Molecular Mechanisms Underlying Nonribosomal Peptide Synthesis: Approaches to New Antibiotics

Stephan A. Sieber and Mohamed A. Marahiel*

Philipps-Universität Marburg, Fachbereich Chemie/Biochemie, Hans-Meerwein-Strasse, 35032 Marburg, Germany

Received March 30, 2004

Contents

1. Introduction

1.1. NRPS Synthesis

Research into bioactive natural products began when A. Fleming discovered the antibiotic activity of the peptide derivative penicillin produced by the fungal host organism *Penicillium notatum.*[1] Since then microorganisms have attracted considerable attention as a new source for pharmaceutical agents, and screening of microbial extracts has afforded a very large number of new compounds with antimicrobial, antiviral, immunosuppressive, and antitumor activities. These secondary metabolites were optimized for their dedicated function over eons of evolution and now represent promising scaffolds for the development of new drug leads.

Among these substances small peptide molecules represent a large subclass of bioactive natural products, which contain unique structural features such as heterocyclic elements, D-amino acids, and glycosylated as well as N-methylated residues. Moreover, in contrast to proteins produced by ribosomal synthesis, small peptide products contain not only the common 20 amino acids but also hundreds of different building blocks, suggesting a nonribosomal origin of biosynthesis. In the 1970s Lipmann et al. reported a nucleic-acid-independent synthesis of the peptide antibiotics gramicidin S and tyrocidine A from *Bacillus* sp. by large enzyme complexes similar to fatty acid synthases.[2] In the following years more and more peptidic natural products were shown to be assembled by such large enzymes, referred to as nonribosomal peptide synthetases (NRPS). Significant progress has been made in the past decades toward understanding the molecular principles of bioactive peptide synthesis in microorganisms, and this has been the subject of extensive recent reviews.[3−6] While research first focused on elucidating the chemical structure of the assembled molecules and characterizing the architecture of the associated multienzymes, later work focused on the identification of their biosynthetic gene clusters. More recent biochemical investigations also revealed high-resolution structures of some of the central catalytic core enzymes. Moreover, genetic and chemoenzymatic approaches were developed to reprogram natural peptide sequences by the combined action of rational enzyme design and chemical peptide synthesis followed by subsequent enzyme catalysis. This review aims to give a global overview of our understanding of natural nonribosomal peptide synthesis and of progress in genetically engineered and chemoenzymatic synthesis of nonribosomal peptide products. In the natural synthesis section major emphasis will be given to recent progress made on structure determination and mechanistic predictions. The genetic and chemoenzymatic section will complement this section by providing approaches to novel peptide antibiotics.

* To whom correspondence should be addressed. Phone: +49-6421-2825722. Fax: +49-6421-2822191. E-mail: marahiel@chemie.uni-marburg.de.

10.1021/cr0301191 CCC: $53.50 © 2005 American Chemical Society
Published on Web 01/20/2005

Stephan A. Sieber was born in Marburg, Germany, in 1976. He studied Chemistry at the Universities of Marburg and Birmingham, England. Fascinated by the great potential of combining enzymatic and chemical synthesis, he joined the group of Professor Mohamed A. Marahiel, where he obtained his diploma in 2001. During his graduate studies in the lab of Professors Marahiel and Christopher T. Walsh at Harvard Medical School, Boston, MA, he worked on a chemoenzymatic approach to cyclic peptides. In 2004 he obtained his Ph.D. degree and received the Friedrich-Weygand award for his thesis work. During his undergraduate and graduate studies he was a fellow of the Studienstiftung des Deutschen Volkes. Soon after graduation he joined the group of Professor Benjamin F. Cravatt at The Scripps Research Institute in La Jolla, CA, for postdoctoral work for which he received a DFG Emmy-Noether Stipend.

Mohamed A. Marahiel studied Chemistry at the Universities of Cairo (Egypt) and Göttingen (Germany). In 1977 he obtained his Ph.D. degree in Biochemistry and Microbiology from the University of Göttingen. 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-Universität 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 (U.K.) and at the Biolabs, Harvard University, respectively. Since 2004 he has been a member of the "Leopoldina"—Deutsche Akademie der Naturforscher. His present research focuses on the structure–function relationship and elucidation of reaction mechanisms of modular peptide synthetases involved in the nonribosomal synthesis of peptide antibiotics. His group is also interested in studying the cold shock response in soil bacteria as well as other stress-induced proteins.

1.2. Product Diversity Assisted by NRPS

Natural peptide products synthesized by NRPSs can be grouped according to their biological activities. A major class comprises antibiotic and antifungal activities including, for example, the peptides tyrocidine, bacitracin, surfactin, pristinamycin, vancomycin, and fengycin.[3,7−15] Their biological functionality is strictly associated with their chemical structure, which constrains the peptide sequence in its biologically active conformation and ensures specific interaction with a dedicated molecular target. This structural rigidity is achieved by either cyclization or oxidative cross-linking of side chains, which contribute to stability. Moreover, the great diversity of chemical modifications, such as incorporation of fatty acid chains, D-amino acids, glycosylated amino acids, and heterocyclic rings, adds much to these specific interactions (Figure 1).

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

Cyclosporin produced by *Tolypocladium niveum* exhibits immunosuppressive and toxic properties due the formation of a specific complex with cyclophilin which inhibits, in turn, the protein phosphatase calcineurin, responsible for T-cell activation.[32−34]

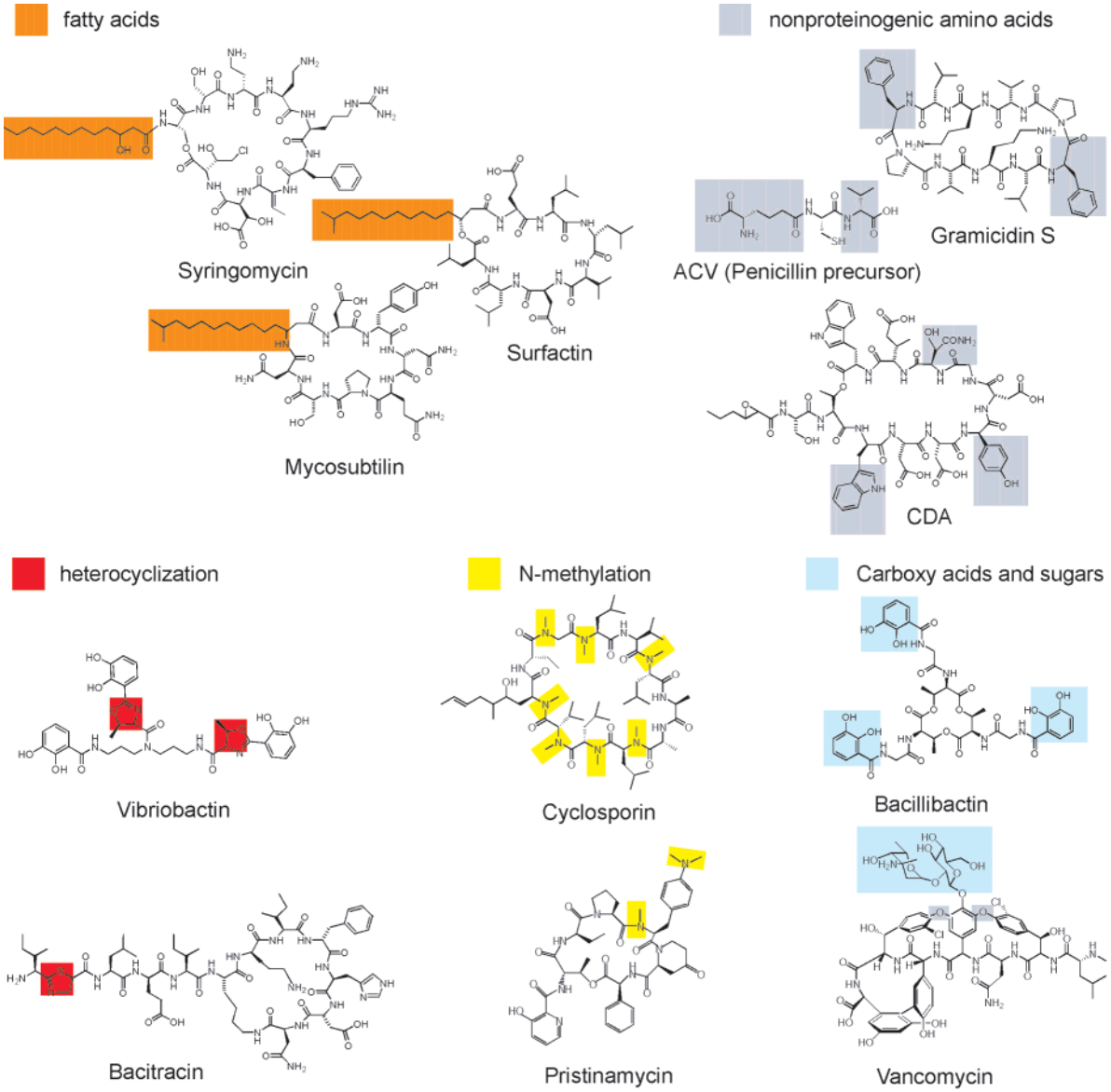

Figure 1. Natural peptidic products. A selection of nonribosomally synthesized peptides. Characteristic structural features are highlighted.

Cyclosporin is highly lipophilic, and 7 of its 11 amino acids are N-methylated. This high degree of methylation protects the peptide from proteolytic digestion but complicates chemical synthesis due to low coupling yields and side reactions.[35] In an iron-deficient environment some bacteria such as *E. coli, B. subtilis,* and *Vibrio cholerae* synthesize and secrete iron-chelating molecules known as siderophores that scavenge Fe^{3+} with picomolar affinity, important for host survival.[36,37] Three catechol ligands derived from 2,3-dihydroxybenzoyl (DHB) building blocks in bacillibactin, enterobactin, and vibriobactin complex iron by forming intramolecular octahedra.

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

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

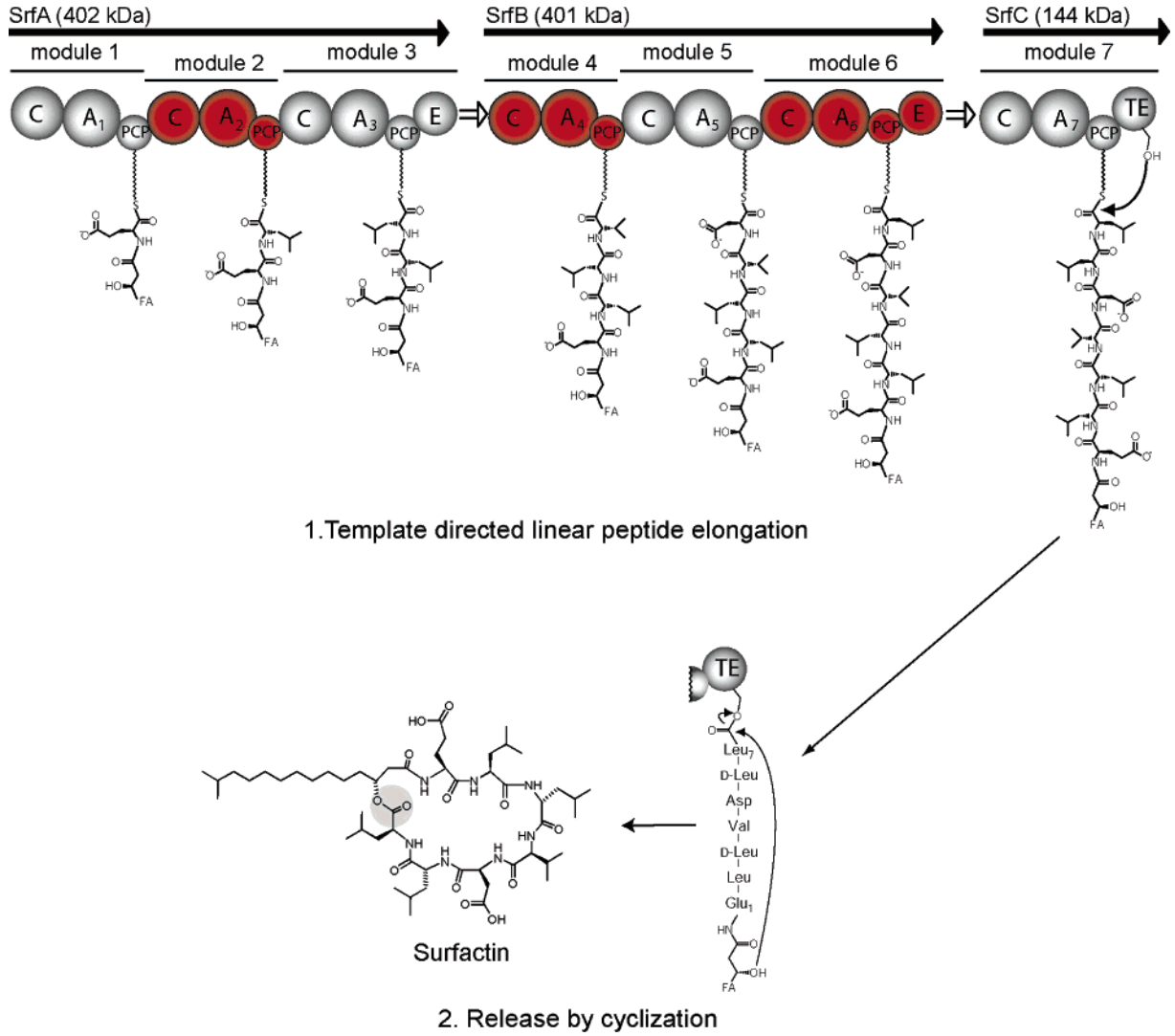

Figure 2. Surfactin assembly line. The multienzyme complex consists of seven modules (grey and red) which are specific for the incorporation of seven amino acids. Twenty-four domains of five different types (C, A, PCP, E, and TE) are responsible for the catalysis of 24 chemical reactions. Twenty-three reactions are required for peptide elongation, while the last domain is unique and required for peptide release by cyclization.

2. NRPS Factory

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

The carboxy group of amino acid building blocks is first activated by ATP hydrolysis to afford the corresponding aminoacyl−adenylate. This reactive intermediate is transferred onto the free thiol group of an enzyme-bound 4′-phosphopantetheinyl cofactor (ppan), establishing a covalent linkage between enzyme and substrate. At this stage the substrate can undergo modifications such as epimerization or N-methylation. Assembly of the final product then occurs by a series of peptide bond formation steps (elongation) between the downstream building block with its free amine and the carboxy−thioester of the upstream substrate. The ppan cofactor facilitates the ordered transfer of thioester substrates between catalytically active units with all intermediates covalently tethered to the multienzyme until the product is released by the action of the C-terminal thioesterase (TE) domain (termination). This strategy minimizes side reactions as well as diffusion times. Type I polyketide synthases (PKS) and fatty acid synthases (FAS) similarly display a multienzymatic

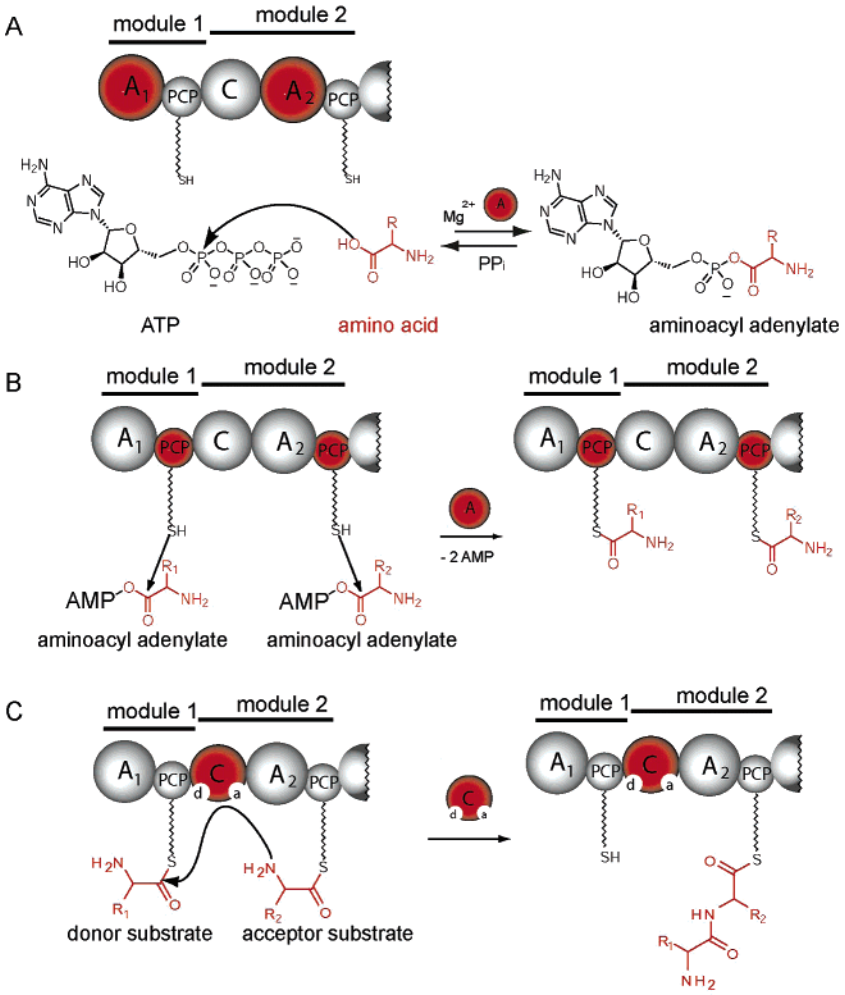

Figure 3. Domain-catalyzed reactions. Domains in action are indicated in red. (A) Recognition and activation of a dedicated amino acid with ATP by the A domain. (B) Covalent attachment of the activated aminoacyl adenylate onto the free thiol group of the PCP-bound ppan cofactor. (C) Peptide elongation by the C domain which catalyzes an attack of the nucleophilic amine of the acceptor substrate onto the electrophilic thioester of the donor substrate.

organization and catalyze a repetitive reaction cycle involving decarboxylative condensation of smaller acyl groups.[43,44]

2.1. Activation by the Adenylation Domain

Each nonribosomal peptide synthesis is initiated by specific recognition and activation of the relevant dedicated amino acid from a pool of substrates by the ca. 550 amino acid long adenylation domain (A domain). For example, each amino acid found in the heptapeptide surfactin is directly selected for incorporation into the growing peptide chain by one of the seven A domains of the surfactin synthetase (Figure 2). Substrate activation is achieved in a two-step chemical reaction. First, after binding of the cognate amino acid, the enzyme catalyzes the formation of an aminoacyl adenylate intermediate at the expense of Mg^{2+}−ATP and release of PP_i (Figure 3A). Second, the amino acid−O−AMP oxoester is converted into a thioester by a nucleophilic attack of the free thiol−ppan cofactor of the adjacent PCP domain, which will be discussed below (Figure 3B). This mechanism resembles amino acid activation catalyzed by aminoacyl−tRNA synthetases, although these enzyme families share neither sequence nor structural similarity.[45] Many A domains can be heterologously expressed in *E. coli,* and their activity as well as substrate specificity can be assayed in vitro by an equilibrium ATP−PP_i exchange assay with radiolabeled PP_i.[46]

Sequence alignments of A domains revealed early on an adenylate (AMP) binding motif that is conserved in a superfamily of so-called adenylate-forming enzymes, which include 4-coumerate−CoA ligases, acetyl−CoA synthetases, and oxidoreductases.[47] Several crystal structures of members of this family have been solved. These include the oxidoreductase luciferase from *Photinus pyralis*,[48] the acetyl−CoA synthetase (Acs) of primary metabolism,[49] the phenylalanine-activating A domain (PheA) of the first module of gramicidin S synthetase of *B. brevis,*[50] and the 2,3-dihydroxybenzoate (DHB) activating A domain (DhbE) of *B. subtilis*.[51] All crystal structures possess low sequence identity but exhibit an almost identical fold. They consist of a large N-terminal subdomain and a small C-terminal subdomain (the term subdomain means a stable tertiary fold within the structure) with the active site at the junction between them. All NRPS A domains share ca. 30−60% sequence identity,[47] which allowed identifi-

cation of 10 "core motifs" that serve as functional anchors. In combination with the A-domain crystal structures it could be shown that many amino acid residues of these motifs are responsible for, i.e., ATP binding, hydrolysis, and adenylation of the substrate carboxy moiety.[52] Cocrystallization of PheA with phenylalanine and AMP revealed specific interactions of the AMP phosphate moiety with a conserved Lys517 (Figure 4A). The substrate binding site has

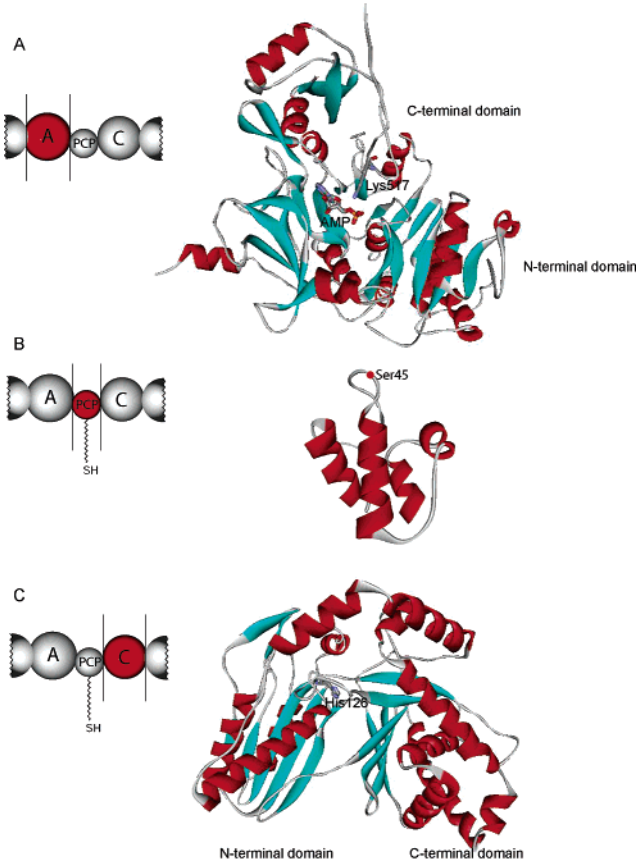

Figure 4. Crystal structures of catalytic core domains. (A) Crystal structure of the phenylalanine-activating A domain (PheA) of the first module of gramicidin S synthetase of *B. brevis*. (B) Structure of the PCP domain derived from the third module of the *B. brevis* tyrocidine synthetase. (C) Crystal structure of VibH, a stand alone C domain of the *V. cholerae* vibriobactin synthetase.

a channel-like entrance, and phenylalanine is bound in a hydrophobic pocket with the carboxy group interacting with Lys517 and the α-amino group with Asp235. Mutation of these residues confirmed their relevance for catalysis.[53,54] Nature uses the important catalytic function of the conserved lysine residue in related acetyl−CoA synthetase (Acs) for regulation of acetyl−CoA production in primary metabolism.[55] Similar to A domains, Acs synthesizes acetyl−CoA from acetate, ATP, and CoA through an acetyl−AMP intermediate. This synthesis is regulated by post-translational acylation of Lys609, which blocks the formation of the adenylate intermediate. Reactivation of the acetylated enzyme requires the NAD-dependent protein deacetylase activity of a CobB Sir2 protein type as shown recently in *S. enterica*.[55] This mode of regulation was speculated to modulate activity for all adenylate-forming enzymes including NRPS A domains.

The large diversity of building blocks found in nonribosomally produced peptides corresponds to the variations in structure and sequence of A-domain binding pockets as seen for the aryl-acid-activating A-domain DhbE.[51] The substrate pocket is shallower than that in PheA, and 2,3-dihydroxybenzoic acid is coordinated by its two hydroxy groups, which interact with Ser240 and Asn235, and by the carboxy moiety, which interacts with two residues Lys517 and His234. This binding mode contrasts with the situation in PheA where a second interaction is provided to the α-amino group and reflects differences in proteinogenic vs nonproteinogenic amino acid activation. By comparing the residues lining the binding pockets of PheA and DhbE with the corresponding residues in other A domains, general rules for substrate specificity were developed. These rules, also referred to as the "nonribosomal code", were tested by mutations in specificity conferring residues in different A domains, which lead to the conclusion that substrate specificity can be predicted with fairly high accuracy.[52,56−58]

The first crystal structures of adenylate-forming enzymes primarily provided insight into the adenylation half reaction, while a recent crystal structure of Acs bound to adenosine-5′-propyl phosphate and coenzyme A (CoA) has also shed light on the second reaction step, thioester formation with coenzyme A.[49] This is particularly interesting since aminoacyl−AMP transfer onto the ppan cofactor of NRPS has largely been the subject of speculation until now.[6,50,51] It was observed that the C-terminal domain of Acs is rotated by ca. 140 ° relative to its conformation in DhbE and PheA. This domain movement exposes a new set of residues to the active site and relocates Lys609, important for the adenylation half reaction, as it coordinates the carbonyl group of acetate. This conclusion was also supported by experiments in which acetylation of Lys609 was shown to inhibit adenylation but had no effect on catalysis of the second half reaction, the acylation of CoA. This result suggests the existence of two sets of catalytic residues corresponding to two different reaction steps and to two different structural orientations.[49,55] Moreover, the structural rearrangement appropriately positions the CoA thiol for nucleophilic attack on the acetyl−AMP intermediate. Since crystal structures only provide snapshots of individual states in a dynamic multiple-step process, the findings suggest that members of the adenylate-forming family of enzymes adopt two different orientations to catalyze their two half reactions.

In contrast to ribosomal protein synthesis, which shows very accurate proof reading, less stringent substrate selection and incorporation is observed for some nonribosomal peptides.[52] However, new studies on the mode of substrate selection of NRPS A domains have revealed an intrinsic ATPase activity which is enhanced in the presence of noncognate amino acid substrates.[59] In turn, less pronounced variations in ATPase activity are observed in A domains with relaxed amino acid specificity. The cyclic

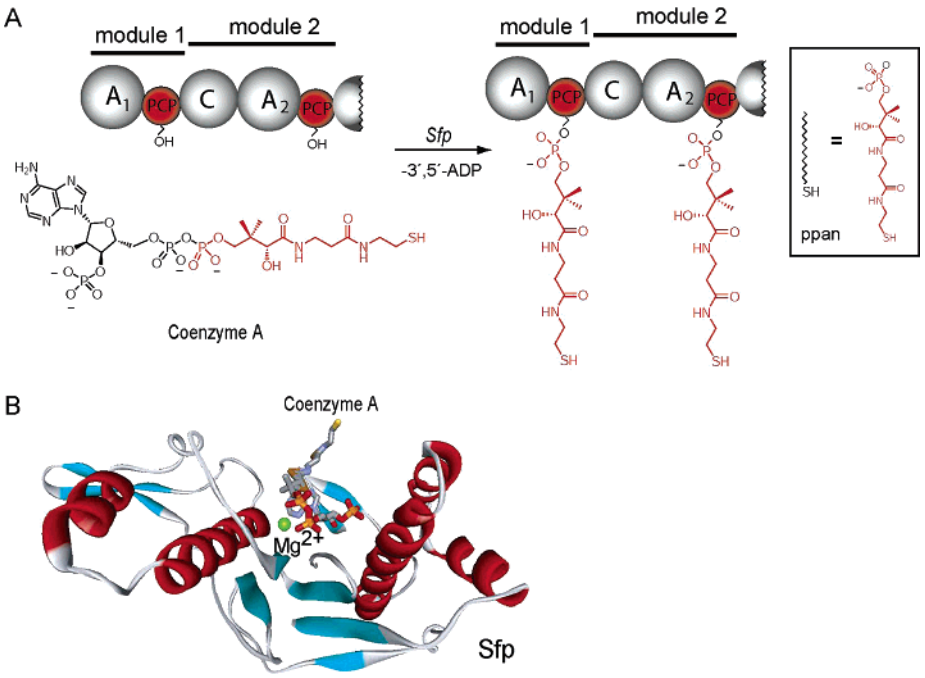

Figure 5. Phosphopantetheinylation: Apo to holo enzyme conversion. (A) The phosphopantetheine moiety of coenzyme A (red) is covalently attached to an invariant serine residue of PCP by Sfp, a dedicated phosphopantheteine transferase. (B) Crystal structure of Sfp with its substrate coenzyme A and Mg^{2+}.

decapeptide tyrocidine consists of a mixture of four compounds that vary in two positions,[60] whereas about 30 variants are known for the immunosuppressive drug cyclosporin.[32] With this simple strategy the diversity of natural products can be readily increased.

2.2. Intermediates Transport by the Peptidyl Carrier Protein

The only NRPS domain without autonomous catalytic activity is the peptidyl carrier protein (PCP), also referred to as thiolation domain (T). The protein comprises ca. 100 amino acids and is located downstream of the A domain. Within the NRPS assembly line PCP is responsible for transportation of substrates and elongation intermediates to the catalytic centers. As discussed in the previous section, the A domain catalyzes the transfer of the activated aminoacyl−adenylate substrate onto the terminal cysteamine thiol group of the ppan cofactor bound to PCP in its second half reaction (Figure 3B). In nonribosomal peptide synthesis the combination of A domain and PCP is defined as an initiation module since both domains are required to activate and covalently tether the first building block for subsequent peptide synthesis. The activity of recombinant A domains with adjacent holo-PCPs can be assayed in vitro by an aminoacylation assay with ATP and radioactive amino acids.[46] In contrast to ribosomal protein synthesis with tRNA-bound ester intermediates, nonribosomal peptide synthetases exploit more reactive PCP−thioesters. This difference in reactivity is due to the lower mesomeric stabilization of the thioester, which forms less stable p_π=p_π double bonds and emphasizes that the initial activation energy provided by the A domain is preserved here for subsequent catalytic reactions such as condensation, hydrolysis, and cyclization (see below).[61]

A 20 Å long prosthetic ppan moiety of coenzyme A is covalently tethered to the side chain of a strictly conserved PCP serine residue and serves as a crane for building-block delivery.[62−65] Transfer of ppan onto apo-PCP is catalyzed by NRPS-specialized 4′-phosphopantetheinyl transferases such as Sfp and Gsp from *B. subtilis* and *B. brevis*, respectively[53,66,67] (Figure 5A). The conversion of inactive apo-PCP into its active ppan−PCP holo form was monitored in vitro with recombinant Sfp and PCPs from the surfactin synthetase. These studies revealed the very low selectivity of Sfp for the carrier proteins.[68] Sfp was shown to efficiently phosphopantetheinylate not only apo-PCPs from various NRPS templates but also acyl carrier proteins from fatty acid and polyketide synthases.[66,69] Furthermore, recent results suggest a broad acyl−CoA tolerance. Various synthetic peptidyl−CoAs were covalently attached to apo-PCPs under catalysis of Sfp, albeit with reduced efficiency. This relaxed specificity of Sfp has proved useful for preparative applications, which will be presented below.[70,71] Insights into how Sfp mediates binding and protein recognition were provided by a crystal structure in complex with its substrate CoA[72] (Figure 5B). The structure of the 224 amino acid comprising Sfp monomer shows a pseudo-2-fold symmetry which divides the protein into two similar folds of almost identical size. The CoA substrate is bound in a bent conformation within a pocket formed by the two Sfp halves. The 3′-phospho-5′-ADP moiety of CoA is well defined in the electron density map and coordinated by several Sfp residues and Mg^{2+}, while the main part of the ppan arm shows no interactions with Sfp and points out into bulk solvent. This unique type of CoA coordination is in complete agreement with the observed binding tolerance for peptidyl−CoA substrates (and acyl−CoA) in which the peptide compo-

nent presumably does not interact with the enzyme (section 3.2.3).[68] PCP cofactor modification was suggested to occur via Sfp−Glu151-mediated deprotonation of the serine hydroxy group of PCP and its subsequent nucleophilic attack on the β-phosphate of CoA, leading to holo-PCP and 3′,5′-ADP as a byproduct. Several mutational analyses of Sfp confirmed this type of CoA binding and revealed several residues in a loop region between β4−α5 in the PCP binding region.[73]

The first solution structure of a PCP was solved using NMR spectroscopy with the PCP of the third module of the *B. brevis* tyrocidine synthetase[74] (Figure 4B). PCP is a distorted four-helix bundle with an extended loop between the first two helices, which is probably important for interaction with Sfp. The invariant serine residue, the site of cofactor binding, is located at the interface between this loop and the second helix. The cofactor shows no interactions with the protein and is accommodated in the solvent. No peptide-binding pocket was observed, which is in agreement with the lack of substrate selectivity. The fold is well defined between residues 8 and 82, and the structural core was defined to be a region spanning 37 amino acids in both directions from the conserved serine. PCPs have a function similar to acyl carrier proteins (ACP) from fatty acid and polyketide synthases. Sequence homologies between PCPs and ACPs only exist in the immediate neighborhood of the invariant serine residue, although all proteins possess an almost identical structural fold.[75−77] The most obvious difference between ACPs and PCPs is the overall charge of the proteins. While ACPs have predominantly acidic side chains on their surface, the PCPs surface is much less polar. This corresponds to the charges of the corresponding ppan−transferases AcpS and Sfp. While AcpS exclusively primes ACP and not PCP, it was observed that Sfp is promiscuous enough to serve both ACP and PCP. A crystal structure of a complex from *B. subtilis* ACP and AcpS[77] throws light on the importance of the charged residues and explains why AcpS only interacts with ACP. Acidic residues of the ACP helix 3 interact specifically by the formation of salt bridges with the first helix of AcpS. This positions the serine residue in the right orientation for the priming reaction. In PCPs one of the interacting acidic residues is exchanged by a basic residue, which might explain the inability of AcpS to prime PCPs. On the basis of these results a novel hybrid PCP was constructed by replacing the *B. brevis* $TycC_3$ PCP helix 2 with the corresponding helix of *B. subtilis* ACP which contains the interacting residues.[78] This hybrid enzyme was stoichiometrically phosphopantetheinylated in vitro by both AcpS and Sfp.

2.2.1. Misacylation and Regeneration

The promiscuity of Sfp also causes undesired misacylation of PCPs within NRPS assembly lines since not only CoA but also acyl−CoAs can serve as cofactor donors. To regenerate these misprimed NRPS templates, nature has developed specific enzymes which catalyze hydrolysis of the undesired acyl group. These so-called thioesterase II domains (TEII) were shown to be specific for acyl−PCPs in in vitro assays, while there was no hydrolysis observed for acyl−ACPs, which are essential in primary fatty acid metabolism.[79] Comparison of the catalytic properties of TEII-mediated aminoacyl− or peptidyl−PCP hydrolysis vs acetyl−PCP showed a strong preference for the latter substrate, which indicates that this proofreading enzyme is important for NRPS activity. Deblocking of misacetylated PCPs was also confirmed by TEII knockout studies.

2.3. Peptide Elongation by the Condensation Domain

After activation and covalent binding of the first amino acid substrate by the A−PCP initiation module, peptide synthesis proceeds by stepwise condensation with amino acid building blocks bound to PCPs of the downstream elongation modules $(C-A-PCP)_n$. Peptide bond formation is mediated by a ca. 450 amino acid long condensation domain (C domain). The C domain catalyzes the nucleophilic attack of the downstream PCP-bound acceptor amino acid with its free α-amino group on the activated thioester of the upstream PCP-bound donor amino acid or peptide[80] (Figure 3C).

Biochemical characterization of different C domains from the *B. brevis* tyrocidine and *E. coli* enterobactin synthetases revealed insights into their substrate specificities. Probing substrate specificity in the natural synthetase is difficult because the upstream donor and downstream acceptor substrates are defined by restrictive A domains. To directly evaluate C-domain specificity, various aminoacyl−CoA substrates were synthesized and attached via an Sfp-catalyzed reaction to the apo-PCPs of a minimal, bimodular NRPS enzyme composed of module 1 from *B. brevis* gramicidine synthetase and module 2 from tyrocidine synthetase. Mischarging of PCPs from both modules in condensation assays revealed that the C domain of tyrocidine module 2 seems to possess an acceptor position for the downstream PCP-bound nucleophile that discriminates against the noncognate D-enantiomer as well as differences in the side chain. By contrast, low side-chain selectivity was observed for the donor position of the upstream PCP-bound electrophile. Interestingly, a preference for cognate D-enantiomers was observed.[80] This was confirmed by further investigations with the C domain of tyrocidine elongation module 5, which revealed that the donor position exclusively selects a tetrapeptide with the cognate D-configuration of the C-terminal residue for condensation reactions.[71] This shows that besides A domains and TEII domains C domains also represent a selectivity filter in nonribosomal peptide synthesis. Selection of the correct downstream nucleophile by the acceptor position prevents the formation of product mixtures and facilitates peptide synthesis in a directed manner. Assays with amino acids attached to *N*-acetylamine (SNAC) support the previously mentioned results and show that these soluble mimics of the ppan arm can serve as substrates for C domains, allowing the first kinetic studies at the acceptor site on C domains. Analogous experiments

at the donor site have not been possible however due to low turnover.[81]

Sequence alignments of several C domains revealed a highly conserved HHXXXDG motif that is also found in acyltransferases such as chloramphenicol acetyltransferase (CAT), NRPS epimerization, and heterocyclization domains.[82,83] Mutations in residues His147 in the conserved motif of the C domain of tyrocidine module 2 and His138 in enterobactin synthetase module F abolished activity in condensation assays,[84,85] providing evidence of an active role of the conserved histidine in catalysis. A recent crystal structure of VibH revealed insights into the architectural organization of this enzyme[82] (Figure 4C). VibH of the *Vibrio cholerae* vibriobactin synthetase is a very unusual NRPS C domain since it catalyzes peptide bond formation between a PCP-bound DHB donor and a freely diffusible norspermidine acceptor substrate.[86] Moreover, VibH represents one of the few NRPS C domains which are not covalently attached to a NRPS module. VibH is a monomer with two pseudodimeric domains consisting of an $\alpha\beta\alpha$ sandwich. The conserved HHXXXDG motif is located in a loop at the interface between the two domains, which provides access to His126 from two different faces of the enzyme. His126 could function as a base to activate the free α-amino-group of the downstream acceptor substrate for nucleophilic attack on the upstream thioester. By analogy to structural data on CAT, it was postulated that the C-terminal face of VibH would bind the donor DHB substrate with its ppan arm extending into a solvent channel. Norspermidine could then nucleophilically attack the thioester from the opposite N-terminal site. In other NRPS C domains with donor and acceptor presented on PCPs, both ppan arms would be extended into the solvent channel from opposing open ends, which would represent binding sites for the two PCP proteins. In contrast to the tyrocidine C domain and EntF C domain, mutation of the same conserved histidine in the C domain VibH resulted in little reduction in catalytic activity, which may reflect its specialized catalytic requirements.[82]

The heterocyclization domain (Cy), which can replace C domains in NRPS templates and catalyzes peptide elongation as well as heterocyclization by a more complex mechanism, is structurally and mechanistically related to the C domain. Five-membered heterocyclic rings such as oxazoline in vibriobactin and thiazoline in bacitracin are common structural features of nonribosomal peptides and important for chelating metals or interaction with proteins, DNA, or RNA[87] (Figure1). The first reaction step promoted by Cy domains is the nucleophilic attack of a PCP-bound cysteine, threonine, or serine acceptor substrate onto the thioester of the donor substrate. Recent mutational studies of the bacitracin Cy domain demonstrated that the free α-amino group is the nucleophile in this step, as observed for C domains.[88] In the next step the side chain hydroxy or thiol group carries out a nucleophilic attack onto the α-carbonyl C atom of the donor amino acid, forming a five-membered heterocyclic ring, which is subsequently dehydrated to form the final oxazoline or thiazoline product. Further insights into the mechanism of catalysis were provided by the vibriobactin synthetase, which contains two adjacent Cy domains. Analysis revealed that Cy domain 2 was responsible for the condensation step while Cy domain 1 carried out the heterocyclization and dehydration steps. This result indicates separate mechanisms for catalysis of condensation and heterocyclization.[86] These investigations as well as more general structural similarities indicate that Cy domains are evolutionarily specialized C domains.

2.4. Editing Domains

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

2.4.1. Epimerization

Almost every nonribosomally synthesized peptide contains D-configurated amino acids to a various extent. NRPSs utilize two different strategies for their incorporation. The most common route involves epimerization of L-amino acids by integrated 450-amino-acid-long epimerization domains (E).[89] The latter promote epimerization of the C_α-carbon of the PCP-tethered aminoacyl substrate to afford a D/L equilibrium.[90] Racemization in vitro can either occur from L to D or D to L. Rapid quench kinetics revealed that this equilibrium is achieved within seconds.[91] Specific incorporation of only the D-amino acid into the growing peptide chain is ensured by the enantioselective donor site of the downstream condensation domain.[80] This principle is also used in the surfactin synthetase in modules 3 and 6 to incorporate D-Leu twice in the final product. The combination of D- and L-amino acids contributes to the unique conformation of surfactin that is important for its biological activity.[10] A second strategy of D-amino acid incorporation is often observed in fungal systems:[32] The A domain of cyclosporin synthetase, for example, exclusively incorporates D-Ala, which is provided by an external racemase.[92]

Biochemical characterization of E-domain substrate specificity also revealed that noncognate amino acids were racemized but with lower efficiency.[93] Further studies showed that artificial E-domain constructs without a preceding C domain (as observed in the native initiation module) could epimerize aminoacyl−PCP. In contrast, identical constructs with a preceding cognate C domain (as in an elongation module) did not show epimerization activity for the bound aminoacyl−S−ppan substrate. This observation led to the conclusion that C domains tightly bind aminoacyl−PCP in the acceptor site until condensation occurs. The resulting peptidyl−PCP has a lower binding affinity for the acceptor site and is then transferred to the subsequent E domain or next C domain.[71,94] These investigations contributed to an

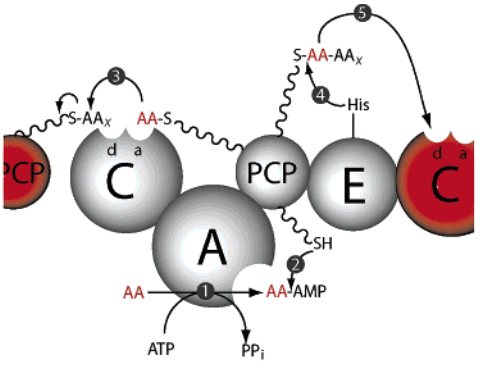

Figure 6. Directionality of peptide elongation. Reaction sequence (1−5) within an NRPS elongation module (shown in gray) containing an epimerization domain. AA = amino acid; AA_X = upstream amino acid or peptidyl chain.

understanding of timing and directionality in nonribosomal peptide synthesis (Figure 6).

Although a crystal structure of an E domain is not yet available, sequence alignments indicate their similarity to C and Cy domains.[82] C and E domains share a conserved HHXXXDG motif, the second His of which seems to be involved in catalysis.[83,90] Similar to C domains, it is assumed that this residue de- and subsequently reprotonates the C_α carbon atom (one base mechanism). Sequence alignments and structural comparison with the VibH C domain revealed that E domains have an insertion at the C-terminal end of the solvent channel, indicating that this face may be blocked and PCP binding may occur from the N-terminal face only.[82]

Moreover, it was shown that E domains play a crucial role in NRPS protein−protein recognition, mainly in bacterial systems.[95] For example, two subunits in surfactin synthetase have terminal E domains which communicate intermolecularly (in trans) with the C domain of the adjacent subunit (Figure 2). Moreover, product formation between two modules of the tyrocidine synthetase in trans was only observed if the free-standing module 1 (A−PCP−E) harbored the cognate E domain. A recent sequence analysis of the operon-encoding linear gramicidin A revealed the presence of 7 E domains, one of which surprisingly attached to a glycine-incorporating module. Since this E domain, which carries mutations in core motifs, is at an in trans junction of the synthetase, it may only be required for mediating protein−protein recognition.[96]

2.4.2. Methylation

Some nonribosomal peptides such as cyclosporin,[97] enniatin,[98] actinomycin,[99] and pristinamycin[11,100] have N-methylated peptide bonds. This modification is introduced by a ca. 420 amino acid comprising N-methylation domain (N-Mt) which is inserted into the accompanying A domain. The N-Mt domain catalyzes the transfer of the *S*-methyl group of *S*-adenosylmethionine (SAM) to the α-amino group of the thioesterified amino acid releasing *S*-adenosylhomocysteine as a reaction byproduct. In comparison to other NRPS domains discussed previously, less is known about N-Mt domains. It was shown for actinomycin synthetase a valine-activating A domain in module 2 could be replaced by a methyl−valine-activating A−N−Mt domain. This construct promoted peptide bond formation with acyl threonine, which is in agreement with the observed relaxed donor site specificity of C domains. In the absence of the appropriate acceptor only methyl−valine was observed, indicating that N-methylation occurs before peptide bond formation.[99]

SAM-dependent C-methylating domains (C-Mt) are also known. In yersiniabactin synthetase a thiazoline ring is C-methylated.[101] Recently, a new type of SAM-dependent methyl transferases was identified. Melithiazol synthetase from the myxobacterium *Melittangium lichenicola* lacks the conserved SAM binding signature sequence of N-Mt domains.[102] This enzyme is involved in an unusual methylation of a carboxy acid to form an ester, which represents the last step in melithiazol biosynthesis.

2.4.3. Further Modifications

Besides epimerization and methylation further modifications can be introduced into the peptide sequence. The oxidation state of oxazoline and thiazoline rings can be altered under catalysis of additional oxidation (Ox) or reduction (R) domains. The ca. 250 amino acids comprising oxidation domains are observed, e.g., in the NRPS modules of bleomycin, epothilone, or myxothiazol synthetases.[103−105] Two different organizations of the flavine−mononucleotide (FMN) containing domains have been reported. In myxothiazol synthetase one Ox domain is C-terminally fused to the PCP while another Ox domain is incorporated within the A domain.[105] Recent biochemical characterization of a recombinant Ox domain from the epothilone synthetase revealed that this enzyme retains autonomous activity, catalyzing the oxidation of thiazoline to thiazole. Molecular oxygen was required in these experiments to reoxidize reduced FMN.[104] Similar experiments were carried out with an Ox domain from bleomycin.[106] Recently, it was shown that an in-frame deletion of an Ox domain from myxothiazol synthetase did not alter the final product from a thiazole into a thiazoline derivative. It was therefore speculated that the other Ox domain of the synthetase would oxidize both thiazolines.[107] An interesting oxidation strategy is also realized in the biosynthesis of melithiazol and myxothiazol.[102,107] Glycine is incorporated as the last amino acid into the myxothiazol precursor and subsequent hydroxylation by a monoOx domain leads to the release of the myxothiazol amide and PCP-bound glyoxylic acid. While the terminal amide of myxothiazol is the final product, the terminal amide of melithiazol is processed further by enzymatic hydrolysis and methylation (see N-methylation).

Nature has also developed an opposite strategy which allows reduction of heterocycles by addition of two electrons as seen in one of the rings in yersiniabactin and pyochelin. Reduction is catalyzed by

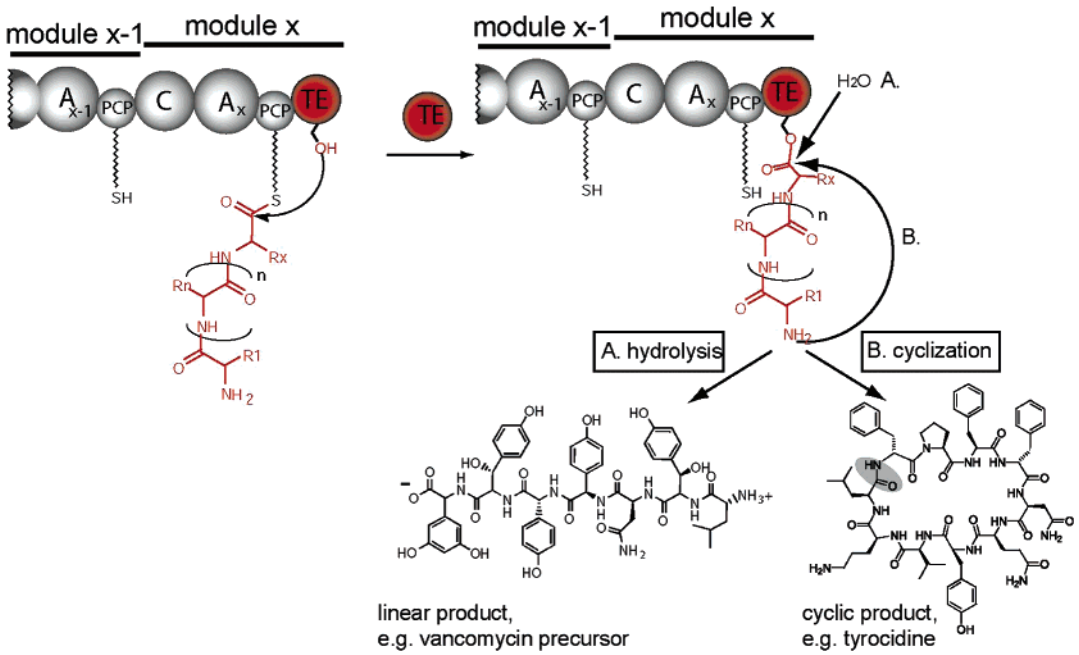

Figure 7. Peptide release by the TE domain. Depending on the identity of the NRPS template, product release can be carried out either by the external nucleophile water to give the linear acid product (A), as observed in case of the vancomycin TE, or by an internal nucleophile to yield a cyclic product as seen for tyrocidine TE (B).

NADPH-dependent R domains which catalyze the reduction of thiazoline into thiazolidine.[108] R domains are also involved in peptide release as discussed later.

A key characteristic of many nonribosomally produced peptides such as surfactin, mycosubtilin, and fengycin is the N-terminally fused lipo acid. Thus far, it is not well understood how this lipoinitiation is mediated. Since all initiation modules in these synthetases start with a C domain it is assumed that an acyl transferase provides the acyl chain for the C domain donor site.[13,42]

Another modification of the N-terminal peptide end is N-formylation catalyzed by a *N*-formyltetrahydrofolate-dependent formyltransferase domain (F) as observed in anabaenopeptilid 90-A and linear gramicidine A.[96,109]

2.5. Peptide Release

All catalytic domains discussed so far are repeating units of the enzymatic template and contribute to the synthesis of a linear peptide molecule tethered to the multienzyme. To reactivate the multienzyme for a next synthesis cycle the mature peptide has to be cleaved once it reaches the end of the assembly line. This reaction is usually accomplished by a ca. 280 amino acid long thioesterase domain (TE domain) fused to the C-terminal module, also referred to as a termination module. In the last step of peptide assembly an active site serine of the TE domain carries out a nucleophilic attack on the PCP−peptidyl thioester to form a covalent acyl−enzyme intermediate. Depending on the NRPS template and hence on the TE domain, this intermediate can either be released by hydrolysis as a linear acid or by an intramolecular reaction with an internal nucleophile to give a cyclic peptide (Figure 7). Hydrolytic release is observed for peptides such as vancomycin, whose peptide backbone is constrained by further postsynthetic oxidative cross-linking reactions, whereas the tyrocidine and surfactin backbones become directly constrained by TE-mediated macrocyclization (Figure 2 and Figure 7A).[12,110]

Alternatively, peptide release can also occur concomitant with reduction of the carboxy group catalyzed by the NADPH-dependent reduction domain (R) to give linear aldehydes or alcohols such as in the yeast biosynthetic pathway for the essential amino acid lysine[111] and in the biosynthesis of linear gramicidin A in *B. brevis*[96] or by head-to-tail condensation mediated by the C domain as seen in cyclosporin synthetase. Many other cyclization strategies have been uncovered which give rise to a large and diverse set of cyclic or cyclic branched molecules with distinct biological activities. Thioesterase domains that catalyze a cyclization reaction are also referred to as peptide cyclases.

The broad variety of cyclization strategies is encoded by the three-dimensional architectures of TE domains. Recent access to the crystal structure of the surfactin TE (Srf TE) domain has afforded deeper insight into the mechanism of cyclization.[112] The crystal structure represents a very prominent α,β-hydrolase fold which is characteristic of a large family of proteases, lipases, and esterases (Figure 8A). The α,β-hydrolase fold also provides a stable scaffold for the active sites of a variety of other TE domains which share only low sequence identities of 10−15%.[4,113] This high flexibility in primary sequence reflects the broad spectrum of activities mediated by these enzymes which need a rigid fold to ensure precise alignment for catalytic action. The crystal structure of Srf TE revealed several catalytic residues and regions which were assumed to play a crucial role in catalysis. A putative PCP domain interaction site ensures docking and presentation of the ppan-bound substrate in the active site via a cleft in the

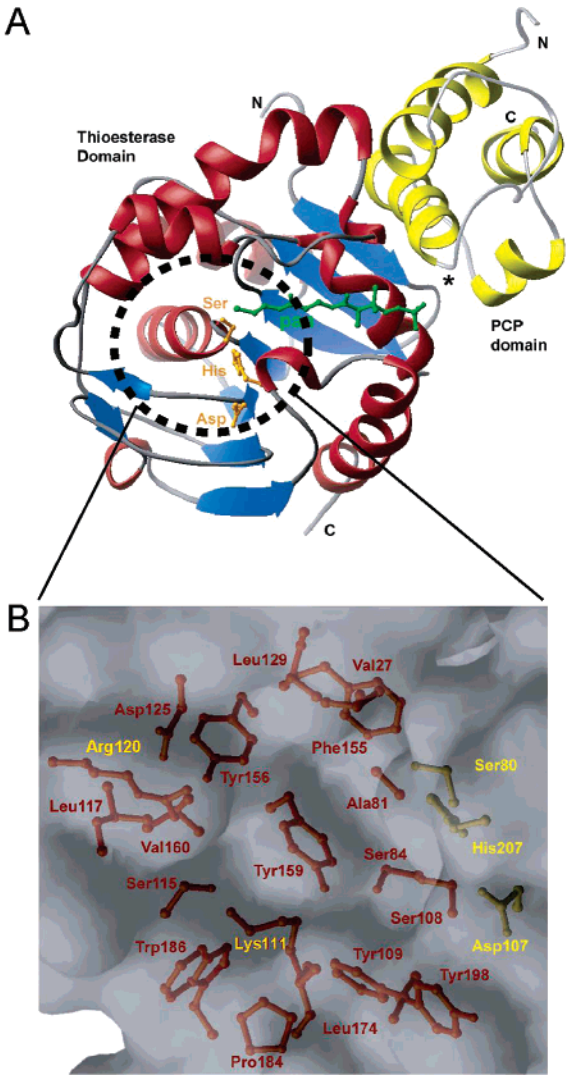

Figure 8. Structure of Srf TE. (A) Crystal structure of Srf TE (α,β-hydrolase fold) together with an upstream PCP domain (shown in yellow) at a putative interaction site. A modeled posphopantetheine arm points from the invariant PCP serine residue (*) into the Srf TE active site cavity (circled). (B) Magnification of the binding pocket reveals the catalytic triad (Ser80, His207, and Asp107) shown in yellow. The cavity is lined predominantly by hydrophobic residues with the exception of two positively charged side chains of Lys111 and Arg120 (yellow).

surface of the TE (Figure 8A). The active site cavity is bowl shaped and lined with predominantly hydrophobic and aromatic residues with the exception of two positively charged residues Lys111 and Arg120 (Figure 8B). This hydrophobic environment matches the hydrophobic peptide sequence of surfactin, while the two positive charges in the active site were postulated to mediate recognition and alignment by coordination of Glu1 and Asp5 in the surfactin sequence. Mutations of these residues to Ala confirmed their role in catalysis.[110] The three residues Ser80, His207, and Asp107 exhibit the right geometry and distance required for a catalytic triad. Mutation of all three residues confirmed that serine forms a covalent acyl−enzyme intermediate after deprotonation by His.[110] This acyl−enzyme intermediate was first identified by mass spectrometry in the enterobactin TE after mutation of the triad residue His to Ala. No such behavior was observed for the same mutation in Srf TE, which indicates different reactivity profiles in different TE domains. Better understanding of the Srf TE substrate-binding mode was then achieved by cocrystallization studies with a boronic acid inhibitor.[110] The cocrystal structure confirmed previous mutational studies and revealed binding of the boronic acid by the triad residue Ser80. Moreover, the structure revealed distinct recognition and binding of the C-terminal residues Leu7 and D-Leu6 of the surfactin peptide in hydrophobic pockets while the rest of the peptide seemed to be less ordered and constrained by the enzyme. Cyclization of lipopeptides may be triggered by a pronounced hydrophobic pocket which is present in the Srf TE active site close to the predicted fatty acid position. Binding of the fatty acid in this pocket might ensure a precise positioning of the β-hydroxy group required for a nucleophilic attack on the serine−peptide oxoester. This assumption is supported by investigations with the CDA TE domain where longer fatty acid chains increased not only cyclization yields dramatically but also the regioselectivity of the cyclization reaction.[70]

Although all members of the α,β-hydrolase enzyme family possess a similar structural fold, the mode by which the acyl−enzyme intermediate breaks down can be substantially different. While some TE domains such as vancomycin TE and lipases catalyze hydrolysis of the enzyme−substrate oxoester to give linear acids, other TE domains catalyze cyclization. A critical stage of the two different paths is encoded by the oxyanion hole. Two amide bonds stabilize the negatively charged tetrahedral intermediate in the release reaction. The structural integrity in the Srf TE domain is ensured by a rigid proline residue adjacent to one of the stabilizing amide bonds. This proline is highly conserved in cyclizing TE domains. In contrast, a flexible glycine is at this position in hydrolyzing lipases, which may ease the entry of water during a critical step of catalysis. Mutation of Pro26 to Gly in the Srf TE confirmed this hypothesis, dramatically increasing the amount of hydrolysis product formed. This residue may therefore represent a switch between cyclic or linear products in different enzyme families.

2.5.1. Diversity by Cyclization

Macrocyclization is a key structural feature of many nonribosomal peptide products which constrains the flexible peptide chain in a biologically active conformation. Rigidity in the peptide backbone facilitates specific interactions with dedicated cellular targets (section 1). In turn, many linear analogues of cyclic peptides display no or only diminished activities.[114] The huge variety of cyclic nonribosomal peptides is achieved by a diverse set of TE domains. Their high degree of specificity allows cyclization reactions to occur in the presence of other nucleophiles without the use of protecting groups. Contrary to the situation in solution where the peptide chain has to find the right conformation for cyclization by various rotations, the TE domain active site guides the folding of the peptide chain. TE domains can

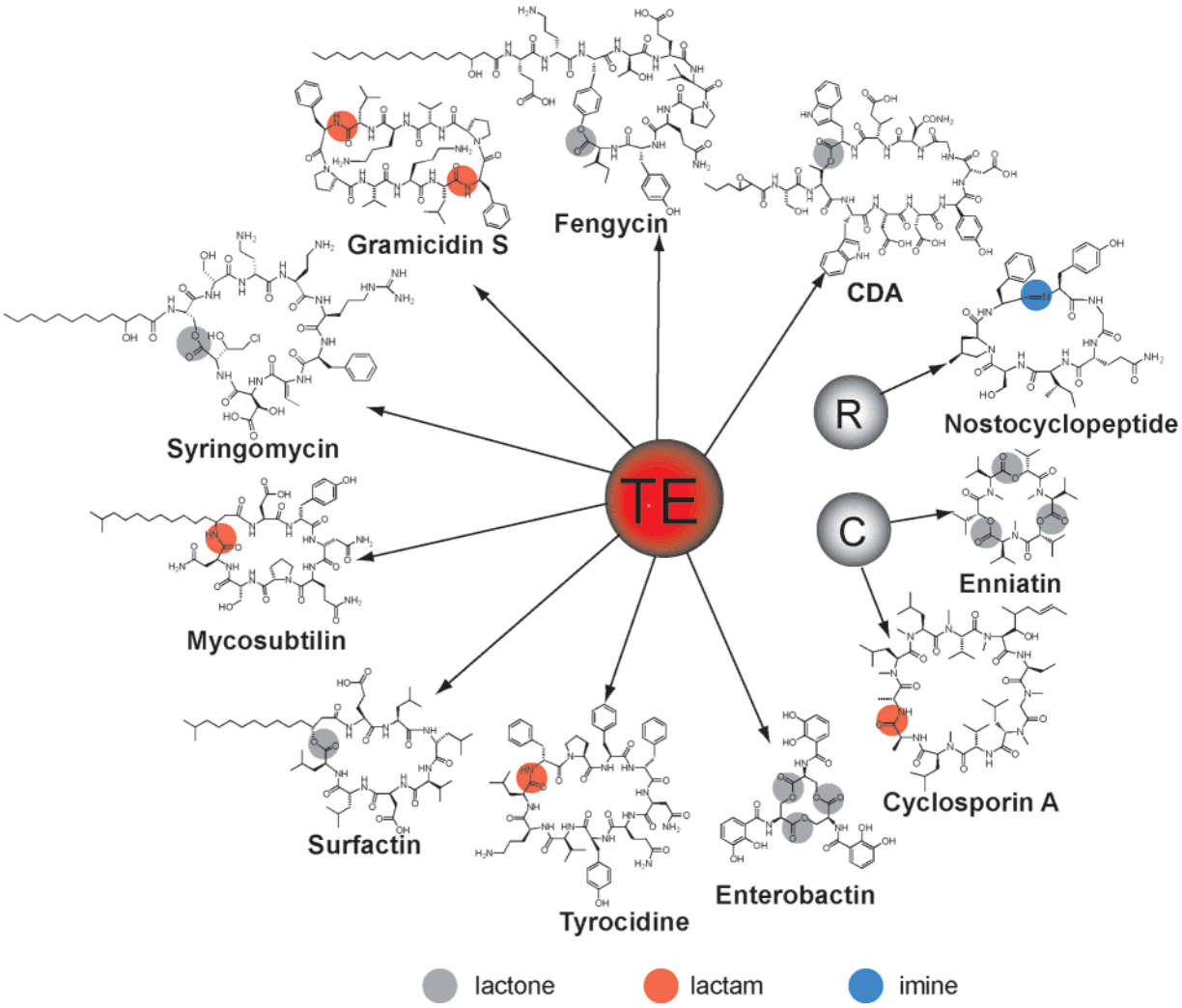

Figure 9. Cyclization strategies. The majority of cyclization reactions within NRPS templates are catalyzed by TE domains (red). In some cases, also C domains and R domains (gray) are involved in cyclic peptide release.

specifically select one residue from a source of nucleophiles to catalyze regio- and stereoselective cyclization reactions. These nucleophiles always attack to the C-terminal end of the peptide; no side-chain-to-side-chain or N-terminal-amine-to-side-chain cyclizations have ever been observed. In basic head-to-tail cyclizations, as seen for tyrocidine (*B. brevis*), the free N-terminal amine is connected to the C-terminus, yielding a lactam product (Figure 9). Surfactin and mycosubtilin (*B. subtilis*) are examples of a branched chain lactone and lactam, respectively (Figure 9). The TE mediates ring closure by connecting a β-hydroxy fatty acid for surfactin and a β-amino fatty acid for mycosubtilin to the peptide C-terminus. Both lipoheptapeptides share similarities in size, activity, and mode of synthesis as well as in the precursor β-keto fatty acid. In the case of surfactin the ketone is reduced to a hydroxy group, while in mycosubtilin several catalytic domains convert the ketone into an amino group. This processing of the same precursor in different ways leads to an increase in structural diversity. A change in the nucleophile from a hydroxy group to an amine seems to alter the chemoselectivity of the corresponding TE. Experiments revealed that it is not possible to cyclize a β-amino analogue of either surfactin or mycosubtilin directly with Srf TE, which demonstrates that alternative nucleophiles are not tolerated.[110] Similar results were observed for CDA TE.[70] Srf TE is specific only for the (*R*)-configured hydroxy fatty acid, emphasizing a high degree of chemo-, stereo-, and regioselectivity also observed for the analogous fengycin, CDA, and syringomycin TE domains.[70,115,116] Besides functionalized fatty acid residues, amino acid side chains can also be involved in cyclization. The TE domains of fengycin, syringomycin, and CDA synthetases specifically select dedicated tyrosine, serine, and threonine side-chain nucleophiles for connection with the C-terminus (Figure 9).

In many NRPSs the modular enzymatic template is collinear with the peptide product sequence. In these linear type A NRPS assembly lines TE domains only catalyze one reaction step, either cyclization or hydrolysis of the linear precursor.[5] However, in iterative NRPS type B templates, the TE domains have an additional function which allows the enzyme to repeat the collinear synthesis once or twice. In this case the TE has to count the monomers stalled at the end of the assembly line and initiates release by cyclic dimer or trimer formation only once the desired length is achieved. This strategy is observed for gramicidin S, enterobactin, and bacillibactin peptides

(Figure 9). Less is known about the mechanism and structure of iterative TEs. Detailed mass spectrometric analysis was carried out for the last module of the enterobactin assembly line (EntF) containing a C-terminal TE domain which catalyzes cyclotrimerization of three 2,3-dihydroxybenzoyl serine (DHB−Ser) units to give the cyclic trilactone enterobactin.[117] It was demonstrated that this TE is involved in two reactions: acyl chain growth and cyclization.

Macrocyclization is not exclusively mediated by TE domains. In cyclosporin A the final peptide bond is formed by a putative condensation domain[118] instead of a TE domain, emphasizing that nature developed additional enzyme species capable of catalyzing product release by cyclization. A PCP−C didomain can also catalyze oligomerization as observed for the trilactone enniatin.[119] Recently, a new type of head-to-tail macrocyclization was reported for nostocyclopeptide from the terrestrial cyanobacterium *Nostoc* sp. The C-terminal residue of the linear peptide is reduced by the action of an R domain to give an aldehyde, which is then intramolecularly captured by the α-amino group of the N-terminal amino acid residue to give a cyclic imine[120] (Figure 9).

2.6. Quaternary Architecture

Many enzymes catalyzing sequential metabolic reactions aggregate by noncovalent linkage of identical or nonidentical subunits to form multienzyme complexes. The fatty acid synthases (FAS) of eukaryotes and the modular polyketide synthases (PKS) are well-known examples of a class of enzymes that forms complexes composed of two identical subunits (Figure 10).[121−123] A double-helical structure model for modular PKS has been proposed.[123] According to this model the enzyme subunits are orientated head-to-head and folded in an interwound helical manner. Ketosynthase (KS), acyl carrier protein (ACP), and acyltransferase (AT) domains form a core which is necessary for the observed interaction between KS and ACP domains from different strands (Figure 10). Optional domains such as the ketoreductase domain (KR) are accommodated in outside loops. The model for FAS enzymes also requires functional interactions between both subunits to explain the results of cross-linking studies.[124,125]

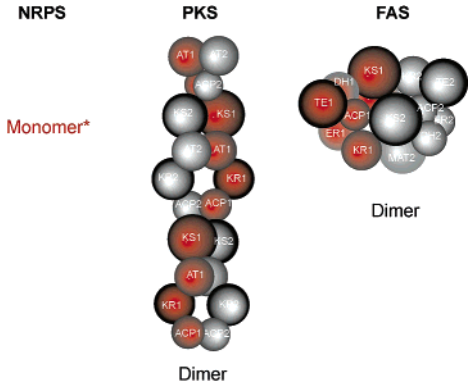

Figure 10. Quaternary structures of NRPS, PKS, and FAS multienzyme complexes. While FAS and PKS fold into homodimeric enzymes with domains communicating across the dimeric interface, no such interaction was found for several NRPSs (* except VibF).

The related organization and chain elongation logic of PKS and NRPS as well as the existence of naturally occurring hybrids that produce natural products such as epothilone has led to the assumption of a similar quaternary structure for both enzyme families.[44] In contrast to PKS and FAS enzymes, not much was known about the quaternary organization of NRPS for a long time. Although high-resolution structures of NRPS core domains are now available, structural insights into multidomain organization have remained elusive. Structural differences in the oligomeric states of NRPS and PKS enzymes were first revealed by comparison of their respective thioesterase domains.[112,113] While the crystal structure of the erythromycin PKS TE domain clearly shows a leucine-rich hydrophobic dimer interface, only a monomeric structure is observed for the surfactin NRPS TE domain. A monomeric organization has been similarly postulated[82] for the NRPS C domain VibH.

The first global analysis of the quaternary architecture of NRPS was carried out with various NRPS multidomain modules derived from tyrocidine, gramicidin S, enniatin, and enterobactin synthetases using strategies which were previously successfully applied to establish a dimeric interaction in FAS and PKS enzymes.[126,127] Biophysical methods such as gel filtration, chemical cross-linking, and analytical ultracentrifugation revealed a monomeric organization of all enzymes investigated. For the larger dimodular enzyme $TycB_{2-3}$ (A−PCP−C−A−PCP−E) a dimeric species was observed during ultracentrifugation at high nonphysiological concentrations, which was speculated to be an evolutionary relic. Biochemical experiments such as mutant complementation and affinity tag labeling also supported the monomeric state of the tested NRPS enzymes.[126] On the basis of these results, an overall monomeric structure of NRPSs was suggested. The existence of NRPS−PKS hybrids lead to the assumption that the PKS portion of the protein dimerizes while the NRPS part is in a monomeric form. In contrast to dimeric FAS and PKS, the monomeric structure of many NRPS enzymes emphasizes that there is no mechanistic requirement for them to function as dimers. In a recent publication Smith and co-workers showed that there is no such mechanistic requirement for FAS[128] either, in contrast to previous belief. Full inactivation of one polypeptide chain in a dimer did not abolish FAS activity, indicating that one functional subunit in a heterodimer is sufficient for product assembly. The dimeric state therefore seems to contribute predominantly to enzyme stability and integrity, since monomeric FASs are functionally inactive.

Recent ultracentrifugation and mutant complementation studies of the NRPS module VibF, which has the unusual domain organization Cy−Cy−A−C−PCP−C, revealed a dimeric state.[129] In contrast to cold labile FAS, the VibF dimers dissociate at elevated temperatures, which suggests a different pattern in dimer breakdown and reformation. Since

the accompanying free-standing C-domain VibH was shown to be monomeric, two oligomeric states in the vibriobactin biosynthetic template can be assumed, similar to the postulated situation in NRPS−PKS hybrid enzymes. This may also indicate that NRPS enzymes are able to display two different modes of structural organization. Knowledge of the quaternary structure of NRPSs is not only important for understanding reactions occurring on or between enzymes but may also contribute to engineering of more efficient new hybrid enzymes. A structural model of intramolecular domain interactions within a monomer must await high-resolution structures of a multidomain system.

2.7. NRPSs in Higher Eucaryotes

Until recently, NRPSs were only observed in bacteria, fungi, and yeast, but two recent publications provide evidence that a NRPS-like assemblage is also formed in higher eukaryotes.[130,131] In *Drosophila* the three-domain multienzyme *Ebony* (A−PCP−AS) was shown to be involved in histamine neurotransmitter metabolism at the photoreceptor synapse of the eye. Its postulated function is to ensure rapid removal of histamine from the synaptic cleft, which is essential to excite the postsynaptic cell by disinhibition. Histamine is trapped by Ebony via peptide bond formation with β-alanine. Experimental data with a 879 amino acid long recombinant form of Ebony revealed a novel two-step reaction mechanism involving amino acid activation by an A domain, covalent attachment to the ppan group of the PCP, followed by peptide bond formation via a C-terminal domain with an as yet unknown mechanism.[130] In vitro assays showed that the Ebony A domain exclusively activates β-alanine and transfers it subsequently to PCP. The novel C-terminal AS domain presumably then catalyzes the nucleophilic attack of primary amines such as histamine onto the β-alanine thioester. Given the existence of Ebony in *Drosophila*, NRPSs in higher organisms cannot be excluded. This conclusion was confirmed by a recent publication of Kato and coworkers, who described a 1100 amino acid long multienzyme from mouse, U26, which contains an A domain, PCP, and seven pyrroloquinoline quinone (PQQ) binding motifs.[131] In mammals U26 seems to be involved in lysine degradation by oxidation of 2-aminoadipic 6-semialdehyde by PQQ to 2-aminoadipic acid. Interestingly, in yeast lysine is synthesized by the reverse pathway from the NRPS enzyme Lys2 and Lys5. Lys2 has NADPH-dependent 2-aminoadipic acid reductase activity, and Lys5 catalyzes 4′-phosphopantetheinylation of Lys2.[111] A homologue of Lys5 has been identified in humans, raising the question as to whether a structural analogue of Lys2 exists in animals. In addition, the human 4′-phosphopantetheinyl transferase was capable of priming prokaryotic PCP and ACP domains, indicating that in humans a single enzyme with broad specificity is responsible for all posttranslational priming reactions.[132] These recent developments indicate that NRPS activity has been preserved throughout evolution to higher eukaryotes.

3. Approaches to New Antibiotics

One of the current challenges in NRPS research is to re-engineer natural products in order to increase or alter their biological activities. Research of the past few years has revealed that NRPS and NRPS−PKS hybrids can produce biologically active compounds and have outstanding potential for new drug discovery. A prominent example is the mixed NRPS−PKS product epothilone which is a promising candidate for combating cancer.[133−135] One goal is to improve natural-product-based drugs by rational protein engineering of these enzymes (e.g., by module and domain swapping). Recently, a second strategy has evolved which exploits a combination of chemistry and enzymology to create libraries of novel compounds. Solid-phase peptide chemistry is a well-established method to produce linear peptides in good yield. Cyclization of these molecules, however, poses some challenges,[38] which can often be solved by enzymatic means. The chemoenzymatic approach can also contribute to a better understanding of NRPS natural product assembly in general and of individual domains. Knowledge gained from such studies is essential for further biological or chemical engineering.

3.1. Genetic Engineering Approaches

Nature utilizes the modular assembly line methodology to produce a large set of small bioactive peptides with a huge variety of building blocks and modifications. The order and domain composition of modules are the result of a careful selection during evolution to synthesize a peptide molecule with the best bioactivity. As a consequence, several hundred NRPS building blocks are known which are incorporated by the same number of specific modules. Once the logic and mechanisms of NRPS assembly had been explored, interest developed in rationally redesigning the NRPS template to synthesize new peptide products.

Redesign was initially attempted at the genetic level, and several strategies have been examined to alter product outcome, including exchange of A−PCP units, artificial fusion of modules, module swaps, and deletions. The first reported genetic re-engineering experiment involved the terminal module of the surfactin synthetase which incorporates leucine in the natural system. This module exhibits the domain composition C−A−PCP−TE. To alter the amino acid specificity, the activating and covalent attachment domains (A−PCP) were exchanged by A−PCP units from bacterial and fungal origin with various amino acid specificities. Novel surfactin variants with aliphatic (Val), charged (Orn), and aromatic (Phe) residues at position 7 were created and confirmed by mass spectrometry.[136] All these new variants displayed the same hemolytic activity as the native surfactin. However, low yields of the peptide products (0.1−0.5% in comparison to the parent strain) were observed probably because of the high selectivity of C domains in the acceptor site for the cognate amino acid substrate.[80,81] Moreover, disruption of essential amino acids at domain borders could cause additional problems, and further attempts to obtain variants by

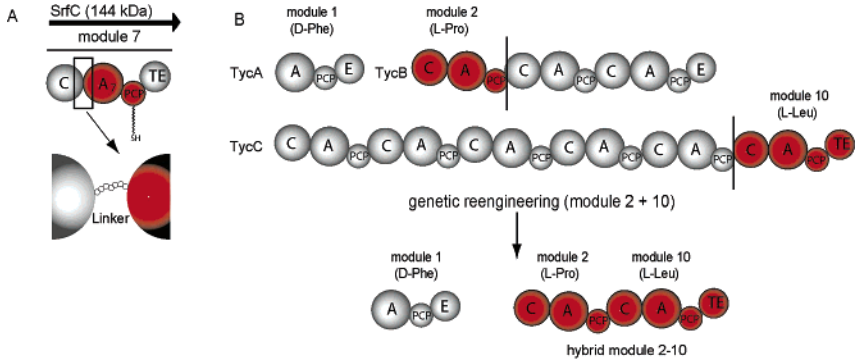

Figure 11. NRPS engineering. (A) Schematic representation of a linker region (ca. 15 amino acids) between individual domains. (B) Construction of a bimodular hybrid NRPS template derived from module 2 and module 10 (shown in red) of the tyrocidine synthetases Tyc B and Tyc C. Modules were defined as C−A−PCP.

domain swapping of surfactin synthetase module 2 were unsuccessful.[137] Domain borders generally seem to be crucial determents of multienzyme activity. Biochemical studies, sequence analysis, and structural information revealed linker regions between NRPS domains[50,74,82,138] which correspond to short nonfunctional stretches of ca. 15 amino acids bearing predominantly small and hydrophilic side chains with an almost random distribution. The sequence flexibility and their location between domains make such linkers suitable targets for artificial fusions without disrupting enzymatic integrity (Figure 11A). This strategy was first tested for tyrocidine synthetase by fusion of Pro-activating module 2 with Orn-activating module 9 or Leu-activating module 10[138] (Figure 11B). These two hybrid enzymes were then incubated with D-Phe-activating module 1 to yield artificial D-Phe-Pro-Orn and D-Phe-Pro-Leu tripeptides. Catalyzed tripeptide release was only observed when a TE domain was present at the C-terminal end of the last module. The catalytic potential of TE domains to increase the rate of product formation in engineered synthetases was explored in more detail by fusion of six different TEs to module 2 of the tyrocidine synthetase. All TEs were active in hydrolyzing dipeptide products from the enzymatic template, albeit with variable turnover rate after incubation with module 1.[139] Interestingly, an artificial NRPS−PKS hybrid enzyme, which was generated by fusion of the DEBS TE, also yielded active protein, demonstrating that engineering within the two enzyme families can be successful. The promiscuity of TE domains at hydrolyzing diverse products in artificial synthetases was further utilized in a recent approach for production of the dipeptide α-Asp-Phe, which is a precursor of the sweetener Aspartame. Asp-Phe represents the first example in which redesign of a NRPS was guided by a specific application. Although many NRPS biosynthetic clusters have been sequenced, no biosynthetic template is known where Asp- and Phe-activating modules are organized one after the other. Consequently, six hybrid NRPS enzymes were created in which A−PCP or A−PCP−C units of Asp-activating module 5 derived from the surfactin synthetase were fused via their linker regions to A−PCP, C−A, A−PCP, or A units of Phe-activating modules 1 or 3 of the tyrocidine synthetase.[140] The best hybrid enzyme composed, of A−PCP from surfactin module 5, C−A from tyrocidine module 3, and PCP−TE from tyrocidine module 10, displayed good chemoselectivity for the desired product α-Asp-Phe. However, the turnover rate of 0.7 min^{-1} limits the practicality of technical applications and requires further optimization.

Many natural peptide products display small heterocyclic elements which are associated with their bioactivity. Synthesis and incorporation of heterocyclic compounds such as thiazoline or oxazoline are therefore desired for the synthesis of new pharmaceutical lead compounds. Recently, a genetic approach to these compounds was introduced by construction of new hybrid NRPSs using heterocyclization domains (Cy).[88] In this study the Ile-activating A−PCP module 1 of bacitracin synthetase was fused either to a Thr-activating Cy−A−PCP module of mycobactin synthetase or to a Cys-activating Cy−A(Ox)−PCP module of myxothiazole synthetase.[88] The latter module carries an oxidation domain which further modifies the thiazoline product to give a thiazole. To ensure product release, both hybrids were again equipped with the tyrocidine TE domain at the C-terminus. As predicted from the selectivity of the corresponding A domains, the two expected products are Ile-Thr-oxazoline and oxidized Ile-Cys-thiazole. In case of the latter, Ile-Ser-oxazole was the preferred product, demonstrating tolerance for different amino acid substrates. The model studies show that heterocyclization can be achieved in engineered synthetases, although low yields of the corresponding products are typically encountered.

In addition to module and domain swapping, insertion and deletion of modules within their defined linker regions could contribute to alterations in the product sequence. To this end, deletion of the Leu-incorporating module 2 from surfactin synthetase was engineered.[141] Genetic manipulation and fermentation in the natural producer strain *B. subtilis* afforded the predicted surfactin product deprived of the second Leu residue in about 10% yield (in comparison to the parent strain). This represents a major improvement compared to initial engineering studies on surfactin module 2 and 7 and points to the importance of precise linker surgery.[136,137] This conclusion is further supported by results obtained from an exchange of Leu-activating module 2 of surfactin synthetase with Leu-activating module 10 of tyrocidine synthetase. Using the same linker strategy as in the deletion studies, surfactin produc-

tion could be improved to a 19% yield (in comparison to the parent strain).

The recombination of whole modules represents a rather drastic intervention in NRPS biosynthesis which usually results in reduced catalytic efficiency and product yield. A more conservative strategy involves manipulating the A domain's specificity through point mutations of substrate-coordinating amino acid residues according to the "nonribosomal code" (see A-domain section). For example, the substrate specificity of the Glu-activating module 1 of surfactin synthetase was rationally altered in this way. A single point mutation changed the specificity from Glu to Gln without a decrease in catalytic efficiency. A second specificity change in module 5 from Asp to Asn yielded the expected surfactin derivative in vivo.[57]

The exploration and definition of linker regions and C- and A-domain specificities has enabled substantial progress in the field of genetic NRPS engineering. However, the ultimate goal of mixing and matching diverse modules to synthesize peptides of any desired sequence has not yet been achieved. High-yield production of bioactive peptides with unusual structural elements such as D-amino acids and heterocyclic elements by fermentation would be an economic alternative to expensive chemical synthesis.

3.2. Chemoenzymatic Approaches

One goal of modern drug design is to identify new pharmacophores by rapid synthesis and bioactivity screens. NRPS peptides are promising scaffolds for such drug leads and have attracted much attention for applications in medicine, as the example of cyclosporin illustrates. Genetic manipulation is one way to create potential new NRPS drug leads, but it requires much labor and effort to generate the desired peptide product. To ensure rapid synthesis of large peptide libraries, chemical solid-phase synthesis seems to be superior. However, a limitation of this technique is the weak tendency of linear peptides to macrocyclize by chemical means due to the high entropic cost of populating the conformation with the right geometry cyclization. As a consequence, chemical synthesis often suffers from low cyclization yields.[35,142,143] A recent chemoenzymatic approach combines the strength of synthetic peptide synthesis with the strength of regio- and stereoselective TE-domain cyclization in high yields.[144,145] Once a new lead drug has been identified by this rapid synthetic method, genetic engineering may provide the enzymatic synthesis template to allow high-yield fermentation of the desired drug.

3.2.1. Chemoenzymatic Potential of TE Domains

To evaluate the chemoenzymatic potential of TE domains, characterization of substrate specificity, tolerance, and enzymatic restrictions have to be performed. These specific TE-domain investigations cannot be performed with the whole multienzyme complex because its large size causes preparative problems. Instead, an easy in vitro assay system was devised by cloning TE domains from the tyrocidine and surfactin synthetases and producing them as isolated enzymes. To test weather these excised TEs are active outside their natural synthetase context, their cognate decapeptide and heptapeptide substrates were synthesized by chemical means. In this approach the complete linear enzymatic peptide synthesis machinery, composed of a repeating set of catalytic domains, was replaced by solid-phase peptide synthesis (Figure 12A). Both methods share a similar strategy of precursor activation and tethering on a solid support, which facilitates rapid and ordered synthesis of desired products in high yields. One advantage of chemical synthesis is the huge diversity which can be incorporated into the linear peptide chain in order to create a variety of substrate analogues for biochemical studies of excised TE domains. In the natural synthetase peptide substrates are activated as ppan thioesters which ensure acylation of the TE domain active site. To provide similar recognition and activation for the artificial system, the thiol component of the natural ppan cofactor *N*-acetylcysteamine (SNAC) was attached to the C-terminus of a synthetic peptide (Figure 12B). Incubation of artificial tyrocidine and surfactin SNAC substrates with the cognate TE domains resulted in cyclization and hydrolysis.[110,115,146] The cyclic product was formed in as high as 85% yield, and the turnover rate for tyrocidine TE was 59 min^{-1}, sufficient for more specific investigations.

A set of experiments was designed to evaluate the general utility of TE domains for catalyzing diverse cyclization reactions. With regard to general utility as a cyclization tool, broad substrate tolerance would be desirable. Enzymatic recognition elements of Srf TE and Tyc TE were first investigated by systematic alteration of each amino acid in the hepta- and decapeptide sequences. In the case of tyrocidine, each amino acid in the substrate was replaced by alanine and the resulting 10 SNAC variants were incubated with the TE. This amino acid scan revealed a broad substrate tolerance. Only amino acids at the C- and N-termini seem to be recognized, leaving space for alterations in the middle of the peptide (Figure 13). Identity and stereochemistry of the N-terminal D-Phe is essential for enzyme activity, although an exchange of the free amine by a hydroxy group was tolerated and led to macrolactonization. Alterations in peptide length gave cyclic hexa- and dodecapeptides, and cyclodimerization of two pentapeptides was also tolerated.[115] Tyc TE thus seems to be a very promiscuous cyclization catalyst, tolerant to changes which are desired in the generation of new peptide products. Alterations of the tyrocidine peptide backbone yielded a minimal recognition model for explaining cyclization activity.[147] Similar investigations on the substrate tolerance of Srf TE also showed the importance of the N- and C-terminal residues in the peptide[110] (Figure 13). Cocrystallization with a boronic acid inhibitor confirmed the biochemical data, providing evidence for two hydrophobic binding pockets which can accommodate the two C-terminal Leu residues of surfactin. In contrast to Tyc TE, however, alterations in the nucleophile and substrate length were not tolerated. The β-sheet content of the two peptides may provide an explanation for the observed

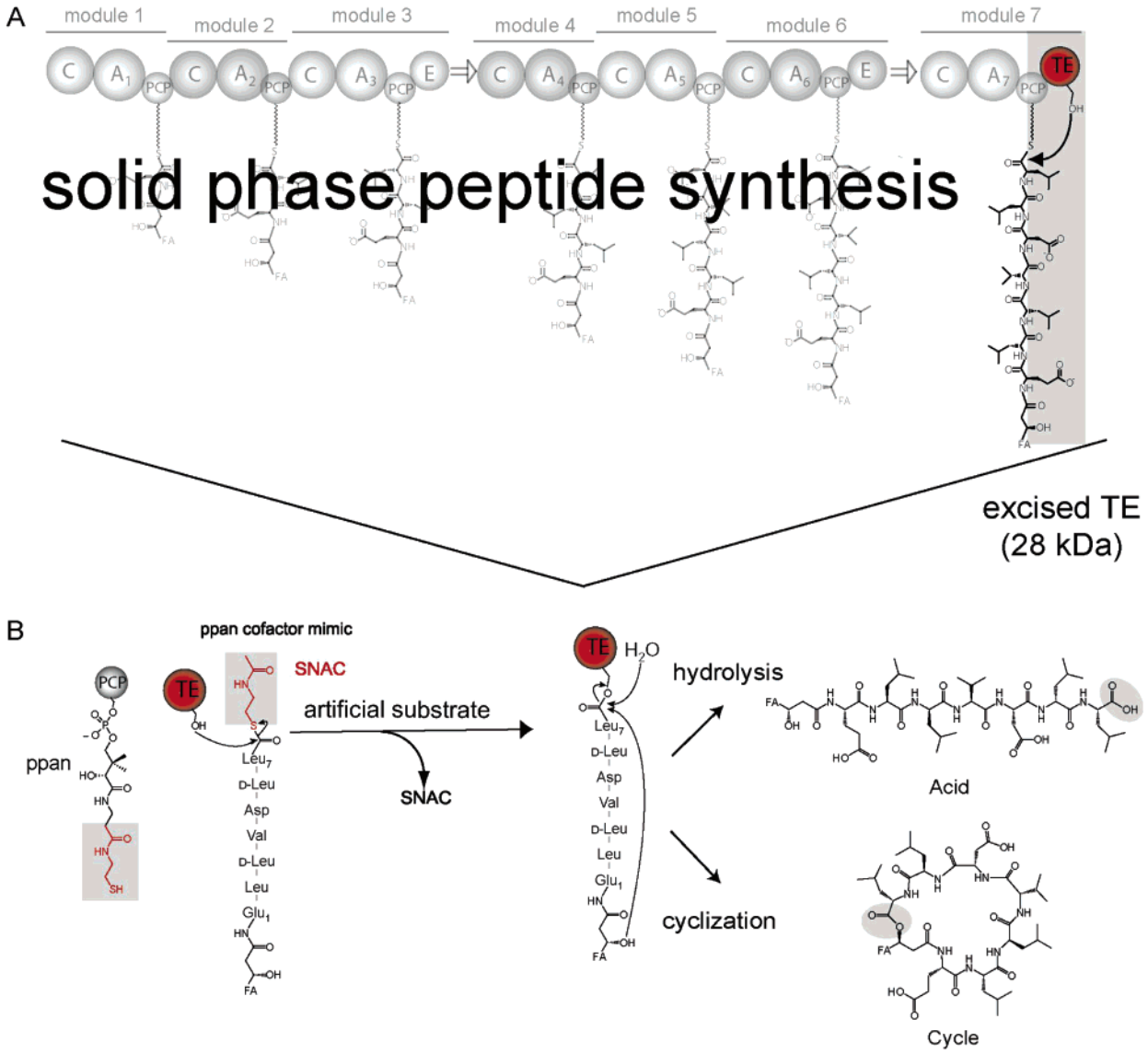

Figure 12. Chemoenzymatic cyclization. (A) The NRPS multienzyme machinery required for peptide elongation is replaced by solid-phase peptide synthesis, and the TE domain is used as an isolated enzyme. (B) Recognition of the artificial substrate by the enzyme is ensured by the ppan cofactor mimic SNAC (highlighted).

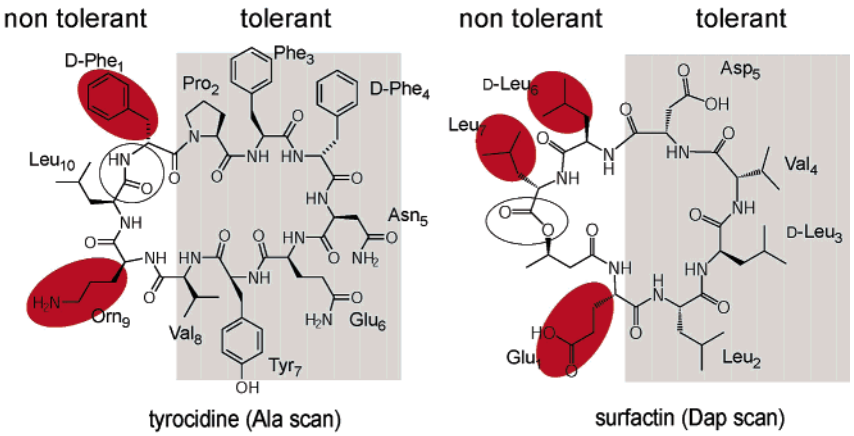

Figure 13. Substrate tolerance of tyrocidine and surfactin TE domains. Residues marked in gray (tyrocidine 2−8 and surfactin 2−5) can be substituted by alanine or diaminopropionic acid, respectively. Those shown in red are essential for substrate recognition by the TEs.

differences. Tyrocidine has a high β-sheet content, and cyclization is promoted by preorganization of the peptide backbone. Cyclization of tyrocidine can therefore also occur without enzymatic catalysis but with lower efficiency.[148] The reduced number of β-sheets in surfactin as well as the branched chain cyclization

mode, which requires larger fatty acids for improved activity,[70] may account for the observed differences in substrate tolerance. The broader specificity of Tyc TE makes it a promising candidate for biosynthetic applications.

3.2.2. Chemoenzymatic Route to New Drugs and Peptide Antibiotics

Many cyclic peptide antibiotics act on bacterial cells by insertion into the membrane, followed by disruption of osmotic and ionic regulation. Studies with small synthetic peptides revealed that alternating D- and L-amino acids as well as the presence of positively charged side chains contribute much to the overall activity against negatively charged prokaryotic cell membranes.[149] Many nonribosomal peptides such as tyrocidine exhibit low selectivity for prokaryotic vs eukaryotic cells, which limits their application as antibiotic drugs. To improve the preference for bacterial targets as well as the spectrum of activity against common bacterial pathogens, Walsh and coworkers introduced a solid-phase combinatorial synthesis of novel tyrocidine analogues.[150] A library of peptides with natural and nonnatural amino acid substituents introduced at two positions was constructed on a solid support (PEGA resin) by parallel synthesis. Enzymatic on-resin cyclization and subsequent analysis of the antibiotic activity of the reaction products against *B. subtilis* revealed potency for those peptides carrying a positively charged D-amino acid at the D-Phe4 position with 30-fold selectivity for bacterial membranes. Two of the best analogues also gained activity against Gram-negative organisms (Figure 14A). The improved tyrocidine variants which have been identified via this combinatorial chemoenzymatic approach can now be translated back into a modular engineered NRPS template for large-scale production by fermentation.

The chemoenzymatic potential of Tyc TE was subsequently used to generate small molecules with different therapeutic potential: peptide inhibitors of integrin receptors and hybrid peptide/polyketides (Figure 14B and C). Many natural ligands for integrins contain an Arg-Gly-Asp (RGD) sequence motif that is believed to be important for receptor interaction. Peptides containing the RGD motif are potent inhibitors and therefore potential therapeutic leads. Because interaction with the receptor is improved by incorporation of β-turns, a large number of changes in Tyc TE was required to produce the desired inhibitor substrate. In the end, only the Tyc TE minimal recognition elements were retained. Although 7 of 10 cognate residues were replaced in some cases with amino acids of opposite stereochemistry, Tyc TE was still capable of cyclizing the substrate, albeit with reduced yield.[114] Products were shown to be inhibitors of ligand binding by integrin receptors with cyclization as an important contributor to nanomolar potency. The broad substrate tolerance of Tyc TE was further utilized to mediate cyclization of hybrid polyketide−tyrocidine substrates as well as (*E*)-alkene−dipeptide isostere peptidomometics. The first approach contributes significantly to the synthesis of novel hybrid peptide/polyketides with therapeutic potency,[151] while the second method allows the peptide backbone to be modified postsynthetically by chemical metathesis.[152] Natural hybrids such as epothilone and bleomycin already exhibit cytostatic activity which could be further optimized via the chemoenzymatic route.

3.2.3. Expanding the TE Tool Box

In previous sections the utility of chemoenzymology was illustrated predominantly through studies with tyrocidine thioesterase. To expand the scope of macrocyclization catalysts, thioesterases from other NRPS enzymes were cloned and overexpressed. Surprisingly, no activity was observed for fengycin, mycosubtilin, and syringomycin TEs upon incubation with SNAC substrates, indicating limitations in the chemoenzymatic potential of these cyclases.[153] Two reasons might account for the observed inactivity of these TE domains. First, the enzymes could have been misfolded and were unable to reconstitute their activity after heterologous expression. Alternatively, substrate presentation by the short SNAC leaving group might have been insufficient. To rule out the second possibility, a new strategy was employed which allowed Sfp-catalyzed loading of peptidyl−CoAs onto apo PCPs mimicking the natural substrate presentation as close as possible (Figure 14A). Remarkably, Sfp was promiscuous enough to tolerate peptidyl−CoA substrates instead of CoA and acetyl−CoA. This observation is explained by the crystal structure, which shows specific interactions with the adenine base but the ppan arm pointing into solution (Figure 5B). This binding mode ensures enough space and freedom for the attached peptide chain. Loading fengycin−CoA onto the fengycin PCP−TE didomain indeed gave rise to both cyclization and hydrolysis activity, which had not been previously observed with SNACs. This result demonstrated that the PCP-tethered ppan arm is necessary to direct the substrate into the enzyme active site and guarantees appropriate alignment for nucleophilic attack of the active site serine.[153] This approach additionally illustrates the high regioselectivity of fengycin TE.

Recently, other C and R domains have been examined with respect to the possibility of bypassing NRPS specificity by directly loading peptidyl−CoA on any apo-PCP within an assembly line.[71,96] However, the single turnover nature of the reaction has proved to be a limitation (Figure 14A). After product release, the cofactor ppan remains attached to the PCP−TE didomain cyclase, which blocks further Sfp-catalyzed transfer of additional peptidyl−CoAs onto ppan−PCP.

To force multiple rather than single turnover cycles a new strategy was developed to expand the utility of peptidyl−CoA loading. It is known that amino acid thioesters undergo trans-thioesterification reactions when exposed to thiol-containing compounds.[154] In principle, such a thioester exchange reaction between the free ppan−PCP thiol and a soluble thioester−peptide substrate could enable both chemical reloading of substrate onto the ppan−PCP−TE didomain[154,155] and natural substrate presentation by the ppan arm. Inspired by expressed protein ligation,[155]

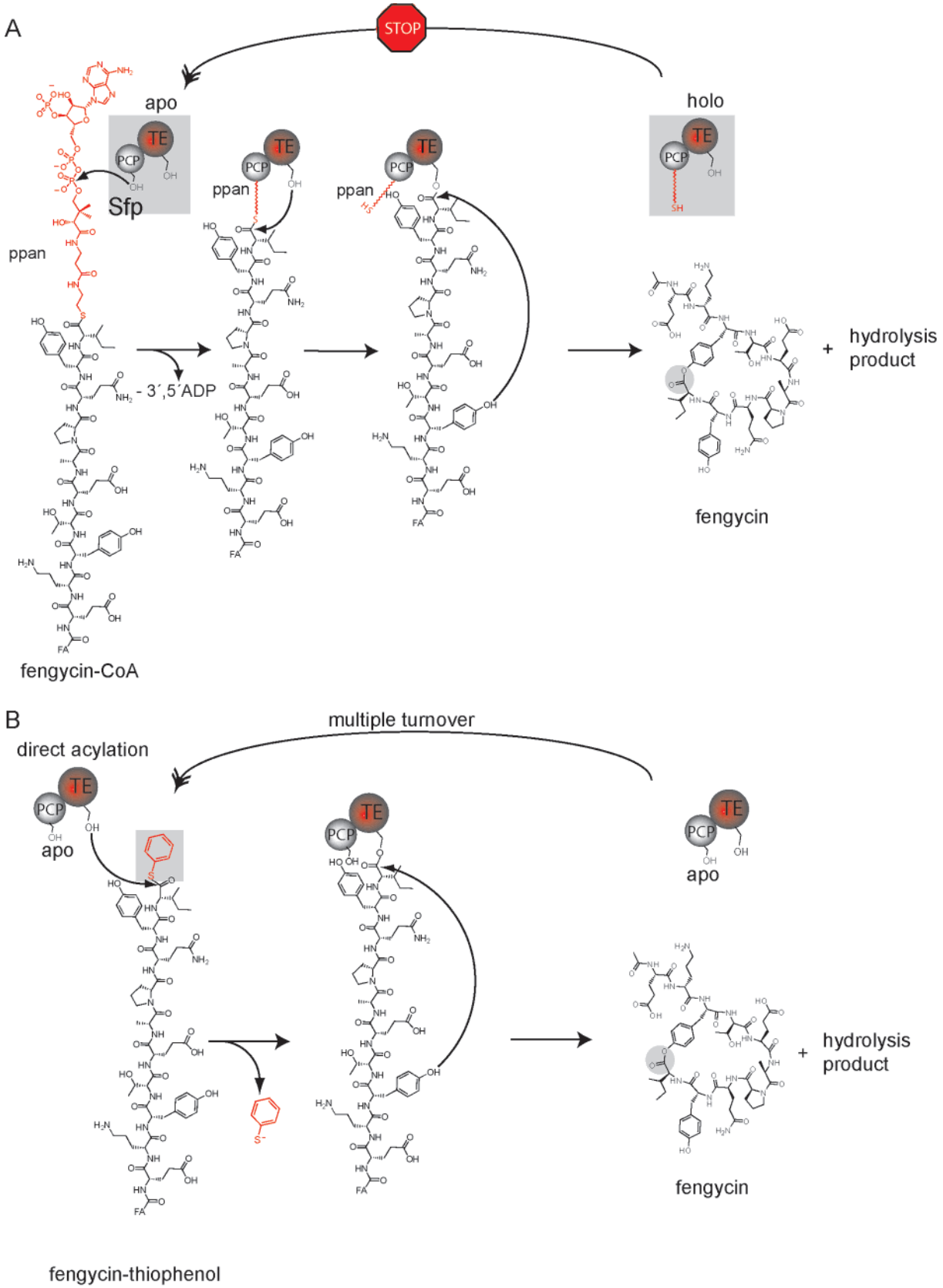

Figure 14. New strategies of substrate presentation. (A) PCP−TE ensures natural substrate interaction. Synthetic peptidyl−CoA (e.g., fengycin) is loaded onto apo PCP−TE by Sfp to give peptidyl−S−ppan−PCP−TE. Subsequently, the peptide substrate is transferred onto an invariant serine residue of the TE active site, which is then released by cyclization (via tyrosine$_3$ in case of fengycin). (B) Activity-based enzyme acyclation. The active site serine of the TE domain is directly acylated by a reactive peptidyl−thiophenol substrate. The acyl−enzyme intermediate is then captured by an intramolecular nucleophilic attack to yield the cyclic product (e.g., fengycin).

the peptide was activated as a thioester with thiophenol, which has favorable leaving-group properties. In the course of these studies it became obvious that soluble fengycin thiophenol directly acylated the TE active site serine rather than the free ppan thiol (Figure 14B).[116] The rapid direct acylation of the active site serine confirmed the autonomous activity of the excised enzyme. Moreover, this result showed that contrary to previous belief, natural cofactor recognition elements as displayed in SNAC or CoA substrates are not necessary for enzyme acylation but can be replaced by a suitable reactive leaving group.

The results of these experiments further suggested that nature might have developed peptide cyclases with different catalytic activities. While tyrocidine TE and surfactin TE show activity directly with SNAC substrates, mycosubtilin, fengycin, and syringomycin TEs appear to be completely inactive. A 15-fold increase in catalytic activity was observed for Srf TE when the SNAC leaving group was replaced by thiophenol. Recently, a similar increase in activity was reported for CDA−thiophenol compared to CDA−SNAC for CDA cyclase.[70] The thiophenol leaving group increases the velocity of the acylation step. While the acylation step depends on substrate presentation, as seen in the peptidyl−CoA experiments, deacylation is an intrinsic property of the acyl−enzyme intermediate, as shown by comparable cyclization-to-hydrolysis ratios irrespective of whether SNAC or thiophenol substrates are used. Undesired hydrolytic byproducts were observed in thiophenol-based in vitro studies, presumably due to spontaneous cleavage of the highly activated substrate in solution. The ratio of cyclization to hydrolysis was most favorable when cognate thiophenol substrates that fit precisely into the enzyme active site were used. Not much is known about the hydrolysis rate of natural NRPS templates, but it is likely that the multienzyme complexes also produce hydrolyzed byproducts to a certain extent. In general, however, selective enzyme acylation can be achieved for dedicated peptide cyclases with substrates activated with a variety of leaving groups.

3.2.4. Synthetic Utility of TEs: Chemical vs Enzymatic Cyclization

To evaluate the utility and potential of enzyme-catalyzed cyclization reactions, a comparison with established chemical methods is necessary. In principle, the chemical formation of macrocyclic rings is difficult because of energetically disfavored ecliptic and transannular interactions.[61] In solution, only a few conformers have the right geometry to allow intramolecular attack of a nucleophilic group on the C-terminal carboxy group. The entropic costs of populating these few productive conformations by several C−C bond rotations are high and therefore disfavored. To ensure regioselective cyclization, undesired competing nucleophiles such as hydroxy or amino groups have to remain protected while the desired nucleophile needs to be deprotected, which requires orthogonal protecting-group strategies. Side reactions, including intermolecular peptide bond formation and subsequent cyclo-oligomerization, may predominate since peptide bonds are usually trans-configured and favor a higher population of linear precursors. To minimize intermolecular reactions, high dilution conditions are applied (10^{-4}−10^{-5} M) which make large-scale reactions difficult. Alternatively, a peptide can be cyclized while it is still attached to the resin. Because the peptide chains are physically separated, intramolecular reactions are favored.[156] Turn-inducing elements such as D-amino acids, proline, glycine, or N-alkyl amino acids can also favor cyclization in solution, illustrating the importance of preorganization of the linear precursor for efficient ring closure, something that has also been observed for some NRPS products.[147] Another problem of chemical cyclization is the activation of the C-terminal carboxy group without amino acid racemization. Coupling reagents such as BOP and TBTU permit rapid cyclization but suffer from C-terminal racemization.[156] Less reactive coupling reagents minimize racemization but prolong reaction times. Better results were achieved with HOAt and DPPA.[142,157−159] Typically, with the above-mentioned chemical cyclization methods a 30−40% yield of cyclic peptide products can be obtained.[35,142,143,157] Reaction times range from several hours to days.

By contrast, enzymatically catalyzed cyclization reactions do not require protecting groups or high dilution conditions due to enzymatic specificity. In the literature, enzymatic methods for head-to-tail peptide cyclization have been predominantly reported. Cyclization of linear peptide esters was first described for the subtilisin mutant subtiligase.[160] Subtiligase cyclizes peptide esters longer than 12 residues with yields of 30−88% in a regioselective head-to-tail fashion. Hydrolysis and dimerization are observed byproducts. Head-to-tail cyclization without byproducts was reported for an intramolecular cyclization using split-inteins, allowing the generation of backbone-cyclized peptides in vitro and in vivo.[161]

Several NRPS enzymes, including surfactin, mycosubtilin, fengycin, and syringomycin TE, were shown to regiospecifically catalyze branched chain cyclization between one dedicated nucleophile and the activated C-terminal peptide residue in the presence of other potential competing nucleophiles. No oligomerization or C-terminal racemization was observed. The advantage of enzymatic vs chemical cyclization is illustrated by the example of tyrocidine A synthesis. While chemical, on-resin cyclizations typically occur in only 30% yield, enzymatic cyclization gives 85% product.[146,162] Because ring formation competes with the production of linear hydrolytic byproducts due to competing nucleophilic attack of water molecules, typical yields of TE-mediated cyclization reactions range from 40% to 91%; observed reaction times are several minutes to hours.

Since these recombinant peptide cyclases are usually embedded in a hydrophobic multienzyme complex, their production as isolated TE or PCP−TE domains may enhance the exposure of the active site to water. Moreover, substrate analogues used in in vitro studies often lack structural features, like long-chain fatty acyl chains, which are important for a

perfect fit into the enzyme binding pocket as seen for CDA but are difficult to incorporate synthetically.[70] The latter point is illustrated by in vivo studies with a genetically engineered surfactin synthetase lacking module 2.[141] While it was not possible to detect hydrolyzed surfactin product in the supernatant of the wild-type producer cells, hydrolysis was observed for the shorter surfactin variant produced by the engineered strain. This result indicates that hydrolysis can also occur in vivo, where the TE domain is embedded in a functional multienzyme complex. Therefore, presentation of the dedicated substrate seems to be a key step to minimize hydrolysis by TE domains. Moreover, artificial thioester leaving groups such as SNAC and thiophenol are short mimics of the natural cofactor ppan−PCP and may therefore be less effective of blocking water from the active site and hence more susceptible to hydrolysis. The nucleophilic side reaction with water could be minimized by enzyme catalysis in organic solvents.[163] The high enzymatic selectivity for cognate substrates can also be a disadvantage when the cyclization of substrate analogues is desired. Substitutions of residues, especially at the C- and N-terminal ends of the peptide sequence, can decrease or completely abolish cyclization yields. Nevertheless, with the current set of active peptide cyclases, diverse cyclization reactions such as head-to-tail and branched chain lactamization and lactonization of various substrates can already be performed.

In comparison to organic synthesis, the chemoenzymatic approach minimizes time and side reactions and maximizes purity and yield of cyclic peptide production. The current chemoenzymatic applications of TEs are summarized in Figure 15. Biochemical prerequisites such as enzyme characterization and leaving-group technology have been solved which will

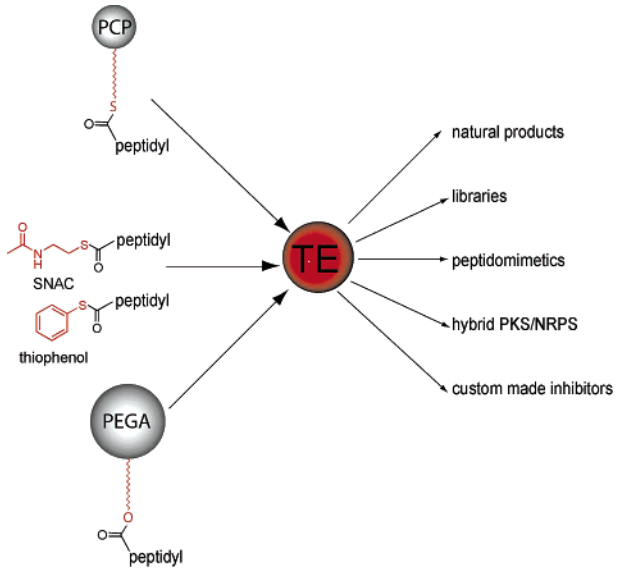

Figure 15. TE domains as versatile catalysts with a potential for the synthesis of new bioactive compounds by a variety of strategies. With recent progress in charging TE domains with activated substrates, a large catalytic toolbox of enzymes awaits chemoenzymatic application for the generation of novel antibiotics. Substrates can be presented to the TE either bound to a PCP or an artificial solid support or by soluble thioester leaving groups.

now allow for applications such as formation of various cyclic peptide libraries.[150] In the future, further exploration of structurally diverse peptide cyclases will help to generate a tool kit of cyclization catalysts which in combination with protein evolution are likely to broaden the applications and increase the utility of enzymatic peptide cyclization. These proteins may rival the utility of lipases as catalysts for stereoselective transformations of diverse substrate molecules.

4. Conclusions

NRPSs are highly sophisticated natural nanomachines that were optimized for the biosynthesis of compounds that cannot be produced by the ribosomal machinery and were selected during evolution for diverse structures and for broad biological activities. Recently, a wealth of information about the three-dimensional structure of several NRPS core domains in combination with detailed biochemical, chemoenzymatic, and genetic studies has not only facilitated the construction of hybrid NRPS but also accelerated the speed by which such bioactive cyclic peptides can be produced. However, some NRPS global structural aspects remain still elusive. One current challenge is the crystallization of modules comprising a multidomain structure that can provide information about domain interaction and the overall architecture within these building blocks of such megaenzymes. Moreover, such structural information could contribute to a precise definition of interdomain linker regions and possible protein−protein interaction sites between the catalytic domains during the concerted action of this assembly line mechanism. This knowledge will have a direct influence on the success of rational engineering attempts, which at present suffer from the lack of this information. In contrast to the well-studied essential domains (A, C, PCP, TE), very little is known about the mechanisms of chemical reactions catalyzed by tailoring domains such as peptide heterocyclization, *N*-methylation, oxidation, reduction, formylation, epimerization, etc. The enzymatic domains carrying out these reactions act within the NRPS assembly line in high precision and efficiency, a fact that makes them attractive for synthetic applications. Chemoenzymatic cyclization by TE domains has already proven this notion and is now established for a set of excised TE domains. Future research will show if this new single-domain catalysis is suitable and potent enough to identify novel drug leads by large cyclic library screens. The utility certainly depends on enzymatic substrate tolerance, turnover, and product yield. In many cases this will need to be optimized by directed protein evolution efforts. Enzyme engineering will further show if other NRPS core and tailoring domains will exhibit the same tolerance in vitro for their desired chemical reactions

5. Acknowledgments

We thank all the members of our group for helpful discussions and support. This work was supported

by the Studienstiftung des deutschen Volkes (SAS) and Deutsche Forschungsgemeinschaft (MAM).

6. References

(1) Flemming, A. *Br. J. Exp. Path.* **1929**, *10*, 226.
(2) Lipmann, F.; Gevers, W.; Kleinkauf, H.; Roskoski, R., Jr. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1971**, *35*, 1.
(3) Walsh, C. T. *Antibiotics: Actions, Origins, Resistance*; ASM Press: Washington, DC, 2003.
(4) Schwarzer, D.; Finking, R.; Marahiel, M. A. *Nat. Prod. Rep.* **2003**, *20*, 275.
(5) Mootz, H. D.; Schwarzer, D.; Marahiel, M. A. *Chembiochem* **2002**, *3*, 490.
(6) Marahiel, M. A.; Stachelhaus, T.; Mootz, H. D. *Chem. Rev.* **1997**, *97*, 2651.
(7) Zasloff, M. *Nature* **2002**, *415*, 389.
(8) Mootz, H. D.; Marahiel, M. A. *J. Bacteriol.* **1997**, *179*, 6843.
(9) Simlot, M. M.; Pfaender, P.; Specht, D. *Hoppe-Seyler's Z. Physiol. Chem.* **1972**, *353*, 759.
(10) Peypoux, F.; Bonmatin, J. M.; Wallach, J. *Appl. Microbiol. Biotechnol.* **1999**, *51*, 553.
(11) de Crecy-Lagard, V.; Blanc, V.; Gil, P.; Naudin, L.; Lorenzon, S.; Famechon, A.; Bamas-Jacques, N.; Crouzet, J.; Thibaut, D. *J. Bacteriol.* **1997**, *179*, 705.
(12) Hubbard, B. K.; Walsh, C. T. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 730.
(13) Steller, S.; Vollenbroich, D.; Leenders, F.; Stein, T.; Conrad, B.; Hofemeister, J.; Jacques, P.; Thonart, P.; Vater, J. *Chem. Biol.* **1999**, *6*, 31.
(14) Steller, S.; Vater, J. *J. Chromatogr., B: Biomed. Sci. Appl.* **2000**, *737*, 267.
(15) Vanittanakom, N.; Loeffler, W.; Koch, U.; Jung, G. *J. Antibiot. (Tokyo)* **1986**, *39*, 888.
(16) Rieber, M. T.; Imaeda, T.; Cesari, I. M. *J. Gen. Microbiol.* **1969**, *55*, 155.
(17) Konz, D.; Klens, A.; Schorgendorfer, K.; Marahiel, M. A. *Chem. Biol.* **1997**, *4*, 927.
(18) Hori, K.; Yamamoto, Y.; Minetoki, T.; Kurotsu, T.; Kanda, M.; Miura, S.; Okamura, K.; Furuyama, J.; Saito, Y. *J. Biochem. (Tokyo)* **1989**, *106*, 639.
(19) Katsu, T.; Kobayashi, H.; Fujita, Y. *Biochim. Biophys. Acta* **1986**, *860*, 608.
(20) Stone, K. J.; Strominger, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 3223.
(21) Storm, D. R.; Strominger, J. L. *J. Biol. Chem.* **1973**, *248*, 3940.
(22) Maget-Dana, R.; Thimon, L.; Peypoux, F.; Ptak, M. *Biochimie* **1992**, *74*, 1047.
(23) Duitman, E. H.; Hamoen, L. W.; Rembold, M.; Venema, G.; Seitz, H.; Saenger, W.; Bernhard, F.; Reinhardt, R.; Schmidt, M.; Ullrich, C.; Stein, T.; Leenders, F.; Vater, J. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 1329.
(24) Maget-Dana, R.; Peypoux, F. *Toxicology* **1994**, *87*, 151.
(25) Hutchison, M. L.; Gross, D. C. *Mol. Plant Microbe Interact.* **1997**, *10*, 347.
(26) Scholz-Schroeder, B. K.; Hutchison, M. L.; Grgurina, I.; Gross, D. C. *MPMI* **2001**, *14*, 336.
(27) Hojati, Z.; Milne, C.; Harvey, B.; Gordon, L.; Borg, M.; Flett, F.; Wilkinson, B.; Sidebottom, P. J.; Rudd, B. A. M.; Hayes, M. A.; Smith, C. P.; Micklefield, J. *Chem. Biol.* **2002**, *9*, 1175.
(28) Alborn, W. E., Jr.; Allen, N. E.; Preston, D. A. *Antimicrob. Agents Chemother.* **1991**, *35*, 2282.
(29) McHenney, M. A.; Hosted, T. J.; Dehoff, B. S.; Rosteck, P. R., Jr.; Baltz, R. H. *J. Bacteriol.* **1998**, *180*, 143.
(30) Canepari, P.; Boaretti, M. *Microb. Drug Resist.* **1996**, *2*, 85.
(31) Di Giambattista, M.; Engelborghs, Y.; Nyssen, E.; Clays, K.; Cocito, C. *Biochemistry* **1991**, *30*, 7277.
(32) Weber, G.; Schörgendorfer, K.; Schneider-Scherzer, E.; Leitner, E. *Curr. Genet.* **1994**, *26*, 120.
(33) Belshaw, P. J.; Schreiber, S. L. *J. Am. Chem. Soc.* **1997**, *119*, 1805.
(34) Liu, J.; Farmer, J. D., Jr.; Lane, W. S.; Friedman, J.; Weissman, I.; Schreiber, S. L. *Cell* **1991**, *66*, 807.
(35) Thern, B.; Rudolph, J.; Jung, G. *Angew. Chem., Int. Ed.* **2002**, *41*, 2307.
(36) Quadri, L. E. *Mol. Microbiol.* **2000**, *37*, 1.
(37) Crosa, J. H.; Walsh, C. T. *Microbiol. Mol. Biol. Rev.* **2002**, *66*, 223.
(38) Sewald, N. *Angew. Chem., Int. Ed.* **2002**, *41*, 4661.
(39) Nicolaou, K. C.; Boddy, C. N. C.; Brase, S.; Winssinger, N. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 2096.
(40) Lipmann, F. *Adv. Microb. Physiol.* **1980**, *21*, 227.
(41) von Döhren, H.; Keller, U.; Vater, J.; Zocher, R. *Chem. Rev.* **1997**, *97*, 2675.
(42) Cosmina, P.; Rodriguez, F.; de Ferra, F.; Grandi, G.; Perego, M.; Venema, G.; van Sinderen, D. *Mol. Microbiol.* **1993**, *8*, 821.
(43) Schwarzer, D.; Marahiel, M. A. *Naturwissenschaften* **2001**, *88*, 93.
(44) Cane, D. E.; Walsh, C. T. *Chem Biol* **1999**, *6*, R319.
(45) Stryer, L. *Biochemistry*, 4th ed.; Spektrum Akademischer Verlag GmbH: Heidelberg-Berlin-Oxford, 1996.
(46) Linne, U.; Marahiel, M. A. *Methods Enzymol.* **2004**, in press.
(47) Turgay, K.; Krause, M.; Marahiel, M. A. *Mol. Microbiol.* **1992**, *6*, 529.
(48) Conti, E.; Franks, N. P.; Brick, P. *Structure* **1996**, *4*, 287.
(49) Gulick, A. M.; Starai, V. J.; Horswill, A. R.; Homick, K. M.; Escalante-Smemerena, J. C. *Biochemistry* **2003**, *42*, 2866.
(50) Conti, E.; Stachelhaus, T.; Marahiel, M. A.; Brick, P. *EMBO J.* **1997**, *16*, 4174.
(51) May, J.; Kessler, N.; Marahiel, M. A.; Stubbs, M. T. *PNAS* **2002**, *99*, 12120.
(52) Stachelhaus, T.; Mootz, H. D.; Marahiel, M. A. *Chem. Biol.* **1999**, *6*, 493.
(53) Gocht, M.; Marahiel, M. A. *J. Bacteriol.* **1994**, *176*, 2654.
(54) Saito, M.; Hori, K.; Kurotsu, T.; Kanda, M.; Saito, Y. *J. Biochem. (Tokyo)* **1995**, *117*, 276.
(55) Starai, V. J.; Celic, I.; Cole, R. N.; Boeke, J. D.; Escalante-Semerena, J. C. *Science* **2002**, *298*, 2390.
(56) Du, L.; Sanchez, C.; Chen, M.; Edwards, D. J.; Shen, B. *Chem. Biol.* **2000**, *7*, 623.
(57) Eppelmann, K.; Stachelhaus, T.; Marahiel, M. A. *Biochemistry* **2002**, *41*, 9718.
(58) Challis, G. L.; Ravel, J.; Townsend, C. A. *Chem. Biol.* **2000**, *7*, 211.
(59) Pavela-Vrancic, M.; Dieckmann, R.; von Döhren, H. *Biochim. Biophys. Acta* **2004**, *1696*, 83.
(60) Ruttenberg, M. A.; Mach, B. *Biochemistry* **1966**, *5*, 2864.
(61) Brückner, R. *Reaktionsmechanismen*; Spektrum Verlag: Heidelberg, 1996.
(62) Stein, T.; Vater, J.; Kruft, V.; Wittmann-Liebold, B.; Franke, P.; Panico, M.; McDowell, R.; Morris, H. R. *FEBS Lett.* **1994**, *340*, 39.
(63) Stein, T.; Vater, J.; Kruft, V.; Otto, A.; Wittmann-Liebold, B.; Franke, P.; Panico, M.; McDowell, R.; Morris, H. R. *J. Biol. Chem.* **1996**, *271*, 15428.
(64) Stachelhaus, T.; Hüser, A.; Marahiel, M. A. *Chem. Biol.* **1996**, *3*, 913.
(65) Schlumbohm, W.; Stein, T.; Ullrich, C.; Vater, J.; Krause, M.; Marahiel, M. A.; Kruft, V.; Wittmann-Liebold, B. *J. Biol. Chem.* **1991**, *266*, 23135.
(66) Lambalot, R. H.; Gehring, A. M.; Flugel, R. S.; Zuber, P.; LaCelle, M.; Marahiel, M. A.; Reid, R.; Khosla, C.; Walsh, C. T. *Chem. Biol.* **1996**, *3*, 923.
(67) Walsh, C. T.; Gehring, A. M.; Weinreb, P. H.; Luis, E. N.; Flugel, R. S. *Curr. Opin. Chem. Biol.* **1997**, *1*, 309.
(68) Quadri, L. E.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P.; Walsh, C. T. *Biochemistry* **1998**, *37*, 1585.
(69) Kealey, J. T.; Liu, L.; Santi, D. V.; Betlach, M. C.; Barr, P. J. *Proc. Natl. Acad. Sci.* **1998**, *95*, 505.
(70) Grünewald, J.; Sieber, S. A.; Marahiel, M. A. *Biochemistry* **2004**, *43*, 2915.
(71) Clugston, S. L.; Sieber, S. A.; Marahiel, M. A.; Walsh, C. T. *Biochemistry* **2003**, *42*, 12095.
(72) Reuter, K.; Mofid, M. R.; Marahiel, M. A.; Ficner, R. *EMBO J.* **1999**, *18*, 6823.
(73) Mofid, R. M.; Finking, R.; Essen, L. O.; Marahiel, M. A. *Biochemistry* **2004**, *43*, 4128.
(74) Weber, T.; Baumgartner, R.; Renner, C.; Marahiel, M. A.; Holak, T. A. *Struct. Fold Des.* **2000**, *8*, 407.
(75) Holak, T. A.; Kearsley, S. K.; Kim, Y.; Prestegard, J. H. *Biochemistry* **1988**, *27*, 6135.
(76) Crump, M. P.; Crosby, J.; Dempsey, C. E.; Parkinson, J. A.; Murray, M.; Hopwood, D. A.; Simpson, T. J. *Biochemistry* **1997**, *36*, 6000.
(77) Parris, K. D.; Lin, L.; Tam, A.; Mathew, R.; Hixon, J.; Stahl, M.; Fritz, C. C.; Seehra, J.; Somers, W. S. *Structure* **2000**, *8*, 883.
(78) Mofid, M. R.; Finking, R.; Marahiel, M. A. *J. Biol. Chem.* **2002**, *277*, 50293.
(79) Schwarzer, D.; Mootz, H. D.; Linne, U.; Marahiel, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 14083.
(80) Belshaw, P. J.; Walsh, C. T.; Stachelhaus, T. *Science* **1999**, *284*, 486.
(81) Ehmann, D. E.; Trauger, J. W.; Stachelhaus, T.; Walsh, C. T. *Chem. Biol.* **2000**, *7*, 765.
(82) Keating, T. A.; Marshall, C. G.; Walsh, C. T.; Keating, A. E. *Nat. Struct. Biol.* **2002**, *9*, 522.
(83) De Crecy-Lagard, V.; Marliere, P.; Saurin, W. *C. R. Acad. Sci. III* **1995**, *318*, 927.
(84) Bergendahl, V.; Linne, U.; Marahiel, M. A. *Eur. J. Biochem.* **2002**, *269*, 620.
(85) Roche, E. D.; Walsh, C. T. *Biochemistry* **2002**, *42*, 1334.
(86) Marshall, C. G.; Hillson. N. J.; Walsh, C. T. *Biochemistry* **2002**, *41*, 244.
(87) Roy, R. S.; Gehring, A. M.; Milne, J. C.; Belshaw, P. J.; Walsh, C. T. *Nat. Prod. Rep.* **1999**, *16*, 249.

(88) Duerfahrt, T.; Eppelmann, K.; Müller, R.; Marahiel, M. A. *Chem. Biol.* **2004**, *11*, 261.
(89) Pfeifer, E.; Pavela-Vrancic, M.; von Dohren, H.; Kleinkauf, H. *Biochemistry* **1995**, *34*, 7450.
(90) Stachelhaus, T.; Walsh, C. T. *Biochemistry* **2000**, *39*, 5775.
(91) Luo, L.; Walsh, C. T. *Biochemistry* **2001**, *40*, 5329.
(92) Hoffmann, K.; Schneider-Scherzer, E.; Kleinkauf, H.; Zocher, R. *J. Biol. Chem.* **1994**, *269*, 12710.
(93) Linne, U.; Marahiel, M. A. *Biochemistry* **2000**, *39*, 10439.
(94) Linne, U.; Doekel, S.; Marahiel, M. A. *Biochemistry* **2001**, *40*, 15824.
(95) Linne, U.; Stein, D. B.; Mootz, H. D.; Marahiel, M. A. *Biochemistry* **2003**, *42*, 5114.
(96) Kessler, N.; Schuhmann, H.; Morneweg, S.; Linne, U.; Marahiel, M. A. *J. Biol. Chem.* **2004**, *279*, 7413.
(97) Velkov, T.; Lawen, A. *Biotechnol. Annu. Rev.* **2003**, *9*, 151.
(98) Billich, A.; Zocher, R.; Kleinkauf, H.; Braun, D. G.; Lavanchy, D.; Hochkeppel, H. K. *Biol. Chem. Hoppe Seyler* **1987**, *368*, 521.
(99) Schauwecker, F.; Pfennig, F.; Grammel, N.; Keller, U. *Chem. Biol.* **2000**, *7*, 287.
(100) de Crecy-Lagard, V.; Saurin, W.; Thibaut, D.; Gil, P.; Naudin, L.; Crouzet, J.; Blanc, V. *Antimicrob. Agents Chemother.* **1997**, *41*, 1904.
(101) Miller, D. A.; Walsh, C. T.; Luo, L. *J. Am. Chem. Soc.* **2001**, *123*, 8434.
(102) Weinig, S.; Hecht, H. J.; Mahmud, T.; Müller, R. *Chem. Biol.* **2003**, *10*, 939.
(103) Du, L.; Chen, M.; Sanchez, C.; Shen, B. *FEMS Microbiol. Lett.* **2000**, *189*, 171.
(104) Schneider, T. L.; Shen, B.; Walsh, C. T. *Biochemistry* **2003**, *42*, 9722.
(105) Silakowski, B.; Schairer, H. U.; Ehret, H.; Kunze, B.; Weinig, S.; Nordsiek, G.; Brandt, P.; Blocker, H.; Hofle, G.; Beyer, S.; Muller, R. *J. Biol. Chem.* **1999**, *274*, 37391.
(106) Shen, B.; Du, L.; Sabchez, C.; Edwards, D. J.; Chen, M.; Murell, J. M. *J. Nat. Prod.* **2002**, *65*, 1262.
(107) Weinig, S.; Mahmud, T.; Müller, R. *Chem. Biol.* **2003**, *10*, 953.
(108) Reimmann, C.; Patel, H. M.; Serino, L.; Barone, M.; Walsh, C. T.; Patel, H. M. *J. Bacteriol.* **2001**, 813.
(109) Rouhiainen, L.; Paulin, L.; Suomalainen, S.; Hyytiainen, H.; Buikema, W.; Haselkorn, R.; Sivonen, K. *Mol. Microbiol.* **2000**, *37*, 156.
(110) Tseng, C. C.; Bruner, S. D.; Kohli, R. M.; Marahiel, M. A.; Walsh, C. T.; Sieber, S. A. *Biochemistry* **2002**, *41*, 13350.
(111) Ehmann, D. E.; Gehring, A. M.; Walsh, C. T. *Biochemistry* **1999**, *38*, 6171.
(112) Bruner, S. D.; Weber, T.; Kohli, R. M.; Schwarzer, D.; Marahiel, M. A.; Walsh, C. T.; Stubbs, M. T. *Structure (Cambridge)* **2002**, *10*, 301.
(113) Tsai, S. C.; Miercke, L. J.; Krucinski, J.; Gokhale, R.; Chen, J. C.; Foster, P. G.; Cane, D. E.; Khosla, C.; Stroud, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14808.
(114) Kohli, R. M.; Takagi, J.; Walsh, C. T. *PNAS* **2002**, *99*, 1247.
(115) Kohli, R. M.; Trauger, J. W.; Schwarzer, D.; Marahiel, M. A.; Walsh, C. T. *Biochemistry* **2001**, *40*, 7099.
(116) Sieber, S. A.; Tao, J.; Walsh, C. T.; Marahiel, M. A. *Angew. Chem.* **2004**, *116*, 499.
(117) Shaw-Reid, C. A.; Kelleher, N. L.; Losey, H. C.; Gehring, A. M.; Berg, C.; Walsh, C. T. *Chem. Biol.* **1999**, *6*, 385.
(118) Weber, G.; Leitner, E. *Curr. Genet.* **1994**, *26*, 461.
(119) Haese, A.; Schubert, M.; Herrmann, M.; Zocher, R. *Mol. Microbiol.* **1993**, *7*, 905.
(120) Becker, J. E.; Moore, R. E.; Moore, B. S. *Gene* **2004**, *325*, 35.
(121) Witkowski, A.; Joshi, A.; Smith, S. *Biochemistry* **1996**, *35*, 10569.
(122) Kao, C. M.; Pieper, R.; Cane, D. E.; Khosla, C. *Biochemistry* **1996**, *35*, 12363.
(123) Staunton, J.; Caffrey, P.; Aparicio, J. F.; Roberts, G. A.; Bethell, S. S.; Leadlay, P. F. *Nat. Struct. Biol.* **1996**, *3*, 188.
(124) Rangan, V. S.; Joshi, A. K.; Smith, S. *J. Biol. Chem.* **1998**, *273*, 34949.
(125) Witkowski, A.; Joshi, A. K.; Rangan, V. S.; Falick, A. M.; Witkowska, H. E.; Smith, S. *J. Biol. Chem.* **1999**, *274*, 11557.
(126) Sieber, S. A.; Linne, U.; Hillson, N. J.; Roche, E.; Walsh, C. T.; Marahiel, M. A. *Chem. Biol.* **2002**, *9*, 997.
(127) Glinski, M.; Urbanke, C.; Hornbogen, T.; Zocher, R. *Arch. Microbiol.* **2002**, 178.
(128) Joshi, A. K.; Vangipuram, S. R.; Witkowski, A.; Smith, S. *Chem. Biol.* **2003**, *10*, 169.
(129) Hillson, N. J.; Walsh, C. T. *Biochemistry* **2003**, *42*, 766.
(130) Richardt, A.; Kemme, T.; Wagner, S.; Schwarzer, D.; Marahiel, M. A.; Hovemann, B. T. *J. Biol. Chem.* **2003**, *278*, 41160.
(131) Kasahara, T.; Kato, T. *Nature* **2003**, *422*, 832.
(132) Joshi, A. K.; Zhang, L.; Rangan, V. S.; Smith, S. *J. Biol. Chem.* **2003**, *278*, 33142.
(133) Bollag, D. M.; McQueney, P. A.; Zhu, J.; Hensens, O.; Koupal, L.; Liesch, J.; Goetz, M.; Lazarides, E.; Woods, C. M. *Cancer Res.* **1995**, *55*, 2325.
(134) Molnar, I.; Schupp, T.; Ono, M.; Zirkle, R.; Milnamow, M.; Nowak-Thompson, B.; Engel, N.; Toupet, C.; Stratmann, A.; Cyr, D. D.; Gorlach, J.; Mayo, J. M.; Hu, A.; Goff, S.; Schmid, J.; Ligon, J. M. *Chem. Biol.* **2000**, *7*, 97.
(135) Tang, L.; Shah, S.; Chung, L.; Carney, J.; Katz, L.; Khosla, C.; Julien, B. *Science* **2000**, *287*, 640.
(136) Stachelhaus, T.; Schneider, A.; Marahiel, M. A. *Science* **1995**, *269*, 69.
(137) Schneider, A.; Stachelhaus, T.; Marahiel, M. A. *Mol. Gen. Genet.* **1998**, *257*, 308.
(138) Mootz, H. D.; Schwarzer, D.; Marahiel, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5848.
(139) Schwarzer, D.; Mootz, H. D.; Marahiel, M. A. *Chem. Biol.* **2001**, *8*, 997.
(140) Duerfahrt, T.; Doekel, S.; Sonke, T.; Quaedflieg, P. J. L. M.; Marahiel, M. A. *Eur. J. Biochem.* **2003**, *270*, 4555.
(141) Mootz, H. D.; Kessler, N.; Linne, U.; Eppelmann, K.; Schwarzer, D.; Marahiel, M. A. *J. Am. Chem. Soc.* **2002**, *124*, 10980.
(142) Davies, J. S. *J. Pept. Sci.* **2003**, *9*, 471.
(143) Sewald, N.; Jakubke, H.-D. *Peptides: Chemistry and Biology*; Wiley-VCH: Weinheim, 2002.
(144) Sieber, S. A.; Marahiel, M. A. *J. Bacteriol.* **2003**, *185*, 7036.
(145) Kohli, R. M.; Walsh, C. T. *Chem. Commun. (Cambridge)* **2003**, 297.
(146) Trauger, J.; Kohli, R.; Mootz, H.; Marahiel, M.; Walsh, C. *Nature* **2000**, *407*, 215.
(147) Trauger, J. W.; Kohli, R. M.; Walsh, C. T. *Biochemistry* **2001**, *40*, 7092.
(148) Bu, X.; Wu, X.; Xie G.; Guo, Z. *Org. Lett.* **2002**, *4*, 2893.
(149) Fernandez-Lopez, S.; Kim, H. S.; Choi, E. C.; Delgado, M.; Granja, J. R.; Khasanov, A.; Kraehenbuehl, K.; Long, G.; Weinberger, D. A.; Wilcoxen, K. M.; Ghadiri, M. R. *Nature* **2001**, *412*, 452.
(150) Kohli, R. M.; Walsh, C. T.; Burkart, M. D. *Nature* **2002**, *418*, 658.
(151) Kohli, R. M.; Burke, M. D.; Tao, J.; Walsh, C. T. *J. Am. Chem. Soc.* **2003**, *125*, 7160.
(152) Garbe, D.; Sieber, S. A.; Bandur, N.; Koert, U.; Marahiel, M. A. *Chembiochem* **2004**, *5*, 1000.
(153) Sieber, S. A.; Walsh, C. T.; Marahiel, M. A. *J. Am. Chem. Soc.* **2003**, *125*, 10862.
(154) Dawson, P. E.; Churchill, M. J.; Ghadiri, M. R.; Kent, S. B. H. *J. Am. Chem. Soc.* **1996**, *119*, 4325.
(155) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci.* **1998**, *95*, 6705.
(156) Chan, W. C.; White P. D. *Fmoc Solid-Phase Peptide Synthesis*; Oxford University Press: Oxford, 2000.
(157) Zhu, J.; Ma, D. *Angew. Chem.* **2003**, *43*, 5506.
(158) Schmidt, R.; Neubert, K. *Int. J. Pept. Protein Res.* **1991**, *37*, 502.
(159) Ehrlich, A.; Heyne, H.-U.; Winter, R.; Beyermann, M.; Haber, H.; Carpino, L. A.; Bienert, M. *J. Org. Chem.* **1996**, *61*, 8831.
(160) Jackson, D. Y.; Burnier, J. P.; Wells, J. A. *J. Am. Chem. Soc.* **1995**, *117*, 819.
(161) Muir, T. W. *Annu. Rev. Biochem.* **2003**, *72*, 249.
(162) Nishino, N.; Xu, M.; Mihara, H.; Fujimoto, T.; Ueno, Y.; Kumagai, H. *Tetrahedron Lett.* **1992**, *33*, 1479.
(163) Klibanov, A. M. *Nature* **2001**, *409*, 241.

CR0301191