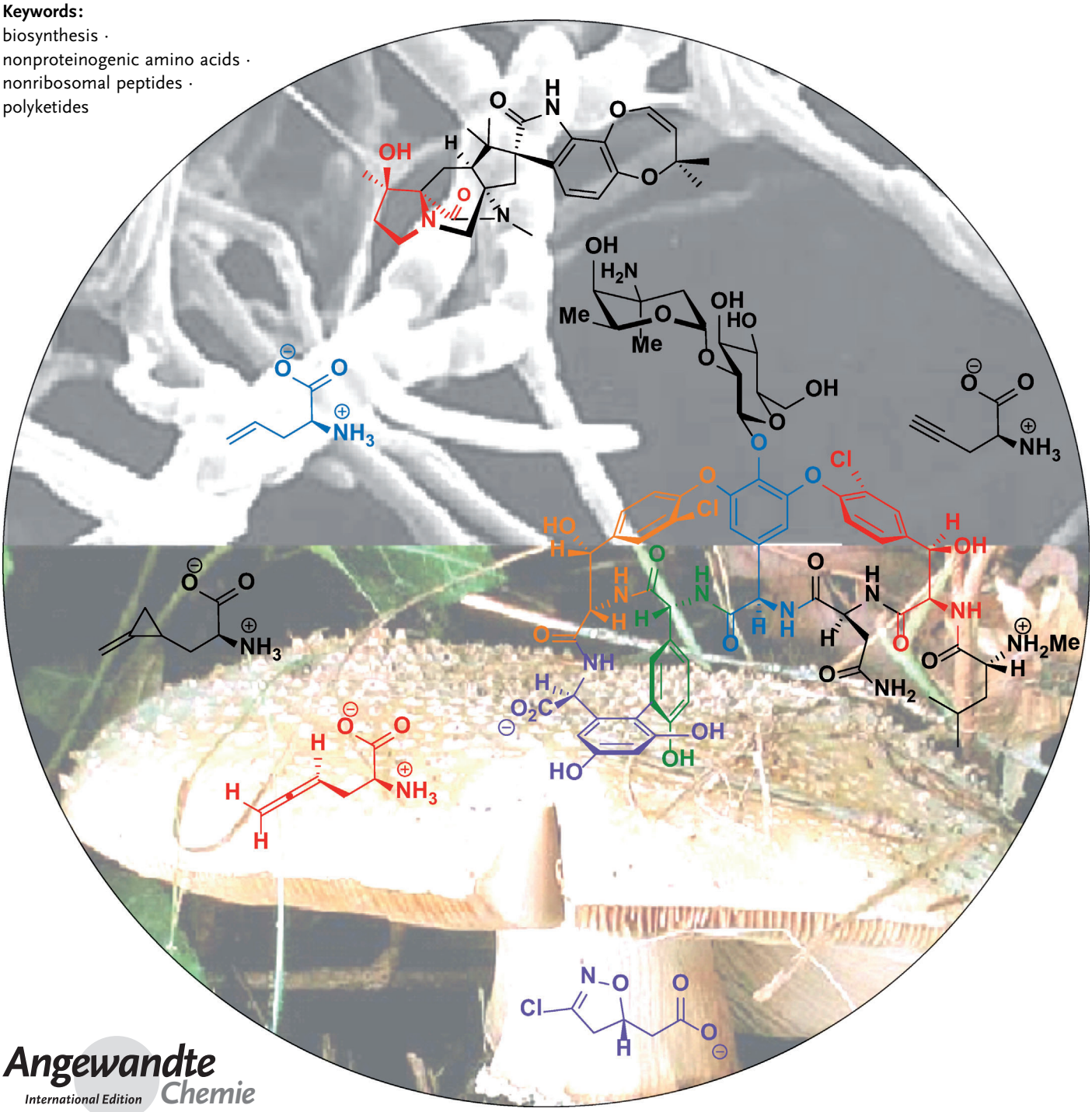
Biosynthesis

DOI: 10.1002/anie.201208344

Nonproteinogenic Amino Acid Building Blocks for Nonribosomal Peptide and Hybrid Polyketide Scaffolds

*Christopher T. Walsh, Robert V. O'Brien, and Chaitan Khosla**

Keywords:
biosynthesis · nonproteinogenic amino acids · nonribosomal peptides · polyketides



Freestanding nonproteinogenic amino acids have long been recognized for their antimetabolite properties and tendency to be uncovered to reactive functionalities by the catalytic action of target enzymes. By installing them regiospecifically into biogenic peptides and proteins, it may be possible to usher a new era at the interface between small molecule and large molecule medicinal chemistry. Site-selective protein functionalization offers uniquely attractive strategies for posttranslational modification of proteins. Last, but not least, many of the amino acids not selected by nature for protein incorporation offer rich architectural possibilities in the context of ribosomally derived polypeptides. This Review summarizes the biosynthetic routes to and metabolic logic for the major classes of the noncanonical amino acid building blocks that end up in both nonribosomal peptide frameworks and in hybrid nonribosomal peptide-polyketide scaffolds.

From the Contents

1. Introduction

Microbes have a remarkable capacity to build amino acid frameworks that are not incorporated into proteins. It has been estimated that about 500 naturally occurring amino acids have been identified to date, leaving the 20 proteinogenic amino acids as the 4% minority. Many of these nonproteinogenic amino acids have been discovered and their biological context established since the last time this subject was reviewed in this journal in 1983.[1] While some of the nonproteinogenic amino acids are utilized as intermediates in primary metabolic pathways (e.g. homoserine, ornithine), most of the unheralded 96% majority serve as building blocks for small bioactive peptide scaffolds. They may represent an under-utilized inventory of building blocks for protein engineers, medicinal chemists, and materials scientists.

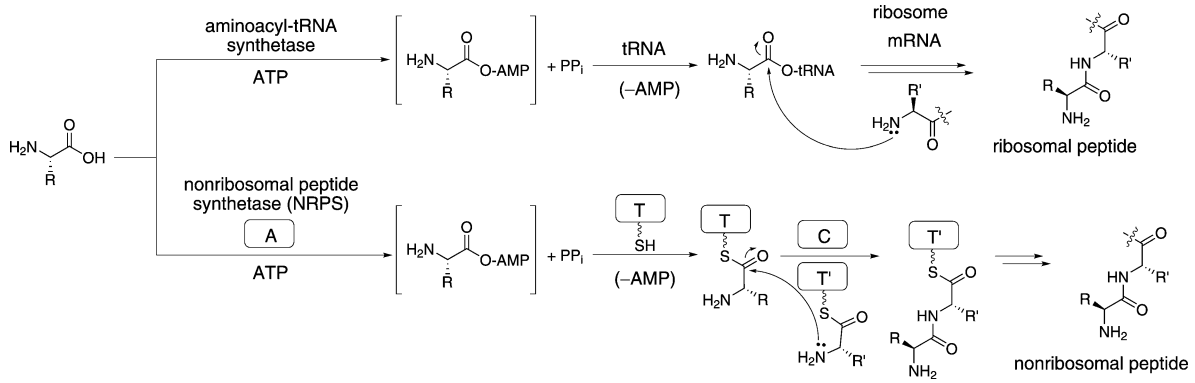
On the one hand, the existence of many hundreds of natural peptides with one or more nonproteinogenic amino acid residues reflects the ability of the side chains of these particular monomers to impart some useful functional property not available from the 20 proteinogenic building blocks. On the other hand, these natural peptide frameworks highlight the existence of biological machineries for the selection, activation, and incorporation of nonproteinogenic amino acid monomers into nonribosomal peptides (NRPs) as well as chain elongation of the resulting intermediates.[2,3] Perhaps many of those machineries could be moved, evolved, and re-engineered to work as building blocks in ribosomal protein biosynthesis. The development of pairs of evolved tRNAs and aminoacyl-tRNA synthetases to recognize synthetic unnatural amino acids, their introduction into growing cells, and the ability of the resulting cells to efficiently synthesize proteins harboring the corresponding unnatural amino acids augur well for the utility of these endogenously generated nonproteinogenic amino acids to the protein engineer.[4,5] To do so requires knowledge of the metabolic pathways to these noncanonical building blocks and the kinds of enzymatic transformations involved.

This Review summarizes the biosynthetic routes to and metabolic logic for the major classes of the noncanonical amino acid building blocks that end up in both nonribosomal peptide frameworks and in hybrid nonribosomal peptide-polyketide scaffolds.

In addition to microbial peptides of 2–22 residues built nonribosomally and, therefore, independent of mRNA mes-

[*] Prof. C. T. Walsh
Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School
Boston, MA 02115 (USA)
Dr. R. V. O'Brien, Prof. C. Khosla
Department of Chemistry, Stanford University
Stanford, CA 94305-5025 (USA)
E-mail: khosla@stanford.edu
Prof. C. Khosla
Chemical Engineering, Stanford University
Stanford, CA 94305-5025 (USA)
Prof. C. Khosla
Biochemistry, Stanford University
Stanford, CA 94305-5025 (USA)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201208344.



Scheme 1. Biosynthesis of ribosomal and nonribosomal peptides. ATP = adenosine triphosphate, AMP = adenosine monophosphate, PP_i = pyrophosphate, A domain = adenylation domain, T domain = thiolation (peptidyl carrier protein) domain, C domain = condensation domain.

sages, there are hundreds of naturally occurring scaffolds that are hybrids of peptide and polyketide (PK) frameworks,[6] with many containing nonproteinogenic amino acid monomers. The rapamycin and FK506 family members,[7,8] sanglifehrin,[9] and the epothilones are examples of therapeutically interesting categories of such NRP-PK hybrid molecular frameworks that harbor nonproteinogenic amino acid units.[10,11]

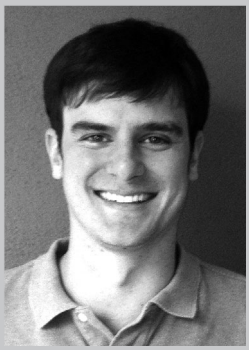
2. Amino Acid Selection, Activation, and Incorporation for Ribosomal versus Nonribosomal Peptide Biosynthesis

In both NRP and hybrid NRP-PK scaffolds, the amino acid monomers—proteinogenic and nonproteinogenic—are selected and activated by 50 kDa proteins, sometimes free-standing, but most often joined in multidomain proteins. These activation domains are termed adenylation (A) domains for the reactions they perform on their amino acid substrates.[2] Each A domain selects a specific amino acid. Once bound to the active site, the enzyme utilizes the carboxylate group of the amino acid to attack the α-phosphate of a bound ATP cosubstrate, thereby yielding a bound aminoacyl-adenylate (aminoacyl-AMP; Scheme 1). This is exactly the mechanism for the first step of aminoacyl-tRNA synthetases that activate the amino acids used in protein biosynthesis: they are also amino acid adenylating enzymes.[12] However, structural analysis suggests, notwithstanding a few exceptions,[13] that the A domains and aminoacyl-tRNA synthetases evolved independently in most cases. Therefore, they are not readily interchangeable, for example, in protein biosynthesis replacement studies.

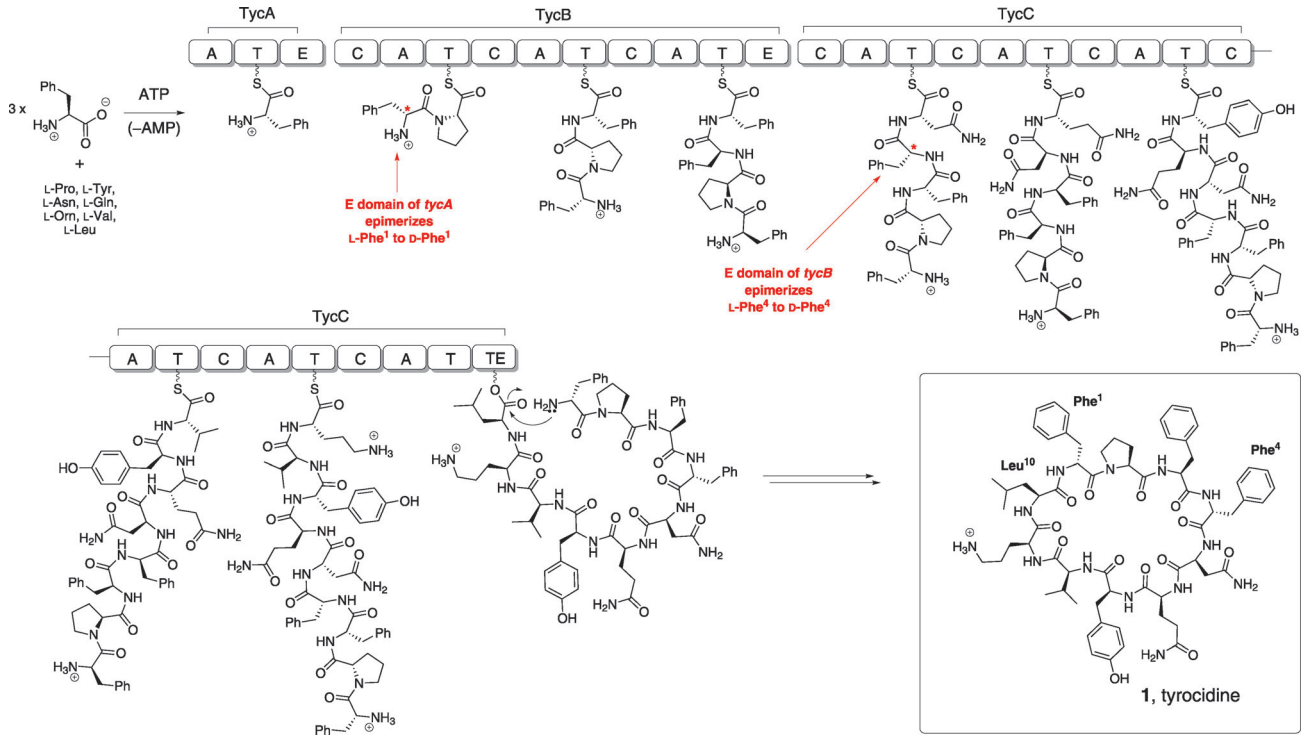
In the second step, the two enzyme classes diverge with respect to the nature of the cosubstrate to which the activated aminoacyl group is now transferred (Scheme 1). For the aminoacyl-tRNA synthetases the cosubstrate is a cognate tRNA and the attacking nucleophilic group is the 3′- or 2′-OH of the ribose ring of the terminal adenine in the CCA tail of each tRNA.[14] The resultant aminoacyl-tRNA is released and is subsequently chaperoned to the large subunit of the ribosome for mRNA-instructed protein biosynthesis.[14] In contrast, for the NRPS adenylation domains, the cosubstrate in the second step is the HS-pantetheinyl arm of the holo form of an 8–10 kDa thiolation (T) domain.[15] The T domains are often in a cis position to the A domains, as a contiguous downstream region of 80–100 residues following the approximately 500 residue A domains (A-T didomains by themselves or as parts of larger proteins). Transfer to the nucleophilic thiolate of the pantetheinyl arm yields aminoacyl thioesters tethered to the T domain.[16]

Chain elongation on the ribosome and on the NRPS assembly lines also has parallels and distinctions (Scheme 1). On the ribosome, an incoming aminoacyl-tRNA moiety sits next to the growing peptidyl-tRNA, with occupancy of the incoming monomer determined by triplet coding between the

Christopher T. Walsh, born in 1944, studied biology at Harvard and completed his PhD in biochemistry in the group of Fritz Lipmann at The Rockefeller University. 1972–1987 he was on the MIT Faculty, and since then has been at Harvard Medical School. He served as Chair of the Department of Chemistry at MIT (1982–1978) and of the Department of Biological Chemistry & Molecular Pharmacology at Harvard Medical School (1987–1995). His research interests focus on the molecular logic and enzymatic machinery of peptide-based natural product biosynthesis.



Robert V. O'Brien, born in 1981, studied at St. John's College in Annapolis and at Seattle University. He received his PhD in Chemistry from Boston College under the guidance of Prof. Amir H. Hoveyda on Ru- and Mo-catalyzed olefin metathesis, and the application of Z-selective cross-metathesis to the synthesis of the antitumor agent KRN7000. His current research at Stanford University involves the production of erythromycin analogues by biosynthesis and chemical modification.



Scheme 2. Tyrocidine and its NRPS. A = adenylation domain, T = thiolation (peptidyl carrier protein) domain, C = condensation domain, E = epimerase domain, TE = thioesterase domain.

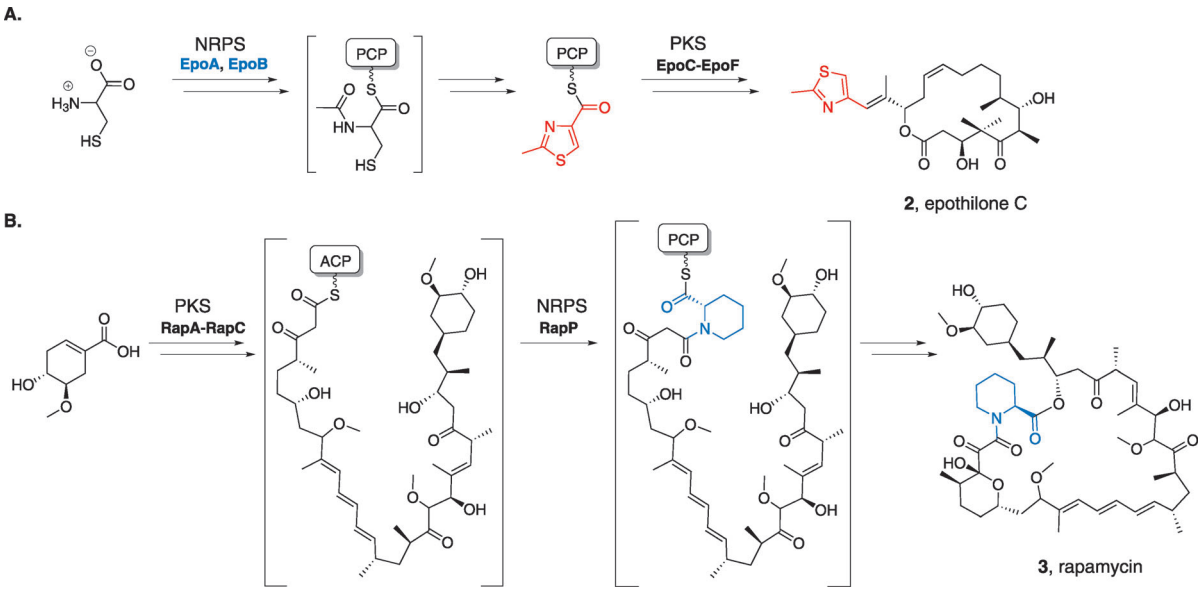
mRNA codon and the tRNA anticodon. Once the proper aminoacyl-tRNA has been recognized and properly situated by the action of conditional GTPases, peptide-bond formation can occur in the peptidyl transferase center of the 50S ribosomal subunit, mediated by the 23S rRNA. The nucleophile is the N-terminal amino group of the peptidyl-tRNA. The electrophile is the activated carboxy oxygen ester of the aminoacyl-tRNA. Peptide-bond formation generates an elongated peptidyl-tRNA and the deacylated tRNA that had delivered the incoming aminoacyl group as a coproduct. In contrast, peptide-bond formation in the mRNA-independent NRPS assembly lines is driven by an approximately 50 kDa condensation (C) domain. Again, these domains are most often encoded in a cis arrangement such that a typical NRPS elongation module has the composition C-A-T, and would be about 110 kDa, and consists of three (semi)-autonomously folded domains.[2,17] The specificity of monomer incorporation into the growing peptide chain is not determined by the mRNA triplet code, but rather by which NRPS modules interact with each other sequentially.

A prototypical NRPS responsible for tyrocidine (**1**) biosynthesis is shown in Scheme 2.[18] As is evident from this example, in many cases 3–5 modules are encoded in a single protein and chain elongation proceeds from the N-C direction.

For hybrid NRP-PK natural products, NRPS modules in particular have evolved to interface with polyketide synthase assembly lines.[17,19] One can find both possible orientations of NRPs and PKS modules in naturally occurring assembly lines. When an NRPS module is upstream of a PKS module, as is the case in epothilone C (**2**, Scheme 3 A) biosynthesis, the aminoacyl-S-T domain must be recognized by and transferred to the first domain of the downstream PKS module.[20] Typically these are ketosynthase (KS) domains, which make C–C bonds by decarboxylative Claisen condensations. In the alternative configuration, where an NRPS module is downstream of a PKS module, as in rapamycin (**3**, Scheme 3 B) and FK506 assembly,[8] the C domain of that NRPS module must be able to recognize the polyketidyl-S-T domain immediately upstream.[21] Now the attacking nucleophile is the amino group of the aminoacyl-S-T domain in the NRPS module. Bond formation creates an amide linkage as the polyketidyl chain is transferred onto the NRPS module.[20,22]

There are examples of hybrid assembly lines where NRPS domains are flanked both upstream and downstream by PKS domains (e.g. epothilone)[20] and the reverse (e.g. bleomycin,[23] yersiniabactin).[24] One of the virtues of hybrid assembly lines is that they can build complex scaffold

Chaitan Khosla, born in 1964, received his PhD in chemical engineering at Caltech in 1990. After postdoctoral studies at the John Innes Centre, UK, he joined the faculty of Stanford University in 1992. He is currently a professor in the Departments of Chemistry and Chemical Engineering, and also directs the Chemical Biology Institute at Stanford. His research focuses on understanding and engineering assembly line polyketide biosynthesis.



Scheme 3. Epothilone C and rapamycin biosynthesis. A) NRPS-PKS pathway. B) PKS-NRPS pathway. Nonproteinogenic amino acids: 2-methylthiazole-4-carboxylate, L-pipecolate. See hyperlinks: EpoA: AAF62880.1; EpoB: AAF62881.1; and Scheme S1 in the Supporting Information.

architectures and mix a variety of oxygen and nitrogen functionalities into the frameworks of the natural products. Advances in the structural determination of domains and whole modules in PKS and NRPS assembly lines has begun to give insights into engineering new PKS-NRPS connectivities and altered specificities of building blocks.[25,26]

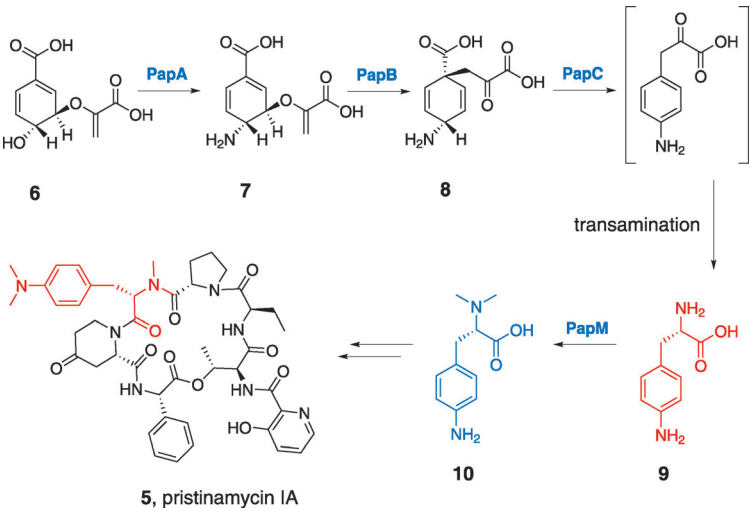
2.1. Para-Amino-Phe Biosynthesis in Streptomyces pristinaespiralis

The noncanonical *para*-dimethylaminophenylalanine occurs in pristinamycin IA, an antibiotic produced by *S. pristinaespiralis*.[27] This building block arises from *p*-NH_2-Phe (by S-adenosylmethionine-mediated dimethylation), which is also a building block for chloramphenicol assembly in *Streptomyces venezuelae*.[28] Phe is not the direct precursor of *p*-NH_2-Phe; instead it arises by diversion of the metabolic flux from chorismate away from prephenate (Scheme 4). In keeping with NRPS assembly logic, four *pap* genes are found in tandem next to the NRPS and PKS genes that build pristinamycin IA (**5**, Scheme 4) scaffolds in the *S. pristinaespiralis* genome.[29,30]

PapA is a glutamine-dependent chorismate amidotransferase, which converts chorismate (**6**, Scheme 4) into 4-amino-4-deoxychorismate (**7**, Scheme 4)—the key intermediate formed in *p*-aminobenzoic acid (PABA) biosynthesis.[31] However, in this case, PapB intervenes and the enolpyruvyl side chain is not eliminated. Instead, PapB is an aminodeoxychorismate mutase that generates 4-amino-4-deoxyprephenate (**8**, Scheme 4), presumably by a 3,3-electrocyclic rearrangement. This undergoes an aromatizing decarboxylation/dehydrogenation by PapC to produce 4-aminophenylpyruvate, which in turn undergoes transamination to yield 4-amino-L-Phe (**9**). PapM can methylate the amino group once or twice to yield the mono- or dimethylamino-L-Phe (**10**) monomer, either of which can be incorporated into pristinamycins.

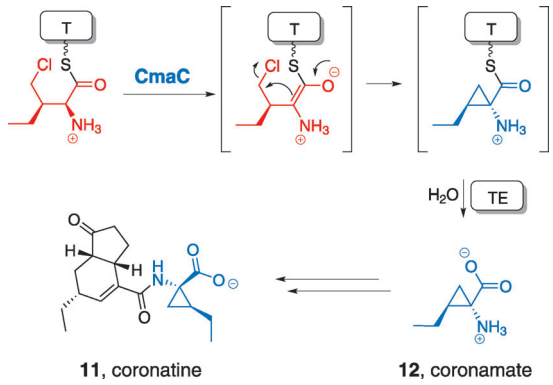
2.2. Incorporation of Nonproteinogenic Amino Acids into Proteins

Schultz and co-workers moved the *papABC* gene cluster from *S. venezuelae* into an *Escherichia coli* strain into which they had inserted an evolved tRNA/tRNA synthetase pair



Scheme 4. *p*-Aminophenylalanine biosynthesis in the pristinamycin IA pathway. Nonproteinogenic amino acids: L-(*p*-amino)-Phe (**9**) and L-(*p*-(*N*,*N*-dimethylamino)-Phe (**10**). See hyperlinks: PapA: AAC44866.1; PapB: AAC44868.1; PapC: AAC44867.1; PapM: AAC44869.1.

that would recognize *p*-NH_2-Phe and incorporate it site-specifically in myoglobin through the amber-suppressor codon.[32] In principle, the strategy for *p*-NH_2-Phe incorporation could be generalized to any of the other nonproteinogenic amino acid building blocks described in this Review, as long as the genes are known and the proteins are expressed in an active form in a bacterial host (e.g. *E. coli*) harboring the engineered tRNA/tRNA synthetase pair to recognize the corresponding building block.[33] In practice, as discussed in the following sections, many nonproteinogenic amino acids required for NRP or mixed PK/NRP biosynthesis are generated as aminoacyl thioesters tethered to a protein-bound pantetheine arm. Once synthesized, these building blocks are directly utilized by the NRPS. Hijacking these unnatural amino acids for ribosomal protein synthesis will, therefore, require an engineered thioesterase. Fortunately, nature harbors a number of these aminoacyl thioesterases: for example, the nonproteinogenic amino acid coronamate is produced as a protein-bound thioester, which is hydrolyzed prior to incorporation into the natural product coronatine (**11**, Scheme 5).[34]



Scheme 5. Coronamate biosynthesis and incorporation in the coronatine pathway. Nonproteinogenic amino acids: **(2*R*)-2-amino-3-(chloromethyl)pentanoic acid**, **coronamate** (**12**). See hyperlink: CmaC: NP 794455.1.

3. Biosynthesis of Nonproteinogenic Amino Acids

3.1. ORFs within NRPS Clusters

In this Review we focus on the unusual amino acid building blocks, the nonproteinogenic ones, which are fashioned by producer microbes for specific incorporation into the known classes of molecular scaffolds. Some of these nonproteinogenic amino acids are fashioned from the canonical 20 proteinogenic amino acids, while others are biosynthesized de novo. For any nonproteinogenic amino acid with a known biosynthetic pathway, a hyperlink is provided within each scheme legend that links to the relevant GenBank entry for each enzyme in the biosynthetic pathway. In cases where the biosynthetic pathway is unknown, a link to the Supporting Information is included, which contains hypothetical biosynthetic pathways for pedagogical purposes; links are also provided for several schemes for known biosynthetic pathways that are contained within the Supporting Information.

In most cases, the genes encoding these dedicated amino acid building blocks are adjacent to genes for the NRPS assembly line modules, thus enabling coordinate regulation to make the building blocks as they are needed. This feature greatly simplifies the task of finding open reading frames (ORFs) that encode enzymes dedicated to nonproteinogenic amino acids and then moving them to production organisms. For example, as discussed above, the genes encoding *p*-dimethylaminophenylalanine are clustered in the pristinamycin IA producer (Scheme 4).[29]

3.2. Major Classes of Reactions That Build Nonproteinogenic Amino Acids

Most of the almost 500 nonproteinogenic amino acids are produced by a small number of enzymatic reactions, either from preexisting proteinogenic amino acids or by de novo construction. Examples discussed below include: 1) racemases that make D-amino acids from the L enantiomers;[35,36] 2) mutases that convert α-amino acids into β-amino acids;[37,38] 3) reactions that produce β-methyl-substituted amino acids and hydroxylases that install OH groups, often as β-OH substituents;[39,40] 4) oxygen-dependent halogenases that install chlorine substituents on both activated and unactivated carbon centers;[41] 5) de novo construction of phenylglycines with both 4-OH- and 3,5-$(OH)_2$ substitution patterns, which are shown to arise by entirely different molecular logic;[42,43] 6) biosynthesis of α,ω-diamino acids, such as 2,3-diaminopropionate (DAP) and 2,4-diaminobutyrate (DAB);[44–46] 7) several distinct types of cyclic nonproteinogenic amino acids are discussed, from three-membered-ring cyclopropanes (e.g. coronomate (**12**), Scheme 5), four-membered-ring tabtoxin β-lactam,[47] to five-membered-ring enduracididine,[48] to six-membered-ring capreomycidine,[49] as well as pipecolate and piperazate rings; and 8) the transformation of tryptophan to kynurenine,[50] and of tyrosine to propyl- and propenylprolines is analyzed.[51]

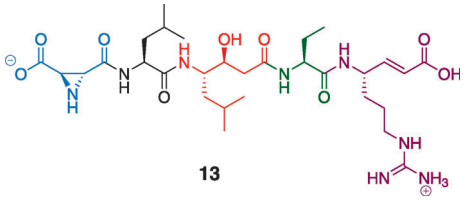
Some of the carbon skeletons of noncanonical amino acids, such as 2-amino-3-hydroxy-4-methyl-6-octenoate, 2-amino-8-oxo-9,10-decanoate, and 3-amino-9-methoxy-2,6,8-trimethyl-10-phenyl-4,6-decadienoate, are fashioned by PKSs as free monomers, which undergo late-stage enzymatic reductive amination.[52] Other amino acids, such as statine and isostatine, are synthesized on hybrid NRPS-PKS assembly lines and directly incorporated into the growing chains.[53] Comparable hybrid logic and enzymatic machinery pertains for building vinyl-Arg and vinyl-Tyr units.[54]

To illustrate particular molecular contexts for major nonproteinogenic amino acid classes, 15 NRP scaffolds and an additional 4 hybrid NRP-PK scaffolds are introduced in the next section. Together, these 19 frameworks contain about 50 of the most common nonproteinogenic building blocks, representative of almost all the major classes of amino acid scaffolds. We then review the details of how those and related nonproteinogenic building blocks are generated within the producer microbes.

4. Nonproteinogenic Amino Acid Residues in NRP Frameworks

4.1. Linear Peptide Frameworks

Nonproteinogenic building blocks are found in linear NRP scaffolds, such as in the sponge-associated miraziridine A (**13**, Figure 1), probably produced by bacteria in the consortium.[55] This linear pentapeptide has four nonproteinogenic residues out of a total of five. In addition to aminobutyrate (Aba),[56] the next longer homologue of Ala, there are three remarkable building blocks: an aziridine dicarboxylate (Azd),[57] statine,[53] and vinyl-Arg.[54] The three-membered aziridine ring is unusual in biology and likely a reactive chemical functionality.[58,59] Statine and vinyl-Arg are synthesized by hybrid NRPS-PKS assembly lines, as explained in Section 6.1.

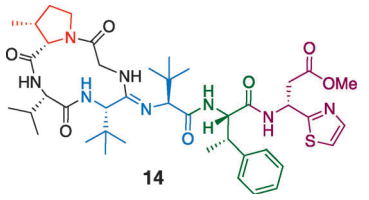


13

Figure 1. Miraziridine A. Nonproteinogenic amino acids: (3*S*, 4*S*)-statine, (2*R*, 3*R*)-aziridine-2,3-dicarboxylate, L-α-aminobutyrate, (*S*)-vinyl-Arg. See Schemes S2–S5 in the Supporting Information.

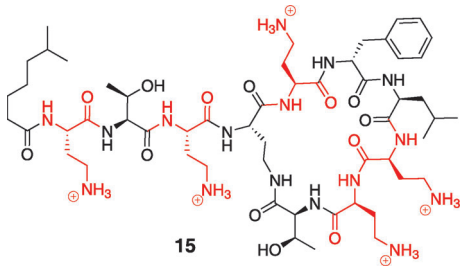
4.2. Cyclic Peptide Frameworks

There is sustained interest in cyclic frameworks whose conformational constraints generate fixed three-dimensional molecular architectures that can produce high-affinity ligands for biological targets.[60] Here we discuss 15 cyclic peptides with 3–15 residues in their macrocyclic frameworks. Ring closing of such cyclic peptides can be achieved through formation of macrolactam (cyclic amides), macrolactone (cyclic esters), cyclic amidine (as in the case of bottromycin A_2 (**14**), Figure 2),[61] or through formation of an aryl ether bridge. As seen in Scheme 2, the decapeptidic macrolactam of tyrocidine (**1**) is formed by head-to-tail condensation of Phe1 and Leu10,[62] whereas the polymyxin nonapeptides (such as polymyxin B2 (**15**), Figure 3) form stem–loop macrolactam structures in which a side-chain amine from 2,4-DAB is the nucleophile.[46,63] Macrolactone linkages (also known as depsipeptides) can arise from the side-chain hydroxy group of Thr, as in callipeltin B (**16**, Figure 4)[64,65] and daptomycin (**17**, Figure 5),[50,66] or Ser, as in enterobactin.[67]



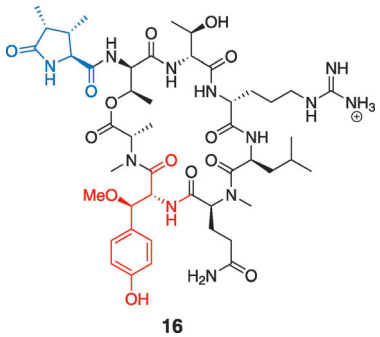
14

Figure 2. Bottromycin A_2. Nonproteinogenic amino acids: (2*S*,3*R*)-3-methyl-Pro, L-*tert*-butyl-Gly, (2*S*,3*S*)-β-methyl-Phe, (*R*)-methyl 3-amino-3-(thiazol-2-yl)propanoate.



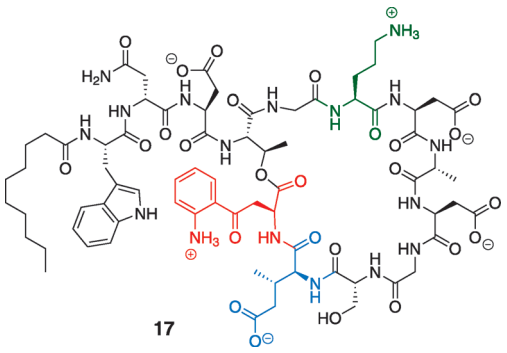
15

Figure 3. Polymyxin B2, a stem–loop macrolactam. Nonproteinogenic amino acid: L-2,4-diaminobutyrate (2,4-DAB). See Scheme S6 in the Supporting Information.



16

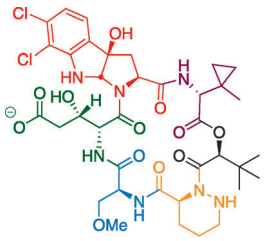
Figure 4. Callipeltin B, a Thr-linked depsipeptide. Nonproteinogenic amino acids: (2*R*,3*R*)-β-methoxy-Tyr, (2*S*,3*S*,4*R*)-3,4-dimethyl-5-oxo-Pro. See Scheme S7 in the Supporting Information.



17

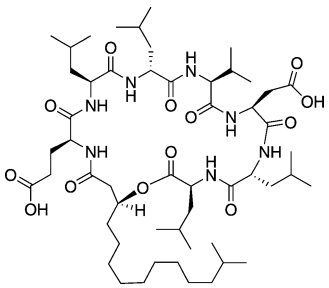
Figure 5. Daptomycin. Nonproteinogenic amino acids: L-kynurenine, (3*S*,4*S*)-β-methylglutamate, L-ornithine. See Scheme S8 in the Supporting Information.

Alternatively, an α-hydroxy acid building block incorporated in place of one of the amino acids can form a macrolactone, as in the kutznerides (e.g. kutzneride 1 (**18**), Figure 6).[68] A variant occurs in the N-acylated lipopeptide surfactin (**19**, Figure 7), where a β-OH fatty acyl chain participates in macrolactone formation.[69]



18

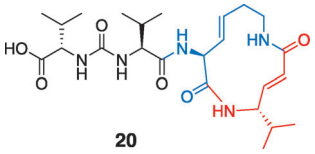
Figure 6. Kutzneride 1, a cyclic hexa(depsi)peptide. Nonproteinogenic amino acids: dichloropyrroloindoline carboxylate, L-*O*-methyl-Ser, (2*R*,3*S*)-β-hydroxy-Glu, methylcyclopropyl-Gly (MeCPGly), piperazate. See Schemes S9–S11 in the Supporting Information.



19

Figure 7. Surfactin, a β-hydroxycarboxylic lactone.

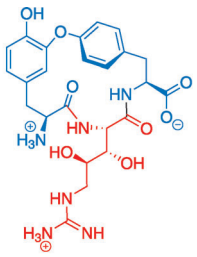
Among the smallest natural peptide macrocycles is the 12-atom macrolactam found in the proteasome inhibitor syringolin A (**20**, Figure 8).[70,71] It has a stem–loop structure with a macrocyclic lactam derived from only two residues, both noncanonical. One is a Δ^3-lysene that may arise from dehydration of a 3-OH-Lys residue.[70] The other is a vinyl-Val formed by an NRPS-PKS hybrid module (see Section 6.1). The conjugated enamide derived from vinyl-Val is the electrophilic moiety of syringolin, which irreversibly inactivates the proteasome by capture of the active-site Thr_1-NH_2.[72]



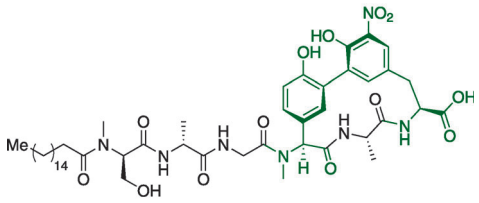
20

Figure 8. Syringolin A. Nonproteinogenic amino acids: L-vinyl-Val, L-Δ-3-Lysene. See Scheme S12 in the Supporting Information.

Eurypamide A (**21**, Figure 9) is a three-residue peptide macrocycle, isolated from a presumed bacterial community in Palauan sponges,[73] containing a Tyr-O-Tyr cross-link similar to the analogous Tyr-C-Tyr cross-link in arylomycin antibiotics (e.g. arylomycin B-C_{16} (**22**), Figure 9).[74,75] The side-chain aryl ether in **21** is thought to arise from a one-electron phenoxyl radical reaction on the two tyrosine side chains of a tripeptide precursor catalyzed by an iron-based oxygenase.[76] These cross-links set the constrained architecture in **21**



21, eurypamide A

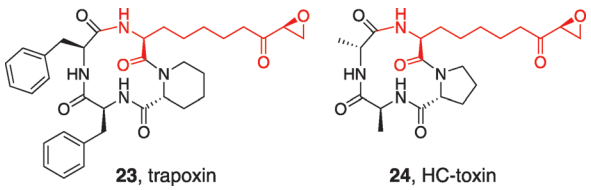


22, arylomycin B-C_{16}

Figure 9. Aryl ethers and aryl–aryl linkages in cyclic peptides. Nonproteinogenic amino acids: (2*S*,3*R*,4*S*)-3,4-dihydroxy-Arg, Aryl ether (L-Tyr/(*R*)-β-Tyr), Aryl ether (L-(*p*-hydroxy)phenyl-Gyl/L-*m*-nitro-Tyr). See Supporting Scheme S13 in the Supporting Information.

and **22** (Figure 9).[75,77] Eurypamide A (**21**) also harbors the unusual 3*R*,4*S*-dihydroxy-L-Arg moiety,[77] although the timing of the bishydroxylation is unknown. It could occur on the free amino acid, the presumed aminoacyl-S-T NRPS assembly line intermediate, or the released peptide.

Macrocyclic tetrapeptides are also found in the histone deactylase (HDAC) inhibitors of the trapoxin (**23**) and HC-toxin class (**24**, Figure 10).[78,79] This compact architecture presents the epoxyketone side chain of the unusual 2-amino-8-oxo-9,10-epoxydecanoate (Aeo) epoxy residue to the active site of HDACs for irreversible capture of the active-site nucleophile.[80] The strategy for construction of the Aeo monomer is discussed in Section 6.2.



23, trapoxin 24, HC-toxin

Figure 10. Natural products containing L-Aeo. Nonproteinogenic amino acids: L-2-amino-8-oxo-9,10-decanoate (Aeo). See Scheme S14 in the Supporting Information.

Cyclic pentapeptide scaffolds are found in the antitubercular drugs capreomycin (**25**, Figure 11) and viomycin (**26**, Figure 11).[49] They are notable for the macrolactam ring closure being achieved with a 2,3-DAP residue,[44] a β-lys residue in the stem of the stem–loop structure, the cyclic arginine derivative capreomycidine, and a ureidodehydroAla residue. These unusual amino acids play important roles in the ability of both antibiotics to bind across the interface of the 30S and 50S subunits of bacterial ribosomes.[81] The

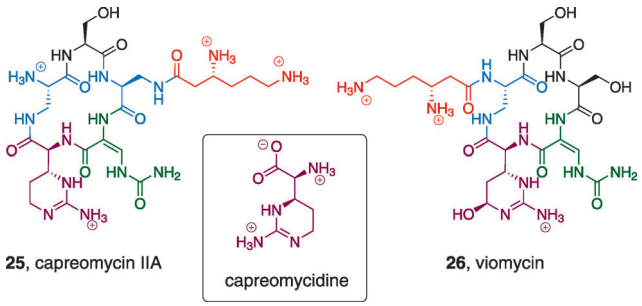
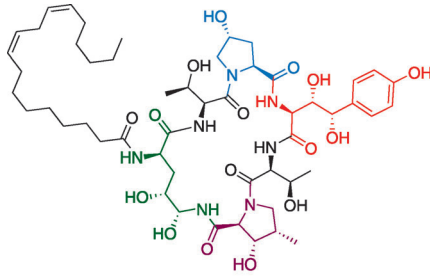


Figure 11. Natural products containing a capreomycidine moiety. Nonproteinogenic amino acids: β-Lys, L-2,3-DAP, (*E*)-ureidodehydro-Ala, capreomycidine and δ-hydroxycapreomycidine. See Scheme S15 in the Supporting Information.

biosynthesis of the β-Lys and capreomycidine building blocks are discussed in Sections 5.1.2 and 5.6.3, respectively.

Cyclic hexa(depsi)peptides include the kutzneride family of hexapeptidolactones (e.g. **18**, Figure 6)[82] with antifungal activity, and the echinocandins (e.g. echinocandin B (**27**), Figure 12),[83,84] which target the fungal 1,3-glucan synthases.

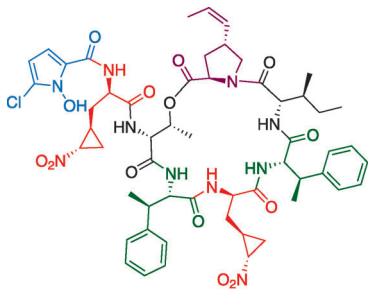


27

Figure 12. Echinocandin B. Nonproteinogenic amino acids: (2*S*,3*S*,4*S*)-3,4-dihydroxyhomo-Tyr, (2*S*,4*R*)-4-hydroxy-Pro, (2*R*,4*R*,5*R*)-4,5-dihydroxy-Orn, (2*S*,3*S*,4*S*)-3-hydroxy-4-methyl-Pro. See Scheme S16 in the Supporting Information.

Kutznerides have five nonproteinogenic amino acid building blocks and an unusual hydroxy acid (*tert*-butylglycolate) that participates in forming the macrolactone framework.[68] The echinocandins have a hexapeptide scaffold with a metastable hemiaminal linkage between N_5 of an N_1-acylated ornithine and the carbonyl group of a 3-methyl-4-hydroxyproline residue. Both the ornithine side chain and an unusual homotyrosine side chain are dihydroxylated; the timing of the hydroxylations is unknown.

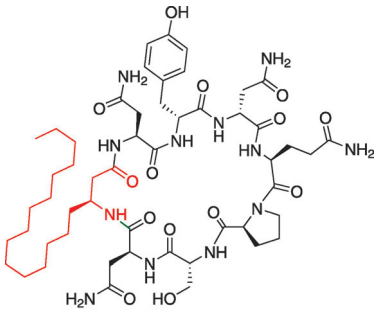
Hormaomycin (**28**, Figure 13) is another hexapeptidolactone that is also replete with noncanonical amino acid units, including two β-methyl-Phe residues, a propenylproline, and a remarkable nitrocyclopropylalanine residue.[85] The OH group of a Thr residue participates in the formation of the hexapeptidolactone macrocycle; its amino group is acylated with a dipeptide moiety containing another nitrocyclopropyl-Ala and a proline-derived *N*-OH-2-chloropyrrole carboxyl residue. A schematic representation of the chain elongation of this remarkable natural product is shown in Scheme S17 in



28

Figure 13. Hormaomycin. Nonproteinogenic amino acids: Nitrocyclopropyl-Ala, 5-chloro-1-hydroxypyrrole-2-carboxylate, (2*S*,3*R*)-β-methyl-Phe, (2*R*,4*R*)-4-((*Z*)-propenyl)-Pro. See Schemes S17–S20 in the Supporting Information.

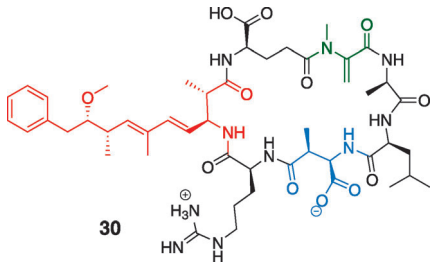
the Supporting Information. Heptapeptidyl frameworks are found in several natural products, including the N-acylated mycosubtilin (**29**, Figure 14),[86] and surfactin lipopeptides



29

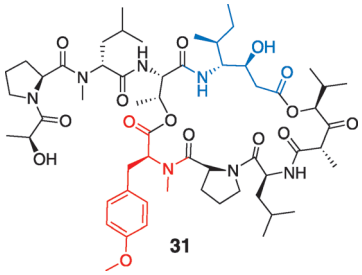
Figure 14. Mycosubtilin. Nonproteinogenic amino acid: (*S*)-β-aminostearate.

from *Bacillus* spp (see Figure 7).[69] The β-hydroxy substituent of the fatty acyl chain engages in macrolactonization in surfactin, while the fatty acyl chain in mycosubtilin instead bears a β-amino group that makes a macrolactam. Callipeltin B (**16**, Figure 4) is a sponge-derived heptadepsipeptide with a D-*allo*-Thr residue and also a β-methoxy-Tyr residue.[64] The antibacterial polymyxins (e.g. **15**, Figure 3) have multiple 2,4-DAB residues, one of which participates in the heptapeptidyl macrolactam loop.[63] These side chains are cationic under physiological conditions and may serve to sequester this antibiotic by interaction with anionic lipopolysaccharides on Gram-negative bacterial pathogens.[87] Cyanobacteria make a variety of toxic cyclic peptides through a nonribosomal strategy;[88] among them are the microcystins (e.g. microcystin LR (**30**), Figure 15), cyclic heptapeptides that target protein phosphatase I for inhibition.[89] More than 70 structural variants are known, thus reflecting the promiscuity in the heptamodule NRPS assembly lines. Among its unusual residues are dehydroalanine, β-Me-D-isoaspartate, the polyketide-derived Adda, and isoglutamate. The didemnins (e.g. didemnin B (**31**), Figure 16) are cyclic heptadepsipeptides isolated from bacterial symbionts of sea squirts,[90] and have



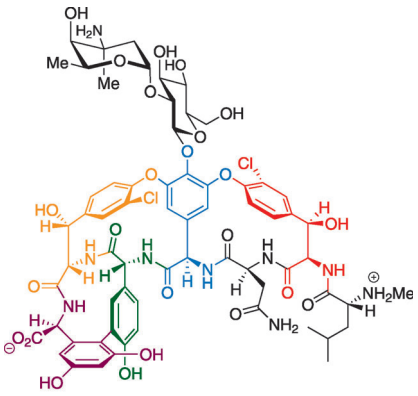
30

Figure 15. Microcystin LR. Nonproteinogenic amino acids: vinylogous β-Ala, (2*R*,3*S*)-β-methylaspartate, dehydro-Ala.

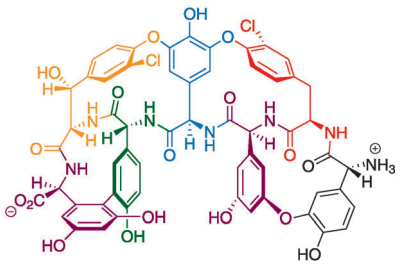


31

Figure 16. Didemnin B. Nonproteinogenic amino acids: L-*O*-Me-Tyr, (3*S*,4*R*,5*S*)-isostatine.



32, vancomycin



33, teicoplanin agylcone

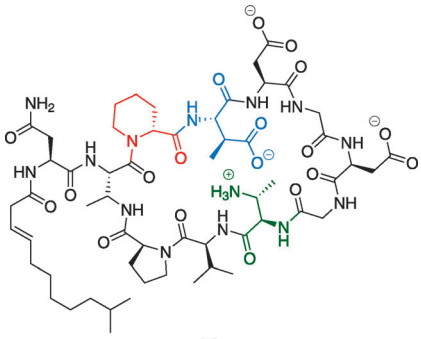
Figure 17. Vancomycin and teicoplanin. Nonproteinogenic amino acids: (2*R*,3*R*)-*m*-chloro-(or β-hydroxy-*m*-chloro)Tyr aryl ether, 3,5-dihydroxy-Tyr aryl ether, *p*-hydroxyphenyl-Gly aryl-aryl, 3,5-dihydroxyphenyl-Gly aryl-aryl or aryl ether, (2*S*,3*R*)-β-hydroxy-*m*-chloro-Tyr aryl ether. See Schemes S21–S23 in the Supporting Information.

been of interest for their immunosuppressive properties.[91] Among the building blocks of special interest is the isostatine residue, with a 3-OH, 4-NH_2 grouping, assembled by a hybrid NRPS-PKS module discussed in Section 6.1.

A distinct mode of converting a linear heptapeptidyl framework into a constrained architecture occurs in the maturation of the T_7-domain-tethered linear heptapeptidyl intermediate in the assembly of the antibiotic vancomycin (**32**, Figure 17).[92,93] Three heme iron-containing monooxygenases act to generate phenoxyl radical intermediates to couple residues 2–4, 4–6, and 5–7.[94,95] Each of these five participating residues is nonproteinogenic, and has an accessible redox potential for a radical coupling reaction. Residues 2 and 6 are β-OH-Tyr. Residues 4 and 5 are 4-OH-phenylglycines, and residue 7 is a 3,5-dihydroxyphenylglycine. In teicoplanin (**33** (aglycone), Figure 17) a fourth cross-link occurs between Tyr_1 and 3,5-$(OH)_2$-$Phegly_3$, such that all seven residues participate in a total of four cross-links. The formation of these unusual aryl amino acids is discussed below.

Daptomycin is a clinically used N-acylated 13-residue NRP antibiotic (**17**, Figure 5).[96] It has a 10-residue macrolactone ring, fashioned from the OH group of Thr_4 and the carbonyl group of residue 13, the nonproteinogenic kynurenine, which arises from the oxidative metabolism of tryptophan.[97] It also harbors a nonproteinogenic 3-methyl-Glu residue 11 as one of several side-chain carboxylates that coordinate Ca^{2+} ions and change the conformation to a membrane-active form.

Friulimicin is another nonribosomally produced lipopeptide (**34**, Figure 18),[45] with a decapeptidyl macrolactam between a 2,3-diaminobutyryl residue at position 2 and Pro_{11}. Friulimicin binds the lipid II carrier required for the

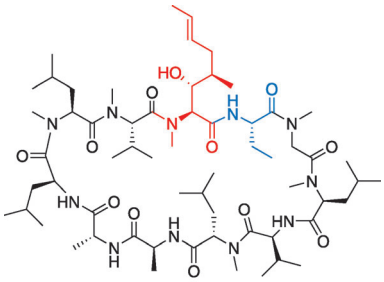


34

Figure 18. Friulimicin. Nonproteinogenic amino acids: D-pipecolate, β-methyl-Asp, 2,3-DAB.

formation of peptidoglycan in the assembly of the bacterial cell wall.[98] Friulimicin also has additional nonproteinogenic amino acid building blocks, including $pipecolate_3$, β-$Me\text{-}Asp_4$, and another 2,3-DAB.

An 11-residue macrolactam is formed in the last step of the biosynthesis of the immunosuppressant undecapeptide cyclosporin A (**35**, Figure 19), when D-Ala_1 is coupled in a head-to-tail fashion to the thioesterified carbonyl group of L-Ala_{11} by the megadalton NRPS cyclosporin synthetase.[99] Seven of the 11 residues of cyclosporin are N-methylated. There is also a nonproteinogenic 2-aminobutyryl residue and

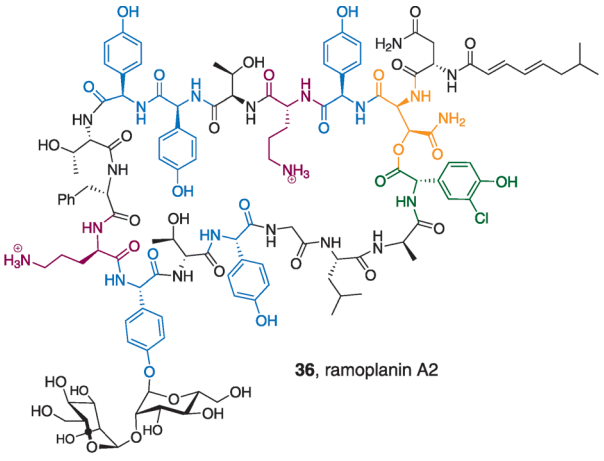


35

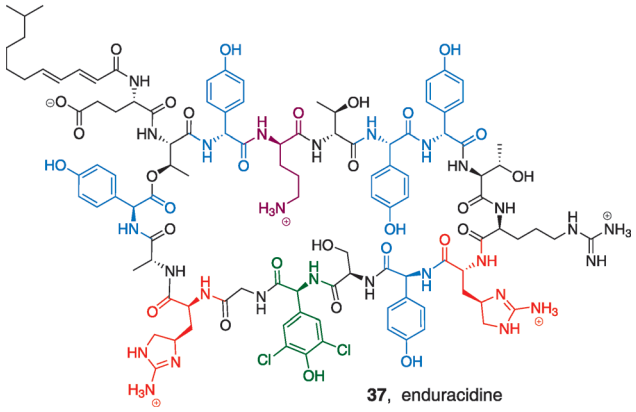
Figure 19. Cyclosporin A. Nonproteinogenic amino acids: (2*S*,3*R*,4*R*)-(*E*)-butenylmethyl-Thr, L-α-aminobutyrate.

a methylbutenylthreonine residue that is generated by a PKS assembly line and a subsequent transamination.[100]

Among the largest nonribosomally generated peptide macrocycles are the ramoplanin A2 (**36**)[101] and related enduracidine (**37**, Figure 20) antibiotics.[48] In both of these lipid II binders, residues 2 and 16 are macrocyclized to yield a 15-residue cyclic framework.[102] Among the noncanonical building blocks are 4-hydroxyphenylglycine residues and, in



36, ramoplanin A2



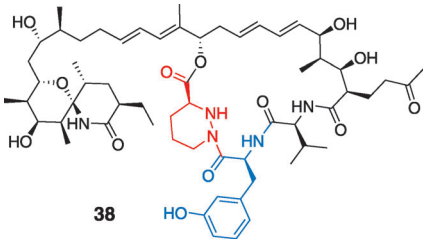
37, enduracidine

Figure 20. Ramoplanin and enduracidine. Nonproteinogenic amino acids: enduracididine, D or L *p*-hydroxyphenyl-Gly, L-3,5-dichloro-4-hydroxyphenyl-Gly or 3-chloro-4-hydroxyphenyl-Gly (in **34**), D-Orn, (2*S*,3*S*)-β-hydroxy-Asp.

enduracidine, two cyclic arginine-derived enduracididine residues.

4.3. Macrocyclic Hybrid Frameworks from Nonribosomal Peptides and Polyketides

Various macrocyclic ring sizes exist in hybrid NRP-PK scaffolds. The PKS machinery can generate fully reduced CH_2 groups, CHOH, HC=CH, or C=O groups from partial processing of tethered intermediates on the PKS assembly lines.[17,103] These three functional groups can mix and match with the amino acid building blocks, as exemplified by the macrocyclic portion of syringolin A (**20**, Figure 8), epothilone C (**2**, Scheme 3), and rapamycin (**3**, Scheme 3). The naturally occurring immunosuppressant sanglifehrin A (**38**, Figure 21) has 22 atoms in the macrocycle, spans 3 amino acids (Val-(*m*-Tyr)-piperazate), and the rest is polyketide.[104] The phytotoxin coronatine (**11**, Scheme 5) is produced by hybrid PKS-NRPS machinery in the plant pathogenic *Pseudomonas syringae* through ligation of a cyclopropyl amino acid to the polyketide fragment coronafacic acid.[34,105,106]



38

Figure 21. Sanglifehrin A. Nonproteinogenic amino acids: L-piperazate, L-*meta*-Tyr.

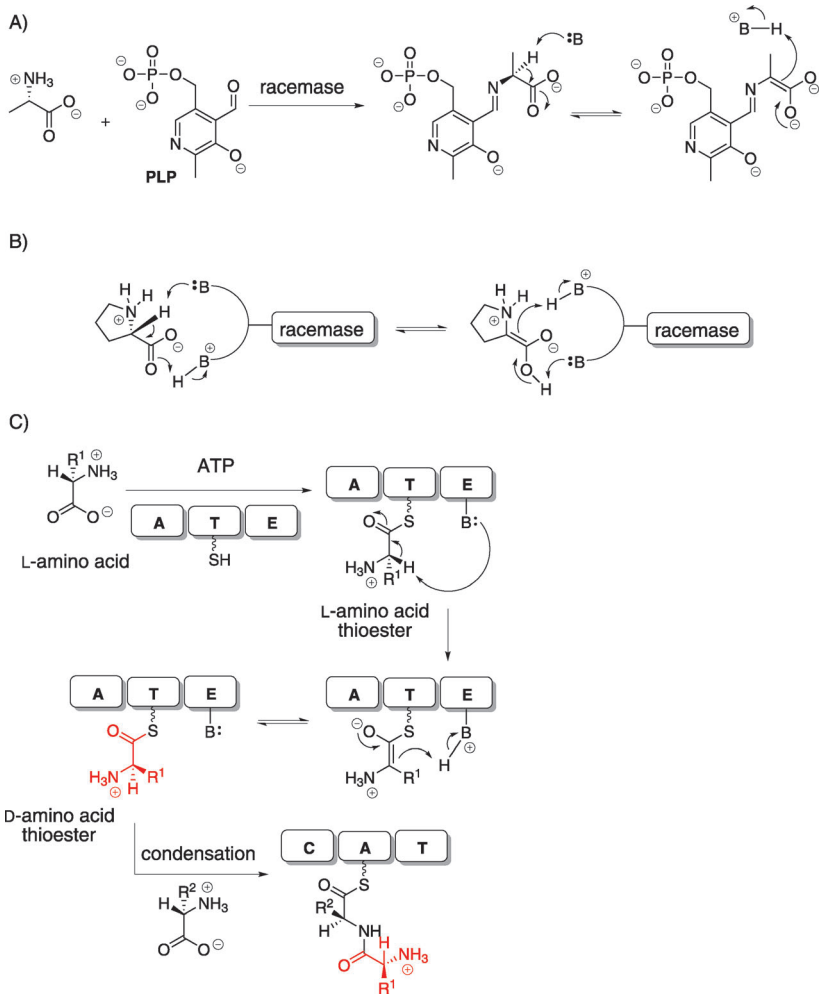
5. Generation of Nonproteinogenic Amino Acid Building Blocks

5.1. D-Amino Acids and β-Amino Acids

As noted above, microcystin LR (**30**, Figure 15) contains a D isomer of β-methyl isoaspartate, and vancomcyin and teicoplanin (**32** and **33**, Figure 17) have D-hydroxyphenylglycine units, which engage in the cross-linking of the rigid scaffold. Cyclosporin has one D-Ala residue (**35**, Figure 19). More generally, many dozens of D-amino acid residues are found in nonribosomal peptide scaffolds. Furthermore, there are β-amino acids, as in the side-chain β-Lys residue in the antitubercular agents viomycin and capreomycin (**25** and **26**, Figure 11), as well as β-Phe in andrimid and β-Tyr in the enediyne antitumor natural products.[107,108]

5.1.1. Racemases and Epimerases

The most common route to D-amino acids is by the action of amino acid racemases and epimerases to convert L isomers into D isomers. For example, D-Ala is generated in bacterial peptidoglycan layers and cyclosporine by prototypic pyri-



Scheme 6. Racemase and epimerase activity. A) PLP-dependent catalysis for racemization. B) Acid/base catalysis for racemization. C) Template-bound catalysis for racemization.

doxal-phosphate (PLP) dependent racemases that generate the stabilized C_2 carbanion from Ala=PLP aldimine that undergoes a stereorandom reprotonation (Scheme 6A).[35,36] On the other hand, the glutamate, proline, and diaminopimelate racemases are PLP-independent and use two active-site bases—one protonated as the conjugate acid, the other deprotonated as the initiating active-site base—to catalyze the stereochemical equilibration at Cα (Scheme 6B).[109] Broad-specificity PLP racemases that act on a range of proteinogenic amino acids have been isolated from *Pseudomonas putida*,[110] as has a threonine epimerase which is representative of diastereomer interconversion (L-Thr to D-*allo* Thr and D-Thr to L-*allo*-Thr).[111]

In many NRPS assembly lines, including those for vancomycin and teicoplanin, there are approximately 50 kDa large epimerization domains which do not act on free amino acids, but only after an L-amino acid has been activated and tethered as the L-aminoacyl thioester.[112] The C_2-H is then kinetically acidic, since the carbanion is stabilized as the thioester enolate (Scheme 6C). The stereorandom return of the hydrogen yields a mix of L- and D-aminoacyl thioesters for subsequent chain elongation by downstream D-specific condensation domains.[113–115]

5.1.2. Aminomutases Interconvert α- and β-Amino Acids

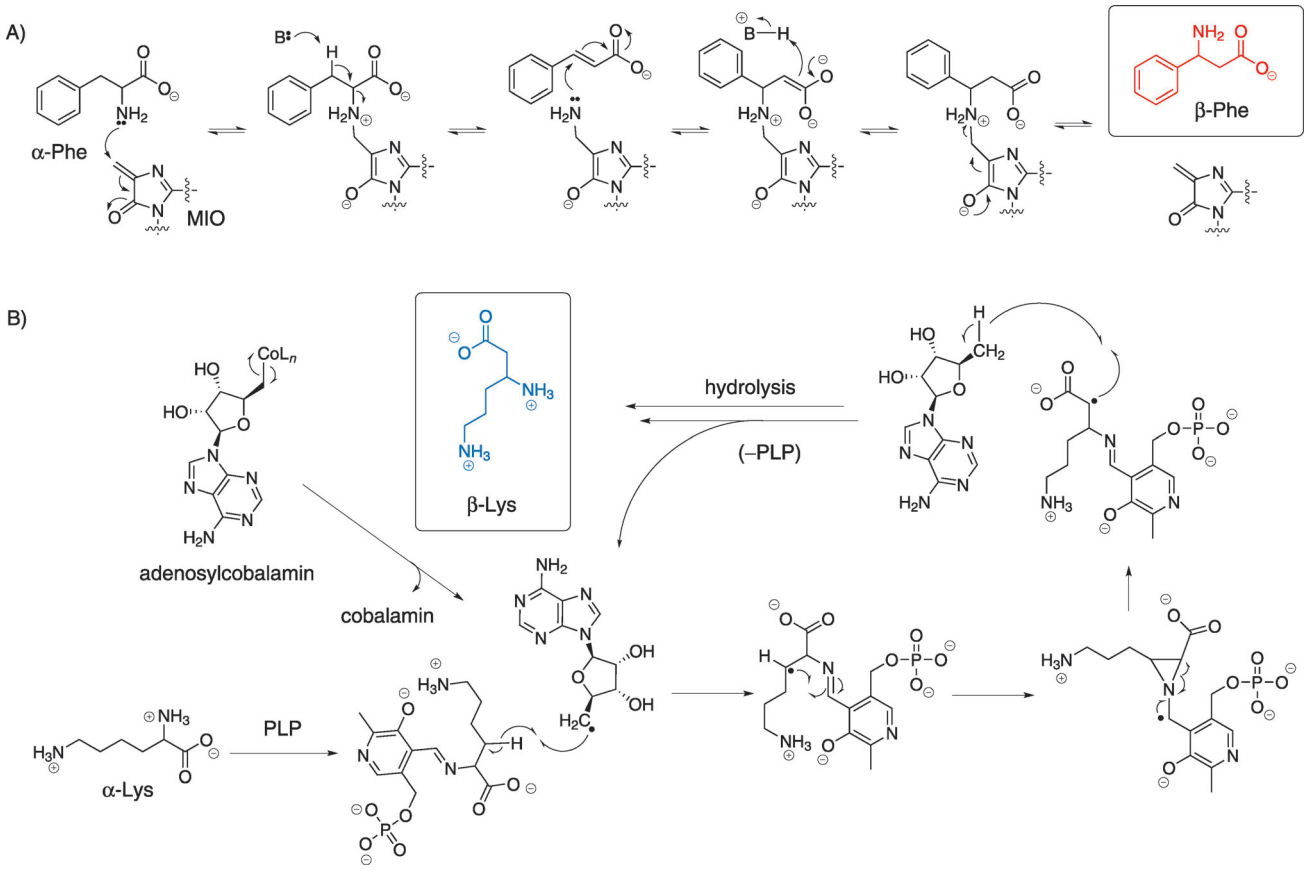
The most general route to β-amino acids is by the action of a family of amino acid mutases on the corresponding α-amino acids. Two types of microbial enzymes have been observed to have such mutase activity.[116,117] One involves conversion of Phe or Tyr into the corresponding β-amino acids. In the case of tyrosine, both 3*R*- and 3*S*-β-tyrosine are produced by mutases from various sources.[37,38] This subfamily of enzymes contain the covalently attached cofactor 4-methylidene-5-imidazole-5-one (MIO), which arises during an automodifying maturation of an Ala-Ser-Gly tripeptide loop in the inactive proenzyme (Scheme 7A). The prosthetic MIO group acts as an electrophilic center for attack by the amino group of the bound amino acid substrate. The resulting adduct can undergo Cβ–H bond cleavage to yield a transient noncovalent olefinic species (e.g. cinnamate when starting from Phe), with the itinerant amino group tethered to the MIO moiety. The amino group is transferred back to the cinnamate by attack at Cβ rather than Cα,[118] followed by elimination of β-Phe and restoration of the starting prosthetic MIO group. All this is catalyzed by a single enzyme.

The second subfamily of amino acid mutases is found in *Clostridium* spp, and encompasses lysine-2,3-aminomutase, glutamate 2,3-aminomutase, and lysine 5,6-aminomutase.[117] This class of enzymes utilizes both pyridoxal phosphate and adenosyl-B_{12} as coenzymes (Scheme 7B). The PLP forms its typical aldimine with the substrates, and the B_{12} coenzyme mediates radical intermediates in the -H for -NH_2 exchanges. Also in this enzyme class are D-ornithine 4,5-aminomutase (which yields (2*R*,4*S*)-diaminopentanoate) and lysine-5,6-aminomutase, each of which acts on the distal rather than the proximal amine of these dibasic amino acids. The second enzyme will also process (3*S*)-β-Lys to (3*S*,5*S*)-diaminohexanoate. Mutant forms of lysine 2,3-aminomutase can produce β-alanine from Lys as a biotechnological alternative to the aspartate decarboxylase route. The β-Lys in capreomycin and viomycin may arise by this kind of enzymatic action (**25** and **26**, Figure 11).

5.2. C-Methyl-, O-Methyl-, and N-Methyl Amino Acids

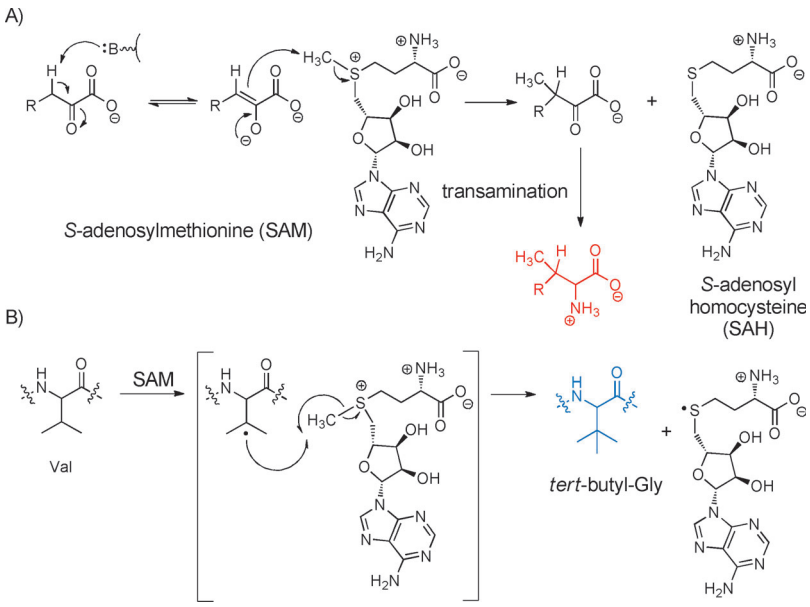
5.2.1. Methylations at Cβ

Friulimicin contains a β-methyl-Asp residue (**34**, Figure 18), daptomycin a β-methyl-Glu residue (**17**, Figure 5), and hormaomycin two β-Me-Phe residues (**28**, Figure 13). In the cases that have been investigated to date for



Scheme 7. Interconversion of α- and β-amino acids. A) MIO-mediated catalysis. B) Catalysis with radical intermediates.

the timing and mechanism, Cβ-methylations on residues incorporated through the NRPS assembly lines occur at the α-keto acid stage.[119] This strategy allows the use of the abundant methyl donor SAM to transfer its "CH_3^+" equivalent to a Cβ carbanion equivalent. Subsequent diastereoselective transamination of the 3-methyl-2-keto acids yields the β-branched amino acids (Scheme 8A). These transformations occur prior to activation of the amino acids by the corresponding NRPS. Multiple methylations on pyruvate can generate the dimethyl and trimethyl versions of the keto acid and yield either the corresponding amino acid (*tert*-butylglycine) or, by reduction, *tert*-butylglycolate, as found in kutznerides (e.g. **18**, Figure 6). In contrast to the de novo assembly of the *tert*-butyl moiety on an amino acid building block, the *tert*-butylglycines in the ribosomal peptide polytheonamide B and bottromycin A_2 (**14**, Figure 2)[120,121] appear to be introduced by post-translational C-methylation of valine side chains at the β carbon atom, through a radical SAM-mediated transfer of $CH_3^{\bullet}$ equivalents (Scheme 8B). An analogous C-methylation of an Ile side chain also occurs to give the corresponding branched side chain. Likewise, the β-methyl-Phe residue in bottromycin is proposed to arise by such a radical SAM-dependent mechanism. Even more remarkable is a related net sequential addition of four C_1 units to build a $CH_2C(CH_3)_3$ moiety, presumably by the same radical SAM mechanism (this provides *tert*-butyl-



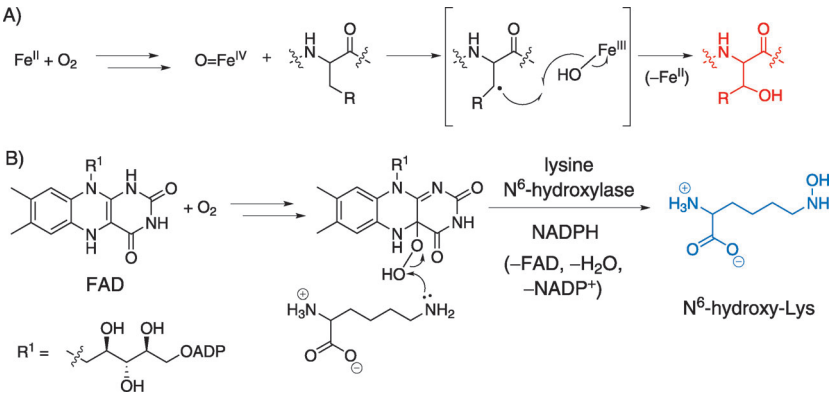
Scheme 8. C-Methylation of amino acids at the β position. A) SAM-mediated C-methylation. B) Radical SAM-mediated C-methylation. See Figure S1 in the Supporting Information.

glycine, which is found in **14**, Figure 2).[61] Some C-methylations of the indole ring of tryptophan involve unreactive carbon sites; those SAM-dependent methyl transfers may also proceed via $CH_3^{\bullet}$ radical equivalents, related to polytheonamide B biosynthesis.

5.2.2. O- and N-Methylations

For O-methylations, as in the O-Me-Ser in kutznerides (Figure 6) and the O-Me-Tyr in didemnin B (**31**, Figure 16), the side-chain OH groups are the proximal nucleophiles to attack the activated methyl group of the cosubstrate SAM in a more conventional transfer of a "CH_3^+" equivalent.[68,90] Whether the O-methylation occurs at the free amino acid or after incorporation into a peptide framework is not known.

A number of NRPs have one or more peptide bonds that have undergone N-methylation. For example, 7 of the 11 peptide bonds in cyclosporin A (**35**, Figure 19) and the N terminus of vancomycin (**32**, Figure 17) have N-methyl substituents. N-Methylation also arises through the use of SAM as a cosubstrate and involves direct capture of the itinerant methyl group by the basic nitrogen atom of the amine acting as a nucleophile.[2] Most of these reactions are catalyzed by methyltransferase (MT) domains that are inserted into A domains in the relevant chain-elongating NRPS modules.[2] Preliminary studies have shown that some *N*-methylamino acids will function in ribosomal protein biosynthesis in *E. coli*.[122,123]



Scheme 9. Hydroxylation of amino acids. A) Metal-dependent hydroxylation. B) FAD-dependent hydroxylation.

5.3. Side-Chain Hydroxylations

There are 11 hydroxylated noncanonical amino acid residues in the 15 nonribosomal peptide scaffolds noted in the previous section. Two of the hydroxy groups, in the statine residue of miraziridine A (**13**, Figure 1) and the isostatine residue of didemnin B (**31**, Figure 16), do not derive from molecular oxygen (O_2), but instead stem from the carboxy group of Val and Ile, respectively, as will be discussed in Section 6.1 on hybrid NRPS-PKS modules. All the others arise from O_2 by the action of hydroxylases, which come in two varieties, depending on the cofactor involved. They either harbor an inorganic redox-active iron center or they use the organic coenzyme FAD. The former generates a high-valent oxoiron intermediate (Scheme 9A), whereas the latter forms a less-reactive FAD-OOH moiety as the oxygen-transfer agent (Scheme 9B).[124] The choice between these two alternative mechanisms is tuned to the reactivity of the amino acid receiving the electrophilic oxygen atom.[125] The six amino acid units in the particular examples (Tyr, Arg, Pro, Glu, Orn, Homo-Tyr) were all hydroxylated at specific carbon sites, and all require iron as a cofactor, either as heme iron from the cytochrome P450 superfamily monooxygenases or as mononuclear nonheme ferrous iron.[126,127] Both forms of these iron-containing oxygenases generate high-valent oxoiron species which can cleave unactivated C–H bonds homolytically and deliver back an $OH^{\bullet}$ equivalent.

The proteinogenic building block *p*-Tyr is normally generated by the nonheme iron-containing phenylalanine hydroxylase, which oxygenates the phenyl ring of Phe regiospecifically. In the biosynthesis of sanglifehrin (**38**, Figure 21)[9] and also in the peptidyl nucleoside pacidamycin antibiotic family,[128] an orthologous Phe hydroxylase acts instead with specificity at the *meta* position to generate the nonproteinogenic *meta*-tyrosine. Enzymatic hydroxylation at the benzylic (C_3) position of the Tyr residues yields the β-OH-Tyr that is found in positions 2 and 6 of vancomycin and teicoplanin.[39,40] The three oxidative cross-linkages in the vancomcycin scaffold (and four in teicoplanin) occur by sequential action of three hemeprotein oxygenases encoded in the biosynthetic gene cluster and link β-OH-Tyr_2 to 4-OH-$Phegly_4$, 4-OH-$Phegly_4$ to β-OH-Tyr_6, and 4-OH-$Phegly_5$ to dihydroxy$Phegly_7$ (see Scheme S21 in the Supporting Information).[129] Importantly, whereas β-OH-Tyr is synthesized as a free amino acid, oxidative cross-linking occurs on the NRPS assembly line.

The glutamate-3-hydroxylase in kutzneride maturation (**18**, Figure 6) is a mononuclear nonheme iron oxygenase,[68] while the oxygenase acting on both the C_4- and C_5-positions of the Orn side chain in the echinocandin scaffold (**27**, Figure 12) is a heme-containing oxygenase.[130] It is not yet clear how to predict whether a given iron-based hydroxylase will modify the free amino acid (a feature that would be most useful for their incorporation into proteins) or instead work only on that residue displayed on the deshydroxy peptide scaffold.

Flavin-dependent monooxygenases utilize FL-OOH as the transfer agent as an electrophilic ("OH^+") equivalent.[125] For example, the formation of the 2-Cl-*N*-OH-pyrrole moiety in the biosynthesis of hormaomycin (**28**, Figure 13)[51] and hydroxylation of a 3-Cl-β-Tyr monomer to produce 3-Cl-5-OH-β-Tyr in the biosynthesis of C-1027 require the activity of FAD-containing monooxygenases.[131] Both oxidative reactions occur on aminoacyl substrates tethered to T domains through thioester linkages. Analogously, for N-hydroxylations at basic side-chain amino groups (e.g. Orn, Lys in siderophore biosynthesis and at the indole NH of Trp-based scaffolds), the nucleophilicity of the nitrogen atom is sufficiently high that

a less-robust oxygenation reagent can suffice. Many such amino acid N-oxygenases are flavin-containing enzymes and act on free-standing amino acid substrates.[125] Further oxygenation of N-OH to nitroso and even nitro groups occurs in some NRP biosynthetic pathways, as proposed in the formation of the nitrocyclopropylalanine residue in hormaomycin (**28**, Figure 13).[51]

5.4. Chlorination as a Variant of Hydroxylation

Amino acid chlorinases can be viewed as variants of the above oxygenase families.[41] In cases where a formal replacement of C–H by C–Cl occurs at unactivated carbon atoms in amino acid side chains, the responsible catalysts are again mononuclear iron enzymes, which consume one equivalent of O_2 and α-ketoglutarate—in close analogy to their oxygenase counterparts. Such halogenases for Pro, Leu, and Ile act on aminoacyl-S-T substrates.[132]

In contrast, the electron-rich phenol side chain of Tyr and the indole ring of Trp residues again require a weaker chlorinating agent, in analogy to the hydroxylation chemistry cited above that is effected by flavoenzymes. Here too, $FADH_2$-containing halogenases are used, this time to deliver a "Cl^+" equivalent.[133] Examples include the 3-Cl-β-OH-Tyr residues of glycopeptide antibiotics (Figure 17),[134] ClmS in chloramphenicol biosynthesis,[135] the 6,7-dichloropyrroloindole residue in the kutznerides (Figure 6),[136] the biosynthesis of dicyostelium differentiating factor Dif-1, and in the biosynthesis of the 2-Cl-*N*-OH-pyrrole carboxy residue of hormaomycin (Figure 13).[51,137] This is also the likely route to 3,5-dichloro-4-OH-Phegly found in enduracidine (**37**, Figure 20).[48] Intriguingly, the known iron-based halogenases recognize aminoacyl-S-T substrates, while the known flavin-dependent ones act upon free amino acids.

5.5. Diamino Acids

While the six carbon atom long dibasic (α,ω) amino acid lysine is a proteinogenic building block, the dibasic C_3, C_4, and C_5 amino acids are nonproteinogenic building blocks. For example, the C_3 2,3-DAP residue is found in capreomycin (**25**, Figure 11), five 2,4-DAB residues are incorporated into the cationic polymyxins (**15**, Figure 3), and the 2,5-diaminopentanoate (ornithine) occurs in echinocandin (**27**, Figure 12; where it is doubly hydroxylated) and daptomycin (**17**, Figure 5), as well as in many siderophores.[138] In addition to the linear 2,4-DAB, the branched 2,3-DAB is also found, as the 2*S*,3*S* diastereomer, in friulimicin (**34**, Figure 18).[45]

The biosynthesis of ornithine has been well-studied, as it is in fact a component of primary amino acid metabolism. 2,3-DAP arises from pyridoxal phosphate dependent enzymatic dehydration of serine to a dehydroAla-PLP intermediate, which is captured by NH_3 (see Scheme S15 in the Supporting Information).[44] The branched 2,3-DAB would arise correspondingly from Thr dehydration and regiospecific C_3 amination, whereas the 2,4-DAB regioisomer presumably arises from enzymatic transamination of the primary metabolite aspartate semialdehyde (see Scheme S6 in the Supporting Information).[45,46]

5.6. Cyclization of Amino Acid Units

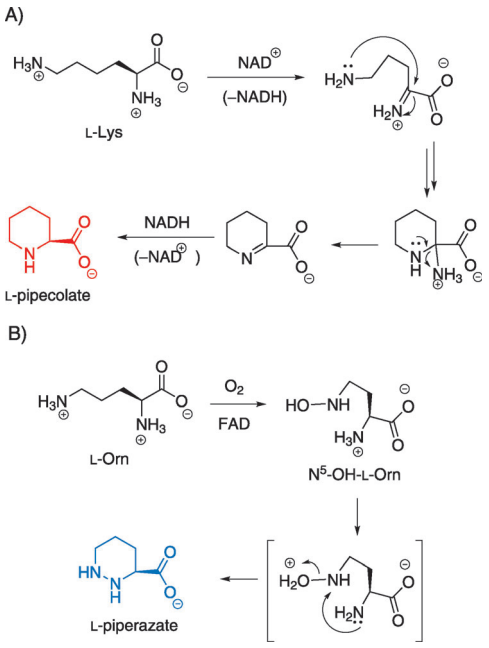
5.6.1. Three-Membered Rings

One of the important functional attributes of nonproteinogenic/noncanonical amino acids is the conformational constraints and rigidifications that are imparted to the structural scaffolds where they are incorporated. One such constraint is ring size. In the monocyclic amino acid category, one can observe three-membered rings, such as 1-amino-1-carboxy-2-ethylcyclopropane in the coronamate moiety of coronatine (**12**, Scheme 5),[105] the aziridine dicarboxylate in miraziridine A (**13**, Figure 1),[55] and in the nitrocyclopropyl ring embedded in the lysine-derived framework of hormaomycin (**28**, Figure 13).[51] The cyclopropane in coronamic acid derives from L-*allo*-Ile (whose direct precursor is unknown, but could be D-Ile). L-*allo*-Ile is activated and tethered on the S-pantetheinyl arm of a thiolation domain of the CmaA enzyme.[34] There it undergoes regiospecific Cγ-chlorination, prior to intramolecular attack of Cα on that chlorine-substituted Cγ to construct the cyclopropane framework of coronamic acid, still tethered through the thioester linkage. Thioesterase action releases the free coronamic acid,[139] which on subsequent ligation to the bicyclic polyketide coronofacic acid, yields the toxin coronatine (**11**, Scheme 5), a mimic of jasmonate plant hormones. The processing of Val in place of Ile through this chlorination/cyclization process yields the methyl-substituted cyclopropane known as norcoronamic acid. The cryptic chlorination chemistry is mediated by an O_2-requiring iron-containing halogenase as discussed above; instead of transferring an $OH^{\bullet}$, CmaB transfers a $Cl^{\bullet}$ from the active site of the high-valent iron intermediate.[34]

Less apparent is the enzymatic pathway from Lys to nitrocyclopropylalanine found in hormaomycin (**28**, Figure 13); the C_6-amino group of Lys has been oxidized to the nitro group and a cyclopropane has formed by a C–C bond-forming step between C_4 and C_6. Two possible variations of the ring opening of an internal lactone following N_6- and C_4-oxygenation have been proposed, but not yet validated.[51]

5.6.2. Orn and Lys Cyclizations

Six-membered-ring cyclic amino acids include the proline analogue pipecolate, found in rapamycin (**3**, Scheme 3) and the 1,2-hydrazo analogue piperazate, seen in sanglifehrin (**38**, Figure 21) and the kutznerides (**18**, Figure 6). Pipecolate is synthesized by lysine cyclodeaminase;[140] this enzyme class contains tightly bound NAD and begins catalysis by oxidation of the α-amine to the imine, thereby yielding NADH as a transient co-product still bound to the enzyme. Intramolecular attack by the C_6-NH_2 generates a cyclic imine after elimination of NH_3. The Δ^1-piperidine carboxylate is finally reduced by hydride transfer from the bound NADH, thereby



Scheme 10. Pipecolate and piperazate biosynthesis. A) Proposed mechanism of pipecolate formation. B) Proposed mechanism of piperazate formation.

regenerating the starting oxidation state of the bound coenzyme (NAD) and releasing pipecolate (Scheme 10A).

The piperazate ring found in kutznerides and sanglifehrin is also a cyclic six-atom framework, but has two nitrogen atoms in a hydrazo (N_1–N_2) linkage. It derives from ornithine, and the first step is N_5-hydroxylation by a dedicated flavoprotein.[141] It has been proposed that further oxygenation occurs on the *N*-hydroxy-Orn to set up an intramolecular attack of the α-NH_2 as the nucleophile on such an electrophilic Nδ atom to form the unusual N–N bond (Scheme 10B).

5.6.3. Cyclization of the Guanidino Moiety of Arg

Five- and six-membered, Arg-derived cyclic guanidines are observed in enduracidine (**37**, Figure 20) and capreomycin (**25**, Figure 11), respectively. The six membered capreomycidine ring is proposed to arise following hydroxylase action to create a 3-OH-Arg.[142,143] Subsequent processing by a pyridoxal phosphate dependent enzyme presumably leads to the elimination of water and the formation of an enamino-PLP adduct. Intramolecular capture by the terminal guanidino nitrogen atom and release from the active site of the PLP enzyme yields the cyclic six-membered ring (Scheme 11A). In viomycin (**26**, Figure 11), the same ring system has undergone an additional hydroxylation to generate a tubericidine residue. The route to the five-membered cyclic guanidine in enduracidine is not as well worked out (the route shown in Scheme 11B is hypothetical), although a similar ring system in guadinomine is formed by intramolecular attack of the terminal nitrogen atom of the guanidine group on a C_3–C_4 unsaturated intermediate.[144] In mannopeptimycins (e.g. **39**, Scheme 11B), a subsequent C_3-hydroxylation yields β-OH enduracididine.[145]

5.6.4. Formation of the Pyrroloindole Tricycle

Tryptophan residues can be converted into tricyclic pyrroloindole cores by the addition of two types of electrophiles across the 2,3-double bond of the pyrrole ring. In the generation of the *Bacillus* competence factor peptides, a geranyl group is added at C_3, while the downstream amide NH group attacks C_2.[146] Alternatively, the delivery of electrophilic oxygen atom as "OH^+" by oxygenases (to generate a transient indole epoxide/hydroxyiminium species) can induce capture by the intramolecular amide to build a pyrroloindole, as in kutzneride biosynthesis (Figure 6).[147] These modifications must occur either on the elongating linear chain or after the macrocyclic peptide scaffold has formed, and are not possible on free Trp. A comparable indole epoxygenation is observed during the fungal biosynthesis of fumiquinazoline A.[148]

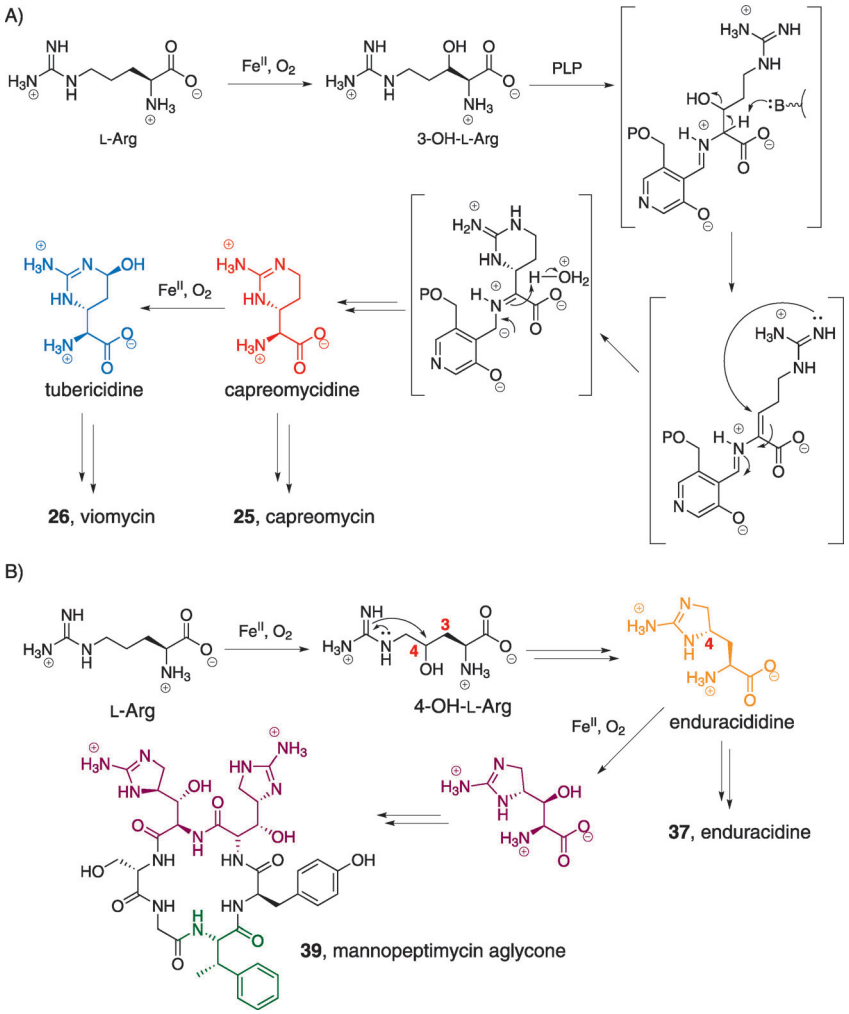
5.7. Kynurenine and Alkyl/Alkenyl Proline Biosynthesis

In daptomycin (**17**, Figure 5), the indole ring of tryptophan is oxidatively cleaved by the action of tryptophan 2,3-dioxygenase to yield *N*-formylkynurenine. Enzymatic deformylation yields free kynurenine, which is selected, activated, and incorporated by the adenylation domain of NRPS module 13 to become the terminal residue in daptomycin.

The analogous oxidative cleavage of the aromatic ring of tyrosine occurs in the formation of the propenylproline monomer of hormaomycin (**28**, Figure 13) as well as in the biosynthesis of anthramycin, suburomycin, and tomayomycin.[149] The pathway starts by conventional *ortho*-hydroxylation of Tyr to L-DOPA and subsequent oxidation of the diphenol to the orthoquinone.[150] Intramolecular attack by the amine on the quinone yields a dihydroxyindole carboxylate that is the substrate for the oxygenative cleavage of the ring by an intradiol-type dioxygenase, thereby yielding a diene-dicarboxylate-substituted proline. Removal of the C_2 enoyl moiety must occur, as well as addition of a C_1 fragment, probably by radical SAM enzymology, to construct the C_3-substituted propenyl group.[51] The propenyl-Pro is presumed to be a free amino acid building block for the hormaomycin NRPS assembly line, with a total of six enzymes required to transform Tyr into propenyl proline.[51]

5.8. Hydroxylated Phenylglycines

The 4-OH-Phegly monomer of vancomycin (**32**, Figure 17), teicoplanin (**33**, Figure 17), and enduracidine (**37**, Figure 20) is synthesized by four enzymes.[42] The first enzyme, a homologue of prephenate decarboxylase, utilizes the primary metabolite prephenate to produce *p*-hydroxyphenyl pyruvate. A second decarboxylase oxygenates this intermediate to form *p*-hydroxy mandelate. Oxidation of the hydroxy group to the ketone yields *p*-hydroxybenzoyl formate. The last enzyme transaminates this keto acid to L-*p*-OH-Phegly. The phenylglycine moiety in pristinamycin (Scheme 4) is also derived from an analogous pathway.[43]



Scheme 11. Biosynthesis of cyclic guanidines. A) Biosynthesis of six-membered cyclic guanidines. B) Biosynthesis of five-membered cyclic guanidines. Nonproteinogenic amino acids: capreomycidine, tubericidine, (2*S*,3*S*)-β-methyl-Phe, β-hydroxyenduracididine, enduracididine.

By contrast, the 3,5-dihydroxyphenylglycine building block in glycopeptides (Figure 17) is assembled by a distinct pathway involving four enzymes. The pathway commences with a type III PKS that uses four molecules of malonyl-CoA and by an iterative decarboxylative Claisen reaction to build the C_8 framework of 3,5-dihydroxyphenylacetyl-CoA.[151,152] The next enzyme performs a remarkable chain-shortening oxidation, dependent on O_2, to generate the 3,5-dihydroxybenzoyl formate. The last enzyme is a transaminase to produce the 2*S* form of 3,5-(OH)$_2$-Phegly.

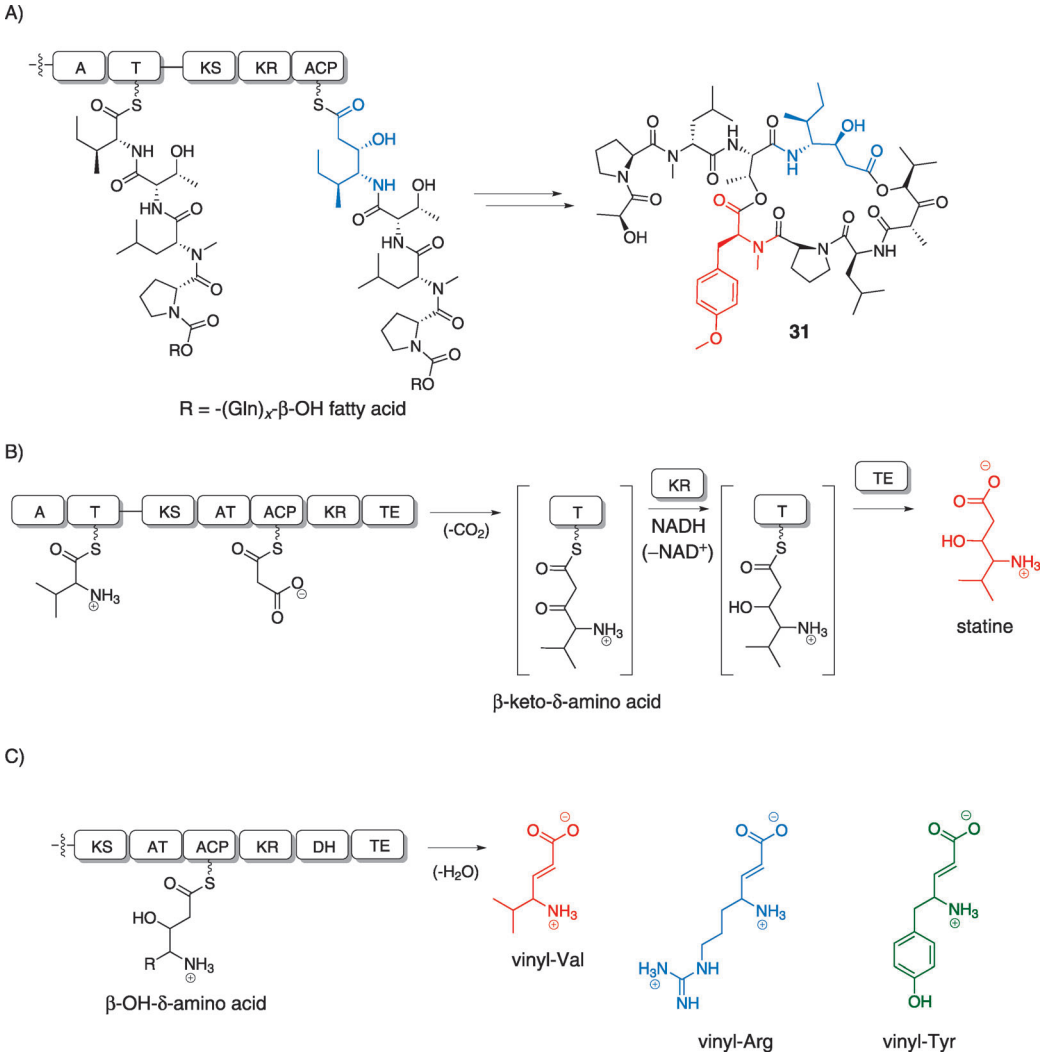
6. Amino Acids Synthesized by Enzymatic Assembly lines

A variety of nonproteinogenic amino acids, or peptides containing them, have been isolated in which the amino group is at C_4 rather than at C_2. In many of these cases it appears that a C_2 extender, derived from a CH_2-COOH equivalent, has been inserted into a proteinogenic amino acid scaffold and now comprises C_2 and C_1, respectively. C_3 in these scaffolds is often substituted with an OH or an OCH_3 group, as seen in the amino acids statine in miraziridine A (**13**, Figure 1) and isostatine in didemnin B (**31**, Figure 16). This is also the case for dolaisoleucine in dolastatin 10.[153] The statine residue with its tetrahedral C_3-OH atom has been a useful inhibitory mimic of tetrahedral intermediates in the action of several aspartyl proteases.[154] Correspondingly, miraziridine has a vinyl-Arg residue at its C terminus (Figure 1) and syringolin A has a vinyl-Val (**20**, Figure 8). These residues are also extended by an acetate-derived C_2 unit, which is sp^2 rather than sp^3 hybridized.

6.1. NRPS-PKS Hybrids

These observations on the connectivity of amino acids could be fully rationalized when biosynthetic gene clusters became available for natural products containing such 3-OH-4-NH_2 frameworks, such as didemnin B (**31**, Figure 16) and andrimid.[107,155,156] The didemnin cluster encodes a PKS module interspersed between NRPS modules. In that case, the growing N-acyl-Pro-D-MeLeu-Thr-Ile-S-tetrapeptide is tethered as a thioester to the S-pantetheinyl arm of the peptidyl carrier protein domain of the fourth NRPS module. Immediately downstream is a PKS module, which tethers a malonyl unit, also with a S-pantetheinyl linkage, on its own acyl carrier protein. Decarboxylation of the malonyl moiety generates the C_2 carbanion for capture of the upstream peptidyl thioester. The C–C bond-forming step moves the peptidyl moiety onto the C_2 unit and sets the peptidyl amino carbon as C_3. The PKS module has a ketoreductase domain so the β-keto group is then reduced to a β-OH group, thereby yielding the isostatyl residue as a ketide-extended N-acylpentapeptidyl thioester. This intermediate is then subjected to further chain elongation by the downstream NRPS modules, with macrocyclizing release, thereby embedding the isostatyl residue within the macrocyclic framework. The mechanism of an isostatyl biosynthesis is shown in Scheme 12 A.

Analogously, if Leu is activated by such an NRPS-PKS hybrid, the resulting chain-extended hybrid building block would be a statyl residue (Scheme 12 B), as in miraziridine. In the andrimid biosynthetic pathway, the comparable NRPS-PKS pair lacks a reductase domain within the PKS module, so the β-ketone functionality is retained.[156] If the chain-elongated intermediate in the statine pathway with its β-OH group were further subjected to typical α,β-dehydration by a dehydratase domain, the resulting product would be vinyl-



Scheme 12. Hybrid NRPS-PKS assembly lines. A) Isostatine assembly into didemnin. B) Mechanism of statine biosynthesis. C) Vinylogous amino acid biosynthesis. Nonproteinogenic amino acids: A) *O*-Me-Tyr, isostatine; B) statine; C): vinyl-Val, vinyl-Arg, vinyl-Tyr.

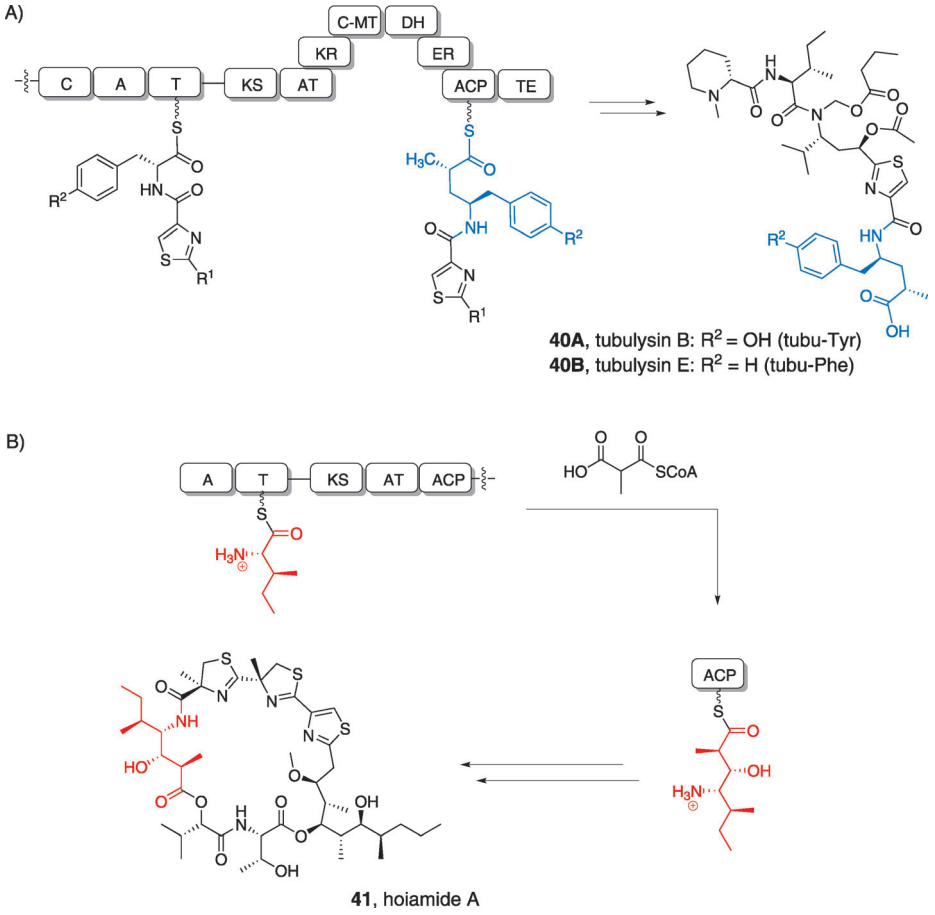
Leu (Scheme 12C). The comparable vinyl-Val residue is known (Figure 8). Alternatively, starting from Arg, comparable logic in an NRPS-PKS bimodule would generate vinyl-Arg for the miraziridine assembly line (Figure 1).[54] Similarly, a vinyl-Tyr can be found in cyclotheonamide.[157] Finally, if the PKS module has an enoyl reductase (ER) domain, the vinyl group would undergo hydrogenation, thus resulting in an amino acid extended by a CH_2COOH group. Such extended Tyr and Phe residues are found in the tubulin inhibitors tubulysins B (**40A**) and E (**40B**, Scheme 13A), where they are known as tubu-tyrosine and tubu-phenylalanine.[158]

Many PKS assembly lines use methylmalonyl-CoA instead of malonyl-CoA as chain-extender units. Indeed the hoiamides (e.g. hoiamide A (**41**), Scheme 13B) harbor a 4-amino-3-hydroxy-2,5-dimethylheptanoate residue that results from an Ile and methylmalonyl-utilizing NRPS/PKS hybrid module.[159] Of course, one can have more than one PKS-mediated chain-extension cycle to build longer alkyl chains. For example, the 3-amino-2,5,7,8-tetrahydroxy-10-methylundecanoate in cyanobacterial pahayokolides may arise from leucine and three malonyl-CoA extender units with additional adjustments of redox states.[160] Here, the β-amino acid would be generated by reductive amination of the β-keto acid produced by the PKS machinery.

It should be noted that the opposite arrangement of modules (i.e. PKS-NRPS) also exists in nature's repertoire of hybrid assembly lines, as in the case of rapamycin (**3**, Scheme 3). However, these configurations typically do not yield an amino acid product; instead, they catalyze the formation of an amide bond.

6.2. Amino Acids Synthesized by PKSs and Reductive Amination

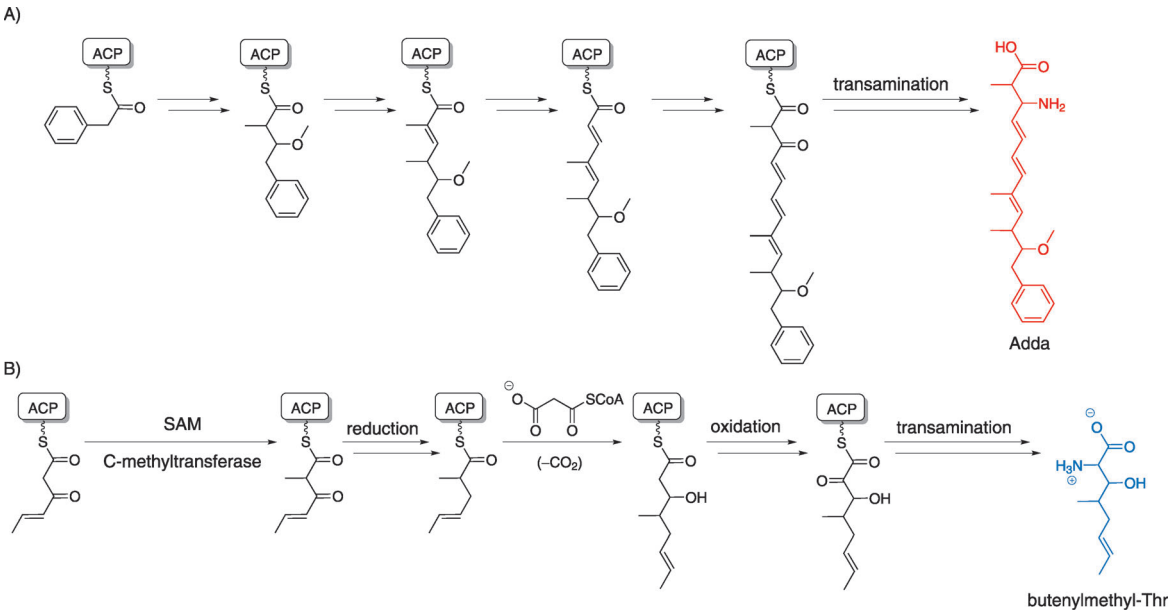
The nonproteinogenic amino acids Aeo in the cyclic tetrapeptide HDAC inhibitors trapoxin and HC-toxin (**23** and **24**, Figure 10),[161] Adda (3-amino-9-methoxy-2,6,8-trimethyl-10-phenyl-4,6-decadienoate) in microcystin LR (**30**, Figure 15),[162] and butenylmethylthreonine in cyclosporin A (**35**, Figure 19)[100] are clearly reminiscent of PKS assembly line origins. For Adda, in particular, a benzoyl-CoA starter unit gets extended by three methylmalonyl-CoA and two malonyl-CoA extender units to yield a β-keto acid that undergoes transamination (Scheme 14A).[79,162] A similar route and PKS complex may be involved in the formation of the Atmu (3-amino-2,5,7,8-tetrahydroxy-10-methylundecanoate) unit in pahayokolides,[160] this time with additional hydroxylation at C_2 and C_6. The long-chain (C_{15-17}) β-amino acid starter units in mycosubtilin (**29**, Figure 14) and iturin are also derived by a similar mechanism. In this case, a β-keto acyl-CoA, from fatty acid biosynthesis, undergoes transthiolation on the NRPS heptapeptide assembly lines and then undergoes enzymatic transamination in situ as the prologue to building the acylated peptide chain.[86b] That N-terminal β-amino fatty acyl group is subsequently the nucleophile in the



Scheme 13. Hoiamide and tubulysin biosynthesis. A) Tubu-tyrosine and tubu-phenylalanine biosynthesis. B) Mechanism of 4-amino-3-hydroxy-2,5-dimethylheptanoate biosynthesis Nonproteinogenic amino acids: 4-amino-3-hydroxy-2,5-dimethylheptanoate, tubu-Tyr or tubu-Phe.

last step of the mycosubtilin assembly to create the heptapeptidyl macrolactam scaffold. Note that this route to β-amino acids is distinct from the action of amino acid mutases on α-amino acids, such as the conversion of α-Phe into β-Phe carried out by MIO-containing mutases (see Section 5.1.2).

The butenylmethylthreonine residue (2*S*,3*R*,4*R*,6*E*-2-amino-3-hydroxy-4-methyl-6-octenoate) in the immunosuppressant undecapeptide cyclosporin A (**35**, Figure 19) is an α- rather than a β-amino acid, but similarly has a PKS origin (Scheme 14B).[163] The C_9 thioester is assumed to undergo oxygenation and oxidation of the CH_2 group at Cα, followed by hydrolysis and transamination to give the α-amino acid. An analogous route is envisioned for the 2-amino-8-oxo-9,10-epoxydecanoate (Aeo) residue in the HC toxin and trapoxin family of cyclic tetrapeptide antagonists of histone deacetylases (Figure 10).[79] The epoxide is thought to arise from a Δ^9 olefin. The saturated 2-amino-8-oxodecanoate residue homologous to Aeo is found in apicidin,[164] and the 9-OH analogue in another member



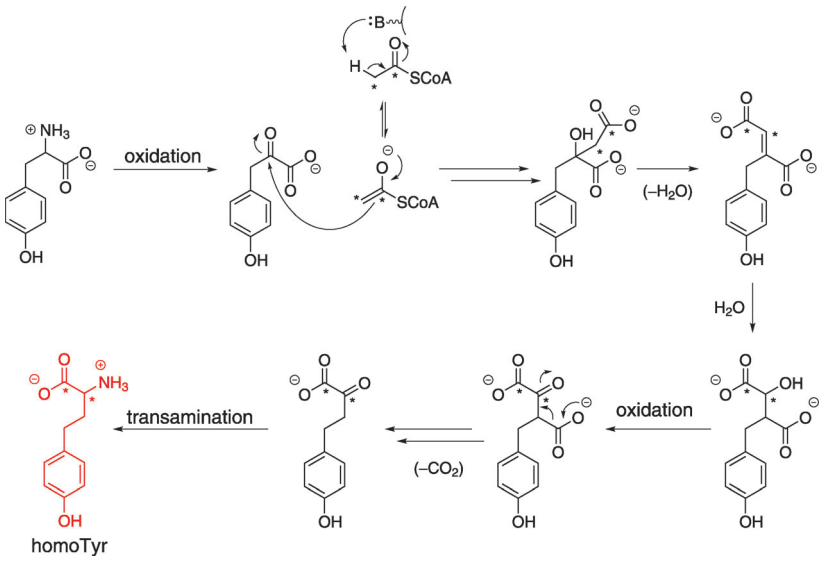
Scheme 14. Biosynthesis of polyketide precursors to amino acids. A) Biosynthesis of 3-amino-9-methoxy-2,6,8-trimethyl-10-phenyl-4,6-decadienoate (Adda). B) Biosynthesis of butenylmethyl-Thr. Nonproteinogenic amino acids: Adda, butenylmethyl-Thr.

of the cyclic tetrapeptide fungal metabolite family, consistent with differential tailoring of the polyketide precursors.[165]

7. HomoTyr and Other Homologated Amino Acids

HomoPhe and homoTyr, where an additional methylene group has been inserted into the chain, arise from the scaffolds of Phe and Tyr, respectively. Homotyr residues are found in some cyanopeptolins and 3,4-dihydroxy-homoTyr is found in echinocandins.[84,166] Early labeling studies had shown that the dihydroxy-homoTyr in echinocandins (e.g. **27**, Figure 12) arises from Tyr and the two carbon atoms of acetate. It is presumed that the carboxylate moiety of the Tyr precursor is lost (two carbon atoms added, one carbon atom lost). A precedent for such a net CH_2 homologation is found in the biosynthesis of leucine.[167] By analogy, for homoTyr, the C_2 carbanion of acetyl-CoA could add into the ketone of *p*-hydroxyphenyl pyruvate (arising from Tyr transamination) to yield a β-OH adduct, which can isomerize to the α-OH,β-COOH adduct through reversible dehydration and rehydration in the opposite regiochemical sense (Scheme 15). Oxidation of the hydroxy group to the ketone and resultant facile decarboxylation of the β-carboxylate yields the homologated α-keto acid. Transamination then gives homoTyr. Dihydroxylation at C_3 and C_4 to complete the framework found in echinocandins is presumed to occur by an iron-catalyzed oxygenase reaction. An analogous pathway from Phe and acetyl-CoA would lead to homoPhe, found in some nonribosomal peptides, including in pahayokolides as the D isomer.[160]

In norvaline biosynthesis, α-ketobutyrate can substitute for *p*-hydroxyphenyl pyruvate in the initial condensation with the acetyl-CoA carbanion. If instead α-keto-β-methyl valerate (derived from the transamination of Ile) is the keto acid condensing partner with acetyl-CoA for isopropylmalate synthase, homoisoleucine will be formed at the final transamination step. The simplest homologated nonproteinogenic amino acid is 2-aminobutyrate, (homologated alanine), which is available from the transamination of the same α-ketobutyrate.



Scheme 15. Biosynthesis of homo-Tyr.

8. Amino Acids with Potentially Reactive Functional Groups

Proteinogenic amino acids do not contain any reactive or potentially reactive electrophilic side-chain functionalities such as olefins, aldehydes, or ketones, thus ruling out unwanted side reactions of these groups once they are installed in proteins. On the other hand, natural nonproteinogenic amino acids contain some potentially reactive functional groups such as olefins (and very rarely alkynes), epoxides (and very rarely aziridines), and, exceedingly rarely, the β-lactam acylation "warheads".[168]

8.1. Olefins and Alkynes

α,β-Olefinic amino acids are enamines and are not stable, because of spontaneous isomerization to the imines, followed by hydrolysis to ammonia and the keto acid. One can observe posttranslational dehydration of Ser and Thr residues in small proteins to yield the α,β-unsaturated amides (dehydroAla and also dehydrobutyrine and also in pahayokolide).[160,169] While these are kinetically stable, they are electrophilic at Cβ through conjugation to the amide and undergo capture by nucleophiles: this is the basis for the formation of lanthionine and methylanthionine residues, notably in the lantipeptides. In capreomycin, the dehydroalanine moiety is further modified with a ureido group (**25**, Figure 11).

In contrast, β,γ-olefinic amino acids are stable.[170] The parent in this category, vinyl-Gly (**42**, Figure 22 A), is a known natural metabolite,[171] although its biosynthesis is not fully clear: a by-product of PLP-containing γ-elimination enzymes would be one route. A related β,γ-olefinic amino acid is dehydro-Val (**43**, Figure 22 A), which occurs in the hexapeptide mycotoxin phomopsin A.[172] The β-alkoxy-substituted vinylglycines, including methoxyvinylglycine (**44**) and rhizobitoxine (**45**, Figure 22 A), are less reactive than vinyl-Gly as free metabolites.[173,174] Once they undergo elimination of the OR group in the active site of PLP enzymes they are subject to capture by other nucleophiles including oxygen and sulfur side chains of amino acids in the enzyme. Notably, these amino acids are synthesized as free metabolites (i.e. not tethered to T domains of NRPS modules).

The simplest γ,δ-olefinic amino acid allylglycine (**46**, Figure 22 B) is also a known metabolite in mushrooms, but its biosynthetic pathway has not been delineated.[175,176] It is a stable molecule, but the olefin can be isomerized to result in conjugation with the carboxyl group and

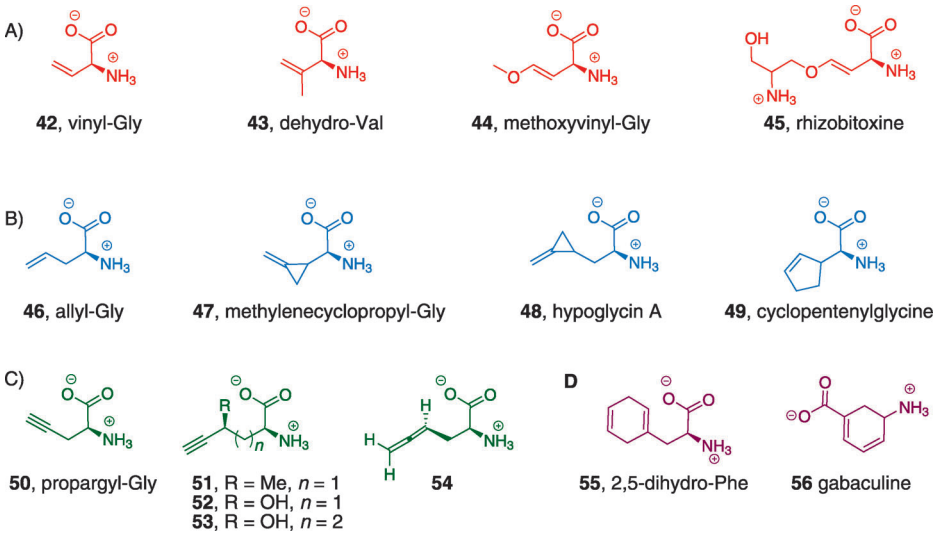


Figure 22. Alkyne- and alkene-containing amino acids. A) β,γ-olefins, B) γ,δ and δ,ε-olefins, C) alkynes and an allene, D) dienes.

become electrophilic by the subset of PLP enzymes that carry out γ-elimination and replacement reactions.[177] Several other olefinic amino acids have also been isolated from a variety of plants, including β,γ-unsaturated methylenecyclopropyl-Gly (**47**) which occurs in the seeds of *Litchi chinensis*;[178] δ,ε-unsaturated amino acid hypoglycin A (**48**), which occurs in the seeds of *Blighia sapida* (ackee) fruit and is known to cause Jamaican vomiting sickness if ingested;[179] and cyclopentenyl-Gly (**49**, Figure 22B), which is known to inhibit the utilization of isoleucine by *E. coli*.[180] The alkyne congener of vinyl-Gly, L-propargylglycine (**50**, Figure 22C), has the same stability and latent reactivity with the same set of PLP enzymes.[181] Several other naturally occurring alkyne-containing amino acids are known, such as 2-amino-4-methyl-5-hexynoic acid (**51**), 2-amino-4-hydroxy-5-hexynoic acid (**52**), and 2-amino-4-hydroxy-6-heptynoic acid (**53**, Figure 22C), all isolated from the seeds of *Euphoria longan*.[182] Additionally, an allene-containing amino acid (2-amino-4,5-hexadienoic acid (**54**), Figure 22C) was isolated from the mushroom *Amanita solitaria*.[183]

2,5-Dihydrophenylalanine (**55**, Figure 22D) is another γ,δ-olefinic amino acid that can act as a bacterial antimetabolite that interferes with Phe metabolism.[184] Another amino acid of particular interest in the natural dienoic dihydro amino acid category is gabaculine (**56**, Figure 22D).[185] Its biosynthesis is not yet resolved, but its mechanism of action with certain PLP-dependent enzymes has been studied.[186] The formation of the gabaculine=PLP aldimine, followed by Cα-H abstraction generates an intermediate that can aromatize, and in so doing capture the PLP coenzyme as a non-hydrolyzable amine adduct.[185]

8.2. Epoxides

Although prevalent in certain terpene and polyketide classes, epoxides are relatively rare in amino acid derived natural products.[168] One amino acid noted above is the 2-amino-8-oxo-9,10-epoxydecanoate, Aeo, which is responsible for the covalent modification of histone deacetylases when embedded in the HC-toxin/trapoxin class of cyclic tetrapeptides (Figure 10).[79] The epoxide almost certainly arises from the T-domain-bound 9,10-olefin precursor by action of an iron-based monooxygenase, as recently proven for dapdiamide C.[187] When bound to active sites of their target enzymes, such epoxides require an acid catalyst to transfer a proton to the oxygen atom and lower the transition-state energy for C–O bond cleavage when a nucleophile attacks the carbon atom.

8.3. Aldehydes and Ketones

Aldehyde and ketone functional groups are absent in proteins, as pointed out by Young and Schultz in their successful efforts to expand the genetic code and incorporate ketone-containing unnatural amino acids into engineered proteins.[33]

However, aldehydic amino acids are central intermediates in the metabolism of the amino acids aspartate and glutamate. Each of the side-chain carboxylates can be phosphorylated enzymatically to yield aspartyl-β-phosphate (**57**) and glutamyl-γ-phosphates (**58**, Figure 23A), respectively. These are then reduced enzymatically by hydride transfer from dihydronicotinamide cofactors. The Asp aldehyde is a key intermediate in the microbial and fungal biosyntheses of lysine, methionine, isoleucine, and threonine.[167] The glutamyl-γ-semialdehyde is a corresponding intermediate which undergoes transamination to ornithine or by intramolecular imine

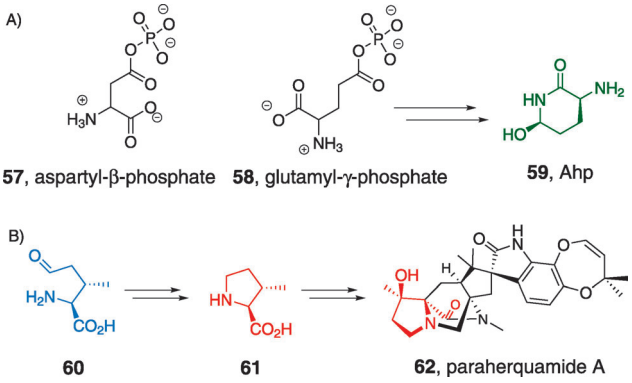


Figure 23. Amino acids accessed through aldehyde intermediates. A) Glutamyl and aspartyl phosphates. B) Paraherquamide A biosynthesis. Nonproteinogenic amino acids: 5-oxo-Ile, β-Me-Pro, Ahp.

formation and NADH-mediated reduction to proline. A cyclized form of the glutamyl semialdehyde is found in some nonribosomal peptides, where it has been termed an Ahp residue (3-amino-6-hydroxypiperidone (**59**), Figure 23 A).[188] The nitrogen atom of the immediately downstream amide in the peptide backbone acts as a nucleophile to form the cyclic hemiaminal linkage, a conformationally restricting cyclization event. Examples of such natural products include stigonemapeptin,[188] symplocamide,[189] and cyanopeptolins.[190]

In contrast to the reduction of acid groups to aldehydes with Asp and Glu, the terminal δ-CH_3 group of Ile is oxidized to the corresponding aldehyde (5-oxo-Ile, **60**), presumably through the alcohol intermediate. Both reactions proceed on the free amino acid on the way to form β-methylPro (**61**), which is incorporated into paraherquamide A (**62**, Figure 23 B).[191] A similar residue is also embedded in the peptide macrocycle of GE37468 to yield the β-methyl-δ-OH-proline residue in that scaffold.[192] Cysteine residues in some proteins, such as sulfatases, can be oxidized posttranslationally to the side-chain aldehyde, thereby yielding formylglycine (**63**, Figure 24) residues, which participate as electrophiles in the hydrolysis of sulfate ester substrates.[118] These unique formylglycine residues can be captured by external nucleophiles to engineer site-specifically modified proteins.

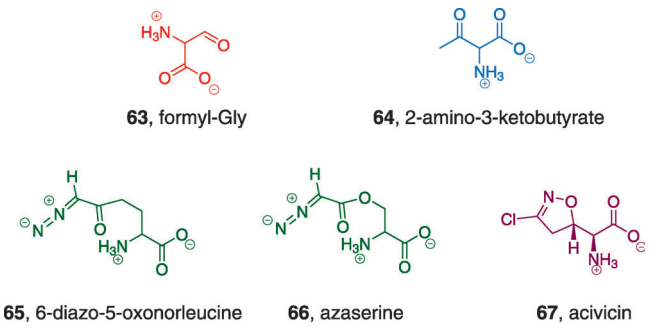


Figure 24. Aldehyde-, imine-, ketone-, and diazo-containing amino acids. Nonproteinogenic amino acids: formyl-Gly, 2-amino-3-ketobutyrate, α-diazoketone or α-diazoester, acivicin.

Ketones are more thermodynamically stable carbonyl groups than aldehydes. A few cases of ketone-containing nonproteinogenic amino acids are known. The most prominent is 2-amino-3-ketobutyrate (**64**, Figure 24), which arises from enzymatic oxidation of the side-chain alcohol of L-threonine.[193] This C_4 α-amino-β-keto acid is subject to retro-aldol cleavage in the presence of CoASH as a cosubstrate to yield acetyl-CoA from $C_{3,4}$ and glycine from $C_{1,2}$. Retro-aldol fragmentation would be a continuing but presumably muted liability if 2-amino-3-keto amino acids were incorporated into protein backbones.

6-Diazo-5-oxonorleucine (DON (**65**), Figure 24) contains a diazomethyl ketone in its compact scaffold.[194] Its biosynthesis has not been reported, even though it was isolated 60 years ago. A related α-diazoester, azaserine (**66**, Figure 24), is also naturally occurring and has antitumor and antibiotic properties.[194,195] The diazoketone group is set up for loss of dinitrogen (N_2) when attacked by nucleophiles, so an engineered tRNA/tRNA synthetase pair for DON might be of use for selective reactivity.

The cyclic amino acid acivicin (**67**, Figure 24) is another example of a heterocyclic electrophile that mimics glutamine residues.[196] Its biosynthesis remains unknown.

8.4. α,β-Unsaturated Amides, Epoxyamides, and Epoxy Ketones

When olefins are positioned α,β to carbonyl groups they become activated as electrophiles for conjugate addition. We noted that the vinyl-Val residue in syringolin (**20**, Figure 8) is the reactive site for covalent modification of the proteasome.[197] Another example is in the dapdiamide family of antibiotics, where the nonproteinogenic amino acid 2,3-DAP has a fumaramoyl group appended to the β-amino functionality.[198] When dapdiamide A (**68**, Figure 25) binds to the active site of the glutaminase domain of a bacterial glucosamine-6-P synthase, its fumaramoyl-DAP moiety mimics a "stretched" glutamine whose electrophilic group can covalently inactivate the active-site Cys.[199] In this case, the conjugated olefinic group is not built de novo into the amino acid but brought in by enzymatic acylation (of a DAP-Val dipeptide intermediate) with fumarate and then converted from the fumaroyl- into the fumaramoyl-DAP (Figure 25).

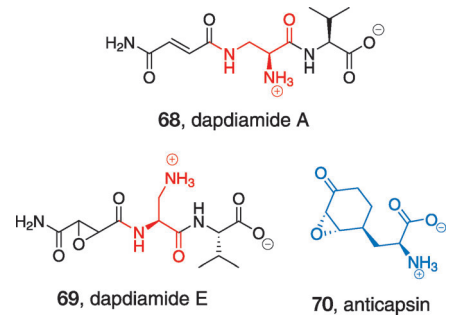


Figure 25. Dapdiamides and anticapsin. Nonproteinogenic amino acids: L-2,3-DAP, anticapsin.

In some members of the dapdiamide antibiotic family noted above, such as dapdiamide E (**69**, Figure 25), the fumaramoyl group is converted by a mononuclear iron enzyme into the epoxysuccinamoyl group, thereby yielding the epoxysuccinamoyl-DAP as the glutamine analogue which again captures the active-site Cys thiolate in the glutaminase domain of glucosamine-6-P synthase.[187]

Another amino acid with an epoxycarbonyl side chain is anticapsin (**70**, Figure 25), the "warhead" in bacilysin.[200] The keto group ultimately derives from the 7-OH group of prephenate, by enzymatic oxidation at a late stage of the diversion of prephenate through to anticapsin.[201] Similar to dapdiamide E, the epoxycyclohexanone side chain can capture the same active-site cysteine of glucosamine-6-P synthase.[202]

9. Challenges and Opportunities for Utilization of Nonproteinogenic Amino Acids as Protein Building Blocks

In this Review, we have introduced the reader to a broad range of nonproteinogenic amino acids. A number of these can be readily biosynthesized as freestanding building blocks, and simply await exploitation by the protein engineer. However, several levels of potential challenges can be envisioned for the ready utilization of some of the other amino acids. Consider, for instance, five examples. The olefinic allylglycine (**46**, Figure 22B) as well as the alkynoic propargylglycine (**50**, Figure 22C) and congeners could prove useful in providing bioorthogonally reactive functional groups in proteins. Likewise, the diazoketone amino acids DON (**65**, Figure 24) and azaserine (**66**, Figure 24) have specific reactivity features, as does aziridine dicarboxylate (Figure 1). A third set of unusual building blocks is represented by the β-oxy-γ-aminodolaisoleucine.[153]

A first level of challenge is whether the biosynthetic genes are known from the producer organism. For none of these five building blocks is that currently the case. This problem may be readily addressable for the DON producer *Streptomyces ambofaciens*, which has been known for decades and whose genome could be readily sequenced, although it is not entirely clear what kind of encoded enzymes to look for bioinformatically.[194] In contrast, the dolastatin 10 producer comes from *Symploca sp. V642* in a complex bacterial community associated with the sea hare *Dolabella auricularia*, so gene identification is substantially more challenging.[203] Finally, for allylglycine (**46**, Figure 22B) and 2-amino-4,5-hexadienoic acid (**54**, Figure 22C), which are mushroom metabolites, those genomes have yet to be sequenced and it is also not clear what the biosynthetic clusters will look like a priori.[175,176,183]

A second level of challenge is to assign function to enzymes encoded by a given gene cluster so that genes necessary and sufficient to make the amino acid building block can be moved to a desired surrogate host for production. Related to this issue is deciphering the chemical logic and the number of protein-mediated steps required to construct the novel amino acid. A corollary is whether the pathway generates the free amino acid or one tethered to a domain on an NRPS or PKS-thiolation domain. In the latter case, as noted earlier, a thioesterase needs to be found and included with the NRPS/PKS genes to release the free amino acid.

The third level of challenge includes engineering an orthogonal tRNA/aminoacyl-tRNA synthetase pair to recognize and activate the specific nonproteinogenic amino acid building block. This has been successfully accomplished for many α-amino acids over the past decade, but β- and γ-amino acids are likely to be substantially more problematic for such evolution. These latter building blocks will also provide problems for ligation by the peptidyl transferase centers of the ribosome. On the other hand, as noted earlier in this Review, it appears that at least some aliphatic *N*-methyl but not *N*-ethyl amino acids can be ligated during ribosomal protein biosynthesis.

Notwithstanding these potential barriers, amino acids are central and versatile building blocks in the chemistry of life, and amino acid metabolism in microbes can go far beyond the familiar 20–22 proteinogenic α-amino acids. Although no single organism shows the full biosynthetic capacity, the inventory of known naturally occurring nonproteinogenic amino acids can be up to 20-fold greater than the proteinogenic set. Although many of the amino acids can only be found in secondary or conditional metabolic pathways, others such as ornithine, homoserine, and aminolevulinate are central players in primary metabolism. The enzymatic pathways from proteinogenic to nonproteinogenic amino acids are often short and efficient, in terms of both the number of steps and atom-economy considerations. The genes encoding for these conversions are almost invariably clustered, and offer straightforward prospects for moving specific modification pathways between microbes.

Freestanding nonproteinogenic amino acids have long been recognized for their antimetabolite properties and tendency to be uncovered to reactive functionalities by the catalytic action of target enzymes. By installing them regiospecifically into biogenic peptides and proteins, it may be possible to usher a new era at the interface between small molecule and large molecule medicinal chemistry. Furthermore, site-selective protein functionalization offers uniquely attractive strategies for posttranslational modification of proteins.[4] Last, but not least, many of the amino acids not selected by nature for protein incorporation offer rich architectural possibilities in the context of ribosomally derived polypeptides.

We are grateful to the National Institutes of Health for financial support (GM20011 and GM49338 for C.T.W. and GM 087934 for C.K.).

Received: October 17, 2012
Published online: May 31, 2013

[1] a) I. Wagner, H. Musso, *Angew. Chem.* **1983**, *95*, 827–839; *Angew. Chem. Int. Ed. Engl.* **1983**, *22*, 816–828; b) for an online database of nonproteinogenic amino acids (norine), visit http://bioinfo.lifl.fr/norine/.

[2] R. Finking, M. A. Marahiel, *Annu. Rev. Microbiol.* **2004**, *58*, 453–488.

[3] C. T. Walsh, *Acc. Chem. Res.* **2008**, *41*, 4–10.

[4] C. C. Liu, P. G. Schultz, *Annu. Rev. Biochem.* **2010**, *79*, 413–444.

[5] J. A. Johnson, Y. Y. Lu, J. A. Van Deventer, D. A. Tirrell, *Curr. Opin. Chem. Biol.* **2010**, *14*, 774–780.

[6] a) C. Hertweck, *Angew. Chem.* **2009**, *121*, 4782–4811; *Angew. Chem. Int. Ed.* **2009**, *48*, 4688–4716; b) E. S. Sattely, M. A. Fischbach, C. T. Walsh, *Nat. Prod. Rep.* **2008**, *25*, 757–793.

[7] T. Schwecke, J. F. Aparicio, I. Molnár, A. König, L. E. Khaw, S. F. Haydock, M. Oliynyk, P. Caffrey, J. Cortés, J. B. Lester, G. A. Böhm, J. Staunton, P. F. Leadlay, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7839–7843.

[8] C. Barreiro, C. Priteo, A. Sola-Landa, E. Solera, M. Martínez-Castro, R. Pérez-Redondo, C. Garcia-Estrada, J. F. Aparicio, L. T. Fernández-Martínez, J. Santos-Aberturas, Z. Salehi-Najafabadi, A. Rodriguez-Garcia, A. Tauch, J. F. Martin, *J. Bacteriol.* **2012**, *194*, 3756–3757.

[9] X. Qu, N. Jiang, F. Xu, L. Shao, G. Tang, B. Wilkinson, W. Liu, *Mol. BioSyst.* **2011**, *7*, 852 – 861.
[10] I. Molnár, T. Schupp, M. Ono, R. Zirkle, M. Milnamow, B. Nowak-Thompson, N. Engel, C. Toupet, A. Stratmann, D. D. Cyr, J. Gorlach, J. M. Mayo, A. Hu, S. Goff, J. Schmid, J. M. Ligon, *Chem. Biol.* **2000**, *7*, 97 – 109.
[11] L. Tang, S. Shah, L. Chung, J. Carney, L. Katz, C. Khosla, B. Julien, *Science* **2000**, *287*, 640 – 642.
[12] J. J. Burbaum, P. Schimmel, *J. Biol. Chem.* **1991**, *266*, 16965 – 16968.
[13] M. Mocibob, N. Ivic, S. Bilokapic, T. Maier, M. Luic, N. Ban, I. Weygand-Durasevic, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14585 – 14590.
[14] M. Ibba, D. Soll, *Annu. Rev. Biochem.* **2000**, *69*, 617 – 650.
[15] T. Stachelhaus, A. Hüser, M. A. Marahiel, *Chem. Biol.* **1996**, *3*, 913 – 921.
[16] 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.
[17] M. A. Fischbach, C. T. Walsh, *Chem. Rev.* **2006**, *106*, 3468 – 3496.
[18] H. D. Mootz, M. A. Marahiel, *J. Bacteriol.* **1997**, *179*, 6843 – 6850.
[19] L. Du, B. Shen, *Curr. Opin. Drug Discovery Dev.* **2001**, *4*, 215 – 228.
[20] H. Chen, S. O'Connor, D. E. Cane, C. T. Walsh, *Chem. Biol.* **2001**, *8*, 899 – 912.
[21] G. J. Gatto, Jr., S. M. McLoughlin, N. L. Kelleher, C. T. Walsh, *Biochemistry* **2005**, *44*, 5993 – 6002.
[22] S. E. O'Connor, H. Chen, C. T. Walsh, *Biochemistry* **2002**, *41*, 5685 – 5694.
[23] L. Du, C. Sánchez, M. Chen, D. J. Edwards, B. Shen, *Chem. Biol.* **2000**, *7*, 623 – 642.
[24] B. A. Pfeifer, C. C. C. Wang, C. T. Walsh, C. Khosla, *Appl. Environ. Microbiol.* **2003**, *69*, 6698 – 6702.
[25] C. Khosla, S. Kapur, D. E. Cane, *Curr. Opin. Chem. Biol.* **2009**, *13*, 135 – 143.
[26] M. Strieker, A. Tanović, M. A. Marahiel, *Curr. Opin. Struct. Biol.* **2010**, *20*, 234 – 240.
[27] Y. Mast, T. Weber, M. Gölz, R. Ort-Winklbauer, A. Gondran, W. Wohlleben, E. Schinko, *Microb. Biotechnol.* **2011**, *4*, 192 – 206.
[28] J. He, N. Magarvey, M. Piraee, L. C. Vining, *Microbiology* **2001**, *147*, 2817 – 2829.
[29] V. Blanc, P. Gil, N. Bamas-Jacques, S. Lorenzon, M. Zagorec, J. Schleuniger, D. Bisch, F. Blanche, L. Debussche, J. Crouzet, D. Thibaut, *Mol. Microbiol.* **1997**, *23*, 191 – 202.
[30] Y. J. Mast, W. Wohlleben, E. Schinko, *J. Biotechnol.* **2011**, *155*, 63 – 67.
[31] J. M. Green, B. P. Nichols, *J. Biol. Chem.* **1991**, *266*, 12971 – 12975.
[32] R. A. Mehl, J. C. Anderson, S. W. Santoro, L. Wang, A. B. Martin, D. S. King, D. M. Horn, P. G. Schultz, *J. Am. Chem. Soc.* **2003**, *125*, 935 – 939.
[33] T. S. Young, P. G. Schultz, *J. Biol. Chem.* **2010**, *285*, 11039 – 11044.
[34] F. H. Vaillancourt, E. Yeh, D. A. Vosburg, S. E. O'Connor, C. T. Walsh, *Nature* **2005**, *436*, 1191 – 1194.
[35] J. P. Shaw, G. A. Petsko, D. Ringe, *Biochemistry* **1997**, *36*, 1329 – 1342.
[36] P. K. Mehta, P. Christen, *Adv. Enzymol. Relat. Areas Mol. Biol.* **2000**, *74*, 129 – 184.
[37] D. Krug, R. Müller, *ChemBioChem* **2009**, *10*, 741 – 750.
[38] S. D. Christenson, W. Wu, M. A. Spies, B. Shen, M. D. Toney, *Biochemistry* **2003**, *42*, 12708 – 12718.
[39] R. D. Süssmuth, W. Wohlleben, *Appl. Microbiol. Biotechnol.* **2004**, *63*, 344 – 350.
[40] O. Puk, D. Bischoff, C. Kittel, S. Pelzer, S. Weist, E. Stegmann, R. D. Süssmuth, W. Wohlleben, *J. Bacteriol.* **2004**, *186*, 6093 – 6100.
[41] F. H. Vaillancourt, E. Yeh, D. A. Vosburg, S. Garneau-Tsodikova, C. T. Walsh, *Chem. Rev.* **2006**, *106*, 3364 – 3378.
[42] B. K. Hubbard, M. G. Thomas, C. T. Walsh, *Chem. Biol.* **2000**, *7*, 931 – 942.
[43] F. Ningsih, S. Kitani, E. Fukushima, T. Nihira, *Microbiology* **2011**, *157*, 3213 – 3220.
[44] For the formation of 2,3-diaminopropionate in zwittermycin biosynthesis, see B. M. Kevany, D. A. Rasko, M. G. Thomas, *Appl. Environ. Microbiol.* **2009**, *75*, 1144 – 1155.
[45] C. Müller, S. Nolden, P. Gebhardt, E. Heinzelmann, C. Lange, O. Puk, K. Welzel, W. Wohlleben, D. Schwartz, *Antimicrob. Agents Chemother.* **2007**, *51*, 1028 – 1037.
[46] C. S. Vandenende, M. Vlasschaert, S. Y. K. Seah, *J. Bacteriol.* **2004**, *186*, 5596 – 5602.
[47] a) T. G. Kinscherf, R. H. Coleman, T. M. Barta, D. K. Willis, *J. Bacteriol.* **1991**, *173*, 4124; b) R. D. Durbin, T. F. Uchytil, J. A. Steele, L. D. Ribeiro, *Phytochemistry* **1978**, *17*, 147 – 148.
[48] X. Yin, T. M. Zabriskie, *Microbiology* **2006**, *152*, 2969 – 2983.
[49] J. J. Barkei, B. M. Kevany, E. A. Felnagle, M. G. Thomas, *ChemBioChem* **2010**, *10*, 366 – 376.
[50] J. N. Steenbergen, J. Alder, G. M. Thorne, F. P. Tally, *J. Antimicrob. Chemother.* **2005**, *55*, 283 – 288.
[51] I. Höfer, M. Crüsemann, M. Radzom, B. Geers, D. Flachschaar, X. Cai, A. Zeeck, J. Piel, *Chem. Biol.* **2011**, *18*, 381 – 391.
[52] J.-H. Ahn, J. D. Walton, *Mol. Plant-Microbe Interact.* **1997**, *10*, 207 – 214.
[53] For the biosynthesis of statine-containing thaliandepsin E and thailandepsin B (also known as burkholdac B), see a) C. Wang, C. J. Flemming, Y.-Q. Cheng, *MedChemComm* **2012**, *3*, 976 – 981; b) C. Wang, L. M. Henkes, L. B. Doughty, M. He, D. Wang, F.-J. Meyer-Almes, Y.-Q. Cheng, *J. Nat. Prod.* **2011**, *74*, 2031 – 2038; c) J. B. Biggins, C. D. Gleber, S. F. Brady, *Org. Lett.* **2011**, *13*, 1536 – 1539.
[54] Miraziridine is the only natural product known to contain a vinylogous arginine moiety. Jamaicamides contain a methyl pyrrolinone moeity derived from vinylogous alanine, see a) D. J. Edwards, B. L. Marquez, L. M. Nogle, K. McPhail, D. E. Goeger, M. A. Roberts, W. H. Gerwick, *Chem. Biol.* **2004**, *11*, 817 – 833. Cyclotheonamides contain vinylogous tyrosine, see b) N. Fusetani, S. Matsunaga, H. Matsumoto, Y. Takebayashi, *J. Am. Chem. Soc.* **1990**, *112*, 7053 – 7054. The vinylogous alanine moiety of Jamaicamides are biosynthesized by a hybrid NRPS-PKS assembly line; vinylogous arginine and vinylogous tyrosine are likely produced by an analogous pathway.
[55] Y. Nakao, M. Fujita, K. Warabi, S. Matsunaga, N. Fusetani, *J. Am. Chem. Soc.* **2000**, *122*, 10462 – 1043.
[56] For studies on the biosynthesis of Aba, see a) O. K. Sebek, H. A. Barker, *J. Bacteriol.* **1968**, *96*, 2094 – 2098; b) Z. Tokarski, H. E. Klei, C. M. Berg, *Biotechnol. Lett.* **1988**, *10*, 7 – 10; c) I. G. Fotheringham, N. Grinter, D. P. Pantaleone, R. F. Senkpeil, P. P. Taylor, *Bioorg. Med. Chem.* **1999**, *7*, 2209 – 2213.
[57] a) Azd was first isolated from *Streptomyces sp.* MD398-A1, see H. Naganawa, N. Usui, T. Takita, M. Hamada, H. Umezawa, *J. Antibiot.* **1975**, *28*, 828 – 829; b) miraziridine is known to inhibit serine proteases, and this activity has been attributed to the aziridine moiety, see N. Schaschke, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 855 – 857.
[58] For other aziridine-containing natural products including azicemicin (an aziridine-containing polyketide isolated from *Kibdelosporangium sp.* MJ126-NF4, which displays antimicrobial activity against Gram-positive bacteria) as well as acinomycin, mitomycin, and ficellomycin, see a) T. Tsuchida, H. Iinuma, N. Kinoshita, T. Ikeda, R. Sawa, Y. Takahashi, H.

Nagawana, T. Sawa, M. Hamada, T. Takeuchi, *J. Antibiot.* **1993**, *46*, 1772 – 1773; and b) Y. Ogasawara, H.-w. Liu, *J. Am. Chem. Soc.* **2009**, *131*, 18066 – 18068 and references therein. See Scheme S24 in the Supporting Information.

[59] Synthetic deriviatives of Azd are utilized as protease inhibitors, see a) A. Ponte-Sucre, R. Vicik, M. Schultheis, T. Schirmeister, H. Moll, *Antimicrob. Agents Chemother.* **2006**, *50*, 2439 – 2447; b) T. Schirmeister, *J. Med. Chem.* **1999**, *42*, 560 – 572; c) T. Schirmeister, M. Peric, *Bioorg. Med. Chem.* **2000**, *8*, 1281 – 1291; d) R. Vicik, V. Hoerr, M. Glaser, M. Schultheis, E. Hansell, J. H. McKerrow, U. Holzgrabe, C. R. Caffrey, A. Ponte-Sucre, H. Moll, A. Stich, T. Schirmeister, *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2753 – 2757.

[60] Y. V. G. Schlippe, M. C. T. Hartman, K. Josephson, J. W. Szostak, *J. Am. Chem. Soc.* **2012**, *134*, 10469 – 10477.

[61] H. Shimamura, H. Gouda, K. Nagai, T. Hirose, M. Ichioka, Y. Furuya, Y. Kobayashi, S. Hirono, T. Sunazuka, S. Omura, *Angew. Chem.* **2009**, *121*, 932 – 935; *Angew. Chem. Int. Ed.* **2009**, *48*, 914 – 917.

[62] a) F. Kopp, M. A. Marahiel, *Nat. Prod. Rep.* **2007**, *24*, 735 – 749; b) J. W. Trauger, R. M. Kohli, H. D. Mootz, M. A. Marahiel, C. T. Walsh, *Nature* **2000**, *407*, 215 – 218; c) R. M. Kohli, C. T. Walsh, M. D. Burkart, *Nature* **2002**, *418*, 658 – 661.

[63] a) D. R. Storm, K. S. Rosenthal, P. E. Swanson, *Annu. Rev. Biochem.* **1977**, *46*, 723 – 763; b) M. Shaheen, J. Li, A. C. Ross, J. C. Vederas, S. E. Jensen, *Chem. Biol.* **2011**, *18*, 1640 – 1648.

[64] R. Krishnamoorthy, L. D. Vazquez-Serrano, J. A. Turk, J. A. Kowalski, A. G. Benson, N. T. Breaux, M. A. Lipton, *J. Am. Chem. Soc.* **2006**, *128*, 15392 – 15393.

[65] For a review on marine cyclic depsipeptides, see a) G. S. B. Andavan, R. Lemmens-Gruber, *Mar. Drugs* **2010**, *8*, 810 – 834; there are other natural products that contain β-methoxy-Tyr, for example, Mirabamide C, see: b) A. Plaza, E. Gustchina, H. L. Baker, M. Kelly, C. A. Bewley, *J. Nat. Prod.* **2007**, *70*, 1753 – 1760; and also neamphamide A, see c) N. Oku, K. R. Gustafson, L. K. Cartner, J. A. Wilson, N. Shigematsu, S. Hess, L. K. Pannell, M. R. Boyd, J. B. McMahon, *J. Nat. Prod.* **2004**, *67*, 1407 – 1411.

[66] V. Miao, M.-F. Coëffet-LeGal, P. Brian, R. Brost, J. Penn, A. Whiting, S. Martin, R. Ford, I. Parr, M. Bouchard, C. J. Silva, S. K. Wrigley, R. H. Baltz, *Microbiology* **2005**, *151*, 1507 – 1523.

[67] a) I. G. O'Brien, G. B. Cox, F. Gibson, *Biochim. Biophys. Acta* **1970**, *201*, 453 – 460; b) K. N. Raymond, E. A. Dertz, S. S. Kim, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3584 – 3588.

[68] D. G. Fujimori, S. Hrvatin, C. S. Neumann, M. Strieker, M. A. Marahiel, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 16498 – 16503.

[69] a) M. S. Yeh, Y. H. Wei, J. S. Chang, *Biotechnol. Prog.* **2005**, *21*, 1329 – 1334; b) F. Peypoux, J. M. Bonmatin, J. Wallach, *Appl. Microbiol. Biotechnol.* **1999**, *51*, 553 – 563.

[70] U. Wäspi, D. Blanc, T. Winkler, P. Rüedi, R. Dudler, *Mol. Plant-Microbe Interact.* **1998**, *11*, 727 – 733.

[71] H. Amrein, S. Makart, J. Granado, R. Shakya, J. Schneider-Pokorny, R. Dudler, *Mol. Plant-Microbe Interact.* **2004**, *17*, 90 – 97.

[72] a) M. Groll, B. Schellenberg, A. S. Bachmann, C. R. Archer, R. Huber, T. K. Powell, S. Lindow, M. Kaiser, R. Dudler, *Nature* **2008**, *452*, 755 – 758; b) B. Schellenberg, L. Bigler, R. Dudler, *Environ. Microbiol.* **2007**, *9*, 1640 – 1650.

[73] M. V. R. Reddy, M. K. Harper, D. J. Faulkner, *Tetrahedron* **1998**, *54*, 10649 – 10656.

[74] P. A. Smith, T. C. Roberts, F. E. Romesberg, *Chem. Biol.* **2010**, *17*, 1223 – 1231.

[75] T. C. Roberts, P. A. Smith, F. E. Romesberg, *J. Nat. Prod.* **2011**, *74*, 956 – 961.

[76] For a related oxidative phenolic coupling in teicoplanin and vancomycin biosynthesis, see P. F. Widboom, S. D. Bruner, *ChemBioChem* **2009**, *10*, 1757 – 1764.

[77] M. Ito, M. Yamanaka, N. Kutsumura, S. Nishiyama, *Tetrahedron Lett.* **2003**, *44*, 7949 – 7952.

[78] M. Kijima, M. Yoshida, K. Sugita, S. Horinouchi, T. Beppu, *J. Biol. Chem.* **1993**, *268*, 22429 – 22435.

[79] J. Walton, *Phytochemistry* **2006**, *67*, 1406 – 1413.

[80] M. Yoshida, S. Horinouchi, T. Beppu, *Bioessays* **1995**, *17*, 423 – 430.

[81] T. Yamada, Y. Mizugichi, K. H. Nierhaus, H. G. Wittmann, *Nature* **1978**, *275*, 460 – 461.

[82] M. Strieker, E. M. Nolan, C. T. Walsh, M. A. Marahiel, *J. Am. Chem. Soc.* **2009**, *131*, 13523 – 13530.

[83] C. Keller-Juslén, M. Kuhn, H. R. Loosli, T. J. Petcher, H. P. Weber, A. von Wartburg, *Tetrahedron Lett.* **1976**, *17*, 4147 – 4150.

[84] For the biosynthesis of 3-hydroxy-4-methylproline, see a) A. A. Adefarati, R. A. Giacobbe, O. D. Hensens, J. S. Tkacz, *J. Am. Chem. Soc.* **1991**, *113*, 3542 – 3545; b) A. A. Adefarati, O. D. Hensens, E. T. T. Jones, J. S. Tkacz, *J. Antibiot.* **1992**, *45*, 1953 – 1957; c) M. Schönewolf, J. Rohr, *Angew. Chem.* **1991**, *103*, 211 – 213; *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 183 – 185.

[85] E. Rössner, A. Zeeck, W. A. König, *Angew. Chem.* **1990**, *102*, 84 – 85; *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 64 – 65.

[86] a) 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; b) D. B. Hansen, S. B. Bumpus, Z. D. Aron, N. L. Kelleher, C. T. Walsh, *J. Am. Chem. Soc.* **2007**, *129*, 6366 – 6367.

[87] R. A. Dixon, I. Chopra, *J. Antimicrob. Chemother.* **1986**, *18*, 557 – 563.

[88] J. C. Kehr, D. Gatte Picchi, E. Dittmann, *Beilstein J. Org. Chem.* **2011**, *7*, 1622 – 1635.

[89] H. Fujiki, M. Suganuma, *Anti-Cancer Agents Med. Chem.* **2011**, *11*, 4 – 18.

[90] K. L. Rinehart, Jr., J. B. Gloer, J. C. Cook, Jr., S. A. Mizsak, T. A. Scahill, *J. Am. Chem. Soc.* **1981**, *103*, 1857 – 1859.

[91] a) D. W. Montgomery, C. F. Zukoski, *Transplantation* **1985**, *40*, 49 – 56; b) B. V. SirDeshpande, P. L. Toogood, *Biochemistry* **1995**, *34*, 9177 – 9184.

[92] D. Bischoff, B. Bister, M. Bertazzo, V. Pfeifer, E. Stegmann, G. J. Nicholson, S. Keller, S. Pelzer, W. Wohlleben, R. D. Süssmuth, *ChemBioChem* **2005**, *6*, 267 – 272.

[93] D. B. Li, K. Woithe, N. Geib, K. Abou-Hadeed, K. Zerbe, J. A. Robinson, *Methods Enzymol.* **2009**, *458*, 487 – 509.

[94] B. Hadatsch, D. Butz, T. Schmiederer, J. Steudle, W. Wohlleben, R. Süssmuth, E. Stegmann, *Chem. Biol.* **2007**, *14*, 1078 – 1089.

[95] E. Stegmann, S. Pelzer, D. Bischoff, O. Puk, S. Stockert, D. Butz, K. Zerbe, J. Robinson, R. D. Süssmuth, W. Wohelleben, *J. Biotechnol.* **2006**, *124*, 640 – 653.

[96] B. I. Eisenstein, F. B. Oleson, Jr., R. H. Baltz, *Clin. Infect. Dis.* **2010**, *50*, S10 – 15.

[97] L. Robbel, M. A. Marahiel, *J. Biol. Chem.* **2010**, *285*, 27501 – 27508.

[98] T. Schneider, K. Gries, M. Josten, I. Wiedemann, S. Pelzer, H. Labischinski, H. G. Sahl, *Antimicrob. Agents Chemother.* **2009**, *53*, 1610 – 1618.

[99] a) G. Weber, E. Leitner, *Curr. Genet.* **1994**, *26*, 461 – 467; b) A. Laupacis, P. A. Keown, R. A. Ulan, N. McKenzie, C. R. Stiller, *Can. Med. Assoc. J.* **1982**, *126*, 1041 – 1046.

[100] M. Offenzeller, G. Santer, K. Totschnig, Z. Su, H. Moser, R. Traber, E. Schneider-Scherzer, *Biochemistry* **1996**, *35*, 8401 – 8412.

[101] D. F. Farver, D. D. Hedge, S. C. Lee, *Ann. Pharmacother.* **2005**, *39*, 863 – 868.

[102] a) P. Cudic, J. K. Kranz, D. C. Behenna, R. G. Kruger, H. Tadesse, A. J. Wand, Y. I. Veklich, J. W. Weisel, D. G. McCafferty, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7384 – 7379; b) X. Fang, K. Tiyanont, Y. Zhang, J. Wanner, D. Boger, S. Walker, *Mol. BioSyst.* **2006**, *2*, 69 – 76.
[103] C. Khosla, Y. Tang, A. Y. Chen, N. A. Schnarr, D. E. Cane, *Annu. Rev. Biochem.* **2007**, *76*, 195 – 221.
[104] L.-H. Zhang, J. O. Liu, *J. Immunol.* **2001**, *166*, 5611 – 5618.
[105] C. L. Bender, F. Alarcón-Chaidez, D. C. Gross, *Microbiol. Mol. Biol. Rev.* **1999**, *63*, 266 – 292.
[106] E. R. Strieter, A. Koglin, Z. D. Aron, C. T. Walsh, *J. Am. Chem. Soc.* **2009**, *131*, 2113 – 2115.
[107] A. Fredenhagen, S. Y. Tamura, P. T. M. Kenny, H. Komura, Y. Naya, K. Nakanishi, K. Nishiyama, M. Sugiura, H. Kita, *J. Am. Chem. Soc.* **1987**, *109*, 4409 – 4411.
[108] Y. Chen, M. Yin, G. P. Horsman, B. Shen, *J. Nat. Prod.* **2011**, *74*, 420 – 424.
[109] M. E. Tanner, *Acc. Chem. Res.* **2002**, *35*, 237 – 246.
[110] K. Kino, M. Sato, M. Yoneyama, K. Kirimura, *Appl. Microbiol. Biotechnol.* **2007**, *73*, 1299 – 1305.
[111] Y. H. Lim, K. Yokoigawa, N. Esaki, K. Soda, *J. Bacteriol.* **1993**, *175*, 4213 – 4217.
[112] T. Stachelhaus, C. T. Walsh, *Biochemistry* **2000**, *39*, 5775 – 5787.
[113] S. L. Clugston, S. A. Sieber, M. A. Marahiel, C. T. Walsh, *Biochemistry* **2003**, *42*, 12095 – 12104.
[114] C. J. Balibar, F. H. Vaillancourt, C. T. Walsh, *Chem. Biol.* **2005**, *12*, 1189 – 1200.
[115] D. B. Stein, U. Linne, M. A. Marahiel, *FEBS J.* **2005**, *272*, 4506 – 4520.
[116] B. Wu, W. Szymański, G. G. Wybenga, M. M. Heberling, S. Bartsch, S. de Wildeman, G. J. Poelarends, B. L. Feringa, B. W. Dijkstra, D. B. Janssen, *Angew. Chem.* **2012**, *124*, 497 – 501; *Angew. Chem. Int. Ed.* **2012**, *51*, 482 – 486.
[117] B. Wu, W. Szymański, M. M. Heberling, B. L. Feringa, D. B. Janssen, *Trends Biotechnol.* **2011**, *29*, 352 – 362.
[118] *Posttranslational modification of proteins* (Ed.: C. T. Walsh), Roberts, Colorado, **2006**.
[119] C. Mahlert, F. Kopp, J. Thirlway, J. Micklefield, M. A. Marahiel, *J. Am. Chem. Soc.* **2007**, *129*, 12011 – 12018.
[120] T. Hamada, S. Matsunaga, G. Yano, N. Fusetani, *J. Am. Chem. Soc.* **2005**, *127*, 110 – 118.
[121] a) W. J. K. Crone, F. J. Leeper, A. W. Truman, *Chem. Sci.* **2012**, *3*, 3516 – 3521; b) J. P. Gomez-Escribano, L. Song, M. J. Bibb, G. L. Challis, *Chem. Sci.* **2012**, *3*, 3522 – 3525.
[122] T. Kawakami, H. Murakami, H. Suga, *Chem. Biol.* **2008**, *15*, 32 – 42.
[123] A. O. Subtelny, M. C. T. Hartman, J. W. Szostak, *J. Am. Chem. Soc.* **2008**, *130*, 6131 – 6136.
[124] a) B. Entsch, D. P. Ballou, V. Massey, *J. Biol. Chem.* **1976**, *251*, 2550 – 2563; b) B. Entsch, D. P. Ballou, M. Husain, V. Massey, *J. Biol. Chem.* **1976**, *251*, 7367 – 7369.
[125] T. Wencewicz, C. T. Walsh, *Nat. Prod. Rep.* **2013**, *30*, 175 – 200.
[126] *Cytochrome P450*: *Structure*, *Mechanism*, *and Biochemistry*, *3rd ed.* (Ed.: P. R. Ortiz de Montellano), Kluwer/Plenum, New York, **2005**.
[127] M. M. Abu-Omar, A. Loaiza, N. Hontzeas, *Chem. Rev.* **2005**, *105*, 2227 – 2252.
[128] W. Zhang, B. D. Ames, C. T. Walsh, *Biochemistry* **2011**, *50*, 5401 – 5403.
[129] M. J. Cryle, A. Meinhart, I. Schlichting, *J. Biol. Chem.* **2010**, *285*, 24562 – 24574.
[130] Y. Tang, C. Walsh, unpublished results.
[131] S. G. Van Lanen, S. Lin, G. P. Horsman, B. Shen, *FEMS Microbiol. Lett.* **2009**, *300*, 237 – 241.
[132] C. S. Neumann, D. G. Fujimori, C. T. Walsh, *Chem. Biol.* **2008**, *15*, 99 – 109.
[133] E. Yeh, S. Garneau, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 3960 – 3965.
[134] O. Puk, P. Huber, D. Bischoff, J. Recktenwald, G. Jung, R. D. Sussmuth, K. H. van Pée, W. Wohlleben, S. Pelzer, *Chem. Biol.* **2002**, *9*, 225 – 235.
[135] K. Podzelinska, R. Latimer, A. Bhattacharya, L. C. Vining, D. L. Zechel, Z. Jia, *J. Mol. Biol.* **2010**, *397*, 316 – 331.
[136] J. R. Heemstra, Jr., C. T. Walsh, *J. Am. Chem. Soc.* **2008**, *130*, 14024 – 14025.
[137] C. S. Neumann, C. T. Walsh, R. R. Kay, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 5798 – 5803.
[138] J. Cheung, F. C. Beasley, S. Liu, G. A. Lajoie, D. E. Heinrichs, *Mol. Microbiol.* **2009**, *74*, 594 – 608.
[139] W. L. Kelly, M. T. Boyne II, E. Yeh, D. A. Vosburg, D. P. Galonić, N. L. Kelleher, C. T. Walsh, *Biochemistry* **2007**, *46*, 359 – 368.
[140] G. J. Gatto, Jr., M. T. Boyne II, N. L. Kelleher, C. T. Walsh, *J. Am. Chem. Soc.* **2006**, *128*, 3838 – 3847.
[141] C. S. Neumann, W. Jiang, J. R. Heemstra, Jr., E. A. Gontang, R. Kolter, C. T. Walsh, *ChemBioChem* **2012**, *13*, 972 – 976.
[142] M. G. Thomas, Y. A. Chan, S. G. Ozanick, *Antimicrob. Agents Chemother.* **2003**, *47*, 2823 – 2830.
[143] V. Helmetag, S. A. Samel, M. G. Thomas, M. A. Marahiel, L. O. Essen, *FEBS J.* **2009**, *276*, 3669 – 3682.
[144] a) M. Iwatsuki, R. Uchida, H. Yoshijima, H. Ui, K. Shiomi, A. Matsumoto, Y. Takahashi, A. Abe, H. Tomoda, S. Omura, *J. Antibiot.* **2008**, *61*, 222 – 229; b) T. C. Holmes, A. E. May, K. Zaleta-Rivera, J. G. Ruby, P. Skewes-Cox, M. A. Fischbach, J. L. DeRisi, M. Iwatsuki, S. Omura, C. Khosla, *J. Am. Chem. Soc.* **2012**, *134*, 17797 – 17806.
[145] N. A. Magarvey, B. Haltli, M. He, M. Greenstein, J. A. Hucul, *Antimicrob. Agents Chemother.* **2006**, *50*, 2167 – 2177.
[146] M. Okada, I. Sato, S. J. Cho, H. Iwata, T. Nishio, D. Dubnau, Y. Sakagami, *Nat. Chem. Biol.* **2005**, *1*, 23 – 24.
[147] A. Broberg, A. Menkis, R. Vasiliauskas, *J. Nat. Prod.* **2006**, *69*, 97 – 102.
[148] B. D. Ames, X. Liu, C. T. Walsh, *Biochemistry* **2010**, *49*, 8564 – 8576.
[149] C. T. Walsh, S. W. Haynes, B. D. Ames, *Nat. Prod. Rep.* **2012**, *29*, 37 – 59.
[150] K. L. Colabroy, W. T. Hackett, A. J. Markham, J. Rosenberg, D. E. Cohen, A. Jacobson, *Arch. Biochem. Biophys.* **2008**, *479*, 131 – 138.
[151] H. Chen, C. C. Tseng, B. K. Hubbard, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 14901 – 14906.
[152] C. C. Tseng, F. H. Vaillancourt, S. D. Bruner, C. T. Walsh, *Chem. Biol.* **2004**, *11*, 1195 – 1203.
[153] R. K. Pettit, G. R. Pettit, K. C. Hazen, *Antimicrob. Agents Chemother.* **1998**, *42*, 2961 – 2965.
[154] D. H. Rich, *J. Med. Chem.* **1985**, *28*, 263 – 273.
[155] Y. Xu, R. D. Kersten, S.-J. Nam, L. Lu, A. M. Al-Suwailem, H. Zheng, W. Fenical, P. C. Dorrestein, B. S. Moore, P.-Y. Qian, *J. Am. Chem. Soc.* **2012**, *134*, 8625 – 8632.
[156] M. Jin, M. A. Fischbach, J. Clardy, *J. Am. Chem. Soc.* **2006**, *128*, 10660 – 10661.
[157] M. Hagihara, N. J. Anthony, T. J. Stout, J. Clardy, S. L. Schreiber, *J. Am. Chem. Soc.* **1992**, *114*, 6568 – 6570.
[158] a) G. Kaur, M. Hollingshead, S. Holbeck, V. Schaeuer-Vukašinović, R. F. Camalier, A. Dömling, S. Agarwal, *Biochem. J.* **2006**, *396*, 235 – 242; b) H. Steinmetz, N. Glaser, E. Herdtweck, F. Sasse, H. Reichenbach, G. Höfle, *Angew. Chem.* **2004**, *116*, 4996 – 5000; *Angew. Chem. Int. Ed.* **2004**, *43*, 4888 – 4892.
[159] H. Choi, A. R. Pereira, Z. Cao, C. F. Shuman, N. Engene, T. Byrum, T. Matainaho, T. F. Murray, A. Mangoni, W. H. Gerwick, *J. Nat. Prod.* **2010**, *73*, 1411 – 1421.
[160] L. Liu, D. W. Bearden, K. S. Rein, *J. Nat. Prod.* **2011**, *74*, 1535 – 1538.

[161] J. S. Scott-Craig, D. G. Panaccione, J. A. Pocard, J. D. Walton, *J. Biol. Chem.* **1992**, *267*, 26044 – 26049.
[162] G. Christiansen, J. Fastner, M. Erhard, T. Börner, E. Dittman, *J. Bacteriol.* **2003**, *185*, 564 – 572.
[163] M. Offenzeller, Z. Su, G. Santer, H. Moser, R. Traber, K. Memmert, E. Schneider-Scherzer, *J. Biol. Chem.* **1993**, *268*, 26127 – 26134.
[164] S. J. Darkin-Rattray, A. M. Gurnett, R. W. Myers, P. M. Dulski, T. M. Crumley, J. J. Allocco, C. Cannova, P. T. Meinke, S. L. Colletti, M. A. Bednarek, S. B. Singh, M. A. Goetz, A. W. Dombrowski, J. D. Polishook, D. M. Schmatz, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 13143 – 13147.
[165] Z. Jiang, M. O. Barret, K. G. Boyd, D. R. Adams, A. S. F. Boyd, J. G. Burgess, *Phytochemistry* **2002**, *60*, 33 – 38.
[166] H. S. Okumura, B. Philmus, C. Portmann, T. K. Hemscheidt, *J. Nat. Prod.* **2009**, *72*, 172 – 176.
[167] J. McMurry, T. P. Begley, *Organic Chemistry of Biological Pathways*, Roberts, Colorado, **2005**.
[168] C. J. Thibodeaux, W. C. Chang, H. W. Liu, *Chem. Rev.* **2012**, *112*, 1681 – 1709.
[169] J. M. Willey, W. A. van der Donk, *Annu. Rev. Microbiol.* **2007**, *61*, 477 – 501.
[170] M. Berkheij, *Synthesis of 1,3-Diene-containing α-amino acids and peptides via N-acyliminium ions*, *Doctoral dissertation*, University of Amsterdam, Amsterdam, **2008**.
[171] G. Dardenne, J. Casimir, M. Marlier, P. O. Larsen, *Phytochemistry* **1974**, *13*, 1897 – 1900.
[172] M. F. Mackay, A. V. Donkelaar, C. C. J. Culvenor, *J. Chem. Soc. Chem. Commun.* **1986**, 1219 – 1221.
[173] U. Sahm, G. Knobloch, F. Wagner, *J. Antibiot.* **1973**, *26*, 389 – 390.
[174] T. Yasuta, S. Okazaki, H. Mitsui, K. Yuhashi, H. Ezura, K. Minamisawa, *Appl. Environ. Microbiol.* **2001**, *67*, 4999 – 5009.
[175] S.-I. Hatanaka, Y. Niimura, K. Takishima, *Trans. Mycol. Soc. Jpn.* **1985**, *26*, 61 – 68.
[176] Y. Yamaura, M. Fukuhara, E. Takabatake, N. Ito, T. Hashimoto, *Toxicology* **1986**, *38*, 161 – 173.
[177] R. R. Rando, *Nature* **1974**, *250*, 586 – 587.
[178] D. O. Gray, L. Fowden, *Biochem. J.* **1962**, *82*, 385 – 389.
[179] K. Tanaka, E. M. Miller, K. J. Isselbacher, *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 20 – 24.
[180] a) L. Andersen, B. Nielsen, J. W. Jaroszewski, *Chirality* **2000**, *12*, 665 – 669; b) R. L. Dennis, W. J. Plant, C. G. Skinner, G. L. Sutherland, W. Shive, *J. Am. Chem. Soc.* **1955**, *77*, 2362 – 2364.
[181] J. P. Scannell, D. L. Pruess, T. C. Demny, F. Weiss, T. Williams, A. Stempel, *J. Antibiot.* **1971**, *24*, 239 – 244.
[182] M.-Y. Sung, L. Fowden, D. S. Millington, R. C. Sheppard, *Phytochemistry* **1969**, *8*, 1227 – 1233.
[183] W. S. Chilton, G. Tsou, L. Kirk, R. G. Benedict, *Tetrahedron Lett.* **1968**, *9*, 6283 – 6284.
[184] a) G. Nass, K. Poralla, H. Zahner, *Naturwissenschaften* **1971**, *58*, 603 – 608; b) J. M. Crawford, S. A. Mahlstedt, S. J. Malcolmson, J. Clardy, C. T. Walsh, *Chem. Biol.* **2011**, *18*, 1102 – 1112.
[185] R. R. Rando, N. Relyea, L. Cheng, *J. Biol. Chem.* **1976**, *251*, 3306 – 3312.
[186] R. R. Rando, F. W. Bangerter, *J. Am. Chem. Soc.* **1976**, *98*, 6762 – 6764.
[187] M. A. Hollenhorst, S. B. Bumpus, M. L. Matthews, J. M. Bollinger, Jr., N. L. Kelleher, C. T. Walsh, *J. Am. Chem. Soc.* **2010**, *132*, 15773 – 15781.
[188] H. S. Kang, A. Krunic, J. Orjala, *J. Nat. Prod.* **2012**, *75*, 807 – 811.
[189] R. G. Linington, D. J. Edwards, C. F. Shuman, K. L. McPhail, T. Matainaho, W. H. Gerwick, *J. Nat. Prod.* **2008**, *71*, 22 – 27.
[190] A. Tooming-Klunderud, T. Rohrlack, K. Shalchian-Tabrizi, T. Kristensen, K. S. Jakobsen, *Microbiology* **2007**, *153*, 1382 – 1393.
[191] E. M. Stocking, J. F. Sanz-Cervera, R. M. Williams, *J. Am. Chem. Soc.* **1996**, *118*, 7008 – 7009.
[192] T. S. Young, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 13053 – 13058.
[193] J. P. Marcus, E. E. Dekker, *Biochem. Biophys. Res. Commun.* **1993**, *190*, 1066 – 1072.
[194] C. C. Nawrat, C. J. Moody, *Nat. Prod. Rep.* **2011**, *28*, 1426 – 1444.
[195] a) J. Ehrlich, M. P. Knudsen, L. E. Anderson, H. J. Koepsell, G. L. Coffey, D. L. Kohberger, A. B. Hillegas, J. E. Oyaas, *Nature* **1954**, *173*, 72; b) Q. R. Bartz, T. H. Haskell, C. C. Elder, D. W. Johannessen, R. P. Frohardt, A. Ryder, S. A. Fusari, *Nature* **1954**, *173*, 72 – 73.
[196] L. Allen, R. Meck, A. Yunis, *Res. Commun. Chem. Pathol. Pharmacol.* **1980**, *27*, 175 – 182.
[197] D. Krahn, C. Ottmann, M. Kaiser, *Nat. Prod. Rep.* **2011**, *28*, 1854 – 1867.
[198] J. Dawlaty, X. Zhang, M. A. Fischbach, J. Clardy, *J. Nat. Prod.* **2010**, *73*, 441 – 446.
[199] M. A. Hollenhorst, I. Ntai, B. Badet, N. L. Kelleher, C. T. Walsh, *Biochemistry* **2011**, *50*, 3859 – 3861.
[200] S. A. Mahlstedt, C. T. Walsh, *Biochemistry* **2010**, *49*, 912 – 923.
[201] a) J. B. Parker, C. T. Walsh, *Biochemistry* **2012**, *51*, 3241 – 3251; b) J. B. Parker, C. T. Walsh, *Biochemistry* **2012**, *51*, 5622 – 5632.
[202] H. Chmara, H. Zähner, E. Borowski, *J. Antibiot.* **1984**, *37*, 1038 – 1043.
[203] H. Luesch, R. E. Moore, V. J. Paul, S. L. Mooberry, T. H. Corbett, *J. Nat. Prod.* **2001**, *64*, 907 – 910.