Cite this: *Nat. Prod. Rep.*, 2011, **28**, 99

www.rsc.org/npr

REVIEW

Fungal cyclooligomer depsipeptides: From classical biochemistry to combinatorial biosynthesis

Roderich Süssmuth,[a] Jane Müller,[a] Hans von Döhren[b] and István Molnár*[cd]

Received 14th July 2010
DOI: 10.1039/c001463j

Covering: up to the end of August 2010

This review surveys the biological activities and the iterative and recursive biosynthetic mechanisms of fungal cyclooligomer depsipeptides, and their structural diversification by various combinatorial biosynthetic methods.

[a]*Technische Universität Berlin, Institut für Chemie, Strasse des 17. Juni 124, 10623 Berlin, Germany. E-mail: suessmuth@chem.tu-berlin.de; Fax: (+49) 030-314-79651*
[b]*Technische Universität Berlin, AG Biochemie und Molekulare Biologie, Franklinstrasse 29, 10587 Berlin, Germany*
[c]*SW Center for Natural Products Research and Commercialization, Office of Arid Lands Studies, School of Natural Resources and the Environment, The University of Arizona, 250 E. Valencia Rd., Tucson, AZ, 85706, USA. E-mail: imolnar@cals.arizona.edu*
[d]*Bio5 Institute, The University of Arizona, 1657 E. Helen Str., Tucson, AZ, 85721, USA*

1 Introduction

Nonribosomal peptides represent an extensive family of small-molecule natural products, including antibiotics, anticancer agents, immunosuppressants, enzyme inhibitors, siderophores, herbicides, antifungals, insecticides, and anthelminthics. Nonribosomal peptides are biosynthesized on giant multi-domain enzymes called nonribosomal peptide synthetases (NRPSs).[1–4] The core domains of NRPSs include those that are responsible

for the recognition and activation of the precursors by adenylation using ATP as the co-substrate (A = adenylation domains); those that capture the precursor adenylates as covalent thioesters *via* a phosphopantetheine linker (T = thiolation domains); and those that catalyze the condensation of the growing peptide chain, covalently bound to the previous T domain, with the precursor thioester by amide or ester bond formation (C = condensation domains). NRPSs may also feature active site domains for the modification of the precursors and/or the intermediates by *N*-methylation (M domains), oxidation/reduction (Ox and KR domains), epimerization (E domains) and heterocycle formation (Cy = cyclization domains); and for the release and the cyclization (if necessary) of the final products from the enzymes (TE = thioesterase, or R = reductase domain, or a terminal C domain). These domains are organized into modules in the NRPS mutlitenzymes. In linear (Type A) NRPSs, each module, and each active site domain is used only once in an assembly line fashion (processive NRPSs). In rare cases violations to this rule may be observed, for example as in module skipping.[5,6] In contrast, iterative (Type B) NRPSs use some of

Roderich Süssmuth

Roderich Süssmuth is the Rudolf Wiechert Professor of Biological Chemistry at TU Berlin. He obtained his Ph.D. degree in chemistry in 1999 from the Eberhard Karls Universität Tübingen with Günther Jung. In 2000, he was a Feodor-Lynen-Fellow of the Alexander von Humboldt-Foundation with Carlos Barbas III and Richard Lerner. Subsequently he became an Assistant Professor in Tübingen with an Emmy-Noether fellowship granted by the DFG, and then moved in 2004 to TU Berlin. His research is focused on antibiotics, biosynthetic assembly lines, enzymes and enzyme inhibitors.

Hans von Döhren

Hans von Döhren received his doctoral degree at TU Berlin with Horst Kleinkauf in 1977, characterizing the structure of gramicidin S synthetase. As a postdoc he initiated work on the applications of peptide synthetases and multienzyme structure–function analysis. From 1983 to the present, he has been a lecturer in biochemistry. His research has focused on peptide biosynthesis, with the current main topics being in the areas of mechanisms of amino acid activation and peptide bond formation, β-lactam antibiotics and fungal peptide diversity, and cyanobacterial toxins.

Jane Müller

Jane Müller received her Master's degree (Diploma) in chemistry from the Technical University in Berlin. During that time she worked in the group of Dr. Rainer Zocher on the enzymatic synthesis of new cyclodepsipeptides as potential drug candidates. She is currently a Ph.D. student in the laboratory of Roderich Süssmuth at TU Berlin and works on the heterologous expression of cyclodepsipeptide synthetases.

István Molnár

István Molnár studied molecular biology at the ELTE University (Budapest, Hungary) before joining the Institute for Drug Research in Budapest to investigate antibiotic biosynthesis in actinomycetes. He went on to conduct graduate studies in microbial sterol transformations (Ph.D. 1993 with Prof. Yoshikatsu Murooka at Hiroshima University, Japan), and postdoctoral research on the biosynthesis of rapamycin (1994–1997 with Prof. Peter Leadlay at Cambridge University, UK). As a staff scientist, and later group leader and principal scientist for Syngenta Biotechnology (formerly Novartis Agribusiness), North Carolina, he worked on the biosynthesis of epothilone and soraphen, and the biocatalytic production of emamectin. Since 2004, he has been an Associate Professor at the Natural Products Center of the University of Arizona, studying fungal polyketide and nonribosomal peptide biosynthesis. His research interests include the engineering of natural product biosynthetic pathways for drug discovery and development, and biofuels research.

their modules more than once in a programmed iterative fashion. Finally, nonlinear (Type C) NRPSs feature unusual arrangements of their core domains and often incorporate precursors that are not tethered on carrier domains.[7]

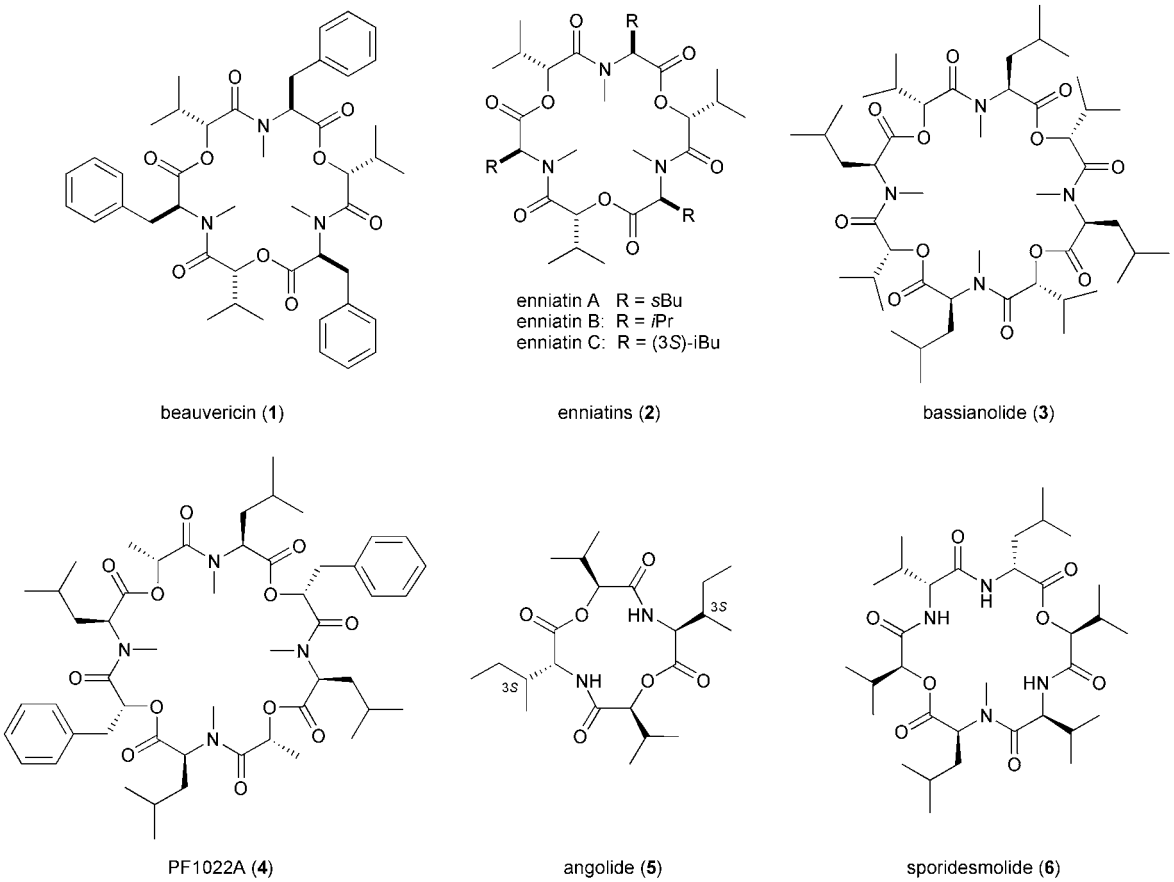
The tremendous structural variety of nonribosomal peptides is based on the flexibility of the biosynthetic programming of the NRPS: the utilization of non-proteinogenic amino acid precursors (more than 300 described); the formation of main-chain heterocycles (thiazole, oxazole and their derivatives); and the construction of linear, macrocyclic or branched macrocyclic structures with amide, ester or even thioester or imino ring closures.[4,8] In the scaffold of the nonribosomal depsipeptides, at least one bond of the peptide backbone is replaced by an ester bond: these connect carboxy groups of amino acids with a 2-hydroxycarboxylic acid, or provide alternative routing of the chain *via* side chain hydroxy groups of amino acids and the C-terminus of the peptide. The structural complexity of nonribosomal (depsi)peptides is further enhanced by the installation of N-terminal aryl or alkyl caps, lipid or glycosyl side chains, and the formation of intramolecular bridges (disulfide bridges, oxidative coupling between side chains), as catalyzed by "decorating" enzymes.[9]

Cyclooligomer nonribosomal peptides consist of oligopeptide or, in the case of cyclooligomer depsipeptides (CODs), oligopeptidol monomer units that undergo recursive head-to-tail condensation, or oligomerization *via* side chains, followed by macrocyclization.[8] The corresponding NRPSs form a subclass of Type B NRPSs. These NRPSs use their modules iteratively for the biosynthesis of several copies of identical or nearly identical peptide/peptidol monomer units that remain covalently bound on the enzyme. These enzymes also evolved mechanisms for the recursive, stepwise intermolecular ligation and final intramolecular cyclization of the monomer units in a concerted cyclooligomerization process.

Fungal CODs, the subject of the current review, are privileged pharmacophores that display a wide variety of bioactivities, including antibiotic, insecticidal, anthelminthic, herbicidal, antiretroviral, cytotoxic, anti-haptotactic, and chemosensitizer activities, as well as inhibition of cholesterol biosynthesis, and repression of amyloid plaque formation in Alzheimer's disease. Fungal COD biosynthesis has been characterized first by isolating and reconstituting active cyclooligomer depsipeptide synthetase (CODS) enzymes from the producer fungi, and later by isolating, characterizing and heterologously expressing the encoding synthetase genes. New fungal CODS genes were discovered by genome mining, and interesting mechanistic differences were noted for CODS of fungal *versus* bacterial origin. Novel analogs of fungal CODs have been generated by a variety of combinatorial biosynthetic methods, including precursor-directed biosynthesis, mutasynthesis, combinatorial mutasynthesis, and total biosynthesis.

2 Cyclooligomer depsipeptides in fungi: Structure, distribution, biological activities

CODs display a rotational symmetry as a result of their oligomeric structure: CODs are biosynthesized as dimers, trimers or tetramers of monomer units, each of which are formed from at



beauvericin (**1**) enniatins (**2**) bassianolide (**3**)

PF1022A (**4**) angolide (**5**) sporidesmolide (**6**)

least one 2-hydroxycarboxylic acid and one 2-amino acid. Beauvericin (**1**), enniatins (**2**), bassianolide (**3**) and PF1022A (**4**) and its congeners contain residues with alternating D and L configuration, similar to 6–8-member cyclic D,L-α-peptide antibacterial agents.[10,11] Variations in the amino acid or the hydroxycarboxylic acid positions of the monomers lead to the production of various COD congeners in a given COD producer fungus. COD production in fungi has hitherto been only documented in the Hypocreomycetidae and the Xylariomycetidae, two subclasses of the Sordariomycetes.

Cyclodepsipeptides showing a pseudo-cyclodimeric structure have also been isolated from several Dothiodeomycete *Pithomyces* spp. (teleomorph: *Leptosphaerulina* spp.). The cyclotetradepsipeptide angolide[12] (**5**) and the cyclohexadepsipeptide sporidesmolides[13] (**6**) are apparent dimers of a dipeptidol (angolide) or a tripeptidol (sporidesmolides). However, the exclusive utilization of D amino acids in one half of these molecules, and L amino acids in the other half, argues against their origin from cyclodimerization (see Section 4.9).

2.1 Cyclohexadepsipeptides: Beauvericin

Beauvericin (**1**) is a cyclic trimer of a dipeptidol monomer. This dipeptidol is formed from (2*R*)-2-hydroxy-3-methylbutanoic acid (D-hydroxyisovaleric acid, D-Hiv) and *N*-methyl-L-phenylalanine (*N*-Me-Phe). Beauvericin is produced by many *Fusarium* species in the *Gibberella fujikuroi* complex, and it is also the main fermentation product of the hypocrealean entomopathogens *Beauveria bassiana*, *Paecilomyces fumoso-roseus* (renamed *Isaria fumosorosea*), and *P. tenuipes* (renamed *I. tenuipes*).[14–17] Beauvericin production was suggested as a diagnostic feature for a clade in the genus *Isaria*, represented by strains of *I. cicadae*, *I. fumosorosea*, *I. japonica*, and *I. tenuipes*, as well as the *Cordyceps* (teleomorph) state of *I. cicadae*.[18] Interestingly, co-production of beauvericin with enniatins H, I and MK 1688 was observed in *Fusarium oxysporum* FB1501.[19] Beauvericin and the enniatins (**2**, see below) transport mono- and bivalent cations across membranes as free carriers and uncouple oxidative phosphorylation, with a 2:1 depsipeptide:cation sandwich as the mobile species.[20] The degree to which this common ion transporter activity is responsible for the disparate biological activities of beauvericin and the enniatins is currently unknown.

Beauvericin is toxic to brine shrimp and to the larvae of insects,[16,21] and acts as an important virulence factor during insect pathogenesis by *B. bassiana*.[22] It also displays moderate antifungal activity, and antibiotic activity against Gram-positive bacteria.[21] Beauvericin is a low-micromolar inhibitor of acyl CoA:cholesterol acyltransferase (ACAT, EC 2.3.1.26): inhibition of this enzyme leads to decreased plasma cholesterol levels.[23,24] ACAT inhibition also suppresses proteolytic processing of the β-amyloid precursor protein, thereby reducing amyloid plaque density in animal models of Alzheimer's disease.[25,26] A recent publication showed that beauvericin, as well as enniatin I and enniatin MK 1688, exhibit strong *in vitro* inhibitory activity against the type-1 human immunodeficiency virus (HIV-1) integrase, but not against the Moloney murine leukemia virus reverse transcriptase.[27]

Beauvericin also reverses the multidrug-resistance (MDR) phenotype in yeast and potentiates the fungicidal activity of fluconazole against fluconazole-resistant *Candida albicans* at concentrations that are not directly fungicidal.[28–30] Synergistic activities amongst different antibiotics are well known (*i.e.* streptogramin A and B components), but compounds that show only weak antibiotic or antifungal activity themselves might also increase the potency of *bona fide* antibiotics or antifungals. These potentiators or sensitizers might prevent the inactivation of the antibiotic or antifungal agents, and facilitate the penetration or inhibit the active efflux of these agents through the cell envelopes of the pathogens.[31]

Beauvericin was also shown to act as a potentiator of cytotoxic drugs in multidrug-resistant (MDR) cancer cell lines. Overexpression of P-glycoprotein (Pgp), an ABC superfamily transporter, is a prominent cause of multidrug resistance in human cancers. Beauvericin was shown to directly bind to purified Pgp with an apparent K_D of 0.36 μM, and to inhibit the drug transport function of Pgp in membrane vesicle preparations at 1 μM. Beauvericin restored daunorubicin accumulation in the Pgp-overexpressing MDR Chinese hamster ovarian cell line CHRC5 at sub-cytotoxic concentrations.[32]

Beauvericin displays potent cytotoxic activity against different human cell lines.[33] Beauvericin increases cytoplasmic Ca^{2+} concentration, causes ATP depletion, and activates calcium-sensitive cell apoptotic pathways.[34,35] At sub-cytotoxic concentrations, beauvericin inhibits the haptotactic motility of cancer cells.[36] Formation of new blood vessels in tumors (angiogenesis), tissue invasion by cancer cells, and metastasis all involve haptotaxis (directional cell motility).[37] In contrast, haptotaxis is rather infrequent in adults under ordinary physiological conditions.[38] Inhibition of angiogenesis is a validated cancer chemotherapeutic strategy as shown by thalidomide and bevacizumab, and is one of the established mechanisms of action of the marketed drugs sunitinib, sorafenib, and paclitaxel.[39,40] Inhibition of tissue invasion and metastasis might restrain new tumor formation, or increase successful containment of solid tumors. The cytotoxic and the anti-haptotactic activities of the taxanes[41] and the *Vinca* alkaloids[42] have distinct mechanisms of actions.

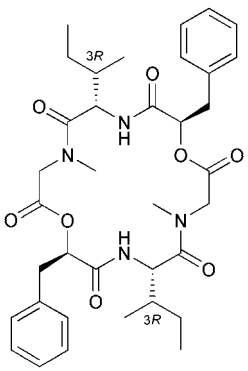
2.2 Cyclohexadepsipeptides: Enniatins

Enniatins (**2**) are prevalent fungal mycotoxins from *Fusarium* spp.[43,44] that are also produced by *Verticillium hemipterigenum*[45,46] and *Halosarpheia* sp.[47] Enniatins feature similar trimeric structures to that of beauvericin, with D-Hiv as the hydroxycarboxylic acid constituent of the dipeptidol monomer. However, the aromatic *N*-Me-Phe moieties featured in the beauvericin dipeptidol monomers are replaced by aliphatic *N*-Me-Ile (enniatin A), *N*-Me-Val (enniatin B) or *N*-Me-Leu (enniatin C). Enniatins act as ionophores by forming freely diffusible sandwiches in biological membranes,[48] and display similar activities to that of beauvericin, including anthelminthic, phytotoxic, antibiotic, antifungal, ACAT inhibitory, cytostatic and cytotoxic activities.[24,33,43,49] Several studies have compared the biological activities of beauvericin and the enniatins, *e.g.* with regard to cytotoxicity and accumulation in cells, the latter of which was found higher for beauvericin than for the enniatins.[50] No resistance was found in two-year continuous exposure studies in human nasopharyngeal carcinoma KB-31 cell lines against beauvericin, and only a low level of resistance developed against

enniatins. Multidrug-resistance exporters were found to be induced in the selected enniatin-resistant cell lines, and these cells also showed cross-resistance to other chemotherapeutics. In contrast, enniatins were found to restore the antifungal activity of cycloheximide or cerulenin against multidrug-resistant *Saccharomyces cerevisiae* at non-toxic concentrations, due to their specific inhibition of the ABC-transporter Pdr5p.[51] Fusafungin, a mixture of various enniatins (enniatin A, B and C), with antimicrobial activity against several Gram-positive and Gram-negative bacteria and *Candida albicans*, is currently used as a topical agent for the treatment of respiratory infections.[52] Enniatins have also been shown to act as virulence factors for the plant pathogen *Fusarium avenaceum*.[53]

2.3 Cyclohexadepsipeptides: Hirsutellide A

Hirsutella kobayashii, a hypocrealean entomopathogenic fungus,[17] produces the antimycobacterial and antimalarial metabolite hirsutellide A (**7**).[54] As opposed to beauvericin (**1**) and the enniatins (**2**), this cyclohexadepsipeptide is a dimer of the tripeptidol monomer D-2-hydroxy-3-phenylpropanoic acid–L-*allo*-isoleucine–*N*-methylglycine, and not a trimer of a dipeptidol monomer. The biosynthesis of this cyclohexadepsipeptide has not yet been characterized.



hirsutellide A (**7**)

2.4 Cyclooctadepsipeptides: Bassianolide

Bassianolide (**3**) is a cyclic octadepsipeptide isolated from the hypocrealean entomopathogens *Beauveria bassiana* and *Lecanicillium* sp. (formerly *Verticillium lecanii*),[17] and from the wood-decaying *Xylaria* sp. BCC1067.[55] Its 24-membered macrolactone ring is formed as a tetramer of the dipeptidol monomer D-Hiv–*N*-Me-Leu. This is the same monomer unit that yields enniatin C upon trimerization in *Fusarium* sp. Bassianolide was shown to be toxic to insect larvae,[56] and represents an important virulence factor of *B. bassiana* during insect pathogenesis.[57] Bassianolide inhibits acetylcholine-induced smooth muscle contraction in a manner that does not involve ionophoric interactions.[58] Bassianolide was moderately or weakly toxic to different cancer cell lines *in vitro*, but showed no anti-haptotactic activity.[55,57]

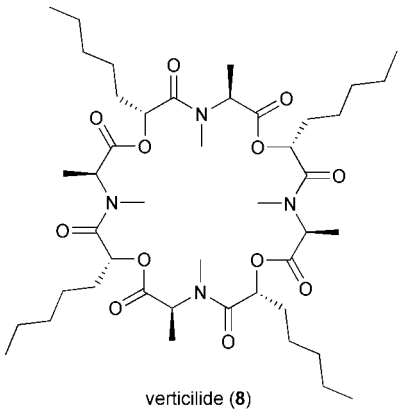
2.5 Cyclooctadepsipeptides: The PF1022 congeners

PF1022A (**4**) and its congeners are produced by *Mycelia sterilia* (an unidentified fungal strain growing only as mycelium with no observed production of conidia),[59] later identified as *Rosellinia* sp. (Xylariaceae).[60] The PF1022s are built from the hydroxycarboxylic acids D-Lac or D-PheLac, and the amino acid *N*-Me-Leu. The main metabolite PF1022A has a C_2-symmetry and is formally composed of two tetradepsipeptides. However, other PF1022 congeners have been described with various combinations of D-Lac and D-PheLac, thus indicating a relaxed incorporation of hydroxycarboxylic acids into dipeptidols that undergo cyclotetramerization during PF1022 biosynthesis.[61] The PF1022s are highly potent anthelminthics with no significant insecticidal activity, and low toxicity to mammals. They cause flaccid paralysis in nematodes in a novel mechanism that involves binding to latrophilin-like receptors at nanomolar IC_{50} concentrations, and the inhibition of calcium-activated potassium channel-dependent (SLO-1) signaling pathways.[62] Some semisynthetic cyclooctadepsipeptides of the PF1022 series, especially those such as emodepside with *para* substitutions in their phenyllactic acid moiety, are even more active than natural PF1022s.[63] Considerable work has been carried out on defining the structure–activity relationships of PF1022 derivatives. In early approaches, a systematic exchange of the amino acid *N*-Me-L-Leu by a series of related *N*-alkyl amino acids was performed, shifting the β-carbon or the γ-carbon substituents to higher hydrophobicities. The resulting activity data strongly suggest that *N*-Me-L-Leu is crucial for high *in vivo* activity.[64] Remarkably, the exchange of the carbonyl oxygen with sulfur resulted in increased activity,[65] and more recent findings point towards biosynthetic options for future introduction of this modification.[66] In addition, the depsipeptide backbone conformation influences anthelminthic properties, as demonstrated by a systematic exchange of the dipeptidols by β-turn mimetics, *e.g.* D-Pro–L-Pro. In this case, increased activity was only observed with derivatives in which the D-PheLac–L-Leu dipeptidols were replaced in a manner that retained the conformation of the PF 1022A parent molecule, whereas other variants led to decreased anthelminthic activity.[67] PF1022A and its semisynthetic derivatives are effective against benzimidazole-, levamisole- and ivermectin-resistant gastrointestinal parasitic nematodes.[68] Since such resistant parasites are rapidly spreading in sheep, cattle, horses and pigs, drugs such as the PF1022 derivatives which display a new mode of action, are urgently needed.[63] Semisynthetic PF1022 derivatives are also considered as next-generation treatments for human parasitic nematode infections such as river blindness, where resistance to the current standard-of-care 'mectin' class of drugs is increasing.[69] Unfortunately, the chemical synthesis of CODs suffers from several shortcomings: the hydroxy groups from the 2-hydroxycarboxylic acids are less nucleophilic towards facile ester formation by methods commonly applied in peptide chemistry; racemization of the amino acids and hydroxycarboxylic acids is frequent; and the coupling efficiency of *N*-methyl amino acids is reduced. As a consequence, synthetic approaches do not at present appear to be an economical means for the commercial production of PF1022 derivatives for use as drugs.[70]

2.6 Cyclooctadepsipeptides: Verticilide

Verticilide (**8**) was isolated from the hypocrealean fungus *Verticillium* sp. FKI-1033, on the basis of its inhibition of the binding

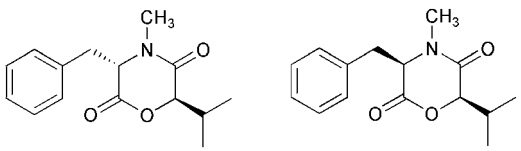
of ryanodine, a plant alkaloid, to the insect ryanodine receptor.[71] This receptor is a Ca^{2+} release channel in the sarcoplasmic reticulum, and has been identified as a potential target for new insecticides. Verticilide is a cyclotetramer formed from the dipeptidol D-2-hydroxyheptanoic acid–*N*-methyl-L-alanine. The biosynthesis of verticilide has not yet been elucidated.



verticilide (**8**)

2.7 Diketomorpholines: Bassiatin and lateritin

Intramolecular cyclization of a single D-Hiv–*N*-Me-L-Phe monomer unit (the same monomer that yields beauvericin upon cyclotrimerization) yields the morpholine-2,5-dione bassiatin (**9**), isolated from *Beauveria bassiana*. Bassiatin inhibits the ADP-induced aggregation of rabbit platelets,[72] but was found to show no significant activity in a cytotoxicity assay with CCRF-CREM human leukemia cells.[34] The diastereoisomer of bassiatin, the acyl-CoA:cholesterol acyltransferase inhibitor lateritin (**10**), was isolated from *Isaria japonica*, *Gibberella lateritium*, and an endophytic *Fusarium* sp.[73–75] Lateritin was also produced in a mixed fermentation involving five different filamentous fungi, including *Fusarium oxysporum* which might be the *de facto* producer of this compound.[76] Lateritin showed antibacterial, antifungal and cancer cell cytotoxic activities.[76] The structure and co-occurrence of these compounds in fungi that also produce CODs raises the possibility that these diketomorpholines are not *bona fide* independent biosynthetic metabolites, but simply derailed shunt products of COD biosynthetic pathways.



bassiatin (**9**) lateritin (**10**)

2.8 CODs in bacteria

Cyclooligomer peptides (for example, gramicidin S) and depsipeptides (for example, enterobactin) are also biosynthesized by bacteria from peptide or peptidol monomer units.[8,77] The ester bonds in the macrocycles of the bacterial CODs are formed either by ligation of peptidol monomers *via* the 2-hydroxy groups of hydroxycarboxylic acids (*e.g.*, valinomycin, cereulide, and serratomolide) as in fungal CODs, or by cyclooligomerization of peptide monomers *via* the side chain alcohols of amino acids (*e.g.*, enterobactin, triostin, echinomycin, and quinomycin).

3 Survey of COD biosynthetic systems in fungi

3.1 Enzyme isolation and stability

Isolation of fungal multienzymes was pioneered in the Kleinkauf laboratory, starting with the enniatin, beauvericin and cyclosporin synthetases.[78–80] The strategy consisted of (1) obtaining a suitable producer strain, possibly improved by random mutagenesis; (2) determining the production phase within the fermentation process employed; (3) disrupting the fresh or frozen mycelial suspension at 4 °C by a French press, or the frozen mycelia by grinding in liquid nitrogen; (4) nucleic acid precipitation by poylethylene imine solution; (5) fractionated salting out of proteins using ammonium sulfate; and (6) further standard protein purification steps. These included gel filtration (as the large multienzymes may be well separated), ion exchange chromatography on DEAE-cellulose, hydrophobic interaction chromatography (most commonly propyl- or butyl-agarose), and sucrose gradient centrifugation. To further concentrate the enzyme, the fractions were passed through size selective filters. The enzyme preparations could be stored frozen at −80 °C in the presence of glycerol. In the case of the ESYN enniatin synthetase, the pure enzyme could be stored at the previously described conditions for several years without significant loss of activity. The specific protocols varied with the strains used, the properties of the mycelia, the modes of cell disruption, and the enzyme content. Detailed protocols were established for enniatin synthetases from *Fusarium equiseti* (synonym: *F. scirpi*, previously described as *F. oxysporum*),[79] *F. sambucinum* and *F. lateritium*;[81] for the beauvericin synthetase from *Beauveria bassiana*;[82] and for the PF1022 synthetase from *Mycelia sterilia* (*Rosellinia* sp.).[61] These CODS have been purified to homogeneity as judged by SDS polyacrylamide electrophoresis, and their molecular masses have been estimated by gel filtration, sucrose gradient ultracentrifugation and electrophoretic mobilities. Since reference proteins in the high mass range (300–400 kDa) were either not available at the time, or their masses were poorly determined, the first estimates for their size (250 kDa) turned out to be well below the accurate masses later confirmed by analytical ultracentrifugation and DNA sequencing.

The purity of the enzyme preparations was monitored by *in vitro* formation of CODs upon supply of substrates (amino acid and hydroxycarboxylic acid), Mg^{2+}, ATP and SAM. The common reaction times were up to one hour at 25 °C, usually with complete loss of activity. The half lives of the enzyme preparations were found to be diverse, ranging from 12 h at 0 °C for the PF1022 synthetase[61] to 50 h at 25 °C for the enniatin synthetase.[83] Generally, protease content was a limiting factor, and proteinase inhibitors were found to be essential for stabilization (H. Peeters, *Doctoral Thesis*, TU Berlin, 1988). Reported catalytic activities range from 2 pkat mg^{-1} (beauvericin synthetase) and 12 pkat mg^{-1} (PF1022 synthetase) to 100 pkat mg^{-1} for the enniatin synthetase. This highest activity corresponds to a turnover number of 2 catalytic cycles/min.

3.2 Reconstitution of COD biosynthesis *in vitro* with purified CODS enzymes ("total biosynthesis")

The *in vitro* reconstitution of enniatin synthesis clearly demonstrated that all reactions required for COD assembly are performed by a single multienzyme, which is a multifunctional synthetase.[84] This has been confirmed for all other CODSs studied to date,[61,79] and can also be concluded independently from their identical domain organization. Substrate requirements reveal the overall reaction

$$3\ \mathrm{AA} + 3\ \mathrm{HA} + 3\ \mathrm{SAM} + 6\ \mathrm{ATP} \rightarrow \mathrm{COD} + 6\ \mathrm{AMP} + 6\ \mathrm{PP_i} + 3\ \mathrm{SAH} \quad (1)$$

where AA = L-amino acid, HA = D-hydroxycarboxylic acid, SAM = *S*-adenosyl-methionine, ATP = adenosine triphosphate, usually as a $MgATP^{2-}$ complex, COD = cyclooligomer depsipeptide, AMP = adenosine monophosphate, PP_i = pyrophosphate, usually as a $MgPP_i^{2-}$ complex, SAH = *S*-adenosylhomocysteine.

The stoichiometry of the overall reaction in relation to the requirement of 1 mole ATP per peptide or ester bond formed has not been proven so far, but can be assumed based on analogous studies in penicillin biosynthesis.[85] The binding stoichiometry of the required methylation cofactor could be determined as 1 mole SAM per mole FeESYN.[86] Reconstitution routinely involves optimization of each of the substrate concentrations, with the determination of apparent K_m values. A complete kinetic description of such complex systems has not been achieved yet. The reaction cycle is considered irreversible, as COD hydrolysis is not observed under normal conditions. However, as partial reactions are indeed reversible (see below), byproducts such as PP_i and SAH act as inhibitors of enniatin synthesis. In this context, Zocher and coworkers showed that SAH acts as a non-competitive inhibitor with respect to the substrates L-Val, D-Hiv and ATP.[86] Upon omission of SAM, *N*-desmethyl enniatins are obtained at about a ten-fold lower synthesis rate (k_{cat} 0.13 s^{-1} for enniatin, k_{cat} 0.019 s^{-1} for desmethyl enniatin).[79,86] Likewise, limited SAM concentrations lead to the synthesis of partially demethylated enniatin analogs.

The overall kinetic parameters of ESYN from *Fusarium oxysporum* ETH 1536/9 were determined from double reciprocal plots.[86] The K_m values of ATP, L-Val and D-Hiv were determined to be 350 μM, 80 μM and 5 μM, respectively, and did not differ significantly in the presence of SAH. The K_m value of SAM was measured to be 10 μM, which is in the range of other methyltransferases.[87] At higher SAH concentrations, the synthesis of desmethyl enniatins was suppressed, in contrast to a continued, albeit decreased enniatin synthesis. This indicates that SAH only has an effect on the rate of COD synthesis, but not on substrate binding, and hence seems to interfere with elongation or product cyclization. From these findings, it can be surmised that suppression of desmethyl enniatin synthesis by SAH may have a regulatory role on enniatin synthesis *in vivo*. This hypothesis is supported by the finding that desmethyl enniatins are rarely found in fermentation broths, while they are produced during enniatin synthesis *in vitro*.[86]

COD formation can be detected by a colorimetric picrate assay following extraction of mycelia, broth or assay mixtures,[88] or quantified by radioactively labeled SAM or amino acids.[84,86] COD analogs may be separated by TLC or HPLC. More recently electrospray mass spectrometry (ESI-MS) combined with MS/MS fragmentation has been established as an alternative and more precise method for the detection and quantitation of enniatins,[89] as well as of PF1022-analogs.[90]

Using the enniatin synthetase as a model, the enzymatic reconstitution of the 350 kDa PF1022 synthetase was also achieved, and PF1022 congeners were produced *in vitro* with the substrates L-Leu, D-Lac, D-PheLac, Mg^{2+}, SAM and ATP.[61] The K_m values for product formation were determined to be K_m(D-Lac) = 0.77 ± 0.15 μM; K_m(D-PheLac) = 0.45 ± 0.12 μM; and K_m(L-Leu) = 20 ± 3 μM. In this *in vitro* approach, it was not only possible to detect all of the naturally occurring PF1022 congeners by mass spectrometry, but the rarely-occurring truncated hexa- and tetradepsipeptides were also observed. The later compounds were predicted to result from premature release from the PF1022 synthetase.

3.3 Cloning of CODS genes

The first CODS gene cloned was the enniatin synthetase from *Fusarium scirpi* ETH 1536/J5 (GenBank CAA79245).[91] A cDNA expression library constructed in phage λgt11 was screened with a polyclonal antibody raised against ESYN from *F. oxysporum* ETH 1536/9 and the identified gene fragment was used to probe a genomic library. A 9393 bp intron-free reading frame was confirmed by Northern hybridization. Identification of the synthetase gene was accomplished indirectly, by disrupting the corresponding enniatin synthetase gene in the plant pathogen *F. avenaceum* by homologous recombination.[53] A 1 kb fragment of the *F. scirpi* CODS gene was cloned into a plasmid carrying the hygromycin B phosphotransferase gene, and used to transform *F. avenaceum* protoplasts. Nonproducer transformants were further characterized and integrants identified which did not transcribe or express the enniatin synthetase gene. Interestingly, no single-copy transformant could be identified among the 19 nonproducers. The subcloned enniatin synthetase fragment of *F. avenaceum* had only a single nucleotide difference compared to the *F. scirpi* fragment used.

To clone the beauvericin and the bassianolide biosynthetic genes from the hypocrealean entomopathogen *Beauveria bassiana* ATCC 7159, Xu *et al.*[22,57] designed several different pairs of degenerate PCR primers against the conserved A3 and A8 motifs of NRPS A domains.[92] Some of these primers were heavily biased towards the D-Hiv-activating A domains of the *F. equiseti* enniatin synthetase (FeESYN).[91] Enniatins, similar to beauvericin and bassianolide, contain D-Hiv as their 2-hydroxycarboxylic acid constituents. Amongst the PCR products amplified using *B. bassiana* total DNA as a template, two distinct amplicons showed high sequence similarity to FeESYN. Fosmids from a genomic DNA library of *B. bassiana* that hybridized to these amplicons as probes were shown to derive from two disparate genomic loci. Sequencing revealed that each of these loci encodes one enniatin synthetase-like CODS. These CODSs were separately knocked out, and several isolates of the strains with the disrupted CODS genes, as well as ectopic integrants, were fermented under beauvericin/bassianolide production conditions. Ectopic integrants produced both beauvericin and bassianolide

at wild-type yields. Production of beauvericin was abrogated in the BbBEAS knockout strains, while BbBSLS knockouts were unable to produce bassianolide.[22,57]

The XsBSLS bassianolide synthetase of the wood-decaying fungus *Xylaria* sp. BCC1067 was identified by a PCR-based NRPS genome scanning strategy.[55] One of the NRPS-encoding amplicons was used as a probe to clone a CODS-encoding genomic locus from a λ phage genomic library. The *nrpsxy* gene encoding this CODS was disrupted on the chromosome of *Xylaria* sp. by directed gene knockout. Comparison of the metabolic profiles of the mycelial extracts of the wild-type and the mutant strains revealed the production of bassianolide in the wild-type strain, and the abrogation of production of this COD in the knockout strain.[55] Prior to this work, bassianolide production had not been described outside the hypocrealean entomopathogens *Beauveria* and *Verticillium*.

The RsPFSYN PF1022 synthetase gene (BD013055) was cloned from the unidentified fungus *Mycelia sterilia*.[93] This '*Fungus imperfectus*', classified later as *Rosellinia* sp., was isolated from the plant *Camellia japonica* in Japan.[94]

Two sequences of putative CODSs are also available from GenBank, although the CODs that these synthetases produce have not been experimentally established. ADB27871 is derived from *Fusarium oxysporum* FB1501, a strain that produces both beauvericin and the enniatins H, I and MK 1688.[19] AAY73200 is present in *F. venenatum* ATCC20334, a producer of enniatin B.[95] Further genome mining has identified a putative CODS (FOXG_11847) in *F. oxysporum* f. sp. *lycopersici*, a producer of enniatins. The current gene model for this CODS ends prematurely at a contig gap, and therefore the encoded CODS is a truncated protein with approximately 140 amino acids missing from the C-terminus. Two additional putative CODS were also found by genome mining from *Trichoderma* sp.: TRIVE1.E_GW1.16.170.1 from *T. virens*, and TRIAT1.E_GW1.1.2949.1 from *T. atroviridae*. To the best of our knowledge, no COD has been isolated from either of these *Trichoderma* species to date.

4 Functional anatomy of fungal CODSs

4.1 Overall structures

The 10 known full-length (or close to full-length) CODS are approximately 3150 amino acids in length and have a deduced molecular weight of ~350 kDa. They share 55–74% identity at the protein level over their entire length. Interestingly, the two CODS from *Fusarium oxysporum*, FolCODS from f. sp. *lycopersici* and FoCODS from strain FB1501, also share only 68% identity.

All identified CODS, regardless of producing hexa- or octadepsipeptides, display the same extended bimodular architecture with an identical domain arrangement: $C_1A_1T_1$–$C_2A_2M_2T_{2a}T_{2b}$–C_3 (Fig. 1). The only exception is the XsBSLS bassianolide synthetase from *Xylaria* sp. which has an additional "reductase" domain attached at its C-terminus (see Section 5.1).

The A_1 domain of the first module of the CODSs activates the D-2-hydroxycarboxylic acid substrate and loads it onto the T_1 domain in the same module, as shown experimentally in the case of the loading of D-Hiv onto the enniatin synthetase.[96,97] The A_2 domain of the second module activates and loads an L-amino acid substrate molecule onto each of the adjacent twin T_2 domains. The second module also features an *N*-methyltransferase domain (M_2) which is inserted into the A_2 domain between core motifs A8 and A9.[86] The flexible loop between these motifs often accommodates different editing domains in various NRPSs.[98,99] The SAM-dependent *N*-methyltransferase tailoring

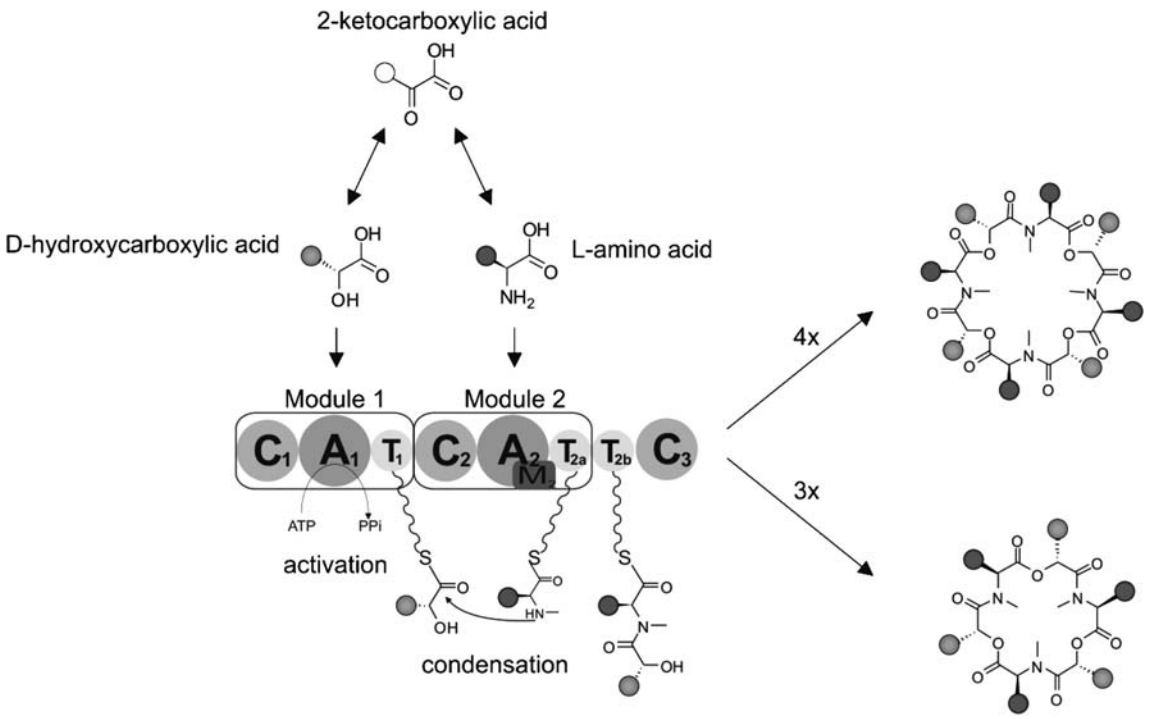


Fig. 1 Biosynthesis of CODs.

domains of NRPSs[82,86] modify the T domain-bound aminoacyl thioesters prior to condensation by the adjacent C domain,[100,101] and were first described and characterized from CODSs.[82,86]

Amide bond formation between the D-2-hydroxycarboxylic acid and *N*-Me-amino acid thioesters is carried out by the C_2 domain. NRPS C domains were shown to form several phylogenetic clades corresponding to functional subtypes.[102] Although the C_2 domains of CODS catalyze a condensation between substrates of D and an L configuration (*i.e.* they are formally $^{D}C_L$ domains), their core motifs are nevertheless more similar to domains of the $^{L}C_L$ subtype.[102] $^{D}C_L$ condensation domains evolved to perform a gating function to select upstream substrates of the D configuration from the racemic mixtures generated by the preceding E domains. CODS A_1 domains however specifically activate only the D enantiomer of the 2-hydroxycarboxylic acid substrate,[97] thus the C_2 domains do not have to select their substrates, nor do they have to interact with E domains. Peptide bond formation between the two substrates generates the dipeptidol monomer, three or four copies of which would then be ligated and finally cyclized in a programmed cyclooligomerization process to generate the cyclohexadepsipeptide or cyclooctadepsipeptide products, respectively.

This deduced assignment is based not only on the colinearity rule of NRPS organization, but is also supported by both limited proteolysis data and information from the expression of CODS fragments.[96,97] During the purification of the enniatin synthetase from *Fusarium scirpi*, endogenous proteolysis was observed, and two main fragments of 200 and 105 kDa were purified.[97] The 200 kDa fragment, comprising the $C_1A_1T_1C_2$ domains, catalyzed hydroxycarboxylic acid adenylation as well as its attachment to the T domain as a thioester. A similar fragment was also obtained from the *F. sambucinum* enniatin synthetase. As the N-terminus was presumably blocked, the N-terminal location of this fragment within the synthetase was inferred from monoclonal antibody (mAb) binding.[103] The 105 kDa fragment did not show catalytic activity, and was mapped by mAb-binding and N-terminal sequencing to the central region, comprising C_2 and a segment of A_2. Active site radiolabeling of ESYN and its fragments with substrates, followed by V8 protease digestion and HPLC separation, identified fragments containing the A_1T_1 didomain for hydroxycarboxylic acid attachment, and the $M_2T_{2a}T_{2b}$ domains for amino acid attachment. Both types of fragments were also shown to contain pantetheine as a required cofactor. Likewise, the SAM-binding site has been localized by radiolabeling and chymotryptic digestion to the M_2-region.[97]

The heterologous expression of various CODS fragments comprising the regions $C_1A_1T_1$ (121 kDa), $A_2M_2T_{2a}T_{2b}$ (158 kDa) and M_2 (65 kDa) of the FeESYN ennitatin synthetase has been achieved in *E. coli*.[96] Catalytic activities of adenylate formation have been demonstrated by D-Hiv and L-Val dependent ATP-PP_i-exchange. Activation of non-cognate substrates was also detected, but this did not exceed 15 to 20% of that with the cognate substrates. As misincorporation varied depending on the fragment size for the A_2-containing fragments, this promiscuity may be an artefact of improper folding. SAM binding by the regions containing the M_2 domain has also been validated by photolabeling with ^{14}C-SAM.

4.2 Adenylation domains

The two adenylation domains activate their respective substrates (D-hydroxycarboxylic acid [HA] for the A_1 domain and L-amino acid [AA] for the A_2 domain) as acyl adenylate intermediates (reactions (2a) and (2b)):

$$\mathbf{A_1} + \mathrm{HA} + \mathrm{ATP} \leftrightarrows \mathbf{A_1}(\mathrm{HA–AMP}) + \mathrm{PP_i} \quad (2a)$$

$$\mathbf{A_2} + \mathrm{AA} + \mathrm{ATP} \leftrightarrows \mathbf{A_2}(\mathrm{AA–AMP}) + \mathrm{PP_i} \quad (2b)$$

These activation reactions have been demonstrated by substrate-dependent ATP–PP_i exchange, relying on the reverse reaction to generate labelled ATP from radiolabelled PP_i and the acyl adenylate intermediate.[104]

Structure-guided alignments of bacterial A domain sequences can be used to predict the residues that line the substrate binding pocket of these enzymes. The predicted amino acid residues can then be used to predict the substrate specificity of the A domains of novel bacterial NRPSs, relying on comparisons with a large database of A domains with known substrate specificities.[3,105,106] A 10-amino acid "specificity code" (the "non-ribosomal code") is routinely used for these predictions.[105] This non-ribosomal code has recently been extended to a set of 34 amino acid residues, modeled to lie within a distance of 8 Å around the substrate.[107]

In contrast to A domains in bacterial NRPS enzymes, an *a priori* prediction of substrate specificity is currently not possible in fungal NRPS A domains. The substrate specificity signature motifs or "specificity codes" are divergent from those of bacteria, and there is a relative lack of A domains with known substrate specificities from fungal sources. Substrate predictions in CODS are thus very imprecise, and even "postdictions" (deriving consensus codes based on product structures) are somewhat equivocal.

4.2.1 Hydroxycarboxylic acid-activating domains (A_1). The A_1 domains of the fungal CODSs all contain a Gly substitution at the highly conserved Asp235 (GrsA numbering).[105] This Asp235 anchors the amino groups of amino acids, and its replacement is a hallmark for A domains incorporating non-amino acid precursors.[99] The 10-amino acids "codes" (Table 1) of the CODS A1 domains are remarkably similar to each other, and to the suggested D-Hiv consensus signature GALx(I/V)VG(S/T)IK.[22,57] The A_1 signature of the RsPFSYN PF1022 synthetase is also similar to the D-Hiv consensus. Although this domain incorporates D-PheLac and D-Lac *in vivo*, it shows remarkably relaxed substrate specificity *in vitro*.[90] Interestingly, no similarity is detected between the hydroxycarboxylic acid-activating A_1 domain signatures of the fungal CODSs and the 2-ketocarboxylic acid-activating A domain signatures derived from bacterial depsipeptide NRPSs.[22,57,108,109]

4.2.2 Amino acid activating domains (A_2). The 10-amino acid specificity codes of the A_2 domains of CODSs show a higher degree of variation (Table 2). A somewhat degenerate fungal Leu signature was proposed[57] that correctly identifies the Leu-specific A_2 domains of XsBSLS, RsPFSYN, and BbBSLS. It would predict Leu specificity for the A_2 domains of both CODSs from *Trichoderma* sp., and for the A_2 domain of FoCODS. However,

Table 1 Specificity-conferring signatures of CODS hydroxycarboxylic acid adenylation domains[a]

Synthetase[b]	Domain	Specificity[c]	235[d]	236	239	278	299	301	322	330	331	517
Fungal D-Hiv signature[e]			G	A	L	x	I/V	V	G	S/T	I	K
BbBEAS	A1	D-Hiv	G	A	L	M	I	V	G	S	I	K
BbBSLS	A1	D-Hiv	G	A	L	M	V	V	G	S	I	K
XsBSLS	A1	D-Hiv	G	A	L	L	V	V	G	T	I	K
FeESYN	A1	D-Hiv	G	A	L	H	V	V	G	S	I	K
FvCODS	A1	?	G	A	L	M	V	V	G	S	I	K
FolCODS	A1	?	G	A	L	M	V	V	G	S	I	K
FoCODS	A1	?	G	A	L	M	I	V	G	S	F	K
TaCODS	A1	?	G	A	V	I	V	V	A	T	I	K
TvCODS	A1	?	G	A	V	I	V	V	A	T	I	K
RsPFSYN	A1	PheLac/Lac	G	A	V	I	V	V	A	T	M	K

[a] Amino acids identical to those in the proposed fungal A domain consensus signature are shown in *white font* on *black background*. Amino acids similar (V = I = L; A = G, S = T, W = Y = F) to those in the proposed A domain consensus signature are shown in *bold type* over a *gray background*. [b] NRPS abbreviations: BbBEAS, *Beauveria bassiana* beauvericin synthetase;[22] BbBSLS, *B. bassiana* bassianolide synthetase;[57] FeESYN, *Fusarium equiseti* enniatin synthetase;[91] FoCODS, *F. oxysporum* cyclooligomer depsipeptide synthetase (ADB27871); FolCODS, *F. oxysporum* f. sp. *lycopersici* cyclooligomer depsipeptide synthetase (FOXG_11847); FvCODS, *F. venenatum* cyclooligomer depsipeptide synthetase (AAY73200); RsPFSYN, *Rosellinia* sp. PF1022 synthetase (BD013055); TaCODS, *Trichoderma atroviridae* cyclooligomer depsipeptide synthetase (TRIAT1.E_GW1.1.2949.1); TvCODS, *T. virens* cyclooligomer depsipeptide synthetase (TRIVE1.E_GW1.16.170.1); XsBSLS, *Xylaria* sp. bassianolide synthetase.[55] [c] A domain specificities: D-Hiv, D-2-hydroxyisovalerate; PheLac/Lac, D-phenyllactate and D-lactate. [d] Amino acid numbering according to the A domain of PheA.[105,106] [e] Fungal D-Hiv signature as proposed previously.[22]

Table 2 Specificity-conferring signatures of CODS amino acid adenylation domains[a]

Synthetase[b]	Domain	Specificity[c]	235[d]	236	239	278	299	301	322	330	331	517
Fungal Leu signature[e]			D	G/A	Y/W	L	I/V	G	G/A	V	x	K
XsBSLS	A2	Leu	D	A	W	L	V	G	A	V	M	K
RsPFSYN	A2	Leu	D	A	W	L	V	G	A	V	I	K
TaCODS	A2	?	D	A	W	L	V	G	A	V	I	K
TvCODS	A2	?	D	A	W	L	V	G	A	V	I	K
BbBSLS	A2	Leu	D	G	Y	I	I	G	G	V	F	K
FoCODS	A2	?	D	G	Y	M	I	A	A	V	F	K
BbBEAS	A2	Phe	D	G	Y	I	M	A	A	V	M	K
FeESYN	A2	Val/Leu/Ile	D	G	W	F	I	G	I	I	I	K
FvCODS	A2	?	D	G	Y	S	F	V	G	V	M	K
FolCODS	A2	?	D	G	Y	C	M	A	G	I	M	K
FsESYN	A2	Ile/Leu/Val	D	G	W	F	A	G	V	M	I	K

[a] Amino acids identical to those in the proposed fungal A domain consensus signature are shown in *white font* on a *black background*. Amino acids similar (V = I = L; A = G, S = T, W = Y = F) to those in the proposed A domain consensus signature are shown in *bold type* over a *gray background*. [b] FsESYN, *Fusarium sambucinum* enniatin synthetase;[81] see Table 1 for further NRPS abbreviations. [c] A domain specificities: Val/Leu/Ile, valine preferred, leucine and isoleucine also accepted; Ile/Leu/Val, isoleucine preferred, leucine and valine also accepted. [d] Amino acid numbering according to the A domain of PheA[105,106] [e] Fungal Leu signature, as proposed previously.[57]

the Phe-specific A_2 domain of BbBEAS is also very similar to this consensus sequence, leading to speculations of "keyhole surgery"-like mutations switching substrate specificity in closely related A domains.[4,6,22] The A_2 signatures of the FeESYN and FsESYN enniatin synthetases which incorporate branched chain amino acids (including Leu) are more distantly related to the Leu consensus sequence.

In a comparative analysis of purified enniatin synthetases from *Fusarium scirpi*, *F. sambucinum* and *F. lateritium*, amino acid specificities were analyzed by estimation of catalytic constants from product formation.[81] The most efficient synthesis was found with L-Val for the *F. scirpi* enzyme, L-Ile and L-Leu for the *F. sambucinum* enzyme, and L-Val for the *F. lateritium* enzyme. These trends are exactly what one would expect from the compositions of the enniatin congener mixtures produced by these strains *in vivo*.

4.3 Carrier domains

Transfer of acyl adenylates (reactions (3a) and (3b)) to the 4′-phosphopantetheine cofactor (S^{p}) attached to the cognate carrier domains (termed T_1 and T_{2a} and/or T_{2b}) has been shown by the isolation of acid stable thiol intermediates, concomitantly demonstrating the presence of *N*-methyl-aminoacyl thioesters preceeding ester bond formation:

$$\mathbf{A_1}(\text{HA–AMP}) + \mathbf{T_1}SH^{p1} \rightarrow \mathbf{T_1}\text{–}S^{p1}\text{–HA} + \text{AMP} + \mathbf{A_1} \quad (3a)$$

$$\mathbf{A_2}(\text{AA–AMP}) + \mathbf{T_2}SH^{p2} \rightarrow \mathbf{T_2}\text{–}S^{p2}\text{–AA} + \text{AMP} + \mathbf{A_2} \quad (3b)$$

Comparison of the T_1 and twin T_2 carrier domains show a relatively high primary sequence divergence. Both T_2 domains might interact with the adjacent *N*-methyl-transferase domain to facilitate *N*-methylation of the amino acid intermediate (reaction (4)). The twin T_2 domains contain more highly conserved charged residues than the T_1 domains, presumably to facilitate multiple docking events with the A, M and C domains, assuming that the primary docking events of these domains rely on electrostatic interactions. In one model proposed for cyclo-oligomerization in fungal CODS (see below), only the T_{2a} domain anchors the amino acid and its *N*-methyl intermediate, while the T_{2b} domain serves as a "waiting position" that holds the resulting dipeptidol intermediate while this awaits cyclo-oligomerization.[110] This scheme implies a lack of interaction between T_{2b} and A_2 and M_2.

4.4 Methyltransferase domains (M_2)

The aminoacyl thioester intermediate is *N*-methylated upon interaction of the T_2 carrier domain with the methyltransferase domain M_2 (reaction (4)), which is integrated between motifs A8 and A9 in the adenylation domain A_2.

$$\mathbf{T_2}\text{–}S^{p2}\text{–AA} + \text{SAM} \rightarrow \mathbf{T_2}\text{–}S^{p2}N\text{MeAA} + \text{SAH} \quad (4)$$

It is presently unknown if both of the twin carrier domains, or only the closely associated T_{2a} domain interact with the M_2 domain. The M_2 domain of the enniatin synthetase from *Fusarium scirpi* has been mapped by both protein chemical methods and fragment expression. The 49 kDa protein has been expressed in *Saccharomyces cerevisiae* with an N-terminal hexahistidine tag and a C-terminal streptavidin II fusion peptide.[100] Contacts with the substrate SAM have been assigned by saturation transfer difference (STD) NMR spectroscopy. Catalytic activity has been demonstrated with L-aminoacyl-*N*-acetylcysteamine thioesters (aminoacyl-SNACs) of substrate-related amino acids. K_m values of L-Val-SNAC and SAM were similar for enniatin synthetase and the expressed fragment, thus indicating correct folding.

In an analysis of methyltransferase domains of multifunctional PKS and NRPS systems of both bacterial and fungal origin, Ansari *et al.*[111] have shown that *N*-, *O*-and *C*-methyltransferase sequences from distinct subgroups. Structure-guided sequence alignments led to the identification of structural motifs in M_2 domains that are similar to those in non-integrated methyl transferases.[112]

4.5 Condensation domains (C)

The carrier domains loaded with thioesters of the hydroxycarboxylic acid and the *N*-methyl amino acid, respectively, will then interact with the condensation domain C_2 to form the dipeptidol intermediate HA-*N*MeAA, which remains bound to the acceptor thiolation domain T_2 (reaction (5)):

$$\mathbf{T_1}\text{–}S^{p1}\text{–HA} + \mathbf{T_2}\text{–}S^{p2}\text{–}N\text{MeAA} \rightarrow \mathbf{T_1}\text{–}SH^{p1} + \mathbf{T_2}\text{–}S^{p2}\text{–}N\text{MeAA–HA} \quad (5)$$

The isolation of the dipeptidol reaction intermediate shows that the peptide bond is formed first,[84] followed later by the ester bond-forming condensation and cyclization reactions. The production of the diketomorpholine bassiatin (**9**) from the beauvericin producer *Beauveria bassiana*, and the isolation of cyclo(Lac-MeLeu) and cyclo(PheLac-*N*-Me-Leu) from the PF1022-producer *Rosellinia* sp. (W. Weckwerth, *Doctoral Thesis*, TU Berlin, 1998) indicate that some dipeptidols may undergo an early cyclization reaction, instead of being used for cyclo-oligomerization. Indeed, such side products show bioactivity,[72] and their formation could be considered as a potential example of multiple product formation from a single NRPS.

The details and the sequence of the cyclooligomerization reactions that follow the formation of the first dipeptidol and finally lead to the release of the finalized COD product remain to be demonstrated. Both the twin T_2 domains and the N- and C-terminal C domains of CODS were proposed to take part in cyclooligomerization, as described in Section 4.6. The N-terminal C_1 and the C-terminal C_3 domains of fungal CODSs show overall sequence similarity to condensation domains, but their core motifs show substantial variation from the canonical forms.[22,57,99] C_3 domains are more conserved, with highly recognizable core motifs C2–C5. The core motif C3 that contains the canonical His active site (H**H**xxDG) is only slightly altered to S**H**ALYDG, and is apparently invariant in all the C_3 domains of CODSs. The C_3 domains show the highest similarity (~42% identity) to the C-terminal C domains of the aureobasidin A1 synthetase from *Aureobasidium pullulans* (ACJ04424) and the cyclosporine synthetase of *Tolypocladium inflatum* (CAA82227). These C-terminal C domains are predicted to catalyze macrocyclization by ester (aureobasidin synthetase) or peptide bond formation (cyclosporine).[114,115] On the other hand, the C_3

domains of CODS exhibit very low similarity to the C-terminal C domains of those bacterial NRPSs which catalyze macrocyclization by ester bond formation during rapamycin, FK506, FK520 and meridamycin biosynthesis, and display negligible similarity to bacterial TE domains that catalyze cyclooligomerization during enterobactin, bacillibactin, valinomycin and cereulide synthesis. The C_1 domains of CODSs are more divergent, with only a variant of the core motif C3 (**SH**xxVD) recognizable. The BbBEAS and the TaCODS C_1 domain active site signatures (HLxxxD and SYxxVD, respectively) even lack the His residue which is considered to be essential for condensation reactions.[22,113] The C_1 domains show no close similarity to any particular group of C domains outside CODSs. The divergence of the C_1 and C_3 domains of CODSs from the canonical amide bond-forming C domains of other NRPSs might be a consequence of their suggested role in the cyclooligomerization process, including the recursive ester bond-forming ligations and the product-releasing cyclization reaction. However, there is no experimental evidence to support the hypothetical functions of either the C_1 or C_3 domains in cyclodepsipeptide synthesis to date.

4.6 Iterative and recursive processes during cylooligomerisation

One of the most interesting processes during COD biosynthesis is the oligomerization and cyclization of the peptidol monomer products on the CODS. According to the classical "parallel model" (Figs. 1 and 2), formulated for enniatin biosynthesis by the Zocher group[91,110] and adapted for all fungal CODSs, the dipeptidol monomer formed from the condensation of the 2-hydroxycarboxylic acid and the *N*-Me-amino acid will remain bound on the T_{2a} domain. The dipeptidol intermediate may then be transferred from carrier domain T_{2a} to T_{2b}, with the latter serving as a dedicated "waiting position" only. Alternatively, the

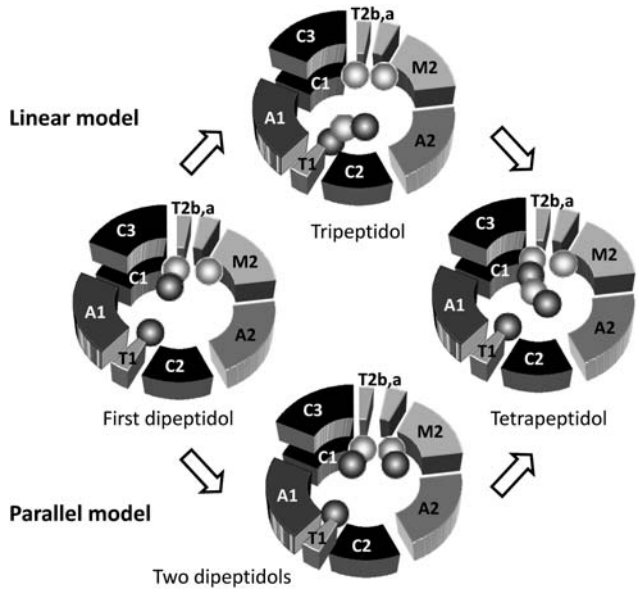


Fig. 2 Models for COD biosynthesis *via* stepwise assembly (Linear model) or oligomerization (Parallel model). Dark grey spheres represent 2-hydroxycarboxylic acid moieties; light grey spheres symbolize amino acid moieties. See text for details.

T_{2b} domain might be available for the sequential assembly of a second peptidol intermediate: in this scenario the T_{2a} and T_{2b} domains would be functionally equivalent, and the "waiting position" would result from the dynamics of the process. Either way, the twin domains T_{2a} and T_{2b}, each loaded with a dipeptidol, were proposed to be involved in the formation of a tetrapeptidol by dimerization (Fig. 2). The formation of this tetrapeptidol may be catalyzed by the condensation domain(s) C_3 and/or C_1 (reaction (6)), and this intermediate remains attached to either T_{2a} or T_{2b}:

$$\mathbf{T_{2a}}\text{–}S^{p2}\text{–}N\text{MeAA–HA} + \mathbf{T_{2b}}\text{–}S^{p3}\text{–}N\text{MeAA–HA} \rightarrow \mathbf{T_{2a}}\text{–}SH^{p2} + \mathbf{T_{2b}}\text{–}S^{p3}\text{–}[N\text{MeAA–HA}]_2 \quad (6)$$

Early release of the enzyme-bound tetrapeptidols in the form of cyclic tetradepsipeptide products was observed *in vitro* in PF1022 synthesis (W. Weckwerth, *Doctoral Thesis*, TU Berlin, 1998). The same truncated compounds with the structures cyclo(D-Lac–*N*-Me-L-Leu–D-Lac–*N*-Me-L-Leu), cyclo(D-Lac–*N*-Me-L-Leu–D-PheLac–*N*-Me-L-Leu), and *cyclo*(D-PheLac–*N*-Me-L-Leu–D-PheLac–*N*-Me-L-Leu) were also found to be formed during fermentation of the wild-type producer strain.

After the tetrapeptidol stage, repetition of reactions (2)–(5) leads to a synthetase with a dipeptidol intermediate anchored at one of the T_2 domains, which could be ligated by the C_1 and/or the C_3 domains with the tetrapeptidol intermediate parked on the other T_2 domain. This leads to the formation of the hexapeptidol intermediate (reaction (7a)).

$$\mathbf{T_{2a}}\text{–}S^{p2}\text{–}N\text{MeAA–HA} + \mathbf{T_{2b}}\text{–}S^{p3}\text{–}[N\text{MeAA–HA}]_2 \rightarrow \mathbf{T_{2a}}\text{–}SH^{p2} + \mathbf{T_{2b}}\text{–}S^{p3}\text{–}[N\text{MeAA–HA}]_3 \quad (7a)$$

In case of octapeptidols such as PF1022, a further dipeptidol assembly and ligation cycle is envisioned. After the appropriate number of recursive intermolecular ligations ($n = 3$ for cyclohexadepsipeptides and $n = 4$ for cyclooctadepsipeptides), the linear oligomer might fold back and become a substrate for the intramolecular cyclization that releases the final cyclooligomer product (equation (7b)). This product release reaction is analogous to that catalyzed by cyclizing C or TE domains, and has been proposed to be carried out by one or both of the C_1 and the C_3 domains.[91,110]

$$\mathbf{T_{2b}}\text{–}S^{p3}\text{–}[N\text{MeAA–HA}]_n \rightarrow \mathbf{T_{2b}}\text{–}SH^{p3} + \text{COD} \quad (7b)$$

An alternative mechanism to the classic "parallel" model of cyclooligomerization would involve the buildup of cyclooligomer depsipeptides on the enzyme by stepwise iterative condensations (Fig. 2). During this "linear" mechanism, dipeptidol formation on the T_2 domains would be followed by condensation with D-Hiv presented on the T_1 domain, catalyzed by either or both the C_1 or C_3 domains, leading to a tripeptidol. The tripeptidol would then be condensed with the *N*-Me-amino acid on the T_2 domain to form the tetrapeptidol. The process would continue in a stepwise manner until the appropriate chain length is achieved and synthesis is terminated by cyclization as catalyzed by either or both of the C_1 or C_3 domains. In this model, the presence of the two copies of the T_2 domains would not be a structural

requirement of COD biosynthesis, and might only increase the turnover of the enzyme, a mechanism which has a precedent in polyunsaturated fatty acid biosynthesis.[116,117]

Experimental evidence for the "parallel" model is limited at the moment, and does not conclusively rule out the "linear" mechanism. The enniatin synthetase was shown to be a monomer,[110] and thus the active sites on a single synthetase should be sufficient for the formation of a COD, excluding the possibility that three synthetase subunits would each contribute one dipeptidol to the final product. Performic acid release of products after a brief *in vitro* condensation reaction with the purified FeESYN enniatin synthetase yielded only the dipeptidol and the tetrapeptidol,[84] supporting the "parallel" mechanism. However, product yield in these pioneering experiments was extremely low. Therefore the formation of tri- and/or pentapeptidols might have gone undetected if synthesis of these species was rate limiting compared with their condensation with amino acids to yield the tetra- and hexapeptidols. Further evidence for the "parallel" mechanism comes from mass spectra of the products of PF1022 fermentations, in which only the even-numbered truncated products (diketomorpholines, tetra- and hexadepsipeptides) could be detected.[61] Further investigations by CODS domain engineering, and Fourier-transform mass spectrometric detection and identification of the enzyme-bound intermediates promises to shed more light on this interesting process.[118,119]

4.7 Cyclooligomerization and the phylogeny of CODSs

Fungal CODSs are programmed to utilize three (hexadepsipeptides) or four (octadepsipeptides) of the dipeptidol monomers during cyclooligomerization. The strict control of the number of monomer ligations catalyzed by these enzymes might thus be expected to be reflected in an appropriate clading of the CODSs, especially their C domains, during multiple sequence alignments. However, this is not the case: alignments with full-size CODS, or their individual C domains, fail to reveal clustering according to the oligomerization state, or even product structure (Fig. 3). Thus, the trimerizing BbBEAS beauvericin synthetase, and the tetramer-forming BbBSLS bassianolide synthetase form a clade that is a sister of a clade containing all the CODS from *Fusarium* spp. These sister clades branch from another clade containing sequences from Xylariaceae (XsBSLS and RsPFSYN) and from *Trichoderma*. The small number of CODS sequences, and the uncertainty as to the products of the *Trichoderma* and some of the *Fusarium* CODSs, does not currently allow us to discern the underlying evolutionary history of these enzymes, and we also cannot correlate oligomerization state with sequence motifs. Programming of the product chain length of other iterative enzymes (for example the fungal polyketide synthases, bacterial Type II polyketide synthases, and the chalcone synthase-like Type III polyketide synthases) also cannot reliably be predicted from primary amino acid sequences, in spite of recent progress.[120] The number of ligations during cyclooligomerization (*i.e.* product oligomerization state) is thus probably determined by the shape and size of the reaction chamber within the CODSs, and might be modulable by a small number of mutations that can transform a CODS from a trimer-forming to a tetramer-forming enzyme (or *vice versa*) after gene duplication and drift.

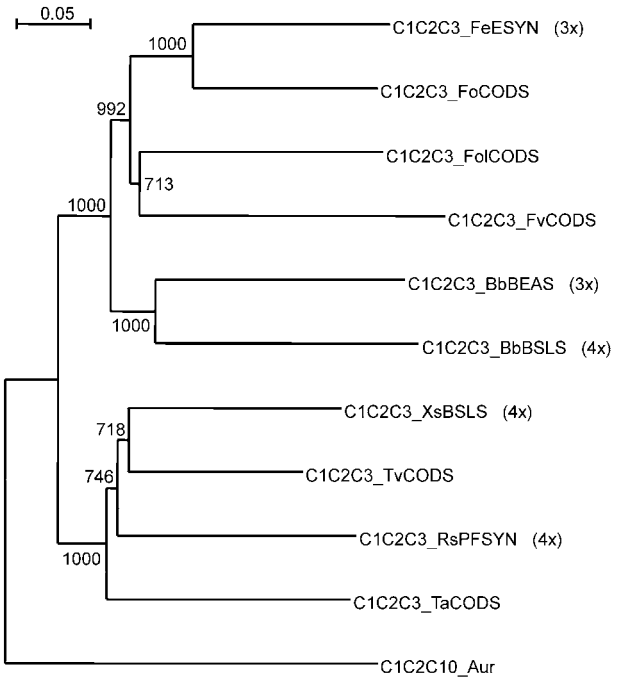


Fig. 3 Phylogenomic analysis of CODSs. The sequences of the C_1, C_2, and C_3 domains of the CODSs were concatenated, a multiple sequence alignment was created in VectorNTI, and bootstrapped trees were calculated in ClustalX with the neighbor-joining method using 1000 repeats. The phylogram was plotted with NJPlot using C1C2C10_Aur, the concatenated sequences of the C_1, C_2 and C_{10} domains of the aureobasidin synthetase,[114] as the outgroup. The scale shows the number of substitutions per site, and significant (>500) bootstrap values are indicated near the forks. (3×) and (4×) indicates the cyclotrimeric or cyclotetrameric nature of the known COD products, respectively.

4.8 Cyclooligomerization during fungal and bacterial COD biosynthesis

Similar to fungal CODs, bacterial cyclooligomer peptides (*e.g.*, gramicidin S) and depsipeptides (*e.g.*, enterobactin, valinomycin, cereulide, and chromodepsipeptides) are also biosynthesized by Type B iterative NRPSs. While the bacterial synthetases also have modules that are responsible for the programmed iterative synthesis of the peptide/peptidol monomer units, their cyclooligomerization mechanisms involve different domains compared to those in the fungal CODSs (Fig. 4). Bacterial cyclooligomerizing synthetases do not have N- and C-terminal C domains, or twin T domains in their second module. Instead, these enzymes feature a recursive TE domain, located at their C-terminus. In the example of the enterobactin synthetase, this TE holds the first monomer unit as an acyl-*O*-TE ester intermediate on its active site Ser, until the next monomer unit is assembled on the adjacent T domain. The recursive TE domain then ligates two monomers to form a dimer that can be released by TE-catalyzed cyclization, or might serve as a partner for further oligomerizations and the cyclization (Fig. 4).[8,118,121–123] Excised recursive TE domains from the gramicidin S and the thiocoraline synthetases have been shown to catalyze cyclodimerization *in vitro*.[124,125] Considering the radically different enzymology of the bacterial *vs.* the fungal cyclooligomerizing synthetases, recursive ligation and cyclization catalysis during COD biosynthesis must have

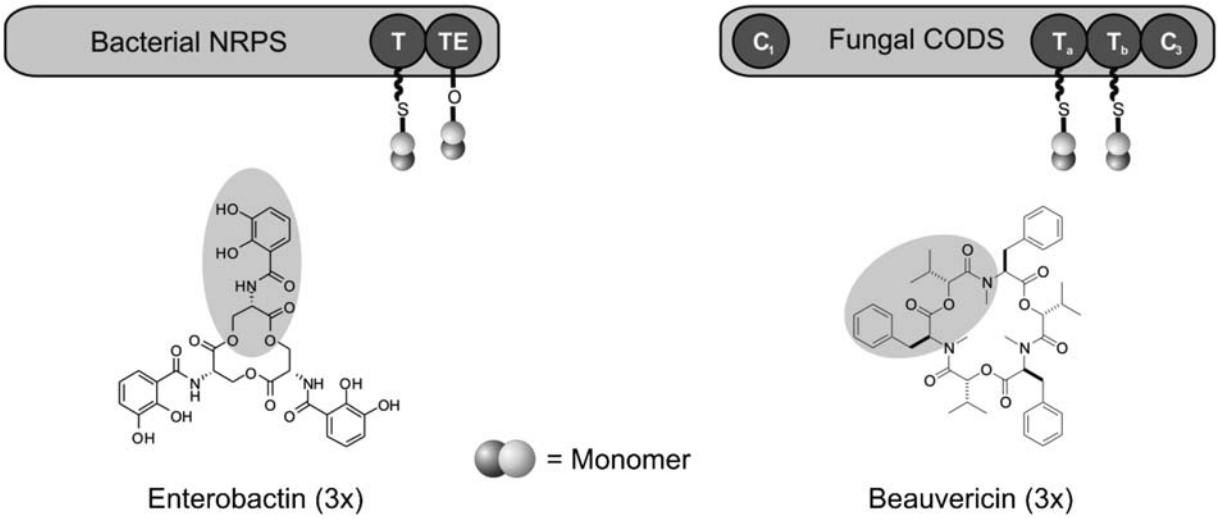


Fig. 4 Cyclooligomerization during COD biosynthesis in bacteria *vs.* fungi. See text for details.

a polyphyletic origin and have developed through convergent evolution.

4.9 Pseudo-cyclodimeric fungal natural products

Fungal cyclodepsipeptides showing a pseudo-cyclodimeric structure have been isolated from several Dothiodeomycete *Pithomyces* spp. (teleomorph: *Leptosphaerulina* spp.). Thus, the cyclotetradepsipeptide angolide (**5**)[12] and the cyclohexadepsipeptide sporidesmolide (**6**)[13] are apparent dimers of a dipeptidol (angolide) or tripeptidol (sporidesmolide). However, the two "monomers" are not strictly equivalent: one of these units contains D and L amino acids, while the other features exclusively L amino acids. Two hypotheses have been advanced to explain this non-equivalence of the two halves of these molecules. For angolide, a cyclic, all-L dipeptide intermediate, possibly a diketopiperazine, was proposed to undergo random enzyme-catalyzed inversion of the α-position of one residue only, followed by *N*-acylation and insertion of the two L-Hiv residues, respectively.[126] An alternative mechanism was also suggested for sporidesmolide and for angolide involving the biosynthesis of two depsipeptidols on two nonequivalent "sites" of a multienzyme, followed by cyclodimerization and multienzyme-controlled specific isomerization.[126,127] We propose that the simplest biosynthetic model for these compounds would involve a four-module (angolide) or six-module (sporidesmolide) type A (processive) NRPS (Fig. 5). The presence of epimerization domains in the modules for D amino acids would account for the incorporation of L amino acid precursors into all amino acid positions, with a cyclizing (but not cyclooligomerizing) C-terminal TE or C domain releasing the cyclic product from the enzyme. Confirmation of this hypothesis awaits sequencing of the respective gene clusters.

5 Precursor supply, regulation and export

5.1 Biosynthesis of 2-hydroxycarboxylic acids in COD-producing fungi

Fungi biosynthesize the D-2-hydroxycarboxylic acid constituents of their CODs by reducing the appropriate free ketocarboxylic acids, derived from amino acid metabolism. COD biosynthesis seems to have recruited enzymes and their encoding genes from two distinct superfamilies for this purpose. Thus, beauvericin biosynthesis in *Beauveria bassiana*, enniatins biosynthesis in *Fusarium* spp., and bassianolide biosynthesis in *Xylaria* sp. utilize ketoisovalerate reductases of the 6-phosphogluconate dehydrogenase superfamily to produce D-Hiv. In contrast, PF1022 biosynthesis in *Rosellinia*, and the biosynthesis of the unknown CODS of *Trichoderma virens* and *T. atroviridae* seems to have

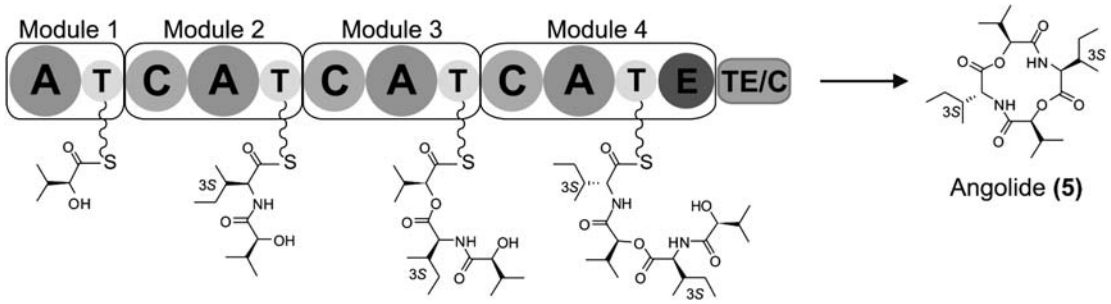


Fig. 5 Proposed biosynthesis of angolide, a pseudo-cyclodimeric fungal natural product. One of the D-Hiv moieties was arbitrarily assigned as the starter unit. The picture shows the growing depsipeptide chain as the appropriate intermediates anchored on the T domains. See text for details.

co-opted dehydrogenases from the D-isomer-specific 2-hydroxyacid dehydrogenase superfamily.

The Zocher lab described the purification and enzymatic characterization of a 53 kDa "D-hydroxyisovalerate dehydrogenase" from *Fusarium sambucinum*.[128–130] This enzyme was shown to catalyze the reversible interconversion of ketoisovalerate and D-Hiv, with high enantioselectivity in an ordered bi-bi kinetic mechanism. The sequential order of the reaction was found to be identical to that of ketopantoate reductases from the 6-phosphogluconate dehydrogenase superfamily. Ketopantoate reductases (E.C. 1.1.1.169) catalyze the NADPH-dependent stereospecific reduction of ketopantoate to D-pantoate in vitamin B_5 biosynthesis.[131] The *F. sambucinum* D-Hiv dehydrogenase displayed high substrate specificity, and was specific for $NADP^+$. However, the sequence of the protein and its encoding gene have not been reported.

The BbBEAS beauvericin synthetase locus of *Beauveria bassiana* was found to contain a gene (*kivr*) encoding a putative protein with a GxGxxGxxxA NAD(P)H-binding signature and high similarity to COG1893 ketopantoate reductases.[132] The predicted KIVR protein had a deduced MW = 51 493 Da, in good agreement with the size of the *Fusarium sambucinum* D-Hiv dehydrogenase enzyme. KIVR was predicted to show a similar secondary structure to those of ketopantoate reductases. It was also predicted to share a Glu-Asn-Lys active site triad architecture with them, as well as key conserved amino acids involved in substrate and product orientation. No similar gene was clustered with the BbBSLS bassianolide synthetase of the same strain. KIVR was expected to supply D-Hiv for the biosynthesis of beauvericin and perhaps bassianolide: accordingly, disruption of the *kivr* gene in the genome of *B. bassiana* abrogated not only the production of beauvericin, but also that of bassianolide. Chemical complementation of the mutant by supplementing the fermentation medium with D-Hiv restored the production of both CODs. Thus, KIVR is the only enzyme that can produce D-Hiv in *B. bassiana* for the biosynthesis of both beauvericin and bassianolide, thereby representing a functional crosstalk between the two COD biosynthetic systems of the strain.[132]

Immediately upstream of the FolESYN enniatin synthetase of *Fusarium oxysporum* f. sp. *lycopersici* is a divergently transcribed gene (FOXG_11846) that encodes a putative ketoisovalerate reductase which is 61% identical and 76% similar to the *B. bassiana* KIVR. The respective predicted protein FOXG_11846.2 was erroneously annotated in a C-terminally truncated form, but this shortened version of 367 amino acids was successfully expressed in *E. coli*, although mostly in the form of inclusion bodies (P. Grzesik, *Diploma Thesis*, TU Berlin, 2009). The N-terminally His-tagged construct of about 43 kDa showed KIV-dependent NADPH-consumption with a K_m of 2.5 mM, compared to 0.2 mM for the 53 kDa dehydrogenase isolated from *F. sambucinum*.

The uncharacterized reductase domain appended to the C-terminus of the XsBSLS bassianolide synthetase of *Xylaria* sp. (amino acids 3136–3546) also shows 25% identity and 43% similarity to KIVR. Both FOXG_11846 and the R domain of XsBSLS retain the $NADP^+$-binding site and the Glu-Asn-Lys active site triad architecture of the *B. bassiana* KIVR and the related ketopantoate reductases from the 6-phosphogluconate dehydrogenase superfamily, and likely supply D-Hiv or similar D-2-hydroxycarboxylic acids for their cognate CODS partners.

In contrast, the putative keto(phenyl)propionate reductase BD105415 which is clustered with the RsPFSYN of *Rosellinia* sp. displays a very low (10%) identity to the KIVR proteins. The Glu-Asn-Lys active-site triad of the KIVR sequences or the ketopantoate reductases from the 6-phosphogluconate dehydrogenase superfamily are not retained in BD105415, nor are the additional residues involved in the stabilization of the substrate or the product. Instead, BD105415 reveals significant similarity to lactate dehydrogenases (COG1052) within the D-isomer-specific 2-hydroxyacid dehydrogenase superfamily, with an $NAD(P)^+$-binding Rossmann fold (cl09931) at the C-terminal half of the protein. Similarly, the *Trichoderma* CODSs are also clustered with putative $NAD(P)^+$-binding, D-isomer-specific 2-hydroxyacid dehydrogenases (*T. virens*: e_gw1.82.328.1, *T. atroviridae*: fgenesh1_pm.contig_27_#_418 and Triat1.e_gw1.1.3874.1). Interestingly, while the *T. virens* e_gw1.82.328.1 dehydrogenase and the *T. atroviridae* Triat1.e_gw1.1.3874.1 enzymes are 86% identical at the protein level, the two *T. atroviridae* dehydrogenase protein sequences, both bordering the same CODS, share only 24% identity.

Biosynthesis-guided purification studies in the PF1022 producer led to the purification of a 38 kDa D-phenyllactate dehydrogenase (W. Weckwerth, *Doctoral Thesis*, TU Berlin, 1998). In these studies, the phenylpyruvate dependence of COD biosynthesis was monitored in a reconstituted reaction system containing the PF1022 synthetase, Leu, ATP, SAM, and NADPH. The D-phenyllactate dehydrogenase enzyme was purified about 5000-fold in 7 steps, and shown to reduce phenylpyruvate with a K_m = 38 μM. Besides phenylpyruvate, *p*-hydroxyphenyl-pyruvate (K_m = 45 μM) and 2-ketoisocaproate (K_m = 53 μM) were also accepted as substrates. Results from gel filtration experiments indicated a dimeric structure. Internal tryptic peptides of the purified enzyme showed some similarity to BD105415 (predicted size 36 470 Da).

5.2 Hydroxycarboxylic acid incorporation in bacterial *vs.* fungal NRPS systems

Incorporation of intra-chain 2-hydroxycarboxylic acids into bacterial cyclodepsipeptide products (*e.g.*, kutznerides, valinomycin, cereulide, hectochlorin and cryptophycin) was shown to proceed by a mechanism different from that in COD-biosynthesizing fungi.[98,108,109,133] Instead of utilizing preformed 2-hydroxycarboxylic acids as substrates, the respective bacterial NRPS A domains activate and load the corresponding 2-ketocarboxylic acids onto the multienzymes. Here, the ketocarboxylic acyl-thioesters undergo stereospecific reduction *in cis* by a ketoreductase (KR) domain, yielding the D- or L-hydroxycarboxylic acyl-thioesters ready for condensation. The on-demand production of hydroxycarboxylic acids in bacteria, and the biosynthesis of a free pool of hydroxycarboxylic acids in fungi both rely on the readily available ketocarboxylic acid pool, derived from amino acid catabolism and anabolism. However, these different hydroxycarboxylic acid supply routes still conceivably represent another example of convergent evolution for the generation of a chiral precursor for natural product biosynthesis in bacteria *vs.* fungi.

5.3 Regulation and export

While most of the genes in the immediate genomic neighborhood of the identified CODS encode hypothetical proteins with no predicted functions, some genes code for predicted proteins with potential functions in the transcriptional regulation of the synthetase, or in the export of the produced COD. However, there is no direct proof for the involvement of any of these predicted proteins in COD biosynthesis.

Genes encoding putative regulatory proteins containing the Gal4-like Zn_2Cys_6 binuclear cluster DNA-binding domain (Smart SM00066, InterPro IPR001138) were found to be clustered with BbBEAS (*orf1*),[22] and FolCODS (FOXG_11849 and FOXG_11859). The BbBSLS bassianolide synthetase cluster also encodes a putative Gal4-like transcriptional regulator (ORF5) and the predicted GTPases ORFs 1 and 4.[57] A predicted Gal4-like transcription factor (gw1.82.211.1) is adjacent to the *T. virens* CODS, while the *T. atroviridae* CODS is clustered with a putative TFIIS-type zinc finger transcription factor (Triat1.e_gw1.1.2849.1).

The *nrpsxy* gene encoding the XsBSLS bassianolide synthetase of *Xylaria* sp. was found to be clustered with the *efxy* gene encoding a putative major facilitator superfamily (MFS) transporter with significant similarity to other MFS transporters encoded in many fungal genomes.[55] The two *Trichoderma* CODSs are also clustered with hypothetical MFS transporters (*T. virens*: fgenesh1_pg.82_#_180; *T. atroviridae*: Triat1.e_gw1.1.3461.1). The *Fusarium oxysporum* f. sp. *lycopersici* FolCODS is clustered with FOXG_11845 which encodes a predicted ATP binding cassette (ABC) multidrug transporter. No transport-related putative proteins were found to be encoded in the sequenced regions of the beauvericin or the bassianolide clusters of *Beauveria bassiana*.[22,57]

6 Heterologous expression of CODSs

6.1 Enniatin synthetase gene fragments

Heterologous expression of sub-fragments of FeESYN was achieved by Zocher and coworkers, aiming at high expression yields for subsequent biochemical characterizations. The first module (the "EA fragment"), including the $C_1A_1T_1$ domains, and the second module (the "EB fragment"), including the $A_2M_2T_{2a}T_{2b}$-and a truncated C_3 domain, were separately cloned into the *E. coli* vectors pBluescript SK+ and pUC8. Expression levels of the ESYN fragments cloned into the pBluescript vector reached 20–40% of the total cellular protein, compared to the pUC8 constructs which yielded 10–20% recombinant protein. Starting from the EB fragment-encoding gene fragment, several truncated constructs (*e.g.* A_2, A_2M_2, M_2) were also made.[96] Unfortunately, mostly insoluble proteins were obtained that needed to be denaturated and refolded, significantly increasing the experimental efforts for obtaining functional proteins. The activities of both adenylation domains of FeESYN (A_1 and A_2) were characterized using the ATP-PP_i exchange reaction.[104] However, further assays that measure the covalent loading of the substrates (D-Hiv and L-Val) on the T domains of the expressed protein fragments as hydroxyacyl or aminoacyl thioesters were unsuccessful. This suggested that the phosphopantetheinyl transferase of the *E. coli* fatty acid synthase does not recognize the FeESYN as its substrate.[22] The truncated constructs were also tested for methyltransferase activity, and UV-induced photoaffinity labeling[134] with [^{14}C]- or [^{3}H]-SAM has also been performed in order to prove specific covalent binding of this cosubstrate to the M_2 domain.[96] Truncated fragments of the FeESYN containing the M_2 domain could be photolabeled, whereas fragments without the predicted M_2 domain were not. Another interesting outcome of the experiments was that a truncated protein composed of the N-terminal subdomain of A_2 with M_2 was not able to activate the substrate L-Val. Thus, the activation is based on both of the N- and C-terminal subdomains of the adenylation domain.[135] In subsequent studies several deletion constructs encoding segments of the M_2 domain were generated, and the expressed recombinant proteins were tested by radiolabeling with SAM. These studies addressed the role in cofactor binding of the four conserved motifs of methyltransferases (I: VLEIGTGSGMIL; II: SYVGLDPS; IV: DLVVFNSVVQYFTPPEYL and V: ATNGHFLAARA).[101] A deletion of 11 amino acids from the N-terminus led to a decrease in the cofactor binding, but surprisingly, M_2 fragments with a 21-amino acid N-terminal deletion displayed an even higher binding activity than the embedded M_2-domain in the wild-type FeESYN. Although the conserved motifs are closer to the N-terminus (whole protein: 558 amino acids, start of motif I: 44 aa; motif II: 91 aa; motif IV: 138 aa and motif V: 180 aa from the N-terminus), the truncation of the C-terminal part led to a more significant decrease of the SAM binding of the enzyme.[101]

Since the attempts to obtain soluble single domains of the FeESYN in *E. coli* were not successful, the M_2 methyltransferase domain was overexpressed in yeast, yielding the desired protein in soluble fraction. The M_2 domain (1.3 kbp) was cloned in the *E. coli–S. cerevisiae* shuttle vector pYEXTHS-BN with an N-terminal His_6-tag and C-terminal strep II fusion peptide. Cofactor binding was demonstrated by photoaffinity labeling[134] and by saturation transfer difference (STD)-NMR spectroscopy under equilibrium conditions, establishing the distance and orientation of enzyme-bound SAM relative to the binding site. The kinetic constants for binding of the cofactor and the substrate were also determined, and were shown to be similar to those of the M_2 domain embedded in FeESYN, indicating that the dissected domain was correctly folded upon heterologous expression. The specificity of the methyltransferase was investigated using *N*-acetylcysteamine thioesters (SNAC) of L-Leu, L-Ile, L-Phe, L-Val and D-Val. Surprisingly, all tested amino acids except D-valine yielded methylated products at similar rates, as detected by radioactive labeling and MALDI-TOF mass spectrometry. Although the M_2 domain accepts only amino acids with an L-configuration, it apparently has a widened substrate tolerance, even for substrates which had previously been shown not to be substrates for the full-length enniatin synthetase, such as L-Phe.[100]

6.2 Heterologous production of beauvericin

The BbBEAS beauvericin synthetase of *Beauveria bassiana* has been functionally expressed in *E. coli* using the expression vector pACYCDuet-1 (Novagen).[22] The Duet system allows the cloning of two ORFs behind two separate T7 promoters for co-expression from the same vector: Xu *et al.*[22] used this feature to

co-express *bbBeas* with the broad-spectrum phosphopantetheinyl transferase *sfp* gene from *Bacillus subtilis*[136] to guarantee the production of holo-BbBEAS. *E. coli* BL21 (DE3) carrying the expression construct produced small amounts of beauvericin upon supplementation of the cultures with D-Hiv. Heterologous production of beauvericin was improved by using Origami B(DE3) as the expression host: this strain facilitates protein folding by promoting cytoplasmic disulfide bond formation. Reduction of the cultivation temperature to 16 °C during protein expression, and feeding both D-Hiv and L-Phe during the beauvericin production stage led to a process with a beauvericin yield of approximately 8 mg/L in *E. coli*. This represents approximately 40% of the yield of the native producer *B. bassiana*, and a reduction in fermentation times. Importantly, beauvericin from *E. coli* was fully *N*-methylated, but remained associated with the cells.

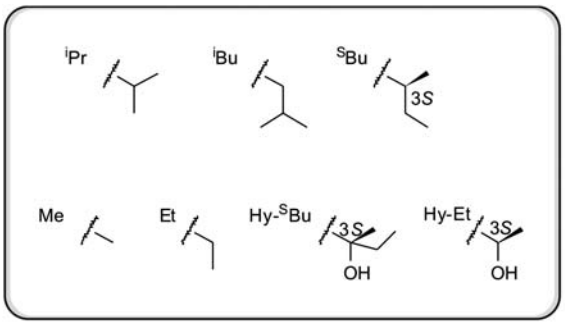
Heterologous expression of CODS, and heterologous production of CODs in strains that are easier to manipulate genetically than the native fungal producers will facilitate the enzymatic characterization of CODS and their engineered variants, and will promote the combinatorial biosynthesis of novel CODs.

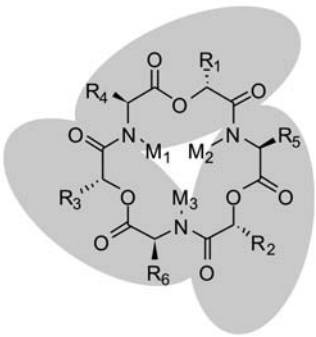
7 Structural diversification of CODs

7.1 Natural COD congeners

Twenty-seven natural enniatin congeners have been isolated and characterized to date from various enniatin producer fungi

Table 3 Naturally occurring enniatin congeners





Enniatin	R_1	R_2	R_3	R_4	R_5	R_6	M_1	M_2	M_3	First isolated from	Ref.
A	iPr	iPr	iPr	sBu	sBu	sBu	Me	Me	Me	*Fusarium orthoceras* var. *enniatinum* ETH 1523 and *F. scirpi* ETH 1536	152
A1	iPr	iPr	iPr	iPr	sBu	sBu	Me	Me	Me	*Fusarium roseum acuminatum*	153
B	iPr	iPr	iPr	iPr	iPr	iPr	Me	Me	Me	*Fusarium* spp. ETH 4363 and ETH 1574	152
B1	iPr	iPr	iPr	sBu	iPr	iPr	Me	Me	Me	*Fusarium roseum acuminatum*	153
B2	iPr	iPr	iPr	iPr	iPr	iPr	H	Me	Me	*Fusarium avanaceum*	33
B3	iPr	iPr	iPr	iPr	iPr	iPr	H	H	Me	*Fusarium avanaceum*	33
C	iPr	iPr	iPr	iBu	iBu	iBu	Me	Me	Me	*Fusarium* spp. ETH 4363 and ETH 1574	152
D (= B4)	iPr	iPr	iPr	iBu	iPr	iPr	Me	Me	Me	*Fusarium* sp. FO-1305	24
E[a]	iPr	iPr	iPr	iBu	iPr	sBu	Me	Me	Me	*Fusarium* sp. FO-1305	24
	iPr	iPr	iPr	iBu	sBu	iPr					
F	iPr	iPr	iPr	iBu	sBu	sBu	Me	Me	Me	*Fusarium* sp. FO-1305	24
G	iPr	iPr	iPr	iPr	iBu	iBu	Me	Me	Me	*Halosarpheia* sp. strain 732	47
H	sBu	iPr	iPr	iPr	iPr	iPr	Me	Me	Me	*Verticillium hemipterigenum* BCC 1449	144
I	sBu	sBu	iPr	iPr	iPr	iPr	Me	Me	Me	*Verticillium hemipterigenum* BCC 1449	144
MK 1688	sBu	sBu	sBu	iPr	iPr	iPr	Me	Me	Me	*Verticillium hemipterigenum* BCC 1449	144
J1	iPr	iPr	iPr	Me	iPr	iPr	Me	Me	Me	*Fusarium* sp. strain F31	142
J2	iPr	iPr	iPr	iPr	sBu	Me	Me	Me	Me	*Fusarium* sp. strain F31	142
J3	iPr	iPr	iPr	iPr	Me	sBu	Me	Me	Me	*Fusarium* sp. strain F31	142
K1	iPr	iPr	iPr	Et	iPr	iPr	Me	Me	Me	*Fusarium* sp. strain F31	142
L	iPr	iPr	hy-sBu	iPr	iPr	iPr	Me	Me	Me	Unidentified fungus (BCC 2629)	154
M1	iPr	sBu	hy-sBu	iPr	iPr	iPr	Me	Me	Me	Unidentified fungus (BCC 2629)	154
M2	iPr	hy-sBu	sBu	iPr	iPr	iPr	Me	Me	Me	Unidentified fungus (BCC 2629)	154
N	sBu	sBu	hy-sBu	iPr	iPr	iPr	Me	Me	Me	Unidentified fungus (BCC 2629)	154
O1	iPr	iPr	sBu	iBu	iPr	iPr	Me	Me	Me	*Verticillium hemipterigenum* BCC 1449	46
O2	iPr	iPr	sBu	iPr	iBu	iPr	Me	Me	Me	*Verticillium hemipterigenum* BCC 1449	46
O3	iPr	iPr	sBu	iPr	iPr	iBu	Me	Me	Me	*Verticillium hemipterigenum* BCC 1449	46
P1	iPr	iPr	iPr	hy-Et	iPr	iPr	Me	Me	Me	*Fusarium acuminatum* (*Gibberella acuminata*)	143
P2	iPr	iPr	iPr	hy-Et	iBu	iPr	Me	Me	Me	*Fusarium acuminatum* (*Gibberella acuminata*)	143

[a] Enniatin E is produced by the organism as a mixture of the two listed diastereomers that were not named separately.[24]

(Table 3). Enniatin A and B were first isolated in *Fusarium orthoceras* var. *enniatinum*, *F. scirpi* and two other *Fusarium* spp., by Plattner and coworkers[137] in a screening for new antibiotics. *Fusarium* sp. FO-1305 produces enniatin D, E and F, identified because of their strong acyl-CoA:cholesterol acyltransferase (ACAT) inhibition (Section 2.2).[24] The enniatin congeners A, A1, B, B1 and B2 display cytotoxic activity against two human cell lines (the hepatocellular carcinoma line Hep G2 and the fibroblast-like fetal lung cell line MRC-5).[33] With more powerful analytical methods, more natural enniatin analogs were discovered in a shorter time, including the recently described enniatins H, I and MK1886, which possess anti-HIV activity.[27]

The cylcooctadepsipeptides of the PF1022 family are produced by *Mycelia sterilia* (*Rosellinia* sp.), with PF1022A as the main metabolite.[61] The PF1022 congeners consist of four L-Leu residues, but differ in their D-hydroxycarboxylic acid content. Thus, PF1022A-D and F differ in the number of D-PheLac and D-Lac residues occupying the hydroxycarboxylic acid positions (Table 4). In contrast, PF1022E and PF1022-202 both contain two D-Lac, but with one or both D-PheLac positions replaced by *p*-hydroxy-D-PheLac.

Beauveria bassiana ARSEF 4122 produces beauvericin A and B, with (2*R*,3*S*)-2-hydroxy-3-methylpentanoate (D-2-hydroxy-3-methylvalerate, D-Hmv) residues replacing one or two D-Hiv residues. These analogs were evaluated in an insecticidal assay.[138] Beauvericins D (L-Phe replacing one *N*-methyl-L-Phe residue), E (L-Leu instead of one *N*-Me-L-Phe residue) and F [(2*R*)-2-hydroxy-4-methylpentanoate instead of one D-Hiv] were isolated from *Beauveria* sp. FKI-1366 and shown to display antifungal activity.[28,29]

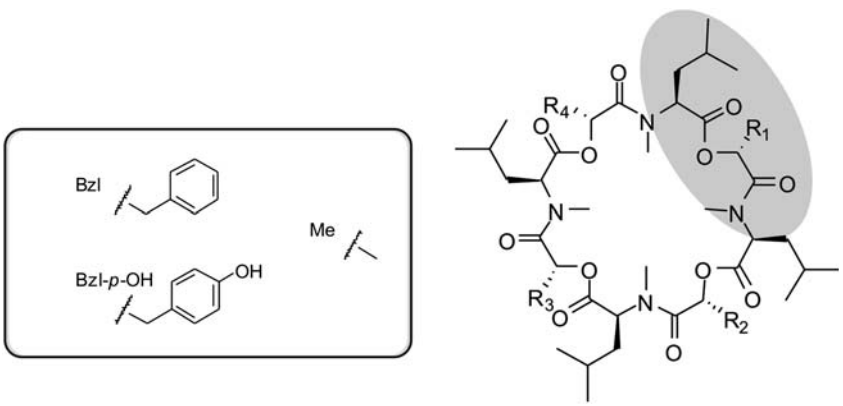
7.2 Unnatural CODs from precursor-directed biosynthesis

Precursor-directed biosynthesis[139] utilizes the power of chemical synthesis to generate precursor analogs that can be processed to complex "unnatural products" by living cells of the producer microorganisms. Precursor-directed biosynthesis relies on the substrate flexibility of native biosynthetic enzymes to recognize and accept the substrate analogs and to fully process the resulting modified intermediates. The substrate analogs are in direct competition with the native, intrinsic substrates of the biosynthetic enzymes. This method also presupposes the successful uptake and compartmentalization of the analogs, and has to contend with the potential toxicity of both the analogs themselves and the modified products.

Precursor-directed biosynthesis introduces added complexity during COD analog production. First, amino acid precursor analogs may incorporate directly to the amino acid positions in the COD, or may be converted to the corresponding D-2-hydroxycarboxylic acid and thus may also replace the hydroxycarboxylic acid constituents of the COD. Further, each precursor is used several times during the iterative assembly of the monomer units. Thus, incorporation of a precursor analog leads to the production of a COD analog family in which 1, 2 or all 3 (trimeric CODs) or 1, 2, 3 or all 4 (tetrameric CODs) of the positions for that substrate are replaced by the analog. Further, when two molecules of the same precursor analog incorporate into a tetrameric COD, two isomeric products are produced (one where the replacements took place in adjoining monomers, and another where the modified and the native monomers are alternating), due to the possibility of circular permutation.

Precursor-directed biosynthesis has been assessed by Zocher and coworkers using the enniatin producers *Fusarium scirpi* and *F. sambucinum*.[140] A small set of radioactively labeled hydroxycarboxylic acids (DL-2-hydroxy-*n*-valeric acid, D-2-hydroxy-3-methyl-*n*-valeric acid, DL-hydroxybutyric acid [DL-Hbu], and D-Lac) and L-amino acids (L-2-amino butyric acid [L-Abu], L-Ala, L-Cys, L-Thr, L-Ser and L-allylglycine) were separately fed in

Table 4 Naturally occurring PF1022 congeners

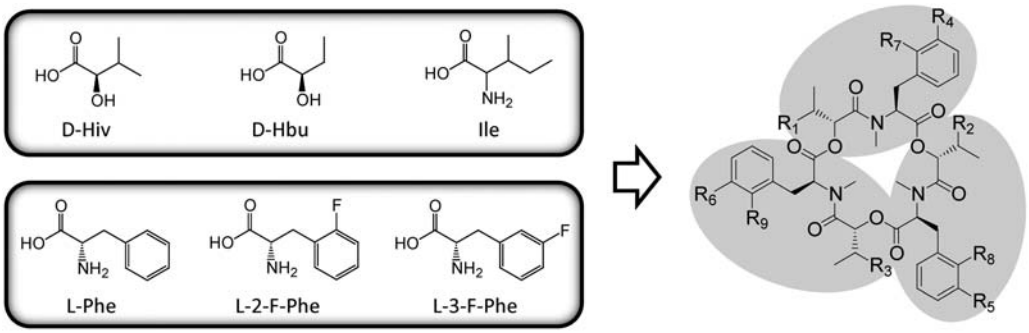


	R_1	R_2	R_3	R_4
PF 1022A	Me	Bzl	Me	Bzl
PF 1022B	Bzl	Bzl	Bzl	Bzl
PF 1022C	Me	Bzl	Bzl	Bzl
PF 1022D	Me	Me	Me	Bzl
PF 1022E	Me	Bzl-*p*-OH	Me	Bzl
PF 1022F	Me	Me	Me	Me
PF 1022-202	Me	Bzl-*p*-OH	Me	Bzl-*p*-OH

a single dose (10 mM, final concentration) to the cultures after 72 h of fermentation. The cultivation was continued for another two days. The formation of enniatin analogs was analyzed by HPLC, mass spectrometry and NMR spectroscopy. Amongst the hydroxycarboxylic acids, formation of product analogs could only be observed with D-Lac and DL-Hbu. In contrast, the amino acids L-Ala, L-Abu, L-Ser and L-Thr all yielded new enniatins. The "unnatural" all-D-Lac enniatin was found to have anthelminthic properties.[141] Subsequently, L-Ala, L-Thr and L-Abu-containing natural enniatins (enniatin J1-3,K, P1/2) have also been isolated from *Fusarium* sp. strain F31 and *Fusarium acuminatum* (*Gibberella acuminata*).[142,143]

Precursor-directed biosynthesis was also used to produce enniatin analogs, using the insect pathogenic fungus *Verticillium hemipterigenum* BCC 1449 as the producing organism.[144] This strain biosynthesizes enniatins B (trimer of D-Hiv–L-Val), B_4 (one L-Leu and two L-Val as the amino acid constituents), H (one D-Hmv and two D-Hiv as the hydroxycarboxylic acids), and I (two D-Hmv and one D-Hiv as the hydroxycarboxylic acid constituents, Table 3). Upon feeding L-Leu, the production of enniatin B4 was increased, and the fermentations also yielded enniatins G (two L-Leu and one L-Val as the amino acids) and minor amounts of enniatin C (three L-Leu as the amino acid constituents). Feeding L-Ile increased the production of enniatins H and I, and led to the production of the new enniatin analog MK1688 (trimer of D-Hmv–L-Val). Thus, L-Leu is readily accepted by the amino acid-activating A_2 domain of the *V. hemipterigenum* enniatin synthetase *in vivo*, but not used as a precursor for the hydroxycarboxylic acid positions in enniatin. Conversely, L-Ile serves as an alternative substrate for D-hydroxycarboxylic acid biosynthesis, and for the subsequent incorporation into enniatin by the hydroxycarboxylic acid-activating A_1 domain. However, L-Ile is apparently not utilized *in vivo* as an alternative amino acid precursor by the CODS. All the isolated enniatin analogs were evaluated for their antiplasmodial, antimycobacterial, and cancer cell cytotoxic activities.

Nilanonta *et al.*[15] have used the hypocrealean entomopathogen *Paecilomyces tenuipes* BCC 1614 to produce beauvericin analogs (Fig. 6). Feeding L-Ile (2*S*,3*S*) or D-*allo*-Ile (2*R*,3*S*) led to the production of beauvericins A, B and C with one, two or all three D-Hiv positions replaced by (2*R*,3*S*)-Hmv. Feeding D-Ile (2*R*,3*R*) or L-*allo*-Ile (2*S*,3*R*) provided *allo*-beauvericins A, B, and C, featuring one, two or three (2*R*,3*R*)-Hmv residues. These experiments are congruent with the conversion of all four Ile diastereomers to the 2-ketocarboxylic acid, and to the stereospecific reduction of this intermediate to the corresponding (2*R*)-hydroxycarboxylic acid, with retention of configuration at the 3-position. Both (2*R*,3*S*)- and (2*R*,3*R*)-Hmv are accepted as alternative substrates by the hydroxycarboxylic acid-activating A_1 domain of the *P. tenuipes* beauvericin synthetase *in vivo*. Conversely, none of the Ile diastereomers are acceptable *in vivo* substrates to replace L-Phe in beauvericin. The new beauvericin analogs showed similar antimycobacterial, antiplasmodial and cancer cell antiproliferative activities to that of beauvericin.



Feeding	Compound	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9
None	Beauvericin	Me	Me	Me	H	H	H	H	H	H
L-Ile of D-allo-Ile	Beauvericin A	(3*S*)-Et	Me	Me	H	H	H	H	H	H
	Beauvericin B	(3*S*)-Et	(3*S*)-Et	Me	H	H	H	H	H	H
	Beauvericin C	(3*S*)-Et	(3*S*)-Et	(3*S*)-Et	H	H	H	H	H	H
D-Ile or L-allo-Ile	*allo*-Beauvericin A	(3*R*)-Et	Me	Me	H	H	H	H	H	H
	allo-Beauvericin B	(3*R*)-Et	(3*R*)-Et	Me	H	H	H	H	H	H
	allo-Beauvericin C	(3*R*)-Et	(3*R*)-Et	(3*R*)-Et	H	H	H	H	H	H
D-Hbu	Beauvericin G_1	H	Me	Me	H	H	H	H	H	H
	Beauvericin G_2	H	H	Me	H	H	H	H	H	H
	Beauvericin G_3	H	H	H	H	H	H	H	H	H
L-3-F-Phe	Beauvericin H_1	Me	Me	Me	F	H	H	H	H	H
	Beauvericin H_2	Me	Me	Me	F	F	H	H	H	H
	Beauvericin H_3	Me	Me	Me	F	F	F	H	H	H
L-2-F-Phe	Beauvericin I_1	Me	Me	Me	H	H	H	F	H	H
	Beauvericin I_2	Me	Me	Me	H	H	H	F	F	H
	Beauvericin I_3	Me	Me	Me	H	H	H	F	F	F

Fig. 6 Precursor-directed biosynthesis of beauvericin analogs.

Precursor-directed biosynthesis was also applied to produce beauvericin analogs with *Beauveria bassiana* ATCC 7159 (Fig. 6), using 30 potential precursor analogs of D-Hiv and L-Phe.[145] Feeding L-Ile afforded beauvericins A, B and C: similar experiments yielded the same products in *Paecilomyces tenuipes*.[144] However, the BbBEAS beauvericin synthetase proved to be rather fastidious *in vivo*, with only a few other precursor analogs accepted. Thus, D-Hiv could be replaced only by (2*R*)-2-hydroxybutyric acid (D-Hbu) to yield beauvericins G_1, G_2 and G_3 featuring one, two or three D-Hbu moieties. As expected, the (2*S*)-isomer (L-Hbu) was not accepted by the system. L-Phe could only be substituted by 2-fluoro or 3-fluoro analogs of Phe to yield the beauvericin I_{1-3} and H_{1-3} series, respectively. Both the L and D isomers of these amino acid precursor analogs were readily utilized by the cells, but the synthetase itself seemed to be stereospecific, as the final beauvericin analogs contained amino acids with only the L configuration. This suggested that the substrates underwent epimerization in the cells prior to incorporation into the COD. The isolated novel beauvericin analogs were evaluated for cancer cell antiproliferative and cell motility inhibitory activities. These two bioactivities were affected to a different degree by the structural changes, suggesting that it might be possible to separately optimize cytotoxicity and haptotaxis inhibition in future beauvericin analogs.[145]

Precursor-directed biosynthesis of unnatural PF1022 derivatives in *Mycelia sterilia* (*Rosellinia* sp.) was only successful with *p*-nitro-PheLac and *p*-nitro-L-Phe (this latter precursor analog undergoes *in vivo* deamination and ketoreduction to *p*-nitro-PheLac). The feeding of 10–70 mM of these precursors yielded up to 40% of PF 1022-268 (the monosubstituted *p*-nitro-PheLac derivative) and up to 10% of the desired disubstituted *p*-nitro-PheLac derivative, PF 1022-220 (the yield of PF 1022A = 100%). PF 1022-220 constitutes a potentially useful intermediate that might significantly simplify the production process for emodepsid, an important semisynthetic anthelminthic agent.[70] However, the low yields of precursor-directed biosynthesis currently prohibit the scale-up of this process to industrial production (W. Weckwerth, *Doctoral Thesis*, TU Berlin, 1998, and M. Krause, *Doctoral Thesis*, TU Berlin, 1998).

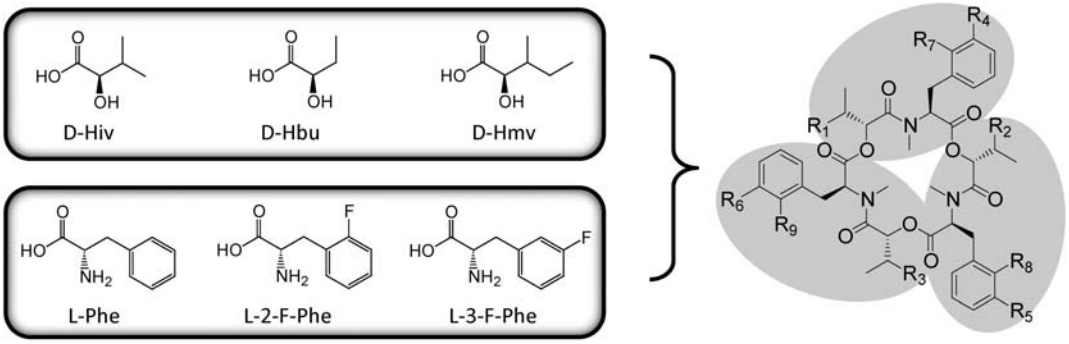
As these examples show, the success of precursor-directed biosynthesis experiments in CODS systems is currently not predictable. Substrate promiscuity does not derive merely from substrate recognition and activation by the A domains. Rather, successful production of unnatural CODs requires correct processing of the precursors and intermediates by the C domains and the subsequent modifying enzymes. The products should also be acceptable for the COD export system, and the unnatural CODs should not be overly toxic to the producer cells. More knowledge and expertise has to be gathered to make precursor-directed biosynthesis of CODs more predictable, reliable, and economical on the industrial scale.

7.3 Mutasynthesis and combinatorial mutasynthesis

Mutasynthesis is one of the most successful methods of combinatorial biosynthesis that has been applied to many natural product classes including polyketides, siderophores, nucleosides, aminocoumarins and nonribosomal peptides.[139,146–148] Mutasynthesis couples the power of chemical synthesis to generate structurally diverse analogs of biosynthetic precursors or intermediates (the so-called "mutasynthons") with the ability of genetic engineering to create appropriate mutants that are blocked in the biosynthesis or utilization of the natural (endogenous) precursors or intermediates. While mutasynthesis was first demonstrated 40 years ago,[149] recent developments in synthetic techniques, the availability of large commercial compound libraries, and advances in genetic and genomic techniques led to a renaissance of this approach in both the biotech industry and the academic community.[139,146–148] Mutasynthesis improves on precursor-directed biosynthesis by removing competition between the synthetic analog and the endogenous precursor, thus allowing the uncontested incorporation of the mutasynthon. It also simplifies the isolation of the resulting unnatural products by eliminating the biosynthesis of the native natural product. For CODs, this unchallenged incorporation means that instead of obtaining a COD analog family in which the native precursor postions are variably replaced by the precursor analog, the result being a single unnatural product bearing substitutions at all expected positions.

Xue *et al.* have used a *kivr* mutant of *Beauveria bassiana* ATCC 7159 for the mutasynthetic production of novel beauvericin analogs.[132] This strain lacks ketoisovalerate reductase, and thus it is unable to produce D-Hiv or similar branched-chain 2-hydroxycarboxylic acids as precursors for beauvericin biosynthesis, leading to a complete block in beauvericin (and bassianolide) biosynthesis. From five commercially available 2-hydroxycarboxylic acids (Hbu, DL-Lac, hydroxyisocaproic acid, mandelate, and cyclohexyllactate), only D-Hbu restored COD biosynthesis in *B. bassiana*, leading to the exclusive and high-titer production of beauvericin G_3. This analog has previously been detected during precursor-directed biosynthesis, but in substantially lower yields.[145] Feeding synthetic DL-2-hydroxy-3-methylvaleric acid (DL-Hmv) to the *kivr* mutant strain was found to support the exclusive production of the known analog beauvericin C in a good yield. This analog had been previously observed in small amounts in precursor-directed biosynthesis with wild-type beauvericin producer strains.[15,145]

To further increase the structural variety of beauvericin analogs, Xu *et al.* have conducted simultaneous feeding of precursor analogs in pairwise combinations using the *kivr* knockout *B. bassiana* strain, in a procedure dubbed "combinatorial mutasynthesis" (Fig. 7).[132] Such scrambling with two precursor analogs would not have been practical using the wild-type strain, as the presence of the competing native precursors and precursor analogs would have led to a very large number of possible combinations and circular permutations along the COD macrocycle. Such a complex product mixture would have been challenging to separate, and each analog might have been present only in minor amounts. In contrast, combinatorial mutasynthesis significantly simplified product profiles: the fed hydroxycarboxylic acids completely substituted the D-Hiv positions of beauvericin, while the Phe analogs replaced 0, 1, 2, or all 3 Phe in the products. To demonstrate this principle, the D-Hiv analogs D-Hbu and DL-Hmv, and the Phe analogs 3-fluoro-L-Phe and 2-fluoro-L-Phe, whose acceptability to this strain had already been shown,[132,145] were used for this study. Combinatorial mutasynthesis with these four precursor analogs yielded 14 new beauvericin analogs belonging to five novel series. Importantly, in



Feeding	Compound	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9
D-Hiv	Beauvericin	Me	Me	Me	H	H	H	H	H	H
D-Hbu	Beauvericin G_3	H	H	H	H	H	H	H	H	H
D-Hmv	Beauvericin C	Et	Et	Et	H	H	H	H	H	H
D-Hbu + L-3-F-Phe	Beauvericin G_3H_1	H	H	H	F	H	H	H	H	H
	Beauvericin G_3H_2	H	H	H	F	F	H	H	H	H
	Beauvericin G_3H_3	H	H	H	F	F	F	H	H	H
D-Hbu + L-2-F-Phe	Beauvericin G_3I_1	H	H	H	H	H	H	F	H	H
	Beauvericin G_3I_2	H	H	H	H	H	H	F	F	H
	Beauvericin G_3I_3	H	H	H	H	H	H	F	F	F
D-Hmv + L-3-F-Phe	Beauvericin CH_1	Et	Et	Et	F	H	H	H	H	H
	Beauvericin CH_2	Et	Et	Et	F	F	H	H	H	H
	Beauvericin CH_3	Et	Et	Et	F	F	F	H	H	H
D-Hmv + L-2-F-Phe	Beauvericin CI_1	Et	Et	Et	H	H	H	F	H	H
	Beauvericin CI_2	Et	Et	Et	H	H	H	F	F	H
	Beauvericin CI_3	Et	Et	Et	H	H	H	F	F	F
D-Hmv + D-Hbu	Beauvericin BG_1	Et	Et	H	H	H	H	H	H	H
	Beauvericin AG_2	Et	H	H	H	H	H	H	H	H

Fig. 7 Mutasynthesis and combinatorial mutasynthesis of beauvericin analogs.

several product analogs, all hydroxycarboxylic acid and amino acid positions of beauvericin were completely replaced by the fed mutasynthons. Nine isolated analogs and three beauvericin analog mixtures that could not be separated under standard conditions were evaluated for cancer cell antiproliferative and cell motility inhibitory activities. As before, variation of the two activities due to the structure changes was not strictly parallel, indicating that more drastic structural alterations of the beauvericin scaffold may help to disconnect the antiproliferative and anti-haptotactic activities.

7.4 *In vitro* biosynthesis ("total biosynthesis") of COD analogs using purified enzymes

The *in vitro* reconstitution of enniatin biosynthesis was first achieved with purified ESYN from *Fusarium oxysporum*, using the natural substrates.[79] In an extended approach, Zocher and coworkers used additional amino acid substrate analogs for the incorporation of amino acids L-Ala, L-Cys, L-Thr and L-Ser into the enniatin structure (Fig. 8A).[140] Among the hydroxycarboxylic acid substrates, D-Hiv could be replaced by D-Hbu and D-Lac.[140]

To learn more about the substrate specificity of the enniatin CODS and to create a bigger library of new derivatives by chemoenzymatic synthesis or total biosynthesis, various hydroxycarboxylic acids with linear, branched and cyclic side chains – up to seven carbon atoms and various functional groups, *e.g.* halogens, hydroxy and thioether groups – were chemically synthesized and tested in an *in vitro* assay with the purified ESYN from *F. oxysporum*. Surprisingly, some of the hydroxycarboxylic acid substrates proved to be as good substrates as D-Hiv. Thus, D-Hiv could be efficiently replaced by D-chlorolactate, D-bromolactate, D-propargyl lactate, and D-Hbu, whereas the extension of the aliphatic side chain decreased product yield (Fig. 8A). From these findings, the binding pocket is proposed to accommodate alkyl chain residues with a minimum of two carbon atoms (D-Hbu), but with a maximum of three carbons in linear and four carbons in branched chains. No substrate activation was found for polar, ionic or aromatic hydroxycarboxylic acid side chains.[89]

Similar to that with the enniatin synthetase FoESYN, the purified RsPFSYN was also used to perform *in vitro* total biosynthesis. The naturally found product distributions in fermentations could be reproduced *in vitro* with the natural substrates.[61] Since the observed substrate spectrum of RsPFSYN includes two sterically and electronically very different hydroxycarboxylic acids, D-Lac and D-PheLac, a certain substrate promiscuity was expected from this enzyme. Various hydroxycarboxylic acids were synthesized and tested with RsPFSYN for the enzymatic assembly of PF1022 analogs, with the range of synthetic substrates extended to >40 aromatic and aliphatic hydroxycarboxylic acids. The results showed that a large variety of aliphatic and aromatic hydroxycarboxylic acids

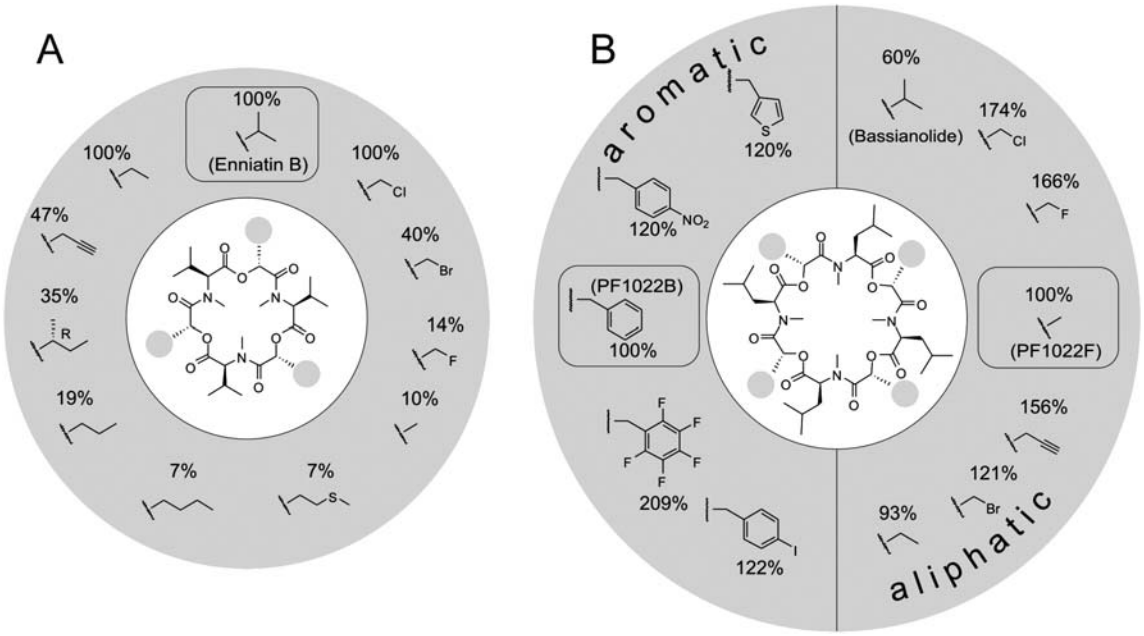


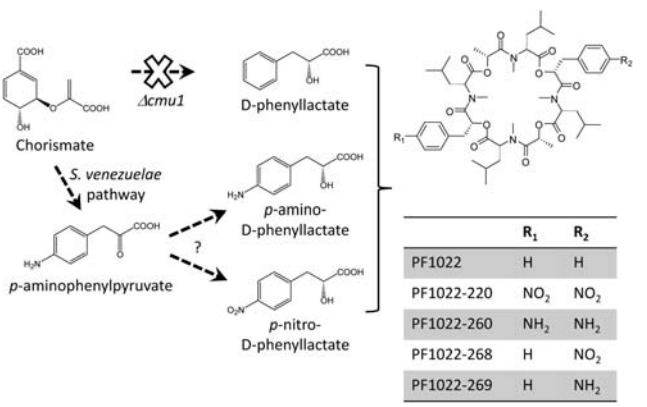
Fig. 8 Total biosynthesis of COD analogs. Incorporation of different synthetic 2-hydroxycarboxylic acids (only the side chains shown), replacing the side chains (represented as small circles) in **A.** enniatin, and **B.** PF 1022. The percentages describe the enzyme activity in $k_{cat,app}$ in comparison to the natural substrate (enniatin: D-Hiv; PF1022: D-PheLac for the aromatic and D-Lac for the aliphatic precursors, respectively).

were indeed acceptable as substrates (Fig. 8B). A strong correlation has been observed between the substrate tolerances of FoESYN and RsPFSYN towards aliphatic hydroxycarboxylic acids. Among the aromatic PF1022 derivatives obtained, the most interesting ones contained heterocycles (*e.g.* thiophene), or functionalized phenyl rings (*e.g.* a perfluorinated analog). While no clear rules for the substrate specificity of RsPFSYN could be deduced, the substrate tolerance seemed to increase for aromatic residues with lesser steric demand, or for precursors with decreased rotational freedom at the β-position.[90]

Overall, a number of truly unnatural enniatins and PF1022 derivatives were generated using *in vitro* biosynthesis. Such derivatives could facilitate further semi-synthetic modification, for example by Sonogashira coupling at the alkyne functionalities.[150] Remarkably, RsPFSYN has an extended substrate spectrum towards aromatic residues compared to FoESYN, but neither of these CODSs is able to accept hydroxycarboxylic acids with polar or charged side chains. Nevertheless, both enzymes display a potential for the generation of large COD libraries, at least *in vitro*. A significant disadvantage of this chemoenzymatic approach is that only small amounts of CODs are obtained in routine experiments, and scale-up of the reactions is difficult. However, as opposed to precursor-directed biosynthesis and mutasynthesis, *in vitro* biosynthesis is not limited by substrate uptake or catabolism by the cell, nor by precursor or product toxicity issues.

7.5 Combinatorial biosynthesis by precursor supply pathway engineering

Modulation of precursor availability has been shown to alter the yields of natural congeners of CODs and to support the biosynthesis of novel analogs during precursor-directed biosynthesis, mutasynthesis, and total biosynthesis, as described above. Modulation of the availability of intrinsic precursor pools during *in vivo* biosynthesis of CODs by metabolic engineering, and the resulting direct biosynthetic production of COD analogs, has only been reported in the case of PF1022s (Fig. 9).[151] Thus, a chorismate mutase-deficient strain of the PF1022 producer *Rosellinia* sp. has been transformed with three genes from *Streptomyces venezuelae* that allow the biosynthesis of *p*-aminophenylpyruvate from chorismate. Knockout of the *cmu1* chorismate mutase gene disrupted the biosynthesis of the PF1022 precursor D-PheLac (and also that of L-Phe) from chorismate. However, this mutation did not completely eliminate D-PheLac biosynthesis: deamination of medium-derived Phe to phenylpyruvate, followed by dehydrogenation, could still supply some of this precursor. Consequently, the *cmu1* knockout strain still produced PF1022A, albeit at a significantly reduced level. The



	R_1	R_2
PF1022	H	H
PF1022-220	NO_2	NO_2
PF1022-260	NH_2	NH_2
PF1022-268	H	NO_2
PF1022-269	H	NH_2

Fig. 9 Biosynthesis of PF1022 analogs by precursor supply pathway engineering.

cmu1 mutation nevertheless allowed the more efficient channeling of chorismate towards *p*-aminophenylpyruvate by the introduced heterologous pathway. The resulting *p*-aminophenylpyruvate was apparently converted to *p*-amino-D-PheLac by an uncharacterized endogenous enzyme of *Rosellinia* sp.: this analog of D-PheLac has been shown to be acceptable for the RsPFSYN PF1022 synthetase.[61,90] The engineered strain was shown to produce PF1022-269 and PF1022-260 (one or both D-PheLac replaced by *p*-amino-D-PheLac, respectively). Surprisingly, the strain also produced PF1022-268 and PF1022-220 (one or both D-PheLac replaced by *p*-nitro-D-PheLac), due to an adventitious oxidation of the amino group to a nitro functionality. Although the yield of the PF1022 analogs was low, the engineered strain secreted substantial amounts of the aromatic hydroxycarboxylic acids into the medium, indicating that the heterologous precursor biosynthetic pathway was relatively efficient in producing the modified substrate. Direct biosynthetic production of PF1022 analogs with *para*-position-specific modifications of the benzene ring may allow a large range of further semi-synthetic chemical modifications. It might also replace a conventional synthetic scheme for the nitration of PF1022A that has a low specificity, uses toxic reagents, and is relatively costly.[151]

8 Conclusions

Cyclooligomer depsipeptide natural products are produced by fungi of the Hypocreomycetidae and the Xylariomycetidae, both in the class Sordariomycetes. These compounds function as mycotoxins and potential virulence factors in their native context, but may be harnessed as antibiotics, insecticides, anthelminthics, herbicides, antitumor agents, and chemosensitizers by the pharmaceutical and the agribusiness industries. They may also serve as cholesterol biosynthesis inhibitors, repress amyloid plaque formation, or arrest the spread of tumors by inhibiting directional cell migration. Fungal COD biosynthesis involves cyclooligomer depsipeptide synthetases (CODSs), Type B NRPS enzymes that conduct the programmed iterative assembly of oligopeptidol monomer units. These CODS also catalyze the recursive ligation and cyclization of the monomers in a concerted cyclooligomerization process that involves the formation of intra- and intermolecular ester bonds. Fungal COD biosynthesis has been studied for over 30 years, starting with the isolation of active CODS enzymes from the producer fungi, and *in vitro* reconstitution of COD synthesis using appropriate substrates. Isolation of the corresponding synthetase genes allowed heterologous expression of CODS and/or their enzymatically active fragments for further biochemical and biosynthetic studies, and for the production of the CODs themselves in a prokaryotic host in a heterologous biocatalytic process. The isolated CODS genes have revealed interesting mechanistic differences between fungal and bacterial CODSs, indicating that the biosynthesis of CODs has a polyphyletic origin. A variety of combinatorial biosynthetic methods, including precursor-directed biosynthesis, mutasynthesis, combinatorial mutasynthesis, and total biosynthesis, as well as genome mining for putative CODS-encoding genes in sequenced fungal genomes, promises to significantly extend the diversity of this interesting natural product family in the near future.

9 Abbreviations

A	adenylation
AA	L-amino acid
ABC	ATP binding cassette
Abu	2-aminobutyric acid
ACAT	acyl-CoA:cholesterol acyltransferase
AMP	adenosine monophosphate
ATP	adenosine triphosphate
B.	*Beauveria*
BEAS	beauvericin synthetase
BSLS	bassianolide synthetase
C	condensation
CoA	coenzyme A
COD	cyclooligomer depsipeptide
CODS	cyclooligomer depsipeptide synthetase
Cy	cyclization
E	epimerization
ESYN	enniatin synthetase
F.	*Fusarium*
G.	*Gibberella*
H.	*Hirsutella*
HA	D-hydroxycarboxylic acid
Hbu	2-hydroxybutyric acid
Hiv	2-hydroxyisovaleric acid
HIV	human immunodeficiency virus
Hmv	2-hydroxy-3-methylvaleric acid
HPLC	high-performance liquid chromatography
I.	*Isaria*
Ile	isoleucine
KIVR	ketoisovalerate reductase
KR	ketoreductase
Lac	lactic acid
Leu	leucine
M	methylation
mAb	monoclonal antibody
MDR	multidrug resistance
NADP	nicotinamide adenine dinucleotide phosphate
NRPS	nonribosomal peptide synthetase
Ox	oxidation
P.	*Paecilomyces*
PFSYN	PF1022 synthetase
Pgp	P-glycoprotein
Phe	phenylalanine
PheLac	phenyllactic acid
PP_i	pyrophosphate
Pro	proline
R	reductase
S.	*Saccharomyces*
SAH	*S*-adenosylhomocysteine
SAM	*S*-adenosyl-L-methionine
T.	*Trichoderma*
T	thiolation
TE	thioesterase
TLC	thin-layer chromatography
V.	*Verticillium*
Val	valine

10 References

1 A. Koglin and C. T. Walsh, *Nat. Prod. Rep.*, 2009, **26**, 987–1000.
2 M. A. Fischbach and C. T. Walsh, *Chem. Rev.*, 2006, **106**, 3468–3496.
3 H. von Döhren, *Adv. Biochem. Engin./Biotechnol.*, 2004, **88**, 217–264.
4 J. L. Meier and M. D. Burkart, *Chem. Soc. Rev.*, 2009, **38**, 2012–2045.
5 S. C. Wenzel, B. Kunze, G. Höfle, B. Silakowski, M. Scharfe, H. Blocker and R. Müller, *ChemBioChem*, 2005, **6**, 375–385.
6 S. C. Wenzel, P. Meiser, T. M. Binz, T. Mahmud and R. Müller, *Angew. Chem., Int. Ed.*, 2006, **45**, 2296–2301.
7 H. D. Mootz, D. Schwarzer and M. A. Marahiel, *ChemBioChem*, 2002, **3**, 490–504.
8 L. Du and L. Lou, *Nat. Prod. Rep.*, 2010, **27**, 255–278.
9 S. A. Samel, M. A. Marahiel and L. O. Essen, *Mol. BioSyst.*, 2008, **4**, 387–393.
10 V. Dartois, J. Sanchez-Quesada, E. Cabezas, E. Chi, C. Dubbelde, C. Dunn, J. Granja, C. Gritzen, D. Weinberger, M. R. Ghadiri and T. R. Parr, Jr., *Antimicrob. Agents Chemother.*, 2005, **49**, 3302–3310.
11 S. Fernandez-Lopez, H.-S. Kim, E. C. Choi, M. Delgado, J. R. Granja, A. Khasanov, K. Kraehenbuehl, G. Long, D. A. Weinberger, K. M. Wilcoxen and M. R. Ghadiri, *Nature*, 2001, **412**, 452–456.
12 D. W. Russell, *J. Chem. Soc.*, 1965, 4664–4668.
13 D. W. Russell, *J. Chem. Soc.*, 1962, 753–761.
14 A. Moretti, G. Mule, A. Ritieni and A. Logrieco, *Int. J. Food Microbiol.*, 2007, **118**, 158–163.
15 C. Nilanonta, M. Isaka, P. Kittakoop, S. Trakulnaleamsai, M. Tanticharoen and Y. Thebtaranonth, *Tetrahedron*, 2002, **58**, 3355–3360.
16 S. Gupta, S. B. Krasnoff, N. L. Underwood, J. A. Renwick and D. W. Roberts, *Mycopathologia*, 1991, **115**, 185–189.
17 I. Molnár, S. B. Krasnoff and D. M. Gibson, *Nat. Prod. Rep.*, 2010, **27**, 1241–1275.
18 J. J. Luangsa-Ard, P. Berkaew, R. Ridkaew, N. L. Hywel-Jones and M. Isaka, *Mycol. Res.*, 2009, **113**, 1389–1395.
19 H.-H. Song, J.-H. Ahn, Y. H. Lim and C. Lee, *J. Microbiol. Biotechnol.*, 2006, **16**, 1111–1119.
20 L. K. Steinrauf, *Metal Ions Biol. Systems*, 1985, **19**, 139–171.
21 R. L. Hamill, C. E. Higgens, M. E. Boaz and M. Gorman, *Tetrahedron Lett.*, 1969, **10**, 4255–4258.
22 Y. Xu, R. Orozco, E. M. K. Wijeratne, A. A. L. Gunatilaka, S. P. Stock and I. Molnár, *Chem. Biol.*, 2008, **15**, 898–907.
23 A. Miyazaki, T. Kanome and T. Watanabe, *Curr. Drug Targets: Cardiovasc. & Haematol. Disord.*, 2005, **5**, 463–469.
24 H. Tomoda, X. H. Huang, J. Cao, H. Nishida, R. Nagao, S. Okuda, H. Tanaka, S. Omura, H. Arai and K. Inoue, *J. Antibiot.*, 1992, **45**, 1626–1632.
25 L. Puglielli, G. Konopka, E. Pack-Chung, L. A. Ingano, O. Berezovska, B. T. Hyman, T. Y. Chang, R. E. Tanzi and D. M. Kovacs, *Nat. Cell Biol.*, 2001, **3**, 905–912.
26 H. J. Huttunen, C. Greco and D. M. Kovacs, *FEBS Lett.*, 2007, **581**, 1688–1692.
27 C.-G. Shin, D.-G. An, H.-H. Song and C. Lee, *J. Antibiot.*, 2009, **62**, 687–690.
28 T. Fukuda, M. Arai, Y. Yamaguchi, R. Masuma, H. Tomoda and S. Omura, *J. Antibiot.*, 2004, **57**, 110–116.
29 T. Fukuda, M. Arai, H. Tomoda and S. Omura, *J. Antibiot.*, 2004, **57**, 117–124.
30 L. Zhang, K. Yan, Y. Zhang, R. Huang, J. Bian, C. Zheng, H. Sun, Z. Chen, N. Sun, R. An, F. Min, W. Zhao, Y. Zhuo, J. You, Y. Song, Z. Yu, Z. Liu, K. Yang, H. Gao, H. Dai, X. Zhang, J. Wang, C. Fu, G. Pei, J. Liu, S. Zhang, M. Goodfellow, Y. Jiang, J. Kuai, G. Zhou and X. Chen, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 4606–4611.
31 O. Lomovskaya and K. A. Bostian, *Biochem. Pharmacol.*, 2006, **71**, 910–918.
32 F. J. Sharom, P. Lu, R. Liu and X. Yu, *Biochem. J.*, 1998, **333**, 621–630.
33 L. Ivanova, E. Skjerve, G. S. Eriksen and S. Uhlig, *Toxicon*, 2006, **47**, 868–876.
34 G.-M. Jow, C.-J. Chou, B.-F. Chen and J.-H. Tsai, *Cancer Lett.*, 2004, **216**, 165–173.
35 B. F. Chen, M. C. Tsai and G. M. Jow, *Biochem. Biophys. Res. Commun.*, 2006, **340**, 134–139.
36 J. Zhan, A. M. Burns, M. X. Liu, S. H. Faeth and A. A. L. Gunatilaka, *J. Nat. Prod.*, 2007, **70**, 227–232.
37 P. Carmeliet, *Nat. Rev. Genet.*, 2003, **4**, 710–720.
38 D. Shan, L. Chen, J. T. Njardarson, C. Gaul, X. Ma, S. J. Danishefsky and X. Y. Huang, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 3772–3776.
39 D. Belotti, V. Vergani, T. Drudis, P. Borsotti, M. R. Pitelli, G. Viale, R. Giavazzi and G. Taraboletti, *Clin. Cancer Res.*, 1996, **2**, 1843–1849.
40 R. S. Herbst, *Expert Opin. Emerging Drugs*, 2006, **11**, 635–650.
41 G. Taraboletti, G. Micheletti, M. Rieppi, M. Poli, M. Turatto, C. Rossi, P. Borsotti, P. Roccabianca, E. Scanziani, M. I. Nicoletti, E. Bombardelli, P. Morazzoni, A. Riva and R. Giavazzi, *Clin. Cancer Res.*, 2002, **8**, 1182–1188.
42 A. Kruczynski, M. Poli, R. Dossi, E. Chazottes, G. Berrichon, C. Ricome, R. Giavazzi, B. T. Hill and G. Taraboletti, *Eur. J. Cancer*, 2006, **42**, 2821–2832.
43 S. Firakova, B. Proksa and M. Sturdikova, *Pharmazie*, 2007, **62**, 563–568.
44 A. E. Desjardins and R. H. Proctor, *Int. J. Food Microbiol.*, 2007, **119**, 47–50.
45 C. Nilanonta, M. Isaka, P. Kittakoop, J. Saenboonrueng, V. Rukachaisirikul, P. Kongsaeree and Y. Thebtaranonth, *J. Antibiot.*, 2003, **56**, 647–651.
46 S. Supothina, M. Isaka, K. Kirtikara, M. Tanticharoen and Y. Thebtaranonth, *J. Antibiot.*, 2004, **57**, 732–738.
47 Y. Lin, J. Wang, X. Wu, S. Zhou, L. L. P. Vrijmoed and E. B. G. Jones, *Aust. J. Chem.*, 2002, **55**, 225–227.
48 M. Kamyar, P. Rawnduzi, C. R. Studenik, K. Kouri and R. Lemmens-Gruber, *Arch. Biochem. Biophys.*, 2004, **429**, 215–223.
49 J. F. Grove and M. Pople, *Mycopathologia*, 1980, **70**, 103–105.
50 R. Dornetshuber, P. Heffeter, M. Sulyok, R. Schumacher, P. Chiba, S. Kopp, G. Koellensperger, M. Micksche, R. Lemmens-Gruber and W. Berger, *Mol. Nutr. Food Res.*, 2009, **53**, 904–920.
51 K. Hiraga, S. Yamamoto, H. Fukuda, N. Hamanaka and K. Oda, *Biochem. Biophys. Res. Commun.*, 2005, **328**, 1119–1125.
52 D. Levy, A. Bluzat, M. Seigneuret and J. L. Rigaud, *Biochem. Pharmacol.*, 1995, **50**, 2105–2107.
53 M. Herrmann, R. Zocher and A. Haese, *Mol. Plant Microbe Interact.*, 1996, **9**, 226–232.
54 N. Vongvanich, P. Kittakoop, M. Isaka, S. Trakulnaleamsai, S. Vimuttipong, M. Tanticharoen and Y. Thebtaranonth, *J. Nat. Prod.*, 2002, **65**, 1346–1348.
55 J. Jirakkakul, J. Punya, S. Pongpattanakitshote, P. Paungmoung, N. Vorapreeda, A. Tachaleat, C. Klomnara, M. Tanticharoen and S. Cheevadhanarak, *Microbiology*, 2008, **154**, 995–1006.
56 M. Kanaoka, A. Isogai, S. Murakoshi, M. Ichinoe, A. Suzuki and S. Tamura, *Agric. Biol. Chem.*, 1978, **42**, 629–635.
57 Y. Xu, R. Orozco, E. M. K. Wijeratne, P. Espinosa-Artiles, A. A. L. Gunatilaka, S. P. Stock and I. Molnár, *Fungal Genet. Biol.*, 2009, **46**, 353–364.
58 S. Nakajyo, K. Shimizu, A. Kometani, A. Suzuki, H. Ozaki and N. Urakawa, *Jpn. J. Pharmacol.*, 1983, **33**, 573–582.
59 T. Sasaki, M. Takagi, T. Yaguchi, S. Miyadoh, T. Okada and M. Koyama, *J. Antibiot.*, 1992, **45**, 692–697.
60 S. Miyadoh, H. Kawasaki, K. Aoyagi, T. Yaguchi, T. Okada and J. Sugiyama, *Nippon Kingakkai Kaiho*, 2000, **41**, 183–188.
61 W. Weckwerth, K. Miyamoto, K. Iinuma, M. Krause, M. Glinski, T. Storm, G. Bonse, H. Kleinkauf and R. Zocher, *J. Biol. Chem.*, 2000, **275**, 17909–17915.
62 M. Guest, K. Bull, R. J. Walker, K. Amliwala, V. O'Connor, A. Harder, L. Holden-Dye and N. A. Hopper, *Int. J. Parasitol.*, 2007, **37**, 1577–1588.
63 A. Harder, H. P. Schmitt-Wrede, J. Krucken, P. Marinovski, F. Wunderlich, J. Willson, K. Amliwala, L. Holden-Dye and R. Walker, *Int. J. Antimicrob. Agents*, 2003, **22**, 318–331.
64 J. Scherkenbeck, A. Harder, A. Plant and H. Dyker, *Bioorg. Med. Chem. Lett.*, 1998, **8**, 1035–1040.
65 P. Jeschke, A. Harder, W. Etzel, W. Gau, G. Thielking, G. Bonse and K. Iinuma, *Pest Manage. Sci.*, 2001, **57**, 1000–1006.
66 S. Banala and R. D. Süssmuth, *ChemBioChem*, 2010, **11**, 1335–1337.

67 H. Dyker, A. Harder and J. Scherkenbeck, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 6129–6132.
68 G. von Samson-Himmelstjerna, A. Harder, N. C. Sangster and G. C. Coles, *Parasitology*, 2005, **130**, 343–347.
69 J. Scherkenbeck, P. Jeschke and A. Harder, *Curr. Top. Med. Chem.*, 2002, **2**, 759–777.
70 A. Harder and G. von Samson-Himmelstjerna, *Parasitol. Res.*, 2001, **87**, 924–928.
71 S. Monma, T. Sunazuka, K. Nagai, T. Arai, K. Shiomi, R. Matsui and S. Omura, *Org. Lett.*, 2006, **8**, 5601–5604.
72 T. Kagamizono, E. Nishino, K. Matsumoto, A. Kawashima, M. Kishimoto, N. Sakai, B. M. He, Z. X. Chen, T. Adachi, S. Morimoto and K. Hanada, *J. Antibiot.*, 1995, **48**, 1407–1412.
73 H. Oh, T. Kim, G.-S. Oh, H.-O. Pae, K.-H. Hong, K.-Y. Chai, T.-O. Kwon, H.-T. Chung and H.-S. Lee, *Planta Med.*, 2002, **68**, 345–348.
74 K. Hasumi, C. Shinohara, T. Iwanaga and A. Endo, *J. Antibiot.*, 1993, **46**, 1782–1787.
75 Z. Du, C. Song, B. Yu and X. Luo, *Zhongguo Yaowu Huaxue Zazhi*, 2008, **18**, 452–456.
76 R. K. Pettit, G. R. Pettit, J.-P. Xu, C. A. Weber and L. A. Richert, *Planta Med.*, 2010, **76**, 500–501.
77 O. Froshov, T. L. Zimmer and S. G. Laland, *FEBS Lett.*, 1970, **7**, 68–71.
78 H. Peeters, R. Zocher, N. Madry, P. B. Oelrichs, H. Kleinkauf and G. Kraepelin, *J. Antibiot.*, 1983, **36**, 1762–1766.
79 R. Zocher, U. Keller and H. Kleinkauf, *Biochemistry*, 1982, **21**, 43–48.
80 R. Zocher, T. Nihira, E. Paul, N. Madry, H. Peeters, H. Kleinkauf and U. Keller, *Biochemistry*, 1986, **25**, 550–553.
81 R. Pieper, H. Kleinkauf and R. Zocher, *J. Antibiot.*, 1992, **45**, 1273–1277.
82 H. Peeters, R. Zocher and H. Kleinkauf, *J. Antibiot.*, 1988, **41**, 352–359.
83 N. Siegbahn, K. Mosbach, K. Grodzki, R. Zocher, N. Madry and H. Kleinkauf, *Biotechnol. Lett.*, 1985, **7**, 297–302.
84 R. Zocher, U. Keller and H. Kleinkauf, *Biochem. Biophys. Res. Commun.*, 1983, **110**, 292–299.
85 W. Kallow, H. von Döhren and H. Kleinkauf, *Biochemistry*, 1998, **37**, 5947–5952.
86 A. Billich and R. Zocher, *Biochemistry*, 1987, **26**, 8417–8423.
87 G. L. Cantoni, H. H. Richards and P. K. Chiang, in *Transmethylation*, eds. E. Usdin, R. T. Borchardt and C. R. Creveling, Elsevier/North-Holland, New York, 1979, pp. 155–164.
88 T. K. Audhya and D. W. Russell, *Anal. Lett.*, 1973, **6**, 265–274.
89 S. C. Feifel, T. Schmiederer, T. Hornbogen, H. Berg, R. D. Süssmuth and R. Zocher, *ChemBioChem*, 2007, **8**, 1767–1770.
90 J. Müller, S. C. Feifel, T. Schmiederer, R. Zocher and R. D. Süssmuth, *ChemBioChem*, 2009, **10**, 323–328.
91 A. Haese, M. Schubert, M. Herrmann and R. Zocher, *Mol. Microbiol.*, 1993, **7**, 905–194.
92 M. A. Marahiel, *J. Pept. Sci.*, 2009, **15**, 799–807.
93 N. Mido, K. Okakura, K. Miyamoto, M. Watanabe, K. Yanai, T. Yasutake, S. Aihara, T. Futamura, H. Kleinkauf and T. Murakami, *World Pat.* WO 0118179-A1, 2001.
94 R. Jeschke, K. Iinuma, A. Harder, M. Schindler and T. Murakami, *Parasitol. Res.*, 2005, **97**(Suppl. 1), S11–16.
95 R. M. Berka, M. W. Rey and W. T. Yoder, *US Pat.* US 6,893,839 B1, 2005.
96 A. Haese, R. Pieper, T. von Ostrowski and R. Zocher, *J. Mol. Biol.*, 1994, **243**, 116–122.
97 R. Pieper, A. Haese, W. Schröder and R. Zocher, *Eur. J. Biochem.*, 1995, **230**, 119–126.
98 N. A. Magarvey, Z. Q. Beck, T. Golakoti, Y. Ding, U. Huber, T. K. Hemscheidt, D. Abelson, R. E. Moore and D. H. Sherman, *ACS Chem. Biol.*, 2006, **1**, 766–779.
99 D. Schwarzer, R. Finking and M. A. Marahiel, *Nat. Prod. Rep.*, 2003, **20**, 275–287.
100 T. Hornbogen, S. P. Riechers, B. Prinz, J. Schultchen, C. Lang, S. Schmidt, C. Mügge, S. Turkanovic, R. D. Süssmuth, E. Tauberger and R. Zocher, *ChemBioChem*, 2007, **8**, 1048–1054.
101 C. Hacker, M. Glinski, T. Hornbogen, A. Doller and R. Zocher, *J. Biol. Chem.*, 2000, **275**, 30826–30832.
102 C. Rausch, I. Hoof, T. Weber, W. Wohlleben and D. H. Huson, *BMC Evol. Biol.*, 2007, **7**, 78.
103 A. Billich, R. Zocher, H. Kleinkauf, D. G. Braun, D. Lavanchy and H. K. Hochkeppel, *Biol. Chem. Hoppe-Seyler*, 1987, **368**, 521–529.
104 S. G. Lee and F. Lipmann, *Methods Enzymol.*, 1975, **43**, 585–602.
105 T. Stachelhaus, H. D. Mootz and M. A. Marahiel, *Chem. Biol.*, 1999, **6**, 493–505.
106 G. L. Challis, J. Ravel and C. A. Townsend, *Chem. Biol.*, 2000, **7**, 211–224.
107 C. Rausch, T. Weber, O. Kohlbacher, W. Wohlleben and D. H. Huson, *Nucleic Acids Res.*, 2005, **33**, 5799–5808.
108 D. G. Fujimori, S. Hrvatin, C. S. Neumann, M. Strieker, M. A. Marahiel and C. T. Walsh, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 16498–16503.
109 N. A. Magarvey, M. Ehling-Schulz and C. T. Walsh, *J. Am. Chem. Soc.*, 2006, **128**, 10698–10699.
110 M. Glinski, C. Urbanke, T. Hornbogen and R. Zocher, *Arch. Microbiol.*, 2002, **178**, 267–273.
111 M. Z. Ansari, J. Sharma, R. S. Gokhale and D. Mohanty, *BMC Bioinformatics*, 2008, **9**, 454.
112 T. Hornbogen and R. Zocher, in *Handbook of Industrial Mycology*, ed. Z. An, Marcel Dekker, New York, 2005, pp. 449–477.
113 V. Bergendahl, U. Linne and M. A. Marahiel, *Eur. J. Biochem.*, 2002, **269**, 620–629.
114 J. L. Slightom, B. P. Metzger, H. T. Luu and A. P. Elhammer, *Gene*, 2009, **431**, 67–79.
115 G. Weber, K. Schorgendorfer, E. Schneider-Scherzer and E. Leitner, *Curr. Genet.*, 1994, **26**, 120–125.
116 H. Jiang, R. Zirkle, J. G. Metz, L. Braun, L. Richter, S. G. Van Lanen and B. Shen, *J. Am. Chem. Soc.*, 2008, **130**, 6336–6337.
117 A. S. Rahman, J. Hothersall, J. Crosby, T. J. Simpson and C. M. Thomas, *J. Biol. Chem.*, 2005, **280**, 6399–6408.
118 P. C. Dorrestein and N. L. Kelleher, *Nat. Prod. Rep.*, 2006, **23**, 893–918.
119 S. B. Bumpus and N. L. Kelleher, *Curr. Opin. Chem. Biol.*, 2008, **12**, 475–482.
120 G. Yadav, R. S. Gokhale and D. Mohanty, *PLoS Comput. Biol.*, 2009, **5**, e1000351.
121 P. C. Dorrestein, J. Blackhall, P. D. Straight, M. A. Fischbach, S. Garneau-Tsodikova, D. J. Edwards, S. McLaughlin, M. Lin, W. H. Gerwick, R. Kolter, C. T. Walsh and N. L. Kelleher, *Biochemistry*, 2006, **45**, 1537–1546.
122 P. C. Dorrestein, S. B. Bumpus, C. T. Calderone, S. Garneau-Tsodikova, Z. D. Aron, P. D. Straight, R. Kolter, C. T. Walsh and N. L. Kelleher, *Biochemistry*, 2006, **45**, 12756–12766.
123 F. Kopp and M. A. Marahiel, *Curr. Opin. Biotechnol.*, 2007, **18**, 513–520.
124 L. Robbel, K. M. Hoyer and M. A. Marahiel, *FEBS J.*, 2009, **276**, 1641–1653.
125 K. M. Hoyer, C. Mahlert and M. A. Marahiel, *Chem. Biol.*, 2007, **14**, 13–22.
126 R. O. Okotore and D. W. Russell, *Can. J. Biochem.*, 1972, **50**, 428–439.
127 D. W. Russell, *Biochim. Biophys. Acta, Gen. Subj.*, 1972, **261**, 469–474.
128 C. Lee, H. Gorisch, H. Kleinkauf and R. Zocher, *J. Biol. Chem.*, 1992, **267**, 11741–11744.
129 C. Lee and R. Zocher, *J. Biochem. Mol. Biol.*, 1996, **29**, 493–499.
130 C. Lee, H. Goerisch and R. Zocher, *J. Biochem. Mol. Biol.*, 2000, **33**, 228–233.
131 A. Ciulli, D. Y. Chirgadze, A. G. Smith, T. L. Blundell and C. Abell, *J. Biol. Chem.*, 2007, **282**, 8487–8497.
132 Y. Xu, E. M. K. Wijeratne, P. Espinosa-Artiles, A. A. L. Gunatilaka and I. Molnár, *ChemBioChem*, 2009, **10**, 345–354.
133 A. V. Ramaswamy, C. M. Sorrels and W. H. Gerwick, *J. Nat. Prod.*, 2007, **70**, 1977–1986.
134 J. H. Hurst, M. L. Billingsley and W. Lovenberg, *Biochem. Biophys. Res. Commun.*, 1984, **122**, 499–508.
135 E. Conti, T. Stachelhaus, M. A. Marahiel and P. Brick, *EMBO J.*, 1997, **16**, 4174–4183.
136 L. E. N. Quadri, P. H. Weinreb, M. Lei, M. M. Nakano, P. Zuber and C. T. Walsh, *Biochemistry*, 1998, **37**, 1585–1595.
137 P. A. Plattner, U. Nager and A. Boller, *Helv. Chim. Acta*, 1948, **31**, 594–602.
138 S. Gupta, C. Montllor and Y.-S. Hwang, *J. Nat. Prod.*, 1995, **58**, 733–738.
139 S. Weist and R. D. Süssmuth, *Appl. Microbiol. Biotechnol.*, 2005, **68**, 141–150.

140 M. Krause, A. Lindemann, M. Glinski, T. Hornbogen, G. Bonse, P. Jeschke, G. Thielking, W. Gau, H. Kleinkauf and R. Zocher, *J. Antibiot.*, 2001, **54**, 797–804.
141 1997.
142 A. Pohanka, K. Capieau, A. Broberg, J. Stenlid, E. Stenstrom and L. Kenne, *J. Nat. Prod.*, 2004, **67**, 851–857.
143 S. Uhlig, L. Ivanova, D. Petersen and R. Kristensen, *Toxicon*, 2009, **53**, 734–742.
144 C. Nilanonta, M. Isaka, R. Chanphen, N. Thong-orn, M. Tanticharoen and Y. Thebtaranonth, *Tetrahedron*, 2003, **59**, 1015–1020.
145 Y. Xu, J. Zhan, E. M. K. Wijeratne, A. M. Burns, A. A. L. Gunatilaka and I. Molnár, *J. Nat. Prod.*, 2007, **70**, 1467–1471.
146 K. J. Weissman, *Trends Biotechnol.*, 2007, **25**, 139–142.
147 A. Kirschning, F. Taft and T. Knobloch, *Org. Biomol. Chem.*, 2007, **5**, 3245–3259.
148 J. Kennedy, *Nat. Prod. Rep.*, 2008, **25**, 25–34.
149 W. T. Shier, K. L. Rinehart, Jr. and D. Gottlieb, *Proc. Natl. Acad. Sci. U. S. A.*, 1969, **63**, 198–204.
150 H. Doucet and J. C. Hierso, *Angew. Chem., Int. Ed.*, 2007, **46**, 834–871.
151 K. Yanai, N. Sumida, K. Okakura, T. Moriya, M. Watanabe and T. Murakami, *Nat. Biotechnol.*, 2004, **22**, 848–855.
152 P. A. Plattner and U. Nager, *Experientia*, 1947, **3**, 325–326.
153 B. S. Deol, D. D. Ridley and P. Singh, *Aust. J. Chem.*, 1978, **31**, 1397–1399.
154 P. Vongvilai, M. Isaka, P. Kittakoop, P. Srikitikulchai, P. Kongsaeree, S. Prabpai and Y. Thebtaranonth, *Helv. Chim. Acta*, 2004, **87**, 2066–2073.

