HHS Public Access

Author manuscript

J Am Chem Soc. Author manuscript; available in PMC 2014 March 20.

Published in final edited form as:

J Am Chem Soc. 2013 March 20; 135(11): 4457–4466. doi:10.1021/ja312572v.

EcdGHK are Three Tailoring Iron Oxygenases for Amino Acid Building Blocks of the Echinocandin Scaffold

Wei Jiang[*,1], **Ralph A Cacho**[*,2], **Grace Chiou**[3], **Neil K Garg**[3], **Yi Tang**[2,3], and **Christopher T Walsh**[1]

[1]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 02115

[2]Department of Chemical and Biochemical Engineering, University of California, Los Angeles, CA 90095

[3]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095

Abstract

The echinocandins are a small group of fungal N-acylated cyclic hexapeptides that are fungicidal for candida strains and fungistatic for aspergilli by targeting cell wall 1,3-β-glucan synthases. The side chains of all six amino acid building blocks have hydroxyl groups, including the nonproteinogenic 4*R*,5*R*-dihydroxy-Orn$_1$, 4*R*-OH-Pro$_3$, 3*S*, 4*S*-dihydroxy-homoTyr$_4$ and 3*S*-OH-4*S*-Me-Pro$_6$. The echinocandin (*ecd*) gene cluster contains two predicted nonheme mononuclear iron oxygenase genes (*ecdG,K*) and one encoding a P450 type heme protein (*ecdH*). Deletion of the *ecdH* gene in the producing *Emericella rugulosa* generates an echinocandin scaffold (echinocandin D) lacking both hydroxyl groups on Orn$_1$. Correspondingly, the Δ*ecdG* strain failed to hydroxylate C$_3$ of the homoTyr residue, and purified EcdG hydroxylated free L-homoTyr at C$_3$. The Δ*ecdK* strain failed to generate mature echinocandin unless supplemented with either 4*R*-Me-Pro or 3*S*-OH-4*S*-Me-Pro, indicating blockage of a step upstream of Me-Pro formation. Purified EcdK is a Leu 5-hydroxylase, acting iteratively at C$_5$ to yield γ-Me-Glu-γ-semialdehyde in equilibrium with the cyclic imine product. Evaluation of deshydroxyechinocandin scaffolds in *in vitro* anticandidal assays revealed up to 3-fold loss of potency for the Δ*ecdG* scaffolds, but a 3-fold gain of potency for the Δ*ecdH* scaffold, in line with prior results on deoxyechinocandin homologs.

INTRODUCTION

The echinocandins are a small group of fungal lipopeptides that act to kill other fungi by targeting the 1,3-glucan synthase that builds the 1,3-glucans that are key structural components of fungal cell walls. [1,2] Three semisynthetic versions (with the natural fatty acyl chains replaced by synthetic hydrophobic groups) of the naturally occurring pneumocandin

Corresponding Authors: Christopher_walsh@hms.harvard.edu, Yitang@ucla.edu.

*These authors contributed equally to this work.

SUPPORTING INFORMATION AVAILABLE

Additional NMR parameters and spectra, primer list, PCR screening for KO confirmation, and chromatographic traces are found in the supplemental information. This material is available free of charge via the Internet at http://pubs.acs.org.

or echinocandin scaffolds have been commercialized and approved for human clinical use as fungicidal and fungistatic agents for the treatment of candidiasis and aspergillosis, respectively. [3,4,5]

For naturally occurring echinocandin B (1, Figure 1) the lipid is the di-olefinic linoleic acid. It has an amide linkage to a modified L-ornithine residue that is part of a cyclic hexapeptide framework (Figure 1). Four of the six residues in the hexapeptide framework are nonproteinogenic: only Thr_2 and Thr_5 are proteinogenic amino acids. The 4*R*,5*R*-dihydroxy-L-Orn_1, 4*R*-OH-Pro_3, 3*S*,4*S*-dihydroxy-homotyrosine$_4$, and 3*S*-OH-4*S*-methyl-Proline$_6$ (3-OH-Me-Pro) residues are typically not found in proteins (although collagen has hydroxyproline residues introduced by posttranslational hydroxylation). Echinocandin C (**2**) has the C_3-OH but lacks the C_4-OH on the homoTyr residue, while echinocandin D (**3**) is a trideoxy form that has the 3-OH-$HomoTyr_4$ but a nonhydroxylated Orn_1. [6]

The unusual building block inventory suggested that echinocandins and the closely related pneumocandin are constructed by fungal nonribosomal peptide synthetase (NRPS) assembly lines. This expectation has been recently validated by the identification of a six module NRPS biosynthetic gene *ecdA* in *A. nidulans var roseus* (recently renamed *Emericella rugulosa*) [7] whose disruption abolishes production of **1**. [8] Among the coding sequences in the vicinity of the *ecdA* gene are *ecdG*, *ecdH*, *ecdK*, two of which are predicted to be α-ketoglutarate-dependent mononuclear, nonheme iron oxygenases genes (*ecdG*, *ecdK*), with the third predicted to be a P450 type of hemeprotein monooxygenase gene (*ecdH*) (Figure 1). Given that all six of the echinocandin B amino acid residues have at least one side chain hydroxyl group (Orn_1 and $HomoTyr_4$ each have two such −OH groups), it is likely that these oxygenases may fulfill tailoring functions at some stage to create the hydrophilic surface of the mature echinocandin B antifungal agent. In addition, it is likely that the Me-Pro framework at residue 6 arises from the proteinogenic amino acid building block L-leucine by hydroxylation at C_5, [9] suggesting another, albeit cryptic, step for hydroxylase action in echinocandin producers.

To address the nature of the involvement of *ecdG*, *ecdH*, and *ecdK* in the maturation of the lipopeptide scaffold of echinocandin B, we report the gene disruptions in each of these three genes and analysis by mass spectrometry and NMR of the deshydroxy forms of echinocandin that accumulate in the media in such fermentations. Studies with purified EcdG and EcdK confirm regiospecific oxygenase action on L-homoTyr and L-Leu, respectively. The results allow assessment of the specificity of each of the three oxygenases and the consequences for biological activity (antifungal action) as the nascent echinocandin scaffold is progressively matured.

RESULTS

Echinocandin Production from wild type *E. rugulosa*

In preparation for analysis of the echinocandin forms produced in gene deletion mutants, we first examined the forms of echinocandin generated by the native *E. rugulosa* producer under echinocandin production conditions. [10] In addition to the fully hydroxylated **1** (*m/z* calcd for $C_{52}H_{81}N_7O_{16}$ $[M+H]^+[-H_2O]$ 1042.5707, found 1042.5700), there is a monodeoxy

form **2** (= echinocandin C) noted previously by earlier workers [6] that corresponds to the lack of a hydroxyl group at the 4-position of the homoTyr residue (*m/z* calcd for $C_{52}H_{81}N_7O_{15}$ $[M+H]^+[-H_2O]$ 1026.5758, found 1026.5757) (Figure 2A). Presumably, conversion of homoTyr to 3,4-dihydroxy-homoTyr, either as the free monomer or as part of a hexapeptide intermediate, is incomplete and either occurs late in maturation or does not prevent hydroxylation at other sites during maturation. The presence of **2** along with **1** suggests hydroxylation at the 3-position of $homoTyr_4$ can occur before hydroxylation at C_4.

Echinocandin Product Profile from a Δ*ecdH Emericella rugulosa* strain

Construction and characterization of the Δ*ecdH* strain is shown in Fig S1 and is representative of the approaches used also for the Δ*ecdG* and Δ*ecdK* genes. Deletion mutants in each of the *ecdGHK* genes were screened for insertion of *bar* knockout cassettes followed by PCR validation of *bar* gene insertion into the desired *ecd* gene. [11,12] Mutant cultures were grown under echinocandin production conditions, [10] compounds were extracted and evaluated by MS and NMR.

The Δ*ecdH* strain loses production of **1** and **2** while yielding dideoxy (*m/z* calcd for $C_{52}H_{81}N_7O_{14}$ $[M+H]^+$ 1028.5914; found 1028.5901) and trideoxy forms (*m/z* calcd for $C_{52}H_{81}N_7O_{13}$ $[M+H]^+$ 1012.5965; found 1012.5965) of **1** (Figure 2B). The trideoxy compound is the major component with yields ~10 mg from 8 L culture. NMR characterization reveals the loss of hydroxyl group at 4-$homoTyr_4$ and both of the 4- and 5-OH of Orn_1 (Table S1). This is identical to echinocandin D, compound **3**, which is one of the naturally occurring minor echinocandins (Table S2). [6,13] The yield of dideoxy compound **4** from Δ*ecdH* was too low for NMR assignment. We thus carried out feeding studies with deuterated amino acid monomers to evaluate the hydroxylation state at certain residues. 10 mg of 3,3-4,4-5,5-perdeutero-(D_6)-L-ornithine was used in 10 mL culture to assess the hydroxylation of the Orn_1 residue. **1** and **2** extracted from the WT strain exhibits an M+4 parent ion (Figure S2), consistent with loss of a deuterium at each of C_4 and C_5 as 4-OH, 5-OH-L-Orn is incorporated into the antibiotic framework. Variants **3** and **4** isolated from Δ*ecdH* strain fed with (D_6)-L-Orn, on the other hand, both had prominent peaks at M+6, indicating that none of the Orn side chain deuterium atoms had been lost (Figure 2C). Therefore, compound **4** must retain all the hydroxylations except C_4 and C_5 at Orn_1.

We also tested feeding with perdeuteroproline (D_7) or perdeuteroTyr (D_7) as a parallel control. While no incorporation was observed in rich medium, in minimal medium, echinocandins were still produced and a low but detectable level of (D_6)-L-proline incorporation (loss of one deuterium at C_4 during conversion to 4-OH-Pro) was obtained in **1** from the native producer (Figure S3). Compound **3** from the Δ*ecdH* strain exhibited the same M+6 parent ion (Figure S3), indicating no structural difference between them on the Pro_4 residue. On the other hand, there was still no detectable incorporation of perdeutero-Tyr. Proline is not expected to be a precursor to 3*S*-OH-4*S*-Me-L-Pro_6 since precedent suggests a route from leucine, [14] and accordingly no incorporation of deuterium from (D_6)-L-proline is found at that residue.

Thus, one can unambiguously assign the hemeprotein P450 EcdH as the Orn oxygenase acting to hydroxylate two adjacent carbons. The *ecdH* gene knockout studies are silent on two aspects of the timing of C_4 vs C_5 hydroxylation of Orn_1: (a) whether the free amino acid undergoes hydroxylation or only the Orn residue after incorporation into the nascent peptide is oxygenated; and (b) whether C_4 or C_5 is oxygenated first. Efforts to isolate soluble EcdH from *E. coli* expression were unsuccessful (overproduction yielded insoluble inactive protein). Because EcdH may be membrane-associated, expression in yeast microsomes was conducted. Neither free L-Orn nor the trideoxy form **3** of echinocandin was detectably oxygenated on incubation with those microsomal preparations, so no conclusion can be drawn about point (a).

However, given that C_5 hydroxylation of Orn_1 in mature **1** is a key feature of the hemiaminal linkage, one expects this would not have happened before Orn was incorporated into the echinocandin peptide backbone. Otherwise a free C_5-OH-L-Orn tetrahedral addition compound would unravel to the aldehyde and in that oxidation state will not be able to form the amide bond with 3-OH-Me-Pro_6. The relative timing of C_5 vs C_4 hydroxylation can be intuited in the Δ*ecdG* studies described below, from the trideoxy echinocandin **7** NMR where the C_5 of Orn has been hydroxylated but C_4 has not. In aggregate, the results suggest EcdH acts on Orn only after incorporation into the nascent echinocandin chain and acts first at C_5, then at C_4. There are precedents for a single P450 enzyme acting iteratively on its substrate. Notably, the P450 side chain cleavage enzyme in the adrenal steroid hormone pathway converts cholesterol to 22*R*-OH-cholesterol, then to the 20*R*, 22*R*-diol, before a third $P450_{scc}$ oxygenative step cleaves between the diol to yield pregnenolone and isocaproic aldehyde. [15]

Echinocandin Product Profile from a Δ*ecdG Emericella rugulosa* strain

The Δ*ecdG* strain (Figure S4) gave a more complex echinocandin metabolite profile, producing a mixture of mono- (**5**), di- (**6**), tri- (**7**) and tetra-deoxy (**8**) compounds, with **6** and **7** as the major components (Figure 3A). Both of them were successfully isolated (4.0 mg **6** and 5.5 mg **7** from 8 L culture) and characterized by NMR. The dideoxy form **6** (*m/z* calcd for $C_{52}H_{81}N_7O_{14}$ $[M+H]^+$ -$[H_2O]$ 1010.5809; found 1010.5819) is missing both hydroxyl groups on 3- and 4-$homoTyr_4$ (Table S1). The trideoxy **7** (*m/z* calcd for $C_{52}H_{81}N_7O_{13}$ $[M+H]^+$ -$[H_2O]$ 994.5859; found 994.5826), in addition to missing two -OH groups on $homoTyr_4$, also is lacking an -OH at C_4 of Orn_1 (Table S1). Consistent with NMR analysis, feeding with (D_6)-L-Orn to the Δ*ecdG* strain led to M+4 and M+5 shifts in compound **6** and **7**, respectively (Figure 3B).

The yield of **8** (*m/z* calcd for $C_{52}H_{81}N_7O_{12}$ $[M+H]^+$ 996.6016; found 996.6035) was too low for NMR analysis, however, its structure can be deduced from feeding results. It exhibits an M+6 peak upon D_6-L-Orn addition, indicating both hydroxyl groups on Orn_1 are absent. Together with the confirmed structure of **6** and **7**, it is reasonable to propose that Orn_1 and $homoTyr_4$ side chains in the tetradeoxy compound **8** are completely unmodified. Consistent with this scheme, **8** shows an M+6 pattern upon (D_7)-L-proline feeding (Figure S3), indicating hydroxylation on the proline moiety. As expected, **7** presents the same M+6 pattern on the proline residue as well (the incorporation level in **5** and **6** is too low to be

identified). The presence of all three variations of the Orn_1 moiety (deshydroxy in **8**, monohydroxy in **7**, dihydroxy in **5** and **6**) suggests that the hydroxylation of Orn does not *require* EcdG but is *affected* upon the absence of EcdG. This could reflect altered conformation of the hexapeptide macrocycle when particular-OH groups are missing. Therefore, the target of EcdG, as suggested by the missing hydroxyl groups on homoTyr in **6** and **7**, should be homoTyr. Because only one hydroxyl group is missing in compound **5**, we expect it to be the specific hydroxyl group (either 3OH- or 4OH on $homoTyr_4$) installed by EcdG. Resolving the structure of **5** would be one key to this question, however, the isolated quantity was too low for detailed NMR analysis. Since 4-OH on $homoTyr_4$ is not ubiquitous in the echinocandins obtained from WT and gene deletion mutants, we favored EcdG as a homotyrosine 3-hydroxylase, and the *in vitro* assay described next confirmed it.

Purified EcdG is a Homotyrosine 3-Hydroxylase

The mononuclear nonheme iron enzyme EcdG could be expressed and purified in soluble form from *E. coli* (Figure 4A) with a yield of ~30 mg/L. Assay with the amino acid building block L-homoTyr (Figure 4B) followed by Fmoc derivatization revealed that EcdG acts to introduce one hydroxyl group on free L-homoTyr (Fmoc-homoTyr: *m/z* calcd for $C_{25}H_{23}NO_5$ $[M+Na]^+$ 440.1468, found 440.1460; Fmoc-hydroxy-homoTyr: *m/z* calcd for $C_{25}H_{23}NO_6$ $[M+Na]^+$ 456.1418, found 456.1411). As has been seen with other nonheme iron oxygenases that generate high valent oxoiron intermediates, [16,17] EcdG inactivates itself after a small number of turnovers. Variation of the amount of pure EcdG relative to excess substrate indicates a proportional amount of product formation (Figure 4C) and indicates that about 3-3.5 turnovers occur per inactivation event. The estimated initial rate of hydroxylation is higher than 5 catalytic events per minute.

Although only 3-3.5 oxygenation events per EcdG enzyme molecule occur before enzyme suicide, it was possible to scale up the monohydroxy homoTyr product and separate it from unreacted L-homoTyr by Hypercarb column using 20 mM perfluoropentanoic acid / acetonitrile as mobile phase (Figure 4D, homoTyr: *m/z* calcd for $C_{10}H_{13}NO_3$ $[M+H]^+$ 196.0968, found 196.0975; hydroxy-homoTyr: *m/z* calcd for $C_{10}H_{13}NO_4$ $[M+H]^+$ 212.0917, found 212.0925). Proton NMR and gCOSY clearly indicated the enzymatic product is a 3-OH-homoTyr (Figure S5 and S6) and correlates with the chemical shifts reported for synthetic (2*S*, 3*R*)-OH-homoTyr. [18]. We have not determined the absolute stereochemical configuration at the C3-OH of the EcdG *in vitro* product but anticipate it will be 3*R*.

Pure EcdG showed no activity with the deshydroxy echinocandin scaffold isolated from the Δ*ecdG* mutant strain (**6** & **7**), confirming that it does not act after the NRPS assembly line has produced the hexapeptide framework. Moreover, EcdG also will not take the 3-OH-homoTyr product and convert it to the 3,4-dihydroxy amino acid (data not shown), making it unlikely to be the hydroxylase for introducing the 4-OH group in addition to the hydroxylation at C3. Therefore, the prediction is that the adenylation domain in module 4 of the six module EcdA would activate 3-OH-homoTyr (or 3,4-dihydroxy-homoTyr but less likely, see discussion below) for incorporation into the growing echinocandin chain.

Deletion of *ecdK* abolishes production of echinocandins

Given the above results from *ecdG* and *ecdH* deletion strains, we anticipated that the remaining putative oxygenase, the nonheme mononuclear iron protein EcdK, could act either on Pro_3, 4*R*-Me-Pro_6, or both, since these are the remaining two hydroxylated nonproteinogenic residues in **1**. It was also possible that EcdK could be involved in the biosynthesis of 4-Me-Pro, which in cyanobacteria is known to involve initial hydroxylation at C_5 of L-leucine. [14] When metabolites of Δ*ecdK* mutants were analyzed by LCMS, indeed, production of **1** was abolished. Feeding experiments with homoTyr or 4*R*-OH-Pro did not restore production of **1**. On the other hand, either 3*S*-OH-4*S*-Me-Pro or 4*R*-Me-Pro did restore robust echinocandin production (Figure 5A). These results clearly implicate EcdK in a step upstream of 4*R*-Me-Pro formation. The first step in Me-Pro construction is proposed in other systems to be hydroxylation of Leu to 5-OH-Leu (Figure 5B). [14] Further oxidation (either by a second hydroxylation at C_5 or via action of an alcohol oxidase) to γ-methyl-glutamic acid-γ-semialdehyde (γ-Me-Glu-γ-semialdehyde) would set up nonenzymatic cyclization to 3-methyl-Δ^1-pyrroline-5-carboxylic acid (MeP5C). Subsequent reduction of the 1-pyrroline by a nicotinamide-utilizing reductase would give 4-Me-Pro (Figure 5B).

EcdK oxygenates C_5 of L-Leucine iteratively

In order to verify these results observed in the knockout of *ecdK*, we sought to reconstitute *in vitro* the oxidation of L-Leu by EcdK. EcdK could be overproduced in and purified from *E. coli* as an N-His_6-tagged enzyme (Figure 5C) as a soluble protein in a yield of 33 mg/L. Assay of pure EcdK with L-leucine followed by analysis of products revealed O_2-dependent conversion to two new products. One was identified as a hydroxy-Leu by LC/MS after Fmoc derivatization (Figure 5D, trace 2 & 3) (Fmoc-OH-Leu: *m/z* calcd for $C_{21}H_{23}NO_5$ $[M+Na]^+$ 392.1468, found 392.1465). A synthetic standard of Fmoc-5-OH-Leu, which has the same retention time, confirms the assignment of the enzymatic product (Figure 5D, trace 1). The second proposed product, MeP5C, could be captured by ortho-aminobenzaldehyde (*o*-AB)[19] to yield the characteristic A_{440} nm chromophore of the dihydroquinazolinium derivative in a pair of products (major to minor = 6/1, Figure 5D, trace 1 & 2) with identical mass (MeP5C: *m/z* calcd for $C_{13}H_{15}N_2O_2$ $[M]^+$ 231.1128, found 231.1121) and UV spectrum (Figure S8). This is consistent with a pair of diastereomers created at C_1 on capture of the C_1 imine (Figure S9). Scale up of the EcdK reaction was challenging, given that EcdK underwent suicidal inactivation during turnover (Figure 5F). This is characteristic of a subset of the nonheme iron oxygenases, as exemplified also by EcdG (Figure 4C), and most probably reflects adventitious, uncontrolled reaction of an Fe^{IV}=O high valent intermediate in the oxygenation process. From the three traces in Figure 5F, there are about 5-6 turnovers/inactivation event for EcdK. Even with the low turnover/inactivation ratio it was possible to use enough pure EcdK to generate sufficient dihydrquinazolinium product adduct. It was purified by HPLC (Figure S8), collected for NMR analysis of the major isomer, and the structure was validated (SI figure S10-S14). In turn, this supports the scheme in Figure 5B that EcdK acts tandemly to oxygenate C_5 of L-Leu. The *gem* diol product is the carbonyl hydration product of γ-Me-Glu-γ-semialdehyde, in turn in equilibrium with the cyclic imine that is trappable with *o*-AB.

To further corroborate the view that EcdK can act at C_5 of L-Leu iteratively, 5-OH-Leu was synthesized from an intermediate in MePro synthesis. [14] The 5-OH-Leu served as a substrate for EcdK yielding the same imine as seen from L-Leu (Figure 5D, trace 4 & 5; Figure 5E, trace 3 & 4). The cyclic imine, the penultimate compound in biosynthesis of the cyclic amino acid Me-Pro, presumably then undergoes reduction (e.g. by an NADH-dependent oxidoreductase) to afford Me-Pro. Recently, a comparable mononuclear iron-dependent L-Leu 5-hydroxylase has been characterized from *Nostoc punctiforme*. [20] Apparently the L-Leu-5-hydroxylase from *N. punctiforme* stops after the first hydroxylation, in contrast to EcdK which performs both oxidation steps.

Anticandidal activities of echinocandin variants

Earlier reports noted that **2** and **3** retained similar activity against *Candida* as **1**, albeit without detailed data. [6,21] Studies on the semisynthetic cilofungin, which has the same hexapeptide core as echinocandin but different fatty acid side chain, indicated a sixteen fold drop in potency from the fully oxygenated hexapeptide scaffold to the tetra-deshydroxy type scaffold. [21] It is known that the nature of the hydrophobic acyl chain in the semisynthetic variants is an important determinant of their antifungal activity and toxicity. [5,22,23] Given that we had access to **1** from wild type and **3** from the Δ*ecdH* strain, we examined their antifungal activity directly side-by-side along with other deshydroxy analogs against *Candida albicans* strain ATCC 90234 with the results shown in Table 1. Under these assay conditions **3** is three-fold more potent, while the compounds lacking the 3-hydroxy on homotyrosine (**5-7**) are 1.3-5.1 fold less active. Zambias et al noted from work on synthetic derivatives of cilofungin that the hydroxy groups on Pro_3 and Me-OH-Pro_6, along with the homoTyr residue, were the most important hydroxylation sites for anticandidal activity. [21] Our results are consistent with observations on cilofungin but we do not yet see production of echinocandin variants lacking hydroxylation on Pro3 or Me-Pro_6.

DISCUSSION AND CONCLUSIONS

The echinocandins undergo a cascade of hydroxylation reactions during the maturation process of the peptide scaffold. Six hydroxyl groups are distributed over the four nonproteinogenic building blocks in **1**. Two additional hydroxylations, at C_5 of L-Leu have been presumed to initiate the pathway to 4*R*-Me-Pro (as a precursor to the 3*S*-OH-4*S*-Me-Pro_6 residue). Thus, at least eight hydroxylations are required for generation of the mature echinocandin B scaffold, albeit the two hydroxyls introduced at C_5 of L-Leu are cryptic, *i.e.* not persisting in the mature echinocandin scaffold.

Among the questions we had at the outset were which of the putative three EcdGHK oxygenases, if any, act iteratively on Orn_1, and $homoTyr_4$ and which may act on Pro_3 and/or 4*R*-Me-Pro_6 or the L-Leu precursor to Me-Pro. The timing of side chain oxygenations in the antifungal scaffold assembly was also unknown: it was not apparent whether the hydroxylases act upon free amino acid, amino acyl- or peptidyl-S-NRPS stage, or on a deshydroxy echinocandin framework. In addition, there was ambiguity about whether hydroxylations must go in a prescribed order and whether independent routes of oxygenative maturation can go in parallel.

This study identifies five oxygenation events by these three iron enzymes. The nonheme iron EcdG works on free L-homotyrosine to generate 3-OH-L-homoTyr. This step occurs at an early stage during echinocandin biosynthesis before the building block is loaded onto the hexamodular NRPS EcdA. On the other hand, the hydroxylase for C_4 of homotyrosine remains unidentified. Knockout of *ecdG* or *ecdH* significantly decreases the hydroxylation efficiency at this position in the final echinocandin scaffold, although neither enzyme directly hydroxylates it. We thus propose that the HomoTyr-4-hydroxylation reaction occurs after EcdH, probably as the last step during echinocandin B biosynthesis (Figure 6). That placement implies this particular hydroxylase takes the partially deshydroxy macrocycle as a substrate and can produce minor compounds **4** and **5**, with a preference for generating the fully hydroxylated echinocandin B scaffold.

The hemeprotein EcdH, which iteratively works at C_5 and then C_4 of Orn_1, probably functions at a late stage after the release of the macrocyclic peptide from the NRPS EcdA, since a free C_5-OH-L-Orn would unravel to the aldehyde and not be competent to be joined to 3*S*-OH-Me-Pro_6. EcdH is sensitive to the conformational change of the hexapeptide macrocycle when particular-OH groups are missing. When *ecdG* is deleted, EcdH becomes less efficient and produces three variants at Orn_1 (compound **5-8**). The EcdH-mediated hydroxylation at C_5 of Orn_1 creates the hydrolytically labile hemiaminal linkage: a subsequent elimination of water to the cyclic imine would be on pathway to hydrolytic linearization and deactivation of the hexapeptide macrocycle. The post NRPS generation of that hemiaminal linkage is in contrast to known routes to imine linkages in other nonribosomal cyclic peptides such as nostocyclopeptide [24] and koranimine. [25] In those cases, the NRPS assembly lines have C-terminal reductase domains that release the linear peptide aldehydes as nascent products. Imine formation with the N-terminal-NH_2 is the cyclization step. In contrast EcdA releases the cyclic hexapeptide apparently via a fungal macrocyclizing C_T domain and the hemiaminal (precursor to imine) is a post assembly line maturation step.

EcdK, in analogy to EcdG, works as a nonheme mononuclear iron oxygenase at the level of a free amino acid. In this case L-Leu is the substrate and undergoes tandem hydroxylation at C_5 in the first and second step in the pathway to 4*R*-Me-proline assembly. Thus, both EcdH and EcdK perform iterative hydroxylations, one on the same carbon (EcdK), the other on adjacent carbons (EcdH) of their substrates. The proposed timing of EcdG, EcdK, and EcdH in echinocandin scaffold maturation is noted in Figure 6.

Three echinocandin oxygenation events are still to be characterized. To date we have not identified the hydroxylase(s) for Pro_3 (oxygenation at C_4) and 4*R*-Me-Pro_6 (oxygenation at C_3) and at what stage they occur. Given the absence of any deshydroxy forms of echinocandins at either of these two residues, hydroxylations on the macrocycle after release by the nonribosomal peptide synthetase EcdA is disfavored. Until the Pro/Me-Pro hydroxylases are found, we cannot tell if the free amino acids or peptidyl-S-NRPS species are substrates. Precedent does suggest conversion of Pro to 4*R*-OH-Pro can occur at the level of the free amino acid. [26] It may be that is also a reasonable expectation for the transformation of 4*R*-Me-Pro to 3*S*-OH-4*S*-Me-Pro. As noted earlier, the homoTyr-4-hydroxylase is also not yet identified. We have noted that two genes annotated as *htyE* and

htyF are adjacent to the *htyA-D* genes required for conversion of Tyr to homoTyr. [8] Whether either HtyE or HtyF are any of the missing three oxygenation catalysts for echinocandin B maturation is a subject for further study.

MATERIALS AND METHODS

Gene knock-out in *Emericella rugulosa*

Procedure for the gene knockout for *E. rugulosa* is based on a protocol established for *A. nidulans* A4 with modification.[12] The linear knockout cassette was constructed as done in previous studies [12] except that the glufosinate resistance gene driven by the *trpC* promoter was used as selection marker.

Transformation of *E. rugulosa* was done as follows: *E. rugulosa* conidiospores were inoculated to 250 mL of glucose minimal media with 10 mM ammonium tartrate as nitrogen source (GMMT)[8] at a concentration of 10^7 conidiospores/mL. The culture was shaken at 28 °C for 16 hrs and 250 rpm. Thereafter, 2 g of germlings was digested at 28 °C while shaking at 100 rpm with 3 g of Vinotaste Enzyme Mixture (Novozyme) in 8 mL GMMT supplemented with 8 mL of 1.2 M KCl and 50 mM citric acid (pH 5.5). After 3 hrs, the digestion mixture was overlaid on top of an equal volume of 1.2 M sucrose and centrifuged at 1800 g and 4 °C. The protoplasts trapped on top of the sucrose cushion were isolated, washed with 0.6 M KCl three times and resuspended in 500 μL 0.6 M KCl and 50 mM $CaCl_2$ solution. Ten micrograms of the linear knockout cassette was added to 100 μL of resuspended protoplasts and the mixture was briefly vortexed for 3 seconds. 50 μL of PEG solution (0.6 M KCl, 50 mM $CaCl_2$, 75 mM polyethylene glycol with average molecular weight of 3250 g/mol, 10 mM Tris-HCl, pH 7.5) was added to the protoplast and the resulting mixture was incubated at 4 °C for 20 minutes. Thereafter, 1 mL of PEG solution was added to the protoplast mixture and incubated at room temperature for an additional 20 minutes and subsequently plated to GMMT plates supplemented with 1.2 M sorbitol and 10 mg/mL glufosinate.

Genomic DNA from the resulting transformants was isolated using ZR Fungal/Bacterial DNA Miniprep kit (Zymo) and was used as template for PCR screening. Primers for the PCR screening are given in Table S3.

Culturing of *Emericella rugulosa* and mutants in large scale

The media and process used for ***Emericella rugulosa*** growth are adapted from US Patent 4,288,549.[10] Wild-type or verified mutant *E. rugulosa* cultures were grown in solid GMMT plates (or solid GMMT plates supplemented with 10 mg/mL glufosinate in the case of the mutants) at 28 °C for 4 days. The harvested conidiospores from the plates were inoculated to 8 L of medium 2 [2.5% (w/v) glucose (Sigma), 1% peptone (BD Biosciences), 1% (w/v) starch (Sigma), 1%(w/v) molasses, 0.4% (w/v) N-Z Amine A (Sheffield Biosciences), 0.2% (w/v) calcium carbonate (Sigma)] and was shaken at 250 rpm and 28 °C for 7 days.

Preparative Scale Isolation of Echinocandins for NMR analysis

The extraction process for the echinocandins was adapted from US Patent 4,288,549. [10] The whole cell culture and broth was mixed with an equal volume of methanol and stirred thoroughly for one hour at 16 °C. It was then filtered through a layer of Miracloth and adjusted to pH 4.0 by the addition of concentrated HCl. The acidified filtrate was extracted twice with equal volumes of chloroform. The chloroform extracts were combined and concentrated *in vacuo* to 5-10 mL with a dark yellow color. The solution was centrifuged and the supernatant was loaded onto a Sephadex LH20 size-exclusion column (400×40 mm). The compounds were eluted by using 1:1 chloroform/methanol (v/v) with a rate of 1.5 mL/min. Fractions with desired compounds were combined and dried *in vacuo*. The mixture was resuspended in methanol, centrifuged, and the supernatant was injected onto a Phenomenex Luna prep-HPLC C18 column (250 × 21.1 mm, 10 micron, part # 00G-4253-P0-AX) using a solvent gradient of 40 to100% B over 30 min (solvent A, 0.1% TFA/H2O; solvent B, 0.1% TFA/MeCN). Elution of Echinocandins was monitored at 280 nm. The fractions with desired compounds were lyophilized for further analysis.

Culturing, feeding and extraction of *Emericella rugulosa* with deuterated amino acids for MS analysis

10 mL *Emericella rugulosa* culture, either in medium 2 described above or GMMT, was inoculated and incubated at 28 °C without shaking. Deuterated amino acids [(D_6)-L-ornithine, (D_7)-L-proline, and (D_7, ^{15}N)-L-tyrosine] were purchased from Cambridge Isotope Laboratories (Andover, MA). Orn and Pro for feeding studies were prepared as neutral aqueous solutions with a concentration of 60 mg /mL and sterile filtered. Tyrosine solution with a concentration of 36 mg / mL was prepared in a similar way but at pH 13 due to tyrosine's poor solubility at neutral pH. 250 μL of compound solutions were added to 10 mL *Emericella rugulosa* culture at Day 4 after inoculation, and harvested at Day 7. For compound extraction, an equal volume of methanol was added to the culture broth, mixed well, incubated at room temperature for 15 min and filtered. An equal volume of chloroform was then added to the mixture. The organic layer was separated, dried with anhydrous Na_2SO_4, and filtered. The crude extract was dried *in vacuo* and stored at -20 °C for further use.

LC-MS methods for natural compound profiles

All LCMS analysis was carried out on an Agilent 1200 Series HPLC coupled to an Agilent 6520 QTOF spectrometer. The crude material was re-dissolved in 200 μL methanol and filtered through a 0.45 μM membrane. 10 μL were injected for each analytical run. Separation was achieved on a 50 × 2.0 mm Phenomenex Gemini C18 column (part # 00B-4435-B0) using a solvent gradient of 20 to 100% B over 30 min (solvent A: 0.1% formic acid in H_2O; solvent B: 0.1% formic acid in acetonitrile). Temperature was 350 °C.

Cloning of EcdG and EcdK

RNA from 7-day old wild-type *E. rugulosa* liquid static culture was isolated using Ribopure-Yeast kit (Ambion). First strand cDNA synthesis was performed with SuperScript III-First Strand Synthesis SuperMix using the RNA from wild-type *E. rugulosa* as a

template and oligo-dT as a reverse-transcriptase primer. EcdG and EcdK cDNA were amplified using primers listed in Table S3 and cloned into p*ET*30- Xa/LIC vector. The resulting constructs were transformed into *E.coli* Xl1 Blue cells (Stratagene) and sent for sequencing to verify correct splicing of the cDNA.

Reconstitution of EcdG/EcdK and hydroxylation assays

Anaerobic reconstitution of EcdG/EcdK with Fe(II), α-ketoglutarate (α-KG), and DTT was conducted as for other nonheme Fe(II)-dependent halogenases. [27] Then the hydroxylation assay was initiated by adding 100 μM or otherwise indicated concentration of reconstituted EcdG or EcdK to a 300 μL solution containing 50 mM HEPES (pH 7.0), 2 mM α-KG, and 1 mM homoTyr (EcdG) or Leu or 5-hydroxyl-Leu (EcdK). The mixture was incubated at ambient temperature and 50 μL aliquot was quenched by addition of 50 μL acetonitrile at various time points. The solution was centrifuged at 13k rpm for 5 min and the supernatant was subjected for further use.

Fmoc derivatization and LC/MS analysis

To 100 μL of quenched hydroxylation assay mixture described above was added 50 μL of 200 mM sodium borate (pH 8.0) and 20 μL of 10 mM 9-fluorenylmethyl chloroformate (Fmoc-Cl). After incubation at room temperature for 5 min, 20 μl of 100 mM 1-adamantanamine was added to stop the reaction. 10 μL of the mixture were injected onto a 50 × 2.0 mm Phenomenex Gemini C18 column (part # 00B-4435-B0) using a solvent gradient of 20 to 100% B over 20 min (solvent A: 0.1% formic acid in water; solvent B: 0.1% formic acid in acetonitrile).

Isolation of 3-OH-homoTyrosine

1.2 mL of the quenched EcdG hydroxylation assay mixture described above was dried *in vacuo* and re-dissolved in 200 μL H_2O. 50 μL was injected for each run to a 100 *2.1 mm Thermo Hypercarb column (5 micron, part # 35005-102130). Solvent A was composed of 20 mM perfluoropentanoic acid in water; solvent B was composed of acetonitrile without any additive. Separation was achieved by using the LC gradient 0% B for 5 min followed by 0-70% B over 30 min.

The process was monitored at 275 nm. Solutions from two major absorption peaks were collected and subjected to MS to confirm that hydroxyl-homoTyr was eluted prior to homoTyr. Hydroxyl-homoTyr from 10 runs was collected, combined, lyophilized and subjected to NMR.

Ortho-aminobenzaldehyde derivatization and LC/MS analysis

The hydroxylation assay containing 50 mM HEPES (pH 7.0), 2 mM α-KG, 1 mM Leu or 5-hydroxyl-Leu, and 0.2 mM reconstituted EcdK proceeded for 15 min at ambient temperature. The solution was filtered through a Microcon YM-3 (Amicon) microcentrifuge filter to remove the enzyme. The flow-through was collected, to which solid ortho-aminobenzaldehyde (*o*-AB) was directly added at 0.5 mg per 100 μL, mixed well and incubated at ambient temperature for 1 hour. Afterwards, the insoluble *o*-AB was removed by filtration and the solution was subjected to LC/MS or HPLC (see below). 10 μL were

injected onto a 50 × 2.0 mm Phenomenex Gemini C18 column (part # 00B-4435-B0) using a solvent gradient of 1 to 100% B over 20 min (solvent A: 0.1% formic acid in water; solvent B: 0.1% formic acid in acetonitrile).

Isolation of *o*-AB-derivatized 3-methyl-Δ^1-pyrroline-5-carboxylic acid (MeP5C) by HPLC

100 μL of *o*-AB derivatized solution described above was injected for each run to a 25 cm *4.6 mm Supelco Discovery C18 HPLC column (5 micron, part # 504971). Separation was achieved by using a very mild solvent gradient of 5-8% B over 20 min (solvent A: 0.1% TFA in water; solvent B: 0.1% TFA in acetonitrile). The process was monitored at 440 nm. Solutions from the major absorption peak (the 2nd) were collected. Material from ~20 runs was combined, lyophilized and subjected to NMR.

Synthesis of 4*R*-methyl proline and 4-hydroxyl-L-leucine

Following the protocol disclosed by Rapoport et al, [28] 4*R*-methyl proline was synthesized starting from commercially available materials. 5*R*-hydroxyl-L-leucine was synthesized by the known method described by Moore et al. [14]

Antifungal susceptibility test by broth microdilution method

The potency of Echinocandin B **1** and its analogs against *Candida albicans* was measured following a published protocol using 96-well plate. [29] MIC_{90} was designated as the lowest concentration that produced an increase of less than 10% in OD over that of the adjacent sterile control.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work is supported by NIH Grants 1R01GM092217 and 1DP1GM106413 to Y. T.; 1R01GM49338 to C.T.W.; NRSA GM08496 to R.A.C; and NSF DGE-0707424 to G.C.

References

1. Benz F, Knüsel F, Nüesch J, Treichler H, Voser W, Nyfeler R, Keller-Schierlein W. Helv Chim Acta. 1974; 57:2459. [PubMed: 4613708]
2. Sawistowska-Schröder ET, Kerridge D, Perry H. FEBS Lett. 1984; 173:134. [PubMed: 6235127]
3. Bouffard FA, Zambias RA, Dropinski JF, Balkovec JM, Hammond ML, Abruzzo GK, Bartizal KF, Marrinan JA, Kurtz MB. J Med Chem. 1994; 37:222. [PubMed: 8295208]
4. TOMISHIMA M, OHKI H, YAMADA A, TAKASUGI H, MAKI K, TAWARA S, TANAKA H. J Antibiot (Tokyo). 1999; 52:674. [PubMed: 10513849]
5. Debono M, Turner WW, LaGrandeur L, Burkhardt FJ, Nissen JS, Nichols KK, Rodriguez MJ, Zweifel MJ, Zeckner DJ. J Med Chem. 1995; 38:3271. [PubMed: 7650681]
6. Traber R, Keller-Juslén C, Loosli H-R, Kuhn M, Von Wartburg A. Helv Chim Acta. 1979; 62:1252.
7. Tóth V, Nagy C, Miskei M, Pócsi I, Emri T. Folia Microbiol. 2011; 56:381. [PubMed: 21858538]
8. Cacho RA, Jiang W, Chooi Y-H, Walsh CT, Tang Y. J Am Chem Soc. 2012; 134:16781. [PubMed: 22998630]
9. Adefarati AA, Giacobbe RA, Hensens OD, Tkacz JS. J Am Chem Soc. 1991; 113:3542.

10. Boeck, LD.; Kastner, RE. US Patent. 4,288,549. 1981.
11. Nayak T, Szewczyk E, Oakley CE, Osmani A, Ukil L, Murray SL, Hynes MJ, Osmani SA, Oakley BR. Genetics. 2006; 172:1557. [PubMed: 16387870]
12. Szewczyk E, Nayak T, Oakley CE, Edgerton H, Xiong Y, Taheri-Talesh N, Osmani SA, Oakley BR. Nat Protocols. 2007; 1:3111.
13. Evans DA, Weber AE. J Am Chem Soc. 1987; 109:7151.
14. Luesch H, Hoffmann D, Hevel JM, Becker JE, Golakoti T, Moore RE. J Org Chem. 2003; 68:83. [PubMed: 12515465]
15. Strushkevich N, MacKenzie F, Cherkesova T, Grabovec I, Usanov S, Park H-W. Proc Natl Acad Sci U S A. 2011; 108:10139. [PubMed: 21636783]
16. Galonić DP, Barr EW, Walsh CT, Bollinger JM, Krebs C. Nat Chem Biol. 2007; 3:113. [PubMed: 17220900]
17. Vaillancourt FH, Yin J, Walsh CT. Proc Natl Acad Sci U S A. 2005; 102:10111. [PubMed: 16002467]
18. Kurokawa N, Ohfune Y. J Am Chem Soc. 1986; 108:6041. [PubMed: 22175372]
19. Mezl VA, Knox WE. Anal Biochem. 1976; 74:430. [PubMed: 962101]
20. Hibi M, Kawashima T, Sokolov P, Smirnov S, Kodera T, Sugiyama M, Shimizu S, Yokozeki K, Ogawa J. Appl Microbiol Biotechnol. 2012; 1
21. Zambias RA, Hammond ML, Heck JV, Bartizal K, Trainor C, Abruzzo G, Schmatz DM, Nollstadt KM. J Med Chem. 1992; 35:2843. [PubMed: 1495015]
22. Boeck LD, Fukuda DS, Abbott BJ. J Antibiot (Tokyo). 1989; 42:382. [PubMed: 2708131]
23. Debono M, Abbott BJ, Fukuda DS, Barnhart M, Willard KE, Molloy RM, Michel KH. J Antibiot (Tokyo). 1989; 42:389. [PubMed: 2708132]
24. Becker JE, Moore RE, Moore BS. Gene. 2004; 325:35. [PubMed: 14697508]
25. Evans BS, Ntai I, Chen Y, Robinson SJ, Kelleher NL. J Am Chem Soc. 2011; 133:7316. [PubMed: 21520944]
26. Lawrence CC, Sobey WJ, Field RA, Baldwin JE, Schofield CJ. Biochem J. 1996; 313:185. [PubMed: 8546682]
27. Vaillancourt FH, Yeh E, Vosburg DA, O'Connor SE, Walsh CT. Nature. 2005; 436:1191. [PubMed: 16121186]
28. Koskinen AMP, Rapoport H. J Org Chem. 1989; 54:1859.
29. Schwalbe, R.; Steele-Moore, L.; Goodwin, AC., editors. Antimicrobial Susceptibility Testing Protocols. CRC Press; 2007.

ABBREVIATIONS

3-OH-Me-Pro	3S-OH-4S-methyl-Proline
Me-Pro	4*R*-methyl-Proline
α-KG	α-ketoglutarate
γ-Me-Glu-γ-semialdehyde	γ-methyl-glutamic acid-γ- semialdehyde
MeP5C	3-methyl-Δ^1-pyrroline-5-carboxylic acid
***o*-AB**	ortho-aminobenzaldehyde
GMMT	glucose minimal media with 10 mM ammonium tartrate

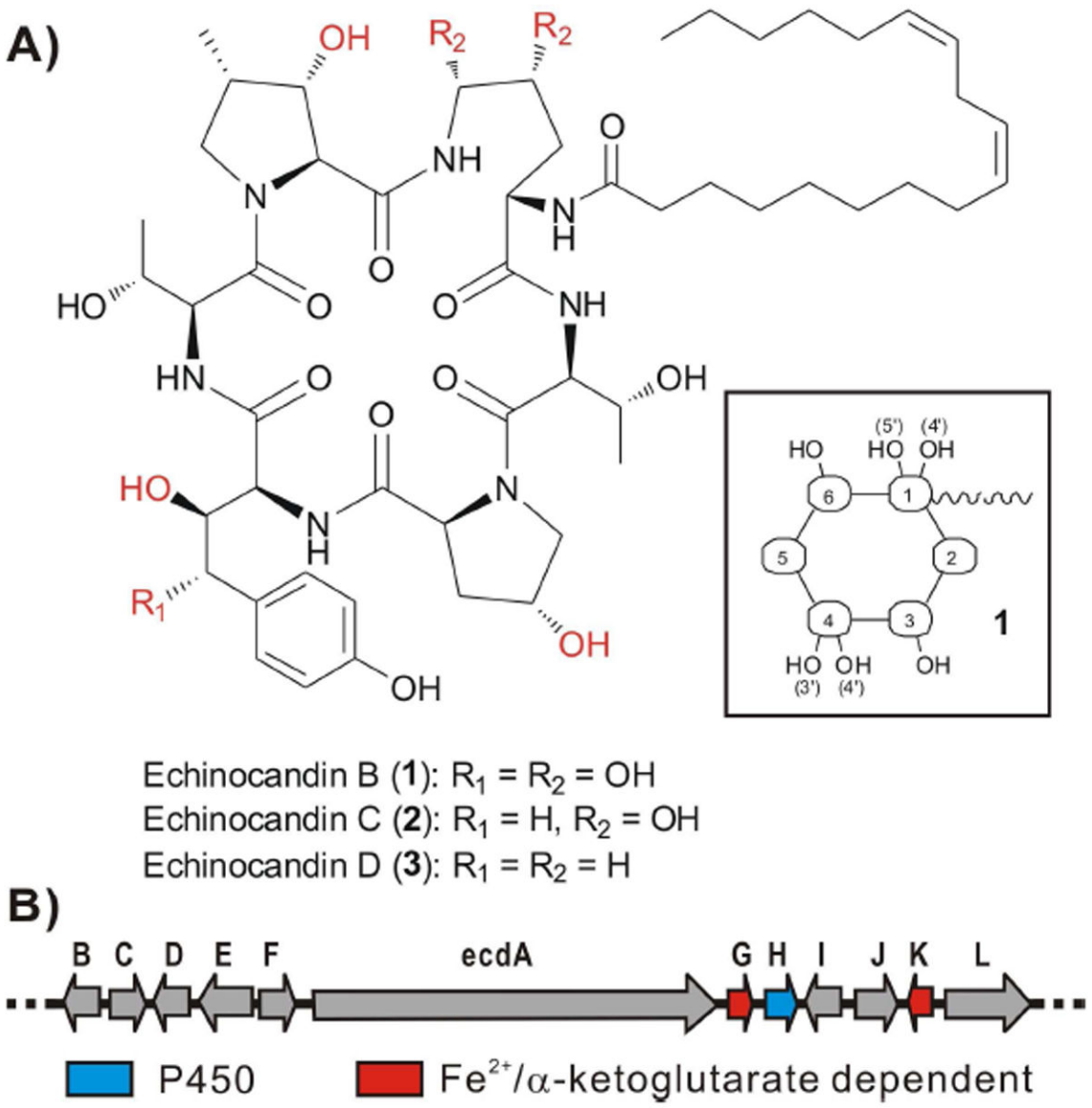

Figure 1.

A) Structure of Echinocandin B (**1**), C (**2**), and D (**3**). A schematic representation of **1** used in Figures 2 and 3 is shown in the inset. B) Proposed nonheme mononuclear iron oxygenase and P450 type heme genes in *ecd* gene clusters.

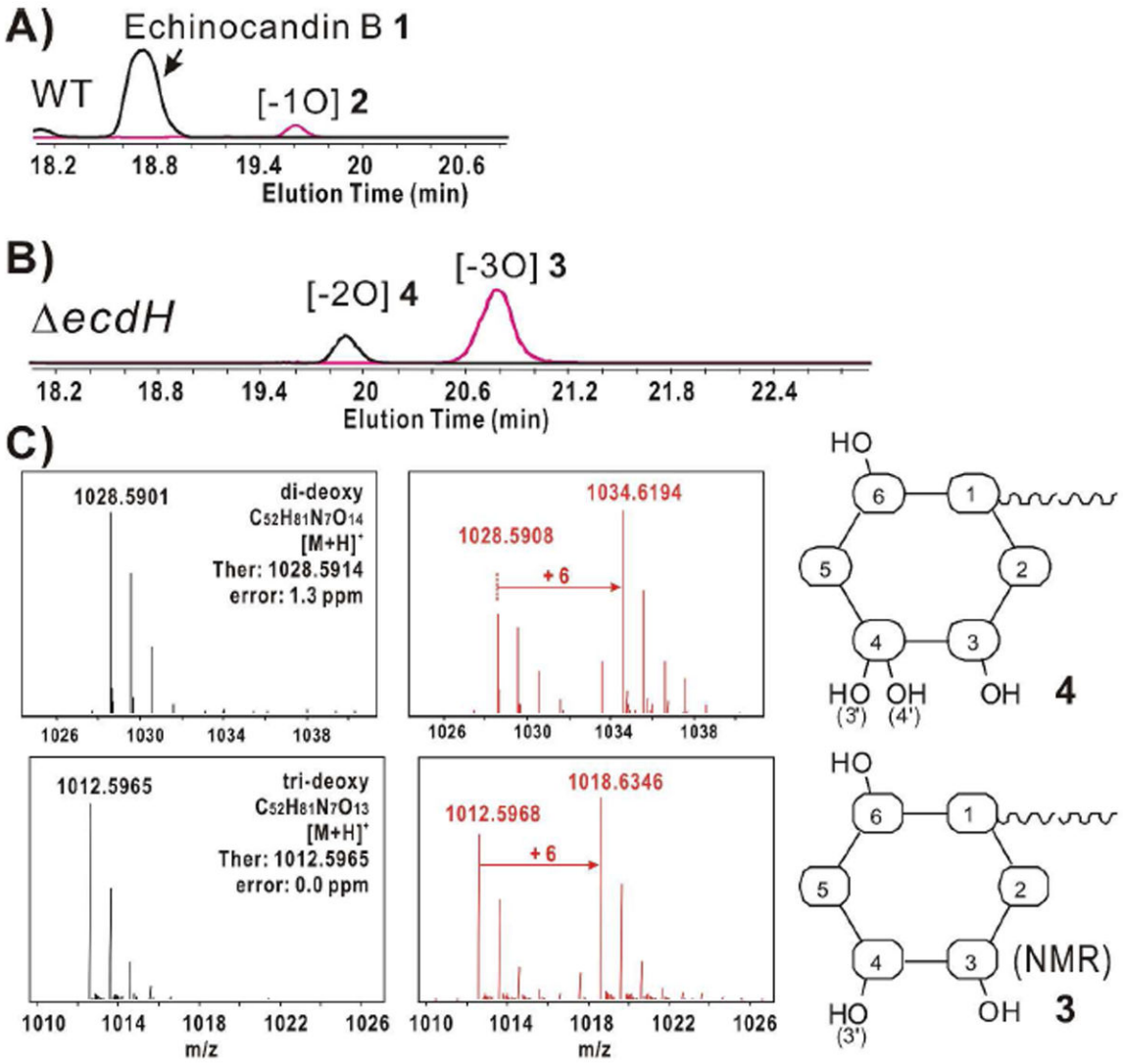

Figure 2.

A) Extracted ion chromatogram of Echinocandin B (**1**) and C (**2**) from *Emericella rugulosa* profile. B) Extracted ion chromatogram of Echinocandin mutants from Δ*ecdH* profile. C) MS of parent ions from Echinocandin mutants produced by Δ*ecdH* (**3** & **4**), either from regular media (black) or fed with (D_6)-L-ornithine (red). The proposed structure of each compound is on the right.

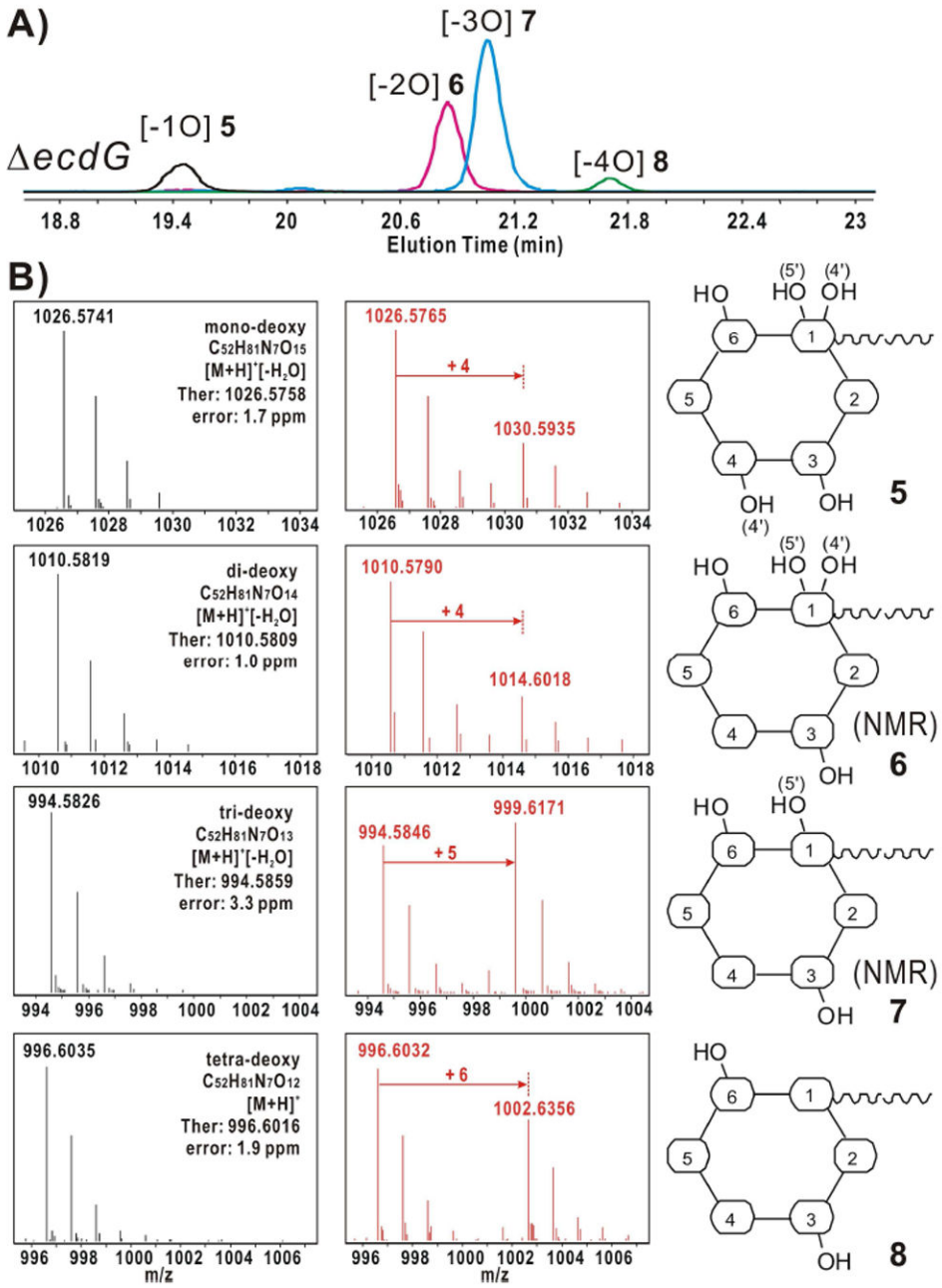

Figure 3.

A) Extracted ion chromatogram of Echinocandin mutants from Δ*ecdG* profile. B) MS of parent ions from Echinocandin mutants produced by Δ*ecdG* (**5 - 8**), either from regular media (black) or fed with (D_6)-L-ornithine (red). The proposed structure of each compound is on the right.

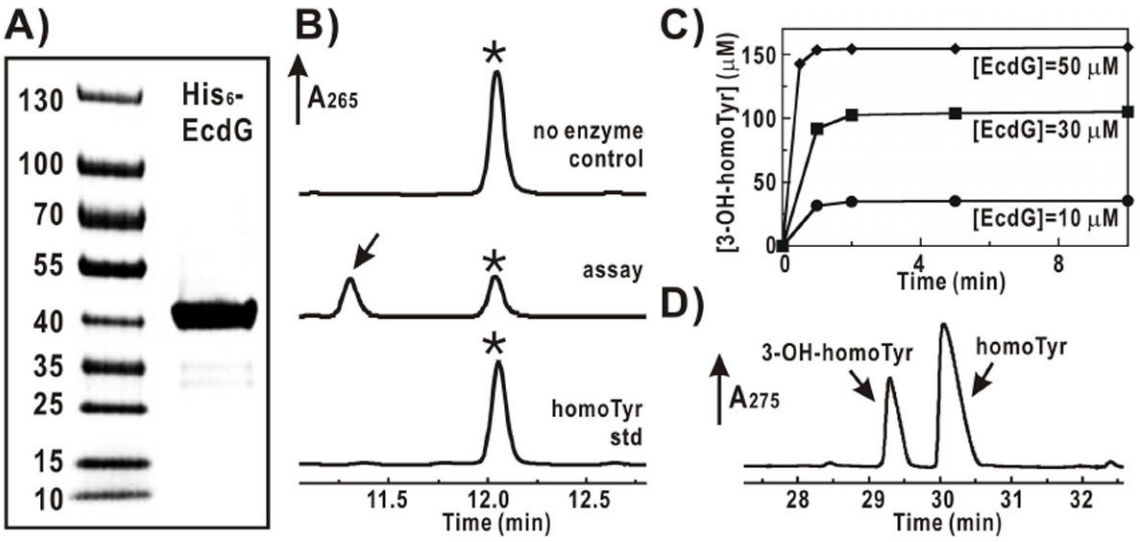

Figure 4.

A) SDS-PAGE gel showing purified His_6-EcdG enzyme (expected molecular weight: 40.0 kDa). B) HPLC-UV (259 nm) traces indicate formation of hydroxy-homoTyr (arrow, middle trace) relative to negative control without enzyme (top) and homoTyrosine standard (star, bottom trace) as Fmoc-derivatives. C) Time course of formation of hydroxy-homoTyr at different enzyme concentrations. D) HPLC-UV (275 nm) trace showing isolation of hydroxyl-homoTyr from EcdG assay solutions for further NMR analysis.

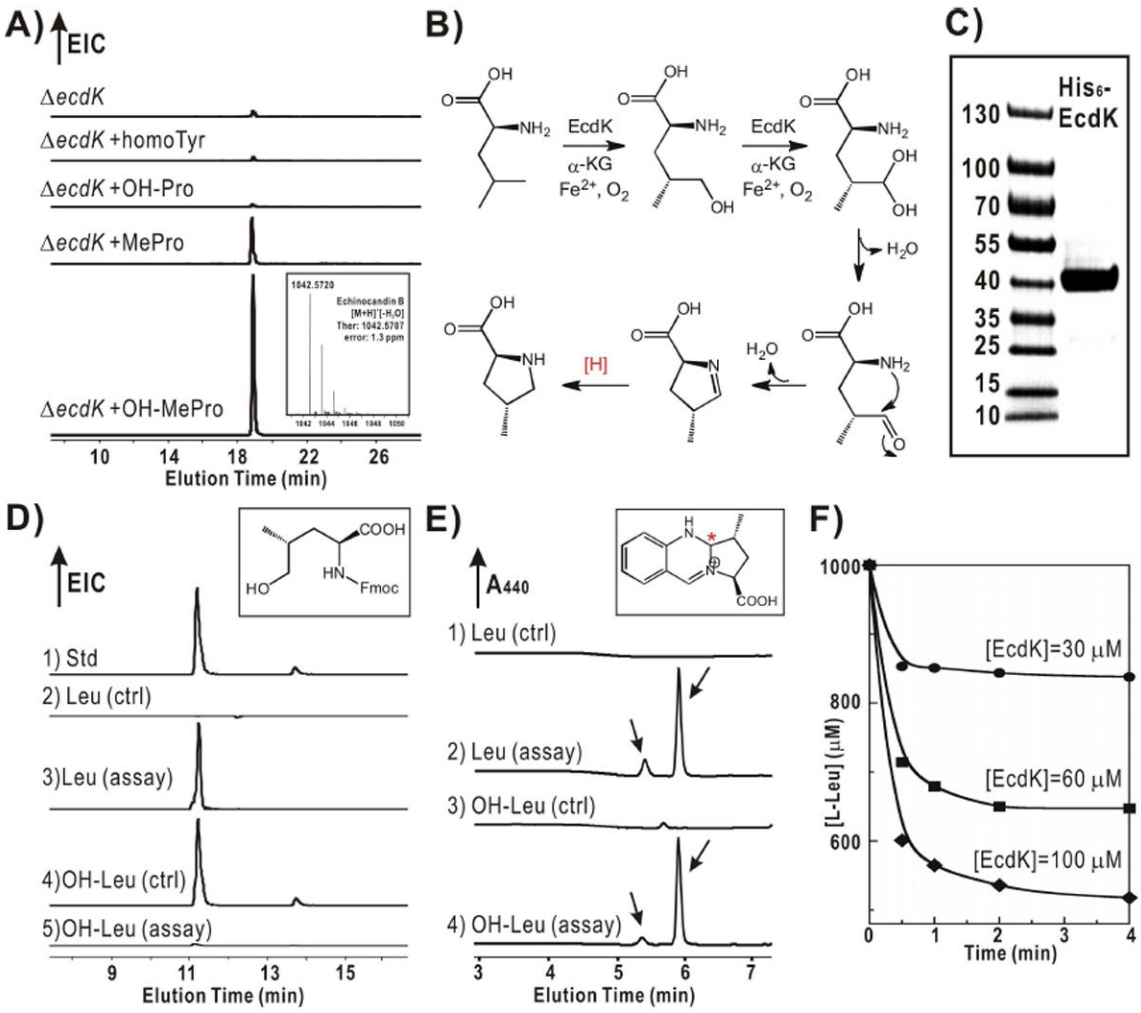

Figure 5.

A) Extracted ion chromatogram of Echinocandin B (**1**) from Δ*ecdK* mutant. The culture media is supplemented with amino acid as labeled (OH-Pro = 4*R*-OH-L-Pro; MePro = 4*R*-Me-L-Pro; OH-Me-Pro: 3*S*-OH-4*S*-Me-L-Pro). **B**) Proposed biosynthesis pathway for 4*R*-methyl-L-proline from L-Leu by tandem hydroxylation at C_5. **C**) SDS-PAGE gel showing purified His_6-EcdK enzyme (expected molecular weight: 39.7 kDa). **D**) Extracted ion chromatogram of Fmoc-derivatized 5-OH-Leu (structure shown in inset). Top to bottom: 1) synthetic standard of Fmoc-5-OH-Leu; 2) assay with Leu, no enzyme; 3) assay with Leu, with enzyme; 4) assay with 5-OH-Leu, no enzyme; 5) assay with 5-OH-Leu, with enzyme, reflecting comlete consumption of 5-OH-Leu. **E**) HPLC-UV (440 nm) traces showing the formation of MeP5C after *o*-AB derivatization (derivatized product structure shown in inset). Arrows indicate a pair a diastereomers with a ratio of ~ 1:6. Top to bottom: 1) assay with Leu, no enzyme; 2) assay with Leu, with enzyme; 3) assay with 5-OH-Leu, no enzyme; 4) assay with 5-OH-Leu, with enzyme. (**F**) Time course of utilization of L-leucine at different EcdK concentrations. Residual Leu is quantified in its Fmoc-derivatized form by UV/Vis absorption.

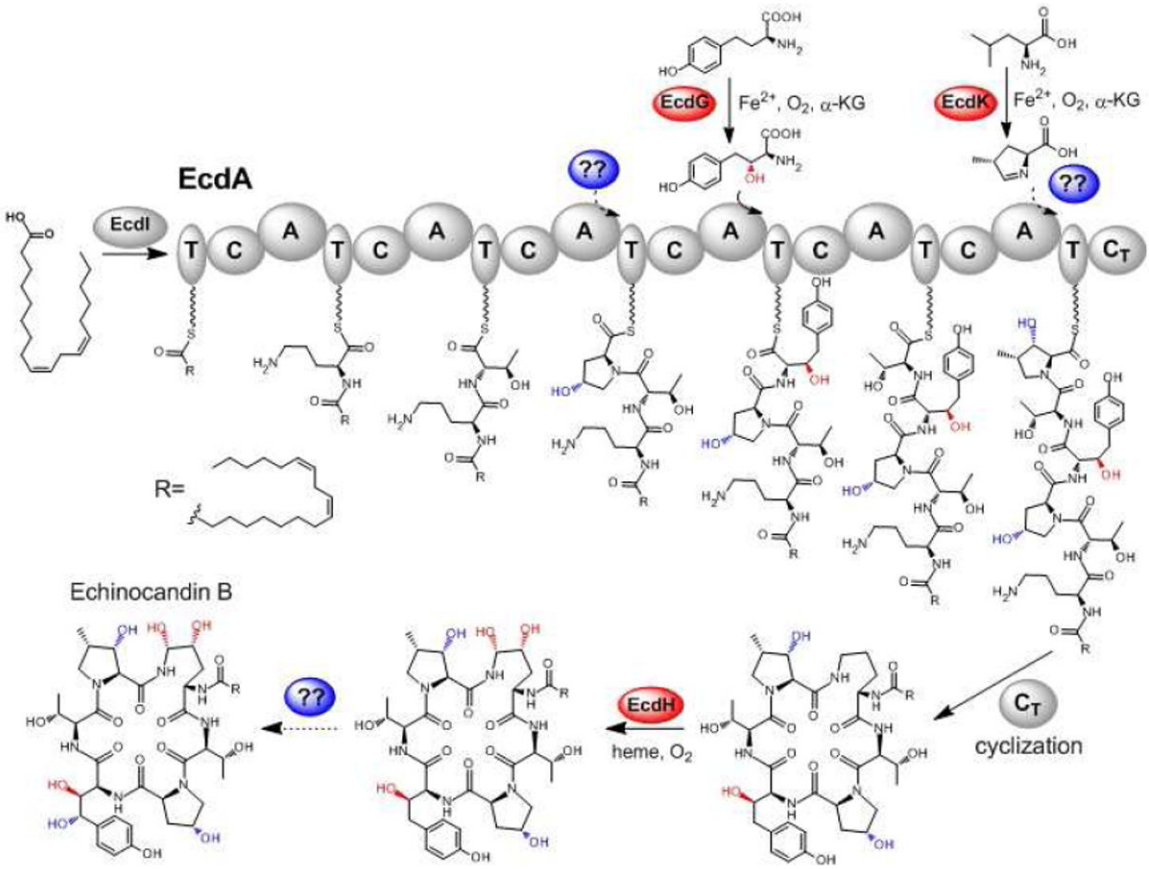

Figure 6.

Proposed timing of post-translational hydroxylations for biosyntheisis of echinocandin B (**1**). Confirmed steps and modifications are highlighted in red while proposed ones are in blue.

Table 1

MIC_{90} measurement of compound echinocandin B (**1**) and its deshydroxy analogs against *Candida albicans* strain ATCC 90234.

Compound	MIC (μg/mL)
1 (wt)	0.6
3 (Δ*ecdH*)	0.2
5 (Δ*ecdG*)	3.1
6 (Δ*ecdG*)	0.8
7 (Δ*ecdG*)	1.6