}}$ -NOESY and indirectly by the proton resonances of signals from the side chain of I74: As these protons are in close proximity to the aromatic of W23, their signals are shifted upfield due to the local magnetic effect of the  $\pi$ -electron system (Figure 33C). In the A state conformation,

these side chains are more than 15 Å apart, so a significant change in the chemical shifts of the resonances from the side chain protons of I74 would be expected, if the PCP undergoes a transition from the A/H to the A state upon binding of Sfp.



**Figure 33:** Mapping of the chemical shift perturbations in  $[\text{H};^{15}\text{N}]$ -HSQC spectra from a titration experiment of PCP (surface) with Sfp (semi-transparent gray) makes sense, no matter if the complexed PCP is in the A/H state (left) or in the A state (right). The magnitude of the perturbation is color-coded from yellow (weak) to red (strong), whereas the active site serine is shown in blue (A). W23 and I74 (red sticks) were seen as a more promising probe for a potential change in the conformation, as they are in close proximity in the A/H state (left) and far apart in the A state (right) (B). The contact between these two side chains in the A/H state was validated by several NOE signals (dashed lines). Furthermore, an effect of the aromatic on the resonances of the side chain of I74 is evident. Only QG2, which is pointing away from the aromatic, is not effected significantly. Resonances of I74 are listed in comparison with the resonances of the same groups in the four other isoleucine residues of TycC3\_PCP(S45A) (C).

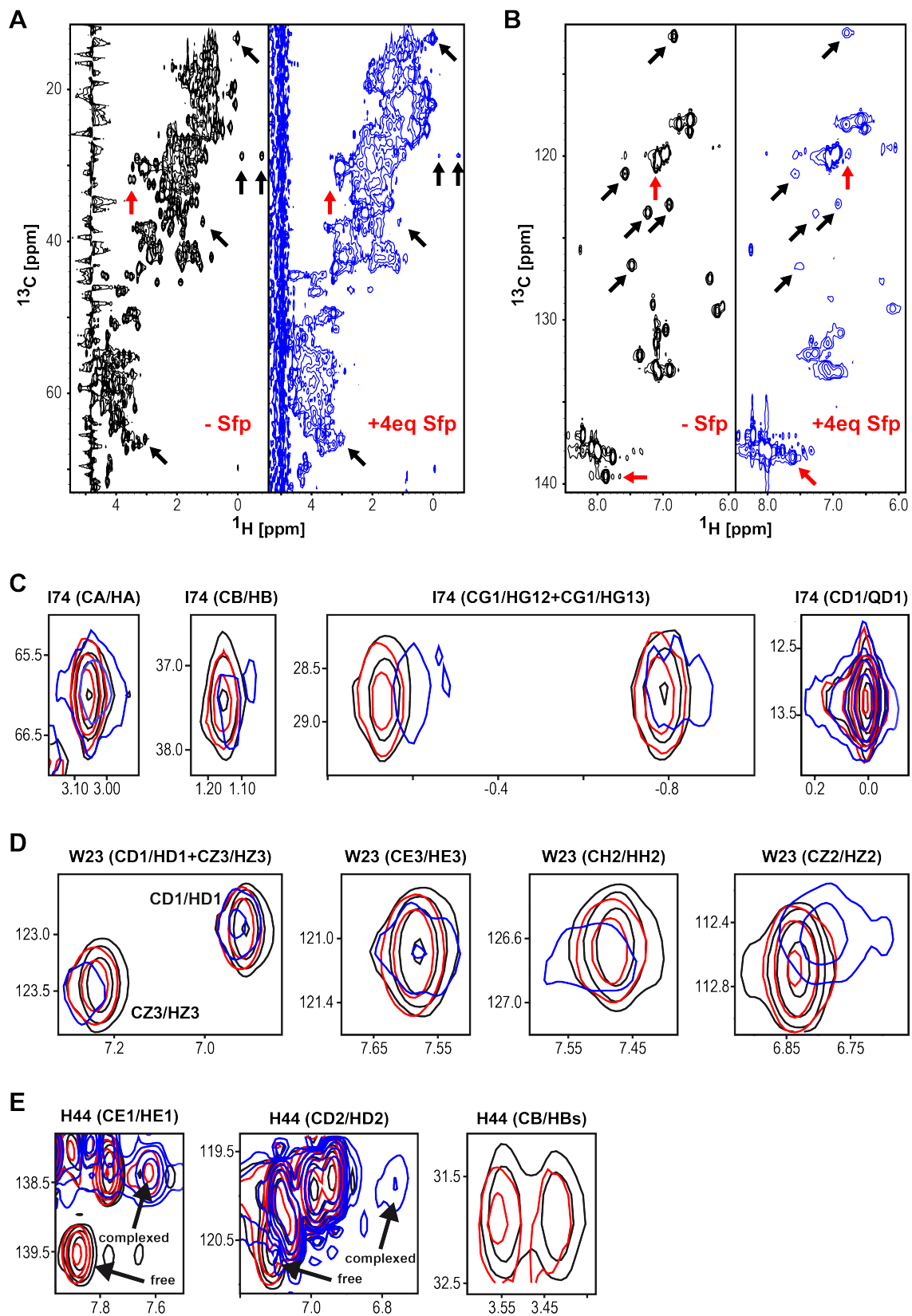


Already, the high similarity of the  $[^1\text{H};^{13}\text{C}_{\text{aliphatic}}]$ - and  $[^1\text{H};^{13}\text{C}_{\text{aromatic}}]$ -HSQC spectra of TycC3\_PCP(S45A) in absence and presence of Sfp (Figure 34A+B) indicates that the PCP does not undergo a large conformational change, as the transition from the A/H to the A state would be. Since, except for the  $\gamma$ -methyl group, all side chain signals of I74 are well separated from other peaks in the  $[^1\text{H};^{13}\text{C}_{\text{aliphatic}}]$ -HSQC spectra, these conformation specific signals could easily be monitored (Figure 34C). Their intensity decreases when Sfp is added, but their positions do not change significantly. While the decrease of their intensity can be explained with line-broadening due to the complex formation, the minor CSPs indicate that the environment of the side chain of I74 does not change when the complex forms. The same is true for the signals of W23 in the  $[^1\text{H};^{13}\text{C}_{\text{aromatic}}]$ -HSQC spectra. Their intensities decrease, but they do not shift dramatically (Figure 34D).

To exclude that the observed signals represent a population of non-complexed PCP, they were compared with the signals of H44 (Figure 34E). Flanking the (mutated) active site serine N-terminally, the chemical environment of H44 changes when PCP binds to Sfp, no matter if the transition from the A/H to the A state occurs, or not. Therefore, double peaks corresponding to the free and the complexed PCP were observed for signals from H44 in presence of 0.5 eq Sfp, but in presence of 4.0 eq Sfp only the signals from the complexed PCP remain. Thus, it is clearly shown that the signals observed in the spectra recorded in presence of 4.0 eq Sfp come exclusively from the complexed PCP.

**Figure 34** (Next page)

Comparison of  $[^1\text{H};^{13}\text{C}_{\text{aliphatic}}]$ - (A) and  $[^1\text{H};^{13}\text{C}_{\text{aromatic}}]$ -HSQC spectra (B) of TycC3\_PCP(S45A) without (black) and with Sfp (blue). The similar overall peak pattern in presence and absence of Sfp indicates that no major conformational change occurs upon Sfp-binding. This finding is confirmed by the fact that the conformation-specific resonances of I74 and W23 (black arrows) are hardly affected, while the signals of H44 (red arrows) change their position significantly. The signals from I74 (C), W23 (D), and H44 (E) in presence of no (black), 0.5 eq (red), and 4.0 eq (blue) Sfp are shown in detail.



### 4.2.3 Production of Mutant TycC3\_PCP and Sfp

Based on the crystal structure of the PCP/Sfp complex two important protein/protein interaction sites were proposed: The intermolecular hydrogen bond between Q40 of PCP and Y36 of Sfp as well as the hydrophobic patch in the C-terminal PPT domain of Sfp, to which helix 2 of the PCP binds to via the side chains of L46 and M49. To investigate the impact of these interactions on the complex stability and the ppan transfer reaction mutants, of both proteins were made.

As the intermolecular hydrogen bond is formed between the backbones of the two proteins with Y36 as hydrogen donor, a Y36P mutation was introduced in Sfp to disrupt the interaction. For L46 it was assumed that its side chain fills a hydrophobic cavity formed by residues of Sfp. Thus, mutations to amino acids with a small hydrophobic (L46A), with a polar (L46N), and with a charged side chain (L46D) were made. The side chain of M49 does not seem to point into such a cavity, but is kinked with the terminal methyl group pointing to the PCP (Figure 30B). Therefore, M49 was only mutated to aspartate. Furthermore, the backbone torsion of G42 was assumed to be important for a correct alignment of the two interaction sites in the PCP. To probe this assumption, G42 was mutated to alanine, which cannot adopt this torsion. The same mutation was introduced for G43 as a positive control, as its torsion can be adopted by alanine, too. Mutant and wild type proteins were analyzed by ITC and a priming assay.

#### 4.2.3.1 Cloning an Expression of the Mutants

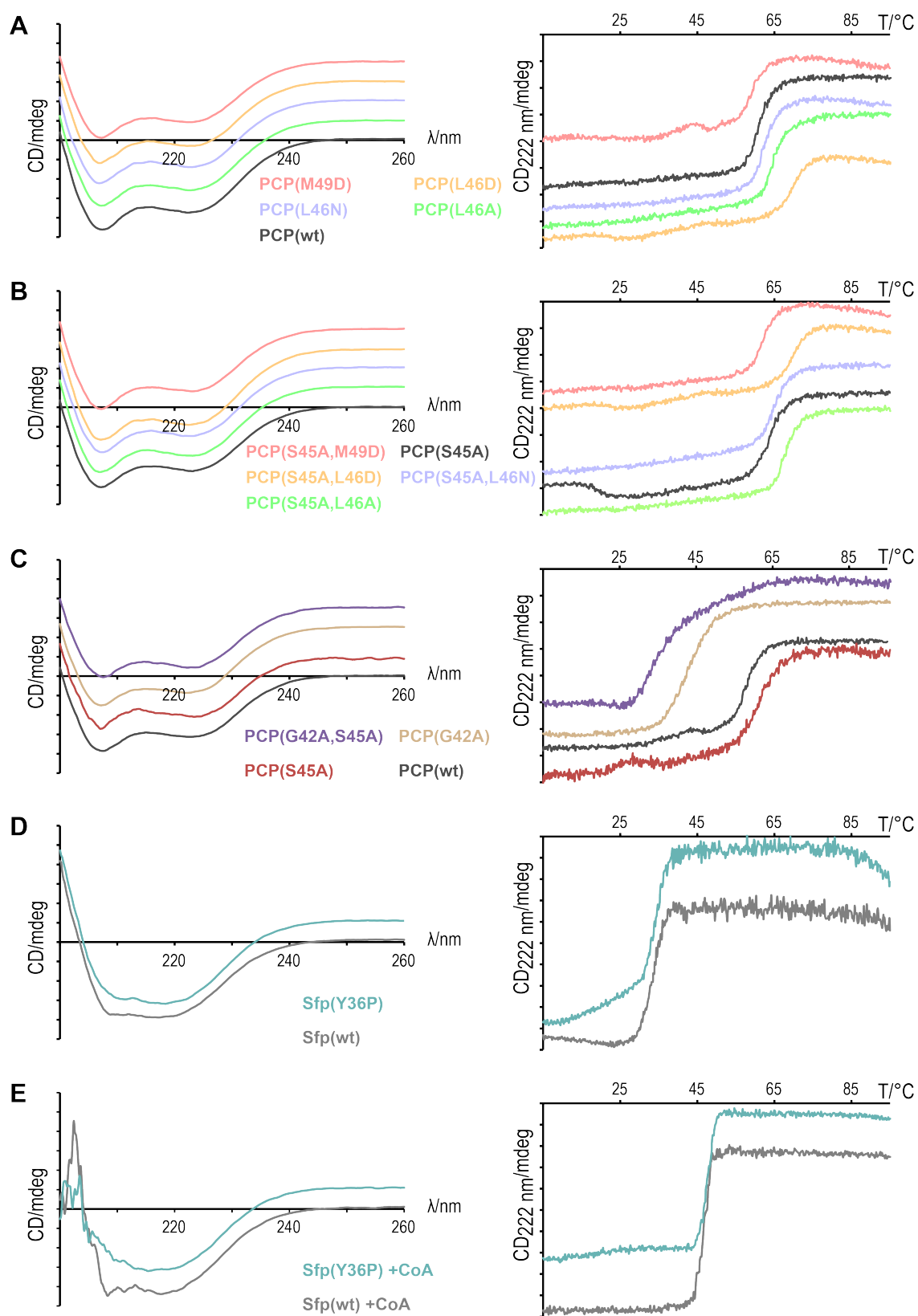
*tycC3\_PCP* in pQE70, *tycC3\_PCP(S45A)* in pQE70, and *sfp* in pQE60 were used as templates for site directed mutagenesis. All mutations could be introduced, and the constructs were used for expression in *E. coli* M15(pREP4), using LB as described for the wild type proteins. Expression, nickel affinity purification and SEC resulted in soluble, monomeric protein with yields of ~60 mg per liter culture medium for Sfp(Y36P) and 5-15 mg for all PCP mutants but those harboring a glycine to alanine mutation. TycC3\_PCP(G42A), TycC3\_PCP(G42A,S45A), TycC3\_PCP(G43A), and TycC3\_PCP(G43A,S45A) could not be expressed soluble

under several tested conditions, and by nickel affinity purification with 6 M urea in the buffers, no protein could be obtained, either. To test a different expression system, the genes were subcloned into pBH4(His6) and used for expression in *E. coli* BL21(DE3). Again, no protein expression was detectable, although *tycC3\_PCP* and *tycC3\_PCP(S45A)* were successfully used for expression in *E. coli* BL21(DE3) when subcloned into pBH4(His6).

As several attempts to express the proteins with a mutated glycine *in vivo* had failed, the constructs in pBH4 were used for expression *in vitro*, using a cell-free expression system designed in our lab<sup>191</sup>. All constructs could be expressed by this approach, and after optimization of the magnesium concentration 3-5 mg protein were obtained from 1 mL reaction mixture. TycC3\_PCP and TycC3\_PCP(S45A) were successfully expressed in the same manner for use in control experiments. Though the proteins were completely insoluble, they could be dissolved in 6 M urea and refolded by dialysis. The resulting proteins were relatively pure, so the his-tag was directly cleaved with TEV protease, followed by reversed nickel affinity purification and a SEC run. During the purification, TycC3\_PCP(G43A) and TycC3\_PCP(G43A,S45A) appeared to be relatively instable even at 4°C with a tendency to precipitate.

#### 4.2.3.2 Quality Control of the Mutant Proteins

Though all proteins could be produced as soluble monomers, they were further analyzed by CD spectroscopy to exclude a different fold induced by the mutation and, in case of the cell-free expressed proteins, by the denaturation and subsequent refolding during the purification. CD spectra of all proteins but TycC3\_PCP(G43A) and TycC3\_PCP(G43A,S45A) were recorded in 10 mM NaP<sub>i</sub> (pH 6.8) at 5°C. The G43A mutants were too instable in this buffer to be used for measurements. In case of Sfp(wt) and Sfp(Y36P), the spectra were measured in presence of 0.5 mM CoA and 1 mM MgCl<sub>2</sub>, too. Afterwards, melting curves of each protein were recorded to check if the mutations had influenced their thermal stability.



**Figure 35:** Figure legend on the following page.

**Figure 35** (previous page)

CD spectra (left panels) of TycC3\_PCP variants with an intact active site serine (A), mutated active site (B) and expressed *in vitro* (C) all look very similar indicating a similar fold of all PCP variants. The same is true for most of the melting curves (right panel). Only the proteins harboring the G42A mutation are less stable.

Wild type and mutant Sfp have comparable CD spectra and melting curves in absence (D) as well as in presence of CoA (E). In the latter case the quality of the CD spectra at lower wavelengths is reduced due to the higher conductivity of the solution whereas the cofactor increases the stability of the proteins.

For clarity reasons the curves are shifted vertically by different offsets. The color code of the curves is given in the picture.

The CD spectra of all variants of TycC3\_PCP looked almost alike independent of the introduced mutations and the strategy for protein preparation. In contrast to that, the melting curves of some mutants differ: While most PCP constructs have a melting temperature of  $\sim 65^\circ\text{C}$ , TycC3\_PCP(G42A) and TycC3\_PCP(G42A,S45A) are less stable with melting temperatures of  $\sim 45^\circ\text{C}$  (Figure 35A-C). The CD spectrum of Sfp seems not to be altered when the Y36P mutation is introduced either. Furthermore, wild type and mutant Sfp have a similar melting point of  $\sim 35^\circ\text{C}$ , which increases by approximately  $10^\circ\text{C}$  in the presence of CoA (Figure 35D+E). It is noteworthy that the proteins were not stable at all temperatures below their melting points as approximated from the CD melting curves. Wild type and mutant Sfp in absence of CoA even precipitated slowly at  $5^\circ\text{C}$ , especially at higher concentrations. Still, none of the used proteins formed soluble aggregates, since SEC runs of solutions clarified by centrifugation always showed monomeric protein exclusively.

To further probe if the binding of CoA is effected by the Y36P mutation in Sfp, ITC experiments with both Sfp variants and CoA were performed at  $5^\circ\text{C}$  in the buffer for NMR supplemented with 2 mM TCEP and 5 mM  $\text{MgCl}_2$ .

**Table 8:** Thermodynamic data for the binding of CoA by Sfp at  $5^\circ\text{C}$  derived by ITC

Sample	$K_D$ [ $\mu\text{M}$ ]	$\Delta H$ [ $\text{kJ mol}^{-1}$ ]	$\Delta S$ [ $\text{J mol}^{-1} \text{K}^{-1}$ ]	N
Sfp(wt) + CoA	$1.49 \pm 0.17$	$-16.4 \pm 0.3$	52.3	$0.78 \pm 0.01$
Sfp(Y36P) + CoA	$1.57 \pm 0.18$	$-16.1 \pm 0.2$	53.2	$1.10 \pm 0.01$

Again, no significant differences could be observed between wild type and mutant Sfp (Table 8).

#### 4.2.4 Impact of Mutations on the Stability of the PCP/Sfp Complex

To get thermodynamic information about the stability of the PCP/Sfp complex, ITC titration experiments of the different PCP mutants with Sfp(wt), and in case of TycC3\_PCP(S45A) also with Sfp(Y36P), were performed. Titrations were carried out in presence of CoA at 5°C and 25°C, though TycC3\_PCP(G42A,S45A) was not stable enough at the higher temperature and was only tested at 5°C.

TycC3\_PCP(S45A) binds to Sfp(wt) with a  $K_D$  of less than 1  $\mu$ M at both temperatures, no matter if the PCP was expressed *in vivo* or *in vitro*. Yet, the contribution of entropy and enthalpy differ at 5°C and 25°C, as the complex formation is endothermic but entropy-driven at the lower temperature and becomes exothermic with a reduced entropic contribution at the higher temperature. By introduction of the Y36P mutation in Sfp, the affinity is reduced by a factor of  $\sim 70$  at 5°C, and no binding was detectable at 25°C. The interaction of TycC3\_PCP(G42A,S45A) with Sfp(wt) showed a significantly reduced binding compared to TycC3\_PCP(S45A)/Sfp(wt), as the  $K_D$  was  $\sim 40$   $\mu$ M. All mutations of L46 and M49 in helix 2 of the PCP resulted in no detectable binding at both temperatures (Table 9).

**Table 9:** ITC data for titration with different PCP and Sfp mutants

Sample	T [°C]	$K_D$ [ $\mu$ M]	$\Delta H$ [kJ mol <sup>-1</sup> ]	$\Delta S$ [J mol <sup>-1</sup> K <sup>-1</sup> ]	N
S45A + Sfp(wt)	5	0.74±0.04	15.6±0.1	173.8	1.23±0.01
S45A + Sfp(wt)	25	0.92±0.12	-10.7±0.1	80.0	1.22±0.01
S45A + Sfp(Y36P)	5	54.3±6.5	18.2±2.3	147.0	1.00±0.10
S45A + Sfp(Y36P)	25	-	-	-	-
S45A + Sfp(wt) <sup>a</sup>	5	0.94±0.08	9.49±0.1	149.5	0.98±0.01
G42A,S45A + Sfp(wt) <sup>a</sup>	5 <sup>b</sup>	39.7±2.2	24.3±1.2	171.7	1.08±0.04
S45A,L46A + Sfp(wt)	5/25	-	-	-	-
S45A,L46N + Sfp(wt)	5/25	-	-	-	-
S45A,L46D + Sfp(wt)	5/25	-	-	-	-
S45A,M49D + Sfp(wt)	5/25	-	-	-	-

<sup>a</sup> expressed cell-free

<sup>b</sup> TycC3\_PCP(G42A,S45A) was not stable at 25°C

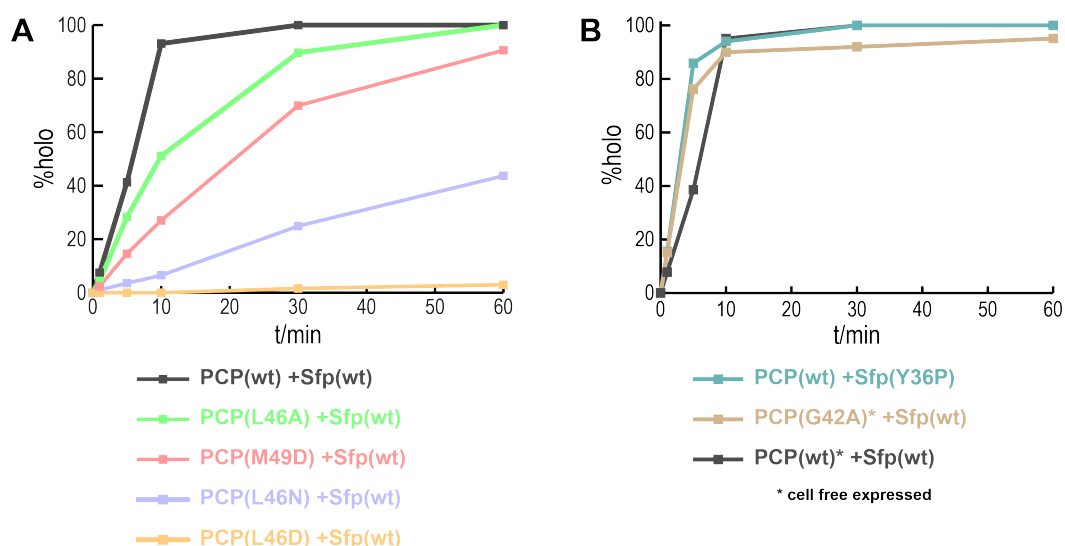
- no binding detectable ( $K_D > 100$   $\mu$ M)

TycC3\_PCP(G43A,S45A) was tested, too, but was not stable enough even at 5°C.

#### 4.2.5 Velocity of the PCP Priming Reaction for Different Mutants

To correlate the thermodynamic data obtained by ITC with kinetic information, a time dependent priming assay was used. Separation of *apo*- and *holo*-PCP for pQE70 constructs was accomplished with a linear gradient from 45-54.5% MeCN at 50°C and from 45-59% MeCN at 50°C for pBH4 constructs.

Complete conversion of TycC3\_PCP(wt) to its *holo*-form catalyzed by Sfp(wt) was observed after 30 min, whereas PCPs with mutation of L46 and M49 showed a reduced velocity of the reaction. The effect of the L46A mutation was not as strong as for L46N, and the introduction of an aspartate at this position almost suppresses the reaction completely, whereas the same mutation for M49 had a less dramatic effect (Figure 36A).



**Figure 36:** Time dependent formation of *holo*-PCP catalyzed by Sfp. The velocity of the reaction is slower for all PCPs harboring a mutation in helix 2 compared to TycC3\_PCP(wt) (A).

Sfp(wt) modifies the wild type PCP with the same efficiency, no matter if the PCP was expressed *in vivo* or *in vitro*. The formation of *holo*-TycC3\_PCP(G42A) catalyzed by Sfp(wt) is even faster with a velocity comparable to the reaction of TycC3\_PCP(wt) with Sfp(Y36P) (B). Due to its thermal instability, a portion of *apo*-TycC3\_PCP(G42A) precipitated before it was modified during the incubation. Error bars were omitted for clarity as they were rather small (max. 2.7%).

Repetition of the reaction with *in vitro* expressed TycC3\_PCP(wt) resulted in a virtually identical reaction profile. When Sfp(wt) was substituted by the Y36P mutant, the velocity was increased and the same effect was observed when TycC3\_PCP(G42A) was incubated with the wild type enzyme (Figure 36B).



## 5. Discussion and Outlook

### 5.1 Interactions in the Peptide Bond Formation

#### 5.1.1 Activity of TycC4\_C

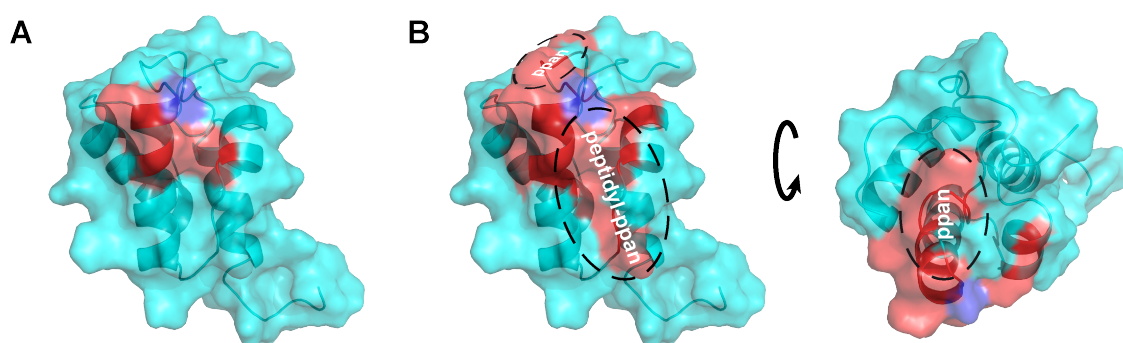
For the investigation of the interaction of a peptidyl-PCP with its downstream C domain the structurally well-characterized TycC3\_PCP was chosen. As a consequence of this choice, its natural interaction partner TycC4\_C was used. Both proteins could be expressed as soluble monomers; however, in the course of the project three experiments have indicated that the excised TycC4\_C is not active: First of all, no differences in the velocity of the hydrolysis of the thioester bond of peptidyl-TycC3\_PCP have been observed in the presence and absence of TycC4\_C (Figure 19A), although it was shown that the C domain in the TycC5-6\_PCP-C bidomain catalyzes the hydrolysis of the PCP-bound peptide<sup>69</sup>. Secondly, TycC4\_C has not effected the hydrolysis of valyl-TycC4\_PCP (Figure 19B). According to the acceptor site model of the C domain<sup>75</sup>, it would be expected that valyl-TycC4\_PCP binds to TycC4\_Cc resulting in a sterical protection of the thioester against hydrolysis. Finally, no product formation was observed in the condensation reaction. Since the C domain was monomeric as determined by SEC, it is unclear why it has not been active. One explanation might be that the binding to the PCPs in *trans* is too weak to have an apparent effect on the hydrolysis or on the catalysis of the condensation reaction, although cases are known in which other C domains have still been active or showed binding to PCPs in *trans* when excised from the NRPS<sup>74, 94</sup>. Lacking an active C domain, the recognition of the PCP-bound peptide by the C domain was not further investigated by NMR-titration experiments with separated domains, as it had been planned initially.

### 5.1.2 The Structure of peptidyl(NH)-TycC3\_PCP

During the activity tests of TycC4\_C, it has been shown that TycC3\_PCP and TycC4\_PCP can be quantitatively loaded with their cognate peptide and amino acid, respectively, so their interaction with the cofactors could potentially be investigated. However, the hydrolysis of the thioester of valyl-TycC4\_PCP was too fast to allow any structural investigation by NMR. Without any reference spectra of the native valyl-TycC4\_PCP at hand, the use of a non-hydrolyzable analog was not considered, as an effect of the substitution on the structure could not be ruled out.

On the other hand, the hydrolysis of peptidyl-TycC3\_PCP was slow enough to record a HSQC spectrum and compare it with the same spectrum of peptidyl(NH)-TycC3\_PCP. The differences in these spectra are rather small (Figure 20B), so it has been concluded that the introduction of the amide bond did not alter the structure. Structure elucidation of peptidyl(NH)-TycC3\_PCP revealed that the PCP is in the A/H state conformation. The small number of NOE contacts between the PCP and the peptidyl(NH)-ppan indicates that the cofactor is very flexible and has no distinct orientation relative to the PCP. In contrast to the flexible peptidyl-ppan moiety, the loop preceding helix 2 of the PCP appears to be tightly attached to the helical core of the domain (Figure 22). Therefore, it seems unlikely that it can open a cavity for the peptide-loaded cofactor. This is a clear difference to ACPs of type II FAS, for which it has been observed that the growing fatty acid chain is buried between the helices of the carrier protein<sup>138, 139</sup>. This finding makes sense, as the growing fatty acid chain has similar hydrophobic properties like the helical core of the ACP, whereas the peptide intermediates differ in their hydrophobicity but are in general more polar. Besides that, a comparable cavity was not observed for an ACP from a type I FAS<sup>135</sup>. In contrast to type II FASs, the different catalytic domains of a type I FAS are not distributed among different proteins, but are fused into one or two polypeptide chains. Thus, it is not necessary to protect the intermediate from hydrolysis while the ACP migrates from one protein to the next, but a good accessibility for the different catalytic domains within the same protein is preferable. As NRPSs are also multi-domain enzymes in which the PCPs interact mainly in *cis*, the similarity to ACPs from type I FASs makes sense, too.

Still, there is some evidence that the loading of the peptide onto the PCP has an effect on its structure: Comparison of the HSQC spectra of *holo*- and peptidyl(NH)-TycC3\_PCP reveals significant CSPs for certain residues (Figure 20A). These residues cluster in two regions around the active site serine when mapped on the structure of the PCP (Figure 37B). Remarkably, these are the region where the ppan-arm is located in the *holo*-PCP<sup>38</sup> and the region showing NOEs to the cofactor in peptidyl(NH)-TycC3\_PCP (Figure 37A).



**Figure 37:** Mapping of the NOEs observed between the protein and the cofactor (red) in peptidyl(NH)-TycC3\_PCP indicated that the cofactor is in loose contact with parts of helix 2 and 3 neighboring the active site serine (blue) (A). Mapping of the shift differences in the  $[^1\text{H};^{15}\text{N}]$ -HSQC spectra of *holo*- and peptidyl(NH)-TycC3\_PCP (red) shows that the same region is effected by loading of the peptide to the ppan-cofactor. A second area showing differences is located at the opposite side of the active site serine. In *holo*-TycC3\_PCP, this part of the loop preceding helix 2 is in contact with the cofactor (B).

Combined, the data suggests that the PCP does not undergo a conformational change concerning its backbone when it is loaded with the peptide, but the ppan-arm is changing its orientation towards the protein: In *holo*-TycC3\_PCP, the ppan-arm is in proximity to the loop preceding helix 2, and when loaded with the peptide the cofactor changes its position. In peptidyl(NH)-TycC3\_PCP, the ppan-arm is in loose contact with helix 2 and 3, whereas the peptide is solvent-exposed. This reorientation of the cofactor might influence the reorientation of the whole peptide-loaded PCP domain towards the donor site of its downstream C domain in the native NRPS. However, in nature there is no direct conversion of a *holo*- to a peptide-loaded PCP and the orientation of the aminoacylated cofactor is still unknown. Besides, to validate this hypothesis further, structural information about the interaction of the peptidyl-PCP with the C domain is needed.

### 5.1.3 Structure of *holo*-TycC5-6\_PCP-C

Due to the lack of an active C domain, structural information about the interaction of a peptide-loaded PCP with the donor site of its cognate C domain could not be derived from NMR titration experiments with the separated domains. Instead, it was intended to get this information from the TycC5-6\_PCP-C bidomain. For this bidomain, it has been shown that the C domain can still act on the PCP-bound peptide, and the *apo*-form of the protein has been crystallized<sup>69</sup>. In the structure of the *apo*-protein, the orientation of the domains cannot represent the donor site complex, as the distance between the active sites of the two domains is too large to be bridged by the ppan-arm. Therefore, it was intended to load the bidomain with its cognate nonapeptide, hoping that the cofactor will drive the bidomain into the correct orientation. Potentially, the desired structure of the donor site complex could give insights not only into the protein/protein contacts between the domains, which are important for this conformation, but also into the recognition of the peptide by the C domain. Still, this attempt failed, as the bidomain could not be expressed in the pure *apo*-form, even though the expression was performed under numerous different conditions (Figure 18). The *in vivo* phosphopantetheinylation of TycC5-PCP-C is likely to be catalyzed by the *E. coli* PPT EntD. This PPT is involved in the production of the NRP enterobactin, which is involved in iron-uptake under iron-limited condition<sup>206</sup>. However, even the additional supplementation of the growth media with iron did not reduce the ppan transfer significantly. Thus, it remains unclear why the PPT is so active when the bidomain is expressed, whereas other PCP constructs can be expressed in their *apo*-form under similar conditions.

Without pure *apo*-TycC5-6\_PCP-C at hand it, was not possible to produce homogeneous peptidyl-TycC5-6\_PCP-C for crystallization. Instead, the *apo/holo*-mixture derived from the expression was further phosphopantetheinylated *in vitro*, and the bidomain was successfully crystallized in its *holo*-form (Figure 24). Unfortunately, the affinity of the C domain for the ppan-arm of the *holo*-PCP was not high enough to induce the conformational change towards the donor site orientation, but the bidomain was still in the same orientation as observed for its *apo*-form.

This finding emphasizes the importance of the peptide for the binding of the PCP to the acceptor site of the C domain. This is accompanied by the need for samples of homogeneous *apo*-protein which can be modified using CoA derivatives *in vitro* for subsequent structural studies. Potential ways to get such a homogeneous *apo*-bidomain are its expression in an *entD* knockout strain, its expression in a cell-free *in vitro* system, or the treatment of the *apo/olo*-mixture with a phosphodiesterase. However, initial tests on the cell-free expression of TycC5-6\_PCP-C did not yield soluble protein, and the acyl carrier protein hydrolase from *Pseudomonas aeruginosa*<sup>207, 208</sup> showed no activity against the *olo*-bidomain (data not shown).

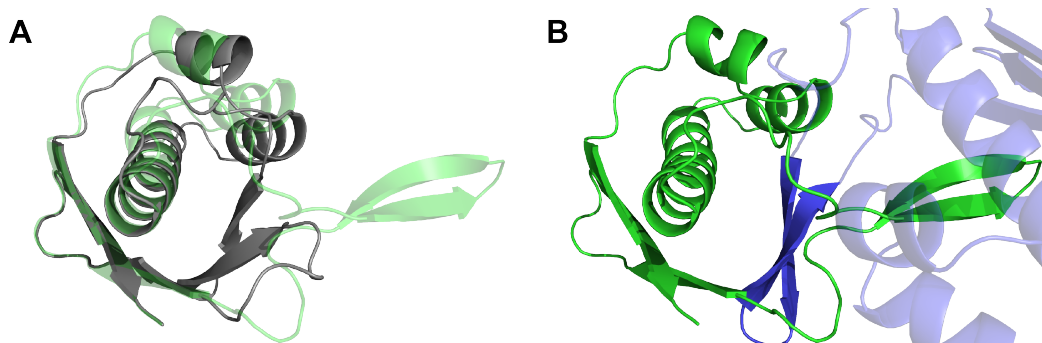
## 5.2 Phosphopantetheinylation of PCPs

The structure of the PCP/Sfp complex derived by X-ray crystallography gives new insights into the structural flexibility of the involved proteins, into the catalysis of the transfer reaction, and into the carrier protein recognition by group II PPTs. Especially with respect to the structural flexibility of the PCP and its recognition by Sfp, further experiments were performed to survey and strengthen the hypotheses made from the crystal structure.

### 5.2.1 Structural Flexibility of Sfp

The most obvious difference between the structures of Sfp in presence (PDB 4MRT) and absence of a PCP (PDB 1QR0) is the orientation of  $\beta 1'/\beta 2'$  in the C-terminal PPT domain. In the structure by Reuter *et al.*<sup>154</sup>, this  $\beta$ -sheet points away from the rest of the protein, whereas it is in close proximity to  $\alpha 2$ ,  $\alpha 3$ , and  $\beta 2$  in the presented structure of the PCP/Sfp complex (Figure 38A).

A closer look at the crystal packing of Sfp in absence of a PCP reveals that always two protein molecules form a dimer<sup>154</sup>. In these dimers,  $\beta 1'/\beta 2'$  of the C-terminal domain of one molecule lies in the same region of the neighboring molecule in which the intramolecular contacts are observed in the structure of the PCP/Sfp complex (Figure 38B).



**Figure 38:** In the C-terminal PPT domain of Sfp, the orientation of  $\beta 1'/\beta 2'$  in presence of a PCP (gray) differs significantly from the one observed in absence of a PCP (semi-transparent green) (A). In the crystal structure of Sfp in absence of a PCP, the position of  $\beta 1'/\beta 2'$  is occupied by a neighboring protein molecule (blue, with  $\beta 1'/\beta 2'$  shown in solid), so always two proteins form a dimer (B).

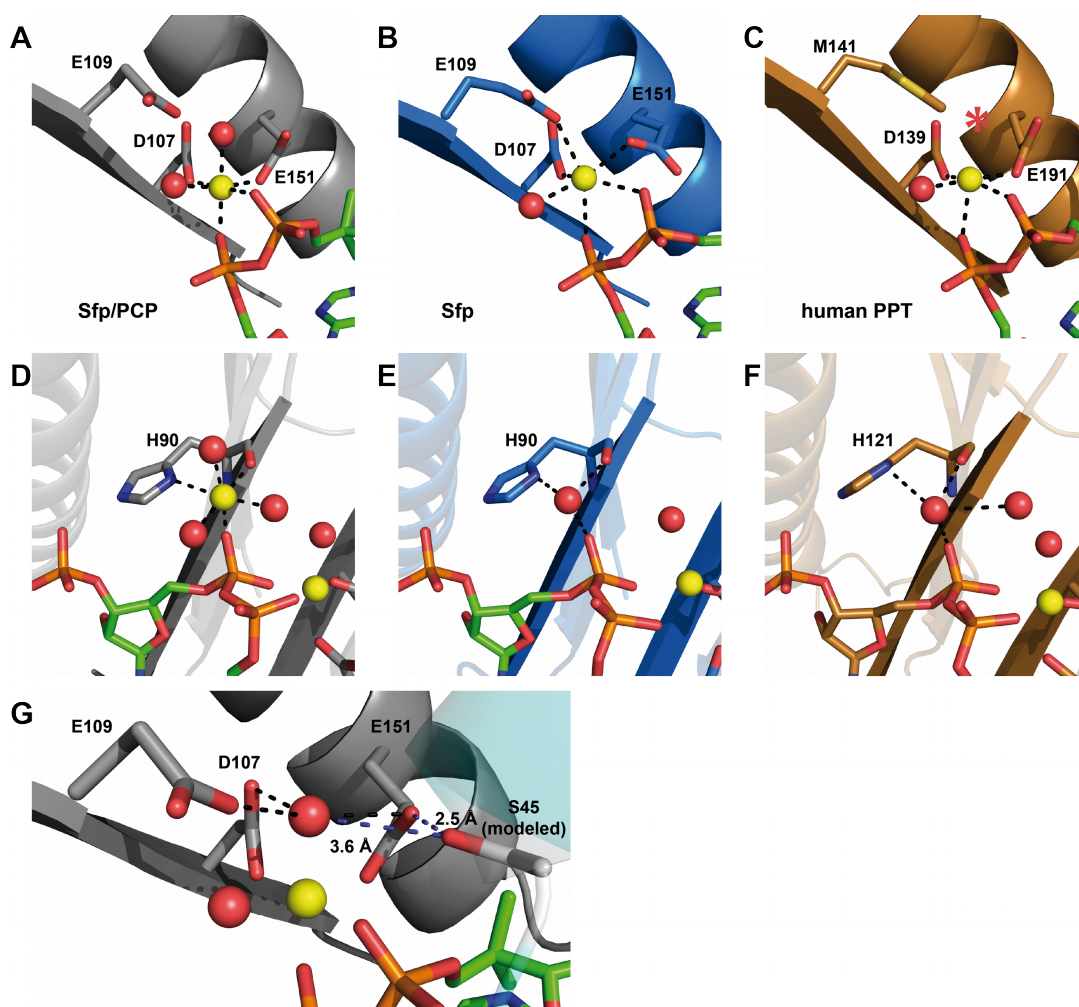
The reorientation of  $\beta 1'/\beta 2'$  upon binding of the PCP might be involved in an activation mechanism, but it is more likely that it is an artifact of the crystallization process, as no evidence for dimeric Sfp in solution was found either in previous studies<sup>35, 154</sup> or in the presented work. Besides, for the human PPT of which the crystal structures of the protein alone (PDB 2BYD), in complex with CoA (PDB 2C43), and in a ternary complex with CoA and an ACP (PDB 2CG5) are known, only the monomeric PPT is observed, and the binding pocket hardly changes its shape in all three forms<sup>205</sup>. Nevertheless, the reorientation of  $\beta 1'/\beta 2'$  indicates a structural flexibility of this region, but the extent of the significance of this flexibility with respect to the activity of Sfp is not known.

## 5.2.2 Mechanism of the Phosphopantetheine Transfer by Sfp

The structure of the PCP/Sfp complex is the first of a carrier protein in complex with a group II PPT, in which both proteins as well as CoA and its ligands are well-resolved.

Comparison of the coordination of the central magnesium ion in the known structures indicates a possible activation switch in the active site. In the structure of Sfp without a carrier protein, the magnesium ion is coordinated by the side chains of D107, E109 and E151 as well as by a water molecule and the 5'  $\alpha$ - and  $\beta$ -phosphate groups of CoA, whereas the side chain of E109 is replaced by a second water

molecule in the PCP/Sfp complex (Figure 39A+B). If the hydroxyl group of S45 in TycC3\_PCP is modeled in the (-) gauche conformation in the structure of the PCP/Sfp complex, it is in close proximity to this second water molecule and the carboxyl group of E151 (Figure 39G), which both could be directly involved in its deprotonation as proposed previously<sup>153, 205, 209</sup>. On the other hand, due to the large distance between them, the previously suggested direct coordination of the serine by the magnesium ion<sup>154</sup> seems unlikely.



**Figure 39:** The coordination of the catalytic magnesium ion (yellow) differs in the structures of the PCP/Sfp complex (gray; A), Sfp (blue; B), and the human PPT (copper; C) where it is coordinated by either one or two water molecules (red; the second water molecule in the structure of the human PPT is shown as an asterisk, as it showed only weak electron density).

In the same structures, a second magnesium ion seems to be present which, lacks parts of its inner coordination shell and was assigned as a water molecule in the structures of Sfp (E) and the human PPT (F), but shows an octahedral coordination in the PCP/Sfp complex (D).

Modeling of the hydroxyl group of S45 of the PCP (cyan) brings it into close proximity to the side chain of E151 of Sfp and a water molecule coordinating the catalytic magnesium ion, which both might deprotonate it (G).

If the replacement of the side chain of E109 by a water molecule that is subsequently involved in the deprotonation of the hydroxyl group of the conserved serine of the PCP is an activation switch, this switch is not conserved among group II PPTs, as the residue is not conserved. For example, a methionine, incapable of coordination to the magnesium ion by its side chain, occupies the corresponding position in the human PPT (Figure 39C). Further insight into the deprotonation of the side chain of the conserved serine could be gained by neutron scattering, as the state of protonation of the water molecule coordinated to the magnesium ion and the side chain of E151 could be resolved with this method.

Beside the known catalytic magnesium ion, a second magnesium ion was found in the structure of the PCP/Sfp complex (Figure 39D). Although it was not reported in other structures so far, electron density was observed at the same position in the structures of Sfp and the human PPT, but it was assigned as water molecules (Figure 39E+F). However, this electron density is coordinated to the carbonyl and the aromatic of H90 and H121, respectively, as well as to the 5'  $\alpha$ -phosphate of CoA which would fit for a positively charged magnesium ion rather than a water molecule. A magnesium ion at this position makes sense not only from an electrostatic point of view but from a mechanistic one, too, as it can compensate the additional negative charge at the 5' phosphate of the 3',5' adenosine bisphosphate after the transfer of the ppan.

As all residues which are involved in the coordination of the two magnesium ions were already subject to previous studies<sup>36, 209</sup>, no further mutational analysis in view of the ppan-transfer reaction was performed. The studies previously made confirmed the invariability of D107 and E151<sup>36, 209</sup>, which was also shown for the corresponding residues in the human PPT<sup>205</sup>. Remarkably, the mutation of the not-conserved E109 to aspartate led to the complete inactivation of Sfp, too<sup>209</sup>. On the other hand, mutation of H90 in Sfp and the corresponding H121 in the human PPT, which are not only coordinating the second magnesium ion, but the 3' phosphate of CoA, too, did not inactivate the transferases<sup>205, 209</sup>. Although the impact of the disruption of the coordination of the 3' phosphate is probably even more severe than the disruption of the coordination of the second magnesium ion, both effects together still do not lead to the inactivation of Sfp,. Therefore, it can be concluded that the second magnesium



ion is not essential for the function of the enzymes. However, to confirm the proposed role of the magnesium ion, the complex of Sfp and 3',5' adenosine bisphosphate could be crystallized to see on the one hand, if the magnesium ion is present in this structure, too, and on the other hand, how it is coordinated to the 5' phosphate group.

In contrast to the structure by Reuter *et al.*, the whole pantetheine moiety of CoA in the structure of the PCP/Sfp complex is in a distinct conformation and shows electron density (Figure 27). Nevertheless, this conformation is likely to be only a snapshot, as previous studies have shown that Sfp accepts CoA derivatives with diverse modifications of the pantetheine moiety like the substitution of the thiol group with an amino group<sup>164, 193</sup> or sulfonyl group<sup>210</sup> as well as the attachment of amino acids<sup>73</sup>, peptides<sup>69, 211</sup>, fatty acids<sup>208</sup>, and other organic molecules like fluorescence labels<sup>161, 163</sup>. The observed promiscuity of Sfp for CoA derivatives suggests that it is not in as close a contact with the pantetheine moiety as the structure might suggest.

### 5.2.3 Carrier Protein Recognition by Sfp

The structure of the PCP/Sfp complex revealed two interaction sites which show direct protein/protein contacts: The intermolecular hydrogen bond between the backbones of both proteins and the hydrophobic interactions between helix 2 of the PCP and the C-terminal PPT domain of Sfp (Figure 30).

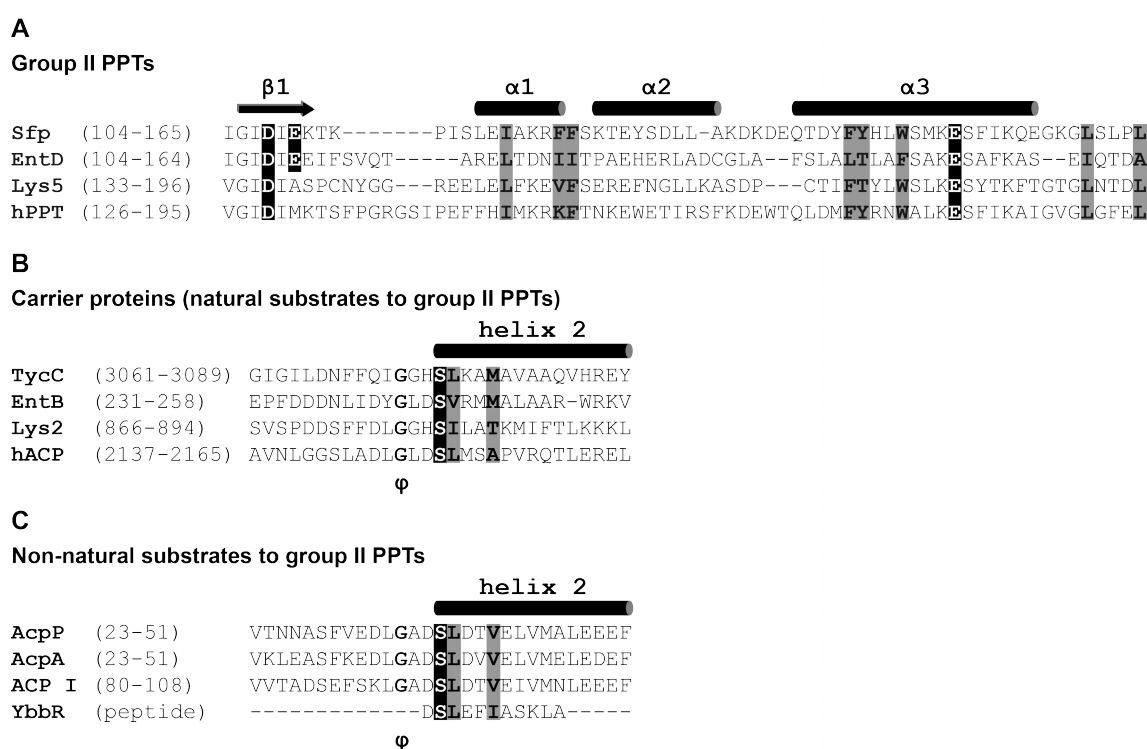
Only two residues of the PCP, L46 and M49, appear to contribute to the hydrophobic interaction. The side chains of both residues are pointing towards a large hydrophobic surface, and especially the side chain of L46 is buried in a cavity. The importance of this interaction was surveyed by mutational analysis of the involved residues of the PCP. To determine the effect of the introduced mutations, the binding affinity of the two proteins and their productive interaction was studied by isothermal calorimetry (Table 9) and a priming assay (Figure 36A). In both assays, the introduced mutations had distinct effects whose magnitude differed depending on the nature of the introduced mutation. All mutants of L46 and M49 showed no apparent binding in the ITC experiments, which indicates a reduction of the affinities to a  $K_D$  of more than 100  $\mu\text{M}$ , whereas the PCP, harboring only the active site S45A

mutation, has a  $K_D$  of less than 1  $\mu\text{M}$ . In the priming assay, it was shown that the reduced affinity is accompanied by a reduction of the velocity of the ppan transfer reaction. Still, the phosphopantetheinylation of TycC3\_PCP(L46A) happens significantly faster than the one of the same protein with the L46N mutation, and when the L46D mutation is introduced, product formation is hardly detectable. This is in agreement with the hydrophobic binding pocket for L46 in the structure of the complex. Alanine has a hydrophobic side chain, but it is considerably smaller, which is why it cannot fill the whole binding pocket, so the binding to Sfp and the velocity of the ppan transfer are reduced. However, this mutation does not cause a repulsive force, which is the case when a polar asparagine or a charged aspartate is introduced at this position. Thus, TycC3\_PCP(L46N) and especially TycC3\_PCP(L46D) show a stronger reduction in the velocity of their phosphopantetheinylation, although these residues have larger side chains to fill the binding pocket. In contrast to L46, M49 is not pointing into a defined binding pocket, so the introduction of an aspartate at this position does not have the same impact on the reaction velocity as the L46D mutation has.

Disruption of the hydrogen bond between the two proteins, either by introduction of the Y36P mutation in Sfp or the G42A mutation in TycC3\_PCP, reduces the affinity of the two proteins towards each other, too, but not as strong as the mutation of L46 or M49, as the affinity can still be determined by ITC at 5°C, at which temperature the  $K_D$  is in both cases  $\sim 50 \mu\text{M}$  (Table 9). Yet, the reduced affinity does not lead to a slower transfer of ppan in the priming assay, but accelerates the velocity of the reaction (Figure 36B). Although the presented data cannot explain this finding, one might speculate that the reduced affinity results in a faster decay of the *holo*-PCP/Sfp complex after the transfer reaction, so the enzyme can undergo the next catalytic cycle faster. On the other hand, a further reduction of the affinity, as observed for the PCPs harboring a mutation in helix 2, results in such a slow formation of the reactive *apo*-PCP/Sfp that the velocity of the overall reaction is reduced. To probe this hypothesis, it would be necessary to apply a method like surface plasmon resonance spectroscopy, so not only the complex stability but the on and off rates of the complex formation could be determined, too. Furthermore, it would be necessary to

consider effects like the affinity of Sfp to the *holo*-PCPs in presence of 3'5' adenosine bisphosphate.

The hydrophobic interactions and the intermolecular hydrogen bond, which are the driving force of the formation of the PCP/Sfp complex, were also observed in the complex of the human ACP/PPT<sup>205</sup>. This finding suggests that the mode of interaction between a group II PPT and its carrier protein substrates might be conserved among a variety of organisms. However, alignments of group II PPTs and their carrier protein substrates reveal that the involved residues are not conserved (Figure 40A+B).



**Figure 40:** Alignments of Sfp with group II PPTs from *E. coli*, *Saccharomyces cerevisiae*, and *Homo sapiens* (A), of TycC3\_PCP with substrate carrier proteins of these PPTs (B), and of non-natural substrates of Sfp (C) reveal that the hydrophobic properties of the residues involved in the complex formation (gray) are conserved. The glycine three residues N-terminal of the conserved serine (indicated by a  $\phi$ ) is conserved, too, whereas the residues forming in the intermolecular hydrogen bond are not. The residues involved in cofactor binding are shown in black, and on top of the alignments, the structural elements of Sfp and TycC3\_PCP(S45A) within the complex are indicated.

Nevertheless, these residues share similar hydrophobic properties, suggesting that the PPTs all have a hydrophobic surface to which the carrier proteins can bind with their second helix, which harbors a large hydrophobic side chain at the position

corresponding to L46 in TycC3\_PCP and a hydrophobic side chain at the position corresponding to M49.

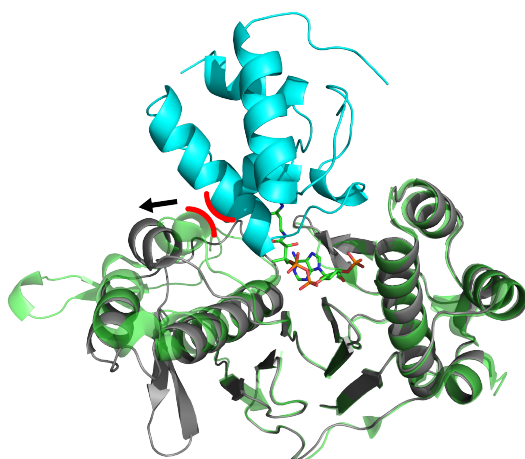
The residues involved in the intermolecular hydrogen bond show no similarities in the different proteins. Yet, as the hydrogen bond is formed between the backbones of the proteins, it is could be possible that the side chains of the involved residues vary, and the conserved glycine that was found to be important for the correct orientation of the two interaction sites in the PCP/Sfp complex indicates that the hydrogen bond is formed in the other complexes, too.

Remarkably, in carrier proteins which are not natural substrates of group II PPTs the residues corresponding to L46 and M49 of TycC3\_PCP have similar properties (Figure 40C). The conservation of these properties is probably the explanation for the promiscuity of Sfp for a large variety of carrier protein substrates. On the other hand, the finding of the small YbbR peptide, which mimics only helix 2 of a carrier protein and can be phosphopantetheinylated by Sfp<sup>212</sup>, emphasizes the finding that the intermolecular hydrogen bond is not essential for the productive interaction.

In the past, efforts have been made to use bacterial group II PPTs as drug targets<sup>213, 214</sup>. Since the presented results indicate that the binding of the carrier protein by a group II PPT is conserved and the same seems to be true for the binding of CoA, these enzymes are not very suitable targets for drugs directly acting on their active site, as such drugs would probably inhibit not only bacterial PPTs but their human homolog, too.

Besides revealing the interactions that are the driving force of the formation of the PCP/Sfp complex, the presented crystal structure shows that the active site mutant TycC3\_PCP(S45A) adopts the A/H state conformation when it is in complex with Sfp (Figure 26), which was confirmed by NMR titration experiments (Figure 34). This result contradicts previous work according to which the PCP adopts the A state conformation when it is in a complex with Sfp<sup>38</sup>. This statement was based on [<sup>1</sup>H;<sup>15</sup>N]-HSQC spectra recorded during a titration experiment of <sup>15</sup>N-labeled TycC3\_PCP(S45A) with unlabeled Sfp. The following docking simulations were performed using the PCP in the A as well as in the A/H state conformation and the crystal structure of Sfp by Reuter *et al.*. In these simulations, the PCP in the A/H state conformation did not fit into the binding pocket of Sfp in a way that brings the active

site serine into close proximity to the 5'  $\beta$ -phosphate of CoA. From this, it was concluded that only the PCP in the A state conformation can bind to Sfp<sup>38</sup>. However, due to the different position of  $\alpha 1$  in the C-terminal PPT domain, the binding pocket of Sfp in the crystal structure of the PCP/Sfp complex is larger than it is in the structure of Sfp in absence of the carrier protein. In this orientation of  $\alpha 1$ , the PCP in the A/H state conformation fits into the binding pocket (Figure 41).



**Figure 41**

An overlay of the structure of the complex of TycC3\_PCP(S45A) (cyan) and Sfp (gray) with the structure of Sfp in absence of a carrier protein (semi-transparent green), which was previously used for modeling of the complex, shows a sterical clash between the PCP in the A/H state and  $\alpha 1$  of the C-terminal PPT domain of Sfp in the structure which was used for modeling (red lines). In the structure of the complex the helix has changed its position (arrow), so the binding pocket is larger and can be occupied by the PCP in the A/H state conformation.

Furthermore, it has also been proposed that free TycC3\_PCP(S45A) exists exclusively in the A state conformation<sup>38</sup>. By resolving the structure of TycC3\_PCP(S45A) in solution, it has been shown that this statement is wrong, too. As the statement was made basing only on the comparison of two [<sup>1</sup>H;<sup>15</sup>N]-HSQC spectra, it seems possible that these spectra were misinterpreted, although it remains unclear what led to this misapprehension. Remarkably, in the other structures of NRPS fragments comprising a PCP with an active site serine to alanine mutation, the A/H state was observed, too<sup>71, 94</sup>.

So far, the A state conformation has only been observed for *apo*-TycC3\_PCP. For this protein, double peaks corresponding to the A state and the A/H state conformation were observed not only in the [<sup>1</sup>H;<sup>15</sup>N]-HSQC but in the <sup>15</sup>N-resolved NOESY spectrum, too. From the latter, the structure of the A state conformation could be determined<sup>38</sup>. However, the only functional interaction partner of a PCP in the *apo*-form is a PPT, which converts the PCP to its active *holo*-form. As no evidence has been found that the A state confirmation plays a role in the interaction of the PCP with Sfp, its biological function needs to be further investigated.

In summary one can say, the presented work proposes a structural model for the interaction of Sfp with a PCP in the A/H state conformation. The interaction of the two proteins is stabilized by hydrophobic interactions between helix 2 of the PCP and the C-terminal domain of Sfp as well as by an intermolecular hydrogen bond between the loop preceding helix 2 and the N-terminal domain of Sfp, although the latter does not seem to be essential. In this arrangement, the side chain of the conserved serine can be deprotonated either by E151 of Sfp or by an activated water molecule coordinated to the magnesium ion in the active site of Sfp. The subsequent nucleophilic attack of the deprotonated side chain might be favored by a second magnesium ion, which is present in the active site, too.

## 6. References

1. Walsh CT. Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* 2004; **303**(5665): 1805-10.
2. Doekel S, Marahiel MA. Biosynthesis of natural products on modular peptide synthetases. *Metabolic engineering* 2001; **3**(1): 64-77.
3. Felnagle EA, Jackson EE, Chan YA, Podevels AM, Berti AD, McMahon MD *et al.* Nonribosomal peptide synthetases involved in the production of medically relevant natural products. *Molecular pharmaceuticals* 2008; **5**(2): 191-211.
4. Ruegger A, Kuhn M, Lichti H, Loosli HR, Huguenin R, Quiquerez C *et al.* Cyclosporin A, a Peptide Metabolite from *Trichoderma polysporum* (Link ex Pers.) Rifai, with a remarkable immunosuppressive activity. *Helvetica chimica acta* 1976; **59**(4): 1075-92.
5. Weber G, Schorgendorfer K, Schneider-Scherzer E, Leitner E. The peptide synthetase catalyzing cyclosporine production in *Tolypocladium niveum* is encoded by a giant 45.8-kilobase open reading frame. *Current genetics* 1994; **26**(2): 120-5.
6. Umezawa H, Maeda K, Takeuchi T, Okami Y. New antibiotics, bleomycin A and B. *The Journal of antibiotics* 1966; **19**(5): 200-9.
7. Du L, Sanchez C, Chen M, Edwards DJ, Shen B. The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces verticillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase. *Chemistry & biology* 2000; **7**(8): 623-42.
8. Johnson BA, Anker H, Meleney FL. Bacitracin: A New Antibiotic Produced by a Member of the B. Subtilis Group. *Science* 1945; **102**(2650): 376-7.
9. Konz D, Klens A, Schorgendorfer K, Marahiel MA. The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multi-modular peptide synthetases. *Chemistry & biology* 1997; **4**(12): 927-37.
10. Fairbrother RW, Williams BL. Two new antibiotics; antibacterial activity of novobiocin and vancomycin. *Lancet* 1956; **271**(6954): 1177-8.

## 6. REFERENCES

---

11. van Wageningen AM, Kirkpatrick PN, Williams DH, Harris BR, Kershaw JK, Lennard NJ *et al.* Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. *Chemistry & biology* 1998; **5**(3): 155-62.
12. Rice LB. The clinical consequences of antimicrobial resistance. *Current opinion in microbiology* 2009; **12**(5): 476-81.
13. McGowan JE, Jr., Tenover FC. Confronting bacterial resistance in healthcare settings: a crucial role for microbiologists. *Nature reviews. Microbiology* 2004; **2**(3): 251-8.
14. Collignon P, Powers JH, Chiller TM, Aidara-Kane A, Aarestrup FM. World Health Organization ranking of antimicrobials according to their importance in human medicine: A critical step for developing risk management strategies for the use of antimicrobials in food production animals. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2009; **49**(1): 132-41.
15. Labella A, Gennari M, Ghidini V, Trento I, Manfrin A, Borrego JJ *et al.* High incidence of antibiotic multi-resistant bacteria in coastal areas dedicated to fish farming. *Marine pollution bulletin* 2013; **70**(1-2): 197-203.
16. Pottinger PS. Methicillin-resistant Staphylococcus aureus infections. *The Medical clinics of North America* 2013; **97**(4): 601-19, x.
17. Willems RJ, Top J, van Santen M, Robinson DA, Coque TM, Baquero F *et al.* Global spread of vancomycin-resistant Enterococcus faecium from distinct nosocomial genetic complex. *Emerging infectious diseases* 2005; **11**(6): 821-8.
18. Kelesidis T, Humphries R, Uslan DZ, Pegues DA. Daptomycin Nonsusceptible Enterococci: An Emerging Challenge for Clinicians. *Clinical Infectious Diseases* 2011; **52**(2): 228-234.
19. Walsh C. Where will new antibiotics come from? *Nature reviews. Microbiology* 2003; **1**(1): 65-70.
20. O'Connell KM, Hodgkinson JT, Sore HF, Welch M, Salmond GP, Spring DR. Combating Multidrug-Resistant Bacteria: Current Strategies for the Discovery of Novel Antibacterials. *Angewandte Chemie* 2013.
21. Gruenewald S, Mootz HD, Stehmeier P, Stachelhaus T. In vivo production of artificial nonribosomal peptide products in the heterologous host Escherichia coli. *Applied and environmental microbiology* 2004; **70**(6): 3282-91.



22. Siewers V, Chen X, Huang L, Zhang J, Nielsen J. Heterologous production of non-ribosomal peptide LLD-ACV in *Saccharomyces cerevisiae*. *Metabolic engineering* 2009; **11**(6): 391-7.
23. Gidijala L, Kiel JA, Douma RD, Seifar RM, van Gulik WM, Bovenberg RA *et al.* An engineered yeast efficiently secreting penicillin. *PloS one* 2009; **4**(12): e8317.
24. Stachelhaus T, Marahiel MA. Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA. *The Journal of biological chemistry* 1995; **270**(11): 6163-9.
25. Hur GH, Vickery CR, Burkart MD. Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. *Natural product reports* 2012; **29**(10): 1074-98.
26. Schneider A, Stachelhaus T, Marahiel MA. Targeted alteration of the substrate specificity of peptide synthetases by rational module swapping. *Molecular & general genetics : MGG* 1998; **257**(3): 308-18.
27. Cane DE, Walsh CT, Khosla C. Harnessing the biosynthetic code: combinations, permutations, and mutations. *Science* 1998; **282**(5386): 63-8.
28. Miao V, Coeffet-Le Gal MF, Nguyen K, Brian P, Penn J, Whiting A *et al.* Genetic engineering in *Streptomyces roseosporus* to produce hybrid lipopeptide antibiotics. *Chemistry & biology* 2006; **13**(3): 269-76.
29. Stachelhaus T, Schneider A, Marahiel MA. Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. *Science* 1995; **269**(5220): 69-72.
30. Eppelmann K, Stachelhaus T, Marahiel MA. Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics. *Biochemistry* 2002; **41**(30): 9718-26.
31. Gilhuus-Moe CC, Kristensen T, Bredesen JE, Zimmer T, Laland SG. The presence and possible role of phosphopantothenic acid in gramicidin S synthetase. *FEBS letters* 1970; **7**(3): 287-290.
32. Stein T, Vater J, Kruff V, Wittmann-Liebold B, Franke P, Panico M *et al.* Detection of 4'-phosphopantetheine at the thioester binding site for L-valine of gramicidinS synthetase 2. *FEBS letters* 1994; **340**(1-2): 39-44.

33. Stachelhaus T, Huser A, Marahiel MA. Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases. *Chemistry & biology* 1996; **3**(11): 913-21.
34. Stein T, Vater J, Kruff V, Otto A, Wittmann-Liebold B, Franke P *et al.* The multiple carrier model of nonribosomal peptide biosynthesis at modular multienzymatic templates. *The Journal of biological chemistry* 1996; **271**(26): 15428-35.
35. Lambalot RH, Gehring AM, Flugel RS, Zuber P, LaCelle M, Marahiel MA *et al.* A new enzyme superfamily - the phosphopantetheinyl transferases. *Chemistry & biology* 1996; **3**(11): 923-36.
36. Quadri LE, Weinreb PH, Lei M, Nakano MM, Zuber P, Walsh CT. Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* 1998; **37**(6): 1585-95.
37. Weber T, Baumgartner R, Renner C, Marahiel MA, Holak TA. Solution structure of PCP, a prototype for the peptidyl carrier domains of modular peptide synthetases. *Structure* 2000; **8**(4): 407-18.
38. Koglin A, Mofid MR, Lohr F, Schafer B, Rogov VV, Blum MM *et al.* Conformational switches modulate protein interactions in peptide antibiotic synthetases. *Science* 2006; **312**(5771): 273-6.
39. Koglin A, Lohr F, Bernhard F, Rogov VV, Frueh DP, Strieter ER *et al.* Structural basis for the selectivity of the external thioesterase of the surfactin synthetase. *Nature* 2008; **454**(7206): 907-11.
40. Gevers W, Kleinkauf H, Lipmann F. The activation of amino acids for biosynthesis of gramicidin S. *Proceedings of the National Academy of Sciences of the United States of America* 1968; **60**(1): 269-76.
41. Gevers W, Kleinkauf H, Lipmann F. Peptidyl transfers in gramicidin S biosynthesis from enzyme-bound thioester intermediates. *Proceedings of the National Academy of Sciences of the United States of America* 1969; **63**(4): 1335-42.
42. Kleinkauf H, Gevers W, Lipmann F. Interrelation between activation and polymerization in gramicidin S biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* 1969; **62**(1): 226-33.

43. Conti E, Stachelhaus T, Marahiel MA, Brick P. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *The EMBO journal* 1997; **16**(14): 4174-83.
44. Stachelhaus T, Mootz HD, Marahiel MA. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chemistry & biology* 1999; **6**(8): 493-505.
45. Challis GL, Ravel J, Townsend CA. Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chemistry & biology* 2000; **7**(3): 211-24.
46. Lautru S, Deeth RJ, Bailey LM, Challis GL. Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. *Nature chemical biology* 2005; **1**(5): 265-9.
47. Lee TV, Johnson LJ, Johnson RD, Koulman A, Lane GA, Lott JS *et al.* Structure of a eukaryotic nonribosomal peptide synthetase adenylation domain that activates a large hydroxamate amino acid in siderophore biosynthesis. *The Journal of biological chemistry* 2010; **285**(4): 2415-27.
48. Chen CY, Georgiev I, Anderson AC, Donald BR. Computational structure-based redesign of enzyme activity. *Proceedings of the National Academy of Sciences of the United States of America* 2009; **106**(10): 3764-9.
49. Guo ZF, Jiang M, Zheng S, Guo Z. Structural change of the enterobactin synthetase in crowded solution and its relation to crowding-enhanced product specificity in nonribosomal enterobactin biosynthesis. *Bioorganic & medicinal chemistry letters* 2010; **20**(13): 3855-8.
50. Conti E, Franks NP, Brick P. Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes. *Structure* 1996; **4**(3): 287-98.
51. Gulick AM, Starai VJ, Horswill AR, Homick KM, Escalante-Semerena JC. The 1.75 Å crystal structure of acetyl-CoA synthetase bound to adenosine-5'-propylphosphate and coenzyme A. *Biochemistry* 2003; **42**(10): 2866-73.
52. Marahiel MA, Stachelhaus T, Mootz HD. Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. *Chemical reviews* 1997; **97**(7): 2651-2674.

## 6. REFERENCES

---

53. Chang KH, Xiang H, Dunaway-Mariano D. Acyl-adenylate motif of the acyl-adenylate/thioester-forming enzyme superfamily: a site-directed mutagenesis study with the *Pseudomonas* sp. strain CBS3 4-chlorobenzoate:coenzyme A ligase. *Biochemistry* 1997; **36**(50): 15650-9.
54. Yonus H, Neumann P, Zimmermann S, May JJ, Marahiel MA, Stubbs MT. Crystal structure of DltA. Implications for the reaction mechanism of non-ribosomal peptide synthetase adenylation domains. *The Journal of biological chemistry* 2008; **283**(47): 32484-91.
55. Gulick AM. Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS chemical biology* 2009; **4**(10): 811-27.
56. Molnar I, Schupp T, Ono M, Zirkle R, Milnamow M, Nowak-Thompson B *et al.* The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So ce90. *Chemistry & biology* 2000; **7**(2): 97-109.
57. Takahashi H, Kumagai T, Kitani K, Mori M, Matoba Y, Sugiyama M. Cloning and characterization of a *Streptomyces* single module type non-ribosomal peptide synthetase catalyzing a blue pigment synthesis. *The Journal of biological chemistry* 2007; **282**(12): 9073-81.
58. Schneider TL, Shen B, Walsh CT. Oxidase domains in epothilone and bleomycin biosynthesis: thiazoline to thiazole oxidation during chain elongation. *Biochemistry* 2003; **42**(32): 9722-30.
59. Baltz RH. Function of MbtH homologs in nonribosomal peptide biosynthesis and applications in secondary metabolite discovery. *Journal of industrial microbiology & biotechnology* 2011; **38**(11): 1747-60.
60. Quadri LE, Sello J, Keating TA, Weinreb PH, Walsh CT. Identification of a *Mycobacterium tuberculosis* gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin. *Chemistry & biology* 1998; **5**(11): 631-45.
61. Felnagle EA, Barkei JJ, Park H, Podevels AM, McMahon MD, Drott DW *et al.* MbtH-like proteins as integral components of bacterial nonribosomal peptide synthetases. *Biochemistry* 2010; **49**(41): 8815-7.

62. Heemstra JR, Jr., Walsh CT, Sattely ES. Enzymatic tailoring of ornithine in the biosynthesis of the *Rhizobium* cyclic trihydroxamate siderophore vicibactin. *Journal of the American Chemical Society* 2009; **131**(42): 15317-29.
63. Imker HJ, Krahn D, Clerc J, Kaiser M, Walsh CT. N-acylation during glidobactin biosynthesis by the tridomain nonribosomal peptide synthetase module GlbF. *Chemistry & biology* 2010; **17**(10): 1077-83.
64. Coeffet-Le Gal MF, Thurston L, Rich P, Miao V, Baltz RH. Complementation of daptomycin dptA and dptD deletion mutations in trans and production of hybrid lipopeptide antibiotics. *Microbiology* 2006; **152**(Pt 10): 2993-3001.
65. Chen H, Hubbard BK, O'Connor SE, Walsh CT. Formation of beta-hydroxy histidine in the biosynthesis of nikkomycin antibiotics. *Chemistry & biology* 2002; **9**(1): 103-12.
66. Stachelhaus T, Mootz HD, Bergendahl V, Marahiel MA. Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain. *The Journal of biological chemistry* 1998; **273**(35): 22773-81.
67. Bergendahl V, Linne U, Marahiel MA. Mutational analysis of the C-domain in nonribosomal peptide synthesis. *European journal of biochemistry / FEBS* 2002; **269**(2): 620-9.
68. Keating TA, Marshall CG, Walsh CT, Keating AE. The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains. *Nature structural biology* 2002; **9**(7): 522-6.
69. Samel SA, Schoenafinger G, Knappe TA, Marahiel MA, Essen LO. Structural and functional insights into a peptide bond-forming bidomain from a nonribosomal peptide synthetase. *Structure* 2007; **15**(7): 781-92.
70. Leslie AG. Refined crystal structure of type III chloramphenicol acetyltransferase at 1.75 Å resolution. *Journal of molecular biology* 1990; **213**(1): 167-86.
71. Tanovic A, Samel SA, Essen LO, Marahiel MA. Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science* 2008; **321**(5889): 659-63.
72. Bloudoff K, Rodionov D, Schmeing TM. Crystal structures of the first condensation domain of CDA synthetase suggest conformational changes during the synthetic cycle of nonribosomal peptide synthetases. *Journal of molecular biology* 2013; **425**(17): 3137-50.

73. Belshaw PJ, Walsh CT, Stachelhaus T. Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science* 1999; **284**(5413): 486-9.
74. Kraas FI, Helmetag V, Wittmann M, Strieker M, Marahiel MA. Functional dissection of surfactin synthetase initiation module reveals insights into the mechanism of lipoinitiation. *Chemistry & biology* 2010; **17**(8): 872-80.
75. Linne U, Marahiel MA. Control of directionality in nonribosomal peptide synthesis: role of the condensation domain in preventing misinitiation and timing of epimerization. *Biochemistry* 2000; **39**(34): 10439-47.
76. Clugston SL, Sieber SA, Marahiel MA, Walsh CT. Chirality of peptide bond-forming condensation domains in nonribosomal peptide synthetases: the C5 domain of tyrocidine synthetase is a (D)C(L) catalyst. *Biochemistry* 2003; **42**(41): 12095-104.
77. Boettger D, Hertweck C. Molecular diversity sculpted by fungal PKS-NRPS hybrids. *Chembiochem : a European journal of chemical biology* 2013; **14**(1): 28-42.
78. Kraas FI, Giessen TW, Marahiel MA. Exploring the mechanism of lipid transfer during biosynthesis of the acidic lipopeptide antibiotic CDA. *FEBS letters* 2012; **586**(3): 283-8.
79. Kopp F, Marahiel MA. Macrocyclization strategies in polyketide and nonribosomal peptide biosynthesis. *Natural product reports* 2007; **24**(4): 735-49.
80. Gao X, Haynes SW, Ames BD, Wang P, Vien LP, Walsh CT *et al.* Cyclization of fungal nonribosomal peptides by a terminal condensation-like domain. *Nature chemical biology* 2012; **8**(10): 823-30.
81. Duerfahrt T, Eppelmann K, Muller R, Marahiel MA. Rational design of a bimodular model system for the investigation of heterocyclization in nonribosomal peptide biosynthesis. *Chemistry & biology* 2004; **11**(2): 261-71.
82. Balibar CJ, Vaillancourt FH, Walsh CT. Generation of D amino acid residues in assembly of arthrofactin by dual condensation/epimerization domains. *Chemistry & biology* 2005; **12**(11): 1189-200.
83. Jin M, Fischbach MA, Clardy J. A biosynthetic gene cluster for the acetyl-CoA carboxylase inhibitor andrimid. *Journal of the American Chemical Society* 2006; **128**(33): 10660-1.

84. Fortin PD, Walsh CT, Magarvey NA. A transglutaminase homologue as a condensation catalyst in antibiotic assembly lines. *Nature* 2007; **448**(7155): 824-7.
85. Eys S, Schwartz D, Wohlleben W, Schinko E. Three thioesterases are involved in the biosynthesis of phosphinothricin tripeptide in *Streptomyces viridochromogenes* Tu494. *Antimicrobial agents and chemotherapy* 2008; **52**(5): 1686-96.
86. Du L, Lou L. PKS and NRPS release mechanisms. *Natural product reports* 2010; **27**(2): 255-78.
87. Roskoski R, Jr., Kleinkauf H, Gevers W, Lipmann F. Isolation of enzyme-bound peptide intermediates in tyrocidine biosynthesis. *Biochemistry* 1970; **9**(25): 4846-51.
88. Debono M, Barnhart M, Carrell CB, Hoffmann JA, Occolowitz JL, Abbott BJ *et al.* A21978C, a complex of new acidic peptide antibiotics: isolation, chemistry, and mass spectral structure elucidation. *The Journal of antibiotics* 1987; **40**(6): 761-77.
89. Kakinuma A, Ouchida A, Shima T, Sugino H, Isono M, Tamura G *et al.* Confirmation of the Structure of Surfactin by Mass Spectrometry. *Agricultural and Biological Chemistry* 1969; **33**(11): 1669-1671.
90. Perez Baz J, Canedo LM, Fernandez Puentes JL, Silva Elipe MV. Thiocoraline, a novel depsipeptide with antitumor activity produced by a marine *Micromonospora*. II. Physico-chemical properties and structure determination. *The Journal of antibiotics* 1997; **50**(9): 738-41.
91. Bruner SD, Weber T, Kohli RM, Schwarzer D, Marahiel MA, Walsh CT *et al.* Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE. *Structure* 2002; **10**(3): 301-10.
92. Samel SA, Wagner B, Marahiel MA, Essen LO. The thioesterase domain of the fengycin biosynthesis cluster: a structural base for the macrocyclization of a non-ribosomal lipopeptide. *Journal of molecular biology* 2006; **359**(4): 876-89.
93. Liu Y, Zheng T, Bruner SD. Structural basis for phosphopantetheinyl carrier domain interactions in the terminal module of nonribosomal peptide synthetases. *Chemistry & biology* 2011; **18**(11): 1482-8.
94. Frueh DP, Arthanari H, Koglin A, Vosburg DA, Bennett AE, Walsh CT *et al.* Dynamic thiolation-thioesterase structure of a non-ribosomal peptide synthetase. *Nature* 2008; **454**(7206): 903-6.

95. Trauger JW, Kohli RM, Mootz HD, Marahiel MA, Walsh CT. Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature* 2000; **407**(6801): 215-8.
96. Trauger JW, Kohli RM, Walsh CT. Cyclization of backbone-substituted peptides catalyzed by the thioesterase domain from the tyrocidine nonribosomal peptide synthetase. *Biochemistry* 2001; **40**(24): 7092-8.
97. Kohli RM, Walsh CT, Burkart MD. Biomimetic synthesis and optimization of cyclic peptide antibiotics. *Nature* 2002; **418**(6898): 658-61.
98. Kohli RM, Takagi J, Walsh CT. The thioesterase domain from a nonribosomal peptide synthetase as a cyclization catalyst for integrin binding peptides. *Proceedings of the National Academy of Sciences of the United States of America* 2002; **99**(3): 1247-52.
99. Jackowski S, Rock CO. Consequences of reduced intracellular coenzyme A content in *Escherichia coli*. *Journal of bacteriology* 1986; **166**(3): 866-71.
100. Schneider A, Marahiel MA. Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in non-ribosomal peptide biosynthesis in *Bacillus subtilis*. *Archives of microbiology* 1998; **169**(5): 404-10.
101. Schwarzer D, Mootz HD, Linne U, Marahiel MA. Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. *Proceedings of the National Academy of Sciences of the United States of America* 2002; **99**(22): 14083-8.
102. Yeh E, Kohli RM, Bruner SD, Walsh CT. Type II thioesterase restores activity of a NRPS module stalled with an aminoacyl-S-enzyme that cannot be elongated. *Chembiochem : a European journal of chemical biology* 2004; **5**(9): 1290-3.
103. Linne U, Schwarzer D, Schroeder GN, Marahiel MA. Mutational analysis of a type II thioesterase associated with nonribosomal peptide synthesis. *European journal of biochemistry / FEBS* 2004; **271**(8): 1536-45.
104. Hoffmann K, Schneider-Scherzer E, Kleinkauf H, Zocher R. Purification and characterization of eucaryotic alanine racemase acting as key enzyme in cyclosporin biosynthesis. *The Journal of biological chemistry* 1994; **269**(17): 12710-4.



105. Stein DB, Linne U, Hahn M, Marahiel MA. Impact of epimerization domains on the intermodular transfer of enzyme-bound intermediates in nonribosomal peptide synthesis. *Chembiochem : a European journal of chemical biology* 2006; **7**(11): 1807-14.
106. Linne U, Doekel S, Marahiel MA. Portability of epimerization domain and role of peptidyl carrier protein on epimerization activity in nonribosomal peptide synthetases. *Biochemistry* 2001; **40**(51): 15824-34.
107. Silakowski B, Schairer HU, Ehret H, Kunze B, Weinig S, Nordsiek G *et al.* New lessons for combinatorial biosynthesis from myxobacteria. The myxothiazol biosynthetic gene cluster of *Stigmatella aurantiaca* DW4/3-1. *The Journal of biological chemistry* 1999; **274**(52): 37391-9.
108. Du L, Chen M, Sanchez C, Shen B. An oxidation domain in the BlmIII non-ribosomal peptide synthetase probably catalyzing thiazole formation in the biosynthesis of the anti-tumor drug bleomycin in *Streptomyces verticillus* ATCC15003. *FEMS microbiology letters* 2000; **189**(2): 171-5.
109. Patel HM, Walsh CT. In vitro reconstitution of the *Pseudomonas aeruginosa* nonribosomal peptide synthesis of pyochelin: characterization of backbone tailoring thiazoline reductase and N-methyltransferase activities. *Biochemistry* 2001; **40**(30): 9023-31.
110. Reimann C, Patel HM, Serino L, Barone M, Walsh CT, Haas D. Essential PchG-dependent reduction in pyochelin biosynthesis of *Pseudomonas aeruginosa*. *Journal of bacteriology* 2001; **183**(3): 813-20.
111. Gaitatzis N, Kunze B, Muller R. In vitro reconstitution of the myxochelin biosynthetic machinery of *Stigmatella aurantiaca* Sg a15: Biochemical characterization of a reductive release mechanism from nonribosomal peptide synthetases. *Proceedings of the National Academy of Sciences of the United States of America* 2001; **98**(20): 11136-41.
112. Li Y, Weissman KJ, Muller R. Myxochelin biosynthesis: direct evidence for two- and four-electron reduction of a carrier protein-bound thioester. *Journal of the American Chemical Society* 2008; **130**(24): 7554-5.

## 6. REFERENCES

---

113. Zerbe K, Pylypenko O, Vitali F, Zhang W, Rousset S, Heck M *et al.* Crystal structure of OxyB, a cytochrome P450 implicated in an oxidative phenol coupling reaction during vancomycin biosynthesis. *The Journal of biological chemistry* 2002; **277**(49): 47476-85.
114. Bischoff D, Bister B, Bertazzo M, Pfeifer V, Stegmann E, Nicholson GJ *et al.* The biosynthesis of vancomycin-type glycopeptide antibiotics--a model for oxidative side-chain cross-linking by oxygenases coupled to the action of peptide synthetases. *ChemBiochem : a European journal of chemical biology* 2005; **6**(2): 267-72.
115. Reimmann C, Serino L, Beyeler M, Haas D. Dihydroaeruginosic acid synthetase and pyochelin synthetase, products of the pchEF genes, are induced by extracellular pyochelin in *Pseudomonas aeruginosa*. *Microbiology* 1998; **144 ( Pt 11)**: 3135-48.
116. Schauwecker F, Pfennig F, Grammel N, Keller U. Construction and in vitro analysis of a new bi-modular polypeptide synthetase for synthesis of N-methylated acyl peptides. *Chemistry & biology* 2000; **7**(4): 287-97.
117. O'Brien DP, Kirkpatrick PN, O'Brien SW, Staroske T, Richardson TI, Evans DA *et al.* Expression and assay of an -methyltransferase involved in the biosynthesis of a vancomycin group antibiotic. *Chemical Communications* 2000; (1): 103-104.
118. Hacker C, Glinski M, Hornbogen T, Doller A, Zocher R. Mutational analysis of the N-methyltransferase domain of the multifunctional enzyme enniatin synthetase. *The Journal of biological chemistry* 2000; **275**(40): 30826-32.
119. Weckwerth W, Miyamoto K, Iinuma K, Krause M, Glinski M, Storm T *et al.* Biosynthesis of PF1022A and related cyclooctadepsipeptides. *The Journal of biological chemistry* 2000; **275**(23): 17909-15.
120. Süssmuth RD, Pelzer S, Nicholson G, Walk T, Wohlleben W, Jung G. New Advances in the Biosynthesis of Glycopeptide Antibiotics of the Vancomycin Type from *Amycolatopsis mediterranei*. *Angewandte Chemie International Edition* 1999; **38**(13-14): 1976-1979.
121. Pfeifer BA, Wang CC, Walsh CT, Khosla C. Biosynthesis of Yersiniabactin, a complex polyketide-nonribosomal peptide, using *Escherichia coli* as a heterologous host. *Applied and environmental microbiology* 2003; **69**(11): 6698-702.

122. Mahlert C, Kopp F, Thirlway J, Micklefield J, Marahiel MA. Stereospecific enzymatic transformation of alpha-ketoglutarate to (2S,3R)-3-methyl glutamate during acidic lipopeptide biosynthesis. *Journal of the American Chemical Society* 2007; **129**(39): 12011-8.
123. Nelson JT, Lee J, Sims JW, Schmidt EW. Characterization of SafC, a catechol 4-O-methyltransferase involved in saframycin biosynthesis. *Applied and environmental microbiology* 2007; **73**(11): 3575-80.
124. Chen H, Walsh CT. Coumarin formation in novobiocin biosynthesis: beta-hydroxylation of the aminoacyl enzyme tyrosyl-S-NovH by a cytochrome P450 NovI. *Chemistry & biology* 2001; **8**(4): 301-12.
125. Bosello M, Mielcarek A, Giessen TW, Marahiel MA. An enzymatic pathway for the biosynthesis of the formylhydroxyornithine required for rhodochelin iron coordination. *Biochemistry* 2012; **51**(14): 3059-66.
126. Schoenafinger G, Schracke N, Linne U, Marahiel MA. Formylation domain: an essential modifying enzyme for the nonribosomal biosynthesis of linear gramicidin. *Journal of the American Chemical Society* 2006; **128**(23): 7406-7.
127. Lazos O, Tosin M, Slusarczyk AL, Boakes S, Cortes J, Sidebottom PJ *et al.* Biosynthesis of the putative siderophore erythrochelin requires unprecedented crosstalk between separate nonribosomal peptide gene clusters. *Chemistry & biology* 2010; **17**(2): 160-73.
128. Ames BD, Walsh CT. Anthranilate-activating modules from fungal nonribosomal peptide assembly lines. *Biochemistry* 2010; **49**(15): 3351-65.
129. Solenberg PJ, Matsushima P, Stack DR, Wilkie SC, Thompson RC, Baltz RH. Production of hybrid glycopeptide antibiotics in vitro and in *Streptomyces toyocaensis*. *Chemistry & biology* 1997; **4**(3): 195-202.
130. Losey HC, Peczu MW, Chen Z, Eggert US, Dong SD, Pelczer I *et al.* Tandem action of glycosyltransferases in the maturation of vancomycin and teicoplanin aglycones: novel glycopeptides. *Biochemistry* 2001; **40**(15): 4745-55.
131. Chang Z, Flatt P, Gerwick WH, Nguyen VA, Willis CL, Sherman DH. The barbamide biosynthetic gene cluster: a novel marine cyanobacterial system of mixed polyketide synthase (PKS)-non-ribosomal peptide synthetase (NRPS) origin involving an unusual trichloroleucyl starter unit. *Gene* 2002; **296**(1-2): 235-47.

132. Vaillancourt FH, Yin J, Walsh CT. SyrB2 in syringomycin E biosynthesis is a nonheme FeII alpha-ketoglutarate- and O<sub>2</sub>-dependent halogenase. *Proceedings of the National Academy of Sciences of the United States of America* 2005; **102**(29): 10111-6.
133. Puk O, Huber P, Bischoff D, Recktenwald J, Jung G, Sussmuth RD *et al.* Glycopeptide biosynthesis in *Amycolatopsis mediterranei* DSM5908: function of a halogenase and a haloperoxidase/perhydrolase. *Chemistry & biology* 2002; **9**(2): 225-35.
134. Xu GY, Tam A, Lin L, Hixon J, Fritz CC, Powers R. Solution structure of *B. subtilis* acyl carrier protein. *Structure* 2001; **9**(4): 277-87.
135. Ploskon E, Arthur CJ, Evans SE, Williams C, Crosby J, Simpson TJ *et al.* A mammalian type I fatty acid synthase acyl carrier protein domain does not sequester acyl chains. *The Journal of biological chemistry* 2008; **283**(1): 518-28.
136. Johansson P, Mulinacci B, Koestler C, Vollrath R, Oesterhelt D, Gringer M. Multimeric options for the auto-activation of the *Saccharomyces cerevisiae* FAS type I megasynthase. *Structure* 2009; **17**(8): 1063-74.
137. Roujeinikova A, Baldock C, Simon WJ, Gilroy J, Baker PJ, Stuitje AR *et al.* X-ray crystallographic studies on butyryl-ACP reveal flexibility of the structure around a putative acyl chain binding site. *Structure* 2002; **10**(6): 825-35.
138. Roujeinikova A, Simon WJ, Gilroy J, Rice DW, Rafferty JB, Slabas AR. Structural studies of fatty acyl-(acyl carrier protein) thioesters reveal a hydrophobic binding cavity that can expand to fit longer substrates. *Journal of molecular biology* 2007; **365**(1): 135-45.
139. Ploskon E, Arthur CJ, Kanari AL, Wattana-amorn P, Williams C, Crosby J *et al.* Recognition of intermediate functionality by acyl carrier protein over a complete cycle of fatty acid biosynthesis. *Chemistry & biology* 2010; **17**(7): 776-85.
140. Edwards DJ, Marquez BL, Nogle LM, McPhail K, Goeger DE, Roberts MA *et al.* Structure and biosynthesis of the jamaicamides, new mixed polyketide-peptide neurotoxins from the marine cyanobacterium *Lyngbya majuscula*. *Chemistry & biology* 2004; **11**(6): 817-33.

141. Chang Z, Sitachitta N, Rossi JV, Roberts MA, Flatt PM, Jia J *et al.* Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*. *Journal of natural products* 2004; **67**(8): 1356-67.
142. Rahman AS, Hothersall J, Crosby J, Simpson TJ, Thomas CM. Tandemly duplicated acyl carrier proteins, which increase polyketide antibiotic production, can apparently function either in parallel or in series. *The Journal of biological chemistry* 2005; **280**(8): 6399-408.
143. Evans SE, Williams C, Arthur CJ, Burston SG, Simpson TJ, Crosby J *et al.* An ACP structural switch: conformational differences between the apo and holo forms of the actinorhodin polyketide synthase acyl carrier protein. *ChemBiochem : a European journal of chemical biology* 2008; **9**(15): 2424-32.
144. Sharma AK, Sharma SK, Surolia A, Surolia N, Sarma SP. Solution structures of conformationally equilibrium forms of holo-acyl carrier protein (PfACP) from *Plasmodium falciparum* provides insight into the mechanism of activation of ACPs. *Biochemistry* 2006; **45**(22): 6904-16.
145. Evans SE, Williams C, Arthur CJ, Ploskon E, Wattana-amorn P, Cox RJ *et al.* Probing the Interactions of early polyketide intermediates with the Actinorhodin ACP from *S. coelicolor* A3(2). *Journal of molecular biology* 2009; **389**(3): 511-28.
146. Haushalter RW, Filipp FV, Ko KS, Yu R, Opella SJ, Burkart MD. Binding and "pKa" modulation of a polycyclic substrate analogue in a type II polyketide acyl carrier protein. *ACS chemical biology* 2011; **6**(5): 413-8.
147. Volkman BF, Zhang Q, Debabov DV, Rivera E, Kresheck GC, Neuhaus FC. Biosynthesis of D-alanyl-lipoteichoic acid: the tertiary structure of apo-D-alanyl carrier protein. *Biochemistry* 2001; **40**(27): 7964-72.
148. Ehmman DE, Gehring AM, Walsh CT. Lysine biosynthesis in *Saccharomyces cerevisiae*: mechanism of alpha-amino adipate reductase (Lys2) involves posttranslational phosphopantetheinylation by Lys5. *Biochemistry* 1999; **38**(19): 6171-7.
149. Wang L, Jil C, Xu Y, Xu J, Dai J, Wu Q *et al.* Cloning and characterization of a novel human homolog\* of mouse U26, a putative PQQ-dependent AAS dehydrogenase. *Molecular biology reports* 2005; **32**(1): 47-53.

## 6. REFERENCES

---

150. Donato H, Krupenko NI, Tsybovsky Y, Krupenko SA. 10-formyltetrahydrofolate dehydrogenase requires a 4'-phosphopantetheine prosthetic group for catalysis. *The Journal of biological chemistry* 2007; **282**(47): 34159-66.
151. Elovson J, Vagelos PR. Acyl carrier protein. X. Acyl carrier protein synthetase. *The Journal of biological chemistry* 1968; **243**(13): 3603-11.
152. Lambalot RH, Walsh CT. Cloning, overproduction, and characterization of the *Escherichia coli* holo-acyl carrier protein synthase. *The Journal of biological chemistry* 1995; **270**(42): 24658-61.
153. Parris KD, Lin L, Tam A, Mathew R, Hixon J, Stahl M *et al.* Crystal structures of substrate binding to *Bacillus subtilis* holo-(acyl carrier protein) synthase reveal a novel trimeric arrangement of molecules resulting in three active sites. *Structure* 2000; **8**(8): 883-95.
154. Reuter K, Mofid MR, Marahiel MA, Ficner R. Crystal structure of the surfactin synthetase-activating enzyme sfp: a prototype of the 4'-phosphopantetheinyl transferase superfamily. *The EMBO journal* 1999; **18**(23): 6823-31.
155. Jenni S, Leibundgut M, Maier T, Ban N. Architecture of a fungal fatty acid synthase at 5 Å resolution. *Science* 2006; **311**(5765): 1263-7.
156. Lomakin IB, Xiong Y, Steitz TA. The crystal structure of yeast fatty acid synthase, a cellular machine with eight active sites working together. *Cell* 2007; **129**(2): 319-32.
157. Fichtlscherer F, Wellein C, Mittag M, Schweizer E. A novel function of yeast fatty acid synthase. Subunit alpha is capable of self-pantetheinylation. *European journal of biochemistry / FEBS* 2000; **267**(9): 2666-71.
158. Finking R, Solsbacher J, Konz D, Schobert M, Schafer A, Jahn D *et al.* Characterization of a new type of phosphopantetheinyl transferase for fatty acid and siderophore synthesis in *Pseudomonas aeruginosa*. *The Journal of biological chemistry* 2002; **277**(52): 50293-302.
159. Nakano MM, Marahiel MA, Zuber P. Identification of a genetic locus required for biosynthesis of the lipopeptide antibiotic surfactin in *Bacillus subtilis*. *Journal of bacteriology* 1988; **170**(12): 5662-8.

160. Praphanphoj V, Sacksteder KA, Gould SJ, Thomas GH, Geraghty MT. Identification of the alpha-aminoadipic semialdehyde dehydrogenase-phosphopantetheinyl transferase gene, the human ortholog of the yeast LYS5 gene. *Molecular genetics and metabolism* 2001; **72**(4): 336-42.
161. Mercer AC, La Clair JJ, Burkart MD. Fluorescent multiplex analysis of carrier protein post-translational modification. *Chembiochem : a European journal of chemical biology* 2005; **6**(8): 1335-7.
162. Roujeinikova A, Baldock C, Simon WJ, Gilroy J, Baker PJ, Stuitje AR *et al.* Crystallization and preliminary X-ray crystallographic studies on acyl-(acyl carrier protein) from *Escherichia coli*. *Acta crystallographica. Section D, Biological crystallography* 2002; **58**(Pt 2): 330-2.
163. La Clair JJ, Foley TL, Schegg TR, Regan CM, Burkart MD. Manipulation of carrier proteins in antibiotic biosynthesis. *Chemistry & biology* 2004; **11**(2): 195-201.
164. Liu Y, Bruner SD. Rational manipulation of carrier-domain geometry in nonribosomal peptide synthetases. *Chembiochem : a European journal of chemical biology* 2007; **8**(6): 617-21.
165. Koglin A, Walsh CT. Structural insights into nonribosomal peptide enzymatic assembly lines. *Natural product reports* 2009; **26**(8): 987-1000.
166. Marahiel MA. Working outside the protein-synthesis rules: insights into non-ribosomal peptide synthesis. *Journal of peptide science : an official publication of the European Peptide Society* 2009; **15**(12): 799-807.
167. Strieker M, Tanovic A, Marahiel MA. Nonribosomal peptide synthetases: structures and dynamics. *Current opinion in structural biology* 2010; **20**(2): 234-40.
168. Condurso HL, Bruner SD. Structure guided approaches toward exploiting and manipulating nonribosomal peptide and polyketide biosynthetic pathways. *Current opinion in chemical biology* 2012; **16**(1-2): 162-9.
169. Finking R, Mofid MR, Marahiel MA. Mutational analysis of peptidyl carrier protein and acyl carrier protein synthase unveils residues involved in protein-protein recognition. *Biochemistry* 2004; **43**(28): 8946-56.
170. Lai JR, Fischbach MA, Liu DR, Walsh CT. Localized protein interaction surfaces on the EntB carrier protein revealed by combinatorial mutagenesis and selection. *Journal of the American Chemical Society* 2006; **128**(34): 11002-3.

171. Zhou Z, Lai JR, Walsh CT. Interdomain communication between the thiolation and thioesterase domains of EntF explored by combinatorial mutagenesis and selection. *Chemistry & biology* 2006; **13**(8): 869-79.
172. Mitchell CA, Shi C, Aldrich CC, Gulick AM. Structure of PA1221, a nonribosomal peptide synthetase containing adenylation and peptidyl carrier protein domains. *Biochemistry* 2012; **51**(15): 3252-63.
173. Zhou Z, Lai JR, Walsh CT. Directed evolution of aryl carrier proteins in the enterobactin synthetase. *Proceedings of the National Academy of Sciences of the United States of America* 2007; **104**(28): 11621-6.
174. Lai JR, Fischbach MA, Liu DR, Walsh CT. A protein interaction surface in nonribosomal peptide synthesis mapped by combinatorial mutagenesis and selection. *Proceedings of the National Academy of Sciences of the United States of America* 2006; **103**(14): 5314-9.
175. Hahn M, Stachelhaus T. Selective interaction between nonribosomal peptide synthetases is facilitated by short communication-mediating domains. *Proceedings of the National Academy of Sciences of the United States of America* 2004; **101**(44): 15585-90.
176. Hur GH, Meier JL, Baskin J, Codelli JA, Bertozzi CR, Marahiel MA *et al.* Crosslinking studies of protein-protein interactions in nonribosomal peptide biosynthesis. *Chemistry & biology* 2009; **16**(4): 372-81.
177. Nazi I, Koteva KP, Wright GD. One-pot chemoenzymatic preparation of coenzyme A analogues. *Analytical biochemistry* 2004; **324**(1): 100-5.
178. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* 2011; **7**: 539.
179. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* 2004; **60**(Pt 12 Pt 1): 2126-32.
180. Guntert P. Automated structure determination from NMR spectra. *European biophysics journal : EBJ* 2009; **38**(2): 129-43.



181. Herrmann T, Guntert P, Wuthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *Journal of molecular biology* 2002; **319**(1): 209-27.
182. Evans P. Scaling and assessment of data quality. *Acta crystallographica. Section D, Biological crystallography* 2006; **62**(Pt 1): 72-82.
183. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR *et al.* Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* 2011; **67**(Pt 4): 235-42.
184. Leslie AW, Powell H. Processing diffraction data with mosflm. In: Read R, Sussman J (eds). *Evolving Methods for Macromolecular Crystallography*, vol. 245. Springer Netherlands, 2007, pp 41-51.
185. Kibbe WA. OligoCalc: an online oligonucleotide properties calculator. *Nucleic acids research* 2007; **35**(Web Server issue): W43-6.
186. Koradi R, Billeter M, Güntert P. Point-centered domain decomposition for parallel molecular dynamics simulation. *Computer Physics Communications* 2000; **124**(2-3): 139-147.
187. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta crystallographica. Section D, Biological crystallography* 2011; **67**(Pt 4): 355-67.
188. Shen Y, Delaglio F, Cornilescu G, Bax A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of biomolecular NMR* 2009; **44**(4): 213-23.
189. Mullis KB, Faloona FA. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in enzymology* 1987; **155**: 335-50.
190. Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT *et al.* Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 1988; **239**(4839): 487-91.
191. Schwarz D, Junge F, Durst F, Frolich N, Schneider B, Reckel S *et al.* Preparative scale expression of membrane proteins in Escherichia coli-based continuous exchange cell-free systems. *Nature protocols* 2007; **2**(11): 2945-57.

192. Schagger H, von Jagow G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical biochemistry* 1987; **166**(2): 368-79.
193. Worthington AS, Burkart MD. One-pot chemo-enzymatic synthesis of reporter-modified proteins. *Organic & biomolecular chemistry* 2006; **4**(1): 44-6.
194. Ogura K, Terasawa H, Inagaki F. An improved double-tuned and isotope-filtered pulse scheme based on a pulsed field gradient and a wide-band inversion shaped pulse. *Journal of biomolecular NMR* 1996; **8**(4): 492-8.
195. Grzesiek S, Bax A. An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *Journal of Magnetic Resonance (1969)* 1992; **99**(1): 201-207.
196. Logan TM, Olejniczak ET, Xu RX, Fesik SW. Side chain and backbone assignments in isotopically labeled proteins from two heteronuclear triple resonance experiments. *FEBS letters* 1992; **314**(3): 413-8.
197. Yamazaki T, Forman-Kay JD, Kay LE. Two-dimensional NMR experiments for correlating carbon-13.beta. and proton.delta./epsilon. chemical shifts of aromatic residues in <sup>13</sup>C-labeled proteins via scalar couplings. *Journal of the American Chemical Society* 1993; **115**(23): 11054-11055.
198. Lohr F, Hansel R, Rogov VV, Dotsch V. Improved pulse sequences for sequence specific assignment of aromatic proton resonances in proteins. *Journal of biomolecular NMR* 2007; **37**(3): 205-24.
199. O'Brien IG, Cox GB, Gibson F. Biologically active compounds containing 2,3-dihydroxybenzoic acid and serine formed by *Escherichia coli*. *Biochimica et biophysica acta* 1970; **201**(3): 453-60.
200. Drake EJ, Nicolai DA, Gulick AM. Structure of the EntB multidomain nonribosomal peptide synthetase and functional analysis of its interaction with the EntE adenylation domain. *Chemistry & biology* 2006; **13**(4): 409-19.
201. Woodrow GC, Young IG, Gibson F. Biosynthesis of enterochelin in *Escherichia coli* K-12: separation of the polypeptides coded for by the entD, E, F and G genes. *Biochimica et biophysica acta* 1979; **582**(1): 145-53.

202. Mootz HD, Marahiel MA. The tyrocidine biosynthesis operon of *Bacillus brevis*: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *Journal of bacteriology* 1997; **179**(21): 6843-50.
203. Mofid MR, Marahiel MA, Ficner R, Reuter K. Crystallization and preliminary crystallographic studies of Sfp: a phosphopantetheinyl transferase of modular peptide synthetases. *Acta crystallographica. Section D, Biological crystallography* 1999; **55**(Pt 5): 1098-100.
204. Busche A, Gottstein D, Hein C, Ripin N, Pader I, Tufar P *et al.* Characterization of molecular interactions between ACP and halogenase domains in the Curacin A polyketide synthase. *ACS chemical biology* 2012; **7**(2): 378-86.
205. Bunkoczi G, Pasta S, Joshi A, Wu X, Kavanagh KL, Smith S *et al.* Mechanism and substrate recognition of human holo ACP synthase. *Chemistry & biology* 2007; **14**(11): 1243-53.
206. Raymond KN, Dertz EA, Kim SS. Enterobactin: an archetype for microbial iron transport. *Proceedings of the National Academy of Sciences of the United States of America* 2003; **100**(7): 3584-8.
207. Murugan E, Kong R, Sun H, Rao F, Liang ZX. Expression, purification and characterization of the acyl carrier protein phosphodiesterase from *Pseudomonas Aeruginosa*. *Protein expression and purification* 2010; **71**(2): 132-8.
208. Kosa NM, Haushalter RW, Smith AR, Burkart MD. Reversible labeling of native and fusion-protein motifs. *Nature methods* 2012; **9**(10): 981-4.
209. Mofid MR, Finking R, Essen LO, Marahiel MA. Structure-based mutational analysis of the 4'-phosphopantetheinyl transferases Sfp from *Bacillus subtilis*: carrier protein recognition and reaction mechanism. *Biochemistry* 2004; **43**(14): 4128-36.
210. Ishikawa F, Haushalter RW, Lee DJ, Finzel K, Burkart MD. Sulfonyl 3-alkynyl pantetheinamides as mechanism-based cross-linkers of acyl carrier protein dehydratase. *Journal of the American Chemical Society* 2013; **135**(24): 8846-9.
211. Sieber SA, Walsh CT, Marahiel MA. Loading peptidyl-coenzyme A onto peptidyl carrier proteins: a novel approach in characterizing macrocyclization by thioesterase domains. *Journal of the American Chemical Society* 2003; **125**(36): 10862-6.

212. Yin J, Straight PD, McLoughlin SM, Zhou Z, Lin AJ, Golan DE *et al.* Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proceedings of the National Academy of Sciences of the United States of America* 2005; **102**(44): 15815-20.
213. Yasgar A, Foley TL, Jadhav A, Inglese J, Burkart MD, Simeonov A. A strategy to discover inhibitors of *Bacillus subtilis* surfactin-type phosphopantetheinyl transferase. *Molecular bioSystems* 2010; **6**(2): 365-75.
214. Leblanc C, Prudhomme T, Tabouret G, Ray A, Burbaud S, Cabantous S *et al.* 4'-Phosphopantetheinyl transferase PptT, a new drug target required for *Mycobacterium tuberculosis* growth and persistence in vivo. *PLoS pathogens* 2012; **8**(12): e1003097.
215. Rogov VV, Rozenknop A, Rogova NY, Lohr F, Tikole S, Jaravine V *et al.* A universal expression tag for structural and functional studies of proteins. *Chembiochem : a European journal of chemical biology* 2012; **13**(7): 959-63.

## 7. Appendix

### 7.1 Protein Sequences and Properties

All protein sequences are given as coded by the genes. Sites for posttranslational modification with ppan are depicted in bold, and the residues that were cleaved off by TEV protease are underlined.

#### Sfp (from pQE60)

```

1   MKIYGIYMDR PLSQEENERF MTFISPEKRE KCRRFYHKED AHRTLLGDVL
51  VRSVISRQYQ LDKSDIRFST QEYGKPCIPD LPDAHFNISH SGRWVIGAFD
101 SQPIGIDIEK TKPISLEIAK RFFSKTEYS D LLAKDKDEQT DYFYHLWSMK
151 ESFIKQEGKG LSLPLDSFSV RLHQDGQVSI ELPDSHSPCY IKTYEVDPGY
201 KMAVCAAHPD FPEDITMVS Y EELLRSHHHH HH

```

Extinction coefficient at 280 nm: 26740 M<sup>-1</sup>cm<sup>-1</sup>

Average molecular weight: 27201.7 Da

#### TycC3\_PCP (from pQE70)

```

1   MPVTEAQYVA PTNAVESKLA EIWERVLGVS GIGILDNFFQ IGGHSLKAMA
51  VAAQVHREYQ VELPLKVLFA QPTIKALAQY VATRSHHHHH H

```

M1 was cleaved *in vivo* in all variants.

Extinction coefficient at 280 nm (for P2-H91): 9530 M<sup>-1</sup>cm<sup>-1</sup>

Average molecular weight (for P2-H91; *apo*): 9973.4 Da

#### TycC3\_PCP (from pBH4)

```

1   MGHHHHHHHDY DIPTTENLYF QGSAQYVAPT NAVESKLAEI WERVLGVSGI
51  GILDNFFQIG GHSLKAMAVA AQVHREYQVE LPLKVLFAQP TIKALAQYVA
101 TSGK

```

Extinction coefficient at 280 nm (for G22-K104): 9530 M<sup>-1</sup>cm<sup>-1</sup>

Average molecular weight (for G22-K104; *apo*): 8897.3 Da

**TycC4\_C (from pET28)**

1 MAMYVLRQFA DTGTVYNMPS ALYIEGDLDR KRFEAAIHGL VERHESLRTS  
51 FHTVNGEPVQ RVHEHVELNV QYAEVTEAQV EPTVESFVQA FDLTKAPLLR  
101 VGLFKLAAGR HLFLLDMHHI ISDGVSAGII MEEFSKLYRG EELPALSVDY  
151 KDFAVWQSEL FQSDVYTEHE NYWLNAFSGD IPVLNLPADF SRPLTQSFEG  
201 DCVSFQADKA LLDDLHKLAQ ESQSTLFMVL LAAYNVLLAK YSGQEDIVVG  
251 TPIAGRSHAD IENVLGMFVN TLALRNYPVE TKHFQAFLEE VKQNTLQAYA  
301 HQDYPFEALV EKLDIQRDLS RNPLFDTMFI LQNLQKAYE LDGLKLEAYP  
351 AQAGNAKFDL TLEAHEDETG IHFALVYSTK LFQRESIERM AGHFLQVLRQ  
401 VVADQATALR LEHHHHHH

Extinction coefficient at 280 nm: 31860 M<sup>-1</sup>cm<sup>-1</sup>

Average molecular weight: 47685.8 Da

**TycC4\_PCP (from pBH4)**

1 MGHHHHHHHDY DIPTTENLYF QGSSAFVAAQ NDTEAKLQOI WQEVLGIPAI  
51 GIHDNFFEIG GHSLKAMNVI TQVHKTFQVE LPLKALFATP TIHELAAHIA  
101 ESAFE

Extinction coefficient at 280 nm (for G22-E105): 5690 M<sup>-1</sup>cm<sup>-1</sup>

Average molecular weight (for G22-E105; apo): 9138.4 Da

**TycC5-6\_PCP-C (from pBH4)**

1 MGHHHHHHHHH HHDYDIPTTE NLYFQGSEYV APRSVWEARL AQVWEQVLNV  
51 PQVGALDDFF ALGGHSLRAM RVISSMHNEY QVDIPLRILF EKPTIQELAA  
101 FIEETAKGNV FSIEPVQKQA YYPVSSAQKR MYILDQFEGV GISYNPSTM  
151 LIEGKLERTR VEAAFQRLIA RHESLRTSFA VVNGEPVQNI HEDVPFALAY  
201 SEVTEQEARE LVSSLVQPFQ LEVAPLIRVS LLKIGEDRYV LFTDMHHSIS  
251 DGVSSGILLA EWVQLYQGDV LPELRIQYKD FAVWQQEFSQ SAAFHKQEAY  
301 WLQTFADDIP VLNLPTDFTR PSTQSFAGDQ CTIGAGKALT EGLHQLAQAT  
351 GTTLYMVLLA AYNVLLAKYA GQEDIIVGTP ITGRSHADLE PIVGMFVNTL  
401 AMRNKPQREK TFSEFLQEVK QNALDAYGHQ DYPFEELVEK LAIARDLSRN  
451 PLFDTVFTFQ NSTEEVMTLP ECTLAPFMTD ETGQHAKFDL TFSATEEREE  
501 MTIGVEYSTS LFTRETMERF SRHFLTIAAS IVQNPHIRLG EIDML

Extinction coefficient at 280 nm (for G26-L545): 50210 M<sup>-1</sup>cm<sup>-1</sup>

Average molecular weight (for G26-L545; apo): 58749.5 Da

## 7.2 Resonance Assignment of peptidyl-TycC3\_PCP

Residue	Group	CS/ppm	Residue	Group	CS/ppm	Residue	Group	CS/ppm			
02	Pro	CA	62.347	08	Tyr	CA	60.243	13	Gln	N	116.058
02	Pro	CB	32.674	08	Tyr	CB	38.142	14	Ala	CA	55.472
02	Pro	CD	49.429	08	Tyr	CD1	133.232	14	Ala	CB	19.012
02	Pro	CG	26.588	08	Tyr	CE1	117.903	14	Ala	H	8.704
02	Pro	HA	4.484	08	Tyr	H	8.528	14	Ala	HA	4.136
02	Pro	HD2	3.428	08	Tyr	HA	4.386	14	Ala	N	120.764
02	Pro	HD3	3.368	08	Tyr	HB2	3.182	14	Ala	QB	1.515
02	Pro	QB	2.504	08	Tyr	HB3	2.763	15	Val	CA	65.161
02	Pro	QG	2.042	08	Tyr	N	124.581	15	Val	CB	31.512
03	Val	CA	63.154	08	Tyr	QD	7.068	15	Val	CG1	21.925
03	Val	CB	32.602	08	Tyr	QE	6.591	15	Val	CG2	21.294
03	Val	CG1	21.337	09	Val	CA	61.301	15	Val	H	7.948
03	Val	CG2	20.679	09	Val	CB	34.772	15	Val	HA	4.191
03	Val	HA	4.152	09	Val	CG1	21.380	15	Val	HB	2.011
03	Val	HB	2.111	09	Val	H	6.758	15	Val	N	120.399
03	Val	QG1	0.960	09	Val	HA	3.746	15	Val	QG1	0.865
03	Val	QG2	0.970	09	Val	HB	1.526	15	Val	QG2	0.740
04	Thr	CA	61.418	09	Val	N	130.087	16	Glu	CA	59.502
04	Thr	CB	70.243	09	Val	QQG	0.780	16	Glu	CB	30.846
04	Thr	CG2	21.829	10	Ala	CA	50.591	16	Glu	CG	37.105
04	Thr	H	8.274	10	Ala	CB	17.659	16	Glu	H	8.481
04	Thr	HA	4.458	10	Ala	H	8.144	16	Glu	HA	3.550
04	Thr	HB	4.287	10	Ala	HA	4.171	16	Glu	HB2	2.029
04	Thr	N	116.963	10	Ala	N	130.286	16	Glu	HB3	1.609
04	Thr	QG2	1.222	10	Ala	QB	1.368	16	Glu	HG2	2.367
05	Glu	CA	56.688	11	Pro	CA	63.813	16	Glu	HG3	1.807
05	Glu	CB	30.259	11	Pro	CB	32.651	16	Glu	N	118.752
05	Glu	CG	36.328	11	Pro	CD	50.151	17	Ser	CA	62.464
05	Glu	H	8.476	11	Pro	CG	28.063	17	Ser	CB	62.680
05	Glu	HA	4.303	11	Pro	HA	4.188	17	Ser	H	8.397
05	Glu	HB2	2.117	11	Pro	HD2	3.748	17	Ser	HA	3.973
05	Glu	HB3	1.973	11	Pro	HD3	3.648	17	Ser	N	112.638
05	Glu	N	122.482	11	Pro	HG2	2.196	17	Ser	QB	3.908
05	Glu	QG	2.287	11	Pro	HG3	1.582	18	Lys	CA	57.370
06	Ala	CA	52.375	11	Pro	QB	1.911	18	Lys	CB	29.985
06	Ala	CB	19.510	12	Thr	CA	61.956	18	Lys	CD	26.609
06	Ala	H	8.339	12	Thr	CB	71.614	18	Lys	CE	41.300
06	Ala	HA	4.294	12	Thr	CG2	21.797	18	Lys	CG	24.076
06	Ala	N	124.889	12	Thr	H	9.711	18	Lys	H	8.107
06	Ala	QB	1.355	12	Thr	HA	4.473	18	Lys	HA	4.290
07	Gln	CA	55.504	12	Thr	HB	4.310	18	Lys	HB2	2.444
07	Gln	CB	29.441	12	Thr	N	113.598	18	Lys	HB3	2.018
07	Gln	CG	33.665	12	Thr	QG2	1.225	18	Lys	HD2	1.839
07	Gln	H	8.358	13	Gln	CA	51.864	18	Lys	HD3	1.632
07	Gln	HA	4.268	13	Gln	CB	39.715	18	Lys	HE2	2.923
07	Gln	HB2	2.037	13	Gln	H	7.418	18	Lys	HE3	2.748
07	Gln	HB3	1.966	13	Gln	HA	4.756	18	Lys	HG2	1.635
07	Gln	N	120.527	13	Gln	HB2	3.105	18	Lys	HG3	1.454
07	Gln	QG	2.343	13	Gln	HB3	2.951	18	Lys	N	124.451

## 7. APPENDIX

Residue	Group	CS/ppm	Residue	Group	CS/ppm	Residue	Group	CS/ppm
19 Leu	CA	58.157	23 Trp	HE1	7.207	28 Gly	QA	3.931
19 Leu	CB	42.147	23 Trp	HE3	7.578	29 Val	CA	59.836
19 Leu	CD1	26.400	23 Trp	HH2	7.459	29 Val	CB	35.505
19 Leu	CD2	24.163	23 Trp	HZ2	6.827	29 Val	CG1	21.270
19 Leu	CG	27.637	23 Trp	HZ3	7.233	29 Val	CG2	20.022
19 Leu	H	8.753	23 Trp	N	120.265	29 Val	H	7.049
19 Leu	HA	4.033	23 Trp	NE1	117.565	29 Val	HA	4.460
19 Leu	HB2	1.906	24 Glu	CA	60.681	29 Val	HB	1.965
19 Leu	HB3	0.867	24 Glu	CB	29.359	29 Val	N	115.115
19 Leu	HG	1.970	24 Glu	CG	37.877	29 Val	QG1	0.896
19 Leu	N	118.589	24 Glu	H	9.059	29 Val	QG2	0.845
19 Leu	QD1	0.737	24 Glu	HA	4.210	30 Ser	CA	56.977
19 Leu	QD2	0.847	24 Glu	HB2	2.270	30 Ser	CB	65.265
20 Ala	CA	56.190	24 Glu	HB3	2.181	30 Ser	H	8.179
20 Ala	CB	16.849	24 Glu	HG2	2.875	30 Ser	HA	4.680
20 Ala	H	8.208	24 Glu	HG3	2.458	30 Ser	N	115.464
20 Ala	HA	3.780	24 Glu	N	119.298	30 Ser	QB	3.751
20 Ala	N	121.064	25 Arg	CA	58.980	31 Gly	CA	46.522
20 Ala	QB	1.408	25 Arg	CB	30.219	31 Gly	H	8.573
21 Glu	CA	59.782	25 Arg	CD	43.733	31 Gly	HA2	3.880
21 Glu	CB	29.886	25 Arg	CG	27.355	31 Gly	HA3	3.756
21 Glu	CG	36.335	25 Arg	H	8.061	31 Gly	N	112.019
21 Glu	H	7.718	25 Arg	HA	4.208	32 Ile	CA	63.173
21 Glu	HA	4.132	25 Arg	N	119.400	32 Ile	CB	38.976
21 Glu	HB2	2.408	25 Arg	QB	2.067	32 Ile	CD1	15.694
21 Glu	HB3	2.360	25 Arg	QD	3.271	32 Ile	CG1	28.485
21 Glu	HG2	2.551	25 Arg	QG	1.819	32 Ile	CG2	18.529
21 Glu	HG3	2.181	26 Val	CA	65.830	32 Ile	H	8.261
21 Glu	N	118.975	26 Val	CB	32.702	32 Ile	HA	3.896
22 Ile	CA	65.862	26 Val	CG1	22.968	32 Ile	HB	1.811
22 Ile	CB	38.704	26 Val	CG2	22.230	32 Ile	HG12	1.691
22 Ile	CD1	13.913	26 Val	H	8.470	32 Ile	HG13	1.092
22 Ile	CG1	29.033	26 Val	HA	3.778	32 Ile	N	121.228
22 Ile	CG2	17.938	26 Val	HB	1.989	32 Ile	QD1	1.096
22 Ile	H	8.413	26 Val	N	119.261	32 Ile	QG2	1.196
22 Ile	HA	3.778	26 Val	QG1	1.055	33 Gly	CA	44.320
22 Ile	HB	2.247	26 Val	QG2	1.014	33 Gly	H	8.767
22 Ile	HG12	2.078	27 Leu	CA	55.764	33 Gly	HA2	4.451
22 Ile	HG13	1.215	27 Leu	CB	42.436	33 Gly	HA3	3.660
22 Ile	N	120.359	27 Leu	CD1	25.401	33 Gly	N	115.058
22 Ile	QD1	0.886	27 Leu	CD2	21.998	34 Ile	CA	64.462
22 Ile	QG2	1.133	27 Leu	CG	27.097	34 Ile	CB	37.720
23 Trp	CA	61.825	27 Leu	H	9.015	34 Ile	CD1	14.498
23 Trp	CB	29.127	27 Leu	HA	4.008	34 Ile	CG1	25.447
23 Trp	CD1	122.932	27 Leu	HB2	1.912	34 Ile	CG2	17.846
23 Trp	CE3	121.107	27 Leu	HB3	1.307	34 Ile	H	8.577
23 Trp	CH2	126.602	27 Leu	HG	1.488	34 Ile	HA	4.015
23 Trp	CZ2	112.625	27 Leu	N	116.087	34 Ile	HB	1.757
23 Trp	CZ3	123.467	27 Leu	QD1	0.504	34 Ile	N	110.067
23 Trp	H	9.558	27 Leu	QD2	0.080	34 Ile	QD1	0.744
23 Trp	HA	4.188	28 Gly	CA	46.703	34 Ile	QG2	0.667
23 Trp	HB2	3.340	28 Gly	H	8.057			
23 Trp	HB3	3.306	28 Gly	N	108.740			
23 Trp	HD1	6.917						



Residue	Group	CS/ppm	Residue	Group	CS/ppm	Residue	Group	CS/ppm
35 Leu	CA	53.243	40 Gln	CG	34.079	46 Leu	HG	1.823
35 Leu	CB	42.064	40 Gln	H	8.280	46 Leu	N	123.071
35 Leu	CD1	25.950	40 Gln	HA	4.365	46 Leu	QD1	0.992
35 Leu	CD2	22.729	40 Gln	N	119.912	46 Leu	QD2	0.974
35 Leu	CG	27.238	40 Gln	QB	2.412	47 Lys	CA	58.576
35 Leu	H	8.197	40 Gln	QG	2.433	47 Lys	CB	32.991
35 Leu	HA	4.689	41 Ile	CA	61.212	47 Lys	CD	29.175
35 Leu	HG	1.367	41 Ile	CB	37.916	47 Lys	CE	42.051
35 Leu	N	117.009	41 Ile	CD1	15.467	47 Lys	CG	26.091
35 Leu	QB	1.740	41 Ile	CG1	25.467	47 Lys	H	7.750
35 Leu	QD1	0.938	41 Ile	CG2	17.900	47 Lys	HA	4.190
35 Leu	QD2	0.851	41 Ile	H	7.148	47 Lys	HB2	1.646
36 Asp	CA	55.251	41 Ile	HA	4.471	47 Lys	HB3	1.342
36 Asp	CB	41.143	41 Ile	HB	2.151	47 Lys	HD2	1.354
36 Asp	H	7.089	41 Ile	HG12	1.279	47 Lys	HD3	1.019
36 Asp	HA	4.362	41 Ile	HG13	1.145	47 Lys	HG2	1.234
36 Asp	HB2	2.782	41 Ile	N	111.094	47 Lys	HG3	1.091
36 Asp	HB3	2.590	41 Ile	QD1	0.954	47 Lys	N	119.398
36 Asp	N	121.488	41 Ile	QD2	0.959	47 Lys	QE	2.907
37 Asn	CA	53.296	42 Gly	CA	44.742	48 Ala	CA	55.122
37 Asn	CB	39.592	42 Gly	H	7.443	48 Ala	CB	18.378
37 Asn	H	9.127	42 Gly	HA2	4.310	48 Ala	H	8.721
37 Asn	HA	4.944	42 Gly	HA3	3.510	48 Ala	HA	3.579
37 Asn	N	123.300	42 Gly	N	106.400	48 Ala	N	122.253
37 Asn	QB	2.836	43 Gly	CA	45.469	48 Ala	QB	1.457
38 Phe	CA	61.975	43 Gly	H	7.840	49 Met	CA	58.789
38 Phe	CB	40.895	43 Gly	HA2	2.657	49 Met	CB	32.514
38 Phe	CD1	131.425	43 Gly	HA3	1.870	49 Met	CE	17.219
38 Phe	CE1	129.485	43 Gly	N	107.945	49 Met	CG	32.515
38 Phe	CZ	127.501	44 His	CA	54.836	49 Met	H	7.854
38 Phe	H	9.334	44 His	CB	31.762	49 Met	HA	4.156
38 Phe	HA	3.404	44 His	CD2	120.486	49 Met	HB2	2.301
38 Phe	HB2	2.758	44 His	CE1	139.519	49 Met	HB3	2.248
38 Phe	HB3	1.789	44 His	H	6.368	49 Met	HG2	2.859
38 Phe	HZ	6.302	44 His	HA	5.031	49 Met	HG3	2.723
38 Phe	N	126.423	44 His	HB2	3.542	49 Met	N	116.814
38 Phe	QD	4.700	44 His	HB3	3.485	49 Met	QE	2.113
38 Phe	QE	6.195	44 His	HD2	7.111	50 Ala	CA	55.250
39 Phe	CA	59.714	44 His	HE1	7.884	50 Ala	CB	18.346
39 Phe	CB	38.115	44 His	N	112.667	50 Ala	H	7.418
39 Phe	CD1	132.213	45 Ser	CA	62.618	50 Ala	HA	4.351
39 Phe	CE1	130.600	45 Ser	CB	65.450	50 Ala	N	123.485
39 Phe	CZ	129.237	45 Ser	HA	4.361	50 Ala	QB	1.640
39 Phe	H	8.053	45 Ser	QB	4.360	51 Val	CA	66.993
39 Phe	HA	3.928	46 Leu	CA	58.586	51 Val	CB	31.265
39 Phe	HB2	3.284	46 Leu	CB	41.582	51 Val	CG1	23.243
39 Phe	HB3	2.955	46 Leu	CD1	24.922	51 Val	CG2	20.959
39 Phe	HZ	6.186	46 Leu	CD2	23.791	51 Val	H	7.979
39 Phe	N	112.237	46 Leu	CG	27.779	51 Val	HA	3.428
39 Phe	QD	7.334	46 Leu	H	8.791	51 Val	HB	2.110
39 Phe	QE	6.972	46 Leu	HA	4.463	51 Val	N	118.719
40 Gln	CA	57.209	46 Leu	HB2	1.945	51 Val	QG1	0.816
40 Gln	CB	29.456	46 Leu	HB3	1.763	51 Val	QG2	0.742

## 7. APPENDIX

Residue	Group	CS/ppm	Residue	Group	CS/ppm	Residue	Group	CS/ppm
52 Ala	CA	56.163	58 Glu	CA	58.355	63 Leu	H	8.961
52 Ala	CB	18.064	58 Glu	CB	30.553	63 Leu	HA	4.839
52 Ala	H	8.456	58 Glu	CG	34.890	63 Leu	HB2	1.862
52 Ala	HA	3.979	58 Glu	H	7.890	63 Leu	HB3	1.196
52 Ala	N	121.598	58 Glu	HA	3.937	63 Leu	HG	1.460
52 Ala	QB	1.640	58 Glu	HB2	1.824	63 Leu	N	127.914
53 Ala	CA	55.465	58 Glu	HB3	1.434	63 Leu	QD1	0.741
53 Ala	CB	18.119	58 Glu	HG2	1.958	63 Leu	QD2	0.625
53 Ala	H	8.288	58 Glu	HG3	1.344	64 Pro	CA	62.077
53 Ala	HA	4.307	58 Glu	N	116.122	64 Pro	CB	32.315
53 Ala	N	121.975	59 Tyr	CA	58.644	64 Pro	CD	51.148
53 Ala	QB	1.622	59 Tyr	CB	40.547	64 Pro	CG	27.833
54 Gln	CA	59.325	59 Tyr	CD1	132.872	64 Pro	HA	4.446
54 Gln	CB	28.685	59 Tyr	CE1	117.759	64 Pro	HB2	2.471
54 Gln	CG	34.198	59 Tyr	H	8.285	64 Pro	HB3	1.722
54 Gln	H	7.933	59 Tyr	HA	4.684	64 Pro	HD2	4.062
54 Gln	HA	4.153	59 Tyr	HB2	3.206	64 Pro	HD3	3.284
54 Gln	HB2	2.285	59 Tyr	HB3	2.575	64 Pro	QG	2.030
54 Gln	HB3	2.186	59 Tyr	N	112.216	65 Leu	CA	58.367
54 Gln	HG2	2.652	59 Tyr	QD	7.116	65 Leu	CB	40.540
54 Gln	HG3	2.523	59 Tyr	QE	6.583	65 Leu	CD1	24.878
54 Gln	N	118.546	60 Gln	CA	57.405	65 Leu	CD2	23.877
55 Val	CA	67.537	60 Gln	CB	25.599	65 Leu	CG	27.481
55 Val	CB	31.547	60 Gln	CG	34.606	65 Leu	H	8.743
55 Val	CG1	24.694	60 Gln	H	8.381	65 Leu	HA	3.714
55 Val	CG2	21.994	60 Gln	HA	3.477	65 Leu	HB2	1.598
55 Val	H	8.914	60 Gln	HB2	2.273	65 Leu	HB3	1.280
55 Val	HA	3.675	60 Gln	HB3	2.211	65 Leu	HG	1.675
55 Val	HB	2.362	60 Gln	HG2	2.204	65 Leu	N	126.260
55 Val	N	121.705	60 Gln	HG3	2.131	65 Leu	QD1	0.907
55 Val	QQG	1.119	60 Gln	N	115.665	65 Leu	QD2	0.799
56 His	CA	60.931	61 Val	CA	60.083	66 Lys	CA	59.400
56 His	CB	30.206	61 Val	CB	35.536	66 Lys	CB	32.116
56 His	CD2	118.366	61 Val	CG1	21.203	66 Lys	CD	29.464
56 His	CE1	138.258	61 Val	H	6.533	66 Lys	CE	41.926
56 His	H	8.306	61 Val	HA	4.344	66 Lys	CG	24.886
56 His	HA	4.325	61 Val	HB	1.697	66 Lys	H	8.561
56 His	HB2	3.461	61 Val	N	114.510	66 Lys	HA	3.898
56 His	HB3	3.237	61 Val	QQG	0.873	66 Lys	HB2	1.800
56 His	HD2	6.583	62 Glu	CA	55.866	66 Lys	HB3	1.707
56 His	HE1	7.767	62 Glu	CB	30.299	66 Lys	N	114.175
56 His	N	120.918	62 Glu	CG	36.304	66 Lys	QD	1.642
57 Arg	CA	59.253	62 Glu	H	8.759	66 Lys	QE	2.918
57 Arg	CB	30.240	62 Glu	HA	4.245	66 Lys	QG	1.353
57 Arg	CD	43.390	62 Glu	HG2	2.177	67 Val	CA	65.617
57 Arg	CG	27.227	62 Glu	HG3	1.929	67 Val	CB	31.434
57 Arg	H	7.879	62 Glu	N	125.922	67 Val	CG1	22.529
57 Arg	HA	3.957	62 Glu	QB	1.911	67 Val	CG2	19.507
57 Arg	HG2	1.780	63 Leu	CA	50.757	67 Val	H	7.159
57 Arg	HG3	1.669	63 Leu	CB	43.306	67 Val	HA	3.595
57 Arg	N	119.328	63 Leu	CD1	26.202	67 Val	HB	1.775
57 Arg	QB	1.983	63 Leu	CD2	24.772	67 Val	N	119.733
57 Arg	QD	3.232	63 Leu	CG	26.880	67 Val	QG1	0.738
						67 Val	QG2	0.092

Residue	Group	CS/ppm	Residue	Group	CS/ppm	Residue	Group	CS/ppm
68 Leu	CA	58.261	73 Thr	CA	59.305	77 Leu	QD2	0.915
68 Leu	CB	41.933	73 Thr	CB	72.331	78 Ala	CA	55.100
68 Leu	CD1	25.197	73 Thr	CG2	21.465	78 Ala	CB	19.617
68 Leu	CD2	24.980	73 Thr	H	8.360	78 Ala	H	9.392
68 Leu	CG	28.136	73 Thr	HA	5.189	78 Ala	HA	3.747
68 Leu	H	6.888	73 Thr	HB	4.691	78 Ala	N	120.479
68 Leu	HA	3.916	73 Thr	N	110.118	78 Ala	QB	1.421
68 Leu	HB2	1.493	73 Thr	QG2	1.071	79 Gln	CA	58.788
68 Leu	HB3	1.279	74 Ile	CA	66.000	79 Gln	CB	28.220
68 Leu	HG	1.278	74 Ile	CB	37.454	79 Gln	CG	33.771
68 Leu	N	120.871	74 Ile	CD1	13.312	79 Gln	H	7.181
68 Leu	QD1	0.506	74 Ile	CG1	28.823	79 Gln	HA	4.067
68 Leu	QD2	0.217	74 Ile	CG2	18.584	79 Gln	HG2	2.535
69 Phe	CA	59.932	74 Ile	H	9.913	79 Gln	HG3	2.370
69 Phe	CB	38.566	74 Ile	HA	3.056	79 Gln	N	117.823
69 Phe	CD1	131.276	74 Ile	HB	1.152	79 Gln	QB	2.187
69 Phe	CE1	130.975	74 Ile	HG12	-0.158	80 Tyr	CA	61.982
69 Phe	CZ	129.356	74 Ile	HG13	-0.798	80 Tyr	CB	38.228
69 Phe	H	8.209	74 Ile	N	125.114	80 Tyr	CD1	133.054
69 Phe	HA	4.259	74 Ile	QD1	0.009	80 Tyr	CE1	118.044
69 Phe	HB2	2.981	74 Ile	QG2	0.683	80 Tyr	H	7.712
69 Phe	HB3	2.917	75 Lys	CA	60.213	80 Tyr	HA	3.987
69 Phe	HZ	7.180	75 Lys	CB	33.095	80 Tyr	HB2	3.295
69 Phe	N	112.790	75 Lys	CD	29.837	80 Tyr	HB3	3.076
69 Phe	QD	7.043	75 Lys	CE	42.218	80 Tyr	N	121.716
69 Phe	QE	7.041	75 Lys	CG	24.820	80 Tyr	QD	6.898
70 Ala	CA	53.991	75 Lys	H	9.036	80 Tyr	QE	6.753
70 Ala	CB	19.626	75 Lys	HA	3.793	81 Val	CA	66.721
70 Ala	H	7.426	75 Lys	HB2	1.717	81 Val	CB	32.018
70 Ala	HA	4.283	75 Lys	HB3	1.531	81 Val	CG1	23.252
70 Ala	N	118.491	75 Lys	HD2	1.550	81 Val	CG2	20.545
70 Ala	QB	1.554	75 Lys	HD3	1.477	81 Val	H	8.543
71 Gln	CA	53.212	75 Lys	HG2	1.396	81 Val	HA	3.255
71 Gln	CB	29.159	75 Lys	HG3	1.208	81 Val	HB	1.807
71 Gln	CG	33.636	75 Lys	N	119.492	81 Val	N	117.874
71 Gln	H	7.495	75 Lys	QE	2.902	81 Val	QG1	0.909
71 Gln	HA	4.834	76 Ala	CA	54.798	81 Val	QG2	0.205
71 Gln	HB2	2.048	76 Ala	CB	19.424	82 Ala	CA	55.139
71 Gln	HB3	1.960	76 Ala	H	7.621	82 Ala	CB	18.582
71 Gln	HG2	2.283	76 Ala	HA	4.216	82 Ala	H	8.411
71 Gln	HG3	2.218	76 Ala	N	119.598	82 Ala	HA	3.926
71 Gln	N	116.078	76 Ala	QB	1.533	82 Ala	N	121.040
72 Pro	CA	64.275	77 Leu	CA	57.608	82 Ala	QB	1.441
72 Pro	CB	30.089	77 Leu	CB	42.187	83 Thr	CA	64.367
72 Pro	CD	49.313	77 Leu	CD1	25.173	83 Thr	CB	69.519
72 Pro	CG	27.667	77 Leu	CD2	25.111	83 Thr	CG2	21.601
72 Pro	HA	4.996	77 Leu	CG	26.659	83 Thr	H	7.554
72 Pro	HB2	2.640	77 Leu	H	8.913	83 Thr	HA	4.082
72 Pro	HB3	2.308	77 Leu	HA	4.100	83 Thr	HB	4.188
72 Pro	HD2	3.583	77 Leu	HB2	1.910	83 Thr	N	109.056
72 Pro	HD3	3.510	77 Leu	HB3	1.494	83 Thr	QG2	1.232
72 Pro	HG2	2.276	77 Leu	HG	2.091			
72 Pro	HG3	1.976	77 Leu	N	121.911			
			77 Leu	QD1	0.990			

Residue	Group	CS/ppm	Residue	Group	CS/ppm	Residue	Group	CS/ppm			
84	Arg	CA	56.240	85	Ser	CA	58.929	87	His	CA	57.329
84	Arg	CB	30.840	85	Ser	CB	63.752	87	His	CB	30.321
84	Arg	CD	42.320	85	Ser	H	7.790	87	His	H	8.147
84	Arg	CG	26.974	85	Ser	HA	4.294	87	His	HA	4.427
84	Arg	H	7.373	85	Ser	HB2	3.904	87	His	HB2	3.203
84	Arg	HA	4.197	85	Ser	HB3	3.760	87	His	HB3	3.074
84	Arg	HB2	1.922	85	Ser	N	114.639	87	His	N	125.437
84	Arg	HB3	1.409								
84	Arg	HD2	2.854	86	His	CA	56.028				
84	Arg	HD3	2.563	86	His	CB	30.104				
84	Arg	N	119.874	86	His	HA	4.606				
84	Arg	QG	1.369	86	His	HB2	3.140				
				86	His	HB3	3.080				

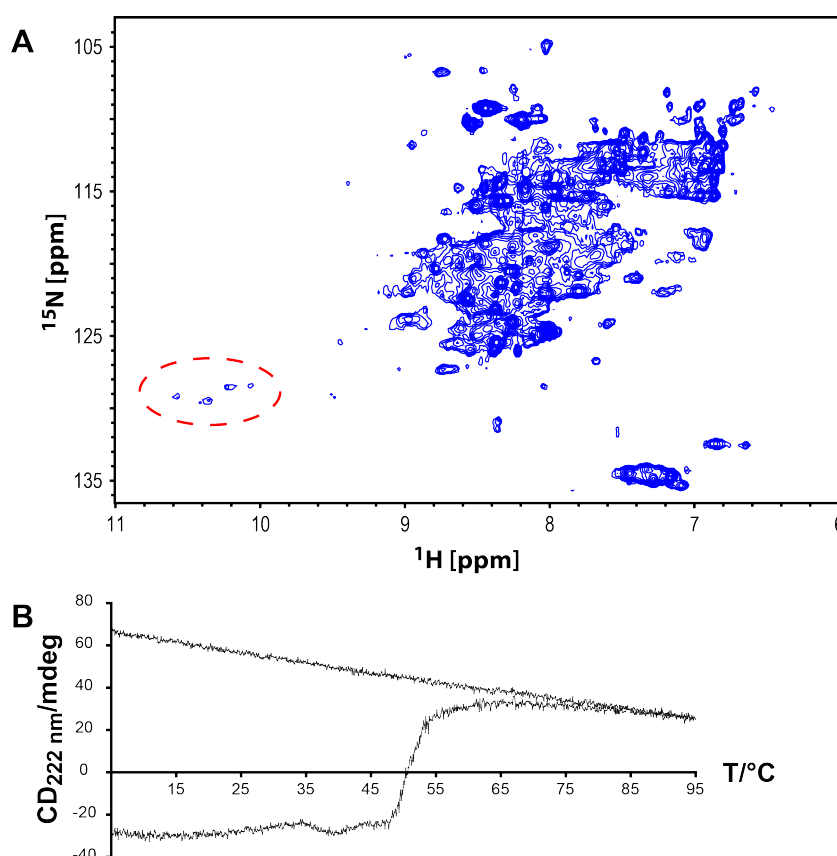
The assignment of TycC3\_PCP(S45A) is deposited at the Biological Magnetic Resonance Bank under accession number 19479.

## 7.3 Results of Side Projects

### 7.3.1 Cyclization Domain

A construct of BacA2\_Cy, the cyclization domain of module 2 of the bacitracin A synthetase subunit A, in pBH4(His6) can be expressed in E coli BL21(DE3) cells in various cultivation media and at temperatures between 16-30°C, with high yields (>50 mg/L). The protein forms soluble aggregates *in vitro* when purified following the standard protocol (nickel affinity purification, treatment with TEV protease, reverse nickel affinity purification and SEC), but addition of 10% (w/v) sucrose to all buffers can prevent aggregation, and the monomeric fraction can be separated from the aggregates formed *in vivo* by SEC. After separation of the aggregates, the monomeric fraction is relatively stable, even when the sucrose is removed by dialysis against 25 mM HEPES, 50 mM KCl (pH 7.2) and subsequent SEC: When stored for 2 days at RT, only ~10% aggregate is formed.

Recording of a [<sup>1</sup>H;<sup>15</sup>N]-HSQC of a uniformly <sup>15</sup>N-labeled sample revealed that the protein is not suitable for structure determination by NMR, due to a poor quality of the spectrum (Figure 42A).

**Figure 42**

A  $[^1\text{H};^{15}\text{N}]$ -HSQC of 250  $\mu\text{M}$  BacA2\_Cy in 50 mM Arg/Glu and 5% (w/v) sucrose at pH 6.8 recorded at 298 K gave a poor overall dispersity of the signals due to the size of the protein although the signals of the tryptophan side chain (dashed line) indicate that the protein is folded (A).

A CD melting curve recorded of the a 10  $\mu\text{M}$  sample in the same buffer as used for NMR confirmed that the protein is folded as it denaturates irreversibly at  $\sim 50^\circ\text{C}$  (B).

Nevertheless, the dispersion of the signals of the tryptophan side chains indicate that the protein is folded, which was confirmed as it can be thermally denaturated (Figure 42B; due to the sucrose the ellipticity could not be recorded below  $\sim 215$  nm. Therefore, no CD spectrum was recorded).

### 7.3.2 *In vivo* Expression of Peptides

The precursors of the lasso peptides astexin-1 and microcin J25 could be expressed *in vivo* without any observable degradation, when an ubiquitin-fusion system introduced by Rogov *et al.*<sup>215</sup> was used.

To facilitate binding of the internal his-tag, the nickel affinity purification was performed in 6 M urea, and the fusion protein could be cleaved off quantitatively with TEV protease after refolding by dialysis. The product of the cleavage was purified by HPLC with a CC 250/2 Nucleodur 100-3C18ec column and a ddH<sub>2</sub>O/MeCN gradient.

Endothelin-1 could be produced in the same way, but showed different variations of disulfide bridging and co-eluted with ubiquitin when applied to HPLC purification.

## Curriculum Vitae

### Personal Details

---

Name: Peter Tufar  
Date of birth: 14.02.1984  
Place of birth: Marburg/Lahn (Germany)  
Nationality: German, Austrian



### Education

---

02/2009 – 01/2014 **Goethe University Frankfurt, Germany**  
PhD thesis in the group of Prof. Dr. Volker Dötsch  
*“Structural and Functional Investigation of the Selectivity in Interactions of Peptidyl Carrier Proteins from Bacillus”*

10/2003-10/2008 **Philipps-University Marburg, Germany**  
Diploma degree in chemistry (Grade 1.4)  
Diploma thesis in the group of Prof. Dr. Mohamed A. Marahiel  
*“Investigation of the Maturation and Structural Stability of the Lasso Peptides Capistruin and Microcin J25 by NMR Spectroscopy”*

09/1994-06/2003 **Elisabethschule Marburg, Germany**  
General qualification for university entrance (Grade 1.6)

### Professional Experience

---

since 02/2009 **Goethe University Frankfurt, Germany**  
PhD thesis in the group of Prof. Dr. Volker Dötsch

10/2008-01/2009 **Philipps-University Marburg, Germany**  
Internship in the group of Prof. Dr. Mohamed A. Marahiel

06/2003-07/2003 **Max-Planck-Institute for Polymer Research Mainz, Germany**  
Internship in the group of Prof. Dr. Hans W. Spiess

05/2001-01/2009 **Elisabethkirche Marburg, Germany**  
Employment (12.5 h/week) as a sexton

## Honors

---

04/2005-09/2008      Scholarship of the Hessian Institution of Scholarships

## Publications

---

**Tufar, P.**, Rahighi, S., Kraas, F.I., Kirchner, D.K., Löhr, F., Henrich, E., Köpke, J., Dikic, I., Güntert, P., Marahiel, M.A., and Dötsch, V. (submitted) *The structural basis for carrier protein posttranslational modification revealed by resolving the crystal structure of a PCP/Sfp complex*

Luh, L.M., Hänsel, R., Löhr, F., Kirchner, D.K., Krauskopf, K., Pitzius, S., Schäfer, B., **Tufar, P.**, Corbeski, I., Güntert, P., and Dötsch, V. (2013). *Molecular crowding drives active pin1 into nonspecific complexes with endogenous proteins prior to substrate recognition* Journal of the American Chemical Society *135*, 13796-13803.

Busche, A., Gottstein, D., Hein, C., Ripin, N., Pader, I., **Tufar, P.**, Eisman, E.B., Gu, L., Walsh, C.T., Sherman, D.H., Löhr, F., Güntert, P., and Dötsch, V. (2012). *Characterization of molecular interactions between ACP and halogenase domains in the Curacin A polyketide synthase* ACS chemical biology *7*, 378-386.

Crystal structure of the TycC3\_PCP(S45A)/Sfp complex at PDB (accession code 4MRT)

Solution NMR structure of TycC3\_PCP(S45A) at PDB (accession code 2MD9) with its assignment at the BMRB (accession number 19479).

## Poster Presentations

---

- 10/2013      **Molecular Life Sciences 2013**, Frankfurt (Germany)  
*Structural and Functional Investigation of the Phospho-pantetheinylation Reaction in Nonribosomal Peptide Synthetases from Bacillus*
- 07/2013      **Euromar 2013**, Hersonissos (Greece)  
*Structural and Functional Investigation of the Phospho-pantetheinylation Reaction in Nonribosomal Peptide Synthetases from Bacillus*
- 09/2008      **30<sup>th</sup> Discussion meeting of the magnetic resonance spectroscopy division of the GDCh**, Regensburg (Germany)  
*Structural Investigation of Lasso Peptides*

## **Oral Presentations**

---

01/2013                    **BMLS retreat**, Riezlern (Austria)  
*Exchange Market in Crisis: PCP Loses A-state Rating*

## **Academic Teachers**

---

Analytical Chemistry U. Pyell, A. Seubert, X. Xie

Biochemistry            V. Dötsch, L.-O. Essen, M. A. Marahiel

Inorganic Chemistry M. Bröring, S. Dehnen, B. Harbrecht, J. Sundermeyer, W. Uhl

Organic Chemistry    A. Geyer, G. Hilt, U. Koert, T. Schrader, M. Schween

Physical Chemistry    G. Germano, N. Hampp, W.-C. Pilgrim, H. Schwalbe,  
H. Uchtmann, K.-M. Weitzel, J. H. Wendorff

Physics                    P. Jakob, C. Schwee



## Acknowledgments

Prof. Dr. Volker Dötsch möchte ich für die Betreuung meiner Doktorarbeit und die Erstellung des Erstgutachtens danken. Durch seine besondere Art der Wertschätzung und Förderung mir gegenüber ermöglichte er mir eine freie Gestaltung und Durchführung meiner Projekte; auch dafür gilt ihm mein Dank.

Prof. Dr. Mohamed A. Marahiel danke ich für die Erstellung des Zweitgutachtens. Außerdem gilt ihm mein besonderer Dank dafür, dass er mich im Laufe meiner Promotion jederzeit in seinem Arbeitskreis willkommen geheißen und meine Arbeit mit Interesse verfolgt und gefördert hat.

Während meiner Doktorarbeit hatte ich das große Privileg, mit ganz besonderen Menschen in „*The Lab*“ zu arbeiten. Daniel Schwarz und Gregor Deutsch haben meiner Laboralltag mit viel Spaß und häufiger Hilfe in allen Lebenslagen bereichert; ich freue mich sehr, dass die daraus entstandenen Freundschaften auch nach ihrem Fortgehen Bestand haben. Auch wenn er mittlerweile seine Doktorarbeit begonnen hat, wird Christian Osterburg für mich immer der „Hiwi der Herzen“ bleiben! Laura Luh und ich haben am gleichen Tag mit der Doktorarbeit angefangen und sind den ganzen Weg über Seite an Seite (oder wegen der Position unsere Schreibtische und Laborplätze wohl besser gesagt: Rücken an Rücken) durch alle Hochs und Tiefs gegangen. Das werde ich nie vergessen!

Auch dem Rest des Arbeitskreises gilt meine tiefempfundene Dankbarkeit für die nette Atmosphäre, die (fast) jeden Tag geherrscht hat, den vielen Kuchen und die schönen Feiern und Ausflüge. Insbesondere Alena Busche, Robert Hänsel, Christopher Hein, Erik Henrich und Sebastian Kehrlöser möchte ich dafür danken, dass sie immer ein offenes Ohr und einen netten Spruch für mich hatten. Außerdem möchte ich Friederike Junge danken, dass sie über lange Phasen meiner Doktorarbeit ihr bestes gegeben hat, um mich sowohl an der Uni als auch privat zu unterstützen.

Einen großen Anteil daran, dass das tägliche Arbeiten überhaupt möglich war, haben Birgit Schäfer und Sigrid Oğuzer-Fachinger, die das von uns veranstaltete Chaos immer irgendwie überblickt und den Laden am Laufen gehalten haben. Dafür gilt ihnen mein Dank ebenso wie mein Respekt. Gleiches gilt auch für Frank „Murph“ Löhr, der für die drei wichtigsten Geräte des Arbeitskreises (in dieser Reihenfolge: Kaffeemaschine, Grill und NMR Spektrometer) zuständig ist. Schade, dass er nie mit in die Nordwestkurve wollte... Und dann ist da noch der Mann, bei dem MacGyver in der Lehre war: Manfred Strupf hat für jedes meiner vielen technischen Problem eine schnelle und funktionierende Lösung gefunden.

Meiner ehemaligen Praktikantin Diana Kolbert und meinem Hiwi Benedikt Kuhn danke ich für die Unterstützung und manch amüsanten Moment.

Auch außerhalb des Arbeitskreises habe ich in Frankfurt viel Unterstützung erfahren. Besonders hervorheben möchte ich Simin Rahighi und Donata Kirchner, die beide wichtige Beiträge zum Gelingen meiner Arbeit geleistet haben und dabei stets ebenso freundlich wie hilfsbereit waren. Außerdem möchte ich Barbara Rathmann und Nina Socorro Cortina für ihre Hilfe danken.

Meine Arbeit hat mich auch oft nach Marburg geführt, weswegen ich dem gesamten AK 66 für seine Gastfreundschaft danken möchte. Ein besonderer Dank an Femke Kraas, Gabi Schimpf-Weiland, Matthias Strieker und Marcel Zimmermann für die nette Zusammenarbeit und häufige Hilfe!

Danke auch an alle, die mir immer wieder gezeigt haben, dass es neben der Doktorarbeit auch ein richtiges Leben gibt! Unvergessen sind die zahlreichen Stadionbesuche *im Herzen von Europa*, bei denen ich besonders oft die Begleitung durch Uta und Thea Müller-Kuller sowie Martin Deutsch genießen durfte.

Außerdem danke ich Matthieu Leinweber, Rainer Schoof und Eva Matejcek für viele schöne Spieleabende, Wanderungen, gemeinsames Kochen und einfach eine schöne Zeit.

Mein größter Dank gilt meiner Familie, die mich von daheim (aus Marburg), sowie über Länder- und kulturelle Grenzen (aus Wien beziehungsweise Offenbach) hinweg immer unterstützt, wenn nötig getröstet und mir unschätzbare Freude bereitet hat.