

# Complexity Generation in Fungal Peptidyl Alkaloid Biosynthesis: Oxidation of Fumiquinazoline A to the Heptacyclic Hemiaminal Fumiquinazoline C by the Flavoenzyme Af12070 from *Aspergillus fumigatus*

Brian D. Ames,<sup>†,⊥</sup> Stuart W. Haynes,<sup>†,⊥</sup> Xue Gao,<sup>‡</sup> Bradley S. Evans,<sup>§</sup> Neil L. Kelleher,<sup>||</sup> Yi Tang,<sup>‡</sup> and Christopher T. Walsh<sup>\*,†</sup>

<sup>†</sup>Department of Biological Chemistry & Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, United States

<sup>‡</sup>Department of Chemical & Biomolecular Engineering, Department of Chemistry & Biochemistry, University of California, Los Angeles, Los Angeles, California 90095, United States

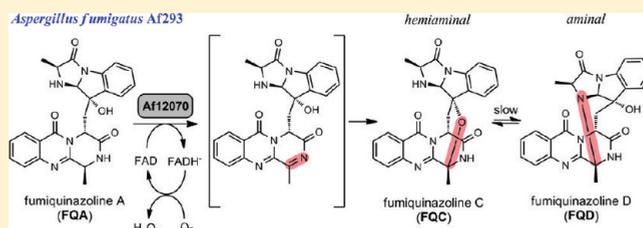
<sup>§</sup>Institute for Genomic Biology, University of Illinois, Urbana, Illinois 61801, United States

<sup>||</sup>Department of Chemistry, Northwestern University, Evanston, Illinois 60208, United States

## Supporting Information

**ABSTRACT:** The human pathogen *Aspergillus fumigatus* makes a series of fumiquinazoline (FQ) peptidyl alkaloids of increasing scaffold complexity using L-Trp, 2 equiv of L-Ala, and the non-proteinogenic amino acid anthranilate as building blocks. The FQ gene cluster encodes two non-ribosomal peptide synthetases (NRPS) and two flavoproteins. The trimodular NRPS Af12080 assembles FQF (the first level of complexity) while the next two enzymes, Af12060 and Af12050, act in tandem in an oxidative annulation sequence

to couple alanine to the indole side chain of FQF to yield the imidazolindolone-containing FQA. In this study we show that the fourth enzyme, the monovalent flavoprotein Af12070, introduces a third layer of scaffold complexity by converting FQA to the spirohemiaminal FQC, presumably by catalyzing the formation of a transient imine within the pyrazinone ring (and therefore acting in an unprecedented manner as an FAD-dependent amide oxidase). FQC subsequently converts nonenzymatically to the known cyclic aminal FQD. We also investigated the effect of substrate structure on Af12070 activity and subsequent cyclization with a variety of FQA analogues, including an FQA diastereomer (2'-*epi*-FQA), which is an intermediate in the fungal biosynthesis of the tremorgenic tryptoquialanine. 2'-*epi*-FQA is processed by Af12070 to *epi*-FQD, not *epi*-FQC, illustrating that the delicate balance in product cyclization regiochemistry can be perturbed by a remote stereochemical center.



The secondary metabolome of the human fungal pathogen *Aspergillus fumigatus* is extensive, including several polyketides, but notably a range of peptidyl alkaloids that include gliotoxin, fumitremorgins, fumigaclavines, fumagillins, and fumiquinazolines.<sup>1,2</sup> The fumiquinazolines (FQs) are signature peptidyl alkaloids produced by all 40 *A. fumigatus* strains that have been examined for natural products, although their physiologic function remains to be elucidated.<sup>1</sup> The FQs are a group of compounds that display varying degrees of structural complexity<sup>3,4</sup> and include the tricyclic FQF, which is modified to the more architecturally elaborate framework seen in FQA and then finally to a third level of complexity in FQC, which contains a heptacyclic scaffold formed via a bridging hemiaminal linkage (Figure 1).

The 29 Mb *A. fumigatus* genome encodes 14 predicted polyketide and 14 non-ribosomal peptide gene clusters.<sup>5</sup> We have recently assigned two NRPS genes (*af12080* and *af12050*) on chromosome 6 of *A. fumigatus* Af293 as being involved in

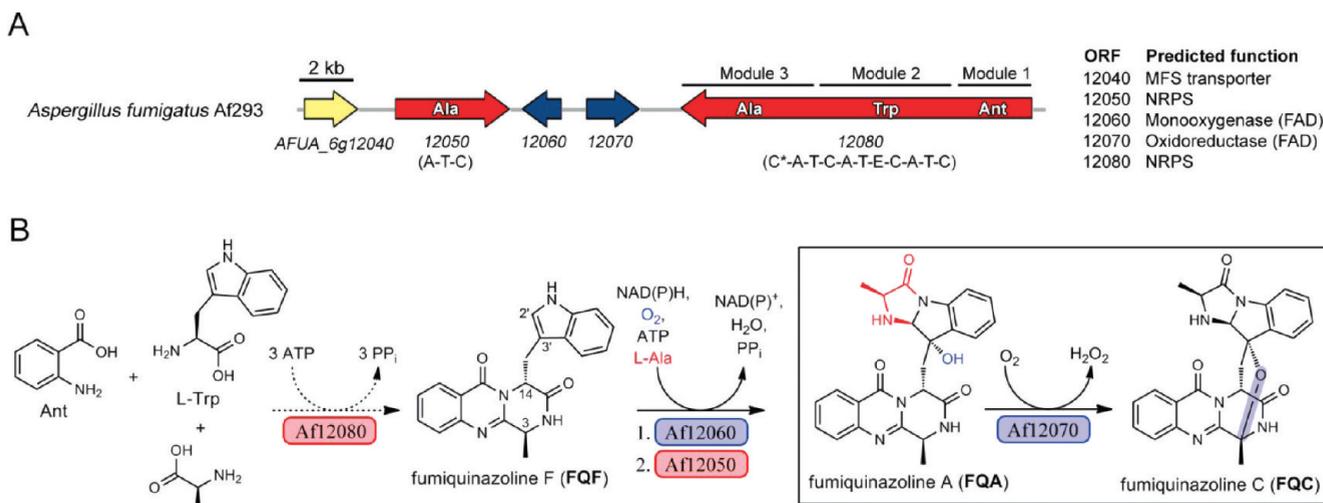
the assembly of the fumiquinazolines (Figure 1A).<sup>6,7</sup> Our initial interest in these NRPS systems came as a result of the identification of fungal adenylation domains that activate anthranilate (a core building block in the fumiquinazoline framework), and we have shown that this activity resides in module one of the trimodular NRPS Af12080.<sup>6</sup> Two adjacent genes in the cluster, *af12060* and *af12050*, have been expressed in *E. coli* and shown to (1) oxygenate the indole side chain of FQF (Af12060) and then (2) activate L-Ala for annulation to the oxidized indole scaffold (Af12050), thereby constructing two new C–N bonds.<sup>7</sup> This set of oxidative annulation operations stereospecifically transforms the bicyclic indole in FQF to the tricyclic imidazoindolone in the product FQA (Figure 1B).

**Received:** August 15, 2011

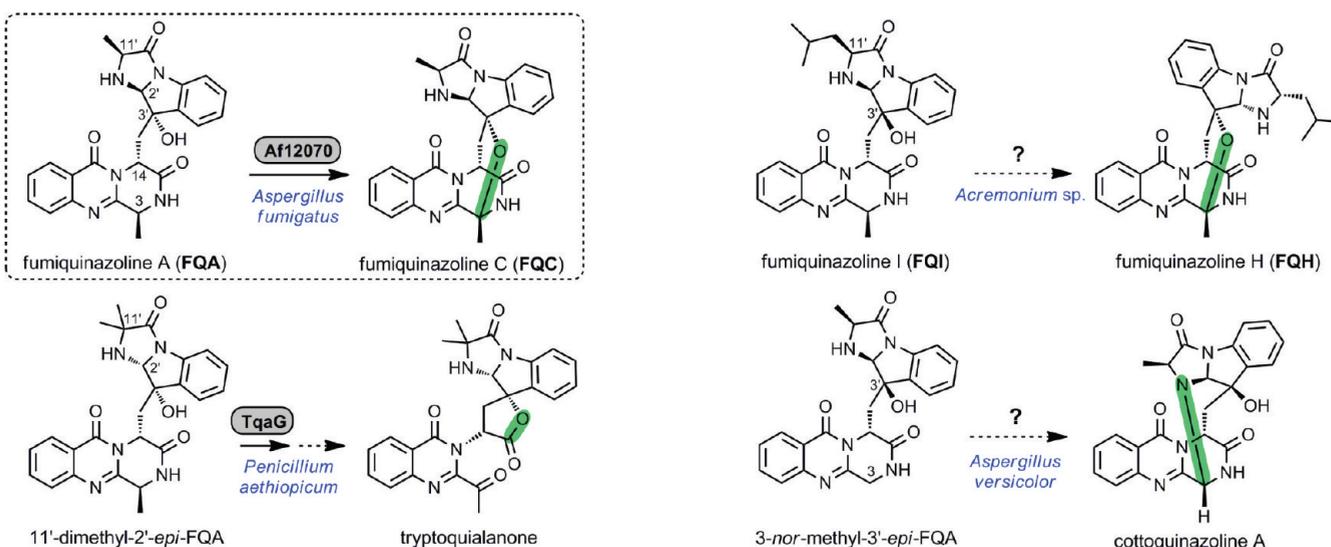
**Revised:** September 6, 2011

**Published:** September 7, 2011





**Figure 1.** Gene cluster and proposed biosynthetic route for the production of fumiquinazoline alkaloids from *Aspergillus fumigatus* Af293.



**Figure 2.** Examples of FQA-like scaffold modification by intramolecular nucleophilic reaction/rearrangement.

In this work we have characterized Af12070, the fourth enzyme encoded in the FQ gene cluster. Af12070 takes a molecule with a fair degree of rotational freedom, which consists of two distinct moieties (a pyrazinoquinazolinone and an imidazoindolone) linked by a single methylene “bridge” and transforms it into the fixed and highly constrained heptacyclic scaffold FQC (Figure 1B). This fascinating piece of biosynthetic origami, though unusual, is not unique to Af12070; instead, it appears to form part of a small group of fungal biosynthetic enzymes that build up complex multicyclic scaffolds. TqaG (which bears 72% amino acid sequence similarity to Af12070) from the recently characterized tryptoquialanine pathway in *Penicillium aethiopicum*,<sup>8</sup> and as yet unassigned enzymes in *Aspergillus versicolor*<sup>9,10</sup> and an *Acremonium* sp.,<sup>11</sup> appear to catalyze potentially analogous and equally intriguing transformations en route to tryptoquialanine, cottoquinazolines, and fumiquinazoline H, respectively (Figure 2). Intriguingly, whole-genome transcriptional profiling has shown that *af12070* and *af12050* gene expression is down-regulated in *A. fumigatus* Af293  $\Delta$ *brlA* or  $\Delta$ *stuA* mutants, suggesting that the production of the fumiquinazoline

metabolites is developmentally regulated and tied to conidiation (asexual sporulation).<sup>12</sup>

This work forms the final piece of our *in vitro* studies of fumiquinazoline biosynthesis, where together the four enzyme pathway of Af12050 through Af12070 recapitulates all three levels of FQ metabolite architectural complexity generation. We show that Af12070 is a flavoprotein amide oxidase (acting on R-[R']-CH-NH-CO-R'') of notable similarity to the more prevalent amine oxidases (which act on R-[R']-CH-NH-R''<sup>13</sup>). The regioselective amide oxidation of FQA by Af12070 yields the architecturally complex FQC as the kinetic product via the formation of an unusual spirohemiaminal linkage. FQC is also shown to slowly convert to the known cyclic aminal metabolite FQD nonenzymatically. We also examine the Af12070 processing of 2'-epi-FQA and 11'-dimethyl-2'-epi-FQA, which are known intermediates in the tryptoquialanine pathway of *P. aethiopicum* that share a large degree of structural similarity to FQA<sup>8</sup> (Figure S1). These studies allowed us to observe the effects of these remote site substitutions on the cyclic aminal/hemiaminal product ratio and investigate the affect of these remote structural changes on product outcome.

## EXPERIMENTAL PROCEDURES

**General Materials and Methods.** Oligonucleotide primers were purchased from Integrated DNA Technologies (Coralville, IA). PCR reactions were carried out using Phusion High-Fidelity PCR MasterMix (New England Biolabs). Plasmid DNA was propagated in *E. coli* XL1 Blue (Stratagene) and prepared using the QIAprep Spin Mini Kit (Qiagen). DNA sequencing to confirm the correct construction of expression vectors was performed by Genewiz (South Plainfield, NJ). Protein concentration was determined spectrophotometrically using theoretical extinction coefficients obtained from the online ProtParam tool.<sup>14</sup> An Agilent Technologies 6520 Accurate-Mass QTOF instrument was used for high-resolution LC-MS analysis, and a Beckman Coulter System Gold instrument equipped with diode-array detection was used for reverse-phase HPLC. NMR data were collected on a Varian 600 MHz spectrometer using the residual solvent peak from incomplete deuteration as internal standard ( $\text{CDCl}_3$ ,  $\delta$  7.26).

**Additional Details for Bioinformatic Analysis.** Sequence alignments were performed with ClustalW<sup>15</sup> and colored using GeneDoc.<sup>16</sup> Homology modeling was performed with HHpred.<sup>17</sup> Superimposition of Af12070 with VAO-family oxidases of known structure allowed for manual docking of FAD into the active-site of Af12070. NetNGlyc 1.0 was used to predict Asn glycosylation sites, and the three sites identified were checked by homology modeling to ensure the residues were surface exposed. Of note is that while these servers are dedicated to analysis of mammalian proteins, the N-X-S/T sequence motif and the transfer of the GlcNAc substrate to the polypeptide are conserved between mammalian and fungal systems.<sup>18</sup>

**Cloning, Expression, and Purification of Af12070.** *A. fumigatus* Af293 cDNA was used for all cloning. A full-length version of Af12070 was cloned into pET-30 Ek/LIC to encode a 536 residue (59 kDa) protein with an N-terminal 6x-His tag using the PCR primers: 5'-GACGACGACAAGatgcagta-catccctctgatctc-3' (forward), 5'-GAGGAGAAGCCCGGT-TAcagactcaacggaatcgagcc-3' (reverse, bases complementary to gene are in lowercase type, underline indicates a stop codon). DNA sequencing revealed that the terminal (3') boundary for the first intron of the cloned gene was different than the 3' boundary identified for AFUA\_6g12070 in the NCBI database (GenBank accession XM\_745993). The first intron of the database entry ends at base 3,020,413 of the genomic DNA for Af293 chromosome 6 (accession NC\_007199), whereas the termination site for the sequenced clone is 12 nucleotides downstream at base 3,020,425. Consequently, the cloned transcript encodes a protein that is four amino acids shorter than the database entry AFUA\_6g12070, and the amino acid numbering subsequent to residue 87 is different.

A truncated version of Af12070 which removes the first 24 amino acids of the full-length construct (Af12070  $\Delta$ 1–24), was cloned into pET-52b 3C/LIC to encode a 511 residue (56 kDa) protein with an N-terminal Strep-tag II and C-terminal 10x-His tag using the PCR primers: 5'-CAGGGACCCGGTgacacaccaggattgatgggaatg-3' (forward), 5'-GGACCA-GAGCGTTcagactcaacggaatcgagcc-3' (reverse).

Protein overproduction was performed in *E. coli* Origami-DE3 cells in order to promote disulfide bond formation. Both constructs were expressed and protein purified in a similar manner: 4 L of cells were grown at 37 °C in LB including the

appropriate antibiotics to an  $\text{OD}_{600}$  between 0.4 and 0.8, and the temperature was lowered to 16 °C prior to induction with 0.1 mM IPTG. Cells were harvested 18–24 h postinduction by centrifugation, suspended in lysis buffer (25 mM Tris-HCl [pH 7.5], 300 mM NaCl, 20% glycerol, 0.1% Tween 20), and lysed using an EmulsiFlex-C5 homogenizer (Avestin). Insoluble material was removed by centrifugation (35000 g), and soluble protein was applied to 1–2 mL of Ni-NTA agarose (Qiagen) equilibrated in lysis buffer. Ni-affinity purification was performed by batch binding protein for 30 min at 4 °C, washing the Ni-resin twice with 20 mL of buffer A (25 mM Tris-HCl [pH 7.5], 300 mM NaCl, 10% glycerol, 0.5 mM EDTA, 25 mM imidazole), and then eluting protein using three 5 mL volumes of buffer A including 250 mM imidazole. The elutions were pooled and concentrated using a centrifugal filtration device (30K MWCO, Amicon). Additional purification of Af12070  $\Delta$ 1–24 was performed using Strep-Tactin Superflow Plus resin (Qiagen). Concentrated protein post Ni-NTA ( $\approx$ 2 mL) was added to 1 mL of resin for batch binding for 1–2 h at 4 °C, the resin was then washed with three 1 mL volumes of buffer A (pH 8.0/NaOH), and bound protein was eluted with five 0.5 mL volumes of buffer A including 2.5 mM desthiobiotin. Elutions containing target protein were concentrated by centrifugal filtration. For both full-length and truncated versions of Af12070, the concentrated protein was either stored at 4 °C for immediate use or flash-frozen in liquid  $\text{N}_2$  and stored at  $-80$  °C.

For all results sections describing the in vitro characterization of Af12070, the version of the protein referred to or utilized in the assays is the N-terminal deletion construct Af12070  $\Delta$ 1–24.

**Identification of the Covalent Attachment Site of FAD to Af12070 by Mass Spectrometry.** Identification of the flavin cofactor as FAD rather than FMN and localization of that modification to His105 was performed using trypsin digestion followed by capillary LC-MS<sup>n</sup>. All data were recorded on a custom 11T LTQ-FT Ultra (Thermo-Fisher Scientific) equipped with an Eksigent 2D-LC nanocapillary LC system and a Picoview 550 nanospray source (New Objective). Data were analyzed manually using the Qualbrowser application of Xcalibur (Thermo-Fisher Scientific). 20  $\mu\text{L}$  of 40  $\mu\text{M}$  Af12070  $\Delta$ 1–24 was denatured using 5  $\mu\text{L}$  of protein extraction reagent 4 (Sigma) for 10 min and then diluted 4-fold with 10 mM  $\text{NH}_4\text{HCO}_3$ , pH 8.0, prior to application to an immobilized trypsin column prepared according to the manufacturer's instructions (Sigma). Digestion was allowed to proceed for 1 h before the digested peptides were eluted. The digested peptides were desalted using diafiltration over an Amicon Ultracel 10K MWCO filter (Millipore). The retentate was then concentrated to dryness in a speedvac. Concentrated peptides were reconstituted in 50  $\mu\text{L}$  of 0.1% formic acid in water, and 20  $\mu\text{L}$  of this sample was injected onto a 75  $\mu\text{m}$   $\times$  80 mm Biobasic C<sub>8</sub> Picofrit column (New Objective). Peptides were eluted using a linear gradient of 5% B to 75% B over 60 min at a flow rate of 300 nL/min (A = 0.1% formic acid in 95%  $\text{H}_2\text{O}$ /5% MeCN, B = 0.1% formic acid in MeCN). MS data were acquired in profile positive ion mode at resolution 50 000 for FT scans and positive centroid mode for ion trap scans. Data dependent or targeted MS<sup>n</sup> scans were used as appropriate, where  $n$  was 2, 3, or 4 and indicates the number of stages of fragmentation used in tandem mass spectrometry.

**Preparation and Purification of FQF.** In order to generate various substrates for the in vitro testing of Af12070,

the precursor FQF was synthesized (as previously described<sup>6</sup>) or obtained by growth and extraction of *P. aethiopicum*  $\Delta$ gsfA/ $\Delta$ tqaH (double knockout reported by Gao et al.<sup>8</sup>). Spores of *P. aethiopicum*  $\Delta$ gsfA/ $\Delta$ tqaH were harvested from a single plate (100 × 20 mm) following 4-day growth at 30 °C on glucose minimal medium (GMM) agar and used to inoculate 3 L of YMEG liquid medium (4 g/L yeast extract, 10 g/L malt extract, 5 g/L glucose) for stationary culture at 25 °C for 7 days (50  $\mu$ L of spore stock to  $\approx$ 100 mL media placed in 150 × 25 mm plates). The resulting fungal mat was separated from the media and extracted three times with 1 L of ethyl acetate. The extract was concentrated under vacuum and washed twice with ethyl acetate:water (2:1), and the combined ethyl acetate layers were dried under vacuum. The residue was purified by silica gel chromatography using the solvent system ethyl acetate:hexane (3:1).

**Preparation and Purification of FQA, 2'-epi-FQA, and 11'-Dimethyl-FQA.** The enzymatic synthesis of FQA was performed in a 30 mL reaction by combining 5  $\mu$ M holo-Af12050, 2.5  $\mu$ M holo-Af12060, 1 mM ATP, 2 mM MgCl<sub>2</sub>, 1 mM L-Ala, 2 mM NADH, and 250  $\mu$ M FQF in NaPi buffer (50 mM sodium phosphate [pH 7.5], 100 mM NaCl and 5% glycerol). 2'-epi-FQA was generated using a similar enzymatic reconstitution procedure, but with 5  $\mu$ M holo-TqaB (the trimodular NRPS of the tryptoquialanine-pathway<sup>19</sup>) in place of Af12050. 11'-Dimethyl-FQA was also prepared similarly, but with 5  $\mu$ M holo-Chimera2 (which contains the A-T domains of TqaB for 2-AIB activation and the C-domain of Af12050<sup>19</sup>) in place of Af12050 and 1 mM 2-AIB in place of L-Ala. Reactions were incubated overnight ( $\approx$ 16 h) at room temperature, quenched by adding an equal volume of MeCN, and centrifuged to remove precipitate. The supernatant following centrifugation was concentrated under vacuum to remove MeCN, flash frozen, and then lyophilized. The resulting residue was taken up in 5 mL of 25% MeCN, filtered (0.2  $\mu$ m PTFE membrane), and purified by reverse-phase HPLC (Phenomenex Luna C18, 250 × 21.2 mm, 10  $\mu$ m) with detection at 254 nm. Solvent systems A (water plus 0.1% TFA) and B (MeCN plus 0.1% TFA) were held at 25% B for 1 min and then run over a linear gradient of 25–53% over 30 min, followed by a gradient of 53–95% B over 1 min before a holding at 95% B for 5 min; the column was then equilibrated back to initial conditions by returning to 25% B and holding for 8 min.

FQA HRMS:  $m/z$  calculated for C<sub>24</sub>H<sub>23</sub>N<sub>5</sub>O<sub>4</sub>: 446.1823 [M + H]<sup>+</sup>. Found: 446.1826.

2'-epi-FQA HRMS:  $m/z$  calculated for C<sub>24</sub>H<sub>23</sub>N<sub>5</sub>O<sub>4</sub>: 446.1823 [M + H]<sup>+</sup>. Found: 446.1824.

11'-Dimethyl-FQA HRMS:  $m/z$  calculated for C<sub>25</sub>H<sub>25</sub>N<sub>5</sub>O<sub>4</sub>: 460.1979 [M + H]<sup>+</sup>. Found: 460.1983.

For spectroscopic characterization and assignments of fumiquinazoline A analogues and comparison to fumiquinazoline A see Table S4.

**O<sub>2</sub> Consumption by Af12070.** A Hansatech oxygen electrode (S1 Clark type) with an Oxygraph control unit was used to monitor O<sub>2</sub> levels in sample reactions reconstituting Af12070 oxidation of FQA. A reaction volume of 300  $\mu$ L was used and contained 200  $\mu$ M FQA and 2  $\mu$ M Af12070 in NaPi buffer. Substrate, buffer, and water were combined in the reaction vessel (with stirring set to 100 rpm and a temperature of 25 °C), and the baseline was allowed to stabilize before initiating reactions by addition of enzyme. Controls were run that omitted either enzyme (with buffer injected in its place) or substrate. Reactions and controls were run in triplicate, and

initial rate data were used for calculation of apparent velocity and turnover number (taking into account 46% holoprotein).

**HPLC and LC-MS Assays for Af12070.** For time course analysis of substrate utilization by Af12070, reactions (typically 400  $\mu$ L) were setup containing 2  $\mu$ M Af12070 and 200  $\mu$ M substrate (FQA, 2'-epi-FQA, 11'-dimethyl-FQA, or 11'-dimethyl-2'-epi-FQA) in NaPi buffer. Reactions were initiated with enzyme and time points taken between 0.1 and 20 h by quenching 50  $\mu$ L aliquots with an equal volume of MeCN. The precipitant was removed from quenched reactions by centrifugation, and 20  $\mu$ L samples of the supernatant were injected for HPLC using an Alltima C18 column (150 × 4.6 mm, 5  $\mu$ m), with detection at 254 nm. Solvent system A (water plus 0.1% TFA) and B (MeCN plus 0.1% TFA) was held at 25% B for 1 min and then run over a linear gradient of 25–55% B over 20 min, followed by a gradient of 55–95% B over 1 min before a holding at 95% B for 2.5 min; the column was then equilibrated back to initial conditions by returning to 25% B and holding for 5 min. To determine the rate of FQA utilization, for each time point the peak corresponding to substrate was integrated and converted to a concentration using a standard curve made by injecting 20  $\mu$ L samples of known concentration. Plotting the concentration of FQA vs time allowed for the calculation of initial rate data. Turnover was calculated based on 46% holo-Af12070.

**Genetic Manipulation and Transformation of *P. aethiopicum* and Characterization of Metabolites.** *tqaB* was deleted in the wild-type *P. aethiopicum* strain using *bar* as a selection marker as previously described.<sup>8</sup> pBARGPE1 (obtained from the Fungal Genetics Stock Center) is a fungal expression vector and contains the *A. nidulans* *gpdA* promoter and *trpC* terminator.<sup>20</sup> The whole *trpC* promoter and *bar* gene fragment were replaced by the *trpC* and *Zeocin* genes using SpeI and PmlI restriction sites to yield the pZeoGPE1. *af12050* was amplified and digested with *Bam*HI and *Afe*I and inserted into pBARGPE1 (digested with *Bam*HI and *Eco*RV) to yield *af12050*-pZeoGPE1 plasmid. The gene for *chimera2* was amplified and digested with *Afe*I and inserted into pBARGPE1 (digested with *Eco*RV) to yield *chimera2*-pZeoGPE1 plasmid.  $\sim$ 20  $\mu$ g of *af12050*-pZeoGPE1 or *chimera2*-pZeoGPE1 plasmid was transformed into the *P. aethiopicum*  $\Delta$ tqaB strain using *Zeocin* as a second selection. Poly(ethylene glycol)-mediated transformation of *P. aethiopicum* was performed essentially as previously described.<sup>8,21</sup> Genomic DNA from *P. aethiopicum* transformants was used for PCR screening. Integration of the *af12050*/*chimera2*-overexpression cassette was confirmed by PCR. For small-scale analysis, the *P. aethiopicum* wild-type and transformants were grown in stationary YMEG liquid culture for 4 days at 25 °C. The cultures were extracted with equal volumes of ethyl acetate and evaporated to dryness. The dried extracts were dissolved in methanol for LC/MS analysis. LC/MS spectra were obtained on a Shimadzu 2010 EV liquid chromatography mass spectrometer using positive and negative electrospray ionization and a Phenomenex Luna 5  $\mu$ m, 2.0 mm × 100 mm C18 reverse-phase column. Samples were separated on a linear gradient of 5–95% CH<sub>3</sub>CN in water (0.1% formic acid) for 30 min at a flow rate of 0.1 mL/min followed by isocratic 95% CH<sub>3</sub>CN in water (0.1% formic acid) for another 15 min.

**Isolation and Characterization of FQC/FQD and Analogues.** (i) FQC and FQD. A 20 mL reaction containing 5  $\mu$ M Af12070 and 200  $\mu$ M FQA was incubated overnight at room temperature, and the sample was quenched, processed, and purified as described above for the scaled production of

FQA-(like) substrates. For FQC and FQD spectroscopic characterization and assignments see Table S1.

FQC HRMS:  $m/z$  calculated for  $C_{24}H_{21}N_5O_4$ : 444.1666  $[M + H]^+$ . Found: 444.1667.

FQD HRMS:  $m/z$  calculated for  $C_{24}H_{21}N_5O_4$ : 444.1666  $[M + H]^+$ . Found: 444.1662.

(ii) 2'-*epi*-FQD. A 25 mL reaction containing 2  $\mu$ M Af12070 and 250  $\mu$ M 2'-*epi*-FQA was incubated overnight at room temperature, and the sample was quenched, processed, and purified as described above for the scaled production of FQA-(like) substrates. For 2'-*epi*-FQD spectroscopic characterization and assignments see Table S2.

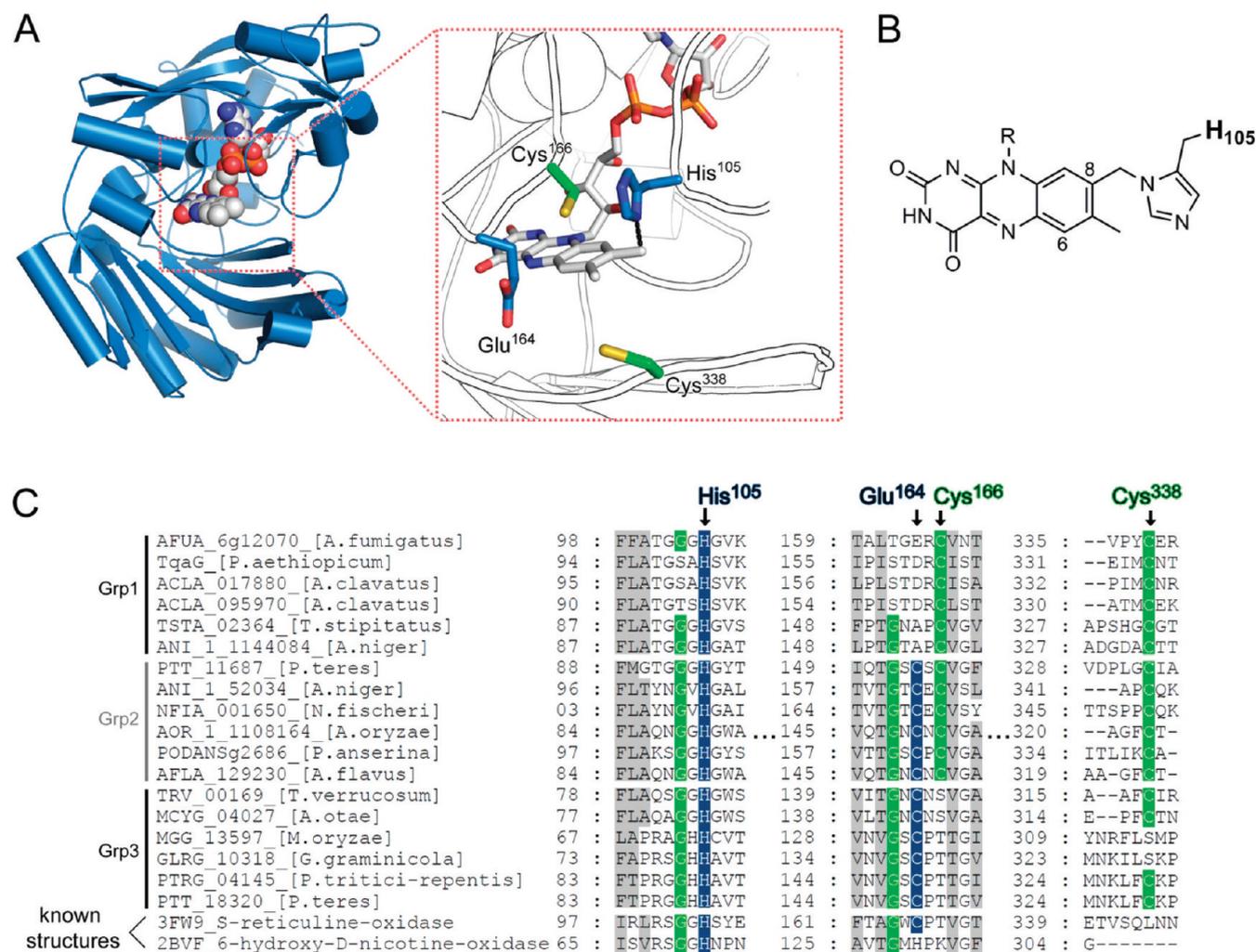
HRMS:  $m/z$  calculated for  $C_{24}H_{21}N_5O_4$ : 444.1666  $[M + H]^+$ . Found: 444.166.

(iii) 1,1'-Dimethyl-FQC. A 25 mL reaction containing 2  $\mu$ M Af12070 and 300  $\mu$ M 1,1'-dimethyl-FQA was incubated overnight at room temperature, and the sample was quenched, processed, and purified as described above for the scaled production of FQA-(like) substrates. For 1,1'-dimethyl-FQD spectroscopic characterization and assignments see Table S3.

HRMS:  $m/z$  calculated for  $C_{25}H_{23}N_5O_4$ : 458.1823  $[M + H]^+$ . Found: 458.1822.

## RESULTS

**Bioinformatic Analysis of Af12070 and TqaG.** To gain insight into the possible function of the remaining ORF in the FQ gene cluster, Af12070, we first turned to bioinformatic analysis. Af12070 and TqaG (46% identity/72% similarity) belong to the vanillyl alcohol oxidase (VAO) family of FAD-dependent oxidoreductases.<sup>22</sup> A characteristic feature of the VAO family is their propensity to form mono- and bivalent linkages to the isoalloxazine of FAD, a self-catalytic process (termed autoflavinylation) that increases the redox potential of the enzyme, enhances structural integrity and protein stability, and protects the coenzyme from modification and enzyme inactivation.<sup>23</sup> Sequence alignment and homology modeling predict that Af12070 and TqaG are monovalent flavoproteins, which form a histidyl linkage to the 8 $\alpha$  methyl of FAD (Figure 3). Homologues of Af12070, identified by a BLASTP search, are predicted to possess both mono- and bivalent



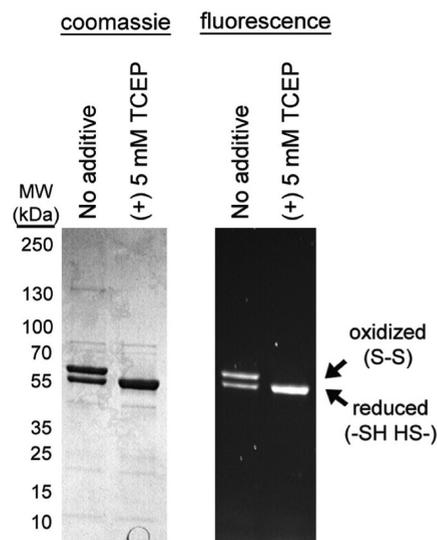
**Figure 3.** Bioinformatic analysis of FAD binding by Af12070 and homologous proteins. (A) Homology model of Af12070 illustrating the predicted covalent attachment of FAD to His105 via an 8 $\alpha$ -N1-histidyl linkage (black dashed line), the position of residue Glu164 (which is replaced by a cysteine in the bivalent FAD proteins in order to generate an additional 6-S-cysteinyll attachment), and two cysteines (residues 156 and 338) which may form a disulfide linkage near the dimethylbenzene of FAD. (B) 2D representation of the predicted 8 $\alpha$ -N1-histidyl FAD linkage of Af12070. (C) Sequence alignment of select vanillyl alcohol oxidase (VAO) family proteins homologous to Af12070. The top 17 sequences (labeled as groups [Grp] 1, 2, and 3) were obtained from a BLASTP search using Af12070 as the query. The bottom two sequences are structural homologues identified by HHpred and represent proteins with bivalent FAD (3FW9) or monovalent FAD (2BVF).

links to FAD (Figure 3C, Grp1 and Grp2). For the putative bivalent flavoprotein homologues (which constitute the majority of the top 100 BLASTP results), the residue Glu164<sup>a</sup> of Af12070 is replaced by a Cys residue, this Cys forms the 6-S-cysteinylyl FAD linkage in the known bivalent flavoprotein (S)-reticuline oxidase (PDB id: 3F9W) (Figure 3C). The BLAST search also shows that two Cys residues (166 and 338 of Af12070), located adjacent to the dimethylbenzene portion of FAD and above the entrance to the active site, are highly conserved within groups 1 and 2 (Figure 3C). Though their inter-residue distance is large in the homology model (7.5 Å), in actuality it is plausible to envision that these Cys residues may form a disulfide bond, thus conferring even greater structural stability on this set of proteins. A small fraction of the top 100 results from BLAST (less than 10%) have a Ser or Thr at the position equivalent to Cys166 of Af12070 (Figure 3C, Grp3), consequently eliminating the possibility of a disulfide at this position and setting this group apart from other Af12070 homologues.

Bioinformatics analyses also indicate that Af12070 may contain an N-terminal signal sequence and multiple N-glycosylation sites. Sequence alignment and modeling identify an unstructured region of Af12070 and TqaG which extends  $\approx 35$  amino acids beyond the FAD-binding domain (Figure S2). PSORT and SignalP analysis predict that the first 17 residues of Af12070 constitute a cleavable hydrophobic signal sequence that directs protein export. A signal peptide is also predicted for TqaG, though the location of the cleavage site is more ambiguous. N-terminal signal sequences have been reported for several VAO-family enzymes, including THCA synthase,<sup>24</sup> (S)-reticuline oxidase (berberine bridge enzyme),<sup>25</sup> cytokinin dehydrogenase (maize isoform ZmCKX1),<sup>26</sup> and aryl alcohol oxidases.<sup>27</sup> Of these, signal peptide directed secretory trafficking has previously been described for THCA synthase and cytokinin dehydrogenase.<sup>24,26</sup> The possibility that the N-terminal signal peptide of Af12070 directs secretion is supported by the prediction that the enzyme is N-glycosylated at up to 3 sites (Asn95, 257, and 284) each of which contains the consensus sequence Asn-X-Ser/Thr and are predicted to be surface exposed based on 3D modeling.

Taken together, the predicted signal peptide, disulfide bond, and Asn glycosylation sites strongly suggest that Af12070 is secreted. Future studies investigating the cellular trafficking of Af12070 will be of interest in determining if, as predicted, the enzyme is exported in order to function biosynthetically, for example by deposition into developing conidia during sporulation. Notably, as we describe below, the prediction of this signal sequence was also experimentally consequential for the heterologous expression of Af12070 as an active, soluble enzyme.

**Recombinant Production and Purification of Af12070 from *E. coli*.** Initial efforts to overproduce the full-length protein encoded by the gene AFUA\_6g12070 (493 amino acids, abbreviated as Af12070) in *E. coli* gave some soluble, active protein, but yields were low and the protein was impure and possessed low activity. Cloning of a construct that deleted the first 24 residues of Af12070 (Af12070  $\Delta 1-24$ ) and expression in *E. coli* Origami(DE3) led to improved solubility and yield, increased purity and stability, and higher enzymatic activity. Af12070  $\Delta 1-24$  was purified as a yellow protein of >90% purity and with a yield of  $\approx 1$  mg/L following dual affinity chromatography (utilizing Ni-NTA and Strep-Tactin resins) (Figure 4). As purified from *E. coli* Origami(DE3) cells,

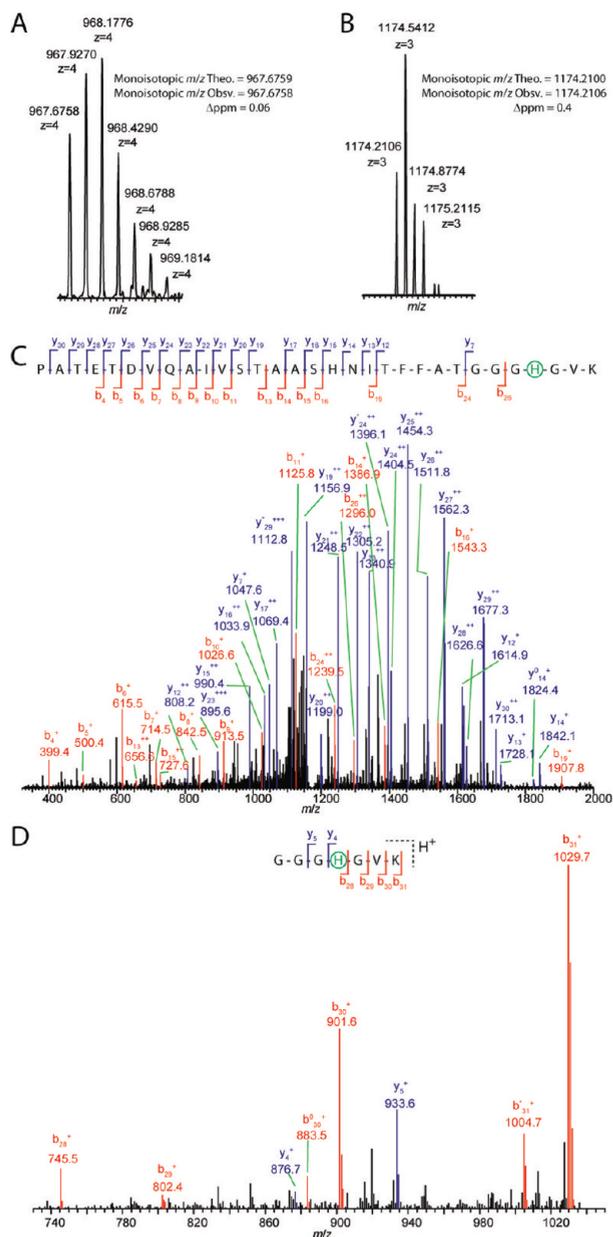


**Figure 4.** SDS-PAGE analysis of purified Af12070  $\Delta 1-24$  (N-terminal truncation of the first 24 amino acids) with staining by Coomassie (left) or visualization by fluorescence imaging (right, 365 nm excitation).

Af12070  $\Delta 1-24$  migrates as two closely spaced bands in nonreducing SDS-PAGE. Upon treatment with reducing agents the top band disappears and the intensity of the bottom band increases, suggesting that the doublet is composed of oxidized (S–S, top band) and reduced (–SH HS–, bottom band) forms of the protein. As there are only three Cys residues present in Af12070, the top band likely represents the Cys166-Cys338 disulfide form of the protein as postulated in the previous section. Though the oxidized species of the protein could be eliminated by adding chemical reducing agents, converting the protein from the reduced to oxidized form proved more difficult; neither prolonged incubation and exposure to air nor chemical treatment (with H<sub>2</sub>O<sub>2</sub> or oxidized glutathione) afforded a significantly increased proportion of the oxidized species. Attempts at overproducing the full-length or an N-terminal deletion construct of the homologous TqaG utilizing either bacterial or yeast expression systems have been unsuccessful; therefore, we have focused on Af12070 for our *in vitro* studies during this work.

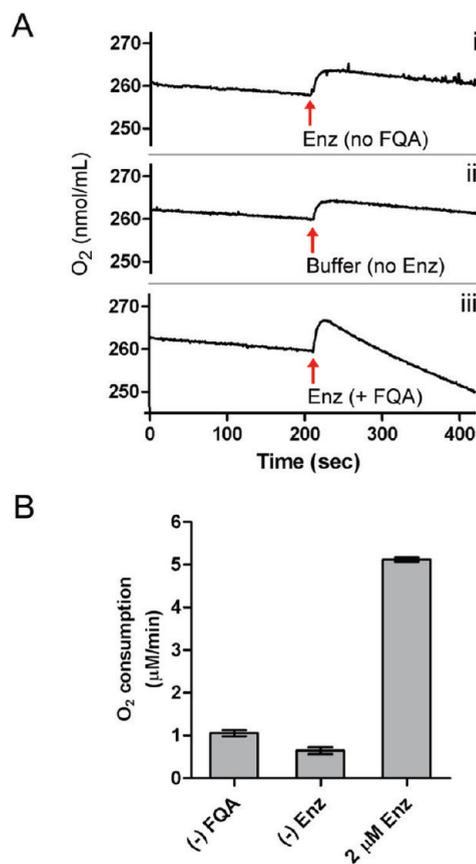
**Af12070 Is a Covalent Flavoprotein and an O<sub>2</sub>-Dependent FQA Oxidase.** Denaturing SDS-PAGE with fluorescence visualization (prior to Coomassie staining) confirmed that Af12070 contained covalently bound flavin (Figure 4). The flavin content of the purified enzyme was calculated to be 46% by UV–vis analysis (Figure S3). Trypsin digestion followed by capillary LC-MS<sup>n</sup> localized the flavin attachment site to His105 and confirmed the identity of the covalent flavin as FAD (Figure 5).

Given that Af12080, 12060, and 12050 act together to generate FQA (Figure 1) and that FQC and FQD are more complex scaffolds predicted to arise from enzymatic modification of FQA, we evaluated if Af12070 would, as predicted, utilize FQA as a substrate. We generated substrate quantities of FQA via incubations of synthetic FQF with Af12060/12050<sup>7</sup> (described in Experimental Procedures) and then turned to evaluate the ability of O<sub>2</sub> to act as a reducible cosubstrate of Af12070. O<sub>2</sub> consumption by Af12070 was determined using an oxygen electrode; combining 2  $\mu$ M Af12070  $\Delta 1-24$  with 200  $\mu$ M FQA (in buffer at pH 7.4) led to a 5-fold increase in



**Figure 5.** LC-MS<sup>n</sup> analysis of Af12070  $\Delta$ 1–24 covalent flavin modification. (A) A  $[M+4H]^{4+}$  ion corresponding to the tryptic peptide Pro74-Lys104 with attached FAD ( $m/z = 967.6758$ ) was found in the LC-MS trace and subsequently targeted for MS<sup>2</sup>. (B) The MS<sup>2</sup> spectrum of the peptide from panel A yielded only a single intense  $[M+3H]^{3+}$  peak corresponding to the Pro74-Lys104 peptide with the neutral loss of AMP from the FAD modification ( $m/z = 1174.2106$ ), and this peak was subsequently targeted for MS<sup>3</sup>. (C) MS<sup>3</sup> analysis localized the FAD modification to the C-terminal region of the peptide. (D) MS<sup>4</sup> analysis of the  $\gamma_7^+$  ion at  $m/z$  1047.6 yielded several fragment ions including the  $y_4^+$  at  $m/z$  876.7 and the internal ion  $[\gamma_7-b_{28}]^+$  at  $m/z$  745.5 that localized the FAD modification precisely to His105.

O<sub>2</sub>-consumption compared to controls lacking enzyme or substrate (Figure 6). This result indicates that O<sub>2</sub> is the electron acceptor for regeneration of oxidized flavin during the oxidative half-reaction of the catalytic cycle.<sup>28</sup> Under the conditions assayed, and taking into account the observed ratio of holo-Af12070 (46%), the apparent turnover number for FQA oxidation was 4.6 min<sup>-1</sup>. This rate is of comparable magnitude to the rate of



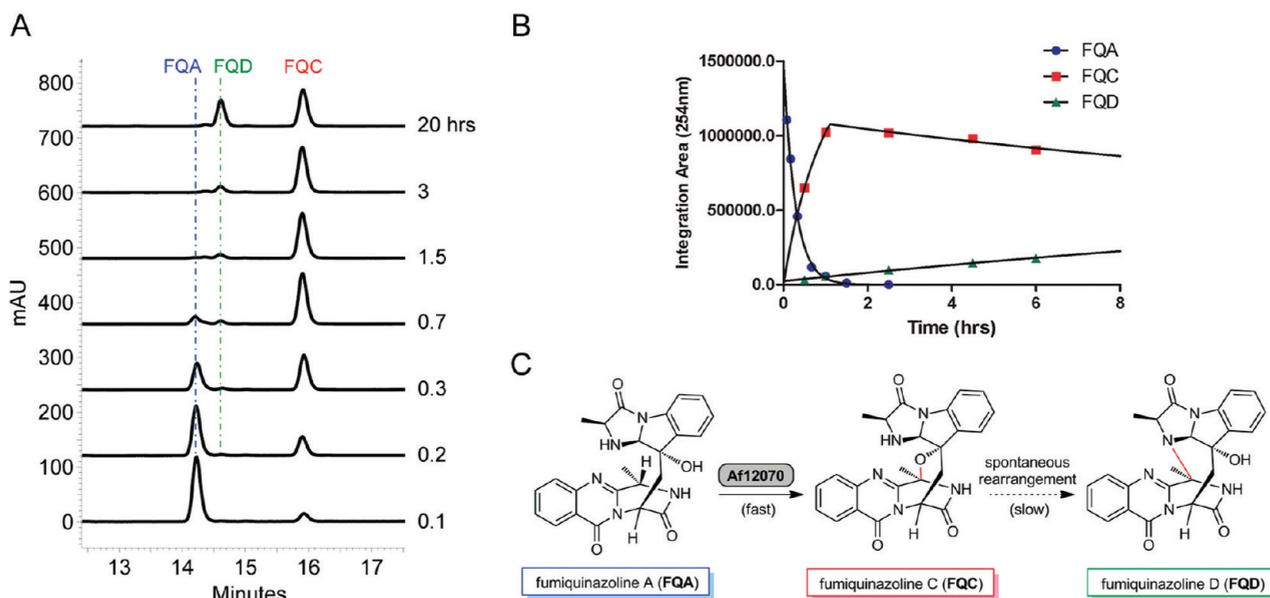
**Figure 6.** Af12070  $\Delta$ 1–24 utilizes O<sub>2</sub> as an electron acceptor in the oxidation of FQA. (A) Oxygen electrode trace data representing (i) no substrate control (buffer plus the addition of 2  $\mu$ M Af12070 as indicated), (ii) no enzyme control (buffer including substrate with the addition of an extra volume of buffer as indicated), and (iii) reconstitution of FQA oxidation by Af12070 (buffer including FQA plus the addition of 2  $\mu$ M Af12070 as indicated). (B) Initial rate data obtained from triplicate measurements of the experiments in panel A.

FQA synthesis catalyzed in tandem by the monooxygenase Af12060 and NRPS Af12050.<sup>7</sup>

### HPLC Analysis of Af12070 Activity toward Its Natural Substrate FQA.

In order to identify the products resulting from FQA oxidation in vitro, we followed the enzymatic activity of Af12070 over time by HPLC. Incubating 2  $\mu$ M Af12070  $\Delta$ 1–24 with 200  $\mu$ M FQA in buffer (pH 7.4) resulted in the disappearance of FQA (retention time  $[t_R] = 14.2$  min) with the initial formation of a new peak at  $t_R = 15.9$  min and then the slow appearance of a second peak at  $t_R = 14.6$  min (Figure 7A,B). Determining the rate of FQA disappearance through peak integration provided an apparent turnover number for FQA oxidation of 5.2 min<sup>-1</sup>, which is consistent with the value obtained by monitoring oxygen consumption with the oxygen electrode. The addition of reducing agents to the in vitro reaction had a variable effect on enzymatic activity, 5 mM TCEP enhanced activity  $\approx$ 2-fold, whereas 5 mM DTT decreased activity  $\approx$ 3-fold, and 5 mM  $\beta$ -ME essentially inactivated the enzyme (Figure S4).

MS analysis indicated that both of the new peaks that appear when Af12070 is incubated with FQA had a mass change of  $-2$  Da compared to FQA, a mass that corresponds to two compounds previously identified from *A. fumigatus* culture extracts, FQC and FQD (FQC,  $[M + H]^+$  observed 444.1667,



**Figure 7.** HPLC-based timecourse data (with detection at 254 nm) for 2  $\mu\text{M}$  Af12070  $\Delta 1$ –24 with 200  $\mu\text{M}$  FQA as substrate. (A) HPLC chromatogram overlay of reaction aliquots quenched at the indicated time by addition of MeCN prior to injection for analysis. (B) Graphical representation of the time course data obtained by integrating the peak corresponding to FQA, FQC, and FQD. (C) Schematic illustrating the observed pattern of product formation when Af12070 was incubated in the presence of substrate FQA.

expected 444.1666; FQD,  $[\text{M} + \text{H}]^+$  observed 444.1666, expected 444.1666).<sup>4</sup>  $^1\text{H}$  NMR characterization and comparison with the literature confirmed the identity of these new peaks as FQC and FQD (Table S1) and led to the peak assignments as provided in Figure 7A. These products arise from the intramolecular cyclization of FQA as a result of the nucleophilic attack of either the free hydroxyl or the secondary amine of the imidazoindolone moiety on C3 of the quinazolinone (Figure 7C). The additional bond formed in FQC generates a bridging 7-membered ring, while in FQD a bridging 8-membered ring is formed. As illustrated in Figure 7, FQA is almost entirely converted to FQC before the appearance of FQD becomes obvious (by  $\approx 0.7$  h), and then FQD slowly grows in (from 0.7 to 20 h) with the concomitant decrease of FQC.

To evaluate if the slow FQC to FQD transformation was catalyzed by Af12070 or was nonenzymatic, a sample of purified FQC was incubated in buffer (pH 7.0) without enzyme. Nonenzymatic, spontaneous rearrangement to FQD occurred, along with the appearance of a substantial amount of a new peak with  $m/z = 462$  which eluted  $\approx 1$  min prior to FQD (Figure S5A). The new peak has a mass ( $[\text{M} + \text{H}]^+$  observed 462.1772, expected 462.1771) and UV–vis spectral properties consistent with the previously reported FQB, a C3-hydroxyl derivative of FQA (3-OH-FQA) that could arise from hydration of FQD/FQC or more probably hydration of a transient N2–C3 imine.<sup>4</sup> The accumulation of the hydrate indicates that the intermolecular addition of water effectively competes with the intramolecular capture of the imine by the hydroxyl or secondary amine groups of the imidazoindolone. Incubation of purified FQD without enzyme also resulted in interconversion to FQC (to low level) and in the accumulation of the putative hydration product 3-OH-FQA (Figure S5A). Comparing the relative amounts of each compound formed (or remaining) following the week-long incubation revealed that spontaneous equilibration favors FQD over FQC, and hydration is favored over either intramolecular product. We

hypothesize that FQC and FQD interconvert through the formation of a transient N2–C3 imine and that this imine intermediate is the nascent product of Af12070 enzymatic action (Figure S5B).

With this knowledge we propose that Af12070 is an FAD-dependent amide oxidase (acting on  $\text{R}-[\text{R}']\text{-CH-NH-CO-R}''$ ) compared to the more prevalent amine oxidases (which act on  $\text{R}-[\text{R}']\text{-CH-NH-R}''$ ).<sup>13</sup> The nascent imine generated by Af12070 undergoes efficient intramolecular capture by the pendant alcohol in FQA to form the 7-member spirohemiaminal FQC as the kinetic product. Slow equilibration ultimately yields the 8-membered aminal containing FQD as the thermodynamic adduct (Figure 8). It is interesting to speculate that the intramolecular attack of the pendant  $-\text{OH}$  to form the spirohemiaminal linkage in FQC may be favored within the active site of Af12070 by the orientation of the  $-\text{OH}$  nucleophile and the nascent imine.

**Af12070 Activity toward Analogues of FQA Reveals That Subtle Structural Differences in Substrate Can Have a Dramatic Impact on End Products.** From our previously reported characterization of the enzymatic elaboration of FQF to FQA by Af12060 and 12050,<sup>7</sup> we had noted that the homologous enzyme pair TqaH and TqaB from *P. aethiopicum* produced 2'-*epi*-FQA from FQF and L-Ala and 11'-dimethyl-2'-*epi*-FQA when 2'-aminoisobutyric acid (AIB) is incorporated in place of L-Ala.<sup>8</sup> An HPLC-based assay was performed to test the ability of Af12070 to utilize these two FQA analogues as substrates and to investigate the effects of structural changes on product outcome. 2'-*epi*-FQA was processed smoothly by Af12070  $\Delta 1$ –24 to give a single new peak (Figure 9A). This compound was isolated and characterized by NMR spectroscopy to be the amine-linked 2'-*epi*-FQD ( $[\text{M} + \text{H}]^+$  observed 444.1662, expected 444.1666) (Table S2), a result that is in contrast to the preference for intramolecular capture by the pendant 3'-OH of FQA to form FQC.

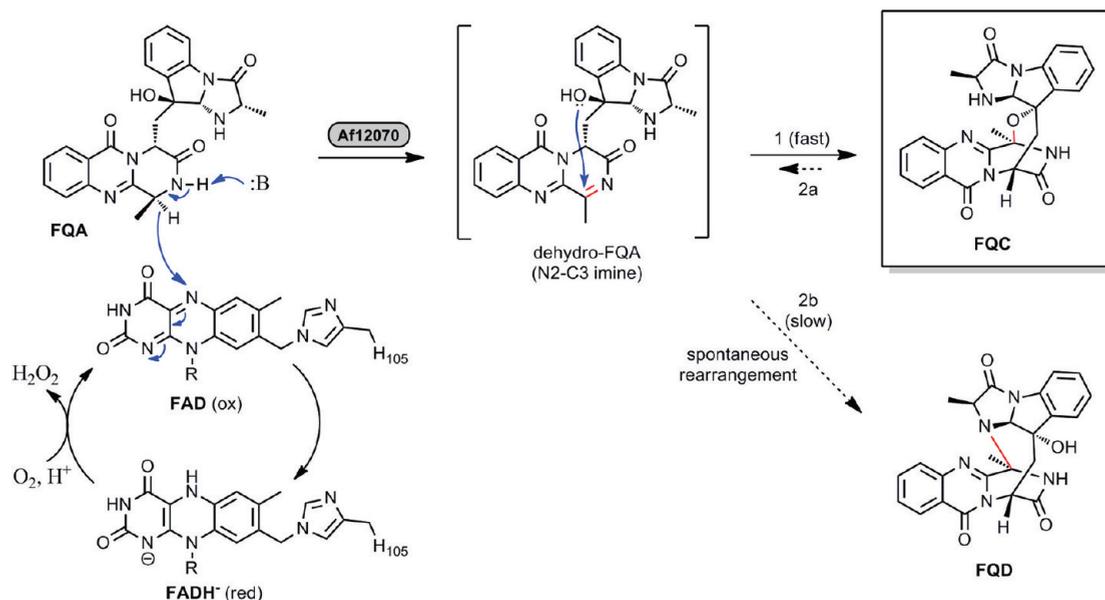


Figure 8. Proposed mechanism for Af12070.

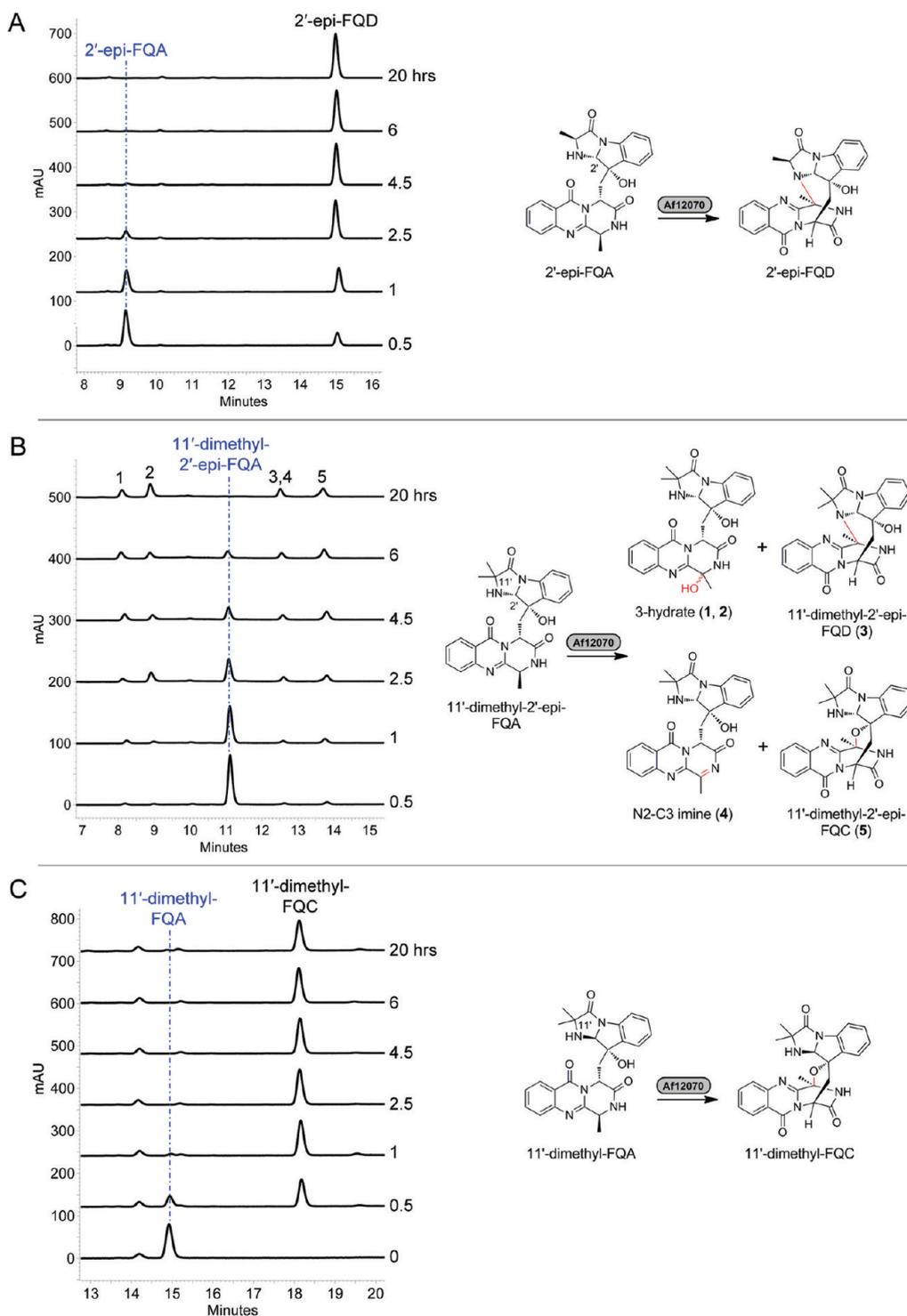
When 11'-dimethyl-2'-epi-FQA ( $m/z = 460$ ) was evaluated with Af12070  $\Delta 1-24$ , a series of five products were detected (Figure 9B and Figures S6 and S7). Two peaks (3 and 4) overlapped during HPLC analysis but separated when analyzed by LC-MS (presumably due to differences in instrument configuration). Peak 1 ( $[M + H]^+$  observed 476.1929, expected 476.1928) and peak 2 ( $[M + H]^+$  observed 476.1930, expected 476.1928) had a mass consistent with imine hydration and are thought likely to be 3-OH diastereomers of the corresponding dimethyl analogues of the previously characterized FQB (3-OH-FQA).<sup>4</sup> Peak 3 ( $[M + H]^+$  observed 458.1826, expected 458.1823) and peak 5 ( $[M + H]^+$  observed 458.1828, expected 458.1823) appear to be the 11'-dimethyl-2'-epi analogues of FQD and FQC, respectively (peak assignments are based on relative retention times compared to the results in Figure 7A).

The mass data obtained for peak 4 indicates the presence of three different compounds with  $m/z$  of 458, 459, and 477 (Figure S8). Interestingly, the retention time of peak 4 matches that of the major peak present in the extract of *P. aethiopicum*  $\Delta tqaI$ , which has been postulated to be a keto acid resulting from hydrolysis of the N2-C3 imine intermediate (TqaI encodes a putative trypsin-like serine protease and is proposed to act immediately downstream of the Af12070 homologue TqaG) (Figure S8).<sup>8</sup> The previous characterization of the keto acid ( $m/z = 477$ ) by Gao et al. was accompanied by spontaneous lactonization to yield deoxytryptoquialanone ( $m/z = 459$ ).<sup>8</sup> Comparison of our in vitro reaction with an authentic standard of deoxytryptoquialanone reveals a small peak in the standard that has a mass corresponding to the keto acid and the same retention time as peak 4 (Figure S8). This small peak presumably results from a degree of deoxytryptoquialanone lactone hydrolysis to form the keto acid during LC-MS analysis. Importantly, the mass spectrum of this standard does not contain a molecular ion with  $m/z = 458$ . It is also important to note that the retention time of (unhydrolyzed) deoxytryptoquialanone does not match that of peak 4 but that we do observe a degree of MS-mediated dehydration of the keto acid to give a species with the same mass and formula as deoxytryptoquialanone; i.e., the presence of the keto acid can explain the species with  $m/z$  of both 459 and 477 in peak 4, but

not that of 458. This indicates that the compound from our in vitro reaction with  $m/z = 458$  is structurally distinct from both the keto acid and deoxytryptoquialanone. The identification of a peak with  $m/z = 458$ , which appears intractable to isolation and characterization, is consistent with the presence of a nascent imine intermediate that hydrolyzes to the observed keto acid under LC-MS analysis and is further masked by a subsequent lactonization.

A third substrate analogue, 11'-dimethyl-FQA, was used to further our analysis of the effects of molecular structure on the fate of the putative enzymatic product of Af12070 (the N2-C3 imine). The substrate 11'-dimethyl-FQA (which is not a known natural product) was obtained using a chimeric monomodular NRPS which combines the A- and T-domains of TqaB (for aminoisobutyric acid selection and incorporation) and the C-domain of Af12050 (to retain the "natural" stereochemistry at C2').<sup>19</sup> When 11'-dimethyl-FQA was incubated with Af12070  $\Delta 1-24$ , we observed the complete consumption of substrate and the formation of a single new peak (Figure 9C); characterization by MS ( $[M + H]^+$  observed 458.1822, expected 458.1823) and NMR spectroscopy revealed that this peak is the hydroxyl-linked 11'-dimethyl-FQC (Table S3). When compared to the results obtained using FQA as substrate (Figure 7A), the absence of other product peaks appearing over time indicates that the additional methyl substituent at C11' discourages the spontaneous interconversion of 11'-dimethyl-FQC to an amine-linked FQD analogue.

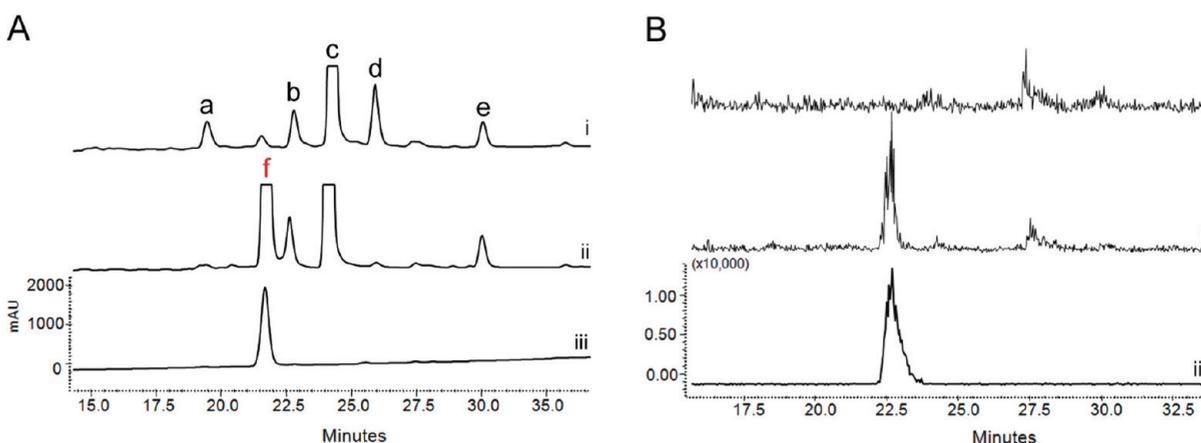
**Genetic Knock-in Studies for the In Situ Production of FQA and Dimethyl-FQA in *P. aethiopicum*.** To provide insight into the substrate specificity of the Af12070 homologue TqaG, and because we could not produce it heterologously in a soluble, active form in vitro, we utilized an in vivo platform previously reported for our characterization of the tryptoquialanine pathway.<sup>8</sup> We developed a *P. aethiopicum* strain that can synthesize FQA derivatives that have the same *trans* stereochemistry across C2'-C3' (C2'-S) as in the *A. fumigatus* fumiquinazoline pathway. To do so, we first deleted the *tqaB* gene in the wild-type *P. aethiopicum* using *bar* as the selection marker. We then knocked-in the functionally homologous *af12050* using the recombination plasmid pZeoGPE1, which



**Figure 9.** HPLC time course data (detection at 254 nm) for 2  $\mu$ M Af12070 with 200  $\mu$ M 2'-epi-FQA (A), 11'-dimethyl-2'-epi-FQA (B), or 11'-dimethyl-FQA (C) as substrates. The structures of the products in panels A and C were determined by NMR spectroscopy, while the structures of products 1–5 in panel B are proposed based on high-resolution MS, UV-vis, and published precedent.

includes the Zeocin resistance gene controlled by the *trpC* promoter and the *af12050* gene controlled by the *gpdA* promoter.<sup>20</sup> The replacement of TqaB with Af12050 is expected to yield the intermediate FQA in vivo and thus expose FQA to potential oxidative catalysis by TqaG. However, experiments utilizing the  $\Delta$ *tqaB::af12050* mutant show efficient in vivo production of FQA, but there were no compounds detected corresponding to TqaG-mediated oxidation of FQA

(Figure 10A). We also used the *P. aethiopicum*  $\Delta$ *tqaB* strain to knock-in a hybrid monomodular NRPS (Chimera2<sup>19</sup>) for the in vivo production of 11'-dimethyl-FQA. Similar to the results obtained with the  $\Delta$ *tqaB::af12050* mutant, 11'-dimethyl-FQA is produced by the  $\Delta$ *tqaB::chimera2* mutant, but oxidation of this compound by TqaG is not observed (Figure 10B). For both knock-in mutants, PCR verification was performed to ensure that the chromosomal copy of *tqaG* gene was not interrupted



**Figure 10.** HPLC and LC-MS traces of metabolite extracts obtained from gene deletion/knock-in strains of *P. aethiopicum* used for the in situ generation of FQA or 11'-dimethyl-FQA. (A) HPLC data (254 nm detection) of (i) extract from  $\Delta tqaB$ , (ii) extract from  $\Delta tqaB::af12050$ , and (iii) an FQA standard. Lowercase letters labeling peaks: a, dechlorogriseofulvin; b, griseofulvin; c, oxy-FQF-dimer; d, viridicatumtoxin; e, FQA. (B) LC-MS data mass filtered for  $m/z$ : 460 (equal to the  $[M + H]^+$  species for dimethyl-FQA): trace i, extract from  $\Delta tqaB$ ; trace ii, extract from  $\Delta tqaB::chimera2$ ; and trace iii, an 11'-dimethyl-FQA standard.

by plasmid insertion through random integration. These results may indicate that Af12070 and TqaG have differing substrate specificities, namely, that TqaG has a strict requirement for a substrate with (*R*)-configuration at C2' (e.g., the naturally occurring 11'-dimethyl-2'-*epi*-FQA or 2'-*epi*-FQA). Alternatively, these negative results could arise from differential subcellular localization of substrate (FQA/11'-dimethyl-FQA) and enzyme (TqaG) and simply reflect the inability of the TqaG to act on a substrate that it does not come into contact with. Obtaining and testing purified TqaG in vitro is one way to distinguish between these possibilities, though our best efforts to overproduce TqaG have been unsuccessful.

## DISCUSSION

The fumiquinolines, as the prefix “fumi” implies, are a signature class of tri- to tetrapeptidyl alkaloid natural products from the human pathogen *Aspergillus fumigatus*. Three levels of increasing complexity of the morphed peptide scaffold can be discerned (Figure 1). The simplest is FQF (and the epimeric FQG), where the tricyclic quinazoline scaffold is constructed from three amino acid building blocks: the proteinogenic *L*-Trp and *L*-Ala and the non-proteinogenic anthranilate (formally a  $\beta$ -amino acid). The next level of complexity is exemplified by conversion of FQF to FQA (and the congener FQI isolated from *Acremonium* sp.<sup>11</sup>) where a fourth amino acid, either *L*-Ala (FQA) or *L*-Leu (FQI), is added in a net annulation to the indole side chain of FQF. The third level of architectural morphing, the focus of this study, is the conversion of FQA to FQC (the corollary process in *Acremonium* is the conversion of FQI to FQH<sup>11</sup>) (Figure 2). The architectural complexity increases dramatically in these three stages of maturation; together they morph a tetrapeptide into a highly constrained small molecule scaffold.

Our entry into this ubiquitous class of *A. fumigatus* metabolites began with a search for the fungal enzymatic machinery that could draw off a portion of the anthranilate pool from tryptophan biosynthesis (and other primary metabolic functions) for use in the production of natural products. We anticipated an NRPS module with an adenylation domain that could recognize and activate anthranilate for activation and incorporation into a trimodular NRPS (Af12080) assembly

line.<sup>6</sup> By those efforts we identified an FQ biosynthetic gene cluster in the *A. fumigatus* Af293 genome which utilizes Af12080 as the synthase for the tripeptidyl-S-enzyme intermediate that cyclizes to give FQF. A dedicated anthranilate synthetase (Af12110) is observed four ORFs downstream of the trimodular NRPS and presumably plays a role in generating enough of this planar aromatic  $\beta$ -amino acid to get the pathway started. Of the remaining predicted three enzymes in the cluster, Af12050, Af12060, and Af12070, we have already shown that the FAD-enzyme Af12060 and Ala-activating monomodular NRPS Af12050 act in tandem to achieve the oxidative annulation involved in the formation of the imidazoindolone tricycle as FQF is processed to FQA.<sup>7</sup>

The remaining ORF Af12070, a predicted FAD-enzyme, was then the likely candidate for introduction of the third stage of scaffold complexity as FQA is converted into the heptacyclic frameworks of FQC and FQD. After trimming off an N-terminal region of the protein, likely to be involved in localization/export in *A. fumigatus*, a soluble, active form of Af12070 was obtained from *E. coli* expression. The enzyme contains FAD covalently attached to His105 and functions as an FQA oxidase. Kinetic analysis and product structure determination reveal that FQA is converted to FQC and that subsequent accumulation of FQD occurs via a nonenzymatic equilibration.

Inspection of the FQC and FQD structures shows that in FQC a new spirohemiaminal linkage results in the formation of a seven-membered ring between the 3'-OH in FQA and the carbon that was originally  $C\alpha$  of the *L*-Ala moiety incorporated into the FQF quinazolinone scaffold (C3 in FQA). In comparison, the aminal linkage observed in FQD forms an eight-membered ring between the nitrogen of the second *L*-Ala (incorporated by Af12050 in the observed indole annulation reaction) and the same C3 carbon on the quinazolinone scaffold as seen in FQC. Both the hemiaminal and aminal linkages are formed by analogous addition reactions: the first between an alcohol and an imine and the second between an amine and the same imine. It is therefore not surprising that the formation of the hemiaminal linkage in FQC could be reversible, and the resultant imine would be available for capture by the competing intramolecular amine to yield FQD.

The time course of product appearance clearly indicates that the initial accumulation of FQC is kinetically favored while FQD is the thermodynamically favored product (Figure 7).

The formation of FQC and FQD, their spontaneous interconversion, and the identification of a hydroxylated FQA analogue (3-OH-FQA) are all consistent with the nascent product of Af12070 catalysis being a N2–C3 imine species. This, together with the fact that several FAD-containing enzymes are known to oxidize carbon–nitrogen bonds,<sup>29</sup> suggests that it is highly likely that Af12070 is a regioselective amide oxidase that acts to dehydrogenate across the N2–C3 bond of the quinazolinone moiety of FQA. Despite the relative abundance of FAD-dependent amine oxidases in biological systems,<sup>13</sup> to the best of our knowledge this is the first characterized example of an amide oxidase. It seems reasonable to envision that the chemistry of hydride abstraction from a carbon alpha to an amide, and thus oxidation, is significantly hindered by the absence of lone-pair stabilization of the resulting cation (as would be seen with an amine). It is therefore intriguing to speculate that in the case of FQA oxidation the extended delocalized system observed in the FQ scaffold (i.e., the quinazolinone which neighbors the oxidized pyrazinone) acts to stabilize the cation resulting from hydride abstraction, such that the FQ system is unusually amenable to the chemistry required for amide oxidation. Recent work describing evidence in favor of amine oxidases mediating transformation via single electron transfer (rather than direct hydride transfer) would also be consistent with the FQ scaffold being more susceptible to oxidation than a typical amide.<sup>30</sup> The presumed imine product resulting from amide oxidation does not accumulate (when FQA is the substrate), and instead the spirohemiaminal FQC is formed due to the intramolecular capture of the imine by the pendant –OH at C3' of the imidazoindolone.

The presumption is that Af12070 acts similar to a canonical FAD-dependent amine oxidase to generate an imine product and reduced flavin (via hydride transfer from substrate to N5 of FAD) and bound imine product in the active site (Figure 8).<sup>13</sup> However, an important distinction in the case of Af12070 is that the –NH group that is enzymatically oxidized is adjacent to a carbonyl within the pyrazinone ring of FQA and thus represents an unusual oxidation of an amide –NH. We have shown that the resulting dihydroflavin is then reoxidized by molecular oxygen in the second (oxidative) half reaction, thus regenerating active enzyme without the requirement for further cosubstrates/enzymes. The regeneration of FAD by O<sub>2</sub> is consistent with an oxidase mechanism yielding H<sub>2</sub>O<sub>2</sub> as a coproduct (as opposed to an N-hydroxylation/dehydration route to the imine which would yield two molecules of water and also would require an external reductant such as NADH).

The biological role and the potential relevance of the interconversion of the kinetically favored FQC and the thermodynamically favored FQD within *A. fumigatus* is yet to be determined. However, we speculate that the vast architectural changes observed in going from FQA to either FQC or FQD would have a significant impact on the interaction of these molecules within biological systems. The retention of Af12070 in the FQ biosynthetic gene cluster and biosynthetic energy exerted in the observed regioselective amide oxidation chemistry suggests that these complex molecules are more than simply architecturally intriguing natural products and must in fact have a tangible value to the

producing fungi within the complex ecological environment of nature.

Four analogous substrates were tested with Af12070 in order to assess substrate tolerance and gain insight into the effects of molecular structure on product formation. Three naturally occurring substrates were used: FQA (the biological substrate for Af12070) and two compounds identified during genetic characterization of the tryptoquialanine pathway from *P. aethiopicum*, 11'-dimethyl-2'-*epi*-FQA and 2'-*epi*-FQA. In addition, a fourth “unnatural” substrate, 11'-dimethyl-FQA, was generated using a genetically engineered monomodular NRPS. Taken together, these four substrates probe the effect of C2' stereochemistry and/or the effect of methyl vs dimethyl substitution at C11'. These experiments enable us to investigate how sites remote to the proposed site of Af12070 enzymatic action influence molecular events, subsequent to catalysis, that dictate the kinetics and thermodynamics of the apparently spontaneous cyclizations that mediate the end-product outcome.

In contrast to FQA, when the 2'-(*R*)-*epi*-FQA molecule is evaluated as an Af12070 substrate, it is the 2'-*epi*-FQD scaffold that forms with none of the 2'-*epi*-FQC hemiaminal product being detected. Thus, the stereochemistry of the annulated indole alters the kinetics of intramolecular capture of the nascent imine product such that, in the 2'-*epi*-scaffold, 2'-*epi*-FQD is both kinetically and thermodynamically favored. The activity of TqaG toward 2'-*epi*-FQA has been tested in vivo by deleting the genes in *P. aethiopicum* which encode TqaM or TqaL,<sup>8</sup> the enzymes responsible for AIB biosynthesis. With AIB synthesis shut down, the monomodular NRPS TqaB selects and couples L-Ala to the FQF framework to give 2'-*epi*-FQA.<sup>8,19</sup> Culture extracts prepared from *P. aethiopicum*  $\Delta$ tqaM or  $\Delta$ tqaL show that 2'-*epi*-FQA is a substrate for TqaG and that 2'-*epi*-FQD is detected along with nortryptoquialanine, the demethyl analogue of the final natural product (Figure S10).<sup>8</sup> While the appearance of 2'-*epi*-FQD as a shunt metabolite clearly illustrates that the kinetics of the processing downstream of TqaG (to give tryptoquialanine) is significantly perturbed by the lack of the 11'-dimethyl group, it also provides clear evidence that Af12070 and TqaG are functionally equivalent with 2'-*epi*-FQA as substrate.

When 11'-dimethyl-FQA was incubated in vitro with Af12070, only 11'-dimethyl-FQC is observed in the reaction mixture even after extended periods of time. Therefore, it appears that the presence of an 11'-dimethyl substitution on FQA produces the opposite result to that observed with 2'-*epi*-FQA. In this case 11'-dimethyl-FQC is both the kinetic and the thermodynamic adducts resulting from intramolecular capture of the nascent imine.

11'-Dimethyl-2'-*epi*-FQA is the proposed natural substrate for TqaG, with its production in nature being more abundant than the analogue 2'-*epi*-FQA due to the strong preference for activation of AIB over L-Ala by TqaB in WT *P. aethiopicum* (Figure S1). When 11'-dimethyl-2'-*epi*-FQA is incubated with Af12070 in vitro, both the hydroxyl- and amine-linked products are detected (11'-dimethyl-2'-*epi*-FQC and 11'-dimethyl-2'-*epi*-FQD), as is the hydrate (consistent with the competitive capture of the nascent imine product by intermolecular addition of water) (Figures S6 and S7). Intriguingly, we can also detect products consistent with the presence of the proposed imine intermediate and that of the keto acid which would be predicted to result from the hydrolysis of this imine (Figures S6, S8, and S9). Overall, we have observed that

2'-*epi*-(*R*) stereochemistry appears to favor FQD formation and that 11'-dimethyl substitution appears to favor FQC formation. It is therefore intriguing to speculate that it is the combination of these competing effects that prolongs the lifetime of the nascent imine product of Af12070 (and TqaG) such that it can be observed when 11'-dimethyl-2'-*epi*-FQA is used as substrate for Af12070 catalysis. The hydrolysis of the imine to yield the keto acid could be direct (i.e., the stepwise hydrolysis of the carbonyl–nitrogen bond of the cyclic imine followed by hydrolysis of the resulting acyclic imine<sup>8</sup>) or indirect following the intramolecular capture of the imine as a hemiaminal/aminal intermediate (Figure S9). If hydrolysis follows intramolecular capture, ring strain could promote spontaneous hydrolysis of the pyrazinone amide, a step which would be accelerated by the trypsin-like serine protease TqaI in nature. In either case, the resultant keto acid could undergo lactonization to give deoxytryptoquialanone en route to the formation of tryptoquialanine in *P. aethiopicum* (Figure S9). The mechanism of this hydrolytic removal of the nitrogen from the tricyclic quinazoline ring is not yet understood. However, comparative analysis of the products resulting from Af12070 and TqaG catalysis clearly suggest that molecular architecture is a critical determinant in the reactivity of the nascent imine and consequently fumiquinazoline vs tryptoquialanine pathway differentiation (see Supporting Information Discussion for details).

The studies described here on the initial characterization of Af12070 raise the question of what are the biological functions of each/any of the fumiquinazolines that make them constitutive metabolites for all the *A. fumigatus* strains so far examined. The complete biochemical characterization of the fumiquinazoline biosynthetic enzymes now potentially opens the door to genetic knockout experiments in *A. fumigatus* in order to search for a phenotypic response to FQF, FQA, or FQC/FQD production (such as reduced virulence or altered patterns of development and/or sporulation) by knocking out *af12080*, *af12060* or *af12050*, and *af12070*, respectively.

From the point of view of biosynthesis of highly morphed and constrained peptide-based architectural scaffolds, the fumiquinazoline biosynthetic pathway comprises an efficient and remarkably impressive set of catalytic machinery. In a short metabolic pathway, with only four enzymes, two NRPS to incorporate four amino acid building blocks and two FAD enzymes to carry out key oxidations, a remarkable pair of heptacyclic frameworks with five nitrogen atoms are constructed.

## ■ ASSOCIATED CONTENT

### ● Supporting Information

Supplementary discussion of fumiquinazoline vs tryptoquialanine pathway differentiation, figures, and tables of NMR data. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [christopher\\_walsh@hms.harvard.edu](mailto:christopher_walsh@hms.harvard.edu). Phone: 617-432-1715. Fax: 617-432-0483.

### Author Contributions

<sup>†</sup>Equally contributing authors.

## Funding

This work is supported in part by National Institutes of Health Grant GM20011 (to C.T.W.), F32GM090475 (to B.D.A.), IR01GM092217 (to Y.T.), and R01GM067725 (to N.L.K.).

## ■ ACKNOWLEDGMENTS

We thank Steven Malcolmson for helpful discussions and Thomas Gerken for providing *A. fumigatus* Af293 cDNA.

## ■ ABBREVIATIONS

2-AIB, 2-aminoisobutyric acid; AMP, adenosine-5'-monophosphate; Ant, anthranilate (2-aminobenzoate); ATP, adenosine-5'-triphosphate; EDTA, ethylenediaminetetraacetic acid; *epi*, epimer; FAD, flavin adenine dinucleotide; FMN, flavin mononucleotide; FT, Fourier transform; FQ, fumiquinazoline; HPLC, high-performance liquid chromatography; HRLC-MS, high-resolution liquid chromatography–mass spectrometry; IPTG, isopropyl- $\beta$ -D-galactopyranoside; LB, Luria–Bertani medium; LIC, ligation-independent cloning; MeCN, acetonitrile; MWCO, molecular weight cutoff; NADH, nicotinamide adenine dinucleotide; Ni-NTA, nickel nitrilotriacetic acid–agarose; NMR, nuclear magnetic resonance; NRPS, non-ribosomal peptide synthetase; PCR, polymerase chain reaction; PDB ID, Protein Data Bank identifier; PTFE, polytetrafluoroethylene; QTOF, quadrupole time-of-flight; SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid; Tris, tris(hydroxymethyl)aminomethane.

## ■ ADDITIONAL NOTE

<sup>a</sup>A discrepancy exists between the database entry for AFUA\_6g12070 and the coding sequence obtained from cDNA cloning due to the predicted/annotated vs experimentally determined boundaries for the first intron. Residue numbering reflects the experimentally determined sequence (see Experimental Procedures for details).

## ■ REFERENCES

- (1) Frisvad, J. C., Rank, C., Nielsen, K. F., and Larsen, T. O. (2009) Metabolomics of *Aspergillus fumigatus*. *Med. Mycol.* 47, 53–71.
- (2) Rementeria, A., et al. (2005) Genes and molecules involved in *Aspergillus fumigatus* virulence. *Rev. Iberoam. Micol.* 22, 1–23.
- (3) Numata, A., et al. (1992) Fumiquinazolines, novel metabolites of a fungus isolated from a saltfish. *Tetrahedron Lett.* 33, 1621–1624.
- (4) Takahashi, C., et al. (1995) Fumiquinazolines A–G, novel metabolites of a fungus separated from a *Pseudolabrus* marine fish. *J. Chem. Soc., Perkin Trans. 1*, 2345–2353.
- (5) Nierman, W.C., et al. (2005) Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature* 438, 1151–1156.
- (6) Ames, B. D., and Walsh, C. T. (2010) Anthranilate-activating modules from fungal nonribosomal peptide assembly lines. *Biochemistry* 49, 3351–3365.
- (7) Ames, B. D., Liu, X., and Walsh, C. T. (2010) Enzymatic processing of fumiquinazoline F: A tandem oxidative-acylation strategy for the generation of multicyclic scaffolds in fungal indole alkaloid biosynthesis. *Biochemistry* 49, 8564–8576.
- (8) Gao, X., et al. (2011) Fungal indole alkaloid biosynthesis: Genetic and biochemical investigation of the tryptoquialanine pathway in *Penicillium aethiopicum*. *J. Am. Chem. Soc.* 133, 2729–2741.
- (9) Fremlin, L. J., Piggott, A. M., Lacey, E., and Capon, R. J. (2009) Cottoquinazoline A and cotteslosins A and B, metabolites from an Australian marine-derived strain of *Aspergillus versicolor*. *J. Nat. Prod.* 72, 666–670.

- (10) Zhuang, Y., et al. (2011) New quinazolinone alkaloids within rare amino acid residue from coral-associated fungus, *Aspergillus versicolor* LCJ-5-4. *Org. Lett.* 13, 1130–1133.
- (11) Gilbert, N. B., Montserrat, A., Paul, R. J., William, F., and Matthias, K. (2000) Oxepinamides A-C and fumiquinazolines H-I: bioactive metabolites from a marine isolate of a fungus of the genus *Acremonium*. *Chem.—Eur. J.* 6, 1355–1360.
- (12) Twumasi-Boateng, K., et al. (2009) Transcriptional profiling identifies a role for BrlA in the response to nitrogen depletion and for StuA in the regulation of secondary metabolite clusters in *Aspergillus fumigatus*. *Eukaryot. Cell* 8, 104–115.
- (13) Fitzpatrick, P. F. (2009) Oxidation of amines by flavoproteins. *Arch. Biochem. Biophys.* 493, 13–25.
- (14) Wilkins, M. R., et al. (1999) Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* 112, 531–552.
- (15) Larkin, M. A. et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948.
- (16) Nicholas, K. B., Nicholas, H. B., and Deerfield, D. W. (1997) GeneDoc: analysis and visualization of genetic variation. *EMBNEW News* 4, 14.
- (17) Soding, J., Biegert, A., and Lupas, A. N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 33, W244–248.
- (18) Deshpande, N., Wilkins, M. R., Packer, N., and Nevalainen, H. (2008) Protein glycosylation pathways in filamentous fungi. *Glycobiology* 18, 626–637.
- (19) Haynes, S. W., Ames, B. D., Gao, X., Tang, Y., and Walsh, C. T. (2011) Unraveling terminal C-domain-mediated condensation in fungal biosynthesis of imidazoindolone metabolites. *Biochemistry* 50, 5668–5679.
- (20) Pall, M., and Brunelli, J. (1993) A series of six compact fungal transformation vectors containing polylinkers with multiple unique restriction sites. *Fungal Genet. Newsl.* 40, 59.
- (21) Szewczyk, E., et al. (2007) Fusion PCR and gene targeting in *Aspergillus nidulans*. *Nature Protoc.* 1, 3111–3120.
- (22) Leferink, N. G. H., Heuts, D. P. H. M., Fraaije, M. W., and van Berkel, W. J. H. (2008) The growing VAO flavoprotein family. *Arch. Biochem. Biophys.* 474, 292–301.
- (23) Heuts, D. P. H. M., Scrutton, N. S., McIntire, W. S., and Fraaije, M. W. (2009) What's in a covalent bond? On the role and formation of covalently bound flavin cofactors. *FEBS J.* 276, 3405–3427.
- (24) Sirikantaramas, S., et al. (2005) Tetrahydrocannabinolic acid synthase, the enzyme controlling marijuana psychoactivity, is secreted into the storage cavity of the glandular trichomes. *Plant Cell Physiol.* 46, 1578–1582.
- (25) Bird, D., and Facchini, P. (2001) Berberine bridge enzyme, a key branch-point enzyme in benzylisoquinoline alkaloid biosynthesis, contains a vacuolar sorting determinant. *Planta* 213, 888–897.
- (26) Šmehilová, M., et al. (2009) Subcellular localization and biochemical comparison of cytosolic and secreted cytokinin dehydrogenase enzymes from maize. *J. Exp. Bot.* 60, 2701–2712.
- (27) Varela, E., Jesús Martínez, M., and Martínez, A. T. (2000) Aryl-alcohol oxidase protein sequence: a comparison with glucose oxidase and other FAD oxidoreductases. *Biochim. Biophys. Acta* 1481, 202–208.
- (28) Mattevi, A. (2006) To be or not to be an oxidase: challenging the oxygen reactivity of flavoenzymes. *Trends Biochem. Sci.* 31, 276–283.
- (29) Fagan, R. L., Palfey, B. A., Lew, M., and Hung-Wen, L. Flavin-Dependent Enzymes, in *Comprehensive Natural Products II*, pp 37–113, Elsevier, Oxford.
- (30) MacMillar, S., Edmondson, D. E., Matsson, O. Nitrogen kinetic isotope effects for the monoamine oxidase B-catalyzed oxidation of benzylamine and (1,1-<sup>2</sup>H<sub>2</sub>)benzylamine: nitrogen rehybridization and CH bond cleavage are not concerted, *J. Am. Chem. Soc.* 133, 12319–12321.