Equisetin biosynthesis in *Fusarium heterosporum*†

James W. Sims, John P. Fillmore, Douglas D. Warner and Eric W. Schmidt*

Received (in Cambridge, UK) 2nd September 2004, Accepted 24th September 2004
First published as an Advance Article on the web 22nd November 2004
DOI: 10.1039/b413523g

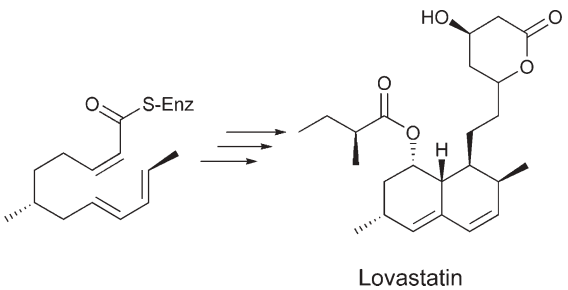
An iterative polyketide synthase–peptide synthetase hybrid assembles the HIV-1 integrase inhibitory tetramic acid, equisetin, in the filamentous fungus *Fusarium heterosporum*.

Fungal polyketides often possess unique carbon skeletons, but the biosynthetic genes for the vast majority of these compounds have yet to be elucidated. We examined the biosynthesis of equisetin,[1,2] a polyketide produced by the filamentous fungus *F. heterosporum*, because of its bioactivity[3,4] and its structural features that relate it to many other fungal compounds. Related polyketides are toxins,[5] biological probes,[6] pharmaceutical candidates,[7] and marketed drugs[8] (Fig. 1). Equisetin inhibits HIV-1 integrase[3] and is also broadly toxic.[4] Key structural features of equisetin include a polyketide-derived decalin and an amino acid-derived heterocycle known as a tetramic acid (2,4-pyrrolidinedione). Tetramic acid metabolites, isolated from fungi, bacteria, and marine sponges, are being developed for pharmaceutical and industrial use, but prior to our study no biosynthetic genes had been reported.[9] The tetramic acid moiety is biosynthetically similar not only to other tetramic acid natural products, but also to a broad array of more distantly related heterocycles. From our work, we hoped to gain insight into the relationships between these pathways that would enable rational engineering of tetramic acids, statins, and chimeric structures.

Equisetin also bears a close similarity to lovastatin, a fungal anti-hypercholesteremic agent that is a top-selling drug and the basis for the synthetic statins (Scheme 1).[8] The lovastatin carbocycle is synthesized by an iterative PKS, the lovastatin nonaketide synthase (LNKS), and an accessory enoylreductase (ER), LovC.[10] LNKS and LovC interact to produce a partially reduced polyketide, and LNKS then appears to catalyze a Diels–Alder cycloaddition. It is still not known how LNKS and LovC control partial reduction and Diels–Alder activities.

Because of the structural homology between equisetin and lovastatin, we designed degenerate PCR primers to amplify LNKS-like polyketide synthase (PKS) sequences from *F. heterosporum*. Using this method, three PKSs sharing homology with the lovastatin synthase were identified.

When grown on corn grit agar (CGA), *F. heterosporum* produced >2 g L^{-1} of equisetin, whereas the compound was not synthesized (<10 ng L^{-1}) on potato dextrose agar (PDA). To determine which of the three PKS sequences could be involved in equisetin biosynthesis, mRNA was extracted from PDA- and CGA-grown *F. heterosporum*. Using RT-PCR, a single PKS sequence, designated *pks2*, was correlated with equisetin production, while the other PKS genes were not expressed in tandem with equisetin biosynthesis. Thus, *pks2* was the most likely candidate for the equisetin PKS gene.



Scheme 1 Diels–Alder route to lovastatin.

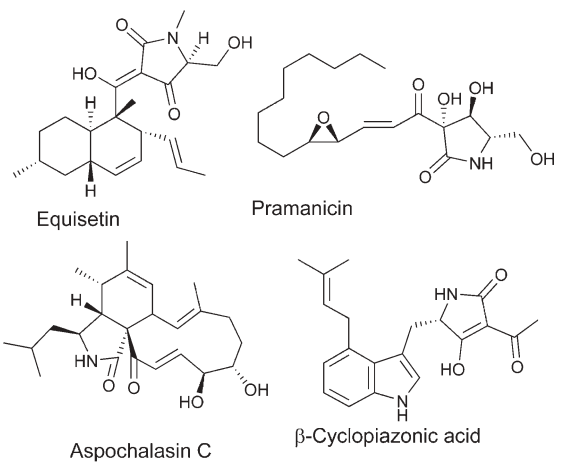


Fig. 1 Fungal tetramic acids and structural relatives.

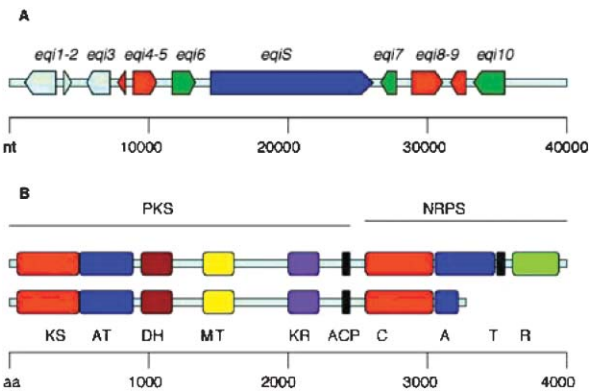
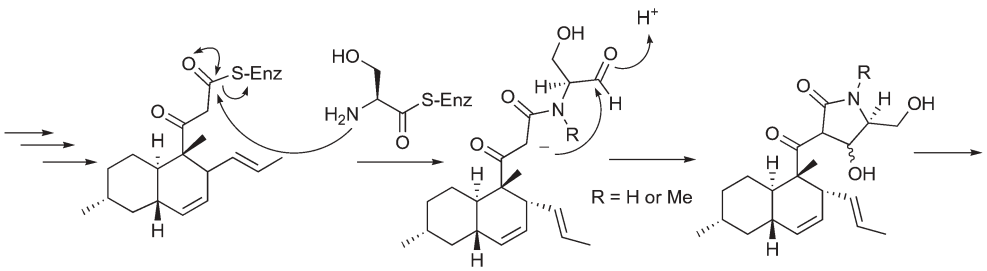


Fig. 2 A. *eqi* cluster. *eqiS*: dark blue; resistance and regulation: green; tailoring: red. B. EqiS (top) compared to LNKS (bottom).

† Electronic supplementary information (ESI) available: experimental procedures. See http://www.rsc.org/suppdata/cc/b4/b413523g/

*ews1@utah.edu



Scheme 2 Proposed biogenesis of equisetin. The structure on the right represents the probable product of EqiS and must be further oxidized to equisetin.

To confirm the role of *pks2*, we employed a knockout mutagenesis technique using a *pks2* fragment flanking a hygromycin resistance cassette. Out of >50 transformants, one was identified by PCR and Southern hybridization to be a specific integrant. The specific integrant, wild-type, and 10 nonspecific integrants were monitored for equisetin production using a bioassay (10 ng L^{-1} detection limit) and TLC techniques. While nonspecific integrants and wild-type synthesized equisetin at wild-type levels, synthesis was completely abolished in the specific integrant. The only hypothesis consistent with the data is that *pks2* plays a role in equisetin biosynthesis. This hypothesis is further supported by RT-PCR and the sequence analysis.

From a cosmid library in vector pMOcosX,[11] 4 *pks2*-containing clones were identified. 39140 bp of the *eqi* cluster were sequenced (GenBank AY700570), revealing 11 open reading frames (*eqi1-10* and *eqiS*), of which up to eight may be involved in equisetin biosynthesis (Fig. 2). *eqiS* is a PKS-nonribosomal peptide synthetase (NRPS) hybrid that is predicted to synthesize most of the molecule, while the remaining genes are likely involved in tailoring, resistance and regulation.

The predicted EqiS is a 3953 amino acid hybrid of an iterative PKS[12] with a modular NRPS[13] and shares similarity with enzymes from fungal genome sequences. The domain order in the EqiS PKS exactly mimics that of LNKS and related enzymes, and the NRPS has similar domain order to several synthetases from fungal genome sequences. All of the domains required for synthesis of the polyketide chain are present in the PKS portion, including β-ketoacyl synthase (KS), acyltransferase (AT), dehydratase (DH), methyltransferase (MT), ketoreductase (KR), and acyl carrier protein (ACP), as found in LNKS. The C-terminus ($\sim$1500 amino acids) of EqiS comprises a complete NRPS module, containing condensation (C), adenylation (A), thiolation (T), and reductive (R) domains. A portion of this NRPS resembles LNKS, which terminates with a C domain and the first several hundred amino acids of an A domain. In LNKS, the NRPS domains are inactive, since the catalytic region[14] of the A domain has been deleted, although the C domain may have a role in lactone formation. The similarity of LNKS and the EqiS N-terminus suggests that LNKS is a degraded tetramic acid biosynthetic gene that has undergone a C-terminal deletion event in the course of pathway evolution. EqiS thus contains all of the domains necessary to produce the completed polyketide structure covalently linked, *via* an amide, to an amino acid (Scheme 2).

The remaining steps, release of the equisetin intermediate and cyclization to the pyrrolidinone, are likely catalyzed by the C-terminal R domain. These domains terminate many NRPS genes, and they are proposed to cleave the covalent protein–small molecule thioester bonds to release aldehydes,[15] which then undergo a variety of metabolic fates.[16] We propose that the R releases an equisetin intermediate as the C-terminal aldehyde, which would then spontaneously cyclize in the cytoplasm. Oxidation would yield the tetramic acid. This represents a great expansion of the products of R domains.

The *eqi* cluster also contains a number of genes involved in resistance, regulation, and oxidation (presumably of the reduced tetramic acid motif). The sequence of the *eqi* genes enables us to predict the biosynthetic pathway to equisetin and related molecules. For example, aspochalasins[6] and pramanicin[7] are likely synthesized by pathways in which the products of EqiS-like proteins are modified by a Diels–Alder reaction or P450 oxidation, respectively.

Phylogenetic analysis of EqiS, in addition to previously reported analyses,[17] indicates that fungal PKS–NRPS hybrids, producing vastly different structures, have diversified by point mutation and deletion, rather than by module swaps within fungi. This contrasts strongly with the strategy of bacteria, in which module and domain swaps and duplications are exploited to generate new carbon skeletons.[18] This model for fungal PKS–NRPS evolution has important consequences for the interrelationship of additional fungal biosynthetic pathways. For example, the cyclopiazonic acids[19] are tetramic acids derived from acetoacetate and tryptophan (Fig. 1). The model implies that the cyclopiazonic acids should have PKS and NRPS portions that are more related to EqiS than to PKSs that make diketides, or to NRPSs that activate Trp. Cyclopiazonic acid is further modified by addition of dimethylallylpyrophosphate (DMAPP) to Trp. Interestingly, the *eqi* cluster contains a DMAPP–Trp transferase, which we do not anticipate to be involved in equisetin biosynthesis but that is homologous to that in the ergot alkaloid pathway.[20]

While this manuscript was under review, a related pathway was reported for fusarin C biosynthesis.[21]

This research was funded by the Herman Frasch Foundation (grant # 514-HF02) and U. Utah. J. P. F. was partially supported by the U. Utah Biosciences Undergraduate Research Program. Fungal vectors were from the FGSC, University of Kansas.

James W. Sims, John P. Fillmore, Douglas D. Warner and Eric W. Schmidt*

Department of Medicinal Chemistry, University of Utah, 30 South 2000 East Rm 201, Salt Lake City, UT, 84112, USA. E-mail: ews1@utah.edu; Fax: 801-585-9119; Tel: 801-534-2348

Notes and references

1 H. R. Burmeister, G. A. Bennett, R. F. Vesonder and C. W. Hesseltine, *Antimicrob. Agents Chemother.*, 1974, **5**, 634–639.

2 N. J. Phillips, J. T. Goodwin, A. Fraiman, R. J. Cole and D. G. Lynn, *J. Am. Chem. Soc.*, 1989, **111**, 8223–8231.
3 S. B. Singh, D. L. Zink, M. A. Goetz, A. W. Dombrowski, J. D. Polishook and D. J. Hazuda, *Tetrahedron Lett.*, 1998, **39**, 2243–2246.
4 R. F. Vesonder, L. W. Tjarks, W. K. Rohwedder, H. R. Burmeister and J. A. Laugal, *J. Antibiot. (Tokyo)*, 1979, **32**, 759–761.
5 R. A. Meronuck, J. A. Steele, C. J. Mirocha and C. M. Christensen, *Appl. Microbiol.*, 1972, **23**, 613–617.
6 G. X. Zhou, E. M. K. Wijeratne, D. Bigelow, L. S. Pierson, H. D. VanEtten and A. A. L. Gunatilaka, *J. Nat. Prod.*, 2004, **67**, 328–332.
7 P. H. M. Harrison, D. W. Hughes and R. W. Riddoch, *Chem. Commun.*, 1998, 273–274.
8 A. Endo, *J. Antibiot. (Tokyo)*, 1979, **32**, 852–854.
9 B. J. L. Royles, *Chem. Rev.*, 1995, **95**, 1981–2001.
10 K. Auclair, A. Sutherland, J. Kennedy, D. J. Witter, J. P. Van den Heever, C. R. Hutchinson and J. C. Vederas, *J. Am. Chem. Soc.*, 2000, **122**, 11519–11520.
11 M. J. Orbach, *Gene*, 1994, **150**, 159–162.
12 L. Hendrickson, C. R. Davis, C. Roach, D. K. Nguyen, T. Aldrich, P. C. McAda and C. D. Reeves, *Chem. Biol.*, 1999, **6**, 429–439.
13 A. Haese, M. Schubert, M. Herrmann and R. Zocher, *Mol. Microbiol.*, 1993, **7**, 905–914.
14 E. Conti, T. Stachelhaus, M. A. Marahiel and P. Brick, *EMBO J.*, 1997, **16**, 4174–4183.
15 T. M. Zabriskie and M. D. Jackson, *Nat. Prod. Rep.*, 2000, **17**, 85–97.
16 Summarized in: J. E. Becker, R. E. Moore and B. S. Moore, *Gene*, 2004, **325**, 35–42.
17 S. Kroken, N. L. Glass, J. W. Taylor, O. C. Yoder and B. G. Turgeon, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 15670–15675.
18 J. H. Crosa and C. T. Walsh, *Microbiol. Mol. Biol.*, 2002, **66**, 223–249.
19 C. W. Holzapfel, *Tetrahedron*, 1968, **24**, 2101–2119.
20 H. F. Tsai, H. Wang, J. C. Gebler, C. D. Poulter and C. L. Schardl, *Biochem. Biophys. Res. Commun.*, 1995, **216**, 119–125.
21 Z. Song, R. J. Cox, C. M. Lazarus and T. J. Simpson, *Chembiochem*, 2004, **5**, 1196–1203.

