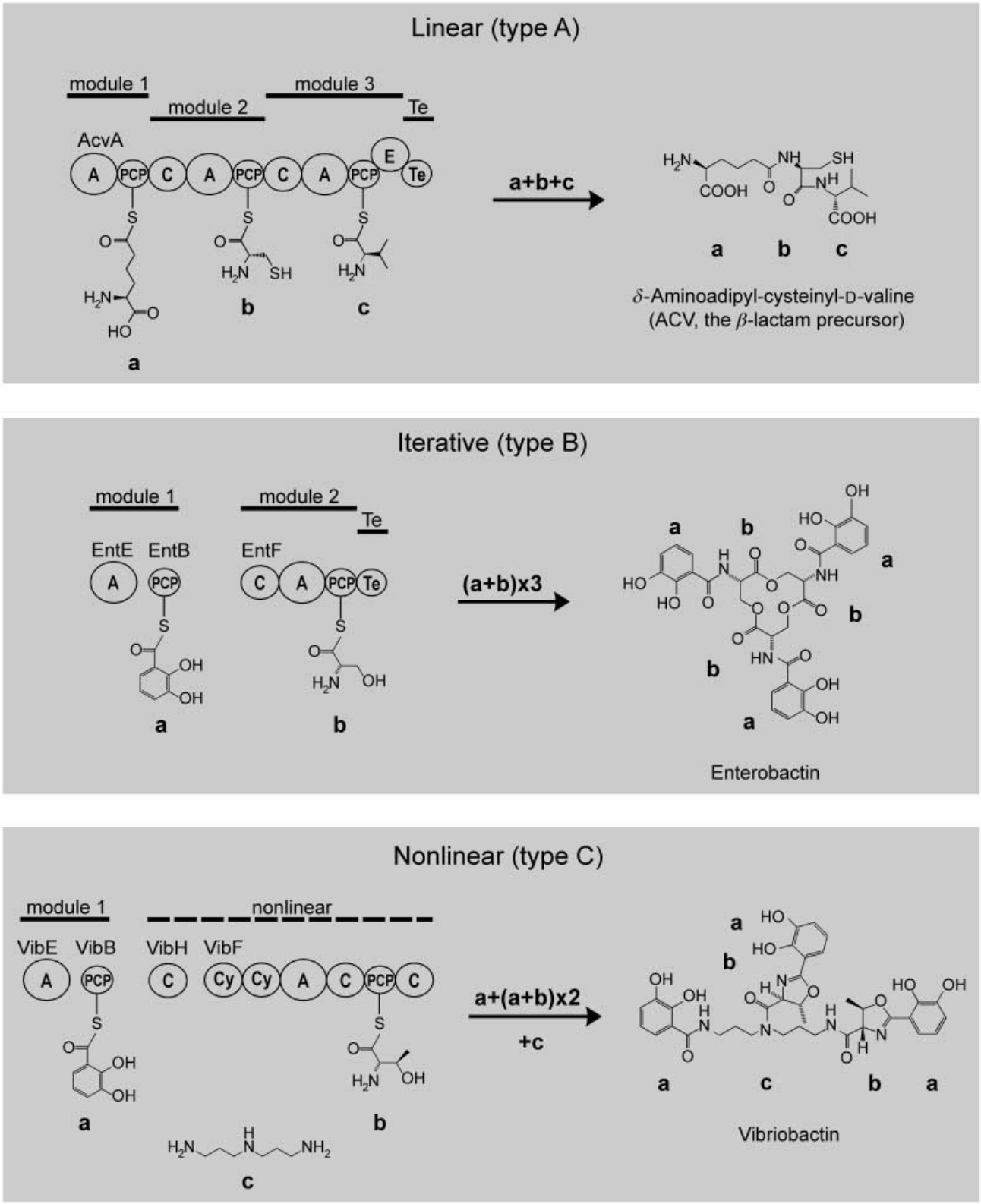
Biosynthetic strategies of nonribosomal peptide synthetases



  1439-4227/02/03/06 $ 20.00+.50/0 ChemBioChem **2002**, 3, 490 – 504

Ways of Assembling Complex Natural Products on Modular Nonribosomal Peptide Synthetases**

Henning D. Mootz, Dirk Schwarzer, and Mohamed A. Marahiel*[a]

Nonribosomal peptide synthetases (NRPSs) catalyze the assembly of a large number of complex peptide natural products, many of which display therapeutically useful activity. Each cycle of chain extension is carried out by a dedicated module of the multifunctional enzymes. A module harbors all the catalytic units, which are referred to as domains, necessary for recognition, activation, covalent binding, and optionally modification of a single building block monomer, as well as for peptide-bond formation with the growing chain. A terminal domain releases the full-length peptide chain from the enzyme complex. Recent characterization of many NRPS systems revealed several examples where the sequence of the product does not directly correspond to the linear arrangement of modules and domains within the enzyme(s). It is now obvious that these systems cannot be regarded as rare exceptions of the common NRPS architecture but rather represent more complicated variations of the NRPS repertoire to increase their biosynthetic potential. In most of these cases unusual peptide structures of the products are observed, such as structures with side-chain acylation, cyclization involving the peptide backbone and/or side chains, and transfer of the peptide chain onto soluble small-molecule substrates. These findings indicate a previously unexpected higher versatility of the modules and domains in terms of both catalytic potential and interaction within the multifunctional protein templates. We propose to classify the known NRPS systems into three groups, linear NRPSs (type A), iterative NRPSs (type B), and nonlinear NRPSs (type C), according to their biosynthetic logic. Understanding the various biosynthetic strategies of NRPSs will be crucial to fully explore their potential for engineered combinatorial biosynthesis.

KEYWORDS:

antibiotics · combinatorial biosynthesis · nonribosomal peptide synthetases · peptides · polyketides

Introduction

Peptide antibiotics represent a large and diverse group of bioactive natural products with a wide range of applications in medicine, agriculture, and biochemical research. They can be classified into different groups according to their way of synthesis. The lantibiotics nisin and subtilin, for example, which contain the thioether amino acid lanthionine, are produced by posttranslational modification and proteolytic processing of ribosomally synthesized precursor peptides.[1] Other peptides are synthesized nonribosomally. The tripeptide glutathione[2] and the bacterial cell-wall precursor dipeptide DAla – DAla[3] are assembled by an enzyme-catalyzed phosphorylation of the carboxyl group of one amino acid to facilitate the nucleophilic attack of the amino group of the consecutive amino acid.

Yet another and very widespread method of peptide assembly is performed on nonribosomal peptide synthetases (NRPSs), which are the subject of this review. They produce secondary metabolite peptides, which can have very complex structures and many of which are used as important pharmaceuticals. Examples are the penicillin precursor ACV (**2**), the antibiotic vancomycin, the microtubule-stabilizing epothilone (**10**), and virulence-conferring siderophores such as enterobactin (**5**) that are secreted by microorganisms to capture iron ions. Producers of nonribosomal peptides are mostly soil-inhabiting microorganisms, such as members of the Gram-positive *Actinomycetes* and *Bacilli*, but also eukaryotic filamentous fungi. Marine microorgansims have also emerged as a rich source for such secondary metabolites.[4] Nonribosomal peptides are usually about 3 – 15 amino acids in length, with the upper limit possibly imposed by the enormous size of the NRPSs (see below). Their structural diversity (see Scheme 1 for examples) exceeds that of ribosomally synthesized peptide antibiotics. This is achieved by the incorporation of proteinogenic and "unnatural" amino acids as well as α-hydroxy and carboxylic acids. Further modification such as *N*-, *C*- and *O*-methylation, acylation, glycosylation, heterocyclic ring formation, and conversion of the building blocks into their stereoisomer can be conducted by the enzymes. The released peptides can be linear, cyclic, or branched-cyclic leading to macrocyclic lactams or lactones. The often high content of hydrophobic residues and the incorporation of long-chain fatty acids render the peptides able to pass or penetrate biological membranes. The peptide-based structures and enormous structural diversity confer the potential to bind to many biological targets. The unusual monomers and

[a] *Prof. Dr. M. A. Marahiel, Dr. H. D. Mootz,[+] Dipl.-Chem. D. Schwarzer*
Philipps-University of Marburg
Fachbereich Chemie/Biochemie
Hans-Meerwein-Strasse, 35032 Marburg (Germany)
Fax: (+49) 6421-282-2191
E-mail: Marahiel@chemie.uni-marburg.de

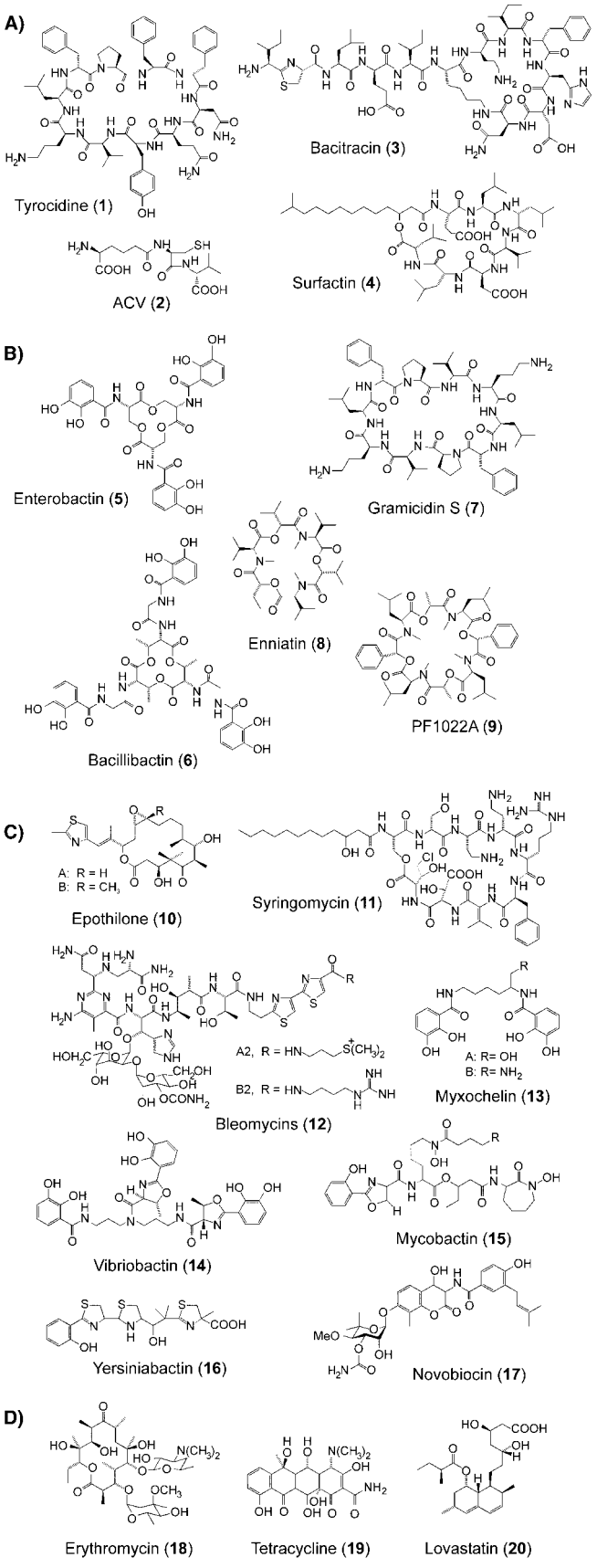
[+] *Current address:*
Laboratory of Synthetic Protein Chemistry
The Rockefeller University
1230 York Avenue, New York, NY 10021 (USA)

[**] A list of abbreviations can be found at the end of the text.

Henning D. Mootz studied chemistry at the Universität-Gesamthochschule Wuppertal (Germany), the University of Kent at Canterbury (UK), and the Philipps-Universität Marburg (Germany); he obtained his Diploma at the latter in 1996. During his PhD work in the group of Mohamed Marahiel he studied the genetics and biochemistry of nonribosomal peptide antibiotic synthesis. He also spent a year in the group of Philippe Marlière at the Institut Pasteur in Paris (France), where he was involved in expanding the genetic code for the incorporation of unnatural amino acids into proteins in vivo. After he obtained his PhD in 1999 from the Philipps-Universität Marburg where he was awarded the prize for the best thesis in the Chemistry Department, he continued as a postdoctoral fellow in Mohamed Marahiel's laboratories. Currently he is enjoying his postdoctoral research in Tom Muir's group at the Rockefeller University in New York (USA), where he is working on protein semisynthesis by using expressed protein ligation. He has received fellowships from the Studienstiftung des deutschen Volkes and the Stiftung Stipendien-Fonds des Verbandes der Chemischen Industrie.

Dirk Schwarzer was born in 1972 in Schleswig (Germany). He studied chemistry at the Philipps-Universität Marburg (Germany). In 1997 he joined the group of Mohamed A. Marahiel and received his Diploma in chemistry in 1998. In his PhD work he is focusing on the study of nonribosomal synthesis of peptide antibiotics and the possibilities for manipulating their biosynthetic pathways.

Mohamed A. Marahiel studied chemistry at the Universities of Cairo (Egypt) and Göttingen (Germany). In 1977 he obtained a PhD in biochemistry and microbiology from the University of Göttingen, were he carried out his research work at the Max Planck Institute for Experimental Medicine. Subsequently he received an assistant professor's position at the Technical University of Berlin, where in 1987 he obtained his Habilitation in biochemistry at the institute of biochemistry and molecular biology. Three years later he moved to the Philipps-Universität Marburg as a professor of biochemistry in the chemistry department. He was a Deutsche Forschungsgemeinschaft fellow in 1978 and 1986 at the John Innes Institute in Norwich (UK) and at the Biolabs, Harvard University (USA), respectively. His present research focuses on the structure – function relationship and on the elucidation of reaction mechanisms of modular peptide synthetases involved in the nonribosomal synthesis of peptide antibiotics, as well as on the rational design of recombinant enzymes for the synthesis of novel bioactive peptides. His group is also interested in studying the function and regulation of the major cold shock proteins in soil bacteria as well as other stress-induced proteins.



Scheme 1. *Products of NRPSs and PKSs. A) Products of linear NRPSs (type A); B) products of iterative NRPSs (type B); C) products of nonlinear NRPSs (type C) and mixed NRPS/PKS systems; D) products of PKSs.*

further chemical modifications are also beneficial for the stability of the products against proteolytic digest. In particular, all nonribosomal peptides contain, to our knowledge, unusual or at least modified amino acids at their N or C termini. Thus, the structures and the diversity of these compounds are preselected for stability and biological activity.

In contrast to the structural diversity of the products, their mode of synthesis, catalyzed by NRPSs, is very conserved.[5, 6] NRPSs are large multifunctional enzymes that can carry out up to several dozen reactions in a coordinated manner. One NRPS for a certain product can be a single polypeptide chain, as is the case for most fungal NRPSs (such as the 1.6 MDa cyclosporin synthetase), or can be organized into a few interacting subunits (each of which is also referred to as an NRPS), as is found for most bacterial systems (for example, the tyrocidine synthetase subunits TycA, B and C (see Figure 2 II). NRPSs are composed of modules, each capable of carrying out one cycle of chain extension. To become active they need to be posttranslationally modified to their holo form by a 4′-phosphopantetheinyl transferase,[7] generally at every module (see below). A key principle in nonribosomal peptide synthesis is the covalent attachment to the enzyme of all monomers and intermediates as thioesters to the enzyme-attached 4′-phosphopantetheine (Ppant) prosthetic groups until the final product is released (see below for the first exceptions to this rule).

A module is composed of domains that catalyze the single reaction steps like activation, covalent binding, optional modification of the incorporated monomer substrate, and condensation with the amino acyl or peptidyl group on the neighboring module. Peptide synthesis proceeds in an N- to C-terminal direction. In a few cases, modification during chain elongation can also be catalyzed by external enzymes. After release from the NRPS complex, the product can undergo further tailoring reactions by independent enzymes to achieve its mature form (see the recent review[8] on tailoring reactions for more details).

Manipulations of NRPSs by genetic engineering of the encoding genes holds the potential for the biosynthesis of useful variants of the well-known pharmaceuticals or even completely new products.[9, 10] A key issue in research on nonribosomal peptide synthesis is how the multifunctional NRPSs manage to specifically assemble these complex natural products from simple monomers. How are the instructions to these complex syntheses stored in the specificity of the single domains and their organization within the enzyme templates? The finding that modules as well as domains, the catalytic units of NRPSs, retain their activity when excised from the multifunctional enzymes supported the model of NRPSs being an array of specifically interacting semiautonomous catalytic units. A colinear relation between the sequence of the product and arrangement of the catalytic domains was confirmed for many biosynthetic systems and can be illustrated with the consecutive work-stations of an assembly-line. Consequently, exchange of domains or modules for counterparts with different substrate specificity resulted in productive NRPSs that synthesized predicted compounds.[9, 11–13] However, many recently discovered NRPSs were found to substantially deviate from the "classical" linear architecture; this makes it difficult or impossible to predict their biosynthetic logic without extensive biochemical characterization.

In this review, we propose a classification of NRPSs into three classes, linear NRPSs (type A), iterative NRPSs (type B), and nonlinear NRPSs (type C), and we will discuss their combinatorial potential. First of all, however, the core domains essential for every NRPS shall be introduced.

The Core Domains of NRPSs: Adenylation, Thiolation, and Condensation

Three essential enzymatic activities are needed for one complete elongation cycle, which is illustrated in Figure 1. They reside in the adenylation (A), thiolation (T; also referred to as the peptidyl carrier domain, PCP), and condensation (C) domains. First, the

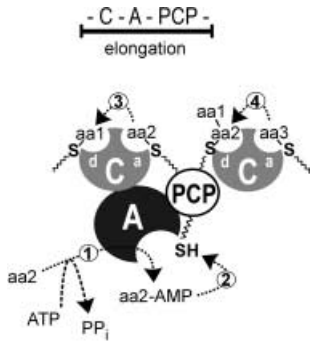


Figure 1. *Domain organization and reaction sequence of an NRPS elongation module. A minimal NRPS elongation module consists of a C domain, an A domain, and a PCP domain. The A domain selects a specific amino acid from the available pool of substrates and catalyzes formation of the amino acyl adenylate under consumption of ATP (reaction 1). The activated acyl moiety is subsequently transferred onto the thiol group of the Ppant prosthetic group of the PCP (reaction 2). Movement of the acyl-S-Ppant facilitates transport to the acceptor or nucleophile position (a) of the upstream C domain, where peptide bond formation with the amino acyl or peptidyl group of the preceeding PCP is catalyzed (reaction 3). Finally, the elongation cycle is completed at the donor or electrophile position (d) of the downstream C domain of the next module, where deacylation of the PCP by nucleophilic attack of the following amino acyl-S-Ppant results in translocation of the growing chain to the following module. The Ppant prosthetic group is represented by the zigzagged lines.*

A domain (about 550 aa in size) selects the cognate amino acid from the pool of available substrates and activates it as the aminoacyl adenylate (reaction 1 in Figure 1). This activation is analogous to the first reaction performed by aminoacyl-tRNA synthetases[14] in the ribosomal system. From the crystal structure of a representative A domain in complex with its substrates, the residues that form the pocket for substrate recognition could be identified[15] and generalized for A domains of diverse selectivity.[16] The analysis of the so-called signature sequences allows the reliable prediction of the cognate substrate(s) of newly identified NRPS sequences and a change of selectivity by site-directed mutagenesis, which might serve as a powerful tool to obtain new products.[16] In the following step, the activated amino acid is transferred onto the thiol moiety of the Ppant prosthetic group attached to the PCP domain (reaction 2 in Figure 1). The Ppant group of each PCP is

posttranslationally introduced onto the side-chain hydroxy group of a conserved serine residue within the PCPs by a cognate Ppant-transferase.[17, 18] Thus, the chemical energy of the phosphoester is used for the formation of the energy-rich thioester bond. The PCP (about 80 aa in size) with its 20 Å Ppant cofactor then serves as a swinging arm to transport the intermediates to the various catalytic centers. This covalent tether, which is the hallmark of the so-called multiple carrier thio template model of nonribosomal peptide synthesis[19] and which is also found in polyketide and fatty acid synthesis, allows "physical channeling"[20] through the multifunctional enzyme. C domains, which are about 450 aa in size, are localized between every consecutive pair of A domains and PCPs and catalyze the formation of the peptide bond between the upstream aminoacyl- or peptidyl-S-PCP moiety and the free amino group of the downstream aminoacyl-S-PCP, thus facilitating the translocation of the growing chain onto the next module (reactions 3 and 4 in Figure 1).

The directionality and ordered initiation of this synthesis is probably achieved by so-called donor and acceptor sites on the C domain for electrophiles and nucleophiles, respectively (illustrated as the two reaction centers on the C domain in Figure 1) The acceptor site binds the aminoacyl-S-PCP with a high affinity until the condensation process with the incoming aminoacyl- or peptidyl-S-PCP is completed.[19, 21, 22] This way, misinitiation at internal modules of the multimodular template is prevented. According to this model, internal elongation modules can be switched into initiation modules simply by deleting the upstream C domain of a C – A – PCP module.[22] Recent biochemical studies indeed revealed an editing function of the C domains at the acceptor site, which was found to discriminate, at least to a certain extent, against amino acids of opposite stereochemistry and with larger side chains, whereas the donor site is more tolerant.[23, 24] Such a specialization for the nucleophile is in agreement with the postulated binding site.

The Thioesterase Domain Can Act as a Macrocyclase

In most NRPS assembly lines a thioesterase-like (Te) domain of about 250 aa is found at the C-terminal end of the last module. After transfer of the linear peptidyl intermediate from the last PCP onto the catalytic serine residue of the Te domain (serine is replaced by cysteine in a few examples), this domain catalyzes release of the product by hydrolysis, cyclization, or oligomerization. To directly assess the catalytic properties of excised Te domains, peptidyl-S-NACs (peptide *N*-acetylcysteamine thioesters) were used as soluble molecule substitutes for the natural peptidyl-S-PCP substrates. The 28 kDa Te domain from tyrocidine NRPS catalyzed the formation of the cyclic antibiotic tyrocidine A from a decapeptidyl-*S*-NAC that corresponded to the tyrocidine sequence.[25] Remarkably, only very few of the ten residues located at the N- and near the C-termini of the substrate decapeptide were found to be important for efficient cyclization. These residues are likely to be involved through hydrogen bonds in a preorganisation of the linear peptide into the antiparallel β-sheet structure that provides the right orientation of the N- and C-termini for ring closure.[26] Other residues in the S-NAC substrates can be substituted, deleted, or even replaced by spacers or insertions of several amino acids leading to an increase in ring size, as was demonstrated by efficient cyclization of various substrate analogues.[25–27] A recent review was dedicated to the different termination reactions catalyzed by NRPSs.[28]

Linear NRPS (Type A)

Examples for linear NRPSs are the tyrocidine (**1**),[29] bacitracin (**3**),[30] surfactin (**4**),[31] actinomycin,[32, 33] ACV (**2**; the penicillin and cephalosporin precursor),[34] cyclosporin,[35] pristinamycin,[36, 37] fengycin,[38, 39] and ergotamine[40] NRPSs and the glycopeptide antibiotics of the vancomycin family, chloroeremomycin,[41] balhimyicin,[42] and complestatin.[43] Figure 2 illustrates linear NRPSs in the example of the biosynthesis of the tripeptide ACV and the module organization of the tyrocidine and surfactin NRPSs. In linear NRPSs, the three core domains are arranged in the order C – A – PCP in an elongation module that adds one amino acid to the growing chain. The first amino acid of the peptide chain is incorporated by the initiation module which lacks a C domain. On the other hand, the terminal module in most cases contains a Te domain to release the full-length peptide chain from the enzyme.[28] In NRPSs of fungal origin the Te domain is often replaced by a specialized C domain, which is believed to catalyze the cyclization reaction.[28] In linear NRPSs, the sequence of the resulting linear peptide chain is entirely determined by the number and order of the modules, in contrast to nonlinear NRPSs, where the domain interplay can be more complicated, as will be shown below. Thus, a typical linear NRPS protein template for a peptide of *n* amino acids consists of *n* modules with the domain organization A – PCP – $(C – A – PCP)_{n-1}$ – Te. Modification domains are inserted into the respective module.

Strategies to Construct Hybrid Enzymes from Linear NRPSs

The arrangement of modules in linear NRPSs suggests a straightforward strategy to create novel products with predicted amino acid changes by domain or module swapping at the respective positions. Several such examples have been reported, both in vivo with the manipulated producer strain secreting the new peptide[9, 11, 44] and in vitro with purified hybrid proteins,[12, 13, 33] to allow detailed studies of turnover rates and possible by-product formation. Repositioning the Te domain from the terminal modules to internal ones resulted in the premature hydrolysis of the correspondingly shortened peptides.[44, 45] Te domains from heterologous NRPSs with a desired cyclization activity could be used to control regioselectivity of the termination reaction.[45] Lichenysin and surfactin (**4**) are structurally almost identical lipoheptapeptides that are used as biosurfactants, but the more powerful lichenysin is produced in only small amounts by *Bacillus licheniformis*. Whereas surfactin (**4**) has a Glu residue at position 1 of the peptide portion, a Gln residue is found in lichenysin. Exchange of the Glu-incorporating

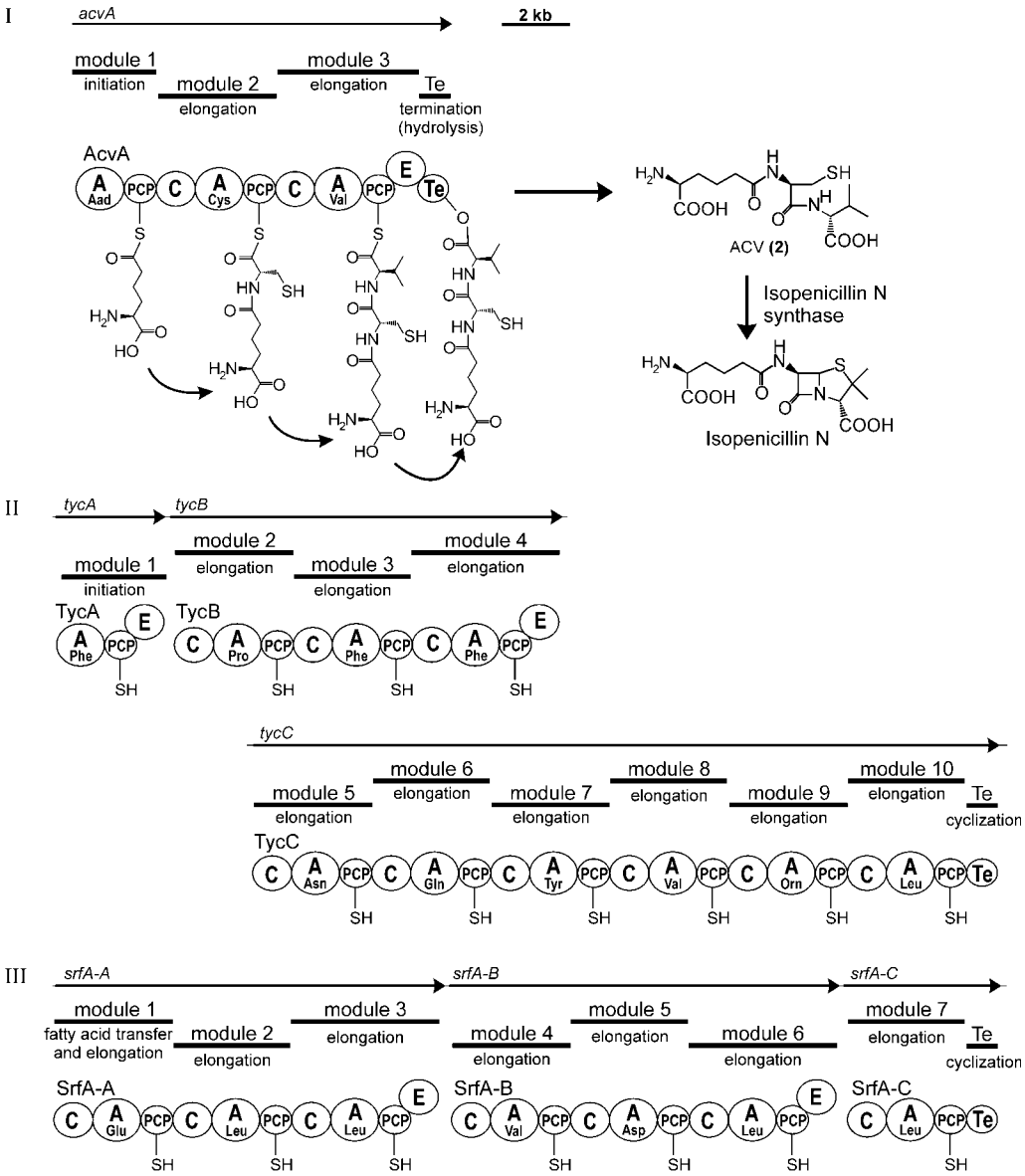


Figure 2. *Organization and mode of synthesis of linear NRPSs (type A). The ACV (I), tyrocidine (II), and surfactin (III) NRPSs are typical examples for linear NRPSs. Peptides of* n *amino acids are assembled on NRPSs with the module organization A – PCP – (C – A – PCP)$_{n-1}$ – Te, thus the number and order of modules determines the peptide sequence. ACV synthetase contains all modules on one polypeptide chain, whereas the tyrocidine and surfactin synthetases are composed of three interacting subunits. The Te domain can catalyze peptide release by hydrolysis, macrolactamization, or macrolactonization. Modules can contain optional domains for the modification of the corresponding amino acid, for example, module 3 of the ACV NRPS and modules 1 and 4 of the tyrocidine NRPS. The N-terminal C domain of module 1 of the surfactin NRPS (III) is responsible for transfer of a fatty acid to the glutamate residue bound at this module. The Ppant prosthetic group of each module is simplified as a single line.*

module (see Figure 2 III) of the surfactin NRPS of *B. subtilis* with the Gln-activating module of the lichenysin NRPS of *B. licheniformis* gave rise to a switch of *B. subtilis* into a high-level lichenysin producer.[11] Thus, engineering a heterologous strain that produces the desired compound in a better yield is an interesting application. In this work, a site on the gene level that corresponds to a highly conserved sequence motif within C domains was chosen for the module exchange. Although a very productive hybrid NRPS was obtained (the recombinant strain produced about 90 % lichenysin compared to the amount of surfactin produced by the parent strain[11]), this approach might be less efficient when applied to heterologous NRPSs (the surfactin and lichenysin NRPSs represent a special case because of their very high sequence identity and because modules were exchanged only at equivalent positions).

It is the prevailing opinion that surgery in linker regions between the independently folding domains without pertubing their integrity is in general the best way to obtain hybrid enzymes. Linker regions, which usually show only little or no conservation and are rich in small and hydrophilic residues, are common in multidomain enzymes and can be identified by limited proteolysis experiments or multiple sequence alignments.[46] Two papers have reported the construction of small hybrid NRPSs by whole-module fusion in the linker regions

between the A domain and the PCP[13] (defining a module as a (PCP – C – A) set of domains), as well as between the PCP and the C domain[12] (defining a module as a (C – A – PCP) set). In both cases the predicted peptides were synthesized by the purified enzymes with turnover rates comparable to those of the wild-type systems. By leaving entire modules intact, these examples also took into account recent results pointing to an editing function at the acceptor position of the C domains (see above).[23, 24] The implication of the latter work for the design of efficient hybrid NRPSs is to avoid fusions between consecutive C and A domains, which are potentially incompatible, but rather to fuse (C – A – PCP) or (PCP – C – A) modules. Interesting in this respect is a work on the construction of a hybrid enzyme derived from the actinomycin NRPS that disregarded this guideline and nevertheless successfully fused between the C and A domains. The A domain and the PCP responsible for valine incorporation were exchanged against a three-domain unit comprising the A domain and the PCP, also valine-specific, and an *N*-methylation (M) domain, which is inserted in the C-terminal end of the A domains.[33] Interestingly, the upstream C domain was also able to accept *N*-methylvaline as the nucleophile for formation of the peptide bond; this indicates a tolerance for this modification. However, the catalytic constants to assess the efficiency of the hybrid enzyme were not determined in this study.

To further refine the rules and explore the limits for the design of artificial NRPSs many domain combinations in hybrid contexts remain to be tested, in particular those involving the numerous modification domains that contribute substantially to the bioactivity of the products. For example, it has been noted that PCPs from modules containing an epimerization domain differ significantly in their primary structure from standard PCPs normally connected to a C domain;[7, 47] this probably reflects specialized communication with this modification domain and suggests possible incompatibilities if exchanged with modules of the other kind.[48] Another important aspect for the full exploration of the biosynthetic potential of NRPSs by combinatorial approaches is to understand the interaction between modules on different subunits of a synthetase complex, that is, interpolypeptide interaction between modules (in trans) as opposed to the intrapolypeptide interaction (in cis) as discussed above. What governs, for example, the ordered interaction of subunits in NRPS complexes comprising more than 2 subunits? Why does SrfA-A donate the intermediate peptide chain only to SrfA-B and not to SrfA-C (see Figure 2 III)? The control and predictable manipulation of these interactions would provide a great combinatorial potential. Recent work on mechanistically and architecturally related multidomain enzymes, the polyketide synthases (PKSs) that assemble another large class of natural products, revealed that short linkers at the N and C termini of the enzymes are necessary and sufficient to mediate the observed protein – protein interactions.[20, 49–51] Exchange of these interpolypeptide linkers can redirect the product flux between the subunits. Comparable linkers can also be identified in NRPSs and NRPS/PKS hybrids,[52] although they are generally shorter and less conserved. It will be an important issue to clarify their role in nonribosomal peptide synthesis.

Iterative NRPSs (Type B)

Iterative NRPSs use their modules or domains more than once in the assembly of one single product. This strategy is employed to build up peptide chains that consist of repeated smaller sequences. For example, the depsipeptide antibiotic enniatin (**8**) of the filamentous fungus *Fusarium scirpi* is a cyclic trimer of the dipeptidol D-2-hydroxyisovaleryl – *N*-methylvaline (D-Hiv – MeVal).[53] The iron-chelating siderophore enterobactin (**5**) produced by *Escherichia coli* is a cyclic trimer of dihydroxybenzoyl-serine units.[54] Structurally very similar is bacillibactin (**6**) of *B. subtilis*, which is a trimerized dihydroxybenzoyl-glycyl – threonine peptide.[55] Gramicidin S (**7**) is assembled by head-to-tail condensation of two identical pentapeptide halves.[27] The NRPSs involved in the synthesis of these compounds contain a reduced number of modules that corresponds to only one set of the repeated sequences, which is then assembled to form the final multimeric product on the terminal PCP and/or Te domain. Thus, the key to the oligomerization process is this latter domain, as illustrated in Figure 3 for enterobactin (**5**). In the synthesis of enterobactin (**5**) and gramicidin S (**7**), the monomer chains are transferred onto the active site serine residue of the C terminal Te domain, thereby deacylating the last PCP so that the next monomer chain can be assembled.[56] Correct oligomerization of the chains takes place on the Te domain until the final product is

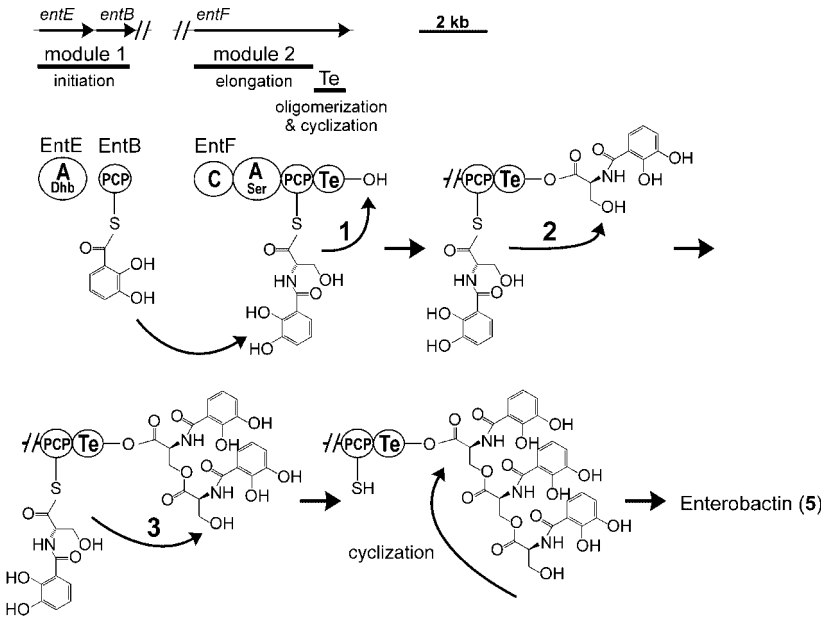


Figure 3. *Enterobactin NRPS as example for iterative NRPS (type B). In iterative NRPSs, modules or domains are used more than once for the assembly of one product molecule. For example, three Dhb-Ser-S-Ppant intermediates are generated on the two modules of the enterobactin NRPS and are oligomerized and cyclized on the Te domain.*

released, usually through cyclization (see Figure 3). In the fungal enniatin NRPS, the Te domain is replaced by a PCP and a C domain. It is believed that the oligomerization intermediates are stalled on this additional PCP with the C domain catalyzing both the transfer and final cyclization.[28] In both examples, Te domain and PCP – C didomain, these domains must control both the number of the iterative cycles and the chemical nature of the bonds between the monomer chains. In enterobactin (**5**) and bacillibactin (**6**), the linkage occurs as ester bonds through the side-chain hydroxy group of the serine and threonine residues, respectively, whereas in gramicidin S the head-to-tail dimerization yields peptide bonds, and in the synthesis of enniatin the hydroxy groups of the α-hydroxyacid D-Hiv form ester bonds with the C-terminal-activated *N*-methyl – valyl residues.

The *N*-methylated cyclooctadepsipeptide PF1022A (**9**), which exhibits anthelmintic properties, is synthesized by an NRPS that is probably of similar domain architecture to the NRPS of the closely related *N*-methylated cyclohexadepsipeptide enniatin.[57] PF1022A (**9**) is assembled by successive condensations of dipeptidol building blocks of D-lactate – MeLeu and D-phenyllactate – MeLeu. Interestingly, by using the purified protein (the encoding gene has not yet been cloned) in assays for in vitro synthesis of PF1022A (**9**),[57] the di-, tetra-, and hexahomologues of PF1022A can be detected as by-products; this indicates partial premature cyclization during the iterative assembly.

Strikingly, we are not yet able to distinguish iterative from linear NRPSs by analysis of the primary sequence. The biochemical knowledge of the enterobactin and gramicidin S NRPSs suggests that all determinants for the multimerization process reside in the terminal Te domain, however, even here multiple sequence alignments with other Te domains point to no clue allowing the prediction of such an iterative function. From the present knowledge, the discussed iterative NRPSs can be regarded as linear NRPSs with an iterative activity at the terminal PCP (Te domain or PCP – C didomain). Thus, if iterativity is desired in artificial hybrid NRPSs, this/these domain(s) must be translocated. We expect, however, that oligomerization/cyclization is a highly specialized process which in hybrid NRPSs will be restricted to monomer peptides that are very similar to the natural substrate in terms of size and sequence of the monomer as well as the nature of the internal nucleophile.

Other examples of iterative NRPSs are those for the biosynthesis of the siderophores yersiniabactin and vibriobactin. These will be discussed with the nonlinear NRPSs in the next section, as they also show an unusual domain organization.

Nonlinear NRPSs (Type C)

Sequencing efforts in the last years have revealed many biosynthetic clusters that deviate in their domain organization from the standard $(C-A-PCP)_n$ architecture of linear NRPSs. These were first regarded as rare exceptions of the colinearity rule.[58, 59] In the meantime it has become clear that most of the first characterized NRPS assembly lines were of the linear type only by chance and that deviations of this type are not an exception but constitute a considerable fraction of the NRPS repertoire that can be found in nature. The hallmark of these systems is at least one unusual arrangement of the core domains C, A, and PCP. Recently described NRPSs that fall into this group are the syringomycin (**11**), yersiniabactin (**16**), vibriobactin (**14**), bleomycin (**12**), and mycobactin (**15**) systems (see Scheme 1 for structures). A deviation from the module arrangement $(C-A-PCP)_n$ of linear NRPSs often goes along with nonlinear peptide products that result from unusual internal cyclizations (for example, bleomycin (**12**)) or branch-point syntheses (for example, vibriobactin (**14**) or mycobactin (**15**)). Detailed biochemical studies are needed to understand the function and the interplay within these enzymes. Another difference to linear NRPSs can be the use of small soluble molecules to be incorporated into the nonribosomally assembled peptide, such as the amines in vibriobactin (**14**) and bleomycin (**12**) synthesis. Since these amines lack a carboxyl group, the prerequisite for a covalent attachment to the enzyme as thioester, their incorporation cannot obey the classic multiple-carrier thiotemplate model in which all substrates and intermediates are covalently bound to the enzyme. Specialized C domains with specificity for peptidyl-S-PCP electrophiles at the donor position and free small molecules as nucleophiles at the acceptor position catalyze these reactions.[60, 61]

Many other biosynthetic clusters encoding putative nonlinear NRPSs whose products are not known can be found in the databases (for example, accession numbers U46488, T50176, and U85909 submitted to the Genbank/EBI Database). Due to their unusual (putative) domain organizations we can at best only very vaguely predict their possible products. Furthermore, inactive domains can complicate predictions based on the primary structures of the enzymes.

We propose to classify all such systems as nonlinear NRPSs (type C), until progress in understanding the underlying biosynthetic logic may call for more precise distinctions.

In syringomycin (**11**) biosynthesis, the two peptide synthetases SyrB and SyrE have been identified.[58, 62] SyrB consists of an A – PCP didomain, whereas SyrE comprises the remaining eight (C – A – PCP) modules and the terminal Te domain for the assembly of the lipononapetide (see Figure 4 I). However, biochemical analysis of the first two A domains of SyrE and the A domain of SyrB clearly showed that SyrE must synthesize the N-terminal lipooctapeptide and SyrB must add L-Thr as the precursor of the ninth amino acid. The difference to linear NRPSs is localized near the C-terminal end of SyrE: between the eighth module and the terminal Te domain a C domain and a PCP are inserted, giving rise to the unsual domain order $(C-A-PCP)_8 - C-PCP_{9a}-Te$ (see Figure 4). As shown in Figure 4 II, it was proposed that SyrB delivers the activated amino acid L-Thr onto the additional PCP_{9a} in trans and that the additional C domain extends the octapeptide to the nonapeptide on PCP_{9a} before release is catalyzed by the Te domain.[58] In this scenario, which awaits biochemical proof, the role of the PCP_{9b} of SyrB remains to be explained.

In the biosynthesis of the siderophore yersiniabactin (**16**; see Figure 5) from *Yersinia* sp. an even more impressive example of

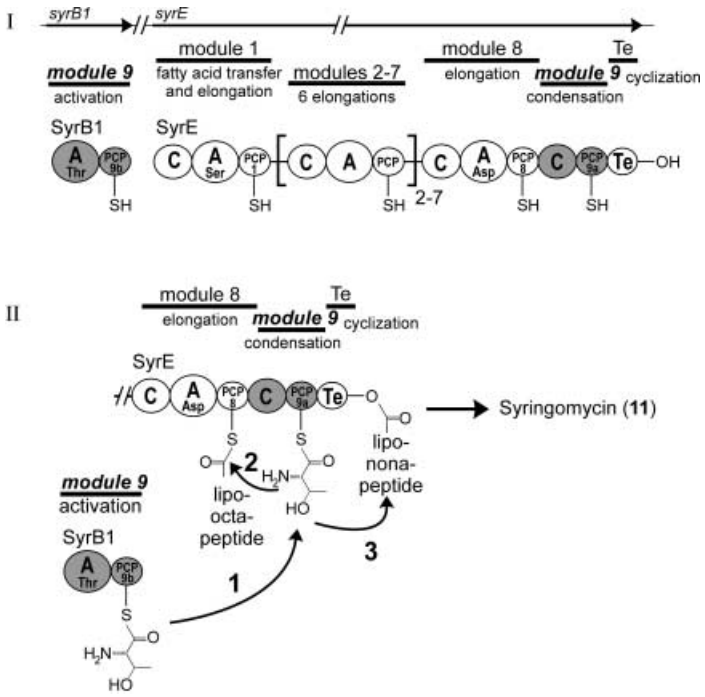


Figure 4. *Syringomycin NRPS is a nonlinear NRPS (type C). The biosynthesis cluster for the lipononapeptide syringomycin encodes the two NRPSs SyrB1 and SyrE (I). SyrB1 recognizes threonine, the ninth amino acid of the syringomycin sequence, and is proposed to complement the incomplete module 9 of SyrE by in trans acylation of PCP9a. It is not clear if PCP9b of SyrB1 is needed for this reaction.*

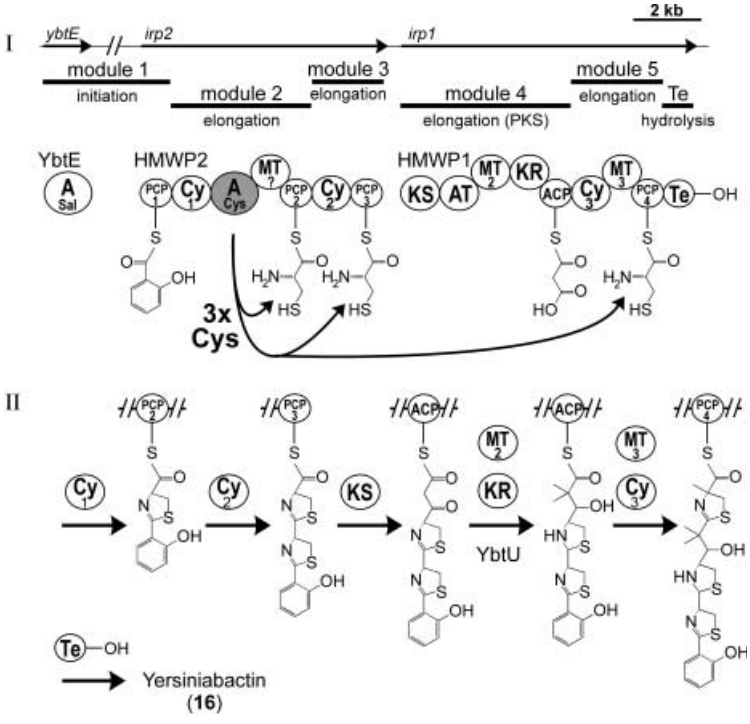


Figure 5. *Yersiniabactin NRPS is a nonlinear NRPS (type C). Yersiniabactin biosynthesis proceeds on a mixed NRPS/PKS system. Panel I shows the binding of the monomer acyl building blocks on the carrier protein domains. The cysteine-specific A domain of HMWP2 acylates three different PCPs on two enzymes. Panel II shows the proposed assembly of the siderophore with the domains that are involved in the single steps highlighted. YbtU is an external NADPH-dependent reductase.*

in trans aminoacylation can be found.[63] The A domain of HMWP2 is specific for cysteine and loads three PCPs with this amino acid:[64–66] one is located in the same unit as the A domain, whereas the second is at unusual distance on the same enzyme, and the third is actually on another enzyme of the complex, HMWP1 (Figure 5 I). This example also highlights the remarkable architectural flexibility that the yersiniabactin NRPS must be able to adopt, beyond the usual ability of a PCP, to interact with an A domain, upstream and downstream C domains, and optionally with modification domains.[67] The necessary spatial proximity of the active sites of the A domain and the three PCP domains could be achieved either by a highly complex three-dimensional arrangement of the domains of yersiniabactin NRPS and/or by large conformational changes. Yersiniabactin and the structurally related siderophore pyochelin[68] are also good examples for tailoring steps carried out by distinct modification enzymes while the growing chain is still attached to the NRPS complex. The NADPH-dependent reductase PchG reduces one of the thiazoline rings to thiazolidine. In yersiniabactin biosynthesis, the homologous YbtU is believed to catalyze the equivalent reaction[69] (see Figure 5 II).

Whereas in yersiniabactin biosynthesis one A domain loads three different PCPs, a C domain is proposed to catalyze formation of two amide bonds during the assembly of the siderophore myxochelin (**13**; see Scheme 1 for structure) from *Stigmatella aurantiaca*.[70] This pathway is illustrated in Figure 6. Obviously, the C domain of the four domain NRPS MxcG (with the domain organisation C – A – PCP – R, where R is a reductase domain) is responsible for acylation of both the α- and ε-side-chain amino groups of the activated lysine residue with dihydroxybenzoyl groups, which are transferred from the PCP of MxcF. According to this hypothesis,[70] the integrated C domain of MxcG must process these two different nucleophiles at its acceptor position (Figure 6). No other C domains have been found in the myxochelin gene cluster. In the subsequent course of the biosynthesis, the bisacylated lysyl-S-PCP intermediate is reduced to the free aldehyde by the R domain of MxcG and is then probably further reduced by an oxidoreductase to myxochelin A or converted by an aldehyde amino-transferase into myxochelin B. Genes encoding candidate enzymes for these reactions are located in the cluster.[70]

The biosynthesis of the siderophore vibriobactin (**14**)[60, 61, 71, 72] is a particularly interesting example as it shows for the first time the use of free soluble molecules to be incorporated into the product without prior binding to the enzyme as a thioester. As shown in Figure 7, the virulence-conferring molecule from *Vibrio cholerae* consists of the triamine norspermidine, with a dihydroxybenzoyl (Dhb) moiety bound to one of the terminal amines and two dihydroxyphenyl oxazoline-carbonyl (Dhp-Oxa) moieties bound to the remaining two amines. Norspermidine has no carboxyl group which could be used for an activation as acyl adenylate. Thus, an A domain for its incorporation is superfluous, instead its entry into the vibriobactin biosynthesis is

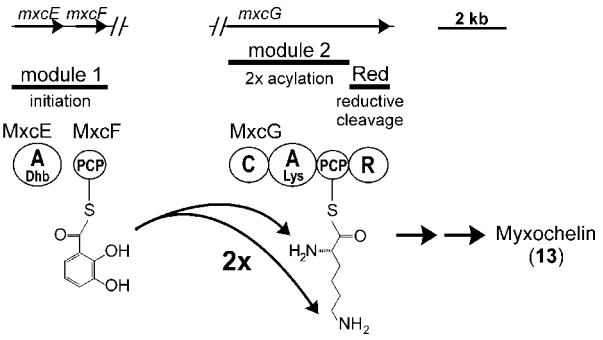


Figure 6. *Myxochelin NRPS is a nonlinear NRPS (type C). The proposed biosynthesis of the siderophore myxochelin involves two acylations of the lysine residue bound to MxcG with Dhb provided by MxcE and MxcF. The C domain of MxcG would carry out both reactions with the α-amino group and the ε-side-chain amino moiety of Lys-S-Pant-MxcG serving as nucleophiles.*

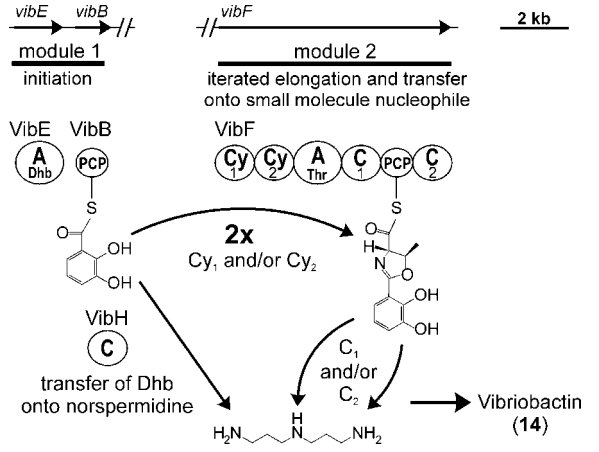


Figure 7. *Vibriobactin NRPS is a nonlinear NRPS (type C). The siderophore vibriobactin is synthesized from one molecule of norspermidine, two molecules of threonine, and three molecules of Dhb. The vibriobactin NRPS is the first biochemically characterized example for the use of free, soluble, small molecules as nucleophiles in nonribosomal peptide synthesis. The stand-alone C domain VibH transfers the first Dhb from Dhb-S-Ppant-VibB onto the triamine norspermidine. The six-domain NRPS VibF acylates the two remaining amine groups with dihydroxyphenyl oxazolinecarbonyl (Dhp-Oxa), which is derived from condensation of Dhb with threonine. It is not yet clear if both or just one of the two Cy and C domains of VibF are involved in these reactions.*

achieved by a stand-alone C domain, VibH, which catalyzes the transfer of a Dhb from VibB onto one of the terminal amines.[60] The two acylations with Dhb-Oxa to yield the final product are performed by VibF[61] (see Figure 7). The Michaelis constant (K_m) value for norspermidine in this reaction was determined to be 1.5 mM whereas those for the mono- and bisacylated intermediates were 1.7 and 25 μM, respectively; this indicates the high-affinity sequestering of the reaction intermediates.[60, 61] There are two striking features about this six-domain enzyme, which can be seen as both an iterative and a nonlinear NRPS. Whereas it contains only one A domain and one PCP to build up two Dhb-threonyl-S-PCP for each product formed (after transfer of Dhb from VibB), there are two Cy domains (for heterocyclic ring formation) and two C domains to give a domain arrangement Cy – Cy – A – C – PCP – C. It is believed that each of the C domains catalyzes transfer of one Dhb-Oxa moiety onto the norspermidine acceptor.[61] This is reasonable since recognition of the middle and the terminal amine might require a dedicated catalyst. However, the reason for the presence of a second Cy domain remains obscure and the question of whether both Cy domains are active must be addressed with mutational studies.[61]

The antitumor drug bleomycin (**12**) shares with vibriobactin the acylation of an amine lacking a carboxyl group for thioester activation, although the attached peptide chain is much larger in this case. Bleomycins with two different terminal amines are known, and two stand-alone C domains can be identified in the sequenced 85 Kb of the biosynthesis cluster[73] (see Figure 8 I). Since the last NRPS modules also lack a Te domain, the knowledge of vibriobactin synthesis would postulate a biosynthetic logic similar to that of VibH of the vibriobactin NRPS:[60] the fully assembled peptide chain would be transferred onto the amine(s) by one or both of the stand-alone C domains that are specific for these small molecule nucleophiles. However, bleomycin also contains other structural elements that are unprecedented in nonribosomal peptide synthesis and for which assumptions on the nature and order within the NRPS of the necessary domains can only be speculative.[73] First, the first amino acid residue is believed to be derived from dehydroalanine, which is then bound to the amino group of Asn 2 by conjugative addition and subsequent aminolysis. The formation of the pyrimidine ring requires cyclization with the side-chain moiety of Asn 3 followed by transamination and oxidation.[73] An aminoacyl ligase at the N terminus of BlmVI was proposed for the aminolysis reaction and a surplus C domain (C′) at the C terminus of BlmV was suggested as candidate for the cyclization reaction.[73] However, one of the two stand-alone C domains could also be involved (see Figure 8 I). The enzymatic activities for transamination and oxidation are probably not integrated in the NRPS. A distinct PCP, encoded by *blmI*[74] (Figure 8 I), might also come into play in a yet unknown fashion. The remaining chain extension would proceed in a relatively standard manner through the NRPS and PKS modules,[73] except for the above-mentioned unusual termination by transfer onto amine acceptors. Therefore bleomycin is a good example about how unusual structural features go along with unusual domain arrangements whose logic is not yet decipherable for us and will require detailed biochemical characterization or mutational studies.

Another very complex nonlinear NRPS is involved in the biosynthesis pathway for the siderophore mycobactin (**15**) of *Mycobacterium tuberculosis*[75] (see Scheme 1 for structure). Although mycobactin (**15**) is on first sight a relatively simple molecule, its assembly is not completely understood and is very difficult to predict from the biosynthesis genes. Nonlinear peptide structures again parallel nonlinear organization of domains in the mixed mycobactin NRPS/PKS hybrid (Figure 8 II). The terminal lysine residue undergoes intramolecular lactamization, which probably occurs in the course of releasing the product from the enzyme complex in a reaction that might be catalyzed by the terminal domain with homologies to the E and C domains.[75] Furthermore, the PKS-derived 3-hydroxybutyrate unit is linked through the hydroxy group as an ester with the first peptide-like half of mycobactin, although PKSs usually catalyze

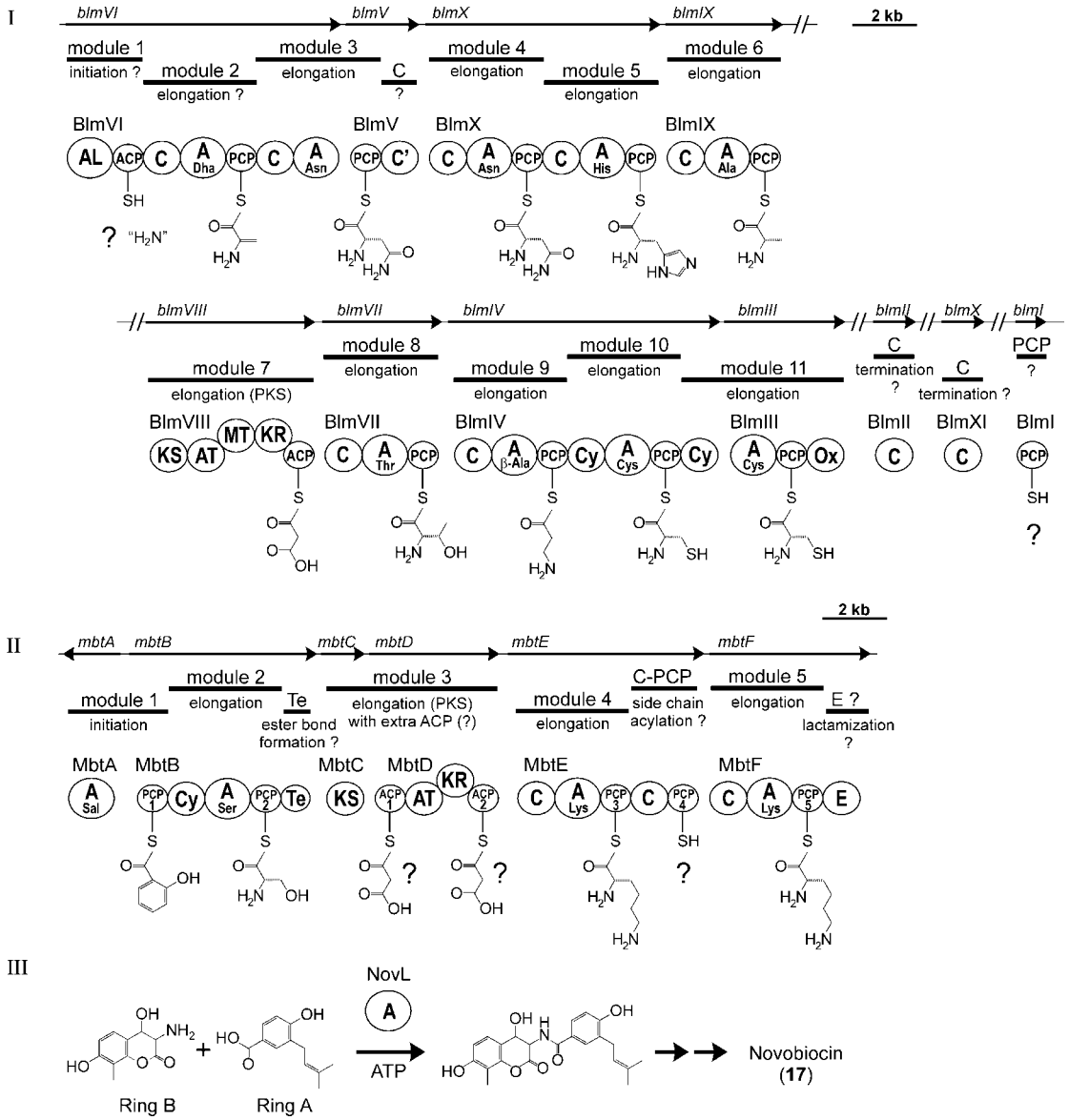


Figure 8. *The bleomycin and mycobactin NRPSs are nonlinear NRPSs (type C). Panels I and II show the NRPSs genes and enzymes of the biosynthetic pathways for bleomycin and mycobactin, respectively, which are only poorly understood. The monomer substrates are drawn to each module as suggested by the authors. An A domain forms a peptide bond in the biosynthesis of Novobiocin. Panel III illustrates the reaction carried out by NovL.*

Claisen-like C–C bond formations. This would mean that mycobactin is assembled from two halves whose synthesis may start independently. The ester bond might be formed by the acyltransferase domain at the end of MbtB.[75] Reminiscent of the syringomycin (**11**) NRPS (see above), insertion of a C–PCP two-domain unit is found between modules 4 and 5 in MbtE (see Figure 8 II). In analogy to the proposed biosynthesis of syringomycin (**11**), it can be hypothesized that this two-domain unit of MbtE might be loaded in trans with a fatty acid by an as yet unidentified factor, for example an acyl-CoA ligase-like enzyme, which is then transferred to the lysine residue side chain. If this hypothesis is true, assembly of mycobactin would proceed in a branch-point synthesis fashion on different parts of the molecule, catalyzed by domains that are integrated within the same NRPS. Detailed biochemical characterization of this interesting pathway would be necessary to unravel the domain interplay. Unfortunately, the production of recombinant mycobactin NRPSs for such studies was hindered by expression and solubility problems in *E. coli*.[75]

Finally, the enzyme NovL involved in the biosynthesis of the antibiotic novobiocin (**17**) from *Streptomyces spheroides*[76, 77] is an example for an unusual peptide antibiotic biosynthesis pathway that does not follow any of the discussed principles. NovL catalyzes formation of the peptide bond between 3-dimethylallyl-4-hydroxybenzoic acid (ring A, see Figure 8 III) and 3-amino-4,7-dihydroxy-8-methyl coumarin (ring B).[77] Strikingly, NovL is a one-domain enzyme only, with homology to acyl-adenylate forming enzymes of the same superfamily as A domains. It activates the carboxy acid of ring A towards the acyl adenylate. Instead of the usual transfer onto a PCP, however, the

acyl adenylate then directly acts as the electrophile for the condensation with amino group nucleophile of ring B. Apparent K_m values were 19 μM for ring A and 131 μM for ring B.[77] Thus, an A domain (or an A-domain-like enzyme) can also use a diffusable small-molecule nucleophile and bypass a PCP. In fact, covalent capture of the highly reactive aminoacyl-adenylate intermediate with diffusable small molecules has been observed as a side reaction for other NRPSs.[64, 78] NovL is an example for the efficient exploitation of this kind of reaction, which can obviously be encountered in many biosynthetic pathways requiring only one condensation step between two precursors of the final product.[79] It can also be compared with the reaction catalyzed by acyl-CoA ligases which belong to the same superfamily as the A domains. It remains a matter of definition as to whether the one-domain NovL should be counted in the family of NRPSs, whose characteristics are the interplay of several domains and the covalent linkage of (at least most of) the biosynthesis intermediates. The potential use of enzymes of the NovL type to create more general peptide bond forming catalysts by in vitro evolution is, however, remarkable.

Comparison with PKS, FAS, and NRPS/PKS Hybrid Systems

It is very helpful to compare NRPSs with PKSs, since both classes of multifunctional enzymes employ a similar synthetic logic in utilizing Ppant-dependent carrier proteins to bind the monomers and the growing chain to the enzyme complex. Naturally occuring NRPS/PKS hybrids underline this close relationship. Thus, insights gained on one system may stimulate research in the other field as well. Instead of amino acids, polyketides are built of acetate and propionate monomers. Malonyl- and methylmalonyl-CoA are the activated precursors used for the extension steps, which proceeds like fatty acid synthesis. In addition, structural diversity of the ketide backbone is achieved by selective reduction of the β-keto group after each condensation step. Other parameters are chain length, macrocyclization and further tailoring through hydroxylation, methylation, or glycosylation (see recent reviews on comparisons between NRPSs and PKSs[52, 80, 81] for a more detailed description). According to Kennedy et al.[82], PKSs can be subdivided into three large groups: modular (type I) and iterative PKSs (type II) are of bacterial origin, iterative type I PKSs are of fungal origin.

The modular PKSs of type I (examples are those for the biosynthesis of erythromycin (**18**; see Scheme 1 for structure)[83] and tylosin[84]) correspond in architecture to the linear NRPSs. Modules and domains are arranged in a linear order according to the stepwise assembly of the ketide chain. A thioesterase domain at the last module cleaves the product by macrocyclization. Like linear NRPSs, modular PKSs are manipulatable in a predictive manner by module or domain swapping, domain insertion, or domain inactivation.[6, 85] Attempts to generate molecular diversity by genetic engineering of linear NRPS-encoding genes may profit from two powerful approaches recently demonstrated for modular type I PKSs. First, as mentioned above, the identification of N- and C-terminal linkers that control the interpolypeptide interaction between modules on different enzymes raises the question of whether similar linkers are operative in NRPSs.[52] Second, a multiplasmid approach was reported for PKSs to produce libraries of variants of the lead compound erythromycin (**18**).[86] To this end, a production strain was transformed with three plasmids, each harboring one of the three genes encoding the PKSs responsible for the synthesis of the erythromycin aglycone precursor. After various domain-swapping experiments in each of the genes, the modified plasmids were used for transformation in all possible combinations (for example, 4 plasmids for each gene give $4 \times 4 \times 4 = 64$ possible combinations). This strategy would be conceivable for all NRPS systems comprising two or more subunits, such as the surfactin, tyrocidine, and bacitracin NRPSs, to produce a large number of variants.

The most convincing argument for the similarity of modular PKSs (type I) and linear NRPSs is the existence of mixed biosynthetic clusters of the two types in nature. In these complexes, NRPS and PKS modules can act together either in trans on distinct enzymes or even in cis connected with each other on one polypeptide chain.[52, 81] Examples of the mixed peptide – polyketide products are the important agents bleomycin (**12**),[73] epothilone (**10**),[87, 88] myxothiazol,[89] microcystin,[90, 91] and rapamycin.[92] (The lipopeptide mycosubtilin is assembled on a hybrid NRPS/fatty acis synthase (FAS) complex[93]). Artificial generation of hybrid NRPS/PKS enzymes would tremendously increase our potential for combinatorial biosynthesis.

In iterative PKSs of type II, the catalytic domains (acyl carrier protein (ACP), ketosynthase, acyltransferase, etc.) are used iteratively to build up a ketide chain that is subsequently cyclized by additional cyclases and aromatases. Tetracycline (**19**) and daunorubicin are examples of the often aromatic products that synthesized in this fashion.[94] The factors that control the chain length are not completely understood. In contrast to the modular PKSs of type I, but in analogy to bacterial FASs, all domains and additional enzymatic activities are localized on distinct polypetide chains. Exchange of enzymes between different systems can lead to new products.[95] There are also examples of distinct NRPS domains, such as A domains in many aryl-capped peptides and siderophores,[96] isolated PCPs in the actinomycin NRPS[32] and associated with the bleomycin cluster,[74] and stand-alone C domains in the vibriobactin[60] and bleomycin clusters.[73] Some authors proposed that these would correspond functionally to iterative type II PKSs. However, they cannot be regarded as such counterparts in terms of biosynthetic logic, since the iterative use of a single set of core domains is missing. In fact, the structural diversity obtained by iterative type II PKSs through various cyclizations of a ketide polymer precursor would not be possible in comparable form with a polymer of a single amino acid. To produce amino acid polymers, nature obviously has selected another kind of peptide synthetases, which operate with a different, Ppant-independent mechanism, such as the cyanophycin synthetase. Cyanophycin is a polymer of L-aspartic acid with L-arginine condensed to the side-chain carboxy acids.[97, 98] The construction of NRPSs from various distinct A domains, PCPs, and C domains that randomly interact with each other in order to produce a peptide library has

been proposed previously[5, 74] but seems out of reach at present time. It would require a general C domain with a broad selectivity on both acceptor and donor sites, as well as control of in trans interactions between A domains and PCPs on one hand and PCPs and C domains on other. Determinants for these issues are at present not at all understood.

Fungal iterative type I PKSs resemble iterative type II PKSs with the difference that all PKS domains are integrated on one polypeptide chain. Accessory enzymes in lovastatin (**20**; see Scheme 1 for structure) biosynthesis were shown to be involved in control of chain length and specific reduction of the extender units;[82] this is, however, only partially understood. This type of PKS can deviate in some aspects from the conserved domain organisation found in the modular type I PKSs or FASs. For example, a two-domain unit of two consecutive ACPs is a typical architectural feature of unknown funtion.[99] Fungal iterative type I PKSs can be regarded as the closest counterpart of iterative NRPSs, especially of those of fungal origin like the enniatin NRPS, although marked differences remain, for example the higher number of modules in iterative NRPSs.

An unsual module composition has been found in the antibiotic TA[100] and albicidin PKSs[101] as well as in a hybrid NRPS/PKS of *B. subtilis* for which the corresponding product is not known.[102] In these cases, the PKS modules lack an AT domain. Acyl transfer with the extender units is believed to be catalyzed by distinct acyl-transferases which are encoded within the clusters. This situation would correspond to the proposed in trans acylation of SyrE by SyrB in syringomycin biosynthesis (see above). Another PKS that corresponds in terms of unpredictable domain interaction to the nonlinear NRPSs is that for the biosynthesis of methymycin and pikromycin.[103] Here, the terminal Te domain can obviously either cyclize a 14-membered ketide assembled on the last module, or, alternatively, can interact with the second last module to cleave a 12-membered ketide. This partial "skipping" of the last module results in formation of a product mixture from one PKS complex.[103] Finally, probably the most exotic modular PKSs, which can be discussed as "nonlinear", have very recently been reported. They are involved in the production of polyunsaturated fatty acids and exhibit unusual domain arrangements with up to nine consecutive ACPs in one enzyme.[104] Thus, also the combinatorial potential of the catalytic domains of PKSs does not seem to be restricted to the solutions of the so far well-characterized systems.

Conclusions

Recent results indicate that the biosynthetic potential of NRPSs is even larger than previously appreciated. Unusual domain organizations within the multifunctional enzyme templates can lead to the assembly of products which are structurally more complex than those derived from linear peptide precursors. We propose the classification into linear NRPSs (type A), iterative NRPSs (type B), and nonlinear NRPSs (type C) to reflect the different mechanisms of synthesis. Detailed biochemical studies of nonlinear NRPSs will be necessary to understand the selectivity and interaction of the involved domains and to exploit these for combinatorial approaches.

Abbreviations

A	adenylation
aa	amino acid
Aad	δ-aminoadipate
ACP	acyl-carrier protein
ACV	α-aminoadipyl-L-cysteinyl-D-valine
AMP	adenosine monophosphate
AT	acyltransferase
ATP	adenosine triphosphate
C	condensation
CoA	coenzyme A
Cy	domain for heterocyclic ring formation
Dhb	dihydroxybenzoyl
Dhp	dihydroxyphenyl
E	epimerization
FAS	fatty acid synthase
KR	ketoreductase
KS	ketosynthase
MT	methyltransferase
NRPS	nonribosomal peptide synthetase
Orn	ornithine
Ox	oxidation
PCP	peptidyl-carrier protein
PKS	polyketide synthase
Ppant	4′-phosphopantetheine
R	reductase
T	thiolation
Te	thioesterase
S-NAC	S-*N*-acetylcysteamine

Work in the M.A.M. laboratory was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

[1] O. McAuliffe, R. P. Ross, C. Hill, *FEMS Microbiol. Rev.* **2001**, *25*, 285 – 308.
[2] M. E. Anderson, *Chem. Biol. Interact.* **1998**, *111 – 112*, 1 – 14.
[3] C. Fan, I. S. Park, C. T. Walsh, J. R. Knox, *Biochemistry* **1997**, *36*, 2531 – 2538.
[4] D. J. Faulkner, *Nat. Prod. Rep.* **1998**, *15*, 113 – 158.
[5] M. A. Marahiel, T. Stachelhaus, H. D. Mootz, *Chem. Rev.* **1997**, *97*, 2651 – 2673.
[6] D. E. Cane, C. T. Walsh, C. Khosla, *Science* **1998**, *282*, 63 – 68.
[7] C. T. Walsh, A. M. Gehring, P. H. Weinreb, L. E. Quadri, R. S. Flugel, *Curr. Opin. Chem. Biol.* **1997**, *1*, 309 – 315.
[8] C. T. Walsh, H. Chen, T. A. Keating, B. K. Hubbard, H. C. Losey, L. Luo, C. G. Marshall, D. A. Miller, H. M. Patel, *Curr. Opin. Chem. Biol.* **2001**, *5*, 525 – 534.
[9] T. Stachelhaus, A. Schneider, M. A. Marahiel, *Science* **1995**, *269*, 69 – 72.
[10] H. D. Mootz, M. A. Marahiel, *Curr. Opin. Biotechnol.* **1999**, *10*, 341 – 348.
[11] M. M. Yakimov, L. Giuliano, K. N. Timmis, P. N. Golyshin, *J. Mol. Microbiol. Biotechnol.* **2000**, *2*, 217 – 224.
[12] H. D. Mootz, D. Schwarzer, M. A. Marahiel, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5848 – 5853.
[13] S. Doekel, M. A. Marahiel, *Chem. Biol.* **2000**, *7*, 373 – 384.
[14] J. G. Arnez, D. Moras, *Trends Biochem. Sci.* **1997**, *22*, 211 – 216.
[15] E. Conti, T. Stachelhaus, M. A. Marahiel, P. Brick, *EMBO J.* **1997**, *16*, 4174 – 4183.
[16] T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem. Biol.* **1999**, *6*, 493 – 505.

[17] 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.
[18] K. Reuter, M. R. Mofid, M. A. Marahiel, R. Ficner, *EMBO J.* **1999**, *18*, 6823 – 6831.
[19] T. Stein, J. Vater, V. Kruft, A. Otto, B. Wittmann-Liebold, P. Franke, M. Panico, R. McDowell, H. R. Morris, *J. Biol. Chem.* **1996**, *271*, 15428 – 15435.
[20] N. Wu, S. Y. Tsuji, D. E. Cane, C. Khosla, *J. Am. Chem. Soc.* **2001**, *123*, 6465 – 6474.
[21] H. D. Mootz, M. A. Marahiel, *Curr. Opin. Chem. Biol.* **1997**, *1*, 543 – 551.
[22] U. Linne, M. A. Marahiel, *Biochemistry* **2000**, *39*, 10439 – 10447.
[23] P. J. Belshaw, C. T. Walsh, T. Stachelhaus, *Science* **1999**, *284*, 486 – 489.
[24] D. E. Ehmann, J. W. Trauger, T. Stachelhaus, C. T. Walsh, *Chem. Biol.* **2000**, *7*, 765 – 772.
[25] J. W. Trauger, R. M. Kohli, H. D. Mootz, M. A. Marahiel, C. T. Walsh, *Nature* **2000**, *407*, 215 – 218.
[26] J. W. Trauger, R. M. Kohli, C. T. Walsh, *Biochemistry* **2001**, *40*, 7092 – 7098.
[27] R. M. Kohli, J. W. Trauger, D. Schwarzer, M. A. Marahiel, C. T. Walsh, *Biochemistry* **2001**, *40*, 7099 – 7108.
[28] T. A. Keating, D. E. Ehmann, R. M. Kohli, C. G. Marshall, J. W. Trauger, C. T. Walsh, *ChemBioChem* **2001**, *2*, 99 – 107.
[29] H. D. Mootz, M. A. Marahiel, *J. Bacteriol.* **1997**, *179*, 6843 – 6850.
[30] D. Konz, A. Klens, K. Schörgendorfer, M. A. Marahiel, *Chem. Biol.* **1997**, *4*, 927 – 937.
[31] P. Cosmina, F. Rodriguez, F. de Ferra, G. Grandi, M. Perego, G. Venema, D. van Sinderen, *Mol. Microbiol.* **1993**, *8*, 821 – 831.
[32] F. Pfennig, F. Schauwecker, U. Keller, *J. Biol. Chem.* **1999**, *274*, 12508 – 12516.
[33] F. Schauwecker, F. Pfennig, N. Grammel, U. Keller, *Chem. Biol.* **2000**, *7*, 287 – 297.
[34] D. J. Smith, A. J. Earl, G. Turner, *EMBO J.* **1990**, *9*, 2743 – 2750.
[35] G. Weber, K. Schörgendorfer, E. Schneider-Scherzer, E. Leitner, *Curr. Genet.* **1994**, *26*, 120 – 125.
[36] V. de Crécy-Lagard, V. Blanc, P. Gil, L. Naudin, S. Lorenzon, A. Famechon, N. Bamas-Jacques, J. Crouzet, D. Thibaut, *J. Bacteriol.* **1997**, *179*, 705 – 713.
[37] V. de Crécy-Lagard, W. Saurin, D. Thibaut, P. Gil, L. Naudin, J. Crouzet, V. Blanc, *Antimicrob. Agents Chemother.* **1997**, *41*, 1904 – 1909.
[38] V. Tosato, A. M. Albertini, M. Zotti, S. Sonda, C. V. Bruschi, *Microbiology* **1997**, *143*, 3443 – 3450.
[39] T. P. Lin, C. L. Chen, L. K. Chang, J. S. Tschen, S. T. Liu, *J. Bacteriol.* **1999**, *181*, 5060 – 5067.
[40] P. Tudzynski, K. Holter, T. Correia, C. Arntz, N. Grammel, U. Keller, *Mol. Gen. Genet.* **1999**, *261*, 133 – 141.
[41] A. M. van Wageningen, P. N. Kirkpatrick, D. H. Williams, B. R. Harris, J. K. Kershaw, N. J. Lennard, M. Jones, S. J. Jones, P. J. Solenberg, *Chem. Biol.* **1998**, *5*, 155 – 162.
[42] S. Pelzer, R. Süssmuth, D. Heckmann, J. Recktenwald, P. Huber, G. Jung, W. Wohlleben, *Antimicrob. Agents Chemother.* **1999**, *43*, 1565 – 1573.
[43] H. T. Chiu, B. K. Hubbard, A. N. Shah, J. Eide, R. A. Fredenburg, C. T. Walsh, C. Khosla, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8548 – 8553.
[44] F. de Ferra, F. Rodriguez, O. Tortora, C. Tosi, G. Grandi, *J. Biol. Chem.* **1997**, *272*, 25304 – 25309.
[45] D. Schwarzer, H. D. Mootz, M. A. Marahiel, *Chem. Biol.* **2001**, *10*, 997 – 1010.
[46] R. Dieckmann, M. Pavela-Vrancic, H. von Döhren, H. Kleinkauf, *J. Mol. Biol.* **1999**, *288*, 129 – 140.
[47] H. von Döhren, R. Dieckmann, M. Pavela-Vrancic, *Chem. Biol.* **1999**, *6*, 273 – 279.
[48] U. Linne, S. Doekel, M. A. Marahiel, *Biochemistry* **2001**, *40*, 15824 – 15834.
[49] R. S. Gokhale, S. Y. Tsuji, D. E. Cane, C. Khosla, *Science* **1999**, *284*, 482 – 485.
[50] S. Y. Tsuji, D. E. Cane, C. Khosla, *Biochemistry* **2001**, *40*, 2326 – 2331.
[51] S. Y. Tsuji, N. Wu, C. Khosla, *Biochemistry* **2001**, *40*, 2317 – 2325.
[52] L. Du, C. Sanchez, B. Shen, *Metab. Eng.* **2001**, *3*, 78 – 95.
[53] A. Haese, M. Schubert, M. Herrmann, R. Zocher, *Mol. Microbiol.* **1993**, *7*, 905 – 914.
[54] A. M. Gehring, I. Mori, C. T. Walsh, *Biochemistry* **1998**, *37*, 2648 – 2659.
[55] J. J. May, T. M. Wendrich, M. A. Marahiel, *J. Biol. Chem.* **2001**, *276*, 7209 – 7217.
[56] C. A. Shaw-Reid, N. L. Kelleher, H. C. Losey, A. M. Gehring, C. Berg, C. T. Walsh, *Chem. Biol.* **1999**, *6*, 385 – 400.
[57] 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.
[58] E. Guenzi, G. Galli, I. Grgurina, D. C. Gross, G. Grandi, *J. Biol. Chem.* **1998**, *273*, 32857 – 32863.
[59] T. A. Keating, C. T. Walsh, *Curr. Opin. Chem. Biol.* **1999**, *3*, 598 – 606.
[60] T. A. Keating, C. G. Marshall, C. T. Walsh, *Biochemistry* **2000**, *39*, 15513 – 15521.
[61] T. A. Keating, C. G. Marshall, C. T. Walsh, *Biochemistry* **2000**, *39*, 15522 – 15530.
[62] J. H. Zhang, N. B. Quigley, D. C. Gross, *J. Bacteriol.* **1995**, *177*, 4009 – 4020.
[63] A. M. Gehring, E. DeMoll, J. D. Fetherston, I. Mori, G. F. Mayhew, F. R. Blattner, C. T. Walsh, R. D. Perry, *Chem. Biol.* **1998**, *5*, 573 – 586.
[64] A. M. Gehring, I. Mori, R. D. Perry, C. T. Walsh, *Biochemistry* **1998**, *37*, 11637 – 11650.
[65] Z. Suo, C. T. Walsh, D. A. Miller, *Biochemistry* **1999**, *38*, 14023 – 14035.
[66] Z. Suo, C. C. Tseng, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 99 – 104.
[67] D. A. Miller, C. T. Walsh, *Biochemistry* **2001**, *40*, 5313 – 5321.
[68] L. E. Quadri, T. A. Keating, H. M. Patel, C. T. Walsh, *Biochemistry* **1999**, *38*, 14941 – 14954.
[69] H. M. Patel, C. T. Walsh, *Biochemistry* **2001**, *40*, 9023 – 9031.
[70] B. Silakowski, B. Kunze, G. Nordsiek, H. Blöcker, G. Höfle, R. Müller, *Eur. J. Biochem.* **2000**, *267*, 6476 – 6485.
[71] J. R. Butterton, M. H. Choi, P. I. Watnick, P. A. Carroll, S. B. Calderwood, *J. Bacteriol.* **2000**, *182*, 1731 – 1738.
[72] E. E. Wyckoff, S. L. Smith, S. M. Payne, *J. Bacteriol.* **2001**, *183*, 1830 – 1834.
[73] L. Du, C. Sanchez, M. Chen, D. J. Edwards, B. Shen, *Chem. Biol.* **2000**, *7*, 623 – 642.
[74] L. Du, B. Shen, *Chem. Biol.* **1999**, *6*, 507 – 517.
[75] L. E. Quadri, J. Sello, T. A. Keating, P. H. Weinreb, C. T. Walsh, *Chem. Biol.* **1998**, *5*, 631 – 645.
[76] M. Steffensky, A. Muhlenweg, Z. X. Wang, S. M. Li, L. Heide, *Antimicrob. Agents Chemother.* **2000**, *44*, 1214 – 1222.
[77] M. Steffensky, S. M. Li, L. Heide, *J. Biol. Chem.* **2000**, *275*, 21754 – 21760.
[78] R. Dieckmann, T. Neuhof, M. Pavela-Vrancic, H. von Döhren, *FEBS Lett.* **2001**, *498*, 42 – 45.
[79] Z. X. Wang, S. M. Li, L. Heide, *Antimicrob. Agents Chemother.* **2000**, *44*, 3040 – 3048.
[80] D. E. Cane, C. T. Walsh, *Chem. Biol.* **1999**, *6*, 319 – 325.
[81] D. Schwarzer, M. A. Marahiel, *Naturwissenschaften* **2001**, *88*, 93 – 101.
[82] J. Kennedy, K. Auclair, S. G. Kendrew, C. Park, J. C. Vederas, C. R. Hutchinson, *Science* **1999**, *284*, 1368 – 1372.
[83] J. Cortes, S. F. Haydock, G. A. Roberts, D. J. Bevitt, P. F. Leadlay, *Nature* **1990**, *348*, 176 – 178.
[84] A. R. Butler, N. Bate, E. Cundliffe, *Chem. Biol.* **1999**, *6*, 287 – 292.
[85] J. Staunton, B. Wilkinson, *Curr. Opin. Chem. Biol.* **2001**, *5*, 159 – 164.
[86] Q. Xue, G. Ashley, C. R. Hutchinson, D. V. Santi, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11740 – 11745.
[87] I. Molnar, T. Schupp, M. Ono, R. Zirkle, M. Milnamow, B. Nowak-Thompson, N. Engel, C. Toupet, A. Stratmann, D. Cyr, J. Gorlach, J. Mayo, A. Hu, S. Goff, J. Schmid, J. Ligon, *Chem. Biol.* **2000**, *7*, 97 – 109.
[88] L. Tang, S. Shah, L. Chung, J. Carney, L. Katz, C. Khosla, B. Julien, *Science* **2000**, *287*, 640 – 642.
[89] B. Silakowski, H. U. Schairer, H. Ehret, B. Kunze, S. Weinig, G. Nordsiek, P. Brandt, H. Blocker, G. Hofle, S. Beyer, R. Muller, *J. Biol. Chem.* **1999**, *274*, 37391 – 37399.
[90] T. Nishizawa, M. Asayama, K. Fujii, K. Harada, M. Shirai, *J. Biochem.* **1999**, *126*, 520 – 529.
[91] D. Tillett, E. Dittmann, M. Erhard, H. von Döhren, T. Börner, B. A. Neilan, *Chem. Biol.* **2000**, *7*, 753 – 764.
[92] T. Schwecke, J. F. Aparicio, I. Molnar, A. Konig, L. E. Khaw, S. F. Haycock, M. Oliynyk, P. Caffrey, J. Cortes, J. B. Lester, G. A. Bohm, J. Staunton, P. F. Leadley, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7839 – 7843.
[93] E. H. Duitman, L. W. Hamoen, M. Rembold, G. Venema, H. Seitz, W. Saenger, F. Bernhard, R. Reinhardt, M. Schmidt, C. Ullrich, T. Stein, F. Leenders, J. Vater, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13294 – 13299.

[94] D. A. Hopwood, *Chem. Rev.* **1997**, *97*, 2465 – 2497.
[95] R. McDaniel, S. Ebert-Khosla, D. A. Hopwood, C. Khosla, *Nature* **1995**, *375*, 549 – 554.
[96] L. E. Quadri, *Mol. Microbiol.* **2000**, *37*, 1 – 12.
[97] H. Berg, K. Ziegler, K. Piotukh, K. Baier, W. Lockau, R. Volkmer-Engert, *Eur. J. Biochem.* **2000**, *267*, 5561 – 5570.
[98] K. Ziegler, A. Diener, C. Herpin, R. Richter, R. Deutzmann, W. Lockau, *Eur. J. Biochem.* **1998**, *254*, 154 – 159.
[99] I. Fujii, A. Watanabe, U. Sankawa, Y. Ebizuka, *Chem. Biol.* **2001**, *8*, 189 – 197.
[100] Y. Paitan, G. Alon, E. Orr, E. Z. Ron, E. Rosenberg, *J. Mol. Biol.* **1999**, *286*, 465 – 474.
[101] G. Huang, L. Zhang, R. G. Birch, *Microbiology* **2001**, *147*, 631 – 642.
[102] A. M. Albertini, T. Caramori, F. Scoffone, C. Scotti, A. Galizzi, *Microbiology* **1995**, *141*, 299 – 309.
[103] Y. Xue, D. H. Sherman, *Nature* **2000**, *403*, 571 – 575.
[104] J. G. Metz, P. Roessler, D. Facciotti, C. Levering, F. Dittrich, M. Lassner, R. Valentine, K. Lardizabal, F. Domergue, A. Yamada, K. Yazawa, V. Knauf, J. Browse, *Science* **2001**, *293*, 290 – 293.

Received: September 6, 2001 [A 292]