ENGINEERING COMBINATORIAL BIOSYNTHESIS OF POLYKETIDE SYNTHASE–PEPTIDE SYNTHETASE PRODUCTS USING NOVEL FUNGAL GENETIC TOOLS

by

Thomas Butanaziba Kakule

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Medicinal Chemistry

The University of Utah

May 2015

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Thomas Butanaziba Kakule**

has been approved by the following supervisory committee members:

Eric W. Schmidt	, Chair	**11/20/2014** Date Approved
Amy M. Barrios	, Member	**11/20/2014** Date Approved
Grzegorz Bulaj	, Member	**11/20/2014** Date Approved
Darrell Davis	, Member	**11/20/2014** Date Approved
Dennis Winge	, Member	**11/21/2014** Date Approved

and by **Darrell Davis** , Chair/Dean of

the Department of **Medicinal Chemistry**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Fungal polyketides are a complex class of natural products with diverse scaffolds. The biosynthesis of these molecules involves the iterative condensation of acetate units using a minimal set of domains on a single polypeptide. The different levels of reduction at each iterative step generates the structural diversity observed via the stuttering action of these domains. For some polyketides, further structural complexity is introduced via amidation by a nonribosomal peptide synthetase (NRPS) module fused to the polyketide synthase (PKS). The NRPS contains a condensation domain (C) that catalyzes the coupling of the polyketide to a specific amino acid. Studies by others have suggested that PKS function is independent of NRPS activity and thus, this system is amenable to combinatorial biosynthesis to generate analogues with different polyketide chains and amino acids by performing module swaps. Forging intermodular interactions and understanding C domain selectivity for substrates is key to successful engineering of these analogues. Herein, I investigate the impact of these factors on the ability to synthesize unnatural products by this route. Studying these components required the overexpression of chimeric gene constructs, which was anticipated to result in low compound yields. Therefore, a novel platform to express fungal genes with yields exceeding 1 g per kg media was developed and validated through the characterization of a silent pathway. Application to the PKS-NRPS combinatorial biosynthesis problem led

to the discovery that C domains are highly selective for closely related substrates in the presence of favorable PKS/NRPS interactions.

To Mom

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

A domain	adenylation domain
ACN	acetonitrile
ACP	acyl carrier protein
AT	acyl transferase
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
C domain	condensation domain
COSY	Correlation Spectroscopy
DAD	diode array detector
ER	enoyl reductase
FPLC	fast protein liquid chromatography
FT-ICR	fourier-transform ion cyclotron resonance
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMBC	Heteronuclear Multiple Bond Correlation spectroscopy
HPLC	high-pressure liquid chromatography
HSQC	Heteronuclear Single Quantum Spectroscopy
IPTG	isopropyl-β-D-1-thiogalactopyranoside

KR	ketoreductase
KS	ketosynthase
LB	Luria–Bertani
LC	liquid chromatography
MS	mass spectrometry
NADPH	nicotinamide adenine dinucleotide phosphate
NMR	nuclear magnetic resonance
NOESY	Nuclear Overhauser Effect Spectroscopy
NRPS	nonribosomal peptide synthetase
PCR	polymerase chain reaction
PKS	polyketide synthase
R domain	reductase domain
ROESY	Rotating-frame Overhauser Effect Spectroscopy
RT	room temperature
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TFA	trifluoroacetic acid
TLC	thin layer chromatography
TOCSY	Total Correlation Spectroscopy

CHAPTER 1

INTRODUCTION

1.1 Fungal Natural Products

The contribution of natural products to the survival of the producing organisms is not clear in most cases, hence their alternate name of secondary metabolites. In some instances, however, they have been found to improve the fitness of the producing organisms within their environments by warding off predators and reducing competitors for nutrients.[1–3] Such biological activities among others have made them invaluable as a potential source of medicines. To date, several of these natural products make up a large percentage of all drugs in use,[4] and fungal natural products have made major contributions to this drug compendium over the years. Popular examples of such drugs include cyclosporine approved for autoimmune suppression for organ transplant patients in 1983; lovastatin for cholesterol management in 1987; and more recently fingolimod, a fungal natural product-inspired drug, for the management of multiple sclerosis. Considering that the currently identified fungal biodiversity is estimated at a meagre 5% and a biologically active molecule is produced by almost every isolated fungal species, fungi remain an underexplored resource for drug discovery.[5] This reality has been brought into greater focus with the advent of the genomics era. Scanning individual fungal genomes for biosynthetic genes has revealed that there are more pathways to natural products than the number of these molecules isolated from these organisms, suggesting that many of these genes have gone silent or are otherwise cryptic under laboratory culture conditions.[5–7] Therefore, vast potential still exists for drug discovery from fungi isolated as far back as 50 years ago. By characterizing newly discovered fungal species and studying silent biosynthetic pathways, it will be possible to tap into the true scientific and medicinal potential of natural products research.

1.2 Cryptic Biosynthetic Pathways

Current understanding of fungal natural product biosynthesis is that all genes involved in the synthesis of a product will be situated at the same locus within the genome, a phenomenon otherwise known as clustering.[6,8–10] Thus, the presence of more biosynthetic gene clusters within a genome than the total number of isolated natural products alludes to the presence of cryptic (or silent) biosynthetic pathways.[7,11] In order to access molecules encoded by these pathways, several strategies have been developed over the years to turn on gene expression. One approach has been to change culture conditions.[7] By altering the nitrogen and carbon sources and mineral composition in growth media, some fungi have produced new metabolites. Another strategy relies on recapitulating the interactions that go on in the natural environment by coculturing the fungi with various microorganisms.[1,12] With the discovery that gene expression control in fungi is also governed by histone methylation and acetylation, epigenetic-modifying small molecules have been added to culture media as another strategy to induce expression of cryptic pathways.[7] Because the aforementioned methods are stochastic processes, they suffer from low success rates and have fallen out of favor since the development of fungal genetic techniques. The appeal of these genetic methods stems from the fact that they allow direct manipulation of cryptic genes. Using common gene cloning techniques, the context of regulation of these genes can be altered, which is seen in the controllable overexpression of target genes.[13] The most common genetic approach is the upregulation of transcription factors found within cryptic gene clusters to modulate expression of coclustered genes to produce new secondary metabolites.[14,15] For instance, aspyridone A from *Aspergillus nidulans* was discovered by increasing the expression of a transcription factor within a cryptic gene cluster with the aid

of a housekeeping gene promoter.[15] Transcription factor overexpression is not a universal solution though; there are many examples where it has not resulted in new metabolite production, and this approach is also not amenable to fungi for which gene transformation methods have not been developed. In addition, general gene transformation techniques for fungi do not exist; such a protocol must be developed for each fungus under study.[7] These shortcomings have led to the development of heterologous expression systems where the cryptic genes are transferred from the investigated fungus to a new host.[16] Gene expression in this host then results in production of the desired metabolite. This approach has the added advantage of allowing assignment of roles to individual genes within the cluster towards synthesis of the final product.

1.3 Heterologous Gene Expression

Typically, the gene of interest is cloned into an expression vector and the resulting plasmid transformed into the host.[16] For expression, the exogenous DNA is either maintained as a plasmid or linearized and integrated into the host genome. Several advances in gene cloning technologies such as restriction enzyme ligation and yeast recombination have made the construction of such expression plasmids a trivial exercise.[17] Even with these advances, successful expressions depend on the suitability of the host. To date, there is a number of heterologous expression hosts that have been developed for secondary metabolite expression.[18–22] First among them is *Escherichia coli*, which has been used extensively to study polyketide biosynthesis.[22–24] While *E. coli* provides a quick expression system because of the short doubling time, it is only suitable for pathways lacking introns. Similarly, the robust expression platform based on *Saccharomyces*

cerevisiae has the advantage of dealing well with eukaryotic codon usage,[25] but falls short by the same measure. Workarounds involve tediously removing the introns by fusion PCR, or extracting the coding sequences from prepared cDNA. The latter method has limited applicability to cryptic genes and unculturable organisms. To address the intron issue, the fungus *Aspergillus oryzae* has been used with great success to study several fungal natural product pathways by direct cloning from genomic DNA.[26] However, the reported compound yields from the *A. oryzae* expression system are low. In this dissertation, we built a high-yielding expression platform to study recombinant genes from various fungi. Unlike previously published systems that use promoters of housekeeping genes, our system is based on using regulatory elements of the biosynthetic pathway to equisetin, which is produced in large amounts (~2 g L^{-1}) by the filamentous fungus *Fusarium heterosporum*. The identification and characterization of the equisetin biosynthetic pathway is described in Chapter 2. How the identified regulatory elements were developed into novel heterologous expression tools is detailed in Chapter 3. The application of these tools to resurrect an extinct natural product from a cryptic biosynthetic pathway is described in Chapters 3 and 4. Furthermore, in Chapter 5, we leveraged this new platform to express recombinant gene fusions designed to explore the combinatorial biosynthesis of a class of polyketides. A brief introduction to polyketide biosynthesis, relevant to work described in this dissertation, is provided below.

1.4 Polyketide Biosynthesis

The polyketide class of natural products is a structurally diverse group of molecules.[27] Even though they arise from decarboxylative condensation of the same simple

building block, malonyl CoA, they exhibit a large range of structure scaffolds from linear aliphatic chains to more decorated quinone-type structures.[28] The core domains of polyketide synthase enzymes include the acyl transferase domain (AT) involved in selecting the extending unit; the ketosynthase domain (KS), which catalyzes the Claisen condensation of the malonyl unit with the growing polymer tethered to the acyl carrier protein domain (ACP), another core domain. Most fungal polyketides are made by synthases that utilize this minimal domain set iteratively to make the backbone polyketide chains. In addition to the aforementioned domains, the fungal PKS may have other domains that introduce greater complexity such as α-C-methylations by a methyl transferase domain; various levels of reduction through the use of a ketoreductase (KR) domain that converts the β-keto group to an alcohol, a dehydratase domain the reduces the ensuing alcohol to an olefin and finally an enoyl reductase (ER) to fully reduce the double bond to saturated aliphatic group. The stuttering activity of these modifying domains during the iterative chain extension process generates the complexity observed for these fungal polyketide scaffolds. In nature, polyketides are often further conjugated with amino acids, sometimes forming tetramic acid rings. These molecules have numerous biological activities including antibacterial, antiviral, and anticancer properties.[29] To date, none of these molecules has been further developed into a drug lead largely due to toxicity.

1.5 Polyketide Synthase–Nonribosomal Peptide Synthetase (PKS-NRPS) Hybrids

Tetramic acids are produced biosynthetically by PKS-NRPS hybrid enzymes.[27,30,31] A PKS-NRPS hybrid consists of the PKS module, having the basic domain architecture

described above, with a nonribosomal peptide synthetase (NRPS) module fused at the C-terminus (exemplified in Figure 1.1).[31] The NRPS module consists of a condensation domain (C), an adenylation domain (A), a thiolation (T), and a reductive domain (R). The A domain selects a specific amino acid and loads it onto the T domain. The C domain then catalyzes the coupling of the polyketide product tethered to the ACP domain of the PKS module to the amino acid on the T domain. The final product is then released by a Dieckmann cyclization catalyzed by the R domain to form the tetramic acid ring. Within this subgroup of fungal polyketides, various amino acid conjugates have been characterized.[30,32] These have been found to contain homoserine, serine, tryptophan, phenylalanine, tyrosine, and alanine, among others (Figure 1.2). This PKS-NRPS hybrid system presents a model where combinatorial rearrangements should be possible to yield analogues of the natural tetramic acids, containing different polyketide chains and amino acid components. However, not much is known about how these modules interact with each other. Recently, two such studies reported hybrid products arising from pairing up a PKS module with a noncognate NRPS module.[33,34] In the first study, the aspyridone synthase (ApdA) PKS was expressed alongside the cyclopiazonic synthase (CpaS) NRPS.[34] Natively, the CpaS NRPS accepts a diketide intermediate; however, the interaction with ApdA PKS synthesized a tetrakide-derived tetramic acid. A similar study by the Cox group in which the tenellin synthase (TenS) PKS module was swapped with DmbS yielded the predicted hexaketide-derived tetramate instead of the pentaketide chain.[33] These results showed that the PKS could act independent of the NRPS component; only after formation of the full-length PKS product did transfer occur. However, work done by the Hertweck group in which the PKS module of the lovastatin nonaketide synthase

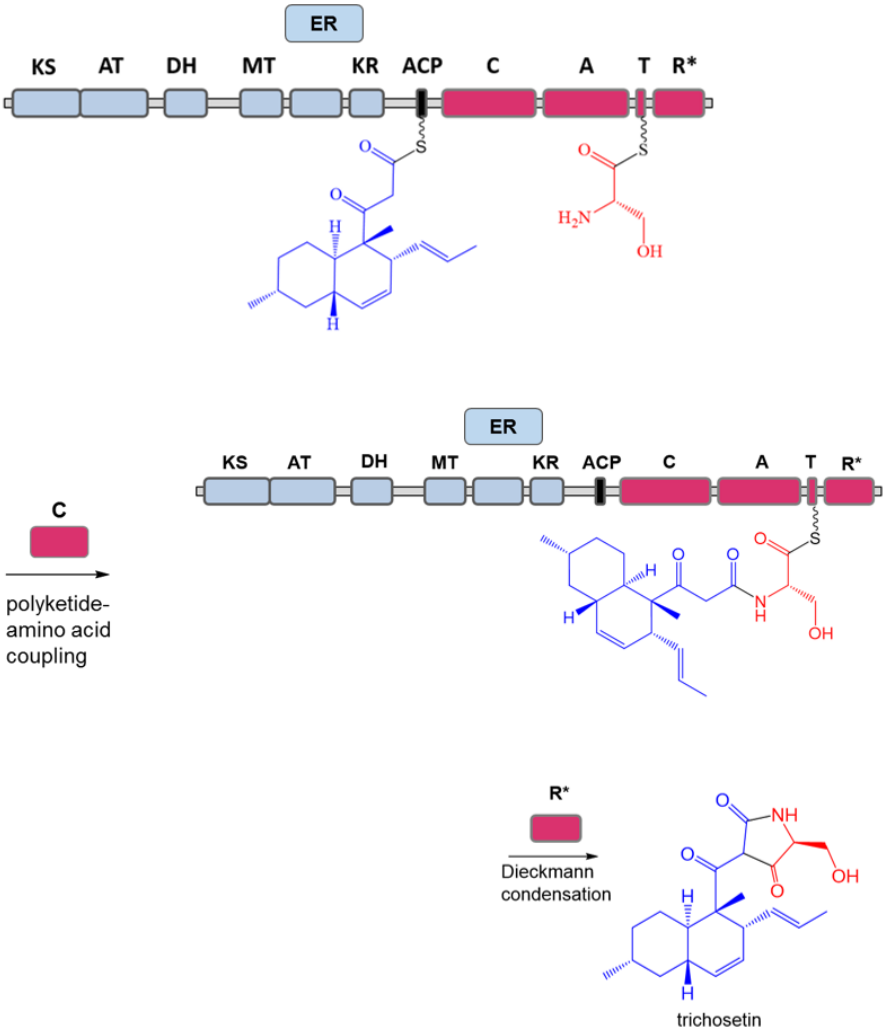


Figure 1.1 PKS-NRPS domain order showing the NRPS module (red) fused to the c-terminus of the PKS module (blue). When the polyketide chain (blue) has been fully extended by the PKS, it is transferred by the acyl carrier protein (ACP) to the NRPS by coupling to the amino acid (red); this reaction is catalyzed by the condensation (C) domain. The intermediate is loaded onto the thiolation (T) domain, and final product release is catalyzed by the R domain via a Dieckmann cyclization to form the tetramic acid ring.

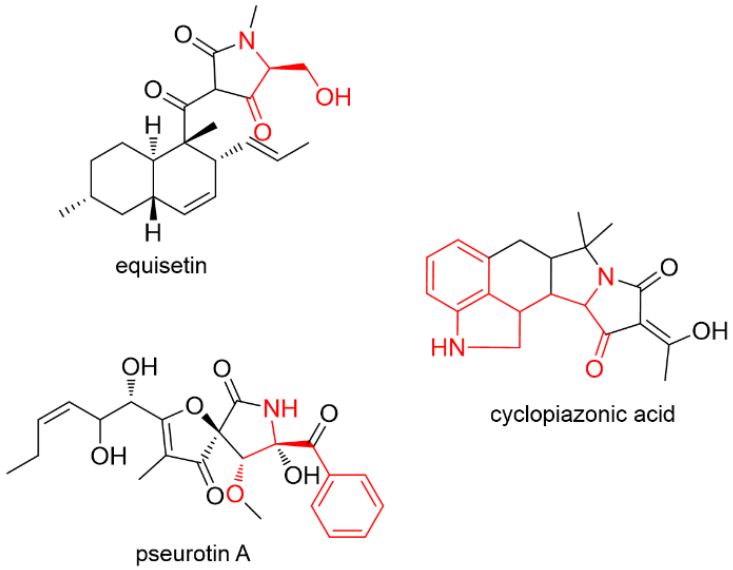


Figure 1.2 Examples of natural products with polyketide scaffolds fused to amino acids (red) and cyclized to form tetramic acid rings.

(LovB) was fused to the chaetoglobosin synthetase (ChaeA) NRPS did not yield the expected hybrid product as would be predicted from the earlier studies.[32] Through relationships inferred from phylogenetic trees, the same study shows that the condensation domains of PKS-NRPSs evolved separately from those of LovB-type PKS enzymes and thus were not able to accept noncognate intermediates. Central to the modular interactions between the PKS and NRPS is the ACP. As mentioned before, the growing polyketide chain remains tethered to the ACP until it is fully extended. For PKS-NRPS enzymes, the polyketide is offloaded from the ACP by way of coupling to the amino acid; this reaction is catalyzed by the C domain. Therefore, productive interaction between the ACP and the C domains is required for successful polyketide transfer, in addition to acceptance of the polyketide intermediate by the C domain. Understanding the overall interactions of ACPs in the context of the PKS-NRPS will allow combinatorial biosynthesis of tetramic acids to access new analogues. In Chapter 5, the results of carefully designed studies are reported that illuminate C domain selectivity in the context of favorable ACP interactions. Based on those findings, ground rules for the combinatorial pairing of PKS and NRPS modules are discussed.

1.6 References

1. Marmann, A.; Aly, A. H.; Lin, W.; Wang, B.; Proksch, P. Cocultivation – A Powerful Emerging Tool for Enhancing the Chemical Diversity of Microorganisms. *Mar. Drugs* **2014**, *12*, 1043–1065.

2. Paul, V. J.; Puglisi, M. P.; Ritson-Williams, R. Marine Chemical Ecology. *Nat. Prod. Rep.* **2006**, *23*, 153–180.

3. Pawlik, J. R. Marine Invertebrate Chemical Defenses. *Chem. Rev.* **1993**, *93*, 1911–1922.

4. Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs Over the Last 25 Years. *J. Nat. Prod.* **2007**, *70*, 461–477.

5. Evans, B. S.; Robinson, S. J.; Kelleher, N. L. Surveys of Nonribosomal Peptide and Polyketide Assembly Lines in Fungi and Prospects for their Analysis in Vitro and in Vivo. *Fungal Genet. Biol.* **2011**, *48*, 49–61.

6. Brakhage, A. A. Regulation of Fungal Secondary Metabolism. *Nat. Rev. Microbiol.* **2013**, *11*, 21–32.

7. Brakhage, A. A.; Schroeckh, V. Fungal Secondary Metabolites – Strategies to Activate Silent Gene Clusters. *Fungal Genet. Biol.* **2011**, *48*, 15–22.

8. Andersen, M. R.; Nielsen, J. B.; Klitgaard, A.; Petersen, L. M.; Zachariasen, M.; Hansen, T. J.; Blicher, L. H.; Gotfredsen, C. H.; Larsen, T. O.; Nielsen, K. F.; Mortensen, U. H. Accurate Prediction of Secondary Metabolite Gene Clusters in Filamentous Fungi. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, E99–107.

9. Walton, J. D. Horizontal Gene Transfer and the Evolution of Secondary Metabolite Gene Clusters in Fungi: An Hypothesis. *Fungal Genet. Biol.* **2000**, *30*, 167–171.

10. Palmer, J. M.; Keller, N. P. Secondary Metabolism in Fungi: Does Chromosomal Location Matter? *Curr. Opin. Microbiol.* **2010**, *13*, 431–436.

11. Bergmann, S.; Funk, A. N.; Scherlach, K.; Schroeckh, V.; Shelest, E.; Horn, U.; Hertweck, C.; Brakhage, A. A. Activation of a Silent Fungal Polyketide Biosynthesis Pathway through Regulatory Cross Talk with a Cryptic Nonribosomal Peptide Synthetase Gene Cluster. *Appl. Environ. Microbiol.* **2010**, *76*, 8143–8149.

12. Schroeckh, V.; Scherlach, K.; Nutzmann, H. W.; Shelest, E.; Schmidt-Heck, W.; Schuemann, J.; Martin, K.; Hertweck, C.; Brakhage, A. A. Intimate Bacterial–Fungal Interaction Triggers Biosynthesis of Archetypal Polyketides in *Aspergillus nidulans*. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 14558–14563.

13. Chiang, Y. M.; Oakley, B. R.; Keller, N. P.; Wang, C. C. Unraveling Polyketide Synthesis in Members of the Genus *Aspergillus*. *Appl. Microbiol. Biotechnol.* **2010**, *86*, 1719–1736.

14. Chiang, Y. M.; Szewczyk, E.; Davidson, A. D.; Keller, N.; Oakley, B. R.; Wang, C. C. A Gene Cluster Containing Two Fungal Polyketide Synthases Encodes the Biosynthetic Pathway for a Polyketide, Asperfuranone, in *Aspergillus nidulans*. *J. Am. Chem. Soc.* **2009**, *131*, 2965–2970.

15. Bergmann, S.; Schuemann, J.; Scherlach, K.; Lange, C.; Brakhage, A. A.; Hertweck, C. Genomics-driven Discovery of PKS–NRPS Hybrid Metabolites from *Aspergillus nidulans*. *Nat. Chem. Biol.* **2007**, *3*, 213–217.

16. Ongley, S. E.; Bian, X.; Neilan, B. A.; Muller, R. Recent Advances in the Heterologous Expression of Microbial Natural Product Biosynthetic Pathways. *Nat. Prod. Rep.* **2013**, *30*, 1121–1138.

17. Ma, H.; Kunes, S.; Schatz, P. J.; Botstein, D. Plasmid Construction by Homologous Recombination in Yeast. *Gene* **1987**, *58*, 201–216.

18. Park, S. R.; Park, J. W.; Jung, W. S.; Han, A. R.; Ban, Y. H.; Kim, E. J.; Sohng, J. K.; Sim, S. J.; Yoon, Y. J. Heterologous Production of Epothilones B and D in *Streptomyces venezuelae. Appl. Microbiol. Biotechnol.* **2008**, *81*, 109–117.

19. Chiang, Y.-M.; Oakley, C. E.; Ahuja, M.; Entwistle, R.; Schultz, A.; Chang, S.-L.; Sung, C. T.; Wang, C. C. C.; Oakley, B. R. An Efficient System for Heterologous Expression of Secondary Metabolite Genes in *Aspergillus nidulans. J. Am. Chem. Soc.* **2013**, *135*, 7720–7731.

20. Gao, L.; Cai, M.; Shen, W.; Xiao, S.; Zhou, X.; Zhang, Y. Engineered Fungal Polyketide Biosynthesis in *Pichia pastoris*: A Potential Excellent Host for Polyketide Production. *Microbial Cell Factories* **2013**, *12*, 77.

21. Stevens, D. C.; Hari, T. P. A.; Boddy, C. N. The Role of Transcription in Heterologous Expression of Polyketides in Bacterial Hosts. *Nat. Prod. Rep.* **2013**, *30*, 1391–1411.

22. Pfeifer, B. A.; Admiraal, S. J.; Gramajo, H.; Cane, D. E.; Khosla, C. Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of *E. coli. Science* **2001**, *291*, 1790–1792.

23. Gao, X.; Wang, P.; Tang, Y. Engineered Polyketide Biosynthesis and Biocatalysis in *Escherichia coli. Appl. Microbiol. Biotechnol.* **2010**, *88*, 1233–1242.

24. Lee, K. K.; Da Silva, N. A.; Kealey, J. T. Determination of the Extent of Phosphopantetheinylation of Polyketide Synthases Expressed in *Escherichia coli* and *Saccharomyces cerevisiae. Anal. Biochem.* **2009**, *394*, 75–80.

25. Tsunematsu, Y.; Ishiuchi, K.; Hotta, K.; Watanabe, K. Yeast-based Genome Mining, Production and Mechanistic Studies of the Biosynthesis of Fungal Polyketide and Peptide Natural Products. *Nat. Prod. Rep.* **2013**, *30*, 1139–1149.

26. Tokuoka, M.; Seshime, Y.; Fujii, I.; Kitamoto, K.; Takahashi, T.; Koyama, Y. Identification of a Novel Polyketide Synthase–Nonribosomal Peptide Synthetase (PKS–NRPS) Gene Required for the Biosynthesis of Cyclopiazonic acid in *Aspergillus oryzae. Fungal Genet. Biol.* **2008**, *45*, 1608–1615.

27. Chooi, Y. H.; Tang, Y. Navigating the Fungal Polyketide Chemical Space: From Genes to Molecules. *J. Org. Chem.* **2012**, *77*, 9933–9953.

28. Crawford, J. M.; Townsend, C. A. New Insights into the Formation of Fungal Aromatic Polyketides. *Nat. Rev. Microbiol.* **2010**, *8*, 879–889.

29. Schobert, R.; Schlenk, A. Tetramic and Tetronic Acids: An Update on New Derivatives and Biological Aspects. *Bioorg. Med. Chem.* **2008**, *16*, 4203–4221.

30. Boettger, D.; Hertweck, C. Molecular Diversity Sculpted by Fungal PKS–NRPS Hybrids. *ChemBioChem* **2013**, *14*, 28–42.

31. Fisch, K. M. Biosynthesis of Natural Products by Microbial Iterative Hybrid PKS–NRPS. *RSC Adv.* **2013**, *3*, 18228–18247.

32. Boettger, D.; Bergmann, H.; Kuehn, B.; Shelest, E.; Hertweck, C. Evolutionary Imprint of Catalytic Domains in Fungal PKS–NRPS Hybrids. *ChemBioChem* **2012**, *13*, 2363–2373.

33. Fisch, K. M.; Bakeer, W.; Yakasai, A. A.; Song, Z.; Pedrick, J.; Wasil, Z.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. Rational Domain Swaps Decipher Programming in Fungal Highly Reducing Polyketide Synthases and Resurrect an Extinct Metabolite. *J. Am. Chem. Soc.* **2011**, *133*, 16635–16641.

34. Xu, W.; Cai, X.; Jung, M. E.; Tang, Y. Analysis of Intact and Dissected Fungal Polyketide Synthase–Nonribosomal Peptide Synthetase in Vitro and in *Saccharomyces cerevisiae*. *J. Am. Chem. Soc.* **2010**, *132*, 13604–13607.

CHAPTER 2

TWO RELATED PYRROLIDINEDIONE SYNTHETASE LOCI IN FUSARIUM HETEROSPORUM ATCC 74349 PRODUCE DIVERGENT METABOLITES

Reprinted with permission from
Kakule, T. B.; Sardar, D.; Lin, Z.; Schmidt, E. W. Two Related Pyrrolidinedione Synthetase Loci in *Fusarium heterosporum* ATCC 74349 Produce Divergent Products. *ACS Chem. Biol.* **2013**, *8*, 1549–1557.

Note: my contribution to this paper was in planning, performing, and analyzing all the experiments with the exception of the biochemical assays.

ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

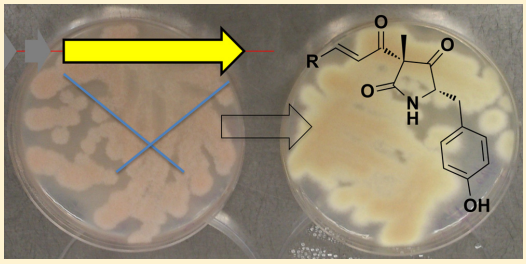
Two Related Pyrrolidinedione Synthetase Loci in *Fusarium heterosporum* ATCC 74349 Produce Divergent Metabolites

Thomas B. Kakule, Debosmita Sardar, Zhenjian Lin, and Eric W. Schmidt*

Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112, United States

Supporting Information

ABSTRACT: Equisetin synthetase (EqiS), from the filamentous fungus *Fusarium heterosporum* ATCC 74349, was initially assigned on the basis of genetic knockout and expression analysis. Increasing inconsistencies in experimental results led us to question this assignment. Here, we sequenced the *F. heterosporum* genome, revealing two hybrid polyketide-peptide proteins that were candidates for the equisetin synthetase. The surrounding genes in both clusters had the needed auxiliary genes that might be responsible for producing equisetin. Genetic mutation, biochemical analysis, and recombinant expression in the fungus enabled us to show that the initially assigned EqiS does not produce equisetin but instead produces a related 2,4-pyrrolidinedione, fusaridione A, that was previously unknown. Fusaridione A is methylated in the 3-position of the pyrrolidinedione, which has not otherwise been found in natural products, leading to spontaneous reverse-Dieckmann reactions. A newly described gene cluster, *eqx*, is responsible for producing equisetin.

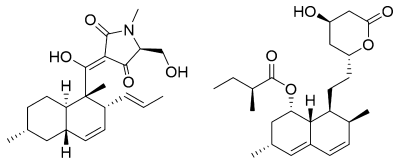


Filamentous fungi are prodigious producers of secondary metabolites. Within this diverse range of natural products, polyketides make up a large class, and their biosynthesis has been the subject of much study over the past decade.[1] A subgroup of polyketides is condensed with amino acids and cyclized to form 3-acyl pyrrolidinediones, also known as tetramic acids. These tetramic acids of natural origin are the product of a polyketide synthase (PKS) fused to a nonribosomal peptide synthetase (NRPS) module. In addition to tetramic acids, fungal PKS-NRPS proteins also lead to the synthesis of tetramic acid-like compounds, such as pyridinones and other derivatives that probably result from redox changes or pericyclic reactions.[2,3]

In such hybrid proteins, the PKS module belongs to the family of highly reducing PKSs (HR-PKS), which iteratively produce complex polyketides through mechanisms that are incompletely understood.[4,5] Typical HR-PKS domains are universally present in fungal hybrids, including ketosynthase (KS), acyltransferase (AT), ketoreductase (KR), dehydratase (DH), *C*-methyltransferase (MT), and acyl carrier protein (ACP). Together, these domains are capable of synthesizing polyketide-derived portions that are α-C-methylated and with carbons in varying reduction states, including ketone, hydroxyl, and olefin. In addition, an enoylreductase (ER) domain, which reduces double bonds, is often present either within the PKS or as an auxiliary protein.[5] In the case of lovastatin biosynthesis, the ER has been shown to collaborate with the core PKS in the production of a decalin ring rather than a linear polyketide.[6]

The NRPS module consists of an adenylation (A) domain that activates a specific amino acid, which is loaded onto the thiolation (T) domain. The covalently loaded amino acid is then condensed with the fully extended polyketide chain from the upstream PKS module by the condensation (C) domain. A Dieckmann cyclization domain (R*) catalyzes the final step that releases the intermediate to form a tetramate, such as pretenellin A, and prepseurotin A.[7−9]

Fusarium heterosporum ATCC 74349 produces a large amount of a tetramate, equisetin **1** (~2 g L^{-1} under some conditions).[10] Equisetin is similar to lovastatin **2** in its polyketide portion, but unlike lovastatin it appends an additional amino acid, serine (Figure 1). This pattern is very similar to products synthesized by fungal hybrid PKS-NRPS enzymes. Because of homology of equisetin to lovastatin, it is



equisetin **1** lovastatin **2**

Figure 1. Decalin-derived fungal polyketides. The decalin substructure is similar for equisetin **1** and lovastatin **2**. The major difference between the biosynthesis of equisetin and lovastatin would be the formation of the NRPS-derived tetramic acid ring of equisetin, which is not found in lovastatin.

Received: March 5, 2013
Accepted: April 24, 2013
Published: April 24, 2013

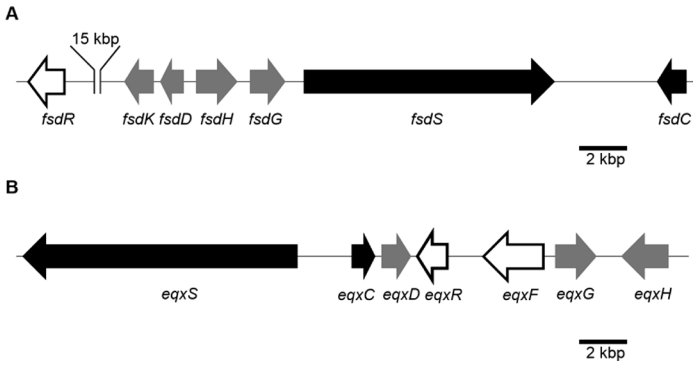


Figure 2. PKS-NRPS gene clusters found in the *F. heterosporum* genome. (A) The *fsd* cluster for fusaridione A (previously known as the *eqi* cluster). (B) The *eqx* cluster for equisetin biosynthesis. Genes including PKS domains are shown in black, including the PKS-NRPS genes *fsdS* and *eqxS* and the *trans*-ER-like genes *fsdC* and *eqxC*. Regulators *fsdR*, *eqxR*, and *eqxF* are shown in white. Other putative enzymes and transporters are shown in gray, including methyltransferases (*fsdD* and *eqxD*), transporters (*fsdG* and *eqxG*), oxidases (*fsdH* and *eqxH*), and a prenyltransferase (*fsdK*). (Note: overexpression of *fsdR* did not alter *fsd* expression, indicating that it may not be part of the pathway; it lies ∼15 kbp from the other pathway genes.).

expected that, in addition to the core PKS-NRPS protein, an auxiliary ER would be required.[5,6] Equisetin is also amide N-methylated, which would most likely require an additional, dedicated MT. Using degenerate PCR and knockout mutation analysis in *F. heterosporum*, a previous study identified three reducing PKS genes, *deg1*, *deg2*, and *deg3*.[10] Through knockout mutagenesis, *deg2* was assigned as being responsible for equisetin production. This PKS-NRPS hybrid gene was thus renamed *eqiS*. The *deg2* cluster also contained genes potentially involved in equisetin biosynthesis, such as additional candidate ER and MT genes.

A key feature of equisetin biogenesis was that equisetin was produced at high levels (2 g L^{-1}) only on corn grit agar (CGA), but no detectable equisetin was produced on many other types of media.[10] In later studies, we sought to exploit this feature for heterologous expression of fungal genes. Several failed attempts led us to question whether the equisetin biosynthetic genes had been correctly identified. Here, we show through genomic, biochemical, genetic, and chemical studies that the previously characterized *eqiS* is not essential for equisetin production but instead produces a novel tyrosine-derived 2,4-pyrrolidinedione, which we named fusaridione A. A second PKS-NRPS gene found in the genome, *deg3*, instead produces equisetin. We renamed the *deg2* cluster *fsd* and the *deg3* cluster *eqx*.

RESULTS AND DISCUSSION

***fsd* Does Not Encode the Equisetin Synthetase.** The *deg2/eqiS/fsdS* knockout vector, TOPO-deg2KO, contained elements of the PKS *fsdS* flanking a phleomycin resistance marker. Homologous recombination in *F. heterosporum* produced a double crossover mutant, FusKO474, as verified by PCR. In addition, the previously described *eqiS/fsdS* hygroymcin resistance-based knockout, JWS19,[10] was resurrected from freezer stocks. In previous experiments, JWS19 was reported to produce no equisetin on CGA after 10 days, while wild-type *F. heterosporum* synthesized abundant equisetin.[10] However, here we found that both FusKO474 and JWS19 produced abundant equisetin equivalent to that produced by the wild-type strain on CGA after 21 days of growth (Supplementary Figure S1). Therefore, *fsdS* is not involved directly in equisetin biosynthesis.

Two PKS-NRPS Clusters Identified in the *F. heterosporum* Genome. To rapidly identify the correct equisetin pathway, the whole genome of *F. heterosporum* ATCC 74349 was sequenced by paired-end Illumina. Automated assembly availed a genome of 39 Mbp in length with a calculated GC content of 47.7% on 120 contigs (75% of the genome was on 17 contigs). These genome statistics are very similar to those for other *Fusarium* species that have been sequenced, such as *F. graminearum* (36 Mbp, 48.3% GC) and *F. verticillioides* (42 Mbp, 48.7% GC).[11] *De novo* annotation software, Maker2, was employed to predict a total of 12,448 proteins using the gene prediction model provided by *F. graminearum*.[12]

BLAST analysis of the whole genome (both assembled and raw reads) identified only three PKS genes with homology to the lovastatin nonaketide synthase, which were identical to the genes previously identified by PCR.[10] In addition, 9 further PKS genes were present in the genome but lacked any features consistent with equisetin production (Supplementary Figure S5). Of these three genes, *fsdS* and *eqxS* were PKS-NRPS genes, whereas *deg1* lacked an NRPS portion. *fsdS* and *eqxS* were very similar in domain order and content, although their protein sequence identity was only 47%. We employed the SMURF program to analyze fungal secondary metabolite genes clustered about these PKS-NRPS genes, revealing that both *fsd* and *eqx* clusters encoded at least one regulatory protein, a MT, a reductase that may function as an ER, and an exporter (Figure 2, Supplementary Tables S5 and S6).[13] Therefore, without functional analysis, these clusters were difficult to differentiate. Since *fsdS* was ruled out as the equisetin synthetase, and since abundant evidence implicates PKS-NRPS genes in tetramic acid synthesis, *eqxS* was a likely candidate.

Pyrophosphate Exchange Assay of FsdS and EqxS. We reexamined the amino acid substrate specificity of both the *fsdS* and *eqxS* proteins. A series of His-tagged, adenylation domain-containing constructs were synthesized to contain varying combinations of NRPS domains and expressed in *E. coli*. The most stable, folded constructs (*eqxS* A and ATR domains; *fsdS* ACP-CATR domains) were purified and used in the pyrophosphate exchange assay[14] using all 20 proteinogenic amino acids and *N*-methyl-L-serine. The major amino acid activated by *fsdS* was L-tyrosine, while that activated by *eqxS*

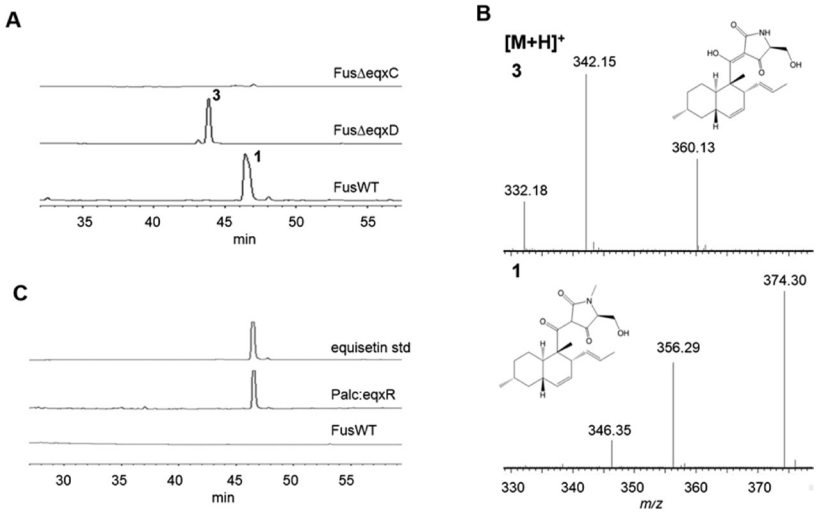


Figure 3. Knockouts and overexpression of *eqx* pathway genes. (A) Analytical HPLC of crude extracts of *eqxC* and *eqxD* knockouts compared to the wild-type strain (monitored at λ 293 nm). The *eqxD* knockout mutant does not produce equisetin **1** but instead produces an earlier eluting compound **3**. However, knocking out *eqxC* abolishes equisetin production, and no new metabolites are observed. (B) LC−MS shows that compounds **1** and **3** differ by 14 Da. (C) Analytical HPLC-DAD analysis of crude extracts of PDB cultures of *F. heterosporum* WT and the Palc:eqxR mutant. Overexpression of *eqxR* induces equisetin production (top trace shows equisetin standard).

was L-serine (Supplementary Figures S3A−D). This implicated *eqx* in equisetin synthesis, since equisetin contains Ser, while *fsdS* would likely produce a tyrosine-derived molecule.

Mutagenesis of *eqx* Cluster. Vectors were constructed to knock out *eqxS* and clustered genes N-MT *eqxD* and *trans*-ER *eqxC*. Initially, all vectors were constructed using native sequences flanking a hygromycin resistance cassette. These vectors were transformed into *F. heterosporum*, and transformants were screened for homologous recombination by PCR (Supplementary Figure S2). While multiple knockout mutants of *eqxC* and *eqxD* were obtained on the first attempt, concomitant knockout of *eqxS* was unsuccessful. Indeed, over 15 attempts were made to knockout *eqxS* using three different vectors, using both hygromycin and phleomycin as selection markers. By comparison, simultaneous attempts to introduce other genes were successful, usually in the first attempt, using the same materials. Speculatively, this may result from an untoward phenotype resulting from manipulation of *eqxS*.

In addition, predicted *eqx* regulators *eqxR* and *eqxF* were cloned into a phleomycin selection vector so that their expression could be controlled by the *Aspergillus nidulans alcA* inducible promoter and inserted into the genome by ectopic integration.[15−17]

***N*-Methyltransferase *eqxD* Plays a Role in Equisetin Biosynthesis.** *eqxD* knockout FusΔeqxD was grown on CGA for 21 days. Extraction with acetone followed by analytical HPLC showed that equisetin production was abolished. Instead, high-level production of an earlier eluting compound was observed (Figure 3A). By DAD-HPLC, the compound had the same chromophore as equisetin, and the LC−MS spectrum showed that the compound was 14 Da lighter than equisetin (Figure 3B). Purification on C_{18} column with a methanol/water gradient afforded pure compound **3** (6 mg; isolated yield 240 mg **3** per kg media). The molecular formula of **3** was determined by FT-ICR ESIMS to be $C_{21}H_{29}NO_4$ (m/z 360.21698 $[M + H]^+$, Δ 0.14 ppm). 1D proton NMR supported the absence of the *N*-methyl singlet at δ 3.07 observed for equisetin.[18] The 1D and 2D NMR spectra were essentially identical with those for the previously reported metabolite trichosetin, a tetramate produced by *Trichoderma harzianum* in dual culture with *Catharanthus roseus*.[19] For example, the ^{13}C spectrum of **3** was essentially identical to that shown in the previous manuscript. Moreover, all of the analyzed 2D data were consistent with the proposed structure for trichosetin.[19] Trichosetin is synonymous with *N*-desmethylequisetin. The chemical specificity of this result provided positive evidence associating the *eqx* cluster with equisetin production.

Equisetin Biosynthesis Requires *trans* ER (*eqxC*) Activity. The *eqxC* knockout, FusΔeqxC, was grown on CGA for 21 days. Extraction followed by analytical HPLC showed that equisetin was not produced, while the wild-type strain produced abundant equisetin (Figure 3A). Since a *trans*-ER is an integral part of fungal HR-PKS proteins,[5] this result strongly implicated the *eqx* PKS in equisetin synthesis.

Altered Regulation of *eqx* Cluster Allows High-Level Production of Equisetin in Broth Culture. Other than CGA culture for equisetin production employed by our lab, one other defined solid-state medium has been reported to induce *F. heterosporum* to biosynthesize equisetin.[20] Our attempts to induce the wild-type strain to produce equisetin in liquid culture by altering both carbon and nitrogen sources were unsuccessful (unpublished data). To achieve controllable expression in liquid culture, we cloned the two *eqx* putative regulators downstream of the inducible *A. nidulans alcA* promoter to make plasmids alcAeqxF and alcAeqxR. Independent transformation of these plasmids into FusWT with subsequent selection on phleomycin led to isolation of transformants Palc:eqxR and Palc:eqxF, which were verified by PCR to carry the expression cassette. Spores were prepared

dx.doi.org/10.1021/cb400159f | *ACS Chem. Biol.* 2013, 8, 1549−1557

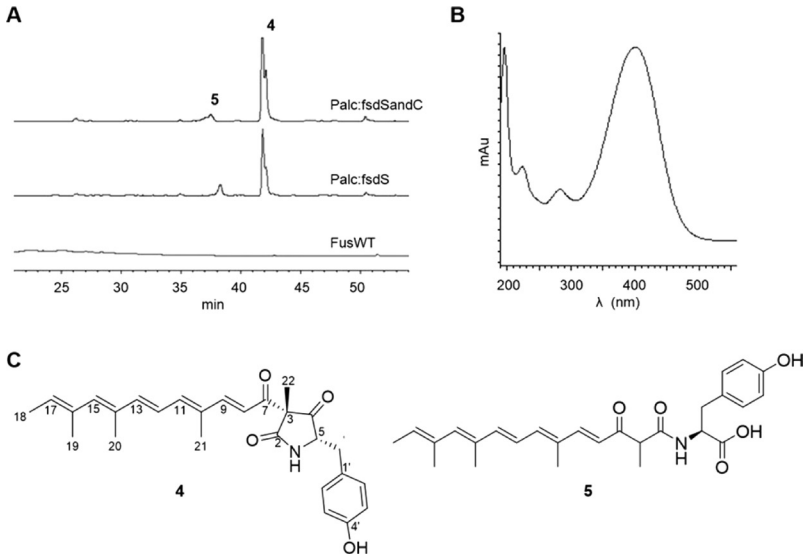


Figure 4. Overexpression of *fsd* pathway genes. (A) Analytical RP HPLC of crude mycelial extracts of 7 day PDB cultures of *F. heterosporum* transformed with *fsdS* and *fsdC* under control of *A. nidulans alcA* promoter (monitored at λ 380 nm). Two major new metabolites (**4** and **5**) were produced by the Palc:fsdS mutant. Introduction of the *fsdC* gene into Palc:fsdS to make the mutant Palc:fsdSandC did not lead to other products. (B) UV spectrum of **4** measured by DAD showing maxima at λ 402 nm, λ 280 nm, and λ 224 nm. (C) Structures of **4** and **5** were determined by 1D and 2D NMR. Metabolite **4** is a novel pyrrolidinedione, and **5** is its ring-opened form.

from each of the mutants and cultured in PDB for 18 h at 30 °C followed by cyclopentanone induction and further incubation for 4 days. DAD-HPLC analysis of the crude extracts from mutant cultures compared to the wild-type strain showed that Palc:eqxR synthesized equisetin at 207 mg L^{-1} under these conditions. By contrast, neither Palc:eqxF nor FusWT produced detectable equisetin in liquid broth (Figure 3C). Therefore, *eqxR* is a positive regulator of the *eqx* cluster. Since regulatory elements are often clustered with fungal biosynthetic pathways, this provided strong positive evidence for the involvement of *eqx* in equisetin production.

***fsd* Pathway Produces a Novel Tyrosine-Derived 2,4-Pyrrolidinedione.** We could not observe any evidence for production of putative *fsd* pathway products in a wide variety of media. In these experiments, we compared the metabolite expression patterns in the *fsdS* knockout (FusKO474) with those of the wild-type strain. This led us to explore another strategy to determine the secondary metabolite associated with potential *fsdS* activity. The promoter of *fsdS* was swapped for the *A. nidulans alcA* promoter to make the fungal expression plasmid alcAfsdS. Transformation of *F. heterosporum* with alcAfsdS led to isolation of a hygromycin-resistant mutant, Palc:fsdS. A possible *trans*-ER, *fsdC*, was also cloned into an *alcA*-controlled expression plasmid to generate alcAfsdC, which was then transformed into Palc:fsdS with phleomycin selection to generate the cotransformed mutant Palc:fsdSandC. Transformation resulted in colonies that were bright yellow on PDA, while wild-type *F. heterosporum* is pink (Supplementary Figure S4). Both mutants were cultivated in PDB with and without cyclopentanone induction. The filtered broth was analyzed by DAD-HPLC. Two new metabolites were produced by both the Palc:fsdS and Palc:fsdSandC mutants, but the metabolites were not detected in the FusWT extract (Figure 4A).

The major new metabolite was purified by C_{18} chromatography to yield a bright yellow compound, fusaridione A (**4**, 8 mg). The molecular formula was determined by FT-ICR ESIMS to be $C_{27}H_{31}O_4N$ (m/z 434.23267 $[M + H]^+$, Δ 0.21 ppm). All proton signals except for two exchangeable protons could be detected in the 1D ^1H NMR spectrum. From the HMQC, HSQC, and HMBC spectra, all 27 carbons were accounted for: 9 quaternary, 5 methyl, 1 methylene, and 12 methine. Consistent with the pyrophosphate exchange assay data showing that FsdS activates tyrosine, typical NMR shifts for tyrosine were identified, including a *p*-substituted phenol (δ 7.06 (d) and δ 6.76 (d)) and a benzylic methylene group (δ 3.06(d) and δ 2.84(d)). The α-proton (H-5) of tyrosine was observed at δ 4.29 and correlated to the phenol ring by TOCSY via W-coupling. UV maxima were observed at λ 280 nm (phenol) and at λ 402 nm for an additional conjugated moiety that was determined to be a pentaene by NMR analysis (Figure 4B). Analysis of the NMR data revealed that the ring is not a tetramic acid, but instead a pyrrolidine-3-methyl-2,4-dione (Figure 4C). The structure was further confirmed by observing ^3J HMBC correlations from the C-22 methyl protons (H-22, δ 1.46) to the carbonyl carbons, the amide carbonyl (C-2, δ 172.0), and C-4 (δ 207.1) (Figure 5). H-5 and H-6 also showed HMBC correlations with ring carbonyls.

The pentaene side chain structure was determined by examination of COSY, TOCSY, and HMBC data, which unambiguously revealed the positions of vinylic methyl groups. The attachment of the polyene side chain to the 3-position of the ring was revealed through an additional ^3J HMBC signal from the C-22 methyl group to the C-7 carbonyl (δ 189.8). The connectivity shown for the polyene chain was determined from the HMBC, COSY and TOCSY spectra. From the NOESY spectra, the *trans* nature of the double bonds was confirmed.

dx.doi.org/10.1021/cb400159f | *ACS Chem. Biol.* 2013, 8, 1549−1557

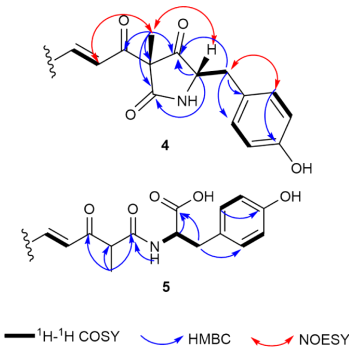


Figure 5. Key NMR data for **4** and **5**.

The methyl group at the 3-position of the pyrrolidinedione ring had NOESY correlations with H-5, fixing it *anti* to C-β of the tyrosine (Figure 5). The tyrosine *S*-configuration was determined by a variation of Marfey's method.[21]

A second major compound present in the *fsdS* expressions was purified as a bright yellow compound, **5** (3 mg), which had a molecular formula of $C_{27}H_{33}NO_5$ (*m*/*z* 452.24335 $[M + H]^+$, Δ 0.44 ppm), differing from **4** by the addition of H_2O. By DAD-HPLC, the compound had UV maxima at λ 280 nm and λ 386 nm, differing from **4** by a blue shift for the conjugated moiety that likely arises from cross-conjugation in comparison to **4**. Indeed, 1D and 2D NMR spectroscopy revealed that **5** was a ring-opened variant of **4** (Figure 4C). For example, chemical shift data of the tyrosine α-carbon (δ 51.6) was consistent with a linear, and not cyclic, molecule. A complicating factor was that the pentaene moiety was present as a mixture of two tautomers, which could not be resolved by changing the solvent or changing the temperature from −20 to +30 °C. HMBC data from both tautomers confirmed the side-chain structure of the molecule. Purification and concentration of **4** using TFA in the chromatography solvents afforded pure **5**, indicating that **5** is a degradation product of **4** that likely results from a reverse-Dieckmann reaction. Upon casual observation, the tautomer problem makes some spectra appear to contain mixtures, when in fact both **4** and **5** were obtained and analyzed as pure materials.

Conclusion. Filamentous fungal genomes contain numerous secondary metabolite gene clusters.[1,22] In recent years, it has become clear that most of these clusters are silent.[1,2,23,24] To fully exploit fungal pathways, currently the best strategy employs genome sequencing, followed by any one of several known methods of eliciting otherwise cryptic pathways.[16,17,25] Here, we sequenced the genome of *F. heterosporum* ATCC 74349 and then used a variety of functional approaches to elucidate the function of two hybrid PKS-NRPS gene clusters. We uncovered a novel PKS-NRPS cluster responsible for equisetin production and reassigned the previously ascribed equisetin synthetase gene to a novel pyrrolidinedione with unusual methylation.

Previously, the equisetin synthetase was assigned by knocking out *fsdS*, which is synonymous with *deg2* and *eqiS*, and finding that equisetin production was abolished in 10-day fermentations.[10] However, the data presented here definitively show that this was a misassignment and that the *eqx* gene cluster is responsible for equisetin biosynthesis. This previous study reinforces the well-known danger of relying on primarily negative data, such as knockout mutagenesis, for functional assignment of genes. It is remarkable that the *fsd* and *eqx* gene clusters are similar in terms of their biosynthetic content. This finding also explains some discrepancies in the literature. Analysis of fungal PKS-NRPS C domain phylogenies showed that the previously identified EqiS fell into a clade that uses aromatic amino acids, whereas EqiS was reported to use serine.[26] This information was interpreted to indicate polyphyly in the NRPS modules of fungal PKS-NRPS proteins. Based upon the evidence here, it is clear that C domain phylogeny is predictive of amino acid specificity.

The following lines of evidence support *eqx* as the equisetin biosynthetic gene cluster. (1) Only two PKS-NRPS genes (*fsdS* and *eqxS*) were present in the genome, and other biosynthetic genes did not have the correct functions for tetramate biosynthesis. (2) Two different *fsdS* knockouts did not inhibit equisetin biosynthesis. (3) Knockout of the PKS component ER *eqxC* abolished equisetin production. (4) Knockout of the MT *eqxD* prevented methylation and led to accumulation of desmethylequisetin. (5) Overexpression of regulator *eqxR* led to efficient synthesis of equisetin in 5 days in liquid culture, whereas otherwise 3 weeks on solid media is required. While activation of a regulator can sometimes activate a remote biosynthetic gene cluster as well as the adjacent cluster,[27] in this case strong induction led to a single product that was attributed to the adjacent cluster. Taken together, this provides both gain- and loss-of-function experiments demonstrating that *eqx* is crucial for equisetin production. Because *eqx* contains broadly similar genes to those found in *fsd*, the biosynthetic scheme is likely identical to that previously proposed (Figure 6).[10]

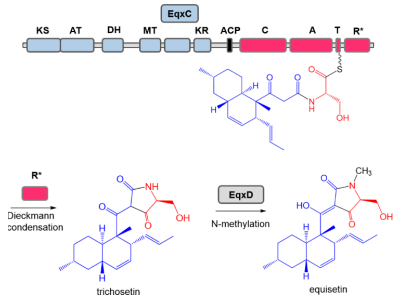


Figure 6. Proposed biogenesis of equisetin. The PKS module of EqxS together with the enoylreductase (EqxC) catalyze the formation of the polyketide unit (blue) which is then conjugated to L-serine (red) by the condensation domain of the NRPS module. Activity of the Dieckmann cyclase domain (R*) results in release of the intermediate as the tetramate, trichosetin. Subsequent N-methylation is carried out by EqxD to give equisetin.

fsdS was apparently silent since no metabolites could be attributed to this pathway, despite controlled experiments in which wild-type was compared to *fsdS* knockouts in multiple conditions. Alternatively, it is possible that the final products of the pathway are unstable or otherwise difficult to isolate. Upon induction under artificial conditions, *fsdS* led to production of the highly unstable product fusaridione A **4**. This product is likely not the natural product, as the other genes in the cluster were not coexpressed, except for the putative ER-like gene, *fsdC*. Other genes located in chromosomal proximity to *fsdS* encode proteins such as a potential prenyltransferase (FsdK)

dx.doi.org/10.1021/cb400159f | *ACS Chem. Biol.* 2013, 8, 1549−1557

and a cytochrome p450 oxidase (FsdH). It is possible that the putative ER does not function as an enzyme in the *fsd* pathway. Alternatively, the products of some of the other proteins (such as FsdK or FsdH) may be the substrates of FsdC. It should be emphasized that FsdC is likely not an ER, as its sequence is only distantly related to other characterized *trans*-ER sequences.

Compound **4** bears structural similarity to products of other fungal PKS-NRPS proteins, such as the proposed prepseurotin A[9] and the known pretenellin A[28] metabolites. In addition, it is similar to the known fungal products, militarinones, for which the biosynthetic genes have not been characterized.[29] Therefore, the proposed biosynthesis of **4** is consistent with what is known about fungal PKS-NRPS metabolism (Figure 7). The

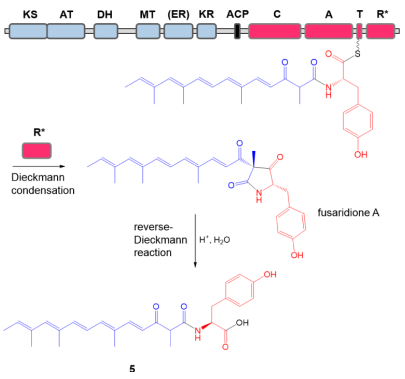


Figure 7. Proposed biogenesis of fusaridione A. The PKS module of FsdS catalyzes the formation of the polyketide unit (blue) which is then conjugated to L-tyrosine (red) by the condensation domain of the NRPS module. Activity of the Dieckmann cyclase domain (R*) results in release of the intermediate as fusaridione A. The unstable pyrrolidinedione ring of fusaridione A is opened through a reverse-Dieckmann reaction to afford **5**.

polyketide is synthesized by incorporation of the equivalent of seven acetate units. Every extension requires participation of the KS, AT, and ACP domains. KR and DH domains produce double bonds at every possible position except for C-7, which remains in the ketone oxidation state. As with other fungal polyketides, methyl groups would be introduced by a dedicated C-methyltransferase domain within the PKS, which acts at C-3, -10, -14, and -16. Subsequently, through the action of C, A, and T domains, tyrosine is activated and appended to the polyketide chain.

The major substantive difference between **4** and previously known compounds is the methyl group at C-3. Because of this methyl group, the product of the Dieckmann condensation **4** is readily degraded via the reverse-Dieckmann reaction in the course of fermentation or when treated with acid. Previously, it was shown that the R* domain in tetramic acid synthetases releases the Dieckmann product.[7,8] In the synthetic literature, it is reported that α-substituted, β-diketones readily undergo the reverse-Dieckmann reaction due to the lack of possible resonance stabilization.[30] Therefore, these results, in tandem with the known activity of R* domains, are consistent with an enzymatic release of **4**, followed by a nonenzymatic hydrolysis to yield the linear product **5**.

In conclusion, we have identified and assigned equisetin production by *F. heterosporum* to a new PKS-NRPS gene cluster, *eqx*. In addition, the product of the PKS-NRPS (FsdS) of the previously reported cluster has been characterized to be a novel compound, (3*R*,5*S*)-5-(4-hydroxybenzyl)-3-methyl-3-((2*E*,4*E*,6*E*,8*E*,10*E*)-4,8,10-trimethyldodeca-2,4,6,8,10-pentaenoyl)pyrrolidine-2,4-dione.

METHODS

Gene Cloning and Analysis Methods. DNA was obtained from *Escherichia coli* and yeast using the QIAprep Spin Miniprep Kit (Qiagen) and from *F. heterosporum* using the DNeasy Plant Mini Kit (Qiagen). DNA was amplified by PCR using high-fidelity enzymes Platinum HiFi Taq (Invitrogen) or Phusion Hot Start II HiFi (Finnzymes). Clones were analyzed using Platinum Taq (Invitrogen). Amplified fragments were cloned into destination vectors using yeast recombination in *Saccharomyces cerevisiae* BY4741, using a previously described lithium acetate method,[31] or by TOPO TA cloning (Invitrogen). The yeast-derived plasmids universally contained an *E. coli* replicon and resistance marker as well as the described[32] yeast plasmid elements. Plasmid rescue from yeast was performed using *E. coli* TOP10 (Invitrogen). Vectors used in this study, their construction methods, and primer sequences are provided in Supporting Information.

Fungal Mutagenesis. Vectors were cloned using hygromycin[33] or phleomycin[34] resistance markers. Gene knockout vectors were obtained by flanking resistance markers with ~2 kbp of *F. heterosporum* target DNA on each side. Overexpression vectors were obtained by fusing genes in frame with the cyclopentanone-responsive promoter *alcA*.[15,17] Linearized plasmids (10 μg) were transformed into *F. heterosporum* following a previously reported method,[10] except that the protoplasting buffer was modified to contain the following (in 20 mL): *Trichoderma* lysing enzyme (1 g), yatalase (100 mg), hemicellulase (30 mg), and β-glucuronidase (5 mg). Transformants were regenerated for 17 h at 30 °C on regeneration broth (1 M sucrose, 0.02% Difco yeast extract) and then plated on regeneration broth/1% agar with the appropriate selection agent (hygromycin 150 μg mL^{-1} or phleomycin 150 μg mL^{-1}). Colonies were screened by PCR to determine whether DNA was inserted via homologous recombination or ectopically. All screens for homologous recombination employed primer sets where the PCR product is ~8-fold larger if homologous recombination has occurred. Screens were employed both on colonies and on mature mycelia after weeks of growth to ensure integrity and purity of recombinant clones.

To knockout *fsdS*, phleomycin-resistance knockout vector TOPO-deg2KO was transformed into *F. heterosporum* by the protoplast method to yield FusKO474. *eqxC* and *eqxD* knockout vectors TOPO-OxoKO and TOPO-MTKO were synthesized similarly, except using the hygromycin resistance marker. Several *eqxS* knockout vectors were synthesized using both the hygromycin and phleomycin markers.

To overexpress proteins in *F. heterosporum* using *alcA*, transcription factors *eqxF* (alcAeqxF) and *eqxR* (alcAeqxR) were transformed into *F. heterosporum* using phleomycin selection, while *fsdS* (alcAfsdS) and *fsdC* (alcAfsdC) were transformed under hygromycin and phleomycin selection, respectively. A double mutant was constructed containing both alcAfsdS and alcAfsdC.

Genome Sequencing and Analysis. *F. heterosporum* DNA was sequenced at the University of Utah Huntsman Cancer Institute sequencing facility on an Illumina HiSeq 2000. A single lane was sequenced in a 100 bp paired-end sequencing run. The resulting raw reads were trimmed to provide reads >40 bp in length and with a PHRED quality score >30. The highest quality assembly was obtained using VELVET with the k-mer value of 61 and 235 x coverage.[35] The sequence was autoannotated with Maker2 using Augustus *de novo* gene prediction parameters, with *Fusarium graminearum* as the model organism.[12,36] BLASTx analysis was performed using a variety of polyketide genes as queries; to detect the closest relatives to lovastatin, residues 9−445 of the lovastatin nonaketide synthase (AAD39830.1)

dx.doi.org/10.1021/cb400159f | ACS Chem. Biol. 2013, 8, 1549−1557

were used in the search. Supplementary Tables S3 and S4 provide genome statistics and e-values for BLAST results.

Pyrophosphate Exchange Assay. The *eqxS* A and ATR domains, as well as the *fsdS* ACP-CATR, were constructed as N-terminal His-tagged sequences, as described in Supporting Information. The *eqxS* constructs were expressed in *E. coli* Rosetta 2DE3 (Novagen). The cultures were grown in 2XYT broth supplemented with ampicillin and chloramphenicol at 225 rpm and 30 °C until an OD_{600} of 0.4 was attained. The temperature was then reduced to 18 °C, IPTG (100 μM) was added, and the cultures were incubated for 18 h. *fsdS* ACP-CATR was expressed in *E. coli* BL21(DE3), to which the chaperone plasmid pG-KJE8 (Takara Bio Inc.) had been added. Protein expression was performed as above, except that the medium was Luria−Bertani (LB) broth, and the inducers were L-arabinose (2 mg mL^{-1}) and tetracycline (2 ng mL^{-1}) in addition to IPTG (100 μM).

Cells were harvested, flash frozen, and stored at −80 °C. Thawed cell pellets were resuspended in lysis buffer (5 mL g^{-1} cell pellet; 50 mM Tris, 200 mM NaCl, 10 mM imidazole, 5% glycerol, pH 7.5) and stirred on ice for 1 h with lysozyme (600 μg mL^{-1}). Cells were sonicated at 30% amplitude for 3 min with 30 s on/off cycles on the VibraCell 750 instrument and then incubated on ice for 30 min with DNase (20 μg mL^{-1}) and 10 mM $MgCl_2$. Following centrifugation at 20,000*g* for 45 min, the supernatant was filtered and loaded onto Ni-NTA resin. The resin was washed twice (1 M NaCl, 30 mM imidazole, pH 8.0), and proteins were eluted (1 M NaCl, 200 mM imidazole, pH 8.0) and dialyzed (25 mM HEPES, 500 mM NaCl, pH 8.0). *fsdS* ACP-CATR was further purified by FPLC on an SD300 column attached to an AKTA purifier, using the dialysis buffer. Purified proteins were analyzed by SDS-PAGE and flash frozen at −80 °C for storage.

The enzymes were used in the pyrophosphate exchange assay following a previously published protocol.[14] Enzymes were used at 1 μM, and all 20 proteinogenic amino acids and *N*-methyl-L-serine were used at 2 mM. A control reaction was performed with no amino acid added. Experiments were carried out in duplicate, and background was subtracted using the no amino acid control.

Chemical Analysis of Fungal Transformants. HPLC was performed using a Hitachi LaChrom Elite System with diode array detection using the Agilent Eclipse XDB C_{18} column (4.6 mm × 150 mm, 5 μm) for analytical purposes and the Discovery HS C_{18} column (25 cm × 10 mm, 5 μm) for preparative purification. Analytical mass spectrometry was performed using the Agilent ZQ, while high-resolution FT-ICR experiments were performed at the University of Utah Mass Spectrometry Core Facility using the LTQ-FT (Thermo-Electron) instrument. NMR data was acquired on a Varian INOVA 500, except for an HMBC spectrum of **4** acquired on a Varian INOVA 600 equipped with a cryoprobe.

Verified transformants were grown on PDA (Difco) for 4 days to generate spores, which were harvested by suspending the spores in sterile water with an inoculating loop followed by filtration through a cotton plug.[10] Fresh spores were used to inoculate media for all fungal expression experiments. In all analytical experiments, wild-type strains were used as simultaneous controls.

Spores from knockout mutants were transferred to CGA[10] and incubated at RT for 21 days prior to harvesting. Cultures were harvested by extracting three times with acetone (200 mL per 100 g media). The acetone was dried by rotary evaporation, leaving an aqueous residual. For analytical HPLC and LC−MS, this residual was desalted over a plug of end-capped C_{18} resin and analyzed directly. To purify compound **3**, a 100 g CGA culture was used, and the residual water was acidified to pH 2.0 with H_2SO_4 (conc) and extracted four times with hexanes (100 mL).[37] The extract was dried by rotary evaporation, resuspended in boiling hexanes (100 mL), and filtered. The hexanes fraction was washed six times with ethanol (2 mL) and dried by rotary evaporation; 25% of the residue was further purified by flash chromatography on end-capped C_{18} using a methanol/H_2O gradient, with TLC analysis employing a ferric chloride stain. Fractions containing **3** (90% methanol elution) were dried by rotary evaporation to give known compound trichosetin **3** (6 mg):[19] ^1H and ^{13}C NMR, see Supplementary Table S7; HRMS, see Supplementary Figure S22.

Spores from overexpression mutants were transferred to PDB (Difco) and incubated for 18 h with shaking at 200 rpm at 30 °C. To induce *alcA*, cyclopentanone (30 mM final concentration) was added, and the cultures were incubated for 6 d. Cultures were filtered through MiraCloth and subsequently extracted. For *eqxF* and *eqxR* overexpression, the filtered broth (250 mL) was extracted with ethyl acetate (2 × 250 mL) containing 1% acetic acid. The solvent was removed by vacuum, and the dried residue was subjected to analytical HPLC. The amount of equisetin **1** in the crude extract of the Palc:eqxR culture was quantified by HPLC using a standard curve generated with pure equisetin. For *fsdS* and *fsdS-fsdC* overexpression experiments, mycelia were extracted with acetone (200 mL), and the dried residue after solvent removal was purified by flash chromatography on end-capped C_{18} using an acetonitrile/water gradient, with TLC analysis employing iodine on silica as a stain. One fraction containing **4** was dried by vacuum to give pure compound fusaridione A **4** (8 mg). The filtered broth (4 L) was extracted with ethyl acetate (2 × 2 L) acidified with 1% acetic acid. The separated organic layer was dried under vacuum, and a portion of the residue was subjected to analytical HPLC. Flash chromatography on end-capped C_{18} using a methanol/H_2O (0.1% TFA) gradient was employed to isolate **5**. Fractions containing **5** were pooled, and solvent was removed under vacuum. Final purification was done with several rounds of preparative HPLC to afford pure compound **5** (3 mg). For ^1H and ^{13}C NMR data of **4** and **5**, see Supplementary Table S8; for HRMS, see Supplementary Figures S23 and S24.

■ ASSOCIATED CONTENT

Supporting Information

This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

The *fsd* and *eqx* gene clusters were deposited in GenBank, accession numbers KC439347 and AY700570.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: ews1@utah.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was funded by NSF 0957791. We thank K. McCluskey (FGSC) for providing plasmids pBC-Phleo and pMT-BFP, J. Kwan for assistance with genome assembly, and J. Skalicky (University of Utah) for help with acquiring NMR data.

■ REFERENCES

(1) von Döhren, H. (2009) A survey of nonribosomal peptide synthetase (NRPS) genes in Aspergillus nidulans. *Fungal Genet. Biol. 46*, S45−S52.

(2) Chooi, Y. H., and Tang, Y. (2012) Navigating the fungal polyketide chemical space: from genes to molecules. *J. Org. Chem. 77*, 9933−9953.

(3) Boettger, D., and Hertweck, C. (2013) Molecular diversity sculpted by fungal PKS-NRPS hybrids. *ChemBioChem 14*, 28−42.

(4) Fisch, K. M., Bakeer, W., Yakasai, A. A., Song, Z., Pedrick, J., Wasil, Z., Bailey, A. M., Lazarus, C. M., Simpson, T. J., and Cox, R. J. (2011) Rational domain swaps decipher programming in fungal highly reducing polyketide synthases and resurrect an extinct metabolite. *J. Am. Chem. Soc. 133*, 16635−16641.

(5) Ames, B. D., Chi, N., Bruegger, J., Smith, P., Xu, W., Ma, S., Wong, E., Wong, S., Xie, X., Li, J. W. H., Vederas, J. C., Tang, Y., and Tsai, S.-C. (2012) Crystal structure and biochemical studies of the

trans-acting polyketide enoyl reductase LovC from lovastatin biosynthesis. *Proc. Natl. Acad. Sci. U.S.A. 109*, 11144−11149.
(6) 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., and Tang, Y. (2009) Complete reconstitution of a highly reducing iterative polyketide synthase. *Science 326*, 589−592.
(7) Sims, J. W., and Schmidt, E. W. (2008) Thioesterase-like role for fungal PKS-NRPS hybrid reductive domains. *J. Am. Chem. Soc. 130*, 11149−11155.
(8) Liu X Fau - Walsh, C. T., and Walsh, C. T. (2009) Cyclopiazonic acid biosynthesis in *Aspergillus* sp.: characterization of a reductase-like R* domain in cyclopiazonate synthetase that forms and releases cyclo-acetoacetyl-L-tryptophan. *Biochemistry 48*, 8746−8757.
(9) Maiya, S., Grundmann, A., Li, X., Li, S.-M., and Turner, G. (2007) Identification of a hybrid PKS/NRPS required for pseurotin A biosynthesis in the human pathogen *Aspergillus fumigatus*. *ChemBioChem 8*, 1736−1743.
(10) Sims, J. W., Fillmore, J. P., Warner, D. D., and Schmidt, E. W. (2005) Equisetin biosynthesis in *Fusarium heterosporum*. *Chem. Commun.*, 186−188.
(11) Ma, L.-J., van der Does, H. C., Borkovich, K. A., Coleman, J. J., Daboussi, M.-J., Di Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., Henrissat, B., Houterman, P. M., Kang, S., Shim, W.-B., Woloshuk, C., Xie, X., Xu, J.-R., Antoniw, J., Baker, S. E., Bluhm, B. H., Breakspear, A., Brown, D. W., Butchko, R. A. E., Chapman, S., Coulson, R., Coutinho, P. M., Danchin, E. G. J., Diener, A., Gale, L. R., Gardiner, D. M., Goff, S., Hammond-Kosack, K. E., Hilburn, K., Hua-Van, A., Jonkers, W., Kazan, K., Kodira, C. D., Koehrsen, M., Kumar, L., Lee, Y.-H., Li, L., Manners, J. M., Miranda-Saavedra, D., Mukherjee, M., Park, G., Park, J., Park, S.-Y., Proctor, R. H., Regev, A., Ruiz-Roldan, M. C., Sain, D., Sakthikumar, S., Sykes, S., Schwartz, D. C., Turgeon, B. G., Wapinski, I., Yoder, O., Young, S., Zeng, Q., Zhou, S., Galagan, J., Cuomo, C. A., Kistler, H. C., and Rep, M. (2010) Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature 464*, 367−373.
(12) Holt, C., and Yandell, M. (2011) MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinf. 12*, 491.
(13) Khaldi, N., Seifuddin, F. T., Turner, G., Haft, D., Nierman, W. C., Wolfe, K. H., and Fedorova, N. D. (2010) SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal Genet. Biol. 47*, 736−741.
(14) Mitchell, C. A., Shi, C., Aldrich, C. C., and Gulick, A. M. (2012) Structure of PA1221, a nonribosomal peptide synthetase containing adenylation and peptidyl carrier protein domains. *Biochemistry 51*, 3252−3263.
(15) Gwynne, D. I., Buxton, F. P., Sibley, S., Davies, R. W., Lockington, R. A., Scazzocchio, C., and Sealy-Lewis, H. M. (1987) Comparison of the cis-acting control regions of two coordinately controlled genes involved in ethanol utilization in *Aspergillus nidulans*. *Gene 51*, 205−216.
(16) Bergmann, S., Schuemann, J., Scherlach, K., Lange, C., Brakhage, A. A., and Hertweck, C. (2007) Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nat. Chem. Biol. 3*, 213−217.
(17) Ahuja, M., Chiang, Y.-M., Chang, S.-L., Praseuth, M. B., Entwistle, R., Sanchez, J. F., Lo, H.-C., Yeh, H.-H., Oakley, B. R., and Wang, C. C. C. (2012) Illuminating the diversity of aromatic polyketide synthases in *Aspergillus nidulans*. *J. Am. Chem. Soc. 134*, 8212−8221.
(18) Phillips, N. J., Goodwin, J. T., Fraiman, A., Cole, R. J., and Lynn, D. G. (1989) Characterization of the *Fusarium* toxin equisetin: the use of phenylboronates in structure assignment. *J. Am. Chem. Soc. 111*, 8223−8231.
(19) Marfori, E. C., Kajiyama, S., Fukusaki, E., and Kobayashi, A. (2002) Trichosetin, a novel tetramic acid antibiotic produced in dual culture of *Trichoderma harzianum* and *Catharanthus roseus* Callus. *Z. Naturforsch. C 57*, 465−470.
(20) Hazuda, D., Blau, C. U., Felock, P., Hastings, J., Pramanik, B., Wolfe, A., Bushman, F., Farnet, C., Goetz, M., Williams, M., Silverman, K., Lingham, R., and Singh, S. (1999) Isolation and characterization of novel human immunodeficiency virus integrase inhibitors from fungal metabolites. *Antivir. Chem. Chemother. 10*, 63−70.
(21) Marfey, P. (1984) Determination of D-amino acids. II. Use of a bifunctional reagent, 1,5-difluoro-2,4-dinitrobenzene. *Carlsberg Res. Commun. 49*, 591−596.
(22) Galagan, J. E., Calvo, S. E., Cuomo, C., Ma, L.-J., Wortman, J. R., Batzoglou, S., Lee, S.-I., Basturkmen, M., Spevak, C. C., Clutterbuck, J., Kapitonov, V., Jurka, J., Scazzocchio, C., Farman, M., Butler, J., Purcell, S., Harris, S., Braus, G. H., Draht, O., Busch, S., D'Enfert, C., Bouchier, C., Goldman, G. H., Bell-Pedersen, D., Griffiths-Jones, S., Doonan, J. H., Yu, J., Vienken, K., Pain, A., Freitag, M., Selker, E. U., Archer, D. B., Penalva, M. A., Oakley, B. R., Momany, M., Tanaka, T., Kumagai, T., Asai, K., Machida, M., Nierman, W. C., Denning, D. W., Caddick, M., Hynes, M., Paoletti, M., Fischer, R., Miller, B., Dyer, P., Sachs, M. S., Osmani, S. A., and Birren, B. W. (2005) Sequencing of Aspergillus nidulans and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature 438*, 1105−1115.
(23) Evans, B. S., Robinson, S. J., and Kelleher, N. L. (2011) Surveys of non-ribosomal peptide and polyketide assembly lines in fungi and prospects for their analysis *in vitro* and *in vivo*. *Fungal Genet. Biol. 48*, 49−61.
(24) Brakhage, A. A. (2013) Regulation of fungal secondary metabolism. *Nat. Rev. Microbiol. 11*, 21−32.
(25) Brakhage, A. A., and Schroeckh, V. (2011) Fungal secondary metabolites - strategies to activate silent gene clusters. *Fungal Genet. Biol. 48*, 15−22.
(26) Boettger, D., Bergmann, H., Kuehn, B., Shelest, E., and Hertweck, C. (2012) Evolutionary imprint of catalytic domains in fungal PKS-NRPS hybrids. *ChemBioChem 13*, 2363−2373.
(27) Bergmann, S., Funk, A. N., Scherlach, K., Schroeckh, V., Shelest, E., Horn, U., Hertweck, C., and Brakhage, A. A. (2010) Activation of a silent fungal polyketide biosynthesis pathway through regulatory cross talk with a cryptic nonribosomal peptide synthetase gene cluster. *Appl. Environ. Microbiol. 76*, 8143−8149.
(28) Heneghan, M. N., Yakasai, A. A., Williams, K., Kadir, K. A., Wasil, Z., Bakeer, W., Fisch, K. M., Bailey, A. M., Simpson, T. J., Cox, R. J., and Lazarus, C. M. (2011) The programming role of trans-acting enoyl reductases during the biosynthesis of highly reduced fungal polyketides. *Chem. Sci. 2*, 972−980.
(29) Schmidt, K., Riese, U., Li, Z. Z., and Hamburger, M. (2003) Novel tetramic acids and pyridone alkaloids, militarinones B, C, and D, from the insect pathogenic fungus *Paecilomyces militaris*. *J. Nat. Prod. 66*, 378−383.
(30) Dieckmann, v. W. (1901) Ueber cyklische beta-ketoncarbon-saureester. *Justus Liebigs Ann. Chem. 317*, 27−109.
(31) Gietz, R. D., Schiestl, R. H., Willems, A. R., and Woods, R. A. (1995) Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast 11*, 355−360.
(32) Bitoun, R., and Zamir, A. (1986) Spontaneous amplification of yeast CEN ARS plasmids. *Mol. Gen. Genet. 204*, 98−102.
(33) Punt, P. J., Oliver, R. P., Dingemanse, M. A., Pouwels, P. H., and van den Hondel, C. A. (1987) Transformation of *Aspergillus* based on the hygromycin B resistance marker from *Escherichia coli*. *Gene 56*, 117−124.
(34) Austin, B., Hall, R. M., and Tyler, B. M. (1990) Optimized vectors and selection for transformation of *Neurospora crassa* and *Aspergillus nidulans* to bleomycin and phleomycin resistance. *Gene 93*, 157−162.
(35) Zerbino, D. R., McEwen, G. K., Margulies, E. H., and Birney, E. (2009) Pebble and Rock Band: heuristic resolution of repeats and scaffolding in the Velvet short-read *de novo* assembler. *PLoS ONE 4*, e8407.
(36) Stanke, M., Schoffmann, O., Morgenstern, B., and Waack, S. (2006) Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinf. 7*, 62.

(37) Burmeister, H. R., Bennett, G. A., Vesonder, R. F., and Hesseltine, C. W. (1974) Antibiotic produced by *Fusarium equiseti* NRRL 5537. *Antimicrob. Agents Chemother. 5*, 634−639.

Supporting Information

Two related pyrrolidinedione synthetase loci in *Fusarium heterosporum* ATCC 74349 produce divergent metabolites

Thomas B. Kakule, Debosmita Sardar, Zhenjian Lin, and Eric W. Schmidt

Department of Medicinal Chemistry, University of Utah, Salt Lake City, UT 84112 USA

Table of Contents

Description of Vector Synthesis

PCR

The general PCR conditions for amplifying the gene pieces were as follows: each primer (2 μL, 20 pmol), HiFi Buffer (2 μL), $MgSO_4$ (0.6 μL), dNTPs (0.8 μL, 2.5 mM each), genomic DNA (1 μL, 100 ng), HiFi Platinum Taq (0.2 μL), and water to 20 μL. Cycles: 95 °C/5 min, (95 °C/30 s, 58 °C/30 s, 68 °C/2.5 min)x34, 68 °C/5 min. Amplicons were purified from a 1% agarose gel with the Millipore DA kit.

All vectors were Sanger sequenced over *F. heterosporum* genes to confirm integrity of vectors.

Knockout cassettes

fsdS. The plasmid pBC-Phleo (FGSC) was digested with NheI, XbaI and DraI, and separated on 1% agarose gel. The 6.4 kb piece containing the sh_ble gene was extracted from the agarose and ligated as insert into AvrII-linearized F4-2.5-pCR2.1 plasmid.[1] PCR screening of colonies was done with primer pair KOPhleoScreen1-F and KOPhleoScreen2-R.

eqxC and eqxD. Homologous ends (approx. 2 kbp) were amplified from *F. heterosporum* genomic DNA with primer pairs pRSMTKO-F/MTKO1-R and MTKO2-F/pRSMTKO-R for *eqxD* and pRSOxoKO-F/OxoKO1-R and OxoKO2-F/pRSOxoKO-R for *eqxC*. The hph gene for the knockout plasmids was PCR-amplified from pHygB-EqiACATR with primer pair KOfuseHygB-F/KO-fuseHygB-R.

Fusion PCR was done in two steps: For PCR#1, 3 μL of each piece (KO1, KO2, and KOfuseHygB) were added to the standard PCR mix described above. Cycles: 95C/5min, (95 °C/30 s, 68 °C/7.5 min)x15, 68 °C/15 min. For PCR#2, 1 μL from PCR#1 was used as template and the number of cycles extended to 34. The knockout fusion cassettes were cloned into the pCR 2.1 TOPO vector (Invitrogen) using the supplier's instructions to make plasmids TOPO-MTKO and TOPO-OxoKO.

***E. coli* expression plasmids**

eqxS ATR. *F. heterosporum* DNA was amplified with primer pair ATR2express-F and ATR2express-R. *S. cerevisiae* BY4741 was transformed with both the amplicon and the BamHI/KpnI-linearized pHygB-EqiACATR plasmid for recombination. Plasmids were screened in TOP10 cells (Invitrogen) and verified by restriction digest.

eqxS A. *F. heterosporum* DNA was amplified with primer pair ATR2express-F and Aden2express-R. As described above, yeast recombination was utilized to construct the plasmid by co-transformation of the amplicon with BamHI/KpnI-linearized pHygB-EqiACATR.

fsdS ACP-CATR. ATR was cloned as previously described.[2] This piece was recombined in yeast with the *fsdS* ACP-C region and N-terminally His-tagged to create plasmid deg2ACP-CATR plasmid.

Fungal expression plasmids

eqxF (alcAeqxF). The *alcA* promoter was amplified from pMT-BFP plasmid (FGSC) with primer pair: alcA2exphleo-F/alcA2exphleo-R. The transcription factor, *eqxF*, was amplified from *F. heterosporum* DNA with primer pair alcA2g10518-F/alcA2g10518-R. Yeast recombination was utilized to construct the expression plasmid by transformation of *S. cerevisiae* BY4741 with both amplicons above and the 7 kbp piece of SacII/AscI-digest of the Exphleo plasmid. The Exphleo plasmid had been previously cloned in our lab to have both the sh_ble and hph selection markers to allow replacement of either marker with the gene of interest for expression driven by the *Aspergillus nidulans gpdA* promoter. The alcAeqxF plasmid was transformed into *F. heterosporum* WT with selection on phleomcyin (150 μg mL^{-1}), and the mutant, Palc:eqxF, was chosen for expression analysis.

eqxR (alcAeqxR). The gene was amplified from *F. heterosporum* DNA with primer pair alcA2g10517-F/alcA2g10517-R. The alcAeqxR plasmid above was digested with SacII and BamHI, and the 9.1 kbp piece used for yeast recombination with the *eqxR* gene amplicon. The plasmid was transformed into *F. heterosporum* WT with phleomcyin selection, and the mutant, Palc:eqxR, was chosen for expression analysis.

fsdC (alcAfsdC). The gene was amplified with primer pair alcA-deg2_ER-F/alcA-deg2_ER-R. Yeast recombination was used to construct the complete plasmid by co-transformation of amplicon with the 9.1 kbp piece from the BamHI/SacII-digest of the alcAeqxF plasmid.

fsdS (alcAfsdS). The *alcA* promoter sequence was amplified from plasmid pMT-BFP (FGSC) with primer pair alcA-F/alcA-R, and the *fsdS* PKS region from genomic DNA with primer pair alcA2deg2-F/EqiPKS-3ex-r. Previously in our lab, the deg2 NRPS region had been cloned into a pRS316-derived plasmid to make pHygB-deg2ACATR. The pHygB-deg2ACATR plasmid was linearized with PmeI, and the linear DNA was co-transformed with above amplicons into yeast BY4741 for recombination to make the alcAfsdS expression plasmid.

Supplementary Tables and Figures

Table S1: Primers used in this study.

Primer	Sequence
pRSOxoKO-F	GCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATGGCTGATGAACG AACCTGGC
pRSOxoKO-R	GAAACAGCTATGACCATGATTACGCCAAGCTCGGAATTCCATAGATCAGC AGCAGCAAG
pRSMTKO-F	GCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATGCTGTTTCAGTG AATCCAGTC
pRSMTKO-R	GAAACAGCTATGACCATGATTACGCCAAGCTCGGAATTGCATCAACAATA GTGCCTGGAA
KOfuseHygB-F	CTAGCAGTTCCAGGTGGAATG
KOfuseHygB-R	CGTCTAGAAAGAAGGATTACCTC
KOPhleoScreen1-F	GCTGACCAGCTCAAGTATCTC
KOPhleoScreen2-R	GGCAGCTTCTGACAAGCGAG
KOxo1-F	TCATTAAGAGCCAGTTCATGGGCGTTGGCATGATGGCCGTGGCTGATGAA CGAACCTGGC
KOxo1-R	ATACAATGCTCATCATAACATTCCACCTGGAACTGCTAGCCAGTCTGACG AGCTGAATG
KOxo2-F	ACAGGTACACTTGTTTAGAGGTAATCCTTCTTTCTAGACGGTTCGCACCAA TCACAACC
KOxo2-R	GGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGCCATAGATCAG CAGCAGCAAG
KOMT1-F	TCATTAAGAGCCAGTTCATGGGCGTTGGCATGATGGCCGTGCTGTTTCAG TGAATCCAGTC
KOMT1-R	AATACAATGCTCATCATAACATTCCACCTGGAACTGCTAGGATCAGTACC AGCGTGTGC
KOMT2-F	ACAGGTACACTTGTTTAGAGGTAATCCTTCTTTCTAGAGCCATAATTGAAG AGACTCCGA
KOMT2-R	GGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCATCAACAATAGT GCCTGGAA
OxoKOScreen-F	GCTTCATGACGACGGGGTT
OxoKOScreen-R	CCGAGTAAGGCTCAAGAGC
MTKOScreen-F	GCCTCAGTATTTACAGACCAATG
MTKOScreen-R	CATCAGTCATAACTTTCACGGTG
ATR2express-F	CACCATCATCACCACAGCCAGGTAGAGAATCTTTATTTTCAGGGCGCACC CGATTATGAATCTAAATG
ATR2express-R	TTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTCAACGCCTTGAAAATA GCTGTG
Aden2express-R	GCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTACGACTGAGTTGTTTGC ACGG
alcA2exphleo-F	ACCGAGATAGGGTTGAGTGTTGGCGCGCCGTTTAAACTTAATTAATCTGC GATCGTCCATAACCG
alcA2exphleo-R	TCTTGGACCTGATCCGCCATTGTTGATTGTTGGTTGATGATGTTG
alcA2g10518-F	CAACATCATCAACCAACAATCAACAATGGCGGATCAGGTCCAAGA
alcA2g10518-R	AACGCGTTTTATTCTTGTTGACATGGAGCTATTAAATCAAAGACCCCATCG CATCTG
alcA2g10517-F	TCTCAGTCTCACCAACATCATCAACCAACAATCAACAATGTCAACCCAAA

	ACAACCAATC
alcA2g10517-R	CGTTTTATTCTTGTTGACATGGAGCTATTAAATCATTAAAGCTTCATCCTT TCCCTTATT
alcAdeg2_ER-F	TCTCAGTCTCACCAACATCATCAACCAACAATCAACAATGGATCGTCCCA ACAAAGCAAT
alcAdeg2_ER-R	CGCGTTTTATTCTTGTTGACATGGAGCTATTAAATCATTAGACGCGATGGT CAATGACTA
alcA-F	ACTTGTTTAGAGGTAATCCTTCTTTCTAGACGGCGGCCGCTCTGCGATCGT CCATAACCG
alcA-R	GCTTTGTAGGAGGGGCAGACATTGTTGATTGTTGGTTGATGATGTTG
alcA2deg2-F	CAACATCATCAACCAACAATCAACAATGTCTGCCCCTCCTACAAAGC
EqiPKS-3ex-r	GACAGTCTTGAGCTGCACTCGC

Table S2: Plasmids used in this study.

Plasmids	Source
pMT-BFP	FGSC[3]
Exphleo	This study
pHygB-EqiACATR	This study
TOPO-MTKO	This study
TOPO-OxoKO	This study
pBC-Phleo	FGSC
alcAfsdS	This study
alcAfsdC	This study
deg2ACP-CATR	This study
deg3A-domain	This study
deg3ATR	This study
alcAeqxF	This study
alcAeqxR	This study
F4-2.5-pCR2.1	Sims J.W. et al.[1]

Table S3: Statistics of the assembled genome of *Fusarium heterosporum* ATCC 74349.

Stats	Contigs	
	Number	Length (bp)
N25	3	2204072
N50	8	1332172
N75	17	739251
Largest	-	6929503
Total size	-	39356893

Table S4: Blastx analysis of *F. heterosporum* using residues 9–445 of lovastatin nonaketide synthase (AAD39830.1) as the query identified *deg1*, *deg2*, and *deg3* as homologues.

Designated gene names	Score(Bits)	E-value
deg3/*eqxS*	126	2e-029
deg2 /*eqiS*/*fsdS*	108	4e-024
deg1	67.4	8e-012

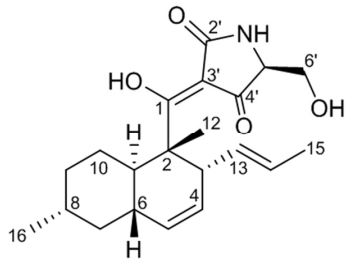
Table S5: BLASTp analysis of *fsd* cluster.

Gene	Deduced function	Closest relative			Other relative		
		Genbank ID, proposed function	Identities/ positives (%)	Organism	Genbank ID, proposed function	Identities/ positives (%)	Organism
fsdS	PKS-NRPS	EGU88865.1, hypothetical protein FOXB_00609	47/65	*Fusarium oxysporum* Fo5176	XP_003022161.1, hybrid PKS-NRPS enzyme	44/62	*Trichophyton verrucosum* HKI 0517
fsdC	enoyl Reductase	EKJ67665.1, hypothetical protein FPSE_12182	77/88	*Fusarium pseudograminearum* CS3096	EKG12691.1, Alcohol dehydrogenase superfamily zinc-containing	33/50	*Macrophomina phaseolina* MS6
fsdD	methyl transferase	EFY96249.1, putative methyltransferase	46/60	*Metarhizium anisopliae* ARSEF 23	XP_002486598.1, methyltransferase SirN-like, putative	34/49	*Talaromyces stipitatus* ATCC 10500
fsdG	exporter	XP_001269052.1, MFS transporter	50/66	*Aspergillus clavatus* NRRL 1	XP_002842733.1, fluconazole resistance protein 1	47/65	*Arthroderma otae* CBS 113480
fsdH	oxidase	XP_001904540.1, hypothetical protein	53/69	*Podospora anserina* S mat+	XP_003716001.1, cytochrome P450 52A13	50/68	*Magnaporthe oryzae* 70-15
fsdK	dimethyl allyl transferase	XP_002149623.1, dimethylallyl tryptophan synthase GliD1	43/58	*Penicillium marneffei* ATCC 18224	EFQ36406.1, aromatic prenyltransferase	35/55	*Glomerella graminicola* M1.001
fsdR	regulator	CCF39140.1, fungal specific transcription factor	41/58	*Colletotrichum higginsianum*	XP_003010764.1, C6 transcription factor	33/51	*Arthroderma benhamiae* CBS 112371

Table S6: BLASTp analysis of *eqx* cluster.

Gene	Deduced function	Closest relative			Other relative		
		Genbank ID, proposed function	Identities/ positives (%)	Organism	Genbank ID, proposed function	Identities/ positives (%)	Organism
eqxS	PKS-NRPS	EGU88865.1, hypothetical protein FOXB_00609	79/88	*Fusarium oxysporum* Fo5176	XP_003010822.1, hybrid PKS-NRPS enzyme	43/62	*Arthroderma benhamiae* CBS 112371
eqxC	enoyl reductase	EGU88867.1, hypothetical protein FOXB_00611	87/92	*Fusarium oxysporum* Fo5176	AAD34554.1, enoyl reductase	44/60	*Aspergillus terreus*
eqxD	methyl transferase	XP_001394032.1, O-methyltransferase	48/69	*Aspergillus niger* CBS 513.88	EKV06090.1, O-methyltransferase	35/57	*Penicillium digitatum* PHI26
eqxF	regulator	EGU88869.1, hypothetical protein FOXB_00613	67/74	*Fusarium oxysporum* Fo5176	XP_003172084.1, fungal specific transcription factor domain-containing protein	54/70	*Arthroderma gypseum* CBS 118893
eqxG	exporter	EGU88870.1, hypothetical protein FOXB_00614	79/85	*Fusarium oxysporum* Fo5176	EGE05778.1, MFS multidrug transporter	60/77	*Trichophyton equinum* CBS 127.97
eqxH	oxidase	EGU87157.1, hypothetical protein FOXB_02335	69/83	*Fusarium oxysporum* Fo5176	CCF39176.1, cytochrome P450	54/71	*Colletotrichum higginsianum*
eqxR	regulator	EGU88868.1, hypothetical protein FOXB_00612	53/65	*Fusarium oxysporum* Fo5176	XP_001268499.1, C6 zinc finger domain protein	22/39	*Aspergillus clavatus* NRRL 1

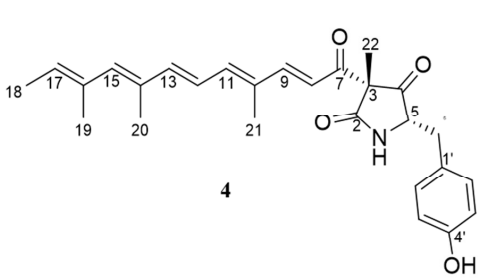
Table S7: NMR data of compound 3. These shifts (including peaks that are not detected) are almost identical to previously reported data for trichosetin.[4]



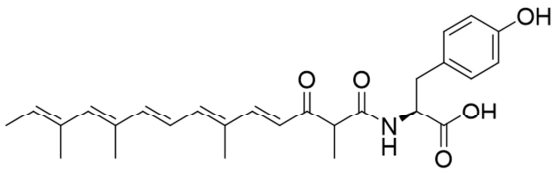
no.	NMR data of 3		Trichosetin data reported by Marfori E.C. et al.[4]	
	δ ^{13}C	**δ ^{1}H, multiplicity, *J* (Hz)**	**δ ^{13}C**	**δ ^{1}H**
1	204.6, C		201.5	
2	51.5, C		50.0	
3	46.2, CH	3.46, br	46.2	3.43
4	133.4, CH	5.15, m	132.3	5.23
5	132.6, CH	5.41, m	131.2	5.44
6	39.9, CH	1.86, m	39.8	1.86
7	43.6, CH_2	1.89, m; 0.89, m	43.4	1.83, 0.86
8	34.8, CH	1.53, m	34.8	1.49
9	36.9, CH_2	1.77, m; 1.11, m	36.9	1.78, 1.10
10	29.5, CH_2	2.01, br; 1.08, br	29.2	2.02, 1.07
11	41.3, CH	1.68, m	41.0	1.64
12	14.2, CH_3	1.45, s	13.7	1.42
13	128.0, CH	5.37, m	127.6	5.38
14	127.9, CH	5.26, m	127.9	5.14
15	17.8, CH_3	1.52, d, 5.7	18.6	1.56
16	23.0, CH_3	0.94, d, 6.8	23.2	0.94
2'	nd[a]		180.9*	
3'	nd		100.7*	
4'	nd		192.9*	
5'	nd	nd	64.5	3.76
6'	62.2, CH_2	3.79,m; 3.74, m	61.9	3.82, 3.79

[a]nd = not detected
*** detected at –80 °C**

Table S8: NMR data of 4 and 5. Data for the major tautomer is tabulated.



4



5

no.	Compound 4		Compound 5	
	δ ^{13}C	δ ^{1}H, multiplicity, *J* (Hz)	δ ^{13}C	δ ^{1}H, multiplicity, *J* (Hz)
1-NH				6.96, s
2	172.0, C		172.6, C	
3	64.8, C		52.5, CH	3.68, m
4	207.1, C		170.9, C	
5	62.9, CH	4.29, m	54.4, CH	4.54, m
6	36.5, CH2	3.06, dd, 13.8, 3.1; 2.84, dd, 13.8, 8.0	36.6, CH2	3.07, m; 2.90, m
7	189.8, C		196.6, C	
8	118.6, CH	6.20, m	123.0, CH	6.29, m
9	148.7, CH	7.32, d, 15.1	148.3, CH	7.31, d, 15.5
10	133.5, C		134.1, C	
11	143.0, CH	6.68, m	141.9, CH	6.67, m
12	123.1, CH	6.65, m	nd	6.52, m
13	145.0, CH	6.60, m	144.5, CH	6.87, m
14	133.1, C		133.7, C	
15	139.8, CH	6.15, m	139.2, CH	6.18, s
16	133.7, C		134.4, C	
17	128.0, CH	5.63, m	128.3, CH	5.59, m
18	12.8, CH3	1.75, d, 5.7	13.7, CH3	1.74, d, 6.9
19	15.1, CH3	1.86, s	16.4, CH3	1.84, s
20	12.4, CH3	2.02, s	13.8, CH3	2.02, s
21	11.2, CH3	1.92, s	12.5, CH3	1.93, s
22	16.1, CH3	1.46, s	14.0, CH3	1.20, s
1'	127.2, C		127.2, C	
2'	130.4, CH	7.06, d, 7.7	131.2, CH	7.04, d, 7.7
3'	114.8, CH	6.76, d, 7.7	115.8, CH	6.72, d, 7.7
4'	156.2, C		156.4, C	

Figure S1: HPLC analysis of crude extracts of fsdS knockouts. Both the fsdS gene knockout generated in this study, FusKO474, and the previously reported knockout, JWS19,[1] produce abundant equisetin **1** when cultured on corn grit agar (CGA) for 21 d. Bottom trace shows expression of wild-type *F. heterosporum*.

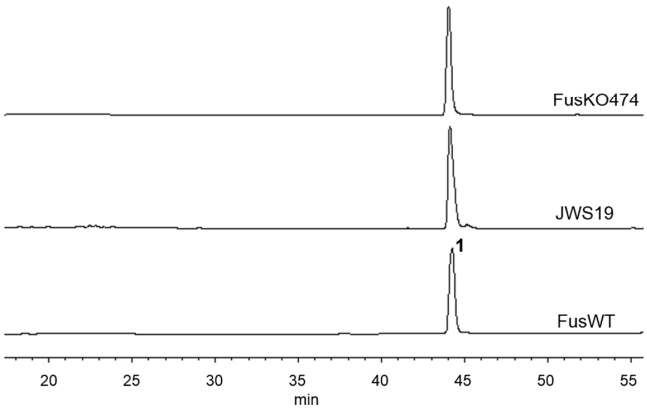


Figure S2: PCR analysis of knockouts. A) PCR on genomic DNA of the FusΔeqxD mutant using primer pair MTKOScreen-F/MTKOScreen-R showing a 400 bp amplicon for FusWT in the middle lane, whereas only the larger 3 kbp piece is amplified from the FusΔeqxD mutant in the right lane. B) PCR on genomic DNA of the FusΔeqxC mutant using primer pair OxoKOScreen-F/OxoKOScreen-R showing a 350 bp amplicon for FusWT in the middle lane, whereas only the larger 3 kbp piece is amplified from the FusΔeqxC mutant in the right lane.

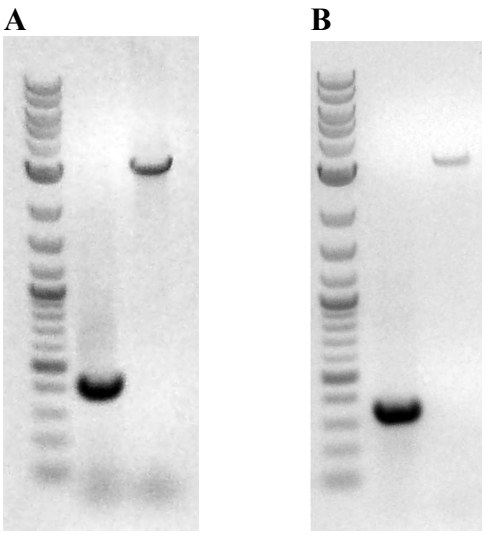


Figure S3A: Pyrophosphate exchange assay with His-FsdS ACP-CATR monomer fraction

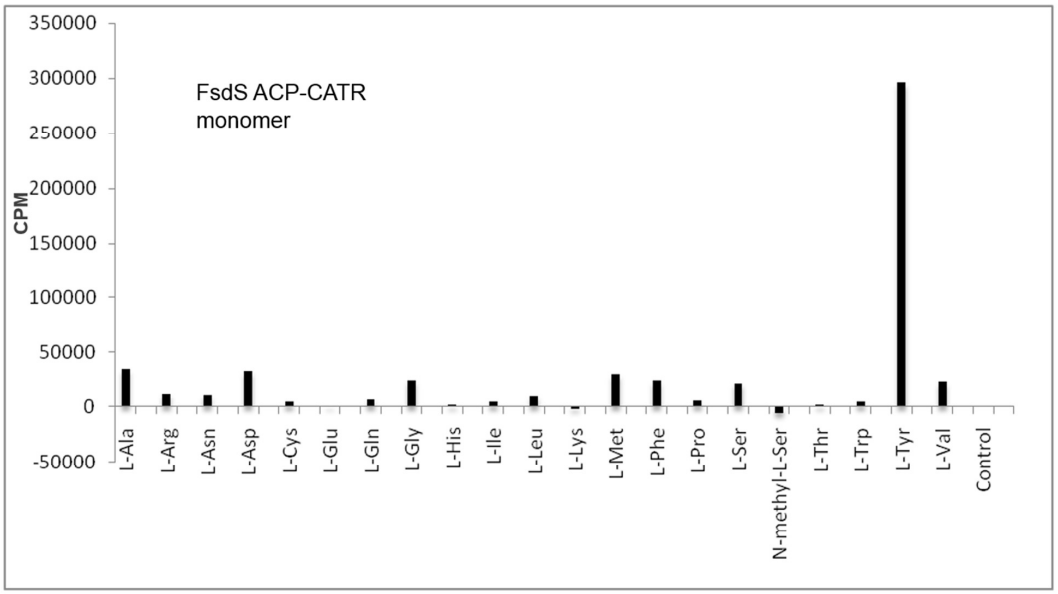


Figure S3B: Pyrophosphate exchange assay with His-FsdS ACP-CATR dimer fraction.

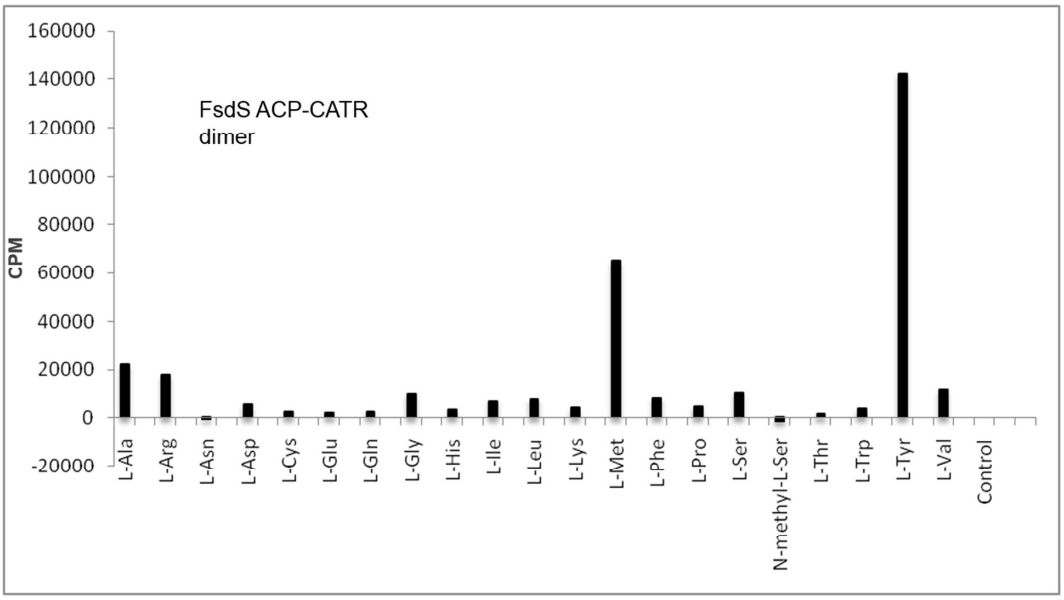


Figure S3C: Pyrophosphate exchange assay with His-EqxS A-domain.

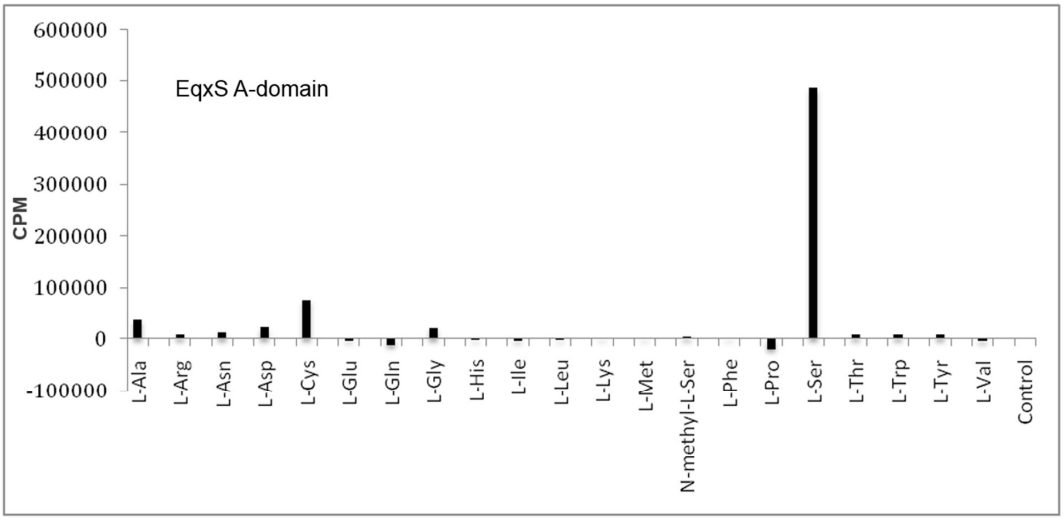


Figure S3D: Pyrophosphate exchange assay with His-EqxS ATR.

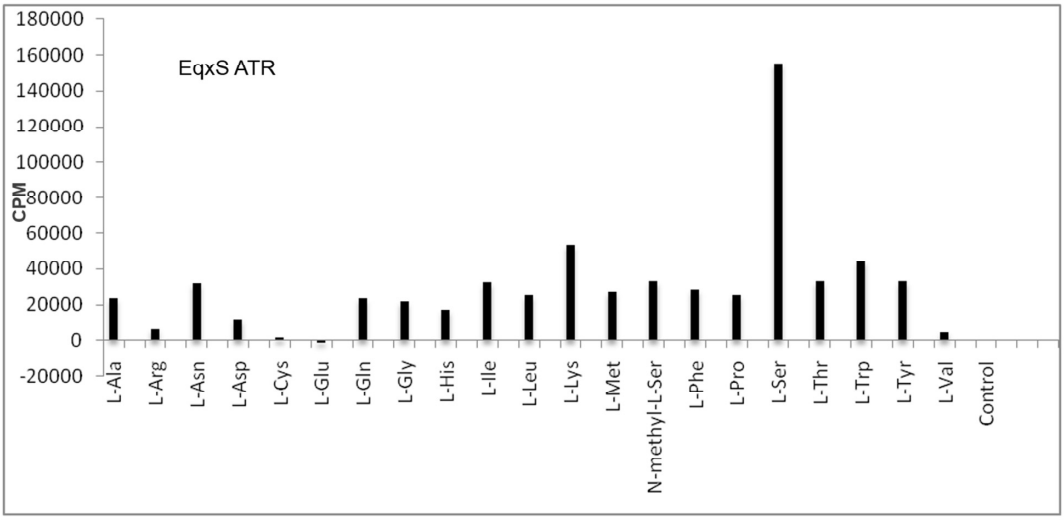


Figure S4: Picture showing impact of yellow pigment production by the Palc:fsdSandC mutant on colony color (grown on PDA). *Fusarium heterosporum* wild-type growth is shown on the right.

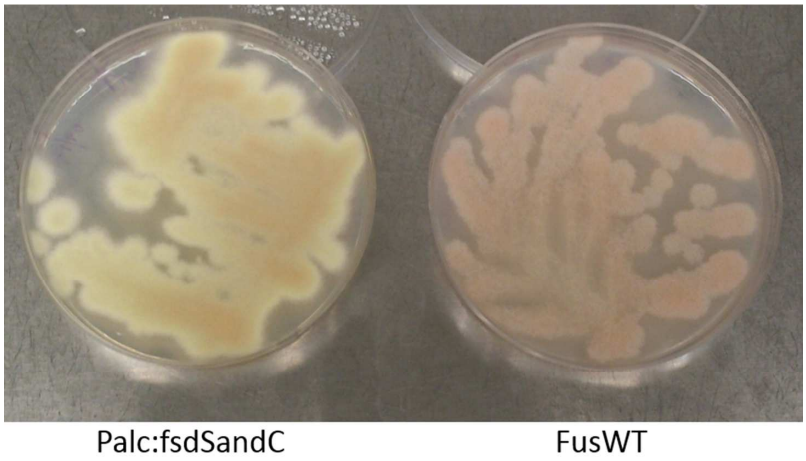
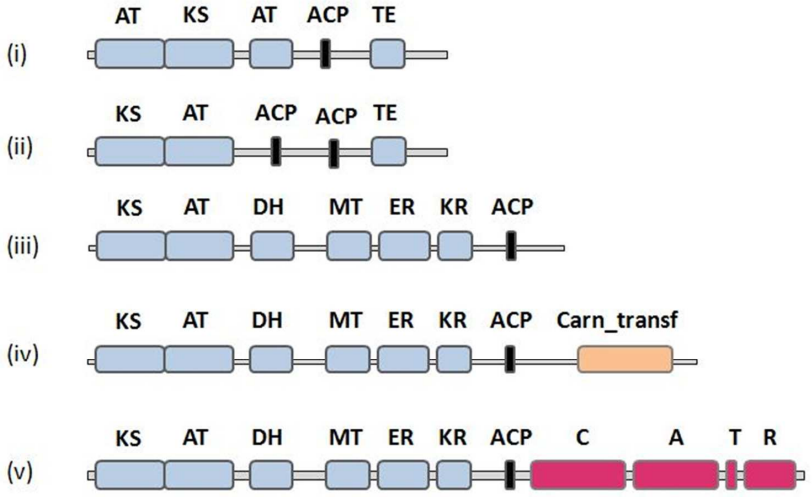


Figure S5: Domain structure of PKS-containing proteins of *F. heterosporum* as identified by BLASTp.



Two PKS proteins were identified to have the non-reducing domain structure (i) and (ii). From the reducing-type PKS group, six contained the domains shown in (iii), and two had a carnitine transferase-like domain after the ACP (iv). Additionally, two PKS-NRPS hybrid proteins (v) were identified and named EqxS and FsdS. Not included in this list are two other AT-containing proteins that were truncated, lacking some domains required for functional PKS enzymes.

Figure S6: 1D ^{1}H spectrum of 3 in CD_3OD

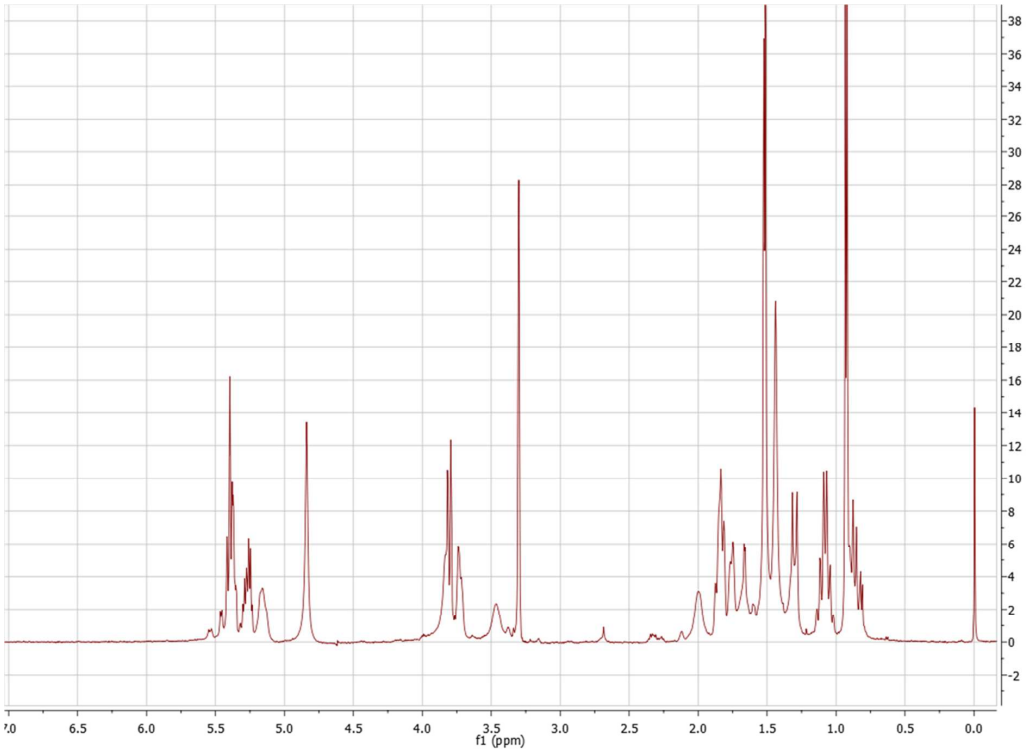


Figure S7: 1D ^{13}C spectrum of 3 in CD_3OD. Compare to an essentially identical ^{13}C spectrum previously reported for trichosetin.[4]

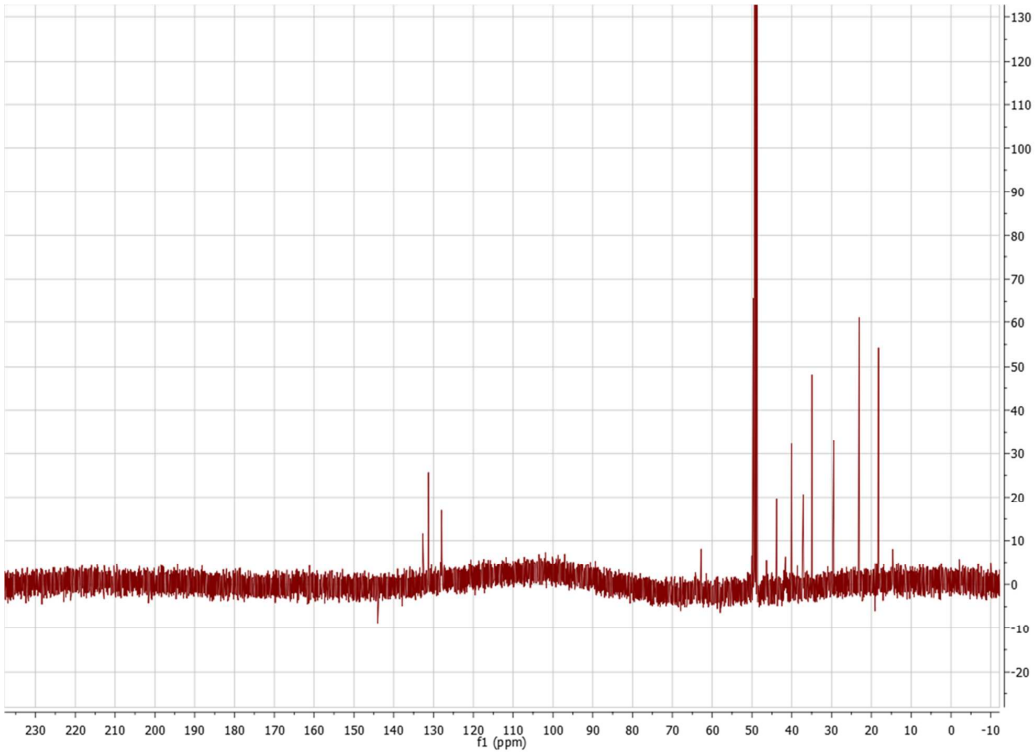


Figure S8: ^{1}H-^{13}C HSQC spectrum of 3 in CD_3OD.

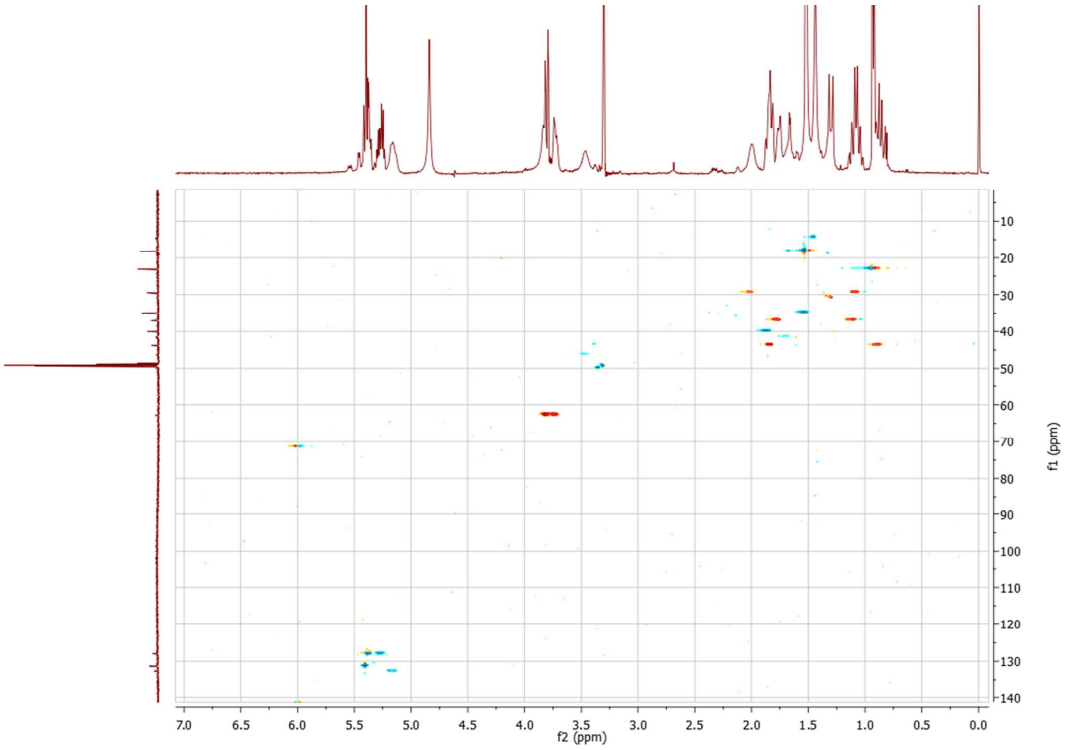


Figure S9: ^{1}H-^{13}C HMBC spectrum of 3 in CD_3OD.

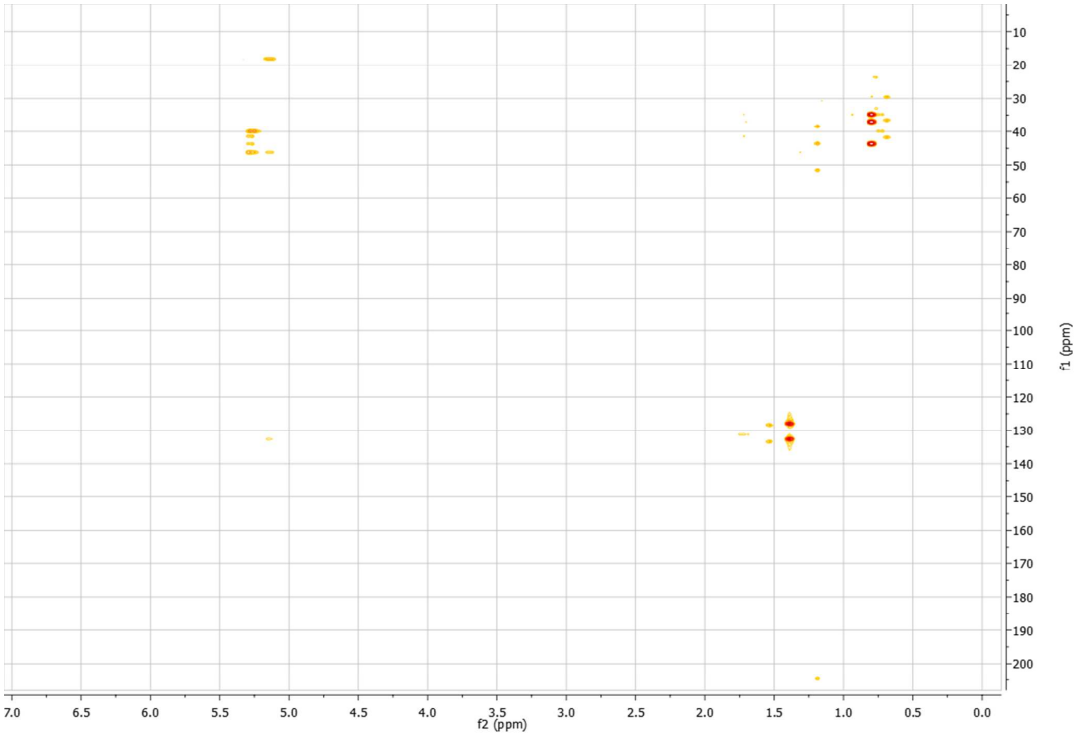


Figure S10: ^{1}H-^{1}H COSY spectrum of 3 in CD_3OD.

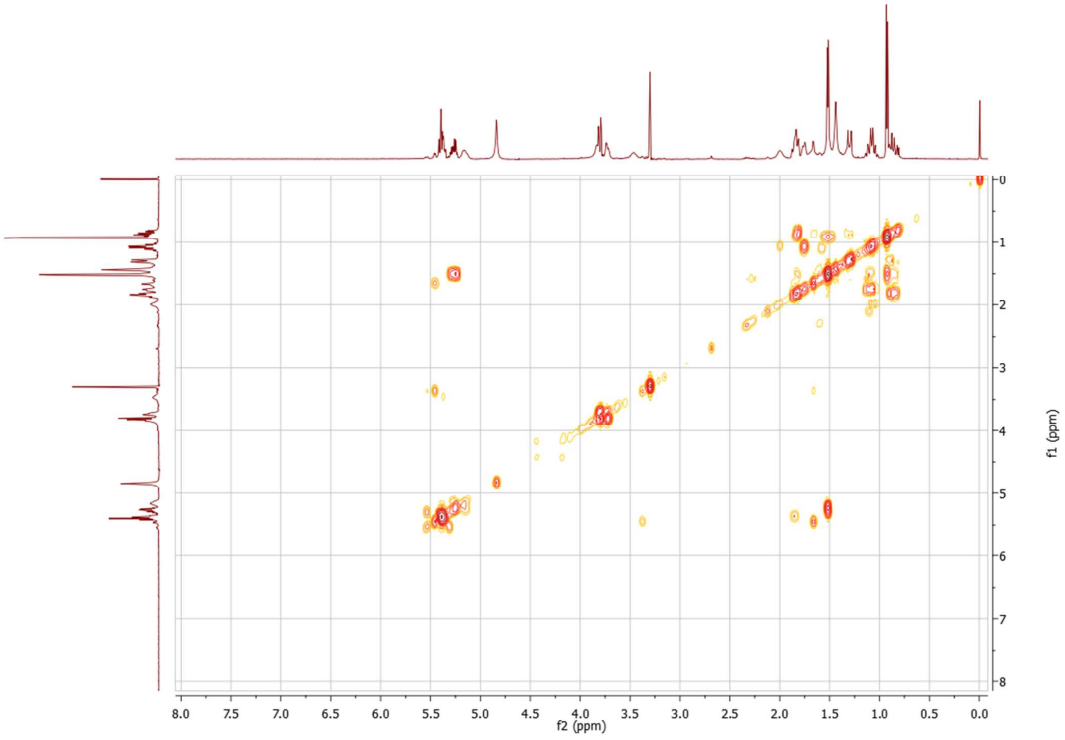


Figure S11: 1D ^{1}H spectrum of 4 in acetonitrile-d_3.

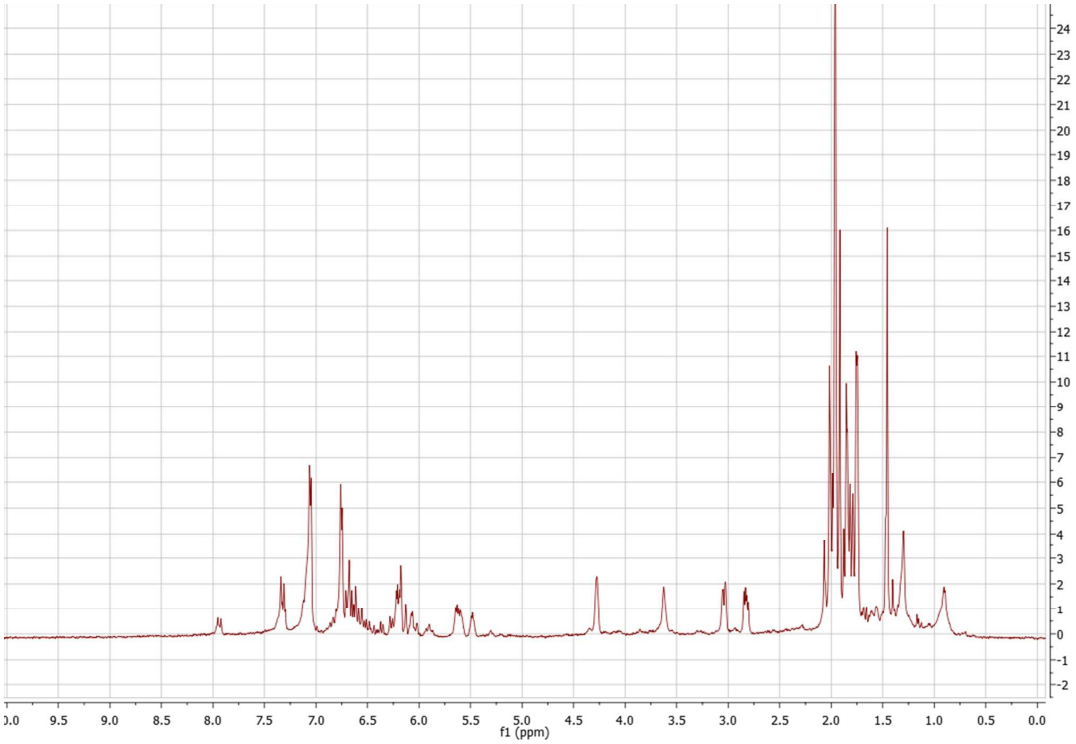


Figure S12: 1D ^{13}C spectrum of 4 in acetonitrile-*d3*

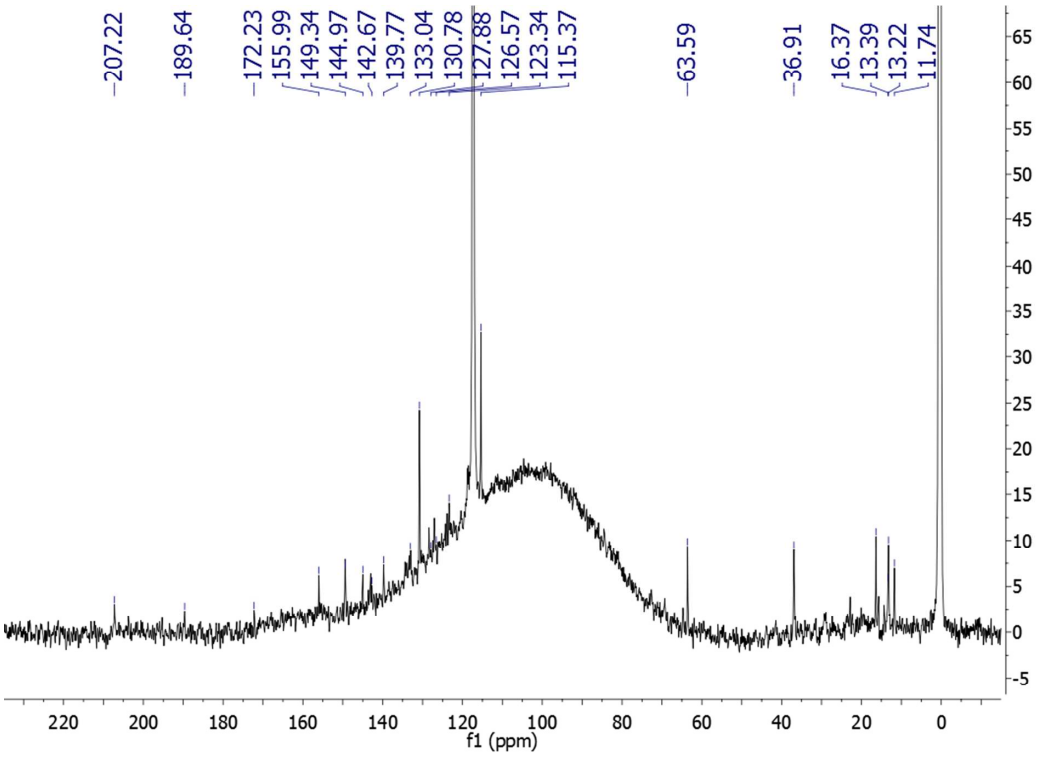


Figure S13: ^{1}H-^{1}H COSY spectrum of 4 in acetonitrile-d_3.

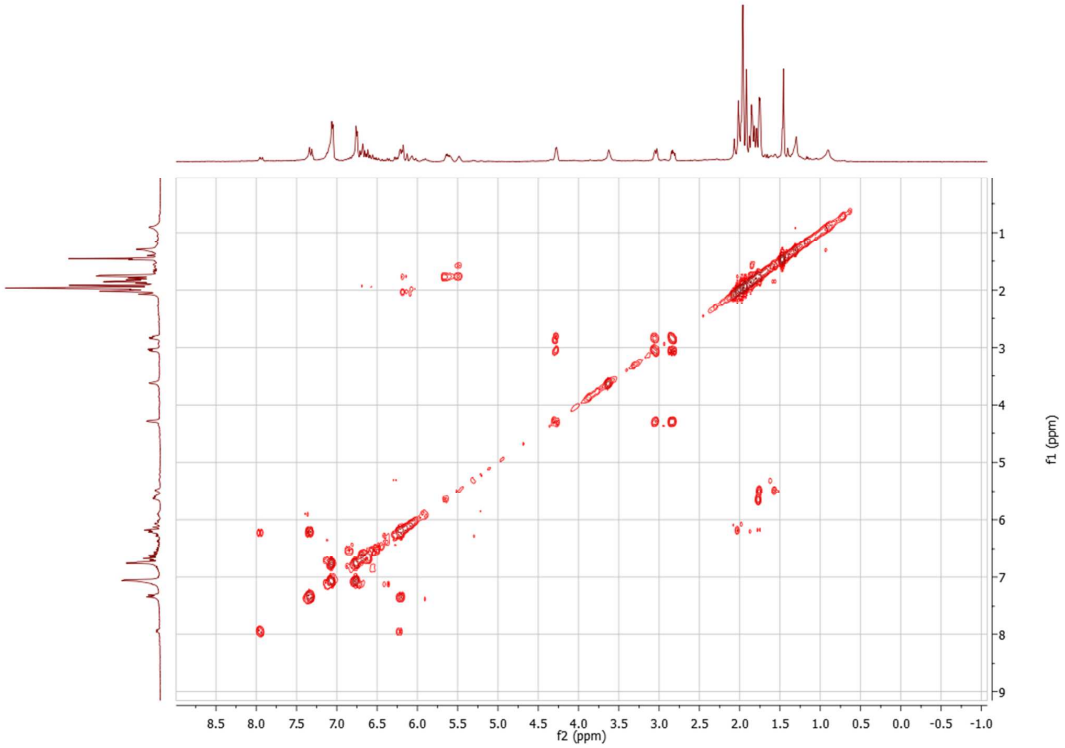


Figure S14: ^{1}H-^{1}H TOCSY spectrum of 4 in acetonitrile-d_3.

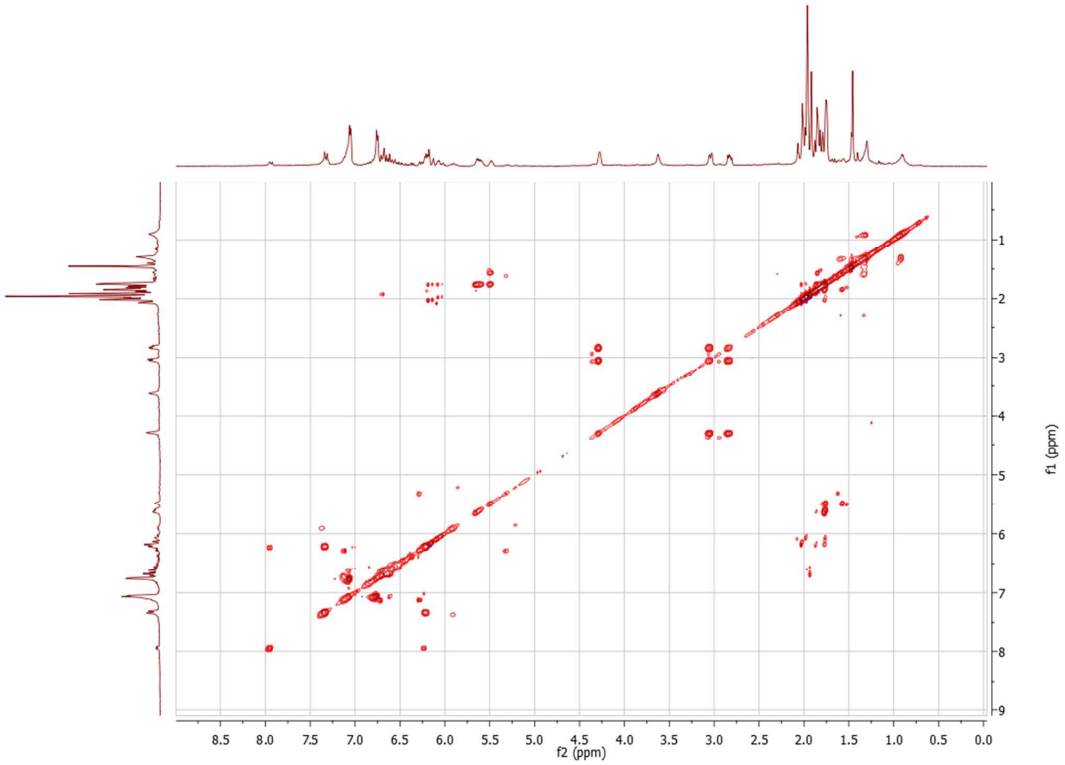


Figure S15: ^{1}H-^{13}C HSQC spectrum of 4 in acetonitrile-d_3.

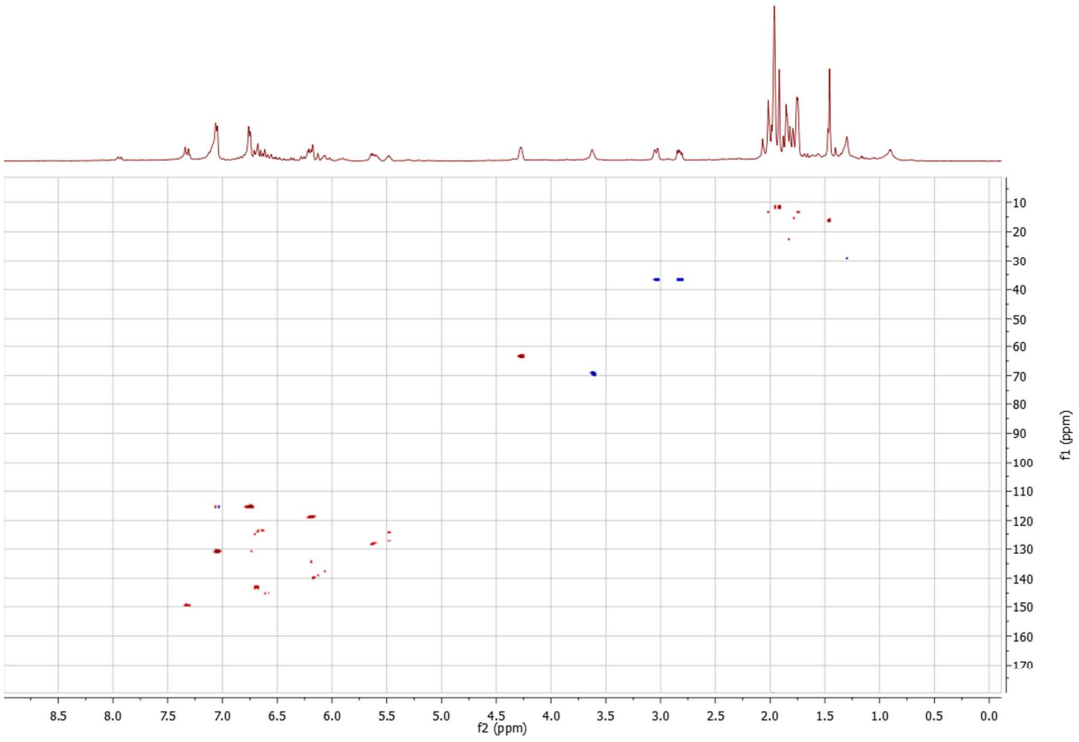


Figure S16: ^{1}H-^{13}C HMBC spectrum of 4 in acetonitrile-d_3.

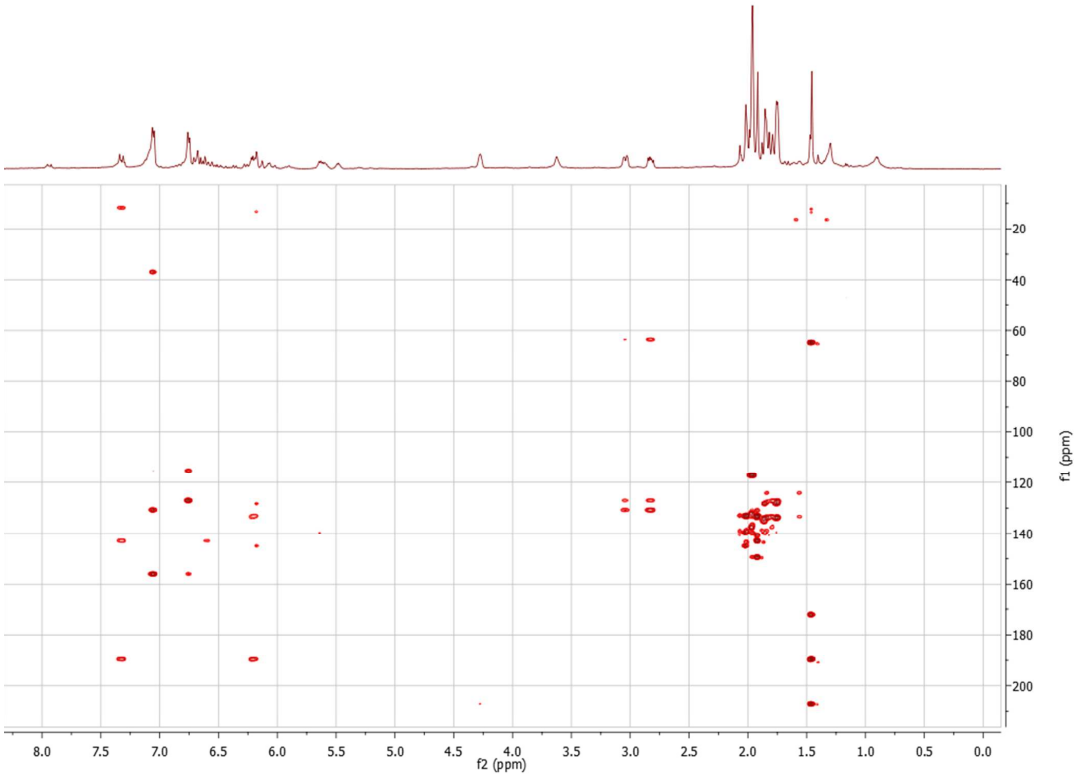


Figure S17: ^{1}H-^{1}H NOESY spectrum of 4 in acetonitrile-*d*$_3$.

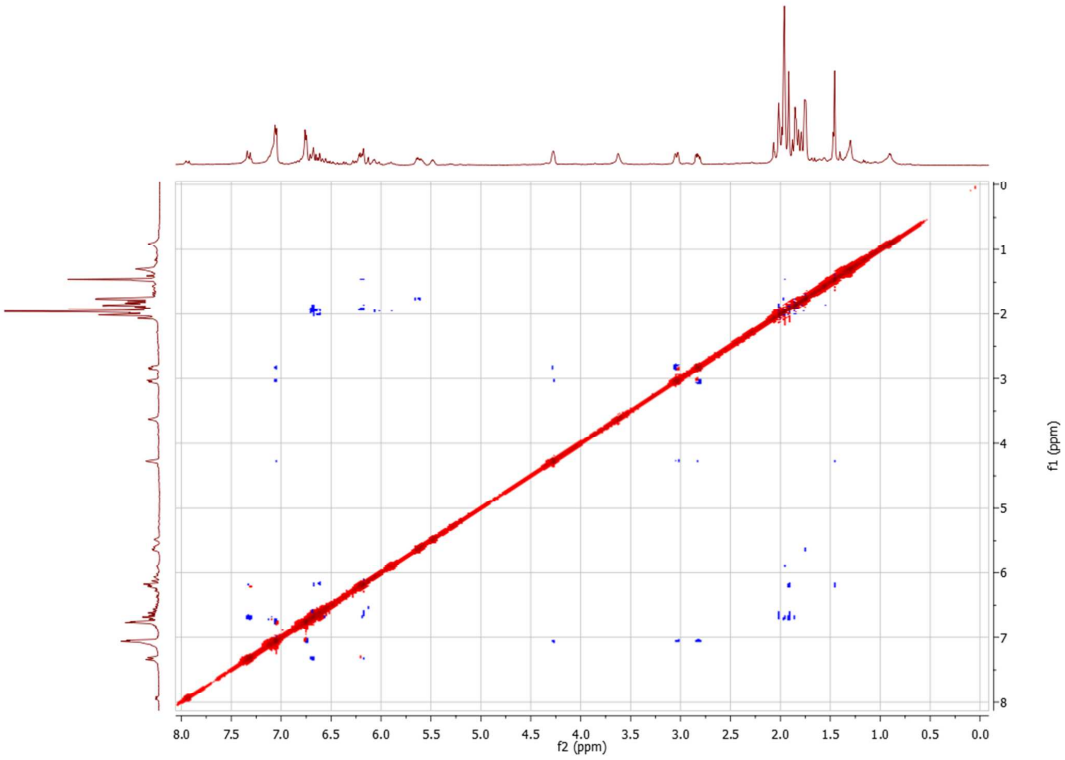


Figure S18: 1D ^{1}H spectrum of 5 in acetonitrile-d_3.

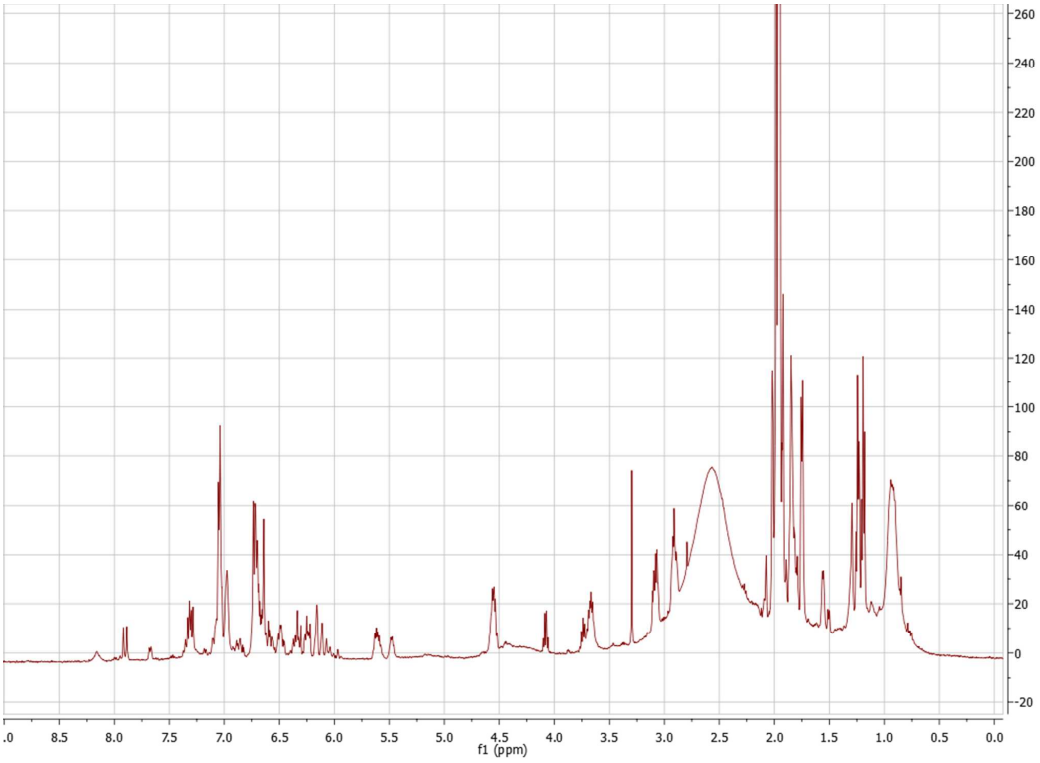


Figure S19: ^{1}H-^{1}H COSY spectrum of 5 in acetonitrile-*d*$_3$.

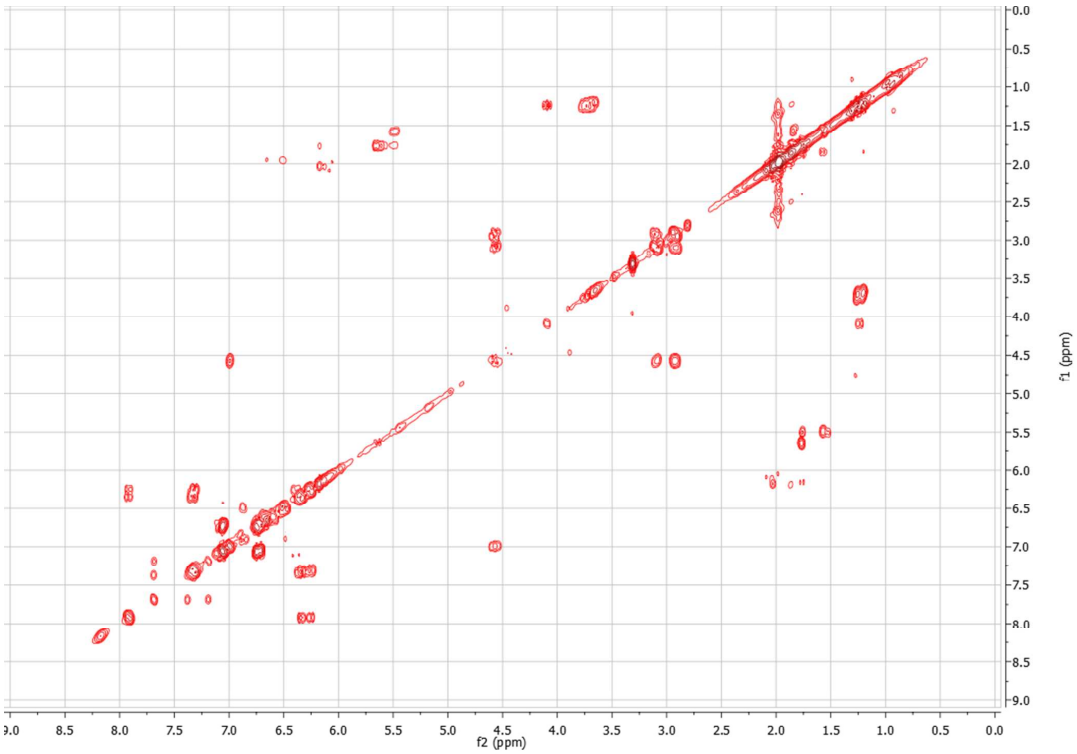


Figure S20: ^{1}H-^{1}H TOCSY spectrum of 5 in acetonitrile-d_3.

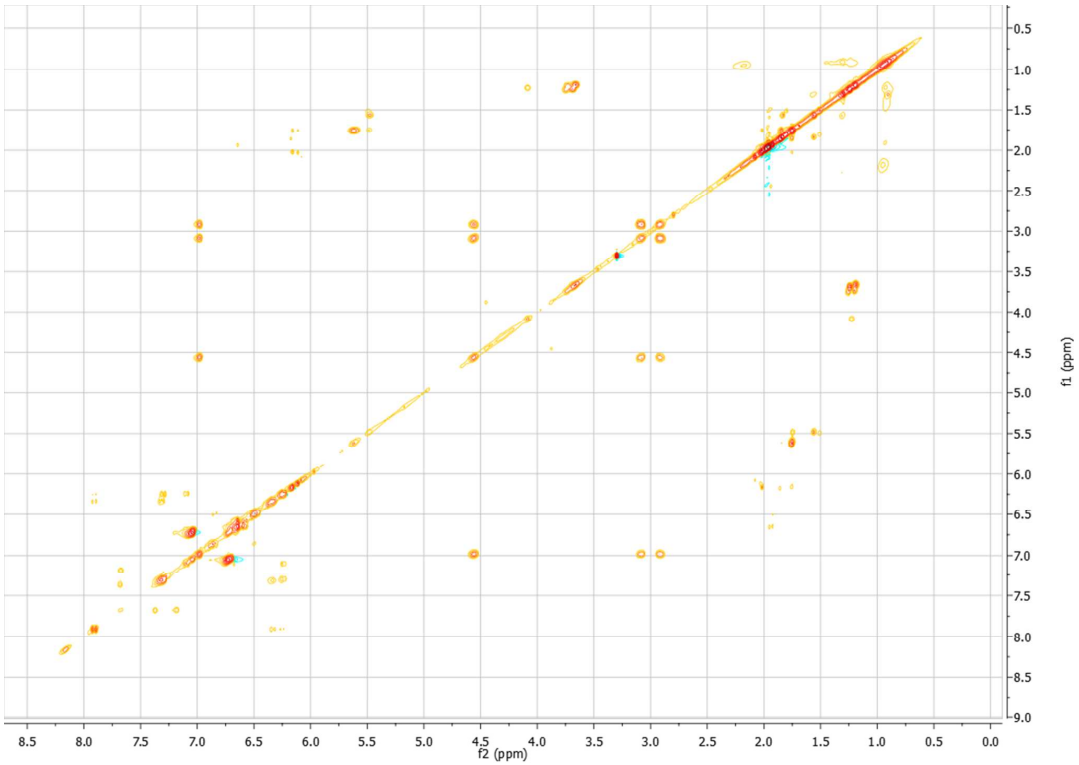


Figure S21: ^{1}H-^{13}C HMBC spectrum of 5 in acetonitrile-d_3.

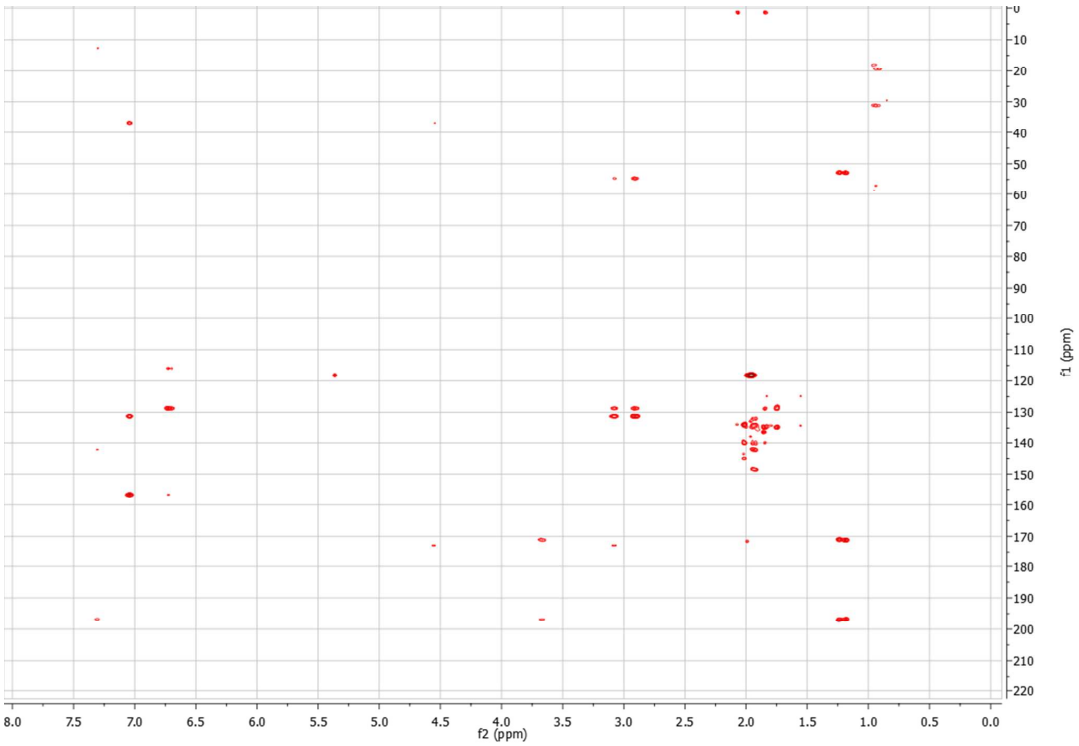


Figure S22: FT-ICR ESIMS data of 3.

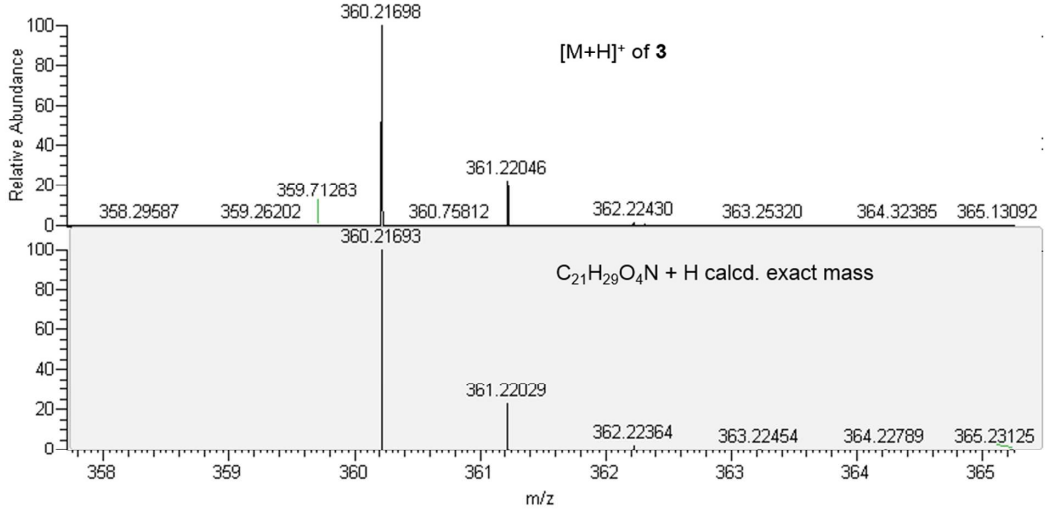


Figure S23: FT-ICR ESIMS data of 4.

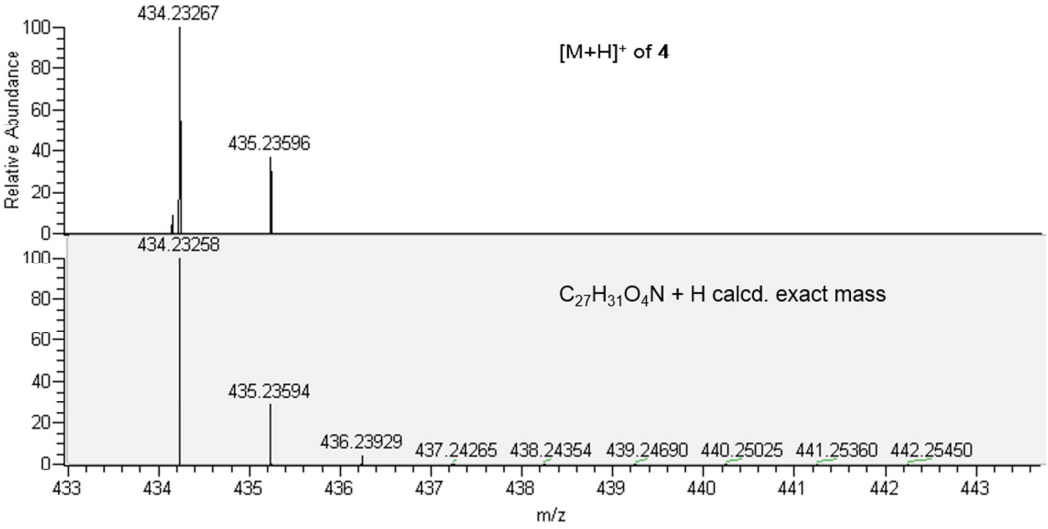
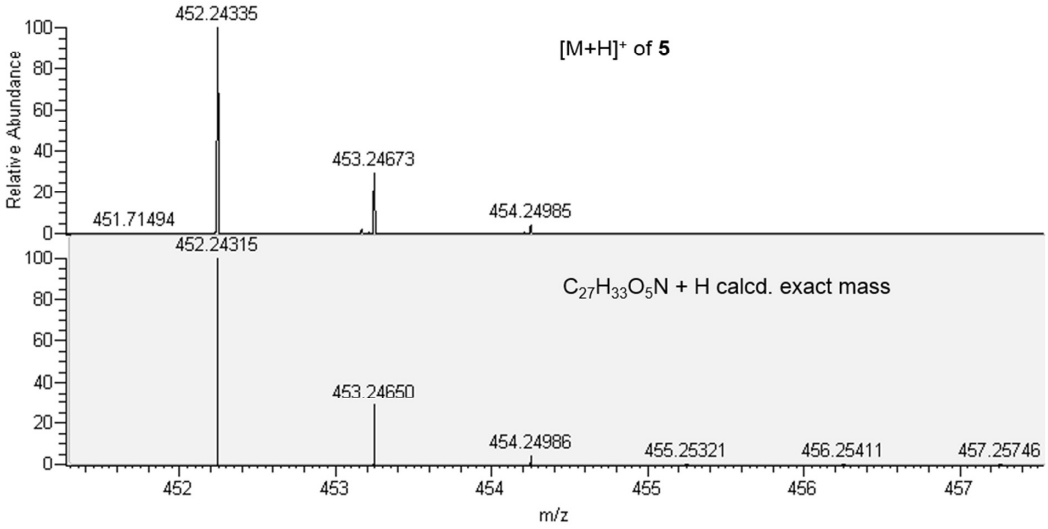


Figure S24: FT-ICR ESIMS data of 5.



REFERENCES

1. Sims, J. W., Fillmore, J. P., Warner, D. D., and Schmidt, E. W. (2005) Equisetin biosynthesis in Fusarium heterosporum, *Chem Commun (Camb)*, 186-188.
2. Sims, J. W., and Schmidt, E. W. (2008) Thioesterase-like role for fungal PKS-NRPS hybrid reductive domains, *J Am Chem Soc 130*, 11149-11155.
3. McCluskey, K. (2003) The Fungal Genetics Stock Center: from molds to molecules, *Adv Appl Microbiol 52*, 245-262.
4. Marfori, E. C., Kajiyama, S., Fukusaki, E., and Kobayashi, A. (2002) Trichosetin, a novel tetramic acid antibiotic produced in dual culture of Trichoderma harzianum and Catharanthus roseus Callus, *Z Naturforsch C 57*, 465-470.

CHAPTER 3

NATIVE PROMOTER STRATEGY FOR HIGH-YIELDING SYNTHESIS AND ENGINEERING OF FUNGAL SECONDARY METABOLITES

Reprinted with permission from
Kakule, T. B.; Jadulco, R. C.; Koch, M.; Janso, J. E.; Barrows, L. R.; Schmidt, E. W. Native Promoter Strategy for High-Yielding Synthesis and Engineering of Fungal Secondary Metabolites. *ACS Synth. Biol.* **2014**, DOI: 10.1021/sb500296p.

Note: my contribution to this paper was in planning, performing, and analyzing all the experiments.

ACS SyntheticBiology

Research Article

pubs.acs.org/synthbio

Native Promoter Strategy for High-Yielding Synthesis and Engineering of Fungal Secondary Metabolites

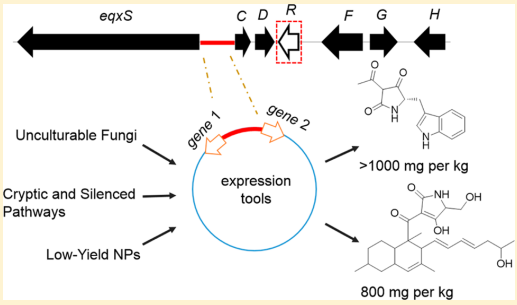
Thomas B. Kakule,[†] Raquel C. Jadulco,[‡] Michael Koch,[‡] Jeffrey E. Janso,[§] Louis R. Barrows,[‡] and Eric W. Schmidt*[,†]

[†]Department of Medicinal Chemistry and [‡]Department of Pharmacology and Toxicology, University of Utah, Salt Lake City, Utah 84112, United States

[§]Natural Products, Worldwide Medicinal Chemistry, Pfizer Worldwide Research and Development, Groton, Connecticut 06355, United States

Supporting Information

ABSTRACT: Strategies are needed for the robust production of cryptic, silenced, or engineered secondary metabolites in fungi. The filamentous fungus *Fusarium heterosporum* natively synthesizes the polyketide equisetin at >2 g L^{-1} in a controllable manner. We hypothesized that this production level was achieved by regulatory elements in the equisetin pathway, leading to the prediction that the same regulatory elements would be useful in producing other secondary metabolites. This was tested by using the native *eqxS* promoter and *eqxR* regulator in *F. heterosporum*, synthesizing heterologous natural products in yields of ∼1 g L^{-1}. As proof of concept for the practical application, we resurrected an extinct pathway from an endophytic fungus with an initial yield of >800 mg L^{-1}, leading to the practical synthesis of a selective antituberculosis agent. Finally, the method enabled new insights into the function of polyketide synthases in filamentous fungi. These results demonstrate a strategy for optimally employing native regulators for the robust synthesis of secondary metabolites.



Because of new technologies in sequencing and bioinformatics, it is now relatively trivial to identify novel biosynthetic pathways to secondary metabolites in cultivated organisms and in the environment.[1] Research in this area has led to a renewed appreciation that many secondary metabolites have yet to be described and that a vast resource awaits discovery.[2,3] Although pathway identification is simple, discovering the compounds produced by these pathways remains much more challenging, especially when the pathways are silent or encoded in the genomes of uncultivated organisms. To obtain the new compound, in many cases, the identified genes must be transferred to a new host and successfully expressed.[4] Despite many advances in technology, this is still nontrivial.

To date, many heterologous expression hosts have been developed for secondary metabolite production.[4−8] Most of these focus on expression of bacterial biosynthetic pathways, for which numerous hosts exist. For eukaryotic pathways, such as those from filamentous fungi, genes have been heterologously expressed in *Escherichia coli*, *Saccharomyces cerevisae*, *Aspergillus oryzae*, and several other yeasts and filamentous fungi.[5−7,9,10] Practical scale production of fungal compounds in bacteria has proven to be challenging.[4] *S. cerevisae* provides a robust platform but with relatively modest purified yields. Additionally, it does not handle introns well, requiring that introns be removed prior to expression. Similarly, heterologous expression platforms in various filamentous fungi have led to relatively modest yields, although in many cases fungal introns are tolerated, allowing genomic DNA to be employed directly.[5,9] These expression systems generally use housekeeping or related primary metabolic promoters that induce robust transcription of the desired genes, indicating that perhaps seeking improvements in level of transcription alone is insufficient to provide high levels of secondary metabolites, and suggesting room for the development of other strategies.

Here, we sought to take advantage of high-titer production of equisetin in the filamentous fungus *Fusarium heterosporum* ATCC 74349.[11,12] *F. heterosporum* produces equisetin at ∼2 g L^{-1} on corn grit agar (CGA), yet production is undetectable (10 ng L^{-1} detection limit) in many other types of media. We thus hypothesized that if an exogenous biosynthetic gene was placed under the control of the equisetin biosynthetic regulon, the heterologous compound would be synthesized in a yield similar to that observed for equisetin in the wild-type strain. Moreover, the highly controlled regulation of the equisetin locus might enable the production of compounds that are natively toxic to *F. heterosporum*. This strategy bears similarity to that previously used in actinomycetes to synthesize

Received: August 8, 2014

   dx.doi.org/10.1021/sb500296p | *ACS Synth. Biol.* XXXX, XXX, XXX−XXX

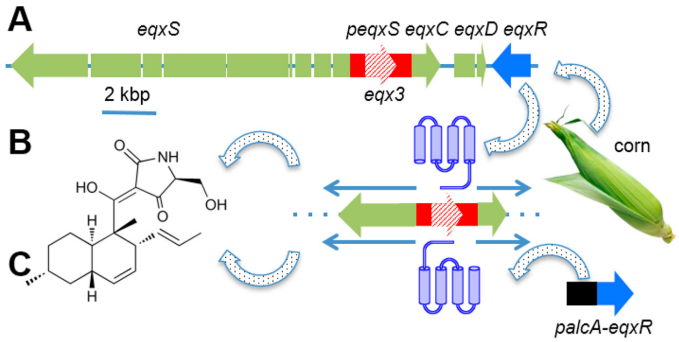


Figure 1. Model of the regulation of equisetin production. (A) The wild-type equisetin biosynthetic gene cluster, *eqx*, contains genes encoding biosynthetic proteins (green), a transcription factor (blue), and a promoter region used in this study (*peqxS*, red). (B) In the wild-type *F. heterosporum*, equisetin is produced when the fungus is grown on corn-derived media. A model is that the transcription factor gene *eqxR* is transcribed, leading to transcription of the pathway via *peqxS*. Shown is transcription of *eqxS* and *eqxC*, which together produce the desmethyl-equisetin analog, trichosetin, which is shown at left. (C) In support of this transcriptional model, when the heterologous promoter *palcA* is placed in front of *eqxR*, production of equisetin becomes constitutive and no longer depends upon growth on corn.

polyketides[13] but stands in contrast to a commonly used strategy in filamentous fungi, wherein promoters or regulators from primary metabolism are used to produce recombinant polyketides.

Using the model described below, we produced several heterologous fungal polyketides in *F. heterosporum*. In fungi, polyketides are biosynthesized by iterative decarboxylative condensation of malonyl units.[14,15] The minimal domains that make up the polyketide synthase (PKS) include acyltransferase (AT) domain, which selects the substrate; acyl carrier protein (ACP), which tethers the growing chain during extension; and the ketosynthase (KS) domain, which catalyzes the condensation reaction.[14−16] In addition, PKS enzymes may contain other modifying domains such as the methyltransferase (MT) domain, which introduces varying patterns of C-methylations along the polyketide backbone. For the reducing-type PKS, varying levels of reduction are achieved by ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER) domains.[14] In general, all domains required to produce a fungal polyketide are contained on a single polypeptide, but for several PKSs, the ER is trans-acting and is translated as a separate polypeptide from the other PKS domains.[17−20] In the latter case, the trans-ER protein is required to produce the natural product. In many fungi, PKS modules exist as hybrids with nonribosomal peptide synthetases (NRPS), resulting in formation of polyketides fused to amino acids.[10,15,21,22] Unlike standalone PKSs that rely on hydrolysis via a thioesterase domain (TE) for product release, several PKS-NRPSs use a terminal reductase (R) domain, which catalyzes a Dieckmann reaction in tandem with product release to form a tetramic acid ring.[23,24]

Here, we describe a platform that enables production of heterologous fungal secondary metabolites in high titers (∼1 g L^{-1} unoptimized yield). The platform was applied to several problems in biosynthesis and drug discovery, including resurrection of a silenced pathway of potential use in tuberculosis.

RESULTS AND DISCUSSION

Design of Expression Strategy. The strategy was based upon prior knowledge of the regulation of equisetin production. Equisetin biosynthesis requires the coordinated action of the PKS EqxS, the auxiliary ER EqxC, and the N-MT EqxD (Figure 1).[11] The first two of these are synthesized from genes that are divergently transcribed from a promoter region, *peqxS*. A regulatory transcription factor, EqxR, drives production of equisetin.[11] On CGA media, a large amount of equisetin is produced, while on other media such as potato dextrose agar (PDA; Difco) or potato dextrose broth (PDB; Difco), equisetin is not produced.[11] Equisetin is constitutively produced on CGA and slowly accumulates over a 21-day period. We thus selected *peqxS* as a platform for producing metabolites, under control of EqxR. In this conception, the *peqxS* region also contains a hypothetical gene, *eqx3*, which is commonly found in equisetin-like pathways. For simplicity, this region was left in place in the current study. In some vectors, we used only one-half of *peqxS* to express genes in one direction; in this series, *eqx3* is disrupted, and the promoter region is referred to as *peqxS′*. In others, the bidirectional promoter region *peqxS* was used; in these cases *eqx3* is not disrupted.

While the wild-type regulator *eqxR* was well controlled to produce equisetin solely on CGA, we envisioned applications in which we might desire more rapid production in liquid media. In previous work, we fused the *alcA* promoter (*palcA*) with *eqxR* (Figure 1C).[11] While *palcA* is inducible, it exhibits significant basal expression. Indeed, *palcA-eqxR* led to constitutive production of equisetin even in broth.[11] However, production yields were lower in broth, and there was a concern that toxic metabolites may delay or halt fungal growth under control of *palcA-eqxR*. By contrast, an advantage of the PDB method was that production could be achieved in 5−7 days, rather than the 21+ days used with CGA. The strategy thus took advantage of this strong control of equisetin production. If toxicity was a concern or if high yields were required, the wild-type *eqxR* construct was used. For a faster assessment of whether a recombinant metabolite could be produced, the *palcA-eqxR* construct could be employed. In this study, we employed two strategies to introduce *palcA-eqxR*: (1) we recently reported a modified *F. heterosporum* strain, Palc:eqxR, which overexpresses the equisetin positive regulator *eqxR* under control of the *alcA* promoter to allow equisetin production in broth culture after 5 days, whereas normally no equisetin is produced in liquid broth;[11] (2) we included the *palcA-eqxR*

dx.doi.org/10.1021/sb500296p | *ACS Synth. Biol.* XXXX, XXX, XXX−XXX

gene in an expression vector that could be transferred to *F. heterosporum*, along with desired secondary metabolic genes.

We had two major concerns in terms of practical genetics: could we repetitively use a single promoter element, and could we process different classes of introns? *F. heterosporum* largely undergoes ectopic recombination in our hands, and it is quite difficult to obtain homologous recombinants. For this reason, we believed that repetitive use of *peqxS* would be tolerated by the strain, enabling production of more than two gene products without recombination. Indeed, in this study, we show that multiple copies of this element are stable in recombination, meaning that the strategy is scalable to recapitulate multigene biosynthetic pathways. We were also concerned about the ability to splice different types of introns. While this will presumably always be a problem in some cases, here, we show that a quite divergent set of introns can be processed. However, our overall strategy mainly uses artificially spliced, intron-free DNA.

Our construction strategy involved using shuttle vectors, in which we could employ yeast recombination in *S. cerevisiae*[25] to build the desired vectors, *E. coli* to amplify those vectors, and *F. heterosporum* for production of compounds. Each vector thus required selection and replication elements for each strain (see Methods). We constructed and tested multiple types of selection markers so that multiple different vectors could be inserted into *F. heterosporum*. Previously, we used hygromycin (*hph*) and phleomycin (*ble*) resistance effectively.[11] Here, we also employed uracil auxotrophy as an additional selection marker.[26] By combining these elements, it becomes possible to use stepwise engineering to insert multiple copies of genes under control of *peqxS*. In addition, it is always possible to insert multiple copies of the promoter into a single vector prior to transformation into *F. heterosporum*, making the strategy highly scalable.

Finally, we used an 8-cutter restriction endonuclease site, either AscI or PacI, to linearize the vectors before fungal transformation. These sites were selected so that the linearized vectors would contain the complete promoter-heterologous gene construct in the correct order. Otherwise, there was a danger of integration with a disrupted reading frame. In addition, we put these sites in some cases between selectable markers and synthetic green fluorescent protein (sGFP) to ensure that the vector was inserted in the correct manner. However, in the event, we never found a rearrangement that necessitated using sGFP. In *F. heterosporum*, linearized vectors integrated ectopically and intact.

Below, we describe a stepwise application of these principles to construct a heterologous expression platform. (1) The basics of transcription using both wild-type *eqxR* and *palcA-eqxR* were examined using sGFP. Although simple chemical analysis of equisetin production previously revealed fundamental aspects of pathway regulation, employing sGFP enabled a direct translational readout that would complement our understanding of *eqxR* and *peqxS*. (2) We desired to test this strategy by adding a single, discrete, heterologous PKS that leads to a known and well characterized product. We selected the *cpaS* gene from the *Aspergillus flavus* cyclopiazonic acid pathway for this purpose. Not only is *cpaS* exceptionally well characterized,[9,27] but *A. flavus* is in a different class (Eurotiomycetes) from *Fusarium* (Sordariomycetes). (3) We next tested the applicability to systems requiring two genes for production of compounds. In the event, we used both the homologous equisetin pathway[11] and the heterologous lovastatin pathway, again from an *Aspergillus*.[19] By examining equisetin production, we hoped to determine whether chromosomal location impacted production; in the event, it did not. The lovastatin PKS genes are well characterized[17,19,28] and biomedically of great importance, making this an interesting target for production. Moreover, only two proteins are required to synthesize the complex core of the molecule. (4) We aimed to produce more than two genes and to test the practical application of the platform to a real-world problem. This involved production of the "extinct" metabolite, pyrrolocin. Taken together, these approaches fully define the application of equisetin regulatory elements to produce diverse secondary metabolites in high yield.

Knockout of the *eqx* Locus. We desired to compare expression in the wild-type *F. heterosporum* with expression in an *eqx* knockout strain. Although *eqxC* has been previously deleted, *eqxS* was difficult to delete. Here, we created a knockout vector in which *eqxS*, *eqxC*, *eqxD*, and the promoter region *peqxS* were deleted (Figure 2). The resulting strain,

Figure 2. Knockout of *eqx* genes. (A) The wild-type gene cluster. (B) Previously, *eqxC* was deleted with *hph*. (C) Here, this *eqxC* knockout was extended to delete *peqxS* and *eqxC* and to disrupt the 5′-regions of *eqxS* and *eqxD*, using the *ble* marker. This knockout required the prior deletion of *eqxC* and did not work in the wild-type strain.

FusΔeqxS was confirmed to be an *eqx* cluster knockout by PCR (Supporting Information Figure S4). Subsequent genome sequencing of FusΔeqxS revealed that the *eqxS* gene was knocked out in the 5′ end and that the deletion vector was inserted into a total of 3 locations in the genome (Supporting Information Figure S5).

***eqxS* Promoter Drives Expression of sGFP in Recombinant Strain.** To test whether the *peqxS* promoter sequence could be used to express exogenous genes, and to determine timing and control of heterologous expression, the *peqxS′* was fused to sGFP and transferred to wild-type *F. heterosporum* in vector FH-1 (Figure 3A; Supporting Information Figure S1). After 21 days on CGA, sGFP was observed in fungal filaments when examined by confocal fluorescence microscopy. Moreover, the fungus was fluorescent over the full expression period on CGA, but it lacked any observable fluorescence on PDA. This was a clear indication that *peqxS′* was sufficient for heterologous expression. It also further supported previous observations that production was constitutive on CGA but completely shut off on PDA and other media. By contrast, when sGFP was added to strains containing the leaky *palcA-eqxR* gene (Figure 3B), the resulting *F. heterosporum* strain was constitutively fluorescent on all media.

Expression of the *cpaS* Gene from *Aspergillus flavus* Produces Expected Metabolite. To determine whether the *peqxS′* could lead to biosynthesis of new compounds, the well-characterized *cpaS* gene from *Aspergillus flavus* was cloned into FH-1 to make pHygB-Cpas. The resulting *F. heterosporum*

dx.doi.org/10.1021/sb500296p | *ACS Synth. Biol.* XXXX, XXX, XXX−XXX

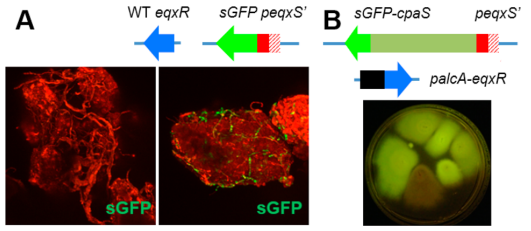


Figure 3. Monitoring *peqxS'*-driven transcription with sGFP. (A) *sGFP* was placed immediately downstream of *peqxS'* and cloned into the wild-type fungus. The wild-type control (left) was compared with the transformed fungus (right) on CGA. Shown is a single corn grit from a 21-day CGA culture, by confocal microscopy. Fluorescence was constitutive on CGA, but not visible on other media. Fungal filaments can be seen glowing green, over the red autofluorescent background. (B) *sGFP* was tethered to the gene, *cpaS*, downstream of the *peqxS'* promoter. The resulting vector also contained the constitutive *palcA-exqR* regulator. Shown here is a photograph of constitutively fluorescent fungal colonies grown on PDA, visualized on a Dark Reader.

mutant, Peqx:Cpas, was then cultured in potato dextrose broth (PDB) for 7 days. The predicted product, cAATrp **1** was identified by high pressure liquid chromatography (HPLC) (Figure 4A) and isolated. Comparison of the ^{1}H NMR spectrum and the molecular formula of **1** with those of the previously reported cAATrp showed that they were identical (Supporting Information Figure S13).[23,27] This confirmed the value of *peqxS'* in producing heterologous compounds.

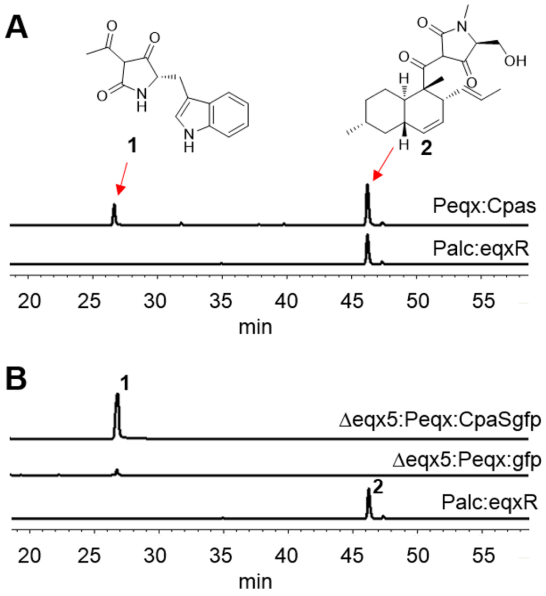


Figure 4. Expression of CpaS and production of cAATrp (**1**). (A) Expression in the wild-type *F. heterosporum* strain. HPLC-diode array detection (DAD) analysis of crude extracts of PDB cultures of Peqx:Cpas mutant and the Palc:eqxR control. The eqxS promoter drives expression to avail expected product **1**. Coproduction of equisetin **2** is observed. (B) Expression in *eqx* knockout strain. Production of **1** is improved, and equisetin **2** is no longer observed.

However, high levels of equisetin **2** were coproduced (Figure 4A).

In other systems, it has been shown that eliminating the production of undesired metabolites increases substrate flux to the target pathway resulting in higher yields of the desired compounds.[5] In our case, we reasoned that deleting the equisetin biosynthetic genes would not only increase flux of the building block, malonyl CoA, to polyketide-type heterologous pathways, but also ease downstream target compound purification and analysis. The *A. flavus cpaS* gene was cloned into vector FH-2 (Supporting Information Figure S2), which contained a copy of *palcA-eqxR* for constitutive expression in broth. An sGFP tag was fused to the C-terminus of *cpaS*. After transformation of this construct into the *eqx* knockout strain FusΔeqx5, we isolated a mutant Δeqx5:Peqx:CpaSgfp that glowed bright yellow on a dark reader (Figure 3B). This strain was cultured in 100 mL PDB for 7 days, extracted, and the extract analyzed by HPLC-DAD to show that **1** was robustly produced, while equisetin was absent (Figure 4B). Production of **1** was measured to be 100 mg L^{-1} by HPLC-DAD in comparison with a standard curve. This is an order of magnitude greater than the previously reported production of the same compound when the *cpaS* was expressed in *A. oryzae* under control of a housekeeping promoter.[27] In addition, this experiment provided some support for our hypothesis, since the yield in the deletion mutant was increased by approximately 4-fold in comparison to production in the wild-type equisetin producer.

When the Δeqx5:Peqx:CpaSgfp mutant was cultured for 21 days on CGA, production of **1** rose to over 1 g kg^{-1} without any optimization of production conditions. *A. flavus* is quite phylogenetically distant to *F. heterosporum*, with the former in Class Eurotiomycetes and the latter in Class Sordariomycetes, indicating that the strategy may be widely applicable to fungal metabolites from different groups.

Use of the Divergent Promoter for Simultaneous Dual Gene Introduction. In initial experiments, only one side of *peqxS* was used, comprising 1.5 kbp of gene sequence. To express two genes, *peqxS* was synthesized, containing the entire ~2.5 kbp of the divergent promoter, to generate vector FH-3 (Figure 5A). Previously, we showed by knockout mutagenesis that *eqxC* was critical for the production of equisetin **2** and trichosetin **3**, and we proposed that it was the trans-acting ER for the equisetin pathway.[11] Here, we showed the direct involvement of *eqxS* in the biosynthesis of trichosetin **3**. Both *eqxS* and *eqxC* were cloned into vector FH-3 under control of the divergent *eqx* promoter. In initial experiments, *eqxS* was fused with a C-terminal sGFP tag so that we could readily confirm protein expression. The resulting vector, hpheqxC +eqxSgfp, was transformed into FusΔeqx5. The isolated transformants were brightly fluorescent, demonstrating appropriate gene expression under control of the divergent promoter. In addition, the fluorescence was constitutively obtained on PDA, further demonstrating the constitutive regulation of the pathway under control of leaky *palcA*.

Dieckmann Cyclase Depends upon Unmodified C-Terminal R Domain. To our surprise, trichosetin was not detected by HPLC-DAD in the crude extracts of the isolated mutants (Figure 5B). Instead, close inspection of the LC/MS trace showed a new product **4** heavier than trichosetin by 18 Da, but with a similar fragmentation pattern (Figure 5C; Supporting Information Figure S11). Compound **4** was characterized by NMR experiments, including ^{1}H, ^{13}C,

dx.doi.org/10.1021/sb500296p | *ACS Synth. Biol.* XXXX, XXX, XXX−XXX

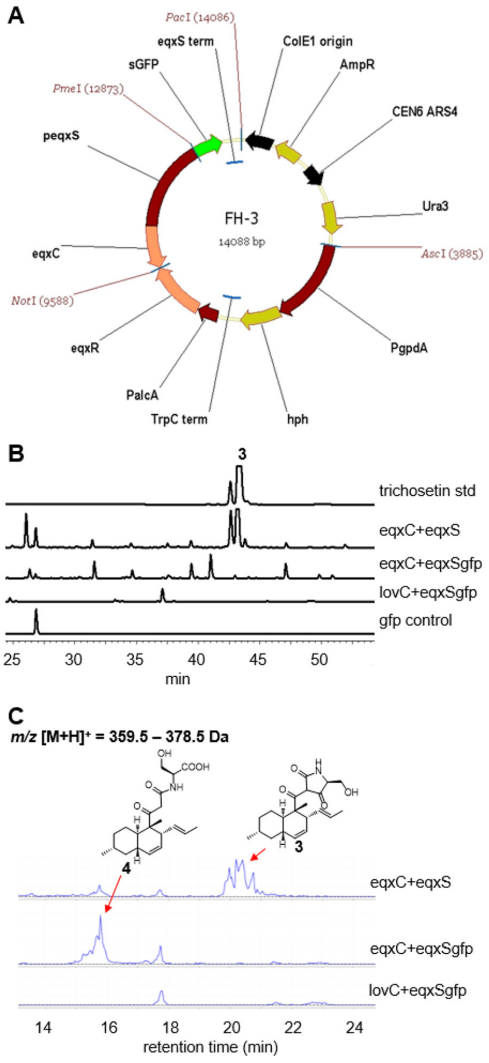


Figure 5. Trichosetin synthesis requires EqxC and unmodified EqxS. (A) Expression vector FH-3 designed with complete intergenic sequence *peqxS*, for dual expression of genes. Also shown are elements that permit cloning by recombination in *S. cerevisiae* and selection in *E. coli*. (B) Analytical HPLC of crude extracts of PDB cultures of FusΔeqx5 transformed with *eqxSgfp* together with either *eqxC* or noncognate *trans*-ER *lovC*. Also shown is nontagged *eqxS* coexpressed with *eqxC* and the trichosetin standard. Trichosetin is only produced in the presence of *eqxC* and unmodified *eqxS*. (C) LC/MS analysis of crude extracts shows that a gfp tag on the C-terminus of EqxS interferes with formation of trichosetin and instead results in production of only the ring-open form **4**.

gCOSY, gHSQC, and gHMBC, to be the ring-opened derivative of trichosetin that had not undergone the terminal Dieckmann reaction.[23,24] This result implied that *eqxS* is indeed the equisetin synthetase, but the gfp tag interferes with the proper functioning of the reductase domain, preventing Dieckmann cyclization and instead promoting water-mediated hydrolysis of the intermediate. This was, in fact, found to be the case because, when *eqxS* was cloned without a C-terminal tag, trichosetin was robustly synthesized at wild-type levels (Figure 5B).

Production of Multiple Gene Products Leads to Synthesis of Lovastatin Precursors in *F. heterosporum*. To establish the utility of this new expression strategy for coexpression of heterologous genes, we cloned the well-studied lovastatin nonaketide synthase (*lovB*) from *Aspergillus terreus* together with its cognate *trans*-ER (*lovC*)[19] to make the hphlovC+lovBgfp plasmid. We also cloned only *lovB* into FH-3 to make the plasmid hpheqxC+lovBgfp. Transformation of these constructs independently into FusΔeqx5 resulted in production of the expected products. Without its cognate *trans*-ER, the expressed LovB synthesized the previously reported truncated intermediates, **5** and **6** (Figure 6A).[19] Coexpression

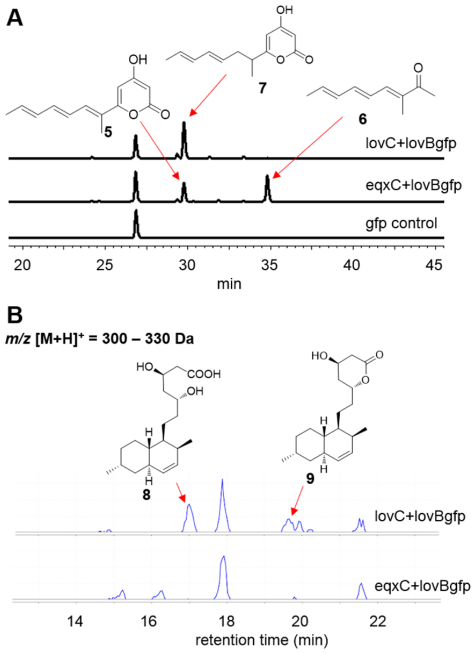


Figure 6. Dual expression with *peqxS* promoter reconstitutes pathway to lovastatin precursor. (A) Analytical HPLC of crude extracts of PDB cultures of FusΔeqx5 transformed with *lovBgfp* together with *lovC* or a noncognate trans-ER *eqxC*. In the presence of *eqxC*, *lovB* produces the polyene pyrone **5** and ketone **6**; and more reduced pyrone **7** with *lovC* coexpression. (B) LC/MS analysis of crude extracts shows formation of the lovastatin precursor, dihydromonacolin L acid **8** and the lactone **9** when *lovBgfp* is coexpressed with *lovC*. EqxC is not able to complement LovB to form **8** or **9**.

of *lovB* with *lovC* produced the expected reduced metabolites **7** (Figure 6A), the dihydromonacolin L acid **8**, and the lovastatin precursor lactone **9** (Figure 6B).[19,29] These metabolites were characterized by comparing their liquid chromatography/mass spectroscopy (LC/MS), ultraviolet (UV), and ^{1}H NMR data to previous reports (Supporting Information Figures S8–S10, S12–S14). This confirmed that, indeed, the intergenic sequence between *eqxS* and *eqxC* could guide transcription divergently to heterologously coexpress two genes. The purified yield of **9** was 130 mg kg^{-1}, and **8** was produced in about equal

dx.doi.org/10.1021/sb500296p | *ACS Synth. Biol.* XXXX, XXX, XXX–XXX

amounts, indicating that the initial unoptimized yield should exceed ~300 mg kg^{-1}.

EqxS Requires Cognate *trans*-ER. In the course of cloning and analyzing *lovC*, we coexpressed *lovC* and *eqxS* under control of the divergent promoter. No trichosetin was produced under these conditions (Figure 5B). LovC could not complement EqxC, indicating that EqxS could not interact with the noncognate ER. Further, the complementary experiment in which EqxC was coexpressed with LovB failed to yield reduced intermediates (Figure 6A). By contrast, the ERs were fully functional when coexpressed with their cognate PKS proteins.

Resurrection of Antituberculosis Agent from a Silenced Biosynthetic Pathway. In collaboration with scientists at Wyeth (now Pfizer), researchers in the Barrows lab cultivated an endophytic fungus, designated strain NRRL 50135, obtained in Papua New Guinea as part of their International Cooperative Drug Discovery Group (ICBG) project. The crude extract was potently active against *Mycobacterium tuberculosis*, and assay-guided purification led to the identification of a novel compound, pyrrolocin A (**10**), as the active principle. In 2007, a tentative structure **10** was assigned based upon HRESIMS and NMR (Figure 7A), but the amount of compound was not sufficient to assign the stereochemistry or to perform further activity tests. Unfortunately, the fungus ceased producing the compound after the initial isolation experiments, which is an extremely common phenomenon in natural products research, so that the project could not be further pursued or published. Numerous attempts at modifying production conditions, using many different published methods, failed to resurrect the synthesis of this potentially important compound. This presented the perfect challenge to test the application of our new production platform. The goal was to resurrect production in *F. heterosporum*, in tandem with greatly increasing the production level so that the compound could be further developed.

We sequenced the NRRL 50135 genome, which after assembly was found to be 54.1 Mbp on 5809 contigs (calculated GC content is 47.5%). Autoannotation software predicted 17722 proteins. From the preliminary structure of **10**, we predicted its biosynthesis would be similar to that of equisetin,[11] requiring a PKS-NRPS, *trans*-ER, and MT. BLAST analysis of the predicted proteins using EqxS as a query produced only 2 PKS-NRPS genes in the entire genome. When the database was queried with EqxC, the top hit was on the same 50.7 kbp contig as the identified top hit obtained with the EqxS query (Figure 7B). These genes were subsequently designated *prlS* and *prlC*, in analogy with the equisetin nomenclature. However, we could not find an MT that coclustered with either of the PKS-NRPS genes. Instead, distantly similar genes (37% identity to the equisetin *N*-methyltransferase, EqxD) were found unclustered with any PKS.

Since production of **10** was predicted to require coexpression of more than two genes, we created a derivative strain of FusΔeqxS, in which the strain was transformed into the uracil auxotroph, FusΔeqxSΔpyrG10, by directed knockout of the orotidine 5′-phosphate decarboxylase gene (*pyrG*). A complement vector for this auxotroph was constructed using the native *F. heterosporum pyrG*, to create expression vector FH-4 (Supporting Information Figure S3). The *prlS* and *prlC* sequences were then cloned into FH-4 and used to transform strain ΔeqxSΔpyrG10 to prototrophy. An isolated mutant ΔpyrG:Peqx:prlS+prlC was cultured in 250 mL PDB for 5

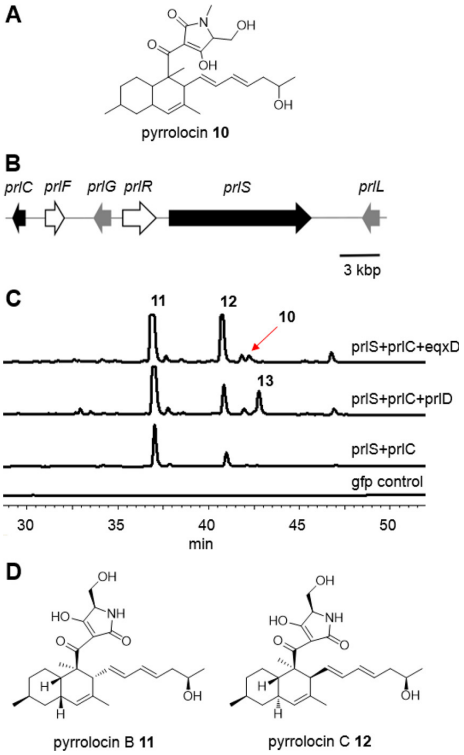


Figure 7. (A) Preliminary structure of antituberculosis agent **10** with mass 458 Da $(M+H)^+$ initially isolated from endophytic fungus NRRL 50135. (B) The identified candidate gene cluster for the biosynthesis of **10** after genome sequencing contains a PKS-NRPS hybrid gene *prlS*, an enoyl reductase *prlC*, two transcription factors *prlF* and *prlR*, and two exporter genes *prlG* and *prL*. (C) Analytical HPLC of crude extract of FusΔeqxSΔpyrG10 transformed with both *prlS* and *prlC* driven by the *eqxS* promoter shows synthesis of two new products **11** and **12** with corresponding mass of 444 Da $(M+H)^+$. Further introduction of methyltransferases *prlD* and *eqxD* led to synthesis of new minor products **13** and **10**, respectively. (D) Compounds **11** and **12** were found to have antituberculosis activity and to have the core structure of **10**.[33]

days. The crude extract contained pyrrolocins B (**11**) and C (**12**), the desmethyl derivatives of **10** (444 Da $[M + H]^+$).[33] Compounds **11** and **12** (Figure 7C) were produced in a 2:1 ratio, totaling >800 mg kg^{-1} on CGA, which is greatly in excess of what was initially found for **10** in the native producer before production was lost (~50 mg L^{-1}).

Two MTs were cloned in attempts to methylate the tetramate ring. The first, PrlD, was the top hit resulting from BLAST analysis of NRRL 50135 when EqxD was used as the query (36% identity). *prlD* was from a contig that was not linked with *prl*, meaning that it might result from a separate biosynthetic pathway. EqxD itself was also coexpressed with PrlS and PrlC. Each MT was cloned into complementary vector FH-1 and separately transformed into *F. heterosporum* ΔpyrG:Peqx:prlS+prlC. The mutants were cultivated for 7 days. Analysis of the crude extracts by HPLC-DAD and LC/MS showed that both PrlD and EqxD led to synthesis of methylated products. Interestingly, the EqxD-expressing strain

dx.doi.org/10.1021/sb500296p | *ACS Synth. Biol.* XXXX, XXX, XXX−XXX

produced a minor amount of authentic pyrrolocin A **10** (<10 mg kg^{-1}) in a background of ~800 mg kg^{-1} of nonmethylated **11** and **12**. The methylated derivative **13** produced by the PrlD-expressing strain did not match the NMR data for **10** and is therefore not the correct MT. No obvious MT exists in the sequenced genome that would be predicted to perform that transformation.

Conclusion. Heterologous expression is potentially the most universal solution to produce natural products from cryptic pathways and unculturable organisms.[4,5] We set out to develop such a production platform based on the equisetin biosynthetic regulon, which would enable production of important metabolites with sufficient yield to ease purification and downstream assays. This platform was designed based upon the hypothesis that high-level transcription alone was insufficient to produce high levels of secondary metabolites. Instead, exploitation of a natively high-producing biosynthetic pathway would provide the ideal environment for secondary metabolite biosynthesis. In developing and testing this platform, we designed derivative strains of *F. heterosporum* and a suite of expression vectors into which various heterologous genes could be cloned. These new tools were validated by producing a range of fungal polyketides from widely different fungi, in unoptimized yields ranging from several hundred mg to >1 g per liter. In fact, we found that cultures on just 50 g of CGA provided enough material for complete characterization of novel compounds. We provided proof of concept by resurrecting an important silenced pathway.

Another advantage of this platform is that it seems to process introns from a broad phylogenetic diversity of fungi. *prlS* and *prlC* contained introns that were properly processed (otherwise interrupting the reading frame of the proteins), demonstrating the breadth of the platform in handling introns from diverse strains. As has been found in other heterologous expression strategies, a disadvantage is that pathway shunt products are often produced along with the major natural product. However, in this case, the abundant amount of desired products compensated for this problem.

We have demonstrated practical application of this platform in reviving antituberculosis activity previously observed from pyrrolocin A, a product of the NRRL 50135 strain. Direct cloning of *prlS* and *prlC* from genomic DNA of strain NRRL 50135 into our expression vectors and subsequent high-level production of related active compounds **11** and **12** demonstrated the ease with which such eukaryotic genes ridden with introns could be characterized. This work enabled the structures of the pyrrolocins to be fully elucidated and the biological activity to be characterized.[33]

This strategy is widely applicable to many types of natural product synthesis in different organisms. The key factors that led us to select *F. heterosporum* as the production host were as follows: (1) the robust synthesis of the native compound, which meant that even if the recombinant yield were greatly reduced it should still be sufficient for chemical analysis; (2) the strict controllability of compound production, where essentially no natural product is produced on most media types, leading to a reduction in problems related to toxicity; (3) the apparent constitutive expression of the natural pathway, where compounds are slowly produced, exported, and accumulated in the medium over the course of weeks on CGA. It is likely that other systems with similar features would be amenable to the same approach.

A few biochemical observations were also enabled by these studies. During the complementation of the *eqx* knockout strain with *eqxC* and *eqxS*, we directly show that both these genes are required for equisetin production. However, we found that R domain function is altered when fused to an sGFP tag. Instead of the expected tetramic acid **3**, we obtained the ring-open **4**. We speculate that sGFP alters the structure of the R domain to prevent either transfer from the T domain or to allow water to enter the R domain active site. The tetramic acid derivative is nearly instantaneously formed from thioesters under neutral buffer conditions in water,[24] making it especially remarkable that the linear form can be obtained from thiotemplated synthesis. Potentially, if desired, this problem might be circumvented by experimenting with different types of linkers that do not disrupt the R domain. A cleavable linker strategy may also be feasible. However, in our hands, we found that sGFP is optional; for example, the complete vector is usually integrated intact in *F. heterosporum*. We also found that the ER proteins cannot be crossed between these pathways but that the wild-type ERs are required at least in these cases. The interaction of these ER proteins is also of interest, since having the correct set of protein partners is essential in the synthesis of the desired natural products.

In the wild-type fungus NRRL 50135, **10** was the major compound produced, with a small amount (<1% estimated from HPLC-MS) of **12** as a side-product.[33] Based upon the sequenced gene cluster and the recombinant expression performed here, it is clear that the combination of PKS-NRPS and auxiliary ER lead to the formation of **12**. Unlike our findings here, the wild-type fungus did not produce any *cis*-decalin product, such as **11**. It is possible that one of the other hypothetical genes in the gene cluster could be responsible for this discrepancy and may act as the pyrrolocin Diels–Alderase, but there are also other interesting possibilities. Of note, no obvious MT was present in the *prl* cluster that might produce **10**. The closest homologue of EqxD from the genome did not produce **10**. By contrast, EqxD itself produced ~10 mg kg^{-1} of **10**. This minor production, using a combination of heterologous and homologous proteins, would likely not have been observable if starting with a less efficient expression system. This emphasizes the value of starting with a high-yielding platform. The yield was still sufficient for biological and chemical characterization from a single 1 kg scale experiment. The reaction with EqxD is also remarkable in that **12** contains D-Ser, while **3** contains L-Ser, and the decalin ring is also enantiomeric between **3** and **12**. This indicates that, perhaps, EqxD has fairly relaxed substrate selectivity.

METHODS

Cloning of Vectors and Expression Plasmids. Standard PCR techniques were employed and plasmid construction carried out as previously described.[11] Details of vector construction can be found in the Supporting Information. Vector images were generated with Vector NTI software (Invitrogen).

Transformation of *Fusarium heterosporum* was done as previously described except that protoplasts were prepared from 8 to 10-h germinating spores.[11] When antibiotic selection was required, hygromycin or phleomycin was added to media at 150 μg mL^{-1}. For uracil auxotroph selection, protoplast regeneration agar lacking uracil was prepared containing 1 M sucrose, 0.02% yeast extract without amino acids, 0.02% BSM supplement, and 1% agar.

dx.doi.org/10.1021/sb500296p | *ACS Synth. Biol.* XXXX, XXX, XXX–XXX

Fungal Mutagenesis. Knockout of *eqxS* FusΔeqxS was constructed by transforming FusΔeqxC with knockout vector ClusterPhleoKO (Supporting Information) and transformants selected on phleomycin. The knockout cassette was made of the phleomycin resistance marker flanked by sequences homologous to regions within *eqxS* and *eqxD* (Figure 2). Isolated transformants were counter-screened for hygromycin sensitivity and verified by colony PCR. The genome of the identified knockout FusΔeqxS was extracted and sequenced. The reads were aligned to the FusWT reference genome[11] with Novoalign and output visualized with the integrative genomics viewer[30] (Supporting Information Figure S5).

Generating the Uracil Auxotroph. The pyrG knockout cassette in the TOPO-pyrGKO plasmid was made by cloning a randomly selected sequence (first exon of *lovC*) into the *F. heterosporum pyrG* sequence. Translation of this sequence results in a truncated, nonfunctional PyrG. FusΔeqxS was transformed with TOPO-pyrGKO, and the protoplasts were regenerated for 72 h at 30 °C before plating on selection medium made of Czapek Dox Broth, 5-FOA (4 g L^{-1}), uracil (1.12 g L^{-1}), uridine (140 mg L^{-1}), and 1.5% agar. Transformants were cross-streaked on Czapek−Dox agar with and without uracil to identify auxotrophs. Diagnostic PCR for homologous integration was done to confirm FusΔeqxSΔ-pyrG10 as a true knockout (Supporting Information Figure S6).

Fluorescence Microscopy. Visualization of gfp expressing mutants on CGA was done on an Olympus FV1000 spectral confocal microscope. Visualization of gfp expressing mutants on PDA was done on a Dark Reader (Clare Chemical).

Genome Sequencing and Analysis. All fungal genomes were extracted with the DNeasy Plant Minikit (Qiagen) and sequenced at the University of Utah Huntsman Cancer Institute sequencing facility on an Illumina HiSeq 2000.

The raw reads from NRRL 50135 genome sequencing were assembled using the SPAdes de novo genome assembler[31] with a kmer range from 31 to 85. Autoannotation of the genome was done using Augustus.[32] To locate the pyrrolocin biosynthetic cluster, BLAST analysis was done using EqxS, EqxC, and EqxD sequences as queries.

Chemical Analysis. Selected transformants were screened for compound production by culture in potato dextrose broth (PDB) for 7 d at 30 °C with shaking at 180 rpm. Spores of compound expressing mutants were then inoculated on to corn grit agar and incubated at room temperature for 21 d. Extraction of filtered PDB broth was done with ethyl acetate containing 1% acetic acid, and CGA was wholly extracted with acetone. Solvents were removed under vacuum, and the crude extracts analyzed by HPLC-DAD and LC/MS using C18 chromatography.

Purification of cAATrp 1. The crude extract from a PDB culture (250 mL) of Δeqx5:Peqx:CpaSgfp was purified by preparative HPLC (4 mL min^{-1}; 5−70% acetonitrile/water-0.05% TFA in 35 min). The pooled fractions were dried under vacuum to afford pure **1** (9 mg), which was then analyzed by ^{1}H NMR and LC/MS (Supporting Information Figures S7 and S13), in comparison to previously published reference data.[23,27] A standard curve was generated using pure **1** by HPLC-DAD and production was quantified from crude extracts of 100 mL PDB cultures and 50 g CGA cultures of Δeqx5:Peqx:Cpasgfp to average 100 mg L^{-1} and 1.25 g/kg (±0.24 g kg^{-1}, n = 3), respectively.

Purification of 4. CGA culture of Δeqx5:Peqx:eqxC +eqxSgfp (50 g) was extracted with acetone. The crude residue was then fractionated by flash chromatography on end-capped C18 with a methanol/water gradient and fractions screened by LC/MS. The fraction containing **4** was further purified by preparative HPLC to afford previously undescribed compound **4** (4.7 mg), which was characterized spectroscopically (see Supporting Information).

Purification of Dihydromonacolin L 9. The crude extract from 50 g CGA culture of Δeqx5:Peqx:lovC+lovBgfp was fractionated by flash chromatography on end-capped C18 with a methanol/water gradient. Fractions were screened by LC/MS and **9** was found to be contained in one fraction. This fraction was dried, and the residue was separated on silica column with 2:1 ethyl acetate/hexanes mobile phase. Further preparative HPLC purification afforded **9** as a white solid (6.6 mg). MS and ^{1}H NMR data were compared with those for the previously reported material.[19]

Purification of 10−13. To purify **11**, the crude extract from 200 g CGA culture was fractionated by flash chromatography on end-capped C18 using a methanol/water gradient. The fractions were analyzed by HPLC-DAD, and a portion (12.5%) of the fraction containing **11** was purified by several rounds of preparative HPLC to obtain pure **11** (8.6 mg; isolated yield 344 mg kg^{-1}). A similar strategy was used to obtain other derivatives.

■ ASSOCIATED CONTENT

Supporting Information

Plasmid construction; plasmids and primers used; NMR data; vector maps; PCR analyses; alignment of FusΔeqx5 illumina reads; LC/MS and UV data; ^{1}H and ^{13}C spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

The *prl* sequence has been deposited in GenBank, accession number KM107910.

■ AUTHOR INFORMATION

Corresponding Author

*Tel.: 801-585-5234; Email: ews1@utah.edu.

Notes

The authors declare no competing financial interest

■ ACKNOWLEDGMENTS

This work was funded by the National Science Foundation (NSF), 0957791. Collection and chemical analysis of the pyrrolocin-producing strain was funded by the National Institutes of Health (NIH), 5UO1TW006671-10.

■ REFERENCES

(1) Challis, G. L. (2008) Genome mining for novel natural product discovery. *J. Med. Chem. 51*, 2618−2628.

(2) Brakhage, A. A. (2013) Regulation of fungal secondary metabolism. *Nat. Rev. Microbiol. 11*, 21−32.

(3) Evans, B. S., Robinson, S. J., and Kelleher, N. L. (2011) Surveys of non-ribosomal peptide and polyketide assembly lines in fungi and prospects for their analysis *in vitro* and *in vivo*. *Fungal Genet. Biol. 48*, 49−61.

(4) Ongley, S. E., Bian, X., Neilan, B. A., and Muller, R. (2013) Recent advances in the heterologous expression of microbial natural product biosynthetic pathways. *Nat. Prod. Rep. 30*, 1121−1138.

(5) Chiang, Y.-M., Oakley, C. E., Ahuja, M., Entwistle, R., Schultz, A., Chang, S.-L., Sung, C. T., Wang, C. C. C., and Oakley, B. R. (2013) An

efficient system for heterologous expression of secondary metabolite genes in *Aspergillus nidulans*. *J. Am. Chem. Soc. 135*, 7720−7731.
(6) Kealey, J. T., Liu, L., Santi, D. V., Betlach, M. C., and Barr, P. J. (1998) Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proc. Natl. Acad. Sci. U.S.A. 95*, 505−509.
(7) Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E., and Khosla, C. (2001) Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science 291*, 1790−1792.
(8) Li, C., Zhang, F., and Kelly, W. L. (2011) Heterologous production of thiostrepton A and biosynthetic engineering of thiostrepton analogs. *Mol. BioSyst. 7*, 82−90.
(9) Tokuoka, M., Seshime, Y., Fujii, I., Kitamoto, K., Takahashi, T., and Koyama, Y. (2008) Identification of a novel polyketide synthase−nonribosomal peptide synthetase (PKS−NRPS) gene required for the biosynthesis of cyclopiazonic acid in *Aspergillus oryzae*. *Fungal Genet. Biol. 45*, 1608−1615.
(10) Xu, W., Cai, X., Jung, M. E., and Tang, Y. (2010) Analysis of intact and dissected fungal polyketide synthase-nonribosomal peptide synthetase *in vitro* and in *Saccharomyces cerevisiae*. *J. Am. Chem. Soc. 132*, 13604−13607.
(11) Kakule, T. B., Sardar, D., Lin, Z., and Schmidt, E. W. (2013) Two related pyrrolidinedione synthetase loci in *Fusarium heterosporum* ATCC 74349 produce divergent metabolites. *ACS Chem. Biol. 8*, 1549−1557.
(12) Singh, S. B., Zink, D. L., Goetz, M. A., Dombrowski, A. W., Polishook, J. D., and Hazuda, D. J. (1998) Equisetin and a novel opposite stereochemical homolog phomasetin, two fungal metabolites as inhibitors of HIV-1 integrase. *Tetrahedron Lett. 39*, 2243−2246.
(13) Stevens, D. C., Hari, T. P. A., and Boddy, C. N. (2013) The role of transcription in heterologous expression of polyketides in bacterial hosts. *Nat. Prod. Rep. 30*, 1391−1411.
(14) Chooi, Y. H., and Tang, Y. (2012) Navigating the fungal polyketide chemical space: From genes to molecules. *J. Org. Chem. 77*, 9933−9953.
(15) Boettger, D., and Hertweck, C. (2013) Molecular diversity sculpted by fungal PKS-NRPS hybrids. *ChemBioChem 14*, 28−42.
(16) Walsh, C. T., and Fischbach, M. A. (2010) Natural Products Version 2.0: Connecting genes to molecules. *J. Am. Chem. Soc. 132*, 2469−2493.
(17) Ames, B. D., Chi, N., Bruegger, J., Smith, P., Xu, W., Ma, S., Wong, E., Wong, S., Xie, X., Li, J. W. H., Vederas, J. C., Tang, Y., and Tsai, S.-C. (2012) Crystal structure and biochemical studies of the trans-acting polyketide enoyl reductase LovC from lovastatin biosynthesis. *Proc. Natl. Acad. Sci. U.S.A. 109*, 11144−11149.
(18) Heneghan, M. N., Yakasai, A. A., Williams, K., Kadir, K. A., Wasil, Z., Bakeer, W., Fisch, K. M., Bailey, A. M., Simpson, T. J., Cox, R. J., and Lazarus, C. M. (2011) The programming role of trans-acting enoyl reductases during the biosynthesis of highly reduced fungal polyketides. *Chem. Sci. 2*, 972−980.
(19) 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., and Tang, Y. (2009) Complete reconstitution of a highly reducing iterative polyketide synthase. *Science 326*, 589−592.
(20) Bergmann, S., Schuemann, J., Scherlach, K., Lange, C., Brakhage, A. A., and Hertweck, C. (2007) Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nat. Chem. Biol. 3*, 213−217.
(21) Fisch, K. M. (2013) Biosynthesis of natural products by microbial iterative hybrid PKS-NRPS. *RSC Adv. 3*, 18228−18247.
(22) Maiya, S., Grundmann, A., Li, X., Li, S.-M., and Turner, G. (2007) Identification of a hybrid PKS/NRPS required for pseurotin A biosynthesis in the human pathogen *Aspergillus fumigatus*. *ChemBioChem 8*, 1736−1743.
(23) Liu, X., and Walsh, C. T. (2009) Cyclopiazonic acid biosynthesis in *Aspergillus* sp.: characterization of a reductase-like R* domain in cyclopiazonate synthetase that forms and releases cyclo-acetoacetyl-L-tryptophan. *Biochemistry 48*, 8746−8757.
(24) Sims, J. W., and Schmidt, E. W. (2008) Thioesterase-like role for fungal PKS-NRPS hybrid reductive domains. *J. Am. Chem. Soc. 130*, 11149−11155.
(25) Ma, H., Kunes, S., Schatz, P. J., and Botstein, D. (1987) Plasmid construction by homologous recombination in yeast. *Gene 58*, 201−216.
(26) Weld, R. J., Plummer, K. M., Carpenter, M. A., and Ridgway, H. J. (2006) Approaches to functional genomics in filamentous fungi. *Cell Res. 16*, 31−44.
(27) Seshime, Y., Juvvadi, P. R., Tokuoka, M., Koyama, Y., Kitamoto, K., Ebizuka, Y., and Fujii, I. (2009) Functional expression of the *Aspergillus flavus* PKS−NRPS hybrid CpaA involved in the biosynthesis of cyclopiazonic acid. *Bioorg. Med. Chem. Lett. 19*, 3288−3292.
(28) Auclair, K., Sutherland, A., Kennedy, J., Witter, D. J., Van den Heever, J. P., Hutchinson, C. R., and Vederas, J. C. (2000) Lovastatin nonaketide synthase catalyzes an intramolecular Diels−Alder reaction of a substrate analogue. *J. Am. Chem. Soc. 122*, 11519−11520.
(29) Xu, W., Chooi, Y. H., Choi, J. W., Li, S., Vederas, J. C., Da Silva, N. A., and Tang, Y. (2013) LovG: The thioesterase required for dihydromonacolin L release and lovastatin nonaketide synthase turnover in lovastatin biosynthesis. *Angew. Chem., Int. Ed. Engl. 52*, 6472−6475.
(30) Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., and Mesirov, J. P. (2011) Integrative genomics viewer. *Nat. Biotechnol. 29*, 24−26.
(31) Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., and Pevzner, P. A. (2012) SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol. 19*, 455−477.
(32) Stanke, M., Steinkamp, R., Waack, S., and Morgenstern, B. (2004) AUGUSTUS: A web server for gene finding in eukaryotes. *Nucleic Acids Res. 32*, W309−W312.
(33) Jadulco, R. C.; Koch, M.; Kakule, T. B.; Schmidt, E. W.; Orendt, A.; He, H.; Janso, J. E.; Carter, G. T.; Larson, E. C.; Pond, C.; Matainaho, T. K.; Barrows, L. R. Isolation of pyrrolocins A−C: *cis*- and *trans*-decalin tetramic acid antibiotics from an endophytic fungal-derived pathway. *J. Nat. Prod.* 2014, *Submitted*.

Supporting Information

Native promoter strategy for high-yielding synthesis and engineering of fungal secondary metabolites

Thomas B. Kakule,[1] Raquel C. Jadulco,[2] Michael Koch,[2] Jeffrey E. Janso,[3] Louis R. Barrows,[2] and Eric W. Schmidt[1*]

[1]Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112, USA

[2]Department of Pharmacology and Toxicology, University of Utah, Salt Lake City, Utah 84112, USA

[3]Natural Products, Worldwide Medicinal Chemistry, Pfizer Worldwide Research and Development, Groton, Connecticut 06355, USA

*Corresponding author: ews1@utah.edu

Table of Contents

Description of plasmid construction

Fungal expression vectors

FH-1. The 1.5 kbp promoter sequence of *eqxS* was amplified from FusWT genomic DNA with primer pair HygB2promoter-F / HygB2promoter-R. The eqxS NRPS sequence together with 500 bp downstream (terminator region) was amplified with primer pair vector2ACATR-F / vector2ACATR-R. Yeast recombination was used to construct FH-1 by transforming *S. cerevisiae* BY4741 with both PCR amplicons together with an 8 kbp NotI/BamHI fragment from pHygB-EqiACATR.

FH-2. This vector was constructed in two parts. First, the *sgfp* sequence was amplified from gGFP with primer pair PmeI-sGFP-F / sGFP-R and cloned by yeast recombination into PmeI-linearized FH-1 to make plasmid pHygB-PmeI-sGFP. Then, the *eqxR* sequence together with the *alcA* promoter were amplified from the alcAeqxR plasmid [1]. This amplicon was transformed into NotI-linearized pHygB-PmeI-sGFP by yeast recombination to make plasmid FH-2.

FH-3. To extend the *eqxS* promoter and include eqxC, PCR was used to amplify the sequence with primer pair deg3prom2-F / ER2vect-R from FusWT genomic DNA. This amplicon was cloned into NotI-digested FH-1 by yeast recombination to form plasmid pHygB-ER+deg3ACATR. This plasmid was then cut with PmeI and joined with sGFP amplicon obtained from gGFP with primer pair PmeI-sGFP-F / sGFP-R by recombination in yeast to make plasmid pHygB-deg3ER+sGFP. FH-3 was constructed through yeast recombination of NotI-digested pHygB-deg3ER+sGFP with PCR product from alcAeqxR with primer pair TrpC2alcA-F / eqxR2eqxC.

FH-4. The *pyrG* sequence was amplified from FusWT genomic DNA with primer pair pyrG2vect-F / pyrG2vect-R and joined with AscI- / NheI-linearized pHygB-deg3ER+sGFP by yeast recombination to make pPyrG-deg3ER+sGFP. This plasmid was then linearized with KpnI and recombined with PCR product amplified from alcAeqxR plasmid with primer pair pyrG2alcA-F/alcAg10517-R[1] to make plasmid FH-4.

Fungal expression plasmids

hphgfp. The sGFP sequence was amplified from plasmid gGFP with primer pair sGFP-F / sGFP-R. The amplicon was fused with PmeI-linearized FH-1 via yeast-mediated recombination to afford plasmid hphgfp.

hphCpas. The *cpaS* sequence was amplified from *Aspergillus flavus* AF293 genome by PCR with primer pair cpaS2deg3-F / cpaS2deg3-R. The amplicon was cloned into PmeI-linearized FH-1 by yeast recombination.

hphCpasgfp. The *cpaS* sequence was amplified by PCR from A. flavus AF293 genomic DNA with primer pair cpaS2deg3-F / cpasFuLL2gfp-R and cloned into PmeI-linearized pHygB-PmeI-sGFP.

hpheqxC+eqxSgfp. The *eqxS* coding sequence was amplified in two fragments. First the NRPS region was amplified with primer pair vector2ACATR-F/ eqxSNRPS-sGfP-R and cloned into FH-3. After linearizing the resultant plasmid with PmeI, the second fragment amplified with primer pair deg3vect_1f / deg3vect_2r was introduced.

hphlovC+eqxSgfp. The *lovC* sequence was amplified from a yeast expression plasmid pMet25-lovC with primer pair lovC2eqxvect-F / lovC2eqxvect-R and subcloned into another plasmid. From this cloning plasmid, the *lovC* containing fragment was obtained by restriction digest. Together with the *eqxR* sequence amplified with primer pair TrpC2alcA-F / eqxR2lovC-R, *lovC* fragment was recombined with

NotI-cut pHygB-PmeI-sGFP. Into this resultant plasmid hphlovC+sGFP, the *eqxS* sequence was cloned as described above for the hpheqxC+eqxSgfp plasmid by yeast recombination.

hpheqxC+lovBgfp. We amplified the coding sequence of *lovB* from YEpADH2p-LovB-His plasmid (from Tsai Lab, UCI) with primer pair deg3LovBPKS-F / LovB2GfP-R, and cloned it into PmeI-cut FH-3 by yeast recombination.

hphlovC+lovBgfp. The hphlovC+sGFP was linearized with PmeI and the *lovB* sequence introduced similarly to the hpheqxC+lovBgfp sequence above.

pyrGprlC+prlS. This plasmid was constructed in two steps. First, the trans-ER, *prlC* was amplified from NRRL 50135 genomic DNA with primer pair g13602-F / g13602-R and cloned into NotI-digested FH-4 by yeast recombination to make plasmid pyrGprlC+sGFP. Then, the new plasmid was linearized with PmeI and recombined with the *prlS* sequence amplified as two fragments with primer pairs g13593-F / g13593-1R and g13593-2F / g13593-R to make the final expression plasmid pyrGprlC+prlS.

hphprlD. The *prlD* sequence was amplified from NRRL 50135 genomic DNA with primer pair NRRL_MT-F / NRRL_MT-R. This amplicon was cloned into PmeI-digested FH-2 to make the hphprlD plasmid.

hpheqxD. The *eqxD* gene sequence was amplified from FusWT genomic DNA with primer pair deg3prom_eqxD-F / deg3prom_eqxD-R and cloned into PmeI-digested FH-2 to make the hpheqxD plasmid.

Knockout plasmids

ClusterPhleoKO. Initially, the *eqxS* PKS sequence was amplified with primer pair pRS2deg3PKS-F / pRS2deg3PKS-R and cloned into vector pRS316 by yeast recombination to make plasmid pRSdeg3PKS. The 12 kbp BlpI / BsrGI fragment of pRSdeg3PKS was fused with a 3.8 kbp ClaI / NdeI fragment from TOPO-PKSPhleoKO, which contains a phleomycin resistance marker, to make plasmid pRSdeg3PhleoKO. ClusterPhleoKO was then constructed by transforming BlpI-linearized pRSdeg3PhleoKO with MTKO2 amplicon[1] into *S. cerevisiae* BY4741 to replace the 5' PKS sequence of *eqxS* with a fragment of *eqxD*.

pRSpyrGKO. The pyrG sequence was amplified in two fragments with primer pairs: PyrKO-XbaI-F / PyrKO-R, and PyrKO-F / PyrKO-XhoI-R. A randomly selected sequence, the first exon of *lovC*, was amplified with primer pair LovCexpress-F/LovCex1-R from the *Aspergillus terreus* genomic DNA. Yeast recombination was utilized to make the pRSpyrGKO by flanking the *lovC* sequence with the two *pyrG* fragments joined with XbaI / XhoI-linearized pRS316 vector.

Table S1: Plasmids used in this study

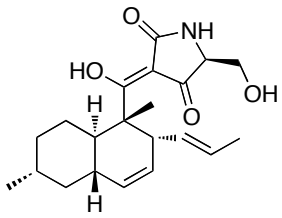
Plasmid	**Source**
FH-1	This study
FH-2	This study
FH-3	This study
FH-4	This study
gGFP	Fungal Genetic Stock Center
hphgfp	This study
hphCpas	This study
hphCpasgfp	This study
pRSdeg3PKS	This study
pRSdeg3PhleoKO	This study
TOPO-PKSPhleoKO	This study
ClusterPhleoKO	This study
pRSpyrGKO	This study
hpheqxC+eqxSgfp	This study
hpheqxC+eqxS	This study
hphlovC+eqxSgfp	This study
hphlovC+lovBgfp	This study
hpheqxC+lovBgfp	This study
pyrGprlC+prlS	This study
hphprlD	This study
hpheqxD	This study
pHygB-EqiACATR	Kakule T. B., et al.[1]
pHygB-PmeI-sGFP	This study
hphlovC+sGFP	This study
YEpADH2p-lovB-His	Tsai Lab, University of California-Irvine

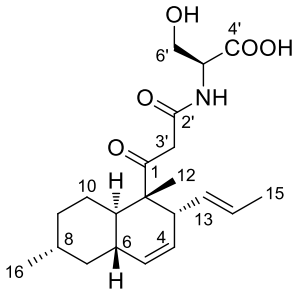
Table S2: Primers used in this study.

Primer name	Sequence
HygB2promoter-F	GTACACTTGTTTAGAGGTAATCCTTCTTTCTAGACGGCGGCCGCCCTTTCCCACAAACTCGACA
HygB2promoter-R	CGGCACAGTAGATGTCTTTCCGTTTAAACTGTCAAGCAAGCTGCAATGGG
vector2ACATR-F	CCCATTGCAGCTTGCTTGACAGTTTAAACGGAAAGACATCTACTGTGCCG
vector2ACATR-R	TTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTAATTAACCTTCCAGATCATCCGCGAG
sGFP-F	TTGCAGATTTTATTACCGTCCCATTGCAGCTTGCTTGACAATGGTGAGCAAGGGCGAGG
sGFP-R	TAATCCATTATACCAAGTTGTGTGCCAACCTTAAATTACTTGTACAGCTCGTCCATG
cpaS2deg3-F	TTGCAGATTTTATTACCGTCCCATTGCAGCTTGCTTGACAATGAGGATGCCAATTGCTGTTATT
cpaS2deg3-R	TAATCCATTATACCAAGTTGTGTGCCAACCTTAAATTATTAATTACCCAAGTAGGGATATCG
pRS2deg3PKS-F	GCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCAGACAGGAGGAAGTGCA
pRS2deg3PKS-R	GGCTCCTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCCTTCCAAGCCAAGCCAAAC
PyrKO-XbaI-F	ACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCACATATCGGATCAATTCCGATG
PyrKO-XhoI-R	TGACCATGATTACGCCAAGCTCGGAATTAACCCTCGGCGCGCCCGCTAGATGTCTCGATAACTC
PyrGKO-F	GTCCGGATATATTCCAATTGCAACATGCTCCCCCCACAGGCTTCTGTTACATCTCTGGC
PyrGKO-R	GGATGGGGGTAATAGAATTGTATCTATGTATCTGACGATCATCCGGGAACCATGTTCACATTG
LovCexpress-F	TCGTCAGATACATAGATACAATTCTATTACCCCCATCCATGGGCGACCAGCCATTCATT
LovCex1-R	TGTGGGGGGAGCATGTTGCAATTGGAATATATCCGGACAGGCGGAGCATTTGCATAGTG
PmeI-sGFP-F	TTTATTACCGTCCCATTGCAGCTTGCTTGACAGTTTAAACATGGTGAGCAAGGGCGAGG
CpasFuLL2gfp-R	ACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACATTACCCAAGTAGGGATATCGCAA
deg3prom2-F	CCATGACTCATTGATTCTGAGTAAG
ER2vect-R	TGTTTAGAGGTAATCCTTCTTTCTAGACGGCGGCCGCTTAAAGAGTGAATACAAGTTTCTTTCC
TrpC2alcA-F	AGGTACACTTGTTTAGAGGTAATCCTTCTTTCTAGACGTCTGCGATCGTCCATAACCGTT
eqxR2eqxC	ATCGGGAAAGAAACTTGTATTCACTCTTTAAGCGGCCGCTCATTAAAGCTTCATCCTTTCCCTTA
pyrG2vect-F	CAAAAGAATAGACCGAGATAGGGTTGAGTGTTGGCGCGCCTACATATCGGATCAATTCCGATG
pyrG2vect-R	GTTTTATTCTTGTTGACATGGAGCTATTAAATCAGGTACCTACCTACTCCAGAGCTATAATGC

pyrG2alcA-F	GATTCAAGATACAGCATTATAGCTCTGGAGTAGGTATCTGCGATCGTCCATAACCGTTCA
g13593-F	GCAGATTTTATTACCGTCCCATTGCAGCTTGCTTGACAATGTCTTCGACAGAGCCTATTG
g13593-1R	GGCACATCTTGCAATTGCACGAT
g13593-2F	GGAGTTCGATCGTGGTGCAC
g13593-R	CTAATCCATTATACCAAGTTGTGTGCCAACCTTAAATTACTAACGCCTGGTGACTAGCTC
g13602-F	ACTATAGACATATTCCAAGACACGGATTACTCAAGTCACTATGGGAAGCCTGGAGAAGGCT
g13602-R	GTACACTTGTTTAGAGGTAATCCTTCTTTCTAGACGTCATTGTGACAGCTTACATACCATCT
deg3prom_eqxD-F	GCAGATTTTATTACCGTCCCATTGCAGCTTGCTTGACAATGTCATCTATCCTTTCGCGTT
deg3prom_eqxD-R	CCATTATACCAAGTTGTGTGCCAACCTTAAATTATCAACTCTGTACAGGTAGCATCTC
NRRL_MT-F	GCAGATTTTATTACCGTCCCATTGCAGCTTGCTTGACAATGGAGCCCCAGCCTCCGCG
NRRL_MT-R	CCATTATACCAAGTTGTGTGCCAACCTTAAATTACATCAAGACCAATTCAAGTGCAGCGA
deg3prom_eqxD-F	GCAGATTTTATTACCGTCCCATTGCAGCTTGCTTGACAATGTCATCTATCCTTTCGCGTT
deg3prom_eqxD-R	CCATTATACCAAGTTGTGTGCCAACCTTAAATTATCAACTCTGTACAGGTAGCATCTC
pyrGseq-F	CGAGACTGAGTATCGAGATAGTG
pyrGseq-R	CGCTAGATGATGTCTCGATAACTC
eqxSNRPS-sGfP-R	ACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACACGCCTTGAAAATAGCTGTGATC
deg3vect_1f	TCCTTCCAAGCCAAGCCAAAC
deg3vect_2r	CCAGACAGGAGGAAGTGCA
deg3LovBPKS-F	ATTTTATTACCGTCCCATTGCAGCTTGCTTGACAATGGCTCAATCTATGTATCCTAATGAG
LovB2sGfP-R	ACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACTGCCAGCTTCAGGGCGGGAT
lovC2eqxvect-F	ATAGACATATTCCAAGACACGGATTACTCAAGTCACTATGGGCGACCAGCCATTCATTC
lovC2eqxvect-R	CACTTGTTTAGAGGTAATCCTTCTTTCTAGACGGCGGCCGCTTACGGCCCCTCGAGCCGAA
eqxR2lovC-R	GAAACTCGTGGTTCGGCTCGAGGGGCCGTAAGCGGCCGCTCATTAAAGCTTCATCCTTTCCCTTA
deg3KOScreen-F	GGACGGTCAGTTTATTTCAGATG
deg3KOScreen-R	CGTCCAAGAGACTCGCTGAA

Table S3: NMR data of **4** (right) compared to **3** (left) in methanol-d_4.





	Previously reported data for Compound 3 [1]		Compound 4	
No.	**δ ^{13}C**	**δ ^{1}H, multiplicity, J (Hz)**	**δ ^{13}C**	**δ ^{1}H, multiplicity, J (Hz)**
1	204.6, C		210.1	
2	51.5, C		54.7	
3	46.2, CH	3.46, br	50.5	2.68, br
4	133.4, CH	5.15, m	132.0	5.12, m
5	132.6, CH	5.41, m	131.5	5.44, m
6	39.9, CH	1.86, m	39.7	1.79, m
7	43.6, CH_2	1.89, m; 0.89, m	43.3	1.84, m; 0.85, m
8	34.8, CH	1.53, m	34.6	1.51, m
9	36.9, CH_2	1.77, m; 1.11, m	36.8	1.75, m; 1.06, m
10	29.5, CH_2	2.01, br; 1.08, br	28.3	1.74, br; 1.01, br
11	41.3, CH	1.68, m	41.0	1.62, m
12	14.2, CH_3	1.45, s	16.9	1.23, s
13	128.0, CH	5.37, m	128.2	5.44, m
14	127.9, CH	5.26, m	127.6	5.38, m
15	17.8, CH_3	1.52, d, 5.7	18.1	1.59, d, 5.7
16	23.0, CH_3	0.94, d, 6.8	22.9	0.93, d, 6.8
2'	nd[a]		169.2	
3'	nd		46.4[b]	3.27[b], d, 16; 3.58[b], d, 16
4'	nd		173.2	
5'	nd	nd	56.0	4.5, dd
6'	62.2, CH_2	3.79,m; 3.74, m	62.9	3.93,dd; 3.83, dd

[a]nd = not detected

[b]These were not detected in methanol-d_4, but were only observed in dmso-d_6 after proton exchange (see Figure S20).

Figure S1: FH-1 vector map

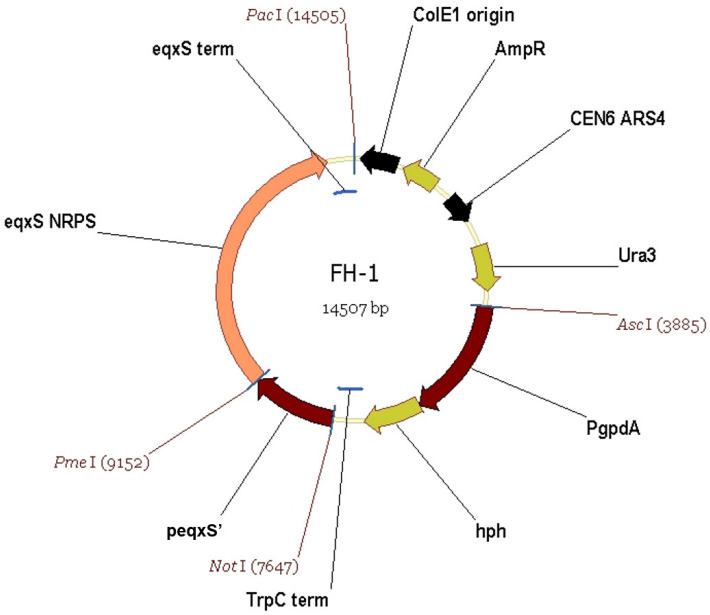


Figure S2: FH-2 vector map

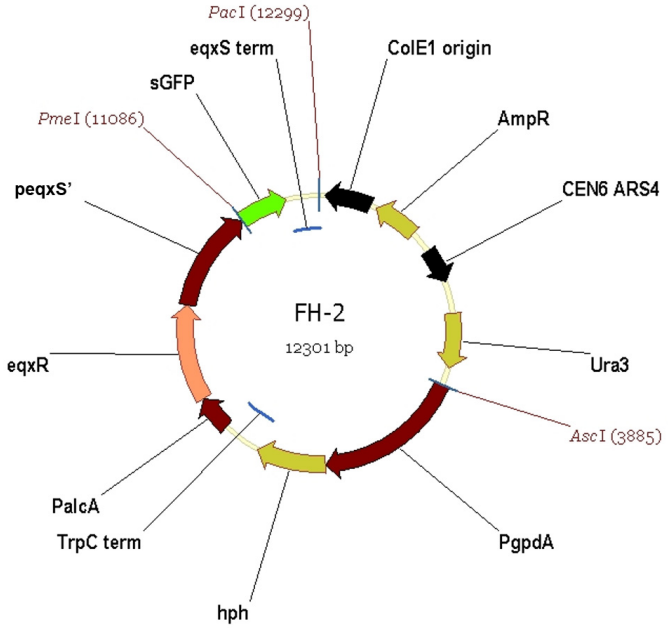


Figure S3: FH-4 vector map

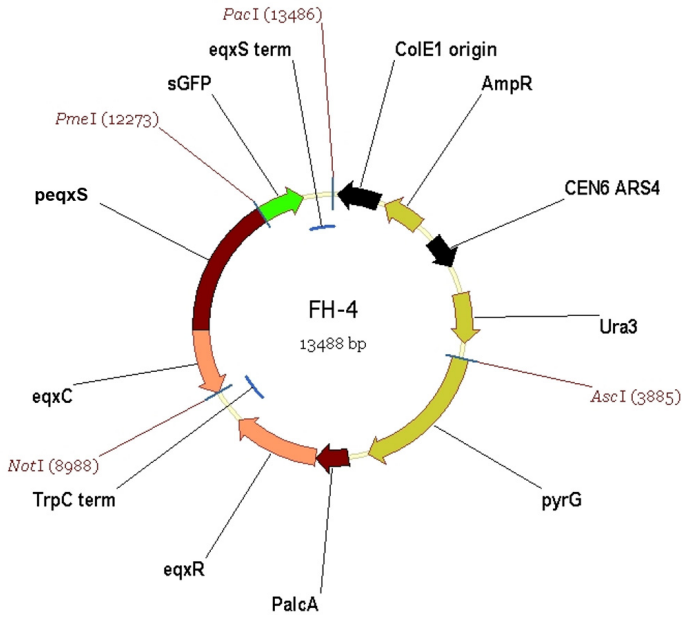


Figure S4: Diagnostic PCR analysis for eqx knockout mutants. Colony PCR done on isolated mutants using primer pairs deg3KOScreen-F / deg3KOScreen-R and MTKOScreen-F / MTKOScreen-R[1] to screen for absence of regions marked red and green, respectively. Lane labelled 1 is the FusWT control; 2 and 3 denote mutant FusΔeqx5 (in duplicate); 4 and 5 denote mutant FusΔeqx8; and lane 6 is the ectopic integrant (resistant to both hygromycin and phleomycin). FusΔeqx5 was sequenced and selected for further experiments.

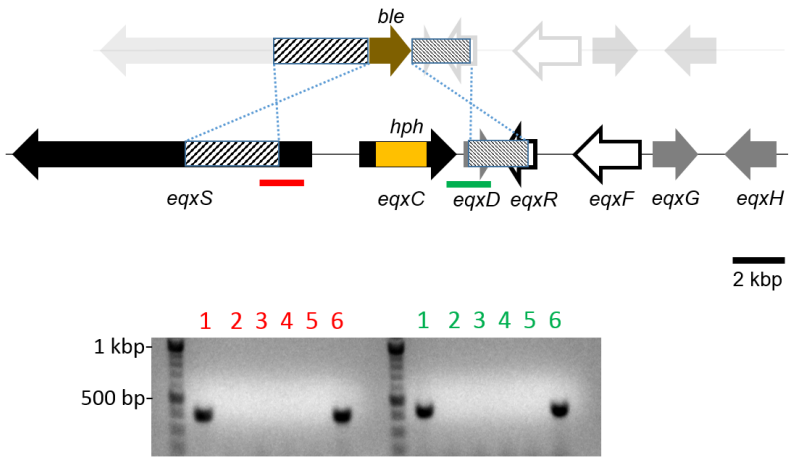


Figure S5: Alignment of Illumina sequencing reads of FusΔeqx5 against FusWT reference genome. As expected, no single read was present within the targeted deletion region. Coverage for the overlapping homologous regions in the knockout cassette was 3 times that for the rest of the genome which corresponds to 3 integration events.

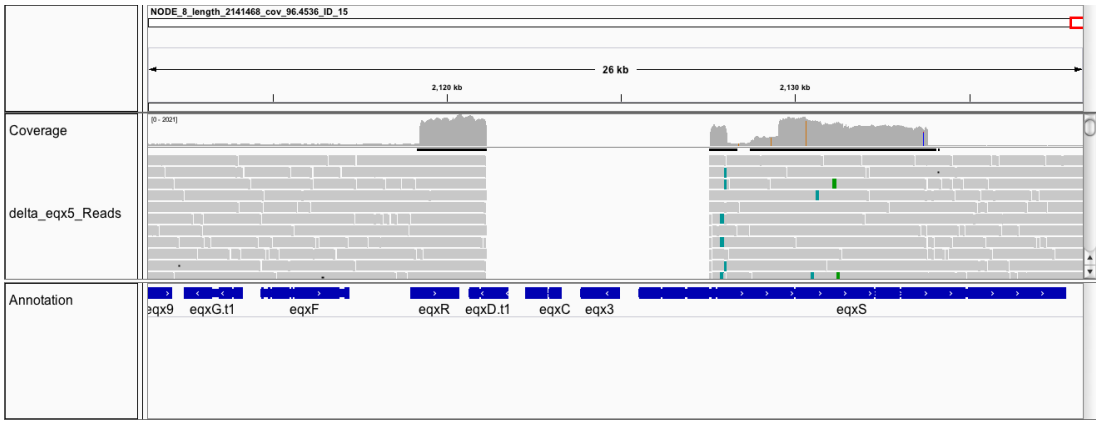


Figure S6: Diagnostic PCR analysis to screen for knockout of *pyrG* from mutant FusΔeqx5. Colony PCR done on transformed mutants that were isolated from 5-FOA containing medium with primer pair pyrGseq-F / pyrGseq-R. Homologous integrants (clones 8, 9, 10, and 11) show a 640 bp band shift from wild-type control. Clone 10 was designated FusΔeqx5ΔpyrG10.

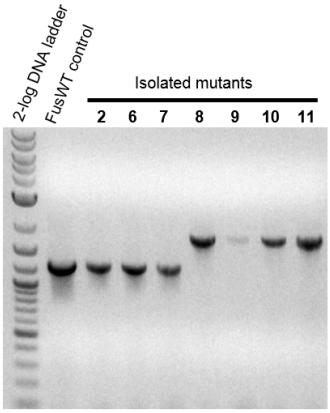


Figure S7: LC/MS and UV data of cAATrp **1**.

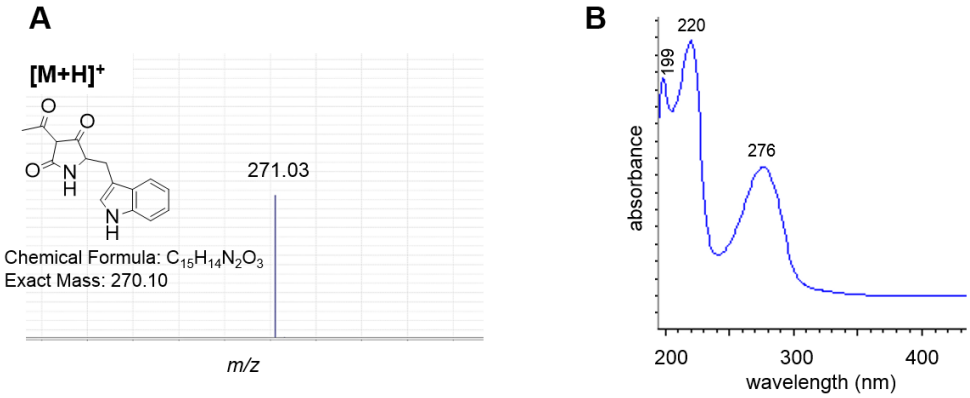


Figure S8: LC/MS and UV data of **5**.

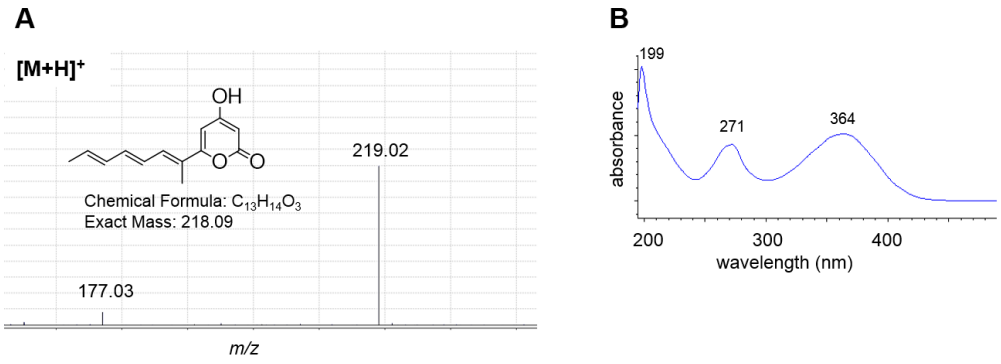


Figure S9: LC/MS and UV data of **6**.

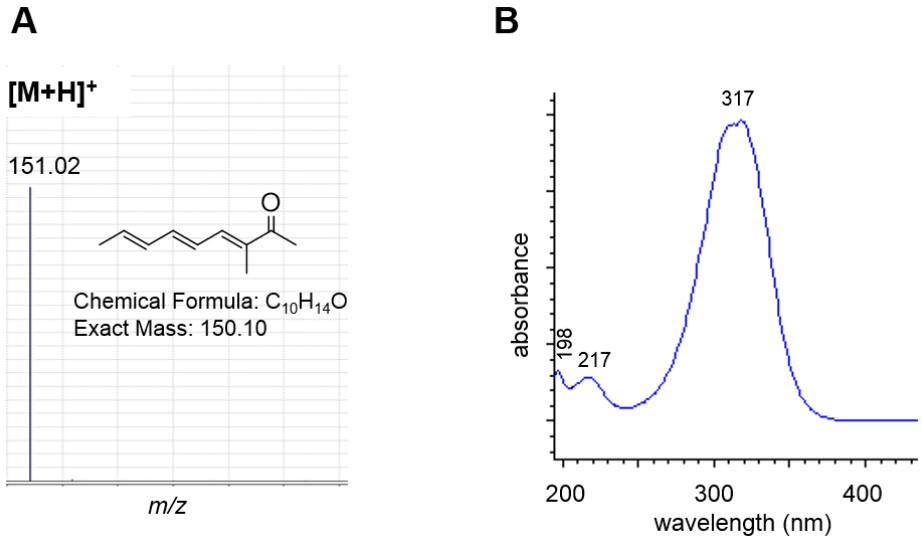


Figure S10: LC/MS and UV data of **7**.

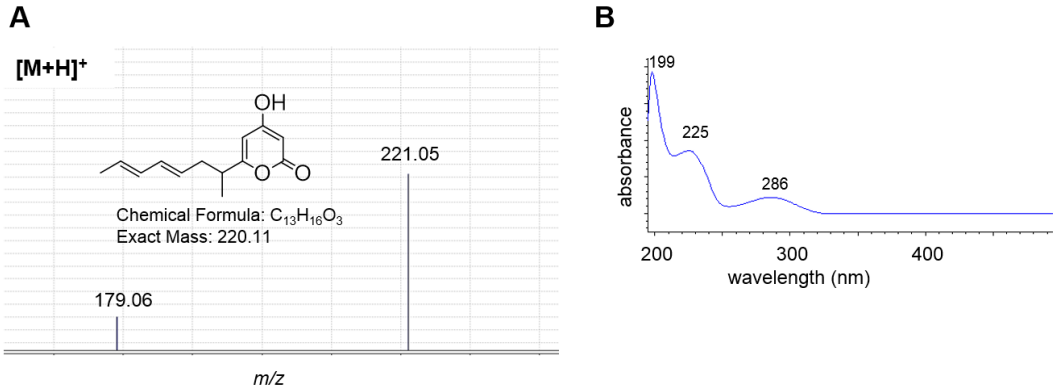


Figure S11: LC/MS data of **4** (A) and trichosetin **3** (B).

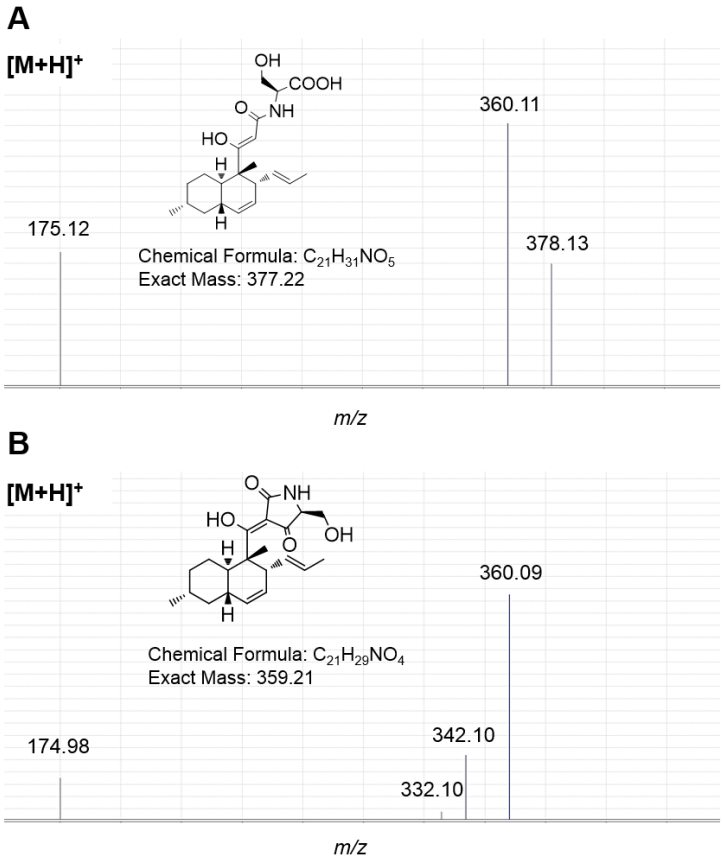


Figure S12: LC/MS data of **9** (A) and **8** (B).

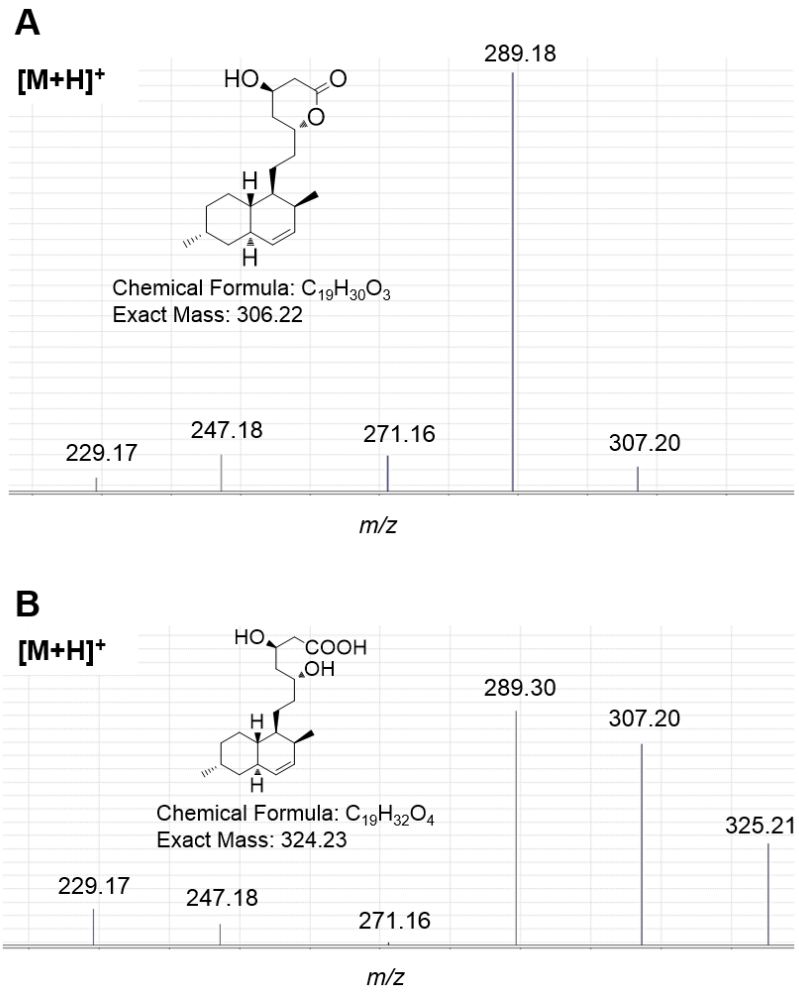


Figure S13: ^{1}H spectrum of cAATrp **1** in $CDCl_3$

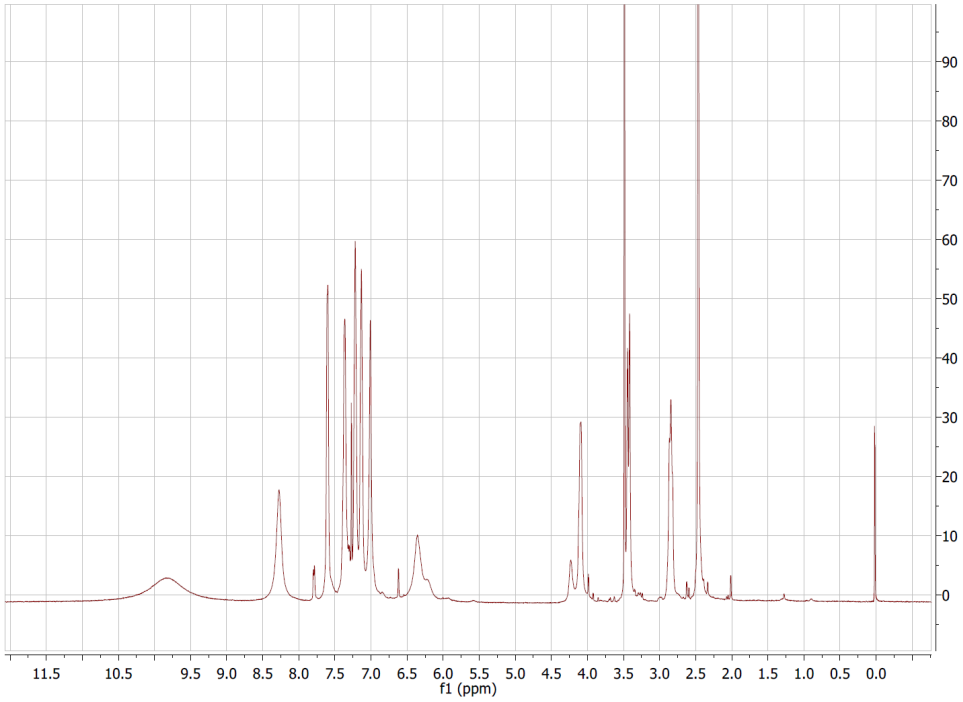


Figure S14: ^{1}H spectrum of **9** in dmso-d_6

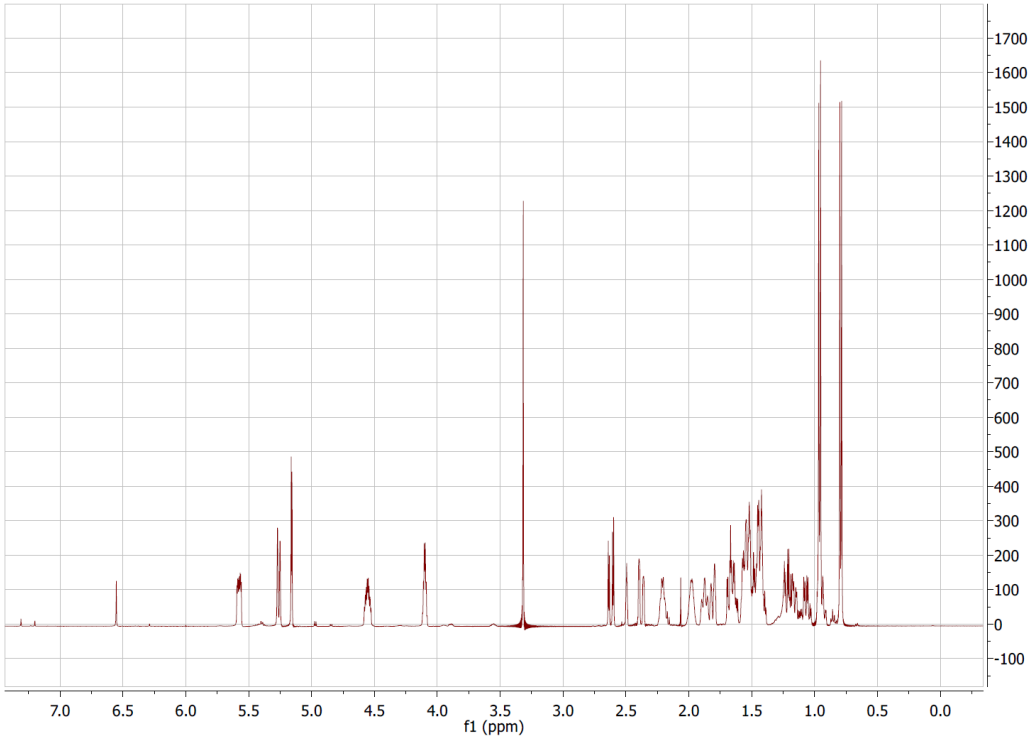


Figure S15: ^{1}H spectrum of **4** in methanol-d_4.

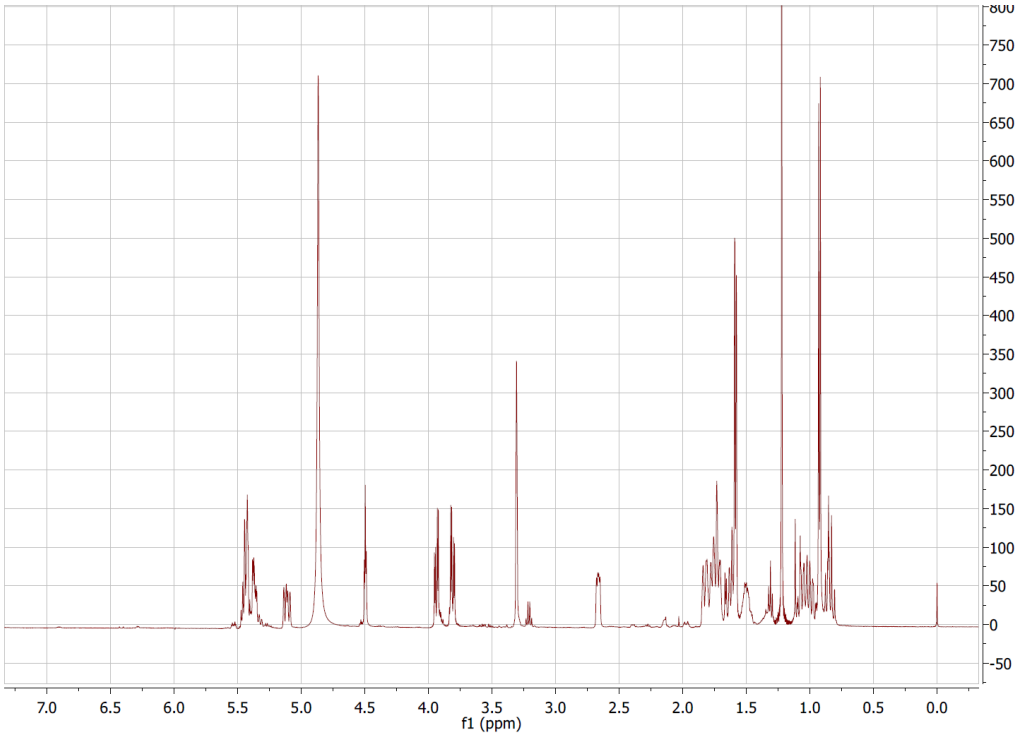


Figure S16: ^{13}C spectrum of **4** in methanol-d_4.

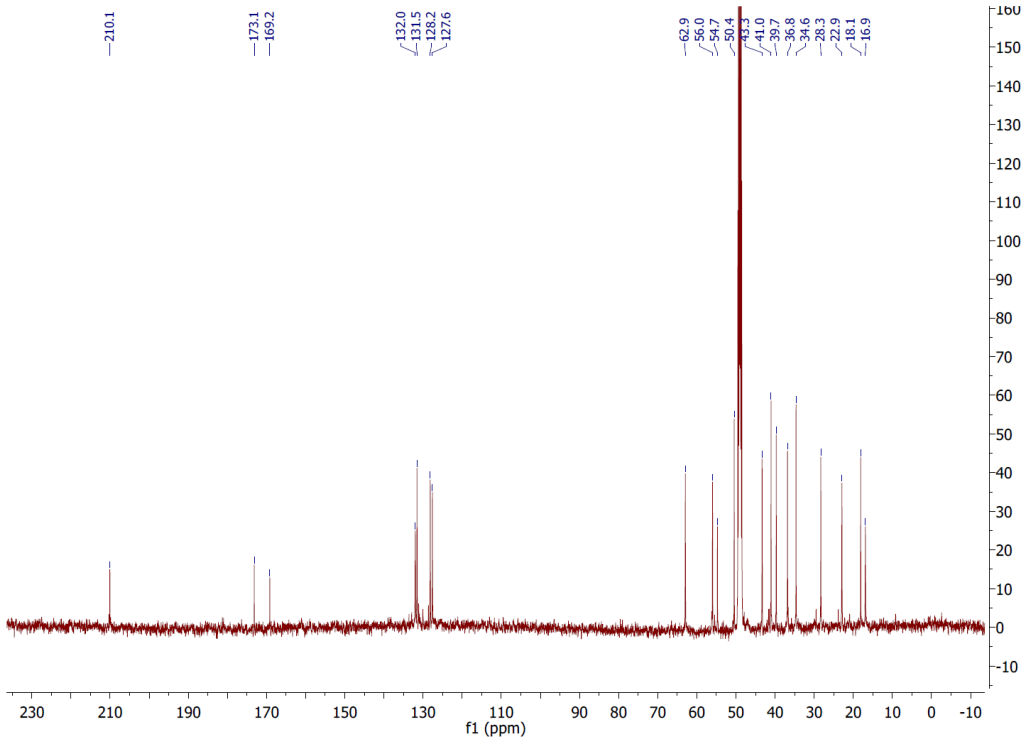


Figure S17: ^{1}H-^{1}H COSY spectrum of **4** in methanol-d_4.

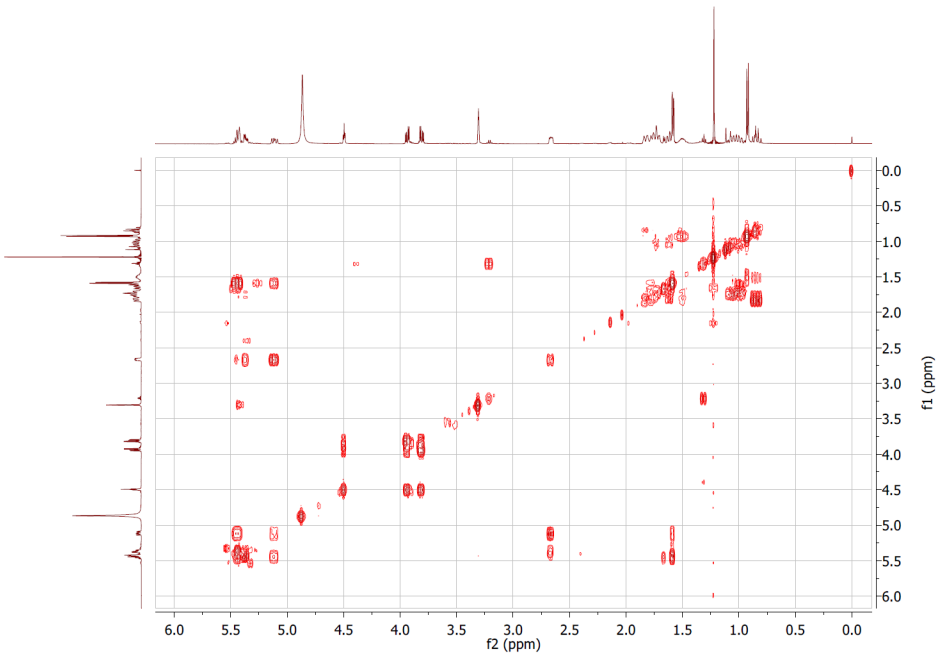


Figure S18: ^{1}H-^{13}C HSQC spectrum of **4** in methanol-d_4.

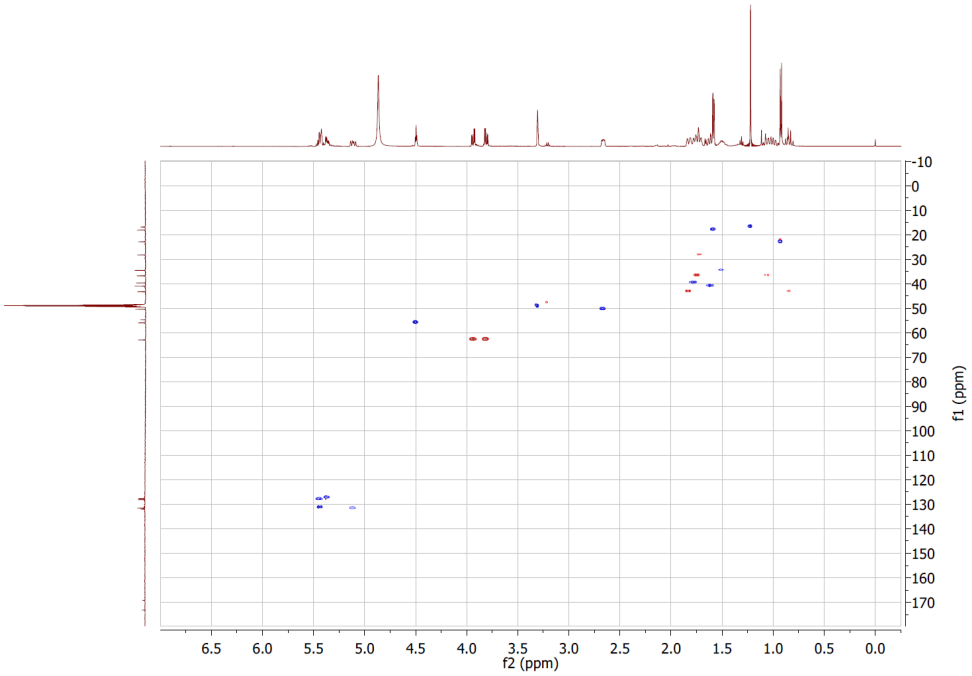


Figure S19: ^{1}H-^{13}C HMBC spectrum of **4** in methanol-d_4.

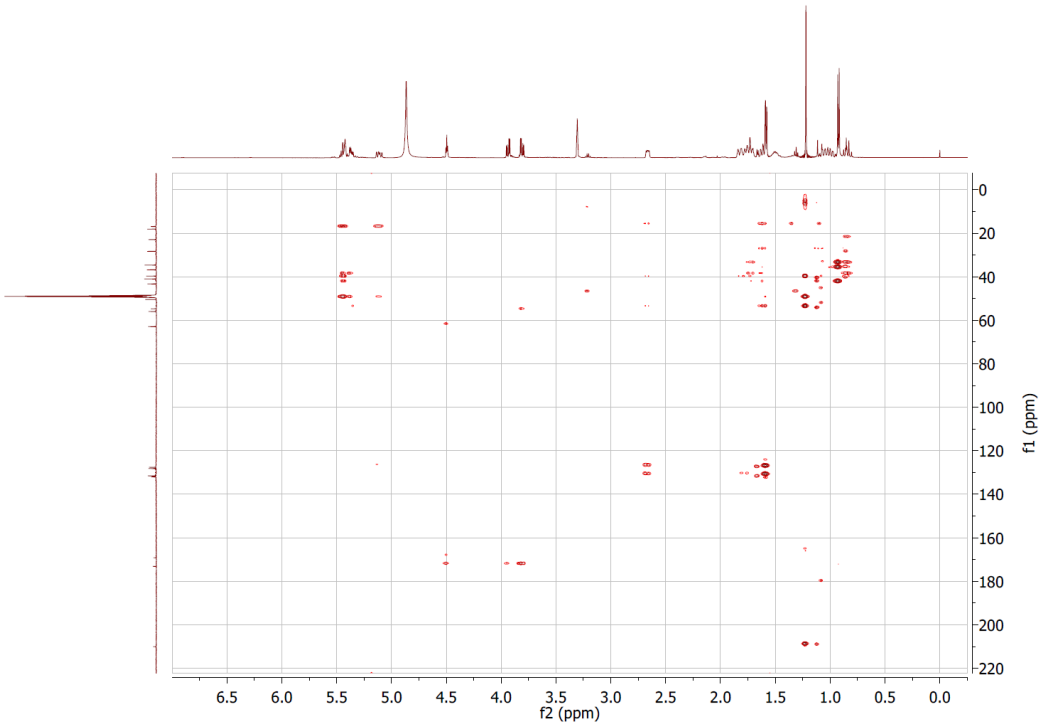
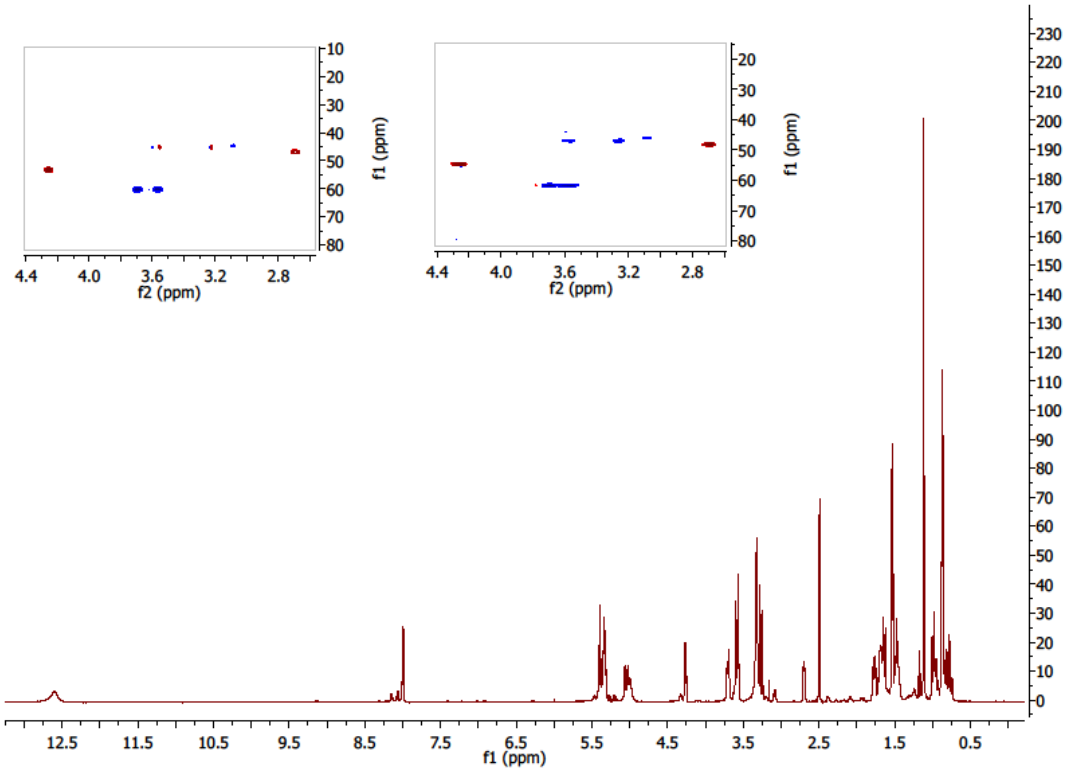


Figure S20: ^{1}H NMR spectrum of **4** in dmso-d_6. Inset are the 2D HSQC NMR spectra in dmso-d_6 after overnight incubation in methanol-d_4 (left), and after a 5 minute exchange in methanol (right). An increase in the signal from the 3' position protons is observed (H-3a', δ 3.27; H-3b', δ 3.58; C-3', δ 46.4).



References

1. Kakule, T. B., Sardar, D., Lin, Z., and Schmidt, E. W. (2013) Two Related Pyrrolidinedione Synthetase Loci in Fusarium heterosporum ATCC 74349 Produce Divergent Metabolites, *ACS Chem. Biol. 8*, 1549-1557.

CHAPTER 4

ISOLATION OF PYRROLOCINS A–C: *cis*- AND *trans*-DECALIN TETRAMIC ACID ANTIBIOTICS FROM AN ENDOPHYTIC FUNGAL-DERIVED PATHWAY

Manuscript reproduced with permission from
Jadulco, R. C.; Koch, M.; Kakule, T. B.; Schmidt, E. W.; Orendt, A.; He, H.; Carter, G. T.; Larson, E. C.; Pond, C.; Matainaho, T. K.; Barrows, L. R. Isolation of Pyrrolocins A–C: *cis*- and *trans*-Decalin Tetramic Acid Antibiotics from an Endophytic Fungal-Derived Pathway. *J. Nat. Prod.* **2014**, DOI: 10.1021/np500617u.

Note: my contribution to this paper was in developing all the recombinant strains and in performing the initial mass spectrometry analysis of the secondary metabolites.

JOURNAL OF NATURAL PRODUCTS

ACS Editors' Choice

Article

pubs.acs.org/jnp

Terms of Use

Isolation of Pyrrolocins A−C: *cis*- and *trans*-Decalin Tetramic Acid Antibiotics from an Endophytic Fungal-Derived Pathway

Raquel C. Jadulco,[†,#] Michael Koch,[†] Thomas B. Kakule,[‡] Eric W. Schmidt,[‡] Anita Orendt,[Δ] Haiyin He,[§] Jeffrey E. Janso,[§] Guy T. Carter,[∥] Erica C. Larson,[†] Christopher Pond,[†] Teatulohi K. Matainaho,[⊥] and Louis R. Barrows*[,†,⊥]

[†]Department of Pharmacology and Toxicology, [‡]Department of Medicinal Chemistry, and [Δ]Center for High Performance Computing, University of Utah, Salt Lake City, Utah United States

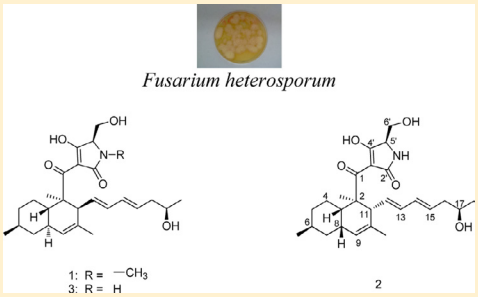
[§]Natural Products − Worldwide Medicinal Chemistry, Pfizer Worldwide Research and Development, 445 Eastern Point Road, Groton, Connecticut 06340, United States

[∥]Carter-Bernan Consulting, 350 Phillips Hill Road, New City, New York 10956, United States

[⊥]School of Medicine and Health Sciences, University of Papua New Guinea, Boroko, NCD, Papua New Guinea

Supporting Information

ABSTRACT: Three new decalin-type tetramic acid analogues, pyrrolocins A (**1**), B (**2**), and C (**3**), were defined as products of a metabolic pathway from a fern endophyte, NRRL 50135, from Papua New Guinea. NRRL 50135 initially produced **1** but ceased its production before chemical or biological evaluation could be completed. Upon transfer of the biosynthetic pathway to a model host, **1**−**3** were produced. All three compounds are structurally related to equisetin-type compounds, with **1** and **3** having a *trans*-decalin ring system, while **2** has a *cis*-fused decalin. All were active against *Mycobacterium tuberculosis*, with the *trans*-decalin analogues **1** and **3** exhibiting lower MICs than the *cis*-decalin analogue **2**. Here we report the isolation, structure elucidation, and antimycobacterial activities of **1**−**3** from the recombinant expression as well as the isolation of **1** from the wild-type fungus NRRL 50135.



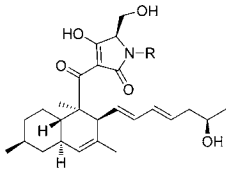
Papua New Guinea is a hot spot for plant biodiversity, harboring an estimated 20 000 individual species of vascular plants. An estimated 60% of these are endemic to Papuasia, one of the highest rates of endemism in the world.[1] Further, it is thought that most tropical plants contain multiple endophytic fungi,[2] providing an immense unexplored reservoir of secondary metabolites.

As part of our International Cooperative Biodiversity Group (ICBG) program, we screened extracts of endophytic fungi isolated from Papua New Guinea terrestrial and marine sources for antibacterial activities. A methanol extract of a phylogenetically novel strain (NRRL accession number 50135) isolated from the stem of an *Asplenium* sp. fern growing in Watunou, Milne Bay Province, Papua New Guinea, was identified as selectively active against Gram-positive bacteria including *Mycobacterium tuberculosis* H37Ra (ATCC 25177), *Staphylococcus aureus*, *Streptococcus pneumoniae*, and *Bacillus subtilis* ATCC6633. We employed an antimycobacterial bioassay-guided isolation to obtain the TB-active metabolite pyrrolocin A (**1**), a *trans*-fused decalin-containing tetramic acid analogue related to equisetin.[3] When NRRL 50135 stopped producing **1**, we used a recombinant expression platform[4] to induce its production. This heterologous expression platform successfully produced **1**, matching the initial molecular weight, UV, NMR, and CD spectra, as well as two additional desmethyl analogues, pyrrolocin B (**2**) and pyrrolocin C (**3**). The three compounds have different configurations in the decalin ring system, with **2** having the *cis*-fused decalin ring and **1** and **3** having the *trans*-decalin configuration. To our knowledge, this is the first report of both *cis* and *trans* configurations being isolated from a single biological source and may be valuable in understanding the biosynthesis of this group of tetramic acid analogues. Furthermore, transplanting, or indeed rescuing, synthetic pathways from one organism into more tractable backgrounds unquestionably furthers drug discovery by magnifying the number and amount of rare natural chemicals with potential pharmacological activities, i.e., accessing dormant biosynthetic potential with a concomitant increase in production. Skillful manipulation of these pathways could theoretically also steer biosynthesis of undesirable intermediates or end products to pharmacologically desirable entities in quantities useful for experimentation or production.

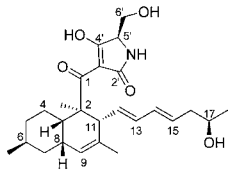
Received: August 1, 2014

dx.doi.org/10.1021/np500617u | *J. Nat. Prod.* XXXX, XXX, XXX−XXX

Tetramic acids are characterized by the presence of a 2,4-pyrrolidinedione moiety in their structures. Tetramic acids that are related to equisetin are particularly important because of their reported antimicrobial activities,[3,5−9] HIV-1 integrase inhibition,[10,11] and microtubule assembly inhibition.[12] The mechanism of their antimicrobial activities has been purported to be due to inhibition of SecA,[13] inhibition of undecaprenyl pyrophosphate synthase,[14] and inhibition of the histidine kinase WalK.[15] Due to their biological activities, the biosynthesis of this group of compounds and related polyketides was also the subject of interest in several reviews.[16−18]



1: R = $-CH_3$
3: R = H



2

RESULTS AND DISCUSSION

Initially, growth of the fungus on semisolid media led to the robust production of **1** at approximately 50 mg L^{-1}. The compound was purified on the basis of promising activity against *M. tuberculosis*. The methanol extract from NRRL 50135 was purified using a preparative C_{18} column to isolate **1**. High-resolution ESIMS determined its molecular formula to be $C_{27}H_{40}NO_5$. The ^{13}C NMR spectra (Table 1) showed 27 carbon signals, while its ^1H NMR spectra presented five olefinic protons, five methyls, four methylenes, and six methines. The carbon connectivities in the decalin ring and the olefinic side chain were established by COSY, HMBC, and HSQC. Their association with equisetin-type tetramic acid was evident from a review of literature data that showed that **1** is related to phomasetin, with the addition of two carbons in the olefinic side chain and with the presence of an *N*-methyl group. In order to establish the relative and absolute configuration of **1** and evaluate its antimycobacterial activity, we fermented NRRL 50135 for additional material. Unfortunately, as is often observed in secondary metabolites isolated from cultivated microorganisms, NRRL 50135 ceased production of **1** before completion of this work. In order to achieve production of **1**, we identified the biosynthetic locus for its production and

Table 1. ^1H and ^{13}C NMR Data for Compound 1 (400 and 100 MHz, respectively, δ in ppm), 2, and 3 (500 and 125 MHz, respectively)

	1[a]		**2**[b]		**3**[a]	
pos	δ_C, type	δ_H (J in Hz)	δ_C, type	δ_H (J in Hz)	δ_C, type	δ_H (J in Hz)
1	195.7, C		204.5, C		198.6,[c] C	
2	48.8, C		52.5, C		49.0, C	
2-Me	13.7, CH_3	1.33, s	19.2, CH_3	1.30, s	13.3, CH_3	1.32, s
3	39.3, CH	1.57, br dd (10.2, 10.2)	38.4, CH	2.72, br d (10.2)	40.4, CH	1.57, m
4	27.7, CH_2	1.89, m	24.6, CH_2	1.38, m	27.5, CH_2	1.93, m
		1.00, m		1.45, m		1.00, m
5	35.4, CH_2	1.70, m	36.6, CH_2	1.67, m	35.4, CH_2	1.71, m
		1.00, m		0.87, dd (13.4, 12.6, 2.9)		1.00, m
6	32.7, CH	1.47, m	29.8, CH	1.40, d (5.3)	32.9, CH	1.49, m
6-Me	22.2, CH_3	0.87, d (6.7)	23.2, CH_3	0.81, d (6.5)	22.1, CH_3	0.88, d (6.5)
7	42.0, CH_2	1.77, m	38.4, CH_2	1.56, d (14.0)	42.0, CH_2	1.77, m
		0.81, m		1.11, dt (12.3, 5.3)		0.82, dd (12.9, 11.7)
8	38.7, CH	1.77, m	36.7, CH	2.09, br s	38.7, CH	1.79, m
9	126.0, CH	5.18, s	126.8, CH	5.09, s	126.0, CH	5.19, s
10	130.9, C		133.7, C		131.6, C	
10-Me	22.2, CH_3	1.52, s	23.3, CH_3	1.63, s	22.4, CH_3	1.52, s
11	48.8, CH	3.14, m	43.5, CH	3.20, d (9.8)	48.4, CH	3.18, m
12	130.7, CH	5.21, dd (15.4, 9.8)	134.1, CH	5.50, dd (14.2, 10.5)	131.0, CH	5.21[d]
13	131.3, CH	5.69, dd (15.4, 10.4)	135.9, CH	6.04, dt (10, 3)	132.1, CH	5.72, dd (15.6, 10.0)
14	130.9, CH	5.86, dd, (15.4, 10.5)	136.1, CH	6.09, dt (10, 3)	131.4, CH	5.90, dd (15.0, 11.2)
15	129.5, CH	5.49, ddd, (15.4, 10, 7)	130.3, CH	5.63, pentet (7.5)	130.3, CH	5.51, dt (15.1, 7.0)
16	42.1, CH_2	2.09, ddd (13.8, 7, 7)	42.1, CH_2	2.24, pentet (6.4)	42.4, CH_2	2.09, m
		1.99, ddd (13.8, 7, 7)		2.18, pentet (6.4)		2.00, pentet (6.8)
17	65.6, CH	3.56, pentet (6.5)	68.7, CH	3.78, m	65.9, CH	3.59, m
18	22.7, CH_3	0.98, d (6.5)	23.1, CH_3	1.14, d (6.20)	23.1, CH_3	0.98, d (6.1)
2′	176.0,[c] C		180.8, C	179.2,[c] C		
3′	100.6,[c] C		100.0,[c] C	100.0,[c] C		
4′	190.6,[c] C		192.9, C	190.5, C		
5′	67.6, CH	3.77, m	62.4, CH	3.87, br s	63.3, CH	3.79, m
6′	58.0, CH_2	3.80, dd (12, 2)	64.3, CH_2	3.83, dd (11.4, 3)	60.4, CH_2	3.62, d (11.0)
		3.68, dd, (12, 1.8)		3.77, m		3.57, dd (12.2, 5.9)
N-Me	26.8, CH_3	2.93, s				

[a]DMSO-d_6. [b]CD_3OD. [c]Broad signal. [d]Overlapped with H-9.

dx.doi.org/10.1021/np500617u | *J. Nat. Prod.* XXXX, XXX, XXX−XXX

moved it into the expression host, *Fusarium heterosporum*, as reported elsewhere.[4] In this context, the major compounds produced were not those found in NRRL 50135, but instead comprised approximately 800 mg of desmethyl pyrrolocin analogues B (**2**) and C (**3**), along with a smaller amount of **1**, per kilogram of fermentation.

Compound **2**, which was isolated as an off-white, amorphous solid, was the major metabolite found in the extract. High-resolution ESIMS determined its molecular formula to be $C_{26}H_{38}NO_5$. Likewise **3**, which was also isolated as an off-white, amorphous solid, has a molecular formula of $C_{26}H_{38}NO_5$, as established by high-resolution ESIMS. Comparison of the UV spectra and ^{1}H, ^{13}C, COSY, HSQC, and HMBC data of **2** and **3** revealed that the two have the same carbon structure and that they differ only in their stereochemistry. Furthermore, the HRESIMS and NMR data showed that **2** and **3** are the *N*-desmethyl analogues of **1**. However, the proton and carbon chemical shift values of **1** were more similar to **3** than to **2** (Table 1).

Relative Configuration of the Decalin Ring Moiety. NMR data showed that compound **2** has a *cis*-fused decalin structure. In the ^{1}H NMR spectrum of **2**, the decalin ring junction signals H-3 (δ 2.72 ppm) and H-8 (δ 2.09 ppm) were deshielded compared to the same protons in **1** and **3**. This is consistent with other *cis*-decalin tetramic acid analogues[5,11,15] when compared to their *trans*-fused counterparts.[3,6−8,10,14] The NMR data also indicated that the fused cyclohexyl ring assumes a chair conformation where H-3 is axial while H-8 is equatorial (Figure 1). The absence of a diaxial coupling between these two

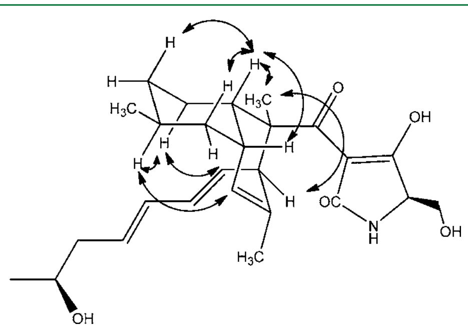


Figure 1. Relevant ROESY correlations in **2**.

protons caused H-8 to be seen as a broad singlet in the ^{1}H NMR spectrum, while H-3 was seen as a broad doublet with J = 10.2 Hz due to its coupling with H-4_{ax} (δ 1.38 ppm). The ROESY spectrum of **2** also showed a correlation between H-3 and H-8, which confirmed that the two are not in a diaxial configuration (Figure 1). H-3 was further observed to correlate with H-5_{ax} (δ 0.87 ppm) and H-7_{ax} (δ 1.11 ppm), while H-8 was also seen to correlate with both H-7_{ax} and H-7_{eq} (δ 1.56 ppm). This was also the case with the equatorial C-6 methyl (δ 0.81 ppm), which exhibited same intensity ROESY correlations with both H-7_{ax} and H-7_{eq} and with H-5_{ax} and H-5_{eq} (δ 1.67 ppm). On the other hand, H-6_{ax} was seen to correlate with H-4_{ax} (δ 1.45 ppm), H-7_{eq}, and H-5_{eq}. The fused cyclohexenyl ring also seems to assume a pseudochair configuration with H-3, C-2 methyl, and H-8 in equatorial positions, which was deduced from their ROESY correlations. Correlations were also observed between H-4_{ax} and H-12 since C-4 and C-12 are in pseudoaxial positions. Likewise, the pseudoaxial H-9 (δ 5.09 ppm) also correlated with H-6_{ax}.

The relative configuration of **3** differs from that of **2** at stereocenters C-8 and C-11. The first indication was the upfield shifts of the proton signals H-3 (δ 1.57 ppm) and H-8 (δ 1.79 ppm) as mentioned above. H-8 was also more deshielded than H-3 compared to *cis*-fused decalins, where H-3 is more deshielded.[5,11,15] The ROESY spectrum established the relative configuration of **3** (Figure 2). The chair conformation of the

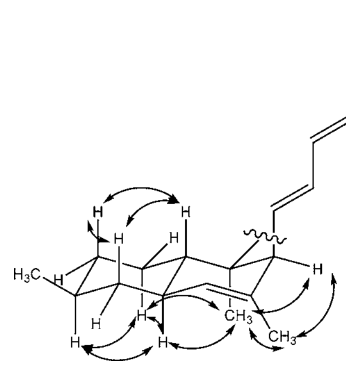


Figure 2. Relevant ROESY correlations in **3**.

cyclohexyl ring was also deduced from the following ROESY correlations: H-3 correlated with H-5_{ax} (δ 1.71 ppm) and H-7_{ax} (δ 0.82 ppm); H-8 correlated with H-4_{ax} (δ 1.00 ppm) and H-6_{ax} (δ 1.49 ppm). This therefore indicated that the C-6 methyl is equatorial. The equatorial C-6 methyl also correlated equally to H-5_{ax} (δ 1.71 ppm) and H-5_{eq} (δ 1.00 ppm) and to H-7_{eq} (δ 1.77 ppm). The fused cyclohexenyl ring in **3** also seems to adopt a pseudochair configuration where H-3 and the C-2 methyl are *anti* to each other in pseudodiaxial positions. H-11 and C-10 methyl appeared to be in pseudoequatorial and pseudoaxial positions, respectively, due to the observed correlation between the C-10 methyl and C-2 methyl.

Absolute Configuration: The Tetramic Acid Ring. The absolute configuration of C-5′ was determined by oxidative bond cleavage of the tetramic ring followed by acid hydrolysis (Figure 3).[19] The crude extract containing both compounds **2** and **3** was reacted with sodium hypochlorite and sodium hydroxide at room temperature for 8 h.[19] The oxidation products were extracted using ethyl acetate and evaporated *in vacuo*. The residue was hydrolyzed using 6 M HCl for 15 h at 110 °C to yield serine. Marfey derivatization[20] of the hydrolytic product was performed using FDLA (1-fluoro-2,4-dinitrophenyl-5-L-alanine amide). The resulting hydrolytic product serine-FDLA (HPS-FDLA) was analyzed by analytical HPLC together with reference standards L- and D-serine, which were derivatized in the same way as the hydrolytic product serine. The HPS-FDLA eluted at 25.00 min, while D-serine-FDLA injected alone and L-serine-FDLA injected alone eluted at 24.83 and 24.63 min, respectively. The HPS-FLDA that was spiked with L-serine-FDLA showed two peaks, at 24.54 and 24.99 min, while the HPS-FDLA that was spiked with D-serine-FDLA showed one peak, at 24.96 min. The HPS-FDLA was also analyzed by LC-MS and showed the expected ion peak at m/z 339 [M + H]$^+$.

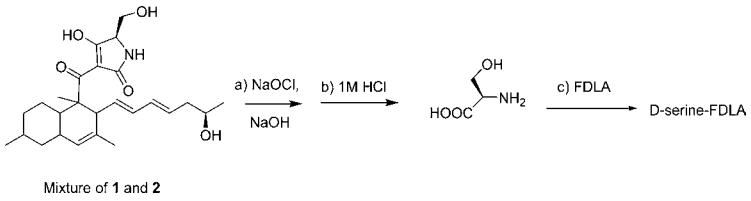
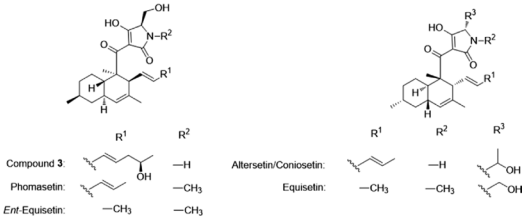


Figure 3. (a) Oxidative bond cleavage;[17] (b) acid hydrolysis; (c) derivatization with Marfey's reagent.



Absolute Configuration: The C-17 Hydroxy Group of the Olefinic Side Chain. Mosher derivatization was performed on **2** and **3** separately. Each compound was reacted with (*R*)- and (*S*)-MTPA chloride to form (*S*)-MTPA and (*R*)-MTPA esters, respectively. ^{1}H NMR analysis of the Mosher esters was performed, and the $\Delta\delta_{S-R}$ calculated (Figure 4).[21] Both **2** and **3** were found to have the 17*R* configuration.

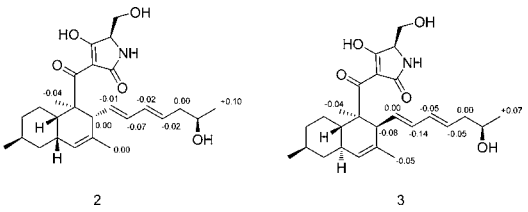


Figure 4. $\Delta\delta_{S-R}$ values obtained in pyridine-d_5 of the MTPA esters of **1** and **2**.

Absolute Configuration: The Decalin Ring of 2. Without any derivatization step, the exciton CD method[22] was used to elucidate the absolute configuration of **2**. This was possible because two existing chromophores can possibly interact and exhibit exciton split CD spectra, which can then be observed as a positive or negative Cotton effect. The positive Cotton effect that is predicted from the configuration of the two chromophores of **2** (Figure 5A) was observed in the experimental CD spectra of **2** (Figure 5B). The tetramic acid ring attached at C-2 and the conjugated diene attached at C-11 are two chromophores that coupled to cause the positive Cotton effect that was observed in the experimental CD spectrum of **2**. The absolute configuration of **2** was therefore assigned as 2*R*, 3*S*, 6*S*, 8*S*, 17*R*, and 5′*R*. This method has been applied to establish the absolute configuration of natural products with pre-existing chromophores, including abscisic acid, quassin, dendryphiellin F, and arnottin II.[22]

Absolute Configuration: The Decalin Ring of 3. The optical rotation $[\alpha]_D$ = +130 and the CD spectrum of **3** (Figure 6) were almost identical to that of phomasetin ($[\alpha]_D$ = +93.9)

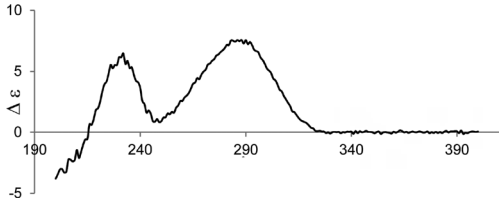


Figure 6. Experimental CD spectrum of **3** in methanol.

and *ent*-equisetin (Table 2) and opposite of that of altersetin/coniosetin (Figure 7).[6,8] The C-5′*R* configuration of the tetramic ring as determined above is also consistent with that of phomasetin. The absolute configuration of **3** was therefore assigned as 2*R*, 3*S*, 6*S*, 8*R*, 17*S*, and 5′*R*.

Absolute Configuration of 1. Compound **1**, which was isolated as an off-white oil, was a minor compound in the recombinant expression. Its HPLC retention time (Figure 8A

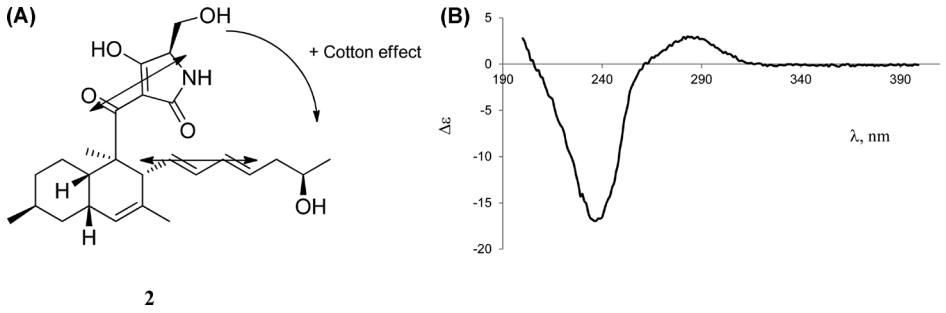


Figure 5. (A) Expected Cotton effect of **2** according to the exciton CD method of determining absolute configuration. (B) Experimental CD spectrum of **2**.

dx.doi.org/10.1021/np500617u | *J. Nat. Prod.* XXXX, XXX, XXX−XXX

Table 2. CD Data of 3 and Related Tetramic Acid Analogues

compound	CD λ_{max} nm ($\Delta\varepsilon$)				
3	225 (+4.8)	232 (+6.5)	260 (+2.6)	288 (+7.5)	330 (0)
phomasetin[10]	225 (+3.2)	232 (+4.4)	260 (+1.0)	290 (+5.2)	330 (0)
ent-equisetin[10]		235 (+6.0)	250 (+4.0)	291 (+12.5)	330 (0)
equisetin[10]	227 (−5.5)	235 (−7.5)	260 (−3.0)	290 (−8.9)	330 (0)
altersetin[6]	212 (−2.9)	232 (−18.1)	253 (−5.0)	282 (−15.3)	330 (0)
coniosetin[8]	227 (−11.4)	232 (−12.0)	251 (−1.4)	283 (−3.0)	330 (0)

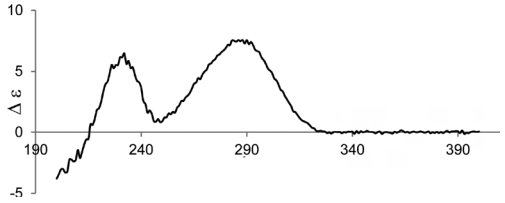


Figure 7. Experimental CD spectrum of coniosetin (reprinted with permission from ref 8).

and B), high-resolution ESIMS, UV, ^{1}H, ^{13}C, COSY, HMBC, and HSQC NMR spectra, and CD data matched those of **1** from the wild-type fungus NRRL 50135. Determination of the absolute configuration of **1** was achieved by comparison of its ROESY and CD spectra with **3**. The ROESY spectrum of **1** was very similar to that of **3**, which indicates that **1** and **3** have the same relative configuration. The similarity of the CD spectra of these two compounds further revealed that **1** and **3** have the same absolute configuration.

Biosynthesis of Pyrrolocins and Related Compounds. The co-occurrence of **1**, **2**, and **3** seems to suggest that **2** and **3** are *cis*- and *trans*-adducts originating from an intramolecular Diels−Alder reaction in the biosynthetic process.[16,17] Chemically, this reaction usually prefers to go through an *endo* transition state, which leads to the *cis*-decalin isomer, while the less favored *exo* transition state leads to the *trans*-decalin. In secondary metabolism, the favored stereochemical route is usually controlled by enzymes, rather than by solution chemistry. For example, the fungal polyketide solanapyrone is cyclized by a fairly novel enzyme after release of the intermediate from the polyketide synthase (PKS).[16,17] In a more relevant example the fungal decalin lovastatin is produced by a PKS that is similar to the pyrrolocin PKS. In lovastatin biosynthesis, it is the PKS itself that controls decalin stereochemistry, leading to strict production of the *trans*-decalin. The relevant biosynthetic intermediate is proposed to contain an ene that is activated by conjugation to an enzyme-bound thioester. By contrast, the synthetic analogue of the proposed intermediate leads to a mixture of *cis*- and *trans*-decalin compounds, reflecting a mixture of *endo*- and *exo*-Diels−Alder reactions, respectively.

Equisetin and its related compounds are thought to be synthesized by a similar mechanism, wherein the PKS itself would dictate whether the Diels−Alder reaction follows the *endo* or *exo* pathway. In the present study, **1** was found as the major product by NRRL 50135, and only trace amounts of **3** and none of **2** were detected (Figure 8A), revealing that only *trans*-decalins were produced in the wild-type fungus. On the other hand, the heterologous expression yielded **2** and **3** in a ∼2:1 ratio, which is what one would expect for a nonenzymatic reaction (Figure 8B). Although there are many reasons for this, the result implies that a separate Diels−Alderase domain might be involved in the biosynthesis of equisetin-like compounds, which in the present study might not be present in the heterologous expression. As postulated by other authors,[17] it seems that LovB (the lovastatin nonaketide synthase) and EqxS

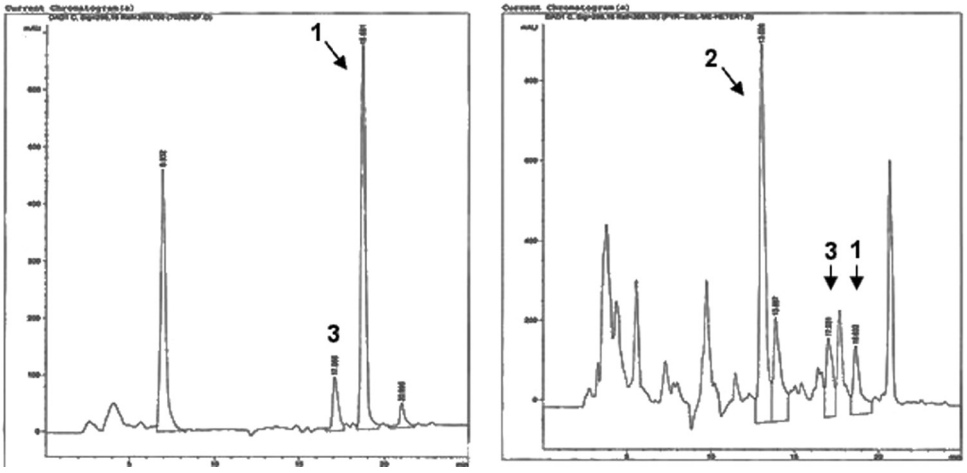


Figure 8. (A) HPLC chromatogram with PDA detection of the original extract of NRRL 50135. (B) HPLC chromatogram with PDA detector of the extract from the expression platform in mutated *F. heterosporum*.

dx.doi.org/10.1021/np500617u | *J. Nat. Prod.* XXXX, XXX, XXX−XXX

(equisetin PKS-NRPS hybrid), at the very least, contain a binding pocket to direct the stereochemical outcome of the reaction. On the other hand, the direct product of the pyrrolocin PKS still contains a carbonyl in the correct position to activate an ene for the Diels–Alder reaction. By contrast, in lovastatin biosynthesis, more advanced PKS intermediates no longer contain an activated ene. This result implies that perhaps pyrrolocins are enzymatically cyclized in the wild-type host, but not in the recombinant host.

Of additional interest, despite the structural similarities between pyrrolocins and compounds such as equisetin, these compounds are often enantiomers of each other. While this fact is very convenient in terms of assigning the absolute configuration of the compound series as we have done here, it is curious. Why is the D-amino acid always found with one decalin absolute configuration and the L-amino acid with the other? We propose two possibilities. First, the presence of the D- or L-amino acid may be controlled enzymatically; the configuration would then dictate the stereochemical course of the Diels–Alder reaction. Second, since the α position is relatively labile in the tetramic acid motif, the amino acid may begin as L-configured and equilibrate to the thermodynamically favored configuration based upon the stereochemistry of the decalin ring. These ideas provide a testable hypothesis that will enable understanding of the timing and biochemical basis of the Diels–Alder reaction in pyrrolocins and the many structural analogues.

Biological Activity. Compounds **1**–**3** were found to be active against *M. tuberculosis*, with **1** and **3** being more potent than **2** (Table 3). The data seem to indicate that a *trans* configuration and *N*-methylation enhance the potency of antimicrobial activity.

Table 3. Antimycobacterial Activities of 1–3

compound	anti-TB IC_{50} (μM)	cytotoxicity IC_{50} (μM)
1	26.3	76.6
2	112.9	167.0
3	56.4	112.9
rifampicin	0.152	not determined

EXPERIMENTAL SECTION

General Experimental Procedures. UV spectra were measured on a Hewlett-Packard 8452A diode array spectrophotometer (Agilent, Santa Clara, CA, USA). IR spectra were recorded using a JASCO (Easton, MD, USA) FT/IR-400 spectrometer. Specific rotations were recorded on a PerkinElmer (Downers Grove, IL, USA) 343 digital polarimeter in MeOH at 22 °C. CD spectra were recorded on an Aviv model 420 CD spectrometer (Aviv Biomedical, Inc., Lakewood, NJ, USA) in methanol at 25 °C. NMR spectra were recorded on a Varian (Palo Alto, CA, USA) INOVA at 500 MHz for ^{1}H and 125 MHz for ^{13}C using vendor-supplied pulse sequences.

Accurate mass measurements were performed by HRESIMS on a Micromass Q-tof Micro (Waters, Milford, MA, USA) using positive ion mode and an FTMS (LTQ-FT, ThermoFisher, Waltham, MA, USA). HPLC was performed on an Agilent 1200 series equipped with a PDA detector (Agilent Technologies). A Luna 5 μm C_{18} column, 250 × 10 mm (Phenomenex, Torrance, CA, USA), was used for the isolation of the three compounds. Supelco Diaion HP20SS was purchased from Sigma-Aldrich (St. Louis, MO, USA). TLC was conducted using Kieselgel 60 F_{254} (Merck, Whitehouse Station, NJ, USA).

Cultivation and Biological Source. The fern *Asplenium* sp. was collected in Watunou, Milne Bay Province, Papua New Guinea, by a team from Wyeth (now Pfizer). A piece of the stem was rigorously surface-sterilized, and the sterile stem was placed on malt extract agar containing antibacterials and 1 μg/mL cyclosporin until fungal hyphae were observed to protrude from the stem. The strain was initially isolated with the name ENDO-0549, and after axenic cultivation, it was deposited in the NRRL culture collection as NRRL 50135. Phylogenetic analysis using molecular markers identified NRRL 50135 as a potentially novel genus in the order Diaporthales. The sequences were deposited in GenBank (KM107910).

Preparation of Inocolum of NRRL 50135. NRRL 50135 was cultured on Difco potato dextrose agar. A 250 mL Erlenmeyer flask containing 50 mL of seed medium, Difco potato dextrose broth, was inoculated with agar-grown culture. The seed was incubated at 22 °C with shaking at 200 rpm until sufficient cell density was achieved to inoculate the production fermentation.

Fermentation of NRRL 50135. The production fermentation medium consisted of 15 g of nongauze milk-filter paper (KenAG Animal Care Group, Ashland, OH, USA) cut into 8 mm wide by 60 mm long strips and 102 g of milled long grain white rice wetted with 125 mL of 0.1% yeast extract solution per 2.8 L Fernbach flask. The fermentation was inoculated with 20 mL of the seed and incubated stationary at 22 °C for 15 days.

Isolation of 1 from NRRL 50135. The cells were extracted three times with methanol (0.6 L). The combined extract was evaporated under reduced pressure to dryness. The methanol extract was then purified by reversed-phase HPLC using a C_{18} column (Phenomenex Luna, 12 μp, 50 × 250 mm) and a gradient solvent of 85–100% acetonitrile in water (40 mL/min, 0 min run), with both solvents containing 0.01% in volume of trifluoroacetic acid (TFA). The peak at 18 min, monitored at 293 nm, was concentrated to give partially purified pyrrolocin. The compound was further chromatographed by HPLC using a different column (YMC ODS-A, 10 μp, 30 × 250 mm) and a gradient solvent of 80–100% acetonitrile in water (20 mL/min, 23 min run), with both solvents containing 0.01% in volume of TFA. The peak at 17 min was collected and evaporated under reduced pressure to afford **1** as a colorless oil (7.2 mg).

Fermentation Procedure for *F. heterosporum*. Mutant strains of *F. heterosporum* carrying pyrrolocin biosynthetic genes were maintained and grown as described.[4]

Isolation of 1–3 from *F. hereosporum*. Three-week cultures of *F. heterosporum* grown on 500 g of corn grit agar at room temperature in 2.5 L Fernbach flasks were extracted twice with acetone for 12 h. Extracts were pooled and dried *in vacuo* using a rotary evaporator (IKA Works, Inc., Wilmington, NC, USA) to afford the crude extract. A portion of the crude extract (6.7%) was dissolved in MeOH, mixed with Diaion HP20SS, and dried. The resin was loaded into a column (8.5 cm × 2.0 cm i.d.) and fractionated using 40 mL each of the following solvents: 100% water, 75% H_2O/25% 2-propanol, 50% H_2O/50% 2-propanol, 25% H_2O/75% 2-propanol, and 100% MeOH to yield five fractions, designated FW, F1, F2, F3, and F4.[23] The fractions were collected and solvents evaporated using a rotary evaporator. Bioassay-guided fractionation showed anti-TB activity for F3. F3 was purified by semipreparative HPLC using a reversed-phase C_{18} column at a 3.5 mL/min flow rate as follows: 60% ACN/20% 0.1% aqueous TFA from 0 to 5 min, linear gradient from 60% to 100% ACN from 5 to 20 min. Compound **1** (1.5 mg) eluted at 18.7 min, **2** (5.0 mg) at 16.0 min, and **3** (11.5 mg) at 13.0 min.

Preparation of (*R*)- and (*S*)-MTPA Ester Derivatives of 2 and 3.[24] **2** and **3** (0.5 mg each) were separately dissolved in pyridine-d_6 (150 μL) and transferred into clean NMR tubes. Under a N_2 gas stream, (*R*)-(+)-α-methoxy-α-(trifluoromethyl)phenylacetyl chloride (5 μL) was added into each sample, and the NMR tubes were then carefully shaken to mix the samples and the MTPA chloride evenly. The NMR tubes were allowed to stand at room temperature and monitored every 2 h by ^{1}H NMR. The reaction was found to be complete after 6 h. ^{1}H NMR data of the (*S*)-MTPA ester derivative of **2** (**2a**) (500 MHz, pyridine-d_6, data were assigned on the basis of its ^{1}H–^{1}H COSY correlations): δ 0.81 (3H, d, J = 6.3 Hz, H_3-6-Me), 1.21 (3H, s, H_3-2-Me), 1.31 (3H, d, J = 6.3 Hz, H_3-18), 2.44 (2H, m, H_2-16), 5.12 (1H, s, H-9). 1.64 (3H, s, H_3-10-Me), 5.58 (1H, p, J = 7.3

dx.doi.org/10.1021/np500617u | *J. Nat. Prod.* XXXX, XXX, XXX–XXX

Hz, H-15), 5.68 (1H, dd, J = 14.7, 10.7 Hz, H-12), 6.11 (1H, dd, J = 14.7, 10.0 Hz, H-13), 6.20 (1H, dd, J = 10.0, 14.0 Hz, H-14). ^{1}H NMR data of the (*S*)-MTPA ester derivative of **3** (**3a**) (500 MHz, pyridine-d_6, data were assigned on the basis of its ^{1}H−^{1}H COSY correlations): δ 0.81 (2H, d, J = 6.6 Hz, H_3-6-Me), 1.17 (3H, d, J = 6.3 Hz, H_3-18), 1.48 (3H, s, H_3-2-Me), 2.22 (2H, m, H_2-16), 3.31 (1H, m, H-11), 5.40 (1H, m, H-15), 5.58 (1H, m, H-12), 5.91 (1H, m, H-14), 6.13 (1H, m, H-13).

In the manner described for **2a** and **3a**, **2** and **3** (0.5 mg each) were separately dissolved in pyridine-d_6 and reacted with (*S*)-(−)-α-methoxy-α-(trifluoromethyl)phenylacetyl chloride (5 μL) in NMR tubes at room temperature for 6 h to afford the (*R*)-MTPA esters of **2** (**2b**) and **3** (**3b**). ^{1}H NMR data of the (*R*)-MTPA ester derivative of **2b** (500 MHz, pyridine-d_6): δ 0.82 (1H, d, J = 6.4 Hz, H_3-6-Me), 1.21 (3H, d, J = 6.4 Hz, H_3-18), 1.25 (3H, s, H_3-2-Me), 1.64 (3H, s, H_3-10-Me), 2.44 (2H, m, H_2-16), 5.09 (1H, s, H-9), 5.60 (1H, p, J = 7.1 Hz, H-15), 5.69 (1H, m, H-12), 6.04 (1H, dd, J = 10.7, 14.7 Hz, H-13), 6.22 (1H, dd, J = 10.6, 14.9 Hz, H-14). ^{1}H NMR data of the (*R*)-MTPA ester derivative of **3b** (500 MHz, pyridine-d_6, data were assigned on the basis of its ^{1}H−^{1}H COSY correlations): δ 0.80 (3H, d, J = 6.3 Hz, H_3-6-Me), 1.10 (3H, d, J = 6.4 Hz, H_3-18), 1.52 (3H, s, H_3-2-Me), 1.80 (3H, s, H_3-10-Me), 2.22 (2H, m, H_2-16), 3.39 (1H, m, H-11), 5.45 (1H, m, H-15), 5.58 (1H, m, H-12), 5.96 (1H, m, H-14), 5.99 (1H, m, H-13).

Oxidative Bond Cleavage of the Tetramic Acid Ring.[19] One *F. heterosporum* sample lacked a methyltransferase and made only **2** and **3**. The crude methanol extract was obtained and dried, and a portion (0.2 g) was dissolved in methanol (6 mL). Then 1 M NaOH (0.75 mL) was slowly added, followed by a NaOCl solution (available chlorine 8.5−13.5%, 3.5 mL). The mixture was stirred at rt for 8 h. Then, 1 M aqueous Na_2SO_2 (3 mL) was added, and the mixture was neutralized by addition of 1 M HCl. After removal of the solvent, the residue was diluted with H_2O, and then the resulting mixture was extracted with ethyl acetate. The organic layer was dried *in vacuo*. The residue was dissolved in 6 N HCl (0.5 mL) and heated in a sealed ampule vial at 110 °C for 15 h. The solvent was removed *in vacuo*. The absolute stereochemistry of the resulting serine was determined by the advanced Marfey method.[20]

The acid hydrolysate was dissolved in 1 mL H_2O. To a 100 μL aliquot were added 1% FDAA (1-fluoro-2,4-dinitrophenyl-5-L-alanine amide) in acetone (200 μL) and 40 μL of 1 N $NaHCO_3$.The mixture was heated at 40 °C for 1 h in a water bath. The vial was then cooled, and 2 M HCl (20 μL) was added. This solution was analyzed by analytical HPLC using the following method: 0 min 5% ACN/95% aqueous TFA (0.1%), 0−2 min 5−25% ACN, 2−3 min 25% ACN, 3−48 min 25−70% ACN, 48−50 min 70% ACN.

Antimycobacterial Assay. *M. tuberculosis* H37Ra growth inhibition was quantified using the colorimetric 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay modified from previously published methods.[25] Plant extracts and control drugs were dissolved in dimethyl sulfoxide (DMSO) to produce test stock solutions, and DMSO served as a negative control. *M. tuberculosis* cultures were dispensed in 200 μL of ADC (Remel, Lenexa, KS, USA)-enriched 7H9 medium into 96-well culture clusters at 100 000 cells per well. Then 1 μL of DMSO (control) or DMSO containing extract or drug dilution was added in quadruplicate wells. After incubation for 4 days at 37 °C, 11 μL of sterile MTT (5 mg/mL in PBS) was added and further incubated for 12 h. The insoluble formazan salt formed was solubilized by the addition of 50 μL of a solubilization solution (5% SDS w/v, 50% DMF v/v, 45% H_2O v/v). A_{570} was measured using a Multiskan FC plate reader (Fisher Scientific, Waltham, MA, USA). All data were corrected against media-only blank wells by subtracting media-only A_{570}. The percent inhibition was calculated as the fraction of the sum of the test wells over the sum of control wells subtracted from 1 and multiplied by 100. The IC_{50} was determined from the percent inhibition as the concentration of test material that inhibited at least 50% of growth of the test organism.

Cytotoxicity Assay. The cytotoxicity assay utilized CEM-TART cells obtained from the NIH AIDS Reagent Reference Program (NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH).[26] Cells were grown in complete RPMI 1640 in a jacketed incubator with 5% CO_2 at saturated humidity and at 37 °C. The assay used round-bottom 96-well culture cluster plates. Compounds to be tested were added in quadruplicate wells, while controls were as follows: solvent only (DMSO) in 16 wells, of which 8 received media only ("media only" blank) and 8 received 20 000 cells/well CEM-TART cells (positive growth control). A further 8 wells received a decreasing amount of doxorubicin in duplicate wells (4 concentrations) and CEM-TART cells serving as positive cytotoxicity control. After 4 days of incubation 11 μL/well of 5 mg/mL MTT in PBS was added, followed by 2 h of further incubation. Cells were pelleted by centrifugation, and the medium in each cell was aspirated off using a vacuum aspirator. DMSO (150 μL/well) was added to dissolve the insoluble formazan salt. To completely dissolve the formazan precipitate, plates were usually agitated for 5 min using a Fisher Scientific MS1 S7 plate shaker. A_{570} was determined using a Scientific Multiskan FC (Fisher Scientific) plate reader. The resulting data were processed with a specially made Excel spreadsheet to calculate the percent inhibition as follows: the average of the media-only blank was subtracted from all wells. The average of the test wells for each compound was subtracted from the average of the positive control wells and divided by the average of the positive growth control, resulting in a fraction of control. To determine the percent fraction of inhibition, the fraction of control was subtracted from unity and multiplied by 100. The IC_{50} was determined from the percent inhibition as the concentration of test material that inhibited at least 50% of TART cell growth.

*Pyrrolocin A (**1**):* brownish oil; $[\alpha]^{22}_D$ +56.1 (*c* 0.03, MeOH); UV (MeOH) λ_{max} nm (log ε 244 (3.48), 295 (3.05); CD λ_{max} nm (Δε) 239 (+5.82), 294 (+2.40); IR (NaCl disk) ν_{max} 3317, 2946, 2832, 2161, 2012, 1977, 1679; ^{1}H NMR (DMSO-d_6, 400 MHz) and ^{13}C NMR (DMSO-d_6, 100 MHz) see Table 1; HRESIMS *m*/*z* 458.2906 $[M + H]^+$ (calculated for $C_{27}H_{40}NO_5$ 458.2906).

*Pyrrolocin B (**2**):* off-white solid; $[\alpha]^{22}_D$ −150 (*c* 0.025, MeOH); UV (MeOH) λ_{max} nm (log ε) 238 (4.4156), 286 (4.0263); CD λ_{max} nm (Δε 236 (−16.9), 283 (3.0); IR (NaCl disk) ν_{max} 3325, 2945, 2834, 2362, 2340, 1683, 1653, 1456; ^{1}H NMR (CD_3OD, 500 MHz) and ^{13}C NMR (CD_3OD, 125 MHz) see Table 1; HRESIMS *m*/*z* 444.2751 $[M + H]^+$ (calculated for $C_{26}H_{38}NO_5$ 444.2750; Δ +0.2 ppm).

*Pyrrolocin C (**3**):* off-white solid; $[\alpha]^{22}_D$ +130 (*c* 0.025, MeOH); UV (MeOH) λ_{max} nm (log ε) 247 (4.4776), 286 (4.1049); CD λ_{max} nm (Δε) 232 (6.46), 288 (7.56); IR (NaCl disk) ν_{max} 3300, 2945, 2916, 2868, 2843, 2155, 1678, 1575, 1453; ^{1}H NMR (DMSO-d_6, 500 MHz) and ^{13}C NMR (DMSO-d_6, 125 MHz) see Table 1; HRESIMS *m*/*z* 444.2762 $[M + H]^+$ (calculated for $C_{26}H_{38}NO_5$ 444.2750; Δ +2.7 ppm).

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

1D and 2D NMR spectra for compounds **1**−**3** and HPLC chromatograms of the FDLA derivatives are available as Supporting Information. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 801 582 4547. Fax: 801 585 5111. E-mail: lbarrows@pharm.utah.edu.

Present Address

[#](R.C.J.) De La Salle Health Sciences Institute College of Pharmacy, Congressional Ave., Dasmariñas, Cavite 4114, Philippines.

Author Contributions

All authors have contributed to the manuscript and have given approval to the final version.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors wish to gratefully acknowledge NIH support through the ICBG 5UO1TW006671-10 R and NSF 0957791. R.C.J. wishes to acknowledge the following persons for their technical assistance: M. Jacobsen for the measurement of the CD spectra, C. Serrano for the IR and OR, T. E. Smith for the LC-MS, and A. Fleming for the ECD calculations.

■ REFERENCES

(1) Beehler, B. M., Swartzendruber, J. F., Eds. In *Papua New Guinea Conservation Needs Assessment*, Vol. *2*; Biodiversity Support Program: Boroko, Papua New Guinea, 1993.
(2) Arnold, A. E.; Lutzoni, F. *Ecology* **2007**, *88*, 541−549.
(3) Vesonder, R. F.; Tjarks, L. W.; Rohwedder, W. K.; Burmeister, H. R.; Laugal, J. A. *J. Antibiot.* **1979**, *32*, 759−761.
(4) Kakule, T. B.; Jadulco, R. C.; Koch, M.; Janso, J. E.; Barrows, L. R.; Schmidt, E. W. Native promoter strategy for high-yielding synthesis and engineering of fungal secondary metabolites. *ACS Synth. Biol.* DOI: 10.1021/sb500296p. [Epub ahead of print].
(5) Boros, C.; Dix, A.; Katz, B.; Vasina, Y.; Pearce, C. *J. Antibiot.* **2003**, *56*, 862−865.
(6) Hellwig, V.; Grothe, T.; Mayer-Bartschmid, A.; Endermann, R.; Geshke, F.; Henkel, T.; Stadler, M. *J. Antibiot.* **2002**, *55*, 881−891.
(7) Ondeyka, J. G.; Smith, S. K.; Zink, D. L.; Vicente, F.; Basilio, A.; Bills, G. F.; Polishook, J. D.; Garlisi, C.; Mcguinness, D.; Smith, D.; Qiu, H.; Gill, C. J.; Donald, R. G. K.; Phillips, J. W.; Goetz, M. A.; Singh, S. B. *J. Antibiot.* **2014**, 1−5.
(8) Segeth, M. P.; Bonnefoy, A.; Brönstrup, M.; Knauf, M.; Schummer, D.; Toti, L.; Vertesy, L.; Wetzel-Raynal, M. C.; Wink, J.; Seibert, G. *J. Antibiot.* **2003**, *56*, 114−122.
(9) Sugie, Y.; Dekker, K. A.; Inagaki, T.; Kim, Y.-J.; Salakibara, T.; Sakemi, S.; Sugiura, A.; Brennan, L.; Duignan, J.; Sutcliffe, J. A.; Kojima, Y. *J. Antibiot.* **2002**, *55*, 19−24.
(10) Signh, S. B.; Zink, D. L.; Goetz, M. A.; Dombrowski, A. W.; Polishook, J. D.; Hazuda, D. J. *Tetrahedron Lett.* **1998**, *39*, 2243−2246.
(11) Singh, S. B.; Zink, D. L.; Heimbach, B.; Genilloud, O.; Teran, A.; Silverman, K. C.; Lingham, R. B.; Felock, P.; Hazuda, D. J. *Org. Lett.* **2002**, *4*, 1123−1126.
(12) Namikoshi, M.; Kobayashi, H.; Yosimoto, T.; Hosoya, T. *J. Antibiot.* **1997**, *50*, 890−892.
(13) Sugie, Y.; Inagaki, S.; Kato, Y.; Nishida, H.; Pang, C.-H.; Saito, T.; Sakemi, S.; Dib-Hajj, F.; Mueller, J. P.; Sutcliffe, J.; Kojima, Y. *J. Antibiot.* **2002**, *55*, 25−29.
(14) Inoshi, J.; Shigeta, N.; Fukuda, T.; Uchida, R.; Nonaka, K.; Masuma, R.; Tomoda, H. *J. Antibiot.* **2013**, *66*, 549−554.
(15) Watanabe, T.; Igarashi, M.; Okajima, T.; Ishii, E.; Kino, H.; Hatano, M.; Sawa, R.; Umekita, M.; Kimura, T.; Okamoto, S.; Eguchi, Y.; Akamatsu, Y.; Utsumi, R. *Antimicrob. Agents Chemother.* **2012**, *56*, 3657−3663.
(16) Campbell, C. D.; Vederas, J. C. *Biopolymers* **2010**, *93*, 755−763.
(17) Kelly, W. L. *Org. Biol. Chem.* **2008**, *6*, 4483−4493.
(18) Stocking, E. M.; Williams, R. M. *Angew. Chem., Int. Ed.* **2003**, *42*, 3078−3115.
(19) Minowa, N.; Kodam, Y.; Harimaya, K.; Mikawa, T. *Heterocycles* **1998**, *48*, 1639−1642.
(20) Marfey, P. *Calsberg Res. Commun.* **1984**, 591−596.
(21) Ohtani, I.; Kusumi, T.; Kashman, Y.; Kakisawa, H. *J. Am. Chem. Soc.* **1991**, *113*, 4092−4096.
(22) Nobuyuki, H.; Nakanishi, K.; and Berova, N. In *Comprehensive Chiroptical Spectroscopy: Applications in Stereochemical Analysis of Synthetic Compounds, Natural Products, And Biomolecules*; Berova, N.; Polavarapu, P. L.; Nakanishi, K.; Woody, R. W., Eds.; John Wiley and Sons, Inc.: NJ, 2012; Vol. *2*, Chapter 4, pp 154−155.
(23) Bugni, T. S.; Harper, M. K.; McCulloch, M. W. B.; Reppart, J.; Ireland, C. M. *Molecules* **2008**, *13*, 1372−1383.
(24) Su, B.-N.; Park, E. J.; Mbwambo, Z. H.; Santarsiero, B. D.; Mesecar, A. D.; Fong, H. H. S.; Pezzuto, J. M.; Kinghorn, A. D. *J. Nat. Prod.* **2002**, *65*, 1278−1282.
(25) Koch, M.; Bugni, T. S.; Sondossi, M.; Ireland, C. M.; Barrows, L. R. *Planta Med.* **2010**, *76*, 1678−1682.
(26) Chen, H.; Boyle, J. T.; Malim, M. H.; Cullen, B. R.; Lyerly, H. K. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7678−7682.

Isolation of Pyrrolocins A-C: *cis*- and *trans*-Decalin Tetramic Acid Antibiotics from an Endophytic Fungal-Derived Pathway

Raquel C. Jadulco,[†] Michael Koch,[†] Thomas B. Kakule,[‡] Eric W. Schmidt,[‡] Anita Orendt,[Δ] Haiyin He,[§] Jeffrey E. Janso,[§] Guy T. Carter,[I] Erica C. Larson,[†] Christopher Pond,[†] Teatulohi K. Matainaho,[⊥] and Louis R. Barrows[†⊥*]

[†]Department of Pharmacology and Toxicology, University of Utah, Salt Lake City, Utah, USA

[‡]Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah, USA

[Δ] Center for High Performance Computing, University of Utah, Salt Lake City, Utah, USA

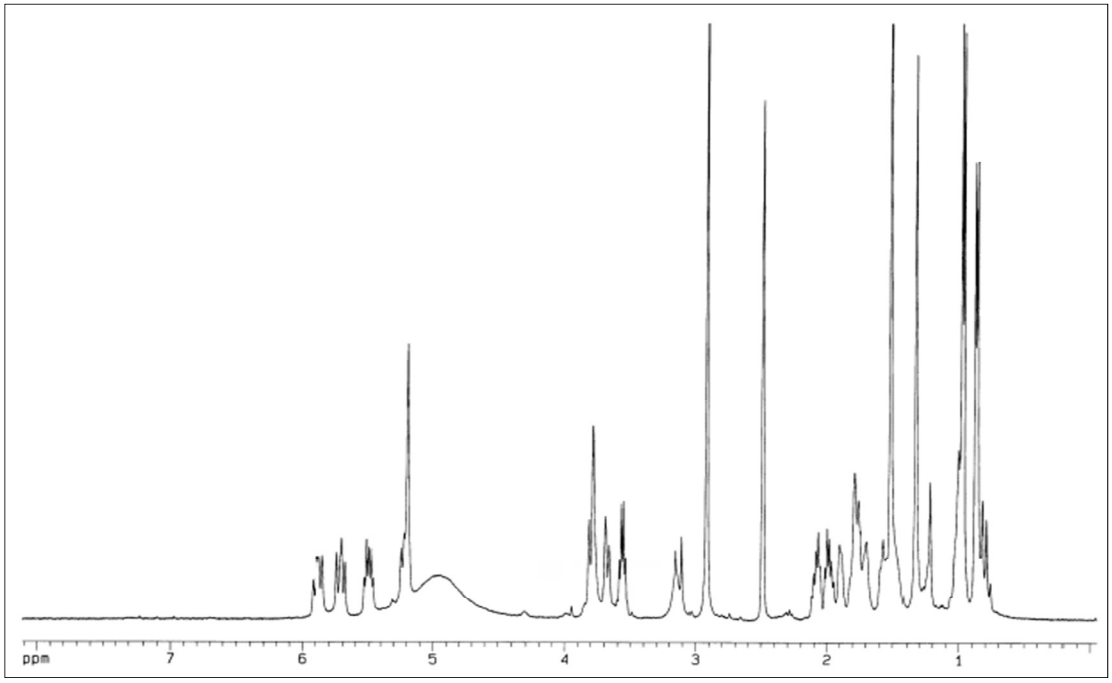
[§]Natural Products – Worldwide Medicinal Chemistry, Pfizer Worldwide Research and Development, 445 Eastern Point Road, Groton, Connecticut 06340, USA

[I]Carter-Bernan Consulting, 350 Phillips Hill Road, New City, NY 10956 , USA

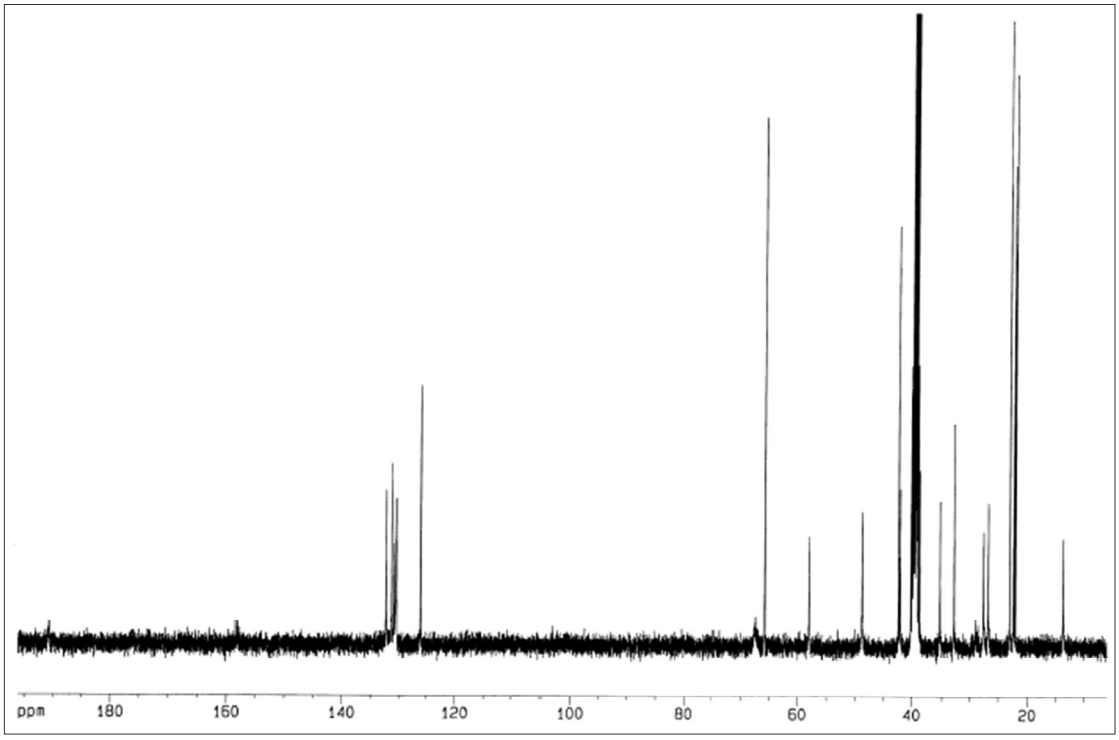
[⊥]School of Medicine and Health Sciences, University of Papua New Guinea, Boroko, NCD, Papua New Guinea

[*]Email: lbarrows@pharm.utah.edu. Phone: (801) 5824547. Fax (801) 5855111

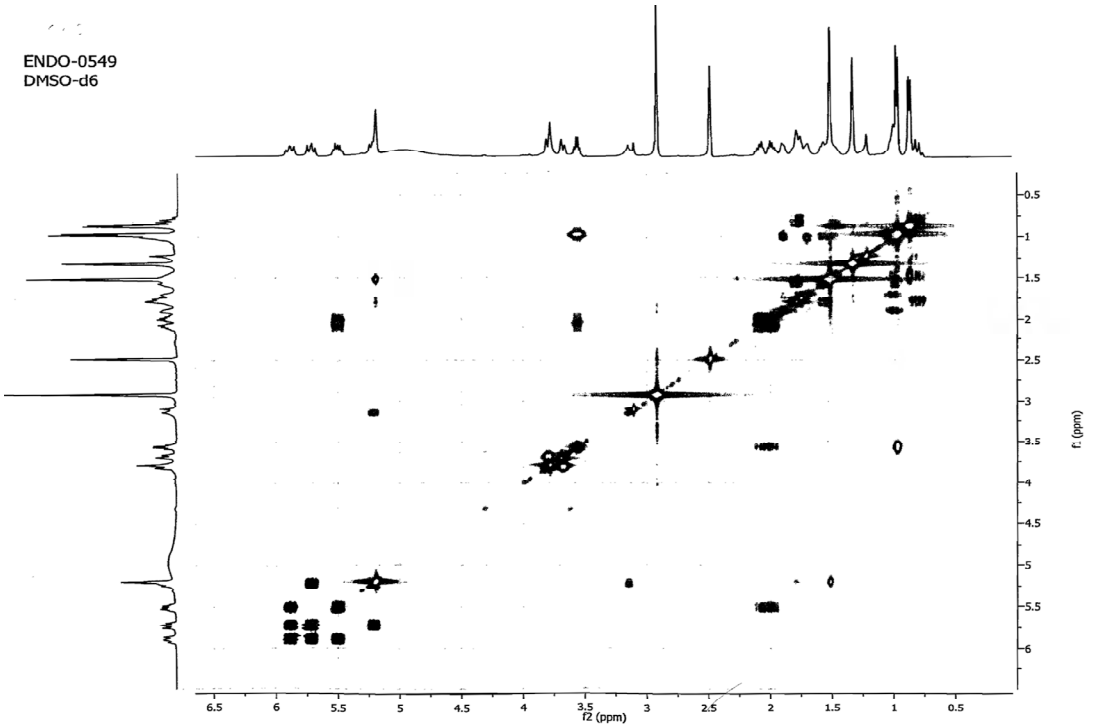
List of supporting information



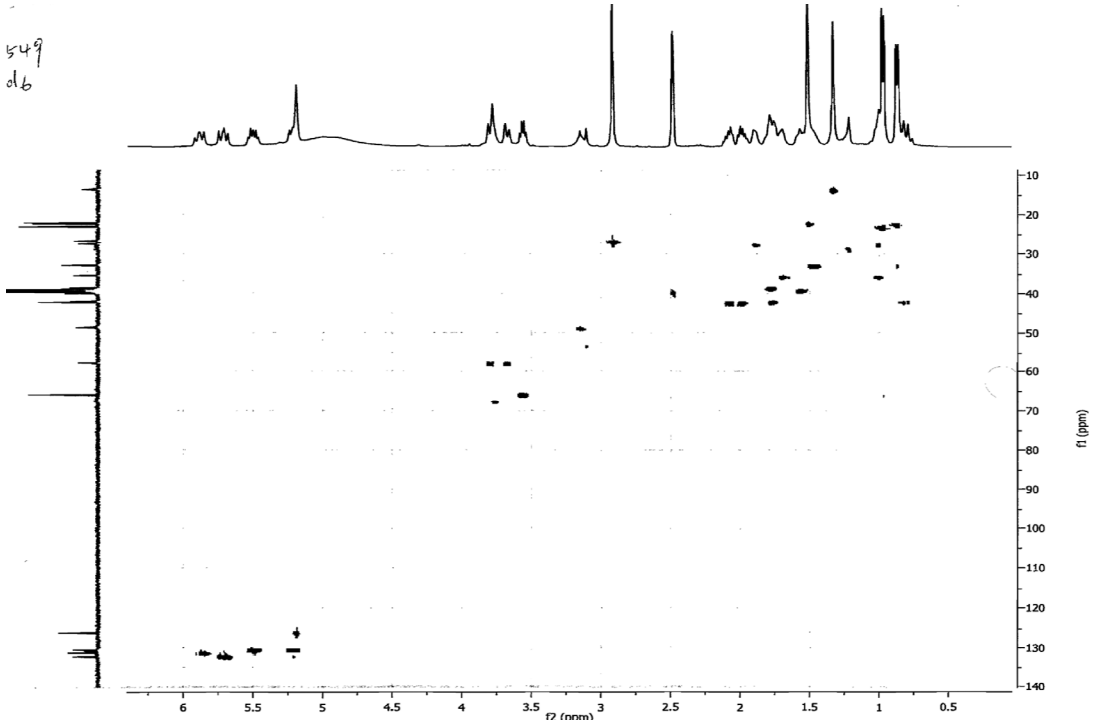
S1. ^{1}H NMR of Compound **1** in DMSO-d_6 at 400 MHz.



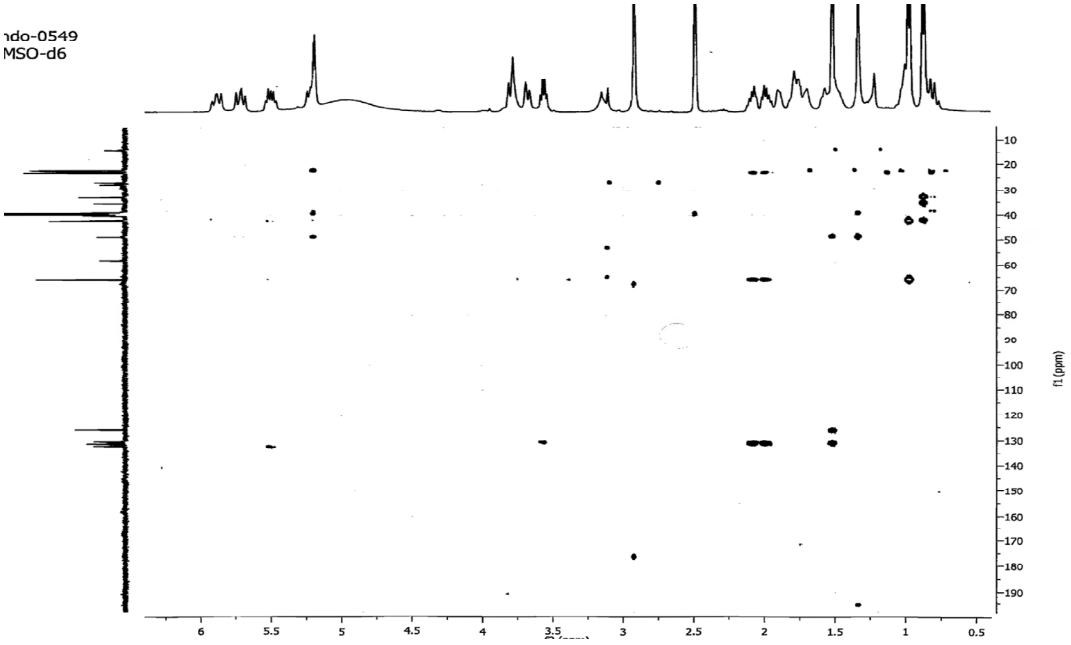
S2. ^{13}C NMR of Compound **1** in DMSO-d_6 at 100 MHz.



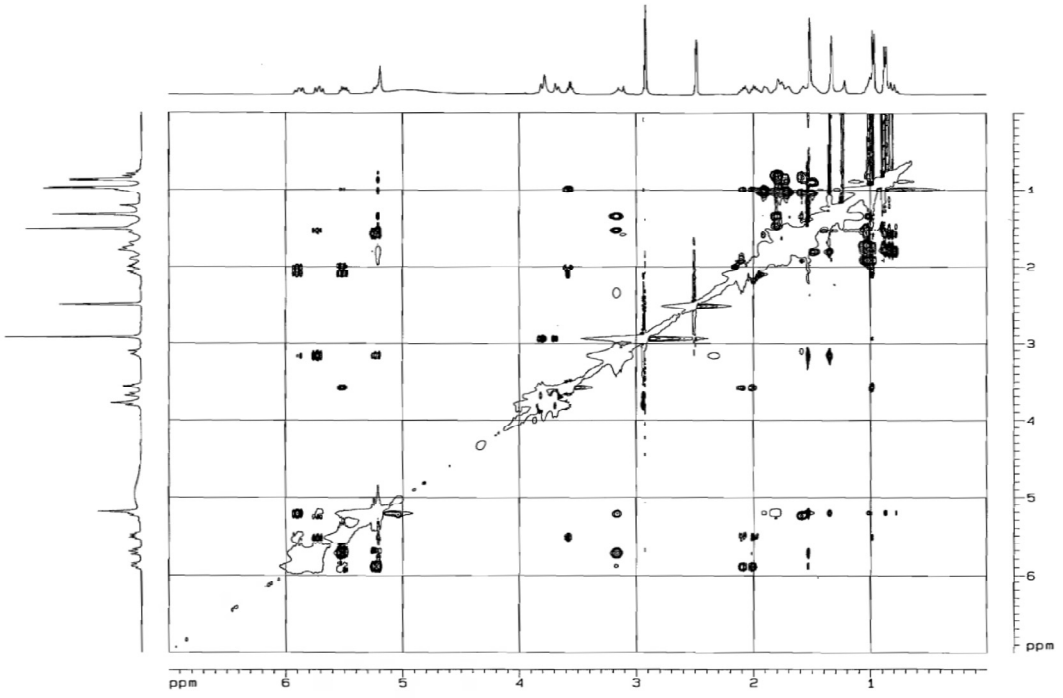
S3. COSY of Compound **1** in DMSO-d_6 at 400 MHz.



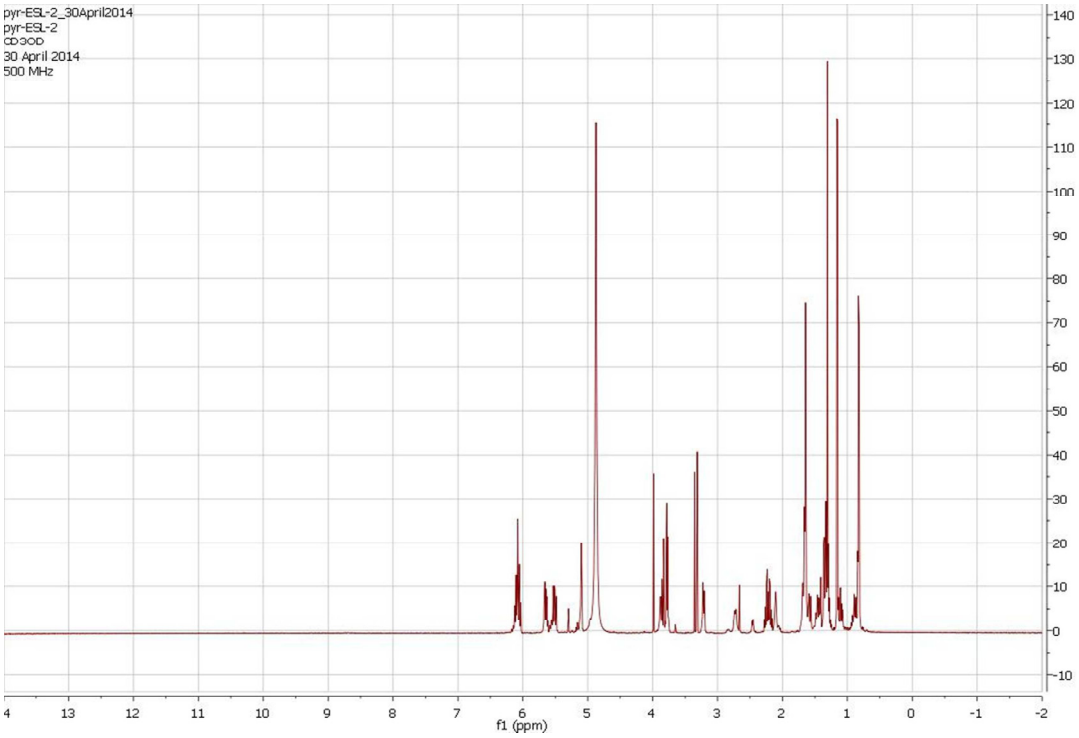
S4. HSQC of Compound **1** in DMSO-d_6 at 400 MHz.



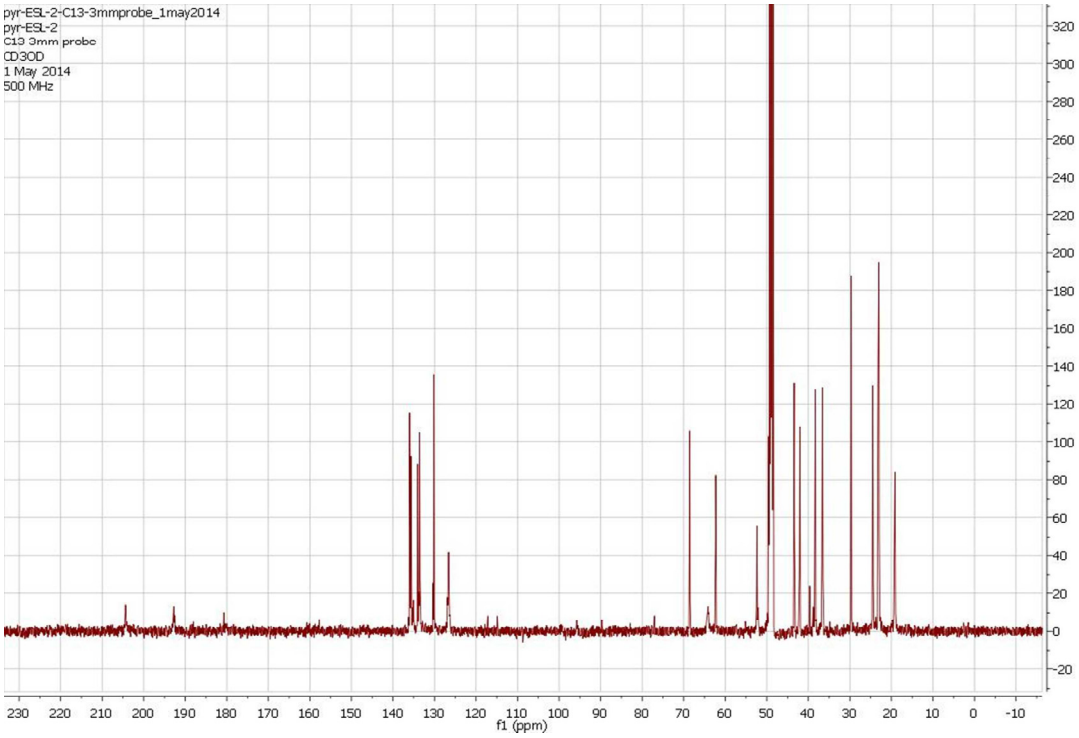
S5. HMBC of Compound **1** in DMSO-d_6 at 400 MHz.



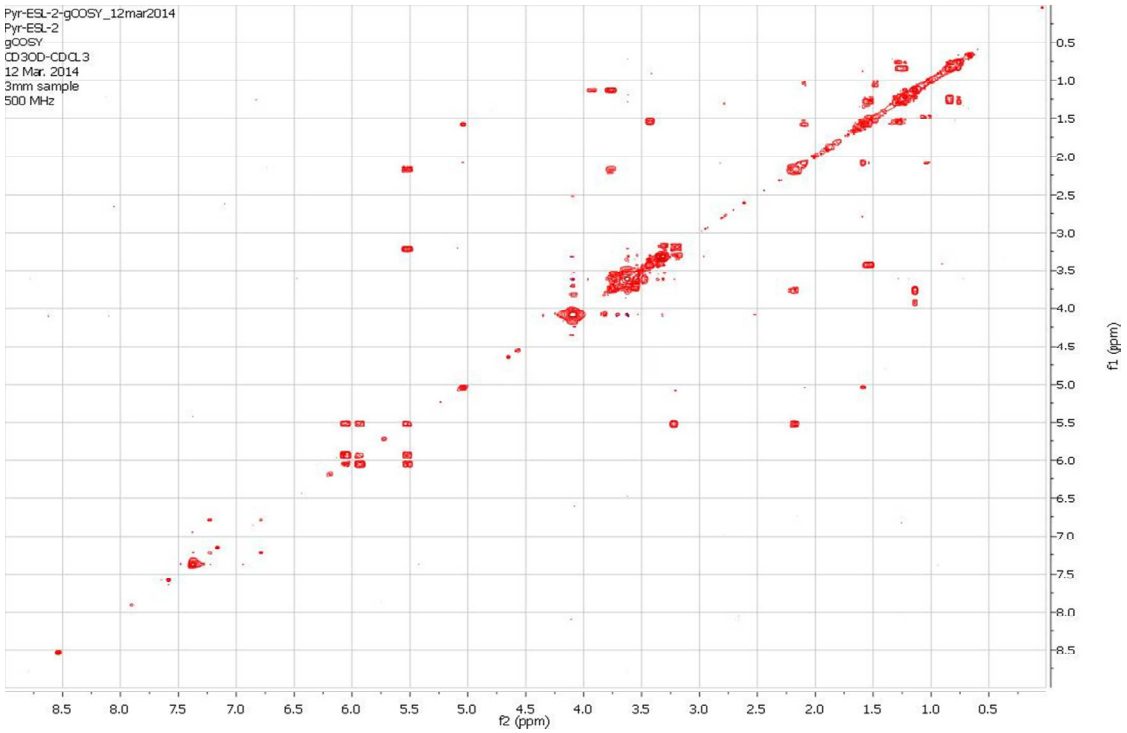
S6. ROESY of Compound **1** in DMSO-d_6 at 400 MHz.



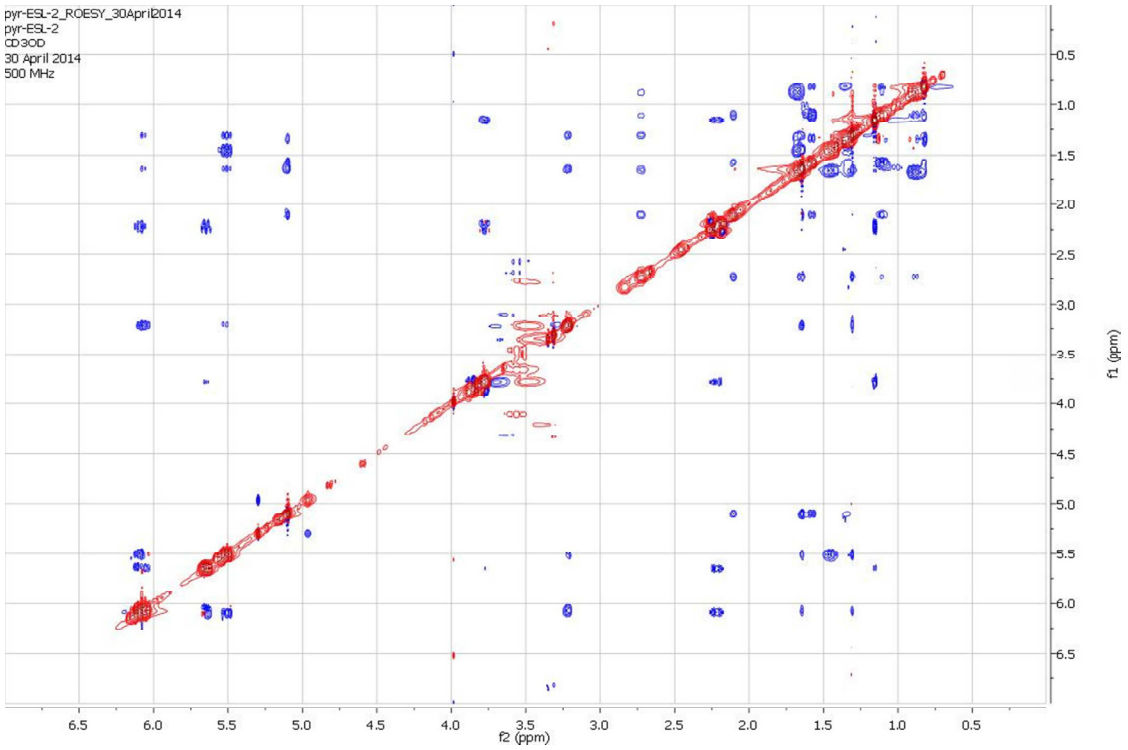
S7. ^{1}H NMR of Compound **2** in CD_3OD at 500 MHz.



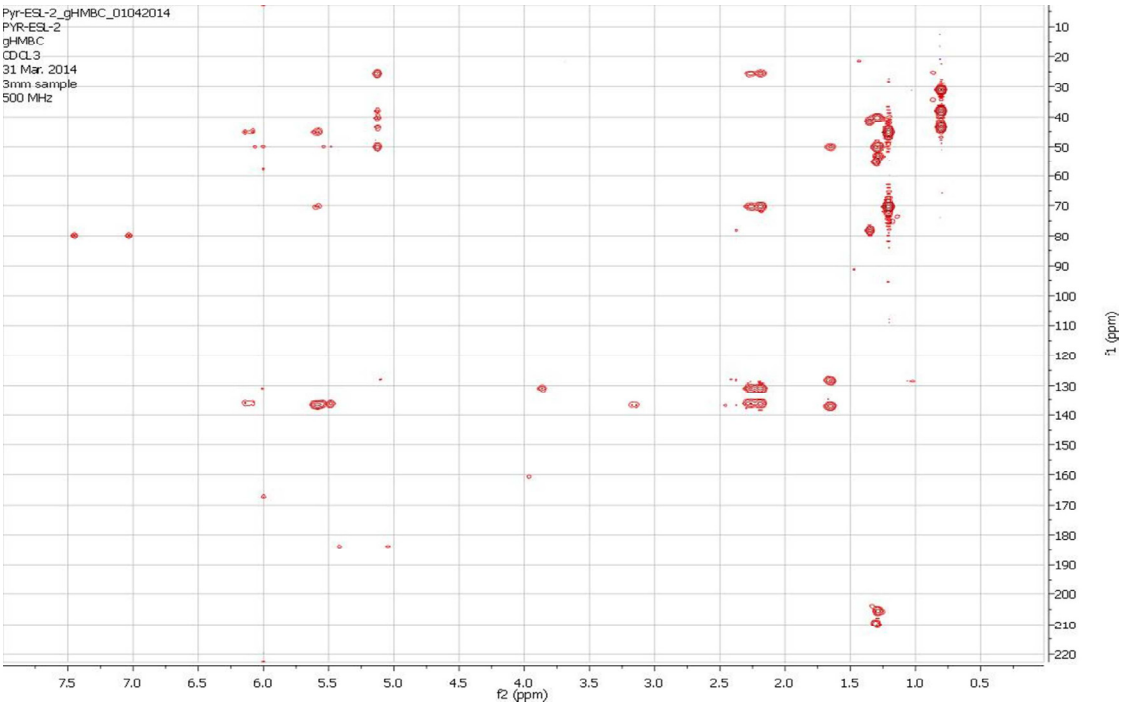
S8. ^{13}C NMR of Compound **2** in CD_3OD at 125 MHz.



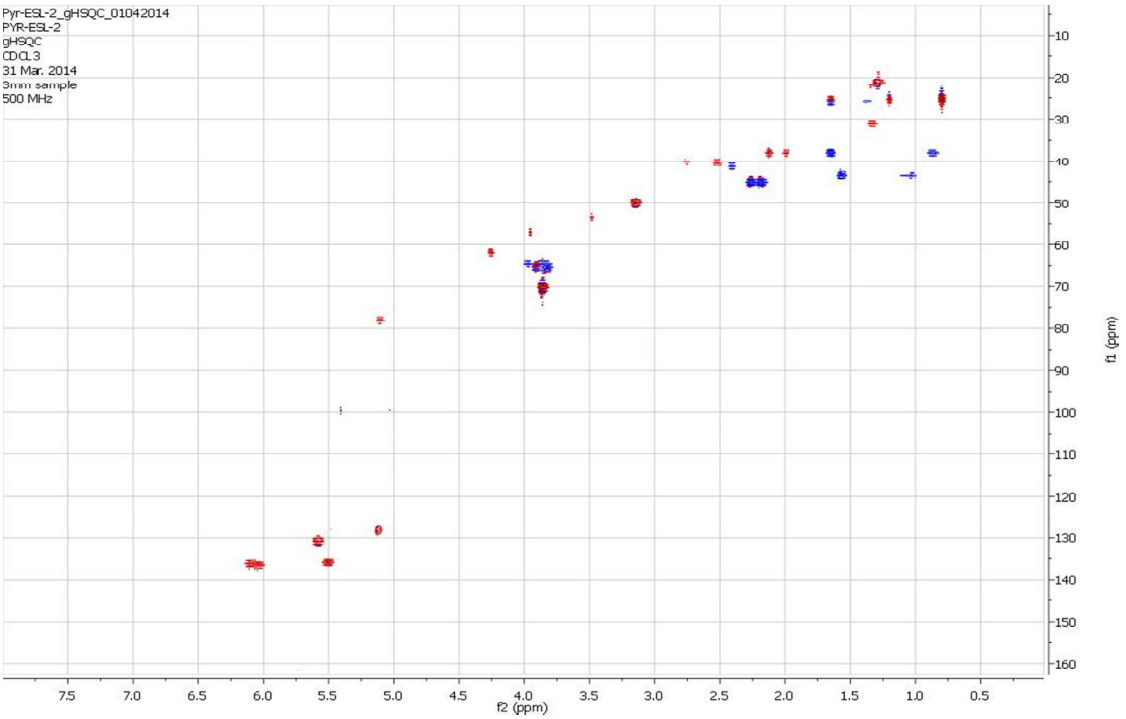
S9. COSY of Compound **2** in DMSO-d_6 at 500 MHz.



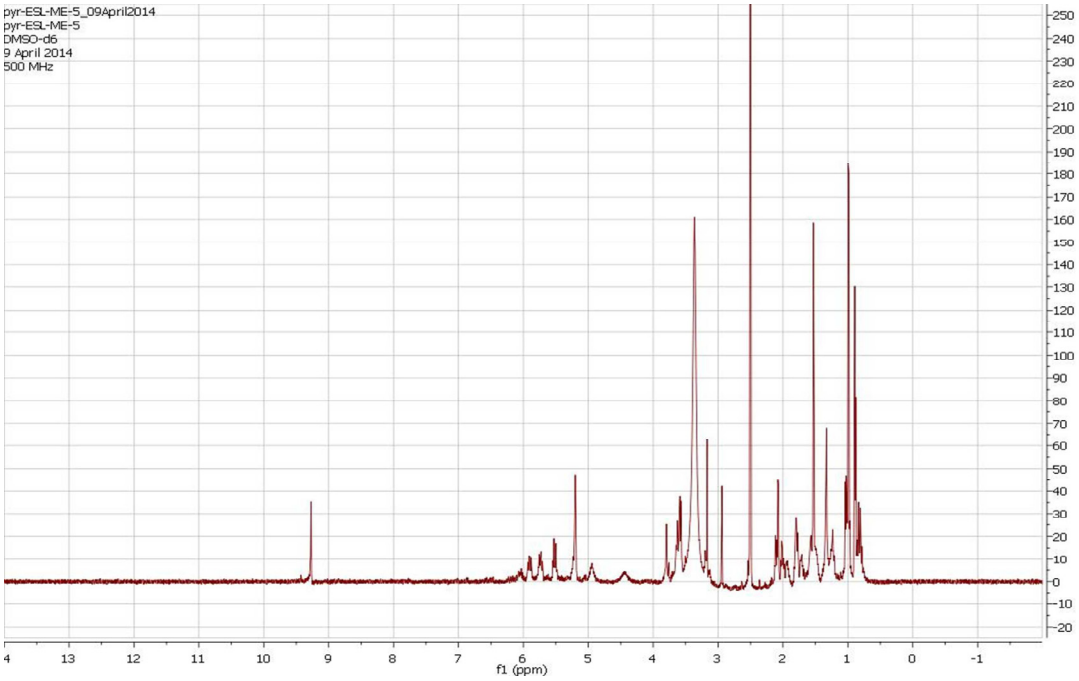
S10. ROESY of Compound **2** in DMSO-d_6 at 500 MHz.



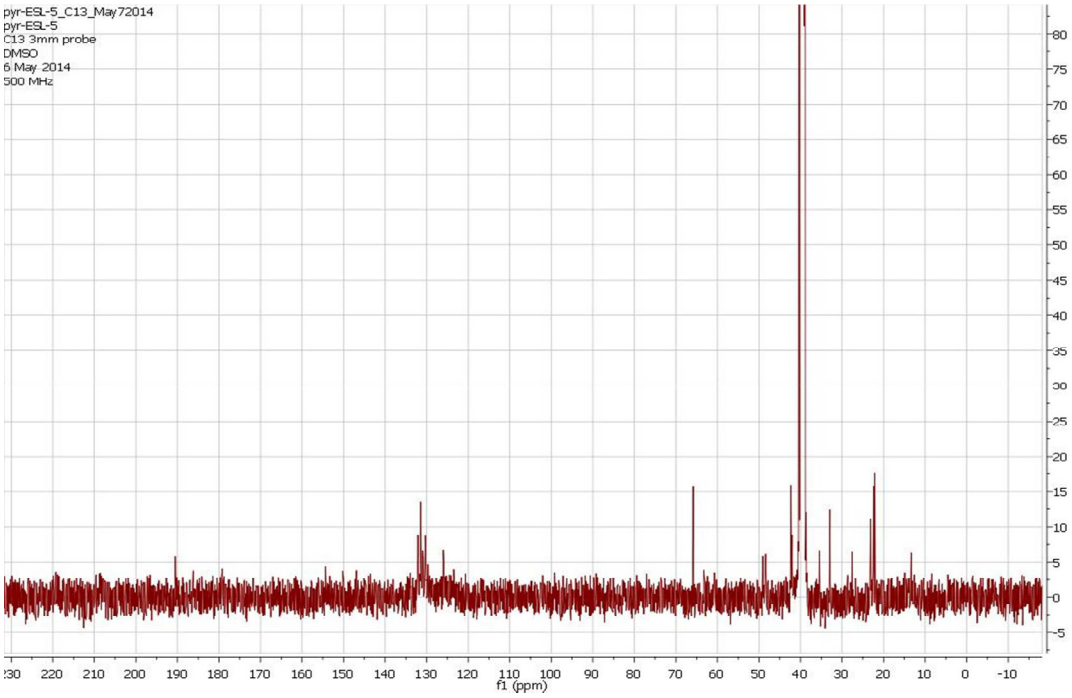
S11. HMBC of Compound **2** in DMSO-d_6 at 500 MHz.



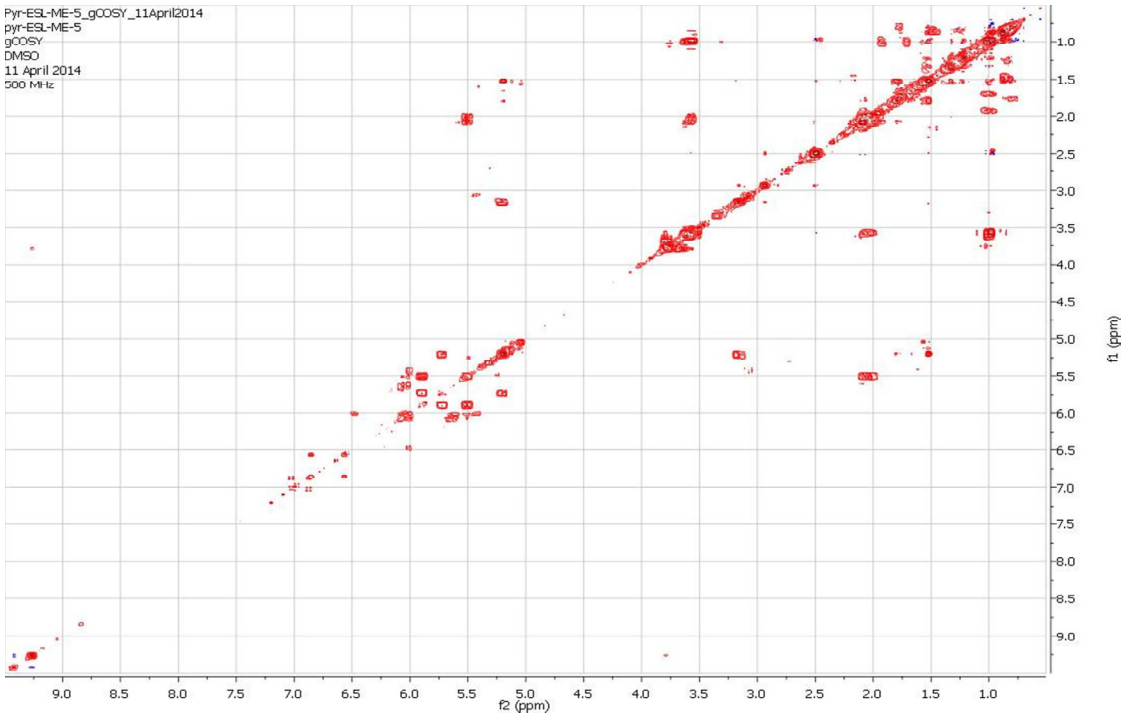
S12. HSQC of Compound **3** in DMSO-d_6 at 500 MHz.



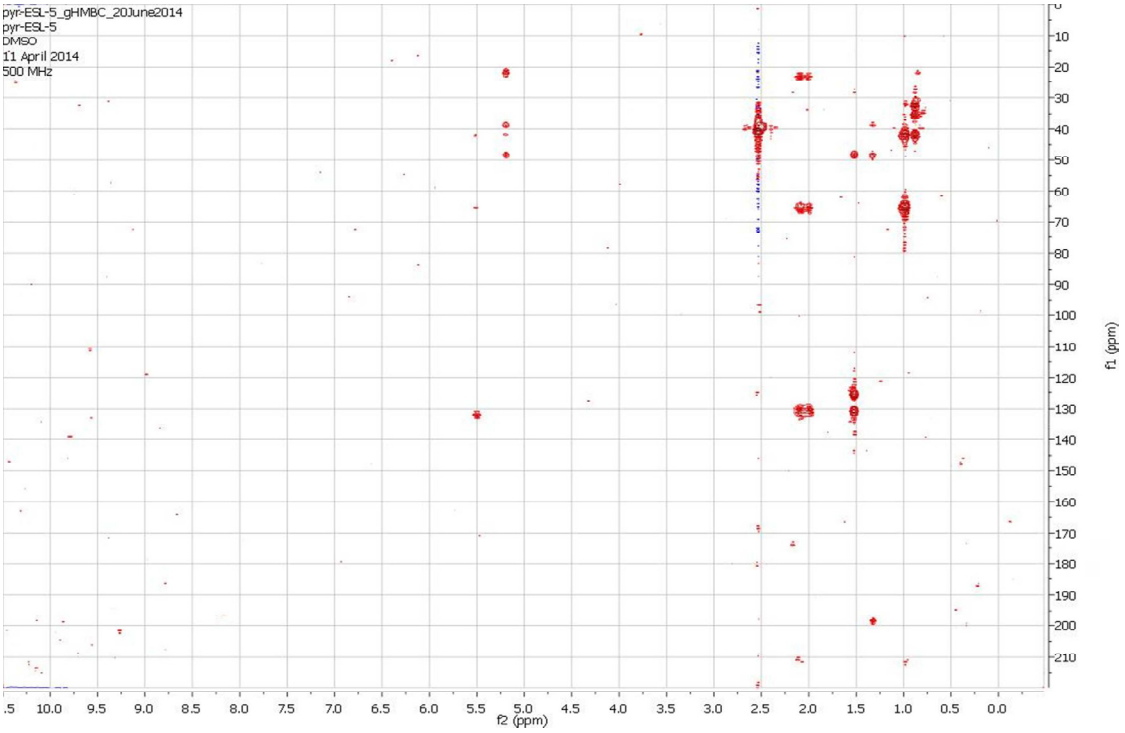
S13. [1]H NMR of Compound **3** in DMSO-d_6 at 500 MHz.



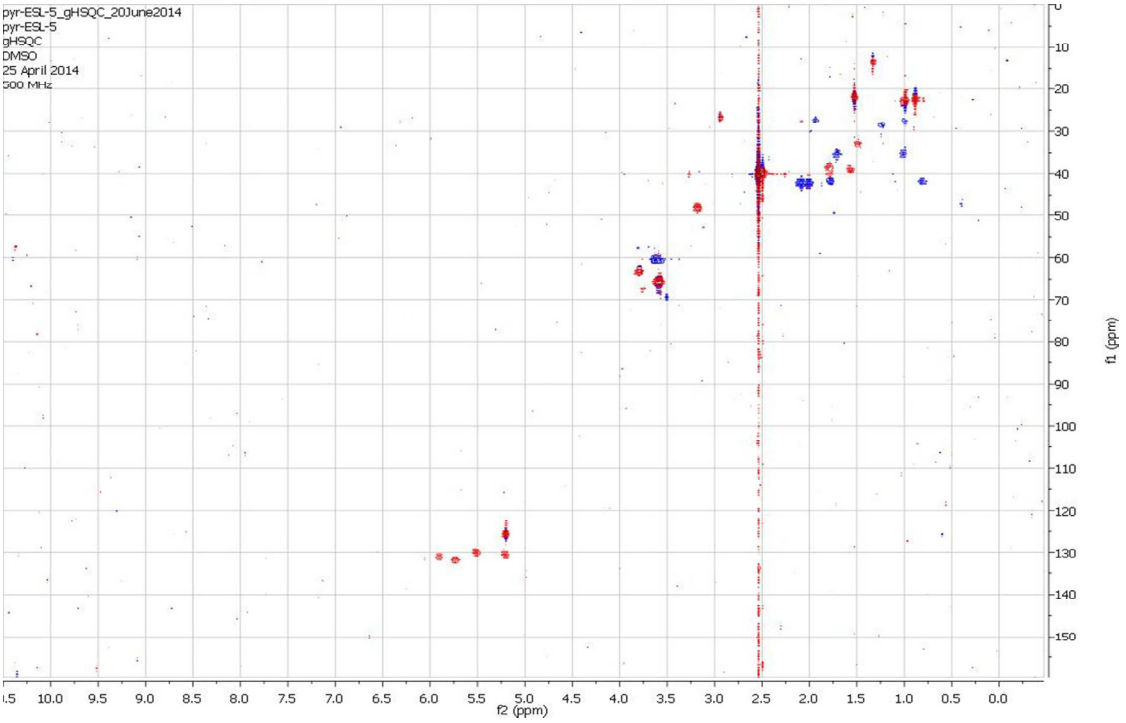
S14. ^{13}C NMR of Compound **3** in DMSO-d_6 at 125 MHz.



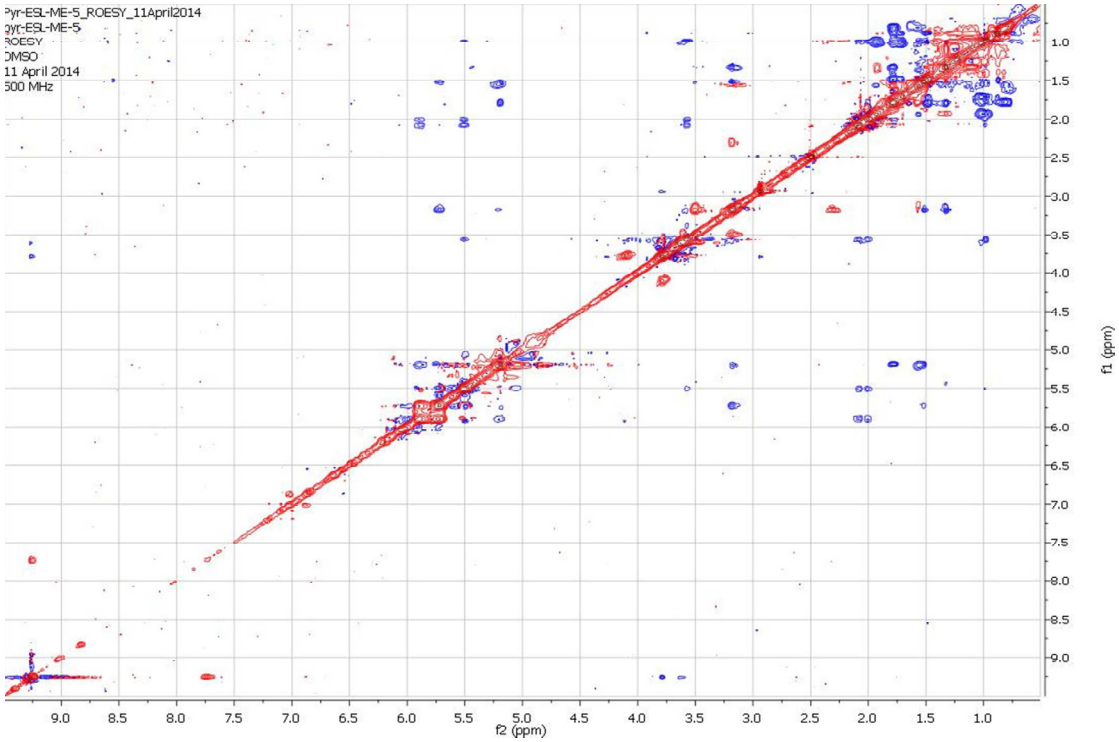
S15. COSY of Compound **3** in DMSO-d_6 at 500 MHz.



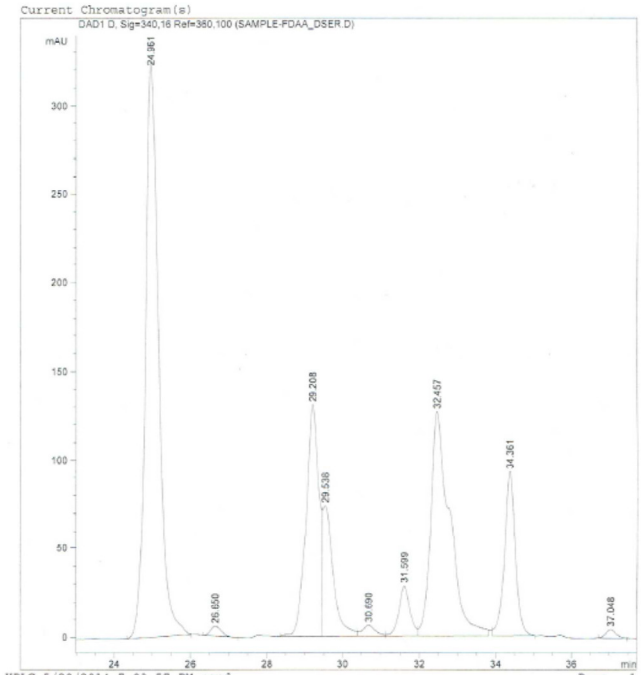
S16. HMBC of Compound **3** in DMSO-d_6 at 500 MHz.



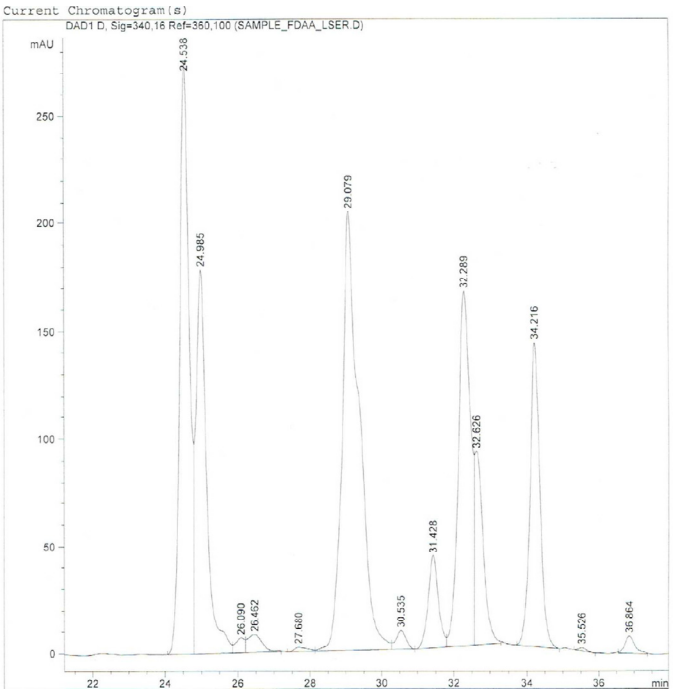
S17. HSQC of Compound **3** in DMSO-d_6 at 500 MHz.



S18. ROESY of Compound **3** in DMSO-d_6 at 500 MHz.



S19. Relevant part of the HPLC chromatogram of hydrolytic product serine-FDLA (HPS-FDLA) co-injected with D-serine-FDLA.



S20. Relevant part of the HPLC chromatogram of hydrolytic product serine-FDLA (HPS-FDLA) co-injected with L-serine-FDLA.

CHAPTER 5

COMBINATORIALIZATION OF FUNGAL POLYKETIDE SYNTHASE–PEPTIDE SYNTHETASE HYBRID PROTEINS

Reprinted with permission from
Kakule, T. B.; Lin, Z.; Schmidt, E. W. Combinatorialization of Fungal Polyketide Synthase–Peptide Synthetase Hybrid Proteins. *J. Am. Chem. Soc.* **2014**, *136*, 17882–17890.

Note: my contribution to this paper was in planning, performing, and analyzing all the experiments.

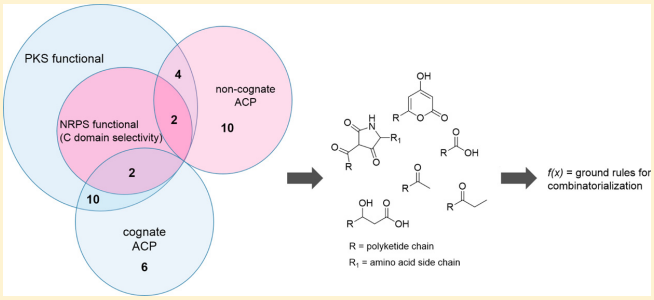
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Combinatorialization of Fungal Polyketide Synthase–Peptide Synthetase Hybrid Proteins

Thomas B. Kakule, Zhenjian Lin, and Eric W. Schmidt*

Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112, United States

Supporting Information



ABSTRACT: The programming of the fungal polyketide synthase (PKS) is quite complex, with a simple domain architecture leading to elaborate products. An additional level of complexity has been found within PKS-based pathways where the PKS is fused to a single module nonribosomal peptide synthetase (NRPS) to synthesize polyketides conjugated to amino acids. Here, we sought to understand the communication between these modules that enable correct formation of polyketide-peptide hybrid products. To do so, we fused together the genes that are responsible for forming five highly chemically diverse fungal natural products in a total of 57 different combinations, comprising 34 distinct module swaps. Gene fusions were formed with the idea of testing the connection and compatibility of the PKS and NRPS modules mediated by the acyl carrier protein (ACP), condensation (C) and ketoreductase (KR) domains. The resulting recombinant gene fusions were analyzed in a high-yielding expression platform to avail six new compounds, including the first successful fusion between a PKS and NRPS that make highly divergent products, and four previously reported molecules. Our results show that C domains are highly selective for a subset of substrates. We discovered that within the highly reducing (hr) PKS class, noncognate ACPs of closely related members complement PKS function. We intercepted a pre-Diels–Alder intermediate in lovastatin synthesis for the first time, shedding light on this canonical fungal biochemical reaction. The results of these experiments provide a set of ground rules for the successful engineering of hr-PKS and PKS-NRPS products in fungi.

INTRODUCTION

Fungal polyketide synthase (PKS)–nonribosomal peptide synthetase (NRPS) hybrid proteins synthesize a diverse array of biomedically and agriculturally important natural products.[1,2] These include the prescribed anticholesterol drug lovastatin, the biologically useful cytochalasins,[3] compounds such as macrocidins[4] that are critical in fungal pathogenesis of plants, and many other agents. The PKS-NRPS hybrid products are among the most common bioactive compounds isolated from filamentous fungi. The first fungal PKS-NRPS to be characterized was the fusarin C synthetase (FUSS),[5] and since that time many related genes have been identified. The many known fungal PKS-NRPS genes are very similar to each other, but their chemical products are not. It has thus been of great interest to learn the biochemical rules governing product formation and to exploit these rules in synthesizing new derivatives via genetic engineering.[6−8] However, success in these endeavors has so far been quite limited.

Fungal PKS-NRPS products are produced via the action of two modules. A single type I, highly reducing (hr) PKS module acts iteratively to synthesize a complex polyketide core.[1] In fungi, iterative hr-PKS proteins exert exquisite control over regiochemistry, synthesizing polyketides with variable reduction at each acetate extension step despite using the same set of reductase domains repeatedly.[9] How each domain can differentially recognize products of different chain elongation steps has remained a mystery. Some fungal iterative PKSs also require auxiliary proteins. For example, several known decalin-containing polyketides require the cooperation of the PKS and an auxiliary enoyl reductase (ER).[10−13] Upon completion of synthesis of the acetate-derived chain on the PKS, the enzyme bound intermediate is transferred to the NRPS module, at which point an amino acid is appended to the compound.[1,2]

Received: October 29, 2014
Published: December 1, 2014

dx.doi.org/10.1021/ja511087p | J. Am. Chem. Soc. 2014, 136, 17882−17890

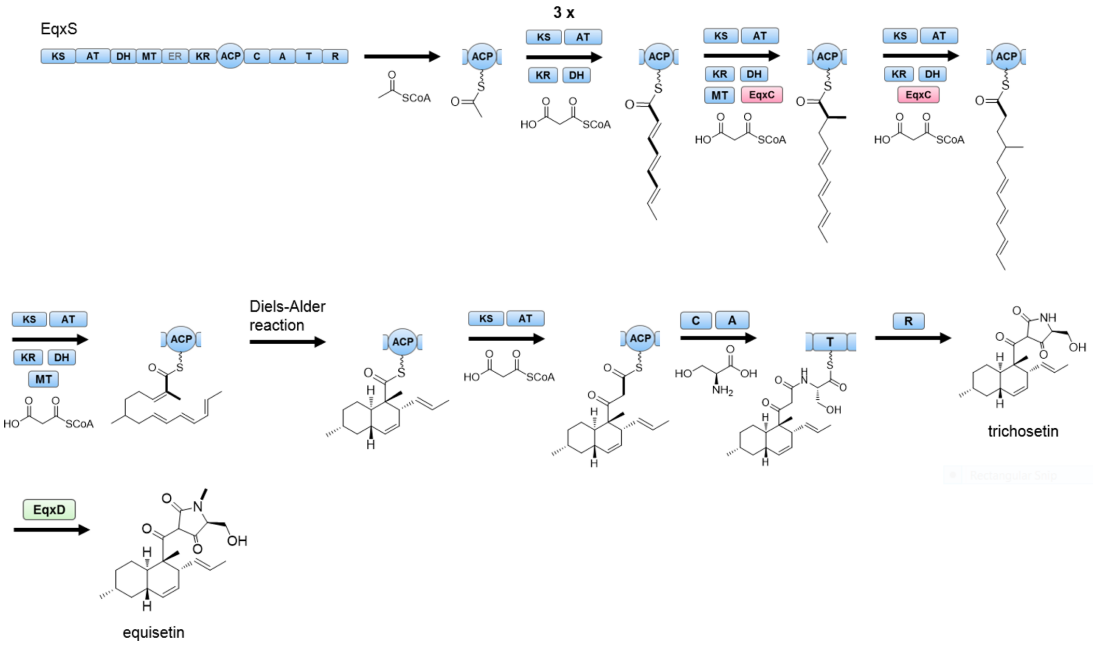


Figure 1. Biogenesis of equisetin showing the complex programming of the PKS where domains on a single polypeptide are utilized iteratively to synthesize the polyketide core. An acyl transferase (AT) domain selects the malonyl units, the ketosynthase (KS) catalyzes the decarboxylative condensation, the C-methyltransferase (MT) performs α-C-methylations, the ketoreductase (KR) reduces keto groups to hydroxyls and the dehydratase domain (DH) eliminates the hydroxyl groups to form olefins, which are then reduced to single bonds by the trans ER EqxC. The synthesized polyketide core is then transferred to the NRPS module, composed of the adenylation domain, which selects the amino acid that gets loaded onto the thiolation domain (T). Conjugation of the polyketide with amino acid is catalyzed by the condensation (C) domain. The R domain catalyzes a Dieckmann reaction to release the product as a tetramic acid, trichosetin.[25,29,30] Further *N*-methylation by EqxD forms equisetin.[11] Whether the Diels–Alder reaction occurs on-enzyme or after product release is still unknown (it has been drawn on-enzyme for clarity).

The NRPS is responsible for selecting the amino acid, synthesizing the peptide bond, and sometimes performing other steps, such as Dieckmann cyclization to produce tetramic acids. The complexity of the PKS-NRPS biosynthetic process is exemplified in Figure 1, showing the biogenesis of equisetin.

Many different types of fungal polyketide-peptide hybrids have been isolated. In the case of lovastatin and relatives, only a fragment of the NRPS is present, and amino acids are not loaded.[10] The remaining NRPS domains are thought to be responsible for potentially catalyzing a Diels–Alder reaction that cyclizes the linear polyketide chain to produce decalins.[6,10] Other products include the cyclopiazonic acid precursor, cAATrp, which contains the shortest possible polyketide chain fused as a tetramic acid to tryptophan (Figure 2).[14] Prepseurotin is derived from a longer, partially reduced polyketide fused to phenylalanine.[15,16] Equisetin has a decalin polyketide structure, like lovastatin, but is fused to serine as a tetramic acid.[11,17] Fusaridione A is a long, linear polyene fused to tyrosine.[11] Tenellin and desmethylbassianin are ring-expanded tetramic acid derivatives; their biosynthetic genes are among the best characterized fungal PKS-NRPSs.[7,18−20] Many other classes of PKS-NRPS products are known.[21]

A key to combinatorializing fungal PKS-NRPS proteins would thus be to hybridize PKS modules, producing diverse acetate-derived scaffolds, with NRPS modules, activating diverse amino acids. Indeed, this has been tried on three occasions with mixed success.[6−8] Two major domains are of

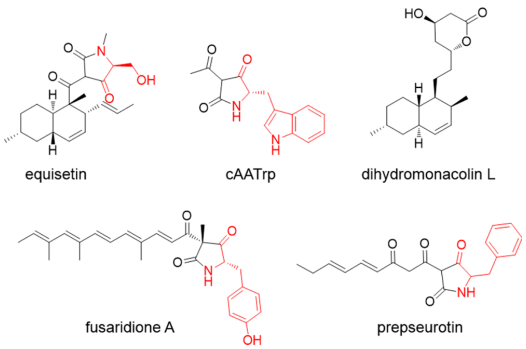


Figure 2. Structures of PKS-NRPS products showing the polyketide chain synthesized by the PKS module (black), and the amino acid appended by the NRPS (red). Dihydromonacolin L is not an amidated product because its biosynthetic enzyme, LovB, possesses a truncated NRPS module.

primary importance in successful fusions: the acyl carrier protein (ACP) from the PKS module, and the peptide bond-forming condensation (C) domain from the NRPS module. It was anticipated that the C domain might exhibit some substrate selectivity, where only certain polyketide products might be acceptable in amide bond formation. The lynchpin was

dx.doi.org/10.1021/ja511087p | *J. Am. Chem. Soc.* 2014, 136, 17882−17890

considered to be the ACP. This small (~70−100 amino acids) but critical domain is covalently tethered to all PKS intermediates and serves to ferry them between different catalytic sites in the PKS.[9] The ACP protein itself must form productive protein−protein interactions with the five or more protein domains in the PKS. The ACP also must form a protein−protein interaction with the C domain, which accepts the incoming polyketide substrate for elongation. Any protein fusions between NRPS and PKS modules therefore pose both a substrate selectivity problem (C domain) and a protein interaction problem (ACP domain).

The major goal of this project was to determine basic programming rules that would enable successful fusion of fungal PKS-NRPS proteins. The study was designed to disentangle the protein−protein interaction and substrate selectivity questions in this highly complex system. To do so, we proposed the hypothesis that fungal ACP domains contain specific sequence elements enabling interaction with PKS and NRPS modules. A series of gene fusions were made using different PKS and NRPS proteins, exploring several types of ACP interactions, with the goal of maintaining normal PKS function and forging productive ACP/C domain interactions. The resulting genes were expressed, and the chemical products were analyzed. In the event, we delineated protein−protein interaction rules and showed that fungal C domains are highly substrate-selective. The secondary goal was to understand key reactions catalyzed by PKS-NRPS proteins, such as the Diels−Alder reaction that leads to decalin products. With the exception of the Diels−Alder reaction, which in the case of lovastatin is clearly catalyzed by the native C domain, we found that fungal PKSs operate independently of NRPS modules and make the products that are expected from the natural PKS-NRPS proteins. Taken together, these experiments provide an integrated view of the function and engineering potential of fungal PKS-NRPS hybrids.

RESULTS

Experimental Design and Mutant Construction. To differentiate the effects of protein−protein interactions and C-domain selectivity, we fused PKS and NRPS proteins at different points, using four different types of ACP connections (Figure 3). First, we left the PKS ACP intact, fusing it directly with the second C domain. Second, we replaced the ACP with the ACP natively fused to the acceptor NRPS. Third, we fused the N-terminal region of the ACP from the PKS with the C-terminal region of the ACP natively fused to the acceptor NRPS. Previous studies with PKS and NRPS proteins have shown that different regions of the ACP are responsible for interacting with donor and acceptor modules.[22,23] While this has not been studied in fungi, we reasoned that a hybrid ACP might overcome potential problems with donor or acceptor recognition. At the least, it would provide crucial information about fungal ACP function. Finally, since multiple ACPs are functional in some bacterial pathways,[24] we placed the ACPs from the PKS and NRPS in tandem.

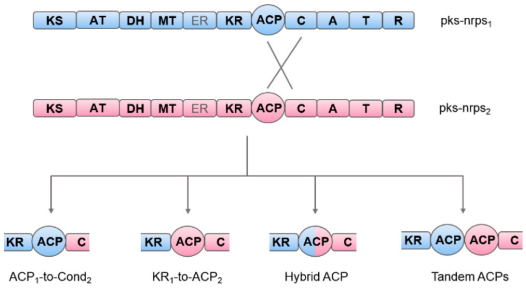


Figure 3. Design of gene fusions representing the different arrangements of the ACPs with the C domain. pksnrps1 (blue) represents *psoA*, *cpaS*, *lovB*, *eqxS* or *fsdS*; and pksnrps2 (red) represents *eqxS* or *fsdS*.

Fusions were performed by yeast recombination to generate expression vectors. The fused genes were then expressed in a model platform recently developed in *Fusarium heterosporum*.[25] The strength of this platform is that it synthesizes heterologous polyketides and polyketide-peptide hybrids in yields of ~100−1000 mg kg^{-1}. We expected that some fusions might lead to decreased yields of products, and therefore, this high yielding starting point would enable even relatively poorly functioning hybrids to be accurately analyzed. This platform uses the native promoters and regulatory elements that normally produce equisetin, but instead redirects them to the production of heterologous products. We used the recently characterized equisetin synthetase EqxS and fusaridione synthetase FsdS proteins as the NRPS modules in all experiments.[11,25] These were then fused to PKS modules from *eqxS* and *fsdS*, as well as from cyclopiazonic acid synthesis (*cpaS*),[26] pseurotin synthesis (*psoA*),[15] and lovastatin synthesis (*lovB*).[12] Where auxiliary ER proteins were required by the PKS, these were coexpressed in *F. heterosporum*. A total of eight PKS-NRPS fusions were created, each with four different ACP connections, for a total of 32 recombinant clones expressed in *F. heterosporum*. Of particular importance, the function of the nonhybridized (wild-type) PKS-NRPS proteins was first assessed in *F. heterosporum*, showing that all of the proteins were functional prior to hybridization.[11,25]

The equisetin and fusaridione synthetases *eqxS/eqxC* and *fsdS* were cloned from *F. heterosporum*, the same strain used as the heterologous production line in this study. *cpaS*, *psoA*, and *lovB* and *lovC* (a trans-ER) were cloned from *Aspergillus flavus*, *A. fumigatus*, and *A. terreus* respectively. Three of the fusion types (comprising 12 total recombinants) failed to function, including *cpaS-eqxS*, *cpaS-fsdS*, and *lovB-fsdS*. An absence of products in these experiments could occur for several different reasons. For example, the polyketide product of *cpaS* is simply acetoacetate, which may not be detectable in the experimental conditions, or protein folding problems may occur. Because of these and other reasons, it is not possible to definitively conclude that absence of product is meaningful in terms of protein interactions or substrate acceptability. Therefore, these failed fusions will not be further analyzed. For the remaining five fusion types (20 total recombinants), including *eqxS-fsdS*, *fsdS-eqxS*, *psoA-fsdS*, *psoA-eqxS*, and *lovB-eqxS*, products were detected (Figure 4).

In addition to these 32 mutants, we also constructed other hybrids to answer specific questions resulting from these initial studies, as described below.

ACP-PKS Interactions. All PKSs were active in making polyketides when fused to their own ACP domains, but only a subset of fusions between a PKS and a noncognate ACP led to polyketide products (Figure 4). All four ACP combinations effectively led to the native PKS products in the reciprocal crosses, *eqxS-fsdS* and *fsdS-eqxS*. This indicated that these ACPs recognized all domains of each other's PKS module. Modest

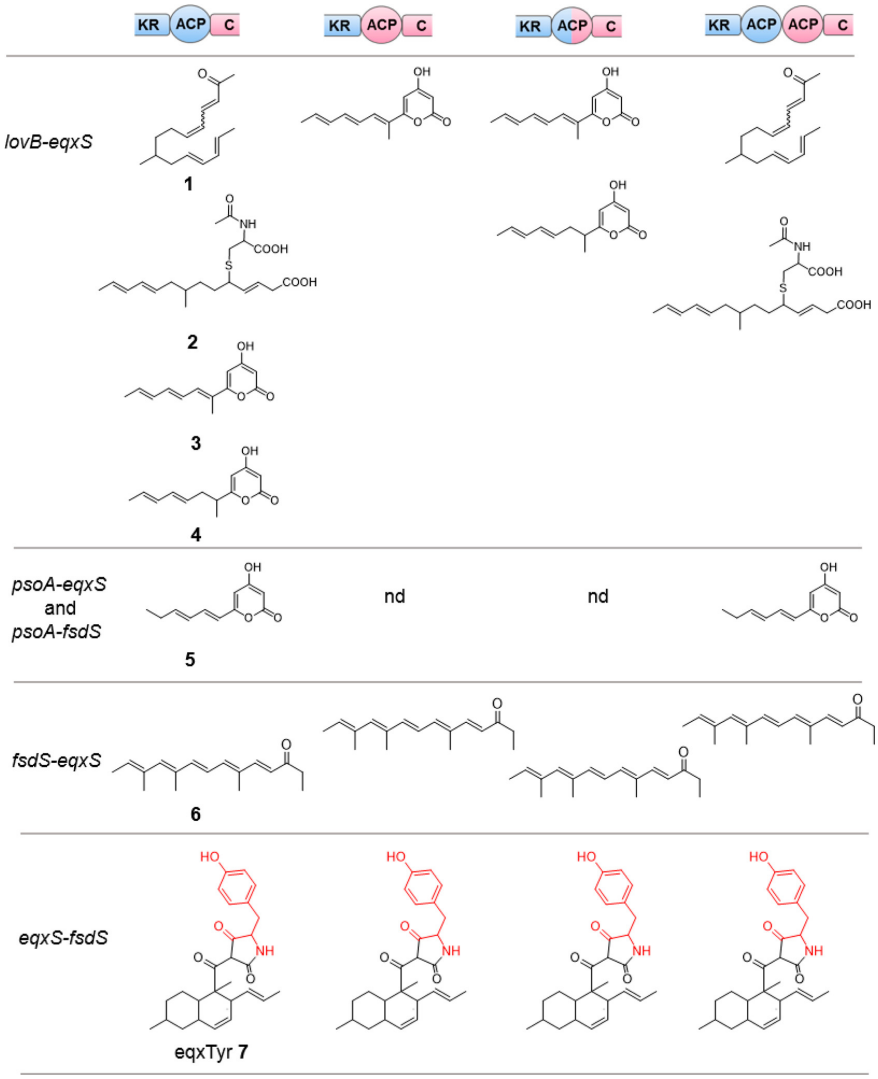


Figure 4. Summary of compounds identified from the different PKS/NRPS fusions. **1**, **2**, **5**, **6** and **7** are new compounds (nd = no new metabolites detected in comparison to *gfp*-expressing control). Note: Compound **1**, all double bonds are trans.

and potentially insignificant yield differences were the only observed effects of swapping ACPs between these systems.

By contrast, in the case of pseurotin fusions *psoA-eqxS* and *psoA-fsdS*, only the cognate ACPs were functional, and neither the fusaridione ACP nor the equisetin ACP could substitute for the pseurotin ACP (Supporting Information Figure S1). Fusions containing either the pseurotin ACP alone, or the tandem ACP system containing one ACP from pseurotin and one from fusaridione or equisetin, both led to PKS product **5** (Figure 4). By contrast, no products were detected in the hybrid ACP or noncognate ACP fusions. Compound **5** contains the complete PKS chain in the correct reduction state as found in prepseurotin. It was likely spontaneously cleaved from the enzyme as the pyrone, as has been proposed for similar products from fungal hr-PKS.[12]

lovB-eqxS fusions also led to PKS products, which differed depending upon which ACP combination was used. Previously, when *lovB* was expressed intact in *F. heterosporum*,[25] in addition to a large amount of the expected product, dihydromonacolin L (Figure 2), we obtained a small amount of pyrone **4** (Figure 4). The cognate *lovB* ACP led to a PKS product **1** that was most similar to the native monacolin product, but lacking Diels–Alder cyclization (Figure 4; Supporting Information Figure S2A). In addition, two pyrone shunt products **3** and **4** were identified, as found with the wild-type enzyme when expressed in *F. heterosporum*. Only these shunt products were observed when noncognate ACPs were used, including from the hybrid

ACP and equisetin-derived ACP fusions. Most interestingly, when *eqxS* provided the ACP, the product was not reduced by the trans-ER *lovC*. The best explanation for this observation is that the equisetin ACP cannot interact with LovC. An exhaustive search for new nitrogen-containing metabolites led to the isolation of a minor product **2**, a linear lovastatin-like polyketide precursor modified by *N*-acetyl cysteine (biochemically analogous to the very similar leukotriene modification[27]). This same product was produced when the ACPs were placed in tandem.

Transfer of PKS Products to the NRPS. Although 16 out of the 32 gene fusions led to functional PKS proteins with detectable products, only 4 of these led to products that were clearly passed along to the NRPS module. These were the *eqxS-fsdS* fusions, in which the expected product was clearly obtained (Figure 4; Supporting Information Figure S3). The sole recombinant natural product, eqxTyr **7**, isolated from this system included the decalin made by the equisetin PKS and the tyrosine tetramic acid made by the fusaridione NRPS. All four *eqxS-fsdS* fusions led to the same product. This indicated that the *fsdS* C domain could accept the equisetin polyketide product and could make functional contacts with the equisetin ACP.

In the reciprocal fusion, *fsdS-eqxS*, the native fusaridione polyketide was isolated, but it was not fused to serine from *eqxS* and was instead decarboxylated. Interestingly, all four ACP fusions led to the same product **6** (Figure 4; Supporting Information Figure S4). This indicated that the *eqxS* C domain could not accept polyketide products synthesized by *fsdS*, since the native *eqxS* ACP was present in some fusions, precluding a role for protein−protein interactions in governing C-domain selectivity. Both the *fsdS* and *eqxS* ACPs could act in concert with the PKS domains, leading to synthesis of the normal *fsdS* PKS product. Decarboxylation of *β*-ketoacids is a spontaneous reaction, so that **6** represents exactly what one would expect from a functional PKS.

Similarly, pseurotin and lovastatin fusions led only to PKS products, with no evident transfer to the *fsdS* or *eqxS* NRPS modules. In the case of the lovastatin fusions, products were still obtained with the construct containing the *eqxS* ACP alone, implicating substrate selectivity of the C domain as the major factor limiting transfer.

Diels−Alder Reaction in Lovastatin Biosynthesis. As noted above, the *lovB-eqxS* fusion led to a polyketide product **1** resembling a pre-Diels−Alder product. Natively, *lovB* includes not just the PKS module, but also an intact C domain and a short piece of the adenylation (A) domain that activates amino acids.[10,12] Previously, it has been speculated that the *lovB* C domain was responsible for Diels−Alderase activity.[6,10,28] This previous idea is strongly supported by the chemistry discovered here, the first time in which the pre-Diels−Alder product has been observed as a *lovB* PKS product. To further probe this issue, we fused *lovB* with *fsdS* at the position in which the *lovB* A domain is fragmented. This construct was expressed in tandem with *lovC*, and it yielded the lovastatin precursor, dihydromonacolin L (Supporting Information Figure S5). Thus, fusions with an intact *lovB* C domain yield the Diels−Alder product, while fusions to heterologous C domains do not and in some cases lead to pre-Diels−Alder products. This provides strong evidence supporting the C domain as the region responsible for the Diels−Alder reaction.

Of note, in the case of equisetin and other tetramic acids, the tetramic acid motif remains adjacent to the decalin ring system, in the perfect position to accelerate a Diels−Alder reaction. By contrast, in lovastatin the final PKS intermediate is chain extended so that a methylene group is adjacent to the decalin ring. Thus, chemically, it is highly likely that the decalin ring in lovastatin must be formed at a chain length in which the enzyme-bound thioester is immediately adjacent to the nascent decalin; it is formed while still bound to the PKS. By contrast, in principle the equisetin decalin ring could still be formed after cleavage from the NRPS. Indeed, in a previous study, we found evidence that the decalin rings in the tetramic acid pyrrolocin may be formed post-PKS.[25] Based upon these results and the ideas of Hutchinson, Vederas, Tang, and others,[6,10,12,28] we speculate that the substrate selectivity and chemical recognition inherent to C domains has been redirected to template the Diels−Alder reaction in lovastatin synthesis.

C Domain Selectivity and KR Domain Activity. In previous studies, the ketoreductase (KR) domain was shown to control chain length in some fungal PKS-NRPS proteins.[7] This provided us with an opportunity to more finely investigate the role of C domain in substrate selectivity, removing protein−protein recognition from consideration and focusing solely on the chain length presented to C domain. We swapped the *eqxS* KR domain for the *lovB* KR domain, in the context of the full-length *eqxS* PKS-NRPS (Figure 5). Since LovB and EqxS natively synthesize nonaketides and octaketides respectively, it was envisioned that slightly different products would be presented to C domain. When this chimera was coexpressed with the trans-ER *eqxC*, several new metabolites were detected. The major product was characterized to be the polyketide **8**

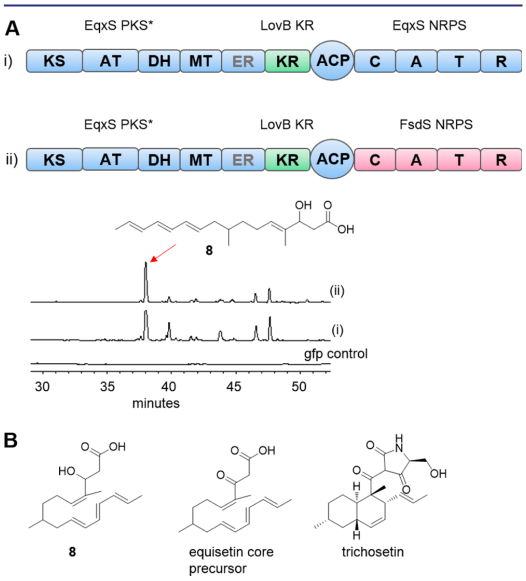


Figure 5. Analysis of chimeric *eqxS* PKS with *lovB* KR. (A) The *eqxS* KR was swapped for the LovB KR domain (green) in *eqxS* (i), and in *eqxS-fsdS* (ACP1-to-Cond2; (ii)). HPLC-DAD analysis of crude extracts of CGA cultures of FusΔeqxS cotransformed with eqxC and either (i) or (ii) shows the production of **8**. Trichosetin and **7** are synthesized in minor amounts by (i) and (ii), respectively. (B) Comparison of **8** with the equisetin polyketide core and trichosetin showing the position of the *β*-hydroxyl group.

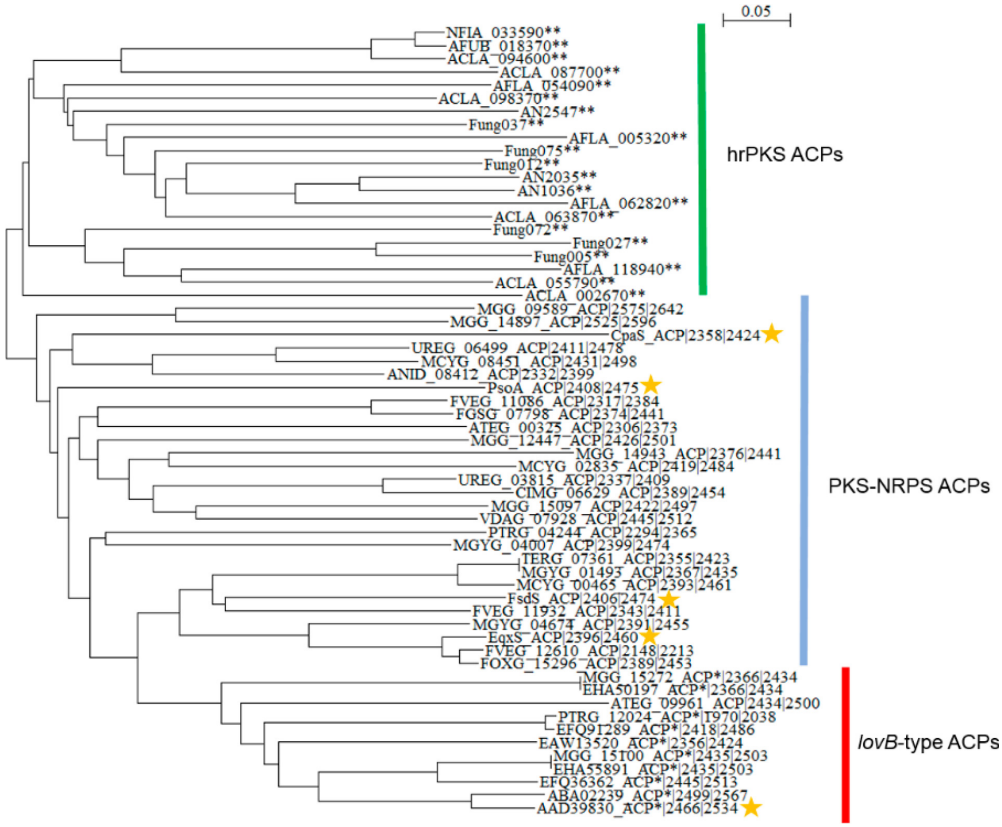


Figure 6. A phylogenetic analysis of ACPs from hrPKSs, PKS-NRPSs, and *lovB*-type PKSs (with truncated NRPS). Starred ACPs were investigated in this study. The *eqxS* ACP and *fsdS* ACP are more closely related to each other than they are to *psoA*, *cpaS*, and *lovB* ACPs.

which is not conjugated to serine (Figure 5A). To our surprise, **8** was the same length as the equisetin pre-Diels–Alder polyketide chain, but with a β-hydroxyl group in place of the keto moiety (Figure 5B). A very minor amount of trichosetin, normally the major product of *eqxS* + *eqxC*, was observed. In contrast to expectation, *eqxC* cooperated with the chimeric protein to perform the two normal reductions observed in equisetin synthesis.

Similarly, a fusion was made in which the *lovB* KR was swapped into the *eqxS-fsdS* chimera. The same major pre-Diels–Alder compound **8** was again observed, with a very minor amount of **7**, the tyrosine analog of equisetin (Figure 5). Both KR swaps provided the same result: simple reduction of the β-keto group absolutely abolished transfer to the NRPS module. Only a small amount of product, escaping the kinetic reduction of the β-keto group, was captured by the equisetin and fusaridione NRPS domain and added to an amino acid. These results serve to strongly reinforce the role of C domain in selecting the precise substrate for chain elongation.

Potential Impact of R Domain on Selectivity. The R domains at the C-terminus of *eqxS* and *fsdS* belong to a subset of proteins that catalyze Dieckmann cyclization to afford tetramic acids; they are not competent reductases and do not bind to NAD(P)H.[29,30] Previously, biochemical experiments with FsdS R and ATR domains showed that, in the case of FsdS, R has broad substrate selectivity that would not be expected to impact the experiments described here.[30] For example, acetoacetyl-alanine was readily accepted by the domains in vitro. To further probe this issue, we examined the products of *lovB-eqxS* and *cpaS-eqxS* fusions, in which the terminal R domain was covalently linked to green fluorescent protein (sGFP). In previous work with full length *eqxS*, we showed that introduction of sGFP in this position blocked activity of R, so that hydrolytic products were obtained instead of Dieckmann cyclase products.[25] The fusion products of *lovB-eqxS-sGFP* were identical to those from *lovB-eqxS*, and as in the previous experiments no products were detected from *cpaS-eqxS-sGFP*. This experiment provides further evidence that the primary chemical selectivity for the *eqxS* NRPS is likely to be at the C domain.

■ DISCUSSION

Here, we answer some of the key unsolved questions about the function and engineering of fungal PKS-NRPS hybrids. In total, we synthesized and tested 57 fusion products, including 32 PKS-NRPS module swaps, 2 KR swaps, 1 *lovB* C-A fusion, and 22 *GFP* fusions to the PKS-NRPS hybrids. By analyzing the chemical products resulting from these fusions, we tested the hypothesis of ACP domain specificity, revealing specificity

elements that will be essential in engineering. We showed that the *eqxS* and *fsdS* C domains are highly substrate selective. Our secondary goal was to better understand fungal hr-PKS function. We provide new insights about the role of specific domains in determination of Diels–Alderase activity and chain-length determination. More fundamentally, we show that the PKS module alone is sufficient to provide the expected polyketide structure, and that C domain does not contribute beyond the Diels–Alder reaction. We demonstrate for the first time a new product resulting from a functional in cis fusion between complete PKS and NRPS modules. The first pre-Diels–Alder lovastatin enzymatic products were discovered, which in tandem with other evidence provide strong support for the role of C domain in lovastatin Diels–Alder reaction.

The biochemical basis of the Diels–Alder reaction has been the source of great interest for the past 15 years, but it has not yet been completely resolved. Elegant work by Vederas and co-workers showed that LovB catalyzes the on-enzyme, stereo-specific Diels–Alder reaction to form the decalin ring found in the lovastatin precursor, dihydromonacolin L.[28] Later studies reported the inability of a truncated LovB, lacking the C domain, to form monacolins when coexpressed with LovC, but instead truncated pyrones are formed.[10] The LovB C domain was however able to restore monacolin production when added in trans to the truncated LovB in vitro. Another study demonstrated that LovB fused post-C domain with the chaetoglobosin A synthetase (CheA) was capable of synthesizing dihydromonacolin L, whereas the direct fusion of the lovB PKS with CheA C domain was nonproductive.[6]

In some ways, our results here are quite similar to the latter two studies. We found that, in the absence of the native C domain, LovB synthesizes polyunsaturated pyrones. When the intact LovB C and a small part of LovB A are added back, rather than tetramate formation, dihydromonacolin L is found. The major difference is that, for the first time, we have also identified a pre-Diels–Alder product **1** from fermentations with LovB lacking its native C domain. The isolation of **1** provides direct evidence showing concretely that all of the domains in LovB and LovC are functional, and that they synthesize an intermediate that would be cyclized if a functional Diels–Alderase were present. When that Diels–Alderase (C domain) is added back, the Diels–Alder reaction takes place. This evidence strongly favors the C domain hypothesis forwarded previously by other researchers. We were able to isolate and characterize this product because we designed our study to use an expression platform that natively produces a high yield of recombinant products.[25]

We thoroughly define the role of the ACP in governing PKS function. In pairing PKS domains with noncognate ACPs, we were first faced with the question of whether PKS domains, including the trans ER, would interact with the noncognate ACP to form the native PKS polyketide intermediate to present to the C domain. This question has been explored in several other types of PKS proteins,[22,31,32] but never before within the fungal hr-PKSs. Within the fungal hr-PKS and PKS-NRPS proteins, the role of ACP has been previously indirectly tested in a series of domain fusions within the closely related bassianin and tenellin PKSs.[7] In these fusions, the tenellin ACP was always in place, and it could interact with several other domains from the bassianin PKS. Our work explicitly and thoroughly explores the role of ACP in PKS interactions. The *eqxS* ACP and *fsdS* ACP are equivalent and interchangeable. The *lovB-eqxS* fusions reveal that the *eqxS* ACP can interact with the LovB PKS, but does not interact with LovC. Interactions with LovC are restored by reintroducing a portion of the *lovB* ACP in the hybrid ACP fusion. By contrast, the *eqxS* and *fsdS* ACPs are incapable of interacting with the *psoA* PKS domains.

A phylogenetic analysis of the ACPs (Figure 6) illuminates the mixed results from these ACP-PKS interaction studies. Strikingly, more closely related ACPs are interchangeable, while more distantly related ACPs are not. The intermediate case, involving the *lovB* swaps, provides an intermediate level of transferability. Therefore, for swaps involving the use of noncognate ACPs, source modules with closely related ACPs may result in production of predicted PKS intermediates.

Transfer of polyketide intermediates to NRPS modules has also been investigated in several cases. Within the fungal enzymes, only three other studies have investigated the compatibility of PKSs with noncognate NRPS modules. (1) The aspyridone synthetase (ApdA) PKS module and the CpaS NRPS module were expressed as separate proteins in *Saccharomyces cerevisiae*.[8] This in trans experiment led to successful production of a small amount of the predicted tetraketide-derived tetramate, whereas natively a diketide intermediate is accepted by the CpaS NRPS module. The ApdA PKS was shown to have the potential to synthesize longer polyketide intermediates, but only the tetraketide was captured by both the ApdA and CpaS C domains. (2) Highly similar proteins desmethylbassianin synthetase DmbS and tenellin synthetase TenS (>85% protein sequence identity) were fused, leading to a successful synthesis of similar products.[7,20] These proteins synthesized hexaketide and pentaketide tetramates with slightly different methylation patterns. (3) The *lovB* PKS was fused to the CheA NRPS, but no products were detected.[6] The incompatibility of LovB PKS with CheA NRPS was attributed to an early evolutionary divergence between the LovB-type PKSs and other PKS-NRPSs as observed from phylogenetic analyses.

A complicating factor in the above studies was that they did not differentiate between the impact of protein–protein and substrate selectivity interactions between modules, and they used only single swaps rather than thoroughly investigating multiple combinations. Here, we tested the effect of C domain selectivity independent of potential unfavorable ACP interactions for a larger set of PKS/NRPS pairings that differ significantly in sequence similarity. In the context of previous work by others, our thorough data here demonstrates that C domains of PKS-NRPS enzymes are often highly chemoselective. The *lovB-eqxS* fusions synthesize several nonamidated polyketide chains. Even in the presence of the *eqxS* ACP which forms interactions with *eqxS* C domain, the polyketides are not transferred to the NRPS. These cases imply an underlying strict *eqxS* C domain substrate selectivity beyond the presence of productive ACP interactions. This is further supported by the inability of similar *fsdS-eqxS* fusions to synthesize tetramate products, but instead form the polyketide **6**. In fact, the experiment with the *lovB* KR domain swap suggests that the *eqxS* C domain possesses a high selectivity for only the equisetin polyketide core because even when it is the disfavored/aberrant polyketide intermediate among others, trichosetin is the only amidated product detected. On the other hand, the *fsdS* C domain displays more relaxed substrate selectivity, accepting the equisetin polyketide core in addition to the more rigid polyene polyketide chain found in fusaridione. In these *eqxS-fsdS* expressions, many other nontetramate products are detected; it is therefore intriguing that eqxTyr **7**

is the only amidated product observed. One explanation for this is that the *fsdS* C domain is not truly accepting of broad substrates, but rather selects for particular attributes of the substrate which may include length, degree of unsaturation and/or methylation pattern. This becomes more apparent when the octaketide intermediate of equisetin is compared to the heptaketide polyketide product of *fsdS* PKS. It is possible that *fsdS* C accepts the linear pre-Diels−Alderase intermediate. The pseudorelaxed selectivity of the *fsdS* C domain for the equisetin core is further demonstrated in the *lovB* KR swap experiment, where trace amounts of eqxTyr were detected, when the equisetin polyketide core is produced as an aberrant product among others. Tenellin synthetase (TenS) produces several tetramates in the absence of its trans ER.[20] Upon the basis of this result, C domains have been speculated to be broad substrate, but our data shows that this is not universal. Even for TenS, the polyketide chains did not differ significantly among the tetramates synthesized; the lengths of these chains are within one ketide unit of the native substrate of the C domain. This difference is exactly what we found for the successful chain transfer in *eqxS-fsdS* fusions.

Our studies also enabled other interesting observations into PKS domain function when the *eqxS* KR was swapped for the *lovB* KR. The synthesis of the same equisetin core polyketide chain albeit with a β-hydroxyl group instead of the anticipated nonaketide chain, in constrast with previous work,[7] shows that iterative PKS programming is much more complex and each system should be evaluated individually until more universal rules can be discovered. The chain length factor for *eqxS* probably lies with the KS domain. This result also shows that, for the case of *eqxS* biosynthesis, capture of the polyketide intermediate by the C domain occurs after polyketide elongation is complete. KR and C domain compete with each other for the last intermediate in chain elongation, with the KR reaction dominating. A caveat in interpreting this result is that the β-hydroxyamide would not be a substrate for Dieckmann cyclization, so that later hydrolysis of the amide bond would provide the detected product even though successful chain transfer had in fact occurred. Arguing against this possibility is the fact that we have previously isolated non-Dieckmann products from experiments with both FsdS and EqxS, and they were perfectly metabolically stable.[11,25]

The results of this study, adding to previous work, will be useful in the engineering of fungal polyketides and polyketide-peptide hybrids. First, the work shows that the expected polyketide products themselves can be directly formed via expression of these fusions; merely swapping out C domains can provide valuable products. Second, it shows that successful engineering strategies will require appropriate selection of ACPs. Currently, it is clear from this work and previous studies that phylogenetically closely related ACPs are empirically likely to function in a heterologous context. However, more research is required to delineate exactly the recognition factors between ACPs and these complex enzymes that may facilitate rational swaps between PKSs. Finally, the substrate selectivity of C domains is a serious hurdle in the generation of recombinant products. The best way forward at this point will involve using C domains that accept somewhat similar PKS products, but that append different amino acids to the compound. The value of this approach is clearly shown by the very conservative previous successes in synthesizing bassianin and cyclopiazonic acid derivatives and in our case of switching the amino acid on equisetin from serine to tyrosine. Fortunately, fungal PKS-NRPS genes are ubiquitous,[33] so that an enormous biodiversity is available for such engineering efforts.

METHODS

Cloning of Fungal Expression Plasmids. Generally, genes were amplified from genomic DNA or subcloned coding sequences using the high-fidelity Phusion polymerase (NEB), and cloned into the expression vectors by yeast recombination.[34,35] Plasmid selection and amplification was then done in *Escherichia coli*. Details of plasmid construction are contained in Supporting Information. Alignments and phylogenetic analysis were generated with ClustalX.

Fungal Transformation. *Fusarium heterosporum* FusΔeqx5 and FusΔeqx5pyrG10 were used as the expression hosts.[25] The plasmids were linearized with PacI/AscI prior to transformation into FusΔeqx5 as previously described.[11,25] Selection was done with hygromycin 150 μg mL^{-1}, and isolated transformants cultivated in potato dextrose broth (PDB) 250 mL for 7 d at 30 °C and corn grit agar (CGA) 50 g for 21 d at rt. *psoA-eqxS* and *psoA-fsdS* mutants were cultured in PDB supplemented with sodium propionate (20 mM) at the 24-h point of the culture because propionyl CoA is the first unit loaded by PsoA.[15] Absent addition of propionate, no new compounds were observed in the *psoA* expressions.

General Procedures for Extraction, Purification, and Characterization of Compounds. PDB cultures were extracted with an equal volume of ethyl acetate, and solvent removed under a vacuum. The crude extracts were analyzed by HPLC using a Hitachi LaChrom Elite instrument equipped with a diode array detector over a Phenomenex Luna C18 column (4.6 × 250 mm, 5 μm). Preparative HPLC was done using the Discovery HS C18 column (25 cm × 10 mm, 5 μm). Extracts were further characterized by LC/MS using the Agilent ZQ to screen for new metabolites. Cultures on CGA were extracted with acetone and the crude extracts treated similarly. Compounds were generally purified by first performing flash chromatography on end-plugged C18 using a methanol/water gradient. The fractions containing the target compounds were then purified by preparative HPLC using an acetonitrile/water solvent system (with or without 0.05% TFA). The collected fractions from several rounds of HPLC were pooled and the solvent removed under a vacuum. HPLC-grade solvents were used as purchased. HRESIMS data for the purified compounds was obtained from a Waters Micromass Q-TOF Micro mass spectrometer. All NMR data was acquired on a Varian INOVA 500, except the ^{1}H, HSQC and HMBC−^{15}N data for **2** which was acquired on a Varian INOVA 600 fitted with a cryoprobe.

Purification of ***1***. Compound **1** was purified by HPLC using 80% acetonitrile/water mobile phase containing 0.05% TFA. Solvent was removed under a vacuum to afford **1** (1.3 mg); HRESIMS m/z 233.1883 $[M + H]^+$ (Calcd for $C_{16}H_{25}O$ 233.1900; Δ −7.3 ppm); 1D and 2D NMR data (Supporting Information Table S5; Figure S6).

Purification of ***2***. Compound **2** was purified by HPLC using 40% ACN/20% water mobile phase containing 0.15% TFA. Removal of solvent under a vacuum yielded **2** (0.4 mg); HRESIMS m/z 398.1994 $[M + H]^+$ (Calcd for $C_{20}H_{32}NO_5S$ 398.1996; Δ −0.5 ppm); 1D and 2D NMR data (Supporting Information Table S2; Figure S7).

Purification of ***6***. The crude extract from a PDB culture (500 mL) was subjected to flash chromatography over end-plugged C18. The fraction containing **6** was dried under a vacuum and the residue further purified by HPLC using an 85% ACN/15% water mobile phase. Solvent was removed under a vacuum to afford **6** (1.1 mg); λ_{max} = 373 nm; HRESIMS m/z 245.1875 $[M + H]^+$ (Calcd for $C_{17}H_{25}O$ 245.1900; Δ −10.2 ppm); 1D and 2D NMR data (Supporting Information Table S4; Figure S8).

Purification of eqxTyr ***7***. A portion of the semipurified fraction from flash chromatography over end-plugged C18 was further purified by HPLC using a 75% ACN/25% water mobile phase containing 0.05% TFA. The collected fractions were pooled and solvent was removed by vacuum to afford an off-white powder (0.9 mg); HRESIMS m/z 436.2492 $[M + H]^+$ (Calcd for $C_{27}H_{34}NO_4$

436.2482; Δ 2.3 ppm); 1D and 2D NMR data (Supporting Information Table S3; Figure S9).

Purification of ***8***. Compound **8** was purified by HPLC using a 67% ACN/33% water mobile phase containing 0.05% TFA. Solvent was removed to afford **8** (1.8 mg); HRESIMS *m*/*z* 291.1963 $[M-H]^-$ (Calcd for $C_{18}H_{27}O_3$ 291.1966; Δ 1.0 ppm); 1D and 2D NMR data (Supporting Information Table S7; Figure S10).

Characterization of ***3***, ***4***, *and Dihydromonacolin L*. These were characterized by comparing UV absorbance data, *m*/*z* values, and MS fragmentation to previously reported data.[10,25]

Characterization of ***5***. Compound **5** was purified by preparative HPLC with a gradient running from 30% to 70% acetonitrile/water in 20 min. Pooled fractions from several rounds of HPLC were dried under a vacuum to afford **5** (1.1 mg); HRESIMS *m*/*z* 193.0863 $[M + H]^+$ (Calcd for $C_{11}H_{13}O_3$ 193.0859; Δ 2.1 ppm). ^{1}H NMR spectra displayed broad signals in both $CDCl_3$ and CD_3OD, and 2D NMR showed weak signals. On the basis of this NMR data, we predicted **5** to be an unsaturated pyrone.[12] After purification, **5** was kept in the dark, under argon. ^{1}H and COSY NMR data indicated an unsaturated side chain as shown in the structure diagram, but like other compounds in this class[12] the compound was not sufficiently well behaved for NMR characterization. Therefore, characterization of **5** was done by comparative gas chromatography-electron impact-MS (GC-EI-MS) with a sample of **4**. To dried samples, methoxyamine hydrochloride (40 μL, 40 mg mL^{-1} in pyridine) was added and the mixture heated at 40 °C for 1 h. After cooling, MSTFA (40 μL) was added and the reaction mixture was left at rt for 12 h. GC/MS analyses were conducted using an HP6890 instrument interfaced with an MSD-HP5973 detector and equipped with a Zebron ZB-5MSi Guardian (30 m × 0.25 mm ID, 0.25 μm film thickness; Phenomenex) column and an HP7682 injector. Helium was used as a carrier gas with a 100:1 split ratio at an injection volume of 1 μL. The injector temperature was set to 250 °C. The oven temperature gradient was programmed as follows: 95 °C held for 1.5 min increased at a rate of 40 °C/min to 118 °C, held for 1 min, increased at a rate of 5 °C/min to 250 °C, increased at a rate of 25 °C/min to 330 °C and held for 12.3 min. MS spectra were obtained in EI mode within a range of *m*/*z* 50−500. Other parameters: MS quad and source temperatures were set to 150 and 230 °C, respectively; solvent cut time was 4 min; and scanning was done at 16 scans/sec. MS fragmentation clearly supported the proposed structure **5**; see Supporting Information Table S8.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Supporting tables and figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

ews1@utah.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was funded by NSF 0957791. We thank Dr. Shiou-Chuan Tsai (UC Irvine) for the generous gift of plasmid YEpADH2p-lovB-His, Dr. David J. Stillman (University of Utah) for plasmid M1192, J.A. Maschek (University of Utah) for GC-EI-MS analysis, J. Skalicky (University of Utah) for help with acquiring NMR data, and T.E. Smith (University of Utah) for technical assistance with acquiring HRESIMS data.

■ REFERENCES

(1) Boettger, D.; Hertweck, C. *ChemBioChem* **2013**, *14*, 28.
(2) Fisch, K. M. *RSC Adv.* **2013**, *3*, 18228.
(3) Carter, S. B. *Nature* **1967**, *213*, 261.
(4) Graupner, P. R.; Carr, A.; Clancy, E.; Gilbert, J.; Bailey, K. L.; Derby, J.-A.; Gerwick, B. C. *J. Nat. Prod.* **2003**, *66*, 1558.
(5) Song, Z.; Cox, R. J.; Lazarus, C. M.; Simpson, T. T. *ChemBioChem* **2004**, *5*, 1196.
(6) Boettger, D.; Bergmann, H.; Kuehn, B.; Shelest, E.; Hertweck, C. *ChemBioChem* **2012**, *13*, 2363.
(7) Fisch, K. M.; Bakeer, W.; Yakasai, A. A.; Song, Z.; Pedrick, J.; Wasil, Z.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. *J. Am. Chem. Soc.* **2011**, *133*, 16635.
(8) Xu, W.; Cai, X.; Jung, M. E.; Tang, Y. *J. Am. Chem. Soc.* **2010**, *132*, 13604.
(9) Chooi, Y. H.; Tang, Y. *J. Org. Chem.* **2012**, *77*, 9933.
(10) 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.
(11) Kakule, T. B.; Sardar, D.; Lin, Z.; Schmidt, E. W. *ACS Chem. Biol.* **2013**, *8*, 1549.
(12) Kennedy, J.; Auclair, K.; Kendrew, S. G.; Park, C.; Vederas, J. C.; Richard Hutchinson, C. *Science* **1999**, *284*, 1368.
(13) Ames, B. D.; Chi, N.; Bruegger, J.; Smith, P.; Xu, W.; Ma, S.; Wong, E.; Wong, S.; Xie, X.; Li, J. W. H.; Vederas, J. C.; Tang, Y.; Tsai, S.-C. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 11144.
(14) Seshime, Y.; Juvvadi, P. R.; Tokuoka, M.; Koyama, Y.; Kitamoto, K.; Ebizuka, Y.; Fujii, I. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 3288.
(15) Maiya, S.; Grundmann, A.; Li, X.; Li, S.-M.; Turner, G. *ChemBioChem* **2007**, *8*, 1736.
(16) Tsunematsu, Y.; Fukutomi, M.; Saruwatari, T.; Noguchi, H.; Hotta, K.; Tang, Y.; Watanabe, K. *Angew. Chem., Int. Ed.* **2014**, *53*, 8475.
(17) Phillips, N. J.; Goodwin, J. T.; Fraiman, A.; Cole, R. J.; Lynn, D. G. *J. Am. Chem. Soc.* **1989**, *111*, 8223.
(18) Halo, L. M.; Heneghan, M. N.; Yakasai, A. A.; Song, Z.; Williams, K.; Bailey, A. M.; Cox, R. J.; Lazarus, C. M.; Simpson, T. J. *J. Am. Chem. Soc.* **2008**, *130*, 17988.
(19) Halo, L. M.; Marshall, J. W.; Yakasai, A. A.; Song, Z.; Butts, C. P.; Crump, M. P.; Heneghan, M.; Bailey, A. M.; Simpson, T. J.; Lazarus, C. M.; Cox, R. J. *ChemBioChem* **2008**, *9*, 585.
(20) Heneghan, M. N.; Yakasai, A. A.; Williams, K.; Kadir, K. A.; Wasil, Z.; Bakeer, W.; Fisch, K. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J.; Lazarus, C. M. *Chem. Sci.* **2011**, *2*, 972.
(21) Schobert, R.; Schlenk, A. *Bioorg. Med. Chem.* **2008**, *16*, 4203.
(22) Kapur, S.; Chen, A. Y.; Cane, D. E.; Khosla, C. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 22066.
(23) Mofid, M. R.; Finking, R.; Marahiel, M. A. *J. Biol. Chem.* **2002**, *277*, 17023.
(24) Gu, L.; Eisman, E. B.; Dutta, S.; Franzmann, T. M.; Walter, S.; Gerwick, W. H.; Skiniotis, G.; Sherman, D. H. *Angew. Chem., Int. Ed.* **2011**, *50*, 2795.
(25) Kakule, T. B.; Jadulco, R. C.; Koch, M.; Janso, J. E.; Barrows, L. R.; Schmidt, E. W. *ACS Synth. Biol.* **2014**, DOI: 10.1021/sb500296p.
(26) Tokuoka, M.; Seshime, Y.; Fujii, I.; Kitamoto, K.; Takahashi, T.; Koyama, Y. *Fungal Genet. Biol.* **2008**, *45*, 1608.
(27) Wang, W.; Ballatori, N. *Pharmacol. Rev.* **1998**, *50*, 335.
(28) Auclair, K.; Sutherland, A.; Kennedy, J.; Witter, D. J.; Van den Heever, J. P.; Hutchinson, C. R.; Vederas, J. C. *J. Am. Chem. Soc.* **2000**, *122*, 11519.
(29) Liu, X.; Walsh, C. T. *Biochemistry* **2009**, *48*, 8746.
(30) Sims, J. W.; Schmidt, E. W. *J. Am. Chem. Soc.* **2008**, *130*, 11149.
(31) Bruegger, J.; Haushalter, B.; Vagstad, A.; Shakya, G.; Mih, N.; Townsend, Craig A.; Burkart, M. D.; Tsai, S.-C. *Chem. Biol.* **2013**, *20*, 1135.
(32) Ma, Y.; Smith, L. H.; Cox, R. J.; Beltran-Alvarez, P.; Arthur, C. J.; Simpson, T. J. *ChemBioChem* **2006**, *7*, 1951.
(33) von Döhren, H. *Fungal Genet. Biol.* **2009**, *46*, S45.
(34) Gietz, R. D.; Schiestl, R. H.; Willems, A. R.; Woods, R. A. *Yeast* **1995**, *11*, 355.
(35) Ma, H.; Kunes, S.; Schatz, P. J.; Botstein, D. *Gene* **1987**, *58*, 201.

Supporting Information

Combinatorialization of Fungal Polyketide Synthase-Peptide Synthetase Hybrid Proteins

Thomas B. Kakule, Zhenjian Lin, and Eric W. Schmidt*

Department of Medicinal Chemistry, University of Utah, Salt Lake City, UT 84112 USA

*Corresponding author: ews1@utah.edu

Table of Contents

Plasmid construction

The general strategy was to amplify the genes from genomic DNA or previously subcloned coding sequences by PCR, and then insert them into expression vectors by yeast recombination. These vectors possessed the C-terminal sequence of either *eqxS* or *fsdS*, starting at the ACP and ending with the R domain. A C-terminal *sgfp* tag was utilized to select for protein-expressing clones by visualization on a Dark Reader (Clare Chemical).[1] The *eqxS-fsdS* and *fsdS-eqxS* clones were not tagged with *sgfp*, while all others were initially tagged for analysis. We found no difference in compound production for the reported constructs with or without the gfp tag, so that all constructs described in the manuscript result from vectors with gfp. The general shuttle vector construction and fungal transformation methods were previously described.[1]

Construction of Vectors.

FH-100 (eqxNRPSgfp): The *eqxS* NRPS (ACP inclusive) was amplified with primer pair vector2ACATR-F / eqxSNRPS-sGfp-R from *eqxS* coding sequence and cloned into vector FH-2.[1]

FH-101 (fsdNRPSgfp): The *fsdS* NRPS (ACP inclusive) was amplified with primer pair d3p-FsdNRPS-F / fsdNRPSgfp-R from previously reported plasmid deg2ACP-CATR[2] and cloned into PmeI-linearized FH-2.

FH-102 (eqxNRPS): The vector FH-1 was digested with NotI and PacI, and the fragment spanning the *eqxS* NRPS was cloned into FH-2 by yeast recombination, replacing the *sgfp* sequence.

FH-200 (eqxC+eqxNRPS): The fragment from the restriction enzyme digest of FH-1[1] with NotI and PacI, spanning the eqxS ACP-C-A-T-R was cloned into PmeI-linearized FH-2 by yeast recombination to replace the *sgfp* sequence.

FH-201 (eqxC+fsdNRPS): The *fsdS* NRPS sequence (ACP inclusive) was amplified from the previously reported deg2ACP-CATR plasmid[2] with primer pair d3p-fsdNRPS-F / d3p-fsdNRPS-R and subcloned into PmeI-linearized FH-1.[1] The resultant plasmid was digested with NotI and PacI, and the *fsdS* NRPS containing fragment transformed into yeast together with PmeI-linearized FH-2 for recombination to replace the *sgfp* sequence.

FH-300 (lovC+eqxNRPSgfp): The *eqxS* NRPS sequence was as described for vector FH100 and cloned into previously reported hphlovC+sGFP plasmid.[1]

FH-301 (lovC+fsdNRPSgfp): This vector was made similarly by cloning the *fsdS* NRPS sequence, amplified as described for FH101, into the hphlovC+sGFP plasmid.[1]

Construction of fungal expression plasmids.

The plasmids for expressing the PKS/NRPS gene fusions were constructed by amplifying the PKS sequence of the target gene by PCR, and then cloning this amplicon into the respective vector which had been linearized with PmeI. The recombination was done by transforming the DNA fragments into *S. cerevisiae* BY4741 to create the desired fusion. Table S2 summarizes the pairings of primers for each fusion construct made, the template DNA used for PCR, and the vector into which the amplicon was cloned.

The fusion of full-length *lovB* with the *fsdS* NRPS within the adenylation domain was obtained by amplifying the *lovB* sequence from plasmid YEp-ADH2p-lovB-His with primer pair LovPKS-F / LovBAden2Aden-r and cloning it into PmeI-linearized FH-1. For the expression, this plasmid was co-

transformed into FusΔeqx5 with plasmid pyrGlovC+sGfP which was made by replacing the *eqxC* sequence in FH-4 with the *lovC* sequence.

KR domain swap: The sequences overlapping the *eqxS* KR were cloned into plasmid M1192 (Stillman Lab, University of Utah) to make plasmid YEpPKSlessKRgfp. The *lovB* KR sequence was amplified with primer pair LovB_KR-F / LovB_KR-R from plasmid YEpADH2p-lovB-His. The SexAI / AflII cut amplicon was ligated into YEpPKSlessKRgfp linearized with the same enzymes. The resultant plasmid YEpLovBKR was cut with AscI / PacI, and 4.7 kbp fragment co-transformed with the large fragment from an AleI digest of hpheqxC+eqxS into yeast for recombination. The resultant plasmid hpheqxC+eqxS*lovBkr was used for fungal expression. To make the *fsdS* NRPS fused construct, the primer pair deg3vect_1f / EqxA2C-d2-R was used to amplify the PKS sequence from hpheqxC+eqxS*lovBkr which was cloned into PmeI-cut FH-201.

Recombinant proteins are translated

Out of the 32 PKS-NRPS gene fusions, 16 led to functional products. A remaining question was whether those 16 failed fusions were actually translated. As alluded to in the manuscript, several different gene fusions were expressed both with and without covalent C-terminal sGFP tags (22 in total, including all 16 that did not lead to functional products). When transferred to *F. heterosporum*, all 22 GFP-tagged mutants were fluorescent under conditions that induce the expression system, indicating that the desired gene products were translated. In addition, both GFP-tagged and untagged versions led to production of identical compounds (or both to no compounds) in *F. heterosporum*, as gauged by comparing their HPLC-DAD and HPLC-MS profiles.

In addition, all of the proteins used are competent catalysts absent their fusion partners. For example, we previously expressed *cpaS*, which produced 1 g kg^{-1} of the expected product.[1] For the remainder of proteins used, all are functional in at least one fusion construct described in this study. All of the genes expressed comprise intron-less DNA, with the exception of *fsdS*, which contains its native introns but which is competently spliced by and natively found in *F. heterosporum*.

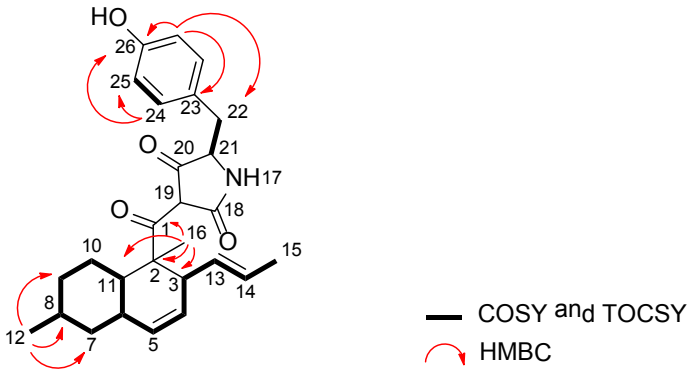
Table S1. Primers used in this study

Primer ID	Sequence
d3p-fsdNRPS-F	ATTTTATTACCGTCCCATTGCAGCTTGCTTGACAGTTTAAACTCTGCATCATCAAAGGTGCC
fsdNRPSgfp-R	ACCACCCCGGTGAACAGCTCCTCGCCCTTGCTCACTCTCCTCGACACCATCTCATC
d3p-fsdNRPS-R	CTAATCCATTATACCAAGTTGTGTGCCAACCTTAAATTATTATCTCCTCGACACCATCTCAT
LovBAden2Aden-r	GGTAATGTCCTCATCAGCCCAAACAGATGGCGCCATGATTGCCAGCTTCAGGGCGGGAT
LovA2C-d3R	CTTGAGTCTTGACGGTCCCAAGATGTCAAGGGAGTGCCTCCCGAAACTTCATTCTCGGAAG
LovHA-d3R	TCTTTGATGAACCAGGAACGGACCTCGACTGCCATCAAAGAGTCCACCCCCTGATCGAT
LovK2A-d3R	GCAAGTTGTACTCTCACCGGCACAGTAGATGTCTTTCCCCCTTCGCCTGCTTTGGC
LovTA-d3R	GCAAGTTGTACTCTCACCGGCACAGTAGATGTCTTTCCTCCCGAAACTTCATTCTCGGAAG
LovBACP2Cond-r	TGAAGAGCCAGTGCCGTCATTGGAAGTGGTTGAGGTGGGTCCCGAAACTTCATTCTCGGAAGT
LovBhybridACP-r	GCTCTTCATGAACCAGGATCGAATGTCCAAAGCAACCAAAGAGTCCACCCCCTGATCGATT
LovBKR2ACP-r	CTTGAGCTGCACTCGCACCGGCACCTTTGATGATGCAGACCCTTCGCCTGCTTTGGCC
LovBACP2ACP-r	TGAGCTGCACTCGCACCGGCACCTTTGATGATGCAGATCCCGAAACTTCATTCTCGGAAGT
CpasA2C-R	TGAAGAGCCAGTGCCGTCATTGGAAGTGGTTGAGGTGGGGGATGACTCCGAACTAGTACCA
CpasHA-R	GCTCTTCATGAACCAGGATCGAATGTCCAAAGCAACCAATGAGTCACAGCCCAGGTCCA
CpasK2A-R	CTTGAGCTGCACTCGCACCGGCACCTTTGATGATGCAGAATCTTCCTGTCCAGCACACTG
CpasTA-R	CTTGAGCTGCACTCGCACCGGCACCTTTGATGATGCAGAGGATGACTCCGAACTAGTACC
CpasA2degC-R	GGCTTGAGTCTTGACGGTCCCAAGATGTCAAGGGAGTGCCTGACTCCGAACTAGTACCATCG
Cpas2deg-HA-R	GCTCTTTGATGAACCAGGAACGGACCTCGACTGCCATCAATGAGTCACAGCCCAGGTCCA
CpasK2degA-R	CGGCAAGTTGTACTCTCACCGGCACAGTAGATGTCTTTCCATCTTCCTGTCCAGCACACTG
Cpas2deg-TA-R	CGGCAAGTTGTACTCTCACCGGCACAGTAGATGTCTTTCCGGATGACTCCGAACTAGTACC
PseuPKS-F	AGATTTTATTACCGTCCCATTGCAGCTTGCTTGACAATGGTCTACACACACTCACCAAAGGAG
PseuA2C-d2-R	TGAAGAGCCAGTGCCGTCATTGGAAGTGGTTGAGGTGGGCACGGGTGCTATAGAGTCCGTTCG
PseuHA-d2-R	TGTCCACATCCAGCTCTTCATGAACCAGGATCGAATGTCGACGGCGACAAGGGAATCGATACC
PseuK2A-d2-R	CTTGAGCTGCACTCGCACCGGCACCTTTGATGATGCAGAAGAGTCCGCACCCGCTTGCG
PseuTA-d2-R	CTTGAGCTGCACTCGCACCGGCACCTTTGATGATGCAGACACGGGTGCTATAGAGTCCGTTCG
d3PseuA2C-R	TTGAGTCTTGACGGTCCCAAGATGTCAAGGGAGTGCCCACGGGTGCTATAGAGTCCGTTCG
d3PseuHA-R	TTTGATGAACCAGGAACGGACCTCGACTGCCATCAAGACGGCGACAAGGGAATCGATACC
d3PseuK2A-R	CGGCAAGTTGTACTCTCACCGGCACAGTAGATGTCTTTCCAGAGTCCGCACCCGCTTGCG
d3PseuTA-R	CAAGTTGTACTCTCACCGGCACAGTAGATGTCTTTCCCACGGGTGCTATAGAGTCCGTTCG
FsdPKS-F	GCAGATTTTATTACCGTCCCATTGCAGCTTGCTTGACAATGTCTGCCCCTCCTACAAAG
FsdA2C-d3-R	CTTGAGTCTTGACGGTCCCAAGATGTCAAGGGAGTGCCAGTCGACGTGCCGTTCAAACT
FsdHA-d3-R	TCTTTGATGAACCAGGAACGGACCTCGACTGCCATCAAAGAATCGACACCTCGCTCGATA
FsdK2A-d3-R	GCAAGTTGTACTCTCACCGGCACAGTAGATGTCTTTCCGGTCTCACCCTTGTTGCAAGA
FsdTA-d3-R	GCAAGTTGTACTCTCACCGGCACAGTAGATGTCTTTCCAGTCGACGTGCCGTTCAAACT
EqxA2C-d2-R	GAAGAGCCAGTGCCGTCATTGGAAGTGGTTGAGGTGGGGATGGGTGAGTGGGACGAGT
EqxHA-d2-R	TCTTTCATGAACCAGGATCGAATGTCCAAAGCAACCAAGGAGTCGACACCTTGCTCAAC
EqxK2A-d2-R	TTGAGCTGCACTCGCACCGGCACCTTTGATGATGCAGAAGTGTATGCCTGAGTATCCAGC
EqxTA-d2-R	TTGAGCTGCACTCGCACCGGCACCTTTGATGATGCAGAGATGGGTGAGTGGGACGAGT
LovB_KR-F	GCGTCGACCAGGTTATTTCCTGGCCCCGGGAG
LovB_KR-R	CGCACTTAAGTCCTCCGCCCGAGTGACGAAC

Table S2. Summary of PKS/NRPS gene fusion expression plasmid design showing primer pairings, template DNA for PCR, and destination vector.

PKS-NRPS type	**Vector**	**Template DNA**	**Forward primer**	**Reverse Primer**			
				KR ACP C	KR ACP C	KR ACP C	KR ACP ACP C
cpas-eqxS	FH-100	hphCpas [1]	cpaS2deg3-F [1]	CpasA2degC-R	CpasK2degA-R	Cpas2deg-HA-R	Cpas2deg-TA-R
cpas-fsdS	FH-101	hphCpas	cpaS2deg3-F	CpasA2C-R	CpasK2A-R	CpasHA-R	CpasTA-R
lovB-eqxS	FH-300	YEpADH2p-lovB-His [1]	LovPKS-F [1]	LovA2C-d3R	LovK2A-d3R	LovHA-d3R	LovTA-d3R
lovB-fsdS	FH-301	YEpADH2p-lovB-His	LovPKS-F	LovBACP2Cond-r	LovBKR2ACP-r	LovBhybridACP-r	LovBACP2ACP-r
fsdS-eqxS	FH-102	alcAfsdS plasmid [2]	FsdPKS-F	FsdA2C-d3-R	FsdK2A-d3-R	FsdHA-d3-R	FsdTA-d3-R
eqxS-fsdS	FH-201	hpheqxC+eqxS [1]	deg3vect-1F [1]	EqxA2C-d2-R	EqxK2A-d2-R	EqxHA-d2-R	EqxTA-d2-R
psoA-eqxS	FH-100	*A. fumigatus* gDNA	PseuPKS-F	d3PseuA2C-R	d3PseuK2A-R	d3PseuHA-R	d3PseuTA-R
psoA-fsdS	FH-101	*A. fumigatus* gDNA	PseuPKS-F	PseuA2C-d2-R	PseuK2A-d2-R	PseuHA-d2-R	PseuTA-d2-R

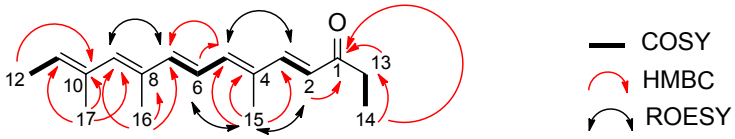
Table S3. NMR data for eqxTyr **7** in methanol-d_4. Accompanying figure shows key 2D NMR data.



Position	δ_C	δ_H (multiplicity, *J* in Hz)
1	204.3 C	
2	51.8 C	
3	43.6 CH	3.34 (m)
4	131.7 CH	5.51 (ddd, 10.0, 4,0, 3.3)
5	130.2	5.27 (d, 10.0)
6	35.8	1.92 (brs)
7	42.0	1.61 (brd, 12.0)
		1.12 (dd, 12.0, 5.4)
8	29.6	1.34 (m)
9	36.7	1.69 (m)
		0.89 (m)
10	24.0	1.61 (brd, 12.0)
		1.46 (dt, 13.0, 12.0)
11	37.9	2.73 (brs)
12	23.1	0.87 (d, 5.9)
13	127.8	5.44 (m)
14	133.5	5.44 (m)
15	17.9	1.66 (d, 4.9)
16	18.9	1.26 (s)
17		nd
18	nd	
19	nd	nd
20	nd	
21	49.2	3.98 (brs)
22	37.3	2.99 (m)
23	127.5	
24	131.6	6.96 (d, 8.0)
25	116.0	6.65 (d, 7.9)
26	157.2	

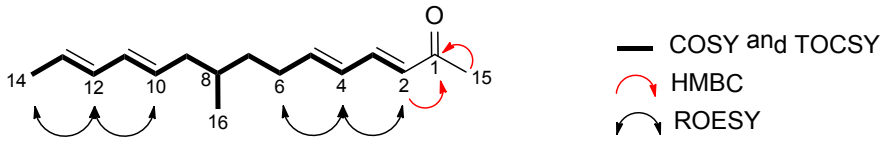
nd = not detected

Table S4. NMR data for **6** in dmso-d_6. Accompanying figure shows key 2D NMR data.



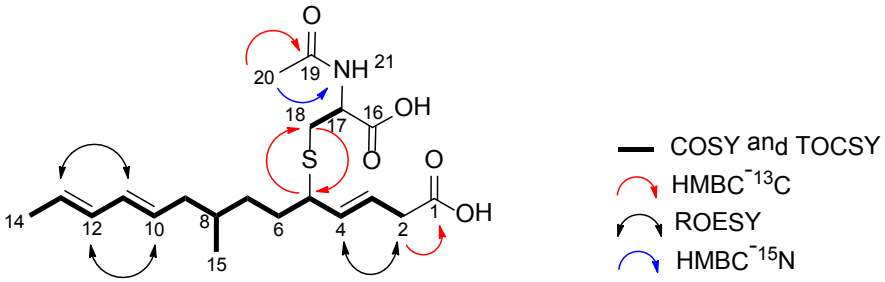
Position	**δ_C**	**δ_H (multiplicity, *J* in Hz)**
1	200.5	
2	124.9	6.18 (d, 15.8)
3	146.7	7.25 (d, 15.8)
4	133.4	
5	140.5	6.64 (m)
6	126.1	6.64 (m)
7	144.0	6.56 (d, 12.5)
8	133.1	
9	139.2	6.12 (s)
10	134.1	
11	128.4	5.57 (q, 6.7)
12	14.6	1.70 (d, 6.7)
13	33.5	2.59 (q, 7.5)
14	8.9	0.98 (t, 7.5)
15	12.8	1.90 (s)
16	14.4	1.95 (s)
17	17.0	1.79 (s)

Table S5. NMR data for **1** in chloroform-*d*. Accompanying figure shows key 2D NMR data.



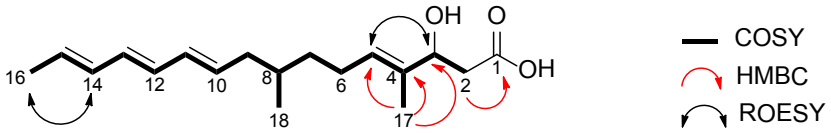
Position	**δ_C**	**δ_H (multiplicity, *J* in Hz)**
1	198.8	
2	128.8	6.08 (d, 15.9)
3	143.7	7.08 (dd, 15.9, 10.0)
4	145.7	6.19 (m)
5	144.0	6.18 (m)
6	30.8	2.23 (m)
		2.18 (m)
7	35.4	1.49 (m)
		1.26 (m)
8	32.9	1.52 (m)
9	40.1	2.05 (ddd, 13.7, 7.3, 6.1)
		1.94 (ddd, 13.7, 7.3, 7.2)
10	129.9	5.52 (dt, 14.6, 7.5)
11	131.7	5.97 (dd, 13.0, 10.5)
12	131.6	6.02 (dd, 13.0, 10.3)
13	127.0	5.61 (dt, 13.6, 6.8)
14	18.1	1.74 (d, 6.8)
15	27.2	2.27 (s)
16	19.2	0.89 (d, 6.6)

Table S6. NMR data for **2** in methanol-d_4. Accompanying figure shows key 2D NMR data.



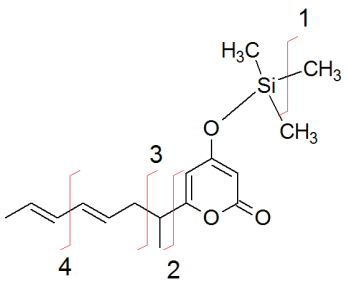
Position	**δ_C**	**δ_H (multiplicity, *J* in Hz)**
1	174.0	
2	36.7	3.09 (m)
3	119.1	5.65 (dt, 15.0, 7.3)
4	143.8	5.37 (dd, 15.0, 9.7)
5	48.1	3.26 (m)
6	31.8	1.60 (m)
		1.49 (m)
7	33.6	1.42 (m)
		1.31 (m)
8	32.9	1.49 (m)
9	39.4	2.05 (m)
		1.90 (ddd, 14.3, 7.2, 6.7)
10	129.5	5.49 (dt, 14.3, 7.2)
11	131.7	5.97 (m)
12	131.6	6.01 (m)
13	126.1	5.57 (dq, 7.2, 6.7)
14	16.5	1.71 (d, 6.8)
15	18.6	0.86 (d, 6.8)
16	172.3	
17	52.6	4.56 (t, 5.9)
18	31.7	2.89 (dd, 13.0, 5.7)
		2.83 (dd, 13.0, 6.6)
19	171.7	
20	21.1	2.00 (s)

Table S7. NMR data for **8** in methanol-d_4. Accompanying figure shows key 2D NMR data.

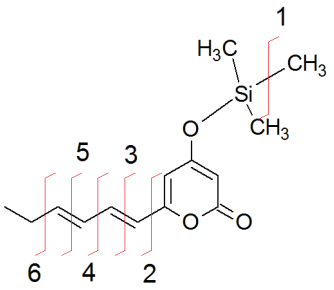


Position	**δ_C**	**δ_H (multiplicity, *J* in Hz)**
1	175.6	
2	42.0	2.47 (m)
3	75.2	4.40 (t, 7.3)
4	137.4	
5	127.8	5.44 (t, 7.3)
6	26.2	2.05 (m)
7	37.4	1.40 (m)
		1.21 (m)
8	34.2	1.54 (m)
9	41.4	2.10 (m)
		1.96 (ddd, 14.7, 7.9, 7.8)
10	133.4	5.61 (m)
11	133.3	6.05 (m)
12	131.9	6.01 (m)
13	132.2	6.07 (m)
14	133.5	6.06 (m)
15	129.6	5.64 (m)
16	18.5	1.75 (d, 6.5)
17	11.8	1.64 (s)
18	20.0	0.90 (d, 6.4)

Table S8. Comparative GC-EI-MS data for compounds **4** and **5** showing the MS fragments observed.



Compound **4**



Compound **5**

Position	Compound **4**		Compound **5**	
	Fragment 1 observed mass (calcd.) Da	Fragment 2 observed mass (calcd.) Da	Fragment 1 observed mass (calcd.) Da	Fragment 2 observed mass (calcd.) Da
1	277 (277)		249 (249)	
2	183 (183)		183 (183)	81 (81)
3	212 (211)	81 (81)	195 (196)	69 (68)
4	249 (251)		207 (209)	
5			221 (222)	
6			235 (235)	

Figure S1. Analytical HPLC-DAD of crude extracts of PDB cultures of FusΔeqx5 transformed with *psoA-eqxS* fusions (monitored at λ 330 nm).

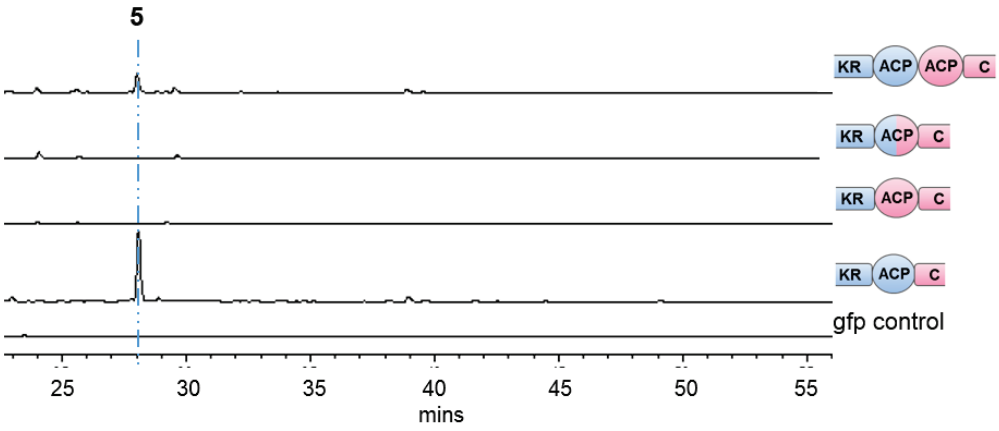


Figure S2. Analysis of crude extracts of CGA cultures of FusΔeqx5 transformed with *lovB-eqxS* fusions. A) HPLC-DAD analysis (monitored at λ 280 nm). B) LC/MS analysis, selected *m/z* [M+H]+ = 397.7–398.7.

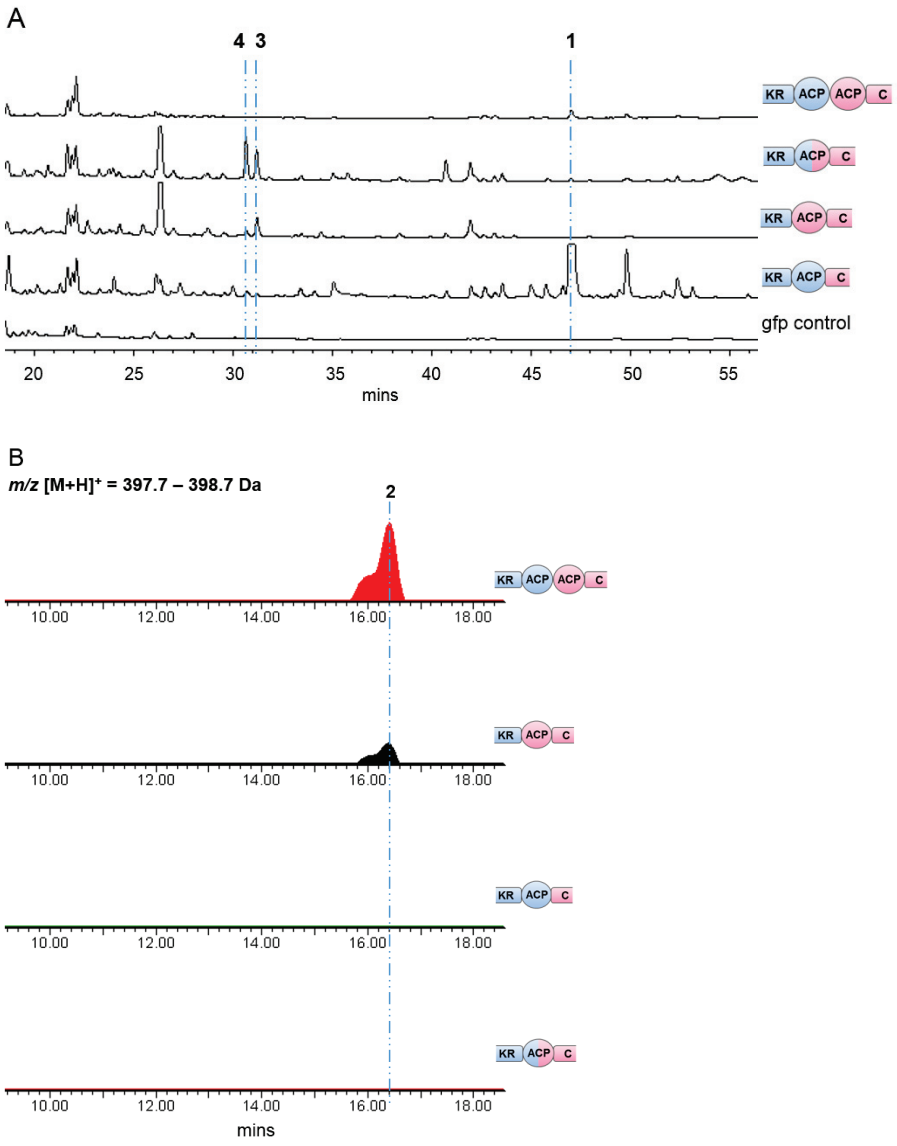


Figure S3. Analytical HPLC-DAD of crude extracts of PDB cultures of FusΔeqx5 transformed with *eqxS-fsdS* fusions (monitored at λ 280 nm).

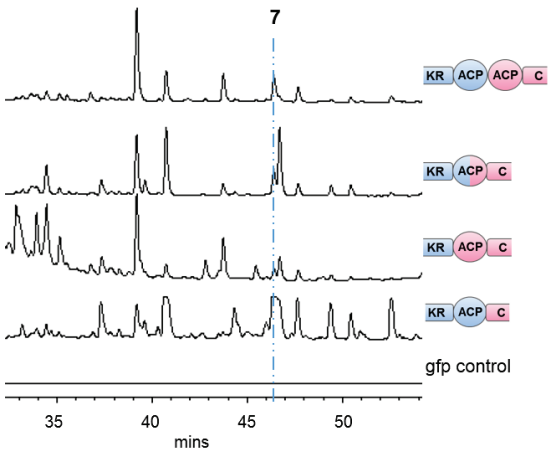


Figure S4. Analytical HPLC-DAD of crude extracts of PDB cultures of FusΔeqx5 transformed with *fsdS-eqxS* fusions (monitored at λ 330 nm).

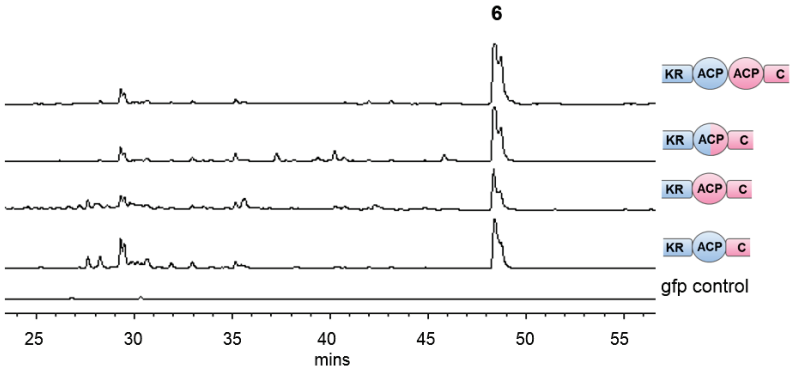


Figure S5. LC/MS analysis of a crude extract of a PDB culture of FusΔeqx5 co-expressiong *eqxC* and the gene fusion of full-length *lovB* with the *fsdS* NRPS within the A domain. Controls were the *gfp*-expressing mutant, and the mutant expressing *lovC* and *lovB*.[1]

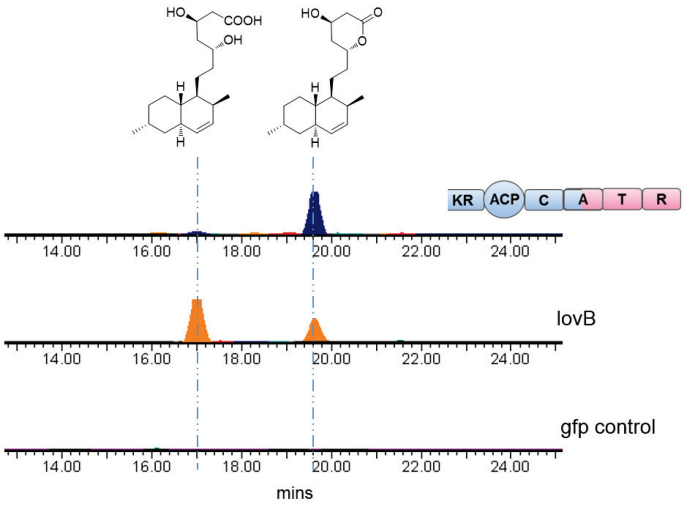


Figure S6. 1D and 2D NMR spectra for **1** in chloroform-*d*. A) ^{1}H B) ^{1}H-^{1}H COSY

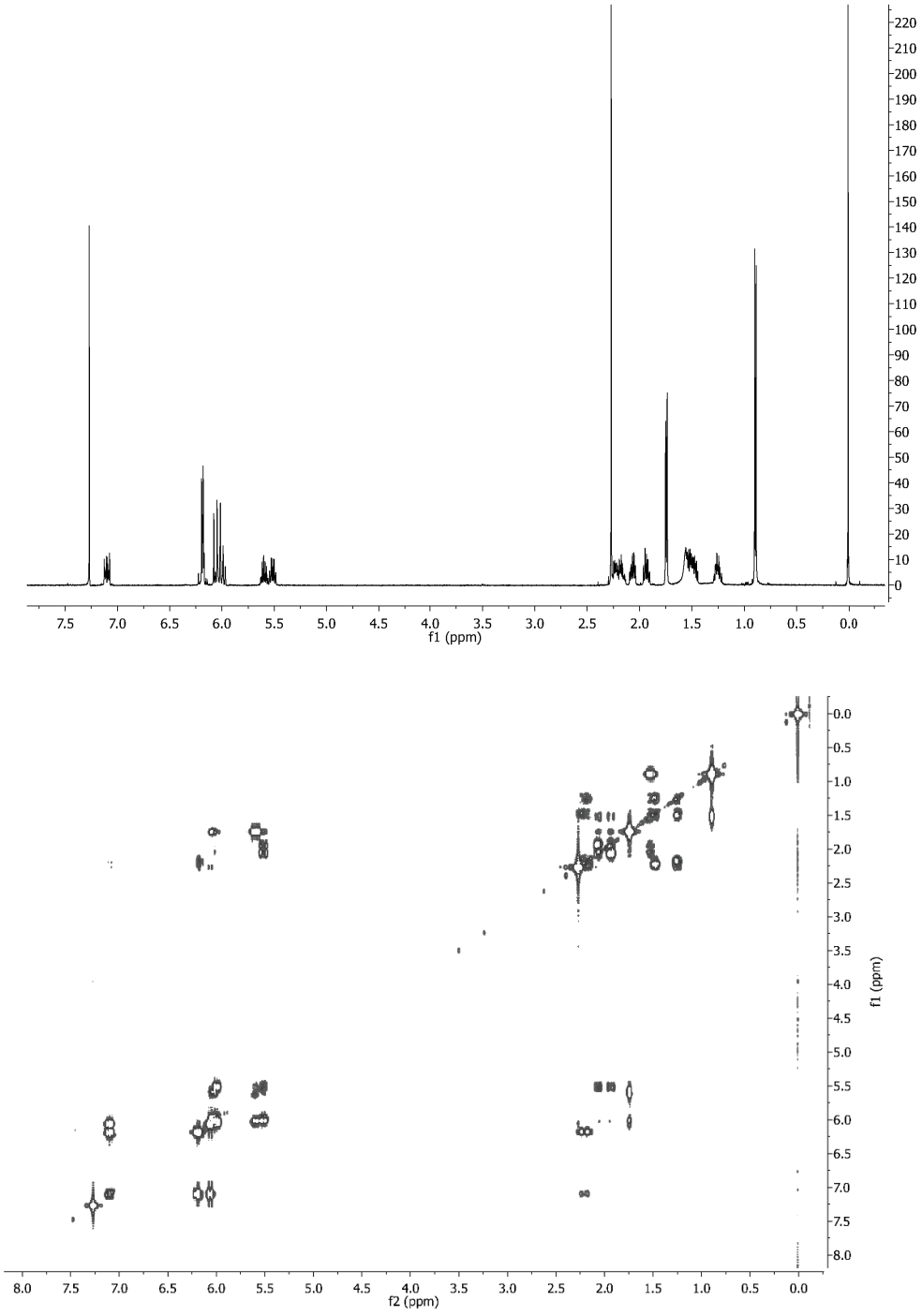


Figure S6. C) ^{1}H-^{13}C HSQC D) ^{1}H-^{13}C HMBC

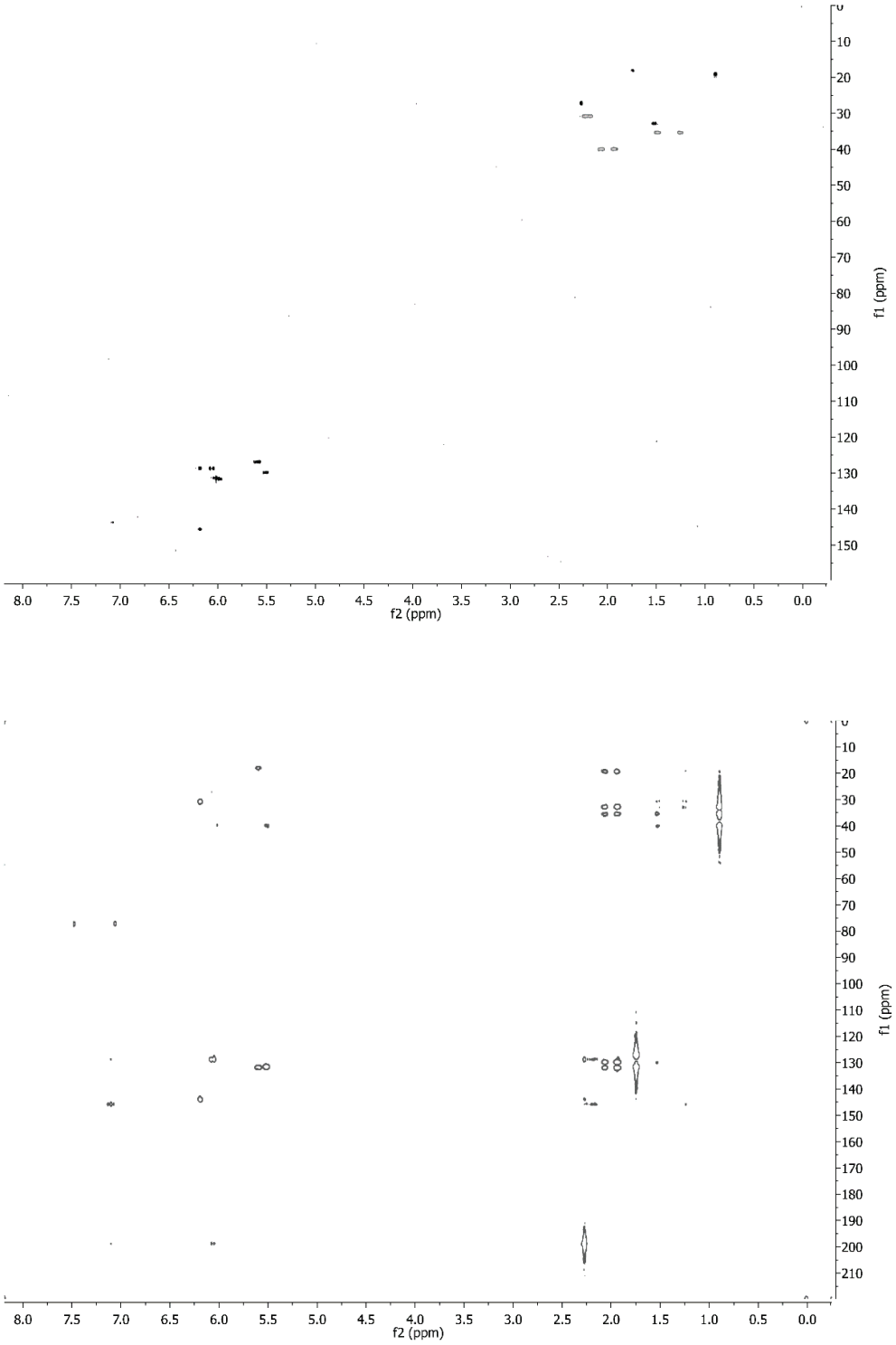
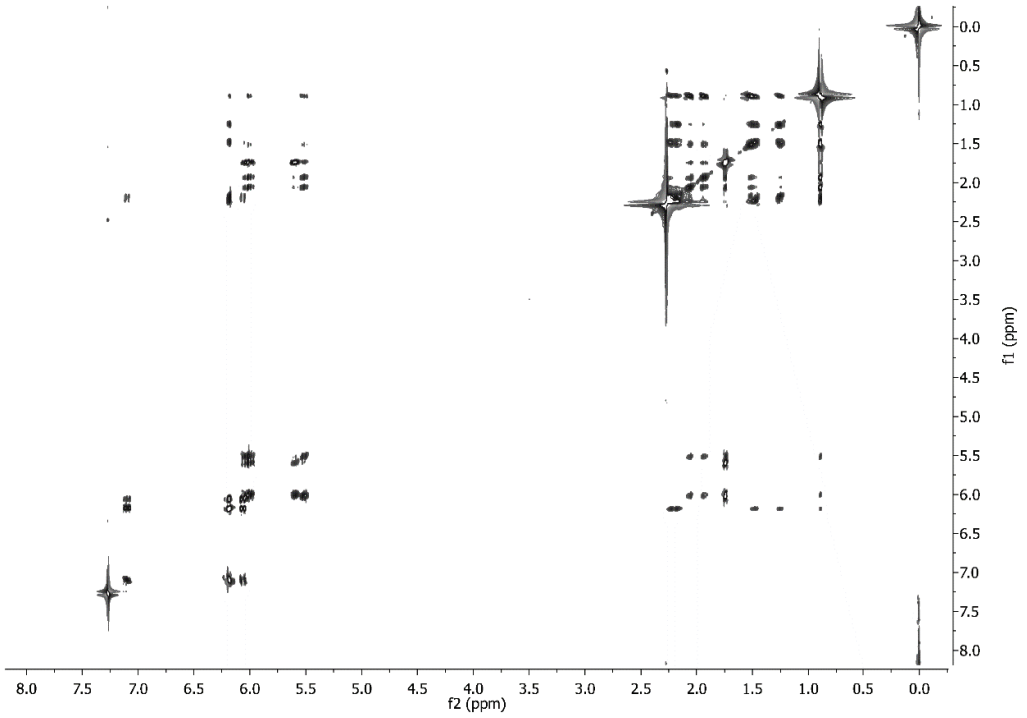


Figure S6. E) ^{1}H-^{1}H TOCSY F) ^{1}H-^{1}H ROESY



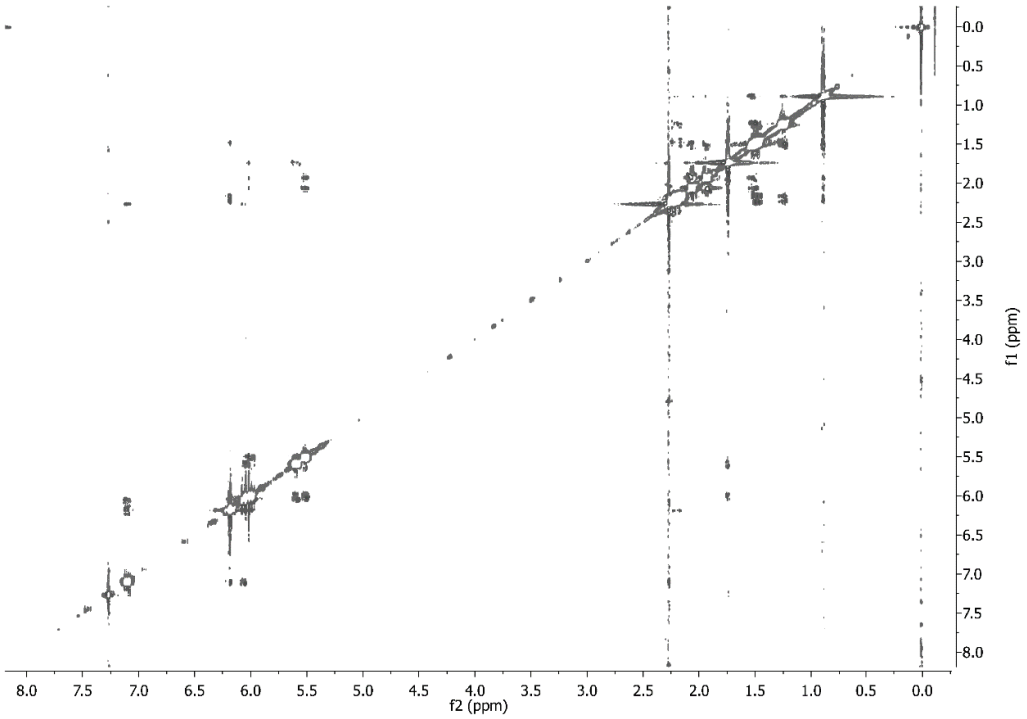


Figure S7. 1D and 2D NMR spectra for **2** in methanol-d_4. A) ^{1}H B) ^{1}H-^{1}H COSY

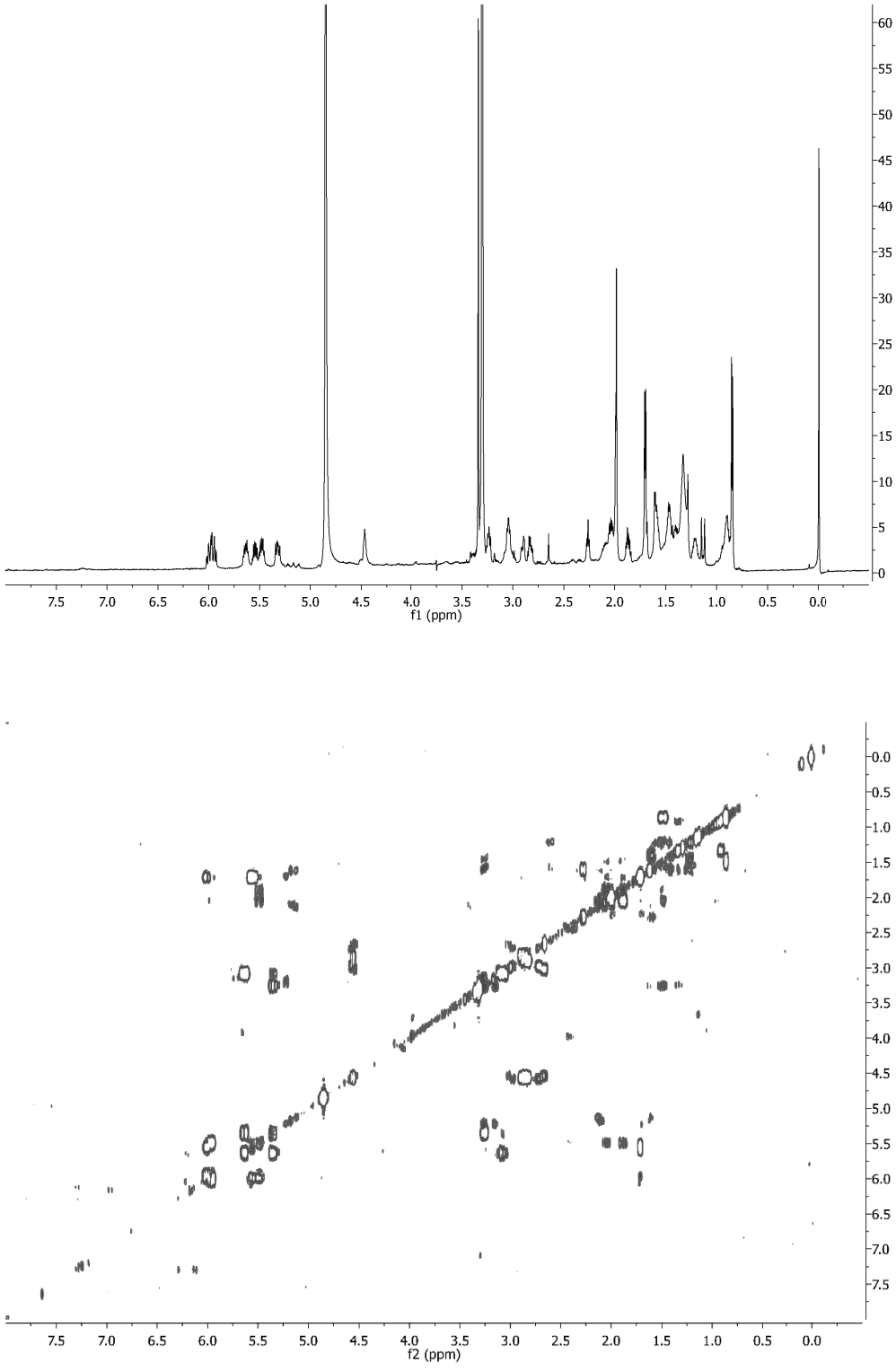
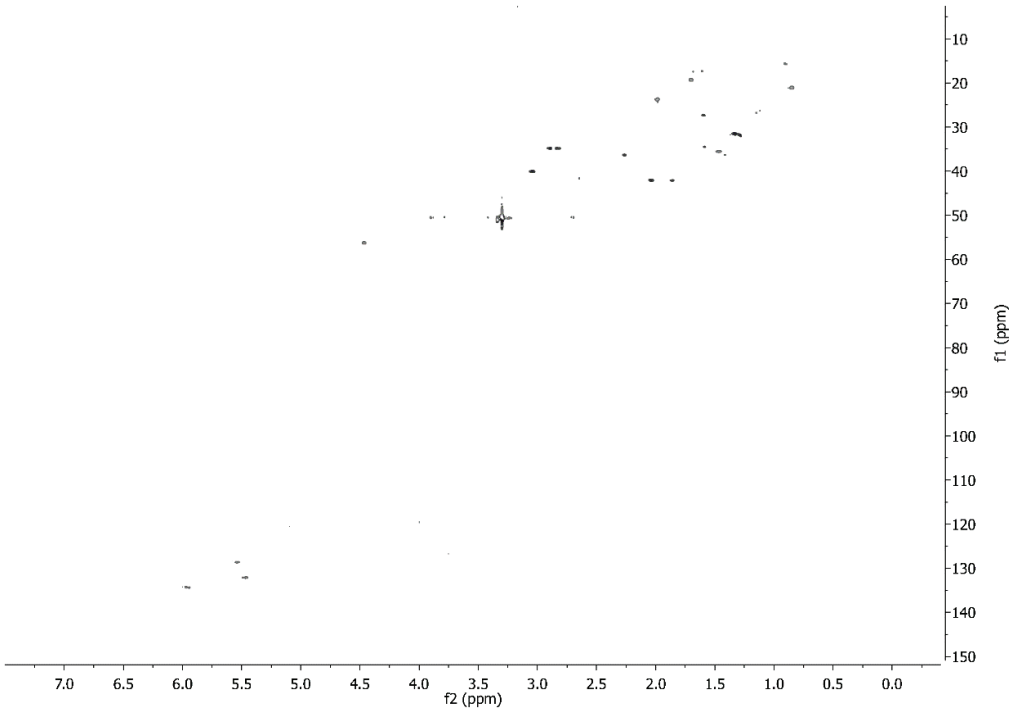


Figure S7. C) ^{1}H-^{13}C HSQC D) ^{1}H-^{13}C HMBC



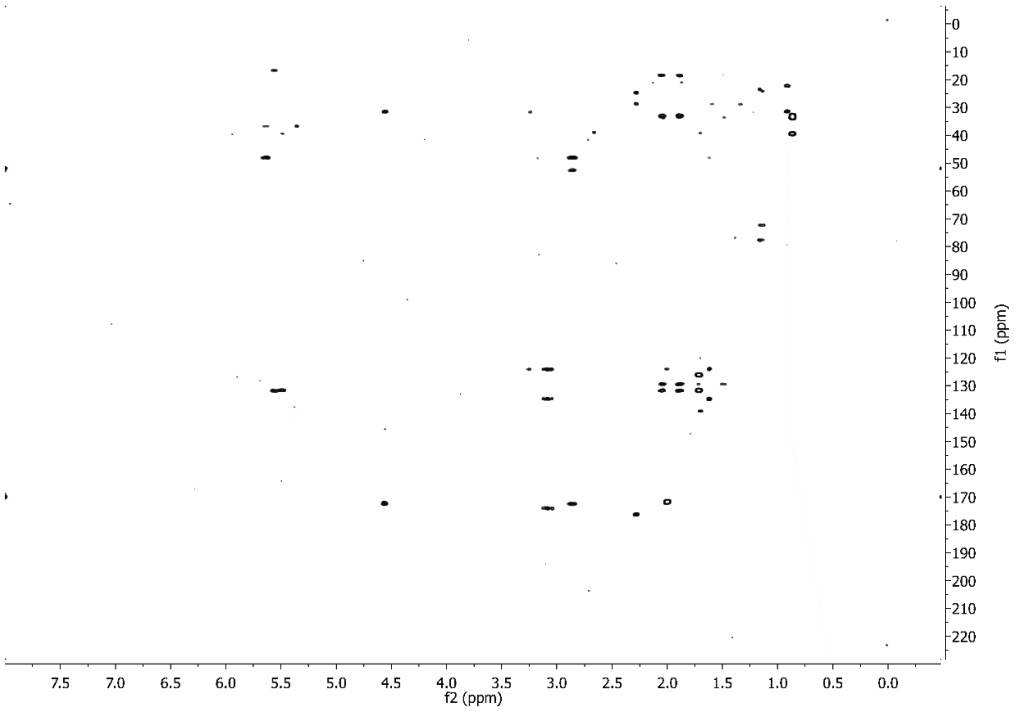
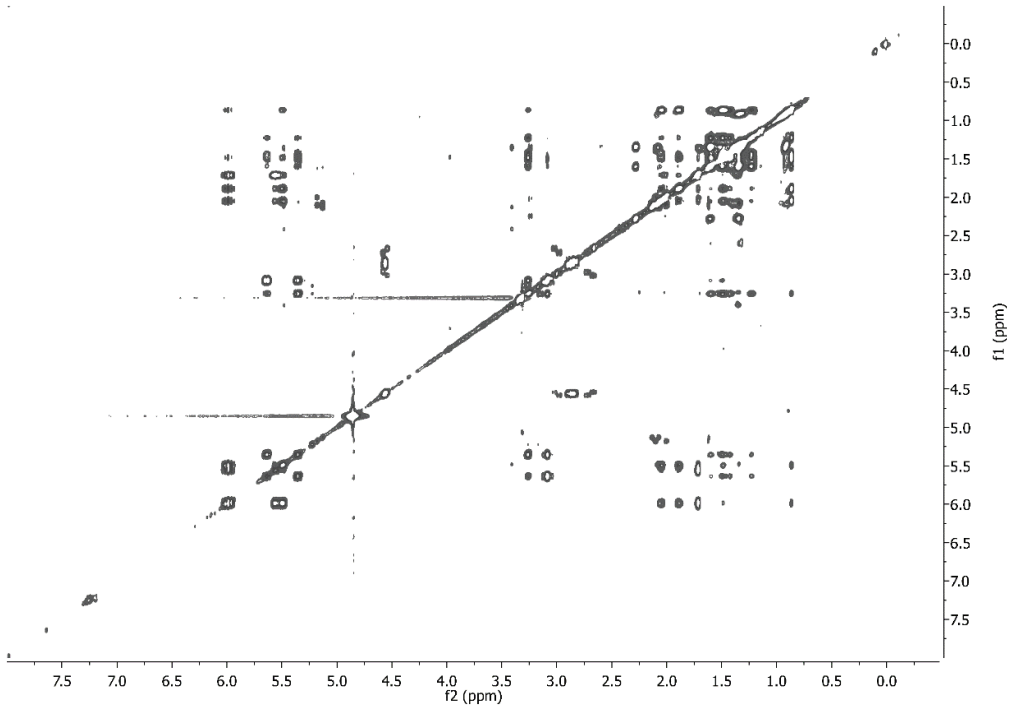


Figure S7. E) ^{1}H-^{1}H TOCSY F) ^{1}H-^{1}H ROESY



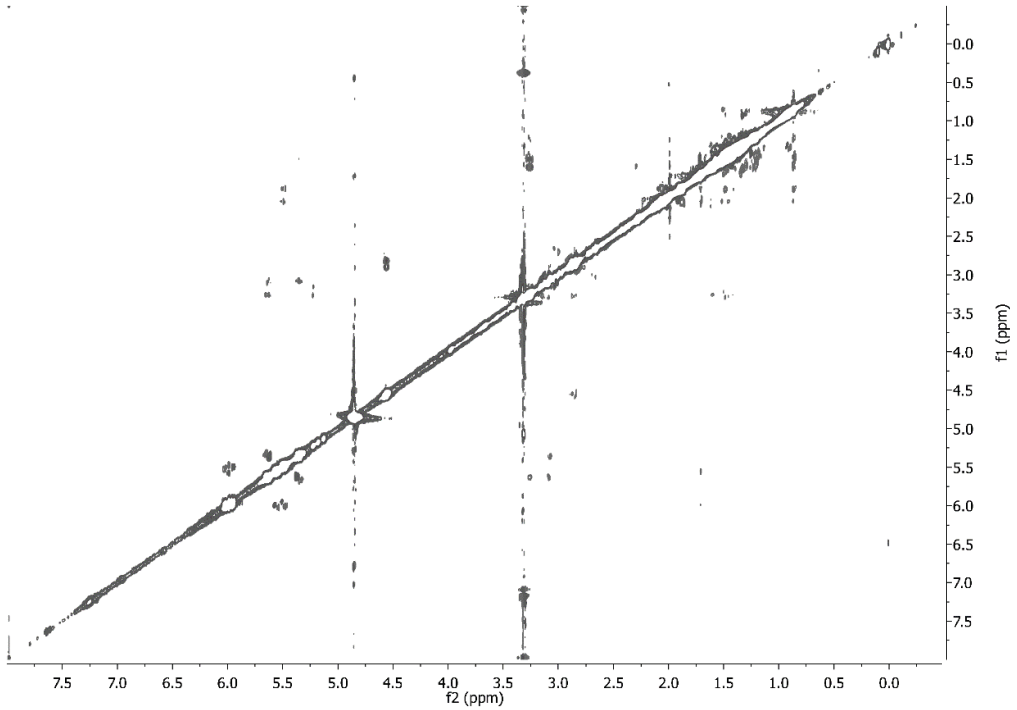


Figure S8. 1D and 2D NMR spectra for **6** in dmso-d_6. A) ^{1}H B) ^{1}H-^{1}H COSY

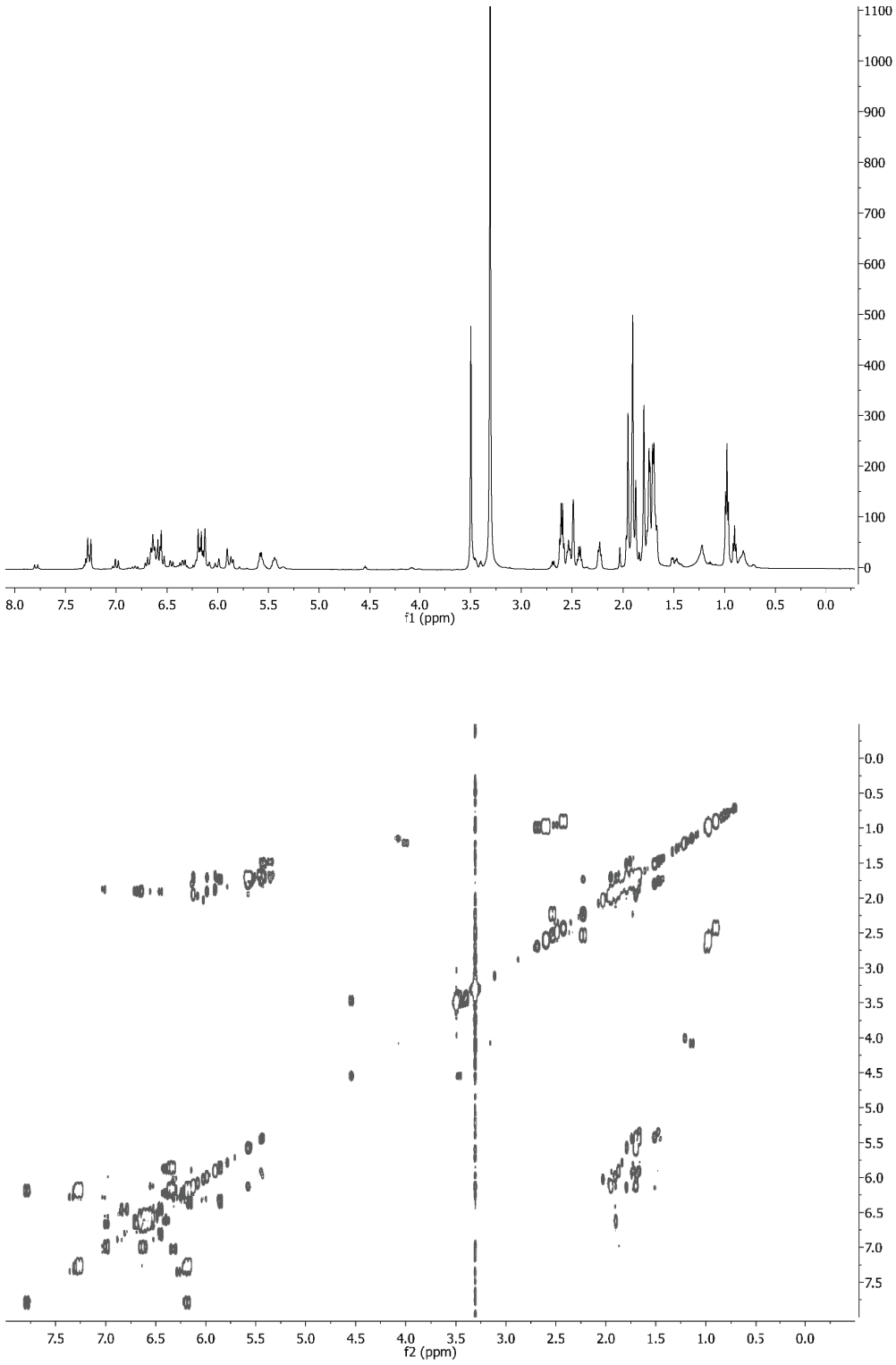
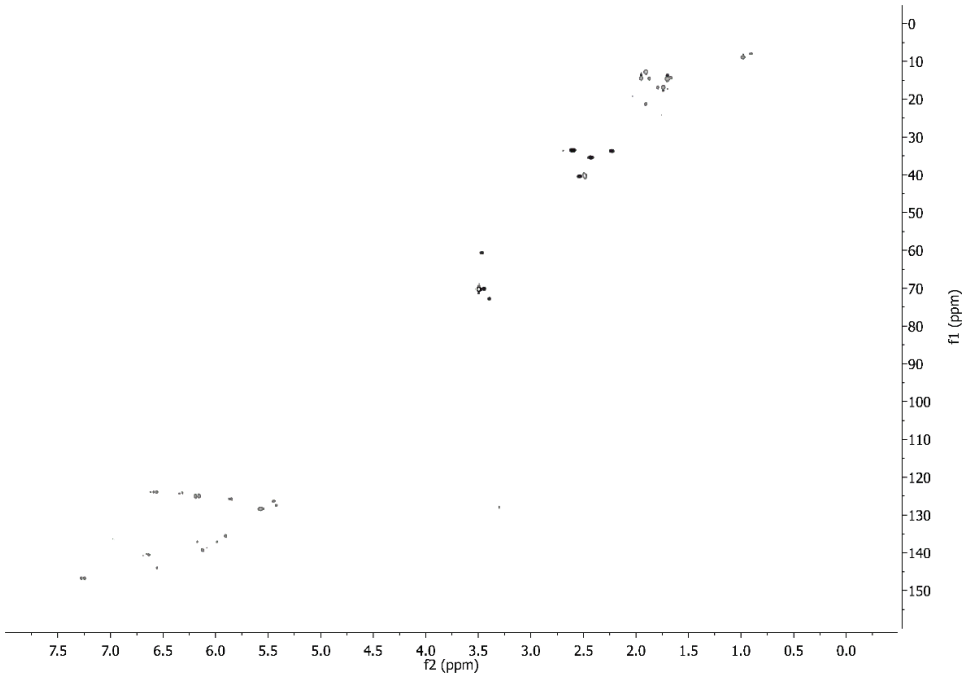


Figure S8. C) ^{1}H-^{13}C HSQC D) ^{1}H-^{13}C HMBC



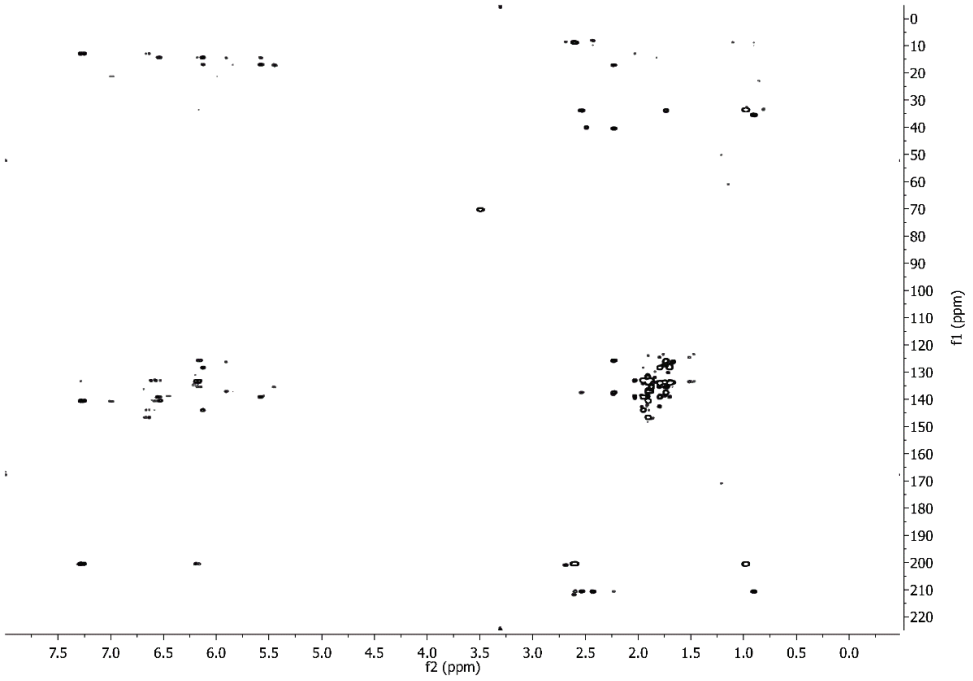


Figure S8. E) ^{1}H-^{1}H ROESY

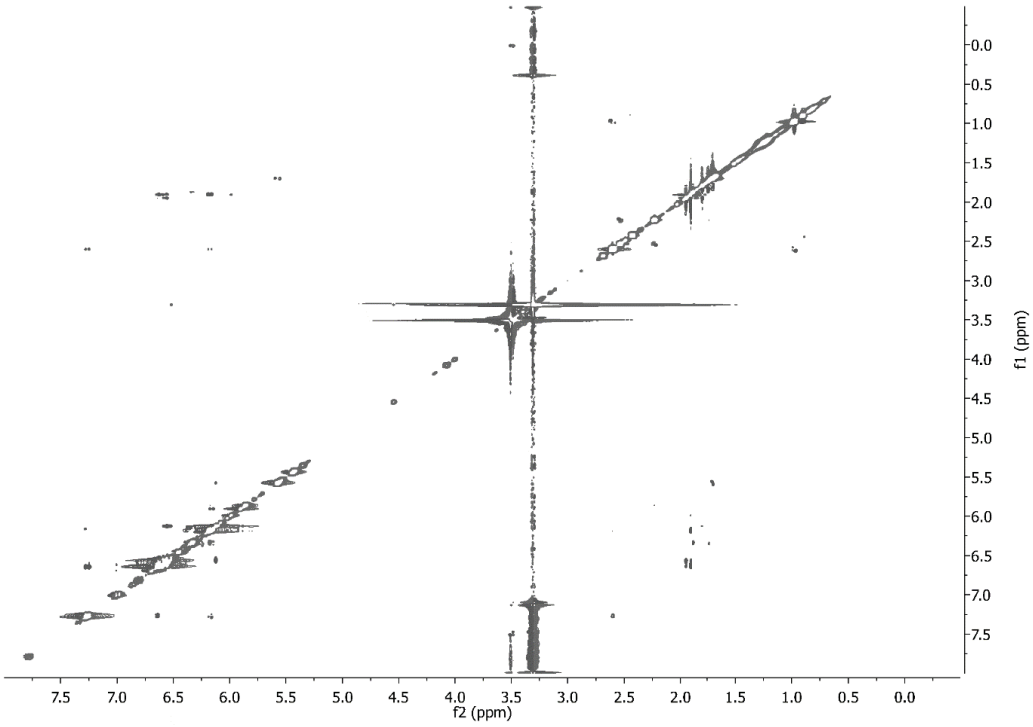


Figure S9. 1D and 2D NMR spectra for eqxTyr **7** in methanol-d_4. A) ^{1}H B) ^{1}H-^{1}H COSY

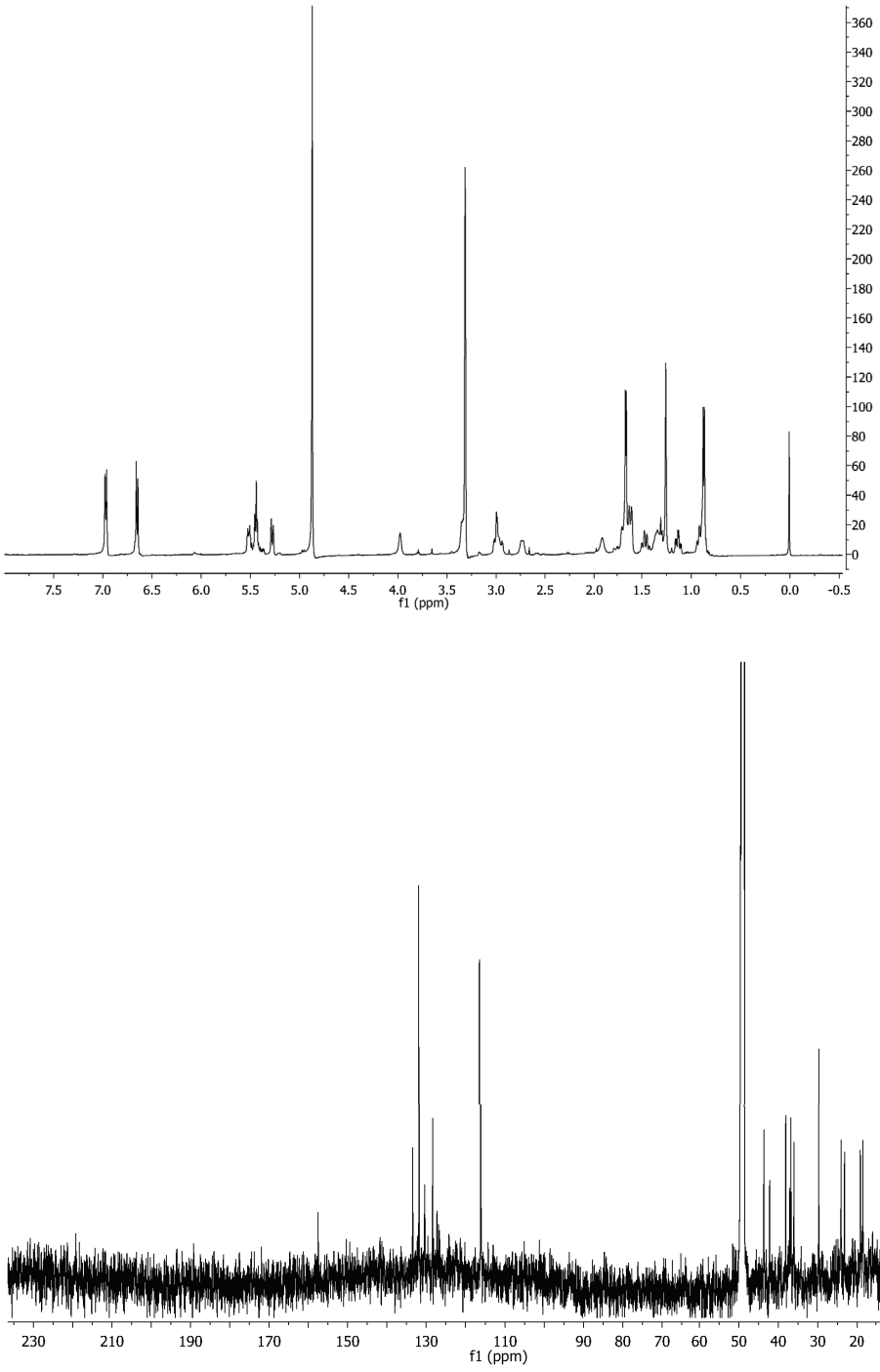
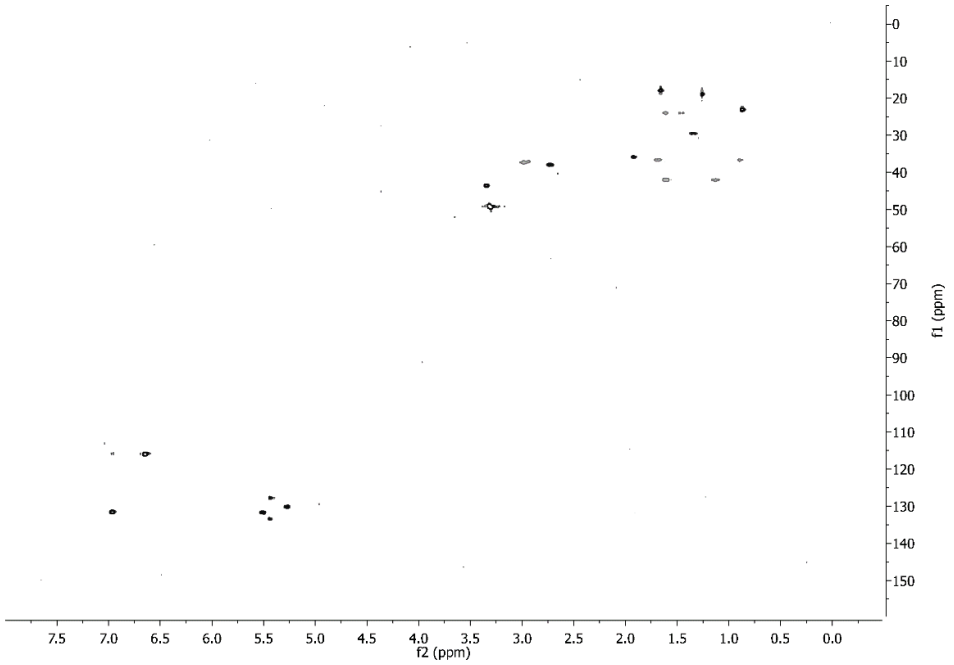


Figure S9. C) ^{1}H-^{13}C HSQC D) ^{1}H-^{13}C HMBC



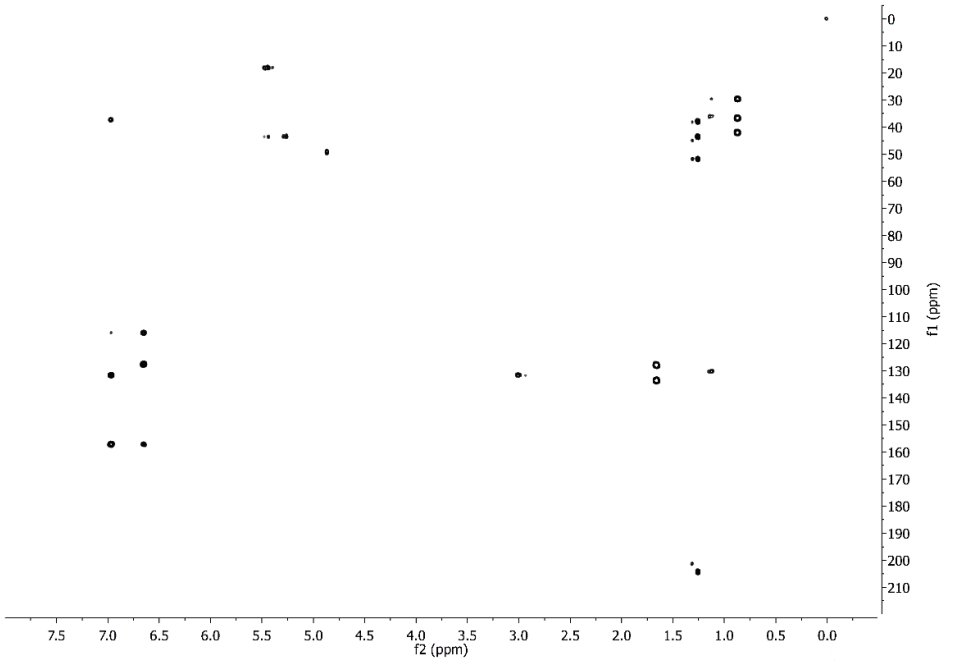
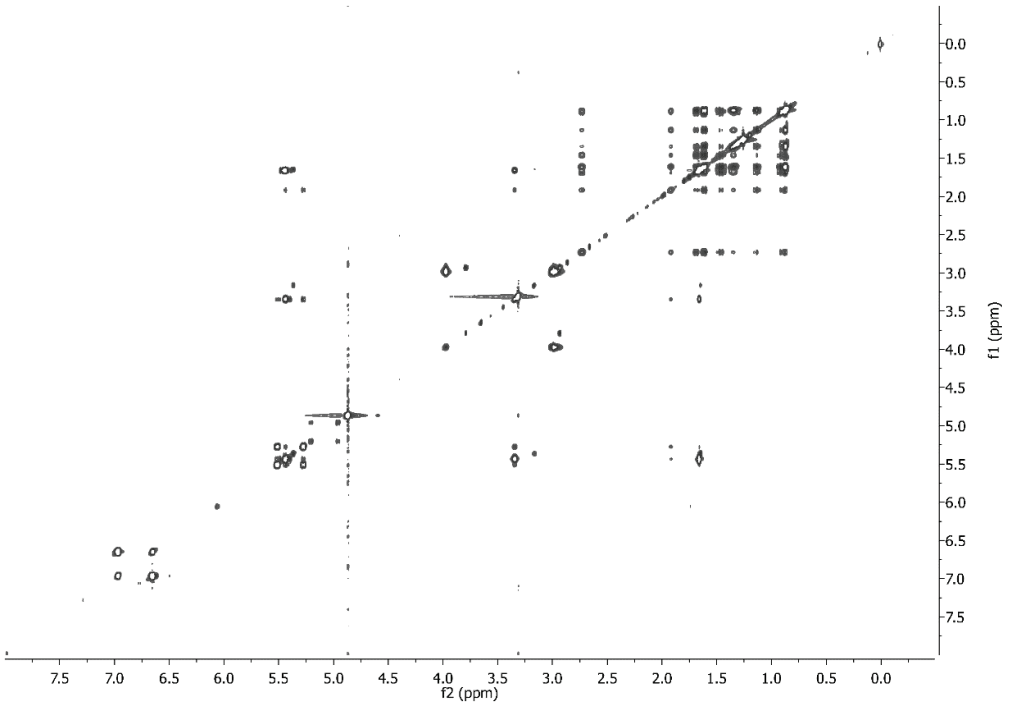


Figure S9. E) ^{1}H-^{1}H TOCSY F) ^{1}H-^{1}H ROESY



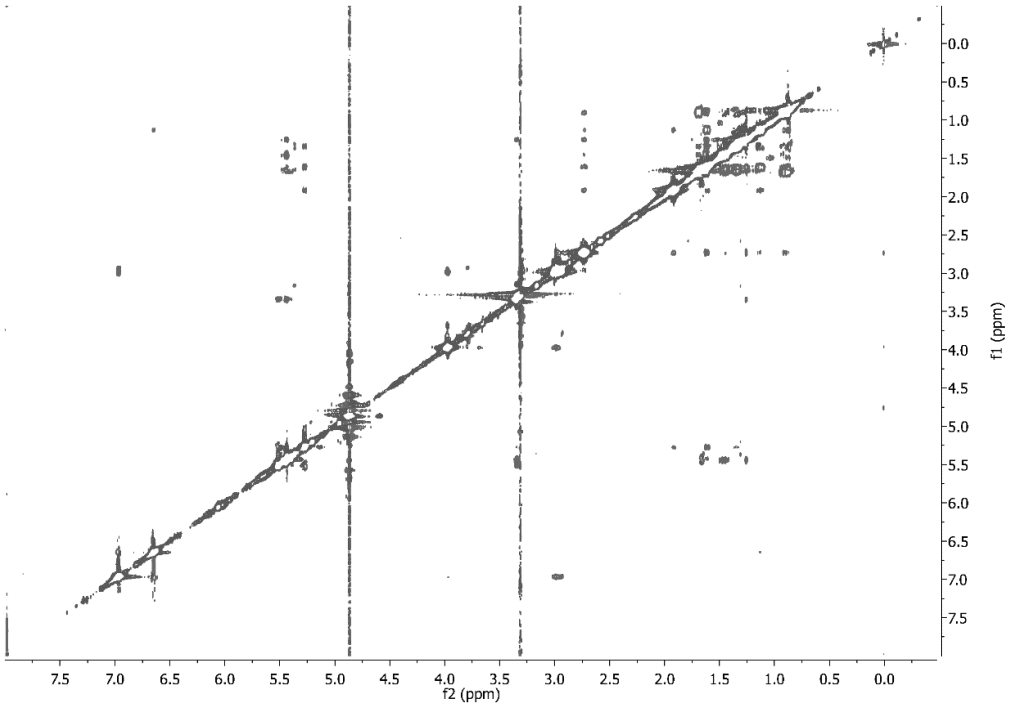


Figure S9. G) ^{1}H-^{1}H COSY

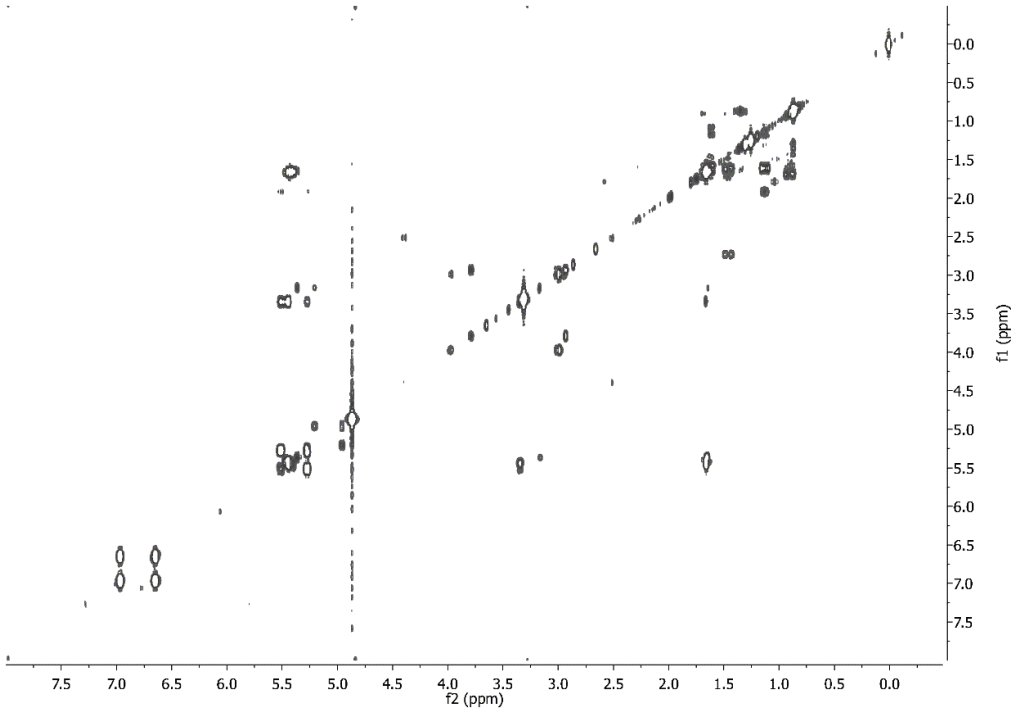
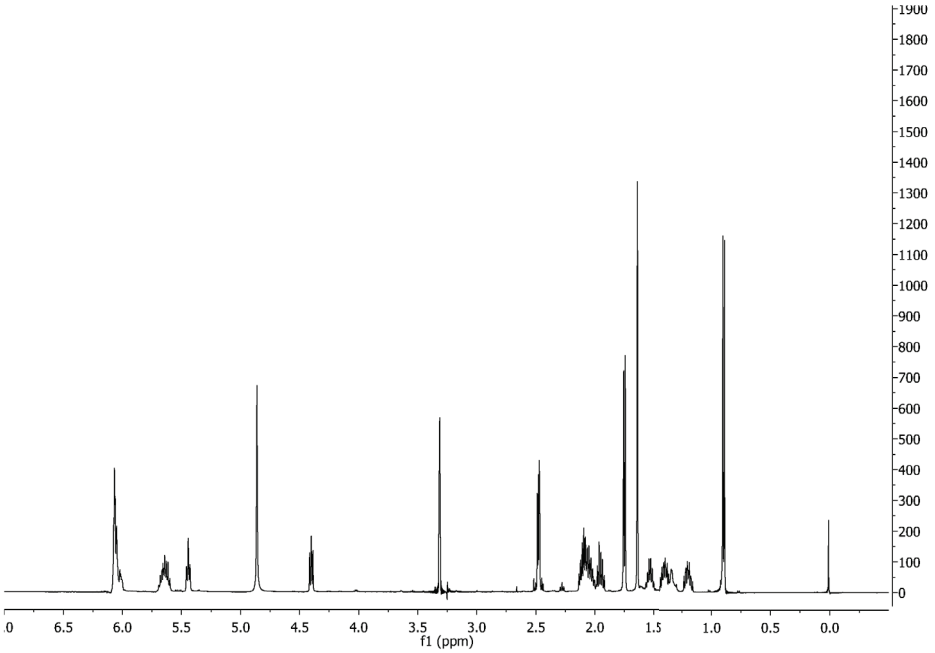


Figure S10. 1D and 2D NMR spectra for **8** in methanol-d_4. A) ^{1}H B) ^{13}C



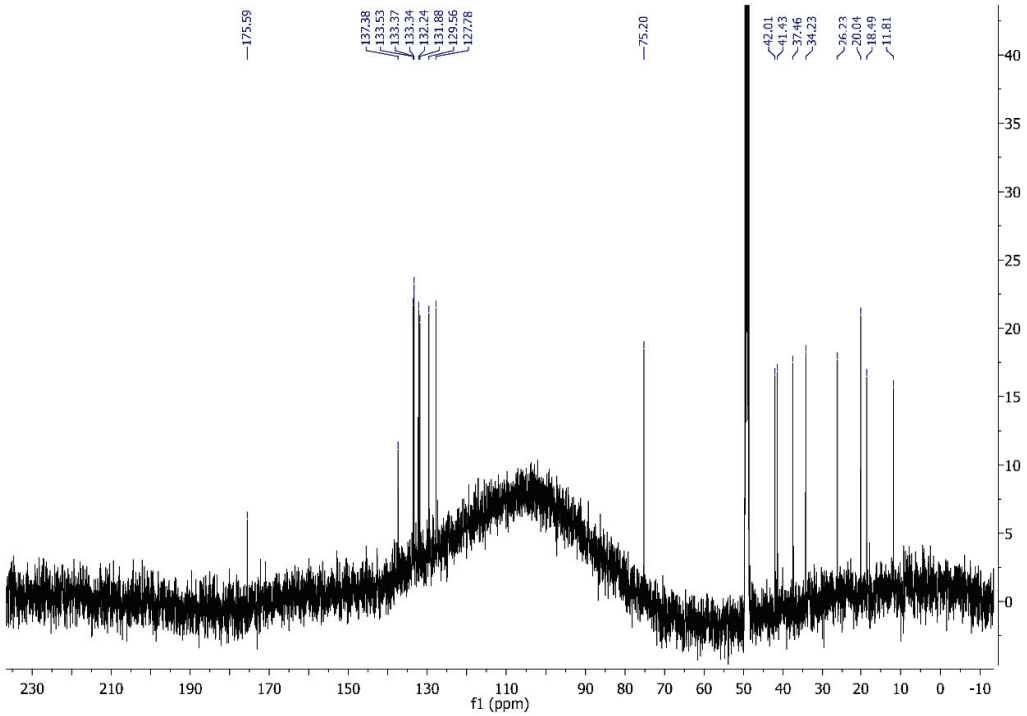
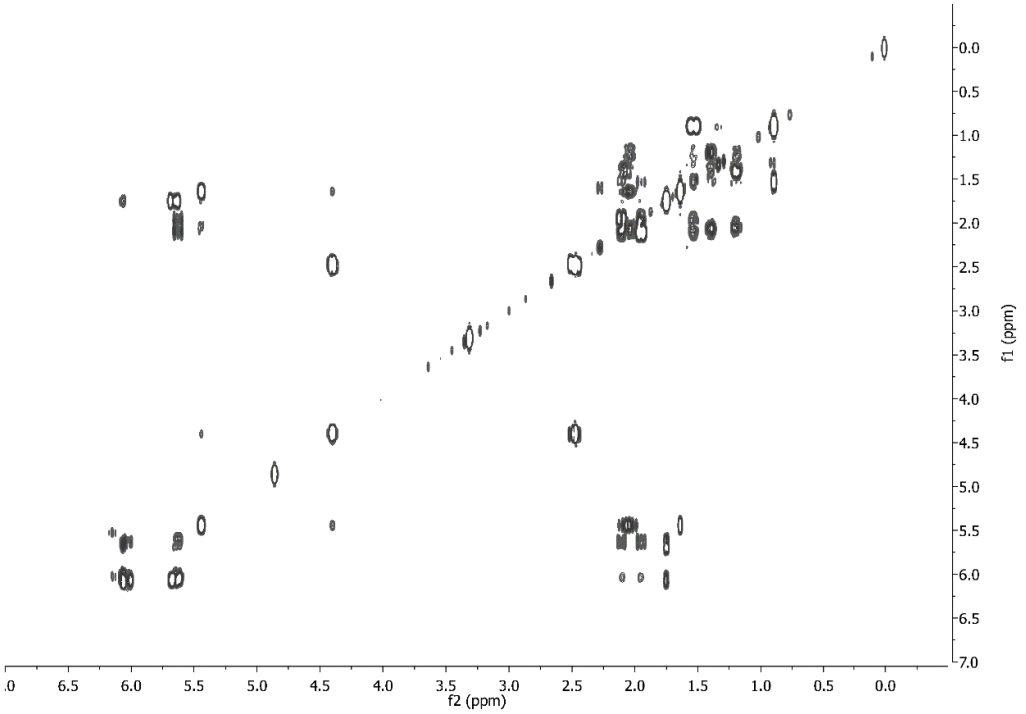


Figure S10. C) ^{1}H-^{1}H COSY D) ^{1}H-^{13}C HSQC



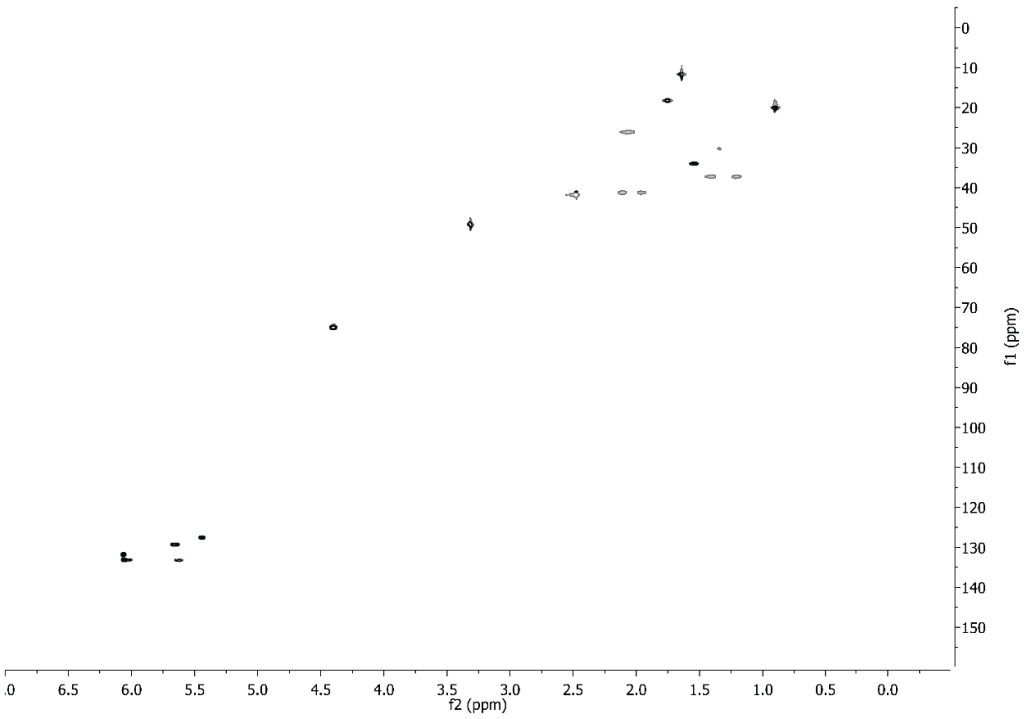
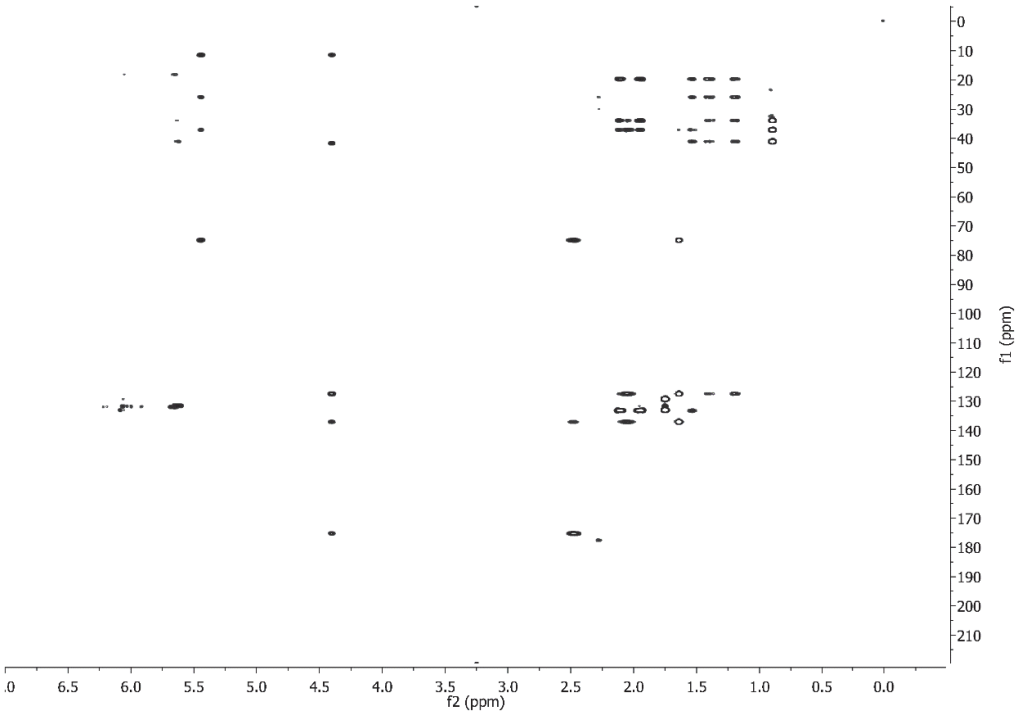
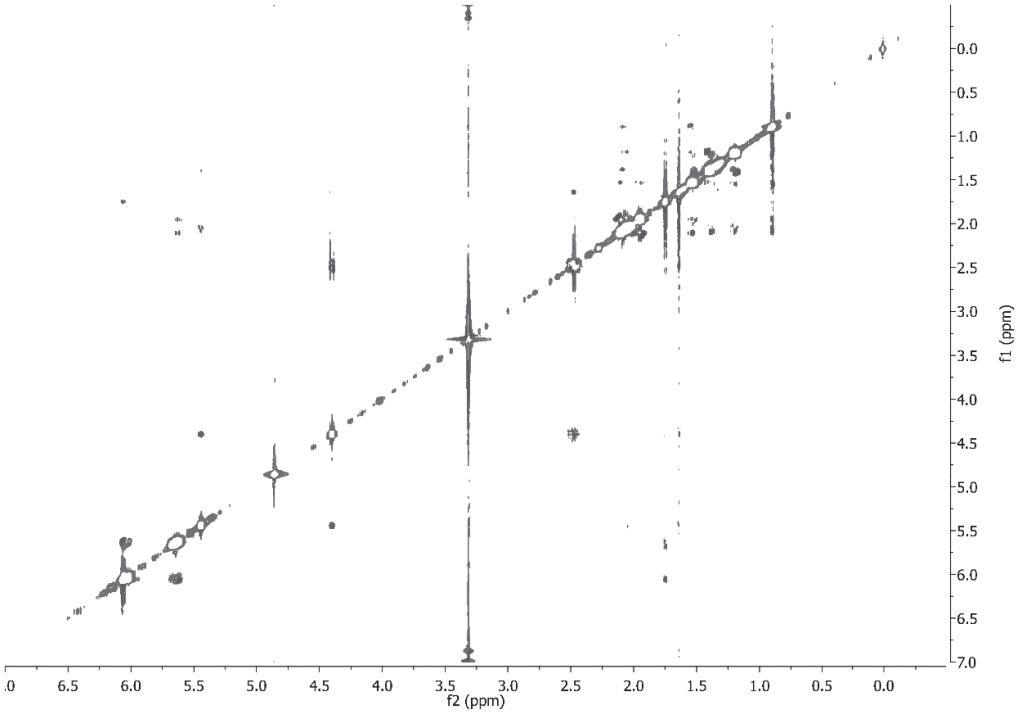


Figure S10. E) ^{1}H-^{13}C HMBC F) ROESY





REFERENCES

(1) Kakule, T. B.; Jadulco, R. C.; Koch, M.; Janso, J. E.; Barrows, L. R.; Schmidt, E. W. *ACS Synth. Biol.* **2014**. DOI: 10.1021/sb500296p. [Epub ahead of print]

(2) Kakule, T. B.; Sardar, D.; Lin, Z.; Schmidt, E. W. *ACS Chem. Biol.* **2013**, *8*, 1549.

CHAPTER 6

FUNGAL TRANSFORMATION

6.1 Introduction

Nontransferability of genetic techniques between fungi is the major hindrance to progress in genome-guided natural product discovery in fungi.[1,2] This has led to the development of heterologous expression systems such as the high-yielding platform described in Chapter 3, which has been applied to different biosynthetic questions in Chapters 3, 4, and 5. To enable the work thus far described, new transformation procedures for *Fusarium heterosporum* had to be developed. A previously reported protocol for this strain was found to be ineffective in my hands.[3] Based on ideas and troubleshooting from reported studies of other filamentous fungi,[2,4–7] a new protocol was crafted. Below is the detailed procedure for transformation of *Fusarium heterosporum* ATCC 74349 (FusWT) and the derivative expression strains FusΔeqx5 and FusΔeqx5ΔpyrG10, the uracil auxotroph.

6.2 Transformation Procedure

Spores are harvested from a 2-week old potato dextrose agar (PDA) culture of FusWT or FusΔeqx5 by adding sterile water (8 mL) to the plate and scraping the fungal growth surface with an inoculating loop. The spore suspension is then filtered through a cotton plug and added to potato dextrose broth (PDB, 100 mL) in a 500 mL Erlenmeyer flask. The spores are then germinated by incubation at 30 °C with shaking at 180 rpm for 8–10 hrs until the 10-cell mycelial stage. It is crucial to monitor growth every 30 mins after the 8-hr mark to ensure that the mycelia do not grow beyond this stage; otherwise protoplasting becomes inefficient. The germinated spores are then transferred to two 50 mL conical tubes and pelleted by centrifuging at 3220 x *g*. The pellets are resuspended in

Buffer A and combined. The germinated spores are then washed twice in Buffer A (50 mL) and then resuspended in protoplasting solution. Protoplasting is then done by incubating this suspension at 30 °C with shaking at 50 rpm for 1.5–3 hrs. It is important to monitor this process every 30 minutes after 1.5 hrs of cell wall digestion, and protoplasting is judged complete after 80% of the mycelia have been digested. The suspension is then filtered through MiraCloth (Calbiochem; sterilized by autoclaving), and the filtrate is centrifuged at 2880 x *g* at 4 °C for 10 mins. The pellet is washed thrice with ice-chilled SC buffer (15 mL), gently resuspended and centrifuged at 2880 x *g* for 5 mins. The protoplasts are resuspended in ice-chilled SPTC buffer (500 μL). The protoplast are then diluted with more ice-chilled SPTC to obtain 10^8 protoplasts per mL with the aid of a hemocytometer. Of importance, protoplasts should never be mixed by pipetting, but rather by gentle tapping on the tube.

To linearized DNA (10–15 μg, not exceeding 20 μL volume) in a 1.7 mL microcentrifuge tube is added protoplasts (100 μL). The mixture is suspended by gently tapping on the tube and then kept on ice for 30 mins. 40% PEG solution (1 volume) is added and after mixing, the transformation mix is incubated at RT for 50 mins. The protoplasts are washed by adding SC buffer (500 μL) at RT. The tube is inverted 8 times to mix and then centrifuged at 2000 x *g* for 5 mins. This wash step is repeated twice more with SC buffer (750 μL) at RT. The final washed pellet is resuspended in SC buffer (100 μL) and to this, regeneration broth (400 μL; 1 M sucrose, 0.02% yeast extract (Difco), filter-sterilized) is added and incubated for 17–24 hrs at 30 °C (static culture). Plating is done by pipetting the regenerated protoplasts into four petri dishes. Molten regeneration agar (1 M sucrose, 0.02% yeast extract (Difco), 1% agar (Difco), autoclaved) at ~42 °C

containing selective agents (hygromycin (Invitrogen, Cat no. 10687-010), 150 μg mL^{-1}; or phleomycin (InvivoGen, Cat no. ant-ph-1), 150 μg mL^{-1}) is poured over the protoplasts, which are then embedded in the agar by gentle swirling of the plates. The agar is allowed to set and cool (~10 mins), and then the plates are sealed with Parafilm and incubated at 30 °C. Transformants that appear after 3–5 days are transferred to PDA containing appropriate selection by picking mycelial agar plugs with pipette tips and placing them on the surface of the destination plate. After the second round of selection, each transformant is cultured to sporulation by streaking PDA with mycelia scraped from the edge of the colony.

Several modified protocols can also be used for different applications. The most useful strain for engineering is FusΔeqx5ΔpyrG10, in which the equisetin cluster has been deleted and uracil is required for growth. This strain lacks any background production of secondary metabolites. Additionally, I describe below a protocol to simultaneously introduce multiple vectors into FusΔeqx5ΔpyrG10. This protocol is necessary for pathway engineering and has so far been used to express up to four genes simultaneously.

The transformation protocol for the FusΔeqx5ΔpyrG10 strain is similar to that described for the FusWT and FusΔeqx5 strains, except that spores are germinated in PDB containing uracil (1.12 g L^{-1}) and uridine (140 mg L^{-1}, freshly prepared and filter-sterilized and added just before culture). Also for this strain, regeneration in broth culture before plating is not necessary. Immediately after washing the PEG out of the protoplasts, the protoplasts are plated by embedding in molten MM regeneration agar (1 M sucrose, 0.02% yeast extract without amino acids, 0.02% BSM supplement, 1% agar), and the plates are incubated at 30 °C. Transformants are passed through a second round of selection on Czapek-Dox agar, and sporulation can then be done by culturing the mutants on PDA.

Cotransformation of the FusΔeqx5ΔpyrG10 strain with two vectors containing *hph* and *pyrG* genes for selection has also been possible with a few modifications of the general procedure. 1) More protoplasts are added to the linearized DNA (150 μL of the 10^8 protoplasts/mL suspension is used). 2) After washing the PEG out of the transformation mix, the protoplasts are embedded in molten MM regeneration agar (15 mL) and incubated at 30 °C for 17 hrs. The agar is then overlayed with MM regeneration agar (15 mL) containing hygromycin 200 μg mL^{-1}, and the incubation continued at 30 °C. Transformants that appear in the top layer after 3–5 days are then transferred to Czapek-Dox agar containing hygromycin 150 μg mL^{-1} for a second round of selection. Sporulation of isolated mutants can then be attained by culturing on PDA.

Of note, 1) Unused protoplasts can be divided into 100 μL aliquots and stored at -80 °C for later use; viability is preserved for up to 2 months. 2) Protoplasting enzymes should not be thawed. Activity is reduced by repeated freeze/thaw cycles. 3) Phleomycin selection is poor with *Fusarium heterosporum*, but is greatly improved if the medium pH is adjusted to within 6.5–7.5. 4) Sporulation of primary cultures started from glycerol stocks can be inefficient. The cause of this is unknown, but we speculate it is due to the glycerol. Instead, inoculating PDA with spores from a 4-day old primary culture produces enough spores for transformations after 2 weeks of culture. 5) The activity of the protoplasting enzymes from the suppliers varies by batch. If protoplasting occurs too quickly, i.e., is complete in under 1.5 hrs, protoplast survival is low in the subsequent steps. It is advised to reduce the amount of Yatalase to 80 mg and β-glucuronidase to 2.5 mg. 6) During protoplasting, the target of 80% digestion is attained when only a fifth of the initial mycelial mass is left when observed by light microscopy. This has been found to be the

optimal level of digestion to obtain viable protoplasts. The viability of the protoplasts reduces when left in the enzyme solution beyond this point. On the other hand, less digestion results in lower protoplast yields after filtration through MiraCloth.

List of Buffers and their composition:

Buffer A (autoclaved)

1.2 M NaCl

114 mM Na_2HPO_4

19.7 mM citric acid

SC buffer (autoclaved)

1.2 M sorbitol

100 mM $CaCl_2$

40% PEG solution (autoclaved)

40% w/v PEG 3350

1.2 M sorbitol

50 mM $CaCl_2$

1% v/v 1 M Tris-Cl pH 8.0 solution

Protoplasting solution (prepared immediately before use; filter-sterilized)

Buffer A (20 mL)

Lysing Enzymes from *Trichoderma harzianum* 1 g (Sigma, Cat no. L1412)

Yatalase 100 mg (Takara Clontech, Cat no. T017)

β-glucuronidase 4 mg (Sigma, Cat no. G0751)

Hemicellulase 20 mg (Sigma, Cat no. H2125)

SPTC buffer

SC buffer (8 mL)

40% PEG solution (2 mL)

DMSO (100 μL)

6.3 References

1. Brakhage, A. A.; Schroeckh, V. Fungal Secondary Metabolites – Strategies to Activate Silent Gene Clusters. *Fungal Genet. Biol.* **2011**, *48*, 15–22.

2. Ruiz-Díez, B. Strategies for the Transformation of Filamentous Fungi. *J. Appl. Microbiol.* **2002**, *92*, 189–195.

3. Sims, J. W.; Fillmore, J. P.; Warner, D. D.; Schmidt, E. W. Equisetin Biosynthesis in *Fusarium heterosporum*. *Chem Commun (Camb)* **2005**, DOI: 10.1039/b413523g.

4. Lopez-Belmonte, F.; García, A. I.; Villanueva, J. R. Observations on the Protoplasts of *Fusarium Culmorum* and on their Fusion. *J. Gen. Microbiol.* **1966**, *45*, 127–134.

5. Proctor, R. H.; Desjardins, A. E.; Plattner, R. D.; Hohn, T. M. A Polyketide Synthase Gene Required for Biosynthesis of Fumonisin Mycotoxins in *Gibberella fujikuroi* Mating Population A. *Fungal Genet. Biol.* **1999**, *27*, 100–112.

6. Skory, C. D.; Horng, J. S.; Pestka, J. J.; Linz, J. E. Transformation of *Aspergillus parasiticus* with a Homologous Gene (pyrG) Involved in Pyrimidine Biosynthesis. *Appl. Environ. Microbiol.* **1990**, *56*, 3315–3320.

7. Vazquez, F.; de Figueroa, L. C. Protoplasts Formation in *Fusarium* species. *Biotechnol. Tech.* **1996**, *10*, 93–98.

CHAPTER 7

CONCLUSIONS

7.1 Conclusions

Limitations to using the fungal polyketide synthases as medicinal chemists in the generation of analogues arise from incomplete understanding of how these enzymes function. Great strides have been made over the years to understand the functioning of bacterial-type PKSs, enabling their application to synthesize new derivatives.[1] The sheer complexity of the fungal PKS has made such undertakings quite risky for this class of natural products and thus the slow progress observed.[2] Rather than a brute force approach to combinatorial biosynthesis as has been attempted by others,[2] the way to mitigate the risks is by taking a synthetic biology perspective,[3] where PKS function is deconstructed into primary components (critical) and secondary components (noncritical) for the task desired. For instance, the work described in Chapter 5 involved the categorization of ACP/C domain interactions and C domain selectivity as primary and other domain activities as secondary components. This reduced the complexity enough to enable the disentanglement of C domain selectivity from PKS/NRPS interactions in the study design. While the results from this work show that analogue synthesis by combinatorial PKS/NRPS pairings is more difficult than anticipated, it establishes ground rules for these attempts and opens up new avenues for inquiry. For example, the characterization of more PKS-NRPSs and linking them to products becomes more appealing. With more members identified, a phylogenetic-guided approach to combinatorial biosynthesis becomes possible. Another possibility is leaving the ACP and C domain intact and modifying the components upstream or downstream. For example, upstream of the ACP, the KR domain was manipulated, but it did not produce the desired outcome. A shift of focus downstream to the A domain, altering its amino acid specificity via a directed

evolution approach, could allow the installation of various amino acids to a fixed polyketide chain to generate tetramic acid analogues. The identified specificity determinants within A domains could then be engineered into already characterized PKS-NRPSs to modify their native products. Beyond combinatorial biosynthesis, the results from the PKS-NRPS study could be applied from a biotechnology perspective to supply crucial intermediates for chemical synthesis. The observation that complete PKS modules fused to highly-selective C domains leads to synthesis of predicted polyketide products native to that PKS allows the decoupling of these polyketides from amino acids in a predictable manner such that the isolated intermediates can be repurposed. Using the model expression platform described in Chapter 3, these intermediates could be supplied in yields high enough for use as starting materials for chemical synthesis protocols. In line with this reasoning, this expression platform allows a semisynthetic approach to doing medicinal chemistry on promising compounds for drug development. For example, the pyrrolocins described in Chapters 3 and 4 were produced in excess of 800 mg kg^{-1} such that chemical modifications can now be pursued to determine structure activity relationships starting from the core structure isolated from the recombinant strain.

In this work, all the research questions posed were answered. More exciting, however, was that more questions arose from observations made on this long arduous journey. One example is the mysterious heterologous production of a mixture of pyrrolocin stereoisomers about the decalin ring, whereas only the *trans*-isomer was produced by the native host. This result raised the possibility of the involvement of a third protein that catalyzes the Diels-Alder reaction, and further suggests that other PKS-NRPS pathways to decalin-containing tetramic acids possess this component. Another is

the observation that Dieckmann reaction capabilities of the R domain are compromised with the attachment of a gfp tag. One hypothesis is that normally, water is excluded from the R domain active site but the tag modifies the structure sufficiently to promote a water-mediated hydrolysis catalyzed by the R domain. Further investigations into this phenomenon guided by R domain structural data will illuminate the sequence features that led to the divergence of these types of R domains from those that actually perform NAD(P)H-reductive release of their substrates. A third unexpected finding was the methylation of pyrrolocin by an equisetin pathway N-methyl transferase (MT). The structure of the native substrate for this MT is quite different from the pyrrolocin structure allowing the possibility that this MT could be repurposed for N-methylation of other small molecules. Structural studies will be required to delineate the active site such that a broad substrate mutant can be designed. Finally, the most intriguing question is why C domains select for a set of closely related substrates. A definitive answer will be difficult to come to, absent a crystal structure of the C domain. Currently, common length of substrates within error of one ketide unit appears to be the most consistent feature that can be attributed to C domain selectivity. This suggests that the size of substrate-binding pocket of the C domain is finely tuned to position the substrate for conjugation to the amino acid. Also the size of the substrate could be critical to exclude water from the active site, akin to the observation for the gfp-tagged R domain. However, based upon the observation that competition for substrates between the KR and C domain exists, it is possible that KR and C domains coevolved to maintain a balanced affinity for intermediates during polyketide elongation. This idea is supported by phylogenetic analyses of PKSs and NRPSs from these hybrid enzymes that show that the family trees

obtained are similar.[2] Beyond further work in combinatorializing PKS-NRPSs, future studies directed towards all the above mentioned observations will push the field forward.

In closing, fungi still remain a promising source of natural products that will impact human health, and elucidation of biosynthetic mechanisms should continue to be a priority to augment the structural diversity found in nature.

7.2 References

1. Hansen, D. A.; Rath, C. M.; Eisman, E. B.; Narayan, A. R. H.; Kittendorf, J. D.; Mortison, J. D.; Yoon, Y. J.; Sherman, D. H. Biocatalytic Synthesis of Pikromycin, Methymycin, Neomethymycin, Novamethymycin, and Ketomethymycin. *J. Am. Chem. Soc.* **2013**, *135*, 11232–11238.

2. Boettger, D.; Bergmann, H.; Kuehn, B.; Shelest, E.; Hertweck, C. Evolutionary Imprint of Catalytic Domains in Fungal PKS–NRPS Hybrids. *ChemBioChem* **2012**, *13*, 2363–2373.

3. Way, J. C.; Collins, J. J.; Keasling, J. D.; Silver, P. A. Integrating Biological Redesign: Where Synthetic Biology Came from and Where it Needs to Go. *Cell* **2014**, *157*, 151–161.