Complete biosynthetic pathways of ascofuranone and ascochlorin in Acremonium egyptiacum

Araki, Yasuko; Awakawa, Takayoshi; Matsuzaki, Motomichi; Cho, Rihe; Matsuda, Yudai; Hoshino, Shotaro; Shinohara, Yasutomo; Yamamoto, Masaichi; Kido, Yasutoshi; Inaoka, Daniel Ken; Nagamune, Kisaburo; Ito, Kotaro; Abe, Ikuro; Kita, Kiyoshi

Published in:
National Academy of Sciences. Proceedings

Published: 23/04/2019

Document Version:
Final Published version, also known as Publisher's PDF, Publisher's Final version or Version of Record

License:
CC BY-NC-ND

Publication record in CityU Scholars:
Go to record

Published version (DOI):
10.1073/pnas.1819254116

Publication details:

Citing this paper
Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights
Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission
Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy
Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.
Complete biosynthetic pathways of ascofuranone and ascochlorin in *Acremonium egyptiacum*

Yasutomo Shinohara, Kotaro Ito, Daniel Ken Inaoka, Shotaro Hoshino, Yasutomo Shinhora, Masaichi Yamamoto, Yasutoshi Kido, Kenichi Kinoshita, Kisaburo Nagamune, Ikuo Abe, and Kiyoshi Kita

*Published online April 5, 2019.*

**Significance**

**Ascofuranone (AF) and ascochlorin (AC) are fungal natural products with similar chemical structures, originally isolated from *Acremonium egyptiacum*. Both have many useful biological properties; in particular, AF is a promising drug candidate against the tropical disease, African trypanosomiasis. However, the difficulty of the synthetic method and the inaccessibility of bioengineering methods have inhibited industrial production. This study identified all of the genes required for the branched biosynthetic pathways of AF/AC, which are clustered at two separate loci in the genome. In addition, we established the *A. egyptiacum* strain selectively producing AF, by genetically blocking the AC biosynthetic pathway. This study benefits the field of combinatorial biosynthesis through presenting biocatalysts and paves the way to cost-effective AF production with bioengineering.**

Ascofuranone (AF) and ascochlorin (AC) are fungal natural products with similar chemical structures, originally isolated from *Acremonium egyptiacum*. Both have many useful biological properties; in particular, AF is a promising drug candidate against the tropical disease, African trypanosomiasis. However, the difficulty of the synthetic method and the inaccessibility of bioengineering methods have inhibited industrial production. This study identified all of the genes required for the branched biosynthetic pathways of AF/AC, which are clustered at two separate loci in the genome. In addition, we established the *A. egyptiacum* strain selectively producing AF, by genetically blocking the AC biosynthetic pathway. This study benefits the field of combinatorial biosynthesis through presenting biocatalysts and paves the way to cost-effective AF production with bioengineering.


**Data deposition:** The data reported in this paper have been deposited in the DNA Data Bank of Japan (accession nos. DRA008136, E-GEAD-282, LC406756, and LC406757).

**Y.A., T.A., and M.M. contributed equally to this work.**

**To whom correspondence may be addressed. Email: matsuzaki.05@alumni.u-tokyo.ac.jp or abei@mol.f.u-tokyo.ac.jp**

This article contains supporting information online at www.pnas.org/cgi/doi/10.1073/pnas.1819254116/DCSupplemental.

Published online April 5, 2019.
Interestingly, these compounds have been isolated from various fungi; for example, 2 from Cylindrocladium ilicicola (15, 16), Cylindrocarpon sp. (17), Fusarium sp. (18), Microcera sp. (19), and Nectria coccinea (20), and 1 from Paecilomyces variotii (21) and Verticillium hemipterigenum (22), indicating the broad distribution of the biosynthetic pathways. Compounds related to 1 and 2 are considered to be synthesized through the prenylation of orsellinic acid (3) and terminal cyclization via epoxidation (17, 18, 23) (Fig. 2A), but the details of their biosyntheses have not been established. Recently, we reported that Stachybotrys bisbyi PYH05-7 encodes the 3-producing polyketide synthase (PKS) StbA, the UbiA-family prenyltransferase StbC, which produces ilicicolinic acid B (4) (Fig. 2A), and the nonribosomal peptide synthetase (NRPS)-like reductase StbB for the synthesis of LL-Z1272 (Ilicicolin B) (5), a putative precursor of 1 and 2 (24). By analogy to the biosynthetic pathways of other fungal meroterpenoids, such as paxilline, and aflatrem (25), the terminal olefin of the prenyl chain of 5 is thought to be epoxidized by a flavin monooxygenase (FMO), and then cyclized by a membrane-bound terpene cyclase (TPC). Considering the different cyclization patterns of 1 and 2, two distinct TPCs were assumed to exist in A. egyptiacum.

Developments in genomic resources and bioinformatics tools have accelerated the elucidation of the biosynthetic genes in various organisms (26). In particular, the biosynthetic genes for fungal secondary metabolites often form functional gene clusters (27), which can be predicted by searching the genome for collocated genes with functional motifs for putative biosynthetic enzymes (26). However, it is difficult to predict the relevance of genes with novel functions and to find biosynthetic clusters separated at multiple loci. For plant metabolites, in contrast, coexpression analyses have been effective in identifying nonclustered biosynthetic genes (28, 29); naturally, such a transcriptome-based approach is also promising in the case of fungal multilocus clusters (30). Since previous studies showed that 1 and 2, as well as other related compounds, suppress respiratory chain activities at multiple targets (31, 32), the fungal production of 1 and 2 is likely to be tightly regulated to avoid inhibiting fungal growth. Therefore, we can reasonably anticipate that determining the gene clusters that are differentially expressed in association with 1 production will lead to the identification of the responsible genes.

In this study, we aimed to find the biosynthetic genes for 1 by using a comparative transcription analysis of A. egyptiacum cultures exhibiting all-or-none production. Taking advantage of heterologous expression systems in Aspergillus spp., in vitro reconstruction assays, and a newly established gene disruption method for A. egyptiacum, we identified the complete biosynthetic enzymes of 2 encoded by the most prominently induced gene cluster. We then discovered another gene cluster at a discrete locus involved in the remaining steps of AF production, by virtue of a combination of motif-based and differential expression approaches. In addition, we analyzed the transcriptional factor regulating the biosynthetic genes of both compounds located in separate chromosomal regions.

**Results and Discussion**

**Discovery of the asc-1 Cluster.** First, we performed a genome-wide transcription analysis using the A. egyptiacum strain F-1392, to identify the differentially expressed gene clusters possibly responsible for the biosynthesis of 1 or 2. The production of 1 and 2 in A. egyptiacum varied dependent on the culture medium, and two media, designated as F1 and AF here, showed the virtual all-or-none difference. We thus prepared poly(A) selected RNAs from mycelia grown in F1 and AF media, which yielded 0.96 and 399 mg of 1 per liter culture, and obtained 611,740 and 846,811 RNA-seq reads, respectively. Although the experiment was preliminary, in terms of read numbers and lacking replicates, mapping the differential expression profiles to the 31-Mb draft genome (SI Appendix, Table S1) shed light on a 32-kb region on
Figure 3. Functional characterizations of AscA-G. (A) Schematic representation of the asc-1 cluster, found by the differential expression analysis. The expression change was indicated with log_{10} value for each gene. (B) HPLC profiles of the in vitro reaction products of ilicicin A (6) as a substrate when incubated with the buffer (i), the homogenate of the A. sojae wild-type strain (ii), the homogenates containing either AscE (iii) or AscG (iv), and the mixed homogenates containing AscE+AscF (v), AscE+AscG (vi), or AscE+AscF+AscG (vii), and authentic ascochlorin (viii). (C) HPLC profiles of authentic ascC (i), authentic ascochlorin (ii), mycelium extracts of A. egyptiacum F-1392 (iii), ΔascE strain (iv), authentic ilicicin A (v), and mycelium extract of ΔascE strain (vi). The yields of the compounds are summarized in SI Appendix, Table S8.
(SI Appendix, Figs. S6 B and C). These results indicated that the AscF-catalyzed terpene cyclization product 8 was oxidized into 2 by the P450 monoxygenase AscG. The absolute configuration of 8 is thought to be identical to that of (14S,15R,19R)-2, which was established by X-ray structure analysis (42).

To experimentally establish the involvement of 7 in the biosynthesis of 1 and/or 2, we constructed A. egyptiacum gene disruptants, using the kat70-deleted strain that we constructed in this study to enhance homologous recombination (43) (SI Appendix, Fig. S9). The ascF-deleted strain (ΔascF) no longer produced 2 (Fig. 3C), but instead accumulated 7 (1.22 g/L; SI Appendix, Table S8), which was isolated and structurally determined by NMR and HR-MS analyses (SI Appendix, Figs. S8 and S10 and Tables S4 and S7). In contrast, the ascE-deleted strain (ΔascE) only accumulated 6 (2.32 g/L; SI Appendix, Table S8), but neither 1 nor 2 (Fig. 3C). These results supported experimental proof that AscE epoxidized the terminal olefin of 6 to produce 7, and AscF cyclized 7 into 8 in AC biosynthesis. For AF biosynthesis, another TPC other than AscF should be involved, since ΔascF still produced 1 (0.50 g/L; SI Appendix, Table S8) (Fig. 3C). Given the fact that ΔascE no longer produced 1, the epoxide 7 was indicated as the last common precursor for the biosyntheses of 1 and 2.

Considering these results, we concluded that the seven genes, ascA-G, in the asc-1 cluster encode the biosynthetic enzymes for the production of 2. AscE is a P450 monoxygenase that catalyzes stereoselective epoxidation of the terminal double bond of the prenyl group (44). In contrast, in most of the cases, fungal meroterpenoid biosynthesis employs FAD-dependent monoxygenases for the formation of epoxide (25). Notably, AscF is the rare meroterpenoid TPC that produces a monocyclic terpene, and the cyclization reaction is proposed to be initiated by protonation of the terminal epoxide of 7 to generate a monocyclic tertiary cation, which is followed by a series of hydride and methyl shifts with abstraction of proton, leading to the formation of the (14S,15R,19R)-trimethylcyclohexanone ring structure of 8 (Fig. 2B). AscA-E are also involved in AF biosynthesis, but additional enzymes are required for the later steps, after the common precursor 7, which should be encoded outside of the asc-1 cluster.

Identification of the Ascofurone Biosynthetic Genes. Once again, the differential expression analysis was exploited to identify the genes responsible for the late stage of AF biosynthesis. Since one additional oxygen atom must be incorporated into the side chain of 7, we focused on P450s, the most abundant oxygenases in fungi. Among the gene clusters induced at least 10-fold in AF medium, only three harbored P450 monoxygenase genes other than the asc-1 cluster (SI Appendix, Table S5). One cluster on scaffold 6 encodes homologs of desmethylbassia biosynthesis enzymes, DmbD, DmbC, and DmbA (45), and thus was not likely to be involved in AF biosynthesis. Another cluster on scaffold 3 was excluded since it encodes an acetyltransferase and a peptidase, which both seemed irrelevant to AF biosynthesis. Combining the function prediction (see below), we targeted the other cluster on scaffold 3 (asc-2 cluster; DDBJ/ENA/GenBank accession no. LC406757) located on the different scaffold from asc-1 cluster (SI Appendix, Fig. S1).

The asc-2 cluster is composed of three genes, designated as ascHIJ (Fig. 4A). All three genes exhibited higher (log_{10} values ≥ 1.5) expression in AF medium (SI Appendix, Table S2). AscH is a P450 sharing 38% amino acid identity with the hexadecane hydroxylase P450-ALK2-A from Candida tropicalis (46). AscJ is classified into the NAD(P)-dependent, short-chain alcohol dehydrogenase family, sharing 33% identity with the aflatoxin biosynthesis protein AhH (47). Finally, AscI is an eight-transmembrane protein deduced by SOSUI program (48) without any Pfam motif; however, it is similar to the citreoviridin biosynthetic protein CtvD (49) and AurD in aurovertin biosynthesis (50), although it shares only 27% and 29% identities, respectively. Both enzymes are regioslective hydroxylases for cyclizing a bisepoxide to give a 3,4-dihydroxy-tetrahydrofururan structure, suggesting that AscI is the TPC for AF biosynthesis.
C_{22}H_{33}ClO_{3} from the HR-MS data (m/z 421.1799, calc. 421.1787; SI Appendix, Fig. S8 and Table S7), indicating that it contains an additional oxygen compared with 7. The NMR data revealed that 7 and 9 are very similar to each other, but one methylene signal (δ H: 2.01/2.05 ppm; δ C: 36.4 ppm) disappeared and one oxymethine signal (δ H: 4.20 ppm; δ C: 75.5 ppm) appeared in 9 (SI Appendix, Fig. S12 and Table S4). Based on the association with the spin system of H-17/H-18, we assigned this signal as H-16. Notably, under acidic condition, 9 was non-enzymatically converted into ascofuranol (10), whose planar structure was identified by HR-MS and NMR analyses, and its absolute configuration was determined as (16S,18S)-10 by the Mosher and NOESY analyses (SI Appendix, Figs. S13 and S14 and Table S6). The absolute configurations of 7 and 9 were thus assigned to be (16S,18S)- and (16R,18S)-9, respectively.

We further conducted in vitro assays to attribute the succeeding reactions to the candidate genes. The microsomal fraction of the yeast expressing asc1 efficiently converted 9 into 10, while that from the yeast harboring a blank vector did not afford any product (Fig. 4C). The $K_m$ and $V_{max}$ values of Asc1 were 50.4 ± 11.1 μM and 129 ± 13 nmol/min, respectively. In addition, we found that the cell-free homogenate of A. sojae expressing asc1 dehydrogenates 10 into 1 (Fig. 4D). Thus, we demonstrated that 7 is hydrolyzed by the P450 monoxygenase AscH, and the resultant 9 is cyclized by Asc1 to 10, which is oxidized into 1 by AscR (Fig. 2A).

Next, we investigated the reaction mechanism of the meroterpenoid TPC Asc1, which lacks significant sequence similarity to any known TPCs. Sequence alignment revealed that Asc1 also harbors several conserved acidic residues that are thought to be essential for the catalysis (36) (SI Appendix, Fig. S15). Indeed, the alanine substitutions of Asc1 D61, E103, D296, E353, and D355 resulted in significant loss of activities (SI Appendix, Fig. S16). Notably, D61, D296, and E353 are also conserved in AurD and CbD (SI Appendix, Fig. S15), suggesting that these residues are thought to be (16S)- and (16R,18S)-9, respectively.

Collectively, the three genes, ascH-J, in the asc-2 cluster encode the late-step biosynthetic enzymes, indicating that AF biosynthesis represents another rare example of a fungal multilocus biosynthetic cluster. Although there have been a few reports (51–54), there is still no established strategy for exploring such clusters. In the cases of prenyl xanthone and austinol, seeking the gene candidates based on the homology of the characteristic fragments, a prenylating PKS and a prenyltransferase, successfully led to the identification of the separate clusters, because there were only a few paralogs in the genome (52, 54). This was not the case for 1, since more than 100 genes encode P450 monoxygenases in the genome of A. egyptiacum, and the TPC, which is usually considered as a characteristic core enzyme, has not been known. The expression analysis was shown to be useful for the motif-independent identification of the fungal biosynthetic gene clusters, although previous studies exploited many sample conditions for the expression analysis and dedicated algorithms (55, 56). This study indicated that, even with a simple comparison of producing and nonproducing conditions, the expression analysis is a powerful method to identify the split gene clusters when combined with a motif-based approach.

**Coregulation of the Multilocus Gene Clusters by AscR.** Last, we analyzed the coregulation of the multilocus AF biosynthetic genes. Since the putative DNA-binding protein AscR was highly expressed in AF medium (SI Appendix, Table S2), it is predicted to be a transcription factor positively regulating the expression of the asc-1 cluster genes. An A. egyptiacum strain constitutively expressing AscR was constructed to verify this, and it produced both 1 and 2 in F1 medium (SI Appendix, Fig. S17A), suggesting the coregulation of the asc-1 and asc-2 clusters by AscR. This was further corroborated by the fact that all of the promoters of ascA-J possessed the shared sequence motif of TCGGCGGNTTW detected by MEME (SI Appendix, Fig. S17B), containing the CGG nucleotide triplet essential for the DNA binding of transcription factors with the same Zn$^{2+}$Cys$_6$ binuclear cluster (57). In fact, this motif was not present in the promoters of the three coregulating genes in both asc-1 cluster, which showed little or no expression in AF medium. We thus consider that AscR positively regulates the expression of both the asc-1 and -2 clusters by binding to this motif.

The evolutionary origin of fungal multifocus biosynthetic clusters has also been a matter of debate. In the cases of the biosyntheses of aflatoxin and austinol, ancestrally single-locus clusters were presumably divided into two separated loci by chromosomal rearrangement, supported by the presence of the relict copy in one cluster and the functional gene in the other cluster (51, 52). However, phylogeny and synteny analyses of trichothecene biosynthesis genes revealed that ancestrally separated genes could work with the major biosynthetic cluster as a coregulated pathway and, depending on the species, were later merged into the major cluster (53). For AF biosynthesis, there was no trace of past duplication suggesting a split into the asc-1 and -2 clusters. BLAST searches revealed partially syntenic clusters of the asc-1 cluster in several related fungi of the class Sordariomycetes, but there is no asc-2 syntenic cluster. Interestingly, the asc-1 syntenic cluster in Coniella lustricola also contains the ascH ortholog (PSR83571; 50% amino acid identity), whereas the fungus does not encode a plausible asc1 ortholog. Since the orthologs of the ascH and asc-1 cluster genes are widely distributed among Sordariomycetes, asc1 is likely to play another role in the biosynthesis of AC-related compounds. Given that the report of AF-producing fungi (13) and the distribution of asc1 orthologs were biased to the class Eurotiomycetes, we assume that asc1 was horizontally transferred and grafted upon the preexisting ascH to make a coregulated asc-2 cluster, although the elucidation will require further detailed investigation including synteny and phylogeny analyses using a wide range of fungal species.

**Conclusions.** We clarified the entire biosynthetic pathways of ascofuranone (1) and ascochlorin (2) in A. egyptiacum (Fig. 2). The biosyntheses of 1 and 2 share the common pathway up to the generation of ilicinol A epoxide (7). Notably, the biosynthetic genes of 1 and 2 are localized at distinct chromosomal regions, but all of their promoters share a common conserved motif, and they are probably regulated by the transcriptional factor, AscR. This study thus contributes to increasing the knowledge on meroterpenoid biosynthesis and demonstrates the power of a differential expression analysis for exploring multilocus biosynthetic clusters. From a clinical viewpoint, the elucidated genes, as well as the established method for the genetic manipulation of the strain F-1392, will be the keys for the drug development of A. egyptiacum. We have already established the ascF-deleted strain, producing exclusively 1, but not 2 (Fig. 3C), and further manipulation will lead to the cost-effective industrial production of ascofuranone.

**Materials and Methods.** A. egyptiacum (synonym: A. sclerotigenum) strain F-1392 (13, 14) is a descendent of the nitrosoguanidine-induced mutant no. 34, characterized in the original paper reporting 1 (2). The differential expression analysis was performed with an Ion PGM system, and the raw sequence reads and the expression profile per gene were deposited in the DDBJ under the accession nos. DRA000136 and DRA000302, respectively. A. egyptiacum gene disruptants were obtained from the newly established Δku70ΔaprG strain by homologous recombination, after transformation by the protoplast–polyethylene glycol method (38). Aspergillus oryzae N2AR1 (niaD $^{-}$, ΔscaB, ΔedaA $^{-}$) was used as the host for the heterologous coexpression of ascA-D, and Aspergillus sojae P6-1 (ΔapG) (39) was used as the host for the high-copy heterologous expression of ascJ. The products from mycelia of each transformant were extracted with acetone, and the in vitro reaction products were extracted with ethyl acetate. The extracts were analyzed by UV-HPLC and liquid chromatography–MS, using ODS columns with standard chromatographic methods. The purified products were monitored by NMR analyses, including $^1$HNMR, $^{13}$CNR, heteronuclear multiple bond coherence,
heteronuclear multiple quantum coherence, and COSY. Full experimental procedures are described in SI Appendix.

ACKNOWLEDGMENTS. We thank Prof. Katsuya Gomi (Tohoku University) and Prof. Katsuhiko Kitamoto (The University of Tokyo) for kindly providing the expression vectors and the fungal strain. We also thank Dr. Keiko Gomi and Dr. Ryoichi Sakaue (Kikkoman Corporation) for their helpful advice. This work is financially supported by Ministry of Education, Culture, Sports, Science, and Technology (METI)/Japan Society for the Promotion of Science (JSPS) KAKENHI Grants JP17H04763 (to T.A.), JP17K70099 (to T.A.), JP17K15767 (to Y.K.), 23117004 (to M.M.), JP16H06443 (to I.A.), JP18K11939 (to I.A.), and 26253025 (to K.K.). Japanese Initiative for Progress of Research on Infectious Disease for Global Epidemic (Grant JP18H0208800 to Y.K.K.; Japan Science and Technology Agency/National Natural Science Foundation of China Strategic International Collaborative Research Program (to I.A.); and Research Program on Emerging and Reemerging Infectious Diseases Grant 17k0108119001 (to K.K.). We also acknowledge support from the Program for Promotion of Basic and Applied Research Projects for Innovation (to N.Y.), and from the Science and Technology Research Promotion Program for Agriculture, Forestry, Fisheries, and Food Industry.