Review

Journal of PeptideScience

Received: 5 August 2009 Revised: 5 August 2009 Accepted: 5 August 2009 Published online in Wiley Interscience: 13 October 2009

(www.interscience.com) DOI 10.1002/psc.1183

Working outside the protein-synthesis rules: insights into non-ribosomal peptide synthesis‡

Mohamed A. Marahiel*

Non-ribosomally synthesized microbial peptides show remarkable structural diversity and constitute a widespread class of the most potent antibiotics and other important pharmaceuticals that range from penicillin to the immunosuppressant cyclosporine. They are assembled independent of the ribosome in a nucleic acids-independent way by a group of multimodular megaenzymes called non-ribosomal peptide synthetases. These biosynthetic machineries rely not only on the 20 canonical amino acids, but also use several different building blocks, including D-configured- and β-amino acids, methylated, glycosylated and phosphorylated residues, heterocyclic elements and even fatty acid building blocks. This structural diversity leads to a high density of functional groups, which are often essential for the bioactivity. Recent biochemical and structural studies on several non-ribosomal peptide synthetase assembly lines have substantially contributed to the understanding of the molecular mechanisms and dynamics of individual catalytic domains underlying substrate recognition and substrate shuffling among the different active sites as well as peptide bond formation and the regio- and stereoselective product release. Copyright © 2009 European Peptide Society and John Wiley & Sons, Ltd.

Keywords: non-ribosomal peptide synthetases; NRPS; peptide antibiotics; structures of NRPS-domains and modules; protein dynamics

Background

Non-ribosomally synthesized peptides which are produced by a broad variety of fungi, bacteria, and lower eukaryotes have been the source for a multitude of therapeutic agents used in different medical fields including infection diseases, cancer, and immunosuppression. In contrast to the ribosomal machinery, where the mRNA templates are translated to proteins, non-ribosomal peptides are assembled on multimodular enzymes that act as a protein-template to generate macrocyclic peptides with an unusual structural diversity.

Assembly Logic of Non-Ribosomal Peptide Synthetases (NRPS)

Although structurally diverse non-ribosomally synthesized peptides (Figure 1) share a conserved mode of synthesis that is realized by the action of the modularly organized NRPS-megaenzymes [1–5], these multienzymes define by the order of their catalytic units the sequence and chemical identity of their final peptide products. In principle, NRPS constitute an array of distinct modular sections, each of which is responsible for incorporation and if necessary modification of one defined monomer into the final product [6–8]. As a consequence, in linear NRPS assembly lines, the number of the catalytic modules and their order exactly matches the number and order of amino acids incorporated in the backbone of the corresponding peptide as shown for the NRPS-template that assembles the branched cyclic dodecapeptide bacitracin [9] (Figure 2). Modules can be further dissected into catalytic domains, each responsible for a specific synthetic step during peptide synthesis [10–12]. In each so called elongation module, at least three essential domains are needed to carry out substrate selection and activation as well as substrates covalent binding to the cofactor 4′-phosphopantetheine (PPan) and peptide bond formation (Figure 3). Within each module, the selection of a specific substrate (amino or carboxy acid) is catalyzed by an adenylation domain (A-domain, ~500 amino acids), followed by the generation of an aminoacyl AMP-mixed anhydride through ATP hydrolysis [13]. This reactive intermediate is further transported onto the thiol moiety of the PPan prosthetic group that is attached to the peptidyl carrier protein (PCP) domain, which is located downstream of the A-domain in the same module [14,15]. The PPan prosthetic arm (about 20 Å) is post-translationally attached onto each PCP domain to the side chain of a highly conserved serine residue. The holo-PCP (80–100 amino acids) is responsible for the transport of the resulting energy-rich thioester bound substrates and all elongation intermediates between the catalytic active sites. Therefore, holo PCP is referred to as the ''swinging arm'' of each NRPS module. In the next step of NRP-synthesis, the peptide bond formation between two bound PCP intermediates, in adjacent modules, is catalyzed by the condensation domain (C-domain, ~450 amino acids). The C-domain catalyzes the nucleophilic attack of the downstream PCP-bound acceptor amino acid with its α-amino group onto the activated upstream PCP-bound donor substrate [15]. The resulting peptidyl intermediate is then translocated down the assembly line for subsequent condensation and further modification steps. The upstream PPan-PCP is now ready for the next reloading reaction. Therefore, an elongation module consists of three domains in the order C-A-PCP, in most cases on

* *Correspondence to: Mohamed A. Marahiel, Biochemistry-Department of Chemistry, Philipps-University Marburg, Hans-Meerwein-Strasse, D-35043 Marburg, Germany. E-mail: marahiel@staff.uni-marburg.de*

Biochemistry-Department of Chemistry, Philipps-University Marburg, Hans-Meerwein-Strasse, D-35043 Marburg, Germany

‡ *The 2008 Max Bergmann Medal Award was assigned to Prof. Mohamed A. Marahiel on occasion of the annual meeting of the Max-Bergmann-Society in Grindelwald, Switzerland, October 5–8, 2008.*

Copyright © 2009 European Peptide Society and John Wiley & Sons, Ltd.

Biography

Mohamed A. Marahiel was born in Gaza, Palestine in 1949. He studied chemistry at the Universities of Cairo (Egypt) and Göttingen (Germany). As a DAAD scholar he obtained 1977 a PhD in biochemistry and microbiology from the University of Göttingen, where he carried out his experimental work on mitochondrial protein synthesis at the Max-Planck Institute for Experimental Medicine in Göttingen with Prof. Hans Küntzel. Subsequently, he received an assistant professor's position at the Technical University of Berlin, where in 1987 he obtained his Habilitation in biochemistry. Three years later he moved to the Philipps-Univerity Marburg as a professor of biochemistry in the chemistry department. He was a DFG fellow in 1978 and 1986 at the John Innes Institute in Norwich (UK) and at the Biolabs, Harvard University (USA), respectively. Since 2004 he is a member of the ''Leopoldina'' – Deutsche Akademie der Naturforscher and since 2009 a fellow of The Royal Society of Chemistry. In 2008 he was awarded the Max-Bergmann-Medal for his work on non-ribosomal peptide synthesis. His present research focuses on the structure-function relationship and on the elucidation of reaction mechanisms of modular peptide synthetases involved in the non-ribosomal synthesis of peptide antibiotics. His group is also interested in studying ribosomally synthesized peptide antibiotics, siderophores, and stress-induced proteins in Bacteria.

a single polypeptide chain. An initiation NRPS-module comprises only an A-PCP didomain. The principle of such an NRPS-assembly line nanomachine is shown for the biosynthesis of the branched cyclic dodecapeptide bacitracin **2** (Figure 2), which is assembled by 3 NRPS large enzymes that constitute 12 modules with 40 catalytic domains [9].

Product release from the NRPS-template is catalyzed by the thioesterase domain (TE-domain, ~280 amino acids), attached to the NRPS *C*-terminal module [16,17]. TE catalyzes the transfer of the full-length peptide chain attached to the ultimate PCP to its active site serine, generating a peptidyl-acyl-enzyme intermediate (acyl-O-TE). Depending on the nature of the final product, TEs catalyze either release of the linear peptides (hydrolysis) or macrocyclization, using an intramolecular nucleophile to generate either cyclic lactams like tyrocidine **1** and cyclosporine **5** or the cyclic lactone as surfactin **6**. TE can also release the product by peptide hydrolysis as for the heptapeptide vancomycin **4**, or even release cyclic branched peptides by using a side chain nucleophile as shown for bacitracin **2** and surfactin **6**. Some TE domains act iteratively as in the case of the tri-lacton siderophor bacillibactin **3** [18,19].

Integration of Tailoring Domains into NRPS Modules

In addition to the essential domains C, A, PCP, and TE, many auxiliary domains [20,21] are directly integrated at different locations within an NRPS elongation module (Figure 4). These *in cis* acting (within the modules integrated) tailoring units include epimerization (E) domains that are located downstream of PCP and catalyze the generation of D-configured amino acids from the corresponding L-isomers when attached to the cofactor PPan. *N*-methyltransferase domains (Mt), on the other hand, use the cofactor SAM to control the methylation state of the peptide products. For example, seven out of the eleven residues in cyclosporine **6** are *N*-methylated. Mt-domains are usually integrated at the *C*-terminal region of the A-domains and act on PCP-S-PPan bound amino acids. Another type of methyltransferases is responsible for *C*-methylation (*C*-Mt). These catalyze the modification of Cα-atoms of cysteine residues that are then converted to the corresponding α-methyl-thiazoline by the action of cyclization (Cy) domains that sometimes substitute the *C*-domains. *C*-Mt-domains are usually located downstream of the PCP-domains (Figure 4). In some NRPS modules that select cysteine, serine, or threonine residues, integrated Cy domains catalyze in addition to peptide bond formation a heterocyclic ring formation to the corresponding thiazoline or oxazoline rings. The mechanism of action of these Cy-domains (Figure 4) that led to peptide bond and ring formation is not fully understood. In close association with some Cy-domains, oxidation (Ox) domains (integrated within the *C*-terminal region of A-domains) catalyze the FMN-dependent formation of thiazoles and oxazoles, whereas reduction (R) domains catalyze the NADPH-dependent formation of thiazolidine and oxazolidine rings (Figure 4). Other modification in NRPS was found to be associated with some initiation reactions, in which the first amino acid gets formylated. For such *N*-terminally formylated NRPS peptides, like linear gramicidin, a special formyltransferase (F) domain is attached to the initiation module (Figure 4). As in the bacterial ribosomal protein synthesis, the F-domain uses N^{10}-formyl-tetrahydrofolate (N^{10}-THF) as the formyl group donor. Figure 4 summarizes some known reactions catalyzed by optional modification domains integrated within NRPS modules.

Post-Assembly Tailoring

In addition to the modifications that occur during NRPS-assembly by the integrated tailoring domains, other modifications can also take place by independent enzymes that temporally associate with the NRPS template. For example, the vancomycin **4** heptapeptide is cross-linked by oxidative cyclization of the phenolic side chain by the action of an independent cytochrome P450-type haem proteins [22,23]. Such a cross-linking causes a fixed peptide architecture that is important for activity. Also independent glycosyltransferase enzymes decorate vancomycin **4** mature aglycon scaffold during its NRPS by sugar residues essential for bioactivity [24,25]. These are only few examples from a large number of possible post-assembly tailoring reactions during the NRPS synthesis. For further reading see Walsh *et al*. [26,27].

Structural Insights into the NRPS Machinery

Crystal- and NMR-structural studies on several essential NRPS domains that were dissected from intact modules and heterologously expressed as active components provided profound insights into the catalytic mechanisms and dynamics of these multidomain enzymes. The discovery of a distinct fold for each of the essential A, PCP, C, and TE domains that constitute initiation, elongation, and termination modules, underlined the individual character of each catalytic unit. However, this provided little structural insights

 Copyright © 2009 European Peptide Society and John Wiley & Sons, Ltd.

Journal of PeptideScience

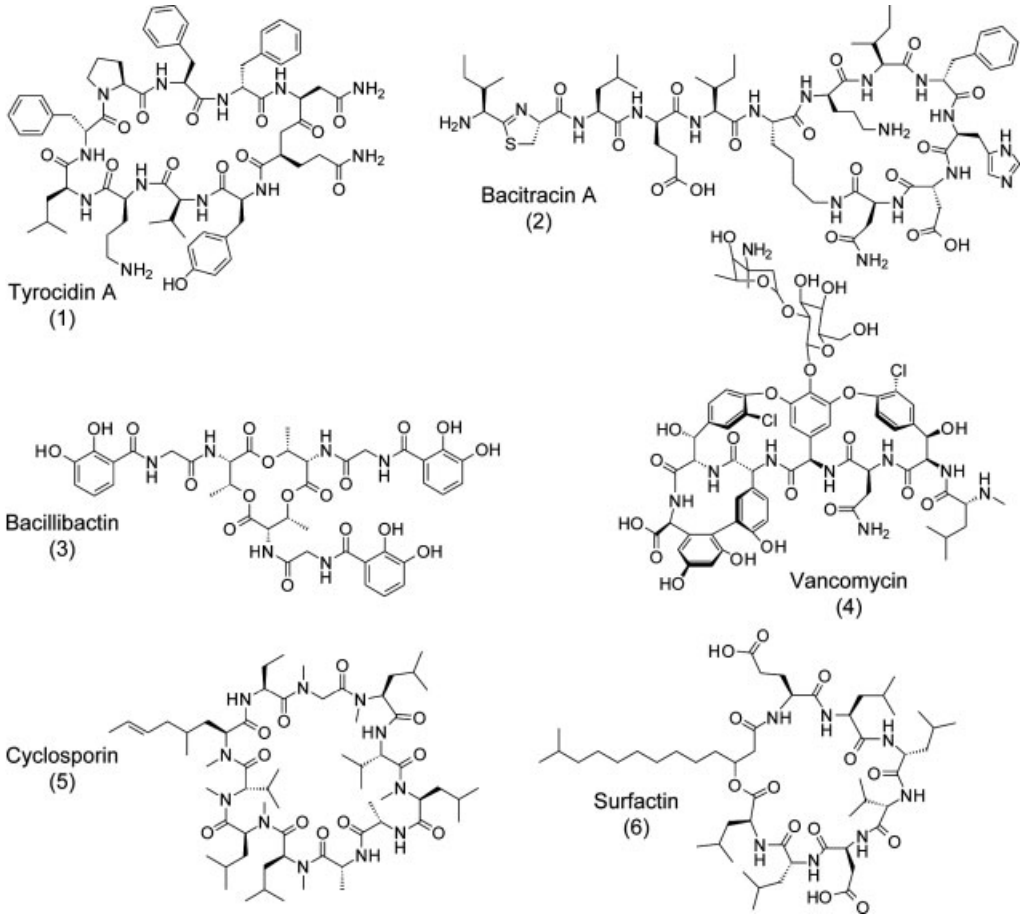

Figure 1. Examples of non-ribosomally synthesized bioactive macrocyclic peptides that comprise unusual structures (D-configured and *N*-methylated amino acids, heterocyclic rings, fatty acids, sugars and non-proteinogenic amino acids) and exhibit antibiotic activity (1, 2, 4), act as siderophore (3), biosurfactant (6), or exhibit immunosuppressive activity (5).

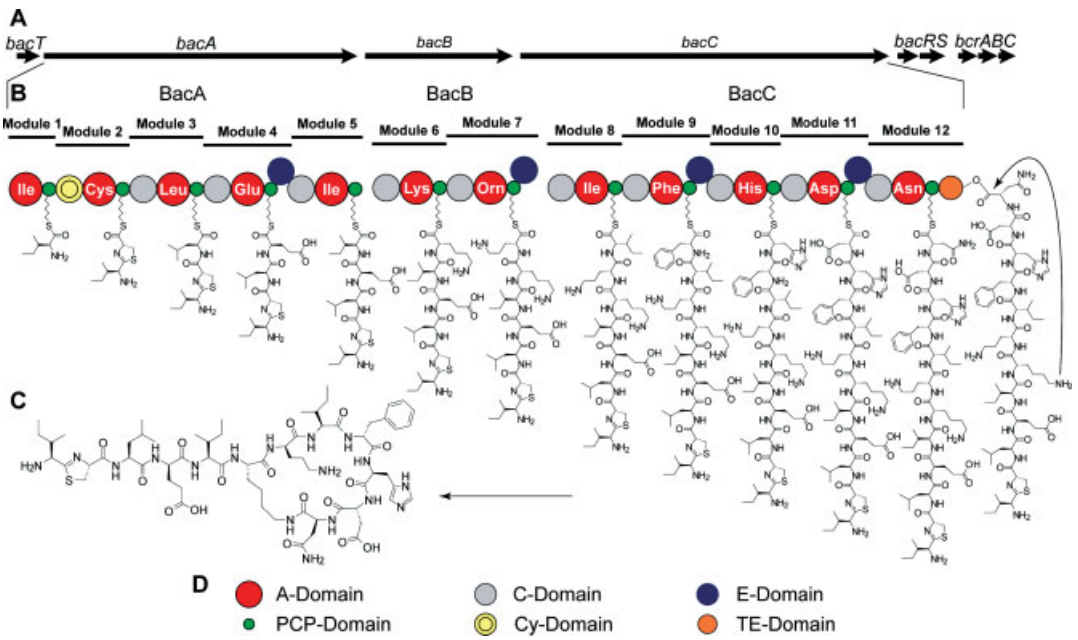

Figure 2. A prototype NRPS assembly line. (A) Biosynthetic gene cluster encoding the bacitracin synthetases (BacA, BacB, and BacC). (B) The three NRPS comprise 12 modules and a total of 40 catalytic domains that assemble the cyclic branched dodecapeptide (C). During elongation all biosynthetic intermediates remain attached as aminoacyl- or peptidyl-S-PPan to the carrier domain PCP until the final product is released from the terminal module by macrocyclization. (D) Color code definition for NRPS domains.

 Copyright © 2009 European Peptide Society and John Wiley & Sons, Ltd.

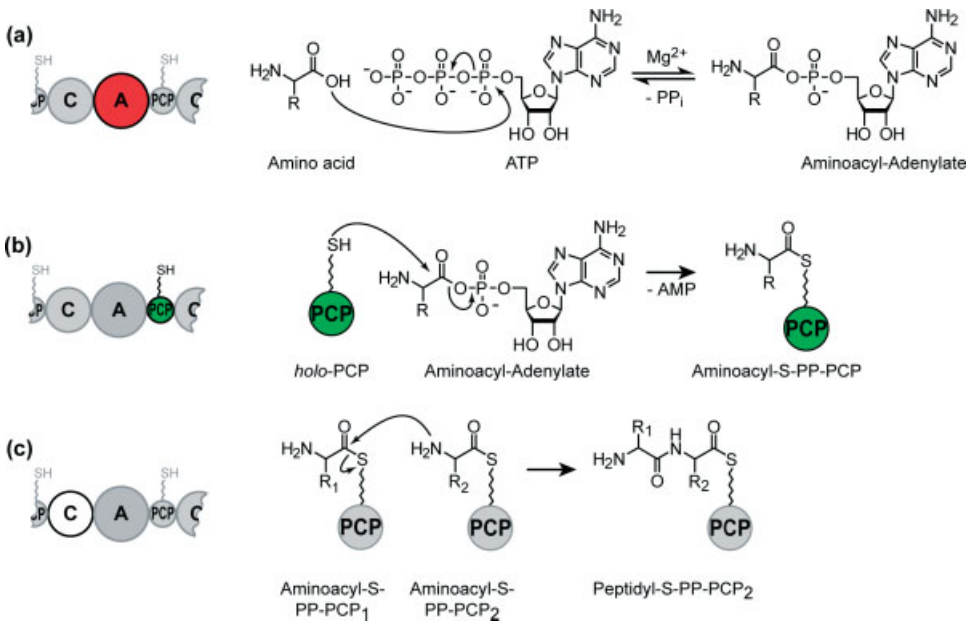

Figure 3. Reactions catalyzed by the three essential NRPS domains A, PCP and C. The cofactor 4′-phosphopantetheine (PP) attached to the carrier domain PCP is shown as a wavy line.

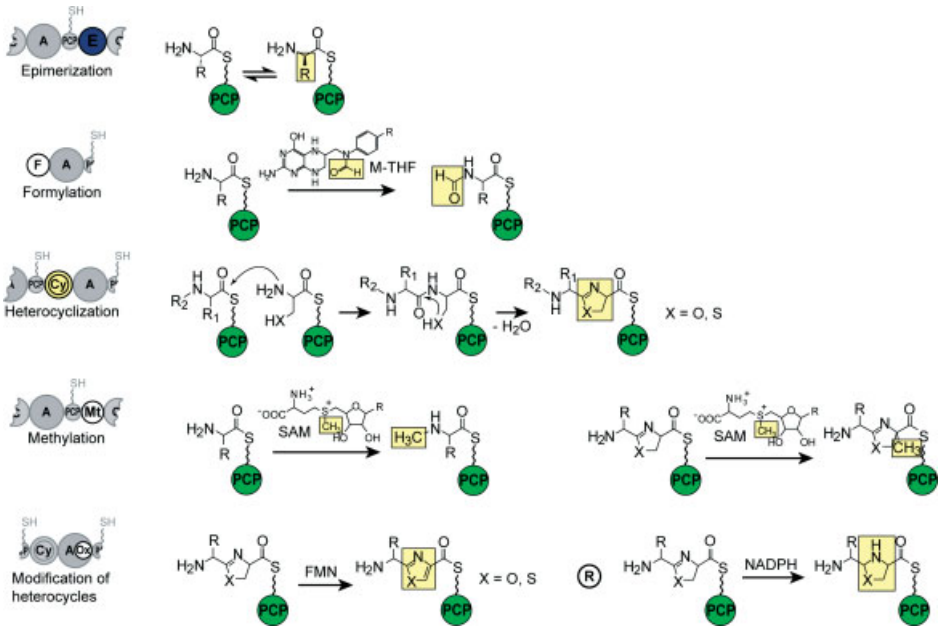

Figure 4. Reactions catalyzed by some NRPS modification domains that act in *cis* during the assembly line on amino acid substrates tethered to the carrier domain PCP. Yellow boxes show the sites of modification and cofactor donor groups. Integration sites of the modifying domains within the NRPS modules and the type of modification are indicated on the left side.

onto how these domains are connected and coordinated inside the intact functional modules [28,29]. In the following, a short description of these essential catalytic folds will be provided and aligned in the context of the recently solved structure of an entire termination module.

The Adenylation (A) Domain

Although A-domains and aminoacyl-tRNA-synthetases catalyze similar reactions (ATP-dependent amino acid activation and subsequent transfer to a suitable carrier molecule), they show structurally no evolutionary relationship to each other. However, A-domains show a significant structural similarity to several members of the adenylate-forming-superfamily. These include in addition to NRPS A-domains [30,31] the firefly luciferase [32], acetyl CoA synthetase [33], and yeast 4-chlorobenzoate: CoA ligase [34]. Despite the low sequence identity, essentially all show a similar fold. A large *N*-terminal subdomain and a small flexible *C*-terminal subdomain (~100 amino acids), with a hydrophobic active site in between. During the catalytic cycle (substrate and ATP binding followed by acyl-AMP formation) the small *C*-terminal subdomain

 Copyright © 2009 European Peptide Society and John Wiley & Sons, Ltd.

Journal of PeptideScience

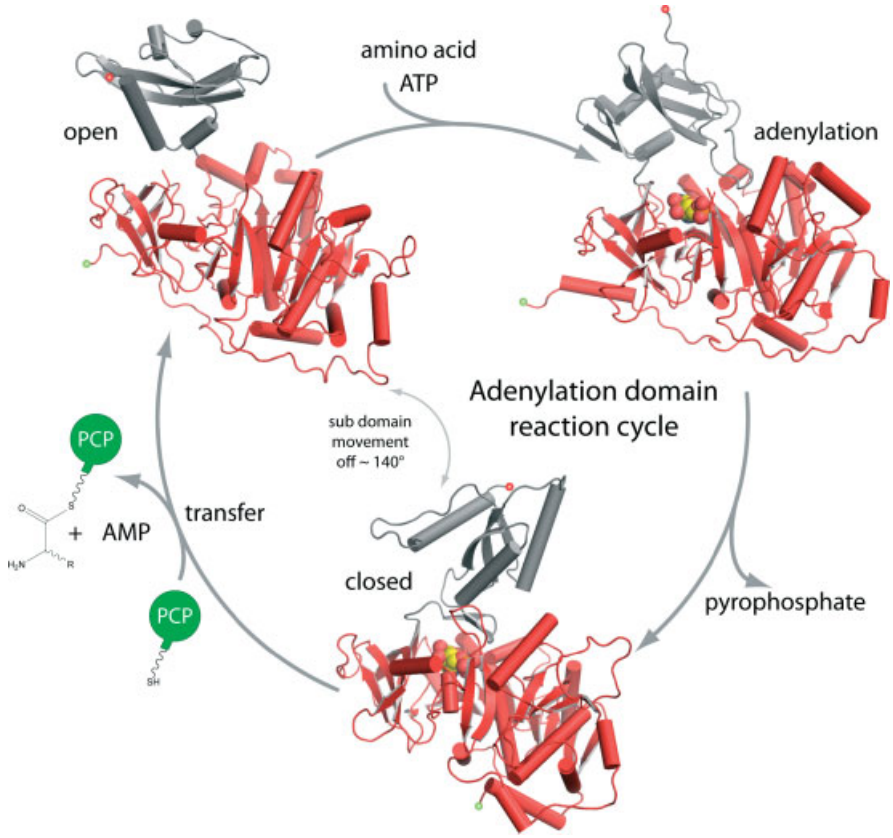

Figure 5. A proposed reaction cycle of an adenylation (A) domain showing the conformational movement of the small *C*-terminal subdomain (gray) upon ATP and substrate amino acid binding. During this initial reaction of the A-domain, aminoacyl-AMP formation and pyrophosphate release, the *C*-terminal subdomain rotates by some 140° from an open into a close conformation. In the second half reaction, through a specific PCP/A interaction, the activated substrate is then transferred to the reactive thiol group of the cofactor bound to the carrier domain PCP (green).

can adopt different orientations relative to the large *N*-terminal subdomain [35]. This motion ranges from 140° rotation in the open state during substrate loading to a closed state that promotes the adenylation reactions (Figure 5). Also, extensive biochemical studies on NRPS A-domains combined with structural and phylogenetic studies revealed the so-called "specificity-conferring code" of the A-domain [36]. This code constitutes 10 amino acid residues, all located in the region of the hydrophobic active site and can be easily identified and extracted out of the sequence of any A-domain and used for *in silico* studies to make a good prediction on the possible substrate without prior biochemical studies. This prediction was also successfully used in several cases for a rational switch of substrate-specificity of several A-domains [37,38].

The Carrier Domain PCP

The small (80–100 amino acids) carrier protein PCP is the site of cofactor binding. Its highly conserved serine residue is post-translationally modified by the 20 Å "swinging arm" 4′-phosphopantetheine (PPan) derived from CoA by the action of a specific PPan-transferase [39,40]. All selected substrate and synthetic intermediates of the NRPS machinery are covalently attached to the holo-carrier-proteins as acyl-S-PCPs (Figure 2). How this PCP molecule in time and space specifically interacts within the different NRPS catalytic units during assembly is unknown, but its newly discovered conformational switches may provide a clue [41]. PCP as other members of the carrier protein superfamily including acyl carrier proteins of fatty acid and polyketide synthases show in solution (NMR) a simple four helix bundle structure. A large loop region (10–15 amino acids) located between helix 1 and helix 2 contains the conserved serine residue to which the PPan-cofactor is attached. Recent high resolution NMR studies on apo-, holo-, and acyl-PCP revealed distinct chemical shifts, suggesting that PCP forms three different and slowly interconverting conformations, depending on its chemical modification states (Figure 6). Three distinct conformational states for PCP were identified: The A (apo), H (holo), and A/H conformers. The A-form and A/H form as well as the H- and A/H forms are slowly interconverting conformations (Figure 6). These conformational switches can either be frozen in the A-form by the exchange of the active site serine into alanine or in the H-form by acylation of holo-PCP. Further studies on PPan-cofactor locations in the frozen A/H and H-form indicate a relocation of the PPan-arm on the surface of PCP by some 100° which corresponds to a motion of 16 Å. This was the first indication that conformational transitions in PCP affect cofactor location and may be also the specific interaction of PCP with different domains within NRPS modules. NMR-titration studies of fully N^{15}-labelled apo- and holo-PCPs with two *in trans* acting proteins, TE II and PPan-transferase Sfp provided support for the idea that the PCP conformation state affects its partner-protein recognition [42,43].

 Copyright © 2009 European Peptide Society and John Wiley & Sons, Ltd.

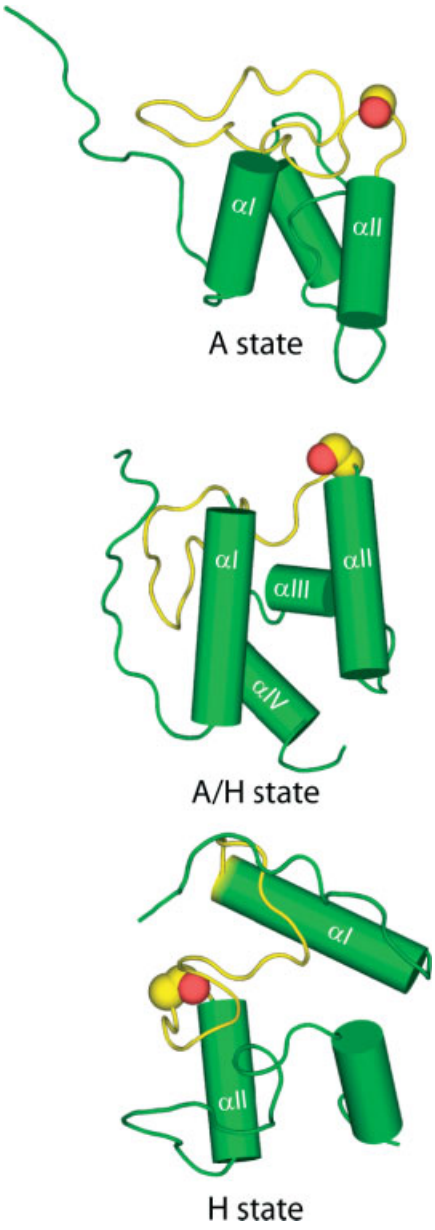

Figure 6. NMR structures of the PCP conformers (A, A/H, and H states). The A/H conformation has the most compact structure, but can also rotate into A or H state. Acylation of the cofactor 4′-phosphopantetheine attached to the serine residue (located within the yellow loop region connecting the helices αI and αII) fixes acyl-holo-PCP in the H state, whereas mutation of this serine residue in apo-PCP into alanine freezes holo-PCP in the A state. These conformations affect PCP partner protein recognition and the location of the acylated cofactor PPan during catalysis (see Ref. 41).

The Condensation (C) Domain

The condensation (C) domains are the control entities of NRPS elongation (C-A-PCP) modules as they catalyze the peptide bond formation between two adjacent PCP molecules loaded with their respective amino acids [44]. An aminoacyl-S-PCP activated amino acid is bound to the acceptor site of the C-domain, whereas its donor site accommodates the incoming peptidyl-S-PCP substrate. Studies with mischarged PCP carrier molecules in *in vitro* peptide condensation assays revealed for the C-domain a high substrate specificity at the acceptor (nucleophile) site and somehow relaxed specificity at its donor (electrophile) site [45]. These biochemical studies were supported by recent structural studies [46]. The overall architecture of the C-domain revealed a V-shaped form with a canyon-like structure in which the two PCP-bound substrates can be positioned from both sides (acceptor and donor sites) in close proximity to the active site histidine located in the conserved HHxxxDG core motif (Figure 7). A model supported by mutational studies suggests that the second histidine of this motif which is located at the floor of the canyon may act as the catalytic base to promote peptide bond formation. However, recent pK value analysis of this active site residue suggests that the peptide bond formation may depend mainly on electrostatic interactions rather than a general acid/base catalysis. The V-shape C-domain structure shows two similar *N*- and *C*-terminal subdomain folds that both belong to the well known chloramphenicol acetyltransferase (CAT) fold. In the C-domain structure the two CAT-folds build the canyon-like active site with the HHxxxDG-motif located on the middle of its floor. In the PCP-C-didomain structure (Figure 7), the PCP domain is attached to the donor site of the C-domain and has an A/H conformation. The flexible linker region between PCP and C is 18 residues in length and show little interaction with both folds. The relative arrangements of PCP and C in this structure place their active sites some 50 Å apart, suggesting a conformational transition prior to peptide bond transfer.

The Thioesterase (TE) Domain

Termination of peptide synthesis on an NRPS assembly line is commonly accomplished by a thioesterase (TE ~280 residues) domain. In contrast to the essential A, C, and PCP domains, which are present in each module, TE domains are only found in termination modules. They catalyze product release by a two step reaction mechanism: Transfer of the full length peptide chain attached to the terminal PCP to a highly conserved serine residue in the active site of TE by the formation of an acyl-O-TE intermediate that is subsequently cleaved by a regio- and stereoselective intramolecular macrocyclization using a peptide internal nucleophile [47,48]. Some TE-domains catalyze product release either by the generation of cyclic or branched-cyclic products (lactones and lactams), whereas others catalyze product release just by hydrolysis (e.g. vancomycin **4**). Other alternative release mechanisms can be achieved by reduction of the peptidyl-S-PCP final product to generate linear aldehyde or alcohol [49]. In such assembly line a *C*-terminal NADPH-dependent reductase (R) domain is attached to the terminal module. However, macrocyclization by TE seems to be the predominant mechanism for product release and regeneration of the NRPS.

Excised TE-domains are now commonly used catalysts for a cell-free synthesis of cyclic peptides when provided by thioester substrate mimics such as peptidyl-SNAC (*N*-acetylcysteamine) or peptidyl-thiophenol [19].

Several structures of dissected TE-domains have been solved, showing that all have the common fold of α/β-hydrolases [50,51], such as esterases and lipases, and all harbor in the active center a catalytic triad of serine, histidine, and aspartate (Figure 8). The serine within this triad is the site of tetrahedral intermediate formation that is stabilized by an oxyanion hole on the way to the acyl-O-TE intermediate. This intermediate breaks down by the nucleophilic attack of a peptide internal nucleophile. How the active site is sufficiently sealed from water to catalyze cyclization rather than hydrolysis is unknown. However, a lid

 Copyright © 2009 European Peptide Society and John Wiley & Sons, Ltd.

Journal of PeptideScience

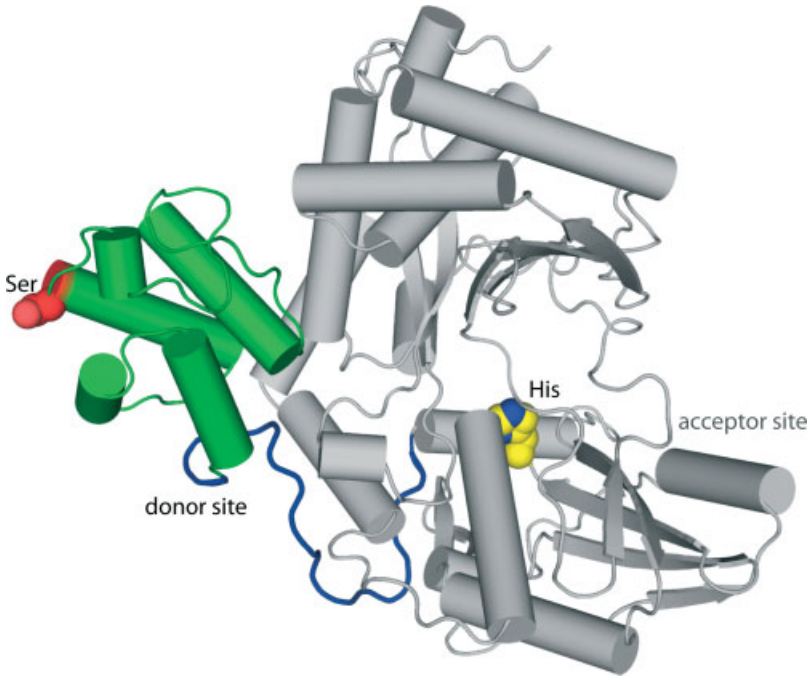

Figure 7. Structure of the PCP-C didomain from the surfactin synthetase. PCP in the A/H state (green) is connected through an 18 residues flexible linker region (blue) to the V-shaped C-domain (gray) at its donor site. The active site histidine located at the floor of the C-domain canyon and the active site serine of PCP (cofactor binding site) are some 50 Å apart (see Ref. 46).

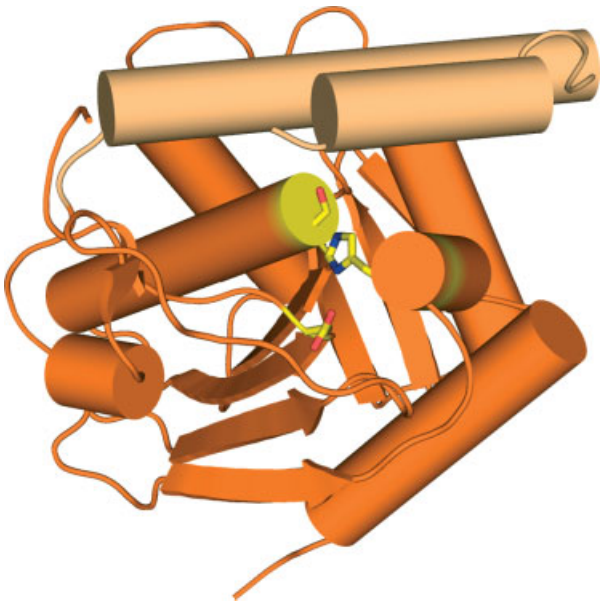

Figure 8. Structure of the dissected Srf-TE domain from the terminal module of the surfactin synthetase SrfA-C. Overall, TE shows a conserved α/β fold with a lid region (light brown) that covers the α/β core (gold) with active site catalytic triad (Ser, His, and Asp). The lid region is flexible and can adapt an open or a closed (shown) conformation, that may prevent water from entering the active site (see Ref. 50,51).

region consisting of a long helix protruding over the globular domain may adopt either an open conformation for substrate entry or a close conformation for excluding water from the active site. Nevertheless, the issue of substrate binding and efficient water exclusion can only be solved by elucidation of a suitable substrate/TE complex structure.

Insights into the Structure of an Entire Termination Module

The four domains comprising the SrfA-C termination module (C-A-PCP-TE; 1274 residues) of the surfactin synthetase catalyze the activation and incorporation of the terminal amino acid leucine and also the release of the final product as a cyclic lipopeptide lactone. A variant of SrfA-C in which the active site serine of PCP was mutated to alanine provided diffracting crystals. Such a mutation in PCP was shown by NMR studies to freeze PCP-conformational switches in the apo(A)-configuration that may contribute to the reduction of domains motion during crystallization. The structure of the SrfA-C variant was solved at 2.6 Å resolution and covered the entire module with its four catalytic domains and the linker regions in between [52]. The folds of the individual domains in SrfA-C module were found to be exactly the same as those solved before for the individually dissected A, C, PCP, and TE domains (Figure 9). However the SrfA-C structure provided unique information on domain–domain interactions within an NRPS module as well as on the nature of the linker regions and on how the essential catalytic domains are oriented and connected in space. Overall, the structural core of the module is a compact rectangular catalytic platform mainly built by the intimate association of the C-domain and the large *N*-terminal part of the A-domain. The two domains are "glued" together by extensive interactions at their interfaces and by the intervening linker region (32 residues) sandwiched in between. Both active sites of the A- and C-domains are arrayed on the same side of the platform. The *C*-terminal lid-region of the A-domain (~100 residues of the *C*-terminal part) and the PCP domains are tethered to each other on top of the platform and connected to the large *N*-terminal region of A by a flexible linker of 15 residues. In this arrangement, PCP and A-lid region can easily move relative to the static C–A-platform. A directed movement of the PCP-domain with its tethered 20 Å PPan-arm during the

 Copyright © 2009 European Peptide Society and John Wiley & Sons, Ltd.

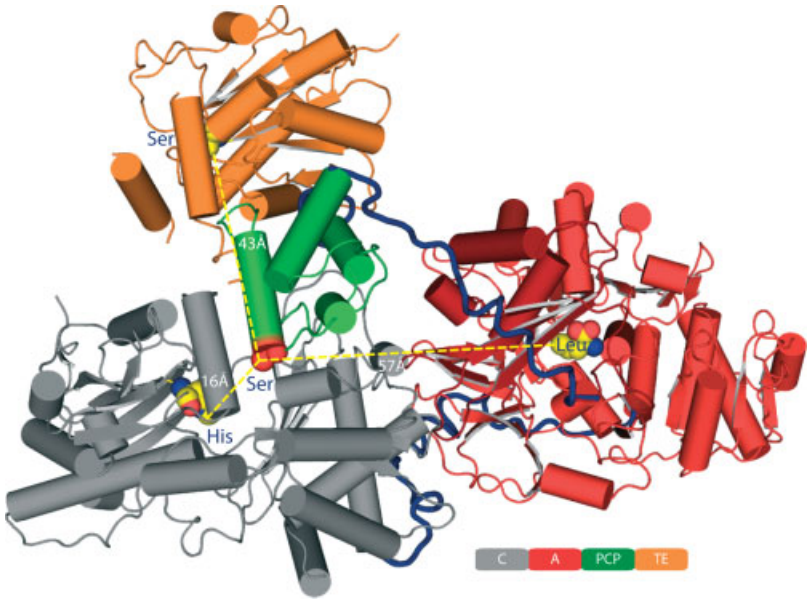

Figure 9. Overall structure of the termination module SrfA- (C-A-PCP-TE, 1274 residues) of the surfactin synthetase. The active site His of the C-domain (gray), the A-domain (red) substrate Leu and the Ser residue (cofactor binding site) of PCP (green) and that of TE (gold) are all shown in space-filling representation. The distances between them are shown in Angstroms and all linker regions between the four domains are indicated in blue (see Ref. 52).

catalytic cycle is clearly required, in order to reach the distant active sites of A, C, and TE, respectively. For example, the active sites of A and C are located some 60 Å apart and those of C and TE more than 40 Å. These distances cannot be bridged by the 20 Å PPan-arm alone without a substantial movement of the entire carrier domain PCP. The PCP domain in SrfA-C structure has a compact A/H configuration [53] that is stalled into the acceptor site of the C-domain. Its cofactor-binding site is located only 16 Å away from the active site histidine of the C-domain, indicating a productive interaction between the two domains. In contrast to the compact C–A-PCP domains association, the terminal TE-domain builds an independent fold connected through a short (9 residues) linker to PCP and shows essentially an independent α/β-fold, exactly identical to that of the dissected TE-domains.

What are the lessons that can be learned out of SrfA-C structural studies concerning reprogramming of NRPS-templates? Certainly the static platform generated by the C–A strong association in each NRPS elongation module suggests that they should be exchanged as a functional pair. This strategy also considers the high substrate specificity at the acceptor site at the C-domain. Also in each module or domain swapping experiment, the PCP interactions with C and A should be optimized.

Acknowledgements

Work in my laboratory was supported by the Deutsche Forschungsgemeinschaft, the EC and Fonds der Chemischen Industrie.

References

1 Marahiel MA. Protein templates for the biosynthesis of peptide antibiotics. *Chem. Biol.* 1997; **4**: 561–567.
2 Marahiel MA, Stachelhaus T, Mootz HD. Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.* 1997; **7**: 2651–2674.
3 Fischbach MA, Walsh CT. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev.* 2006; **106**: 3468–3496.
4 Grünewald J, Marahiel MA. Chemoenzymatic and template-directed synthesis of bioactive macrocyclic peptides. *Microbiol. Mol. Biol. Rev.* 2006; **70**: 121–146.
5 Finking R, Marahiel MA. Biosynthesis of nonribosomal peptides. *Annu. Rev. Microbiol.* 2004; **58**: 453–488.
6 Konz D, Marahiel MA. How do peptide synthetases generate structural diversity? *Chem. Biol.* 1999; **2**: 39–48.
7 Sieber SA, Marahiel MA. Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. *Chem. Rev.* 2005; **105**: 715–738.
8 Caboche S, Pupin M, Leclère V, Fontaine A, Jacques P, Kucherov G. NORINE: a database of nonribosomal peptides. *Nucl. Acids Res.* 2008; **36**: 326.
9 Konz1 D, Klens A, Schörgendorfer K, Marahiel MA. The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multi-modular peptide synthetases. *Chem. Biol.* 1997; **12**: 927–937.
10 Mootz HD, Schwarzer D, Marahiel MA. Ways of assembling complex natural products on modular nonribosomal peptide synthetases. *Chembiochem.* 2002; **6**: 490–504.
11 Schwarzer D, Finking R, Marahiel MA. Chembiochem. Nonribosomal peptides: from genes to products. *Nat. Prod. Rep.* 2003; **20**: 275–287.
12 Felnagle EA, Jackson EE, Chan YA, Podevels AM, Berti AD, McMahon MD, Thomas MG. Nonribosomal peptide synthetases involved in the production of medically relevant natural products. *Mol. Pharm.* 2008; **5**: 191–211.
13 Mootz HD, Marahiel MA. The tyrocidine biosynthesis operon of *Bacillus brevis*: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J. Bacteriol.* 1997; **21**: 6843–6850.
14 Stachelhaus T, Hüser A, Marahiel MA. Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases. *Chem. Biol.* 1996; **11**: 913–921.
15 Stachelhaus T, Mootz HD, Bergendahl V, Marahiel MA. Peptide bond formation in nonribosomal peptide biosynthesis. Catalytic role of the condensation domain. *J. Biol. Chem.* 1998; **35**: 22773–22781.
16 Trauger JW, Kohli RM, Mootz HD, Marahiel MA, Walsh CT. Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature* 2000; **6801**: 215–218.
17 Kohli RM, Trauger JW, Schwarzer D, Marahiel MA, Walsh CT. Generality of peptide cyclization catalyzed by isolated thioesterase domains of nonribosomal peptide synthetases. *Biochemistry* 2001; **24**: 7099–7108.

 Copyright © 2009 European Peptide Society and John Wiley & Sons, Ltd.

18 Kohli RM, Walsh CT. Enzymology of acyl chain macrocyclization in natural product biosynthesis. *Chem. Commun. (Camb)* 2003; **3**: 297–307.
19 Kopp F, Marahiel MA. Macrocyclization strategies in polyketide and nonribosomal peptide biosynthesis. *Nat. Prod. Rep.* 2007; **4**: 735–749.
20 Linne U, Marahiel MA. Reactions catalyzed by mature and recombinant nonribosomal peptide synthetases. *Methods Enzymol.* 2004; **388**: 293–315.
21 Walsh CT. The chemical versatility of natural-product assembly lines. *Acc. Chem. Res.* 2008; **41**: 4–10.
22 Zerbe K, Woithe K, Li DB, Vitali F, Bigler L, Robinson JA. An oxidative phenol coupling reaction catalyzed by oxyB, a cytochrome P450 from the vancomycin-producing microorganism. *Angew. Chem. Int. Ed. Engl.* 2004; **43**: 6709–6713.
23 Woithe K, Geib N, Zerbe K, Li DB, Heck M, Fournier-Rousset S, Meyer O, Vitali F, Matoba N, Abou-Hadeed K, Robinson JA. Oxidative phenol coupling reactions catalyzed by OxyB: a cytochrome P450 from the vancomycin producing organism. Implications for vancomycin biosynthesis. *J. Am. Chem. Soc.* 2007; **129**: 6887–6895.
24 Losey HC, Jiang J, Biggins JB, Oberthur M, Ye XY, Dong SD, Kahne D, Thorson JS, Walsh CT. Incorporation of glucose analogs by GtfE and GtfD from the vancomycin biosynthetic pathway to generate variant glycopeptides. *Chem. Biol.* 2002; **9**: 1305–1314.
25 Losey HC, Peczuh MW, Chen Z, Eggert US, Dong SD, Pelczer I, Kahne D, Walsh CT. Tandem action of glycosyltransferases in the maturation of vancomycin and teicoplanin aglycones: novel glycopeptides. *Biochemistry* 2001; **40**: 4745–4755.
26 Vaillancourt FH, Yeh E, Vosburg DA, Garneau-Tsodikova S, Walsh CT. Nature's inventory of halogenation catalysts: oxidative strategies predominate. *Chem. Rev.* 2006; **8**: 3364–3378.
27 Nolan EM, Walsh CT. How nature morphs peptide scaffolds into antibiotics. *Chembiochem* 2009; **1**: 34–53.
28 Marahiel MA, Essen LO. Nonribosomal peptide synthetases: mechanistic and structural aspects of essential domains. *Methods Enzymol.* 2009; **458**: 337–351.
29 Koglin A, Walsh CT. Structural insights into nonribosomal peptide enzymatic assembly line. *Nat. Prod. Rep.* 2009; **8**: 987–1000.
30 Conti E, Stachelhaus T, Marahiel MA, Brick P. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J.* 1997; **14**: 4174–4183.
31 May JJ, Kessler N, Marahiel MA, Stubbs MT. Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases. *Proc. Natl. Acad. Sci. U.S.A.* 2002; **9**: 12120–12125.
32 Nakatsu T, Ichiyama S, Hiratake J, Saldanha A, Kobashi N, Sakata K, Kato H. Structural basis for the spectral difference in luciferase bioluminescence. *Nature* 2006; **7082**: 372–376.
33 Jogl G, Tong L. Crystal structure of yeast acetyl-coenzyme A synthetase in complex with AMP. *Biochemistry* 2004; **6**: 1425–1431.
34 Gulick AM, Lu X, Dunaway-Mariano D. Crystal structure of 4-chlorobenzoate : CoA ligase/synthetase in the unliganded and aryl substrate-bound states. *Biochemistry* 2004; **27**: 8670–8679.
35 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. *J. Biol. Chem.* 2008; **283**: 32484–32491.
36 Stachelhaus T, Mootz HD, Marahiel MA. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* 1999; **6**: 493–505.
37 Stachelhaus T, Schneider A, Marahiel MA. Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. *Science* 1995; **269**: 69–72.
38 Eppelmann K, Doekel S, Marahiel MA. Engineered biosynthesis of the peptide antibiotic bacitracin in the surrogate host *Bacillus subtilis*. *J. Biol. Chem.* 2001; **276**: 34824–34831.
39 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. *EMBO J.* 1999; **18**: 6823–6831.
40 Mootz HD, Finking R, Marahiel MA. 4′-phosphopantetheine transfer in primary and secondary metabolism of *Bacillus subtilis*. *J. Biol. Chem.* 2001; **40**: 37289–37298.
41 Koglin A, Mofid MR, Löhr F, Schäfer B, Rogov VV, Blum MM, Mittag T, Marahiel MA, Bernhard F, Dötsch V. Conformational switches modulate protein interactions in peptide antibiotic synthetases. *Science* 2006; **312**: 273–276.
42 Lai JR, Fischbach MA, Liu DR, Walsh CT. Localized protein interaction surfaces on the EntB carrier protein revealed by combinatorial mutagenesis and selection. *J. Am. Chem. Soc.* 2006; **34**: 11002–11003.
43 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; **28**: 8946–8956.
44 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**: 10439–10447.
45 Belshaw PJ, Walsh CT, Stachelhaus T. Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science* 1999; **284**: 486–489.
46 Samel SA, Schönafinger 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**: 781–792.
47 Kopp F, Marahiel MA. Macrocyclization strategies in polyketide and nonribosomal peptide biosynthesis. *Nat. Prod. Rep.* 2007; **24**: 735–749.
48 Trauger JW, Kohli RM, Mootz HD, Marahiel MA, Walsh CT. Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature* 2000; **407**: 215.
49 Kopp F, Marahiel MA. Where chemistry meets biology: the chemoenzymatic synthesis of nonribosomal peptides and polyketides. *Curr. Opin. Biotechnol.* 2007; **18**: 513–520.
50 Bruner SD, Weber T, Kohli RM, Schwarzer D, Marahiel MA, Walsh CT, Stubbs MT. Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE. *Structure* 2002; **10**: 301–310.
51 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. *J. Mol. Biol.* 2006; **359**: 876–889.
52 Tanovic A, Samel SA, Essen LO, Marahiel MA. Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science* 2008; **322**: 659–663.
53 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**: 407–418.

 Copyright © 2009 European Peptide Society and John Wiley & Sons, Ltd.