Chain Termination Steps in Nonribosomal Peptide Synthetase Assembly Lines: Directed Acyl-*S*-Enzyme Breakdown in Antibiotic and Siderophore Biosynthesis

Thomas A. Keating, David E. Ehmann, Rahul M. Kohli, C. Gary Marshall, John W. Trauger, and Christopher T. Walsh*[a]

KEYWORDS:

biosynthesis · cyclizations · enzyme catalysis · nonribosomal peptide synthesis · peptides

1. Introduction

A large number of therapeutically useful natural peptides are produced nonribosomally by assembly line enzymology, involving multidomain and multimodular catalysts that activate and assemble constituent amino acid monomers into oligopeptides. The peptide chains released (Scheme 1) can be final products such as cyclosporin (**1**)[1] and bacitracin (**2**)[2] or further tailored enzymatically, such as the tripeptide aminoadipoyl-cysteinyl-valine (ACV; **3**)[3] that is cyclized to penicillins and cephalosporins, and the vancomycin heptapeptide precursor (**4**) that is then oxidatively cyclized and glycosylated[4] to the active antibiotic **5**. A wide variety of nonproteinogenic and D isomers of amino acids are utilized by these assembly lines.[5, 6]

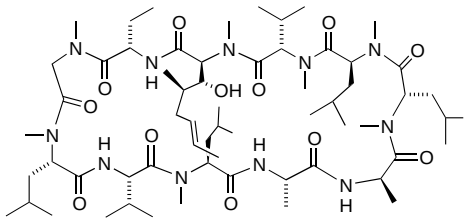
These catalytic proteins make peptide bonds without mRNA templates, instead using a thiotemplate mechanism,[7] in which each amino acid is activated as an aminoacyl-AMP by an adenylation domain (A domain) of 50 kDa, then transferred to a thiol side chain in an adjacent thiolation (T) domain of 10 kDa to tether the aminoacyl moiety as a thioester. The thiol side chain is presented by a 4′-phosphopantetheine moiety introduced posttranslationally on every T domain by dedicated phosphopantetheinyl transferases.[8] Two aminoacyl-*S*-T domains or a growing peptidyl-*S*-T domain and a downstream aminoacyl-*S*-T domain are substrates for amide bond formation and directional chain transfer by the third core domain, the condensation (C) domain, also of about 50 kDa. Thus, a functional module of a nonribosomal peptide synthetase (NRPS) is a C-A-T three-domain module, and two such modules are needed to make a peptide bond (Figure 1 A).[9] For example, the ACV tripeptide synthetase[3] has three modules, one for aminoadipate (Aad), one for cysteine, and one for valine, and the order of these three modules (Aad-Cys-Val) controls the sequence of the tripeptide assembled (Figure 1 B).

In addition to a core set of $(C\text{-}A\text{-}T)_n$ modules for chain elongation steps, the NRPS assembly lines may have specialized initiation modules, for example an A-T two-domain starting module[10] (see Figure 1 B), and invariably have a specialized C-terminal domain adapted for chain termination and release of the full-length natural peptide product.[11] The diversity and function of the chain-releasing C-terminal domains of NRPS assembly lines is the focus of this Minireview.

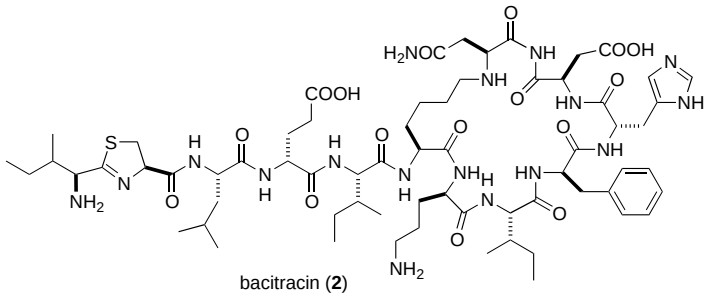
2. Specialized C-terminal domains to effect release of covalently bound acyl-*S*-enzymes from the end of the NRPS assembly lines

All NRPS enzymatic assembly lines face the problem of how to chemically disconnect the full-length acyl chain that has arrived at the most downstream T domain (Figure 2). All the chain elongation steps involve elongations of peptidyl-*S*-pantetheinyl-T domain thioesters by C-domain action: All the growing intermediates are covalently tethered to the assembly line during translocations as an elongating series of acyl-*S*-enzymes. This paradigm imposes several requirements on the NRPS enzymes. Premature hydrolysis of the cascade of acyl-*S*-enzyme intermediates or of the peptide monomers is to be studiously avoided or incomplete fragments are generated and energy in the form of coupled ATP hydrolysis is wasted. It is likely that all the acyl-*S*-T domain intermediates in NRPS assembly lines are kinetically sequestered from water. To release the mature chain the acyl – *S*-enzyme bond must be broken. Thus at the end of the assembly line, the full-length chain is brought out of sequestration; this occurs by transfer from the last carrier protein domain to a specialized C-terminal domain (Figure 2). This final domain must accept the mature acyl chain product so that the assembly lines do not become "stalled" with backed-up intermediates, and must also recognize only the full-length chain so that incomplete chains are not indiscriminately hydrolyzed. As we shall note below, a common architectural answer is a C-terminal thioesterase (TE) domain of 30 – 35 kDa (Figures 2 and 3), but there are

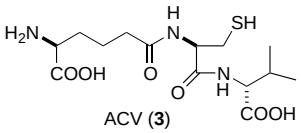
[a] *Prof. C. T. Walsh, Dr. T. A. Keating, Dr. D. E. Ehmann, R. M. Kohli, Dr. C. G. Marshall, Dr. J. W. Trauger*
Harvard Medical School
Department of Biological Chemistry and Molecular Pharmacology
240 Longwood Avenue, Boston, MA 02115 (USA)
Fax: (+1) 617-432-0438
E-mail: walsh@walsh.med.harvard.edu



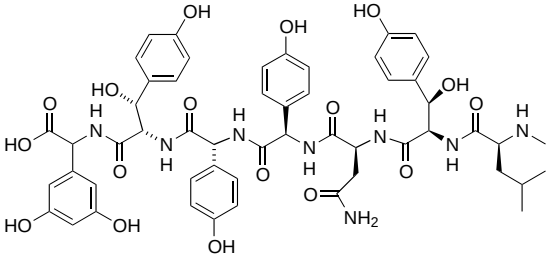
cyclosporin (**1**)



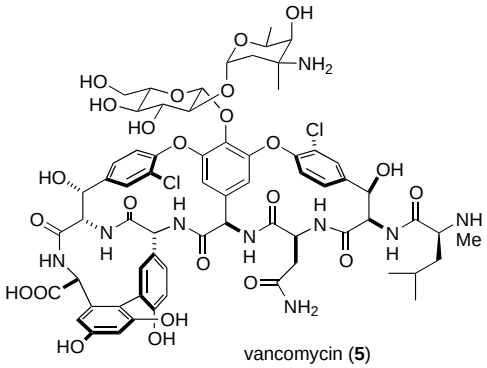
bacitracin (**2**)



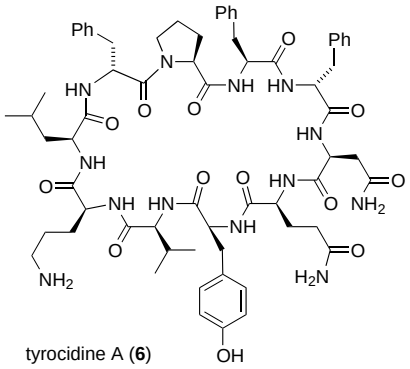
ACV (**3**)



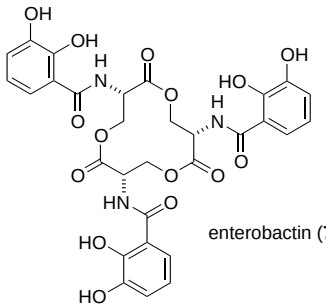
vancomycin heptapeptide precursor (**4**)



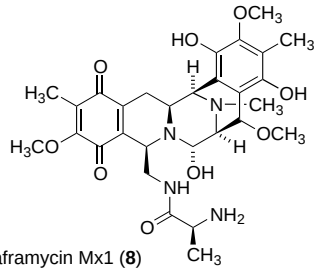
vancomycin (**5**)



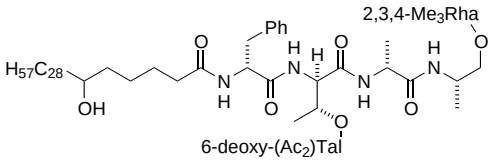
tyrocidine A (**6**)



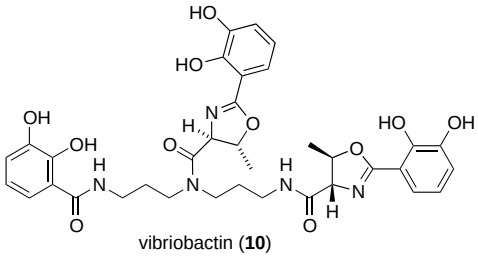
enterobactin (**7**)



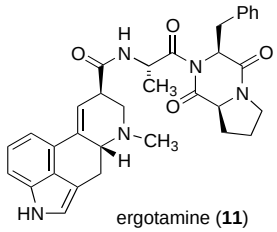
saframycin Mx1 (**8**)



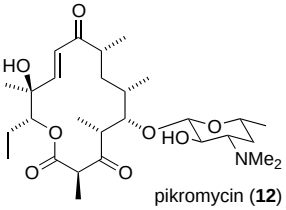
mycobacterial glycopeptidolipid (**9**)



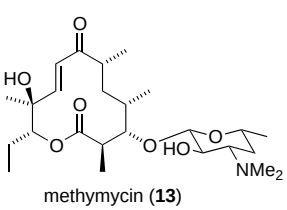
vibriobactin (**10**)



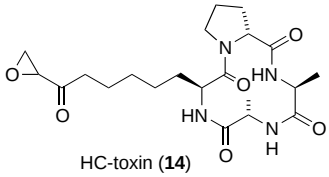
ergotamine (**11**)



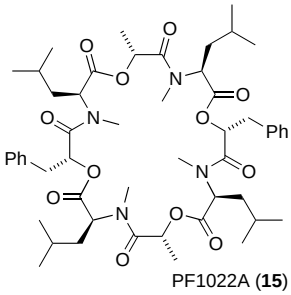
pikromycin (**12**)



methymycin (**13**)



HC-toxin (**14**)



PF1022A (**15**)

Scheme 1. *Naturally occurring peptides produced by nonribosomal peptide synthetases. Rha = rhamnose, Tal = talose.*

also C-terminal reductase domains and C-terminal condensation domain variants that yield alternate products (Figure 3) such as peptide aldehydes and C-terminal amides.

3. C-terminal thioesterase domains in NRPS assembly lines that produce linear carboxylates and cyclic lactones and lactams

The most common C-terminal domain in NRPS enzymatic assembly lines has homology to thioesterases (TEs) found in fatty acid metabolism,[12] and have thus been termed TE domains. These are variants of active-site serine hydrolases in which a His residue is required as general base and a Ser residue as catalytic nucleophile, producing a covalent acyl-*O*-Ser-TE from an acylthioester substrate, which is then hydrolyzed in a second step (Figure 4).[11, 13]

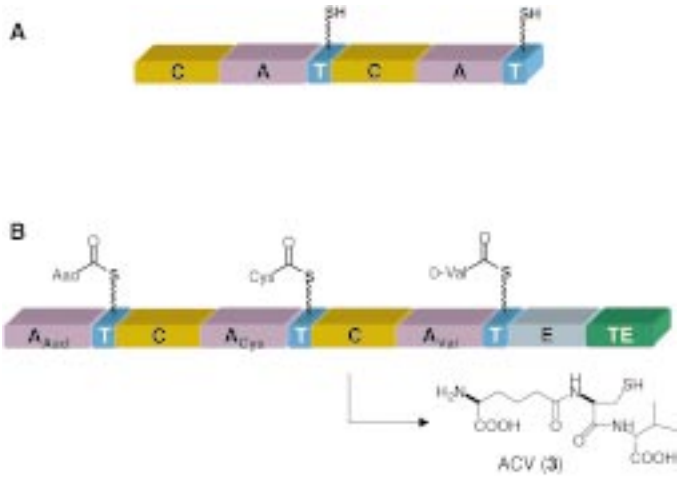


Figure 1. *A) A two-module nonribosomal peptide synthetase (NRPS), the minimum necessary for amide bond formation. C = condensation, A = adenylation, T = thiolation domain (also termed PCP = peptidyl carrier protein). B) α-Aminoadipoyl-cysteine-valine synthetase (three modules) condenses the three amino acid monomers to ACV (**3**), the precursor to isopenicillin N.*

Figure 2. *The thioesterase (TE) domain of an NRPS. TEs are variants of serine hydrolases. An active-site serine accepts the mature peptide chain which is transferred from the final T domain.*

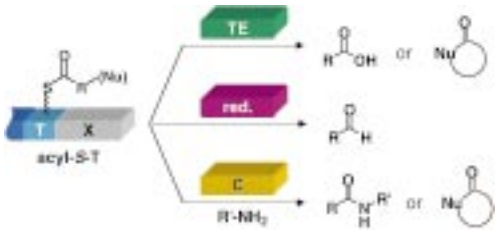


Figure 3. *Alternative termination strategies found in NRPS. The final domain (X) can be a thioesterase (TE), a reductase (red.), or a condensation (C) domain. TEs are found to hydrolyze or cyclize the mature chain, reductases release an aldehyde through NAD(P)H-coupled reduction, whereas C domains can use an inter- or intramolecular nucleophile (Nu) to attack the scissile thioester bond, with similar results to the TEs.*

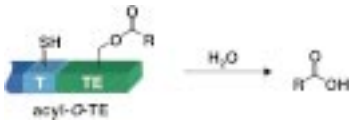


Figure 4. *Hydrolysis of the acyl-O-TE intermediate by water, resulting in the release of the carboxylic acid and the regeneration of the TE for another cycle.*

Many NRPS assembly lines, such as the ACV synthetase in penicillin biogenesis, release a peptide carboxylate, for example **3** (Figure 1 B),[3] and the net acyl-*S*-enzyme hydrolysis paradigm is demonstrated. The full-length peptidyl-*S*-T_n species (T_n = final T domain) is transferred to the TE domain to yield an acyl-*O*-TE,

Editorial Advisory Board Member:[*]

Christopher T. Walsh,
born in 1944, majored in biology at Harvard College and then obtained a Ph.D. in biochemistry in the laboratory of Fritz Lipmann at the Rockefeller University, New York. After postdoctoral research with R. H. Abeles from 1970 – 1972, he joined the faculty of MIT, in both the chemistry and biology departments, to study enzymatic reaction mechanisms, with an emphasis on inhibitor design in both the mechanism-based and slow-binding classes. His group also studied enzymatic catalysts involved in Baeyer – Villiger oxidations, cyclopropane fragmentations, light-induced repair of thymine dimers, nickel-containing enzymes involved in methanogenesis, and the enzymatic detoxification of mercurials. After fifteen years at MIT he left to chair a combined department of biological chemistry and molecular pharmacology at Harvard Medical School. At Harvard his group has studied enzymes of therapeutic relevance, including the peptidyl-prolyl isomerase that is the binding protein for cyclosporin A. Most of the recent focus has been on the mechanism of action of antibiotics, including the antibacterial drug of last resort, vancomycin. He and his group have determined the function of five enzymes that are necessary and sufficient to cause high-level clinical resistance to vancomycin by pathogenic enterococci. The interest in vancomycin, a nonribosomally produced heptapeptide, has extended to studies on the biosysnthesis of other peptide antibiotics and bacterial iron-chelating siderophores by multimodular enzyme assembly lines with the intent of decoding the mechanisms and utilizing them for combinatorial biosynthesis.

[*] *Members of the Editorial Advisory Board will be introduced to the readers with their first manuscript.*

which although thermodynamically less activated than the antecedent thioester, must be kinetically more labile to hydrolysis (Figure 4). The factors that control the timing of interdomain acyl chain transfer from T_n to the TE are not yet known and it is unclear if either interdomain transfer or the TE domain thioesterase activity limits the catalytic throughput of NRPS assembly lines. If the TE domains are mutated at the essential His or Ser residues, turnover ceases.[14, 15]

There are many NRPS assembly lines and cognate polyketide synthase (PKS) assembly lines terminating in homologous TE domains[16, 17] that do not release free carboxylate products but rather release cylic lactones or lactams as noted in Scheme 1. These must arise by intramolecular capture of the acyl-*O*-TE intermediate, in kinetic competition with intermolecular attack by water. Two such NRPS cyclases have been dissected in some mechanistic detail. One is the tyrocidine A synthetase, making the cyclic decapeptide tyrocidine A (**6**), and the other is the *E. coli* siderophore-generating enterobactin synthetase which cyclotrimerizes a dihydroxybenzoylseryl-*S*-T-acyl enzyme to the iron-chelating trilactone **7**.

In the tyrocidine synthetase case the decapeptidyl-*S*-T_{10} upstream acyl enzyme species can be simulated by a synthetic decapeptidyl-*S*-(*N*-acetylcysteamine) (NAC) thioester substrate, and the excised 28-kDa tyrocidine synthetase TE domain on its own retains potent cyclase activity[18] to make the cyclic decapeptide tyrocidine A (**6**; Figure 5A) with a k_{cat} value of 60 min^{-1}. This TE domain autonomously possesses all the instructions for regio- and stereospecific cyclization to the 30-membered cyclic peptide lactam. A model for tyrocidine synthetase chain termination is shown in Figure 5B. Hydrolysis of the decapeptidyl-*S*-NAC substrate is also observed; however, the cyclization/hydrolysis flux is 6:1[18] and this ratio is presumably higher in the intact synthetase assembly line.

In the enterobactin synthase assembly line the four-domain (C-A-T-TE) EntF subunit, in concert with the EntE and EntB subunits, makes the trilactone enterobactin (**7**) with a 12-membered ring at rates of 100 – 200 min^{-1} [15] (Figure 6A). A Ser 1138 →Ala mutation reduced enterobactin formation to less than 1/4500 of the wild-type activity, as this TE domain mutant is unable to accept covalent transfer of DHB-Ser (DHB = 2,3-dihydroxybenzoic acid). When a single mutant, His 1271 →Ala, was generated, V_{max} dropped dramatically but now this mutant EntF accumulated acyl-*O*-TE domain intermediates, DHB-Ser-*O*-TE and also DHB-Ser-*O*-DHB-Ser-*O*-TE, as evaluated by Fourier transform mass spectrometry (FT MS).[14] The TE domain *O*-acylated with a dimeric form of DHB-Ser is a key intermediate; presumably there is also a TE domain *O*-acylated with a linear trimeric form that is captured by intramolecular cyclization as shown in Figure 6B. There is no detectable hydrolase activity for wild-type EntF with its natural substrate. With salicyl-Ser-*S*-EntF a slow hydrolysis is detected after elongation to the dimeric form.[19]

A second example of a TE acting as both a cyclase and a final waystation for accumulation of identical units from the upstream NRPS modules is gramicidin S synthetase.[6] Gramicidin is a macrocycle of ten amino acids, (DPhe-Pro-Val-Orn-Leu)$_2$. As the synthetase possesses only five modules, the initial linear pentapeptide must transfer to the TE to await reloading of monomers and condensation to a second, identical pentapeptide before condensation of the two and cyclization to release. This additional TE function as holding position for oligomerizing product subunits is an efficient method to reuse NRPS modules for biosynthesis.

The switch from release of peptides with a free C-terminal carboxylate to release of cyclic peptide lactams and lactones can be as simple as the kinetic control of the fate of the acyl-*O*-TE intermediates towards capture by water in the second half-reaction (Figure 4). If water is completely excluded then the acyl-*O*-TE intermediate has time to fold in a conformation allowing intramolecular capture exclusively (Figure 7). It will be of interest to determine the structure of TE domains with hydrolase activity and of those with only cyclase activity and to examine the structural basis of the product-determining steps. TE domain swaps could be used purposefully for combinatorial biosynthetic rerouting of NRPS assembly line acyl-enzyme intermediates.

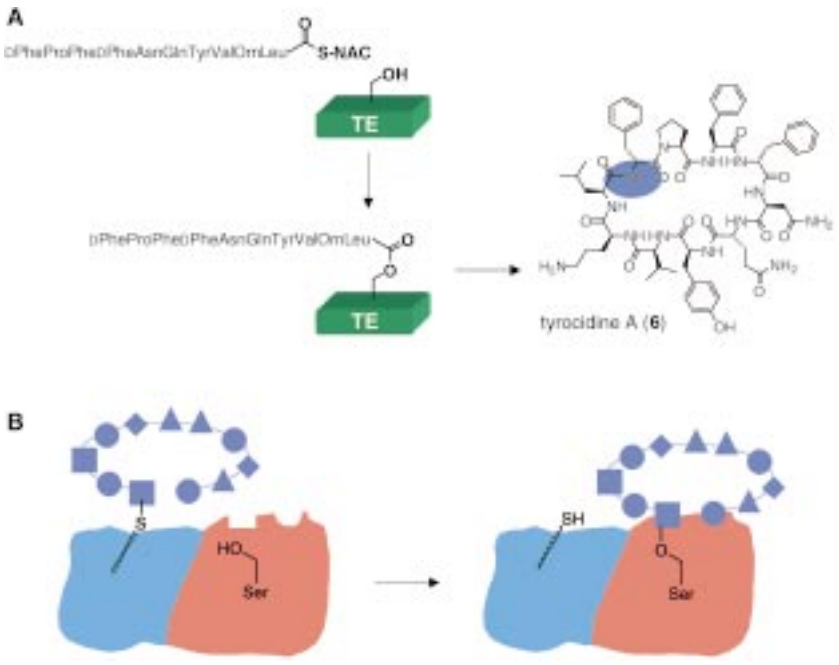


Figure 5. *A) Macrocyclization activity of the tyrocidine A thioesterase domain (as an independently expressed protein) on the* N*-acetylcysteamine (NAC) thioester of the linear tyrocidine decapeptide. The TE accepts the linear* S*-NAC-decapeptide as a substrate, and produces cyclized (→**6**;* $k_{cat} \approx 60\ min^{-1}$*) and hydrolyzed products in a 6:1 ratio in vitro. B) Model of the T-TE didomain terminus of tyrocidine synthetase and proposed cyclization mechanism of the TE. The first, ninth, and tenth residues of the decapeptide have been shown to have the greatest effect on activity, whereas the TE is relatively insensitive to the remaining internal amino acids. The TE is thus represented as having binding sites for the first and last residues, with the remainder lacking specific recognition.*

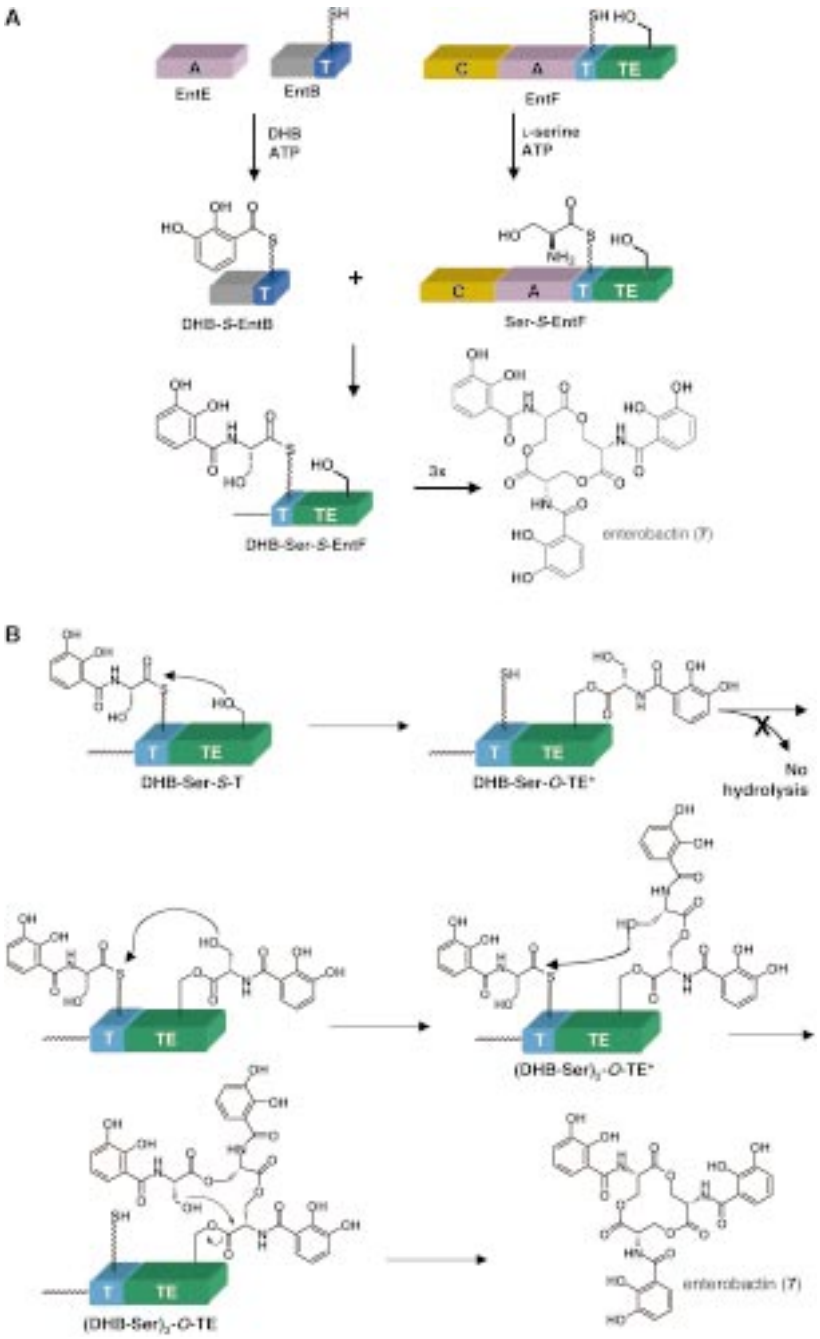


Figure 6. *Schematic representation of the biosynthesis of enterobactin (**7**) by the synthetase components EntE, EntB, and EntF. EntE loads EntB with 2,3-dihydroxybenzoic acid (DHB), while EntF self-acylates with L-serine. After condensation to DHB-Ser-S-EntF, two additional iterations and cyclization yield enterobactin (**7**). B) Minimal mechanism for EntF TE as a waystation for DHB-Ser accumulation and as a macrocyclizing catalyst. Starred (*) species have been identified by mass spectrometry. Mutations of the Ser residue of the TE abolish activity, while mutations of the His general base allow slow accumulation of the monomeric and dimeric forms of DHB-Ser.*

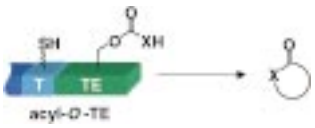
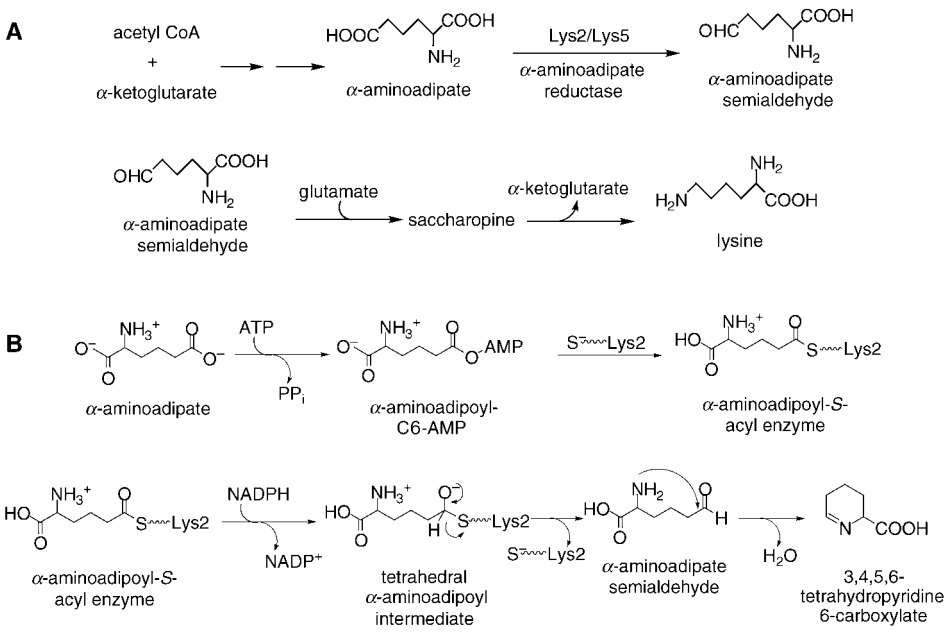


Figure 7. *Cyclization activity of thioesterases. An upstream nucleophile is employed to macrolactonize (X = O) or macrolactamize (X = NH) the substrate with concomitant release from the TE.*

4. C-terminal reductase domains in NRPS assembly lines that release aminoacyl or peptidyl aldehydes

While the C-terminal TE domain motif is most common at the end of NRPS assemblages, there are now several examples where TE domains are absent and instead are replaced by domains of 40 – 50 kDa with homology to NAD(P)-dependent reductases, such as in the yeast biosynthetic pathway for the essential amino acid lysine,[20] in the generation of the tetrapeptide antitumor agent saframycin (**8**),[21] and in mycobacterial peptidolipids (**9**) that may be involved in gliding properties of these pathogens.[22] The system that has been examined in detail and for which a functional reductive release mechanism has been validated is the Lys2/Lys5 pair of proteins in yeast.[20] Yeast carries out lysine biosynthesis from aminoadipate with regiospecific reduction of its C6 carboxylate to the aminoadipate semialdehyde (Scheme 2 A). The aldehyde subsequently undergoes net transamination to lysine via saccharopine. Lys2 is a 155-kDa, three-domain protein with the organization A-T-reductase, while Lys5 is a Lys2-specific phosphopantetheinyl transferase that converts the apo-T to the holo-T domain. Holo-Lys2 is then catalytically competent to activate aminoadipate as the C6-AMP mixed anhydride, and then generate the covalent aminoadipoyl-*S*-T-acyl-enzyme intermediate like a conventional A-T pair of an NRPS assembly line.[20] The third domain utilizes NADPH as a cosubstrate for hydride transfer to reduce the thioester bond of the aminoadipoyl-*S*-T intermediate to a hemithioaminal linkage (Scheme 2 B), which spontaneously decomposes to release the free thiol form of Lys2 and the C6 aldehyde product. This aldehyde cyclizes spontaneously to a tetrahydropyridine carboxylate as the accumulating product.

Analogous reductive cleavage of a tetrapeptidyl-*S*-enzyme intermediate (Ala-Gly-Tyr-Tyr-*S*-T domain) by an NADPH-mediated hydride transfer is predicted to yield the tetrapeptide aldehyde that subsequently cyclizes to the six-membered-ring hemiaminal in saframycin Mx1 (**8**).[21] Reductive cleavage is not likely to be a common chain cleavage mechanism unless the resultant reactive aldehyde can be captured intramolecularly in a stable hemiaminal or imine form as noted in the two examples described above, or reduced further to the alcohol, as implied by the terminal alaninol residue of the glycopeptidolipids like **9**.[22]



Scheme 2. *A) Overview of lysine biosynthesis in yeast. The key NRPS-like step is the reduction of α-aminoadipate to the semialdehyde by the Lys2/Lys5 pair. B) Mechanism of α-aminoadipate reduction by Lys2. After activation (with ATP) and covalent T-domain tethering, the α-aminoadipoyl thioester is reduced by NADPH. The resultant hemithioaminal breaks down to release the semialdehyde product and regenerate holo-Lys2.*

5. Amide bond formation as chain termination strategy by condensation domains

Some NRPS assembly lines apparently terminate not with thioesterase or reductase domains but instead with variants of the condensation (C) domains that are the amide-bond-forming, chain-elongating catalytic domains that act iteratively to grow the peptide chain. Cyclosporin synthetase, the 1.7-MDa polypeptide responsible for assembly of the fungal cyclic undecapeptide immunosuppressant drug **1**, is such an example (Figure 8).[1] The C-terminal 50-kDa condensation domain has been assumed to function as the amide synthase that couples the free amino group of D-Ala 1 to the carbonyl group of L-Ala 11 to produce the cyclic amide product in the chain release step, but there is no direct evidence for this function. A key mechanistic difference between a terminal C domain and a TE is the lack of evidence for an acyl-*O*-C domain intermediate. Indeed, analogy to acyl transferases such as the histone acetyl transferase Hat1[23, 24] indicates that C domains may catalyze direct nucleophilic attack on the scissile thioester bond. Other NRPSs that are predicted[5] to follow this method of C-terminal condensation-domain-mediated cyclization and release are those of enniatin,[25, 26] HC-toxin (**14**),[27] and PF1022A (**15**).[28] Enniatin, which is a cyclic trimer of ([D-2-hydroxyisovaleric acid]-[*N*-Me-valine]) consists of two modules followed by a didomain T-C unit. It is postulated that enniatin, as well as the cyclooctadepsipeptide PF1022A (**15**), may be assembled analogously to enterobactin, with the T-C unit replacing the enterobactin TE and allowing the successive buildup of oligomers on the T domain before C-domain-mediated cyclization.

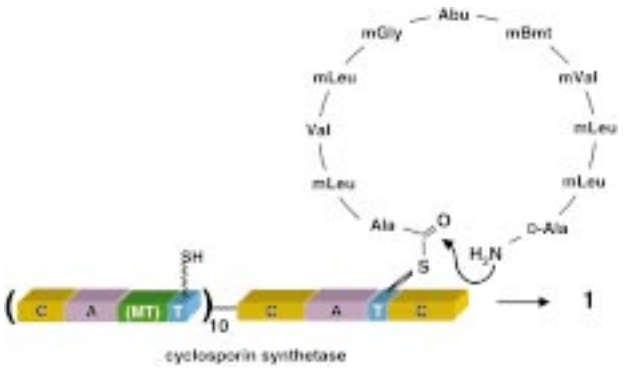


Figure 8. *Schematic representation of the final cyclization and release of cyclosporin (**1**) from its synthetase. The C-terminal condensation domain is proposed to catalyze attack of the N-terminal free amino group of the linear peptide on the thioester group without prior covalent transfer to the C domain, in contrast to the situtation with TEs. Abu = L-α-aminobutyric acid; Bmt = (4R)-4-[(E)-2-butenyl]-4-methyl-L-threonine; m = N-methyl; MT = N-methyltransferase domain (which occurs in several of the upstream modules).*

Recently the NRPS assembly line for the production of the siderophore vibriobactin (**10**) from the pathogenic bacterium *Vibrio cholerae* has been reconstituted in vitro from the pure protein components VibE, VibB, VibF, and VibH (Figure 9).[29, 30] Vibriobactin (**10**) has some design analogy to enterobactin (**7**) in that each compound has three catechol moieties as ligands to chelate ferric ion, but while enterobactin is a cyclic trilactone, the three catechol-containing acyl chains in vibriobactin are appended through three amide linkages to the triamine norspermidine.[31] One of the acyl groups is a simple 2,3-dihydroxybenzoyl (DHB) moiety while the other two represent DHB-threonyl moieties (akin to the DHB-seryl moieties of enterobactin) that have been cyclized and dehydrated to dihydroxyphenyloxazolinyl (DHPO) acyl groups. VibF, the functional homologue of EntF in the enterobactin synthetase system, does not have a C-terminal TE domain, but instead brackets the single T domain with a pair of C domains (Figure 9). Additionally, VibH is a free-standing condensation protein. VibH makes DHB-norspermidine from DHB-*S*-VibB and norspermidine and this is then the preferred substrate for bis-acylation with the DHPO groups from VibF (Figure 9). The VibF condensation domains are the most obvious candidates for DHPO-chain-releasing amide synthases to acylate the remaining primary and secondary amino groups of the DHB-norspermidine product of VibH. All three of these terminating amide synthase activities involve intermolecular bond formation, in contrast with the postulated intramolecular cyclizing activity of the C domain of cyclosporin

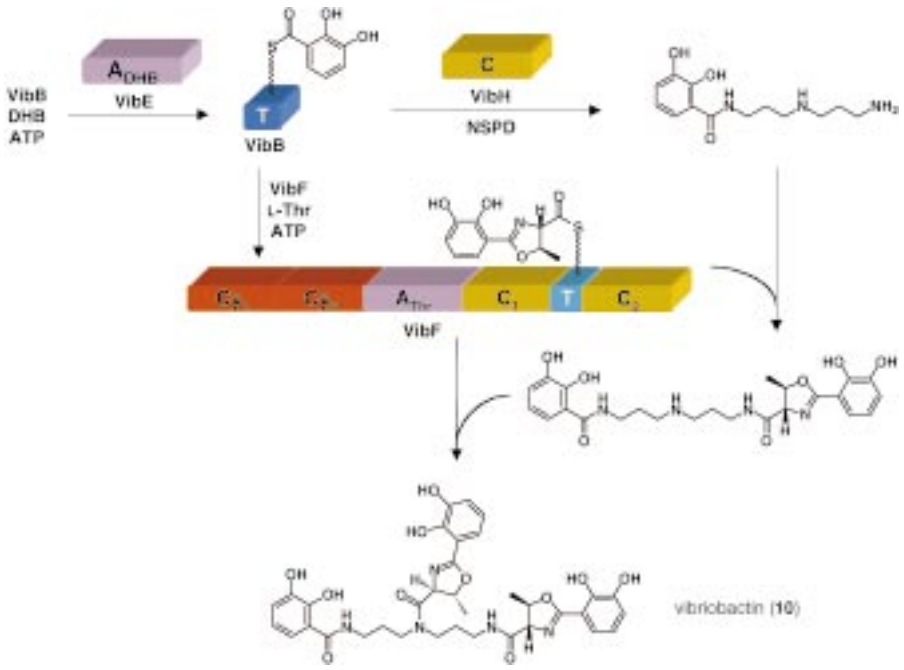


Figure 9. *Mechanistic scheme for vibriobactin (**10**) biosynthesis by VibB, VibE, VibH, and VibF. VibE loads VibB with 2,3-dihydroxybenzoate, while VibF self-acylates with L-Thr. VibH first condenses DHB-S-VibB with norspermidine (NSPD). DHB-S-VibB also serves as a donor for VibF in the formation of the dihydroxyphenyloxazolinyl (DHPO) thioester. VibF then caps the two remaining amine functionalities of DHB-NSPD with DHPO moieties, presumably mediated by the C_1 and C_2 domains, to yield vibriobactin. Cy = cyclization domain.*

synthetase. Each of the acceptor amide nucleophiles in the vibriobactin synthesis sequence is a soluble, small-molecule amine. Thus, C domains that use soluble substrates or have cyclization capacity appear to be a third way of effecting mature chain release from the downstream termini of NRPS assembly lines.

6. Release of the acyl chain by diketopiperazine cyclization in lysergic acid biosynthesis

The ergot fungus *Claviceps purpurea* produces peptide alkaloids responsible for outbreaks of ergotism. Typically these are composed of a nonpeptidic moiety of D-lysergic acid connected by an amide linkage to tripeptide moieties. The tripeptide is assembled by a two-subunit NRPS D-lysergyl-peptide synthetase (LPS),[32, 33] activating Ala, Phe, and Pro to produce ergotamine (**11**). LPS2 activates the D-lysergic acid as donor thioester and LPS1 activates Ala, Phe, and Pro as three aminoacyl-*S*-T-acyl enzymes (Figure 10). There is as yet no sequence reported for the gene encoding LPS1 but the protein subunit appears not to have thioesterase activity, as judged by lack of inhibition of chain release by phenylmethylsulfonyl fluoride (PMSF), a reagent shown to inhibit the TE domain of fatty acid synthase.[34] The final cyclic product is the bis-lactam ergotamine (**11**), a diketopiperazine (DKP) formed from the Phe and Pro moieties. Walzel et al.[32] have proposed that the DKP-forming cyclization is the chain-release mechanism (Figure 10), particularly appropriate for the C-terminal prolyl group since DKP formation with the cyclic amine proline is more rapid than with the primary amino acids. This route of chain release was slow, with a k_{cat} value around 1 min^{-1}, which may indicate noncatalytic cyclization and release. Experiments on truncated dimodule constructs that rely upon noncatalytic DKP formation by proline amide attack are consistent with both this mechanism and this rate.[35, 36]

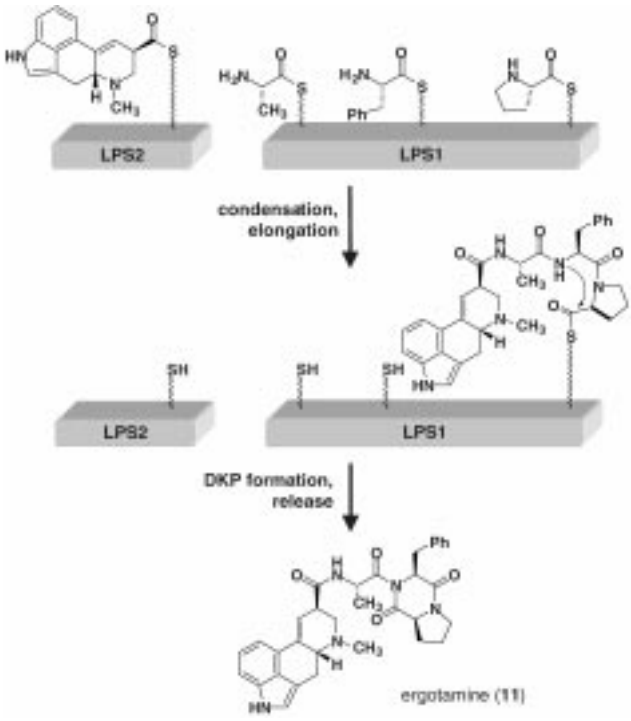


Figure 10. *Biosynthesis and release of ergotamine (**11**). The two protein products involved, LPS2 and LPS1, have not been sequenced, so domain assignments cannot be made. Condensation and elongation of the four monomers leads to the linear tetrapeptide, which is proposed to spontaneously cyclize via diketopiperazine (DKP) formation, promoted by the proline conformation. Rates of DKP formation are consistent with nonenzymatic cyclization.*

7. Examples of novel chain termination in polyketide synthases

While this review has concentrated on chain release strategies in NRPS systems, a few examples of unusual termination in polyketide synthases (PKSs) will illustrate the diversity of these systems. Like NRPSs, PKSs most often terminate in hydrolytic or cyclizing TEs.[37] However, the final PKS module of rifamycin synthase ends

with an acyl carrier protein (ACP) domain instead of a cyclizing TE, as the product structure would suggest.[38] Termination and cyclization to proansamycin X (**16**) (Figure 11 A) are performed by a separate protein, RifF, which is an amide synthase without homology either to TEs or C domains,[39] and which is most similar to arylamine-acetyl-CoA acetyltransferases. Likewise, in the biosynthesis of lovastatin (**17**),[40, 41] a novel transesterase protein, LovD, is proposed to employ the released, soluble product of one PKS subunit as a hydroxy group nucleophile to couple to the terminal thioesterified product of a second PKS subunit, thus forming and releasing the mature molecule **17** (Figure 11 B). In contrast to RifF, LovD appears by homology to be related to carboxypeptidases and esterases. Interestingly, the PKS component lovastatin nonaketide synthase (LNKS) that synthesizes the immediate precursor of monacolin J (Figure 11 B) possesses a C-terminal NRPS condensation domain;[41] the function of this domain in biosynthesis remains obscure. Finally, the pikromycin/methymycin PKS presents an interesting example of a branch point in a PKS assembly line that leads to two products, depending on the timing and location of termination. The six-module pikromycin PKS can not only produce the eponymous 14-membered macrolide **12**, but can also terminate and cyclize the product of the fifth module, yielding the 12-membered macrolide methymycin (**13**).[42] Recent work has demonstrated that methymycin results from an N-terminally truncated form of module 6 (PikAIV) arising from an internal, alternate translational start site.[43] This truncation prevents elongation to pikromycin (**12**), but preserves intact the internal TE that has been shown to be necessary for both methymycin and pikromycin biosynthesis. The TE is thus able to interact with and cyclize the products of both module 5 (in trans) and module 6 (in cis).

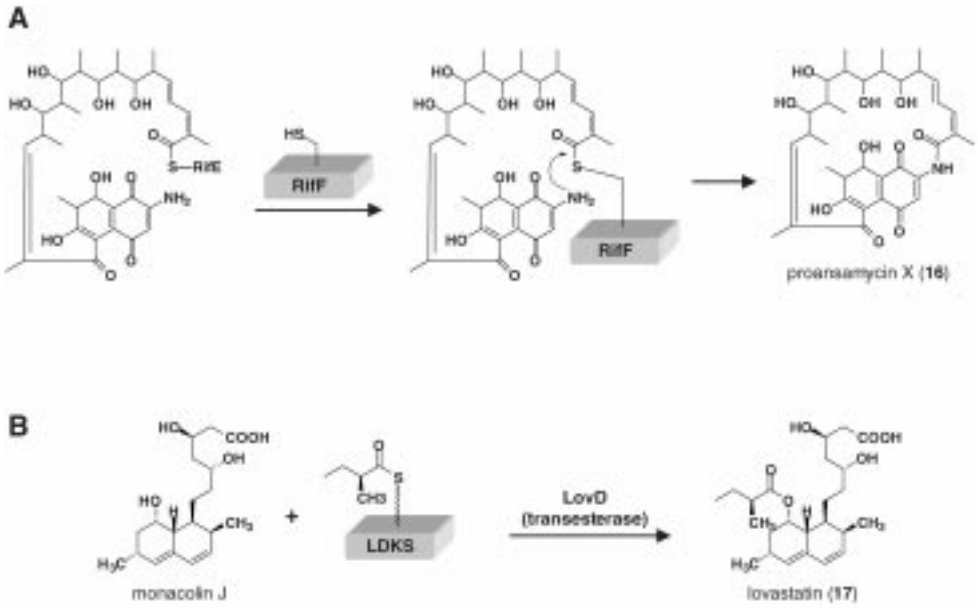


Figure 11. *A) Proposed mechanism of cyclization of the rifamycin linear polyketide to proansamycin X (**16**) by RifF. The polyketide synthase (PKS) RifE terminates in a carrier protein domain, so the separate protein RifF (similar to arylamine-acetyl-CoA acetyltransferases) accepts chain transfer and catalyzes cyclization. B) Proposed mechanism for the final transesterification in lovastatin biosynthesis. Monacolin J, the product of a separate PKS, is acylated by LovD (similar to carboxypeptidases) with an LDKS-protein-bound 2-methylbutyrate (which is a PKS product) to yield lovastatin (**17**). LDKS = lovastatin diketide synthase.*

8. Conclusion

The fundamental use of a thiotemplate mechanism by NRPS assembly lines to activate, tether, and add amino acid monomers to growing chains of peptidyl-*S*-enzymes establishes the thermodynamic activation for chain growth (acyl thioesters in downhill condensations to amides) and imposes constraints of directionality as the attacking nucleophilic aminoacyl-*S*-T domain monomer is always in the downstream position. The additional mechanistic outcome is that the antibiotic or siderophore acyl chain grows as a cascade of elongating, covalently tethered acyl-*S*-enzyme intermediates, moving from upstream holo-T domains to downstream ones. No acyl intermediates diffuse away or are lost to solvent to leave truncated or incomplete chains, provided the elongating series of acyl-*S*-T domains are protected from adventitious hydrolysis. The templating is set by the order of the sequential C-A-T modules and the specificity of each A domain and C domain for activation and condensation, respectively.[7, 10]

This strategy for both NRPS and PKS catalysis means that for the assembly lines to function catalytically rather than stoichiometrically, the full-length antibiotic or siderophore acyl chain, once it has reached the most downstream module and is docked on the last T domain, T_n, must be promptly cleaved to regenerate the assembly line for another biosynthetic iteration. There must be a specialized catalytic machinery at the C termini of the enzymatic assembly lines to deacylate the peptidyl-*S*-T_n domain but none of the earlier peptidyl-*S*-T intermediates. We have noted three general solutions: TE domains, reductase domains, and amide-synthesizing C domains appended to the end of the NRPS assembly lines. Understanding acyl chain transfer mechanism and selectivity from T_n domain to the attacking nucleophile of each of these three product-determining catalytic domains will help explain the diversity possible in these natural product classes and suggest engineering strategies for rerouting acyl chains in combinatorial biosynthesis efforts to create new natural products.[44, 45]

Work in this laboratory has been supported by grants from the National Institutes of Health and from Schering-Plough. T.A.K. is a Fellow of the Cancer Research Fund of the Damon Runyon – Walter Winchell Foundation (DRG-1483); D.E.E. is a National Science Foundation Predoctoral Fellow; C.G.M. is a Postdoctoral Fellow of the National Science and Engineering Research Council of Canada. J.W.T. is a National Institutes of Health Postdoctoral Fellow.

[1] G. Weber, K. Schörgendorfer, E. Schneider-Scherzer, E. Leitner, *Curr. Genet.* **1994**, *26*, 120 – 125.
[2] D. Konz, A. Klens, K. Schörgendorfer, M. A. Marahiel, *Chem. Biol.* **1997**, *4*, 927 – 937.
[3] M. F. Byford, J. E. Baldwin, C.-Y. Shiau, C. J. Schofield, *Chem. Rev.* **1997**, *97*, 2631 – 2649.
[4] A. M. A. van Wageningen, P. N. Kirkpatrick, D. H. Williams, B. R. Harris, J. K. Kershaw, N. J. Lennard, M. Jones, S. J. M. Jones, P. J. Solenberg, *Chem. Biol.* **1998**, *5*, 155 – 162.
[5] D. Konz, M. A. Marahiel, *Chem. Biol.* **1999**, *6*, R39-R48.
[6] M. A. Marahiel, T. Stachelhaus, H. D. Mootz, *Chem. Rev.* **1997**, *97*, 2651 – 2673.
[7] 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*, 15 428 – 15 435.
[8] C. T. Walsh, A. M. Gehring, P. H. Weinreb, L. E. N. Quadri, R. S. Flugel, *Curr. Opin. Chem. Biol.* **1997**, *1*, 309 – 315.
[9] T. A. Keating, C. T. Walsh, *Curr. Opin. Chem. Biol.* **1999**, *3*, 598 – 606.
[10] M. A. Marahiel, *Chem. Biol.* **1997**, *4*, 561 – 567.
[11] A. Schneider, M. A. Marahiel, *Arch. Microbiol.* **1998**, *169*, 404 – 410.
[12] M. Pazirandeh, S. S. Chirala, W.-Y. Huang, S. J. Wakil, *J. Biol. Chem.* **1989**, *264*, 18 195 – 18 201.
[13] M. Pazirandeh, S. S. Chirala, S. J. Wakil, *J. Biol. Chem.* **1991**, *266*, 20 946 – 20 952.
[14] C. A. Shaw-Reid, N. L. Kelleher, H. C. Losey, A. M. Gehring, C. Berg, C. T. Walsh, *Chem. Biol.* **1999**, *6*, 385 – 400.
[15] A. M. Gehring, I. Mori, C. T. Walsh, *Biochemistry* **1998**, *37*, 2648 – 2659.
[16] J. Krätzschmar, M. Krause, M. A. Marahiel, *J. Bacteriol.* **1989**, *171*, 5422 – 5429.
[17] H. D. Mootz, M. A. Marahiel, *J. Bacteriol.* **1997**, *197*, 6843 – 6850.
[18] J. W. Trauger, R. M. Kohli, H. D. Mootz, M. A. Marahiel, C. T. Walsh, *Nature* **2000**, *407*, 215 – 218.
[19] D. E. Ehmann, Ph.D. thesis, Harvard Medical School (Boston, MA), **2000**.
[20] D. E. Ehmann, A. M. Gehring, C. T. Walsh, *Biochemistry* **1999**, *38*, 6171 – 6177.
[21] A. Pospiech, J. Bietenhader, T. Schupp, *Microbiology* **1996**, *142*, 741 – 746.
[22] H. Billman-Jacobe, M. J. McConville, R. E. Haites, S. Kovacevic, R. L. Coppel, *Mol. Microbiol.* **1999**, *33*, 1244 – 1253.
[23] V. de Crécy-Lagard in *Comprehensive Natural Product Chemistry, Vol. 4* (Ed.: J. W. Kelly), Elsevier, Oxford, **1999**, pp. 221 – 238.
[24] R. N. Dutnall, S. T. Tafrov, R. Sternglanz, V. Ramakrishnan, *Cell* **1998**, *94*, 427 – 438.
[25] R. Pieper, A. Haese, W. Schroder, R. Zocher, *Eur. J. Biochem.* **1995**, *230*, 119 – 126.
[26] A. Haese, M. Schubert, M. Herrmann, R. Zocher, *Mol. Microbiol.* **1993**, *7*, 905 – 914.
[27] J. S. Scott-Craig, D. G. Panaccione, J. A. Pocard, J. D. Walton, *J. Biol. Chem.* **1992**, *267*, 26044 – 26049.
[28] W. Weckwerth, K. Miyamoto, K. Iinuma, M. Krause, M. Glinski, T. Storm, G. Bonse, H. Kleinkauf, R. Zocher, *J. Biol. Chem.* **2000**, *275*, 17 909 – 17 915.
[29] T. A. Keating, C. G. Marshall, C. T. Walsh, *Biochemistry* **2000**, *39*, 15 513 – 15 521.
[30] T. A. Keating, C. G. Marshall, C. T. Walsh, *Biochemistry* **2000**, *39*, 15 522 – 15 530
[31] G. L. Griffiths, S. P. Sigel, S. M. Payne, J. B. Neilands, *J. Biol. Chem.* **1984**, *259*, 383 – 385.
[32] B. Walzel, B. Riederer, U. Keller, *Chem. Biol.* **1997**, *4*, 223 – 230.
[33] B. Riederer, M. Han, U. Keller, *J. Biol. Chem.* **1996**, *271*, 27 524 – 27 530.
[34] L. Rogers, P. E. Kolattukudy, M. DeRenobales, *J. Biol. Chem.* **1982**, *257*, 880 – 886.
[35] P. J. Belshaw, C. T. Walsh, T. Stachelhaus, *Science* **1999**, *284*, 486 – 489.
[36] T. Stachelhaus, H. D. Mootz, V. Bergendahl, M. A. Marahiel, *J. Biol. Chem.* **1998**, *273*, 22 773 – 22 781.
[37] C. Khosla, *Chem. Rev.* **1997**, *97*, 2577 – 2590.
[38] H. G. Floss, T.-W. Yu, *Curr. Opin. Chem. Biol.* **1999**, *3*, 592 – 597.
[39] T.-w. Yu, Y. Shen, Y. Doi-Katayama, L. Tang, C. Park, B. S. Moore, C. R. Hutchinson, H. G. Floss, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9051 – 9056.
[40] J. Kennedy, K. Auclair, S. G. Kendrew, C. Park, J. C. Vederas, C. R. Hutchinson, *Science* **1999**, *284*, 1368 – 1372.
[41] L. Hendrickson, C. R. Davis, C. Roach, D. K. Nguyen, T. Aldrich, P. C. McAda, C. D. Reeves, *Chem. Biol.* **1999**, *6*, 429 – 439.
[42] L. Tang, H. Fu, M. C. Betlach, R. McDaniel, *Chem. Biol.* **1999**, *6*, 553 – 558.
[43] Y. Xue, D. H. Sherman, *Nature* **2000**, *403*, 571 – 575.
[44] D. E. Cane, C. T. Walsh, C. Khosla, *Science* **1998**, *282*, 63 – 68.
[45] H. D. Mootz, M. A. Marahiel, *Curr. Opin. Biotechnol.* **1999**, *10*, 341 – 348.

Received: August 21, 2000 [M 118]