JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Combinatorial Generation of Complexity by Redox Enzymes in the Chaetoglobosin A Biosynthesis

Kan'ichiro Ishiuchi,[†] Takehito Nakazawa,[†] Fumitoshi Yagishita,[‡] Takashi Mino,[‡] Hiroshi Noguchi,[†] Kinya Hotta,[§] and Kenji Watanabe*[,†]

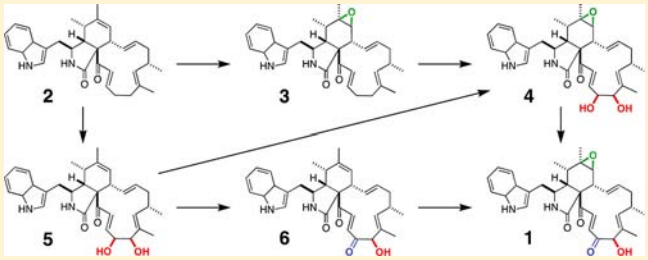
[†]Department of Pharmaceutical Sciences, University of Shizuoka, Shizuoka 422-8526, Japan

[‡]Department of Applied Chemistry and Biotechnology, Graduate School of Engineering, Chiba University, Chiba 263-8522, Japan

[§]School of Biosciences, The University of Nottingham Malaysia Campus, Semenyih, Selangor Darul Ehsan 43500, Malaysia

Supporting Information

ABSTRACT: Redox enzymes play a central role in generating structural complexity during natural product biosynthesis. In the postassembly tailoring steps, redox cascades can transform nascent chemical scaffolds into structurally complex final products. Chaetoglobosin A (**1**) is biosynthesized by a hybrid polyketide synthase−nonribosomal peptide synthetase. It belongs to the chaetoglobosin family of natural products, comprising many analogs having different degrees of oxidation introduced during their biosynthesis. We report here the determination of the complete biosynthetic steps leading to the formation of **1** from prochaetoglobosin I (**2**). Each oxidation step was elucidated using *Chaetomium globosum* strains carrying various combinations of deletion of the three redox enzymes, one FAD-dependent monooxygenase, and two cytochrome P450 oxygenases, and in vivo biotransformation of intermediates by heterologous expression of the three genes in *Saccharomyces cerevisiae*. Five analogs were identified in this study as intermediates formed during oxidization of **2** to **1** by those redox enzymes. Furthermore, a stereochemical course of each oxidation step was clearly revealed with the absolute configurations of five intermediates determined from X-ray crystal structure. This approach allowed us to quickly determine the biosynthetic intermediates and the enzymes responsible for their formation. Moreover, by addressing the redox enzymes, we were able to discover that promiscuity of the redox enzymes allowed the formation of a network of pathways that results in a combinatorial formation of multiple intermediate compounds during the formation of **1** from **2**. Our approach should expedite elucidation of pathways for other natural products biosynthesized by many uncharacterized enzymes of this fungus.



INTRODUCTION

Fungal natural products often exhibit biological activities of medicinal importance. Among them is chaetoglobosin A[1] **1**, which has a unique inhibitory activity against actin polymerization in mammalian cells.[2] Recent sequencing projects have revealed that many fungal genomes contain large numbers of clusters of genes that are predicted to encode for enzymes that can constitute natural product biosynthetic pathways.[3,4] The gene cluster responsible for biosynthesizing **1** was predicted and identified using an siRNA technology in *Penicillium expansum*.[5] The core structure of **1** is biosynthesized by a polyketide synthase−nonribosomal peptide synthetase (PKS−NRPS) hybrid megasynthetase and a stand-alone enoyl reductase found within the gene cluster. Studies have revealed that the tetramic acid moiety in equisetin[6] and cyclopiazonic acid[7] is formed via either a Dieckmann-type cyclization release or a reductive release catalyzed by the reductive (R) domain located at the C-terminal of PKS−NRPS hybrid synthetase. In the latter pathway, a Knoevenagel condensation and another reduction step is thought to follow for the formation of tetramic acid.[8] Most interestingly, the core biosynthesis of **1** is thought to proceed via a Diels−Alder reaction[9−12] to yield prochaetoglobosin I[13] **2**. The same mechanism is thought to apply to the formation of the same core scaffolding structure of **1** isolated from another fungus, *Chaetomium globosum* (Figure 1).

Many variants of chaetoglobosin-type of natural products have been isolated from *C. globosum*.[1b] On the basis of the chemical structures of isolated products, prochaetoglobosin IV[13] **3**, 20-dihydrochaetoglobosin A[14] **4**, cytoglobosin D[15] **5**, chaetoglobosin J[1b] **6**, these four compounds were thought to be intermediates formed by redox enzymes during the biosynthesis of **1** from **2**. However, it is frequently difficult to resolve the biosynthetic pathway that employs many enzymes to generate various intermediates for the formation of a complex final product. One viable approach takes advantage of the use of knockout strains lacking each of the genes in the gene cluster to identify missing intermediates for elucidation of the details of the pathway. To employ this approach, the ability to perform targeted homologous recombination for knocking out specific

Received: March 31, 2013
Published: April 23, 2013

   dx.doi.org/10.1021/ja402828w | *J. Am. Chem. Soc.* 2013, 135, 7371−7377

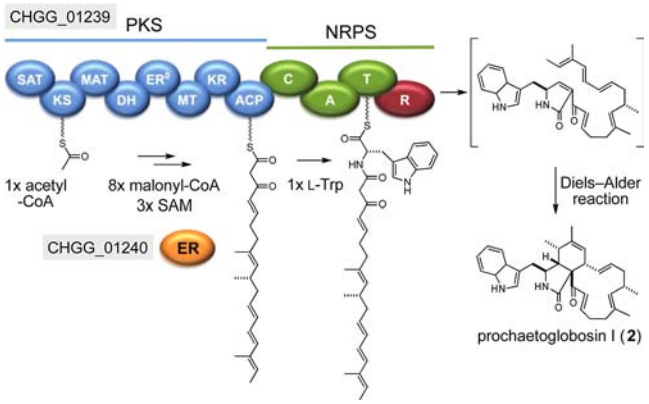


Figure 1. Modular organization of the uncharacterized iterative PKS–NRPS encoded by CHGG_01239 (14.1 kb) and a proposed biosynthetic pathway for the formation of **2** via a Diels–Alder reaction catalyzed by the hybrid synthetase. The iterative PKS–NRPS is shown with its predicted core domains. Abbreviations: SAT, starter unit:ACP transacylase; KS, ketosynthase; MAT, malonyl-CoA acyltransferase; DH, dehydratase; ER^0, inactive enoyl reductase; MT, methyltransferase; KR, ketoreductase; ACP, acyl carrier protein; C, condensation; A, adenylation; T, thiolation; R, reductase; ER, enoyl reductase; SAM, *S*-adenosyl-L-methionine.

genes is essential. However, the activity of *ligD* or its homologues that code for DNA ligase,[16] which is responsible for random nonhomologous recombination, hampers successful targeted recombination in many fungal strains. The deficiency of *ligD* has been established in Aspergilli to circumvent this obstacle.[17] To facilitate progress in understanding the detailed biosynthetic pathway of **1**, we have recently established a *ligD*-deficient strain of *C. globosum* capable of highly efficient targeted gene disruption.[18] Using this system, we chose to examine the details of the biosynthetic mechanism involved in the formation of chaetoglobosins in *C. globosum* to identify the precise role of various redox enzymes in the formation of **1** and its associated intermediates.

RESULTS

Identification of the PKS–NRPS Hybrid Synthetase and Enoyl Reductase Responsible for Biosynthesizing 1.

The core structure of **2** suggested that it was formed by a PKS–NRPS that could assemble a polyketide chain containing a tryptophan residue (Figure 1). There is a hybrid synthetase CheA from *Penicillium expansum* known to exhibit an identical activity.[5] BLASTP[19] search of the *C. globosum* genome sequence identified CHGG_01239, which exhibited an amino acid sequence identity and similarity of 34.3% and 52.2% to CheA, respectively. Thus, we speculated that CHGG_01239 might code for the PKS–NRPS responsible for the formation of **2**. Additionally, this class of hybrid synthetase often accompanies a stand-alone enoyl reductase for reduction of backbone double bond to install desired geometry into the chemical structure of the product. For example, an enoyl reductase CheB works with CheA in the chaetoglobosin biosynthesis, CcsC with CcsA in cytochalasin biosynthesis[20] and ApdC with ApdA in the aspyridone biosynthesis (Figure 2).[21] We found that CHGG_01240 showed high homology to *cheB*, *ccsC*, and *apdC*, hence proposed it to be the enoyl reductase involved in the chaetoglobosin biosynthesis. To verify the function of these two genes, CHGG_01239 and CHGG_01240, we performed targeted gene deletion by double

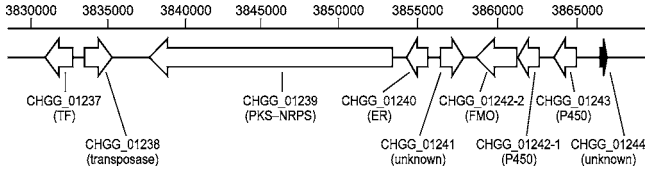


Figure 2. The organization of the chaetoglobosins biosynthetic gene cluster in *C. globosum*. Predicted function of the translation product of each of the gene found in the cluster is given in parentheses. Abbreviation: TF, transcription factor.

homologous gene replacement using an *hph* (hygromycin B phosphotransferase gene) cassette on the *C. globosum* Δ*CgpyrG*/Δ*CgligD* strain, CGKW14 (Figure S2 of the Supporting Information, SI). The CGKW14 strain was generated in a previous study by knocking out *CgligD* in order to suppress random integration of DNA fragments into the genome of *C. globosum*.[18] Thus, like *Aspergillus nidulans*,[22−25] CGKW14 can serve as a convenient strain for analysis of natural product biosynthesis in *C. globosum*. Transformants having *hph* integrated into target genes were successfully obtained after confirming the deletion of CHGG_01239 and CHGG_01240 by diagnostic PCR (Figures S2–S5 of the SI). Five-day mycelia of ΔCHGG_01239/CGKW14 grown on oatmeal agar were extracted with MeOH, and the dried extract was subjected to metabolite analysis by liquid chromatography mass spectrometry (LC–MS). As expected, deletion of CHGG_01239 abolished the production of **1** as well as **4** (Figure 3A i vs ii). Similarly, deletion of ΔCHGG_01240 led to the loss of production of **1** and **4** (Figure 3A i vs iii). The chemical structures of both of those compounds were characterized with high-resolution (HR) electrospray ionization (ESI) LC–MS, ^{1}H NMR and ^{13}C NMR to establish that **1** and **4** are chaetoglobosin A and 20-dihyrochaetoglobosin A, respectively. Moreover, we were able to determine for the first time the absolute configuration of **1** (SI).

Targeted Deletion of CHGG_01242-1, CHGG_01242-2, and CHGG_01243, and Identification of Intermediates 2, 3, 4, 5 and 6. The annotation given in the genome sequence database indicated that CHGG_01242 codes for a single polypeptide. However, closer analysis suggested that CHGG_01242 could in fact consist of two separate open reading frames (ORFs). Bioinformatics analysis indicated that the first ORF likely codes for a cytochrome P450 oxygenase (P450) which we named CHGG_01242-1, while the second ORF codes for a FAD-dependent monooxygenase (FMO), CHGG_01242-2 (Figure S6 of the SI). To verify this revision to gene annotation and the function of three oxygenase genes present in this gene cluster, namely CHGG_01242-1, CHGG_01242-2, and CHGG_01243, we performed targeted deletion of these genes in CGKW14 (see the Materials and Methods section and the SI for the details of deletion mutant constructions.). We initiated the functional analysis of these genes by looking for the biosynthesis of the starting precursor **2** and the absence of the final product **1**. The triple mutant ΔCHGG_01242-1/ΔCHGG_01242-2/ΔCHGG_01243 was successfully obtained, and the deletions were confirmed by diagnostic PCR (Figures S2–S5 of the SI). Dried MeOH extract of five-day mycelia of ΔCHGG_01242-1/ΔCHGG_01242-2/ΔCHGG_01243/CGKW14 grown on oatmeal agar was subjected to metabolite analysis by LC–MS. As expected, deletion of three genes did not affect the biosynthesis

dx.doi.org/10.1021/ja402828w | J. Am. Chem. Soc. 2013, 135, 7371–7377

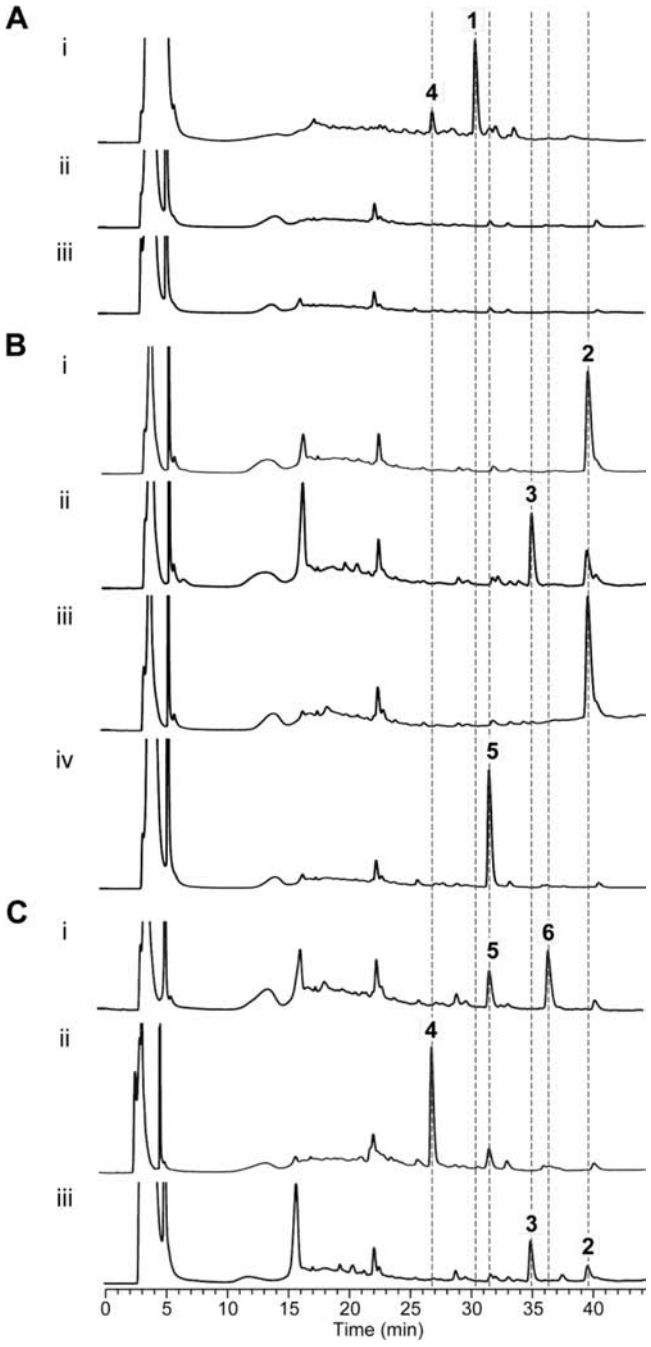
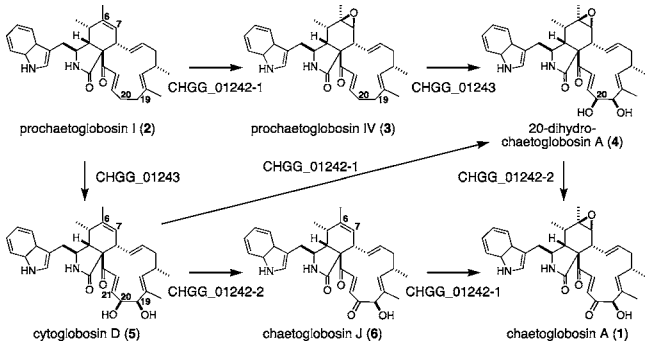


Figure 3. HPLC traces of metabolic extracts from the cultures of various *C. globosum* strains to identify the genes responsible for the oxidation steps involved in the formation of **1** during the chaetoglobosin biosynthesis. All deletion strains were prepared in CGKW14 by introducing *CgpyrG* into the target gene via homologous recombination. All cultures were grown in oatmeal agar medium supplemented with 20 mM uridine and 0.18 mM uracil. All HPLC traces were monitored at 280 nm. (A) Extract of the culture of (i) CGKW14 as a wild-type control, (ii) ΔCHGG_01239 strain, and (iii) ΔCHGG_01240 strain. (B) Extract of the culture of (i) ΔCHGG_01242-1/ΔCHGG_01242-2/ΔCHGG_01243 strain, (ii) ΔCHGG_01242-2/ΔCHGG_01243 strain, (iii) ΔCHGG_01242-1/ΔCHGG_01243 strain, and (iv) ΔCHGG_01242-1/ΔCHGG_01242-2 strain. (C) Extract of the culture of (i) ΔCHGG_01242-1 strain, (ii) ΔCHGG_01242-2 strain, and (iii) ΔCHGG_01243 strain.

of **2** but abolished the formation of **1** (Figure 3B i vs A i). The chemical structure of **2** was characterized by HRESIMS, ^{1}H NMR, and ^{13}C NMR (SI) to confirm it to be prochaetoglobosin I, and the absolute configuration of **2** was determined from X-ray crystal structure (Figure S7 of the SI). This result allowed us to revise the previously proposed biosynthetic steps[5] leading to the formation of **1**, where the oxygen atom at C19 was proposed to be of polyketide origin (Figure 1).

To understand the function of CHGG_01242-1, we analyzed a double mutant strain, ΔCHGG_01242-2/ΔCHGG_01243/CGKW14. The UV traces from HPLC analysis of the extract showed accumulation of **3** besides a decrease in **2** (Figure 3B ii), indicating the involvement of CHGG_01242-1 in the conversion of **2** to **3**. The chemical structure of **3** elucidated by HRESIMS, ^{1}H NMR, and ^{13}C NMR (SI) along with the absolute configuration of **3** determined from X-ray crystal structure (Figure S8 of the SI) revealed that CHGG_01242-1 converted **2** directly into **3** by the stereoselective epoxidation of the olefin at C6−C7. This epoxide-forming activity agrees with the prediction that CHGG_01242-1 codes for a P450 (Scheme 1).

Scheme 1. Biosynthetic Pathway for the Transformation of 2 to 1 Involving Multiple Oxidation Steps



Next, to understand the function of CHGG_01242-2, we analyzed another double mutant strain, ΔCHGG_01242-1/ΔCHGG_01243/CGKW14. On the basis of the UV trace from the HPLC analysis of the extracts, **2** was not converted into other metabolites (Figure 3B iii). This observation indicated that CHGG_01242-2 does not accept **2** as its substrate.

We also examined a double mutant strain ΔCHGG_01242-1/ΔCHGG_01242-2/CGKW14 to understand the function of CHGG_01243. The UV trace from the HPLC analysis of the extract showed accumulation of **5** and loss of **2** (Figure 3B iv), indicating that CHGG_01243 is responsible for the direct formation of **5** from **2**. The chemical structure of **5** elucidated by HRESIMS, ^{1}H NMR and ^{13}C NMR (SI) identified that **5** is a dihydroxyl derivative of **2** (Scheme 1). The absolute configuration of **5** was determined from X-ray crystal structure (Figure S10 of the SI). These results revealed that this P450 is responsible for the stereoselective dihydroxylation of C19 and C20 in **2**.

For a complete characterization of the activity of each of the three redox enzymes and elucidation of the whole picture of the oxidation pathway, we individually analyzed three single mutant strains, ΔCHGG_01242-1, ΔCHGG_01242-2, and ΔCHGG_01243. The analysis of ΔCHGG_01242-1 strain showed accumulation of **5** and **6** with loss of **2** (Figure 3C i vs B i). The isolated product **6** was characterized by HRESIMS,

dx.doi.org/10.1021/ja402828w | J. Am. Chem. Soc. 2013, 135, 7371−7377

^{1}H NMR, and ^{13}C NMR to be chaetoglobosin J (SI). The absolute configuration of **6** was determined from X-ray crystal structure (Figure S11 of the SI). These findings indicated that **5** and **6** are substrates of CHGG_01242-1. Comparison of the structures of the compounds identified thus far allowed us to determine that CHGG_01242-1 performed epoxidation of not only **2** into **3**, but also **5** and **6** into **4** and **1**, respectively.

Next, analysis of the ΔCHGG_01242-2 strain showed that it accumulated **4** in exchange for the loss of **2** (Figure 3C ii), indicating that **4** is a direct substrate of CHGG_01242-2. The isolated product **4** was characterized by HRESIMS, ^{1}H NMR, and ^{13}C NMR to be 20-dihydrochaetoglobosin A (SI). The absolute configuration of **4** was determined from X-ray crystal structure (Figure S9 of the SI). Comparison of the structure of **4** with other compounds allowed us to infer that CHGG_01242-2, a predicted FMO, catalyzes the formation of **1** from **4** through oxidation of the C20 hydroxyl group into a ketone. In addition, the accumulation of **5** and **6** in the strain lacking only CHGG_01242-1 and the structural difference between **5** and **6** suggest that this FMO also accepts **5** as its substrate to form **6**.

Lastly, the analysis of the ΔCHGG_01243 strain showed accumulation of **3** with significant reduction of **2** (Figure 3C iii). The earlier experiment with the ΔCHGG_01242-1/ΔCHGG_01242-2/CGKW14 strain identified that CHGG_01243 accepts **2** as its substrate. However, accumulation of **3** in this CHGG_01243 single mutant strain indicated that CHGG_01243 also takes **3** as its substrate to perform dihydroxylation at C19 and C20. Combining all of the results obtained from the mutant studies, we were able to establish the complete oxidation pathway (Scheme 1) for the transformation of **2** into **1** involving the formation of four intermediate compounds, all catalyzed by the three redox enzymes, CHGG_01242-1, CHGG_01242-2, and CHGG_01243.

Table 1. Deduced Functions of the ORFs in the Chaetoglobosin Biosynthetic Gene Cluster from *C. globosum*[a]

gene	amino acids (no.)	deduced function (homologue, NCBI accession number, species)	identity/similarity (%)
CHGG_01237	536	C_6 zinc finger protein	
CHGG_01238	409	transposase	
CHGG_01239	4699	PKS−NRPS hybrid	
CHGG_01240	377	enoyl reductase	
CHGG_01241	423	hypothetical protein	
CHGG_01242-1	512	P450	
CHGG_01242-2	616	FAD-dependent oxidoreductase	
CHGG_01243	500	P450 (GliF, AAW03300, *Aspergillus fumigatus*)[27]	27/60
CHGG_01244	187	hypothetical protein	

[a]Deduced function of the ORFs was determined based on the sequence similarity/identity to known proteins as determined by Protein BLAST (BLASTP) search[19] against the NCBI non-redundant database.

To address the biochemical function of the three redox enzymes, heterologous production of those enzymes and in vivo biotransformation were performed using *Saccharomyces cerevisiae* BY4741 as a host (see the Materials and Methods section and SI for details.). Each oxidation step was elucidated using a set of strains carrying each of the three genes encoding for one FAD-dependent monooxygenase (CHGG_01242-2) and two cytochrome P450 oxygenases (CHGG_01242-1 and CHGG_01243). The engineered *S. cerevisiae* strain with CHGG_01242-1 was able to convert **2**, **5**, and **6** into **3**, **4**, and **1**, respectively (Figure 4A−C and Figure S12 of the SI). Feeding **4** and **5** to the CHGG_01242-2 strain resulted in the formation of **1** and **6**, respectively (Figure 4D,E). Similarly, the CHGG_01243 strain produced **5** and **4** from fed **2** and **3**, respectively (Figure 4F,G). On the basis of the absolute configurations of five intermediates **2**−**6** and the final product **1** (see SI for details), and in vivo biotransformation of the intermediates, we were able to determine clearly the stereochemical course of the biosynthesis of **1** from **2** that is catalyzed by the three redox enzymes.

DISCUSSION

The homologous recombination-mediated targeted deletion approach that we have developed in *C. globosum* permitted the identification of the chaetoglobosin biosynthetic gene cluster. Deletion strains confirmed that the PKS−NRPS CHGG_01239 and the stand-alone enoyl reductase CHGG_01240 are responsible for the formation of **2**, the scaffolding compound of the chaetoglobosins. Furthermore, the study also allowed us to identify all of the intermediates formed and the enzymes responsible for the oxidation steps taken during the transformation of **2** to **1** in the chaetoglobosin biosynthesis. This represents a major step in elucidating the complex network of post-PKS and post-NRPS modifications, primarily oxidation reactions, that are regularly involved in the fungal natural product biosynthesis. Through the current study, the precise oxidation steps in the biosynthesis of **1** have been revealed, and some important conclusions can be drawn from our results. The functions of CHGG_01242-1, CHGG_01242-2, and CHGG_01243 in the biosynthesis of **1** were confirmed unequivocally by gene deletion and characterization of the blocked intermediates **2**, **3**, **4**, **5**, and **6**. CHGG_01242-1 was shown to catalyze a stereospecific epoxidation on **2**, **5**, and **6**. CHGG_01242-1 is a 512-amino acid enzyme exhibiting 43% sequence identity and 76% similarity with a P450 from *Colletotrichum higginsianum* and containing a conserved cytochrome P450 domain based on the NCBI CD-Search.[26] CHGG_01242-2 was shown to perform dehydrogenation of the C-20 hydroxyl group in **4** and **5**. CHGG_01242-2 coding for a 616-amino acid enzyme with 37% identity and 69% similarity to a FAD/FMN-containing dehydrogenase from *Aspergillus niger* ATCC1015. CHGG_01243 was shown to catalyze a stereospecific dihydroxylation of **2** and **3** at C-19 and C-20, respectively. This dioxygenase exhibits 27% sequence identity and 60% similarity with a P450 from *Aspergillus fumigatus*.[27]

Previous studies have reported that many chaetoglobosin derivatives can be isolated from *C. globosum*.[1] However, it has been very difficult to untangle the biosynthetic pathway from those products. Our current study has revealed that this difficulty partly stems from the fact that those postassembly redox enzymes are capable of accepting more than one compound as their substrate. Indeed, much more time and effort was expended in analyzing the products obtained from the mutant strains than creating them. Through the identification of a set of redox enzymes responsible for the synthesis of a number of key analogs of **1** in *C. globosum*, it became clear that relaxed substrate specificity of these redox enzymes is responsible for the formation of a complex network

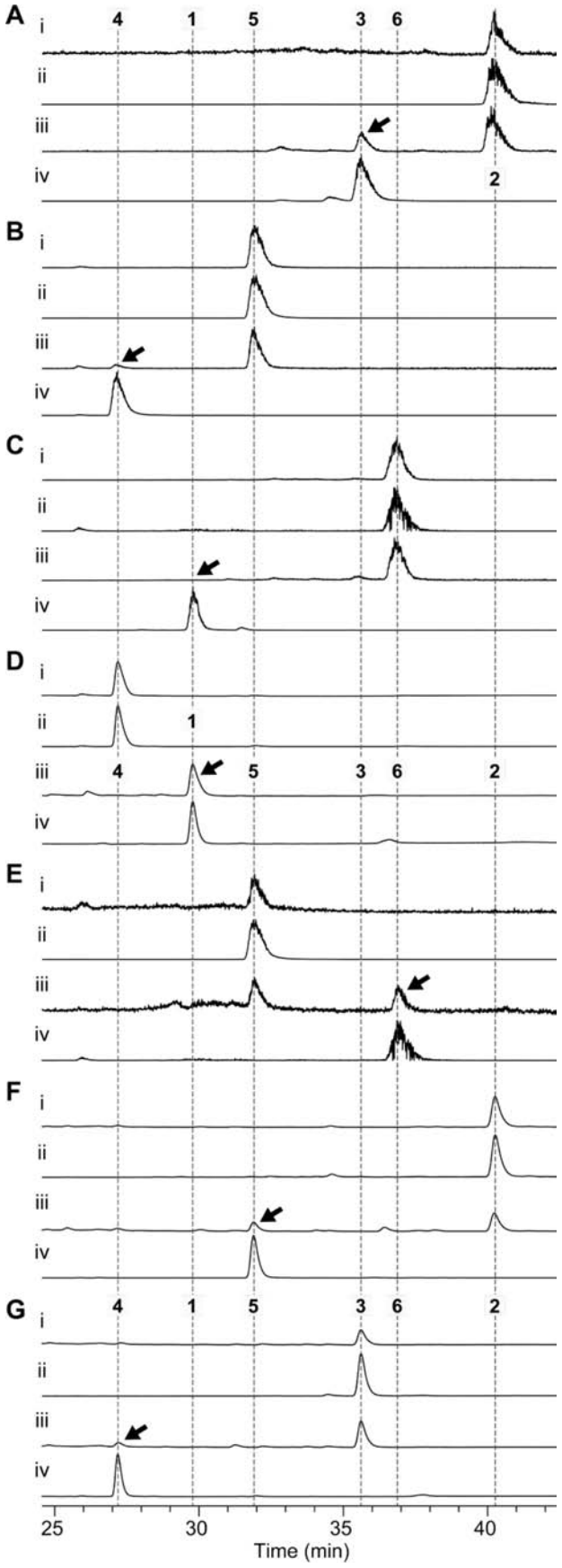


Figure 4. In vivo biotransformation of intermediates **2**–**6** by engineered *S. cerevisiae* BY4741 strains carrying each of the genes encoding for the redox enzymes CHGG_01242-1, CHGG_01242-2, and CHGG_01243 involved in the chaetoglobosin biosynthesis. All (i) plots are negative control obtained with an *S. cerevisiae* BY4741 strain carrying an empty expression vector. (A) (iii) Extract of **2**-supplemented culture of a strain with CHGG_01242-1. The peak corresponding to the expected product **3** is indicated by an arrow. Authentic reference of (ii) **2** and (iv) **3**. The MS data were analyzed for the range of *m*/*z* 475–505. (B) (iii) Extract of **5**-supplemented culture of a strains with CHGG_01242-1. The peak corresponding to the expected product **4** is indicated by an arrow. Refer to Figure S10A of the SI for mass spectra for confirmation of the presence of **4** in this sample. Authentic reference of (ii) **5** and (iv) **4**. The MS data were analyzed for the range of *m*/*z* 505–540. (C) (iii) Extract of **6**-supplemented culture of a strain with CHGG_01242-1. The peak corresponding to the expected product **1** is indicated by an arrow. Refer to Figure S10B of the SI for mass spectra for confirmation of the presence of **1** in this sample. Authentic reference of (ii) **6** and (iv) **1**. The MS data were analyzed for the range of *m*/*z* 475–505. (D) (iii) Extract of **4**-supplemented culture of a strain with CHGG_01242-2. Authentic reference of (ii) **4** and (iv) **1**. All HPLC traces were monitored at 280 nm. (E) (iii) Extract of **5**-supplemented culture of a strain with CHGG_01242-2. The peak corresponding to the expected product **6** is indicated by an arrow. Authentic reference of (ii) **5** and (iv) **6**. The MS data were analyzed for the range of *m*/*z* 500–520. (F) (iii) Extract of **2**-supplemented culture of a strains with CHGG_01243. The peak corresponding to the expected product **5** is indicated by an arrow. Authentic reference of (ii) **2** and (iv) **5**. All HPLC traces were monitored at 280 nm. (G) (iii) Extract of **3**-supplemented culture of a strain with CHGG_01243. The peak corresponding to the expected product **4** is indicated by an arrow. Authentic reference of (ii) **3** and (iv) **4**. All HPLC traces were monitored at 280 nm.

of biosynthetic pathways that results in a combinatorial production of a series of intermediate compounds.

CONCLUSIONS

In this study, we identified the chaetoglobosins biosynthetic pathway in *C. globosum* and the function of three genes coding for redox enzymes, an FMO and two P450s. The knowledge obtained here allowed us to gain deeper understanding of the source of structural variations present in this class of natural products, and reiterated the notion[28] that redox enzymes play a vital role in introducing complexity into the chemical structures of natural products. The current study clearly established that the three redox enzymes are sufficient to transform prochaetoglobosin I (**2**), the nascent product of PKS–NRPS, into chaetoglobosin A (**1**). Moreover, we were able to determine that those redox enzymes possess a considerable substrate tolerance that leads to the formation of various intermediates.

Detailed investigation of a complex network of metabolic pathways described here was made possible through facile preparation of targeted gene disruption mutants of *C. globosum*. Many fungal species are often not amenable to molecular genetic manipulations due to frequent occurrence of random nonhomologous recombination in those organisms. Moreover, deletion of each and every genes in large gene clusters would typically require the development of additional selectable markers or techniques to recycle markers. However, the use of Δ*CgpyrG*/Δ*CgligD* strain, which allowed for highly efficient targeted gene disruption and orotidine-5′-phosphate decarboxylase gene *CgpyrG*-mediated recyclable 5-fluoroorotic acid selection, made it straightforward to prepare multiple gene

dx.doi.org/10.1021/ja402828w | *J. Am. Chem. Soc.* 2013, 135, 7371–7377

deletions in a fungus that has been resistant to reliable molecular genetic manipulations to date. When applied to secondary metabolite biosynthetic gene clusters, this method promises to be valuable in translating the genomic sequence information into biochemical information and ultimately chemical structures. The approach established here should be generally applicable to investigating biosynthetic mechanisms of complex natural products, and useful in examining other fungal species for which established molecular genetic methods are not yet available.

MATERIALS AND METHODS

Strains and General Techniques for DNA Manipulation. For the construction of a disruption cassette and confirmation of the modified genotype, the genomic DNA isolated from CGKW14 or the transformants was analyzed by PCR. Genomic DNA from above strains was prepared using the CTAB isolation buffer as described elsewhere.[29] The gene-specific primers are listed in Table S1 of the SI. PCR was performed using KOD FX Neo (TOYOBO Co., Ltd.). Sequences of PCR products were confirmed through DNA sequencing (Macrogen Japan Corporation). *Escherichia coli* XL1-Blue (Agilent Technologies) was used for plasmid propagation. DNA restriction enzymes were used as recommended by the manufacturer (Fermentas).

In Silico Analysis of Genome Sequence. Protein sequences obtained by translating the genes CHGG_01242-1, CHGG_01242-2, and CHGG_01243 were used for BLASTP[19] query analyses performed against the NCBI GenBank database (http://www.ncbi.nlm.nih.gov) and the Broad Institute database (http://www.broadinstitute.org).

Spectroscopic Analyses. NMR spectra were obtained on a Bruker BioSpin AVANCE 400 MHz spectrometer (^{1}H 400 MHz, ^{13}C 100 MHz). ^{1}H NMR chemical shifts are reported in parts/million (ppm) using the proton resonance of residual solvents as references: $CDCD_3$ δ_H 7.26 and DMSO-d_6 δ_H 2.50. ^{13}C NMR chemical shifts are reported relative to $CDCl_3$ δ_C 77.0 and DMSO-d_6 δ_C 39.5. LC−MS was conducted with a Thermo SCIENTIFIC LCQ Fleet and Exactive liquid chromatography mass spectrometer by using positive electrospray ionization. Samples were separated for analysis on an Alltech 3 μm, 2.1 × 100 mm C18 reversed-phase column using a linear gradient of 5−95% (v/v) CH_3CN in H_2O supplemented with 0.05% (v/v) formic acid at a flow rate of 125 μL/min.

Preparation of the Deletion Strains. Details of the method used to prepare deletion strains are given in the SI. Briefly, deletion of target gene in *C. globosum* was carried out by homologous recombination using the *C. globosum* strain CGKW14 whose random nonhomologous recombination activity has been disabled by the disruption of *CgligD*.[18] A disruption cassette composed of a selection marker *hph* flanked on both sides by a 1500-base pair fragment that is homologous to the site of recombination in the *C. globosum* genome was introduced to CGKW14 to replace the target gene with the selection marker. Disruption of the target gene was confirmed by amplifying the disrupted segment from the genomic DNA by PCR.

Compounds Purification and Structural Characterization. To purify **1** (Scheme 1) for structural analysis, *C. globosum* (wild-type) was cultured in 20 × 20 mL oatmeal agar plates at 30 °C for 15 days. Grown mycelium and medium were extracted with acetone (2L × 2) and concentrated in vacuo to give an oily residue, which was then fractionated by silica gel flash column chromatography with $CHCl_3$/MeOH (1:0 → 0:1). The fraction eluted with $CHCl_3$/MeOH (50:1) was further purified by a reversed-phase HPLC (Nacalai Tesque Inc., COSMOSIL 5C18 MS-II, 20 × 250 mm) on an isocratic elution system of 45% CH_3CN (v/v) in H_2O at a flow rate of 8.0 mL/min to afford **1** (100 mg/L). To purify **2** and **3** (Scheme 1) for structural analysis, ΔCHGG_01243/CGKW-14 was cultured in 100 × 20 mL oatmeal agar plates at 30 °C for 10 days. Grown mycelium was extracted with MeOH (2L × 2) and concentrated in vacuo to give an oily residue, which was partitioned between EtOAc and H_2O. EtOAc-soluble materials were then fractionated by silica gel flash column chromatography with $CHCl_3$/MeOH (1:0 → 0:1) to afford **3** (25.2 mg/L) in the fraction eluted with $CHCl_3$/MeOH (100:1). The fraction eluted with $CHCl_3$/MeOH (1:0) was further purified by a silica gel flash column chromatography with *n*-hexane/EtOAc (1:0 → 0:1) to afford **2** (8.8 mg/L) in the fraction eluted with *n*-hexane/EtOAc (1:1). To purify **4** (Scheme 1) for structural analysis, ΔCHGG_01242-2/CGKW14 was cultured in 100 × 20 mL oatmeal agar plates at 30 °C for 10 days. Grown mycelium was extracted with MeOH (2L × 2) and concentrated in vacuo to give an oily residue, which was partitioned between EtOAc and H_2O. EtOAc-soluble materials were then fractionated by silica gel flash column chromatography with $CHCl_3$/MeOH (1:0 → 0:1) to afford **4** (15.8 mg/L) in the fraction eluted with $CHCl_3$/MeOH (30:1). To purify **5** and **6** (Scheme 1) for structural analysis, ΔCHGG_01242-1/CGKW14 was cultured in 100 × 20 mL oatmeal agar plates at 30 °C for 10 days. Grown mycelium was extracted with MeOH (2L × 2) and concentrated in vacuo to give an oily residue, which was partitioned between EtOAc and H_2O. EtOAc-soluble materials were then fractionated by silica gel flash column chromatography with $CHCl_3$/MeOH (1:0 → 0:1) to afford **5** (3.5 mg/L) and **6** (6.9 mg/L) in the fraction eluted with $CHCl_3$/MeOH (100:1) and (30:1), respectively. Chemical structures of the isolated compounds were identified from the spectroscopic data given in the Supporting Information.

In Vivo Biotransformation Studies with *S. cerevisiae*. Details of the procedures for preparing the expression plasmids used to produce the FMO and P450s, CHGG_01242-1, CHGG_01242-2, and CHGG_01243, in *S. cerevisiae* BY4741, and performing in vivo biotransformation of the pathway intermediates with CHGG_01242-1, CHGG_01242-2, and CHGG_01243 are given in the Supporting Information.

ASSOCIATED CONTENT

Supporting Information

Data from the NMR, MS, and X-ray crystal structure determination of the compounds, including cif files for the crystal structures, and additional experimental information. This material is available free of charge via Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

kenji55@u-shizuoka-ken.ac.jp

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We would like to thank Professor Hideaki Oikawa at the Hokkaido University for providing us with an ^{1}H NMR spectrum of **2**. We would like to express our appreciation for financial support from Japan Society for the Promotion of Science (JSPS) through the "Funding Program for Next Generation World-Leading Researchers," initiated by the Council for Science and Technology Policy (No. LS103) (K.W.) and by Industrial Technology Research Grant Program in 2009 (No. 09C46001a) from New Energy and Industrial Technology Development Organization (NEDO) of Japan (K.W.). These works were also supported in part by the Uehara Memorial Foundation (K.W.), by Mochida Memorial Foundation for Medical and Pharmaceutical Research (K.W.), by The Hokuto Foundation for Bioscience (K.W.), and by the Naito Foundation Japan (K.W.). A postdoctoral fellowship to T.N. from the Institution of Fermentation at Osaka is gratefully acknowledged.

REFERENCES

(1) (a) Sekita, S.; Yoshihira, K.; Natori, S.; Kuwano, H. *Tetrahedron Lett.* **1973**, 2109−2112. (b) Sekita, S.; Yoshihira, K.; Natori, S.; Kuwano, H. *Chem. Pharm. Bull.* **1982**, *30*, 1629−1638. (c) Silverton, J. V.; Akiyama, T.; Kabuto, C.; Sekita, S.; Yoshihira, K.; Natori, S. *Tetrahedron Lett.* **1976**, 1349−1350.

(2) (a) Scherlach, K.; Boettger, D.; Remme, N.; Hertweck, C. *Nat. Prod. Rep.* **2010**, *27*, 869−886. (b) Low, I.; Jahn, W.; Wieland, T.; Sekita, S.; Yoshihira, K.; Natori, S. *Anal. Biochem.* **1979**, *95*, 14−18.

(3) Galagan, J. E.; Calvo, S. E.; Cuomo, C.; Ma, L. J.; Wortman, J. R.; Batzoglou, S.; Lee, S. I.; Bastürkmen, 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*, 1105−1115.

(4) Keller, N. P.; Turner, G.; Bennett, J. W. *Nat. Rev. Microbiol.* **2005**, *3*, 937−947.

(5) Schümann, J.; Hertweck, C. *J. Am. Chem. Soc.* **2007**, *129*, 9564−9565.

(6) Sims, J. W.; Schmidt, E. W. *J. Am. Chem. Soc.* **2008**, *130*, 11149−11155.

(7) Liu, X.; Walsh, C. T. *Biochemistry* **2009**, *48*, 8746−8757.

(8) Boettger, D.; Hertweck, C. *ChemBioChem* **2013**, *14*, 28−42.

(9) Kennedy, J.; Auclair, K.; Kendrew, S. G.; Park, C.; Vederas, J. C.; Hutchinson, C. R. *Science* **1999**, *284*, 1368−1372.

(10) Watanabe, K.; Mie, T.; Ichihara, A.; Oikawa, H.; Honma, M. *J. Biol. Chem.* **2000**, *275*, 38393−38401.

(11) Ose, T.; Watanabe, K.; Mie, T.; Honma, M.; Watanabe, H.; Yao, M.; Oikawa, H.; Tanaka, I. *Nature* **2003**, *422*, 185−189.

(12) Ma, S. M.; Li, J. W. H.; Choi, J. W.; Zhou, H.; Lee, K. K. M.; Moorthie, V. A.; Xie, X.; Kealey, J. T.; Da Silva, N. A.; Vederas, J. C.; Tang, Y. *Science* **2009**, *326*, 589−592.

(13) (a) Oikawa, H.; Murakami, Y.; Ichihara, A. *Tetrahedron Lett.* **1991**, *32*, 4533−4536. (b) Oikawa, H.; Murakami, Y.; Ichihara, A. *J. Chem. Soc. Perkin Trans. 1* **1992**, 2949−2953.

(14) Oikawa, H.; Murakami, Y.; Ichihara, A. *Biosci. Biotech. Biochem.* **1993**, *57*, 628−631.

(15) Cui, C. M.; Li, X. M.; Li, C. S.; Proksch, P.; Wang, B. G. *J. Nat. Prod.* **2010**, *73*, 729−733.

(16) Wei, Y. F.; Robins, P.; Carter, K.; Caldecott, K.; Pappin, D. J.; Yu, G. L.; Wang, R. P.; Shell, B. K.; Nash, R. A.; Schär, P.; Barnes, D. E.; Haseltine, W. A.; Lindahl, T. *Mol. Cell. Biol.* **1995**, *15*, 3206−3216.

(17) Ishibashi, K.; Suzuki, K.; Takakura, C.; Inoue, H. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 14871−14876.

(18) Nakazawa, T.; Ishiuchi, K.; Sato, M.; Tsunematsu, T.; Sugimoto, S.; Gotanda, Y.; Noguchi, H.; Hotta, K.; Watanabe, K., manuscript submitted.

(19) Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis, S.; Madden, T. L. *Nucleic Acids Res.* **2008**, *36*, W5−9.

(20) Qiao, K.; Chooi, Y. H.; Tang, Y. *Metab. Eng.* **2011**, *13*, 723−732.

(21) Bergmann, S.; Schümann, J.; Scherlach, K.; Lange, C.; Brakhage, A. A.; Hertweck, C. *Nat. Chem. Biol.* **2007**, *3*, 213−217.

(22) Brown, D. W.; Adams, T. H.; Keller, N. P. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14873−14877.

(23) Bok, J. W.; Hoffmeister, D.; Maggio-Hall, L. A.; Murillo, R.; Glasner, J. D.; Keller, N. P. *Chem. Biol.* **2006**, *13*, 31−37.

(24) Bok, J. W.; Chiang, Y. M.; Szewczyk, E.; Reyes-Dominguez, Y.; Davidson, A. D.; Sanchez, J. F.; Lo, H. C.; Watanabe, K.; Strauss, J.; Oakley, B. R.; Wang, C. C. C.; Keller, N. P. *Nat. Chem. Biol.* **2009**, *5*, 462−464.

(25) Chiang, Y. M.; Szewczyk, E.; Davidson, A. D.; Keller, N.; Oakley, B. R.; Wang, C. C. C. *J. Am. Chem. Soc.* **2009**, *131*, 2965−2970.

(26) Marchler-Bauer, A.; Bryant, S. H. *Nucleic Acids Res.* **2004**, *32*, W327−331.

(27) Gardiner, D. M.; Howlett, B. J. *FEMS Microbiol. Lett.* **2005**, *248*, 241−248.

(28) Wang, P.; Gao, X.; Tang, Y. *Curr. Opin. Chem. Biol.* **2012**, *16*, 362−369.

(29) Zolan, M. E.; Pukkila, P. J. *Mol. Cell. Biol.* **1986**, *6*, 195−200.