Wolfgang Hüttel*

Structural diversity in echinocandin biosynthesis: the impact of oxidation steps and approaches toward an evolutionary explanation

DOI 10.1515/znc-2016-0156
Received July 29, 2016; revised July 29, 2016; accepted August 28, 2016

Abstract: Echinocandins are an important group of cyclic non-ribosomal peptides with strong antifungal activity produced by filamentous fungi from Aspergillaceae and Leotiomycetes. Their structure is characterized by numerous hydroxylated non-proteinogenic amino acids. Biosynthetic clusters discovered in the last years contain up to six oxygenases, all of which are involved in amino acid modifications. Especially, variations in the oxidation pattern induced by these enzymes account for a remarkable structural diversity among the echinocandins. This review provides an overview of the current knowledge of echinocandin biosynthesis with a special focus on diversity-inducing oxidation steps. The emergence of metabolic diversity is further discussed on the basis of a comprehensive overview of the structurally characterized echinocandins, their producer strains and biosynthetic clusters. For the pneumocandins, echinocandins produced by *Glarea lozoyensis*, the formation of metabolic diversity in a single organism is analyzed. It is compared to two common models for the evolution of secondary metabolism: the 'target-based' approach and the 'diversity-based' model. Whereas the early phase of pneumocandin biosynthesis supports the target-based model, the diversity-inducing late steps and most oxidation reactions best fit the diversity-based approach. Moreover, two types of diversity-inducing steps can be distinguished. Although incomplete hydroxylation is a common phenomenon in echinocandin production and secondary metabolite biosynthesis in general, the incorporation of diverse hydroxyprolines at position 6 is apparently a unique feature of pneumocandin biosynthesis, which stands in stark contrast to the strict selectivity found in echinocandin biosynthesis by Aspergillaceae. The example of echinocandin biosynthesis shows that the existing models for the evolution of secondary metabolism can be well applied to parts of the pathway; however, thus far, there is no comprehensive theory that could explain the entire biosynthesis.

Keywords: antifungals; gene clusters; metabolic diversity; non-ribosomal peptide biosynthesis; secondary metabolism; filamentous fungi.

***Corresponding author: Wolfgang Hüttel,** Institute of Pharmaceutical Sciences, University of Freiburg, Albertstrasse 25, 79104 Freiburg, Germany,
E-mail: wolfgang.huettel@pharmazie.uni-freiburg.de

1 Introduction

Echinocandins are fungal non-ribosomal cyclic hexapeptides with a fatty acid side chain attached to a dihydroxyornithine residue. As they are specific noncompetitive inhibitors of the β-1,3-glucan synthase involved in fungal cell wall biosynthesis, they have a pronounced antifungal bioactivity. Although natural echinocandins are not of clinical use due to their toxicity and low solubility, chemical derivatives such as caspofungin, anidulafungin, and micafungin are most important drugs for the treatment of invasive mycoses. The pharmacological use of echinocandins has been reviewed in several overview articles [1–4]. In Table 1, some properties of echinocandins applied in therapy are summarized.

Besides the pharmacological properties of the echinocandins, their extraordinary structure is very interesting for biosynthetic studies (Figure 1). Thus, most amino acids in echinocandins are non-proteinogenic, the ring is closed via an unusual *N*-acyl-hemiacetal, and in some biosynthetic clusters, a polyketide synthase (PKS) for the synthesis of a branched chain fatty acid side chain is included.

The history of exploring echinocandin biosynthesis has been featured in previous review articles [1, 2, 7–9]. In brief, the biosynthetic origin of the non-proteinogenic amino acids and the dimethylmyristate side chain was elucidated in the early 1990s by ^{13}C-labeling experiments with the pneumocandin producer *G. lozoyensis* [10, 11]. In 2003, an α-ketoglutarate (αKG)-dependent L-proline hydroxylase (PH) activity, which was thought to be involved in pneumocandin biosynthesis, was found in crude protein extracts of *G. lozoyensis* [12]. However, nothing was known about the genes encoding the biosynthesis of the pharmaceutically important compounds until 2012 when the

Table 1: In vitro bioactivity and pharmacological data of semisynthetic echinocandins used for the treatment of systemic fungal infections.

	Anidulafungin	Caspofungin	Micafungin
In vitro bioactivity against *Candida* isolates, average MIC_{90} (range) (μg/mL)			
Candida albicans, 2867 clinical isolates	0.06 (0.007–1.0)	0.06 (0.007–0.5)	0.03 (0.007–1.0)
Candida spp., 5346 clinical isolates	2.0 (0.007– ≥ 8.0)	0.25 (0.007– ≥ 8.0)	1.0 (0.007–2.0)
Pharmacological data			
Route of administration dosage (adults) (mg/day)	As the oral bioavailability of echinocandins is very poor, they are generally administered by intravenous infusion.		
	50–150	50–70	50–150
Protein binding (%)	84	97	99
Volume of distribution (L/kg)	0.7–0.9	n.a.	0.24
Elimination half-life (h)	24–26	9–11	11–17
interactions with CYP450	no	Weak	Weak
Metabolism	Chemical degradation to inactive metabolites	Peptide hydrolysis and N-acylation	Degradation via arylsulfatase and catechol-*O*-methyl-transferase
Elimination	Mainly feces, 1% urine (< 10% unchanged)	35% feces, 41% urine (1.4% unchanged)	40% feces, < 15% urine (< 1% unchanged)
Tolerance	Due to their specific mode of action, echinocandins are comparatively well tolerated. For all echinocandins the most common adverse effects are rather unspecific, such as nausea, fever, phlebitis, or indigestion.		
Resistances	Echinocandin resistances are relatively rare (approximately 2%–3% with *C. albicans* and most other *Candida* spp.) [5]. Nevertheless, resistant strains are known from most pathogenic *Candida* and *Aspergillus* species [5, 6].		

Unless otherwise indicated data were taken from Groll et al. [4] and the literature cited therein.

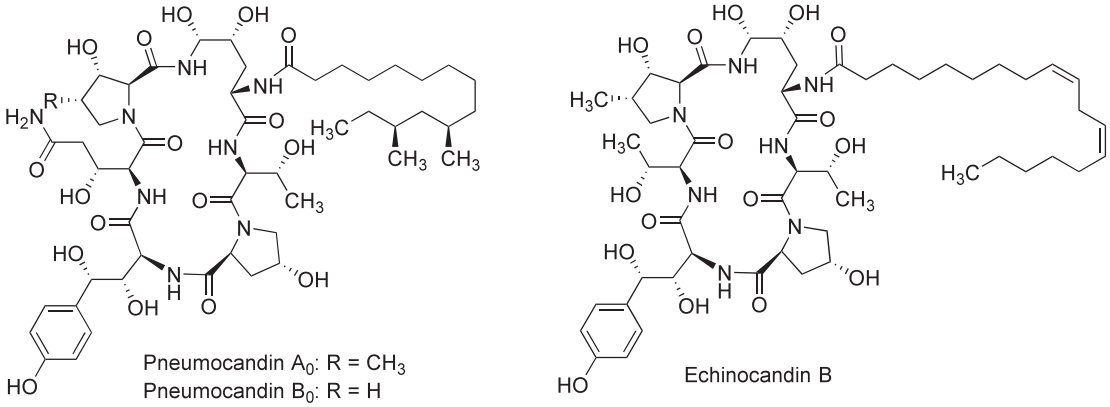

Figure 1: Examples of echinocandin structures. An overview of all compounds and their producer strains can be found in Tables 5 and 6, respectively.

groups of Tang and Walsh disclosed the genes encoding echinocandin B biosynthesis in the genome of *Aspergillus pachycristatus* NRRL 11440 (formerly *Emericella rugulosa*) [13]. According to this study, the genes are arranged in two separate partial clusters (*Ecd* and *Hty*) located on different contigs of the assembly. However, we showed recently that *Ecd* and *Hty* can be aligned to a single cluster (*Ecd/Hty*) and the disjunct genomic location of the cluster fragments was likely an artifact of genome misassembly in the absence of a reference genome [14]. It was also found that *Ecd/Hty* was virtually identical with a putative echinocandin B biosynthetic cluster *AE* deposited at the NCBI as a sequence of the producer strain *Aspergillus delacroxii* NRRL 3860 (formerly *Emericella nidulans* var. *echinulatus*) (cf. Table 5). Sequence analysis of PCR samples from the genomes of both strains, showed that the sequence of *AE* cluster is identical with that of *A. pachycristatus*, but not with *A. delacroxii* [14]. Moreover, a comparison of the calmodulin genes, a common taxonomic marker for fungi, suggested that strain NRRL 3860 belongs to the *Aspergillus*

pachychrystatus/*Aspergillus rugulosa* group rather than to *A. delacroxii*. However, a more detailed investigation will be necessary to clarify the species affiliation [14, 15].

The central gene of the *Ecd/Hty* cluster codes for a non-ribosomal peptide synthetase (NRPS) with a (T_0 CAT CAT CAT CAT CAT CAT C_T)-domain structure (T = thiolation, C = condensation, A = adenylation), which allows the assembly of six amino acids. It was shown that the initial thioesterase domain T_0 is loaded with linoleic acid activated by the fatty acyl-AMP ligase. The first CAT module is selective for ornithine, which is initially N-acylated with the linoleic acid side chain. After assembly of the six amino acids, the peptide is cyclized by the terminal condensation domain (C_T) [13].

Another group of enzyme genes in the *Hty* section of the cluster encodes the biosynthesis of homotyrosine according to a mechanism already postulated by Adefarati et al. in 1991 based on labeling experiments [10, 11]. First, acetate is added to the keto group of hydroxyphenylpyruvate by an isopropylmalate synthase (IPMS) (cf. Figure 3). Then, the hydroxyl group of the product is shifted from C-3 to C-2 by an aconitase (ACN). After oxidative decarboxylation, an α-keto acid is obtained, which is finally transaminated to provide homotyrosine. In a subsequent 2013 study, Tang, Walsh, and colleagues revealed the activity of the three oxygenases in the *Ecd* section of *Ecd*/*Hty*, which will be discussed in the next section [16].

In the same year, the genome of *G. lozoyensis* ATCC 20868 was reported, including the biosynthesis of pneumocandin encoded in a single cluster (*GL*) [17]. The most obvious difference to the echinocandin biosynthetic cluster was the presence of a PKS, which is thought to be responsible for the synthesis of the dimethylmyristate side chain. Four non-heme dioxygenases were identified. For one of them, thought to be a glutamine hydroxylase (GH), no homologue has been found in *Ecd*/*Hty*. Another dioxygenase was characterized as an αKG-dependent PH [18]. More recently, this and three other oxygenases were deleted in *G. lozoyensis* and the production of pneumocandins was analysed [19, 20]. As the oxidation steps are very complex in echinocandin biosynthesis and play a key role in the formation of structural diversity, they are discussed in more detail in the following section.

2 Oxidation reactions in echinocandin biosynthesis

By a combination of predicted activities for the putative proteins with experimental results, a fairly clear picture of the origin and assembly of the building blocks in echinocandin biosynthesis can be obtained (cf. Figure 3) [16]. Nevertheless, the order of biosynthetic steps, especially of the oxidations catalyzed by up to six dioxygenases, remains enigmatic. For most of these enzymes, it is not fully clear whether they accept the free amino acid, an amino acyl or peptidyl-*S*-NRPS precursor, or a dehydroxyechinocandin framework as substrate.

2.1 Oxygenases in *A. pachycristatus*

After the discovery of the echinocandin biosynthetic cluster in *A. pachycristatus*, the groups of Tang and Walsh revealed the activity of the three oxygenases in the *Ecd* section of the *Ecd*/*Hty* cluster, a L-homotyrosine hydroxylase (hT3H), an ornithine hydroxylase (OrnH), and a L-leucine dioxygenase (LDO)-producing (*S*)-1-pyrroline-5-carboxylate, a precursor of *trans*-4-methyl-L-prolin (cf. Figure 3 and Table 4) [16]. Deletion of the αKG-dependent LDO EcdK resulted in a breakdown of echinocandin B biosynthesis; however, a minimal activity was detected in vitro with the heterologously expressed enzyme: the *pro*-(*S*)-methyl group of leucine was first hydroxylated and then further oxidized to the aldehyde, which undergoes spontaneous condensation to the cyclic imine. It is assumed that, in vivo, the imine is reduced to *trans*-4-methylproline by a dehydrogenase whose gene is not included in the cluster.

Genomic deletion and in vitro experiments have consistently shown that the αKG-dependent non-heme dioxygenase EcdG is a homotyrosine 3-hydroxylase [16]. It was found that EcdG acts upon the free amino acid, but not upon the homotyrosine residue in echinocandin D (cf. Table 6). Subsequent hydroxylation to 3,4-dihydroxyhomotyrosine was not observed. Cytochrome P450 monooxygenase EcdH (OrnH) is responsible for both hydroxyl groups at the ornithine residue in echinocandins, which was discovered by the analysis of echinocandin products produced by the corresponding deletion mutant. It should be noted that hydroxylation at C-5 of the free ornithine would result in a terminal hemiaminal, which undergoes rapid hydrolysis to the aldehyde. Hence, a cyclization of the linear echinocandin precursor via formation of an *N*-acyl-hemiacetal is no longer possible. Thus, it has been concluded that ornithine hydroxylation at C-5 must occur after cyclization of an echinocandin precursor and is therefore a very late step in biosynthesis.

Further information about the order of oxidation steps was deduced from deletion of *ecdG* and *ecdH* [16]. Echinocandins produced by these strains not only lack the

hydroxyl groups introduced by the corresponding gene, the hydroxylation at other positions was also incomplete. This effect was explained by a reduced activity of hydroxylases catalyzing subsequent steps because a hydroxyl group is missing in the substrate. In this way, the order of hydroxylation steps could be deduced. Proline or methylproline hydroxylation was not affected by any of the deletions, supporting the hypothesis that this hydroxylation occurs at an earlier stage of echinocandin B biosynthesis, most likely of the free amino acid. In the case of mutant Δ*ecdG*, the missing hydroxyl group at homotyrosine C-3 strongly influences the insertion of the hydroxyl group at homotyrosine C-4 and ornithine hydroxylation at C-4 and C-5 by EcdH (Figure 2) is also affected.

Consequently, both hydroxylation events are thought to occur after homotyrosine C-3-hydroxylation. The ΔEcdG mutant produced echinocandin B derivatives with non-hydroxylated, C-5-monohydroxylated, and C-4,C-5-dihydroxylated ornithine, but there was no hydroxylation solely at C-4. From that, it was deduced that EcdH hydroxylates ornithine preferentially first at C-5, so that the hemiaminal is formed and then at C-4. In both mutants, the hydroxylation of homotyrosine at C-4 was significantly reduced indicating that this is presumably

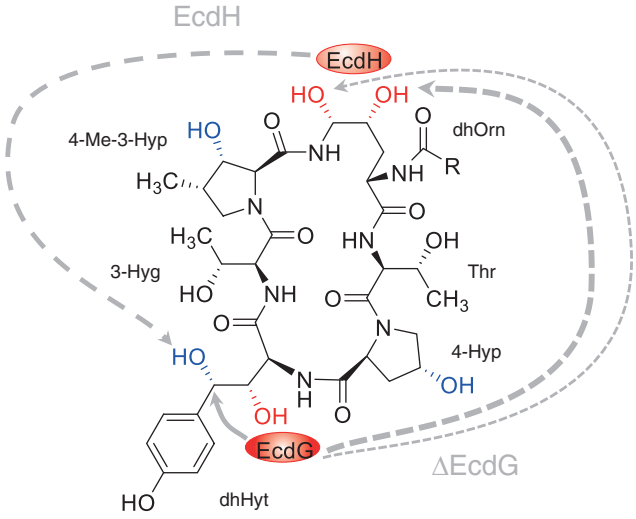

Figure 2: Secondary effects on hydroxylations with deletion mutants ΔEcdG and ΔEcdH [16]. The missing hydroxyl group at homotyrosine C-3 hampers the hydroxylation of homotyrosine C-4 and ornithine C-4 and C-5. The line width of the blue arrows symbolizes the strength of the effect. A lack of ornithine hydroxylation with mutant ΔEcdH affected only the hydroxylation at homotyrosine C-4. Abbreviations of non-proteinogenic amino acids (all in L-configuration): 4-Hyp, *trans*-4-hydroxyproline; 3,4dhHyt, (3*S*,4*S*)-3,4-dihydroxyhomotyrosine; 3-Hyg, (3*S*)-3-hydroxyglutamine; 4-Me-3-Hyp, (3*S*,4*S*)-3-hydroxy-4-methylproline; dhOrn, (4*R*,5*R*)-4,5-dihydroxyornithine (N_β-acylated with fatty acid).

the last oxidation step of echinocandin B biosynthesis. From these results, a general scheme for echinocandin biosynthesis can be derived (Figure 3). For two biocatalytic steps, the reduction 4-methylpyrroline-5-carboxylate to 4-methylproline and the *ortho*-sulfation of the homotyrosine, no corresponding genes are found in the clusters. Imine reduction is most likely performed by a 1-pyrroline-5-carboxylate reductase (EC 1.5.1.2) or a proline dehydrogenase (EC 1.5.5.2). Aromatic hydroxylations and sulfations are common processes in catabolism. Clustering of the genes is apparently not necessary.

2.2 Proline hydroxylation in *G. lozoyensis*

As early as 2003, Petersen et al. detected an αKG-dependent PH activity in the crude protein extract of *G. lozoyensis* producing *trans*-4-hydroxyproline and also the *trans*-3-hydroxy isomer in substantially smaller amounts [12]. About a decade later, our group characterized the corresponding PH (GloF) [18] in the genome of *G. lozoyensis* mutant strain ATCC 74030 [21] (GloF is identical with GLOXY2 in the genome of wild-type *G. lozoyensis* ATCC 20868 [17, 20]). It is distantly related to a group of fungal pipecolic acid hydroxylases discovered only recently [22]. In vitro experiments with the heterologously expressed and purified enzyme showed that proline is converted into *trans*-4-hydroxyproline and a minor amount of the *trans*-3-hydroxy isomer. Additionally, GloF hydroxylates *trans*-4-methylproline to provide (*S*,*S*)-4-methyl-3-hydroxyproline, so that all three hydroxyproline building blocks for the biosynthesis of pneumocandins A and B are provided by this enzyme [18]. The ratio of *trans*-4/*trans*-3 hydroxyproline was about 8:1. This corresponds well with the approximately 7:1 demand for hydroxyprolines for the production of pneumocandins A_0 and B_0, as found for wild-type *G. lozoyensis*. For the selectivity of GloF to actually be correlated with the ratio of pneumocandins A_0 and B_0 products, two prerequisites must be met: (i) the *trans*-3/*trans*-4 selectivity of GloF should be similar or identical in vitro and in vivo; (ii) the adenylation (A)-domain in module 6 of the NRPS must accept 3-hydroxyproline very well (besides 4-methyl-3-hydroxyproline), so that it is quantitatively consumed. Although some indirect evidence for this concept can be derived from other biosynthesis studies (see Section 2.3), further experimental confirmation is required. Proline residues in small peptides were not accepted as substrates by GloF. An activity with acyl carrier protein-bound proline could not be excluded; however, the relatively high activity with the free amino acids suggests that these are the native substrates.

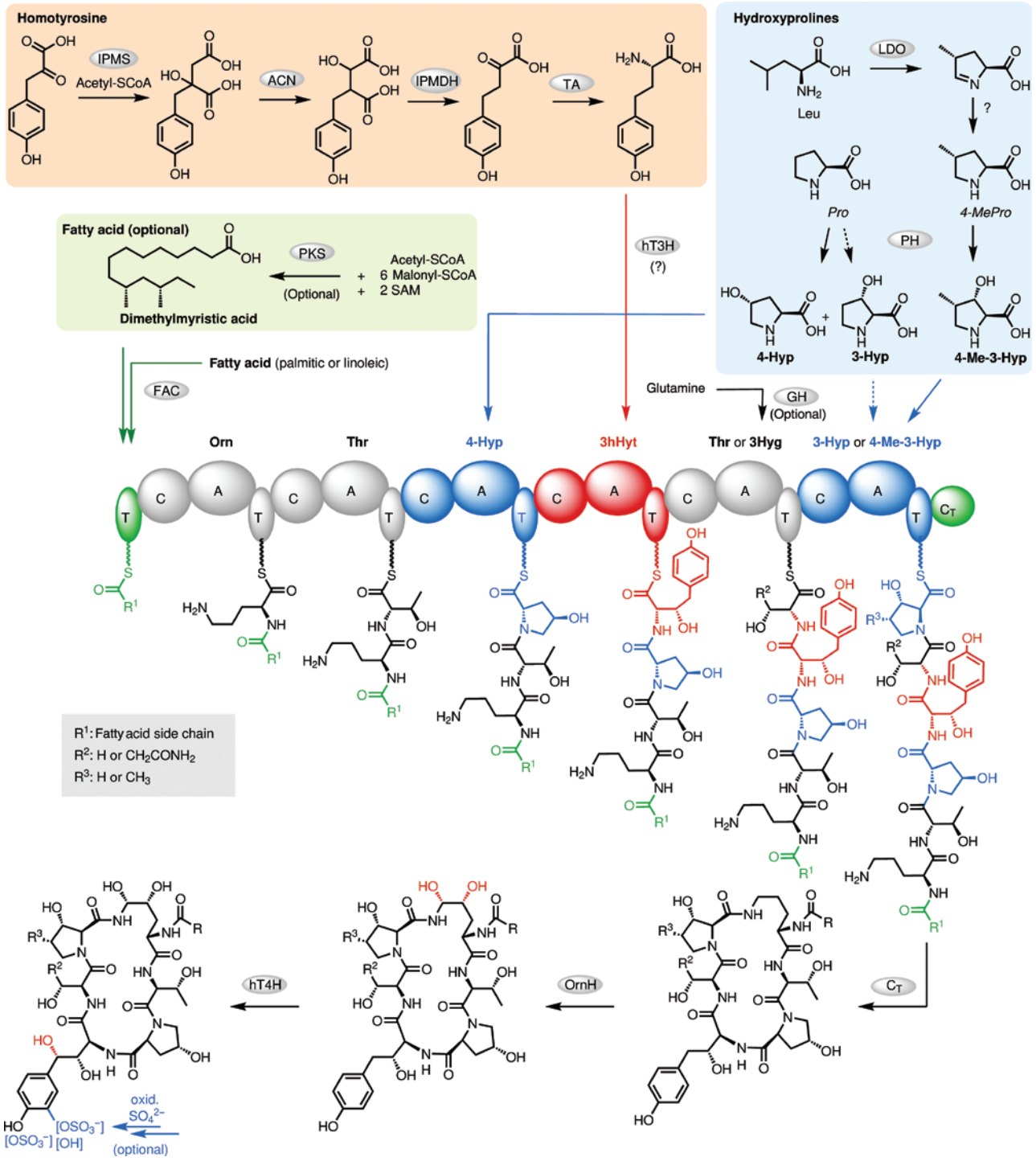

Figure 3: Overview of echinocandin biosynthesis. Abbreviations: FAC, long-chain fatty acyl-CoA ligase; IPMS, isopropylmalate synthase; ACN, aconitase; IPMDH, isopropylmalate dehydrogenase; TA, transaminase; LDO, L-leucine dioxygenase; PH, L-proline hydroxylase; hT3H, L-homotyrosine 3-hydroxylase; GH, putative L-glutamine hydroxylase; OrnH, L-ornithine 4,5-hydroxylase; hT4H, putative L-homotyrosine 4-hydroxylase; PKS, polyketide synthase (see also Table 4); non-proteinogenic amino acids (all in L-configuration): 3-Hyp, *trans*-3-hydroxyproline; 4-Hyp, *trans*-4-hydroxyproline; 3hHyt, (3*S*)-3-hydroxyhomotyrosine; 3-Hyg, (3*S*)-3-hydroxyglutamine, 4-Me-3-Hyp, (3*S*,4*S*)-3-hydroxy-4-methylproline.

Whereas GloF obviously meets the demand for pneumocandin production in wild-type *G. lozoyensis*, this does not apply to the mutant strain. For example, for mutant strain ATCC 74030, which produces pneumocandin B_0 almost exclusively, a *trans*-4/*trans*-3-hydroxyproline ratio of 1:1 would be ideal. With the PH

activity of GloF, a shortage of *trans*-3-hydroxyproline and an overflow of the *trans*-4-hydroxy isomer would be expected. Feeding experiments with hydroxyprolines suggest that this is indeed the case [9, 12]. Addition of *trans*-3-hydroxyproline (0.13 M) to a culture of *G. lozoyensis* ATCC 74030 resulted in a substantial increase in pneumocandin B_0 production (+39% relative to the non-supplemented culture, Figure 4). The concentrations of byproducts with proline derivatives other than *trans*-3-hydroxyproline (pneumocandins C_0, D_0, and E_0) were drastically reduced. In contrast, feeding of *trans*-4-hydroxyproline provided only a relatively small increase in pneumocandin B_0 production (+9%). As expected, the concentrations of pneumocandins C_0 and D_0 were strongly increased. The feeding of proline itself effected a strongly increased intracellular concentration of this amino acid, as proven by a 349% increase in pneumocandin E_0 in the product mixture. Moreover, the production of hydroxyprolines should be enhanced, too. According to the intrinsic selectivity of PH GloF (*trans*-4/*trans*-3 = 8:1), a large quantity of *trans*-4- and a minor amount of *trans*-3-hydroxyproline should be produced. However, the metabolic spectrum of the feeding experiment was primarily affected by an increased supply of *trans*-3-hydroxyproline rather than of the *trans*-4-hydroxy isomer. The concentrations of pneumocandins C_0 and D_0 were less than half of those from cultures without proline feeding, even though a substantially increased intracellular concentration of *trans*-4-hydroxyproline can be assumed. The pronounced effect of *trans*-3-hydroxyproline on the production of pneumocandins is a strong indicator that this compound is, as predicted, a limiting factor for pneumocandin B_0 biosynthesis in *G. lozoyensis* ATCC 74030.

2.3 Other oxygenases involved in pneumocandin biosynthesis

With the protein sequences of the oxygenases from echinocandin B biosynthesis as templates, it was easy to predict the functions of the other oxygenases in the pneumocandin biosynthetic cluster, as shown in Table 2. An additional gene for an αKG/Fe(II)-dependent dioxygenase was found in pneumocandin biosynthesis, which has no homologue in echinocandin B biosynthesis. As residue 6 in the pneumocandins, but not in echinocandin B, is hydroxyglutamine it can be concluded that the enzyme is a GH [18].

Currently, two genome sequences of the pneumocandin producer *G. lozoyensis* are available: one of the wild-type strain ATCC 20868 [17], which is a producer

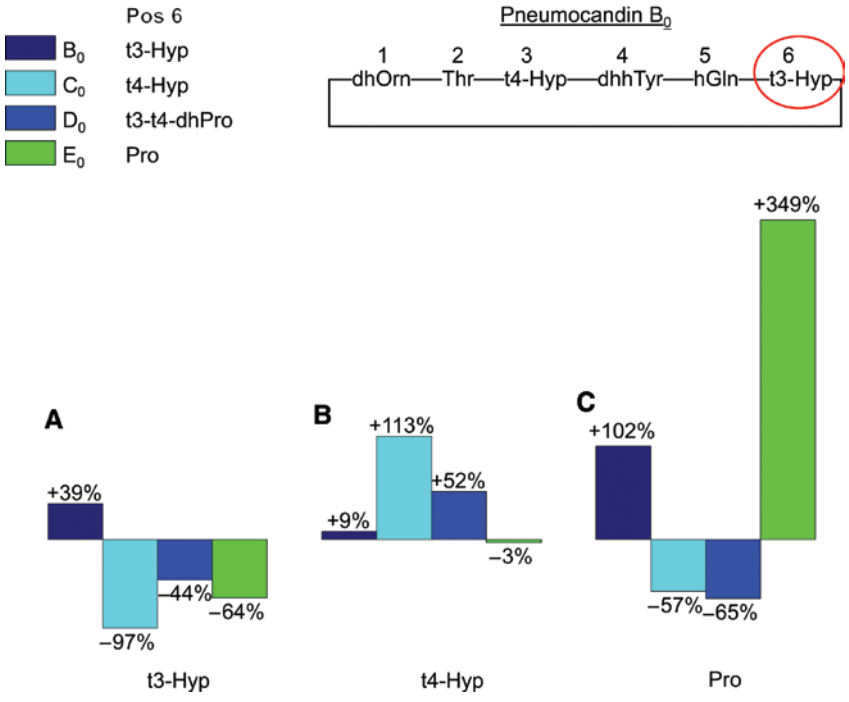

Figure 4: Relative effect on the production of pneumocandin B_0, C_0, D_0, and E_0 after feeding (0.13 M) of (A) *trans*-3-hydroxyproline (t3-Hyp), (B) *trans*-4-hydroxyproline (t4-Hyp), and (C) proline.

Table 2: Homologous dioxygenases from pneumocandin (*Glo*) and echinocandin B (*Ecd*/*Hty*) biosynthesis.

G. lozoyensis[a]	Experimentally determined activity	Type	Sequence identity	*A. pachychristatus*
GLOXY4 (GloC)	Leucine dioxygenase	αKG/Fe(II)	62%	EcdK
GLP450-2 (GloO)	Ornithine 3,4-hydroxylase	CYP	48%	EcdH
GLP450-1 (GloP)	Homotyrosine 4-hydroxylase	CYP	50%	HtyF
GLOXY-3 (GloE)	Putative glutamine hydroxylase	αKG/Fe(II)	–[b]	–
GLOXY2 (GloF)	Proline hydroxylase	αKG/Fe(II)	64%	HtyE
GLOXY1 (GloM)	Homotyrosine 3-hydroxylase	αKG/Fe(II)	58%	EcdG

Modified from [18].
[a]Protein nomenclature for wildtype [20] and in brackets: mutant strain ATCC 74030 [18]. [b]The highest sequence identity to a protein of the *A. pachychristatus* cluster was 33% (EcdG).

of pneumocandins A and B, and the other from mutant strain ATCC 74030 which produces pneumocandin B_0 predominantly [21]. Sequence comparison of the clusters showed that only one enzyme, leucine dioxygenase GloC (= GLOXY4 in wild type [20]), is modified at two sites in the mutant strain: T98I and A294T [20]. Although not located in the active site, it is assumed that these variations hamper the activity of the enzyme, which is involved in methylproline biosynthesis and thus essential for pneumocandin A_0 production (cf. Figure 2). Recently, the targeted deletion of GLOXY4 in wild-type *G. lozoyensis* has been reported [20]. As expected, mutant strain ΔGLOXY4 did not produce pneumocandin A. Instead, the production of pneumocandin B_0 was increased by a factor of 9.5. In a further study, three other oxygenases were deleted in wild-type *G. lozoyensis*, and the effects on pneumocandin production were investigated [19]. Deletion of the cytochrome P450 monooxygenase GLP450-1 gave pneumocandins without hydroxylation at homotyrosine C-4 (pneumocandin F and G), while the inactivation of GLP450-2 resulted in products with unmodified ornithine (pneumocandins A_2 and B_2). In extracts of mutant strain ΔGLOXY1, in which the putative αKG/Fe(II)-dependent homotyrosine C-3 oxygenase was deleted, a complex mixture of nine products was found. It consisted of A- and B-type pneumocandins with various hydroxylation patterns at ornithine C-4 and C-5, as well as at homotyrosine C-4; however, there was no hydroxylation at homotyrosine C-3. These results nicely support the model for echinocandin biosynthesis proposed by the groups of Tang and Walsh [13, 16]. They also show that the principle of echinocandin biosynthesis is well conserved, even in distantly related fungi.

3 Biosynthetic clusters

To date, the sequences of nine echinocandin biosynthetic clusters are available from the NCBI (Table 3). The corresponding gene maps are depicted in Figure 5 and the genes are explained in Table 4. Basically, there are two groups of clusters corresponding to two different classes of fungi, the Leotiomycetes (*GL*/*Glo*, *PH*, *CE_1*, *CE_2* and

Table 3: Echinocandin biosynthetic clusters from sequence databases.

Cluster	Organism	NCBI-ID
GL/Glo	GL: *G. lozoyensis* wild-type strain ATCC 20868	NW007360987
	Glo: *G. lozoyensis* mutant strain ATCC 74030	AGUE01000179
PH	*Phialophora* cf. *hyalina* No. 16616 (previously classified as *Tolypocladium parasiticum*)	AB720726
CE_1	*Coleophoma empetri* F-11899	AB723722
CE_2	*Coleophoma empetri*	AB720725
CC	*Coleophoma crateriformis*	AB720076
Ecd/*Hty*=*AE*	*Ecd*/*Hty*: *A. pachycristatus* NRRL 11440 (= *Emericella rugulosa*) [virtually identical with the *AE* cluster, originally deposited as a sequence from *A. delacroxii* NRRL 3860 (= *Emericella nidulans* var. *echinulatus*) [14]]	Alignment of *Ecd*: JX421684 and *Hty*: JX421685 *AE*: AB720074
AM	*Aspergillus muludensis* DSMZ 5745	KP742486
Ani	*Aspergillus nidulans* NRRL 8112	KT806042
AA	*Aspergillus aculeatus* ATCC 16872	GOLD Project ID: Gp0010055

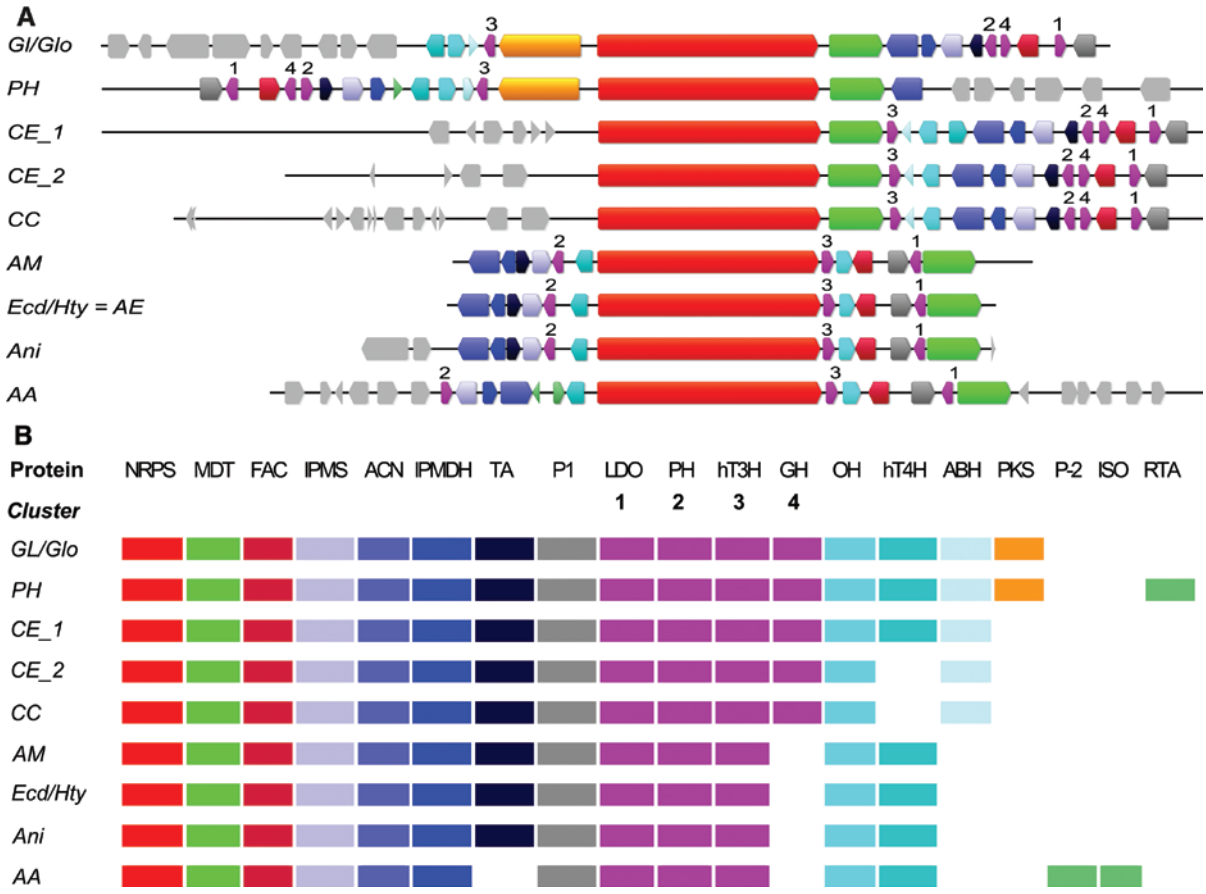

Figure 5: (A) Maps of echinocandin biosynthetic clusters. In the case of incompletely annotated clusters, the missing genes were identified with the annotation software AUGUSTUS [23] and corrected manually. (B) Genetic composition of echinocandin biosynthetic clusters (for abbreviations of clusters and proteins, see Tables 3 and 4, respectively).

CC) and the Eurotiomycetes (*Ecd*/*Hty*, *Ani*, *AE* and *AA*). Clusters from Leotiomycetes may be subdivided into *CE_1*, *CE_2* and *CC*, which are essentially collinear, and *GL*/*Glo* and *PH*. Although the latter appear to be different at first sight, they can formally be interconverted in a single DNA rearrangement if the reverse complement of the last eight genes from the C-terminus of *GL*/*Glo* is shifted to the N-terminus.

The genetic composition of the clusters, shown in Figure 5B, allows a relatively precise prediction of the structure of the corresponding products. Only the sulfate groups at the homotyrosine residue in some of the products are not encoded in the clusters. The presence of a PKS indicates a branched chain fatty acid residue, typically dimethylmyristate (*GL*/*Glo*, *PH*). Found in all Leotiomycetes clusters is an additional non-heme dioxygenase, most likely a GH. In the clusters *CE_2* and *CC*, the heme-dependent homotyrosine 4-hydroxylase (hT4H) is missing. Consequently, there is no hydroxyl group at homotyrosine C-4 in metabolites produced by such strains. The aculeacin biosynthetic cluster *AA* from *Aspergillus aculeatus* ATCC 16872 lacks the gene for the transaminase (TA), which is thought to catalyze the last step in homotyrosine biosynthesis. A BLASTP analysis of the proteome of this strain deposited at the JGI genome portal [24] with the aminotransferase (AT) from *Aspergillus nidulans* NRRL 8112 as template resulted to no putative AT with significantly increased sequence identity (>60%). This should be expected for an orthologous enzyme involved echinocandin biosynthesis. Therefore it is doubtful whether this strain is still able to produce aculeacin. In all clusters from Leotiomycetes, a putative gene for an α/β-hydrolase (ABH) was identified downstream from ornithine 4,5-hydroxylase (OrnH), which has not been described previously. Conclusions about its function are highly speculative and lack experimental evidence. The same applies to the relatively large (≈75 kbp) putative protein P1 of unknown function, which is found in all clusters and is always located downstream from the LDO, in reverse orientation.

Table 4: List of (putative) proteins identified in echinocandin biosynthetic clusters.

Name	Protein function	Evidence	Remarks
NRPS	Non-ribosomal peptide synthetase	E, P	
MDT	Multidrug transporter	P	
FAC	Long-chain fatty acyl-CoA ligase	P	
IPMS	Isopropylmalate synthase	P	Homotyrosine biosynthesis
ACN	Aconitase	P	Homotyrosine biosynthesis
IPMDH	Isopropylmalate dehydrogenase	P	Homotyrosine biosynthesis
TA	Transaminase	P	Homotyrosine biosynthesis, missing in AA
P1	Putative protein, unknown function		
1 LDO	αKG-dependent L-leucine dioxygenase	E	Very low activity in vitro
2 PH	αKG-dependent L-proline hydroxylase	E	Activity in vitro
3 hT3H	αKG-dependent L-homotyrosine 3-hydroxylase	E	Very low activity in vitro
4 GH	αKG-dependent putative L-glutamine hydroxylase	P, C	Facultative, predicted as αKG-dependent dioxygenase
OrnH	CYP, L-ornithine 4,5-hydroxylase	E	
hT4H	CYP, putative L-homotyrosine 4-hydroxylase	P, C	Facultative, predicted as CYP enzyme
ABH	Putative α/β-hydrolase (thioesterase)	P	Uncertain, facultative, function unclear
PKS	Polyketide synthase	P	Facultative, dimethylmyristate biosynthesis
P-2	Putative protein	P	Uncertain, putative proteins found in AA (P-2, ISO) and PH (RTA) exclusively
ISO	Putative NAD-dependent isomerase		
RTA	Putative RTA1 protein		

A proline dehydrogenase activity is expected for echinocandin biosynthesis, but is not found in the cluster. Likewise, the sulfation of some echinocandins is also not encoded in the gene clusters. E, experimental evidence; P, predicted by sequence homology; C, correlation between presence of gene and the occurrence of a particular functional group.

4 Fungal strains and echinocandin structures

During the search for antifungal echinocandin compounds, a number of producer strains have been identified. Some of these strains have also been cultivated on a technical scale [3], so that echinocandin byproducts produced in low concentrations could be isolated and structurally characterized. This section provides an overview of the echinocandin producer strains described so far and their diverse products. For discussion, it is necessary to distinguish between the main product and byproducts. The main product is the primarily produced product. As a rule, the activity of all synthetic enzymes of the biosynthetic cluster, and sometimes even external enzymes, is required to produce the main metabolite. The enzymes convert their preferred substrate with the preferred selectivity. As most organisms only produce one main product, structural diversity can be observed by comparing different species. Such diversity is always due to different genetic constitutions of the producers. In contrast, byproducts emerge from enzymes with relaxed substrate specificity or incomplete product selectivity. Some biosynthetic steps yield more than one product or are simply skipped. Furthermore, congeners may arise from biotransformations catalyzed by promiscuous external enzymes from other metabolic pathways.

Finally, as for other secondary metabolites, it should be noted that echinocandin production can be highly dependent on environmental conditions affecting gene expression and activity of enzymes in many ways.

4.1 Echinocandin producer strains

The current collection of echinocandin producers includes 24 different wild-type fungal strains from more than a dozen different species (Table 5). Fifteen strains belong to the class Leotiomycetes, mostly from the order Helotiales, and nine to the Aspergillaceae family (Eurotiomycetes).

The metabolic diversity produced by a single species is best seen for pneumocandin biosynthesis in *G. lozoyensis* (Helotiales). It is mainly characterized by incomplete hydroxylations of homotyrosine and ornithine C-4 and C-5. A characteristic feature of pneumocandin biosynthesis is the incorporation of diverse hydroxyprolines and proline at residue 6, which is normally reserved for 4-methyl-3-hydroxyproline in echinocandins (Table 6).

In Aspergillaceae, even from the best-characterized producers of echinocandin B or aculeacin, only a portion of the metabolite structures has been established (see

Table 5: Fungal strains producing echinocandins (see also [25]).

Abbreviation	Species	Accession no.	Comment	References
Wild-type strains				
Aa_1	*Aspergillus aculeatus*	M-4845, NRRL 11270		[26]
Aa_2	*Aspergillus aculeatus*	M-4124, NRRL 8075		[27, 28]
Aa_3	*Aspergillus aculeatus* (= *Aspergillus japonicus* var. *aculeatus*)	ATCC 16872, NRRL 5094	Cluster: *AA* (lacks AT) genome available under GOLD sequencing ID: Gp0010055	[24]
Al	*Acrophialophora limonispora* nov. spec. 'Dreyfuss + Muller'	NRRL 8095	Taxonomy of mitosporic *Ascomycota* is uncertain	[29]
An_1	*Aspergillus nidulans*	NRRL 8112, ATCC 58396	Cluster: *Ani*	[30]
An_2	*Aspergillus nidulans*	ZBJ-0817, CCTCC M 2010275		[31]
Ad	*A. delacroxii NRRL 3860* (= *Emericella nidulans* var. *echinulata)*	A 32204, NRRL 3860	Stated cluster: *AE* (identical with *Ecd*/*Hty*)	[32]
Ap	*A. pachycristatus* (= *Emericella rugulosa*, *Aspergillus nidulans* var. *roseus*)	NRRL 11440, ATCC 58397	Products: '*A-42355 antibiotic mixture*' cluster: *Ecd*/*Hty* (identical with *AE*)	[7, 33]
Ar_1	*Aspergillus rugulosus*	NRRL 8113	Factors A–G	[34]
Ar_2	*Aspergillus rugulosus*	NRRL 8039		[35, 36]
Am	*Aspergillus muludensis* (= *A. sydowii* Thorn and Church var. nov. *mulundensis* Roy)	Cult. No. Y-30462, DSMZ 5745	Cluster: *AM*	[7, 37–39]
Cc	*Coleophoma crateriformis*	FERM BP 5796, No. 738,	Cluster: *CC*	[40]
Ce_1	*Coleophoma empetri*	FERM BP 6252, F-11899	Cluster: *CE_1*	[41]
Ce_2	*Coleophoma empetri*	No. 14573	Cluster: *CE_2*	[42, 43]
Cq	*Cryptosporiopsis* cf. *quercina*			[44]
Csp_1	*Cryptosporiopsis* sp.	'Isolate P30A'		[45]
Csp_2	*Cryptosporiopsis* sp.	NRRL 12192, ATCC 20594		[46]
Csp_3	*Cryptosporiopsis* sp.	S 41 062/F		[47, 48]
Csp_4	*Chalara* sp. (= *Pochonia parasitica*, *Tolypocladium parasiticum*)	FERM BP 5553, No. 22210		[42, 49]
Gl	*G. lozoyensis* (= *Zalerion arboricola*)	ATCC 20868, F-160870	Pneumocandins $A_0/B_0 \approx 7{:}1$ Cluster: *GL*	[19, 50–55]
Gl_1	*G. lozoyensis*	F-226836		[50]
Gl_2	*G. lozoyensis*	F- 226838		[50]
Gl_3	*G. lozoyensis*	F- 239379		[50]
Psp	*Pezicula* sp.			[45]
Ph	*Phialophora* cf. *hyalina* (= *Tolypocladium parasiticum*)	No. 16616	Cluster: *Tpa*	[8, 42, 49]
Mutant strains				
An_M1	*Aspergillus nidulans*	ULN-59		[31, 56]
An_M2	*Aspergillus nidulans*	ZBJ12037, CCTCC M 2015677		[56]
Gl_M1	*G. lozoyensis*	ATCC 20957	Pneumocandins $A_0/B_0 \approx 1{:}10$	[54]
Gl_M2	*G. lozoyensis*	ATCC 20958	Main product: pneumocandin A_2	[54]
Gl_M3	*G. lozoyensis*	ATCC 20988	Main products: pneumocandins A_2 and A_4	[54]
Gl_M4	*G. lozoyensis*	ATCC 74030	Pneumocandin $A_0/B_0 \approx 1{:}80$ Cluster: *Glo*	[54]
Gl_M5	*G. lozoyensis*	CGMCC 2933	Developed from strain ATTC 20957 increased Pneumocandin B_0 production Cluster: *Glo*	[57]
Gl_ΔGLOXY1	*G. lozoyensis*	ATCC 20868 ΔGLOXY1	9 Pneumocandins without hydroxylation at hTyr C-3	[19]
Gl_ΔGLOXY4	*G. lozoyensis*	ATCC 20868 ΔGLOXY4	100% pneumocandin B_0	[20]
Gl_ΔGLP450-1	*G. lozoyensis*	ATCC 20868 ΔGLP450-1	Pneumocandins F and G	[19]
Gl_ΔGLP450-2	*G. lozoyensis*	ATCC 20868 ΔGLP450-2	Pneumocandins A_2 and B_2	[19]
Ph_M1	*Phialophora* cf. *hyalina* (= *Tolypocladium parasiticum*)	No. 16616 mutant strain	Increased production of WF16616	[58]

Table 6: Echinocandin structures [24, 33, 41, 45, 58].

Amino acid no.		1		Fatty acid	2	4				5	6		Strain[b] (Cluster)
Compound[a]	Synonyms	Orn			Thr Ser	homo-*Tyr*				Thr hGln Ser	Hyp		
		R^1 C4	R^2 C5	R^3		R^5 C3	R^6 C4	R^7 *meta*	R^8 *para*		R^{10} C3	R^{11} C4	
Pneumocandin A_0	L-671,329	OH	OH	Dimethylmyristoyl	Thr	OH	OH	H	OH	hGln	OH	CH_3	Gl (*GL*), Gl_1, Gl_2, Gl_3, Csp_1, Psp
Pneumocandin A_1						H							**Gl** (*GL*), **Gl_1**, **Gl_2**, **Gl_3**, Gl_ΔGLOXY1
Pneumocandin A_2		H	H										Gl_M2, Gl_M3, Gl_ΔGLP450-2
Pneumocandin A_3		H				H	H						**Gl** (*GL*), Gl_ΔGLOXY1
Pneumocandin A_4		H	H			H	H						**Gl** (*GL*), Gl_M3, Gl_M2, Gl_ΔGLOXY1
Pneumocandin G	Pneumocandin A_G[c]						H						Gl_ΔGLP450-1
Compound 9	Pneumocandin A_9[c]	H				H							Gl_ΔGLOXY1
Compound 13	Pneumocandin A_{13}[c]		H			H	H						Gl_ΔGLOXY1
Pneumocandin B_0												H	**Gl** (*GL*), **Gl_1**, **Gl_2**, **Gl_3**, Gl_M1, Gl_M4 (*Glo*), Gl_M5, Gl_ΔGLOXY4
Pneumocandin B_1						H						H	**Gl** (*GL*), Gl_ΔGLOXY1
Pneumocandin B_2		H	H									H	**Gl** (*GL*), Gl_ΔGLP450-2
Pneumocandin B_5		H										H	**Gl** (*GL*)
Pneumocandin B_6			H									H	**Gl** (*GL*)
Pneumocandin F	Pneumocandin B_F[c]						H					H	Gl_ΔGLP450-1
Compound 10	Pneumocandin B_{10}[c]	H				H	H					H	Gl_ΔGLOXY1
Compound 11	Pneumocandin B_{11}[c]		H			H	H					H	Gl_ΔGLOXY1
Compound 14	Pneumocandin B_{14}[c]	H	H			H	H					H	Gl_ΔGLOXY1
Pneumocandin $B_{0\text{-Ser}}$	Sporiofungin B				Ser							H	**Gl** (*GL*), **Csp_3**
Pneumocandin $B_{5\text{-Ser}}$		H			Ser							H	**Gl** (*GL*)
Pneumocandin C_0											H	OH	**Gl** (*GL*)
Pneumocandin D_0												OH	Gl_M1
Pneumocandin D_2		H	H									OH	Gl (wild-type?, Ref. 6)
Pneumocandin E_0											H	H	**Gl** (*GL*)
Sporiofungin A					Ser								Csp_3
Sporiofungin C			OCH_3		Ser								Csp_3
Antibiotics 31794/F-1				Myristoyl									Al
S 41062/F-1				'Branched C-16 fatty acid'	Ser		H						Csp_2[d]
Cryptocandin				Palmitoyl						Gln		CH_2OH	Cq
WF11899A	FR901379[d]			Palmitoyl				OSO_3^-					**Ce_1** (*CE_1*)
WF11899B	FR901381[e]			Palmitoyl			H	OSO_3^-					**Ce_1** (*CE_1*)
WF11899C	FR901382[e]	H		Palmitoyl			H	OSO_3^-					**Ce_1** (*CE_1*)
WF14573B[f]	FR220897[e]			Palmitoyl			H	OH	OSO_3^-				**Ce_2** (*CE_2*)

Table 6 (continued)

Amino acid no.		1		Fatty acid	2	4				5	6		Strain[b] (Cluster)
Compound[a]	**Synonyms**	**Orn**			**Thr Ser**	**homo-*Tyr***				**Thr hGln Ser**	**Hyp**		
		R^1 C4	**R^2 C5**	**R^3**		**R^5 C3**	**R^6 C4**	**R^7 *meta***	**R^8 *para***		**R^{10} C3**	**R^{11} C4**	
WF14573A[f]	FR220899[e]			Palmitoyl			H	OH	OSO_3^-			H	**Ce_2** (*CE_2*)
WF16616	FR190293[e]							OSO_3^-					**Tp** (*PH*), Tp_M1
WF22210	Fr227673[e]			Dimethylpalmitoyl				OSO_3^-					Csp_4
WF738A	FR209602[e]			Palmitoyl	Ser		H	OSO_3^-					**Cc** (*CC*)
WF738B	FR209603[e]			Palmitoyl	Ser		H	OSO_3^-				H	**Cc** (*CC*)
WF738C	FR209604[e]	H		Palmitoyl	Ser		H	OSO_3^-					**Cc** (*CC*)
Echinocandin B	A-30912 Factor A, Antibiotic A-22082 SL7810/F			Linoleoyl						Thr			**Ad** (*AE*), **An_1** (*Ani*), **An_2**, **Ap** (*Ecd*/*Hty*), **Ar_1**[d], **Ar_2**, An_M1, An_M2, see also [25]
Echinocandin C	A-30912 Factor B, SL 7810/F-II			Linoleoyl			H			Thr			**Ap** (*Ecd*/*Hty*), **Ar_1**[d], **Ar_2**
Echinocandin D	A-30912 Factor D, SL 7810/F-III	H	H	Linoleoyl			H			Thr			**Ap** (*Ecd*/*Hty*), **Ar_1**[d], **Ar_2**
A-30912 Factor H			OCH_3	Linoleoyl						Thr			**Ap** (Ecd/Hty)
Aculeacin A or A_y				Palmitoyl						Thr			**Aa_1**, **Aa_2**[d], **Aa_3** (*AA*)
Aculeacin A_α				Myristoyl						Thr			Aa_1
Aculeacin D_γ				Palmitoyl			H			Thr			Aa_1
Aculeacin D_α				Myristoyl			H			Thr			Aa_1
Mulundocandin				12-Methylmyristoyl						Ser			Am (*AM*)
Deoxymulundocandin				12-Methylmyristoyl			H			Ser			Am (*AM*)

For the general echinocandin structure with the corresponding residues (R), see Figure 6.

The pneumocandins in the shaded fields have only been identified in mutant strains.

[a]Main products are shown in bold. [b]wild-type strains are shown in bold, [c]suggested name according to common pneumocandin nomenclature (the capital letter is determined by the residue at position 6). [d]Further metabolites of unknown structure have been described: **Aa_2**: aculeacins B–G [27]; **Csp_2**: 41062/F-6, F-7 [46]; **Ar_1**: A-30912 Factors C, E, F, G [34]. [e]For an overview of 'FR' compounds, see Hashimoto [59]; for 'WF' classification, see Hino et al. [42]. [f]The production ratio of these compounds is not fully clear. Metabolite WF14573B was selected as the main product because its biosynthesis requires all genes from the cluster.

footnote d in Table 6). Therefore, it is somewhat speculative to describe metabolic diversity for this group of fungi. In general, the structures of metabolites produced by Aspergillaceae appear to be more consistent. The main products differ, if at all, in the fatty acid side chain or amino acid 5, which is serine instead of threonine in mulundocandins. Based on available data, most byproducts occur through incomplete hydroxylation with hT4H. In echinocandin D, the ornithine residue has not been hydroxylated by OrnH at all. In aculeacin biosynthesis, both palmitic and myristic acids are accepted as fatty acids, resulting in γ- and α-forms of aculeacin, respectively. No variations of threonine 3 and 4-methyl-3-hydroxyproline 6 have been documented for echinocandins from Aspergillaceae. Despite the diversity within echinocandin structures, some structural elements are strictly conserved (Figure 6). It is likely that these are essential for biological activity or required for the conformational stability of the molecule. Most notable among these elements is *trans*-4-hydroxyproline 3, which is fully conserved in all known echinocandin structures. The uptake of *trans*-4-hydroxyproline by A-domain 3 of the NRPS must be strictly substrate specific, since similar building blocks, such as 4-methyl-3-hydroxyproline, proline or *trans*-3-hydroxyproline, are available during biosynthesis.

A second element found in all echinocandins described to date is the fatty acid R^3 attached to the α-amino group of ornithine 1. This is explained by its essential function as an anchor for membrane binding in target cells. In products of clusters equipped with a PKS for biosynthesis of a branched chain fatty acid like dimethylmyristic acid, this is incorporated exclusively; otherwise, specific fatty acids from primary metabolism are utilized. A further conserved element is found in residues 2 and 5. Apart from one exception (cryptocandin), these amino acids have a hydroxyl group at C-3 (threonine, serine, hydroxyglutamine), which, if chiral, has the (*R*)-configuration. In pneumocandins and related structures, residue 5 is occupied exclusively by 3-hydroxyglutamine which is generated by hydroxylation of glutamine with a non-heme GH. The complete absence of glutamine 5 derivatives among pneumocandins and related echinocandins suggests that the hydroxylated amino acid is formed before the specific incorporation by A-domain 5 into the peptide chain. Otherwise, incomplete hydroxylation should lead at least to trace amounts of glutamine 5 variants. This hypothesis, however, lacks experimental evidence. As noted, cryptocandin as the main metabolite of *Cryptosporiopsis* cf. *quercina* is a unique exception among the echinocandins, not only due to the glutamine residue at residue 5,

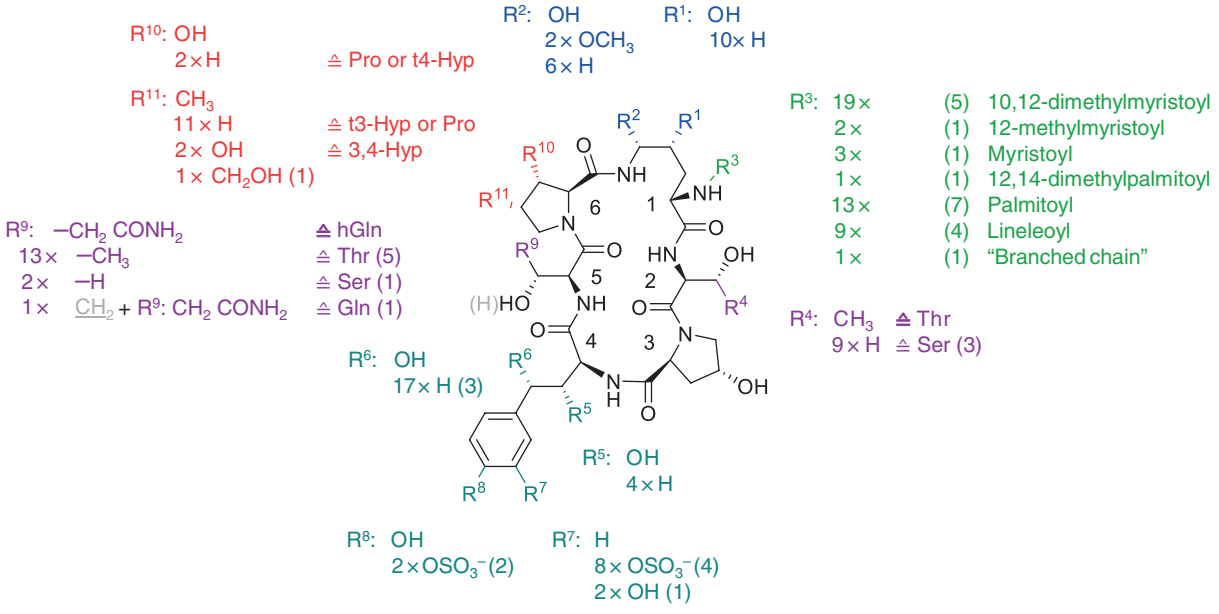

Figure 6: The structures of all characterized echinocandins isolated from wild-type fungi. The conserved echinocandin backbone is shown in black. The colored substituents R^1–R^{11} are varied in at least one echinocandin structure. In the list of substituents at the amino acid side chains, the commonest group is given in bold. For each variation, the number of compounds in which the substituent has been found is given. The number in brackets indicates the number of main metabolites with the substituent. To take into account that some echinocandin structures, such as echinocandin B or pneumocandin A_0, are more widespread in fungal species, each occurrence in a different wild-type species was counted as an individual hit (e.g. pneumocandin A_0 was counted three times, because it has been found in *G. lozoyensis*, *Cryptosporiopsis* sp., and *Pezicula* sp.) [25].

but also because a 4-hydroxymethyl group (R^{11}) occurs in 3-hydroxyproline 6. It would be interesting to determine whether the additional hydroxyl group is introduced by the PH or if an external enzyme is involved. Finally, it is striking that homotyrosine is an integral element of all echinocandins. Incorporations of tyrosine or phenylalanine have yet to be reported.

No variations in the configuration of the up to 17 stereocenters in echinocandins (cf. Figure 1) have been described. Although it is not clear to what extent the configuration of byproducts has been determined, the stereochemistry appears to be strictly conserved.

Of the 11 variable residues R^1–R^{11} in the general structure depicted in Figure 6, only four contain variations in which no oxidation reactions are involved: the fatty acid side chain (R^3), threonine or serine at residue 2 (R^4), homotyrosine sulfate (R^8), and the different amino acid side chains at residue 5 (R^9). The methyl group in proline 6 (R^{11}) is biosynthetically derived from a dioxygenase-catalyzed cyclization of leucine and, thus, the result of an oxidation step. This indicates that metabolic diversity in echinocandin biosynthesis is largely induced by (incomplete) activity of the oxygenases. Structural diversity and especially the generation of byproducts through unselective or incomplete tailoring steps are not limited to echinocandin biosynthesis, but rather are a general feature of secondary metabolism. Nevertheless, the detailed knowledge of echinocandin biosynthesis, at least for some species, provides a basis for a more detailed discussion of the evolutionary background of this phenomenon. Based on current models for secondary metabolite evolution, the significance of metabolic diversity for an organism is discussed in the next section using the example of pneumocandin biosynthesis in *G. lozoyensis*.

5 Comparison of pneumocandin biosynthesis with models for the evolution of secondary metabolism

It is a common phenomenon in secondary metabolism that biosynthetic pathways produce families of closely related compounds [60, 61]. Several reasons for this phenomenon have been discussed [61–68]. For a profound understanding of metabolic evolution, the concrete functions of metabolites in the organism's complex natural environment must be understood. However, these functions are notoriously difficult to investigate. Thus, current knowledge in this field is largely based on more general models explaining the characteristics of secondary metabolite biosynthesis. Two more recent and fundamentally different approaches are introduced briefly. In 1989, Williams et al. proposed a model that emphasized the elaborate biosynthesis pathways for secondary metabolites and their sophisticated modes of action with their targets [68]: the complex structures of such metabolites, which allow an optimal binding to the target, and their elaborate biosyntheses could not have arisen by chance; instead, pathways have evolved in a stepwise manner strictly according to the Darwinian principle, driven by the bioactivity of the product. Only mutations with a positive effect for the organism can prevail and, consequently, secondary metabolites only evolve when this is accompanied by a distinct advantage for the organism. Therefore, for each secondary metabolite a specific physiological function can be expected. Metabolic diversity typically occurs in a late stage of biosynthesis. It has developed in a more recent phase of evolution on the basis of the vast number of hazardous species a microorganism has to cope with in its natural environment. Since each metabolite is supposed to be designed by evolution in order to interact with a specific target, Fischbach and Clardy termed this theory the 'target-based' model [61].

Based on the exploding number of secondary metabolites of unknown function, Firn and Jones introduced a 'diversity-based' model for plant secondary metabolism in 1991 [67]. The idea was further developed [63] and about a decade later Firn and Jones introduced the 'screening' model for secondary metabolism in general [62]. In brief, the likelihood that an accidentally created new metabolite has an advantageous bioactivity for the organism is generally very low. Therefore, the probability that an organism confronted with a threat (e.g. another hazardous organism) has a compound suitable for defense is much higher, from a statistical point of view, when the organism is equipped with a large library of metabolites. To produce such a library at low costs, enzymes with increased substrate promiscuity and reduced substrate selectivity need to be employed. Thereby, metabolic pathways can branch and combine in multiple ways so that a 'matrix grid' of secondary metabolism is formed [62]. Another positive effect of this metabolic network for the organism is an increased metabolic stability. A biosynthetic pathway with strictly substrate-specific enzymes will collapse as soon as one of the enzymes is inactivated or a specific substrate is exhausted. More promiscuous enzymes, however, also accept structurally similar substrates, such as products from other steps in the same biosynthetic route or metabolites from other pathways. Consequently, the production

via a pathway permitting diversity does not necessarily cease when a particular catalytic step breaks down. In summary, diversity-inducing enzymes allow an organism to create metabolic libraries for defense, and can stabilize biosynthetic routes and simplify the evolution of novel biosynthesis pathways. A consequence of this model is that numerous metabolites formed by combinatorial biosynthetic routes can be found in organisms, which have no immediate physiological effects. Both the target-based model and the screening model (also referred to as the diversity-based model [61]) were designed to describe secondary metabolism in its entirety, and not single metabolic pathways such as echinocandin biosynthesis. However, Fischbach and Clardy successfully applied both models in a discussion of the extremely diverse biosynthesis of the gibberellin diterpenoids [61].

With this background, the example of pneumocandin biosynthesis in *G. lozoyensis* (Helotiales), which is by far the best-documented system among the echinocandin producers [2, 7, 9, 17–20] is now examined in an effort to ascertain if one of the models can be applied to echinocandin biosynthesis. During pneumocandin fermentation development at Merck & Co. Research Laboratories, over two dozen different pneumocandins were identified [9], with the chemical structures being reported for 16 of these derivatives. Recently, a total of seven new pneumocandins were isolated, and chemically characterized, from three *G. lozoyensis* mutants (GL_ΔGLP450-1, GL_ΔGLP450-2, GL_ΔGLOXY1; cf. Tables 5 and 6) [19, 20]. The biosynthetic routes of the 23 reported pneumocandins are plotted as lines in a schematic representation of the involved enzymes, according to the current biosynthesis model (Figure 7). This graphic clearly shows that in the earlier steps enzymatic conversions are apparently selective and only very few byproducts are formed: more precisely, byproducts which are incorporated into detectable amounts of pneumocandin derivatives. During the last steps, however, the number of derivatives increases dramatically. This increase begins with the exceptional substrate promiscuity of A-domain 6 in the last module of the NRPS, which accepts at least five different proline derivatives. One of them, *trans*-3-hydroxyproline (*trans*-3-Hyp), originates from the incomplete regioselectivity of the PH and forms the basis of the pneumocandin B family.

In total, not less than eight pneumocandins are finally released from the NRPS. All of them are accepted by the subsequent two cytochrome P450 oxygenases (OrnH and hT4H), which obviously have a relaxed substrate specificity. However, none of the three hydroxylations catalyzed by these enzymes is performed quantitatively. Due to different hydroxylation patterns, the total number of congeners has almost tripled in the end. The combinatorial nature of these transformations suggests that further pneumocandins with other combinations of functional

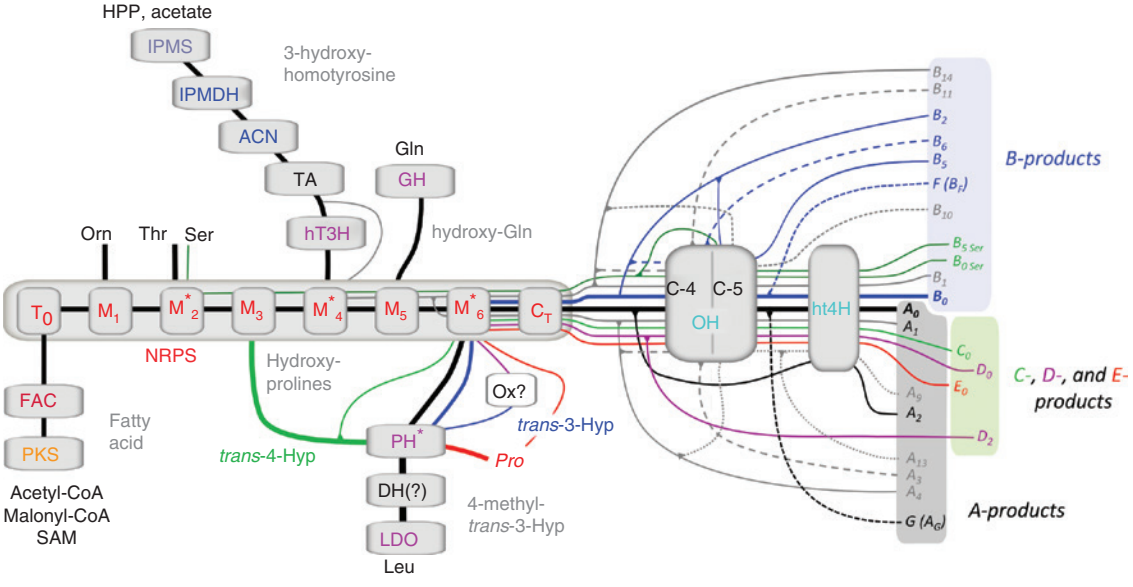

Figure 7: Putative biosynthetic pathways for pneumocandins in *G. lozoyensis*. The grey fields represent the synthetic enzymes involved in echinocandin biosynthesis (for full names of the enzymes and their activity, see Figure 3 and Table 4; the pneumocandin products are defined in Table 6). The modules of the NRPS are depicted as separate active centers (T_0, $M_{1–6}$, and C_T). Each line represents the biosynthesis of a pneumocandin product branching from that of pneumocandins A_0 (black bold line) and B_0 (semibold blue line). Note that some pneumocandins have only been found in mutant strains.

groups should exist; however, their concentrations might be very low, with resulting difficulties in detection and characterization. Interestingly, a blockade of main pathways by deletion of certain oxygenases unveiled several new combinatorial pathways not found in the wild-type strain of *G. lozoyensis* [19, 54]. Considering all data on pneumocandin biosynthesis, elements of both models for secondary metabolism can be identified:

(i) The target-based model: Despite all variations (Figure 7), it can be seen that a sharply defined amino acid backbone is conserved in all echinocandins. (If this is restricted to pneumocandin biosynthesis, this is even more distinct, since there are no variations at R^3, R^7, R^8 and R^9.) The biosynthesis of the echinocandin core structure requires a complex interaction of many diverse enzymes and it can be assumed that this is the result of a long evolutionary process. A key precondition of the target-based model is that all metabolites have a distinct physiological function. In fact, virtually all resulting echinocandins have a pronounced antifungal activity. However, most of them are only produced in very low amounts and the manner in which they are produced does not appear to be targeted nor well controlled. Finally, it is striking that in all known echinocandin biosyntheses, the most oxidized compound is also the main product. In other words, despite all possibilities to omit biosynthetic steps and to generate congeners, the preferred product is that produced by the use of all synthetic enzymes.
Since a targeted biosynthesis of the pneumocandins produced in minor concentrations is rather unlikely, a variation of the target-based model is briefly discussed. This variation postulates that only the main products, pneumocandin A_0 and possibly pneumocandin B_0, have evolved to fit a target; all minor compounds are tolerated as a 'side effect' of biosynthesis. Since these byproducts also exhibit bioactivity, and secondary metabolism generally has a relatively low throughput, the loss of substance is only at little cost, possibly even less than the expression of a fully selective enzymatic machinery. Yet, the metabolic diversity of pneumocandin biosynthesis appears to constitute an evolutionary optimum and, in principle, byproducts could be easily decreased or avoided, probably at minimal cost: A-domain 3 of *G. lozoyensis* NRPS is strictly specific for *trans*-4-hydroxyproline and A-domain 6 of *A. pachycristatus* NRPS is specific for 4-methyl-trans-3-hydroxyproline. This shows that hydroxyprolines can be incorporated specifically. In contrast, A-domain 6 of *G. lozoyensis* NRPS accepts at least five different (hydroxy)prolines. If this was unfavorable, it can be expected that an evolution would have created a substrate-specific domain within a reasonable period. In addition, the amount of partially hydroxylated congeners could be reduced by a simple increase in hydroxylase activity, for instance, via a slight increase in enzyme production.

(ii) The diversity-based model: Whereas production of the echinocandin core structure can be readily explained with the target-based model, there are inconsistencies concerning the diversity-generating late steps of the biosynthesis; rather, the seemingly uncontrolled production of diverse pneumocandins, with the branching and converging pathways forming a complex biosynthetic network, corresponds excellently with the screening or diversity-based model. A key notion of this theory is that organisms producing a large number of secondary metabolites have an advantage in defense against hazardous organisms; however, it is debatable whether this assumption can be applied to the limited diversity provided by a single pathway such as pneumocandin biosynthesis. In principle, it can be surmised that a mixture of antibiotic congeners has a broader applicability against diverse hazardous organisms than a single compound. Nevertheless, most pneumocandins are only produced in very low concentrations. Even more important, they are all derived from a highly specialized core structure, so that their mode of action is probably (but not necessarily) always the same. Thus, a fungus with a fundamental echinocandin resistance will not be affected by any of the derivatives.
The second benefit proposed for an organism with diversity-based secondary metabolism is metabolic flexibility, which also supports evolutionary processes. Thus, a second prediction of the diversity-based model is an increase in metabolic stability resulting from enzymes with reduced substrate specificity. Again, experimental data disclosed by Merck & Co. are instructive. Although all experiments focused on an improved pneumocandin B_0 production and the fermentation conditions were far removed from the natural environment of *G. lozoyensis*, the data provide an interesting insight into the range of variations acceptable for pneumocandin production. Besides variations in the fermentation conditions, excellently reviewed by Connors and Pollard [9], mutants of *G. lozoyensis* were generated by means of chemical mutagenesis [54]. Fermentation experiments with a descendant of pneumocandin B_0 producer strain ATCC 74030 revealed that, in particular, the composition of the medium influenced the metabolite

composition rather than temperature or pH. First, the relaxed substrate specificity of A-domains allows a flexible incorporation of alternative amino acids, depending on their availability in the substrate. Specifically, this was shown by feeding experiments with serine, hydroxyproline and proline [9, 12]. Surprisingly, the feeding of threonine had a slightly inhibitory effect on pneumocandin B_0 biosynthesis (–27%). Other modifications of the medium, such as addition of transition-metal ions or osmotic stress induced by fructose, had a strong impact on pneumocandin production. Notably, the effect differed drastically between individual pneumocandins and, in some cases, the reduced concentration of pneumocandin B_0 was partly offset by an increased titer of a less hydroxylated congener. For example, the production of pneumocandin B_0 was slightly decreased in the presence of nickel and cobalt; however, the production of byproduct pneumocandin B_1 increased three- and fivefold, respectively [10]. The metabolic flexibility in pneumocandin biosynthesis is even more pronounced in deletion mutants. By knockout of the LDO in wild-type *G. lozoyensis*, a mutant was created which was no longer able to produce the main product, pneumocandin A_0 [20]; however, this loss was compensated by a ninefold increase in pneumocandin B_0 production. In cultures with fermentation medium H, the pneumocandin B_0 production strain ATCC 74030, in which the same enzyme is impaired by two point mutations [20], produced even more pneumocandin B_0 than the wild-type production of pneumocandin A_0 [69]. Another mutant derived by classical mutagenesis (ATCC 20958) produced high amounts of the previously unknown pneumocandin A_2 (cf. Table 6) instead of pneumocandin A_0 [54], which can now be explained by an inactivation of OrnH. A derivative strain, mutant ATCC 20988, additionally synthesized pneumocandin A_4 in which both aliphatic hydroxyl groups at the dihydroxyhomotyrosine side chain are missing. More recently, Li et al. reported the knockouts of three oxygenases in wild-type *G. lozoyensis* (ΔhT3H, ΔhT4H, ΔOrnH) [19]. The mutants readily produced the dehydroxypneumocandins (A_1, B_1), (A_G, B_F) and (A_2, B_2), which were expected according to the current model for echinocandin biosynthesis (cf. Figure 3 and Table 6). Notably, from the homotyrosine 3-hydroxylase knockout strain (ΔhT3H), seven additional pneumocandins were isolated, five of which had not been described previously. One compound, pneumocandin A_4, had already been characterized before and is the most reduced pneumocandin main product produced by a *G. lozoyensis* strain (ATCC 20988). Even more degenerate is 'compound 14' (suggested name: pneumocandin B_{14}) isolated from *G. lozoyensis* ΔhT3H. Compared to pneumocandin A_0, its biosynthesis requires nine of 14 synthetic enzymes, including just two of the six oxygenases (PH, GH). Despite the degenerate structure, the antifungal activity of compound 14 was only slightly below average [19].

The strongly increased diversity in the last steps of pneumocandin biosynthesis, largely determined by stochastic processes and the availability of substrates accepted by promiscuous enzymes, is much better explained using the diversity-based model than the targeted oriented approach. However, it is questionable if all compounds synthesized are on standby for screening events, as the diversity-based model implies. As in other complex biosynthesis pathways, only few pneumocandins are produced in considerable concentrations; most are found only in trace amounts. Although a physiological function at such low concentrations should not be generally excluded, it is doubtful that these pneumocandins are potent enough to ensure a reasonable antibiotic activity. Thus, their biosynthesis appears to be more important here, not the final products themselves. In one way, pneumocandin biosynthesis as depicted in Figure 7 resembles a river with tributaries, which forms a large delta of pathways at its end. Only few pathways have a considerable substance flow; however, if the main stream is blocked, for example by inactivation of an enzyme or shortage of a substrate, others are ready to accept the metabolic flux. Consequently, in such a case, pneumocandin biosynthesis is not disrupted, but simply shifted toward a derivative, often without loss in overall production. Such a system not only provides biosynthetic stability in the case of an inactive enzyme, it also allows a maximum flexibility in evolution. Given a hazardous organism against which a minor byproduct has optimal activity, and not the main product, it requires only a few point mutations (possibly only one) or some regulatory effects to increase the byproduct production to effective concentrations. From that perspective, a compound library in the narrow sense is not presented for screening, rather a highly flexible biosynthetic system which is able to modify the product(s) rapidly through minor evolutionary events. Even so, some findings are not fully consistent with a diversity-based model. First, as mentioned before, in all wild-type echinocandin-producer strains analyzed so far, the main product is also the most oxidized metabolite, whose biosynthesis requires

all synthetic enzymes. Thus, the genetic equipment of the clusters is focused on the production of a defined product. Furthermore, the diversity among echinocandin main products is limited, which suggests that a few structures are privileged by evolution. Second, there is no explanation as to why, for example, the incorporation of hydroxyproline 6 is very promiscuous in *G. lozoyensis*, while being very strict in other species. For instance, echinocandin production in *A. pachycristatus* breaks down when methylproline biosynthesis is disrupted [16], which clearly shows that module 6 of the NRPS is strictly specific for 4-methyl-3-hydroxyproline.

In summary, there is a remarkable structural diversity among the pneumocandins induced by distinct biocatalytic steps, which, for example, convert more than one substrate, generate more than one product from a single substrate or are simply omitted. Meanwhile, there are many steps that do not contribute to structural diversity and thus allow the setup of the highly conserved echinocandin backbone. Although these observations are not unusual for biosynthetic routes of secondary metabolites, the evolutionary models discussed here can only account for parts of pneumocandin biosynthesis. In particular, the diversity-based model provides a fundamental explanation for the otherwise barely reasonable formation of many pneumocandin congeners, even though this approach originally referred to secondary metabolism in general rather than specific pathways. Currently, there seems to be no theoretical model that combines the selective biosynthesis of core structures with the emerging diversity induced by certain enzymes typically involved into late 'tailoring' steps. A critical limiting point in the discussion of secondary metabolite evolution in greater depth and the development of new theoretical concepts is the poor knowledge of the actual function of such metabolites in the natural environment. Mostly, this is completely unknown; sometimes, as in the case of echinocandins, a pronounced bioactivity has been identified. The utilization of a secondary metabolite by its producer in the natural habitat, however, has been documented only in very few cases. Nevertheless, detailed information on the chemical ecology of an organism is necessary to draw more specific conclusions on the evolution of its secondary metabolites.

Although a number of questions are still open, pneumocandin biosynthesis belongs to the best studied in fungi. As the bioactivity of the products can be well examined, it constitutes an excellent model system for future studies on ecological aspects, which can also help to understand better the general principles of secondary metabolism.

Acknowledgment: I would like to thank Prof. Michael Müller for critical discussions and for reading the manuscript.

Declaration of interest: The author reports no declarations of interest.

References

1. Stan CD, Tuchilus C, Stan CI. Echinocandins – new antifungal agents. Rev Med Chir Soc Med Nat Iasi 2014;118:528–36.
2. Balkovec JM, Hughes DL, Masurekar PS, Sable CA, Schwartz RE, Singh SB. Discovery and development of first in class antifungal caspofungin (Cancidas®) – a case study. Nat Prod Rep 2014;31:15–34.
3. Emri T, Majoros L, Tóth V, Pócsi I. Echinocandins: production and applications. Appl Microbiol Biotechnol 2013;97:3267–84.
4. Groll AH, Schrey D, Walsh TJ. Echinocandins. In: Kauffman CA, Pappas PG, Sobel JD, Dismukes WE, editors. Essentials of clinical mycology. Springer New York, 2011:95–112.
5. Perlin DS. Mechanisms of echinocandin antifungal drug resistance. Ann NY Acad Sci 2015;1354:1–11.
6. Walker LA, Gow NA, Munro CA. Fungal echinocandin resistance. Fungal Genet Biol 2010;47:117–26.
7. Yue Q, Chen L, Zhang X, Li K, Sun J, Liu X, et al. Evolution of chemical diversity in echinocandin lipopeptide antifungal metabolites. Eukaryot Cell 2015;14:698–718.
8. Bills G, Li Y, Chen L, Yue Q, Niu X-M, An Z. New insights into the echinocandins and other fungal non-ribosomal peptides and peptaibiotics. Nat Prod Rep 2014;31:1348–75.
9. Connors N, Pollard D. Pneumocandin B_0 production by fermentation of the fungus *Glarea lozoyensis*. In: An Z, editor. Handbook of industrial mycology. Boca Raton, FL: CRC Press, 2004:515–38.
10. Adefarati AA, Hensens OD, Jones ETT, Tkacz JS. Pneumocandins from *Zalerion arboricola*. V. Glutamic acid-derived and leucine-derived amino-acids in pneumocandin A_0 (L-671,329) and distinct origins of the substituted proline residues in pneumocandins A_0 and B_0. J Antibiot 1992;45:1953–57.
11. Adefarati AA, Giacobbe RA, Hensens OD, Tkacz JS. Biosynthesis of L-671,329, an echinocandin-type antibiotic produced by *Zalerion arboricola* – origins of some of the unusual amino-acids and the dimethylmyristic acid side-chain. J Am Chem Soc 1991;113:3542–45.
12. Petersen LA, Hughes DL, Hughes R, DiMichele L, Salmon P, Connors N. Effects of amino acid and trace element supplementation on pneumocandin production by *Glarea lozoyensis:* impact on titer, analogue levels, and the identification of new analogues of pneumocandin B_0. J Ind Microbiol Biotechnol 2001;26:216–21.
13. Cacho RA, Jiang W, Chooi Y-H, Walsh CT, Tang Y. Identification and characterization of the echinocandin B biosynthetic gene cluster from *Emericella rugulosa* NRRL 11440. J Am Chem Soc 2012;134:16781–90.
14. Hüttel W, Youssar L, Grüning BA, Günther S, Hugentobler KG. Echinocandin B biosynthesis: a biosynthetic cluster from *Aspergillus nidulans* NRRL 8112 and reassembly of the subclusters *ecd* and *hty* from *Aspergillus pachycristatus* NRRL 11440 reveals a single coherent gene cluster. BMC Genomics 2016;17:570.

15. Samson RA, Visagie CM, Houbraken J, Hong SB, Hubka V, Klaassen CHW, et al. Phylogeny, identification and nomenclature of the genus *Aspergillus*. Stud Mycol 2014;78:141–73.
16. Jiang W, Cacho RA, Chiou G, Garg NK, Tang Y, Walsh CT. EcdGHK are three tailoring iron oxygenases for amino acid building blocks of the echinocandin scaffold. J Am Chem Soc 2013;135:4457–66.
17. Chen L, Yue Q, Zhang X, Xiang M, Wang C, Li S, et al. Genomics-driven discovery of the pneumocandin biosynthetic gene cluster in the fungus *Glarea lozoyensis*. BMC Genomics 2013;14:339.
18. Houwaart S, Youssar L, Hüttel W. Pneumocandin biosynthesis: involvement of a *trans*-selective proline hydroxylase. ChemBioChem 2014;15:2365–69.
19. Li Y, Chen L, Yue Q, Liu X, An Z, Bills GF. Genetic manipulation of the pneumocandin biosynthetic pathway for generation of analogues and evaluation of their antifungal activity. ACS Chem Biol 2015;10:1702–10.
20. Chen L, Yue Q, Li Y, Niu X, Xiang M, Wang W, et al. Engineering of *Glarea lozoyensis* for exclusive production of the pneumocandin B_0 precursor of the antifungal drug caspofungin acetate. Appl Environ Microbiol 2015;81:1550–8.
21. Youssar L, Grüning BA, Erxleben A, Günther S, Hüttel W. Genome sequence of the fungus *Glarea lozoyensis*: the first genome sequence of a species from the helotiaceae family. Eukaryot Cell 2012;11:250.
22. Hibi M, Mori R, Miyake R, Kawabata H, Kozono S, Takahashi S, et al. Novel enzyme family found in filamentous fungi catalyzing *trans*-4-hydroxylation of L-pipecolic acid. Appl Environ Microbiol 2016;82:2070–7.
23. Hoff KJ, Stanke M. WebAUGUSTUS – a web service for training AUGUSTUS and predicting genes in eukaryotes. Nucleic Acids Res 2013;41:W123–W28.
24. Nordberg H, Cantor M, Dusheyko S, Hua S, Poliakov A, Shabalov I, et al. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res 2014;42:D26–D31.
25. de la Cruz M, Martín J, González-Menéndez V, Pérez-Victoria I, Moreno C, Tormo JR, et al. Chemical and physical modulation of antibiotic activity in Emericella species. Chem Biodivers 2012;9:1095–113.
26. Otani M, Satoi S, Takaka M. Antibiotic aculeacin – A. Alpha,-A. Gamma,-D.Alpha and -D.Gamma and their production, 1978. Canada Patent CA 1110561, 1978-04-18.
27. Satoi S, Yagi A, Asano K, Mizuno K, Watanabe T. Studies on aculeacin. II. Isolation and characterization of aculeacins B, C, D, E, F and G. J Antibiot 1977;30:303–7.
28. Mizuno K, Yagi A, Satoi S, Takada M, Hayashi M. Studies on aculeacin. I. Isolation and characterization of aculeacin A. J Antibiot 1977;30:297–302.
29. Dreyfuss MM, Tscherter H. Antibiotic s 31794/F-1, 1979. US Patent 4173629, 1978-06-28.
30. Higgens CE, Michel KH. Antibiotic A-22082 and process for production thereof, 1977. US Patent 4024246, 1977-08-25.
31. Zou SP, Zhong W, Xia CJ, Gu YN, Niu K, Zheng YG, et al. Mutagenesis breeding of high echinocandin B producing strain and further titer improvement with culture medium optimization. Bioprocess Biosyst Eng 2015;38:1845–54.
32. Benz F, Knüsel F, Nüesch J, Treichler H, Voser W, Nyfeler R, et al. Stoffwechselprodukte von mikroorganismen 143. Mitteilung. Echinocandin B, ein neuartiges polypeptid-antibioticum aus *Aspergillus nidulans* var. *Echinulatus*: Isolierung und bausteine. Helv Chim Acta 1974;57:2459–77.
33. Boeck LVD, Kastner RE. Method of producing the A-30912 antibiotics, 1981. US Patent 4288549, 1981-09-08.
34. Hoehn MM, Michel KH. Antibiotic A-30912, 1977. US Patent 4024245, 1977-05-17.
35. Traber R, Keller-Juslén C, Loosli H-R, Kuhn M, Von Wartburg A. Cyclopeptid-antibiotika aus *Aspergillus*-arten. Struktur der echinocandine C und D. Helv Chim Acta 1979;62:1252–67.
36. Keller-Juslén C, Kuhn M, Loosli HR, Petcher TJ, Weber HP, von Wartburg A. Struktur des cyclopeptid-antibiotikums sl 7810 (= echinocandin B). Tetrahedron Lett 1976;17:4147–50.
37. Bills GF, Yue Q, Chen L, Li Y, An Z, Frisvad JC. *Aspergillus mulundensis* sp. Nov., a new species for the fungus producing the antifungal echinocandin lipopeptides, mulundocandins. J Antibiot 2016;69:141–48.
38. Roy K, Mukhopadhyay T, Reddy GC, Desikan KR, Ganguli BN. Mulundocandin, a new lipopeptide antibiotic. I. Taxonomy, fermentation, isolation and characterization. J Antibiot 1987;40:275–80.
39. Mukhopadhyay T, Ganguli BN, Fehlhaber HW, Kogler H, Vertesy L. Mulundocandin, a new lipopeptide antibiotic. II. Structure elucidation. J Antibiot 1987;40:281–9.
40. Kanasaki R, Sakamoto K, Hashimoto M, Takase S, Tsurumi Y, Fujie A, et al. Fr209602 and related compounds, novel antifungal lipopeptides from *Coleophoma crateriformis* no. 738. J Antibiot 2006;59:137–44.
41. Iwamoto T, Fujie A, Sakamoto K, Tsurumi Y, Shigematsu N, Yamashita M, et al. Wf11899a, wf11899b and wf11899c, novel antifungal lipopeptides 1. Taxonomy, fermentation, isolation and physicochemical properties. J Antibiot 1994;47:1084–91.
42. Hino M, Fujie A, Iwamoto T, Hori Y, Hashimoto M, Tsurumi Y, et al. Chemical diversity in lipopeptide antifungal antibiotics. J Ind Microbiol Biotechnol 2001;27:157–62.
43. Kanasaki R, Abe F, Kobayashi M, Katsuoka M, Hashimoto M, Takase S, et al. Fr220897 and fr220899, novel antifungal lipopeptides from *Coleophoma empetri* no. 14573. J Antibiot 2006;59:149–57.
44. Strobel GA, Miller RV, Martinez-Miller C, Condron MM, Teplow DB, Hess WM. Cryptocandin, a potent antimycotic from the endophytic fungus *Cryptosporiopsis* cf. *quercina*. Microbiology 1999;145:1919–26.
45. Noble HM, Langley D, Sidebottom PJ, Lane SJ, Fisher PJ. An echinocandin from an endophytic *Cryptosporiopsis* sp. And *Pezicula* sp. In: *Pinus sylvestris* and *Fagus sylvatica*. Mycol Res 1991;95:1439–40.
46. Dreyfuss MM, Tscherter H. New metabolites, processes for their production and their use, 1982. WIPO Patent WO 1982000587, 1982-03-04.
47. Dreyfuss M. Neue Erkenntnisse aus einem pharmakologischen Pilz-screening. Sydowia 1986;39:22–36.
48. Balkovec JM. Section review: Anti-infectives: lipopeptide antifungal agents. Expert Opin Invest Drugs 1994;3:65–82.
49. Kanasaki R, Kobayashi M, Fujine K, Sato I, Hashimoto M, Takase S, et al. Fr227673 and fr190293, novel antifungal lipopeptides from *Chalara* sp. No. 22210 and *Tolypocladium parasiticum* no. 16616. J Antibiot 2006;59:158–67.
50. Peláez F, Collado J, Platas G, Overy DP, Martín J, Vicente F, et al. Phylogeny and intercontinental distribution of the pneumocandin-producing anamorphic fungus *Glarea lozoyensis*. Mycology 2011;2:1–17.

51. Bills GF, Platas G, Peláez F, Masurekar P. Reclassification of a pneumocandin-producing anamorph, *Glarea lozoyensis* gen. Et sp. Nov., previously identified as *Zalerion arboricola*. Mycol Res 1999;103:179–92.
52. Morris S, Schwartz R, Sesin D, Masurekar P, Hallada T, Schmatz D, et al. Pneumocandin D_0, a new antifungal agent and potent inhibitor of *pneumocystis carinii*. J Antibiot 1994;47:755–64.
53. Schwartz RE, Sesin DF, Joshua H, Wilson KE, Kempf AJ, Goklen KA, et al. Pneumocandins from *Zalerion arboricola*. I. Discovery and isolation. J Antibiot 1992;45:1853–66.
54. Masurekar PS, Fountoulakis JM, Hallada TC, Sosa MS, Kaplan L. Pneumocandins from *Zalerion arboricola*. II. Modification of product spectrum by mutation and medium manipulation. J Antibiot 1992;45:1867–74.
55. Schwartz RE, Giacobbe RA, Bland JA, Monaghan RL. L-671,329, a new antifungal agent. 1. Fermentation and isolation. J Antibiot 1989;42:163–67.
56. Hu Z-C, Peng L-Y, Zheng Y-G. Enhancement of echinocandin B production by a UV- and microwave-induced mutant of *Aspergillus nidulans* with precursor- and biotin-supplying strategy. Appl Biochem Biotechnol 2016;179:1–14.
57. Xu J, Chen Y, Ji X, Gao X, Liu S, Zhang Z. High yield antibiotics producing fungus strain, preparation method and use thereof, 2012, US Patent 20120258498, 11-10-2012.
58. Haifeng H, Lijun W, et al. Method for breeding high-WF16616-yield *Tolypocladium parasiticum* mutant strain, 2013. China Patent CN 103087937, 2013-05-08.
59. Hashimoto S. Micafungin: a sulfated echinocandin. J Antibiot 2009;62:27–35.
60. Roberts TC, Smith PA, Romesberg FE. Synthesis and biological characterization of arylomycin B antibiotics. J Nat Prod 2011;74:956–61.
61. Fischbach MA, Clardy J. One pathway, many products. Nat Chem Biol 2007;3:353–55.
62. Firn RD, Jones CG. Natural products? A simple model to explain chemical diversity. Nat Prod Rep 2003;20:382–91.
63. Firn RD, Jones CG. The evolution of secondary metabolism – a unifying model. Mol Microbiol 2000;37:989–94.
64. Demain A, Fang A. The natural functions of secondary metabolites. In: Fiechter A, editor. History of modern biotechnology I. Berlin, Heidelberg: Springer, vol. 69, 2000:1–39.
65. Vining LC. Secondary metabolism, inventive evolution and biochemical diversity – a review. Gene 1992;115:135–40.
66. Stone MJ, Williams DH. On the evolution of functional secondary metabolites (natural products). Mol Microbiol 1992;6:29–34.
67. Jones CG, Firn RD, Malcolm SB. On the evolution of plant secondary chemical diversity [and discussion]. Philos Trans R Soc B 1991;333:273–80.
68. Williams DH, Stone MJ, Hauck PR, Rahman SK. Why are secondary metabolites (natural products) biosynthesized? J Nat Prod 1989;52:1189–208.
69. Tkacz J, Giacobbe R, Monaghan R. Improvement in the titer of echinocandin-type antibiotics: a magnesium-limited medium supporting the biphasic production of pneumocandins A_0 and B_0. J Ind Microbiol 1993;11:95–103.