JACS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Identification and Characterization of the Echinocandin B Biosynthetic Gene Cluster from *Emericella rugulosa* NRRL 11440

Ralph A. Cacho,[†] Wei Jiang,[§] Yit-Heng Chooi,[†] Christopher T. Walsh,[*,§] and Yi Tang[*,†,‡]

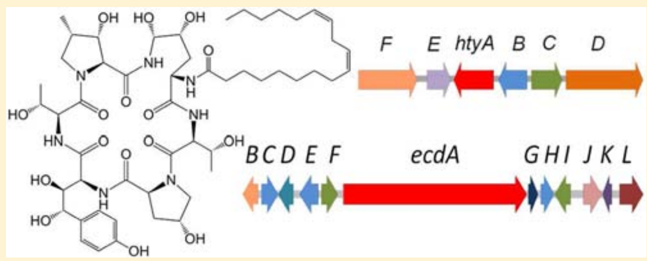
[†]Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, 420 Westwood Plaza, Los Angeles, California 90095, United States

[‡]Department of Chemistry and Biochemistry, University of California, Los Angeles, 607 Charles E. Young Drive East, Los Angeles, California 90095, United States

[§]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 200 Longwood Ave, Boston, Massachusetts 02115, United States

Supporting Information

ABSTRACT: Echinocandins are a family of fungal lipidated cyclic hexapeptide natural products. Due to their effectiveness as antifungal agents, three semisynthetic derivatives have been developed and approved for treatment of human invasive candidiasis. All six of the amino acid residues are hydroxylated, including 4*R*,5*R*-dihydroxy-L-ornithine, 4*R*-hydroxyl-L-proline, 3*S*,4*S*-dihydroxy-L-homotyrosine, and 3*S*-hydroxyl-4*S*-methyl-L-proline. We report here the biosynthetic gene cluster of echinocandin B **1** from *Emericella rugulosa* NRRL 11440 containing genes encoding for a six-module nonribosomal peptide synthetase EcdA, an acyl-AMP ligase EcdI, and oxygenases EcdG, EcdH, and EcdK. We showed EcdI activates linoleate as linoleyl-AMP and installs it on the first thiolation domain of EcdA. We have also established through ATP−PP_i exchange assay that EcdA loads L-ornithine in the first module. A separate *hty* gene cluster encodes four enzymes for de novo generation of L-homotyrosine from acetyl-CoA and 4-hydroxyphenyl-pyruvate is found from the sequenced genome. Deletions in the *ecdA*, and *htyA* genes validate their essential roles in echinocandin B production. Five predicted iron-centered oxygenase genes, *ecdG*, *ecdH*, *ecdK*, *htyE*, and *htyF*, in the two separate *ecd* and *hty* clusters are likely to be the tailoring oxygenases for maturation of the nascent NRPS lipohexapeptidolactam product.



■ INTRODUCTION

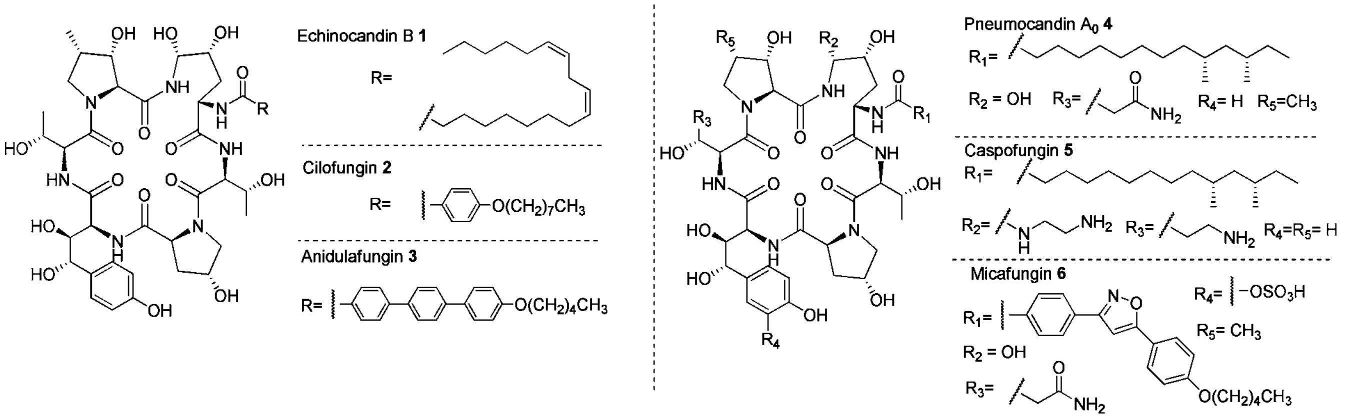
Invasive candidiasis caused by opportunistic pathogenic strains of genus *Candida* accounts for 17% of ICU-related infections, third highest after *Staphylococcus aureus* and *Pseudomonas* spp.-related infections.[1] Moreover, there has been a steady increase in the incidence of invasive candidiasis correlating with the increased use of immunosuppressants, broad-spectrum antibiotics, intravenous catheters and prosthetics, and invasive clinical procedures.[2,3] Echinocandins, a family of lipohexapeptides that prevent fungal wall synthesis through noncompetitive inhibition of 1,3-β-glucan synthase, rapidly rose to the top ranks of antifungal agents due to their activity against a wide range of *Candida* spp., in particular azole-resistant strains, and are significantly less toxic compared to amphotericin B.[4,5]

As the first echinocandin discovered, echinocandin B (**1**) was isolated from *Aspergillus nidulans* var. *echinulatus*[6] and *Aspergillus nidulans* var. *roseus* NRRL 11440.[7] While **1** and family members subsequently isolated, such as pneumocandin A_0 (**4**),[8] aculeacin,[9] cryptocandin,[10] and mulundocandin,[11] show anti-*Candida* activity, the hemolytic properties of natural echinocandins prevented their use as therapeutics. Derivatization of **1** and **4**, especially in the fatty acid moiety, led to the development of cilofungin (**2**),[12,13] anidulafungin (**3**) (Eraxis, Pfizer),[14] caspofungin (Cancidas, Merck and Co.) (**5**),[15] and micafungin (**6**) (Mycamine, Astellas Pharma)[16] (Scheme 1), which have reduced hemolytic properties while the bioactivity of their parent compound is maintained. For example, **3**, a semisynthetic derivative of **1** containing a substituted terphenyl acyl chain, was approved by the FDA in 2006.[14,17]

In addition to their antifungal activity, echinocandins reflect interesting biosynthetic features in their structures. Aside from the long chain fatty acyl amide, the presence of non-proteinogenic amino acids 4*R*,5*R*-dihydroxyl-L-ornithine, 3*S*-hydroxyl-4*S*-methyl-L-proline, 4*R*-hydroxyl-L-proline, and 3*S*,4*S*-dihydroxyl-L-homotyrosine suggests that echinocandins are synthesized by nonribosomal peptide synthetases (NRPSs). An intriguing feature of echinocandins is the presence of multiple alcohol and diol groups within the scaffold as a result of the incorporation of these unnatural residues; the most distinct of these hydroxyl groups is at Cδ of L-ornithine, which creates a hydrolytically labile hemiaminal in the macrocycle[18,19] and the vicinal diol found in 3*S*,4*S*-dihydroxy-L-homotyrosine. It is noted that enzymatic Cδ-hydroxylation of L-ornithine is a

Received: July 26, 2012
Published: September 21, 2012

Scheme 1. Echinocandin B (1), Its Semisynthetic Derivatives Cilofungin (2) and Anidulafungin (3), and the Related Natural Compound Pneumocandin A_0 (4) and Its Semisynthetic Derivatives Caspofungin (5) and Micafungin (6)



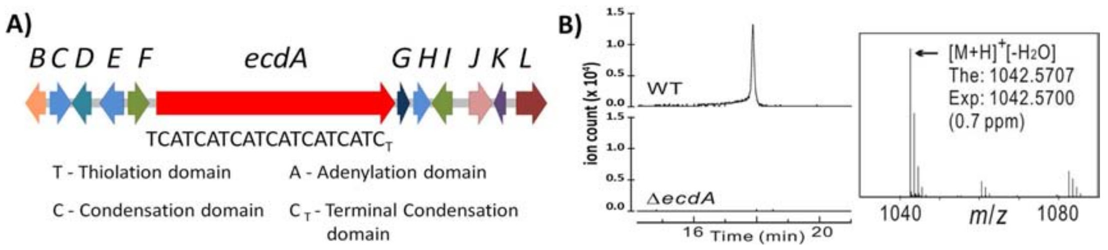


Figure 1. Identification of the *ecd* gene cluster. (A) Organization of the *ecd* gene cluster: *ecdA* encodes for the six-module nonribosomal peptide synthetase and *ecdI* encodes for the linoleyl-AMP ligase responsible for lipoinitiation of EcdA; also found in the cluster are *ecdG*, *ecdH*, and *ecdK*, encoding for proposed hydroxylases. (B) Metabolic profile of Δ*ecdA* mutant shows the loss of production of **1**.

particularly novel and challenging reaction due to the presence of the neighboring amine. Because of this unstable linkage, the total synthesis of **1** has not been demonstrated so far. Only simpler versions without the hydroxyl groups in the L-ornithine position, such as echinocandin D, have been synthesized and used in SAR studies.[20−22]

Despite the medical importance and intriguing structural motifs in echinocandin, the genetic and molecular basis for the biosynthesis of this family has remained unknown to date. The enzymology (sequence, specificity, type of oxygenase) behind this multitude of hydroxylation steps is also unresolved. *A. nidulans* var. *roseus* NRRL 11440 (ATCC 58397) is an industrially important strain that produces **1**, which is a precursor for the semisynthetic **3** (Scheme 1). A polyphasic characterization showed that this strain should belong to the *Emericella rugulosa* species, and henceforth, we use that nomenclature here.[23] In this study, we report the discovery of the gene cluster of **1** in this strain through gene deletion of a multimodular NRPS and biochemical characterization of the enzymes involved in the lipoinitiation process. Through gene deletion and chemical complementation, we have also uncovered the separate gene cluster responsible for the biosynthesis of L-homotyrosine in the peptide scaffold of **1**.

RESULTS

Whole Genome Sequencing and Analysis of NRPS Gene Clusters of *E. rugulosa* NRRL 11440. Whole genome shotgun sequencing of *E. rugulosa* NRRL 11440 was performed using Illumina HiSeq2000 to generate ∼17 Gbp of sequence. Assembly of the sequence reads generated 433 contigs with N50 length of 235 313 bases (See Supporting Information, Table S1). The total length of the 433 contigs amounts to 32 224 016 bases, which is slightly larger than the genome size of the previously sequenced *A. nidulans* A4 strain of ∼31 Mb.[24]

Using the first adenylation (A) domain of the enzyme TqaA (accession number ADY16697) from the tryptoquialanine pathway[25] as BLAST query, we were able to find 26 putative NRPS genes in *E. rugulosa* (Supporting Information, Table S2). Four out of the 26 NRPSs, which we denoted as ErNRPSs, are NRPSs with five or more modules. Two of these four ErNRPSs, annotated as ErNRPS99 and ErNRPS57, are also found in the model fungi *A. nidulans* A4 that does not produce **1**. ErNRPS99 is likely a homologue of EasA found in the biosynthesis of emericellamide[26] due to the high shared sequence identity between the two. ErNRPS57, on the other hand, contains a terminal reductive domain (R) similar to what is found in peptaibol synthetases.[27] Thus, this leaves ErNRPS284 and ErNRPS123 as echinocandin synthetase candidates. ErNRPS284 (Supporting Information, Table S2) contains five modules, which are insufficient for catalyzing the formation of the hexalipopeptide scaffold of **1** based on the colinearity hypothesis.[28] On the other hand, the six-module ErNRPS123 (799 kDa), which is annotated as *ecdA*, has the correct number of modules necessary for the assembly of **1**. However, performing A domain selectivity prediction using NRPS predictor[29] offered little confirmation that this is the correct echinocandin NRPS, with only the third A domain prediction (L-proline) matching the corresponding amino acid in **1** (L-proline or 4*R*-hydroxyl-L-proline) (Supporting Information, Table S3). Further bioinformatics analysis of the domain architecture of EcdA, as well as the genes within the *ecd* cluster, gave indications that this is most likely the correct gene cluster that is consistent with some of the expected biosynthetic transformations required for the assembly of **1**. First, EcdA has

Table 1. Echinocandin B Biosynthetic Gene Cluster

gene	length (aa)	conserved domain/function	nearest BLAST hit (identity, similarity)
ecdB	545	fungal transcription factor	*A. fumigatus* AFUA_048990 (72%, 80%)
ecdC	556	transporter (MFS)	*N. fischeri* NFIA_042010 (70%, 79%)
ecdD	541	transporter (MFS)	*N. fischeri* NFIA_042010 (87%, 93%)
ecdE	703	glycosyl hydrolase	*N. fischeri* NFIA_042050 (78%, 88%)
ecdF	508	glycosidase	*P. purpurogenum* BAA12320 (71%, 84%)
ecdA	7260	NRPS (TCATCATCATCATCATCATC)	*T. reesei* EGR45389 (33%, 52%)
ecdG	338	non-heme iron, α-ketoglutarate dependent dioxygenase	*C. militaris* CCM_03049 (31%, 50%)
ecdH	503	cytochrome P450 heme-iron-dependent oxygenase	*N. hematococca* 100691 (28%, 48%)
ecdI	559	fatty-acyl-AMP ligase	*A. nidulans* A4 AN3490 (52%, 67%)
ecdJ	668	hypothetical protein	*A. capsulatus* 05345 (34%, 47%)
ecdK	332	non-heme iron, α-ketoglutarate dependent dioxygenase	*T. reesei* TRIREDRAFT_58580 (43%, 63%)
ecdL	1479	multidrug transporter (ABC)	*M. anisoplae* MAA_01638 (40%, 59%)

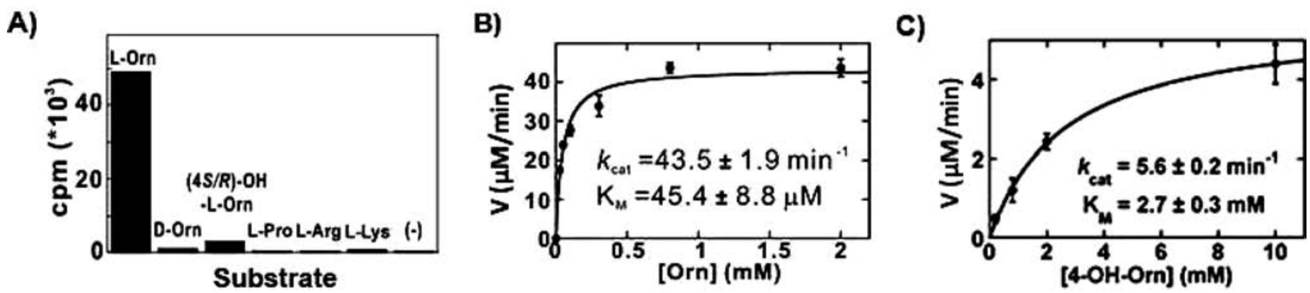


Figure 2. ATP–[^{32}P]PP$_i$ exchange assay of EcdA-M1 (A). The adenylation (A) domain of EcdA-M1 shows preference toward L-ornithine. (B) Michaelis–Menten plot for the A domain of EcdA-M1 with L-ornithine as substrate. (C) Michaelis–Menten plot for the A domain of EcdA-M1 with 4(*R*/*S*)-OH-L-ornithine as substrate.

a terminal condensation domain (C_T) that has been shown to catalyze the cyclization of NRPS products in fungi,[30] in agreement with the anticipated macrocyclization of the hexapeptide. Furthermore, in proximity to *ecdA* is *ecdI* (Figure 1A and Table 1), an acyl-AMP ligase homologue gene, giving a plausible route for lipoinitiation of the NRPS. In addition, other genes adjacent to *ecdA* encode putative non-heme iron, α-ketoglutarate dependent oxygenases (*ecdG* and *ecdK*), and a cytochrome P450 heme-iron-dependent oxygenase (*ecdH*) (Table 1), indicating that the nascent NRPS product may undergo several oxygenation steps, as anticipated for the biosynthesis of **1**. Other genes flanking the putative *ecd* cluster encode a fungal transcription factor gene (*ecdB*), three transporter protein genes (*ecdC*, *ecdD* and *ecdL*), a glycosyl hydrolase gene (*ecdE*), a glycosidase gene (*ecdF*), and a gene encoding for a protein with no predicted conserved domain (*ecdJ*). Interestingly, genes encoding the biosynthesis of unnatural amino acids such as L-homotyrosine are not present in the vicinity of *ecd* gene cluster, but are found to reside elsewhere on the genome (see below).

Verification of the Role of *ecdA* in the Biosynthesis of 1. To confirm the production of **1**, *E. rugulosa* was grown in medium 2 and the metabolite extracted from the fermentation broth in the same manner as previously described.[7] A high-resolution LCMS trace of the extracted metabolites shows a peak at 17.6 min with *m*/*z* of 1042.5700 $[M + H - H_2O]^+$ corresponding to the theoretical mass-to-charge ratio of **1** (*m*/*z* = 1042.5707 $[C_{52}H_{81}N_7O_{16} + H - H_2O]^+$) (Figure 1B). Purification and subsequent characterization by NMR (Supporting Information, Table S3 and Figures S13–S16) and comparison with authentic standard (Supporting Information, Figure S1) verified that the compound is indeed **1**.

To verify whether the *ecd* cluster is responsible for the production of **1**, we developed a gene deletion method for *E. rugulosa* based on previous methods for *A. nidulans* A4,[31] using the glufosinate resistance gene *bar* as selection marker (see Materials and Methods).[32] A gene deletion cassette containing portions of *ecdA* internally disrupted by the *bar* gene driven by the *trpC* promoter was introduced into *E. rugulosa* protoplasts via PEG-mediated transformation, and the resulting glufosinate-resistant strains were selected. A bioassay-guided knockout screening was developed in which individual fungal colonies are spotted on plates preinoculated with *C. albicans*. Approximately 100 mutants were screened by both loss of anti-*Candida* activity, and PCR-based screening using a *bar* gene primer and a primer found outside the knockout cassette. One mutant (Δ*ecdA* I-16) was isolated that lost the ability to inhibit the growth of *C. albicans* (Supporting Information, Figure S2) and also showed the correct PCR-amplified product (Supporting Information, Figure S3). LCMS analysis of the extracted metabolites after 7 days of growth showed no production of **1** (Figure 1B), suggesting that EcdA is the NRPS responsible for the biosynthesis of **1**.

Characterization of the Adenylation Domain (A) of the First Module of EcdA. Guided by the results of the gene knockout of *ecdA*, we proceeded to determine the amino acid specificity of the first A domain of EcdA by cloning the initiation module of EcdA, which is fused to a thiolation domain (T_0) that is likely the site of lipid attachment (EcdA-M1, T_0CAT_1). The 130 kDa protein was expressed in *E. coli* BL21(DE3) cells to a final titer of ~20 mg/L and purity of >95% (Supporting Information, Figure S4). Amino acid-dependent ATP–[^{32}P]PP$_i$ exchange assay with 2 μM of EcdA-M1 at ambient temperature and incubation time of 30 min showed that A_1 could activate L-ornithine (L-Orn) and

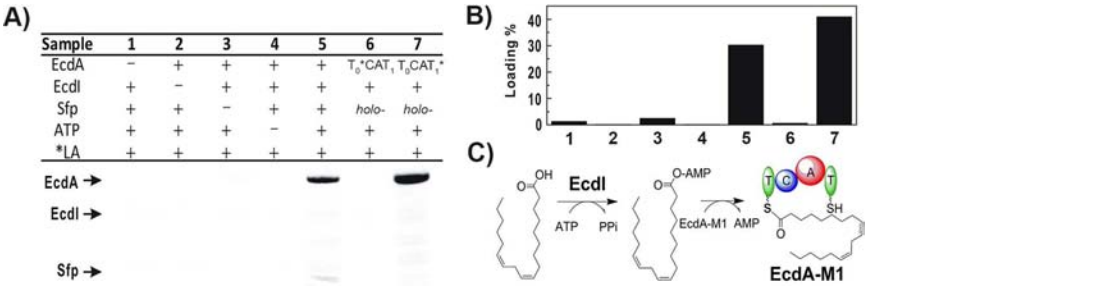


Figure 3. Loading assay of EcdA-M1. (A) EcdA-M1, converted to the holo form in vitro by Sfp, is loaded with [^{14}C]linoleic acid only in the presence of the fatty-acyl-AMP ligase EcdI and ATP, as shown in the autoradiogram (lane 5). Loading of EcdA-M1 variants with [^{14}C]linoleic acid under the same condition as sample 5. (B) Quantification of the percentage of EcdA-M1 loaded with radioactive linoleic acid. Samples 1−7 carried out under identical conditions as in A. (C) Linoleic acid is activated by EcdI to form linoleyl-AMP which is subsequently transferred to EcdA-M1.

(4*R*/*S*)-4-hydroxyl-L-ornithine (4-OH-L-Orn). D-Ornithine and other amino acids with basic side chains such as L-lysine and L-arginine were not activated (Figure 2A). Full kinetic analysis shows that L-Orn (k_{cat} = 43.5 ± 1.9 min^{-1}, K_M = 45.4 ± 8.8 μM) is about a 500-fold better substrate than 4-OH-L-Orn (k_{cat} = 5.6 ± 0.2 min^{-1}, K_M = 2.7 ± 0.3 mM) as judged by k_{cat}/K_M ratios (Figure 2B,C), suggesting that oxidation at Cγ of L-ornithine most likely occurs after loading of L-Orn to T_1. The activation of L-Orn by the first module of EcdA further supports the link between the NRPS and the biosynthesis of **1**.

EcdI-Catalyzed Lipoinitiation of EcdA. The addition of the lipid chain in the biosynthesis of lipopeptides, also known as lipoinitiation,[33,34] is a critical step not only because it connects the fatty acid pool and NRPS biosynthesis but also due to the importance of the lipid to the antimicrobial activities.[14] Several mechanisms exist in activating and transferring the lipid chain to the NRPS.[34] These include fusion of a fatty acid synthetase (FAS)-like module to the N-terminus of a NRPS (mycosubtilin)[35] and transfer of the lipid chain to a dissociated T domain (daptomycin)[36,37] or coenzyme A (surfactin)[33] by a fatty acyl ligase, followed by condensation with an aminoacyl adenylate catalyzed by the first C domain. Since EcdA only contains six A domains, one for each of the residues in the hexapeptide scaffold, we reasoned that EcdI, containing an AMP-binding domain, might be responsible for the formation of the activated form of linoleic acid and its transfer to the T_0 domain of EcdA. To test this hypothesis, we cloned EcdI into an *E. coli* expression vector and expressed the *N*-terminal His-tagged protein in BL21(DE3) with a titer of ~30 mg/L and purity of ~95% (Supporting Information, Figure S5). Using the purified phosphopantetheinyl transferase Sfp from *Bacillus subtilis*,[38] we converted apo EcdA-M1 into its holo form in vitro. After coincubation of holo EcdA-M1 with EcdI, ATP, and [^{14}C]linoleic acid, the assay mixture was analyzed using SDS−PAGE and autoradiography. The autoradiogram shows a strong radiolabeling of the 130 kDa band, indicating that the [^{14}C]linoleic acid is covalently bound to EcdA-M1 (Figure 3A). In contrast, in the absence of EcdI, ATP, or Sfp, nearly no labeling of EcdA-M1 can be detected, confirming the proposed mechanism (Figure 3C). Quantification of the radiolabel suggests that ~30% of EcdA-M1 is loaded with the labeled substrate (Figure 3B), consistent with fractional stoichiometries seen in other labeling studies.[39,40]

To pinpoint which of the two thiolation domains in EcdA initiation module (T_0CAT_1) is acylated by the linoleoyl starter unit in the presence of EcdI, single mutations to the active site serines in the two thiolation domains were made. The S47A mutant (T_0*CAT_1) prevents phosphopantetheinylation of the initiation T domain but leaves the T_1 domain available for conversion into the holo form that should be capable of undergoing covalent loading with [^{14}C]-L-ornithine. Conversely, the corresponding S1127A mutant (T_0CAT_1*) will not be phosphopantetheinylated at T_1 but should be able to load [^{14}C]linoleate. Both the S47A and the S1127A mutants were expressed from *E.coli* BAP1,[41] which coexpresses Sfp for thiolation domain phosphopantetheinylation (thus no Sfp preincubation is required) (Supporting Information, Figure S6). As shown in Figure S7 (Supporting Information), [^{14}C]linoleate is loaded onto the T_0CAT_1* but not the T_0*CAT_1 variant of the EcdA initiation module, as assessed by both autoradiography and by radioactive counting of protein precipitated via addition of trichloroacetic acid. The complementary result is seen for [^{14}C]-L-ornithine covalent loading in which T_0*CAT is labeled but T_0CAT_1* is not (Supporting Information, Figure S8). The relatively low amount of L-Orn radioactivity may reflect the lability of the thioester due to the attack by the Nδ of the ornithine side chain to the activated carbonyl carbon followed by the subsequent release of the cyclic δ-lactam, a known propensity of ornithine thioesters.[42]

In previously characterized lipopeptide NRPSs, the incorporation of fatty acids into the assembly line requires the formation of either acyl-CoA[33] or acyl-AMP[36] prior to loading onto the NRPS. To explore the mechanism of linoleic acid activation by EcdI, we removed excess CoA after conversion of apo EcdA-M1 to its holo form, followed by addition of 0, 0.5, or 5.0 mM CoA to the assay. The reaction rates for the acyl loading step, as determined by the fraction of EcdA-M1 loaded with labeled substrate at different time points, are essentially the same (Supporting Information, Figure S9), ruling out the CoA-dependent mechanism and indicating that the acyl-AMP is directly transferred to EcdA-T_0 by EcdI (Figure 3C).

In order to probe the substrate specificity of EcdI, we coincubated holo-EcdA-M1 with alternative acyl substrates. Palmitic acid (C-16) showed a similar degree of loading to EcdA-M1 compared to linoleic acid (27% vs 30% for linoleic acid). Decreasing the length of the acyl chain to C-10 ([^{14}C]decanoic acid), on the other hand, dramatically reduced the loading of EcdA to ~8% at time points where the loading of linoleic acid reaches saturation. Nevertheless, this range of fatty acyl chain lengths that can be transferred to EcdA-M1 indicates there is room for incorporation of alternative lipid starter molecules into the structure of **1**. Incubation of EcdI and EcdA-M1 with non-fatty acid substrate such as benzoic acid, an aryl carboxylate, did not show covalent loading onto T_0 (Supporting Information, Figure S10).

Discovery of the L-Homotyrosine Biosynthetic Gene Cluster. One of the characteristic features of the echinocandin family is the presence of nonproteinogenic 3*S*,4*S*-dihydroxyl-L-homotyrosine in the fourth amino acid position of the cyclic hexapeptide. This residue is derived from the dihydroxylation of L-homotyrosine either prior to or after incorporation into the peptide scaffold. It was previously shown that L-homotyrosine in the scaffold of **4** is derived from the condensation of 4-hydroxyphenyl-pyruvate and acetate to form 2-(4-hydroxybenzyl)-malate.[43] Likewise, the biosynthesis of L-homophenylalanine in watercress plants is proposed to follow a pathway analogous to leucine biosynthesis,[44] beginning with the condensation of acetyl-CoA and phenylpyruvate to form benzylmalic acid. In order to find putative 2-(4-hydroxybenzyl)-malate synthase (HBMS) that may be involved in L-homotyrosine biogenesis, we searched the *E. rugulosa* genome for genes encoding the functionally analogous isopropyl-malate synthase (IPMS), the enzyme that catalyzes the condensation of α-ketovalerate with acetyl-CoA in leucine biosynthesis. We used the IPMS gene from *Mycobacterium tuberculosis* (accession number MT3813) as BLAST query. Two significant hits were found, ErIPMS48 and ErIPMS66, both of which are found outside the contig containing the *ecd* gene cluster. ErIPMS48 shares 99% protein sequence identity to the predicted *A. nidulans* A4 housekeeping IPMS (AN0804) as well as >85% identity to predicted IPMS in other ascomycetes, suggesting that this gene encodes for the actual IPMS involved in leucine biosynthesis in *E. rugulosa*. On the other hand, the ErIPMS66 protein sequence has a lower similarity to AN0804 (43% protein identity), while such a second IPMS homologue is not present in *A. nidulans* A4. Downstream of ErIPMS66 (designated as *htyA*) are genes putatively encoding a transaminase (*htyB*), a 3-isopropyl-malate dehydrogenase homologue (*htyC*), and an isopropyl-malate isomerase homologue (*htyD*), all possibly involved in biosynthesis of L-homotyrosine (Table 2 and Figure 4A). Moreover, the presence of immediately upstream genes encoding for a predicted non-heme iron, α-ketoglutarate dependent oxygenase (*htyE*), and cytochrome P450 oxygenase (*htyF*) is consistent with the requirement of Cβ and Cγ hydroxylation of L-homotyrosine in **1**. Thus, we reasoned that the gene cluster, here designated as *hty*, may be responsible for the biosynthesis of L-homotyrosine in *E. rugulosa*.

Table 2. Genes Found within the *hty* Cluster

gene	length (aa)	conserved domain/function	nearest BLAST hit (identity, similarity)
htyF	683	heme-dependent P450 oxygenase	*T. stipitatus* TSTA_09270 (45%, 64%)
htyE	329	non-heme iron, α-ketoglutarate dependent dioxygenase	*A. terreus* ATEG09098 (50%, 66%)
htyA	584	isopropyl malate synthase	*A. alternata* BAI44742 (51%, 68%)
htyB	379	transaminase	*A. alternata* BAI44740 (55%, 71%)
htyC	366	isopropyl malate dehydrogenase	*A. alternata* BAI44741 (63%, 73%)
htyD	877	aconitase	*A. alternata* BAI44743 (57%, 68%)

To confirm this hypothesis, we genetically disrupted the *htyA* gene in the same manner as for construction of Δ*ecdA*. Screening of ~100 colonies yielded three PCR-positive mutants (Supporting Information, Figure S11). All three Δ*htyA* mutants lost the ability to inhibit the growth of *C. albicans* under screening conditions (Supporting Information, Figure S2), accompanied by the loss of production of **1** (Figure 4B). To chemically complement the Δ*htyA* mutant, 0.1 mg/mL of L-homotyrosine was supplemented to the growth media. As expected, adding free homotyrosine restored the ability of the mutant to inhibit *Candida* (Supporting Information, Figure S2), as well as the production of **1** (Figure 4B) to wild-type levels. As a negative control, feeding of L-homotyrosine to Δ*ecdA* I-16 mutant did not restore the production of **1** (Supporting Information, Figure S1). Therefore, based on whole genome sequencing, we were able to identify a separately located gene cluster that is responsible for the biosynthesis of an unnatural amino acid building block for the *ecd* pathway.

On the basis of the putative functions of of HtyA-D (Table 2), the biosynthesis of L-homotyrosine is predicted to be as follows (Figure 4C): 4-hydroxyphenyl-pyruvate undergoes an aldol-type condensation by HtyA with the C-2 of acetyl-CoA followed by the release of CoA to form 2-(4-hydroxybenzyl)-malate. This is followed by isomerization of 2-(4-hydroxybenzyl)-malate to 3-(4-hydroxybenzyl)-malate by HtyD. Thereafter, 3-(4-hydroxybenzyl)-malate undergoes decarboxylation and oxidation to form 2-oxo-4-(4-hydroxybenzyl)butanoic acid, coupled to reduction of NAD^+ to NADH by HtyC. The product then undergoes transamination catalyzed by HtyB to form L-homotyrosine. The closest homologues of HtyA-D is found in a four-gene cassette from *Alternaria alternata* that is predicted to catalyze the formation of 2-amino-4-phenylvaleric acid (APVA) in AM-toxin.[45] Interestingly, this suggests that the HtyA-D homologues in the AM-toxin gene cluster must perform two cycles of the α-ketoacid elongation to afford APVA.

DISCUSSION

Echinocandins are a family of antifungal cyclic lipopeptides from ascomycetes. Through the sequencing of the genome of *E. rugulosa* NRRL 11440 and subsequent bioinformatics analysis of NRPS genes, we have identified *ecdA*, encoding a 799 kDa six-module NRPS, which is confirmed by gene deletion to be required for production of **1**. The *ecd* gene cluster is flanked by two microsyntenic blocks belonging to two different chromosomes in *A. nidulans* A4. Upstream of the *ecd* cluster are a group of genes that are syntenic to genes found in chromosome VII of *A. nidulans* A4 while the genes found downstream of the cluster is syntenic to genes found in *A. nidulans* A4 chromosome V (Supporting Information, Figure S12). This suggests that a chromosomal translocation occurred when *E. rugulosa* NRRL 11440 and *A. nidulans* A4 diverged from their common ancestor.

Another interesting result revealed by the study is the separation of the *ecd* gene cluster and the L-homotyrosine biosynthetic genes. By adding the distance to the nearest end of their respective contigs, the minimum distance between the two clusters is ~42.5 kb, assuming that both clusters are located in the same chromosome (Supporting Information, Figure S12). Moreover, similar to the phenomenon seen in the *ecd* gene cluster, the *hty* gene cluster is flanked by microsyntenic blocks from two different *A. nidulans* A4 chromosomes; upstream of *hty* cluster are genes that are syntenic to genes from chromosome VII of *A. nidulans* A4 while downstream are genes that are syntenic to genes found in the *A. nidulans* A4 chromosome VI (Supporting Information, Figure S12). This

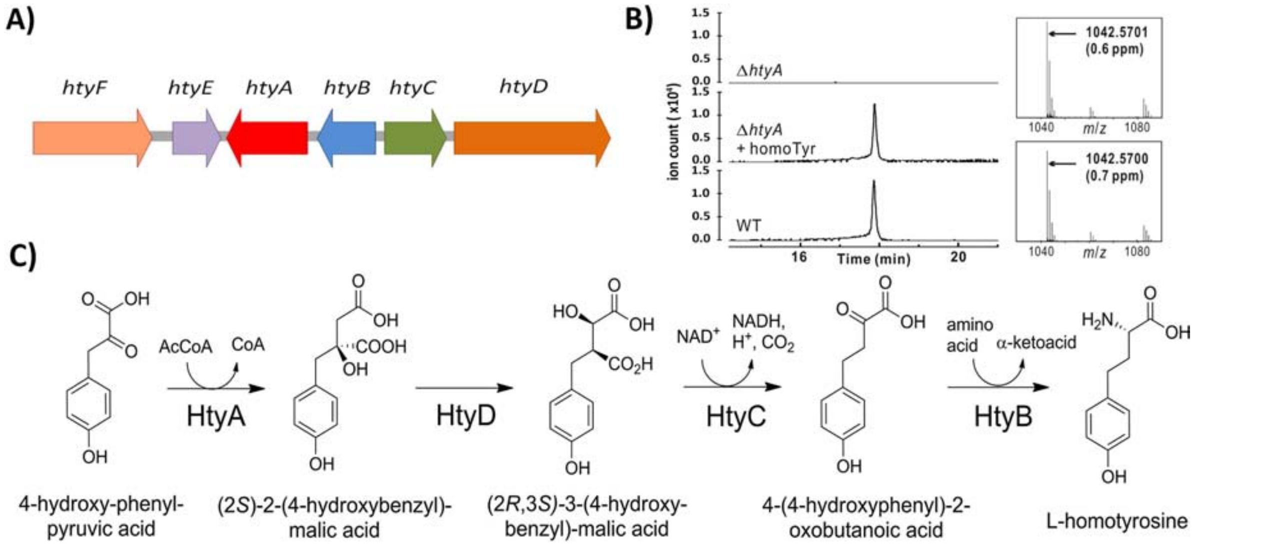
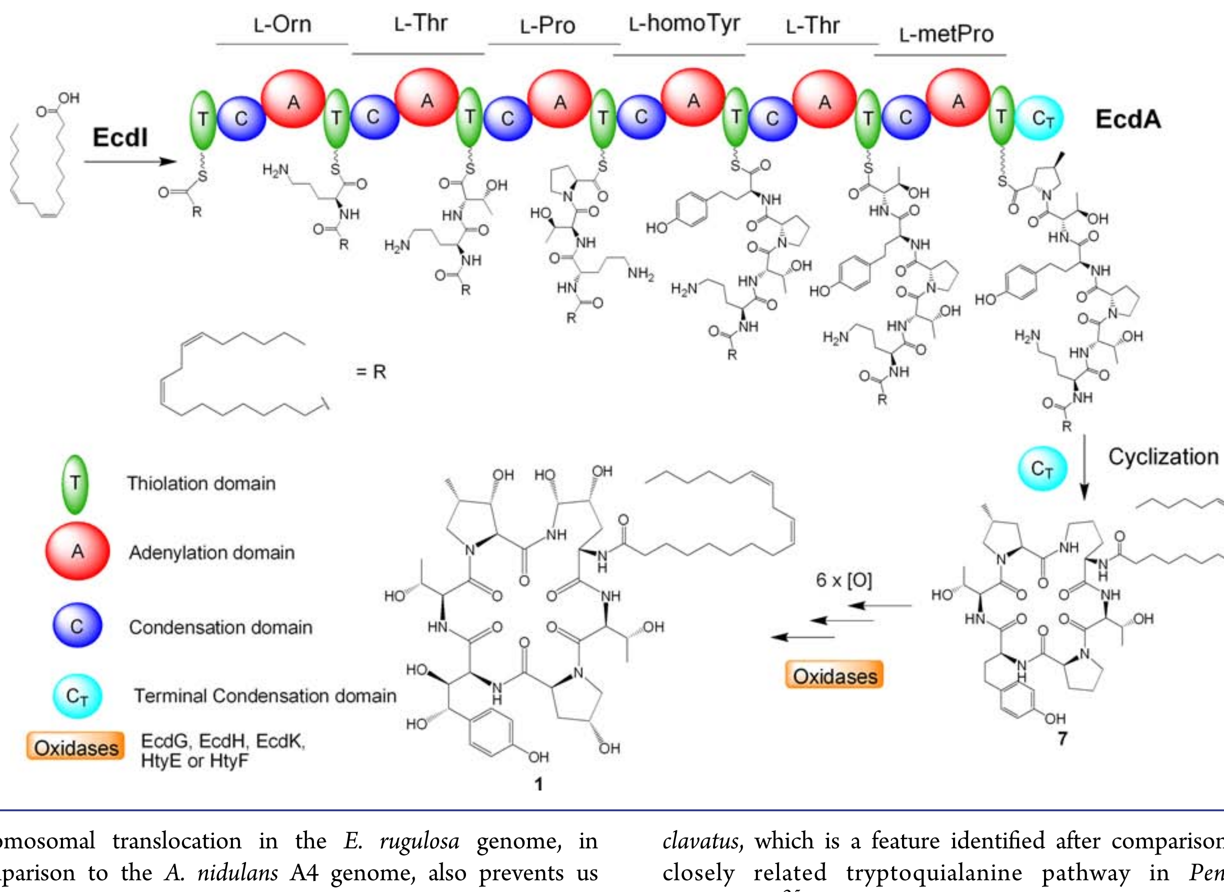


Figure 4. Biosynthesis of L-homotyrosine. (A) Organization of the *hty* gene cluster. (B) Disruption of *htyA* led to the loss of production of **1**, which was restored upon addition of L-homotyrosine to the culture. (C) Putative biosynthesis of L-homotyrosine.

Scheme 2. Putative Biosynthetic Pathway of Echinocandin B 1



chromosomal translocation in the *E. rugulosa* genome, in comparison to the *A. nidulans* A4 genome, also prevents us from mapping the *hty* and *ecd* gene cluster in relation to the *A. nidulans* A4 chromosomes. While separation of the biosynthetic genes for **1** is unusual, it is not unprecedented. Examples of the separation of secondary metabolite biosynthetic genes in fungi are found in the dothistromin pathway in *Dothistroma septosporum*,[46] the convergence of the orsellinic acid and anthrone biosynthetic pathways in *A. nidulans* to form spiroanthrones,[47] the austinol and dehydroaustinol pathways in *A. nidulans*,[48] and the putative tryptoquivaline pathway in *A. clavatus*, which is a feature identified after comparison to the closely related tryptoquialanine pathway in *Penicillium aethiopicum*.[25]

Due to the size of the full-length EcdA, which is a formidable challenge for in vitro evaluation, we dissected the EcdA initiation module to establish the predicted specificity for L-ornithine activation by A_1. The ~130 kDa, four-domain (T_0CAT_1) EcdA initiation module (EcdA-M1) was expressed from *E. coli*. Determination of selectivity of A_1 by amino acid-dependent ATP−[^{32}P]PP_i exchange assay established L-ornithine as the most likely substrate. Heterologous expression of EcdA-M1 and EcdI, an acyl-AMP ligase homologue, allowed us to investigate the mechanism of lipoinitiation of EcdA. Co-incubation of EcdA-M1 with EcdI, ATP, and [^{14}C]linoleic acid indicated loading of linoleic acid to the initiation T_0 domain of EcdA. While much is known regarding the lipoinitiation strategies of the bacterial NRPS's,[33−37] this study presents the first in vitro characterization of the lipoinitiation of fungal NRPS. In comparison to the bacterial NRPS systems, EcdI acts analogously to DptE in daptomycin biosynthesis:[36] each adenylates the fatty acid substrate, which is subsequently transferred to the initiation T_0 domain. The difference between the two systems is that the acceptor thiolation domain is separate from the NRPS DptA in daptomycin biosynthesis but is fused at the N-terminal of EcdA for the biosynthesis of **1**. Hence it is expected that the smaller EcdI requires interaction with a thiolation domain that is part of a multimodular megasynthetases. This appears to be a commonly employed strategy among fungal PKSs and NRPSs, in which smaller enzyme partners are recruited to interact with the thiolation and acyl carrier protein domains.[49,50] Currently, semisynthesis of echinocandins used for medical use such as **3** and **6** require the biological deacylation of their parent compound via feeding to separate cultures of *Actinoplanes* spp. followed by chemical reacylation using protective chemistry.[12,14] Thus, deciphering the lipoinitiation strategy may also enable the engineered biosynthesis of approved derivatives of **1** containing alternative lipid groups.

The presence of acyl-AMP ligase homologue gene *ecdI* within the *ecd* gene cluster led us to investigate whether other organisms, in particular filamentous fungi, also have clustering of genes encoding an acyl-AMP ligase and an NRPS with an N-terminal T_0 domain. A search in the NCBI database revealed five genes for fatty-acyl-AMP ligase that share ≥39% identity with EcdI and are clustered with genes of NRPS with an initiation T_0 domain (Supporting Information, Table S6). Moreover, four of the five EcdI homologue genes are also clustered with highly reducing PKS genes in addition to NRPS genes, further hinting that these gene clusters may encode for lipopeptide biosynthetic enzymes. Furthermore, *easD* from *A. nidulans*, the only one characterized out of the five fatty acyl-CoA ligase genes, is shown to be involved in the biosynthesis of emericellamide. The emericellamide synthetase EasA, however, requires an additional acyltransferase EasC for lipoinitiation.[26] Such an acyltransferase is notably missing in the *ecd* gene cluster and is now proven not to be required for the linoleic acid loading in our in vitro assays.

On the basis of the results of the lipoinitiation by EcdI and determination of the selectivity of A_1, we can propose one possible pathway for the biosynthesis of **1** (Scheme 2): linoleoyl-AMP, produced by EcdI, is transferred to the initiation T_0 of EcdA. The linoleoyl-S-phosphopantetheinyl-T_0 is sequentially extended with L-ornithine, L-threonine, L-proline, L-homotyrosine, L-threonine, and 4*R*-methyl-L-proline to form the linear hexapeptide. Thereafter, the terminal condensation (C_T) performs macrocyclization of the NRPS product[30] and the cyclic scaffold **7** is released from EcdA. In this pathway, in which all the hydroxylation reactions are proposed to occur following completion of the cyclic peptide, the unhydroxylated precursor **7** will undergo six rounds of hydroxylation. In congruence to modification of the residues found in **1**, five hydroxylase genes (*ecdG*, *ecdH*, *ecdK*, *htyE* and *htyF*) are embedded within the *ecd* and *hty* clusters. At this point, it is not possible to assign the hydroxylases based on sequence alone, as all are proposed to act on sp^3 hybridized carbon atoms. It was previously shown that L-proline hydroxylation to 4*R*-hydroxyl-L-proline in protein scaffolds is catalyzed by an non-heme iron, α-ketoglutarate dependent oxygenase.[51] Thus, it is likely that the hydroxylation of L-proline in **1** might be catalyzed by any of EcdG, EcdK, or HtyE. However, the possibility that a P450 oxidase such as EcdH or HtyF can catalyze the reaction cannot be excluded. On the other hand, the formation of vicinal diols to give the 4*R*,5*R*-dihydroxyl-L-ornithine and 3*S*,4*S*-dihydroxyl-L-homotyrosine residue are more novel compared to that of the modified proline and may each require two hydroxylases to separately install the two hydroxyl groups. Due to the lability of the resulting hemiaminal,[18] we anticipate that Cδ hydroxyl group in L-ornithine must be installed after peptide macrocycle formation and is likely the last step in the oxidative tailoring cascade. An equally likely pathway to **1** is that some of the amino acids are hydroxylated prior to incorporation into the hexapeptide. The most plausible candidate for this scenario is 4*R*-hydroxyl-L-proline, which is a commonly observed unnatural amino acid in different organisms.[52−54] The exact timing and substrate of this plethora of hydroxylation enzymes will be determined in subsequent efforts through a combination of A domain activation assays, genetic knockouts of the candidate oxygenases, and in vitro biochemical investigation

In addition to L-homotyrosine and L-ornithine, **1** contains the nonproteinogenic amino acid 3*S*-hydroxyl-4*S*-methyl-L-proline, which is presumably derived from 4*R*-methyl-L-proline. Previous studies of **4**,[43] nostopeptolide,[55] and nostocyclopeptide[56] biosyntheses showed that that 4-methyl-L-proline originates from Cδ oxidation and subsequent cyclization of L-leucine. Oxidation of L-leucine to 5-hydroxyl-L-leucine was recently identified to be catalyzed by a non-heme iron, α-ketoglutarate dependent oxygenase in *Nostoc punctiforme*.[57] Thus, it is probable that one of the non-heme iron, α-ketoglutarate dependent oxygenase such as EcdG, EcdK, and HtyE can perform this reaction. However, neither the *ecd* nor the *hty* clusters harbor the genes encoding enzymes involved in the reactions downstream to oxidation of L-leucine. A genome-wide search for genes for pyrroline-5-carboxylate (P5C) reductase homologue, proposed to catalyze the final step of 4-methyl-L-proline biosynthesis,[55] revealed the presence of four candidate genes in *E. rugulosa*. However, an ortholog for each of the four candidate genes is also present in *A. nidulans* A4 (Supporting Information, Table S5), so at present it is not yet clear which, if any, is involved in synthesis of the 4*R*-methyl-L-proline for **1**.

In conclusion, we report the discovery of the biosynthetic gene cluster of **1**, the first such cluster for a member of the medically relevant fungal lipopeptide family of compounds. This study has also uncovered the genetic basis for the biosynthesis of the nonproteinogenic L-homotyrosine. The mechanism and timing of the hydroxylation of the residues L-homotyrosine, L-proline, 4*R*-methyl-L-proline, and L-ornithine are intriguing features of the biosynthesis of **1** and are currently under investigation in our groups.

■ MATERIALS AND METHODS

General Methods and Materials. *E. rugulosa* (*A. nidulans* var. *roseus*) strain NRRL 11440 was obtained from Agricultural Research Services Culture Collection (Peoria, IL). Authentic standard for **1** was purchased from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA). Primers used in this study were ordered from Integrated DNA

Technologies and are listed in Table S7 (Supporting Information). Sequencing of heterologous expression constructs and knockout cassettes was performed by Laragen, Inc. (Culver City, CA). RNA for cDNA amplification was isolated using a RiboPure-Yeast Kit from Ambion. First-strand cDNA synthesis was performed using SuperScript III-First Strand Synthesis SuperMix (Invitrogen Corp.).

Illumina Hiseq2000 Sequencing and Bioinformatic Analysis. The genomic DNA for sequencing was isolated as described elsewhere from stationary liquid cultures.[58] Shotgun sequencing was performed at Ambry Genetics (Aliso Viejo, CA) using Illumina Hiseq 2000 with a read size of 157 bases, resulting in a total of ~17 Gbp reads. The Illumina sequencing reads were assembled using a hierarchical assembly method using the assembly programs SOAPdeNOVO[59] and Geneious and performed by the UCLA Hoffman Cluster. First, the ~17 Gbp total reads were assembled using SOAPdeNOVO using a *k*-mer size of 87 bp. The first-tier contigs were then assembled using Geneious to generate 433 s tier contigs with a N50 of 235 313 and total length of ~32 Mb.

The second-tier contigs were converted to BLAST database format for local BLAST search using standalone BLAST software (v. 2.2.18). Gene predictions were performed using the FGENESH program (Softberry) and manually checked by comparing with homologous gene/proteins in the GenBank database. Functional domains in the translated protein sequences were predicted using Conserved Domain Search (NCBI) or InterproScan (EBI). The nucleotide sequences for the *ecd* and *hty* gene clusters are deposited to Genbank database with accession numbers JX421685 and JX421684, respectively.

Fungal Transformation and Gene Disruption. Polyethylene glycol-mediated transformation of *E. rugulosa* NRRL 11440 was performed as done previously[31] with the following modifications: the spores from two plates were grown in 250 mL of GMM with 10 mM ammonium tartrate as sole nitrogen source for 16 h at 250 rpm and 28 °C, and 1 g of the germlings from the culture was used for digestion using 3 g of Vinotaste Pro enzyme mixture (Novozyme). The glufosinate resistance gene *bar* with an upstream *trpC* promoter was amplified from pBARGPE1 plasmid obtained from the Fungal Genetic Stock Center (Kansas City, MO). Glufosinate used for the selection of transformants was prepared by extracting twice with an equal volume of 1-butanol from commercial herbicide Finale (Bayer), which contains 11.33% (w/v) glufosinate ammonium.[60] Construction of the knockout cassette was performed by fusion PCR as described elsewhere.[25] Fungal genomic DNA from the transformants was isolated using a ZR-Fungal/Bacterial DNA Miniprep Kit (Zymo). Primers used for PCR screening are listed in Table S7 (Supporting Information).

Extraction and Characterization of 1. The echinocandin extraction method is based on US Patent 4,288,549.[7] Briefly, the strain was grown in 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)] at 28 °C for 7 days in 10 mL cultures for screening of transformants and 2 L for large-scale compound extraction. An equal volume of methanol was added to the whole fermentation broth and the mixture was shaken at 16 °C for 1 h. The mixture was then filtered and the pH adjusted to 4.0. The filtrate was extracted twice with equal volumes of chloroform. The concentrated extract was purified using Sephadex LH-20 resin using 1:1 methanol and chloroform solvent system. The fractions containing **1**, determined by LCMS, was further purified using HPLC with a C-18 column and a 40–95%, 20 min water:acetonitrile gradient. Purified **1** was further characterized via Agilent 6520 high-resolution Q-TOF/LCMS, and its 1D-^{1}H NMR, 2D-COSY, HSQC, and HMBC spectra were recorded using a Varian 600 MHz NMR spectrometer.

Anti-*Candida* Assay. Δ*ecdA* and Δ*htyA* mutants were grown in 10 mL of liquid medium 2 at static conditions for 7 days at 28 °C. Disks of 10 mm in diameter were cut from the fungal mat and transferred to yeast extract–peptone–dextrose (YPD) agar plates that were previously inoculated with *C. albicans* ATCC 90234. The cocultures were grown at 28 °C overnight. For chemical complementation of Δ*htyA* (Supporting Information, Figure S1), L-homotyrosine (0.1 mg/mL) was fed to the static liquid cultures at day 4 before transferring the mycelial disks from individual clones to the YPD plate preinoculated with *C. albicans* at day 7.

Cloning of EcdA-M1. The C-terminal boundary of EcdA-T_1 was predicted through alignment of gramicidin synthetase PCP (accession 1DNY) and fungal NRPSs with EcdA by ClustalW. EcdA-M_1 was amplified from *E. rugulosa* cDNA sample using the primer pair *Nde*I-EcdA-T_0 and *Eco*RI-EcdA-T_1 (see Supporting Information, Table S7). The amplicon was gel-purified, digested with *Nde*I and *Eco*RI, and ligated into pET28a expression vector to create pET28a-EcdA-M1. QuikChange I Site-Directed Mutagenesis Kit (Agilent Technologies) was used to clone the EcdA-M1 variants using pET28A-EcdA-M1 as template and ecdA-S47A-F and ecdA-S47A-R as primer pair to create S47A variant and ecdA-S1127A-F and ecdA-S1127A-R primer pair to create S1127A variant.

Cloning of EcdI. EcdI cDNA was amplified from *E. rugulosa* cDNA sample using primers EcdI-*Nde*I-F and EcdI-*Eco*RI-R. The amplicon was ligated into PCR-blunt vector and was transformed into TOP10 cells. The plasmid was sequenced to confirm correct splicing of the transcript. The plasmid bearing the correct sequence was digested with *Nde*I and *Eco*RI and the insert was cloned into pET28a vector to create pET28a-EcdI.

Heterologous Expression of EcdA-M1 and EcdI. pET28a-EcdA-M1 or pET28a-EcdI was transformed into BL21 (DE3), and the cells were grown in 500 mL of LB at 37 °C and 250 rpm. When the OD600 reading reached 0.4, the cultures were cooled to 16 °C and protein expression was induced by addition of 60 μM IPTG. After 16 h of shaking at 16 °C, the cells were pelleted and resuspended in buffer A (50 mM Tris-HCl, pH 7.9, 5 mM NaCl, 1 mM DTT) with 20 mM imidazole. The cells were lysed via sonication and centrifuged at 4 °C at 15 000 rpm. Nickel–NTA resin was added to the supernatant and was gently stirred at 4 °C for 2 h. The protein/resin mixture was loaded into a gravity flow column, and the His-tagged proteins were purified with increasing concentration of imidazole in buffer A. Additional anion exchange column purification for EcdI was performed using a 80 min gradient of 0–100% buffers A and B (50 mM Tris-HCl, 1 M NaCl, 2 mM DTT) using a MonoQ 10/100 L anion exchange column (GE Healthcare Life Sciences). The fractions containing the desired protein were concentrated using an ultrafiltration column (100 kDa cutoff, for EcdA-M1 and 30 kDa cutoff for EcdI). Protein concentration was determined via by UV absorbance at λ = 280 nm.

ATP–[^{32}P]PP$_i$ Exchange Assays. A typical reaction mixture (500 μL) contained 1.0 μM EcdA, 2 mM amino acid substrate (unless specified), 5 mM ATP, 10 mM $MgCl_2$, 5 mM Na[^{32}P]pyrophosphate ($NaPP_i$) (~1.8 × 10^6 cpm/mL), and 50 mM Tris-HCl (pH 8). Mixtures were incubated at ambient temperature for regular time intervals (e.g., 5 min), and 150 μL aliquots were removed and quenched with 500 μL of a charcoal suspension (100 mM $NaPP_i$, 350 mM $HClO_4$, and 16 g/L charcoal). The mixtures were vortexed and centrifuged at 13 000 rpm for 3 min. Pellets were washed twice with 500 μL of wash solution (100 mM $NaPP_i$ and 350 mM $HClO_4$). Each pellet was resuspended in 500 μL wash solution and added to 10 mL of Ultima Gold scintillation fluid. Charcoal-bound radioactivity was measured using a Beckman LS 6500 scintillation counter.

Loading of [^{14}C]Substrate onto NRPS. The assay was carried out in two steps. First, a 50 μL reaction containing 10 μM EcdA-M1, 20 μM Sfp, 1 mM CoA, 10 mM $MgCl_2$, 1 mM TCEP, and 50 mM HEPES (pH 7.0) was incubated at ambient temperature for 30 min to convert apo EcdA-M1 to its holo form. Afterward, 8 μM EcdI, 5 mM ATP, and ~40 μM [^{14}C]substrate (~4.4 × 10^6 cpm/mL) were added, and the mixture was incubated for 30 min. The reaction was quenched by 600 μL of acetonitrile for the assays with [^{14}C]acyl substrate or 600 μL of 10% trichloroacetic acid for [^{14}C]-L-ornithine with addition of 100 μL of 1 mg/mL BSA. The mixture was vortexed and centrifuged at 13 000 rpm for 3 min. The pellet was then washed twice with 600 μL of acetonitrile, dissolved in 250 μL of formic acid, added into 10 mL of Ultima Gold scintillation fluid, and subjected to a Beckman LS 6500 scintillation counter.

■ ASSOCIATED CONTENT

Supporting Information

Additional bioinformatics data, in vitro assay results, and spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

yitang@ucla.edu; christopher_walsh@hms.harvard.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work is supported by NIH Grants 1R01GM092217 and 1DP1GM106413 to Y.T., 1R01GM49338 to C.T.W., and NRSA GM08496 to R.A.C. Genome assembly of *E. rugulosa* NRRL 11440 was performed on the UCLA Academic Technological Services (ATS) Hoffman2 cluster. We would like to thank Qiyang Hu for his assistance in the use of the UCLA Hoffman cluster, Dr. Jaclyn M. Winter for assistance in RNA extraction, and Prof. Scott G. Filler from the UCLA Harbor Medical Center for providing the *C. albicans* strain.

■ REFERENCES

(1) Vincent, J. L.; Rello, J.; Marshall, J.; Silva, E.; Anzueto, A.; Martin, C. D.; Moreno, R.; Lipman, J.; Gomersall, C.; Sakr, Y.; Reinhart, K. *JAMA, J. Am. Med. Assoc.* **2009**, *302*, 2323.

(2) Arendrup, M. C. *Curr. Opin. Crit. Care* **2010**, *16*, 445.

(3) Mayr, A.; Aigner, M.; Lass-Florl, C. *Mycoses* **2012**, *55*, 27.

(4) Andes, D. R.; Safdar, N.; Baddley, J. W.; Playford, G.; Reboli, A. C.; Rex, J. H.; Sobel, J. D.; Pappas, P. G.; Kullberg, B. J. *Clin. Infect. Dis.* **2012**, *54*, 1110.

(5) Kett, D. H.; Shorr, A. F.; Reboli, A. C.; Reisman, A. L.; Biswas, P.; Schlamm, H. T. *Crit. Care* **2011**, *15*, R253.

(6) Benz, F.; Knüsel, F.; Nüesch, J.; Treichler, H.; Voser, W.; Nyfeler, R.; Keller-Schierlein, W. *Helv. Chim. Acta* **1974**, *57*, 2459.

(7) Boeck, L.; Kastner, R. (Eli Lilly and Co.) U.S. Patent 4,288,549, 1981.

(8) Schwartz, R. E.; Sesin, D. F.; Joshua, H.; Wilson, K. E.; Kempf, A. J.; Goklen, K. A.; Kuehner, D.; Gailliot, P.; Gleason, C.; White, R.; Inamine, E.; Bills, G.; Salmon, P.; Zitano, L. *J. Antibiot. (Tokyo)* **1992**, *45*, 1853.

(9) Mizuno, K.; Yagi, A.; Satoi, S.; Takada, M.; Hayashi, M. *J. Antibiot. (Tokyo)* **1977**, *30*, 297.

(10) Strobel, G. A.; Miller, R. V.; Martinez-Miller, C.; Condron, M. M.; Teplow, D. B.; Hess, W. M. *Microbiology* **1999**, *145* (Pt 8), 1919.

(11) Roy, K.; Mukhopadhyay, T.; Reddy, G. C.; Desikan, K. R.; Ganguli, B. N. *J. Antibiot. (Tokyo)* **1987**, *40*, 275.

(12) Boeck, L. D.; Fukuda, D. S.; Abbott, B. J.; Debono, M. *J. Antibiot. (Tokyo)* **1989**, *42*, 382.

(13) Debono, M.; Abbott, B. J.; Fukuda, D. S.; Barnhart, M.; Willard, K. E.; Molloy, R. M.; Michel, K. H.; Turner, J. R.; Butler, T. F.; Hunt, A. H. *J. Antibiot. (Tokyo)* **1989**, *42*, 389.

(14) Debono, M.; Turner, W. W.; LaGrandeur, L.; Burkhardt, F. J.; Nissen, J. S.; Nichols, K. K.; Rodriguez, M. J.; Zweifel, M. J.; Zeckner, D. J.; Gordee, R. S.; Tang, J.; Parr, T. R. *J. Med. Chem.* **1995**, *38*, 3271.

(15) Bouffard, F. A.; Zambias, R. A.; Dropinski, J. F.; Balkovec, J. M.; Hammond, M. L.; Abruzzo, G. K.; Bartizal, K. F.; Marrinan, J. A.; Kurtz, M. B.; McFadden, D. C.; Nollstadt, K. H.; Powles, M. A.; Schmatz, D. M. *J. Med. Chem.* **1994**, *37*, 222.

(16) Tomishima, M.; Ohki, H.; Yamada, A.; Takasugi, H.; Maki, K.; Tawara, S.; Tanaka, H. *J. Antibiot. (Tokyo)* **1999**, *52*, 674.

(17) Joseph, J. M.; Kim, R.; Reboli, A. C. *Expert Opin Pharmacother* **2008**, *9*, 2339.

(18) Hensens, O. D.; Liesch, J. M.; Zink, D. L.; Smith, J. L.; Wichmann, C. F.; Schwartz, R. E. *J. Antibiot. (Tokyo)* **1992**, *45*, 1875.

(19) Leonard, W. R., Jr.; Belyk, K. M.; Conlon, D. A.; Bender, D. R.; DiMichele, L. M.; Liu, J.; Hughes, D. L. *J. Org. Chem.* **2007**, *72*, 2335.

(20) Kurokawa, N.; Ohfune, Y. *J. Am. Chem. Soc.* **1986**, *108*, 6043.

(21) Evans, D. A.; Weber, A. E. *J. Am. Chem. Soc.* **1987**, *109*, 7151.

(22) Kurokawa, N.; Ohfune, Y. *Tetrahedron* **1993**, *49*, 6195.

(23) Toth, V.; Nagy, C. T.; Miskei, M.; Pocsi, I.; Emri, T. *Folia Microbiol. (Praha)* **2011**, *56*, 381.

(24) Galagan, J. E.; Calvo, S. E.; Cuomo, C.; Ma, L. J.; Wortman, J. R.; Batzoglou, S.; Lee, S. I.; Basturkmen, M.; Spevak, C. C.; Clutterbuck, J.; Kapitonov, V.; Jurka, J.; Scazzocchio, C.; Farman, M.; Butler, J.; Purcell, S.; Harris, S.; Braus, G. H.; Draht, O.; Busch, S.; D'Enfert, C.; Bouchier, C.; Goldman, G. H.; Bell-Pedersen, D.; Griffiths-Jones, S.; Doonan, J. H.; Yu, J.; Vienken, K.; Pain, A.; Freitag, M.; Selker, E. U.; Archer, D. B.; Penalva, M. A.; Oakley, B. R.; Momany, M.; Tanaka, T.; Kumagai, T.; Asai, K.; Machida, M.; Nierman, W. C.; Denning, D. W.; Caddick, M.; Hynes, M.; Paoletti, M.; Fischer, R.; Miller, B.; Dyer, P.; Sachs, M. S.; Osmani, S. A.; Birren, B. W. *Nature* **2005**, *438*, 1105.

(25) Gao, X.; Chooi, Y. H.; Ames, B. D.; Wang, P.; Walsh, C. T.; Tang, Y. *J. Am. Chem. Soc.* **2011**, *133*, 2729.

(26) Chiang, Y. M.; Szewczyk, E.; Nayak, T.; Davidson, A. D.; Sanchez, J. F.; Lo, H. C.; Ho, W. Y.; Simityan, H.; Kuo, E.; Praseuth, A.; Watanabe, K.; Oakley, B. R.; Wang, C. C. *Chem. Biol.* **2008**, *15*, 527.

(27) Wiest, A.; Grzegorski, D.; Xu, B. W.; Goulard, C.; Rebuffat, S.; Ebbole, D. J.; Bodo, B.; Kenerley, C. *J. Biol. Chem.* **2002**, *277*, 20862.

(28) Marahiel, M. A.; Stachelhaus, T.; Mootz, H. D. *Chem. Rev.* **1997**, *97*, 2651.

(29) Röttig, M.; Medema, M. H.; Blin, K.; Weber, T.; Rausch, C.; Kohlbacher, O. *Nucleic Acids Res.* **2011**, *39*, W362.

(30) Gao, X.; Haynes, S. W.; Ames, B. D.; Wang, P.; Vien, L.; Walsh, C. T.; Tang, Y. *Nat. Chem. Biol.* **2012**, *8*, 823.

(31) Szewczyk, E.; Nayak, T.; Oakley, C. E.; Edgerton, H.; Xiong, Y.; Taheri-Talesh, N.; Osmani, S. A.; Oakley, B. R. *Nat. Protoc.* **2006**, *1*, 3111.

(32) Nayak, T.; Szewczyk, E.; Oakley, C. E.; Osmani, A.; Ukil, L.; Murray, S. L.; Hynes, M. J.; Osmani, S. A.; Oakley, B. R. *Genetics* **2006**, *172*, 1557.

(33) Kraas, F. I.; Helmetag, V.; Wittmann, M.; Strieker, M.; Marahiel, M. A. *Chem. Biol.* **2010**, *17*, 872.

(34) Chooi, Y. H.; Tang, Y. *Chem. Biol.* **2010**, *17*, 791.

(35) Hansen, D. B.; Bumpus, S. B.; Aron, Z. D.; Kelleher, N. L.; Walsh, C. T. *J. Am. Chem. Soc.* **2007**, *129*, 6366.

(36) Wittmann, M.; Linne, U.; Pohlmann, V.; Marahiel, M. A. *FEBS J.* **2008**, *275*, 5343.

(37) Baltz, R. H.; Miao, V.; Wrigley, S. K. *Nat. Prod. Rep.* **2005**, *22*, 717.

(38) Mootz, H. D.; Finking, R.; Marahiel, M. A. *J. Biol. Chem.* **2001**, *276*, 37289.

(39) Stachelhaus, T.; Huser, A.; Marahiel, M. A. *Chem. Biol.* **1996**, *3*, 913.

(40) Stachelhaus, T.; Mootz, H. D.; Bergendahl, V.; Marahiel, M. A. *J. Biol. Chem.* **1998**, *273*, 22773.

(41) Pfeifer, B. A.; Admiraal, S. J.; Gramajo, H.; Cane, D. E.; Khosla, C. *Science* **2001**, *291*, 1790.

(42) Gadow, A.; Vater, J.; Schlumbohm, W.; Palacz, Z.; Salnikow, J.; Kleinkauf, H. *Eur. J. Biochem.* **1983**, *132*, 229.

(43) Adefarati, A. A.; Giacobbe, R. A.; Hensens, O. D.; Tkacz, J. S. *J. Am. Chem. Soc.* **1991**, *113*, 3542.

(44) Underhill, E. W. *Can. J. Biochem.* **1968**, *46*, 401.

(45) Harimoto, Y.; Hatta, R.; Kodama, M.; Yamamoto, M.; Otani, H.; Tsuge, T. *Mol. Plant−Microbe Interact.* **2007**, *20*, 1463.

(46) Zhang, S.; Schwelm, A.; Jin, H.; Collins, L. J.; Bradshaw, R. E. *Fungal Genet. Biol.* **2007**, *44*, 1342.

(47) Scherlach, K.; Sarkar, A.; Schroeckh, V.; Dahse, H.-M.; Roth, M.; Brakhage, A. A.; Horn, U.; Hertweck, C. *ChemBioChem* **2011**, *12*, 1836.

(48) Lo, H.-C.; Entwistle, R.; Guo, C.-J.; Ahuja, M.; Szewczyk, E.; Hung, J.-H.; Chiang, Y.-M.; Oakley, B. R.; Wang, C. C. C. *J. Am. Chem. Soc.* **2012**, *134*, 4709.
(49) Xie, X.; Meehan, M. J.; Xu, W.; Dorrestein, P. C.; Tang, Y. *J. Am. Chem. Soc.* **2009**, *131*, 8388.
(50) Ames, B. D.; Nguyen, C.; Bruegger, J.; Smith, P.; Xu, W.; Ma, S.; Wong, E.; Wong, S.; Xie, X.; Li, J. W.-H.; Vederas, J. C.; Tang, Y.; Tsai, S.-C. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 11144.
(51) Hutton, J. J., Jr.; Kaplan, A.; Udenfriend, S. *Arch. Biochem. Biophys.* **1967**, *121*, 384.
(52) Petersen, L.; Olewinski, R.; Salmon, P.; Connors, N. *Appl. Microbiol. Biotechnol.* **2003**, *62*, 263.
(53) Lawrence, C. C.; Sobey, W. J.; Field, R. A.; Baldwin, J. E.; Schofield, C. J. *Biochem. J.* **1996**, *313* (Pt 1), 185.
(54) Hausinger, R. P. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 21.
(55) Luesch, H.; Hoffmann, D.; Hevel, J. M.; Becker, J. E.; Golakoti, T.; Moore, R. E. *J. Org. Chem.* **2003**, *68*, 83.
(56) Becker, J. E.; Moore, R. E.; Moore, B. S. *Gene* **2004**, *325*, 35.
(57) Hibi, M.; Kawashima, T.; Sokolov, P.; Smirnov, S.; Kodera, T.; Sugiyama, M.; Shimizu, S.; Yokozeki, K.; Ogawa, J. *Appl. Microbiol. Biotechnol.* **2012**, DOI: 10.1007/s00253-012-4136-7.
(58) Carlson, J. E.; Tulsieram, L. K.; Glaubitz, J. C.; Luk, V. W. K.; Kauffeldt, C.; Rutledge, R. *Theor. Appl. Genet.* **1991**, *83*, 194.
(59) Li, R.; Zhu, H.; Ruan, J.; Qian, W.; Fang, X.; Shi, Z.; Li, Y.; Li, S.; Shan, G.; Kristiansen, K.; Yang, H.; Wang, J. *Genome Res.* **2010**, *20*, 265.
(60) Hays, S.; Selker, E. *Fungal Genet. Newsl.* **2000**, *47*, 107.