Nonribosomal Peptides

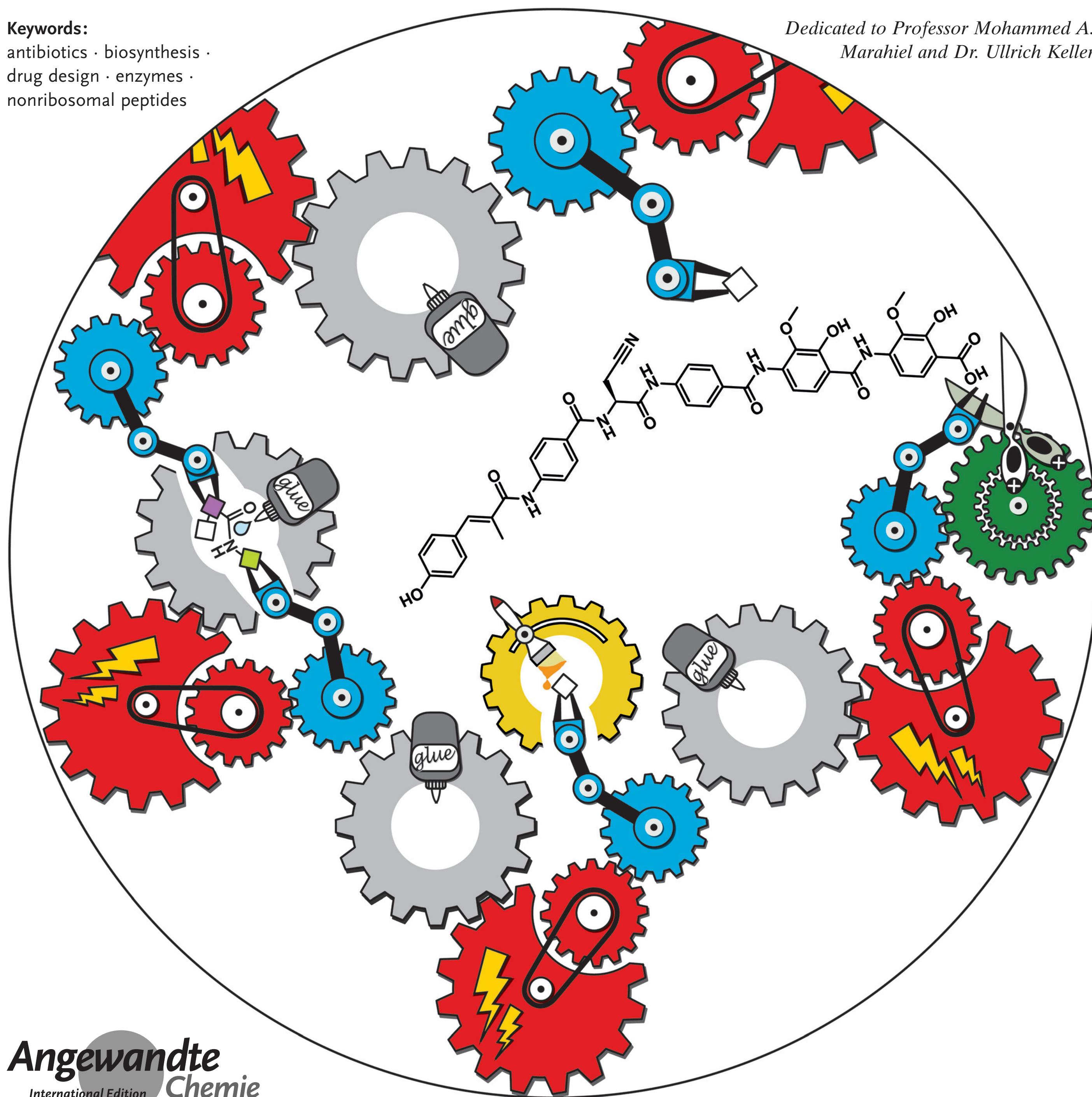
International Edition: DOI: 10.1002/anie.201609079
German Edition: DOI: 10.1002/ange.201609079

Nonribosomal Peptide Synthesis—Principles and Prospects

Roderich D. Süssmuth and Andi Mainz*

Keywords:
antibiotics · biosynthesis · drug design · enzymes · nonribosomal peptides

Dedicated to Professor Mohammed A. Marahiel and Dr. Ullrich Keller



 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Nonribosomal peptide synthetases (NRPSs) are large multienzyme machineries that assemble numerous peptides with large structural and functional diversity. These peptides include more than 20 marketed drugs, such as antibacterials (penicillin, vancomycin), antitumor compounds (bleomycin), and immunosuppressants (cyclosporine). Over the past few decades biochemical and structural biology studies have gained mechanistic insights into the highly complex assembly line of nonribosomal peptides. This Review provides state-of-the-art knowledge on the underlying mechanisms of NRPSs and the variety of their products along with detailed analysis of the challenges for future reprogrammed biosynthesis. Such a reprogramming of NRPSs would immediately spur chances to generate analogues of existing drugs or new compound libraries of otherwise nearly inaccessible compound structures.

From the Contents

1. Introduction 3771
2. Amino Acid Building Blocks—The Basis for Structural Diversity 3779
3. Architecture and Mechanisms of NRPSs 3784
4. Exploitation of NRPS-Based Pathways 3806
5. Outlook 3814

1. Introduction

Ribosomal synthesis is a fundamental process for the synthesis of peptides and proteins. However, alternative solutions for the formation of amide bonds exist in nature: Ligase-mediated reactions involved in the formation of glutathione, in the ubiquitinylation of proteins triggering protein degradation,[1,2] or in early steps of the synthesis of bacterial cell walls.[3,4] A more recently discovered pathway is the tRNA-dependent biosynthesis of some diketopiperazines.[5,6] In the past 50 years researchers have established another major alternative biosynthesis pathway, namely the nonribosomal peptide (NRP) synthesis performed by dedicated nonribosomal peptide synthetases (NRPSs) that are mainly found in bacteria und fungi.

1.1. History

In the early 1960s, when the ribosomal code had been deciphered, researchers investigated how certain cyclic peptides containing D-amino acids were synthesized by *Bacillus* species. A study by Tatum and co-workers[7] showed that the cell-based biosynthesis of tyrocidine was not affected by ribosome inhibitors such as aureomycin (chlorotetracyclin), and from this they hypothesized a mechanism distinct from protein synthesis. The field gained momentum through contributions from the group of the Nobel Prize Laureate Fritz Lipmann, and of Søren Laland,[8] which gave fundamental biochemical and mechanistic insights into NRPSs, including specific ATP-dependent activation of amino acids, thioester-mediated 4′-phosphopanthetheine (Ppant) binding of activated amino acids,[9–12] and the directionality of the peptide synthesis.[11,13,14] Interestingly, subsequent research on bacterial and fungal antibiotics was mostly focused on peptides of nonribosomal origin. Only in 1988 it was shown through the example of the lantibiotic epidermin that peptide antibiotics containing unusual structural modifications are also synthesized ribosomally by microorganisms.[15]

Initially, the work with NRPSs used classical biochemical purification methods developed for spores, mycelia, and cellular extracts. The characterization of their enzymatic activity was achieved by using radiolabeled substrates. Technologies that revolutionized the work with NRPSs were the cloning of genes and gene fragments, their expression as proteins, and the in vitro reconstitution of enzymatic steps or even of the entire biosynthesis. In parallel, techniques for the directed inactivation of genes were developed, predominantly for Actinomycetes, which facilitated the isolation of biosynthetic intermediates. Nowadays, DNA sequencing techniques enable whole microbial genomes to be sequenced almost routinely.

1.2. Origin

The producers of NRPS-based metabolites are mostly bacteria and fungi. Higher-order organisms, for example, sponges, were also considered, but contaminations from symbiotic microorganisms can lead to false assumptions. Nevertheless, the NRPS Ebony from *Drosophila melanogaster* ("fruit fly") as well as the nemamide synthetase from the nematode *Caenorhabditis elegans* seem to be proven examples outside of bacteria and fungi.[16,17]

Screening efforts and more recently genome sequencing projects followed by bioinformatic analyses have already led to quite an insightful picture into the distribution and occurrence of NRPS pathways and their products.[18] Among bacteria, the most prolific contributors are the phyla Actinobacteria, Firmicutes, and classes α-/β-/γ-Proteobacteria, but Cyanobacteria and the class δ-Proteobacteria have received increased focus more recently (Table 1). Fungal NRPS-based metabolites mostly derive from Ascomycota (Table 2), whereas Basidiomycota are hardly represented. Studies of

[*] Prof. Dr. R. D. Süssmuth, Dr. A. Mainz
Technische Universität Berlin, Institut für Chemie
Strasse des 17. Juni 124, 10623 Berlin (Germany)
E-mail: roderich.suessmuth@tu-berlin.de

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 1: Important bacterial phyla (classes) and genera containing NRPS genes.[a]

Bacteria	Phylum (class$^+$)	Genus	Representative compounds
G+	Actinobacteria	*Streptomyces* *Mycobacterium**	various, e.g. glycopeptide antibiotics mycobactin (siderophore)
	Cyanobacteria	*Microcystis*, *Planktothrix*, *Anabaena*, *Oscillatoria*, *Nostoc*	various cyanotoxins, e.g. microcystins, cyanopeptolins, cryptophycin
	Firmicutes	*Bacillus* *Staphylococcus** *Streptococcus**	lipocyclopeptides, e.g. surfactin aureusimine mutanobactin
G−	β-Proteobacteria$^+$	*Burkholderia**	malleobactin (siderophore)
	γ-Proteobacteria$^+$	*Pseudomonas** *Escherichia*/Salmonella*/Yersinia*/Vibrio**, *Serratia*, *Erwinia* *Photorhabdus*	syringomycin, pyoverdin siderophores, e.g. enterobactin, salmochelin, yersiniabactin, vibriobactin various linear and cyclic peptides
	δ-Proteobacteria$^+$	*Myxobacterium* (order)	argyrin, PK-NRP hybrids (tubulysin, epothilone)

[a] G+ = Gram-positive; G− = Gram-negative; * = genus contributing significant human pathogens.

Table 2: Important fungal phyla (subkingdom Dikarya) and genera containing NRPS genes.

Phylum	Genus	Representative compounds
Ascomycota	*Tolypocladium*	cyclosporine A
	Fusarium	enniatins
	Penicillium	penicillin V
	Acremonium	cephalosporin C
	Claviceps	ergopeptins, e.g. ergotamine
	Trichoderma	peptaiboles, e.g. alamethicin
Basidiomycota	*Ustilago*	ferrichrome

fungal NRP biosynthesis lag somewhat behind that of bacteria: fungi are less explored due to their often larger genome sizes, the scattered presence of introns in the gene clusters, and a less-established molecular-biology toolbox.

1.3. Gene Clusters and Biosynthesis

In addition to the easier identification of NRPS genes by modern genome mining tools, NRPS genes are comparatively easy to detect due to their large multidomain organization. In bacteria, the biosynthesis genes of secondary metabolites are commonly found in so-called gene clusters, which is also often the case for fungi. While NRPS genes are considered to be the core of the clusters, they are accompanied by genes for the synthesis of building blocks, product decoration, self-resistance, and peptide export. Advanced genome sequencing techniques have enabled genome mining[19,20] approaches, which are supported by a variety of bioinformatic tools (e.g. AntiSMASH,[21,22] PRISM,[23] and SMURF)[24] for the in silico discovery and analysis of NRPS pathways.[25]

1.4. Structural Complexity

The currently known NRP structures reflect the complexity and abundance of certain structural classes: The largest group is probably head-to-tail-cyclized peptides of various ring sizes (e.g. gramicidin S, cyclosporine) as well as lipocyclopeptides with different linking patterns (e.g. surfactin, iturin, fengycin). Linear peptide structures are also abundant and range from tripeptides (e.g. sevadicin, bialaphos) to 20-mer peptides (peptaiboles, e.g. alamethicin). Apparently, the

Roderich D. Süssmuth studied chemistry at the University of Tübingen and received his PhD with Prof. Günther Jung. After a postdoctoral stay with Carlos Barbas and Richard Lerner at The Scripps Research Institute (San Diego, CA, USA), he habilitated in Organic Chemistry and Biochemistry at Tübingen University. Since 2004, he has been the Rudolf-Wiechert-Professor of Biological Chemistry at the Technical University of Berlin. His research interests are the discovery and biosynthetic investigations of secondary metabolites, biological and chemical peptide synthesis, and medicinal chemistry.

Andi Mainz studied biochemistry at the University of Greifswald and completed his diploma in 2007 in the structural biology department of Sanofi-Aventis Deutschland GmbH in Frankfurt a.M. In 2012 he received his PhD in biophysics, with a focus on biomolecular NMR spectroscopy under the supervision of Prof. Bernd Reif at the Leibniz-Institute of Molecular Pharmacology in Berlin and the Helmholtz-Zentrum München. After a postdoctoral stay with Prof. Reif at TU Munich, he joined the group of Prof. Roderich D. Süssmuth in 2014 to study the biosynthetic machineries of secondary metabolites.

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

current upper size limit for NRPs is 25 amino acids (syringopeptin 25A; Scheme 3). Some NRPs undergo massive structural modifications by tailoring enzymes, and the structurally most complex compounds known are probably the β-lactams, the glycopeptide antibiotics, the ergopeptins, and the bleomycins (for examples see Schemes 3 and 4).

1.5. Function

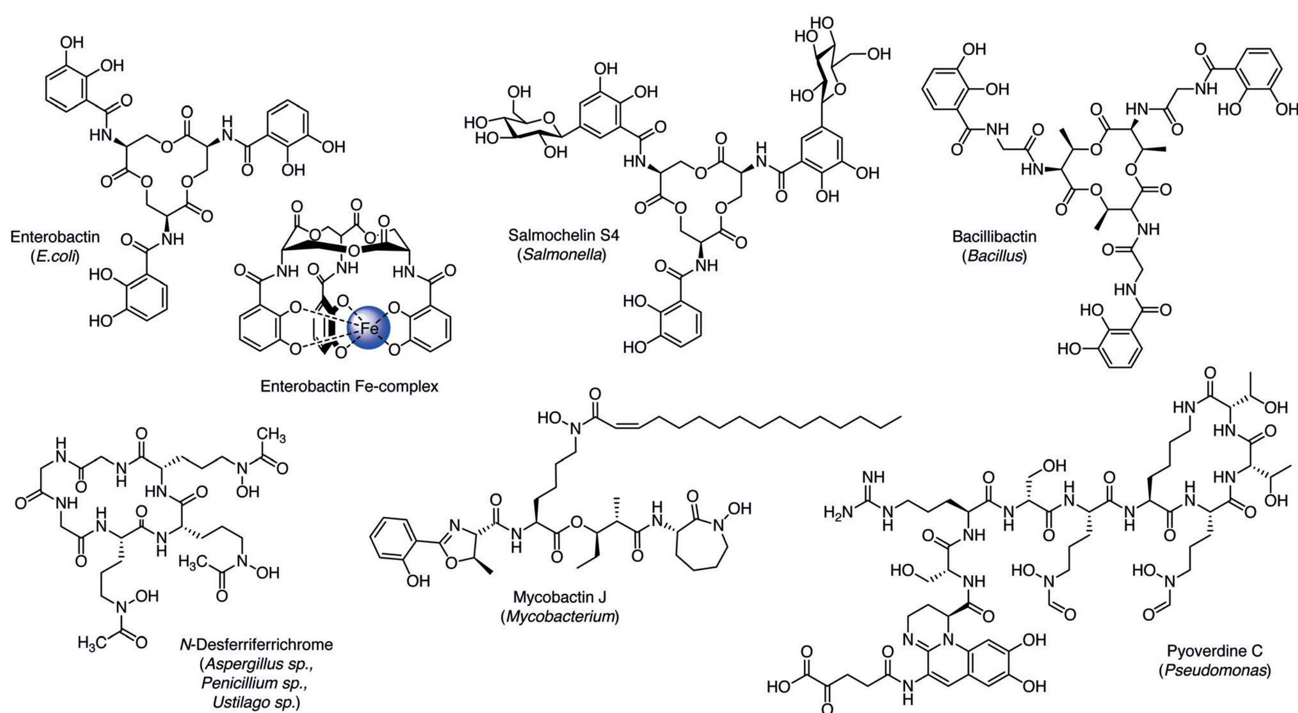
The function of secondary metabolites and their use for the producing organism is the subject of scientific debate.[26] For those from nonpathogenic bacterial strains—for example, various Actinomycetes—a communication or signaling function seems likely. Although it is known that various fungal products are mycotoxins, examples of bacterial NRPs playing a distinct role in pathogenesis have more recently accumulated.[27–29]

1.6. Virulence Factors and Toxins

A large and structurally diverse compound group that serves as virulence factors is the siderophores, which are also synthesized by nonpathogenic bacteria and fungi. Siderophores are exported under low iron levels into the surrounding and are reimported as their Fe^{III} complexes to secure an iron supply for cellular processes. Prominent representatives are enterobactin (*E. coli*) and salmochelin (*Salmonella*),[30] bacillibactin (*Bacillus anthracis*, *B. subtilis*), pyoverdine (*P. aeruginosa*), and mycobactin (*Mycobacterium* sp.; Scheme 1). Aryl acid adenylates as present in enterobactin were used as biosynthesis inhibitors to interfere with the growth of pathogenic bacteria,[31,32] but this approach was not pursued. Interestingly, the mushroom *Ustilago maydis*—which causes corn smut and is a delicacy in Mexico, where it is known as *huitlacoche* (Atztec language)—contains a ferrichrome biosynthesis gene cluster.[33]

Although the physiological role of toxins for the producing organism may not always be clear, the consequences of ingestion by animals or humans can be either acute or chronic, and range from irritating, allergenic, neurotoxic, or hepatotoxic to carcinogenic and mutagenic effects. From a historic perspective, an important group of fungal toxins is the ergot alkaloids, for example, ergotamine, synthesized by the ergot fungus *Claviceps purpurea*. In medieval Europe, moist seasons led to the massive growth of ergot fungus predominantly on rye. The harvest of grains together with the sclerotia (fruiting bodies) and the subsequent consumption of porridge and breads caused severe intoxications, also known as St. Anthony's Fire (lat. *ignis sacer*; Figure 1). Major symptoms of ergotism were convulsions (spasms, psychosis) or gangrene (necrosis of extremities as a result of vasoconstrictive effects). In this context, it is worth mentioning that lysergic acid diethylamide (LSD, Scheme 2), a derivative of the ergoline family, has an infamous history as a psychedelic drug.

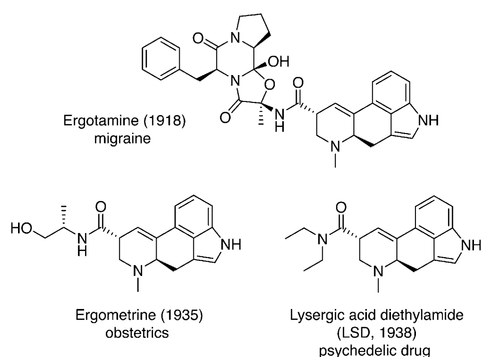
The relationship between fungal infections of plants, particularly of crops, and the production of mycotoxins is evident for *Aspergillus*, *Penicillium*, *Alternaria*, and *Fusarium* species. A major class of toxins comprises diketopiperazine-type peptides, for example, chaetomin, gliotoxin, roquefortin, verruculogen, and fumitremorgin A (Scheme 3).[34] Gliotoxin (*Aspergillus* sp.) has immunosuppressive activities and siro-



Scheme 1. Structures of bacterial and fungal siderophores.

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Figure 1. Gangrene in European art: Left) Piece ("The Temptation of St. Anthony") of the Isenheimer Altar by Matthias Grünewald (Colmar, Alsace). The man shows symptoms of ergotism. Right) "The Cripples" by Pieter Bruegel the Elder (1568). The loss of the lower extremities has been assigned to gangrene.



Scheme 2. Structures of ergot alkaloids.

desmin PL is a non-host-selective phytotoxin produced by the fungus *Leptosphaeria maculans*, which causes blackleg disease of canola (*Brassica napus*). Larger cyclopeptides from plant-pathogenic fungi include trapoxin A (*Helicoma ambiens*), chlamydocin (*Diheterospora chlamydosporia*), alternariolide (AM-toxin, *Alternaria alternate* pv. *Mali*), cyclochlorotine (*Penicillium islandicum* sp.), victorin (*Cochliobolus victoriae*), and apicidin (*Fusarium* sp.). Molecular targets have been determined for some mycotoxins, for example, for the histone deacetylase inhibitor HC-toxin (*Cochliobolus carbonum*), for tentoxin (*Alternaria alternata*) and another tetrapeptide affecting chloroplast development, and for the cytochalasins (*Phoma exigua*, *Zygosporium masonii*) that inhibit actin polymerization.

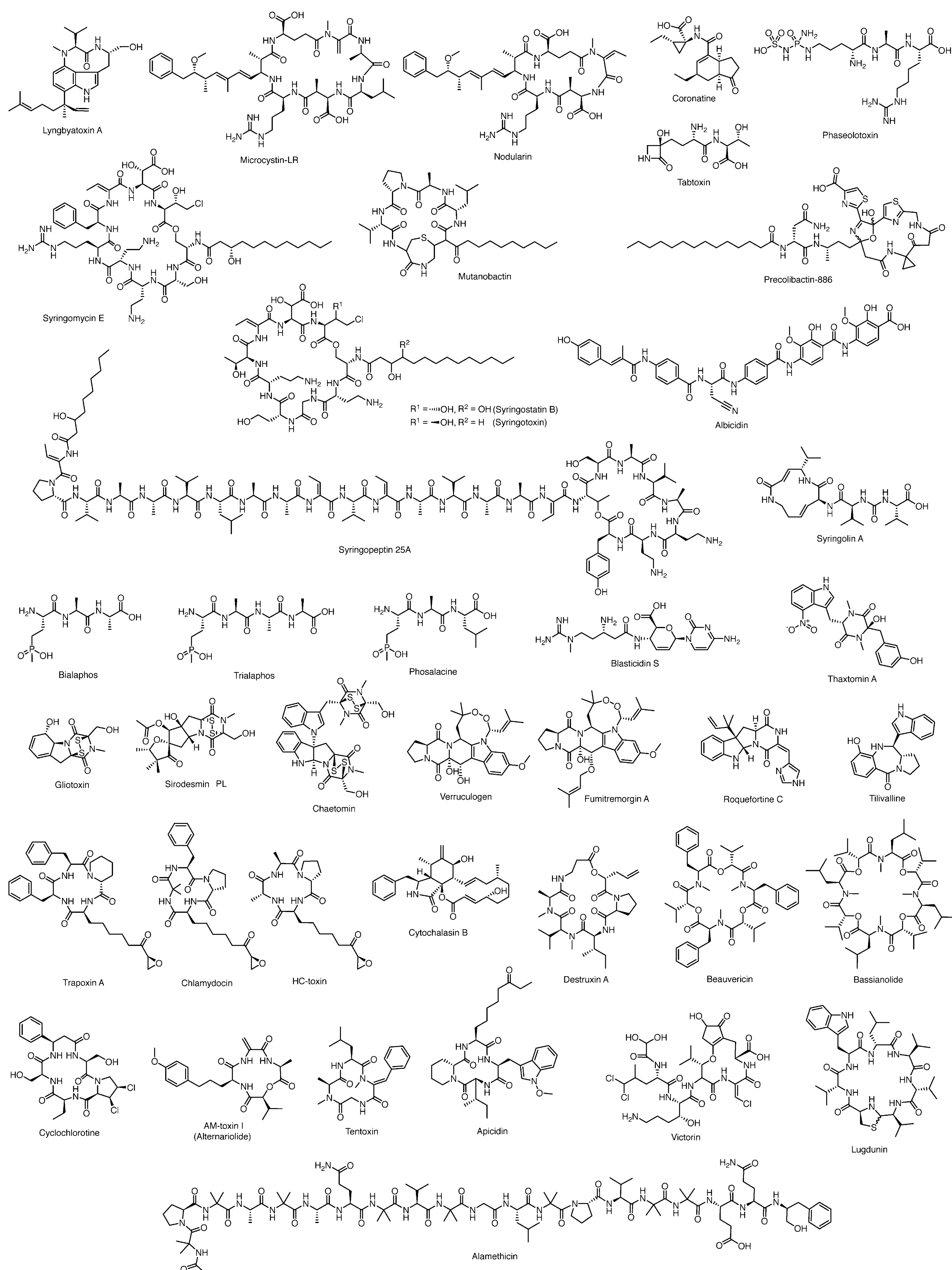
Bacterial infections of plants[34] are mediated by virulence factors also termed as toxins. The inactivation of genes that synthesize these virulence factors leads, in most cases, to a significantly reduced pathogenicity of the producing strain. In this context, the Pseudomonads produce various plant-targeted virulence factors, for example, the pore-forming syringomycin, syringostatin, syringopeptin, and the proteasome inhibitor syringolin (*P. syringae*).[35] Coronatine and the monobactam tabtoxin (*P. syringae*)[36] are further Pseudomonas toxins of some pathovars. As a protoxin, tabtoxin is presumably hydrolyzed to generate the glutamine synthetase inhibitor tabtoxinine β-lactam.[37] Similarly, the hydrolysis product of the tripeptide protoxin phaseolotoxin (*P. syringae*) is an ornithine decarboxylase inhibitor. Mechanistically, there is an apparent analogy to bialaphos (L-alanyl-L-alanyl-phosphinothricin) from *Streptomyces hygroscopicus*, which is also synthesized as a protoxin. The nontoxic protoxin is hydrolyzed to generate the glutamine synthetase inhibitor phosphinothricin also known under the name glufosinate, one of the most successful commercial herbicides. More recently, the gyrase inhibitor albicidin (*Xanthomonas albilineans*) which causes leaf scald disease in sugar cane has been structurally elucidated.[27] Blastidin S, a nucleoside mimic and peptidyltransferase inhibitor from *Streptomyces griseochromogenes*, has phytopathogenic and also fungicidal activity. Finally, the diketopiperazine-type compound thaxtomin A produced by the bacterium *Streptomyces scabies* is an inhibitor of cellulose synthesis and causes the common potato scab.

Cyanobacteria are well-known for the production of various peptidic cyanotoxins: lyngbyatoxin A (dermatoxic), cyanopeptolins (ecotoxic), nodularins (hepatotoxic), and the microcystins (hepatotoxic) are mostly synthesized by the genera *Anabena*, *Microcystis*, *Nodularia*, *Planktothrix* (*Oscillatoria*), and *Lyngbya*. These toxins are particularly relevant to aquatic animals. In addition, seasonal algae blooms can cause severe intoxications to humans, for example, shellfish poisoning and even poisonings from drinking-water reservoirs. More recently, it was found that NRPs from bacterial genera which are part of the human flora (skin, mucosa, intestine) may also have an influence on pathogenesis: mutanobactin A (*Staphylococcus mutans*),[38] colibactin (*Escherichia coli*),[39] tilivalline (*Klebsiella oxytoca*),[40] and lugdunin (*Staphylococcus lugdunensis*).[28] Fungal infections of humans are less common and are mostly an indicator of a severely immunocompromised health status. Interestingly, insecticidal NRPs also exist, for example, destruxin synthesized by the Ascomycete *Metarhizium anisopliae* causes the green muscardine disease in insects. Likewise, the fungus *Beauveria bassiana* produces beauvericin and bassianolide (white muscardine disease) and is used as a biopesticide.

1.7. Human Use

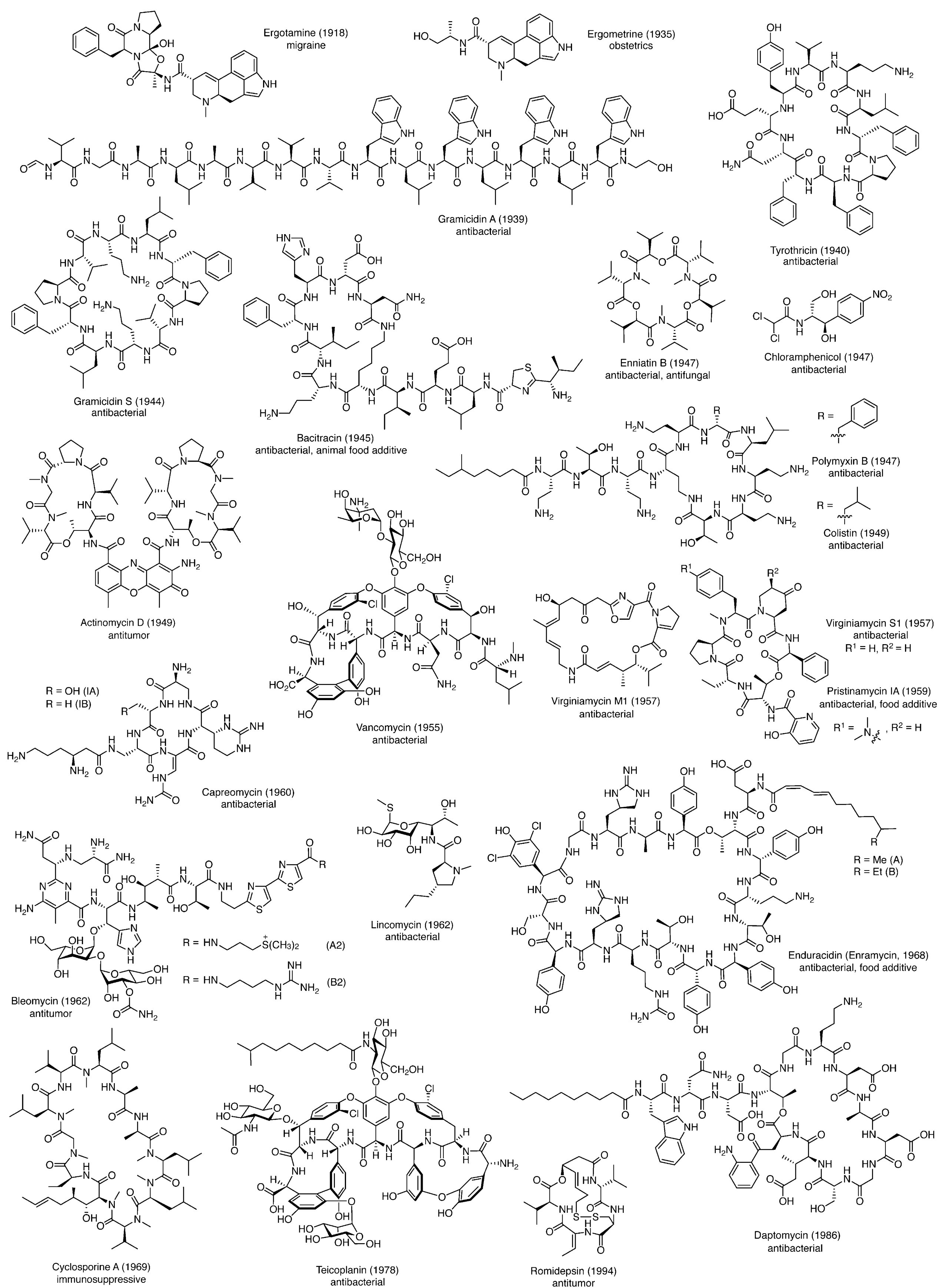
The usefulness of NRPs as drugs is evident. A survey of currently marketed drugs shows nearly 30 NRP (core) structures, which contribute to sales of $/€ billions in the chemical and pharmaceutical industry. Their predominant use is as systemic and topical antibacterials, followed by antitumor drugs, antifungals, and animal feed additives (Scheme 4 and Table 3). There are also important applications as immunosuppressants (cyclosporine), or in obstetrics (ergometrine) and pain treatment (ergotamine). Although the use of antibiotics as additives in animal feed was banned in the European Union in 2006, virginiamycin and bacitracin are still used in other parts of the world. A noteworthy application is the use of emodepside as a semisynthetic anthelmintic peptide in pet care, which is currently under consideration for the treatment of human infections by parasitic worms,[41] foremost onchocerciasis (river blindness) and elephantiasis, which affects hundreds of millions of people.

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 3. Bacterial and fungal virulence factors and toxins produced by NRPS.

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 4. Marketed NRP drugs and important representatives of structural classes and the year they were reported.

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 3: Overview of marketed NRP drugs.

Agent	Origin, producing organism(s)[a]	Marketed[b]	Properties and area of application	Mode of action
actinomycin D (dactinomycin)	*Actinomyces antibioticus* (B), *Streptomyces chrysomallus* (B)	1964	antitumor, (antibacterial: high toxicity)	DNA intercalator, inhibition of transcription
bacitracin	*Bacillus subtilis group* (B), *Bacillus licheniformis* (B)	1948	antibacterial (topical; Gram-positive), animal health feed	bacterial cell-wall biosynthesis (peptidoglycan)
bialaphos	*Streptomyces hygroscopicus*, *S. viridochromogenes* (B)	ca. 1984	herbicide; (phosphinothricin (=glufosinate): synthetic herbicide)	tripeptide prodrug, inhibitor of glutamine synthetase
bleomycin A2,B2	*Streptomyces verticillus* (B)	1969	antitumor (Hodgkin's lymphoma, testicular, ovarian, cervical cancers)	metal-dependent oxidative cleavage of DNA in presence of molecular oxygen
carbapenems[c]	synthetic thienamycin (*Streptomyces cattleya* (B)) analogues, e.g. imipenem	1985	antibacterial (multidrug resistant)	bacterial cell-wall biosynthesis (peptidoglycan; β-lactamase inhibition)
capreomycin IA+IB	*Streptomyces capreolus* (B)	1971	antituberculous (nephrotoxic, ototoxic)	inhibition of the ribosomal protein synthesis (16S and 23S-rRNA)
carfilzomib[c]	synthetic derivative of epoxomycin (*Actinomyces* sp. (B))	2012	anticancer (multiple myeloma)	proteasome inhibitor
caspofungin[d] (MK-0991)	*Glarea lozoyensis* (F), semisynthetic from pneumocandin; further derivatives: micafungin[d]/ anidulafungin[d]	2001 2005/2006	antifungal (candidiasis, aspergillosis)	fungal cell-wall integrity ((1→3)-β-D-glucan synthase)
cephalosporins[d],[e]	*Acremonium chrysogenum* (F), >50 marketed derivatives	1964	antibacterial	bacterial cell-wall biosynthesis (peptidoglycan)
chloramphenicol	*Streptomyces venezuelae* (B); synthetic; further derivatives: thiamphenicol[c], florfenicol[c]	1949	antibacterial (human and veterinary use; florfenicol veterinary use)	inhibition of ribosomal protein synthesis
colistin (polymyxin E)	*Paenibacillus polymyxa* var. colistinus (B)	1958	antibacterial	binding to lipopolysaccharide (outer membrane), interaction with the cytoplasmic membrane
cyclosporine A	*Tolypocladium inflatum* (F)	1983	immunosuppressive (inhibition of transplant rejection), autoimmune diseases	cyclophilin binding, inhibition of IL-2 expression (inhibition of T-cell activation)
dalbavancin	semisynthetic teicoplanin derivative	2014	antibacterial (Gram-positive)	membrane anchoring; disruption of cell membrane and inhibition of bacterial cell-wall biosynthesis (peptidoglycan)
daptomycin (LY146032)	*Streptomyces roseosporus* (B)	2003	antibacterial (Gram-positive)	cell-membrane disruption, aggregation to form holes, membrane depolarization
emodepside[d] (BAY44-4400)	*Mycelia sterilia* (F); semisynthetic from PF1022A	2005	anthelmintic	Slo-1 receptor (K^+ channel)
enduracidin (Enramycin)	*Streptomyces fungicidicus* (B)	1974	antibacterial, food additive	inhibition of MurG (essential for cell-wall biosynthesis in Gram-positive bacteria), inhibition of the transglycosylation step of peptidoglycan biosynthesis
enniatins (fusafungine)	*Fusarium lateritium* (F), *Fusarium scirpi* (F), *Fusarium* sp. (F)	1963	antibacterial (topical), antifungal, anti-inflammatory	ionophore (NH_4^+), membrane depolarization

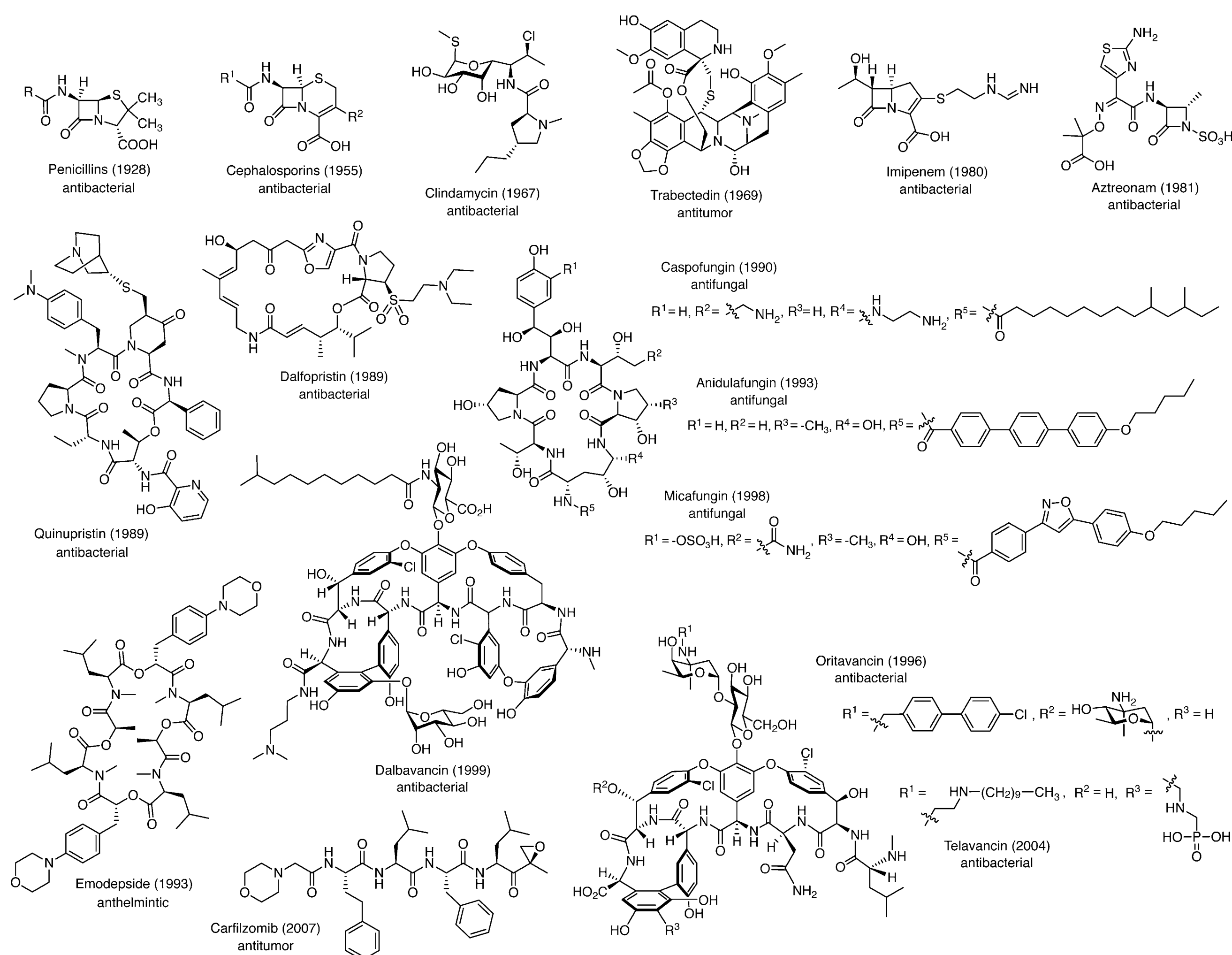
© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 3: (Continued)

Agent	Origin, producing organism(s)[a]	Marketed[b]	Properties and area of application	Mode of action
ergometrine (ergonovine)	*Claviceps purpurea* (F); further derivatives: methylergometrine[d]	1947	obstetrics (therapy as uterus stimulant and vasoconstrictor)	interaction with α-adrenergic, dopaminergic, and serotonin receptors
ergotamine	*Claviceps purpurea* (F)	1921	migraine	vasoconstrictive (5-HT_{1B} receptor, but also dopamine and noradrenaline receptors)
gramicidin A, B, and C	*Bacillus brevis* (B); part of an antibiotic mixture	1952	antibacterial (topical)	ion-channel formation, increasing the permeability of the membrane
gramicidin S	*Bacillus brevis* (B)	1942	antibacterial (topical), antifungal	disruption of the lipid membrane
lincomycin	*Streptomyces lincolnensis* (B) further derivatives: clindamycin[d]	1964 1968	antibacterial (patients allergic to penicillin)	inhibition of the ribosomal protein synthesis (50S-subunit, dissociation of peptidyl-tRNA from the ribosome)
monobactams[e]	*Chromobacterium violaceum* (B); synthetic e.g. aztreonam[c]	1986	antibacterial (Gram-negative)	bacterial cell-wall biosynthesis (peptidoglycan)
oritavancin[d] (LY333328)	*Amycolatopsis orientalis* (B); semisynthetic from vancomycin	2014	antibacterial (Gram-positive; MRSA)[f]	disruption of cell membrane and inhibition of bacterial cell-wall biosynthesis (peptidoglycan), transpeptidation, and transglycosylation
penicillins[d],[e]	*Penicillium* sp. (F) e.g. *Penicillium chrysogenum*	1942	antibacterial	bacterial cell-wall biosynthesis (peptidoglycan)
polymyxin B	*Bacillus polymyxa* (B)	1952	antibacterial (Gram-negative)	binding to lipopolysaccharide (outer membrane), interaction with cytoplasmic membrane
pristinamycin (Ia + IIa)	*Streptomyces pristinaespiralis* (B); quinopristin[d]/dalfopristin[d]: semisynthetic from pristinamycin	1972 1999	antibacterial (Gram-positive), pristinamycin: antibacterial and growth promotor of livestock	ribosomal biosynthesis (50S-subunit, peptidyl transfer, and elongation of protein synthesis)
romidepsin (FR901228)	*Chromobacterium violaceum* (B)	2009	antitumor (cutaneous and other peripheral T-cell lymphomas)	histone deacetylase inhibitor (inducing apoptosis)
teicoplanin	*Actinoplanes teichomyceticus* (B); compound mixture	1988	antibacterial (Gram-positive, MRSA)	membrane anchoring; bacterial cell-wall biosynthesis (peptidoglycan)
telavancin[d]	*Amycolatopsis orientalis* (B), semisynthetic from vancomycin	2009	antibacterial (Gram-positive)	disruption of cell membrane and inhibition of bacterial cell-wall biosynthesis (peptidoglycan)
trabectedin[d] (ET-743)	bacterial symbiont of Ecteinascidia turbinata (sea squirt)	2007	antitumor (antiproliferative, treatment of soft tissue sarcoma)	DNA binder, blocks binding of transcription factors
tyrothricin	*Bacillus brevis* (B), peptide mixture: tyrocidines + gramicidins	1940s	antibacterial (topical; Gram-positive)	disruption of cell membrane
vancomycin	*Amycolatopsis orientalis* (B)	1955	antibacterial (Gram-positive)	bacterial cell-wall biosynthesis (peptidoglycan)
virginiamycin (S1 + M1)	*Streptomyces virginiae* (B)	1959	antibacterial (decontaminant in EtOH production, antibacterial and growth promotor of livestock)	ribosomal biosynthesis (50S-subunit, peptidyl transfer, and elongation of protein synthesis)

[a] Bacterial (B) or fungal (F) producer. [b] Year of approval by regulatory authorities of Europe, US, or Japan. [c] Synthetic drug. [d] Semisynthetic drug. [e] Family of marketed drugs. [f] MRSA = methicillin-resistant *Staphylococus aureus*.

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 5. Marketed semisynthetically modified NRPs and fully synthetic variants and the year they were reported.

Some of the above peptide drugs have been the subject of thorough structure–activity relationship studies and, in consequence, have reached the market as semisynthetic or synthetic compounds in multiple variations (Scheme 5). The most prominent examples are β-lactam antibiotics (penicillins, cephalosporins, penems, and monobactams), of which various derivatives are used as life-saving drugs. On the other hand, for many peptide drug families only a few derivatives (semisynthetic caspofungins) or even only one (cyclosporine A) are/is medically used.

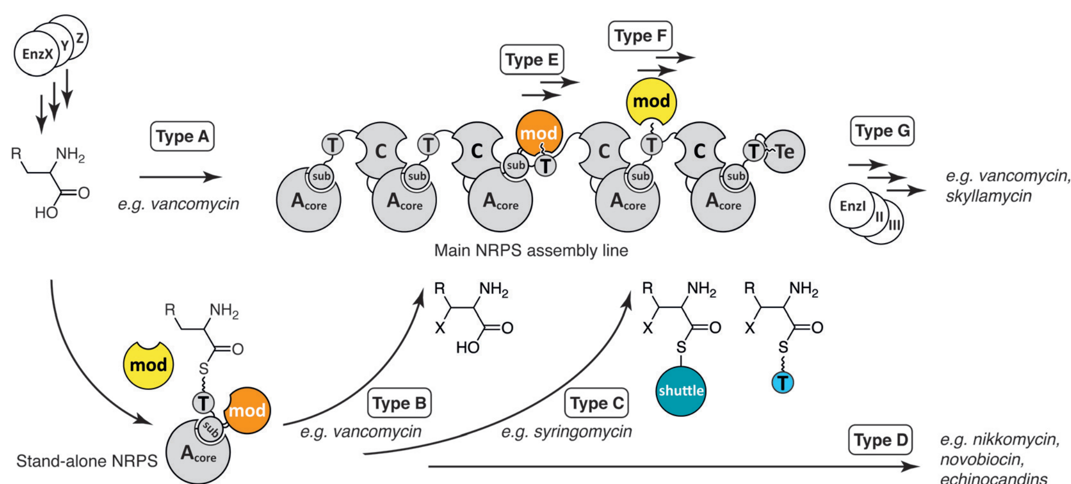
1.8. Microorganisms as Biological Control Agents

In addition to defined drug compositions, as used in chemotherapy, bacterial and fungal strains or preparations thereof find use in agriculture and horticulture.[42] For example, Pseudomonads (G−),[43] Streptomycetes (G+), and Bacilli (G+),[44,45] are used as plant-strengthening or biocontrol agents to enhance growth and crop yields.[46] Likewise, various fungi, that is, *Trichoderma* and *Gliocladium* species as well as *Ampelomyces quisqualis*, have been used for plant protection in forestry and horticulture.[42] As mentioned above, entomopathogenic fungi are used in insect control against various pests, for example, thrips, termites (*Metarhizium anisopliae*), whitefly, and aphids (*Beauveria bassiana*). Lastly, not only because of the ban on the use of antibiotics in animal feed, microorganisms for biocontrol are of great interest in the meat-producing industry. Since the gut microbiome influences states of growth and disease, this can be modulated by probiotic microorganisms.[47,48] As implied by the example of *Bacillus amyloliquefaciens*,[45] the underlying principles contributing to those beneficial effects are, amongst others, the biosynthetic natural product repertoire.

2. Amino Acid Building Blocks—The Basis for Structural Diversity

Amino acids deserve particular attention as they represent the starting material for the biosynthesis of NRPs: the 20 proteinogenic amino acids are complemented by additional building blocks for which nature has developed specific biosynthesis pathways. Hence, NRP biosynthesis generally

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 6. General principles of amino acid supply and modification: Amino acid synthesis by a sequence of cytoplasmic enzymes (type A). Loading and modification by stand-alone NRPSs, followed by hydrolytic release (type B) or transfer onto a shuttle protein or T domain (type C). Substrates may also be fully processed on the stand-alone system and released without further channeling to other NRPSs (type D). NRPS-dependent in-*cis* processing by integral domains (type E) or in-*trans* processing (type F) may be followed by post-NRPS modifications (type G).

occurs in three main phases: 1) building-block assembly, 2) NRPS-mediated peptide assembly followed by 3) post-NRPS modification and decoration.

In general, cytoplasmic enzymes, commonly encoded in the respective NRPS gene clusters, use proteinogenic amino acids or other substrates from primary metabolism to generate nonproteinogenic amino acids (Scheme 6, type A).[49] An alternative NRPS-dependent mechanism is the temporal loading of a proteinogenic amino acid onto a dedicated stand-alone NRPS module (see Section 3.7.4), modification by *trans*-acting enzymes, and subsequent release (type B). Alternatively, the activated building block can be transferred to a shuttle protein or a stand-alone T domain for translocation to the main assembly line (type C). However, in some cases, the stand-alone NRPSs represent the main assembly line and their tethered substrates can be directly liberated as mature products (type D). Another mechanism involves the loading of the amino acid onto a multimodular NRPS assembly line and modification by one or more *cis*-acting NRPS domains (type E). An NRPS may also recruit additional tailoring enzymes that act in *trans* (type F). Finally, post-NRPS modification may occur after release of the peptide from the NRPS (type G). As exemplified by glycopeptide antibiotics, almost all of these mechanisms can be found unified in one biosynthetic pathway.[50]

2.1. Biosynthesis and Structural Features of Amino Acids

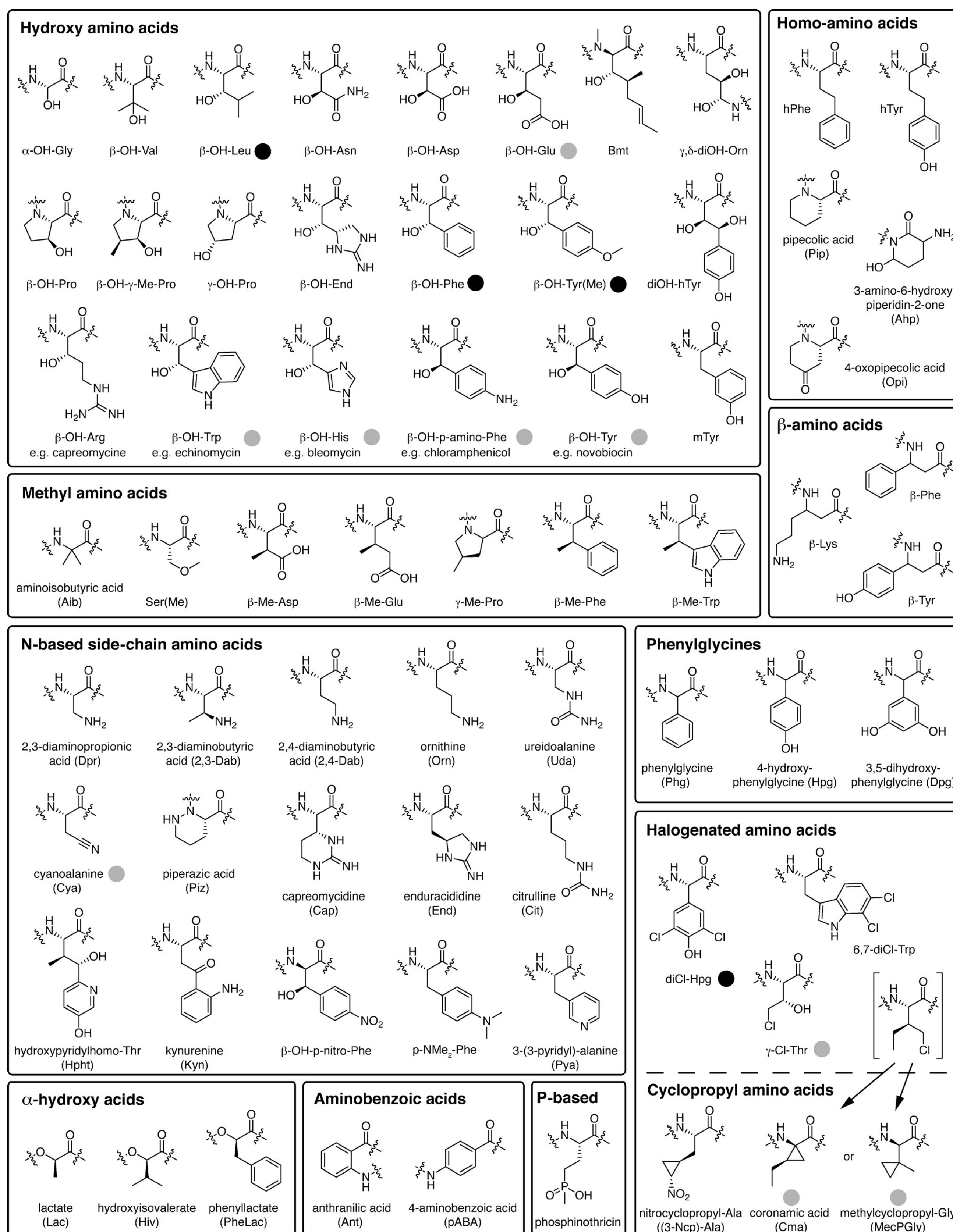
Probably the most common functional groups in nonproteinogenic amino acids (Scheme 7) are hydroxy and methyl groups, which often are localized at the β-positions. Another large family comprises various amino acids with N-based side chains, of which the aliphatic versions can be regarded as structural analogues of Lys and Arg. Further structural classes are β-amino acids, phenylglycines, heterocyclic amino acids, and even aminobenzoic acids (Scheme 7).

Finally, the halogenation of amino acids is a structural modification often required to attain the full bioactivity of the NRP. Subsequent modification (tailoring) of halogenated aliphatic amino acids gives rise to cyclopropyl variants. Other structural modifications, for example, D-amino acids and *N*-methyl amino acids are instead generated on the NRPS (see Section 3.6).

A significant number of amino acids are modified through hydroxylations. Hydroxy groups can serve as handles for further modification, for example, glycosylations, to increase the solubility or binding to the molecular target. Of particular importance are β-hydroxylations (Scheme 7), for example, in type-I glycopeptide antibiotics (β-OH-Tyr),[51,52] ramoplanin (β-OH-Asn), and mannopeptimycin (β-OH-End). Lysobactin (β-OH-Asn/Leu/Phe), skyllamycin ((2*S*,3*S*)-OH-Tyr-(OMe)/(2*S*,3*S*)-OH-Phe/(2*R*,3*S*)-OH-Leu),[53,54] and the fungal echinocandins ((4*R*)-OH-Pro/(3*S*)-OH-(4*S*)-Me-Pro/(3*S*,4*S*)-diOH-hTyr/(4*R*,5*R*)-diOH-Orn/(3*R*)-OH-Gln) are probably among the most hydroxylated peptide antibiotics known.

The most common mechanism for β-hydroxylation is the aforementioned loading of a stand-alone NRPS (Scheme 6, types B and C), which is the platform for a subsequent *in-trans* modification by an oxygenase. In the case of the aromatic amino acids Tyr,[51,52] Trp,[55] and His,[56] these are P450 monooxygenases. In contrast, the hydroxylation of aliphatic side chains, for example, of Glu,[57] is predominantly performed by non-heme Fe^{II}/α-ketoglutarate (KG) dependent oxygenases. Subsequent release of the modified amino acid can occur by a hydrolase, for example, a thioesterase (vancomycin).[51,52] In the biosynthesis of the antibiotic chloramphenicol, which also follows this principle,[58] the substrate *p*-aminophenylalanine (PAPA) is β-hydroxylated by an oxygenase with a dinuclear iron center[59] and the product is reductively cleaved from the NRPS (see Section 3.6.3). Alternatively, nature makes use of a shuttle protein, an aminoacyltransferase, for substrate transfer between a stand-

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 7. Structural classes of nonproteinogenic amino acids that occur in NRPs. Building blocks delivered by stand-alone NRPSs or modified at the main NRPS assembly line are indicated with gray and black circles, respectively.

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

alone NRPS and its partner NRPSs (e.g. syringomycin).[60] The transformation of amino acids into products that do not resemble amino acids (type D) subsequent to β-hydroxylation has been described for novobiocin (assembly of ring A after oxidation of β-OH-Tyr),[61] nikkomycin X (imidazolone base from β-OH-His),[56] and triostin/echinomycin (β-OH-Trp as a precursor of quinoxaline-2-carboxylic acid (QXC)).[55,62] Only in rare cases has an NRPS-independent oxidation (type A) by a non-heme Fe^{II}/α-KG-dependent dioxygenase been observed, for example, for β-OH-Asn from CDA[63,64] or *trans*-3-OH-Pro and the (3*S*)-OH-(4*S*)-Me-Pro from the echinocandin pathway.[65] To date, no clear rules exist to predict the configuration at the hydroxylated carbon atoms. Interestingly, the β-hydroxy amino acid (4*R*)-4-[(*E*)-2-butenyl]-4-methyl-L-threonine (Bmt) that occurs in cyclosporine is not generated by an oxygenase: instead a PKS performs a chain extension followed by a reductive step of the β-keto intermediate.[66]

More recently, evidence arose that some oxygenations occur during peptide assembly on the main NRPS, for example, in skyllamycin biosynthesis (type F). According to this mechanism, the β-hydroxylation of Leu, Phe, and Tyr(OMe) occurs stereospecifically, thus resulting in the same configuration for all β-positions (see Section 3.6.11).[67,68] Finally, post-NRPS hydroxylation (type G) is also a mechanism suggested for some steps of pneumocandin-tailoring to afford (3*S*,4*S*)-diOH-hTyr and (4*R*,5*R*)-diOH-Orn,[55,69,70] as well as for the biosynthesis of aureobasidin (β-OH-Val)[71] and skyllamycin (α-OH-Gly).[53,54] Likewise, 3-amino-6-hydroxy-2-piperidone (Ahp), which occurs in several cyanobacterial cyanopeptolins and in the myxobacterial crocapeptin, is most likely synthesized after NRPS assembly. The oxidation of a proline-containing precursor peptide by a P450 monooxygenase ultimately leads to ring rearrangement under formation of Ahp.[72]

Methyl groups are commonly installed by methyltransferases, with *S*-adenosylmethinonine (SAM) used as a cosubstrate. Methyltransferases, which result in N-, O-, or even C-methylation, contain characteristic specificity-determining signatures. The precursor for β-MeGlu (CDA, daptomycin)[73] is α-KG, and analogous α-ketoacid precursors have been suggested for β-MePhe (hormaomycin;[74] Scheme 8c) and for β-MeTrp (telomycin).[75] Aminoisobutyric acid (Aib) is a characteristic amino acid of the large group of fungal peptaibols. Its biosynthesis is still unknown, although structural homology suggests methylation of Ala by a C-methyltransferase.[76] In contrast, the Glu-analogous β-methyl-Asp (friulimycin) is synthesized by isomerization of Glu mediated by a Glu mutase.[77] Remarkably, the biosynthesis of (2*S*,4*R*)-MePro (e.g. in echinocandin) is not based on a methylation reaction, rather it is based on an α-KG-dependent oxidation of Leu (Leu 5-hydroxylase; Scheme 8d), followed by a cyclization and subsequent reduction of the imine.[69] N-Methylations of the peptide backbone are commonly performed by integral methylation domains on the NRPS (see Section 3.6.2), which also applies to various side-chain O- and N-methylations (e.g. paenilamicin).[29]

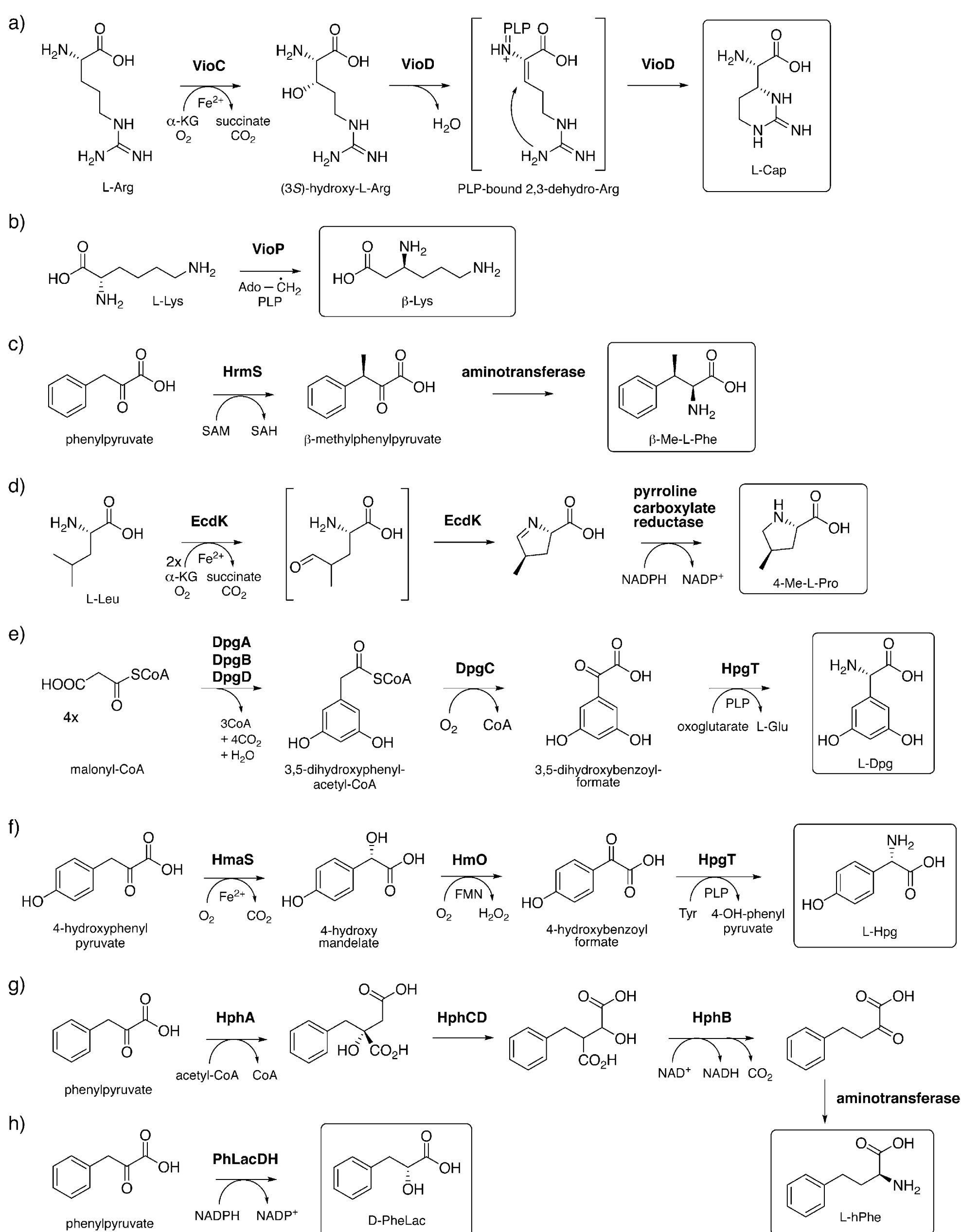
Lys and Arg are the proteinogenic representatives of basic amino acids. Remarkably, various NRPs also contain truncated and cyclic analogues with amino or guanidine functionalities. Diaminopropionic acid (Dap) is synthesized from Ser or Ser(OAc) and Orn (amino donor) by Dpr synthase/Orn cyclodesaminase,[78–80] and is abundant in a variety of aminopolyol peptides, for example, stenothricin, capreomycin, edeine, zwittermycin, and paenilamycin. 2,3-Diaminobutanoic acid (2,3-Dab; friulimicin, pacidamycin) is possibly synthesized by an ATP-grasp ligase from Thr with Asn as the nitrogen source.[81] The related 2,4-diaminobutanoic acid (2,4-Dab) is a constituent of polymyxins (colistin) and originates from Asp-semialdehyde.[82] Ornithine (Orn), which occurs in ramoplanin, originates from Arg (or Glu), and citrulline (Cit) is a constituent of enduracidin. Ureidoalanine (Uda) is a short citrulline analogue of the zwittermycin structure and is likely synthesized by carbamoylation of Dap.[78] The cyclic Arg analogues enduracididine (End) known from mannopeptimycin[83] and capreomycidine (Cap), an amino acid found in tuberactinomycins, for example, viomycin, are structurally unusual. Both amino acids, End[84,85] and Cap (Scheme 8a),[79,86,87] are suggested to originate from Arg. Even cyclic amino acids containing N−N bonds are represented, for example, piperazic acid (Piz),[88] which occurs in kutznerides and himastatin. Although hydroxy-Orn is a biosynthetic intermediate,[89] the mechanism of N−N bond formation has not been ultimately resolved.

The aromatic amino acid kynurenine (Kyn; a component of daptomycin)[90] is a well-known degradation product of Trp and is synthesized by a tryptophan dioxygenase. The amino acid PAPA is a constituent of the pristinamycins and synthesized from chorismate via an aminoprephenate intermediate.[91] The corresponding *p*-nitrophenylalanine is not known and the establishment of the nitro group by an arylamine oxygenase is the final step in the biosynthesis of chloramphenicol.[92] Five genes in the pyridomycin gene cluster have been assigned to the biosynthesis of 3-(3-pyridyl)-alanine (Pya) starting from Asp.[93]

Pipecolic acid (Pip) and its 4-oxo derivative, which occurs in streptogramin, friulimicin, and apicidin, can be considered as the methylene-extended analogue of Pro. Pip is synthesized from Lys by a cyclodeaminase.[94,95] The assembly of pyridylhomothreonine (Pht) in the biosynthesis of nikkomycin is performed via an intermediate picolinic acid, which forms the pyridyl ring.[96] Homophenylalanine (hPhe) and homotyrosine (hTyr) are methylene-extended versions of the corresponding proteinogenic amino acids (Scheme 7). With the exception of the fungal compound echinocandin, which contains hTyr,[97,98] their occurrence has been reported mostly for cyanobacterial peptides, for example, the pahayokolides,[99] cyanopeptolins, and anabaenopeptins.[100] Both hTyr and hPhe are suggested to be assembled from the respective phenylpyruvates and acetyl-CoA (Scheme 8g).[101,102]

Amino acids worth mentioning are the arylglycines, which can be viewed as shortened versions of the proteinogenic aromatic amino acids. The family consists of phenylglycine (Phg) found in streptogramins as well as 4-hydroxyphenylglycine (Hpg)[103,104] and 3,5-dihydroxyphenylglycine (Dpg), which are both constituents of a large number of NRPs, for example, the glycopeptide antibiotics.[105] Whereas Phg and

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 8. Biosynthesis pathways of selected nonproteinogenic amino acids used in NRPS assembly.

Hpg[103] are synthesized through the shikimate pathway, Dpg is assembled by a chalcone synthase from malonyl-CoA (Scheme 8e,f).[49]

Nature also uses β-amino acids such as aliphatic β-Lys (a building block of viomycin and streptothricin), which is synthesized by a radical SAM enzyme (lysine 2,3-aminomutase) that shifts the α-amino group into the β-position (Scheme 8b). In contrast, aromatic 2,3-mutases contain 4-methylideneimidazole-5-one (MIO)[106] as a cofactor and generate β-Phe and β-Tyr (which occur in andrimid and chondramide).[107] Hence, peptide linkages through the β- or γ-positions of such amino acids are particularly abundant in some cyanobacterial NRPs and constitute alternatives for chain elongation during NRPS assembly.

Fungal cyclodepsipeptides of the enniatin-type contain α-hydroxy acids, which originate from the reduction of α-ketoacids (primary metabolism) by the corresponding keto-reductases (Scheme 8h). Further unusual building blocks are the aminobenzoic acids: anthranilic acid (Ant) and various *p*-aminobenzoic acids (pABA) are constituents of sibiromycin and albicidin, respectively.[27] Ant and pABA are both products of the shikimate pathway.

The halogenation of amino acids is sometimes wrongly described simply as decoration of NRPs. This type of modification has a rather significant influence on the bioactivity of NRPs, or in some cases generates an intermediate for further processing. The halogenation of aromatic side chains occurs in a myriad of NRPs. These are introduced by $FADH_2$-

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

dependent halogenases, and an NRPS-dependence of the process has been suggested or proven in many cases, for example, glycopeptide antibiotics of the vancomycin type (3-Cl-β-OH-Tyr)[51,108,109] or kutzneride (6,7-diCl-Trp).[110] In contrast, halogenations of aliphatic and, thus, comparatively non-activated residues are performed by non-heme-FeII α-KG-dependent halogenases. Mechanistically the reactions occur on stand-alone NRPSs (types B and C). In the syringomycin pathway, Thr is thus loaded onto the NRPS SyrB1 and halogenated by SyrB2 in the presence of α-KG, O_2, and Cl^-.[111] The product is then transferred to a shuttle protein for further processing on the main NRPS assembly line.[60] Remarkably, the product of the halogenation of *allo*-Ile has been reported to be a precursor to the cyclopropyl amino acid coronamic acid (Cma).[112] Further cyclopropyl motifs occur in kutzneride and hormaomycin, but for the latter a different assembly has been suggested.[74] Among all the halogenated NRPs, chlorination is clearly dominating, whereas bromination can be achieved in some cases by providing bromine in the growth media.[113] Although a fluorinase has been described to synthesize 4-fluorothreonine from Thr and fluoroacetaldehyde,[114] no naturally occurring NRPs have been reported so far.

Only recently was the biosynthetic origin of the Ile isomer *allo*-Ile elucidated, which implicated a two-step isomerization process from Ile.[115] The Tyr isomer 3-OH-Tyr (pacidamycin, sanglifehrin) is likely synthesized by a phenylalanine hydroxylase.[116,117] Sulfur-containing amino acids, with the exception of Cys and Met, are very rare among NRPS products. The same is true for phosphorus, with the exception of the bialaphos family, with phosphinothricin (glufosinate) as the most prominent representative.[118] There are several structurally complex amino acids such as ADDA[119] in microcystins and aziridino[1,2*a*]pyrrolidinyl amino acids in azinomycin[120,121] that await further biosynthetic investigations.

2.2. N-Terminal Modifications of NRPs

Many NRPs carry N-terminal modifications, foremost acylations. These modifications, which can be considered as some sort of "end cap", have various functions, for example, to protect the N terminus from degradation, to modulate polarity, or to confer specific properties such as membrane insertion.

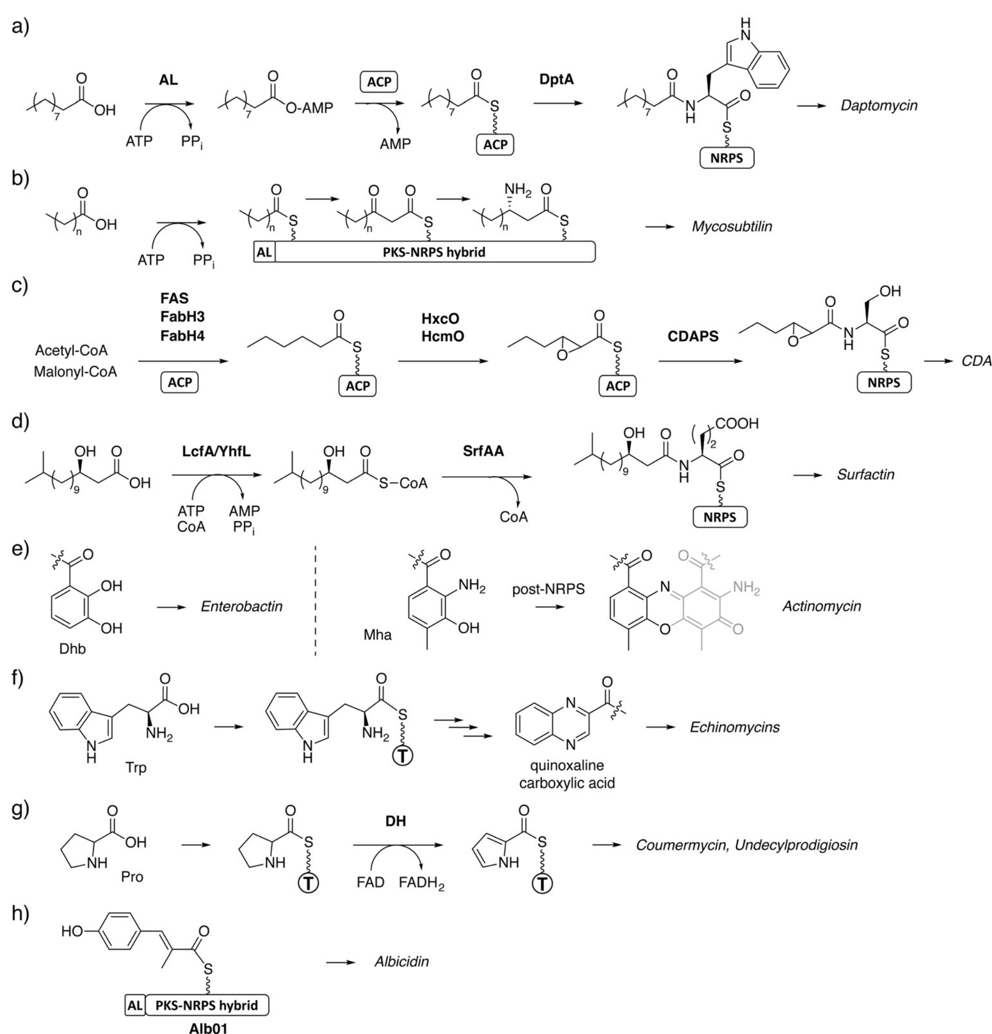
The lipocyclopeptides are a diverse group with a vast number of representatives. The N-terminal acylation is commonly achieved by coupling of an activated linear or branched fatty acid, that is, bound to an acyl carrier protein (ACP) or CoA, to the starter amino acid of an NRPS assembly line. The nature of the fatty acids ranges from short- (epoxomicin), medium- (CDA, eponemycin), to long-chained, and may contain various degrees of unsaturation (echinocandin) and branching (daptomycin). Formylations are installed by discrete NRPS domains (see Section 3.6.4). In cases where the origin is not attributed to primary metabolism, the gene cluster may contain cognate fatty acid biosynthesis (*fab*) genes. Additional functional groups such as -OH or $-NH_2$ in the β-position of the acyl residue allow for macrocyclizations (surfactin, fengycin, mycosubtilin, and bacillomycin). Different pathways have been described for the attachment of an acyl residue to the N-terminal NRPS-bound amino acid (Scheme 9). In the biosynthesis of compounds of the A21978C family and of daptomycin, an individual fatty acid acyl ligase (AL) and an ACP perform the activation (acyl adenylate) and transfer without employing CoA-thioester intermediates. This has also been suggested for echinocandin, where linoleyl-adenosine monophosphate (linoleyl-AMP) is transferred.[101] Likewise, mycosubtilin biosynthesis employs an AL and ACP, but as an integral part of the NRPS assembly line.[122] In contrast, a 3-OH fatty acyl-CoA thioester is taken from the primary metabolism and as such is directly coupled to the N-terminal amino acid on the surfactin NRPS.[123] The CDA gene cluster encodes fatty acid biosynthesis genes that have been suggested to use acetyl-CoA and malonyl-CoA to synthesize the hexanoyl-ACP precursor, which, upon construction of the epoxide, is transferred to the NRPS.[124]

Aromatic acyl residues are also found in a variety of NRPs (Scheme 9). Important modifications comprise those with 2,3-dihydroxybenzoic acid (Dhb), derived from the chorismate pathway,[125] and constituent of catecholate-type siderophores, or 4-methyl-3-hydroxyanthranilic acid (4-Mha), derived from Trp and a precursor of the DNA-intercalating phenoxazine moiety of actinomycin.[126] Dhb and 4-Mha are activated by dedicated NRPS domains (see Section 3.1). Some acyl residues are synthesized from amino acid precursors, for example, quinoxaline-2-carboxylic acid (QXC),[55,62] which originates from Trp and is a constituent of the chromodepsipeptides triostin and echinomycin. Other examples are the conversions of Pro into pyrrolcarboxylic acid (a building block of pyoluteorin and coumermycin)[127] and of Lys to 3-hydroxypicolinic acid (a building block of virginiamycin).[128] Other N-terminal acylations may also be integral parts of PKS-NRPS hybrids, as suggested for the cinnamoylation of albicidin.[27]

3. Architecture and Mechanisms of NRPSs

Nature exploits a modular concept for the synthesis of NRPs, in which each module of an NRPS assembly line performs the activation and coupling of a single amino acid to a growing peptide chain. According to this principle, which is also known under the name collinearity rule, the biosynthesis of a heptapeptide requires seven such modules (see Section 3.7). The modules themselves comprise distinct protein domains that harbor the catalytic centers required for peptide synthesis, that is, 1) the adenylation (A) domain for selection, activation, and loading of the amino acid onto 2) the thiolation (T) domain, also referred to as peptidyl carrier protein (PCP) domain, which bears a 4′-phosphopantetheine (Ppant) prosthetic group in its holoform (Figure 2). The tethered amino acid is then shuttled to 3) the condensation (C) domain, where coupling to the upstream nascent peptide chain is established. Whilst attached to the holo-T domain, building blocks can be shuttled to optional protein domains, either incorporated in the respective module, for example,

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 9. Examples of N-terminal acyl modifications of NRPs: a) ACP-mediated and b) AMP-mediated transfer to NRPSs; c) and d) CoA-mediated loading onto the NRPSs; e) dihydroxybenzoic acid (Dhb) and 4-methyl-3-hydroxyanthranilic acid (4-Mha) are directly recognized as substrates by adenylation domains of NRPSs; f) and g) amino acid tailoring by stand-alone NRPSs; h) PKS-mediated attachment of N-terminal cinnamate.

epimerization (E), formylation (F), methylation (M), heterocyclization (Cy), reduction (R), and oxidation (Ox) domains, or to *trans*-acting tailoring enzymes that install additional modifications. 4) Finally, a thioesterase (Te) domain disconnects the mature oligopeptide from the NRPS machinery and often mediates macrocyclization during this release step.

Recent structural biology approaches on excised domains and intact modules have provided mechanistic details on the catalytic cycle of NRPSs. The first structure of an intact NRPS was the terminal module SrfA-C of the surfactin synthetase.[129] With its C-A-T-Te topology, it revealed for the first time the general domain architecture of an NRPS module. Concurrently, this structure illustrated that a flexible 18 Å Ppant arm alone cannot transit the substrate to all the reaction centers within the module (a trajectory of more than 100 Å), and that long-range domain movements are indispensable for a full catalytic cycle. Structural information on individual domains helped to develop mechanistic inhibitors which tether interacting domains and thereby freeze the inherent dynamics that hinder protein crystallization.[130–132] As outlined in the following sections, recent X-ray structures of entire modules trapped in catalytically relevant states gave atomistic insights into how substrate translocation is coupled to these concerted domain–domain rearrangements and interactions (Figure 3).[131,133]

3.1. Adenylation Domains and MbtH-Like Proteins

The initial step of NRP synthesis is the selection and activation of amino acid substrates, first as mixed anhydride derivatives, namely aminoacyl-AMP, and subsequently as aminoacyl-thioesters covalently attached to the NRPS. These functions are fulfilled by the A domain (ca. 60 kDa), which belongs to the ANL (**A**cyl-CoA synthetases, **N**RPS adenylation domains, and **L**uciferase enzymes) superfamily of adenylating enzymes.[130] All members of this superfamily catalyze an initial adenylation of carboxylate substrates using Mg·ATP. As the aminoacyl-AMP is prone to nonproductive hydrolysis, the A domain protects the high-energy intermediate from bulk water to subsequently catalyze its loading onto the Ppant arm of the holo-T domain in a second reaction step.

A domains are organized in two subdomains: the approximately 50 kDa N-terminal core domain (A_{core}) and the approximately 10 kDa C-terminal subdomain (A_{sub}), which are flexibly linked by a hinge region of about five residues. Consensus motifs of A domains (A1–A10) have been rationalized by several X-ray structures and play structural as well as functional roles.[134,135] The specific binding of an amino acid and Mg·ATP occurs within the A_{core} domain close to the A_{core}–A_{sub} interface. The positioning of an α-amino acid is assured by a highly conserved Asp (A4 motif in A_{core}) and Lys residue (A10 motif in A_{sub}), which stabilize the amino and carboxylate moieties, respectively (Figure 4). Whereas the Lys residue is essential for the adenylation reaction and thus strictly conserved,[130,134,136] the Asp residue is subject to customization and can be found replaced or repositioned within the binding pocket for optimal interaction with

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

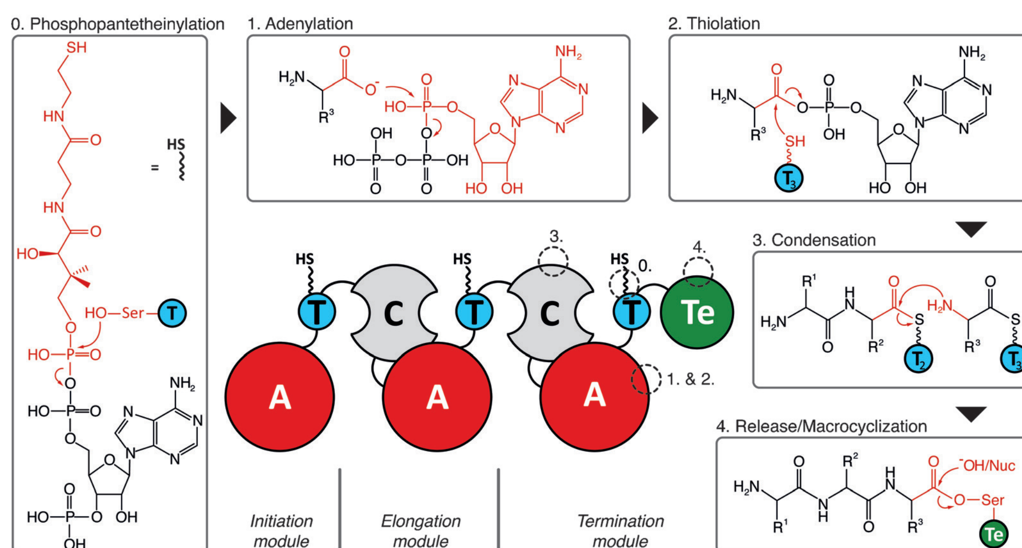


Figure 2. Domain arrangement of bacterial NRPSs and their catalyzed reactions. 0) NRPS priming: PPTase-mediated installment of Ppant at a conserved serine of the apo-T domain. 1) Selection and adenylation of the amino acid by the A domain generates a high-energy aminoacyl-AMP species and PP_i. 2) Subsequent thiolation of the activated amino acid and release of AMP yields an aminoacyl thioester attached to the Ppant of the holo-T domain. This step is catalyzed by the A domain using the same catalytic pocket as in the adenylation partial reaction. 3) Formation of a peptide bond by the C domain couples the activated amino acid to the amino acid or nascent peptide which is attached to the upstream module. 4) Release of the oligopeptide is achieved by formation of an intermediate ester bond between the C terminus of the peptide and a conserved serine of the Te domain. Hydrolysis or intramolecular attack of a nucleophilic moiety yields a linear or macrocyclic product, respectively. The product of each reaction is implicated in red. Nuc = nucleophile.

substrates other than α-amino acids, such as β-amino acids, α-hydroxy acids, α-keto acids, or aminobenzoic acids (Figure 4b).

The conserved Asp residue is located at the entrance of the substrate binding pocket, which is further decorated with various residues to optimally accommodate the side chain of the cognate substrate. Up to eight residues have been identified to be involved in side-chain recognition, which has led to the establishment of a specificity-conferring code for A domains,[134,136,137] also known as the nonribosomal code (Figure 4a,b). Bioinformatic algorithms have since then been developed for the prediction of potential substrates and, thus, product structures of NRPS in genome mining approaches.[21–23,138] The quality of the predictions steadily improves as more structural and biochemical data on NRPS A domains become available. The specificity-conferring code was initially deduced from the cocrystal structure of the A domain of GrsA (gramicidin S synthetase 1) in complex with Mg·AMP and its rather large substrate Phe.[134] In contrast, smaller side chains do not penetrate the binding pocket to such an extent, thereby reducing the effective substrate interaction site. For example, only five residues located at the pocket's entrance play a role in the substrate recognition of Gly,[131] which explains the higher evolutionary variability of the other non-interacting residues (Figure 4f).

A frequent observation in studies of the NRPS A domain is relaxed substrate specificity, which appears to be a strategy used by NRP-producing organisms to increase natural product diversity with a single NRPS.[142] This promiscuity arises mainly from the degenerate nonribosomal code itself, since appropriate combinations of residues in the substrate binding pocket allow for a certain plasticity towards chemically similar (Arg/Lys)[143] or even chemically distinct substrates (Arg/Tyr).[139] Beyond that, recent investigations have indicated that the C domain may directly affect the substrate-specificity profile of its neighboring A domain, presumably by (de-)stabilizing specific conformational states in the NRPS catalytic cycle.[142] This observation underlines the potential risk of perturbing the C-A as well as C-T domain–domain interfaces by using excised A-T protein constructs in adenylation assays.

Structural analysis of various adenylating enzymes as well as intact NRPS modules has established the concept of domain alternation[130–132]—a strategy by which the A domain reorganizes the A_{core}–A_{sub} interface and thereby links domain–domain reorientation to a catalytic switch. After the adenylation reaction, pyrophosphate (PP_i) is released and the well-conserved A8 hinge motif allows a rigid-body torsion of the A_{sub} domain of approximately 140° with respect to A_{core}. The new orientation of the A_{sub} domain not only switches off the adenylation mode (the conserved Lys residue in motif A10 is oriented away from the substrate), but also concerts the relocalization of the holo-T domain such that its Ppant arm is able to approach the aminoacyl-AMP intermediate in the A_{core} binding pocket for the thiolation half-reaction (Figure 3). Likewise, a 180° rotation and 21 Å translocation of the A_{sub} domain synchronized with a 75° rotation and 61 Å translocation of the aminoacylated holo-T domain allows the substrate to traverse the 50 Å distance between the catalytic centers of the A domain and formylation (F) domain in the initiation module of linear gramicidin synthetase LgrA (Figure 3).[133] Hence, the A_{sub} domain functions as a flexible hinge whose rotation entails a pull-and-push motion of the adjacent T domain towards and away from the A domain.

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

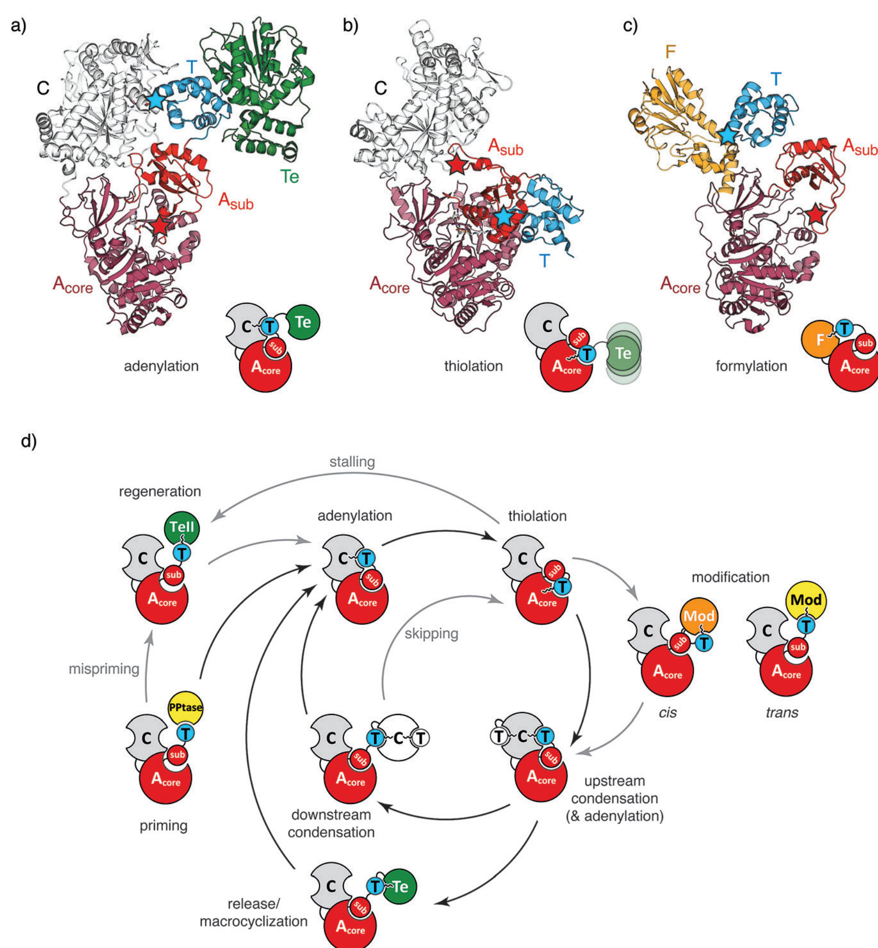


Figure 3. Domain–domain arrangements and catalytic cycle of NRPS modules. a) Adenylation state of the termination module holo-AB3403:[131] Asterisks denote the locations of the Ppant attachment site (blue) and the invariable Lys residue of the A_{sub} domain (red), which is essential for adenylation activity. Notably, the T domain is correctly positioned for simultaneous upstream condensation. b) Thiolation state of the termination module holo-EntF:[131] A rigid-body rotation of the A_{sub} domain triggers a large translocation of the holo-T domain towards the A_{core} domain, thus enabling the thiolation half-reaction. No electron density has been observed for the Te domain, thus indicating its conformational heterogeneity. c) Formylation state of the initiation module of linear gramicidin synthetase:[133] Both, the A_{sub} and T domains adopt an extended arrangement to bridge the distant catalytic centers for adenylation/thiolation (A_{core}) and N-formylation (F domain). Schematic representations of NRPS modules are shown below. The color coding used for the domains is used throughout this Review. Based on these observations and previous knowledge, a model of the catalytic cycle of an NRPS module is illustrated in (d). Priming of the T domain initiates the catalytic cycle. Type II Te domains free the Ppant from the blocking in the case of mispriming with short-chain acyl-CoAs or stalling because of incorrect substrate activation. After initial adenylation, rotation of the A_{sub} domain relocates the holo-T domain and enables penetration of the Ppant into the A_{core} substrate pocket for thiolation. Once the amino acyl-thioester intermediate has formed, it may be subject to modification in *cis* (indicated in orange) or in *trans* (indicated in yellow). The substrate-loaded T domain then migrates to the acceptor site of the C domain, while the A_{sub} domain re-adopts its adenylation state. Hence, adenylation and upstream condensation states formally appear to be equivalent. For downstream condensation, the peptidyl-loaded T domain reorients towards the donor site of the downstream C domain. It is not clear whether A_{sub} retains its conformation, similar to that in the upstream condensation reaction. In principal, the A_{sub} domain could still adopt the adenylation-competent state and, thus, the catalytic cycle could directly proceed with thiolation (skipping). Instead of downstream condensation, the peptidyl-T domain of a termination module translocates towards the Te domain to release/macrocyclize the mature product.

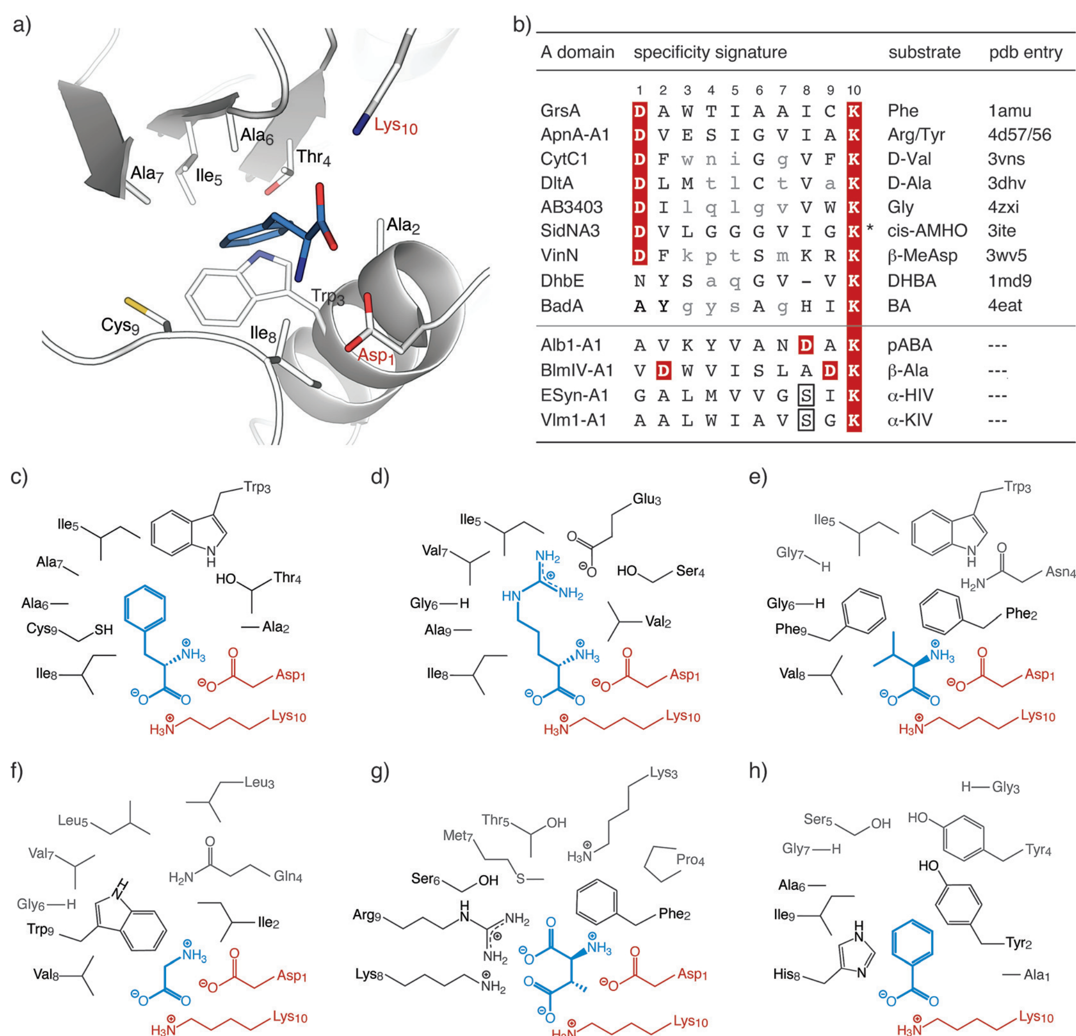
Intriguingly, mutagenesis of the conserved A8 hinge residue (most commonly Asp) with Pro virtually impedes thiolation, but still allows for adenylation activity,[144] thus highlighting the importance of domain alternation for the catalytic cycle.

Sequence analysis identified an LPxP motif downstream of the A10 motif that interacts with the A_{sub} domain and appears to rigidify the A10 loop harboring the catalytically essential Lys residue.[145] A_{sub} domain alternation does not impair this ternary contact, as the binding mode is observed in the crystal structures of both the adenylation and thiolation states.[131,132] Disruption of these hydrophobic interactions between the LPxP motif and the A_{sub} domain, however, has been shown to impede productive adenylation by the enterobactin NRPS EntF.[145] The LPxP motif should, thus, be viewed as an additional core motif A11 of the A domain with structural importance for core motif A10 and the affiliated T domain.

Interestingly, a loop region between core motifs A8 and A9 of the A_{sub} domain serves as an evolutionary insertion point of additional domains that incorporate various modifications in situ, for example, methylation, oxidation, or dehydration of thiolated building blocks (see Section 3.6).[146] None of these domains have been structurally characterized so far, but based on the current knowledge of domain arrangements[129,131,133] the approximate positions of these extra domains can be postulated to be lateral to the T domain in both the adenylation and thiolation states. As domain alternation of A_{sub} appears to be central to the NRPS catalytic cycle, one may anticipate the anchored modification domain to accompany A_{sub} during its 140° rotation. Given the fact that modification domains are much larger than A_{sub} (30–45 kDa versus 10 kDa), it is intriguing that nature exploits this dynamic hinge to incorporate an additional tailoring domain, again illustrating that A_{sub} is a centerpiece of NRPS machines.

Various A domains of bacterial origin have been reported to depend on so-called MbtH-like proteins (MLPs).[19] The naming stems from a small protein encoded in the NRPS gene cluster responsible for the production of the siderophore mycobactin, which confers virulence to the tuberculosis-

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



A domain	1	2	3	4	5	6	7	8	9	10	substrate	pdb entry
GrsA	D	A	W	T	I	A	A	I	C	K	Phe	1amu
ApnA-A1	D	V	E	S	I	G	V	I	A	K	Arg/Tyr	4d57/56
CytC1	D	F	w	n	i	G	g	V	F	K	D-Val	3vns
DltA	D	L	M	t	l	C	t	V	a	K	D-Ala	3dhv
AB3403	D	I	l	q	l	g	v	V	W	K	Gly	4zxi
SidNA3	D	V	L	G	G	G	V	I	G	K *	cis-AMHO	3ite
VinN	D	F	k	p	t	S	m	K	R	K	β-MeAsp	3wv5
DhbE	N	Y	S	a	q	G	V	-	V	K	DHBA	1md9
BadA	A	Y	g	y	s	A	g	H	I	K	BA	4eat
Alb1-A1	A	V	K	Y	V	A	N	D	A	K	pABA	---
BlmIV-A1	V	D	W	V	I	S	L	A	D	K	β-Ala	---
ESyn-A1	G	A	L	M	V	V	G	S	I	K	α-HIV	---
Vlm1-A1	A	A	L	W	I	A	V	S	G	K	α-KIV	---

Figure 4. Substrate specificity of A domains. a) Close-up view of the binding pocket of the Phe-activating A domain of GrsA.[134] Ten residues participate in Phe binding (blue) and give rise to the nonribosomal code (summarized in (b)).[136,137] Signatures for various substrates are listed, for example, α-amino acids with small or bulky side chains, D-amino acids, β-amino acids, benzoic acid derivatives, as well as α-hydroxy/-keto acids. Conserved Asp and Lys residues are highlighted in red. The conserved Ser (or in some cases Thr) residues could potentially replace the Asp residue to stabilize the α-oxygen atom of α-hydroxy and α-keto acids (black box). Only residues shown in capital letters are involved in substrate binding. The asterisk denotes the much larger binding pocket of the fungal A domain SidNA3.[135] c)–h) Schematic representation of the A domain binding pockets which have been experimentally determined.[131,134,139–141] Substrates are shown in blue, Asp1 and Lys10 are highlighted in red, and residues without contact to the substrate are in gray.

causing pathogen *Mycobacterium tuberculosis*.[147] The occurrence of *mbtH*-like genes in NRPS gene clusters hints at MLP-dependent A domains in the respective biosynthetic machinery. Genetic knock-out experiments have indicated that MLPs are often required for the efficient production of NRPs.[148,149] Strikingly, MLP paralogues from other NRPS gene clusters within the genome (up to seven in actinomycetes) can partially complement each other such that NRP production is not compromised.[148] This complementation has even been observed in vitro for MLPs of different species.[150,151]

In a remarkable biochemical study, Boll et al. investigated the effects of MLPs on various tyrosine-activating A domains from the biosynthesis pathways of novobiocin (NovH), clorobiocin (CloH), simocyclinone (SimH), and vancomycin (Pcza361.18).[150] They found that tyrosine adenylation activities of CloH, SimH, and Pcza361.18 in vitro were strongly dependent on the presence of the MLPs CloY, SimY, and Orf1van, respectively. In contrast, NovH from the novobiocin gene cluster, which is devoid of an *mbtH*-like gene, exhibited moderate tyrosine adenylation activity in the absence of MLPs. Surprisingly, NovH was markedly stimulated by the presence of the noncognate MLP CloY by enhancing the catalytic turnover. Moreover, the MLPs used in this study were interchangeable, with no preference towards their cognate interaction partner. An equimolar stoichiometry has been observed for the CloY/CloH and SimY/SimH complexes. Pivotal for this functional study of MLP-A domain–domain interactions was the use of the *E. coli* strain Δ*ybdZ*, in which the MLP YbdZ from the enterobactin gene

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

cluster was inactivated. YbdZ has been shown to co-purify with recombinant A domains and to stimulate their adenylation activity, albeit at low efficiency. The deletion of the *ybdZ* gene caused a significant drop in protein yields, which is a consistent observation that MLPs enhance the soluble expression of their binding partner.

Likewise, the A domains CmnN and VioN of the biosynthesis pathways of the antituberculosis antibiotics capreomycin and viomycin were only active when complexed in an equimolar ratio with their cognate MLPs CmnO and VioO, respectively.[152] When coexpressed in *E. coli*, CmnN and CmnO were co-purified as a functional complex.

Structural data have been obtained on stand-alone MLPs (ca. 8 kDa) by both X-ray crystallography[153] and NMR spectroscopy.[154] MLPs share a conserved fold comprising an α-helix which packs onto a three-stranded antiparallel β-sheet. From multiple sequence alignments of various MLPs, a signature sequence of 15 invariant amino acids, among them three Trp residues, has been deduced.[19] Most of these invariant and mostly hydrophobic residues are clustered on one face of the MLP. The structure of SlgN1-ΔA_{sub}, a naturally occurring MLP-A didomain fusion protein involved in the biosynthesis of the antibiotic streptolydigin, has provided first insights into the binding mode.[155] In this structure, the MLP domain uses its hydrophobic patch to pack against a small α-β motif, which is flanked by loop regions and the core motifs A6 and A7. The three-residue β-strand of this small α-β docking motif terminates with an Ala residue (A433), the side chain of which inserts into a hydrophobic cleft of the MLP domain formed by conserved residues S23, W25, P32, and W35. As expected, the mutations S23Y (MLP) as well as A433E (A_{core}) disabled the MLP-A_{core} interaction and significantly reduced the adenylating activity of SlgN1-ΔA_{sub}.[155] A possible impact of MLP binding on the conformational and, thus, catalytic state of A_{sub} could not be deduced, since the protein construct used in this study was devoid of the flexible A_{sub} domain for crystallization purposes.[155] Very recently, the complex structures of the MLPs YbdZ (*E. coli*) and PA2412 (*P. aeruginosa*) bound to intact EntF (C-A-T-Te topology) in the thiolation state have been reported.[156] Both MLPs occupy a locus of the EntF-A domain juxtapositioned to the C-A_{core} interface and equivalent to that observed for the fusion protein SlgN1-ΔA_{sub}. In these crystal structures, MLP binding does not alter the conformation of the EntF-A domain, thus obscuring the underlying mechanism of MLPs. This may be partially ascribed to the fact that MLP binding merely enhances the adenylation activity of EntF, and that its A domain is not strictly MLP-dependent.[156] MLPs have been postulated to simply stabilize their binding partner (chaperone function), but there is growing evidence that they influence its catalytic properties (allosteric regulation).[155,156] It should be mentioned that the α-β docking motif of the A_{core} domain bears conserved residues that extend away from the MLP towards the Mg·ATP binding site and that are in proximity to a highly conserved loop region that resembles the P loop of many ATPases and GTPases (part of core motif A3, 190-TSGTTGNPGK-199 in GrsA).[134] This loop is dynamic and presumed to interact with the leaving PP_i group. The X-ray structures of the thiolation states of LgrA[133] and EntF[131,156] illustrate that the P loop is sandwiched between helix α1 of A_{sub} and the α-β docking motif of the A_{core} domain. The proximal binding of the MLPs possibly hints at an allosteric regulation of domain alternation, that is, PP_i release, initiation of the thiolation half-reaction, and thus catalytic turnover. It should be emphasized that so far only adenylation activity assays have been performed in the presence of MLPs, thus leaving the second half-reaction unexplored in terms of MLP-mediated catalytic turnover.

Multiple sequence alignments suggest that MLP-dependent A domains possess rather hydrophobic α-β docking motifs, and the corresponding residues are replaced by more hydrophilic or bulkier side chains in MLP-independent A domains.[155,156] With more biochemical and structural data becoming available, predictions of MLP-(in)dependence will become more reliable. Importantly, MLPs do not appear in fungal NRPS gene clusters; they are exclusive to bacterial systems.[19] SidN-A3 is involved in the biosynthesis of a hydroxamate siderophore and is the only fungal A domain whose structure has been elucidated to date.[135] Besides the fact that the α-β docking motif deploys rather polar/charged residues towards the protein exterior (similar to bacterial MLP-independent A domains), the common α-helix is absent. The described arrangement would thus interfere with, if not impede, MLP binding. The question is then: How do fungal A domains achieve a more efficient catalytic turnover if they are strictly independent of MLPs? The answer to this question not only concerns enzymatic properties of bacterial versus fungal A domains, but it may directly relate to the mechanism of MLPs in general. So far there are no reports that fungal A domains have been tested in vitro with respect to their MLP susceptibility.

Interestingly, nature has engineered several MLP-fusion proteins of the type MLP-A didomain (SlgN1), MLP-A-T tridomain (NikP1 of nikkomycin), as well as a unique MLP-P450 oxygenase didomain (LtxB of lyngbyatoxin).[19] The last example is striking as it implies another possible function for MLPs, that is, recruitment of tailoring enzymes for substrate modification on the NRPS assembly line. To date, the mechanism of MLPs and their role in NRP synthesis in general still remain vague—their importance in future biotechnological approaches, however, is already evident.

3.2. T Domains and PPTases

Conformational flexibility is an inherent trait and an essential requirement for the communication and choreography of NRPS domains. Paradigms for this functional flexibility are the A_{sub} domain and the T domain—in other words, the NRPS control center orchestrating the substrate shuttle system. The holo-T domain with its Ppant extension (ca. 18 Å) can be viewed as the flexible robot arm of the NRPS assembly line that covalently sequesters and transfers the amino acyl-/peptidyl-thioester intermediates to all the catalytic centers that are required for modification, condensation, or liberation.[157]

The T domain (ca. 10 kDa) adopts a four-helix bundle,[158,159] with the N terminus of the second helix α2

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

harboring the highly conserved serine residue (GxxS core motif) that becomes posttranslationally modified by a phosphopantetheine transferase (PPTase). As the attachment of Ppant is a functional prerequisite, PPTases play an essential role in NRP synthesis.[160,161] Studies on the PPTase Sfp from *Bacillus* have shown that it binds CoA in a bent conformation with the terminal thiol oriented towards the protein exterior.[162] The resulting tolerance towards terminal modifications has been extensively used to load recombinant apo-T domains with various aminoacyl-CoA analogues and thus to bypass the A domain in in vitro experiments.[133,163–165] On the other hand, the low selectivity of PPTases coupled with the high abundance of short-chain acyl-CoA species in the cell, in particular acetyl-CoA,[166] may cause frequent mispriming of T domains and thereby stall the NRPS assembly line (see below).

The structures of intact NRPS modules display the Ppant arm in extended conformations pointing away from its attachment site to pervade the Ppant binding tunnels of the A and C domains.[131,133] NMR experiments using the excised holo-$T7_{tei}$ domain from the teicoplanin-producing NRPS loaded with a paramagnetic spin label indicated that there is no distinct interaction between the Ppant arm and the T domain in solution.[167] However, it could well be that the spin label itself prohibited the interaction. Indeed, recent NMR studies on substrate-loaded holo-T domains suggest a transient resting position for cognate substrates. When equipped with its pyrrole-*N*-CoA cofactor, the T domain PltL from pyoluteorin biosynthesis revealed that the Ppant arm is folded back and that the pyrrole moiety is accommodated in a hydrophobic cleft between helices α2 and α3.[168] A similar resting state has been demonstrated for the salicyl-*S*-CoA-loaded state of an aryl carrier protein domain from yersiniabactin synthetase.[169] As a consequence, the protein surface near this cleft changes markedly upon substrate sequestration and could, thereby, modulate binding affinities to partner domains, as these are mediated by the α2/α3 interface (see below). It should be emphasized that the interactions between the substrate and the T domain have been described as transient in nature, thus suggesting that conformational equilibria determine the catalytic route.[169] Initial NMR studies of T domains have proposed the absence of a distinct substrate-binding pocket and ruled out any substrate specificity of the T domain.[158] However, given the current knowledge, T domains may in fact exhibit at least some degree of substrate selectivity. Whilst the transient stabilization of acyl intermediates may protect the substrate from potential side reactions in the cell, a dynamic flap motion of the amino acyl-Ppant arm would still allow for its sufficient exposure to catalytic partner domains. Clearly, this fine-tuned equilibrium can be considered very sensitive and needs to be preserved when designing NRPS hybrids—for example, immoderate affinity between the substrate and T domain may otherwise interrupt a productive interaction with the catalytic centers and stall the NRPS machinery. Given the size of the substrate-binding cleft, it can be assumed that peptidyl intermediates may not be subject to such sequestration events, but rather promptly transmitted to the donor sites of C domains for efficient NRP elongation.

Recent structural studies have shown that the holo-T domain virtually retains the conformation of its apo-state,[131,133,153,167–169] which was originally designated as the A/H state.[159] Given the location of the prosthetic group at the beginning of helix α2, it is not surprising that this helix and the preceding loop as well as helix α3 form the major interface between holo-T and its various binding partners, including A,[131,133,153] C,[129,131] Cy,[170] E,[171] and Te domains,[172,173] modification domains such as the F domain,[133] as well as PPTases.[162] Even the recruitment of the tailoring enzyme $P450_{sky}$ monooxygenase involved in the biosynthesis of the cyclodepsipeptide skyllamycin is mediated by hydrophobic contacts with helices α2 and α3 of the $T_{7,sky}$ domain.[174] In all these cases, the loop-α2/α3 interface deploys different patches and helical alignments to adjust its docking mode. Conformational dynamics earlier observed for the T domain of tyrocidine A synthetase may inherently facilitate the interaction with this variety of binding partners, albeit by more subtle rearrangements in the helical bundle than originally proposed.[159]

These aspects consequently lead to the questions of how NRPSs control substrate trafficking with productive directionality and by which mechanisms does the amino acyl-holo-T domain select the correct catalytic domain in an ordered process? In the simplest futile scenario, the amino acyl-holo-T domain of an elongation module could directly donate its activated substrate to the downstream C domain instead of receiving the upstream nascent peptide chain. One may argue that the T domain exhibits higher affinity to the acceptor site of the C domain, which might be more defined and specifically accommodating the amino acyl substrate. Once upstream condensation is completed, the larger peptidyl cargo may reduce the affinity to the acceptor site such that the larger donor site of the downstream C domain can effectively compete for binding. The scenario gets more complicated with more catalytic stations between the two coupling reactions (see Section 3.6). The principle of competitive binding and scanning for certain substrate intermediates has recently been shown to apply to the biosynthesis of glycopeptide antibiotics (see Section 3.6.11).[175] On the other hand, the observation that the adenylation state of an NRPS module is compatible with simultaneous upstream condensation[131] (Figure 3) possibly hints at a strategy by which the catalytic switch of the A_{sub} domain directly channels the newly established amino acyl-holo-T domain towards the intramodule C domain. At the end, a combination of competitive binding, conformational selection, and synchronized domain movements may guide substrate migration and catalytic turnover. Further biophysical investigations of the dynamics in substrate-loaded T domains and their interaction modes with partner domains are required.

3.3. C Domains

C domains (ca. 50 kDa) catalyze the central coupling reaction of the amino acyl or nascent peptidyl intermediate of module $n-1$ to the α-amino group of the building block attached to module n. The C domain is a V-shaped

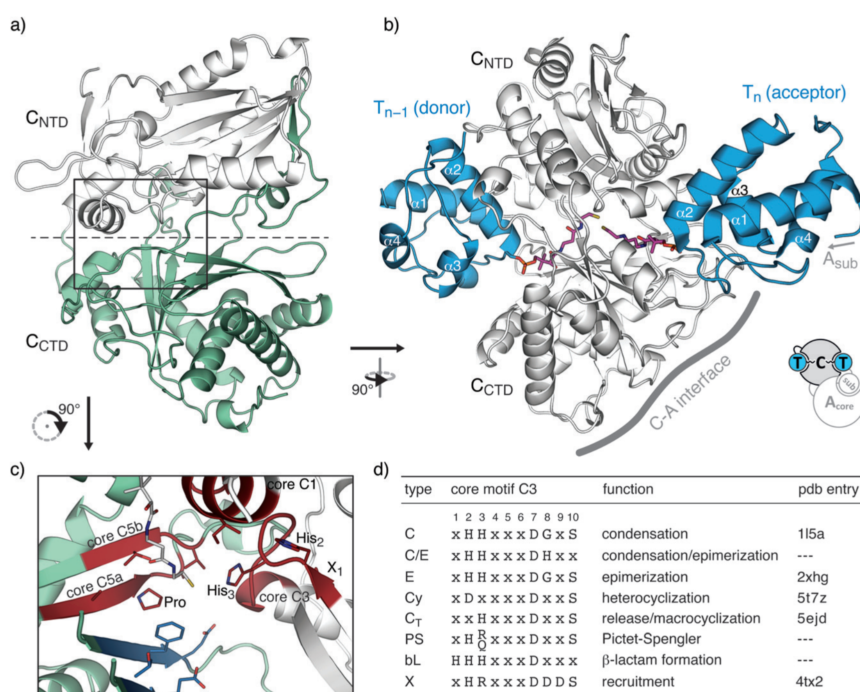
 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

pseudodimer of an N-terminal (C_{NTD}) and a C-terminal subdomain (C_{CTD}).[176] As members of the chloramphenicol acetyltransferase (CAT) superfamily, the C_{NTD} and C_{CTD} subdomains form a central cleft at their interface, which the donor and acceptor Ppant arms have to penetrate from opposite sides to reach the conserved active-site motif HHxxxDG (Figure 5).[177] The second histidine in this motif has been proposed to act as the general base to promote nucleophilic attack of the α-amino group on the thioester[178–180] or to stabilize the tetrahedral transition state.[181,182] The mechanism of C domains is still under debate, since the catalytic impact of the His residue varies markedly for different C domains. An alternative mechanism was proposed recently, after the 1.6 Å crystal structure of an engineered cysteine variant of the first C domain of the calcium-dependent antibiotic synthetase (CDA-C1-E17C) had been solved.[183] An acceptor substrate mimic was covalently tethered to the E17C mutation site and revealed hydrogen bonding between the α-amino group of the substrate and the ε-nitrogen atom of H157 in the catalytic motif as well as the backbone carbonyl oxygen atom of S386. The authors concluded that the hydrogen-bonding pattern involving H157 might simply constrain and correctly align the α-amino nucleophile. More structural data on substrate-bound C domains is required to fully understand the catalytic mechanism.

Comparison of all currently available C domain structures indicates that there are opening and closing dynamics between the C_{NTD} and C_{CTD} lobes (up to 25° in amplitude).[129,131,176,181,184] This aspect is interesting in terms of the extent to which the nascent peptide chain is possibly accommodated and recognized at the donor site of C domains. It has been suggested that C domains play the role of a second selectivity filter during NRP synthesis.[185] In the case of incorrect substrate selection by the A domain, a second proofreading at the C domain minimizes the error rate of an NRPS. As it becomes successively more complicated to control the sequence accuracy of a growing peptide chain, the acceptor site of the C domain can be considered as its major selectivity filter.[185] Nevertheless, as a consequence of the deficit of substrate-bound C domain structures, efforts to deduce a specificity-conferring code equivalent to that of A domains have proved challenging.

C domains have three major interaction partners, that is, the intramodule A and T domains as well as the donor T domain of the upstream module (Figure 3). The C-terminally located A domain forms an interface with the C_{CTD} subdomain,[129,131] which varies in size depending on the catalytic state of the module. Since both the A_{core} and A_{sub} domains are involved in the interaction, this interface contracts upon rotation of the A_{sub} domain from approximately 1050 $Å^2$ (adenylation state) to approximately 800 $Å^2$ (thiolation state).[131] Given the catalytic role of the A_{sub} domain rotation, the C_{CTD}-A_{sub} interface may emerge as an important determinant of NRPS efficiency.[142]

Equally important is the intramodule interaction of the C domain with its acceptor T domain. There is a consistent picture of how the T domain docks to the acceptor site of the C domain to deliver the activated amino acid. Crystal structures of two different termination modules with C-A-T-Te topology, namely apo-SrfA-C-S1003A[129] and holo-AB3403,[131] show a similar binding interface between the T and C domains. Major contributions to this interface arise from α-helices of both C_{NTD} and C_{CTD}. In the structure of holo-AB3403,[131] the Ppant arm adopts an extended conformation and penetrates a tunnel in the C domain to approach the catalytically active His residue (Figure 5). Intriguingly, this binding mode of the acceptor T domain appears to not only promote the condensation reaction, but basically represents the adenylation state (Figure 3). It has, therefore, been suggested that an NRPS module enhances its efficiency by



type	core motif C3 (1 2 3 4 5 6 7 8 9 10)	function	pdb entry
C	x H H x x x D G x S	condensation	1l5a
C/E	x H H x x x D H x x	condensation/epimerization	---
E	x H H x x x D G x S	epimerization	2xhg
Cy	x D x x x x D x x S	heterocyclization	5t7z
C_T	x x H x x x D x x S	release/macrocyclization	5ejd
PS	x H R/Q x x x D x x S	Pictet-Spengler	---
bL	H H H x x x D x x x	β-lactam formation	---
X	x H R x x x D D D S	recruitment	4tx2

Figure 5. Scaffold of C domains and derivatives. a) Overall architecture of the pseudo-dimeric C domain[131] with its central cleft formed by the two lobes C_{NTD} (gray) and C_{CTD} (green). b) Structural model of the condensation state: the complex of the C domain and its acceptor T domain (termination module AB3403) in the adenylation state (pdb 4zxi)[131] is superimposed with the T-E didomain structure (only the donor T domain is shown) of tyrocidine synthetase (pdb 5isx).[171] The Ppant arms (magenta) of the opposing holo-T domains approach in the central cleft close to the catalytic His residue. The C-A domain interface is indicated. c) Close-up view of the catalytic pocket seen from the acceptor site (pdb 4zxi).[131] C_{NTD} and C_{CTD} are color-coded as in (a). The conserved core motifs C1, C3, and C5[177,186] (red) are juxtapositioned to the incoming thiol group of Ppant (white). Additional residues that could potentially interact with the substrate side chain are indicated in blue (downstream of core motif C7). In the case of E domains, the Pro residue in core motif C5a is replaced by Glu, thereby opposing the catalytic His_3 residue of core motif C3.[171,182] A His_1 residue in motif C3 indicates C domains are capable of catalyzing the formation of β-lactams.[165] This His_1 residue would be well positioned to interact with His_2 and/or His_3 to alter the catalytic properties. The C3 core motifs of C domains ($^{L}C_L$, $^{D}C_L$, and $C_{starter}$)[186] and other homologues are summarized in (d).

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

utilizing one conformational state to catalyze two reactions synchronously, that is, preparing the next amino acid for thiolation while upstream condensation is still ongoing.[131] This has further implications for the conformation of the A_{sub} domain during the subsequent downstream condensation—the catalytic state whose structure has still not been elucidated. Since the newly generated amino acyladenylate is prone to unproductive hydrolysis, the A_{sub} domain would need to rest in the adenylation state, or a similarly protecting conformation, until both the upstream and downstream condensation reactions have been performed (Figure 3). According to this scenario, the peptidyl-holo-T domain may attain its binding locus on the downstream C domain by an independent reorientation that does not compromise the conformation of the A_{sub} domain. The only crystal structure of this intermodule interaction between C domains and donor T domains is of the excised apo-T-C didomain of the multi-modular tyrocidine synthetase TycC.[181] In the observed binding mode, the Ppant attachment site of the donor apo-T domain shows an insurmountable distance of about 49 Å to the active site His residue of the C domain, and the structure thus appears to represent a nonproductive interaction state. Given the structural homology between C and E domains (see Section 3.6.1), the recent X-ray structure of the excised holo-T-E didomain of gramicidin S synthetase[171] may serve as a model for the T-C domain–domain interaction and the intermodular substrate transfer (Figure 5). In this structure, the T domain is correctly oriented towards the donor site of the E domain and the Ppant arm penetrates a tunnel between the two lobes of the E domain to position the thiol group approximately 3 Å away from the catalytic His residue. The validity of this model was strengthened very recently, when a similar interaction mode was observed for the holo-T-C_T didomain from the fungal NRPS TqaA (see Section 3.5).[187]

Interestingly, nature utilizes the scaffold of the C domain and modulates its core motifs[186] as well as its protein surface to generate structural homologues with diverse NRPS-relevant functions, for example, epimerization,[182] cyclization,[170] β-lactam formation, or recruitment[188] of auxiliary enzymes (see Section 3.6). Moreover, C domains can be differentiated according to their stereoselectivity: $^{L}C_{L}$ domains promote the coupling of two L-amino acids, whereas $^{D}C_{L}$ domains[189] catalyze the condensation of an upstream D-amino acid and a downstream L-amino acid. The $^{D}C_{L}$ domain subtype is commonly located downstream of an E domain, which itself is a descendant of the C domain.[182] Biochemical studies even reported on dual E/C domains with epimerase and $^{D}C_{L}$ activity,[190] thereby explaining D-configured building blocks in NRPs despite the absence of E domains in the corresponding NRPS modules.[191] Another example of functional diversity is C domains capable of forming ester instead of amide bonds, for example, in the biosynthesis of the antitumor antibiotic C-1027.[192] Likewise, cyclodepsipeptides such as enniatin and beauvericin contain amide and ester bonds, and the respective NRPSs harbor different types of C domains that perform these two coupling reactions in an alternating fashion.[193] Moreover, Rausch et al. identified a C domain subtype distinct from the other clades in their phylogenetic analysis: as a result of its localization in NRPS initiation modules, where it catalyzes the acylation of the first incorporated amino acid, this subtype has been termed a starter C domain ($C_{starter}$).[186] Members of this subtype can be found in the NRPS gene clusters of lipopeptides, for example, surfactin, lichenysin, and fengycin (all bearing a 3-hydroxy fatty acid at the N terminus) as well as pristinamycin, enterobactin, and actinomycin (all bearing a 2-hydroxybenzoate moiety at the N terminus). Biochemical studies by Kraas et al. proved experimentally that the first C domain of the initiation module of the surfactin assembly line SrfAA-C1 indeed functions as a $C_{starter}$ domain and is inherently competent to promote condensation between 3-hydroxymyristic acid-*S*-CoA and the T-domain-tethered glutamate with high selectivity.[123] Finally, terminal C_T domains in fungal NRPS systems appear to functionally replace Te domains in terms of macrocyclization during the detachment process (see Section 3.5).[187,194] From an enzymatic point of view, the C domain and its derivatives can be regarded as the most versatile class of domains in NRPSs.

3.4. COM Domains

The modules of an NRPS assembly line are commonly distributed on several polypeptide chains, for example, A, B, and C (type I NRPS) rather than being incorporated in a single protein (type II NRPS). This leads to the question how nature guarantees a specific communication and intermolecular substrate transfer between partner modules (A-B and B-C), and at the same time precludes biosynthetically futile interactions of nonpartner modules (A-C), consecutive interactions between modules of the same type (B-B), or even cross-talk between different NRPS systems. From sequence alignments, Hahn and Stachelhaus identified stretches of 20–30 amino acids at the N and C termini of NRPSs and assigned a putative communication-mediating (COM) function.[195] The two partner COM domains were later designated as the C-terminal donor COM^{D} domain and the N-terminal acceptor COM^{A} domain.[196] Indeed, C-terminal truncations of more than three amino acids at the COM^{D} domain of TycA in tyrocidin synthetase rendered TycA incompetent to recognize its downstream partner module TycB1.[195] Moreover, charge inversion at the COM^{A} domain of module TycC1 reduced the productive interaction with its cognate donor module TycB3, but facilitated communication with the non-cognate donor system TycA.[196] The portability of the COM system was demonstrated by extensive in vitro and in vivo experiments, in which rational COM^{D}/COM^{A} domain swapping was employed to redirect substrate flux within the tyrocidin NRPS assembly line as well as between modules of different NRPSs (tyrocidin, bacitracin, and surfactin synthetases).[196–198] In fact, excision of the C-terminal E domain of TycA generated an A-T-COM^{D} construct that retained the capability to interact with its cognate COM^{A} domain.[196] These findings imply that intermolecular communication between NRPSs is primarily regulated by compatible COM domain pairs, with only small contributions from the domains they are anchored to. Sequence analysis predicted an α-helical conformation for COM domains. Accordingly,

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

coiled-coil-like interactions with mostly polar and electrostatic contributions have been proposed.[196] Although the N-terminal residues of SrfA-C indicate some propensity for α-helical structures,[129] further structural data of COM domains are needed to substantiate these models.

In all the above-mentioned examples, COM^D and COM^A domains are consistently localized at the C terminus of E or Ox domains and at the N terminus of C domains, respectively. However, there are many NRPSs that do not show such arrangement, thus indicating that the COM domain might be just one recognition system employed by NRPSs. Indeed, sequence comparison of PKS-NRPS hybrids from myxobacteria and cyanobacteria revealed, in addition to rather variable C-terminal extensions, two distinct families of N-terminal recognition elements.[199] Whilst terminal elements at NRPS-PKS junctions appear variable and possibly represent a third recognition system, PKS-NRPS and NRPS-NRPS interfaces share a common docking mechanism. A representative of those N-terminal docking domains, TubCdd (73 amino acids (aa)) of the tubulysin assembly line (NRPS-NRPS junction), has been structurally analyzed.[199] The solution NMR structure of TubCdd identified a unique αββαα-fold with an exposed β-hairpin mediating the homodimerization. The solvent-exposed face of this highly charged β-hairpin emerged as the main determinant for binding to the excised C-terminal docking domain, TubBdd (25 aa), of the upstream NRPS TubB. As confirmed by site-specific mutagenesis, conserved charged residues mediate the interaction between TubBdd and TubCdd (low-micromolar K_d). TubBdd was found to be unstructured in solution and proposed to adopt a distinct conformation upon binding to TubCdd. Such characteristics would assign TubBdd to a member of the intrinsically disordered protein domains, which are frequently involved in protein–protein recognition.[200] Very recently, Dowling et al. reported the X-ray structure of the cyclization (Cy) domain from epothilone biosynthesis (EpoB), which bears an N-terminal docking domain (EpoBdd) of about 55 residues.[170] The two domains are connected by a 20-residue linker that confers conformational flexibility to EpoBdd and enables its movement relative to the Cy domain.[170] After recognition of the partner docking domain EpoAdd, this flexibility may help in directing the upstream carrier protein of the PKS module EpoA to the donor site of the Cy domain for efficient substrate transfer. Similar to TubCdd, EpoBdd adopts an αββαα-fold which, however, is monomeric in the crystal. Most importantly, EpoBdd makes no substantial contacts to its C-terminal Cy domain, thus implicating that the swapping of docking domains for NRPS reengineering may not compromise the integrity of neighboring domains.

To date, neither structural information exists on the different types of docking complexes nor are the selectivity rules of the acceptor domains (KS and C domains) fully understood. This has resulted in recent efforts to program new assembly lines a challenging endeavor with only modest success. The NRPS communication system with all its variants of recognition elements and the influence of their flanking core domains, thus, needs to be further investigated.

3.5. Te Domains

Once NRP synthesis is completed, the T domain of a termination module transfers the mature peptide to the C-terminally located Te domain (ca. 30 kDa), which catalyzes peptide release from the NRPS. As an α/β-hydrolase, the Te domain features a conserved catalytic triad of which the Ser residue, or in some cases Cys, (GxS/CxG core motif) attacks the peptidyl-T-domain thioester and thus temporarily anchors the oligopeptide through an intermediate ester bond to the Te domain (Figure 6a). Peptide release from the Te domain occurs either by hydrolysis (water as a nucleophile) or aminolysis (amine as a nucleophile), thereby liberating linear products. Furthermore, Te domains frequently act as cyclases by constraining the peptide's conformation such that it undergoes intramolecular formation of a lactone or lactam. Intriguingly, nature generates a myriad of head-to-tail (e.g. tyrocidin A, gramicidin S) and side-chain-to-tail (e.g. bacitracin A, daptomycin) macrocycles of various ring size. During the release process, Te domains may even function as epimerases, as seen in the biosynthesis of nocardicin.[201]

The X-ray structure of the surfactin Te domain Srf-Te has revealed two conformers in the asymmetric unit which are designated as the open and closed states.[202] These states refer to a lid usually composed of two to three α-helices and a structurally disordered loop region. NMR investigations on the apo-T-Te didomain of the enterobactin synthetase EntF confirmed the dynamic nature of this lid and described the T-Te domain interaction itself.[172] In contrast, an X-ray study on the same system in its holoform provided a detailed picture of the Ppant binding tunnel.[173] The structurally fluctuating lid opens to accommodate the peptidyl-Ppant arm in a tunnel that bridges the conserved serine residues of the holo-T and Te domains. Varying degrees of lid closure may seal the active site to control unwanted access of exogenous nucleophiles. In addition to the structural properties of the peptidyl substrate, the size, shape, and composition of the substrate cavity of the Te domain as well as conformational dynamics of its lid can be considered to dictate the release mechanism.[203,204] This may, in particular, apply in the case of Te domains of iteratively working NRPSs, for example, enterobactin synthetase[205] or gramicidin synthetase,[206] where multiple loading of building blocks onto the Te domain occurs prior to macrocyclization (see Section 3.7.2).

Experimental data point towards a considerable tolerance of Te domains for substrate loading—a fact that has been exploited for the generation of more than a hundred tyrocidin A analogues (see Section 4.1.2).[207–210] However, some reports have implicated that Te domains can be more stringent in substrate approval in the case of maturation events at the donor T domain. One example is the Te domain of nocardicin synthetase NocB, for which the β-lactam ring is a structural prerequisite for substrate recognition and covalent loading.[201] Similarly, tailoring reactions at the terminal module of pyochelin synthetase involve the formation and subsequent N methylation of tetrahydrothiazole and appear to dictate Te activity.[211] Hence, Te domains can, in principle, fulfill some degree of proofreading to ensure complete substrate processing prior to the irreversible liberation step.

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

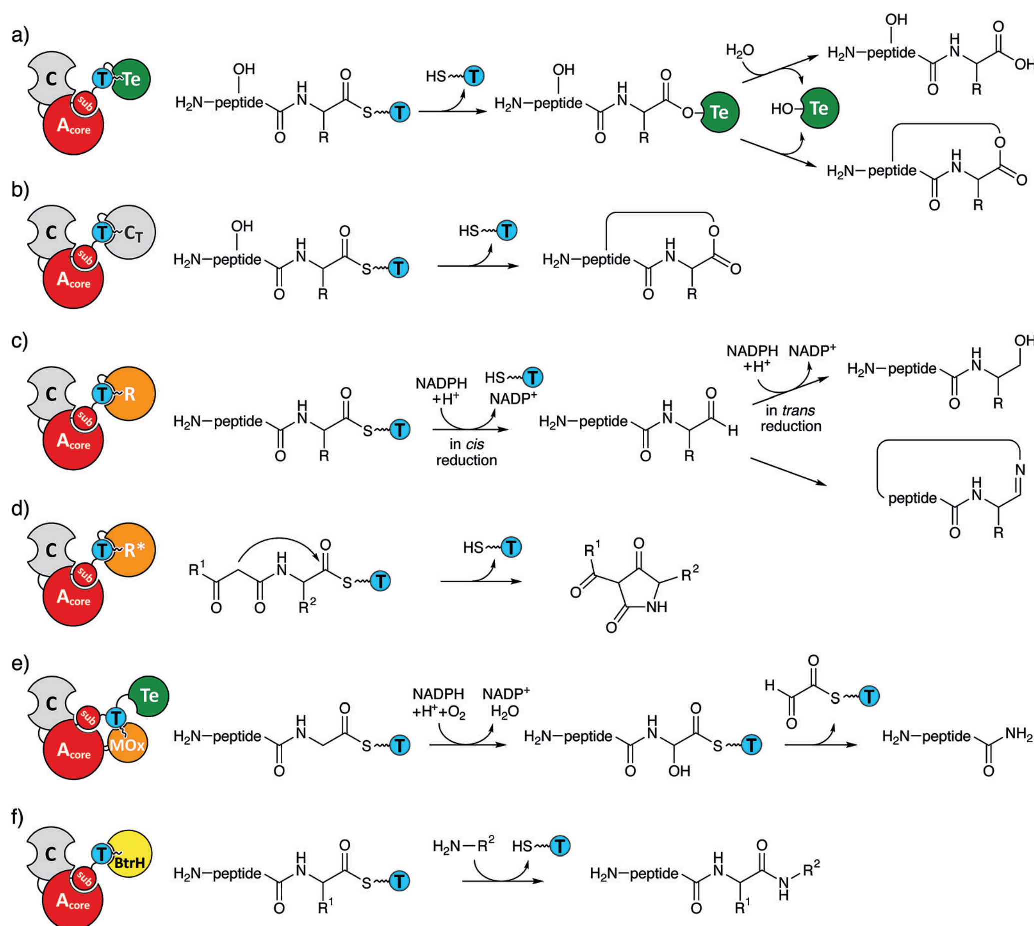


Figure 6. Release mechanisms in NRP biosynthesis. a) Te-mediated release via a transient peptidyl-*O*-Te ester intermediate, which is either hydrolyzed or subject to intramolecular cyclization to the corresponding macrolacton (or macrolactam). b) In fungal NRPSs, terminal C_T domains disconnect the oligopeptide by macrocyclization without formation of an enzyme-bound intermediate. c) Terminal R domains catalyze the reductive release of the oligopeptide to afford aldehyde species, which can undergo either intramolecular cyclization or further reduction to the corresponding alcohols. d) R* domains mediate Dieckmann-type cyclizations of PK-NRP hybrids to yield tetramate moieties. e) In melithiazol biosynthesis,[220] the terminal module incorporates an α-OH-Gly residue, which undergoes spontaneous decomposition and thereby releases an amide product. The T domain is stalled with the residual glyoxylate and needs to be regenerated, presumably by the terminal Te domain. f) Terminal T domains may recruit transferases that disconnect the peptidyl chain by catalyzing attack of exogenous nucleophiles, for example, amino sugars or polyamines, thereby yielding oligopeptides with C-terminal attachments.

The recent crystal structures of termination modules apo-SrfA-C and holo-AB3403 revealed two different orientations of their Te domains despite the arrangements of their C-A-T tridomains being similar.[129,131] Furthermore, there was no electron density for the Te domain in the intact termination module holo-EntF.[131] This finding indicates a high degree of conformational mobility for the Te domain in the thiolation state. Structural heterogeneity has been visualized by negative-stain electron microscopy of EntF, which has demonstrated multiple orientations of the Te domain relative to the other domains.[131] The independent movement of the Te domain presumably reflects the dynamic interaction with its partner T domain.[172] According to the structure of holo-AB3403 in its adenylation state, the T domain would have to rotate by about 180° away from the C domain to present the peptidyl-Ppant arm to the active site of the Te domain.[131] As outlined before, such a movement of the T domain could be analogous to that anticipated for downstream condensation. However, the missing snapshots of both downstream processes—condensation and liberation—will have to be addressed in future structural investigations.

The world of NRPSs is full of alternative routes and, consequently, several Te-independent release mechanisms have been observed (Figure 6). As outlined in Section 3.6.3, NAD(P)H-dependent reductive (R) domains liberate oligopeptides by reducing the peptidyl-*S*-Ppant thioester (Figure 6c).[212–214] Furthermore, fungi have found a functional substitute for cyclization-competent Te domains by employing terminal C-like domains (C_T) as the final constituent of their NRPSs, for example, cyclosporine A or apicidin synthetases (Figure 6b). Gao et al. demonstrated that the C_T domain of the trimodular NRPS TqaA from *Penicillium aethiopicum* is indeed responsible for the macrocyclization of the T_3-domain-tethered tripeptide anthranilate-L-Trp-L-Ala to the tricyclic peptidyl alkaloid fumiquinazoline F.[194] Analogous to C domains, the catalytic activity of the TqaA C_T domain was

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

dependent on the His residue of its core motif SHxxxD, whereas mutagenesis of the Ser residue did not affect the macrocyclization efficiency. It was, therefore, postulated that the catalytic mechanism of C_T domains follows that of canonical C domains with no covalent substrate–enzyme intermediate. Importantly, the C_T domain required interaction with its cognate T domain for proper recognition of the substrate—a feature which may stem from the evolutionary established functional interplay between peptide-elongating C domains and their donor T domains. Very recently, Zhang et al. reported the X-ray structures of the C_T domain and the holo-T-C_T didomain of TqaA.[187] Although sharing the same overall protein fold with canonical C domains, TqaA-C_T is characterized by a compaction of its two lobes and a blocking of its (nonfunctional) acceptor site. The holo-T-C_T didomain structure demonstrates how the upstream T domain associates with the donor site of C_T and inserts its Ppant arm into a narrow channel between the two lobes to position the terminal thiol near the catalytic His residue of the C_T domain. A hydrophobic pocket lined with various aromatic residues has been proposed to structurally confine the aromatic substrate for efficient macrocyclization.

Finally, the discovery of the PK-NRP hybrid paenilamicin of the bee pathogen *Paenibacillus larvae* indicated the presence of a so-called BtrH-like domain downstream of the terminal T domain.[143] This terminal module of the paenilamicin assembly line incorporates a glycine residue, which becomes C-terminally decorated with the polyamine 4,3-spermidine in the mature product. Spermidine has been proposed to be recruited by the terminal BtrH-like domain and to act as an exogenous nucleophile to detach the linear peptide. The BtrH enzyme has been previously described in the biosynthesis of the aminoglycoside antibiotic butirosin from *Bacillus circulans* and functionally assigned as an aminoglycoside *N*-acyltransferase, which mediates release of the ACP-bound γ-L-Glu-4-amino-2-hydroxybutyrate by nucleophilic attack of the exogenous aminoglycoside (Figure 6 f).[215] Understanding the mechanism and substrate binding properties of BtrH-like systems has the potential to allow this alternative release mechanism to be exploited to couple various moieties to oligopeptides and thereby to further expand their chemical and functional diversity.

As an integral part of the NRPS assembly line, type I Te domains (TeI) act in *cis* to specifically recognize, macrocyclize, and release mature products. By contrast, type II thioesterases (TeII) are autonomous repair enzymes (Figure 3). They target stagnating T domains, which have been misprimed with highly abundant short-chain acyl-CoAs, for example, acetyl-CoA, as a result of the low selectivity of PPTases.[216] Similarly, once proofreading at the C domain acceptor site has successfully identified a wrongly activated amino acid, the substrate shuttle system needs to be revived by releasing the aminoacyl moiety that blocks the Ppant arm.[217] In contrast to TeI domains, TeII enzymes possess an easily accessible substrate-binding cavity such that they are less capable of accommodating or constraining peptidyl chains.[218] According to their repair function, TeII enzymes exhibit low substrate specificity. A screening process for low affinity substrates has been proposed that would allow recognition of intermediates with aberrant life times and, thus, to kinetically discriminate between correct and incorrect building blocks.[217]

Strikingly, Hou et al. described an unusual tandem arrangement of two terminal Te domains in the NRPS assembly line of the antibacterial depsipeptide lysobactin from *Lysobacter* sp. ATCC 53042.[219] Biochemical studies identified the first Te domain to be solely responsible for the macrocyclization and product liberation, whereas the second Te domain exclusively catalyzed hydrolysis of the linear peptide and relieved misprimed T domains.

Given the plethora of macrocyclic natural products, Te domains are of central importance for future efforts to develop new cyclic peptides with high rigidity and proteolytic stability. Whilst TeII enzymes ensure efficient biosynthesis of natural products in heterologous hosts, stand-alone TeI domains may prove useful for the semisynthesis of diverse macrocycles.[207,210]

3.6. Introducing Structural Complexity

It is the diversity of chemical traits that renders NRPs so versatile in their biological functions and likewise successful in resisting the defense mechanisms of hosts. The variety of chemical structures arises from the ability of NRPSs to incorporate various building blocks as well as to install additional modifications during and after NRP assembly. The responsible modification domains and the underlying principles are briefly summarized in this section.

3.6.1. E Domains

The majority of NRPs are equipped with D-amino acids which confer beneficial properties, for example, predetermining bioactive conformations or lowering their proteolytic susceptibility. Although some A domains are capable of activating D-amino acids which were generated by cytosolic racemaces,[221] most stereochemical transformations are performed in situ on T-domain-tethered substrates by NRPS-integrated E domains (ca. 50 kDa).[222] These specialized domains are embedded in the NRPS assembly line between T and ${}^{D}C_L$ domains (Figure 7 b).

As descendants of the C domain,[186] E domains share conserved sequence motifs with their evolutionary ancestor, as well as the same overall protein fold.[182] For example, the E domain from the initiation module TycA shares 18% sequence identity with the C domain of elongation module TycC-C6, both from tyrocidine synthetase. The excised TycA-E domain adopts the characteristic pseudodimer structure, with a central cleft formed by the N-terminal (E_{NTD}) and C-terminal subdomains (E_{CTD}).[182] In the TycA-E structure, the invariant residues His743 of E_{NTD} (HHxxxD motif) and Glu882 of E_{CTD} (EGHGRE motif) encounter each other in this central cleft at a distance of approximately 7 Å. This opposing arrangement complies with a general acid-base catalytic mechanism of E domains independent of exogenous cofactors. Samel et al. postulated that the Glu residue abstracts the α-proton, while the positively charged His

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

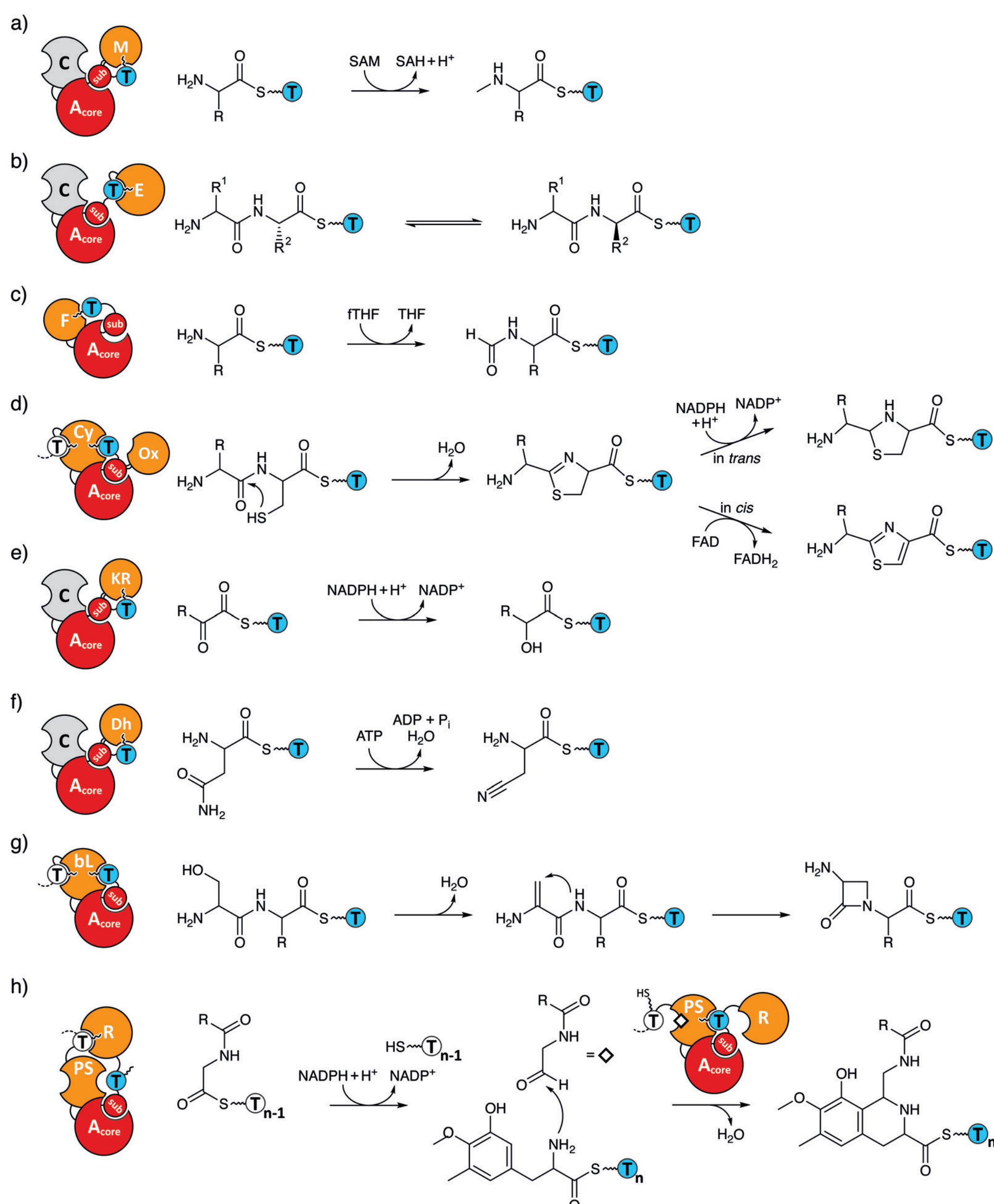


Figure 7. Integral modification domains and their reactions. a) methylation, b) epimerization, c) N-formylation in initiation modules, d) heterocyclization of Ser/Cys residues with subsequent in *cis* oxidation or in *trans* reduction, e) ketoreduction to afford α-hydroxy acids, f) ATP-dependent dehydration of Asn to Cya, g) β-lactam formation through dehydration of Ser/Thr residues, and h) Pictet–Spengler cyclizations involving untethered aldehyde intermediates (only the first reaction cycle is shown for clarity). For b), d), and g) the donor substrates can be either aminoacyl (shown here) or peptidyl intermediates. Typical module topologies are depicted on the left.

residue and the dipole moment of helix α4 potentially stabilize the transient enolate species.[182] This scenario is reminiscent of that proposed for C domains, in which the oxyanion of the tetrahedral transition state could be analogously stabilized, and might, therefore, explain why nature has evolved C domains to catalyze isomerization at the α-position.[181] Strikingly, dual E/C domains lack the catalytically important Glu residue of E domains (replaced with a hydrophobic residue),[186,190] which suggests that these bifunctional catalysts have evolved a distinct mechanism that allows two subsequent reactions in the same catalytic pocket. Dual E/C domains harbor a second HHxxxD motif in helix α1,[186] which covers the active site His residue and thus may modulate its catalytic properties.

In contrast to the C domain, which inherently needs two substrates, the E domain merely awaits one substrate for epimerization. This is reflected in the structural conservation of the "donor site", whereas the opposite entry site—formerly the acceptor site of C domains—shows alterations that 1) occlude the opening to the reaction center and 2) would interfere in T-domain binding.[182] Very recently, the X-ray structure of the excised T-E didomain of the gramicidin

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

initiation module GrsA has been solved in its apo- and holo-forms.[171] The thiol group of Ppant resides between the opposing side chains of His743 and Glu882. As mentioned in Section 3.3, the structural state of the TycA-T-E didomain can be regarded as a representative snapshot of the intermodule interaction of donor T domains with their downstream C domains (Figure 5b).

How are domain–domain interactions regulated in an NRPS elongation module to maintain productive directionality between the catalytic events adenylation, thiolation, upstream condensation, epimerization,[223] and finally downstream condensation? Although E domains tolerate a broad range of substrates, there is experimental evidence that E domains can be discriminated between those that are located in an initiation and those in an elongation module, as they are specialized for epimerization of either aminoacyl or peptidyl substrates, respectively.[224,225] For example, kinetic studies have demonstrated that TycA-E (initiation module) converts both aminoacyl- and dipeptidyl-*S*-Ppant intermediates with similar efficiency.[225] On the other hand, $TycB_3$-E (elongation module) is competent to epimerize Phe-*S*-Ppant, although at about half the rate of its cognate peptidyl-*S*-Ppant substrate.[226] This observation was rationalized to guarantee a preferred transfer route of the aminoacyl intermediate, first to the acceptor site of the C domain for upstream condensation and subsequently to the E domain for epimerization of the peptidyl intermediate, which is then stereoselectively recognized at the adjacent $^{D}C_{L}$ domain. It has been shown that the presence of an N-terminal C domain (acceptor site) is a strong competitor for the aminoacyl-*S*-Ppant intermediate and that it dictates catalytic directionality.[224,226] As Stein et al. have pointed out, an adequate selection and positioning of E domains might be crucial to avoid unfavorable competition between E and C domains in bioengineered NRPS systems.[225]

3.6.2. M Domains

In mammalian epigenetics, N-methylation of Lys and Arg side chains is employed as a transient marker to reversibly regulate histone compaction and gene expression.[227] The N-methylation of backbone amide groups during the biosynthesis of NRPs represents a stable modification that has a significant impact on polarity, hydrogen-bonding capabilities, proteolytic resistibility, and conformational freedom (see review by Chatterjee et al.).[228]

Stand-alone methyltransferases, for example, the structurally characterized glycopeptide *N*-methyltransferase MtfA, are known to tailor NRPs during and after their assembly.[229] However, the majority of methylations in NRP biosynthesis are catalyzed by methyltransferase (M) domains (ca. 45 kDa) which are integrated into the A domain (Figure 7a). The most frequent insertion point is the region between core motifs A8 and A9 located in the flexible A_{sub} hinge domain, but an interruption between core motifs A2 and A3 of the A_{core} domain has also been reported.[146] The lack of structural information on integral M domains has complicated the assignment of their structural core and flanking regions. However, photolabeling studies have confirmed the putative SAM binding region of non-nucleic acid *N*-methyltransferases,[230] which are strongly inhibited by the reaction product *S*-adenosyl homocysteine (SAH) or other derivatives such as sinefungin.[231] Bioinformatic analysis has identified conserved signature motifs useful in genome mining approaches and to delineate the type of methylation.[232] Although N-methylation of the peptide backbone is by far the most abundant type, there are some cases of O-, S-, and even C-methylation. The most prominent example of N-methylation is the cyclic undecapeptide cyclosporine A, which contains seven N-methylated amide groups.[233] Other examples include the cyclodepsipeptides enniatin and PF1022, the microcystins, micropeptins, actinomycin, echinomycin, complestatin, anabaenopeptilide, and bouvardin. All the corresponding NRPSs have their M domains incorporated between core motifs A8 and A9 of the A_{sub} domain.[146] This also applies for the M domain embedded in the NRPS responsible for the biosynthesis of pyochelin.[211] However, in this last case, the tethered Cys substrate of the siderophore becomes N-methylated subsequent to upstream condensation, heterocyclization to thiazoline, and reduction to thiazolidine, that is, at the final stage of all the catalytic events within the module. This is in contrast to the catalytic route described for the second module in the biosynthesis of enniatin and PF1022,[231,234] which involves Nα-methylation of the thiolated substrate prior to condensation, and most probably represents the prevailing mechanism in a minimal modification module. The current data suggest that some NRPSs are capable of producing partly non-methylated derivatives in the absence of SAM,[231,233] while others are strictly dependent on the cofactor for catalytic flux.[234] Domain swapping of a canonical A domain by an A(M) domain in a bimodular actinomycin NRPS caused a dramatic decrease in the catalytic efficiency of the upstream condensation.[235] In this study, it has also been shown that a downstream-positioned E domain is catalytically disabled by N-methylation of the tethered substrate and, thus, appears incompatible with A(M)-T topologies. Ironically, methylation-incompetent M domains are used in pyochelin and yersiniabactin biosynthesis to epimerize aryl-Cys-*S*-thioester intermediates via the transient α-carbanion, which otherwise attacks the methyl group of SAM.[236]

Side-chain N-methylation, for example, of diaminopropionic acid in the biosynthesis of paenilamicin, appears to be catalyzed by separated M domains localized upstream of the respective A domain.[143] The NRPS KtzH is involved in the biosynthesis of the cyclic depsipeptides kutznerides and harbors an M domain (A8–A9 interruption) that generates an *O*-Me-Ser residue.[237] Similarly unprecedented is the S-methylation of Cys side chains by the stand-alone A domain TioN during the biosynthesis of thiocoraline.[238] TioN acts in *trans* to transfer the adenylated substrate to the T_4 domain of TioS. Intriguingly, the M domain of TioN is inserted into the A_{core} domain between motifs A2 and A3. This region is located opposite the docking site of the A_{sub} and T domains and, given the current structural knowledge, represents a so far unexplored interface. However, it is noted here that the docking site for MLPs could be lateral to such domains inserted between motifs A2 and A3. Moreover, the binding partner TioS-A_4 itself bears an embedded M domain

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

between motifs A8 and A9, and catalyzes Nα-methylation of S-Me-L-Cys-S-TioS.[146] It is intriguing that nature achieves both S- and N-methylation of Cys by employing a *trans*-acting A domain with its M domain inserted in the A_{core} (TioN) to complement the acceptor system with its M domain embedded in A_{sub} (TioS). It can be speculated that both M domains might otherwise collide and thus impede formation of a functional complex between TioN-A_{core} and TioS-T_4 for substrate transfer as well as processing.

Finally, NRPS-mediated C-methylation is exemplified by yersiniabactin biosynthesis, in which the 350 kDa PKS-NRPS hybrid HMWP1 performs cyclization of Cys to the corresponding thiazoline species and subsequently incorporates a methyl moiety at its Cα-position.[239] In this case, the M domain is directly embedded between the Cy and T domains.

3.6.3. ***R Domains***

Some linear NRPs and PK-NRP hybrids bear an aldehyde (e.g. α-aminoadipate semialdehyde, saframycin A precursor) or an alcohol function at their C terminus (e.g. glycopeptidolipid, lyngbyatoxin, myxochelin A, myxalamide S, peptaibols), since their biosynthetic assembly lines employ an alternative strategy for product release.

The two- or four-electron reduction of the tethered peptidyl thioester to the aldehyde or the primary alcohol, respectively, is catalyzed by an NAD(P)H-dependent reductase (R) domain (ca. 45 kDa), which replaces the Te domain of a termination module (Figure 6c). Over the past few years, several excised R domains have been biochemically and structurally characterized, for example, the R_{GPL} domain from *Mycobacterium smegmatis* (glycopeptidolipid biosynthesis),[213] the R_{NRP} domain from *Mycobacterium tuberculosis* (unidentified acyl tetrapeptide),[213] the R_{AusA} domain from *Staphylococcus aureus* (aureusimine biosynthesis),[240] and the R_{MxaA} domain from *Stigmatella aurantiaca* (myxalamid biosynthesis).[241] These R domains are organized into an N-terminal subdomain (R_{NTD}, ca. 30 kDa), which adopts a Rossmann fold with a conserved NADPH binding motif (TGxxGxxG), and a partially flexible C-terminal subdomain (R_{CTD}, ca. 15 kDa). R_{CTD} is most likely involved in substrate recognition and positioning by deploying a hydrophobic helix-turn-helix motif. As members of the short-chain dehydrogenase/reductase (SDR) protein family, R domains feature a catalytic triad (T1283, Y1311, and K1315 in R_{MxaA}) proximal to the nicotinamide moiety of the cofactor.

In analogy to Te domains, R domains do not necessarily require tethered substrates to be offered by their preceding T domains. Read and Walsh reported that the T-R_{LtxA} didomain of the lyngbyatoxin gene cluster accepts the *N*-MeVal-Trp-SNAC analogue as a substrate for two consecutive hydride transfer steps to yield the terminal alcohol.[212] Likewise, the dissected R_{GPL} and R_{NRP} domains have been demonstrated to convert valeryl-Phe-Thr-Ala-Ala-*S*-CoA into the corresponding alaninal and alaninol products.[213] The detection of both species in vitro has led to the hypothesis that the aldehyde intermediate dissociates from the R domain and reassociates for a second reduction cycle. Moreover, aldehyde dissociation has been suggested to be necessary for cofactor exchange. This sequential mechanism has been described for other SDRs and is supported by the fact that R_{NRP} and R_{GPL} domains efficiently reduced the aldehyde substrate valeryl-Phe-Thr-Ala-alaninal.[213] Comparison of the k_{cat}/K_M values for the two half reactions of the dissected R_{GPL} domain revealed that aldehyde reduction was at least an order of magnitude more efficient than thioester reduction.[213] However, it needs to be determined whether this difference in catalytic efficiency is an inherent property or a result of the missing interaction with the donating T domain. To date, there is no experimental data illuminating the interaction between R domains and donor T domains. The X-ray structure of the apo-T-R_{AusA} didomain was devoid of electron density for the T domain, thus indicating multiple structural states.[240] Very recently, the stand-alone R domain SpsM was discovered in the biosynthetic gene cluster of the ribosome inhibitor sparsomycin from *Streptomyces sparsogenes*.[214] This finding highlights the ability of the peptidyl-Ppant-T domain to specifically recruit the *trans*-acting R domain for reductive release.

The first reductive step from the tethered peptidyl-Ppant thioester to the aldehyde represents the actual product-release step (Figure 6c), whereas the second round of reduction involves binding of a free aldehyde species. The high catalytic turnover of the second half-reaction observed for R_{GPL}[213] would prevent accumulation of aldehyde intermediates in the cell. It should be stated that some R domains mask the aldehyde species by catalyzing macrocyclization through the N-terminal α-amino group (Figure 6c). In nostocyclopeptide biosynthesis, the terminal R domain catalyzes a two-electron reduction to the aldehyde and mediates subsequent formation of the head-to-tail macrocyclic imine.[242] A similar mechanism has been described for aureusimine biosynthesis, in which the cyclic imine arising from the dipeptide Val-Tyr finally undergoes spontaneous oxidation in air to yield the pyrazinone product.[243]

A further variation to chain-termination mechanisms of NRPSs is represented by so-called R-like domains, or R* domains, which are homologues of R domains, but lack the essential Tyr residue of the Ser/Thr-Tyr-Lys catalytic triad of SDR proteins. R* domains have been shown to catalyze the nonreductive intramolecular Dieckmann cyclization that generates the tetramate scaffold of fungal secondary metabolites such as fusarin, equisetin, (pre)tenellin, and cyclopiazonic acid (Figure 6d). Kinetic experiments with the excised CpaS-T domain and wild-type as well as variant CpaS-R* domains from the cyclopiazonic acid pathway have identified residues S3707 (catalytic triad of SDR proteins), D3803, and H3843 as indispensable for catalyzing the Dieckmann cyclization.[244] Based on the recent structural information on R domains, residues D3803 and H3843 would be structurally positioned in two antiparallel cross-overs between R^*_{NTD} and R^*_{CTD} with a Cα-Cα distance of about 7 Å and within reach of the conserved residue S3707 (ca. 10 Å). Whether these residues indeed direct the generation of the carbanion that attacks the thioester needs to be experimentally validated.

NRPSs utilize R and R* domains as a strategy to avoid negative charges at the C terminus of linear peptides (e.g.

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

linear gramicidin), to further decorate NRPs (e.g. *O*-glycosylated glycopeptidolipid), or to generate cyclic structures (e.g. aureusimine, saframycin, cyclopiazonic acid). Understanding the underlying principles will offer the potential to dictate the release mechanism and thus the chemical properties of liberated NRPs.

3.6.4. F Domains

Although very rare in its occurrence, there is a prominent and well-characterized example of N-formylation in NRP biosynthesis: the N-terminal valine residue of linear gramicidin becomes decorated with a formyl group, thereby masking the otherwise positively charged N terminus (Figure 7c).[245] In view of its head-to-head dimerization during the formation of membrane pores, the uncharged state of its N terminus is very critical for its antibacterial activity. The crucial attachment of the formyl group is catalyzed by the formylation (F) domain (ca. 20 kDa) located at the N terminus of the initiation module of linear gramicidin synthetase (LgrA).[245] Other examples of NRPS initiation modules bearing an N-terminal F domain include kolossin A (*N*α-formyl-D-Leu) and anabaenopeptilide synthetases (*N*α-formyl-Gln). Recent X-ray structures of the F-A-T tridomain module of LgrA trapped in key conformational states revealed that the F and A domains associate in an elongated, although rigid arrangement.[133] In the formylation state, the A_{sub} and T domains undergo extensive movement to translocate the valinyl thioester from the A_{core} subdomain to the catalytic center of the formyltetrahydrofolate (fTHF) dependent F domain (Figure 3). Thereby, the T domain is so positioned that its Ppant attachment site is approximately 21 Å away from the formyl carbon atom of the cofactor. The F domain has been demonstrated to require T-domain-mediated substrate supply and positioning.[245] As a consequence of the separate localization of the small F domain at the N terminus of an initiation module, it is well-suited for implementation in newly designed NRPSs for N-terminal capping. However, the rather strict substrate specificities reported for the F domain as well as the donor site of the downstream C domain[245] demand further insights for efficient utilization of the F domain.

3.6.5. Cy and Ox Domains

Several NRPs including siderophores such as vibriobactin, pyochelin, and yersiniabactin, as well as antitumor agents such as bleomycin, epothilone, and thiocoraline feature characteristic thiazoline- or oxazoline-based heterocycles. The occurrence of these five-membered ring systems arises from the gain of function of a C-domain descendant to catalyze not only peptide-bond formation, but also heterocyclization and dehydration of cysteine, serine, or threonine residues (Figure 7d). This homologue—the cyclization (Cy) domain (ca. 50 kDa)—can be identified on the basis of its signature motif DxxxxDxxS, which corresponds to the catalytic motif HHxxxDGxS of C domains (core motif C3).[246] Mutagenesis studies on the Cy1 domain of HMWP2 (yersiniabactin synthetase) have implicated that both Asp residues but not the Ser residue are essential for condensation as well as heterocyclization.[247] However, recent structural studies on the EpoB-Cy domain (epothilone biosynthesis) have demonstrated that both Asp residues play structural rather than catalytic roles, since they are engaged in two salt bridges with neighboring Arg residues.[170] Instead, the fairly polar substrate pocket of EpoB-Cy appears to be reorganized for the three-step catalytic reaction, and key residues have been identified by mutagenesis and in vitro activity assays. In particular the mutations D449A and N335A caused a dramatic decrease in the rates of product formation compared to the wild-type protein. Residue D449 has been proposed to act as the catalytic base, which may deprotonate the α-amino group of the acceptor substrate to initiate condensation. The catalytic importance of the Asn residue had already been found by Marahiel and co-workers, who employed the first two modules of the bacitracin synthetase BacA fused with the Te domain of tyrocidin synthetase (A-T-Cy-A-T-Te topology) to investigate the sequence of catalytic steps within the Cy domain.[248] The corresponding N900A variant of BacA-Cy was still competent to catalyze condensation, but was unable to heterocyclize Cys.[248] These and several other polar residues of the substrate pocket originate from the C-terminal subdomain of the Cy domain (Cy_{CTD}), thus illustrating the displacement of the catalytic center, usually found in the C_{NTD} subdomain, which allows modulation and expansion of its catalytic spectrum.

Walsh and co-workers studied the six-domain NRPS VibF, which is involved in vibriobactin biosynthesis.[247] VibF contains an unusual tandem arrangement of two Cy domains (Cy_1-Cy_2-A-C_1-T-C_2 topology), with both Cy domains comprising the signature motif DxxxxDxxS. Nevertheless, the authors could show that VibF-Cy_2 exclusively catalyzes the condensation of DHB and Thr, without any ability to produce the methyloxazoline, whereas the condensation-incompetent VibF-Cy_1 performs the heterocyclization reaction. The condensation-critical Asp residue of EpoB-Cy (D449) is conserved in both VibF-Cy_1 (D387) and VibF-Cy_2 (D832), whereas the Asn residue that is critical for heterocyclization in EpoB-Cy (N335) is only found in VibF-Cy_1 (N272) but replaced in VibF-Cy_2 (G719). This further supports the Asn residue playing an important role in the heterocyclization/dehydration reaction of Cy domains. These in vitro observations not only illustrate the loss-of-function of both the Cy_1 and Cy_2 domains in VibF, but also show that this division of labor is only possible due to kinetically separate catalytic events in the Cy domains. Although the exact catalytic residues responsible for the different catalytic steps are still to be identified, the existing data suggest the following catalytic pathway: 1) peptide-bond formation between acyl donors and cysteine/serine/threonine acceptors, 2) cyclization through attack of the thiol or hydroxy side chains at the newly established peptide bond, and, subsequently 3) dehydration of the corresponding hemiaminals.

The potential to employ Cy domains for the production of thiazoline- and oxazoline-containing NRPs is tempting, but several mechanistic details still have to be resolved. In particular, substrate specificity has been shown to be quite different for the NRPS systems tested so far. However, there

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

is consensus that the identity of the T domains and their interaction interfaces play a pivotal role in the stabilization of condensed intermediates to achieve full catalytic conversion at the Cy domain, that is, to the point of heterocyclization.[249]

Further maturation of thiazoline entities may proceed either by two-electron reduction to thiazolidine or by two-electron oxidation to thiazole residues (Figure 7d). The reductive step is catalyzed in *trans* by stand-alone NADPH-dependent reductases (ca. 40 kDa) such as PchG and YbtU in pyochelin and yersiniabactin biosynthesis, respectively. PchG has been demonstrated to recognize its bisthiazoline substrate while it is tethered to the T domain of its partner NRPS PchF and to regioselectively catalyze the reduction of the C-terminal thiazoline residue.[211]

By contrast, the oxidation of thiazoline to thiazole is performed by oxidase (Ox) domains (ca. 30 kDa) that are integral components of the NRPS (Figure 7d). As in the case of canonical M domains, these flavin mononucleotide (FMN) dependent Ox domains are commonly embedded between core motifs A8 and A9 of the A_{sub} domain.[146] This type of domain insertion is exemplified by the NRPSs EpoB and MtaD of the epothilone and myxothiazol biosynthesis pathways, respectively.[250,251] A highly homologous variant of the MtaD-Ox domain is the partner NRPS MtaC, in which the Ox domain resides on the C-terminal side of the T domain (Cy-A-T-Ox topology).[251] Similarly, the NRPS module BlmIII involved in bleomycin production features an $A_{inactive}$-T-Ox domain arrangement.[252,253] Apart from the different loci within NRPSs, these Ox domains share approximately 40% sequence identity and two highly conserved signature motifs Ox1 and Ox2, which relate them to a large group of FMN-dependent oxidoreductases. Schneider et al. have demonstrated that the excised EpoB-Ox domain is functional and stereoselectively recognizes L-thiazolinyl-SNAC analogues as substrates.[250] However, K_M values in the low millimolar range again underline the importance of T-domain-mediated protein–protein interactions for substrate supply and processing. Interestingly, both the excised EpoB-Ox domain as well as the entire module BlmIII were capable of converting oxazolinyl-SNAC into the corresponding oxazole analogue, although with reduced catalytic turnover.

3.6.6. KR Domains

Ketoacyl reductase (KR) domains are prevalent components of PKSs. In rare cases, NRPS systems bear functional insertions that act as α-KR domains (ca. 45 kDa). The cyclic dodecadepsipeptides cereulide from *Bacillus cereus* and valinomycin from several *Streptomyces* strains are potassium-selective ionophors consisting of alternating α-hydroxy acid and α-amino acid building blocks.[254,255] In fact, the responsible A domains initially activate α-keto acids, which are then shuttled to the NADPH-dependent KR domains embedded in the A_{sub} domain (between motifs A8 and A9) for stereospecific reduction to the α-hydroxy acids (Figure 7e).[256] This progressive mechanism in bacterial NRPSs contrasts to fungal cyclodepsipeptide synthetases, for example, for biosynthesis of enniatin, PF1022A, and bassianolide, the A domains of which have evolved to directly utilize α-hydroxy acids and, therefore, dispense with integral KR domains.[193,234,257]

Other examples of in-*cis* reduction of α-ketoacyl-*S*-T domain species can be found during the biosynthesis of kutznerides from actinomycetes and antimycins from streptomyces. KtzG and AntC are the corresponding NRPS modules that catalyze the activation and reduction of the α-ketoacids.[258] These KR domains do not interrupt the A_{sub} domain but reside between the A and T domains. This alternative insertion once more illustrates the plasticity of NRPSs and their ability to incorporate new functions at variable sites.

3.6.7. MOx Domains

Various NRPS pathways are known to involve exogenous monooxygenases (MOx) acting in *trans* on NRPS-tethered intermediates or tailoring released NRP products.[59,259] To date, there are only two examples of NRPS-embedded MOx domains, namely the related MtaG and MelG modules from myxothiazol and melithiazol biosynthesis.[220,251] These termination modules are characterized by a typical C-A-T-Te domain arrangement, but the A domain features an unprecedented integration of the MOx domain (ca. 35 kDa) between core motifs A4 and A5. Myxothiazol A carries an amide moiety at its C terminus. It has been suggested that MtaG incorporates a terminal Gly residue and that its MOx domain then hydroxylates the α-carbon atom of Gly, thereby rendering this residue prone to spontaneous decomposition (Figure 6e).[220] The released myxothiazol A with its C-terminal amide group might be further processed to the carboxylic acid or its methyl ester, whereas the T domain of MtaG needs to be regenerated by its adjacent Te domain, as the remnant glyoxylate still blocks the Ppant arm. The glyoxylyl thioester represents a rather atypical substrate for the type I Te domain, and it needs to be determined whether an exogenous TeII enzyme is involved in the recycling mechanism.

3.6.8. Dehydration

Very recently, structure elucidation of the potent DNA gyrase inhibitor albicidin from *Xanthomonas albilineans* has revealed a central β-cyano-L-alanine (Cya) residue in the otherwise nonchiral oligoaromatic PKS-NRPS hybrid.[27] Strikingly, the responsible stand-alone module Alb04 has been shown to preferentially activate L-Asn as the substrate, thereby hinting at a possible dehydration of the side-chain amide to a nitrile group. Indeed, Alb04 features an inserted sequence between core motifs A8 and A9 (ca. 35 kDa), which is unprecedented in NRPSs, but homologous to members of the adenosine nucleotide α-hydrolase (α-ANH-like III) superfamily with the highly conserved ATP-binding motif SGGKD. It has been postulated that the side-chain amide group of tethered L-Asn becomes O-phosphorylated and formally dehydrated by subsequent dephosphorylation to afford the nitrile moiety (Figure 7f). Although the Asn-to-Cya conversion would represent a unique chemical trans-

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

formation in the world of NRPSs, it still needs to be demonstrated experimentally.

3.6.9. *Pictet–Spengler Reaction*

As outlined in Section 3.6.3, R domains are able to reductively remove peptidyl thioesters from terminal T domains to generate aldehyde species, which may undergo further cyclization reactions (Figure 6c). A remarkable variation of this mechanism is the proposed Pictet–Spengler type formation of the tetrahydroisoquinoline scaffold of antitumor antibiotics such as saframycin, safracin, ecteinascidin 743, and quinocarcin.[260,261] Koketsu et al. proposed an unprecedented seven-step transformation of tethered acyl-dipeptidyl and tyrosyl thioesters into the pentacyclic core structure of saframycin. The iterative mechanism involves the terminal module SfmC with PS-A-T-R topology, in which the Pictet–Spengler (PS) domain represents another homologue of the C domain. In particular, core motifs C1, C3, and C5 of C domains show significant alterations in PS domains, for example, the His motif HHxxxDG of C domains is replaced with the motif HxxxxD (Figure 5d). According to the proposed biosynthetic model,[260] the terminal R domain undergoes a large movement towards the penultimate T domain of the upstream module SfmB to reductively disconnect the donor substrate (acyl-dipeptidyl thioester). The donor site of the PS domain is suggested to recruit the liberated aldehyde species to catalyze imine formation and subsequently the first Pictet–Spengler reaction with the downstream acceptor, an SfmC-tethered Tyr analogue (Figure 7h). A second reduction, at this stage occurring at the terminal T domain of SfmC, releases a bicyclic aldehyde, which is again sequestered by the donor site of the PS domain and yields a tetracyclic scaffold after a second Pictet–Spengler reaction. Final reduction of the thioester allows the amine to attack the aldehyde to form the pentacyclic scaffold of saframycin. This iterative mechanism, thus, requires two essential recruitment steps of already-liberated intermediates. Such a strategy is surprising as it repeatedly diverges from the fundamental concept of NRP synthesis, namely to maintain covalent anchoring of the reaction intermediates to the biosynthetic machinery. Further investigations are required for a deeper understanding of this unprecedented biosynthetic mechanism.

3.6.10. *β-Lactam Formation*

The biosynthesis of monocyclic β-lactam antibiotics such as the nocardicin family has been recently uncovered. Gaudelli et al. unequivocally showed that catalysis of β-lactam ring formation is an inherent feature of the NRPS NocB (A-T-C-A-T-Te topology), which is able to convert T-domain-tethered donor (L-Hpg-L-Arg-D-Hpg-L-Ser-*S*-T_4) and acceptor substrates (L-Hpg-*S*-T_5) to pro-nocardicin G with its characteristic β-lactam core structure.[165] The authors pointed out an additional His residue in the catalytic motif HHHxxxDG of the NocB-C_5 domain and linked this signature to a possible mechanistic route of 1) serine dehydration, 2) amine addition at the generated dehydroalanine under inversion of the configuration, and 3) subsequent attack of the amine on the upstream thioester to yield the T_5-tethered β-lactam species (Figure 7g). The NocB-C_5 domain demonstrates once more that C domains can acquire additional catalytic abilities, in this case to dehydrate the seryl moiety of the donor thioester. Based on the structural knowledge of C domains, the additional His residue (H790 in NocB) can be expected to orient towards the catalytic cleft, within reach of the condensation-critical H792 (Figure 5c). Mutagenesis of H790 indeed impaired the production of pro-nocardicin G, and it has thus been argued that the HHHxxxDG signature in C domains in combination with an upstream serine-/threonine-activating module may prove diagnostic for NRP antibiotics bearing the β-lactam scaffold.[165]

3.6.11. *Trans-Acting Enzymes and Their Recruitment*

Besides their own enzymatic versatility, NRPSs further augment the chemical space of their products by recruiting auxiliary enzymes that perform diverse in-*trans* modifications on NRPS-bound intermediates. Such modifications are often essential for bioactivity.[259,262] Alternatively, final products disconnected from the assembly line may also be subject to chemical tailoring, which frequently represents a form of decoration of an already functional peptide. A recent example of this post-NRPS decoration and potentiation of an antibiotic is the carbamoylation of albicidin by the ATP-dependent *O*-carbamoyl transferase Alb15.[263] Another prominent case is the glycosylation of peptides, in particular that observed for glycopeptide antibiotics. Such glycosylation reactions have been investigated by gene inactivation[259,264,265] and in vitro reconstitution[266,267] of biosynthetic pathways. In particular, gene-inactivation experiments have shown that the NRP core structures were still functionally assembled, despite the inactivation of the corresponding glycosyl transferases.

On the other hand, it also became apparent that the inactivation of *trans*-acting enzymes fully stalls the NRPS assembly line, thereby hinting at a well-coordinated communication between NRPS machinery and auxiliary domains. A good example is the synthesis of glycopeptide antibiotics of the vancomycin-type. Three P450 monooxygenases (OxyA, B, C) have been assigned to the oxidative cross-linking of aromatic side chains of a linear peptide precursor to a cup-shaped tricyclic aglycon. In an ideal reaction, the in vitro reconstitution of all three oxygenases together with various readily synthesized linear peptide precursors would yield novel vancomycin aglycons to counter the vancomycin resistance of bacteria. However, extensive gene-inactivation studies on the balhimycin system revealed a strict orchestration of oxidative cross-linkings catalyzed by 1) OxyB (C-*O*-D ring), 2) OxyA (D-*O*-E ring), and 3) OxyC (AB ring).[259,268] These studies also indicated that the linear peptide is still attached to the NRPS during P450 processing.[269] This in-*trans* hypothesis has been proven by extensive in vitro experiments by the Robinson group,[270–273] who observed the establishment of the C-*O*-D ring by OxyB only for the corresponding peptidyl-*S*-T domain thioesters, while no conversion was observed for the free peptidyl acids.[270]

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Very recently, Cryle and co-workers succeeded in elucidating the structural basis for P450-monooxygenase processing on the NRPS template. They were able to obtain the crystal structure of OxyB (ca. 45 kDa) in complex with an unusual C-domain derivative, termed the X domain (ca. 50 kDa), both from teicoplanin biosynthesis (Figure 8b).[188] Alterations in the core motifs of the C domains, for example, a HRxxxDD motif instead of a canonical HHxxxDG motif (Figure 5d), render the X domain catalytically silent. The X Domain is embedded in the termination module between the T_7 and Te domains. Several loop regions obstruct its nonfunctional acceptor site. OxyB binds near to the donor site of the X domain, primarily through a hydrophilic interface and with low micromolar affinity. The rigid body docking to the X_{NTD} subdomain with only subtle rearrangements in both binding partners underlines the recruitment function of the X domain, that is, to attract the exogenous enzyme and to properly align its active site towards the peptidyl-loaded T_7 domain (see recent review by Cryle and co-workers).[274] Peschke et al. found that the three P450 monooxygenases OxyA, OxyB, and OxyC of the teicoplanin biosynthesis pathway compete for the same interface of the X domain, but that it is the substrate cyclization state that dictates which of those three enzymes initiates catalysis.[175] To expand this substrate-scanning mechanism even further, one may anticipate that the terminal Te domain similarly selects its mature substrate to ensure that only the fully processed aglycon is liberated from the NRPS.

How do NRPS modules, which are devoid of the X domain, recruit tailoring enzymes? Biochemical and structural studies have shown that the P450 monooxygenase from skyllamycin biosynthesis ($P450_{Sky}$) forms a transient complex with its partner $T_{7,Sky}$ domain to catalyze the β-hydroxylation of the tethered *O*-Me-L-Tyr substrate.[275] For crystallization studies, this interaction was trapped by loading the $T_{7,Sky}$ domain with an imidazole-CoA analogue, which ligates the heme group of $P450_{Sky}$ and thereby stabilizes the complex (Figure 8a).[174] The binding interface has a pronounced hydrophobic character. Skyllamycin bears three sites of β-hydroxylations (β-OH-L-Phe_5, β-OH-*O*-Me-L-Tyr_7, β-OH-D-Leu_{11}), and sequence analysis of the respective T domains indicated that these were different from those T domains at which no β-hydroxylation occurs.[275] However, we are far from understanding all the principles of NRPS recruitment because of a lack of structural data. Considering the equivalent position of X domains and C domains downstream of their "donor" T domain, it is not far-fetched to also assume certain recruitment abilities for the donor face of the C domains. One could even speculate that nature deploys an inactive C domain (the X domain) in a termination module, since no downstream module (C domain) is available to stabilize a productive complex between the donor T domain and the exogenous enzyme. However, the termination module of skyllamycin synthetase (canonical C-A-T-Te topology) and in particular stand-alone NRPSs (see below) show that such stabilization is not consistently required. The issue of which domains contribute to NRPS recruitment, and to what extent, should be addressed in future studies, as this holds the potential to allow in-*trans* modifications to be selectively redirected for the generation of new bioactive peptides.

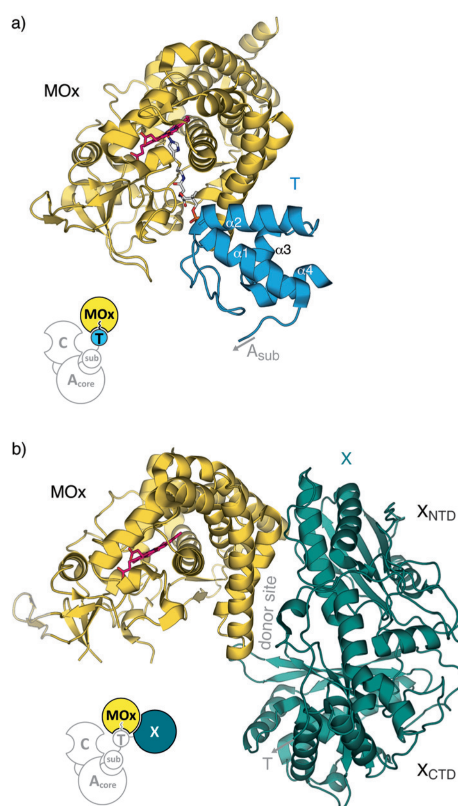


Figure 8. Recruitment of exogenous enzymes. a) T-domain-mediated recruitment of P450 monooxygenases: the T_7 domain of skyllamycin synthetase (blue) in complex with $P450_{Sky}$ (yellow).[174] The binding of the imidazole-derivatized Ppant arm (white) to the heme group (magenta) of $P450_{Sky}$ stabilizes the otherwise transient interaction. b) X-domain-mediated recruitment of the monooxygenase OxyB (yellow).[188] The excised X domain (teal) of the termination module of teicoplanin NRPS employs its donor face (mainly X_{NTD}) to bind OxyB. This arrangement would position the upstream T domain adjacent to the opening of OxyB, similar to the scenario depicted in (a).

3.7. Types of NRPSs

3.7.1. Type A—Linear NRPSs

As introduced in Section 3, the collinearity rule is based on a linear biosynthetic logic, whereby each NRPS module incorporates one designated building block into the propagating peptide chain (Figure 9a). Hence, the number of modules in linearly operating NRPSs (type A NRPSs) equals the number of amino acids found in their products. These numbers vary significantly and so does the dispersion of modules on separate proteins. For example, sevadicin from *Paenibacillus*[276] and bialaphos from *Streptomyces*[277] both represent small linear tripeptides, but the corresponding three modules are distributed on two (SevAB) or three NRPS proteins (PhsABC), respectively. The synthetases of heptameric vancomycin from *Amycolatopsis*,[278] decameric tyrocidine from *Bacillus*,[279] and tridecameric feglymycin from *Streptomyces*[280] likewise consist of three NPRS proteins,

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

namely CepABC, TycABC, and FegQRS. Linear gramicidin from *Bacillus* comprises 15 building blocks, which are assembled by four NRPS proteins (LgrABCD).[281] Kolossin from *Photorhabdus*[191] and syringopeptin from *Pseudomonas*[282] represent the upper extremes in this series. Kolossin A is a linear pentadecamer, which is synthesized by a single giant NRPS (Kol) with a molecular weight of 1.8 MDa. Similarly impressive is the synthetase of syringopeptin 25A: only 3 NRPS proteins (SypABC) assemble 25 building blocks to compose the largest NRP known to date. The rareness of such extended multienzyme assembly lines may simply be ascribed to a decline in catalytic efficiency during progressive cycles of peptide transfer and elongation.

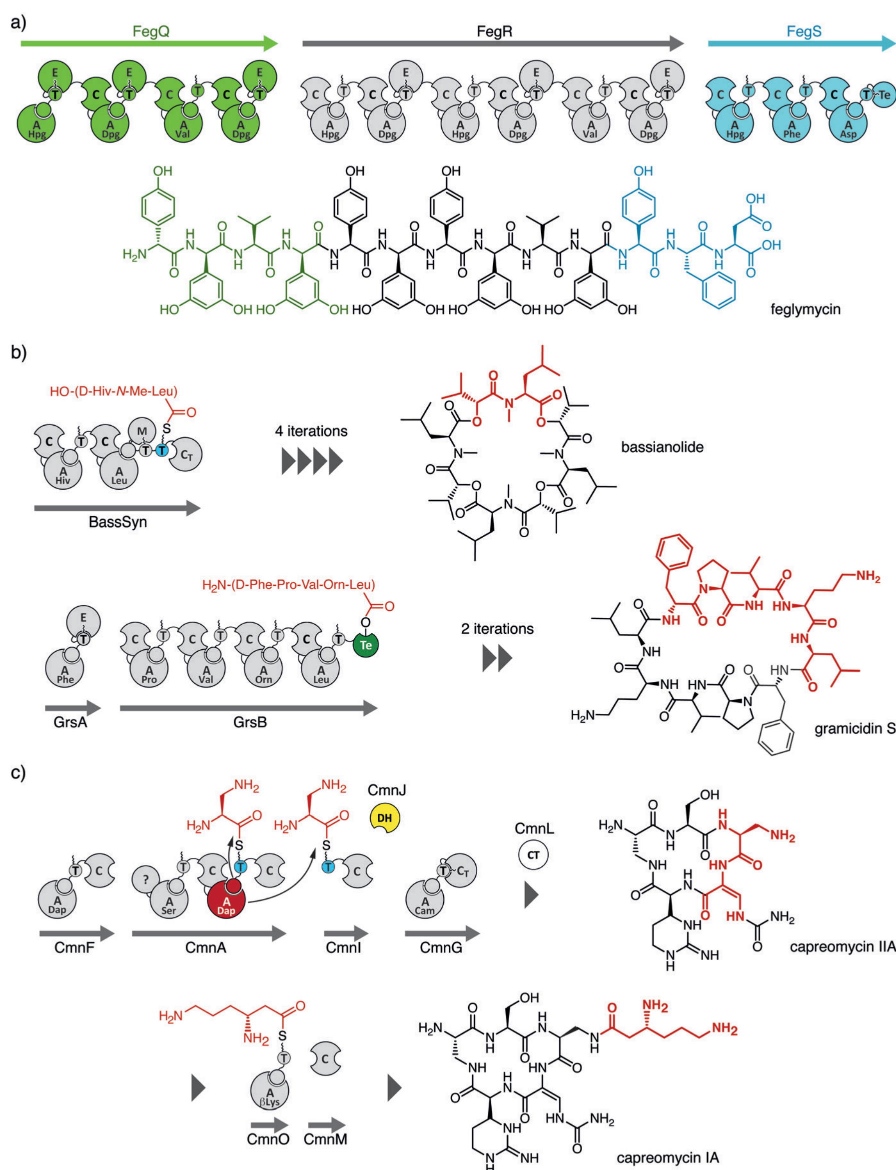


Figure 9. Types of NRPSs and underlying principles. a) In linear NRPSs (type A), each module is used once during a full biosynthetic cycle. The assembly line of tridecameric feglymycin is shown as an example. The modules and their contributions to the mature product are color-coded. b) Iterative NRPSs (type B) reuse the entire assembly line or certain modules and store product intermediates either on T domains (e.g. T_3 of bassianolide synthetase) or Te domains (e.g. in gramicidin S synthetase). These repeating units (highlighted in red) are progressively coupled and macrocyclized once a certain peptide length is achieved. c) Nonlinear NRPSs (type C) are often distributed on several proteins and reuse certain domains during the biosynthesis of a single peptide. In the case of capreomycin biosynthesis,[283] the second A domain of CmnA activates two molecules of L-Dap and loads them onto two distinct T domains: the terminal T domain of CmnA and the T-C didomain CmnI. This mechanism yields a repetition of L-Dap in the final product (red). After processing of Dap4, the stand-alone CmnO and CmnM attach β-Lys (red) to the side chain of Dap3, thereby generating capreomycin IA.

3.7.2. *Type B—Iterative NRPSs*

Some NRPS machineries reuse dedicated modules multiple times during a full biosynthetic cycle (Figure 9b). In many cases, this iterative mechanism (type B NRPSs) causes molecular symmetry of the product. The gramicidin S synthetase is a prominent example of iterative NRPSs.[206,281] Other iteratively synthesized natural products include catecholate siderophores of the enterobactin-type, based on a triseryllactone.[205] Further important examples comprise the family of quinoxaline antibiotics (chromodepsipeptides), such as echinomycin and thiocoraline, which mostly act as DNA intercalators.[284–286] Cyclodepsipeptides (CDPs) represent a structurally distinct group with dipeptides of α-hydroxy acids and α-amino acids as the repeating units. CDPs are synthesized by bacteria (valinomycin, cereulide)[254,255,287] as well as fungi (enniatin, beauvericin, bassianolide, and PF1022).[193,234,257,288] Interestingly, an iterative mechanism of the synthetases of congocidine and asperlicin, which employ a pyrrole and an anthranilic acid substrate, respectively, cannot be easily deduced from the product structures.[289–292]

Mechanistically, the iterative biosynthesis affords a repetitive utilization of modules and requires a so-called "waiting position" for the reaction intermediates. This storage function can be either fulfilled by T domains, for example, for congocidine (loading T domain)[289,290] and enniatin (terminal T domain),[193,293] or by terminal Te domains as in the case of echinomycin,[284] gramicidin S,[206] and enterobactin.[205] The thioesterified or esterified intermediates can be monomeric, as hypothesized for the stand-alone T domain Cgc19 of congocidine biosynthesis, or oligomeric in nature, depending on

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

the number of upstream modules. Interestingly, T domains destined for the storage of intermediates can be arranged very differently in their modules. For example, tandem T domains can be found in the fungal enniatin synthetase (C*-A-T-C-A-**T_2**-**T_3**-C_T topology), where the terminal T_3 domain is presumed to store the elongating depsipeptide.[193,293]

Once tethered to their designated waiting positions, the repeating units may undergo 2–5 cycles of elongation until a critical peptide length is reached, thereby triggering release from the NRPS template by Te (bacteria) or C_T domains (fungi). The binding capacity of these domains may thus dictate the number of repeating units in the product. As macrocyclization is the predominant release mechanism in iterative systems (except for congocidine), the conformational characteristics of the peptide itself may also affect the optimal length that can be accommodated and positioned for intramolecular attack.

A reasonable explanation why nature makes use of an iterative mode may be that oversized and energetically more costly NRPSs with their high susceptibility to misfolding and proteolysis are preferably avoided. For example, the iterative PF1022 synthetase (350 kDa) produces an octameric depsipeptide, whereas a linearly operating NRPS system would either require a single protein with a molecular weight of almost 1 MDa or separate proteins with specific intermolecular recognition elements.

3.7.3. *Type C—Nonlinear NRPSs*

A further variation of the iterative mechanism is the nonlinear mode (type C), which relies on the fact that one domain (not an entire module) is used more than once during NRP biosynthesis (Figure 9c). In mannopeptimycin assembly, for example, one A domain (MppB-A_2) provides β-hydroxyenduracididine-AMP not only for its intramodule T_2 domain, but also for the T_3 domain of the downstream module that is devoid of an A domain (C_2-**A_2**-**T_2**-C_3-**T_3**-E topology of the last two modules).[83] The result is a repetition of the enduracididine residue in mannopeptimycin. A similar mechanism applies to the biosynthesis of viomycin and capreomycin, during which one A domain has to supply activated substrates for its flanking T domain and a *trans*-acting T-C didomain before peptide elongation can occur (Figure 9c).[79,283] The biosynthetic route to bleomycin similarly resorts to a *trans*-acting T domain (BlmIII with $A_{inactive}$-T-Ox topology), which loads Cys for subsequent heterocyclization and oxidation.[250,253] Hence, the reuse of the Cys-activating A domain of the upstream module (BlmIV) in conjunction with its flanking Cy domains yields the terminal bisthiazole fragment of bleomycin. A recurring feature of type C NRPSs is the dissociation of assembly lines into various stand-alone enzymes. This nonmodular trait is discussed in Section 3.7.4.

3.7.4. *Stand-Alone NRPSs*

Although NRPSs are commonly described as large multienzyme machineries that follow a modular logic, many dissociated NRPSs have been reported, that is, entire modules or even single domains operate as stand-alone enzymes. In fact, comprehensive bioinformatic analysis has demonstrated that dissociated nonmodular NRPSs are highly abundant in bacteria.[294]

Such stand-alone enzymes or modules are frequently involved in the generation and delivery of unusual building blocks to the main assembly line (see Section 2). In nikkomycin biosynthesis, stand-alone NikP1 (MLP-A-T tridomain) activates His and recruits the monooxygenase NikQ for β-hydroxylation to yield the precursor β-OH-His.[56] In chloramphenicol biosynthesis, the protein CmlP (A-T-R tridomain) is responsible for the activation of *p*-aminophenylalanine and the recruitment of tailoring enzymes.[295] Two stand-alone A domains (β-Lys activation) and one T-C didomain are involved in streptothricin assembly.[296] The stand-alone proteins PltF (A domain) and PltL (T domain) involved in pyoluterin biosynthesis, activate and load L-Pro.[127] After oxidation and halogenation, the corresponding dichloropyrrolyl-*S*-PltL is shuttled to the downstream PKS assembly line. The uncommon cyclopropane moiety in coronamic acid (e.g. coronatine) arises from chlorination of the γ-methyl group of L-*allo*-Ile-*S*-CmaD (stand-alone T) and subsequent attack of the α-carbanion.[112] A similar mechanism applies during the conversion of L-Ile into 2-(1-methylcyclopropyl)-Gly in kutzneride biosynthesis, with KtzB and KtzC representing stand-alone A and T domains, respectively.[258] In siderophore biosynthesis, aromatic carboxylic acids such as 2,3-dihydroxybenzoate and salicylate are commonly activated by stand-alone A domains. They load the activated starter units either onto stand-alone T domains (designated as aryl-carrier proteins, ArCPs), for example, in enterobactin assembly,[298] or directly onto NRPS-integral ArCPs, for example, in pyochelin and yersiniabactin biosynthesis.[211,299]

The building block β-cyano-L-Ala of albicidin originates from L-Asn, which becomes activated and converted by *trans*-acting Alb04 (A(Dh)-T tridomain).[27] The Alb04-bound β-cyano-L-Ala is most likely delivered to the main assembly line (T_2 domain) by transthiolation by the shuttle protein Alb11. Such a mechanism has been described for the homologues SyrC and CmaE from syringomycin and coronamic acid biosynthesis, respectively.[60] These aminoacyl transferases employ a conserved Cys residue to form an aminoacyl-*S*-enzyme complex and transmit tethered building blocks from one holo-T domain (site of activation) to their destination holo-T domain (site of condensation).

Even stand-alone C domains have been identified, for example, VibH of the dissociated vibriobactin synthetase[176] and Zmsk of the zeamine pathway.[300] Both enzymes perform the condensation of carrier-bound donor substrates (2,3-hydroxybenzoate and octanoate) and nontethered acceptor polyamines (norspermidine and zeamine II), which indicates that correct carrier-mediated positioning of the thioester moiety at the donor site is required for efficient coupling.

An extreme case of assembly line fragmentation has been observed for the PK-NRP hybrid andrimid, the biosynthesis of which is orchestrated by several stand-alone A, T, and A-T proteins.[301] Moreover, the gene cluster is devoid of C domains, but encodes the transglutaminase homologue AdmF that is responsible for the formation of peptide

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

bonds.[302] Another highly fragmented assembly line is that of the peptidyl nucleoside antibiotic pacidamycin, which includes various stand-alone NRPSs of no more than three domains.[116]

These and many other examples illustrate that basically all core domains of NRPSs and various combinations thereof can operate as autonomous proteins in a nonmodular mode. Specific channeling of intermediates most likely relies on established domain–domain interfaces of NRPSs, compatible docking domains, and substrate specificity.

Given their rather simple topology and reduced size, stand-alone NRPSs may emerge as valuable starting points to redesigned biosynthesis pathways. In particular the delivery of unusual building blocks at defined stages of NRP assembly might become an attractive strategy once we have gained a molecular understanding of substrate specificity and the complex protein–protein communication network.

3.7.5. PKS-NRPS Hybrids

A multitude of natural products appear as hybrids of NRPs and PKs (Figure 10). In fact, recent genome mining initiatives have found that approximately one third of more than 3300 identified gene clusters represent PKS-NRPS hybrids.[294] The corresponding biosynthetic machineries are thus capable of communicating and collectively orchestrating the biosynthesis of highly complex structures. As a result of the common thiotemplate mechanism, the structurally similar carrier proteins/domains clearly represent the adaptors between PKS and NRPS modules.

Important examples of mixed PKS-NRPS gene clusters are those of streptogramin A,[303] the electron-transport inhibitor myxothiazol,[251] as well as the aminopolyols zwittermicin[304] and paenilamicin.[143] Another prominent example is the cytostatic bleomycin, which incorporates a single PKS module into its otherwise NPRS-type assembly line.[253]

On the other hand, compounds exist for which the involvement of amino acids as building blocks is only evident from the knowledge of the gene cluster or classical feeding studies. Such examples are rapamycin and epothilone, which are mostly PKs in nature, but borrow a single NRPS module for their assembly.[305] For example, EpoB (Cy-A(Ox)-T topology) is responsible for the incorporation of the thiazole building block of epothilone. Functionally filed between the two PKS modules EpoA and EpoC, EpoB bears N- (55 aa) and C-terminal (8 aa) extensions that have been allocated as recognition elements at the two PKS-NRPS interfaces, namely EpoA-ACP/EpoB-Cy and EpoB-T/EpoC-KS.[305] The N-terminal docking domain EpoBdd has been shown to adopt a monomeric state in the crystal.[170] By contrast, the biophysical characterization of the homologous N-terminal docking domain TubCdd of the tubulysin PKS-NRPS hybrid showed that the exposed β-hairpin that is rich in basic residues not only mediates the intermolecular recognition of its upstream partner module TubB, but also triggers homodimerization of the excised TubCdd.[199] This is an important finding, since PKSs have been shown to operate in dimeric states, whereas NRPSs appear to function as monomers.[306] Any coupling of the two machineries would thus require a compatible oligomerization state, and dimeric TubCdd implies that such an adaption may occur for the NRPS component of hybrid assembly lines. However, it is an open question whether a single PKS module, for example, in bleomycin biosynthesis, could impose its quaternary organization onto the entire NRPS assembly line. Future structural investigations, in particular cryo-electron microscopy, will hopefully shed light on the quaternary architecture and substrate flux in pure NRPS and hybrid PKS-NRPS systems.

3.8. Post-NRPS Dimerization

An interesting phenomenon in NRP biosynthesis is post-NRPS maturation by dimerization, with typical examples being actinomycin and himastatin. Actinomycin synthetase releases the half-molecule 4-methyl-3-hydroxyanthranilic acid (4-MHA) pentapeptide lactone. In a subsequent step, dimerization of the anthranilic acids to the phenoxazine-core occurs either spontaneously in the presence of molecular oxygen or by the action of a phenoxazinone synthase.[307] The inactive himastatin monomer undergoes a biaryl aromatic coupling at the Trp side chains catalyzed by the P450 monooxygenase HmtS.[308]

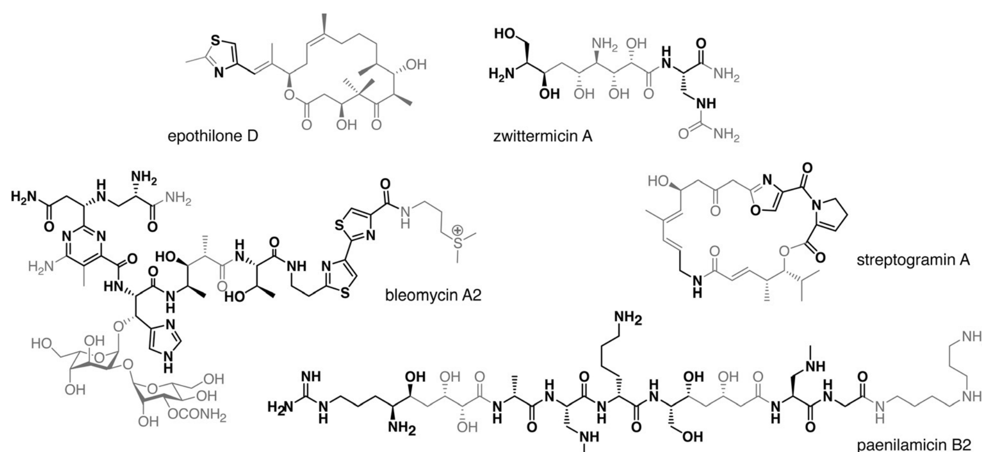


Figure 10. Structures of PK-NRP hybrids. The examples demonstrate the various NRP (black) portions in the hybrids: epothilone D (one NRPS module), zwittermicin A (two NRPS modules), streptogramin A (three NRPS modules), paenilamicin B2 (seven NRPS modules), and bleomycin A2 (only one PKS module).

3.9. Prodrug Mechanisms

The prodrug principle, that is, the biosynthesis of an inactive NRP precursor that is later transformed into the bioactive species, has already been described for

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

the protoxins bialaphos and phosalacine.[309] In both cases, proteolytic cleavage of tripeptide precursors liberates the nonproteinogenic amino acid phosphinothricine, thereby exerting toxicity through inhibition of glutamine synthetase.

More recently, another prodrug mechanism has been discovered. Investigations of the hybrid PKS-NRPS gene cluster of the antibiotic xenocoumacin from *Xenorhabdus nematophila* revealed that the first module of the NRPS XcnA ($C_{starter}$-A-T-E-C-A-T topology) is responsible for the incorporation of an N-terminal *N*-acyl-D-Asn residue to afford the nontoxic pre-xenocoumacin.[310] Bode and co-workers further demonstrated that the transmembrane protease XcnG generates the actual antibiotic species by detachment of the self-resistance *N*-acyl-D-Asn moiety.[310] Meanwhile, several other NRPS pathways have been identified which employ the *N*-acyl-D-Asn motif for self-protection and cellular export, for example, in zwittermicin,[78] paenilamicin,[143] and colibactin biosynthesis.[311,312] Further variations can be found for didemnin (*N*-acyl-D-Gln_n)[313] or even saframycin (acyl-chain).[260] Interestingly, in the fatty acid (FA)-NRP-PK hybrid pre-zeamine, it is the peptide fragment which becomes discarded by an as yet unidentified hydrolase to liberate the antibiotic polyamine PK zeamine.[314]

3.10. Coupling of Ribosomal Peptide Synthesis and NRPS

As outlined in Section 3.7.4, a highly dissociated NRPS assembly line is responsible for the biosynthesis of the tetrapeptide core structure of the pentapeptidyl nucleoside antibiotic pacidamycin.[116] While delineating this network of *trans*-acting NRPS components, Walsh and co-workers identified the tRNA-dependent aminoacyltransferase PacB, which utilizes alanyl-tRNA to couple the activated alanyl residue to the N terminus of the tetrapeptide core structure of pacidamycin.[315] Genetic inactivation of PacB resulted in accumulation of the tetrapeptide in vivo, and enzymatic studies demonstrated that the enzyme exclusively acts on the tetrapeptidyl thioester intermediate anchored to the stand-alone T domain PacH. The observed interaction between PacB and PacH not only illustrates another example of recruitment for tailoring NRPS-bound intermediates, but also showcases an unprecedented intersection between the ribosomal and nonribosomal pathway.

4. Exploitation of NRPS-Based Pathways

An important challenge for current biosynthesis research is to reprogram NRPS systems to design assembly lines that ideally provide any peptide structure of choice. If this approach was successful, it would be a serious alternative to the chemical synthesis of peptides and, apart from possible economic advantages, could contribute to the development of sustainable and environmentally benign biotechnological processes. Over the past few decades there have been a considerable number of attempts to challenge biosynthetic functions or pathways for the synthesis of altered or new NRP structures. These attempts and developments have parallels particularly to PKS systems,[316] and are mainly based on four methods: chemoenzymatics (CHE), precursor-directed biosynthesis (PDB), mutasynthesis (MBS), and combinatorial biosynthesis (CBS; Figure 11).

4.1. Chemoenzymatics

CHE techniques aim at the synthesis of structural variations or newly constructed natural products by using enzyme preparations from a wild-type strain or a heterologous expression strain. The enzymes may consist of single domains, modules, or intact NRPSs as well as tailoring enzymes, and require, apart from their amino acid or peptide substrates, additional cofactors.

4.1.1. NRPSs

In the early years of NRPS research CHE has already been used to accomplish the assembly of new peptides, for example, with enniatin synthetase[193] and cyclosporine synthetase. Supplementation of non-natural amino acids or α-hydroxy acids together with the required cosubstrates afforded a considerable number of new derivatives (Table 4).[317–320] Initially, the producing strains had to deliver sufficient amounts of synthetases, and the use of radiolabeled tracers in many cases was the method of choice to prove the identity of the reaction product. Over the past decades, methodological advances such as recombinant protein technologies and ESI mass spectrometry have significantly facilitated the in vitro reconstitution of pathways and NRP characterization. Recent examples come from CDP synthetases, where CDP analogues have been synthesized using enniatin,[321] PF,[322] and beauvericin synthetase.[323] Of particular note is the pronounced substrate tolerance of their α-hydroxy acid activating A domains. Hence, a whole series of novel α-hydroxy acid analogues of enniatin (aliphatic substrates) and PF1022 (aliphatic and aromatic substrates), for example, containing alkyne, halogen, or thiophene side chains, could be generated (Figure 12). Recently, hybrid PKS-NRPS systems have even been used in the synthesis of preaspyridones.[324]

4.1.2. Te Domain Directed Macrocyclizations

A considerable number of NRPs are macrocycles, and their chemical synthesis can pose an enormous synthetic challenge, since the chemist has to work in high dilution to favor cyclization.[325,326] Accordingly, an appealing idea is to use CHE with Te domains and activated peptide esters to assemble various macrolactams and macrolactones (Figure 12). Initial studies on NRP cyclization employed the Te domain of tyrocidine synthetase (Te_{Tyc}) and peptidyl-SNAC analogues.[207] This approach was expanded to macrocyclize various gramicidin S-like peptides, primarily by Te_{Tyc},[209] and was further applied to a solid-phase peptide library of linear 10-mer tyrocidine precursors bound as ester mimics.[327] Te_{Tyc} has also been used to assemble tyrocidine-based PK-NRP hybrids[328] and for a positional scan that

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

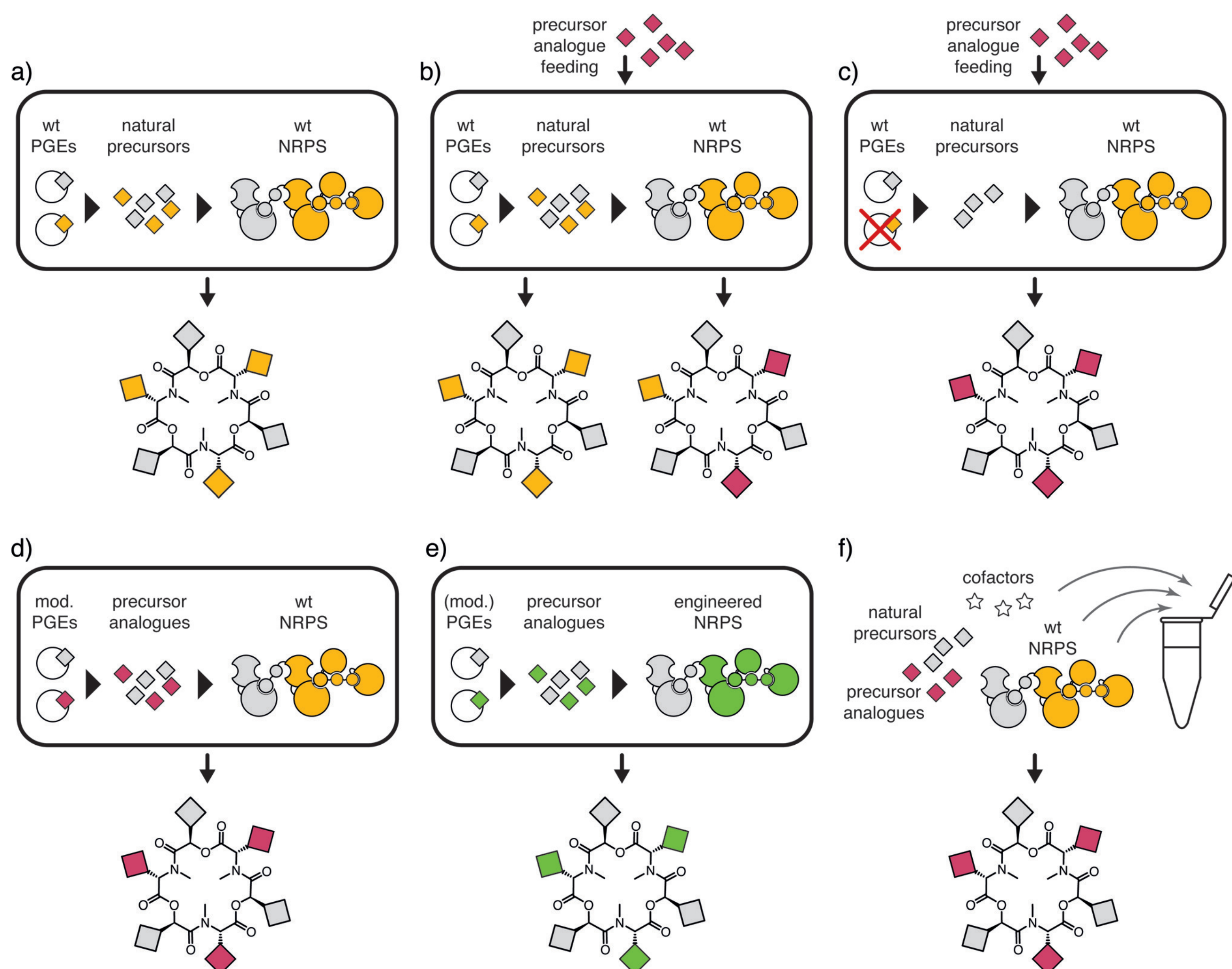


Figure 11. a) Production of NRPs using the wild-type (wt) microorganism, b) precursor-directed biosynthesis (PDB), c) and d) different approaches for mutasynthesis (MBS), where d) exploits relaxed substrate specificity of the NRPS, e) combinatorial biosynthesis (CBS), and f) chemo-enzymatics (CHE). The color coding relates precursor-generating enzymes (PGEs) and their products to the corresponding specificities of NRPSs, with red/orange indicating promiscuity.

tested a large number of synthetic decapeptidyl-SNAC libraries.[329] Other approaches[330] involved peptidyl-SNAC analogues containing non-natural esters, PEG chains, alkynes for click reactions, RGD motifs, and carbohydrate modifications which had been cyclized with Te_{Tyc}[208,331–333] or the Te domain of surfactin synthetase (Te_{Srf}).[334] Te constructs from other NRPS assembly lines were also exploited for the cyclization of peptidyl-SNAC analogues, for example, the use of T-Te_{Grs} didomains with 2–10mer gramicidin-SNAC analogues to install various ring sizes,[206] Te/T-Te_{CDA} from CDA synthetase to assess amino acid substitutions[335] and acetylations,[336] and T-Te didomains from A54145 and daptomycin synthetases to study amino acid substitutions, alteration of ring size, as well as macrolactamization (Thr→Dab).[337] Ala scans of peptide substrates were performed to monitor the effects on cyclization when using, for example, Te of streptogramin B synthetase,[338] T-Te of bacitracin synthetase,[339] and echinomycin synthetase,[340,341] as well as the cyanobacterial (*Nostoc* sp.) Te_{Crp} of the anticancer PK-NRP hybrid cryptophicin.[342] In an original approach, the consecutive reaction of Te_{Crp} and a P450 epoxidase provided a new type of "tandem chemoenzymatics".[343] Although some T-Te didomain constructs were not able to cyclize peptidyl-SNAC substrates, the use of peptide thiophenolesters could enable the cyclization by T-Te didomains of fengycin, mycosubtilin, and syringomycin synthetase,[344] and showed the best results for Te_{CDA}.[335]

In conclusion, Te_{Tyc} is probably the best investigated and most versatile enzyme for macrocyclizations.[208,209,327] However, an inherent drawback of all Te-based macrocyclizations is the requirement of SNAC or alternative thioester derivatives, which commonly involve a laborious synthesis. This obstacle may be alleviated by ester homologues on a solid support;[327] nevertheless, these suffer from scalability of the approach. In addition, studies with thorough reaction analytics have revealed linear peptides as hydrolysis products.[342,343,345] The extent of hydrolysis strongly depends on the nature of the peptide substrate, that is, stereochemical and

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 4: Compilation of CHE, PDB, and MBS studies aiming at structural diversification of NRP-based natural products.

Compound class[a]	CHE	PDB	MBS	Substrates	Products (CHE catalyst)	Ref.
cyclosporine (F)	+	+	−	nonpolar amino acids	analogues (NRPS)	[317–320, 389, 390]
cyclodepsipeptides (F) (e.g. enniatin, beauvericin)	+	+	+	α-hydroxy acids, amino acids	analogues (NRPS)	[288, 321–323, 352, 366, 384]
tyrocidine (B)	+	−	−	peptide S(O)NAC/thiophenol esters; PEGA resin-bound substrates	peptide macrocycles (T and T-Te)	[207–209, 327–329, 331, 333, 334]
gramicidin (B)	+	−	−	peptide SNAC esters	peptide macrocycles (T-Te)	[206]
fengycin (B)	+	−	−	peptide-CoA/T	peptide macrocycles (T-Te)	[391]
CDA, daptomycin (B)	+	−	+	peptide CoA/SNAC/Ppant/thiophenol esters, aza-Trp, F-Trp, CF_3Glu, phenylglycines	peptide macrocycles (T and T-Te), analogues	[335–337, 377, 378, 392]
streptogramin (B)	+	−	−	peptide SNAC esters	peptide macrocycles (Te)	[338]
bacitracin (B)	+	−	−	peptide SNAC esters	peptide macrocycles (T-Te)	[339]
syringomycin, mycosubtilin, fengycin (B)	+	−	−	peptide thiophenol esters	peptide macrocycles (T and T-Te)	[344]
asperlicin (F)	−	+	−	Trp/anthranilate/Leu	analogues	[361]
aureobasidin (F)	−	+	−	(nonpolar) amino acids	peptide analogues	[360]
glycopeptide antibiotics (B) (e.g. balhimycin, vancomycin, teicoplanin)	+	+	+	bromine salts, fluoro-β-hydroxytyrosines, phenylglycines; UDP/UTP carbohydrates	peptide analogues, carbohydrate analogues (NRPS, glycosyltransferases)	[113, 347, 375, 376]
echinomycins (B) (triostin)	+	+	−	peptide SNAC esters; quinoxaline-2-carboxylic acid analogues	analogues (Te); N-terminally substituted echinomycins	[340, 341, 367, 368]
cryptophicin (B)	+	+	−	SNAC esters, Phe analogues, bromine and iodine salts	analogues (NRPS, Te, P450)	[342, 343, 345]
pacidamycin (B)	−	+	+	amino acids (derivatives of Trp, Phe)	analogues	[371, 393, 394]
antimycin (B)	+	−	+	F-Trp, carboxylic acids, SNAC ester derivatives	356 analogues (PKS-NRPS)	[351, 395]
aminocoumarin (B)	+	−	+	pyrrole-carboxylic acid analogues, benzoic acids (ring A)	post-NRPS analogues (acyltransferase, carbamoyltransferase)	[348–350, 385]
roquefortin (F)	−	−	+	diketopiperazines	post-NRPS analogues	[386]
ergot peptides (F)	−	+	+	Phe/Leu analogues, thiaproline	analogues e.g. thiaergosine	[355, 396]
salinosporamide (B)	−	+	+	bromine salts, 5’-fluoro-5’-deoxyadenosine, nonpolar amino acids	bromo/fluoro/cycloalkyl analogues	[380, 381, 388]
diazepinomycin (B)	−	+	−	indoles, Trp, anthranilate	fluorinated derivatives	[369]
iturin (B)	−	+	−	fluoro-Tyr	fluoro-Tyr-iturin	[370]
chaetoglobosin (F)	−	+	−	halogenated Trp	analogues	[365]
nikkomycin (B)	−	+	+	benzoic acids, nucleobases	analogues	[372, 373]
pyochelin (B)	−	−	+	benzoic acids	analogues	[374]

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 4: (Continued)

Compound class[a]	CHE	PDB	MBS	Substrates	Products (CHE catalyst)	Ref.
sibiromycin (B)	–	–	+	anthranilic acids	analogues	[383]
aspyridone (F)	+	–	–	amino acids (Trp, F-Trp, Phe)	analogues	[324]

[a] (B): bacterial producer; (F): fungal producer.

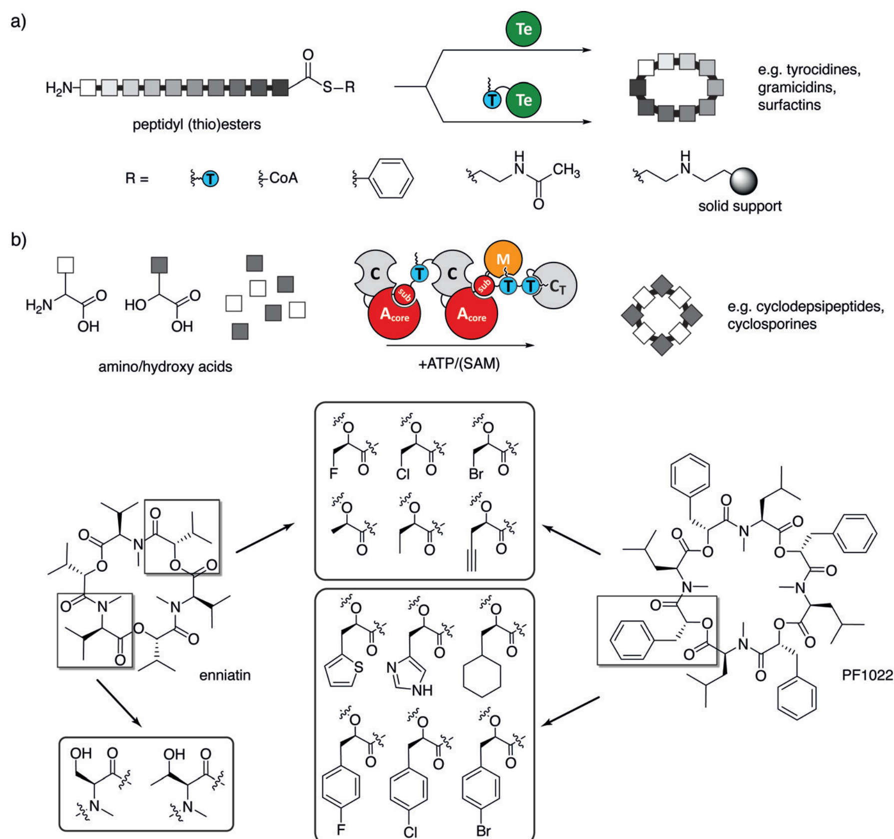


Figure 12. Chemoenzymatic synthesis with catalytic entities from NRPSs: a) Te- and T-Te-catalyzed macrocyclization using various types of activated substrates. b) NRPS-catalyzed synthesis of derivatives of enniatin[321,352] and PF1022[322] with aliphatic and aromatic α-hydroxy acids.

steric features can drive the hydrolysis to become the main reaction.

4.1.3. Tailoring Enzymes and Post-NRPS Assembly

Another application of CHE makes use of tailoring enzymes. This exciting approach promises the generation of otherwise nearly inaccessible peptide structures. As mentioned earlier, P450 monooxygenases would be ideally exploited for the generation of novel glycopeptide backbones, but the strict regulation of ring assembly through NRPS-P450 interactions will prevent this goal from being attained in the near future.[259,269,270] A more accessible approach is the decoration of peptides by post-NRPS enzymes, for example, glycosyltransferases in the case of glycopeptide antibiotics,[346,347] amide synthetases,[348] carbamoyltransferases,[349] and acyltransferases[350] acting on novobiocin precursors as well as promiscuous acyltransferases for modifying antimycin.[351]

4.2. Precursor-Directed Biosynthesis and Mutasynthesis

The precursor-directed biosynthesis (PDB) and mutasynthesis (MBS) approaches are appealing concepts for providing artificial NRP structures. The concepts were introduced by Birch and Rinehart,[353,354] and an early example on NRPSs is that of ergot peptides of the fungus *Claviceps purpurea*.[355]

The idea originates from the awareness that natural products which are difficult to synthesize chemically could be produced by channeling synthetic building blocks into biosynthesis pathways. Although several early reports exist on PDB and MBS, these must be handled with some care if they date back to times before HPLC-MS. Nowadays, an analytical characterization by HPLC-ESI-MS and increasingly by NMR spectroscopy sets the standard for product characterization.

From previous discussions of PDB and MBS,[356,357] it is important to note that PDB and MBS are distinct yet related techniques (Figure 11): in PDB, a synthetic building block is supplied to wild-type strains with an intact biosynthesis assembly line, whereas MBS makes use of mutant strains ("auxotrophs") with impaired building block biosynthesis. Other approaches such as biotransformations have the character of structural tailoring rather than de novo assembly and most importantly make use of nonrelated microorganisms to perform structural modifications on a compound of interest.

Inherent requirements for building blocks in PDB and MBS are: 1) uptake into the cell, 2) resistance against efflux, 3) metabolic stability against degradation or modification, and 4) acceptance by the biosynthetic machinery.

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

4.2.1. Precursor-Directed Biosynthesis

Probably the simplest PDB experiment is its application to halogenated compounds through varying the media. PDB has already been applied to rather complex molecules such as the glycopeptide balhimycin[113] to generate bromobalhimycin (Scheme 10). The incorporation of bromine depends on the specificity of the halogenase,[358,359] whereas fluoride and iodide salts can cause toxic effects.

More complex examples of PDB involve seven fungal systems (Table 4): the synthetases of ergot peptides (*Claviceps purpurea*),[355] aureobasidin,[360] asperlicin,[361] the immunosuppressant cyclosporine (*Tolypocladium inflatum*; Scheme 10),[320,362–364] chaetoglobosin (*Chaetomium sp.*),[365] as well as enniatin (*Fusarium sp.*)[352] and beauvericin (*Beauveriva bassiana*).[366] Among the bacterial systems, new analogues of echinomycin (*Streptomyces echinatus*),[367,368] diazepinomicin (*Micromonospora* sp.),[369] iturin (*Bacillus* sp.),[370] and pacidamycin have been described.[371] In a remarkable example, chemoauxotrophic growth of cyanobacteria facilitated the incorporation of various Phe analogues into the cryptophicin scaffold as well as the bromination and iodination of Tyr (Scheme 10), with the latter case indicating an unusual iodinase activity of the halogenase CrpH.[343]

Commonly substituted analogues of aliphatic and aromatic (preferably Phe, Tyr, Trp) amino acids are used in PDB experiments. For reasons of low steric demand, H/F exchanges have been widely applied. Generally, bacterial systems seem more restrictive towards noncognate substrates. However, direct competition between supplemented and endogenous substrates generally limits the applicability of PDB. This also concerns practical aspects such as purification of desired compounds from mixtures, which makes MBS appear more favorable.

4.2.2. Mutasynthesis

Early examples of MBS are the feeding of nucleobase analogues[372] and benzoic acids[373] to mutants of the nikkomycin producer *Streptomyces tendae* TÜ901. The feeding of benzoic acids also afforded derivatives of the siderophore pyochelin.[374] Probably the most complex structure that MBS has been applied to is the vancomycin-type glycopeptide balhimycin (*A. balhimycina*) by feeding β-OH-Tyr analogues to a β-OH-Tyr-deficient mutant strain (Scheme 10).[375] This study was followed by supplementing a Dpg-deficient mutant strain with phenylacetic acids, mandelic acids, and phenylglycines, with the phenolic hydroxy groups replaced by H or OCH_3.[376] Among the lipocyclodepsipeptides, mutants of the CDA biosynthesis (*S. coelicolor*) have been the subject of extensive MBS studies with auxotrophs of the amino acids Hpg, MeGlu, and Trp. Feeding of 4-F/Cl/OMe-mandelate, glyoxylates, and phenylglycines resulted in only the F-derivatives being accepted as substrates.[377] MeGlu was replaced with CF_3Glu and EtGlu together with natural Glu analogues,[378] and both 5-OH-Trp and 5-F-Trp could replace Trp.[379] Various cycloalkyl amino acid analogues were generated[380,381] as well as fluorosalinosporamide by feeding 5′-fluoro-5′deoxyadenosine (5′-FDA) to mutants of the salinosporamide producer *Salinispora tropica* (Scheme 10).[382] Finally, a recent example of MBS with an anthranilic acid yielded a new deoxysibiromycin with diminished cardiotoxicity but increased antitumor activity.[383]

A knock-out of the ketoreductase Kvir, which synthesizes α-hydroxy acids in the fungal beauvericin producer *Beauveria bassiana*, facilitated the incorporation of non-natural α-hydroxy acids together with fluorinated Phe to give 14 “scrambled” beauvericins.[384] Furthermore, substrate acceptance and production rates of this mutant were compared to a heterologous host (*E. coli*) with beauvericin synthetase.[323]

There are also examples of MBS which involve post-NRPS enzymes, for example, amide synthetase mediated coupling of supplemented benzoic acids to yield 32 new antibacterial aminocoumarins.[348,385] In a subsequent two-stage MBS approach, the assembly of 25 new aminocoumarins has been achieved using two *S. coelicolor* mutant strains in two sequential cultivations.[385] Another post-NRPS example comes from the fungus *Penicillium chrysogenum* where synthetic diketopiperazines have been modified by a dehydrogenase and a dimethylallyl transferase.[386]

A very recent approach termed multiplexing was performed with the bacterial PK-NRP hybrid antimycin to generate 356 new antimycins.[351] In this study, fluoroanthranilate and various malonates were incorporated by MBS followed by a chemoenzymatic diversification by an acyltransferase using various CoA-ester substrates.

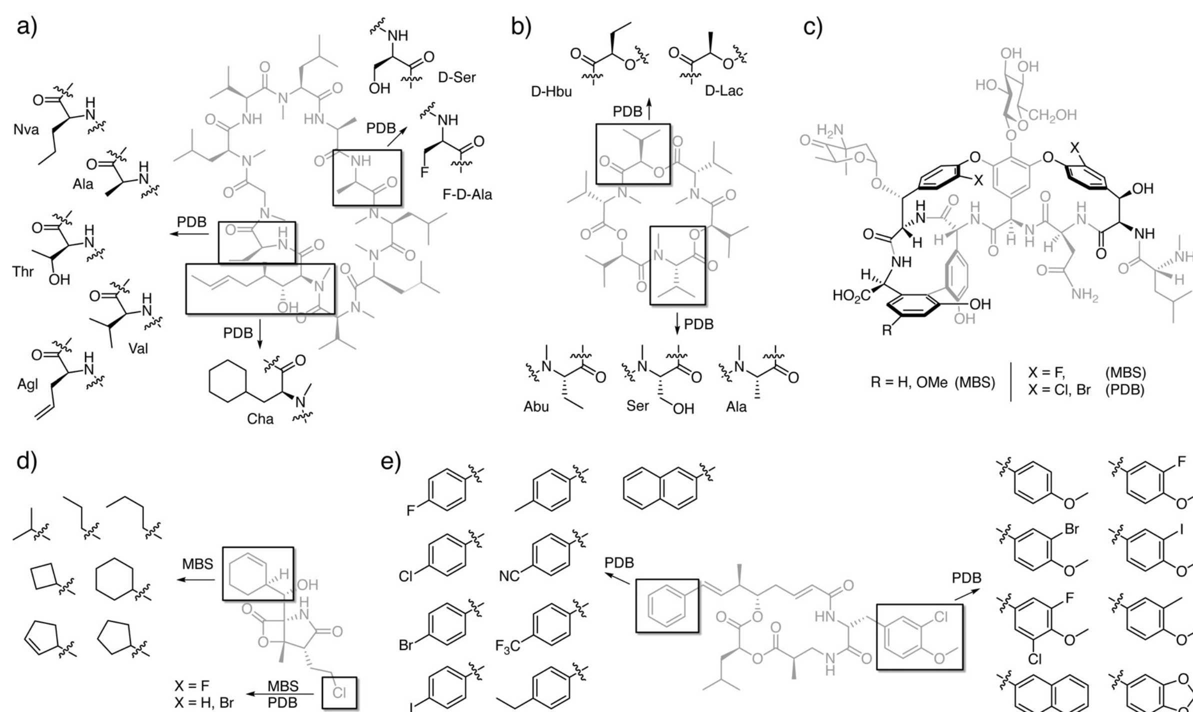
Similar to PDB, the NRPS assembly line only barely tolerates extensive steric and electronic alterations of the mutasynthons. Examples from bacteria dominate those from fungi for reasons of genetic accessibility, which is about to change due to methodological advances in this field. An appealing alternative to MBS is the heterologous expression of only core NRPSs, which function upon feeding of appropriate building blocks (“dependent host”). A related example represents the heterologous expression of three *p*-aminophenylpyruvate biosynthesis genes from the chloramphenicol biosynthesis (*Streptomyces venezuelae*) in a chorismate-deficient mutant of the fungal PF1022 producer *Rosellinia* sp.[387] The thus-generated mutant strain produced nitro-PF1022 derivatives, although only in μg amounts, which are precursors to the anthelmintic drug emodepside.

4.3. Design of NRPSs and Combinatorial Biosynthesis

The modular principle of NRPS assembly lines inspires reprogramming their domains and modules to generate new peptide structures. Such an approach, however, must also be seen in comparison with ribosomal systems,[397] where site-directed mutagenesis as well as more recently supplementation-based incorporation[398] and stop-codon suppression[399,400] promise a seemingly straightforward realization of structural modifications.

NRPS engineering can be performed at different hierarchical levels of the assembly line (Table 5): the exchange of complete NRPS genes leaves the internal structure of the assembly line intact, whereas replacements and fusions of

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 10. Examples of PDB and MBS applications: a) cyclosporine,[363,364] b) enniatin, c) balhimycin,[375,376] d) salinosporamide,[381,382,388] and e) PK-NRP hybrid cryptophicin.[343]

modules or domains may interfere with the sequential and spatial context of the synthetase. Therefore, site-directed mutagenesis constitutes a minimal intervention in the system and, thus, a serious alternative to the above approaches (Figure 13).

Initial studies to realize a combinatorial concept were contributed by the Marahiel group, for example by using surfactin synthetase from *Bacillus subtilis*.[401] Its module 7 (activating Leu) was replaced with modules of gramicidin synthetase (activating Phe, Orn, Val) and the fungal ACV synthetase (activating Cys, Val). In a similar approach, module 2 was replaced with an Orn-activating module.[402] On the biochemical level, that is, the in vitro reconstitution of enzymes, a considerable number of experiments have been performed to fuse modules or domains to generate dipeptides[403–405] and tripeptides,[406] mostly using *E. coli* as the expression system (Table 5).

Conceptually, an important in vivo proof of module-based approaches has been provided by 1) module deletion, 2) module exchange, and 3) module insertion (Figure 13). In the module-deletion experiment, Marahiel and co-workers excised a complete module of surfactin synthetase.[407] The generated Bacillus strain carrying a ΔLeu^2-surfactin synthetase produced Δ2-surfactin. Initial experiments on module exchange were also provided by the Marahiel group.[401] Based on this study, Yakimov et al.[408] replaced module 1 (Glu^1) with a Gln-activating module to produce Gln^1-surfactin. Süssmuth and co-workers performed the insertion of a module into balhimycin synthetase.[409] This Hpg-activating module was constructed from C-A and T-E didomains of two neighboring Hpg-activating modules and only had a single non-natural junction between the A and T domains. The mutant strain of *Amycolatopsis balhimycina* indeed synthesized octapeptides; however, P450 monooxygenase tailoring was disturbed and yielded mainly a linear rather than the expected multicyclic product (Figure 13).

The most extensive studies were probably performed by the Baltz group on the antibacterial drug daptomycin.[410] In initial experiments with the daptomycin (Dpt) synthetase from *Streptomyces roseosporus*, they replaced modules of the type C-A-T (D-Ala^8→D-Ser/D-Lys; D-Ser^{11}→D-Ala), leaving the cognate E domains untouched.[411] Replacement of modules of the type C-A-T-E resulted in lower production yields. Various fusion sites in the linker region of DptD (C-A-T or C-A-T-Te topologies) showed great flexibility to C-terminally install Trp (CDA) and Ile (A54145) instead of Kyn.[412] Interestingly, only one of two fusion constructs (C-A_{Asn}-/T-Te versus C-A_{Asn}-T-/Te) was able to incorporate Asn (LptC from A54145 synthetase) in the terminal position of daptomycin.[412] Even a four-module exchange coding for D-Lys^8-Asp^9-Gly^{10}-D-Asn^{11} instead of D-Ala^8-Asp^9-Gly^{10}-D-Ser^{11} (LptC) was successful, but with a significant drop in the production yield (ca. 1 mg L^{-1}). The group reported 30 hybrid NRPS-containing strains, of which 21 strains produced lipopeptides in yields of 1–100 mg L^{-1}. Module exchanges in the related A54145 synthetase generated 14 analogues bearing amino acid variations in positions 2, 3, 8, 11, and 12.[413] The exchange of a complete NRPS (DptD: mGlu^{12}-Kyn^{13}) exploited the high homology, with genes from CDA and A54145 synthetase installing Trp^{13} and Val^{13}/Ile^{13}, respectively.[414,415] The daptomycin analogues were produced in appreciable amounts (> 20 mg L^{-1}), but multiple module

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Table 5: Compilation of combinatorial biosynthesis studies performed in vitro and in vivo with NRPSs.[a]

	Experiment	In vitro	in vivo	Engineered peptide (synthetase, host)	Ref.
NRPS	exchange	−	+	daptomycin analogues; pos. 2–8; 12/13 (daptomycin, CDA, A54154; *Streptomyces*)	[410, 413, 415]
	exchange	−	+	echinomycin analogue (echinomycin; *E. coli*)	[341, 433]
module	exchange (/A-T/)	+	+	surfactin analogues; pos. 2/7 (surfactin; *E. coli* + *Bacillus*)	[401, 402]
	exchange (/C-A-T/)	−	+	$Gln^1 \rightarrow Glu^1$ lipocyclopeptide (lichenysin; *Bacillus*)	[408]
	module fusion (/C-A-T-/Te)	+	−	tripeptides (tyrocidine; *E. coli*)	[406]
	module fusion (A-T-/C-/A-/T-Te)	+	−	Asp-Phe dipeptide (tyrocidin/surfactin; *E. coli*)	[405]
	exchange (/C-A-T-/E/)	−	+	daptomycins; pos. 2/8/11 (daptomycin, A54154; *Streptomyces*)	[410, 413]
	exchange (/C-A-/T-/Te)	−	+	daptomycins; pos. 13 (daptomycin; *Streptomyces*)	[412]
	deletion (/C-A-T/)	−	+	surfactin 7mer→6mer (surfactin; *Bacillus*)	[407]
	insertion (/C-A-T-E/)	−	+	peptide 7mer→8mer; loss of P450 tailoring functions (balhimycin; *Amycolatopsis*)	[409]
	recombination (C-A-T/)	−	+	chimera of enniatin, beauvericin, and PF1022 (fungal synthetases; *E. coli*, *Aspergillus*)	[418]
	recombination (C-A-T/)	−	+	chimera of bassianolide and beauvericin (fungal synthetases; *Saccharomyces*)	[434]
domain	fusion (A/)	+	−	dipeptides (tyrocidine, bacitracin; *E. coli*)	[403]
	shift (/Te/)	−	+	peptide truncations (surfactin; *Bacillus*)	[435]
	exchange (/A/ and /C-A/)	−	+	pyoverdin and analogues (pyoverdin; *Pseudomonas*)	[416, 436]
	exchange (/C-A/ and /T-C-A/)	−	+	pyoverdin (pyoverdin; *Pseudomonas*)	[417]
	exchange (COM domains)	+	+	di-, tripeptide, diketopiperazines (tyrocidin, gramicidin, surfactin; *E. coli*, *Bacillus*)	[195, 197, 419]
	fusion (A-T-/E; A-/T-E)	+	−	epimerization of amino acids (tyrocidin, bacitracin; *E. coli*)	[411]
amino acid(s)	A domain (site-directed mutagenesis NRPS code)	+	+	$Asp^5 \rightarrow Asn^5$ (surfactin; *E. coli*, *Bacillus*),	[420]
		−	+	$Asp^7 \rightarrow Asn^7$ (CDA; *Streptomyces*)	[421]
		−	+	$Glu^{10} \rightarrow Gln^{10}/(MeGln^{10})$ (CDA; *Streptomyces*)	[422]
	A domain modeling	+	+	diketopiperazine (gramicidin; *E. coli*)	[426]
	A domain (site-directed mutagenesis NRPS code)	+	−	modulation of A domain specificity (tyrocidin; *E. coli*)	[423]
	A domain; random PCR	−	+	andrimids (andrimid; *E. coli*)	[427]
	A domain; focused PCR	−	+	andrimid analogues (andrimid; *Pantoea*)	[428]
	exchange (A→A-M)	+	−	methylated acyldipeptides (actinomycin; *Streptomyces*)	[404]
	M domain (site-directed mutagenesis)	−	+	desmethyl analogue (echinomycin, *E. coli*)	[437]

[a] "//" denotes cleavage sites for module/domain exchanges, fusions and deletions.

substitutions again led to a significant drop in yields ($<1\ \mathrm{mg\,L^{-1}}$). Further domain substitutions have been performed with pyoverdine synthetase (*P. aeruginosa*) by employing A domain, C-A didomain[416,417] as well as T-C-A tridomain substitutions.[417] Recently, two iterative hybrid NRPSs were designed through a swap between the highly homologous fungal NRPSs of beauvericin, enniatin and PF1022, which were heterologously expressed in *E. coli* and *A. niger*.[418] Although the C-A-T fusion rule was applied, production yields of the six new-to-nature NRPs dropped by several orders of magnitude.

Another interesting concept deals with the exchange of COM domains (see Section 3.4) to generate dipeptide and tripeptide combinations in vitro[195,419] as well as tetrapeptides in a *Bacillus* producer.[197]

With the knowledge of the nonribosomal code of A domains,[136,137] it appeared conceivable that single subsitutions of coding residues could suffice to reprogram the specificity of the NRPSs. This was shown by surfactin synthetase from *B. subtilis* producing [Asn5]-surfactin, as well as considerable amounts of wild-type [Asp5]-surfactin.[420] A similar approach was pursued with CDA synthetase, for which two point mutations afforded [Asn7]-CDA[421] and [Gln(MeGln)10]-CDA, respectively, both accompanied by linear peptides.[422]

More recently, several in vitro approaches dealt with the reprogramming of A-domain specificities. Directed evolution by saturation mutagenesis of the specificity-conferring residues of an A domain of tyrocidine synthetase shifted the activation from Phe to Ala.[423] Likewise, transplanting segments of the A domains (hormaomycin synthetase) altered the specificity and adenylation activity of the thus-generated domain.[424] In the same area falls the subdomain-swapping strategy by Hilvert and co-workers for constructing in vitro a *cyclo*(D-Val-Pro) synthetase from a *cyclo*(D-Phe-Pro) synthetase.[425] The same group rationally designed a W239S mutation in the specificity-conferring pocket of the Phe-activating A domains of gramicidin S/tyrocidine synthetase,

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

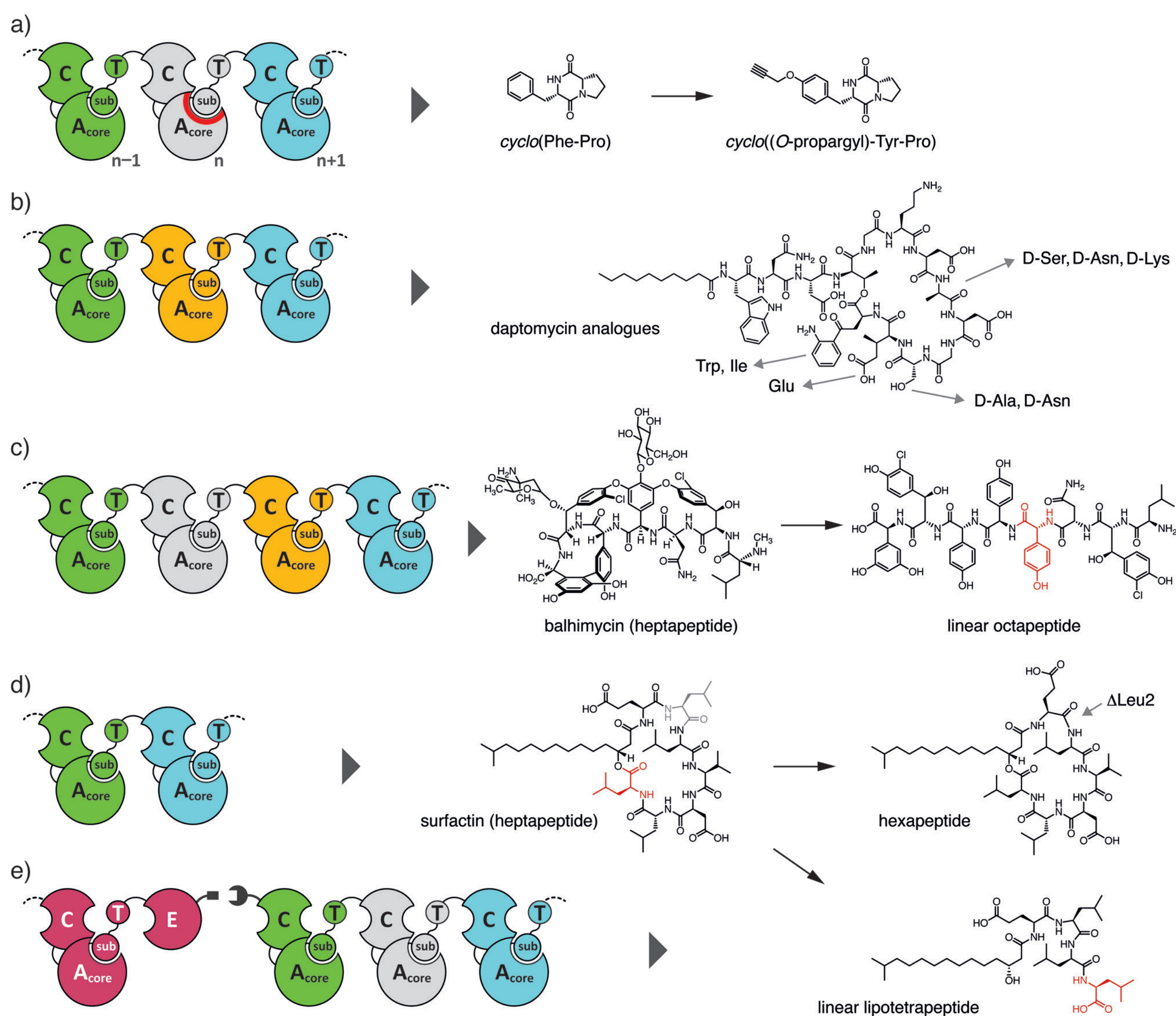


Figure 13. Combinatorial biosynthesis performed in vivo with NRPSs and substructure elements. a) Site-directed mutagenesis in the A-domain substrate pocket, b) module exchange, c) module insertion, d) module excision, and e) COM domain-mediated exchange of NRPSs.

which led to the biosynthesis of Phe(O-*p*-propargyl)-containing diketopiperazine in *E. coli*.[426] Most importantly, in this case, experiments were also verified under in vivo conditions (*E. coli*), with production yields (10 mgL^{-1}) comparable to the wild-type strain (Figure 13).

The functionality of NRPS variants under in vivo conditions has been further shown by two studies using directed evolution of a Val-activating A domain (AdmK from the PKS-NRPS andrimid gene cluster; *Pantoea agglomerans*) as a scaffold. The corresponding A domain from an andrimid-producing *E. coli* was deleted and complemented with an A domain activating 2-aminobutyrate (Abu). After several rounds of mutagenesis and selection from a random DNA library (> 5600 clones), a strain was selected that produced Abu-andrimid and andrimid in a 1:1 ratio.[427] Likewise, *admK* in the natural producer was engineered in three positions of the nonribosomal code, and four new andrimid variants (Ala, Phe, Leu, and Ile) were obtained from a library of > 14000 members.[428]

4.4. Perspectives for Peptide Design

Basic demands on all CHE approaches are robustness and versatility of the biocatalyst combined with high production yields. Although Te-based macrocyclizations could in principal provide very interesting products, they have been, from a contemporary perspective, considered too optimistic. Likewise, the substrate tolerance of NRPSs, with the exception of some fungal systems, for example, cyclosporine and CDPs, still needs further engineering to more readily generate larger compound libraries. In any case, preparative applications have to be considered with some care because of demands on cofactors and substrates, although the currently attainable scale suffices for screening purposes. Hence, although CHE approaches might be useful, a sustainable biotechnological production will sooner or later require transfer into cell-based settings.

PDB and MBS are both valuable in vivo techniques, but strongly depend on the substrate tolerance of the biosynthetic

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

machinery. They can achieve preparative scales in industrial processes as shown by the example of the PK doramectin,[429] a marketed anthelmintic. Hence, decoupling the specificity limits of the biosynthetic machinery is still one of the major tasks in NRPS research.

Although in vitro combinatorial approaches offer verification of basic concepts, experiments often lack ultimate proof by cell-based approaches. The exchange of modules as well as the directed alteration of A-domain specificities in in vivo experiments have been worked out in more detail. Even if successful, such attempts are still accompanied by the production of wild-type compounds,[420,427] truncations through premature release from the NRPS,[401,407,409,421,422] and often a significant drop in production yields.[409,410,418] These indicators demonstrate that interference with the assembly lines is a delicate enterprise, which requires a deeper understanding through future investigations.

Further aspects for combinatorial biosynthesis are the use of heterologous production hosts to provide an established genetic toolbox and optimized methods for mutagenesis and gene expression. As a rule of thumb, the heterologous system should come from the same genus as the designated biosynthesis genes. Remarkably, *E. coli* appears as a viable solution, at least for bacterial NRPSs from Actinobacteria: the multicistronic expression (T7 promotor, a ribosome binding site, and a terminator) of genes of the triostin gene cluster from *Streptomyces lasaliensis* enabled the biosynthesis of triostin in *E. coli*.[62,430] Yeasts or *Aspergillus* strains may constitute appropriate systems for fungal NRPSs.[431,432]

From an economics point of view, the synthesis of linear peptides by NRPSs is, in most cases, not competitive with chemical synthesis. The situation is different for cyclic peptides, however, which are synthetically more difficult to access, and is considerably more in favor of NRPSs and may become more significant in the future. Of particular interest for future engineering purposes are peptides which contain unusual building blocks that are difficult to couple (e.g. racemization-prone amino acids, α-hydroxy acids) and those which underlie tailoring reactions, for example, by oxygenases. For exploitation in biotechnological processes, production yields have to be superior to those of chemical synthesis.

5. Outlook

As outlined in this Review, NRPS assembly lines contribute important drugs for human use and are invaluable sources for future drugs. The expectations for developments are founded on the discovery of new bioactive NRPs, a deeper understanding of biosynthetic principles, and rational pathway engineering as the main pillars. This will firstly require unlocking new sources of biodiversity in terms of microorganisms as well as technical advancements in discovery methods. Progress may come from metagenomics approaches or accessibility to difficult-to-cultivate microorganisms, as recently exemplified for the antibacterial NRP teixobactin,[438,439] for example, using the iChip technology.[440] The enormous developments in DNA sequencing techniques together with bioinformatic tools and a continuously improved predictive power will aid in the discovery process and the ready exploration of entire genomes, which at some point may replace classical screening approaches.

In the past few years, we have gained a structural understanding of NRPS assembly, in particular by recent X-ray structures of intact modules. With the advent of cryo-electron microscopy,[441] structural insights into larger components or even entire NRPS machineries are tangible for further refinement of our picture of these giant enzyme architectures. Studies on NRPSs and their substrate flux will also benefit from method developments in solution- and solid-state NMR spectroscopy,[442–444] as these techniques hold the key to identify the determinants of transient domain–domain interactions within these highly dynamic complexes. Knowledge derived from such integrative structural biology approaches will have an impact on the directed engineering of NRPS pathways, for example, by rewriting the nonribosomal code or by modulating the intermolecular communication between different NRPS components. This may allow us to derivatize known NRPs or to assemble designed peptide structures, which will facilitate a sustainable peptide supply competitive with chemical synthesis approaches. Finally, it remains to be seen whether derivatives of recently discovered NRPs such as albicidin, griselimycin, or teixobactin will make it into clinical applications.

Acknowledgements

This work was supported by the Cluster of Excellence "unifying concepts in catalysis" (UniCat) granted by the Deutsche Forschungsgemeinschaft (DFG). We thank Dr. Ullrich Keller for helpful discussions, as well as Charlotte Steiniger and Tam Dang for their strong support in NRPS model development and the preparation of schemes.

How to cite: *Angew. Chem. Int. Ed.* **2017**, *56*, 3770–3821
Angew. Chem. **2017**, *129*, 3824–3878

[1] A. Hershko, A. Ciechanover, *Annu. Rev. Biochem.* **1998**, *67*, 425–479.
[2] R. J. Deshaies, C. A. P. Joazeiro, *Annu. Rev. Biochem.* **2009**, *78*, 399–434.
[3] A. L. Lovering, S. S. Safadi, N. C. J. Strynadka, *Annu. Rev. Biochem.* **2012**, *81*, 451–478.
[4] W. Vollmer, U. Bertsche, *Biochim. Biophys. Acta Biomembr.* **2008**, *1778*, 1714–1734.
[5] P. Belin, M. Moutiez, S. Lautru, J. Seguin, J.-L. Pernodet, M. Gondry, *Nat. Prod. Rep.* **2012**, *29*, 961–979.
[6] M. Gondry et al., *Nat. Chem. Biol.* **2009**, *5*, 414–420.
[7] B. Mach, E. Reich, E. L. Tatum, *Proc. Natl. Acad. Sci. USA* **1963**, *50*, 175–181.
[8] T. L. Berg, L. O. Froholm, S. G. Laland, *Biochem. J.* **1965**, *96*, 43–52.
[9] R. Roskoski, W. Gevers, H. Kleinkauf, F. Lipmann, *Biochemistry* **1970**, *9*, 4839–4845.
[10] R. Roskoski, H. Kleinkauf, W. Gevers, F. Lipmann, *Biochemistry* **1970**, *9*, 4846–4851.
[11] W. Gevers, H. Kleinkauf, F. Lipmann, *Proc. Natl. Acad. Sci. USA* **1969**, *63*, 1335–1342.
[12] E. Stoll, Ø. Frøyshov, H. Holm, T. L. Zimmer, S. G. Laland, *FEBS Lett.* **1970**, *11*, 348–352.

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[13] W. Gevers, H. Kleinkauf, F. Lipmann, *Proc. Natl. Acad. Sci. USA* **1968**, *60*, 269–276.
[14] H. Kleinkauf, W. Gevers, F. Lipmann, *Proc. Natl. Acad. Sci. USA* **1969**, *62*, 226–233.
[15] N. Schnell, K. D. Entian, U. Schneider, F. Götz, H. Zähner, R. Kellner, G. Jung, *Nature* **1988**, *333*, 276–278.
[16] Q. Shou, L. Feng, Y. Long, J. Han, J. K. Nunnery, D. H. Powell, R. A. Butcher, *Nat. Chem. Biol.* **2016**, *12*, 770–772.
[17] A. Richardt, T. Kemme, S. Wagner, D. Schwarzer, M. A. Marahiel, B. T. Hovemann, *J. Biol. Chem.* **2003**, *278*, 41160–41166.
[18] S. Hartwig, C. Dovengerds, C. Herrmann, B. T. Hovemann, *FEBS J.* **2014**, *281*, 5147–5158.
[19] R. H. Baltz, *J. Ind. Microbiol. Biotechnol.* **2011**, *38*, 1747–1760.
[20] A. Hornung et al., *ChemBioChem* **2007**, *8*, 757–766.
[21] T. Weber et al., *Nucleic Acids Res.* **2015**, *43*, W237–W243.
[22] M. H. Medema, K. Blin, P. Cimermancic, V. de Jager, P. Zakrzewski, M. A. Fischbach, T. Weber, E. Takano, R. Breitling, *Nucleic Acids Res.* **2011**, *39*, W339–W346.
[23] M. A. Skinnider, C. A. Dejong, P. N. Rees, C. W. Johnston, H. Li, A. L. H. Webster, M. A. Wyatt, N. A. Magarvey, *Nucleic Acids Res.* **2015**, *43*, 9645–9662.
[24] N. Khaldi, F. T. Seifuddin, G. Turner, D. Haft, W. C. Nierman, K. H. Wolfe, N. D. Fedorova, *Fungal Genet. Biol.* **2010**, *47*, 736–741.
[25] N. Ziemert, M. Alanjary, T. Weber, *Nat. Prod. Rep.* **2016**, *33*, 988–1005.
[26] J. Davies, *J. Ind. Microbiol. Biotechnol.* **2006**, *33*, 496–499.
[27] S. Cociancich et al., *Nat. Chem. Biol.* **2015**, *11*, 195–197.
[28] A. Zipperer et al., *Nature* **2016**, *535*, 511–516.
[29] S. Müller, E. Garcia-Gonzalez, A. Mainz, G. Hertlein, N. C. Heid, E. Mösker, H. van den Elst, H. S. Overkleeft, E. Genersch, R. D. Süssmuth, *Angew. Chem. Int. Ed.* **2014**, *53*, 10821–10825; *Angew. Chem.* **2014**, *126*, 10998–11002.
[30] B. Bister, D. Bischoff, G. J. Nicholson, M. Valdebenito, K. Schneider, G. Winkelmann, K. Hantke, R. D. Süssmuth, *Biometals* **2004**, *17*, 471–481.
[31] M. Miethke, P. Bisseret, C. L. Beckering, D. Vignard, J. Eustache, M. A. Marahiel, *FEBS J.* **2006**, *273*, 409–419.
[32] J. A. Ferreras, J.-S. Ryu, F. Di Lello, D. S. Tan, L. E. N. Quadri, *Nat. Chem. Biol.* **2005**, *1*, 29–32.
[33] W. M. Yuan, G. D. Gentil, A. D. Budde, S. A. Leong, *J. Bacteriol.* **2001**, *183*, 4040–4051.
[34] S. O. Duke, F. E. Dayan, *Toxins* **2011**, *3*, 1038–1064.
[35] M. Groll, B. Schellenberg, A. S. Bachmann, C. R. Archer, R. Huber, T. K. Powell, S. Lindow, M. Kaiser, R. Dudler, *Nature* **2008**, *452*, 755–758.
[36] T. G. Kinscherf, D. K. Willis, *J. Antibiot.* **2005**, *58*, 817–821.
[37] M. D. Thomas, P. J. Langston-Unkefer, T. F. Uchytil, R. D. Durbin, *Plant Physiol.* **1983**, *71*, 912–915.
[38] P. M. Joyner, J. Liu, Z. Zhang, J. Merritt, F. Qi, R. H. Cichewicz, *Org. Biomol. Chem.* **2010**, *8*, 5486–5489.
[39] J.-P. Nougayrède, S. Homburg, F. Taieb, M. Boury, E. Brzuszkiewicz, G. Gottschalk, C. Buchrieser, J. Hacker, U. Dobrindt, E. Oswald, *Science* **2006**, *313*, 848–851.
[40] G. Schneditz et al., *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 13181–13186.
[41] P. Olliaro, J. Seiler, A. Kuesel, J. Horton, J. N. Clark, R. Don, J. Keiser, *PLoS Neglected Trop. Dis.* **2011**, *5*, e1138. See also http://www.dndi.org/diseases-projects/filarial-diseases/filaria-dndi-strategy/.
[42] T. C. Paulitz, R. R. Bélanger, *Annu. Rev. Phytopathol.* **2001**, *39*, 103–133.
[43] D. Haas, G. Défago, *Nat. Rev. Microbiol.* **2005**, *3*, 307–319.
[44] E. A. Emmert, J. Handelsman, *FEMS Microbiol. Lett.* **1999**, *171*, 1–9.
[45] X. H. Chen et al., *Nat. Biotechnol.* **2007**, *25*, 1007–1014.
[46] S. P. Chowdhury, A. Hartmann, X. Gao, R. Borriss, *Front. Microbiol.* **2015**, *6*, 780.
[47] D. L. Caly, R. D′Inca, E. Auclair, D. Drider, *Front. Microbiol.* **2015**, *6*, 1336.
[48] F. Chaucheyras-Durand, H. Durand, *Benefic. Microbes* **2010**, *1*, 3–9.
[49] V. Pfeifer, G. J. Nicholson, J. Ries, J. Recktenwald, A. B. Schefer, R. M. Shawky, J. Schröder, W. Wohlleben, S. Pelzer, *J. Biol. Chem.* **2001**, *276*, 38370–38377.
[50] W. Wohlleben, E. Stegmann, R. D. Süssmuth, *Methods Enzymol.* **2009**, *458*, 459–486.
[51] O. Puk, D. Bischoff, C. Kittel, S. Pelzer, S. Weist, E. Stegmann, R. D. Süssmuth, W. Wohlleben, *J. Bacteriol.* **2004**, *186*, 6093–6100.
[52] O. Puk, P. Huber, D. Bischoff, J. Recktenwald, G. Jung, R. D. Süssmuth, K. H. van Pée, W. Wohlleben, S. Pelzer, *Chem. Biol.* **2002**, *9*, 225–235.
[53] S. Pohle, C. Appelt, M. Roux, H.-P. Fiedler, R. D. Süssmuth, *J. Am. Chem. Soc.* **2011**, *133*, 6194–6205.
[54] V. Schubert, F. Di Meo, P.-L. Saaidi, S. Bartoschek, H.-P. Fiedler, P. Trouillas, R. D. Süssmuth, *Chem. Eur. J.* **2014**, *20*, 4948–4955.
[55] C. Zhang, L. Kong, Q. Liu, X. Lei, T. Zhu, J. Yin, B. Lin, Z. Deng, D. You, *PLoS One* **2013**, *8*, e56772.
[56] H. Chen, B. K. Hubbard, S. E. O'Connor, C. T. Walsh, *Chem. Biol.* **2002**, *9*, 103–112.
[57] M. Strieker, E. M. Nolan, C. T. Walsh, M. A. Marahiel, *J. Am. Chem. Soc.* **2009**, *131*, 13523–13530.
[58] J. He, N. Magarvey, M. Piraee, L. C. Vining, *Microbiology* **2001**, *147*, 2817–2829.
[59] T. M. Makris, M. Chakrabarti, E. Münck, J. D. Lipscomb, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 15391–15396.
[60] G. M. Singh, F. H. Vaillancourt, J. Yin, C. T. Walsh, *Chem. Biol.* **2007**, *14*, 31–40.
[61] H. Chen, C. T. Walsh, *Chem. Biol.* **2001**, *8*, 301–312.
[62] K. Watanabe, K. Hotta, A. P. Praseuth, K. Koketsu, A. Migita, C. N. Boddy, C. C. C. Wang, H. Oguri, H. Oikawa, *Nat. Chem. Biol.* **2006**, *2*, 423–428.
[63] J. M. Neary, A. Powell, L. Gordon, C. Milne, F. Flett, B. Wilkinson, C. P. Smith, J. Micklefield, *Microbiology* **2007**, *153*, 768–776.
[64] M. Strieker, F. Kopp, C. Mahlert, L.-O. Essen, M. A. Marahiel, *ACS Chem. Biol.* **2007**, *2*, 187–196.
[65] S. Houwaart, L. Youssar, W. Hüttel, *ChemBioChem* **2014**, *15*, 2365–2369.
[66] M. Offenzeller, G. Santer, K. Totschnig, Z. Su, H. Moser, R. Traber, E. Schneider-Scherzer, *Biochemistry* **1996**, *35*, 8401–8412.
[67] S. Uhlmann, R. D. Süssmuth, M. J. Cryle, *ACS Chem. Biol.* **2013**, *8*, 2586–2596.
[68] K. Haslinger, C. Brieke, S. Uhlmann, L. Sieverling, R. D. Süssmuth, M. J. Cryle, *Angew. Chem. Int. Ed.* **2014**, *53*, 8518–8522; *Angew. Chem.* **2014**, *126*, 8658–8662.
[69] W. Jiang, R. A. Cacho, G. Chiou, N. K. Garg, Y. Tang, C. T. Walsh, *J. Am. Chem. Soc.* **2013**, *135*, 4457–4466.
[70] Y. Li, L. Chen, Q. Yue, X. Liu, Z. An, G. F. Bills, *ACS Chem. Biol.* **2015**, *10*, 1702–1710.
[71] J. L. Slightom, B. P. Metzger, H. T. Luu, A. P. Elhammer, *Gene* **2009**, *431*, 67–79.
[72] K. Viehrig, F. Surup, K. Harmrolfs, R. Jansen, B. Kunze, R. Müller, *J. Am. Chem. Soc.* **2013**, *135*, 16885–16894.
[73] C. Mahlert, F. Kopp, J. Thirlway, J. Micklefield, M. A. Marahiel, *J. Am. Chem. Soc.* **2007**, *129*, 12011–12018.
[74] I. Höfer, M. Crüsemann, M. Radzom, B. Geers, D. Flachshaar, X. Cai, A. Zeeck, J. Piel, *Chem. Biol.* **2011**, *18*, 381–391.
[75] C. Fu et al., *J. Am. Chem. Soc.* **2015**, *137*, 7692–7705.

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[76] B. Van de Poel, D. Van Der Straeten, *Front. Plant Sci.* **2014**, *5*, 640.
[77] E. Heinzelmann, S. Berger, O. Puk, B. Reichenstein, W. Wohlleben, D. Schwartz, *Antimicrob. Agents Chemother.* **2003**, *47*, 447–457.
[78] B. M. Kevany, D. A. Rasko, M. G. Thomas, *Appl. Environ. Microbiol.* **2009**, *75*, 1144–1155.
[79] M. G. Thomas, Y. A. Chan, S. G. Ozanick, *Antimicrob. Agents Chemother.* **2003**, *47*, 2823–2830.
[80] E. A. Felnagle, M. R. Rondon, A. D. Berti, H. A. Crosby, M. G. Thomas, *Appl. Environ. Microbiol.* **2007**, *73*, 4162–4170.
[81] L. M. Iyer, S. Abhiman, A. M. Burroughs, L. Aravind, *Mol. BioSyst.* **2009**, *5*, 1636–1660.
[82] P. Louis, E. A. Galinski, *Microbiology* **1997**, *143*, 1141–1149.
[83] N. A. Magarvey, B. Haltli, M. He, M. Greenstein, J. A. Hucul, *Antimicrob. Agents Chemother.* **2006**, *50*, 2167–2177.
[84] A. M. Burroughs, R. W. Hoppe, N. C. Goebel, B. H. Sayyed, T. J. Voegtline, A. W. Schwabacher, T. M. Zabriskie, N. R. Silvaggi, *Biochemistry* **2013**, *52*, 4492–4506.
[85] L. Han, A. W. Schwabacher, G. R. Moran, N. R. Silvaggi, *Biochemistry* **2015**, *54*, 7029–7040.
[86] J. Ju, S. G. Ozanick, B. Shen, M. G. Thomas, *ChemBioChem* **2004**, *5*, 1281–1285.
[87] X. Yin, T. M. Zabriskie, *ChemBioChem* **2004**, *5*, 1274–1277.
[88] A. J. Oelke, D. J. France, T. Hofmann, G. Wuitschik, S. V. Ley, *Nat. Prod. Rep.* **2011**, *28*, 1445–1471.
[89] C. S. Neumann, W. Jiang, J. R. Heemstra, E. A. Gontang, R. Kolter, C. T. Walsh, *ChemBioChem* **2012**, *13*, 972–976.
[90] V. Miao et al., *Microbiology* **2005**, *151*, 1507–1523.
[91] V. Blanc et al., *Mol. Microbiol.* **1997**, *23*, 191–202.
[92] T. M. Makris, V. V. Vu, K. K. Meier, A. J. Komor, B. S. Rivard, E. Münck, L. Que, J. D. Lipscomb, *J. Am. Chem. Soc.* **2015**, *137*, 1608–1617.
[93] T. Huang, Y. Wang, J. Yin, Y. Du, M. Tao, J. Xu, W. Chen, S. Lin, Z. Deng, *J. Biol. Chem.* **2011**, *286*, 20648–20657.
[94] G. J. Gatto, Jr., M. T. Boyne, N. L. Kelleher, C. T. Walsh, *J. Am. Chem. Soc.* **2006**, *128*, 3838–3847.
[95] G. E. Tsotsou, F. Barbirato, *Biochimie* **2007**, *89*, 591–604.
[96] C. Bruntner, C. Bormann, *Eur. J. Biochem. FEBS* **1998**, *254*, 347–355.
[97] C. Keller-Juslén, M. Kuhn, H. R. Loosli, T. J. Petcher, H. P. Weber, A. von Wartburg, *Tetrahedron Lett.* **1976**, *17*, 4147–4150.
[98] F. Benz, F. Knüsel, J. Nüesch, H. Treichler, W. Voser, R. Nyfeler, W. Keller-Schierlein, *Helv. Chim. Acta* **1974**, *57*, 2459–2477.
[99] T. An et al., *J. Nat. Prod.* **2007**, *70*, 730–735.
[100] H. S. Okumura, B. Philmus, C. Portmann, T. K. Hemscheidt, *J. Nat. Prod.* **2009**, *72*, 172–176.
[101] R. A. Cacho, W. Jiang, Y.-H. Chooi, C. T. Walsh, Y. Tang, *J. Am. Chem. Soc.* **2012**, *134*, 16781–16790.
[102] K. Koketsu, S. Mitsuhashi, K. Tabata, *Appl. Environ. Microbiol.* **2013**, *79*, 2201–2208.
[103] B. K. Hubbard, M. G. Thomas, C. T. Walsh, *Chem. Biol.* **2000**, *7*, 931–942.
[104] O. W. Choroba, D. H. Williams, J. B. Spencer, *J. Am. Chem. Soc.* **2000**, *122*, 5389–5390.
[105] R. S. Al Toma, C. Brieke, M. J. Cryle, R. D. Süssmuth, *Nat. Prod. Rep.* **2015**, *32*, 1207–1235.
[106] F. Kudo, A. Miyanaga, T. Eguchi, *Nat. Prod. Rep.* **2014**, *31*, 1056–1073.
[107] S. Rachid, D. Krug, K. J. Weissman, R. Müller, *J. Biol. Chem.* **2007**, *282*, 21810–21817.
[108] D. Bischoff, B. Bister, M. Bertazzo, V. Pfeifer, E. Stegmann, G. J. Nicholson, S. Keller, S. Pelzer, W. Wohlleben, R. D. Süssmuth, *ChemBioChem* **2005**, *6*, 267–272.
[109] P. C. Schmartz, K. Zerbe, K. Abou-Hadeed, J. A. Robinson, *Org. Biomol. Chem.* **2014**, *12*, 5574–5577.
[110] J. R. Heemstra, C. T. Walsh, *J. Am. Chem. Soc.* **2008**, *130*, 14024–14025.
[111] F. H. Vaillancourt, J. Yin, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 10111–10116.
[112] F. H. Vaillancourt, E. Yeh, D. A. Vosburg, S. E. O'Connor, C. T. Walsh, *Nature* **2005**, *436*, 1191–1194.
[113] B. Bister, D. Bischoff, G. J. Nicholson, S. Stockert, J. Wink, C. Brunati, S. Donadio, S. Pelzer, W. Wohlleben, R. D. Süssmuth, *ChemBioChem* **2003**, *4*, 658–662.
[114] C. D. Murphy, D. O'Hagan, C. Schaffrath, *Angew. Chem. Int. Ed.* **2001**, *40*, 4479–4481; *Angew. Chem.* **2001**, *113*, 4611–4613.
[115] Q. Li, X. Qin, J. Liu, C. Gui, B. Wang, J. Li, J. Ju, *J. Am. Chem. Soc.* **2016**, *138*, 408–415.
[116] W. Zhang, B. Ostash, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 16828–16833.
[117] X. Qu, N. Jiang, F. Xu, L. Shao, G. Tang, B. Wilkinson, W. Liu, *Mol. BioSyst.* **2011**, *7*, 852–861.
[118] W. W. Metcalf, W. A. van der Donk, *Annu. Rev. Biochem.* **2009**, *78*, 65–94.
[119] M. Welker, H. V. Döhren, *FEMS Microbiol. Rev.* **2006**, *30*, 530–563.
[120] Q. Zhao et al., *Chem. Biol.* **2008**, *15*, 693–705.
[121] W. Ding, W. Deng, M. Tang, Q. Zhang, G. Tang, Y. Bi, W. Liu, *Mol. BioSyst.* **2010**, *6*, 1071–1081.
[122] D. B. Hansen, S. B. Bumpus, Z. D. Aron, N. L. Kelleher, C. T. Walsh, *J. Am. Chem. Soc.* **2007**, *129*, 6366–6367.
[123] F. I. Kraas, V. Helmetag, M. Wittmann, M. Strieker, M. A. Marahiel, *Chem. Biol.* **2010**, *17*, 872–880.
[124] F. I. Kraas, T. W. Giessen, M. A. Marahiel, *FEBS Lett.* **2012**, *586*, 283–288.
[125] K. N. Raymond, E. A. Dertz, S. S. Kim, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3584–3588.
[126] U. Keller, M. Lang, I. Crnovcic, F. Pfennig, F. Schauwecker, *J. Bacteriol.* **2010**, *192*, 2583–2595.
[127] M. G. Thomas, M. D. Burkart, C. T. Walsh, *Chem. Biol.* **2002**, *9*, 171–184.
[128] W. Namwat, H. Kinoshita, T. Nihira, *J. Bacteriol.* **2002**, *184*, 4811–4818.
[129] A. Tanovic, S. A. Samel, L.-O. Essen, M. A. Marahiel, *Science* **2008**, *321*, 659–663.
[130] A. M. Gulick, *ACS Chem. Biol.* **2009**, *4*, 811–827.
[131] E. J. Drake, B. R. Miller, C. Shi, J. T. Tarrasch, J. A. Sundlov, C. L. Allen, G. Skiniotis, C. C. Aldrich, A. M. Gulick, *Nature* **2016**, *529*, 235–238.
[132] C. A. Mitchell, C. Shi, C. C. Aldrich, A. M. Gulick, *Biochemistry* **2012**, *51*, 3252–3263.
[133] J. M. Reimer, M. N. Aloise, P. M. Harrison, T. M. Schmeing, *Nature* **2016**, *529*, 239–242.
[134] E. Conti, T. Stachelhaus, M. A. Marahiel, P. Brick, *EMBO J.* **1997**, *16*, 4174–4183.
[135] T. V. Lee, L. J. Johnson, R. D. Johnson, A. Koulman, G. A. Lane, J. S. Lott, V. L. Arcus, *J. Biol. Chem.* **2010**, *285*, 2415–2427.
[136] G. L. Challis, J. Ravel, C. A. Townsend, *Chem. Biol.* **2000**, *7*, 211–224.
[137] T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem. Biol.* **1999**, *6*, 493–505.
[138] M. Röttig, M. H. Medema, K. Blin, T. Weber, C. Rausch, O. Kohlbacher, *Nucleic Acids Res.* **2011**, *39*, W362–W367.
[139] H. Kaljunen, S. H. H. Schiefelbein, D. Stummer, S. Kozak, R. Meijers, G. Christiansen, A. Rentmeister, *Angew. Chem. Int. Ed.* **2015**, *54*, 8833–8836; *Angew. Chem.* **2015**, *127*, 8957–8961.
[140] A. Miyanaga, J. Cieślak, Y. Shinohara, F. Kudo, T. Eguchi, *J. Biol. Chem.* **2014**, *289*, 31448–31457.

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[141] C. K. Thornburg, S. Wortas-Strom, M. Nosrati, J. H. Geiger, K. D. Walker, *Biochemistry* **2015**, *54*, 6230 – 6242.
[142] S. Meyer, J.-C. Kehr, A. Mainz, D. Dehm, D. Petras, R. D. Süssmuth, E. Dittmann, *Cell Chem. Biol.* **2016**, *23*, 462 – 471.
[143] S. Müller, E. Garcia-Gonzalez, A. Mainz, G. Hertlein, N. C. Heid, E. Mösker, H. van den Elst, H. S. Overkleeft, E. Genersch, R. D. Süssmuth, *Angew. Chem. Int. Ed.* **2014**, *53*, 10821 – 10825; *Angew. Chem.* **2014**, *126*, 10998 – 11002.
[144] R. Wu, A. S. Reger, X. Lu, A. M. Gulick, D. Dunaway-Mariano, *Biochemistry* **2009**, *48*, 4115 – 4125.
[145] B. R. Miller, J. A. Sundlov, E. J. Drake, T. A. Makin, A. M. Gulick, *Proteins Struct. Funct. Bioinf.* **2014**, *82*, 2691 – 2702.
[146] K. J. Labby, S. G. Watsula, S. Garneau-Tsodikova, *Nat. Prod. Rep.* **2015**, *32*, 641 – 653.
[147] L. E. Quadri, J. Sello, T. A. Keating, P. H. Weinreb, C. T. Walsh, *Chem. Biol.* **1998**, *5*, 631 – 645.
[148] M. Wolpert, B. Gust, B. Kammerer, L. Heide, *Microbiology* **2007**, *153*, 1413 – 1423.
[149] S. Lautru, D. Oves-Costales, J.-L. Pernodet, G. L. Challis, *Microbiology* **2007**, *153*, 1405 – 1412.
[150] B. Boll, T. Taubitz, L. Heide, *J. Biol. Chem.* **2011**, *286*, 36281 – 36290.
[151] W. Zhang, J. R. Heemstra, C. T. Walsh, H. J. Imker, *Biochemistry* **2010**, *49*, 9946 – 9947.
[152] E. A. Felnagle, J. J. Barkei, H. Park, A. M. Podevels, M. D. McMahon, D. W. Drott, M. G. Thomas, *Biochemistry* **2010**, *49*, 8815 – 8817.
[153] E. J. Drake, J. Cao, J. Qu, M. B. Shah, R. M. Straubinger, A. M. Gulick, *J. Biol. Chem.* **2007**, *282*, 20425 – 20434.
[154] G. W. Buchko, C.-Y. Kim, T. C. Terwilliger, P. J. Myler, *Tuberculosis (Edinb.)* **2010**, *90*, 245 – 251.
[155] D. A. Herbst, B. Boll, G. Zocher, T. Stehle, L. Heide, *J. Biol. Chem.* **2013**, *288*, 1991 – 2003.
[156] B. R. Miller, E. J. Drake, C. Shi, C. C. Aldrich, A. M. Gulick, *J. Biol. Chem.* **2016**, *291*, 22559 – 22571.
[157] T. Kittilä, A. Mollo, L. K. Charkoudian, M. J. Cryle, *Angew. Chem. Int. Ed.* **2016**, *55*, 9834 – 9840; *Angew. Chem.* **2016**, *128*, 9988 – 9995.
[158] T. Weber, R. Baumgartner, C. Renner, M. A. Marahiel, T. A. Holak, *Structure* **2000**, *8*, 407 – 418.
[159] A. Koglin, M. R. Mofid, F. Löhr, B. Schäfer, V. V. Rogov, M.-M. Blum, T. Mittag, M. A. Marahiel, F. Bernhard, V. Dötsch, *Science* **2006**, *312*, 273 – 276.
[160] R. H. Lambalot, A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, R. Reid, C. Khosla, C. T. Walsh, *Chem. Biol.* **1996**, *3*, 923 – 936.
[161] J. Beld, E. C. Sonnenschein, C. R. Vickery, J. P. Noel, M. D. Burkart, *Nat. Prod. Rep.* **2014**, *31*, 61 – 108.
[162] P. Tufar et al., *Chem. Biol.* **2014**, *21*, 552 – 562.
[163] L. E. Quadri, P. H. Weinreb, M. Lei, M. M. Nakano, P. Zuber, C. T. Walsh, *Biochemistry* **1998**, *37*, 1585 – 1595.
[164] J. J. La Clair, T. L. Foley, T. R. Schegg, C. M. Regan, M. D. Burkart, *Chem. Biol.* **2004**, *11*, 195 – 201.
[165] N. M. Gaudelli, D. H. Long, C. A. Townsend, *Nature* **2015**, *520*, 383 – 387.
[166] D. S. Vallari, S. Jackowski, C. O. Rock, *J. Biol. Chem.* **1987**, *262*, 2468 – 2471.
[167] K. Haslinger, C. Redfield, M. J. Cryle, *Proteins Struct. Funct. Bioinf.* **2015**, *83*, 711 – 721.
[168] M. J. Jaremko, D. J. Lee, S. J. Opella, M. D. Burkart, *J. Am. Chem. Soc.* **2015**, *137*, 11546 – 11549.
[169] A. C. Goodrich, B. J. Harden, D. P. Frueh, *J. Am. Chem. Soc.* **2015**, *137*, 12100 – 12109.
[170] D. P. Dowling, Y. Kung, A. K. Croft, K. Taghizadeh, W. L. Kelly, C. T. Walsh, C. L. Drennan, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 12432 – 12437.
[171] W.-H. Chen, K. Li, N. S. Guntaka, S. D. Bruner, *ACS Chem. Biol.* **2016**, *11*, 2293 – 2303.
[172] D. P. Frueh, H. Arthanari, A. Koglin, D. A. Vosburg, A. E. Bennett, C. T. Walsh, G. Wagner, *Nature* **2008**, *454*, 903 – 906.
[173] Y. Liu, T. Zheng, S. D. Bruner, *Chem. Biol.* **2011**, *18*, 1482 – 1488.
[174] K. Haslinger, C. Brieke, S. Uhlmann, L. Sieverling, R. D. Süssmuth, M. J. Cryle, *Angew. Chem. Int. Ed.* **2014**, *53*, 8518 – 8522; *Angew. Chem.* **2014**, *126*, 8658 – 8662.
[175] M. Peschke, K. Haslinger, C. Brieke, J. Reinstein, M. J. Cryle, *J. Am. Chem. Soc.* **2016**, *138*, 6746 – 6753.
[176] T. A. Keating, C. G. Marshall, C. T. Walsh, A. E. Keating, *Nat. Struct. Biol.* **2002**, *9*, 522 – 526.
[177] M. A. Marahiel, T. Stachelhaus, H. D. Mootz, *Chem. Rev.* **1997**, *97*, 2651 – 2674.
[178] T. Stachelhaus, H. D. Mootz, V. Bergendahl, M. A. Marahiel, *J. Biol. Chem.* **1998**, *273*, 22773 – 22781.
[179] V. Bergendahl, U. Linne, M. A. Marahiel, *Eur. J. Biochem. FEBS* **2002**, *269*, 620 – 629.
[180] E. D. Roche, C. T. Walsh, *Biochemistry* **2003**, *42*, 1334 – 1344.
[181] S. A. Samel, G. Schoenafinger, T. A. Knappe, M. A. Marahiel, L.-O. Essen, *Structure* **2007**, *15*, 781 – 792.
[182] S. A. Samel, P. Czodrowski, L.-O. Essen, *Acta Crystallogr. Sect. D* **2014**, *70*, 1442 – 1452.
[183] K. Bloudoff, D. A. Alonzo, T. M. Schmeing, *Cell Chem. Biol.* **2016**, *23*, 331 – 339.
[184] K. Bloudoff, D. Rodionov, T. M. Schmeing, *J. Mol. Biol.* **2013**, *425*, 3137 – 3150.
[185] P. J. Belshaw, C. T. Walsh, T. Stachelhaus, *Science* **1999**, *284*, 486 – 489.
[186] C. Rausch, I. Hoof, T. Weber, W. Wohlleben, D. H. Huson, *BMC Evol. Biol.* **2007**, *7*, 78.
[187] J. Zhang, N. Liu, R. A. Cacho, Z. Gong, Z. Liu, W. Qin, C. Tang, Y. Tang, J. Zhou, *Nat. Chem. Biol.* **2016**, *12*, 1001 – 1003.
[188] K. Haslinger, M. Peschke, C. Brieke, E. Maximowitsch, M. J. Cryle, *Nature* **2015**, *521*, 105 – 109.
[189] S. L. Clugston, S. A. Sieber, M. A. Marahiel, C. T. Walsh, *Biochemistry* **2003**, *42*, 12095 – 12104.
[190] C. J. Balibar, F. H. Vaillancourt, C. T. Walsh, *Chem. Biol.* **2005**, *12*, 1189 – 1200.
[191] H. B. Bode et al., *Angew. Chem. Int. Ed.* **2015**, *54*, 10352 – 10355; *Angew. Chem.* **2015**, *127*, 10493 – 10496.
[192] S. Lin, S. G. Van Lanen, B. Shen, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 4183 – 4188.
[193] R. Zocher, U. Keller, H. Kleinkauf, *Biochemistry* **1982**, *21*, 43 – 48.
[194] X. Gao, S. W. Haynes, B. D. Ames, P. Wang, L. P. Vien, C. T. Walsh, Y. Tang, *Nat. Chem. Biol.* **2012**, *8*, 823 – 830.
[195] M. Hahn, T. Stachelhaus, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15585 – 15590.
[196] M. Hahn, T. Stachelhaus, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 275 – 280.
[197] C. Chiocchini, U. Linne, T. Stachelhaus, *Chem. Biol.* **2006**, *13*, 899 – 908.
[198] S. Gruenewald, H. D. Mootz, P. Stehmeier, T. Stachelhaus, *Appl. Environ. Microbiol.* **2004**, *70*, 3282 – 3291.
[199] C. D. Richter, D. Nietlispach, R. W. Broadhurst, K. J. Weissman, *Nat. Chem. Biol.* **2008**, *4*, 75 – 81.
[200] V. N. Uversky, *J. Biol. Chem.* **2016**, *291*, 6681 – 6688.
[201] N. M. Gaudelli, C. A. Townsend, *Nat. Chem. Biol.* **2014**, *10*, 251 – 258.
[202] S. D. Bruner, T. Weber, R. M. Kohli, D. Schwarzer, M. A. Marahiel, C. T. Walsh, M. T. Stubbs, *Structure* **2002**, *10*, 301 – 310.
[203] S. A. Samel, B. Wagner, M. A. Marahiel, L.-O. Essen, *J. Mol. Biol.* **2006**, *359*, 876 – 889.
[204] R. M. Kohli, C. T. Walsh, *Chem. Commun.* **2003**, 297 – 307.

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[205] C. A. Shaw-Reid, N. L. Kelleher, H. C. Losey, A. M. Gehring, C. Berg, C. T. Walsh, *Chem. Biol.* **1999**, *6*, 385–400.
[206] K. M. Hoyer, C. Mahlert, M. A. Marahiel, *Chem. Biol.* **2007**, *14*, 13–22.
[207] J. W. Trauger, R. M. Kohli, H. D. Mootz, M. A. Marahiel, C. T. Walsh, *Nature* **2000**, *407*, 215–218.
[208] J. W. Trauger, R. M. Kohli, C. T. Walsh, *Biochemistry* **2001**, *40*, 7092–7098.
[209] R. M. Kohli, J. W. Trauger, D. Schwarzer, M. A. Marahiel, C. T. Walsh, *Biochemistry* **2001**, *40*, 7099–7108.
[210] R. M. Kohli, C. T. Walsh, M. D. Burkart, *Nature* **2002**, *418*, 658–661.
[211] H. M. Patel, C. T. Walsh, *Biochemistry* **2001**, *40*, 9023–9031.
[212] J. A. Read, C. T. Walsh, *J. Am. Chem. Soc.* **2007**, *129*, 15762–15763.
[213] A. Chhabra, A. S. Haque, R. K. Pal, A. Goyal, R. Rai, S. Joshi, S. Panjikar, S. Pasha, R. Sankaranarayanan, R. S. Gokhale, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 5681–5686.
[214] Z. Rui, W. Huang, F. Xu, M. Han, X. Liu, S. Lin, W. Zhang, *ACS Chem. Biol.* **2015**, *10*, 1765–1769.
[215] N. M. Llewellyn, Y. Li, J. B. Spencer, *Chem. Biol.* **2007**, *14*, 379–386.
[216] D. Schwarzer, H. D. Mootz, U. Linne, M. A. Marahiel, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14083–14088.
[217] E. Yeh, R. M. Kohli, S. D. Bruner, C. T. Walsh, *ChemBioChem* **2004**, *5*, 1290–1293.
[218] A. Koglin et al., *Nature* **2008**, *454*, 907–911.
[219] J. Hou, L. Robbel, M. A. Marahiel, *Chem. Biol.* **2011**, *18*, 655–664.
[220] S. Weinig, H.-J. Hecht, T. Mahmud, R. Müller, *Chem. Biol.* **2003**, *10*, 939–952.
[221] K. Hoffmann, E. Schneider-Scherzer, H. Kleinkauf, R. Zocher, *J. Biol. Chem.* **1994**, *269*, 12710–12714.
[222] T. Stein, B. Kluge, J. Vater, P. Franke, A. Otto, B. Wittmann-Liebold, *Biochemistry* **1995**, *34*, 4633–4642.
[223] A. Stindl, U. Keller, *Biochemistry* **1994**, *33*, 9358–9364.
[224] L. Luo, R. M. Kohli, M. Onishi, U. Linne, M. A. Marahiel, C. T. Walsh, *Biochemistry* **2002**, *41*, 9184–9196.
[225] D. B. Stein, U. Linne, M. Hahn, M. A. Marahiel, *ChemBioChem* **2006**, *7*, 1807–1814.
[226] U. Linne, M. A. Marahiel, *Biochemistry* **2000**, *39*, 10439–10447.
[227] T. Kouzarides, *Cell* **2007**, *128*, 693–705.
[228] J. Chatterjee, F. Rechenmacher, H. Kessler, *Angew. Chem. Int. Ed.* **2013**, *52*, 254–269; *Angew. Chem.* **2013**, *125*, 268–283.
[229] R. Shi, S. S. Lamb, B. Zakeri, A. Proteau, Q. Cui, T. Sulea, A. Matte, G. D. Wright, M. Cygler, *Chem. Biol.* **2009**, *16*, 401–410.
[230] T. Velkov, A. Lawen, *J. Biol. Chem.* **2003**, *278*, 1137–1148.
[231] A. Billich, R. Zocher, *Biochemistry* **1987**, *26*, 8417–8423.
[232] M. Z. Ansari, J. Sharma, R. S. Gokhale, D. Mohanty, *BMC Bioinf.* **2008**, *9*, 454.
[233] T. Velkov, J. Horne, M. J. Scanlon, B. Capuano, E. Yuriev, A. Lawen, *Chem. Biol.* **2011**, *18*, 464–475.
[234] W. Weckwerth, K. Miyamoto, K. Iinuma, M. Krause, M. Glinski, T. Storm, G. Bonse, H. Kleinkauf, R. Zocher, *J. Biol. Chem.* **2000**, *275*, 17909–17915.
[235] F. Schauwecker, F. Pfennig, N. Grammel, U. Keller, *Chem. Biol.* **2000**, *7*, 287–297.
[236] H. M. Patel, J. Tao, C. T. Walsh, *Biochemistry* **2003**, *42*, 10514–10527.
[237] O. E. Zolova, S. Garneau-Tsodikova, *J. Antibiot.* **2014**, *67*, 59–64.
[238] A. H. Al-Mestarihi, G. Villamizar, J. Fernández, O. E. Zolova, F. Lombó, S. Garneau-Tsodikova, *J. Am. Chem. Soc.* **2014**, *136*, 17350–17354.
[239] D. A. Miller, C. T. Walsh, L. Luo, *J. Am. Chem. Soc.* **2001**, *123*, 8434–8435.
[240] M. A. Wyatt, M. C. Y. Mok, M. Junop, N. A. Magarvey, *ChemBioChem* **2012**, *13*, 2408–2415.
[241] J. F. Barajas, R. M. Phelan, A. J. Schaub, J. T. Kliewer, P. J. Kelly, D. R. Jackson, R. Luo, J. D. Keasling, S.-C. Tsai, *Chem. Biol.* **2015**, *22*, 1018–1029.
[242] F. Kopp, C. Mahlert, J. Grünewald, M. A. Marahiel, *J. Am. Chem. Soc.* **2006**, *128*, 16478–16479.
[243] D. J. Wilson, C. Shi, A. M. Teitelbaum, A. M. Gulick, C. C. Aldrich, *Biochemistry* **2013**, *52*, 926–937.
[244] X. Liu, C. T. Walsh, *Biochemistry* **2009**, *48*, 8746–8757.
[245] G. Schoenafinger, N. Schracke, U. Linne, M. A. Marahiel, *J. Am. Chem. Soc.* **2006**, *128*, 7406–7407.
[246] D. Konz, A. Klens, K. Schörgendorfer, M. A. Marahiel, *Chem. Biol.* **1997**, *4*, 927–937.
[247] C. G. Marshall, N. J. Hillson, C. T. Walsh, *Biochemistry* **2002**, *41*, 244–250.
[248] T. Duerfahrt, K. Eppelmann, R. Müller, M. A. Marahiel, *Chem. Biol.* **2004**, *11*, 261–271.
[249] D. A. Miller, C. T. Walsh, *Biochemistry* **2001**, *40*, 5313–5321.
[250] T. L. Schneider, B. Shen, C. T. Walsh, *Biochemistry* **2003**, *42*, 9722–9730.
[251] O. Perlova, J. Fu, S. Kuhlmann, D. Krug, A. F. Stewart, Y. Zhang, R. Müller, *Appl. Environ. Microbiol.* **2006**, *72*, 7485–7494.
[252] L. Du, M. Chen, C. Sánchez, B. Shen, *FEMS Microbiol. Lett.* **2000**, *189*, 171–175.
[253] L. Du, C. Sánchez, M. Chen, D. J. Edwards, B. Shen, *Chem. Biol.* **2000**, *7*, 623–642.
[254] M. Ehling-Schulz, N. Vukov, A. Schulz, R. Shaheen, M. Andersson, E. Märtlbauer, S. Scherer, *Appl. Environ. Microbiol.* **2005**, *71*, 105–113.
[255] A. M. Matter, S. B. Hoot, P. D. Anderson, S. S. Neves, Y.-Q. Cheng, *PLoS One* **2009**, *4*, e7194.
[256] N. A. Magarvey, M. Ehling-Schulz, C. T. Walsh, *J. Am. Chem. Soc.* **2006**, *128*, 10698–10699.
[257] Y. Xu, R. Orozco, E. M. K. Wijeratne, P. Espinosa-Artiles, A. A. L. Gunatilaka, S. P. Stock, I. Molnár, *Fungal Genet. Biol.* **2009**, *46*, 353–364.
[258] D. G. Fujimori, S. Hrvatin, C. S. Neumann, M. Strieker, M. A. Marahiel, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 16498–16503.
[259] D. Bischoff, S. Pelzer, B. Bister, G. J. Nicholson, S. Stockert, M. Schirle, W. Wohlleben, G. Jung, R. D. Süssmuth, *Angew. Chem. Int. Ed.* **2001**, *40*, 4688–4691; *Angew. Chem.* **2001**, *113*, 4824–4827.
[260] K. Koketsu, K. Watanabe, H. Suda, H. Oguri, H. Oikawa, *Nat. Chem. Biol.* **2010**, *6*, 408–410.
[261] K. Koketsu, A. Minami, K. Watanabe, H. Oguri, H. Oikawa, *Curr. Opin. Chem. Biol.* **2012**, *16*, 142–149.
[262] B. Bister, D. Bischoff, G. J. Nicholson, S. Stockert, J. Wink, C. Brunati, S. Donadio, S. Pelzer, W. Wohlleben, R. D. Süssmuth, *ChemBioChem* **2003**, *4*, 658–662.
[263] D. Petras et al., *ACS Chem. Biol.* **2016**, *11*, 1198–1204.
[264] S. Pelzer, R. Süssmuth, D. Heckmann, J. Recktenwald, P. Huber, G. Jung, W. Wohlleben, *Antimicrob. Agents Chemother.* **1999**, *43*, 1565–1573.
[265] B. Hadatsch, D. Butz, T. Schmiederer, J. Steudle, W. Wohlleben, R. Süssmuth, E. Stegmann, *Chem. Biol.* **2007**, *14*, 1078–1089.
[266] H. C. Losey, M. W. Peczuh, Z. Chen, U. S. Eggert, S. D. Dong, I. Pelczer, D. Kahne, C. T. Walsh, *Biochemistry* **2001**, *40*, 4745–4755.
[267] M. Oberthür, C. Leimkuhler, R. G. Kruger, W. Lu, C. T. Walsh, D. Kahne, *J. Am. Chem. Soc.* **2005**, *127*, 10747–10752.
[268] D. Bischoff, S. Pelzer, A. Höltzel, G. J. Nicholson, S. Stockert, W. Wohlleben, G. Jung, R. D. Süssmuth, *Angew. Chem. Int. Ed.* **2001**, *40*, 1693–1696; *Angew. Chem.* **2001**, *113*, 1736–1739.

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[269] D. Bischoff, B. Bister, M. Bertazzo, V. Pfeifer, E. Stegmann, G. J. Nicholson, S. Keller, S. Pelzer, W. Wohlleben, R. D. Süssmuth, *ChemBioChem* **2005**, *6*, 267–272.
[270] K. Zerbe, K. Woithe, D. B. Li, F. Vitali, L. Bigler, J. A. Robinson, *Angew. Chem. Int. Ed.* **2004**, *43*, 6709–6713; *Angew. Chem.* **2004**, *116*, 6877–6881.
[271] K. Woithe et al., *J. Am. Chem. Soc.* **2007**, *129*, 6887–6895.
[272] K. Woithe, N. Geib, O. Meyer, T. Wörtz, K. Zerbe, J. A. Robinson, *Org. Biomol. Chem.* **2008**, *6*, 2861–2867.
[273] P. C. Schmartz, K. Wölfel, K. Zerbe, E. Gad, E. S. El Tamany, H. K. Ibrahim, K. Abou-Hadeed, J. A. Robinson, *Angew. Chem. Int. Ed.* **2012**, *51*, 11468–11472; *Angew. Chem.* **2012**, *124*, 11635–11639.
[274] M. Peschke, M. Gonsior, R. D. Süssmuth, M. J. Cryle, *Curr. Opin. Struct. Biol.* **2016**, *41*, 46–53.
[275] S. Uhlmann, R. D. Süssmuth, M. J. Cryle, *ACS Chem. Biol.* **2013**, *8*, 2586–2596.
[276] E. Garcia-Gonzalez, S. Müller, P. Ensle, R. D. Süssmuth, E. Genersch, *Environ. Microbiol.* **2014**, *16*, 1297–1309.
[277] J. A. Blodgett, J. K. Zhang, X. Yu, W. W. Metcalf, *J. Antibiot.* **2016**, *69*, 15–25.
[278] J. Recktenwald, R. Shawky, O. Puk, F. Pfennig, U. Keller, W. Wohlleben, S. Pelzer, *Microbiology* **2002**, *148*, 1105–1118.
[279] F. Lipmann, W. Gevers, H. Kleinkauf, R. Roskoski, *Adv. Enzymol. Relat. Areas Mol. Biol.* **1971**, *35*, 1–34.
[280] M. Gonsior, A. Mühlenweg, M. Tietzmann, S. Rausch, A. Poch, R. D. Süssmuth, *ChemBioChem* **2015**, *16*, 2610–2614.
[281] N. Kessler, H. Schuhmann, S. Morneweg, U. Linne, M. A. Marahiel, *J. Biol. Chem.* **2004**, *279*, 7413–7419.
[282] B. K. Scholz-Schroeder, J. D. Soule, D. C. Gross, *Mol. Plant-Microbe Interact.* **2003**, *16*, 271–280.
[283] E. A. Felnagle, M. R. Rondon, A. D. Berti, H. A. Crosby, M. G. Thomas, *Appl. Environ. Microbiol.* **2007**, *73*, 4162–4170.
[284] M. Sato, T. Nakazawa, Y. Tsunematsu, K. Hotta, K. Watanabe, *Curr. Opin. Chem. Biol.* **2013**, *17*, 537–545.
[285] F. Lombó, A. Velasco, A. Castro, F. de la Calle, A. F. Braña, J. M. Sánchez-Puelles, C. Méndez, J. A. Salas, *ChemBioChem* **2006**, *7*, 366–376.
[286] A. S. A. Mady, O. E. Zolova, M. Á. S. Millán, G. Villamizar, F. de la Calle, F. Lombó, S. Garneau-Tsodikova, *Mol. BioSyst.* **2011**, *7*, 1999–2011.
[287] Y.-Q. Cheng, *ChemBioChem* **2006**, *7*, 471–477.
[288] H. Peeters, R. Zocher, H. Kleinkauf, *J. Antibiot.* **1988**, *41*, 352–359.
[289] M. Juguet, S. Lautru, F.-X. Francou, S. Nezbedová, P. Leblond, M. Gondry, J.-L. Pernodet, *Chem. Biol.* **2009**, *16*, 421–431.
[290] A. H. Al-Mestarihi, A. Garzan, J. M. Kim, S. Garneau-Tsodikova, *ChemBioChem* **2015**, *16*, 1307–1313.
[291] S. W. Haynes, X. Gao, Y. Tang, C. T. Walsh, *J. Am. Chem. Soc.* **2012**, *134*, 17444–17447.
[292] X. Gao, W. Jiang, G. Jiménez-Osés, M. S. Choi, K. N. Houk, Y. Tang, C. T. Walsh, *Chem. Biol.* **2013**, *20*, 870–878.
[293] M. Glinski, C. Urbanke, T. Hornbogen, R. Zocher, *Arch. Microbiol.* **2002**, *178*, 267–273.
[294] H. Wang, D. P. Fewer, L. Holm, L. Rouhiainen, K. Sivonen, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9259–9264.
[295] L. T. Fernández-Martínez, C. Borsetto, J. P. Gomez-Escribano, M. J. Bibb, M. M. Al-Bassam, G. Chandra, M. J. Bibb, *Antimicrob. Agents Chemother.* **2014**, *58*, 7441–7450.
[296] C. Maruyama, J. Toyoda, Y. Kato, M. Izumikawa, M. Takagi, K. Shin-ya, H. Katano, T. Utagawa, Y. Hamano, *Nat. Chem. Biol.* **2012**, *8*, 791–797.
[297] S. Garneau-Tsodikova, P. C. Dorrestein, N. L. Kelleher, C. T. Walsh, *J. Am. Chem. Soc.* **2006**, *128*, 12600–12601.
[298] D. E. Ehmann, C. A. Shaw-Reid, H. C. Losey, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 2509–2514.
[299] B. A. Pfeifer, C. C. C. Wang, C. T. Walsh, C. Khosla, *Appl. Environ. Microbiol.* **2003**, *69*, 6698–6702.
[300] Y. Cheng et al., *Mol. Plant-Microbe Interact.* **2013**, *26*, 1294–1301.
[301] M. Jin, M. A. Fischbach, J. Clardy, *J. Am. Chem. Soc.* **2006**, *128*, 10660–10661.
[302] P. D. Fortin, C. T. Walsh, N. A. Magarvey, *Nature* **2007**, *448*, 824–827.
[303] N. Pulsawat, S. Kitani, T. Nihira, *Gene* **2007**, *393*, 31–42.
[304] E. A. Stohl, J. L. Milner, J. Handelsman, *Gene* **1999**, *237*, 403–411.
[305] F. Liu, S. Garneau, C. T. Walsh, *Chem. Biol.* **2004**, *11*, 1533–1542.
[306] K. J. Weissman, *Nat. Chem. Biol.* **2015**, *11*, 660–670.
[307] U. Keller, W. Schlumbohm, *J. Biol. Chem.* **1992**, *267*, 11745–11752.
[308] J. Ma, Z. Wang, H. Huang, M. Luo, D. Zuo, B. Wang, A. Sun, Y.-Q. Cheng, C. Zhang, J. Ju, *Angew. Chem. Int. Ed.* **2011**, *50*, 7797–7802; *Angew. Chem.* **2011**, *123*, 7943–7948.
[309] N. Grammel, D. Schwartz, W. Wohlleben, U. Keller, *Biochemistry* **1998**, *37*, 1596–1603.
[310] D. Reimer, K. M. Pos, M. Thines, P. Grün, H. B. Bode, *Nat. Chem. Biol.* **2011**, *7*, 888–890.
[311] C. A. Brotherton, E. P. Balskus, *J. Am. Chem. Soc.* **2013**, *135*, 3359–3362.
[312] X. Bian, J. Fu, A. Plaza, J. Herrmann, D. Pistorius, A. F. Stewart, Y. Zhang, R. Müller, *ChemBioChem* **2013**, *14*, 1194–1197.
[313] Y. Xu, R. D. Kersten, S.-J. Nam, L. Lu, A. M. Al-Suwailem, H. Zheng, W. Fenical, P. C. Dorrestein, B. S. Moore, P.-Y. Qian, *J. Am. Chem. Soc.* **2012**, *134*, 8625–8632.
[314] J. Masschelein et al., *PLoS One* **2013**, *8*, e54143.
[315] W. Zhang, I. Ntai, N. L. Kelleher, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 12249–12253.
[316] A. Kirschning, F. Hahn, *Angew. Chem. Int. Ed.* **2012**, *51*, 4012–4022; *Angew. Chem.* **2012**, *124*, 4086–4096.
[317] A. Lawen, R. Traber, D. Geyl, R. Zocher, H. Kleinkauf, *J. Antibiot.* **1989**, *42*, 1283–1289.
[318] A. Lawen, J. Dittmann, B. Schmidt, D. Riesner, H. Kleinkauf, *Biochimie* **1992**, *74*, 511–516.
[319] A. Lawen, R. Traber, R. Reuille, M. Ponelle, *Biochem. J.* **1994**, *300*, 395–399.
[320] A. Billich, R. Zocher, *J. Biol. Chem.* **1987**, *262*, 17258–17259.
[321] S. C. Feifel, T. Schmiederer, T. Hornbogen, H. Berg, R. D. Süssmuth, R. Zocher, *ChemBioChem* **2007**, *8*, 1767–1770.
[322] J. Müller, S. C. Feifel, T. Schmiederer, R. Zocher, R. D. Süssmuth, *ChemBioChem* **2009**, *10*, 323–328.
[323] D. Matthes, L. Richter, J. Müller, A. Denisiuk, S. C. Feifel, Y. Xu, P. Espinosa-Artiles, R. D. Süssmuth, I. Molnár, *Chem. Commun.* **2012**, *48*, 5674–5676.
[324] W. Xu, X. Cai, M. E. Jung, Y. Tang, *J. Am. Chem. Soc.* **2010**, *132*, 13604–13607.
[325] J. S. Davies, *J. Pept. Sci.* **2003**, *9*, 471–501.
[326] C. J. White, A. K. Yudin, *Nat. Chem.* **2011**, *3*, 509–524.
[327] R. M. Kohli, C. T. Walsh, M. D. Burkart, *Nature* **2002**, *418*, 658–661.
[328] R. M. Kohli, M. D. Burke, J. Tao, C. T. Walsh, *J. Am. Chem. Soc.* **2003**, *125*, 7160–7161.
[329] G. Xie, M. Uttamchandani, G. Y. J. Chen, X. Bu, S. S. Lin, K. M. Wong, W. Yan, S. Q. Yao, Z. Guo, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 989–992.
[330] J. Grünewald, M. A. Marahiel, *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 121–146.
[331] R. M. Kohli, J. Takagi, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1247–1252.
[332] H. Lin, C. T. Walsh, *J. Am. Chem. Soc.* **2004**, *126*, 13998–14003.

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[333] H. Lin, D. A. Thayer, C.-H. Wong, C. T. Walsh, *Chem. Biol.* **2004**, *11*, 1635 – 1642.
[334] D. Garbe, S. A. Sieber, N. G. Bandur, U. Koert, M. A. Marahiel, *ChemBioChem* **2004**, *5*, 1000 – 1003.
[335] J. Grünewald, S. A. Sieber, M. A. Marahiel, *Biochemistry* **2004**, *43*, 2915 – 2925.
[336] J. Grünewald, S. A. Sieber, C. Mahlert, U. Linne, M. A. Marahiel, *J. Am. Chem. Soc.* **2004**, *126*, 17025 – 17031.
[337] F. Kopp, J. Grünewald, C. Mahlert, M. A. Marahiel, *Biochemistry* **2006**, *45*, 10474 – 10481.
[338] C. Mahlert, S. A. Sieber, J. Grünewald, M. A. Marahiel, *J. Am. Chem. Soc.* **2005**, *127*, 9571 – 9580.
[339] B. Wagner, D. Schumann, U. Linne, U. Koert, M. A. Marahiel, *J. Am. Chem. Soc.* **2006**, *128*, 10513 – 10520.
[340] K. Koketsu, H. Oguri, K. Watanabe, H. Oikawa, *Chem. Biol.* **2008**, *15*, 818 – 828.
[341] K. Watanabe, H. Oguri, H. Oikawa, *Curr. Opin. Chem. Biol.* **2009**, *13*, 189 – 196.
[342] Z. Q. Beck, C. C. Aldrich, N. A. Magarvey, G. I. Georg, D. H. Sherman, *Biochemistry* **2005**, *44*, 13457 – 13466.
[343] N. A. Magarvey, Z. Q. Beck, T. Golakoti, Y. Ding, U. Huber, T. K. Hemscheidt, D. Abelson, R. E. Moore, D. H. Sherman, *ACS Chem. Biol.* **2006**, *1*, 766 – 779.
[344] S. A. Sieber, J. Tao, C. T. Walsh, M. A. Marahiel, *Angew. Chem. Int. Ed.* **2004**, *43*, 493 – 498; *Angew. Chem.* **2004**, *116*, 499 – 504.
[345] Y. Ding, C. M. Rath, K. L. Bolduc, K. Håkansson, D. H. Sherman, *J. Am. Chem. Soc.* **2011**, *133*, 14492 – 14495.
[346] X. Fu, C. Albermann, C. Zhang, J. S. Thorson, *Org. Lett.* **2005**, *7*, 1513 – 1515.
[347] H. C. Losey, J. Jiang, J. B. Biggins, M. Oberthür, X. Y. Ye, S. D. Dong, D. Kahne, J. S. Thorson, C. T. Walsh, *Chem. Biol.* **2002**, *9*, 1305 – 1314.
[348] U. Galm, M. A. Dessoy, J. Schmidt, L. A. Wessjohann, L. Heide, *Chem. Biol.* **2004**, *11*, 173 – 183.
[349] H. Xu, L. Heide, S.-M. Li, *Chem. Biol.* **2004**, *11*, 655 – 662.
[350] M. Fridman, C. J. Balibar, T. Lupoli, D. Kahne, C. T. Walsh, S. Garneau-Tsodikova, *Biochemistry* **2007**, *46*, 8462 – 8471.
[351] Y. Yan, J. Chen, L. Zhang, Q. Zheng, Y. Han, H. Zhang, D. Zhang, T. Awakawa, I. Abe, W. Liu, *Angew. Chem. Int. Ed.* **2013**, *52*, 12308 – 12312; *Angew. Chem.* **2013**, *125*, 12534 – 12538.
[352] M. Krause, A. Lindemann, M. Glinski, T. Hornbogen, G. Bonse, P. Jeschke, G. Thielking, W. Gau, H. Kleinkauf, R. Zocher, *J. Antibiot.* **2001**, *54*, 797 – 804.
[353] A. J. Birch, *Pure Appl. Chem.* **1963**, *7*, 527 – 538.
[354] K. L. Rinehart, *Pure Appl. Chem.* **1977**, *49*, 1361 – 1384.
[355] E. Beacco, M. L. Bianchi, A. Minghetti, C. Spalla, *Experientia* **1978**, *34*, 1291 – 1293.
[356] R. Thiericke, J. Rohr, *Nat. Prod. Rep.* **1993**, *10*, 265 – 289.
[357] S. Weist, R. D. Süssmuth, *Appl. Microbiol. Biotechnol.* **2005**, *68*, 141 – 150.
[358] K.-H. van Pée, E. P. Patallo, *Appl. Microbiol. Biotechnol.* **2006**, *70*, 631 – 641.
[359] K.-H. van Pée, C. Dong, S. Flecks, J. Naismith, E. P. Patallo, T. Wage, *Adv. Appl. Microbiol.* **2006**, *59*, 127 – 157.
[360] K. Takesako, S. Mizutani, H. Sakakibara, M. Endo, Y. Yoshikawa, T. Masuda, E. Sono-Koyama, I. Kato, *J. Antibiot.* **1996**, *49*, 676 – 681.
[361] D. R. Houck, J. Ondeyka, D. L. Zink, E. Inamine, M. A. Goetz, O. D. Hensens, *J. Antibiot.* **1988**, *41*, 882 – 891.
[362] H. Kobel, R. Traber, *Eur. J. Appl. Microbiol. Biotechnol.* **1982**, *14*, 237 – 240.
[363] R. Traber, H. Hofmann, H. Kobel, *J. Antibiot.* **1989**, *42*, 591 – 597.
[364] O. D. Hensens, R. F. White, R. T. Goegelman, E. S. Inamine, A. A. Patchett, *J. Antibiot.* **1992**, *45*, 133 – 135.
[365] H. M. Ge, W. Yan, Z. K. Guo, Q. Luo, R. Feng, L. Y. Zang, Y. Shen, R. H. Jiao, Q. Xu, R. X. Tan, *Chem. Commun.* **2011**, *47*, 2321 – 2323.
[366] Y. Xu, J. Zhan, E. M. K. Wijeratne, A. M. Burns, A. A. L. Gunatilaka, I. Molnár, *J. Nat. Prod.* **2007**, *70*, 1467 – 1471.
[367] D. Gauvreau, M. J. Waring, *Can. J. Microbiol.* **1984**, *30*, 439 – 450.
[368] D. Gauvreau, M. J. Waring, *Can. J. Microbiol.* **1984**, *30*, 730 – 738.
[369] A. S. Ratnayake, J. E. Janso, X. Feng, G. Schlingmann, I. Goljer, G. T. Carter, *J. Nat. Prod.* **2009**, *72*, 496 – 499.
[370] S. Moran, D. K. Rai, B. R. Clark, C. D. Murphy, *Org. Biomol. Chem.* **2009**, *7*, 644 – 646.
[371] S. Grüschow, E. J. Rackham, B. Elkins, P. L. A. Newill, L. M. Hill, R. J. M. Goss, *ChemBioChem* **2009**, *10*, 355 – 360.
[372] J. Delzer, H. P. Fiedler, H. Müller, H. Zähner, R. Rathmann, K. Ernst, W. A. König, *J. Antibiot.* **1984**, *37*, 80 – 82.
[373] C. Bormann, A. Kálmánczhelyi, R. Süssmuth, G. Jung, *J. Antibiot.* **1999**, *52*, 102 – 108.
[374] R. G. Ankenbauer, A. L. Staley, K. L. Rinehart, C. D. Cox, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 1878 – 1882.
[375] S. Weist, B. Bister, O. Puk, D. Bischoff, S. Pelzer, G. J. Nicholson, W. Wohlleben, G. Jung, R. D. Süssmuth, *Angew. Chem. Int. Ed.* **2002**, *41*, 3383 – 3385; *Angew. Chem.* **2002**, *114*, 3531 – 3534.
[376] S. Weist, C. Kittel, D. Bischoff, B. Bister, V. Pfeifer, G. J. Nicholson, W. Wohlleben, R. D. Süssmuth, *J. Am. Chem. Soc.* **2004**, *126*, 5942 – 5943.
[377] Z. Hojati et al., *Chem. Biol.* **2002**, *9*, 1175 – 1187.
[378] A. Powell, M. Al Nakeeb, B. Wilkinson, J. Micklefield, *Chem. Commun.* **2007**, 2683 – 2685.
[379] B. Amir-Heidari, J. Thirlway, J. Micklefield, *Org. Biomol. Chem.* **2008**, *6*, 975 – 978.
[380] R. P. McGlinchey, M. Nett, A. S. Eustáquio, R. N. Asolkar, W. Fenical, B. S. Moore, *J. Am. Chem. Soc.* **2008**, *130*, 7822 – 7823.
[381] M. Nett, T. A. M. Gulder, A. J. Kale, C. C. Hughes, B. S. Moore, *J. Med. Chem.* **2009**, *52*, 6163 – 6167.
[382] A. S. Eustáquio, B. S. Moore, *Angew. Chem. Int. Ed.* **2008**, *47*, 3936 – 3938; *Angew. Chem.* **2008**, *120*, 4000 – 4002.
[383] I. T. Yonemoto, W. Li, A. Khullar, N. Reixach, B. Gerratana, *ACS Chem. Biol.* **2012**, *7*, 973 – 977.
[384] Y. Xu, E. M. K. Wijeratne, P. Espinosa-Artiles, A. A. L. Gunatilaka, I. Molnár, *ChemBioChem* **2009**, *10*, 345 – 354.
[385] C. Anderle, S. Hennig, B. Kammerer, S.-M. Li, L. Wessjohann, B. Gust, L. Heide, *Chem. Biol.* **2007**, *14*, 955 – 967.
[386] K. Ouchaou et al., *ChemBioChem* **2015**, *16*, 915 – 923.
[387] K. Yanai, N. Sumida, K. Okakura, T. Moriya, M. Watanabe, T. Murakami, *Nat. Biotechnol.* **2004**, *22*, 848 – 855.
[388] K. S. Lam, G. Tsueng, K. A. McArthur, S. S. Mitchell, B. C. M. Potts, J. Xu, *J. Antibiot.* **2007**, *60*, 13 – 19.
[389] A. Lawen, R. Traber, *J. Biol. Chem.* **1993**, *268*, 20452 – 20465.
[390] R. Traber, H. Hofmann, H. Kobel, *J. Antibiot.* **1989**, *42*, 591 – 597.
[391] S. A. Sieber, C. T. Walsh, M. A. Marahiel, *J. Am. Chem. Soc.* **2003**, *125*, 10862 – 10866.
[392] B. Amir-Heidari, J. Thirlway, J. Micklefield, *Org. Biomol. Chem.* **2008**, *6*, 975 – 978.
[393] S. Grüschow, E. J. Rackham, R. J. M. Goss, *Chem. Sci.* **2011**, *2*, 2182.
[394] R. J. M. Goss, S. Shankar, A. A. Fayad, *Nat. Prod. Rep.* **2012**, *29*, 870 – 889.
[395] M. Sandy, Z. Rui, J. Gallagher, W. Zhang, *ACS Chem. Biol.* **2012**, *7*, 1956 – 1961.
[396] A. Baumert, D. Erge, D. Gröger, *Planta Med.* **1982**, *44*, 122 – 123.
[397] P. G. Arnison et al., *Nat. Prod. Rep.* **2013**, *30*, 108 – 160.

© 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[398] F. Oldach, R. Al Toma, A. Kuthning, T. Caetano, S. Mendo, N. Budisa, R. D. Süssmuth, *Angew. Chem. Int. Ed.* **2012**, *51*, 415–418; *Angew. Chem.* **2012**, *124*, 429–432.
[399] R. S. Al Toma, A. Kuthning, M. P. Exner, A. Denisiuk, J. Ziegler, N. Budisa, R. D. Süssmuth, *ChemBioChem* **2015**, *16*, 503–509.
[400] M. D. B. Tianero, M. S. Donia, T. S. Young, P. G. Schultz, E. W. Schmidt, *J. Am. Chem. Soc.* **2012**, *134*, 418–425.
[401] T. Stachelhaus, A. Schneider, M. A. Marahiel, *Science* **1995**, *269*, 69–72.
[402] A. Schneider, T. Stachelhaus, M. A. Marahiel, *Mol. Gen. Genet.* **1998**, *257*, 308–318.
[403] S. Doekel, M. A. Marahiel, *Chem. Biol.* **2000**, *7*, 373–384.
[404] F. Schauwecker, F. Pfennig, N. Grammel, U. Keller, *Chem. Biol.* **2000**, *7*, 287–297.
[405] T. Duerfahrt, S. Doekel, T. Sonke, P. J. L. M. Quaedflieg, M. A. Marahiel, *Eur. J. Biochem.* **2003**, *270*, 4555–4563.
[406] H. D. Mootz, D. Schwarzer, M. A. Marahiel, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5848–5853.
[407] H. D. Mootz, N. Kessler, U. Linne, K. Eppelmann, D. Schwarzer, M. A. Marahiel, *J. Am. Chem. Soc.* **2002**, *124*, 10980–10981.
[408] M. M. Yakimov, L. Giuliano, K. N. Timmis, P. N. Golyshin, *J. Mol. Microbiol. Biotechnol.* **2000**, *2*, 217–224.
[409] D. Butz, T. Schmiederer, B. Hadatsch, W. Wohlleben, T. Weber, R. D. Süssmuth, *ChemBioChem* **2008**, *9*, 1195–1200.
[410] K. T. Nguyen, D. Ritz, J.-Q. Gu, D. Alexander, M. Chu, V. Miao, P. Brian, R. H. Baltz, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 17462–17467.
[411] U. Linne, S. Doekel, M. A. Marahiel, *Biochemistry* **2001**, *40*, 15824–15834.
[412] S. Doekel, M.-F. C.-L. Gal, J.-Q. Gu, M. Chu, R. H. Baltz, P. Brian, *Microbiology* **2008**, *154*, 2872–2880.
[413] K. T. Nguyen, X. He, D. C. Alexander, C. Li, J.-Q. Gu, C. Mascio, A. Van Praagh, L. Mortin, M. Chu, J. A. Silverman, et al., *Antimicrob. Agents Chemother.* **2010**, *54*, 1404–1413.
[414] M.-F. C.-L. Gal, L. Thurston, P. Rich, V. Miao, R. H. Baltz, *Microbiology* **2006**, *152*, 2993–3001.
[415] V. Miao et al., *Chem. Biol.* **2006**, *13*, 269–276.
[416] M. J. Calcott, J. G. Owen, I. L. Lamont, D. F. Ackerley, *Appl. Environ. Microbiol.* **2014**, *80*, 5723–5731.
[417] M. J. Calcott, D. F. Ackerley, *BMC Microbiol.* **2015**, *15*, 162.
[418] S. Zobel, S. Boecker, D. Kulke, D. Heimbach, V. Meyer, R. D. Süssmuth, *ChemBioChem* **2016**, *17*, 283–287.
[419] M. Hahn, T. Stachelhaus, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 275–280.
[420] K. Eppelmann, T. Stachelhaus, M. A. Marahiel, *Biochemistry* **2002**, *41*, 9718–9726.
[421] G. C. Uguru, C. Milne, M. Borg, F. Flett, C. P. Smith, J. Micklefield, *J. Am. Chem. Soc.* **2004**, *126*, 5032–5033.
[422] J. Thirlway, R. Lewis, L. Nunns, M. Al Nakeeb, M. Styles, A.-W. Struck, C. P. Smith, J. Micklefield, *Angew. Chem. Int. Ed.* **2012**, *51*, 7181–7184; *Angew. Chem.* **2012**, *124*, 7293–7296.
[423] B. Villiers, F. Hollfelder, *Chem. Biol.* **2011**, *18*, 1290–1299.
[424] M. Crüsemann, C. Kohlhaas, J. Piel, *Chem. Sci.* **2013**, *4*, 1041–1045.
[425] H. Kries, D. L. Niquille, D. Hilvert, *Chem. Biol.* **2015**, *22*, 640–648.
[426] H. Kries, R. Wachtel, A. Pabst, B. Wanner, D. Niquille, D. Hilvert, *Angew. Chem. Int. Ed.* **2014**, *53*, 10105–10108; *Angew. Chem.* **2014**, *126*, 10269–10272.
[427] M. A. Fischbach, J. R. Lai, E. D. Roche, C. T. Walsh, D. R. Liu, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 11951–11956.
[428] B. S. Evans, Y. Chen, W. W. Metcalf, H. Zhao, N. L. Kelleher, *Chem. Biol.* **2011**, *18*, 601–607.
[429] C. J. Dutton, S. P. Gibson, A. C. Goudie, K. S. Holdom, M. S. Pacey, J. C. Ruddock, J. D. Bu′Lock, M. K. Richards, *J. Antibiot.* **1991**, *44*, 357–365.
[430] A. P. Praseuth, M. B. Praseuth, H. Oguri, H. Oikawa, K. Watanabe, C. C. C. Wang, *Biotechnol. Prog.* **2008**, *24*, 134–139.
[431] L. Richter, F. Wanka, S. Boecker, D. Storm, T. Kurt, Ö. Vural, R. Süßmuth, V. Meyer, *Fungal Biol. Biotechnol.* **2014**, *1*, 4.
[432] M. N. Heneghan, A. A. Yakasai, L. M. Halo, Z. Song, A. M. Bailey, T. J. Simpson, R. J. Cox, C. M. Lazarus, *ChemBioChem* **2010**, *11*, 1508–1512.
[433] K. Watanabe, K. Hotta, M. Nakaya, A. P. Praseuth, C. C. C. Wang, D. Inada, K. Takahashi, E. Fukushi, H. Oguri, H. Oikawa, *J. Am. Chem. Soc.* **2009**, *131*, 9347–9353.
[434] D. Yu, F. Xu, D. Gage, J. Zhan, *Chem. Commun.* **2013**, *49*, 6176–6178.
[435] F. de Ferra, F. Rodriguez, O. Tortora, C. Tosi, G. Grandi, *J. Biol. Chem.* **1997**, *272*, 25304–25309.
[436] D. F. Ackerley, I. L. Lamont, *Chem. Biol.* **2004**, *11*, 971–980.
[437] K. Watanabe, K. Hotta, A. P. Praseuth, M. Searcey, C. C. C. Wang, H. Oguri, H. Oikawa, *ChemBioChem* **2009**, *10*, 1965–1968.
[438] F. von Nussbaum, R. D. Süssmuth, *Angew. Chem. Int. Ed.* **2015**, *54*, 6684–6686; *Angew. Chem.* **2015**, *127*, 6784–6786.
[439] L. L. Ling et al., *Nature* **2015**, *517*, 455–459.
[440] D. Nichols, N. Cahoon, E. M. Trakhtenberg, L. Pham, A. Mehta, A. Belanger, T. Kanigan, K. Lewis, S. S. Epstein, *Appl. Environ. Microbiol.* **2010**, *76*, 2445–2450.
[441] R. Fernandez-Leiro, S. H. W. Scheres, *Nature* **2016**, *537*, 339–346.
[442] R. Rosenzweig, L. E. Kay, *J. Am. Chem. Soc.* **2016**, *138*, 1466–1477.
[443] A. Mainz et al., *Nat. Struct. Mol. Biol.* **2015**, *22*, 898–905.
[444] J. Hennig, M. Sattler, *Protein Sci.* **2014**, *23*, 669–682.

Manuscript received: September 19, 2016
Accepted Article published: January 11, 2017

 © 2017 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim