Title
Genetic and Biochemical Characterization of Natural Tetracycline Biosynthesis

Author
Wang, Peng

Publication Date
2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Genetic and Biochemical Characterization of Natural Tetracycline Biosynthesis

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Chemical Engineering

By

Peng Wang

2013

ABSTRACT OF THE DISSERTATION

Genetic and Biochemical Characterization of Natural Tetracycline Biosynthesis

By

Peng Wang

Doctor of Philosophy in Chemical Engineering

University of California, Los Angeles, 2013

Professor Yi Tang, Chair

Tetracyclines are a group of natural products produced by soil-borne *Actinobacteria*. Their broad-spectrum biological activities such as antibiotic, anticancer, and novel activity against tetracycline resistant bacteria are attributed to their signature linearly fused four ring structure. However, the extensive use of tetracyclines during the last sixty years has led to emergence of resistance mechanisms among microorganism communities, resulting in dramatically decreased effectiveness of tetracyclines as first line antibiotic agents. Therefore, new generation of tetracyclines is highly demanded to overcome current resistance mechanisms. To obtain new generation of tetracyclines, fundamental understanding of tetracycline biosynthesis and establishment of robust manipulation platform are required. The biosynthetic pathways of three natural tetracycline, including oxytetracycline, SF2575, and dactylocycline, were investigated by using genetic, biochemical, and protein structure based analysis.

Oxytetracycline represents an important example of natural tetracyclines featuring a signature C5 hydroxyl group. Unveiling of oxytetracycline biosynthetic pathway will establish a cornerstone to understand and engineer tetracycline biosynthesis. After a decade investigations, the biosynthetic pathway of a key intermediate anhydrotetracycline has been revealed; however, the longstanding missing link involved in the final transformations from anhydrotetracycline to oxytetracycline remains elusive since the first discovery of oxytetracycline in 1950. Two redox enzymes OxyS and OxyR were unravelled to catalyze the mysterious final transformations by using flavin adenine dinucleotide (FAD) and F_{420} as coenzymes respectively. The protein structure of OxyS has been determined and provided valuable insights into the enzymology unique to tetracycline biosynthesis.

SF2575 stands out from tetracycline family due to its fully substituted tetracycline aglycone and novel potent anticancer activity. Cascade transformations are catalyzed by a set of new tetracycline tailoring enzymes involved in SF2575 biosynthesis, including four methyltransferases, a glycosyltransferase, and three redox enzymes. To decipher the biosynthesis and understand the chemical reactions catalyzed by these dedicated tailoring enzymes, the corresponding encoding genes were inactivated in heterologous expression host *Streptomyces lividans* K4-114. The functions of the tailoring enzymes were elucidated by isolation and structural characterization of intermediates accumulated from the corresponding gene-inactivation mutants. The production of fourteen tetracycline analogs from ten mutants led to functional assignment of SF2575 tailoring enzymes and elucidation of SF2575 biosynthetic pathway. Most interestingly the redundancy of the methyltransferase SsfM3 demonstrates an evolutionary event in which combinatorial biosynthesis strategy has been employed by nature to generate novel tetracycline compounds.

Dactylocycline offers us the third example of natural tetracycline with novel activity against tetracycline resistant bacteria. This promising activity is due to a unique hydroxylamino sugar modification at C6 hydroxyl group of dactylocyclinone which shows cross resistance with tetracycline. Identification of dactylocycline biosynthetic gene cluster led to a proposed biosynthetic pathway and expanded enzymatic tools to synthesize both dactylocycline aglycone and the hydroxylamino sugar moiety. To validate this gene cluster, the *dac* gene cluster was heterologously expressed in *Streptomyces lividans* K4-114 resulting in the production of dactylocyclinone.

Given three natural tetracycline biosynthetic pathways and an accommodating heterologous host, we are able to generate new tetracycline analogs by using combinatorial biosynthesis approaches.

The dissertation of Peng Wang is approved.

Neil Garg

James Liao

Tatiana Segura

Yi Tang, Committee Chair

University of California, Los Angeles

2013

TABLE OF CONTENTS

1 Introduction……..1

1.1 Polyketide biosynthesis……1

1.2 Redox tailoring enzymes …….....13

1.2.1 Pre-assembly modification ……13

1.2.2 Post-assembly modification……... 15

1.3 Tetracycline biosynthesis ……19

2 Results and Discussion ……22

2.1 Enzymatic basis of oxytetracycline final transformations……... 22

2.1.1 Introduction……... 22

2.1.2 Results and Discussion…….. 24

2.1.3 Materials and Methods…….. 39

2.2 Heterologous expression and manipulation of tetracycline biosynthesis 44

2.2.1 Heterologous expression of oxytetracycline biosynthesis …… 44

2.2.2 Identification of SF2575 tailoring enzymes ……..47

2.2.2.1 Introduction ……...47

2.2.2.2 Results ……...50

2.2.2.3 Discussion ……68

2.2.2.4 Materials and Methods ……..75

2.3 Cloning and identification of dactylocycline biosynthetic gene cluster 78

2.3.1 Introduction……78

2.3.2 Results and Discussion…….. 80

2.3.2.1 Genes involved in backbone biosynthesis …… 80

2.3.2.2 Genes involved in post modification steps…… 81

2.3.2.3 Genes involved in sugar biosynthesis ……... 82

2.3.2.4 Genes involved in regulation and others…… 83

2.3.2.5 Heterologous expression of dactylocyclinone…….... 84

3 Conclusions…….. 86

4 Appendices……... 88

5 References……..... 151

LIST OF FIGURES

Figure 1. Chemical structures of four natural polyketide drugs. 3

Figure 2. General mechanism involved in polyketide biosynthesis..... 4

Figure 3. Biosynthesis of 6-dEB and erythromycin..... 6

Figure 4. Biosynthesis of landomycin..... 7

Figure 5. Biosynthesis of flavonoids, stilbenoids, and curcuminoids..... 9

Figure 6. Three layers of secondary metabolite production regulations..... 12

Figure 7. Complexity generation during natural product biosynthesis can take place at three different stages. 13

Figure 8. Representative pathways to generate natural complexity in polyketide biosynthesis...18

Figure 9. Final transformations from ATC to OTC remain elusive..... 20

Figure 10. Established oxytetracycline biosynthetic pathway in previous studies. 21

Figure 11. Redox enzymes OxyS and OxyR catalyze the final transformations involved in oxytetracycline biosynthesis. 22

Figure 12. The organizations of the tetracycline biosynthetic gene clusters. Two essential genes *oxyS* and *oxyR* and the homologs are highlighted in red. 24

Figure 13. In vivo experiments showed that OxyS and OxyR are able to convert **11** to **8** in *S. lividans*. 25

Figure 14. HPLC analysis of OxyR cofactor. 26

Figure 15. Analysis of OxyR and OxyS functions..... 29

Figure 16. The proposed mechanisms of the transformations of **11** to **8** and **9**. 30

Figure 17. Structure-based sequence alignment of OxyS and other homologs. 35

Figure 18. Overall structural alignment of OxyS with RdmE. 36

Figure 19. Reaction catalyzed by RdmE with Akalavinone **15**. (i) alternative view of **15** modeled in the active site of OxyS and (ii) Modeled view of **11** in the active site of OxyS based on the position of **15**. 37

Figure 20. OxyS-only in vitro assay quenched and analyzed after ~1 min. 38

Figure 21. SDS-PAGE analysis of the recombinant proteins in this study..... 38

Figure 22. (i) UV and mass analysis of the new compound produced by OxyS His47Ala mutant assay and (ii) of **9**. 39

Figure 23. Three tetracycline biosynthetic pathways were overexpressed and manipulated in the heterologous host *S. lividans* K4-114. 44

Figure 24. Scheme representation of heterologous expression of *oxy* biosynthetic pathway in *S. lividans* K4 46

Figure 25. HPLC analysis of *oxy* heterologous expression in *S. lividans* K4. 46

Figure 26. Heterologous expression of *ssf* gene cluster in *S. lividans* K4. 51

Figure 27. SsfO1 and OxyS in vitro assay with substrate **33**..... 59

Figure 28. Overall structure of SsfO1 shows three domains including FAD binding domain, Middle domain, and C-terminal thioredoxin-like domain. 62

Figure 29. Structural alignment of SsfO1 with OxyS. 62

Figure 30. HPLC profiles of extracts from cultures of K4-114 integrated with different knockouts of the *ssf* pathways..... 63

Figure 31. Biosynthetic pathways of A) oxytetracycline; B) SF2575. 64

Figure 32. HPLC analyses of *ssf* gene complementary mutants. 67

Figure 33. Sequence alignment of SsfM3 with five different methyltransferases. 74

Figure 34. HPLC analysis of *dac* heterologous expression in *S. lividans* K4 84

Figure 35. Proposed biosynthetic pathway of dactylocycline A. 85

Figure 36. Proposed biosynthetic pathway of hydroxylaminosugar in dactlyocycline biosynthesis. 85

LIST OF TABLES

Table 1. Data collection and refinement statistics of OxyS 34

Table 2. Data collection and refinement statistics of SsfO1 61

Table 3. Minimum inhibitory concentration (MIC) values of compounds in this study 66

Table 4. Putative functions of ORFs in the *dac* biosynthetic gene cluster 79

ACKNOWLEDGMENTS

Section 1.1 contains material which is reprinted (adapted) with permission from "Gao, X., Wang, P., Tang, Y.,"Engineered polyketides biosynthesis and biocatalysis in Escherichia coli." *Appl. Microbiol. Biotechnol*. **2010**, 88, 1233-1242.". Copyright The Author(s) 2010. This article is published with open access at Springerlink.com

Section 1.2 contains material which is reprinted (adapted) with permission from "Wang, P., Gao, X., Tang, Y."Complexity Generation during Natural Product Biosynthesis using Redox Enzymes." *Curr. Opin. Chem. Biol*. **2012**, 16, 362-369.". Copyright (2012) Elsevier Ltd.

Section 2.1 contains material which is reprinted (adapted) with permission from "Wang, P., Bashiri, G., Gao, X., Sawaya, M. R. Tang, Y. "Uncovering the Enzymes that Catalyze the Final Steps in Oxytetracycline Biosynthesis." *J. Am. Chem. Soc.* **2013**, 135 (19), 7138–7141. Copyright (2013) American Chemical Society.

Section 2.2 and **2.3** contain material which is reprinted (adapted) with permission from "Wang, P., Kim. W., Pickens, L. B., Gao, X., Tang, Y. "Heterologous Expression and Manipulation of Three Tetracycline Biosynthetic Pathways." *Angew. Intl. Chem. Ed*. **2012**, 51, 11136–11140.". Copyright (2012) WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

The work in this dissertation was funded by a NSF CBET 1159759 and a David and Lucile Packard Foundation Fellowship to Y. T.

In addition, I would like to deeply thank my supervisor Professor Yi Tang who has supported my Ph. D. program in the Department of Chemical and Biomolecular Engineering at UCLA by providing academic and professional guidance. I am grateful to Professor Tang for his insightful advices and continual encouragement to keep my projects moving forward. I am thankful to my committee members Professor Neil Garg, Professor James Liao, and Professor Tatiana Segura for their wonderful supports of my Ph. D. candidacy. I appreciated and enjoyed their insightful discussions on my projects and supports of my academic career.

I would like to thank Dr. Michael R. Sawaya and Dr. Ghader Bashiri for their contributions to my research: I would not complete my research project without their intellectual and technical supports. I am grateful to Dr. Sawaya and Dr. Duilio Cascio for their advices on protein crystallization of OxyS and SsfO1. To our *Streptomyces* subgroup, I would like to thank Dr. Lauren B. Pickens and Dr. Woncheol Kim for their tremendous contributions in tetracycline projects. I am thankful to Dr. Wenjun Zhang and Dr. Jixun Zhan for the training when I joined Professor Tang's laboratory. Daniel Xue is thanked for his assistance as an undergraduate student. I am grateful to all the former and current members in Professor Tang's laboratory who foster an exciting academic environment and turn a laboratory into a big family.

Pursuing a Ph. D. has been a long and unforgettable experience which is the most exciting and sometimes the most "boring" journey. However, non-academic activities and events have made my daily research life much more enjoyable. Common interests in sports have established another layer of collaborations between Dr. Tang and me. I would like to thank Dr. Tang for his magic spotting when I struggled with bench press and his three-point shots plus creative assists in basketball court to make our team win all the time. I am thankful to Professor

Yunfeng Lu for his Thanksgiving parties and advices for my career. My friends and colleagues Angelica Zabala, Anuradha Biswas, and Dr. Zhen Gu are thanked for their generous helps and encouragement when things did not go well.

I am deeply grateful to my parents Peimin Wang and Huifang Zhao for their love over years. Without their supports and encouragement, I would not be able to pursue my Ph. D. at UCLA Los Angeles thousands of miles away from my hometown. I am thankful for their firm belief in education and unconditional supports of my long Ph. D. journey.

Finally, I am greatly thankful to my beloved wife Xue Gao for being with me since we met each other in LAX. Her unconditional love completes me and inspires me to be a better man. I am grateful to her for taking care of my daily life and sometimes my research projects. Her suggestions and supports have helped me a lot to overcome tons of difficulties. She is my priceless treasure and my family.

VITA

2004	B. S., Life Science East China Normal University Shanghai, China
2004-2009	Research Assistant School of Life Sciences and Biotechnology Shanghai Jiao Tong University, China
2009-2013	Research Assistant or Teaching Assistant Department of Chemical and Biomolecular Engineering University of California, Los Angeles, USA

PUBLICATIONS

1. **Wang P.,** Bashiri, G., Gao, X., Sawaya, M. R. Tang, Y.* "Uncovering the Enzymes that Catalyze the Final Steps in Oxytetracycline Biosynthesis." *J. Am. Chem. Soc.* **2013**, 135 (19), 7138–7141
2. **Wang P.,** Kim W., Pickens L. B., Gao X., and Tang Y.* "Heterologous Expression and Manipulation of Three Tetracycline Biosynthetic Pathways." *Angew. Chem. Int. Ed. Engl.* **2012**, 51, 11136–11140
3. **Wang, P.**†, Gao., X.†, Tang, Y.* "Complexity Generation during Natural Product Biosynthesis using Redox Enzymes." *Curr. Opin. Chem. Biol.* **2012**, 16, 362-369.
4. **Wang, P.**, Gao., X., Chooi, Y., Deng, Z.,* Tang, Y.* "Genetic Characterization of Enzymes Involved in the Priming Steps of Oxytetracycline Biosynthesis in *Streptomyces rimosus*." *Microbiology*, **2011**, 157, 2401-2409.
5. **Wang, P.**†, Zhang, W.†, Zhan, J., Tang, Y.* "Identification of OxyE as an Ancillary Oxygenase During Tetracycline Biosynthesis." *ChemBioChem*, **2009**, 10, 1554-1550.
6. Gao, X.†, **Wang, P.**†, Tang, Y.*, "Engineered polyketides biosynthesis and biocatalysis in Escherichia coli." *Appl. Microbiol. Biotechnol*. **2010**, 88, 1233-1242.

7. Gao, X., Haynes, S. W., Ames, B. D., **Wang, P.**, Vien, L. P., Walsh, C. T.*, Tang, Y.* "Cyclization of Fungal Nonribosomal Peptides Catalyzed by a Terminal Condensation-Like Domain." *Nat. Chem. Biol.* **2012** Oct; 8(10):823-30.

8. Chooi, Y., **Wang, P.**, Fang, J., Li, Y., Wu, K., Wang, P., Tang, Y.* "Discovery and characterization of a group of fungal polycyclic polyketide prenyltransferases." *J. Am. Chem. Soc.* **2012**, 134, 9428-9437.

9. Gao, X., Chooi, Y., Ames, B. D., **Wang, P.**, Walsh, C. T., Tang, Y.* "Fungal Quinazoline Alkaloid Biosynthesis: Genetic and Biochemical Investigation of the Tryptoquialanine Pathway in Penicillium aethiopicum." *J. Am. Chem. Soc.* **2011**, 133, 2729-2741.

10. Pickens, L. B., Kim, W., **Wang, P.**, Zhou, H., Watanabe, K., Gomi, S., Tang, Y.* "Biochemical Analysis of the Biosynthetic Pathway of an Anticancer Tetracycline SF2575." *J. Am. Chem. Soc*. **2009**, 131, 17677–17689.

† Co-first authors * Corresponding authors

1. Introduction

1.1 Polyketide biosynthesis

This section contains material which was published as: Gao, X., Wang, P., Tang, Y.,"Engineered polyketides biosynthesis and biocatalysis in *Escherichia coli*." *Appl. Microbiol. Biotechnol*. **2010**, 88, 1233-1242.

Polyketides constitute an important family of structurally diverse natural products, which include many clinically useful drugs such as the antibiotics erythromycin and tetracycline, anticancer epothilone, and the anti-hypercholesterolemia lovastatin (Figure 1). The pharmaceutical values of many polyketides have led to intense efforts in recent decades towards understanding and engineering the corresponding biosynthetic pathways. These concerted efforts from disciplines such as biochemistry, structural biology, genetics and metabolic engineering have enabled many examples of rational and combinatorial biosynthesis of "unnatural" natural products, as well as enzymatic tools that can be used towards in the synthesis of modification of natural product-derived targets [1-3].

Polyketides are synthesized by a family of multifunctional enzymes known as polyketide synthases (PKSs). PKSs assemble the core structures (or aglycons) of polyketides via the sequential Claisen-like condensations of extender units derived from caboxylated acyl-CoA precursors, in a head-to-tail fashion [4]. Based on the structures of the polyketide products, as well as biochemical features of the PKSs, PKSs are currently classified into type I, type II and type III subgroups. Type I PKSs are megasynthases in which catalytic domains are typically found in a single polypeptide. A modular type I PKS, such as the 6-deoxyerythronolide B synthase (DEBS) [5], consists of multiple modules and each module catalyzes one round of

chain elongation and modification. Linear juxtaposition of modules facilitates unidirectional transfer of the growing polyketide from the upstream to the downstream modules in assembly line-like fashion [6]. An iterative type I PKS, such as the lovastatin nonaketide synthase (LovB) [7], is a monomodule megasynthase in which a single set of catalytic domains are used repeatedly in a highly programmed fashion. Type II PKSs (also known as bacterial aromatic PKSs) are composed of mostly dissociated, monofunctional enzymes that function repeatedly in the synthesis of a poly-ketone backbone [8]. Type II PKSs are involved in the synthesis of aromatic polyketides, such as the aglycons of actinorhodin [9], daunorubicin [10], and landomycin [11]. In both type I and type II PKSs, the minimal PKS components that are required to perform one round of decarboxylative condensation consist of a β-ketoacylsynthase (KS), an acyltransferases (AT) and a phosphopantethienylated acyl carrier protein (ACP) (Figure 2). The elongated polyketide product synthesized by the minimal PKS can be tailored by enzymes such as ketoreductase (KR), dehydratase (DH), enoylreductase (ER) or methyltransferase (MT), etc. Type III PKSs, such as chalcone synthase [12] are homodimeric KSs that synthesize smaller aromatic compounds in bacteria, fungi and plants. Products synthesized by the various types of PKSs can each undergo a different set of post-PKS modifications by decorative enzymes encoded in the biosynthetic pathways, such as cyclases, oxygenases, glycosyltransferases, etc to afford the structurally diverse natural products [13].

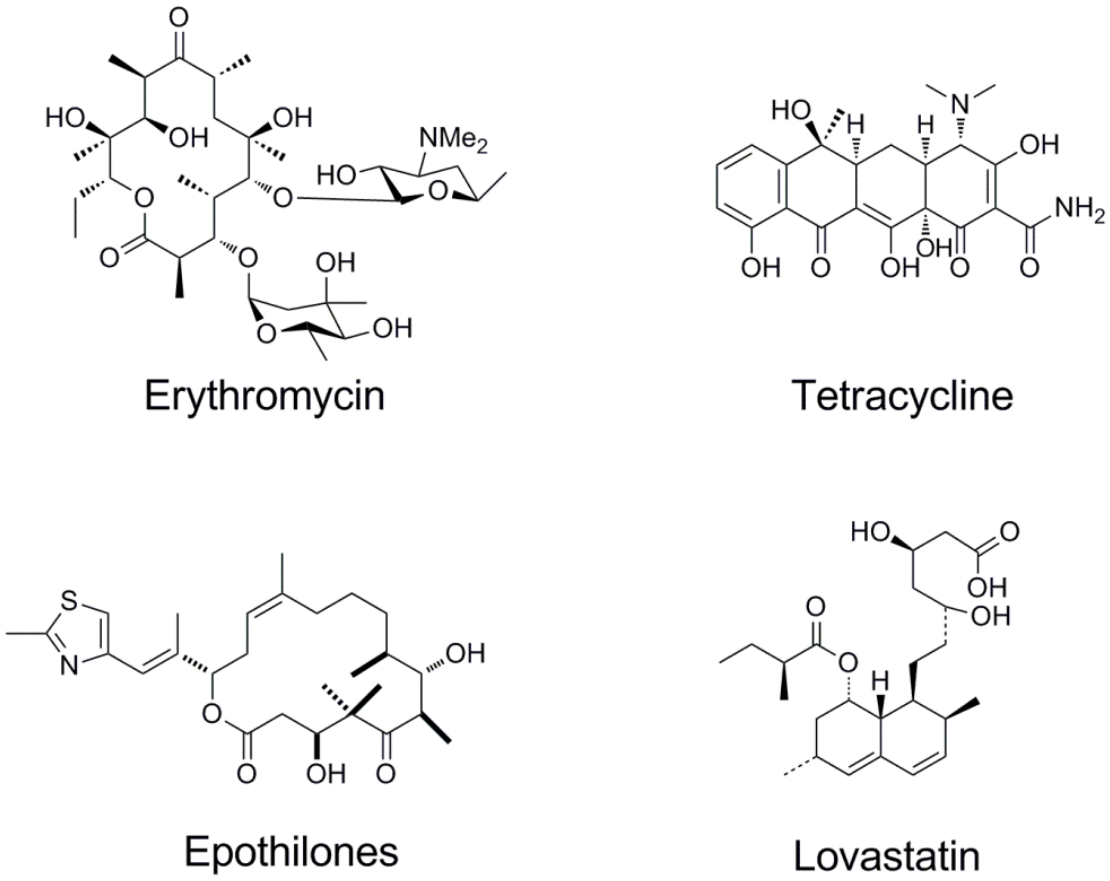


Figure 1. Chemical structures of four natural polyketide drugs.

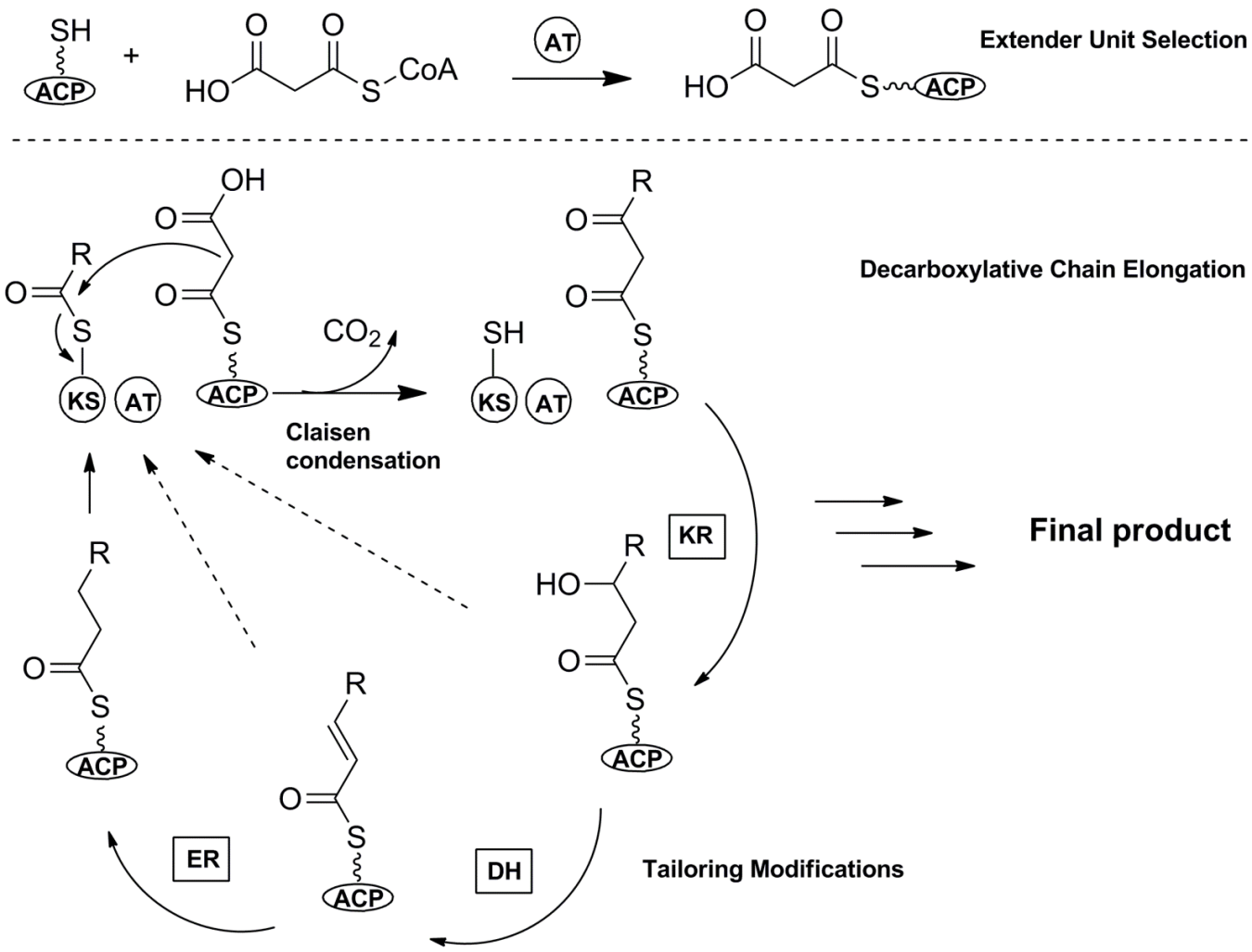


Figure 2. General mechanism involved in polyketide biosynthesis. KS, ketosynthase; AT, acyl transferase; ACP, acyl carrier protein; KR, ketoreductase; ER, enoylreductase; DH, dehydratase.

Erythromycin precursor 6-deoxyerythronolide B (6-dEB) is the 14-membered macrocyclic core of the antibiotic erythromycin synthesized by *Saccharopolyspora erythraea*. The PKS that synthesizes 6-dEB is DEBS, which consists of three large polypeptides, each exceeding 300 kDa in molecular weight. Together, DEBS contains one loading module and six extension modules, utilizes propionyl-CoA as a starter unit and six *(2S)*-methylmalonyl-CoA as extender units (Figure 3). Each module (DEBS module 1, module 2, *etc*) is responsible for one chain extension cycle, as well the reductive tailoring of resulting β-keto product. At the end of the biosynthetic assembly, a thioesterase (TE) domain fused to the C-terminus of module 6 catalyzes the macrocyclization of the linear polyketide to yield 6-dEB [14]. The cyclized 6-dEB is then modified by a series of post-PKS enzymes, including hydroxylation and glycosylation, to yield the bioactive natural products erythromycin A-D.

Since the initial discovery of the *ery* genes that encodes DEBSs [15-16], DEBS has served as the model system to study and engineer modular type I PKSs. The linear arrangement of domains and modules inspired the colinearity rule, which allows precise prediction and manipulation of the product structure from PKS protein sequence for most of the modular PKSs. Numerous 6-dEB, erythromycin and other analogs were synthesized through domain/module deletions, insertions and replacements [17]. These efforts were first performed in the natural producer *S. erythraea* [18] and model *Streptomyces* hosts [19-20] with great success.

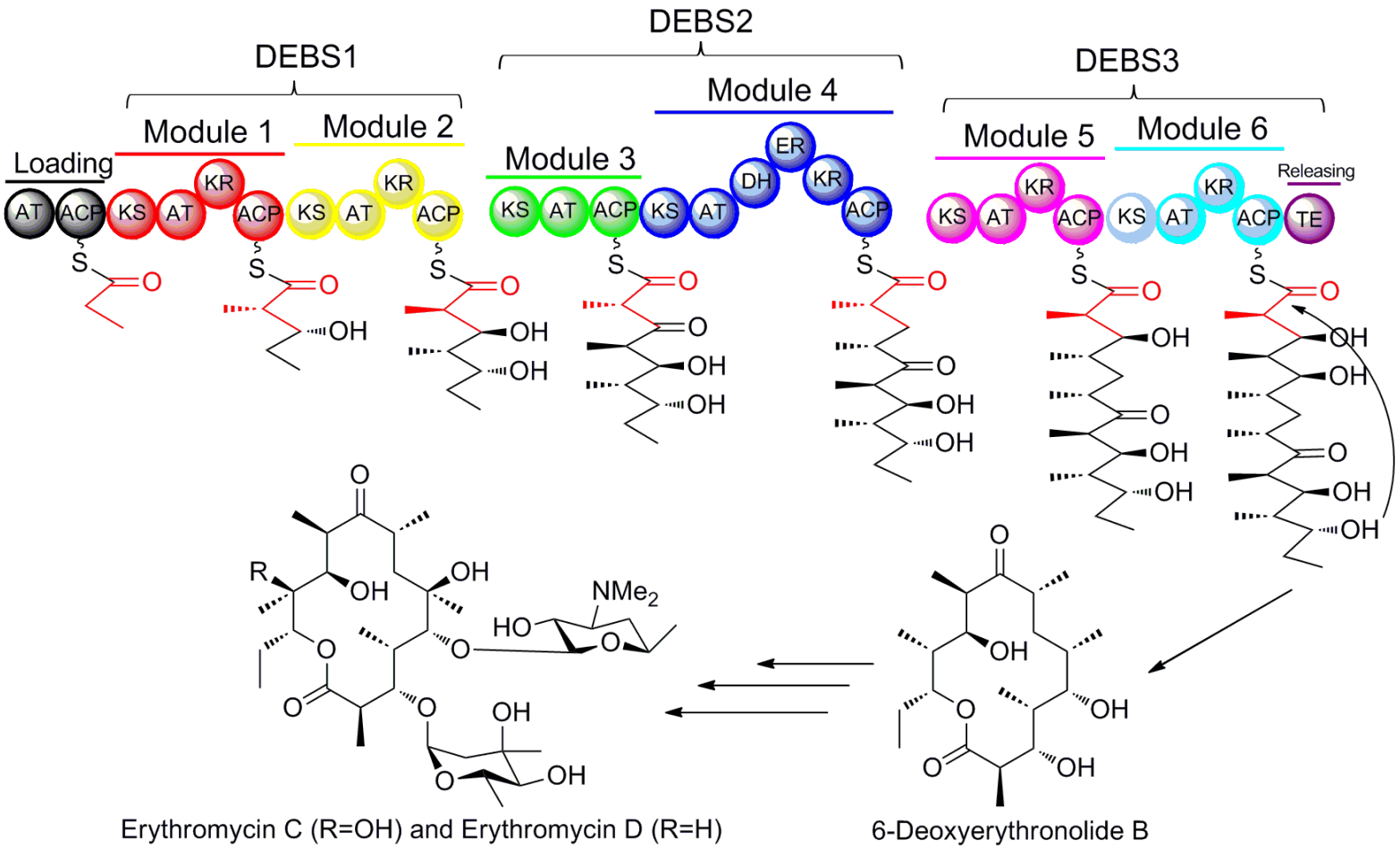


Figure 3. Biosynthesis of 6-dEB and erythromycin.

Type II polyketide landomycins represent a well-known group of natural product named as angucyclines [21]. Feeding experiments showed that the all carbons of landomycin aglycone (landomycinone) are acetate-originated indicating a decaketide intermediate involved in the biosynthesis of landomycins [22]. As expected, a type II polyketide synthase unit LanABC is encoded in landomycin gene cluster for generation of the decaketide poly-ketone intermediate in an iterative manner [23]. The decaketide intermediate then undergoes signature angular cyclization in the third ring leading to the bending of the forth ring in all the angucycline chemical structures [21]. The post tailoring enzymes especially redox enzymes further modify the angular four ring structure to afford the 11-deoxylandomycinone and ultimately landomycins by adding a number of sugar moieties (Figure 4) [21].

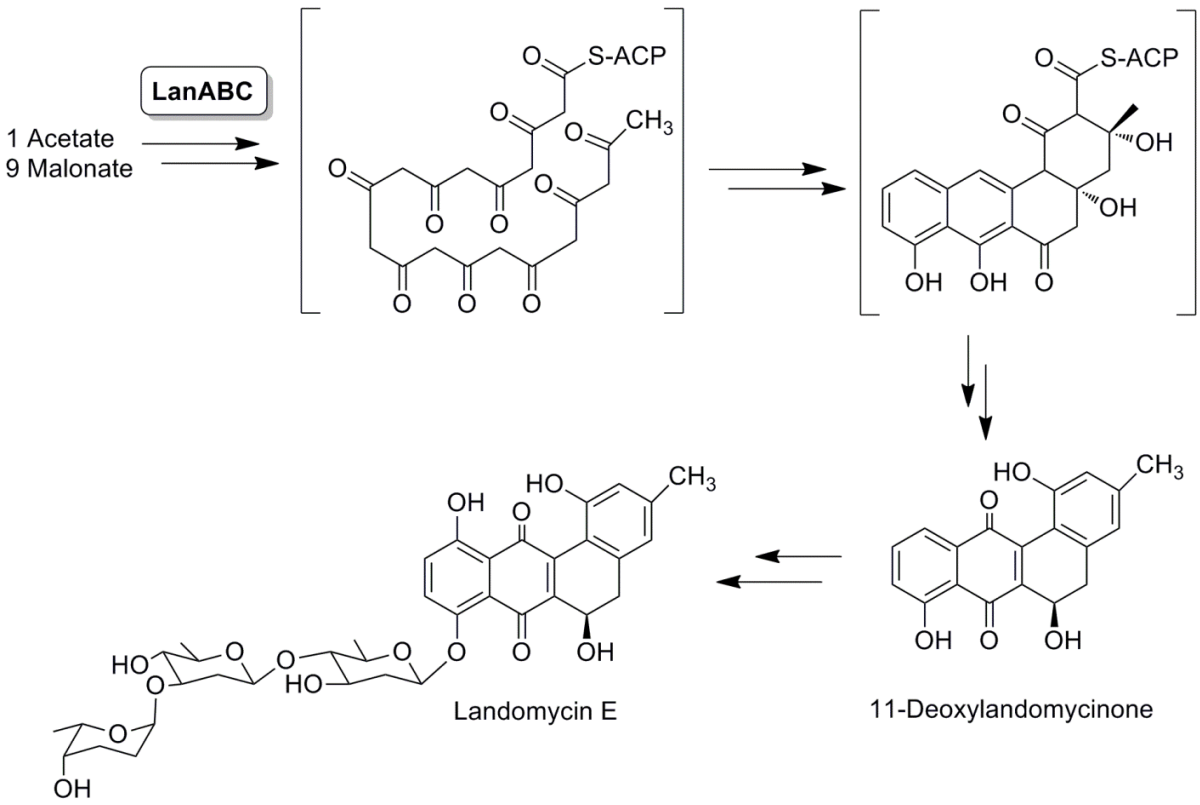


Figure 4. Biosynthesis of landomycin.

The type III PKSs provide the key structural scaffolds of a variety of plant secondary metabolites by catalyzing decarboxylative condensation between the starter unit, such as *p*-coumaroyl-CoA or cinnamoyl-CoA, and the extended unit malonyl-CoA. Depending on the specific activities of the type III PKSs, the product can be cyclized into different structures. The Chalcone synthases (CHS) cyclize the products into flavanones, which can be further modified into to a variety of flavonoids [24]. Stilbene synthase (STS) can cyclize the polyketide via a different regioselectivity to produce the stilbene backbone, which is the key intermediate in the biosynthesis of stilbenoids [25]. Curcuminoid synthase (CUS) represents the non-cyclization type of type III PKSs that only catalyzes condensation reactions to generate curcuminoids [26].

Horinouchi and co-workers used *E. coli* as a heterologous host for different plant type III PKSs and generated libraries of plant-specific polyketides from simple amino acid precursors by using different type III PKSs. First, an *E. coli* strain was engineered to overexpress phenylalanine ammonia lyase (PAL) from Rhodotorula rubra and 4-coumarate:coenzyme A ligase (4CL) from *S. coelicolor* A3(2). When supplemented with phenylalanine or tyrosine, the required starter units cinnamoyl-CoA or p-coumaroyl-CoA can be produced, respectively. Using this host, coexpression of CHS from the licorice plant *Glycyrrhiza echinata* resulted in the productions of pinocembrin chalcone and naringenin chalcone (Figure 5) [27-28]. Similarly, resveratrol and pinosylvin were synthesized by replacing CHS with Arachis hypogaea STS and overexpressing Corynebacterium glutamicum acetyl-CoA carboxylase (ACC) (Figure 5) [29]. Finally, the coexpression of PAL, 4CL, CUS, and ACC in *E. coli* produced bisdemethoxycurcumin and dicinnamoylmethane (Figure 5) [30]. Hence, once a pathway for the biosynthesis of precursors is established, the *E. coli* host can be directed towards the synthesis of different type III polyketides by varying the choice of the key PKS.

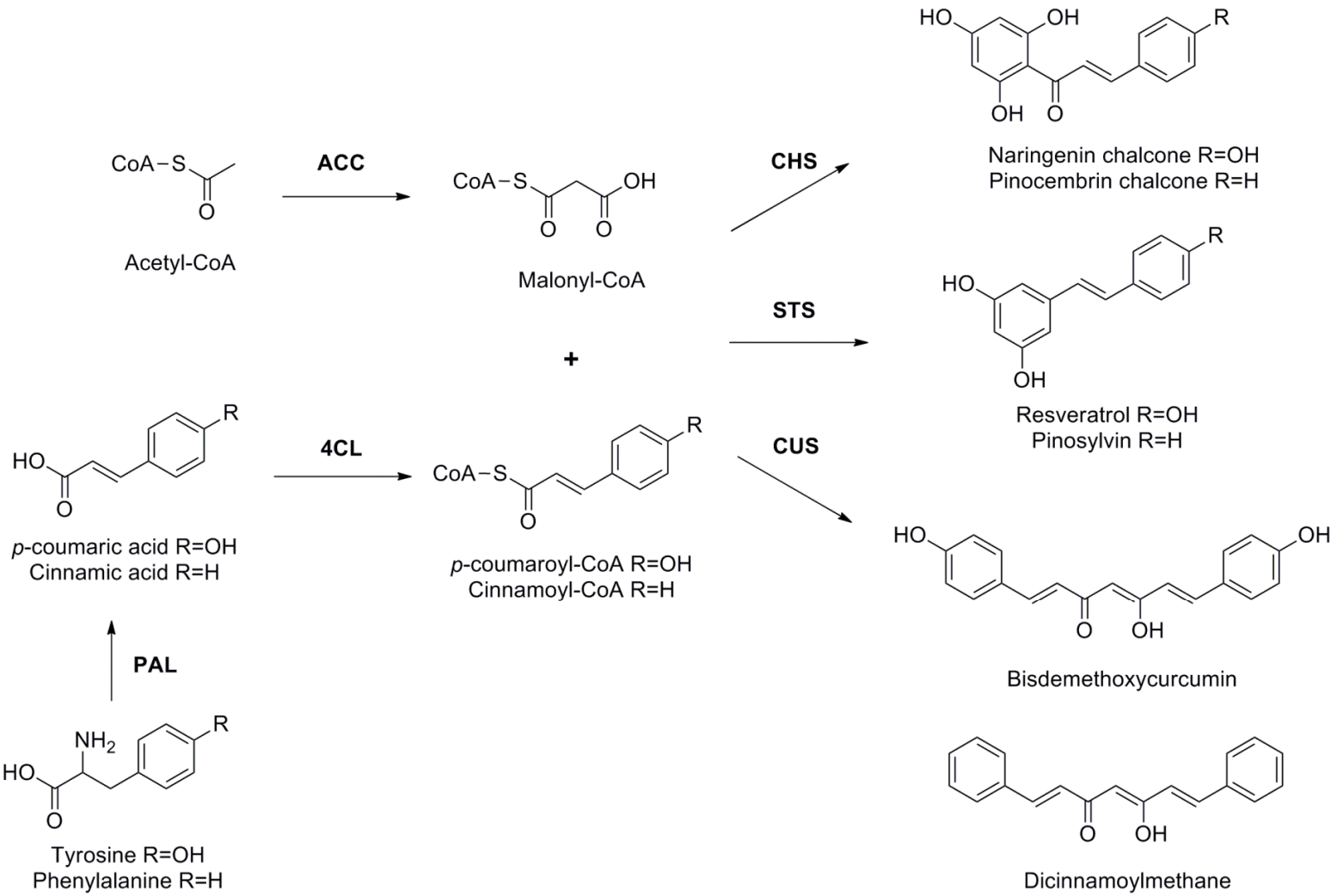


Figure 5. Biosynthesis of flavonoids, stilbenoids, and curcuminoids.

Polyketides are synthesized mostly by the soil-borne or marine actinomycetes bacteria and filamentous fungi. Type III polyketides can also be synthesized by plants. While these organisms can be prolific and impressive natural product chemists, many of them are unfortunately difficult to work with in both laboratory and industrial settings. The difficulties can be attributed to one of more of the following reasons: 1) the strain is difficult to culture (long doubling times) or domesticate; 2) the strain is genetically intractable and refractory towards common molecular biology tools; and 3) the polyketide biosynthetic pathways are weakly expressed or silent under laboratory culturing conditions. To overcome these limitations, one important goal towards engineered biosynthesis of polyketides is the establishment of robust heterologous hosts [31].

Two closely related actinomycete *S. coelicolor* and *S. lividans* have been used extensively as a powerful heterologous expression hosts for the understanding and engineering of polyketide biosynthesis recently due to the well-developed genetic tools and the available genome sequences of the two strains [32]. Three signature natural product biosynthetic gene clusters of actinorhodin (Act), undecylprodigiosin (Red), and calcium-dependent antibiotic (CDA) are harboured in both genomes but with essentially different expression levels [33]. In *S. lividans,* the three gene clusters remain silent while *S. coelicolor* steadily produces all the three natural products without further activation manipulation [34]. To reduce the possible interferences between Act and transplanted polyketide biosynthetic pathways, *S. lividans* K4-114 was constructed by deleting *act* gene cluster [35]. Compared to *S. coelicolor*, the natural silence of the three natural product biosynthesis and the highly efficient uptake of methylated exogenous DNA advantage *S. lividans* as a superior host for the reconstitution, manipulation, and optimization of polyketide biosynthesis [35].

The soil-borne filamentous bacterial specie *Streptomyces* features a complex development circles including spores germination, vegetative mycelium growth, and aerial hyphae compartment-spore formations [36]. Production of secondary metabolites is usually detected during the development of aerial hyphae and spores formation associating with nutrient limitations [37]. The synchronization between the specific development phase and the secondary metabolites production required a sophisticated regulatory system in *Streptomyces*. Basically, there are three layers of controllers to regulate secondary metabolite production and coordinate it with cellular development (Figure 6). First, the global regulators direct both secondary metabolite production and cellular development. For instance, BldA is in charge of both the production of Act and Red, and the formation of aerial hyphae and spores in *S. coelicolor* by encoding rare leucine codon UUA tRNA [38]. Second, the pleiotropic regulators usually govern two or more secondary metabolite biosynthesis. A representative of this regulator family is AsfR which regulates Act, Red, and other calcium dependent metabolites in *S. coelicolor* [39]. Third, the pathway specific regulators control a single matching secondary metabolite biosynthetic pathway. For example, *Streptomyces* antibiotic regulatory protein (SARP) family is ubiquitous in *Streptomyces* species playing indispensable roles on secondary metabolite regulations. SARP regulators usually consist of a N-terminus DNA binding domain and an activation domain [40]. By recruiting RNA polymerase in the specific promoter region, SARP regulators initiate the activated transcription of biosynthetic genes and sequentially highly enhance the production of the regulated secondary metabolites.

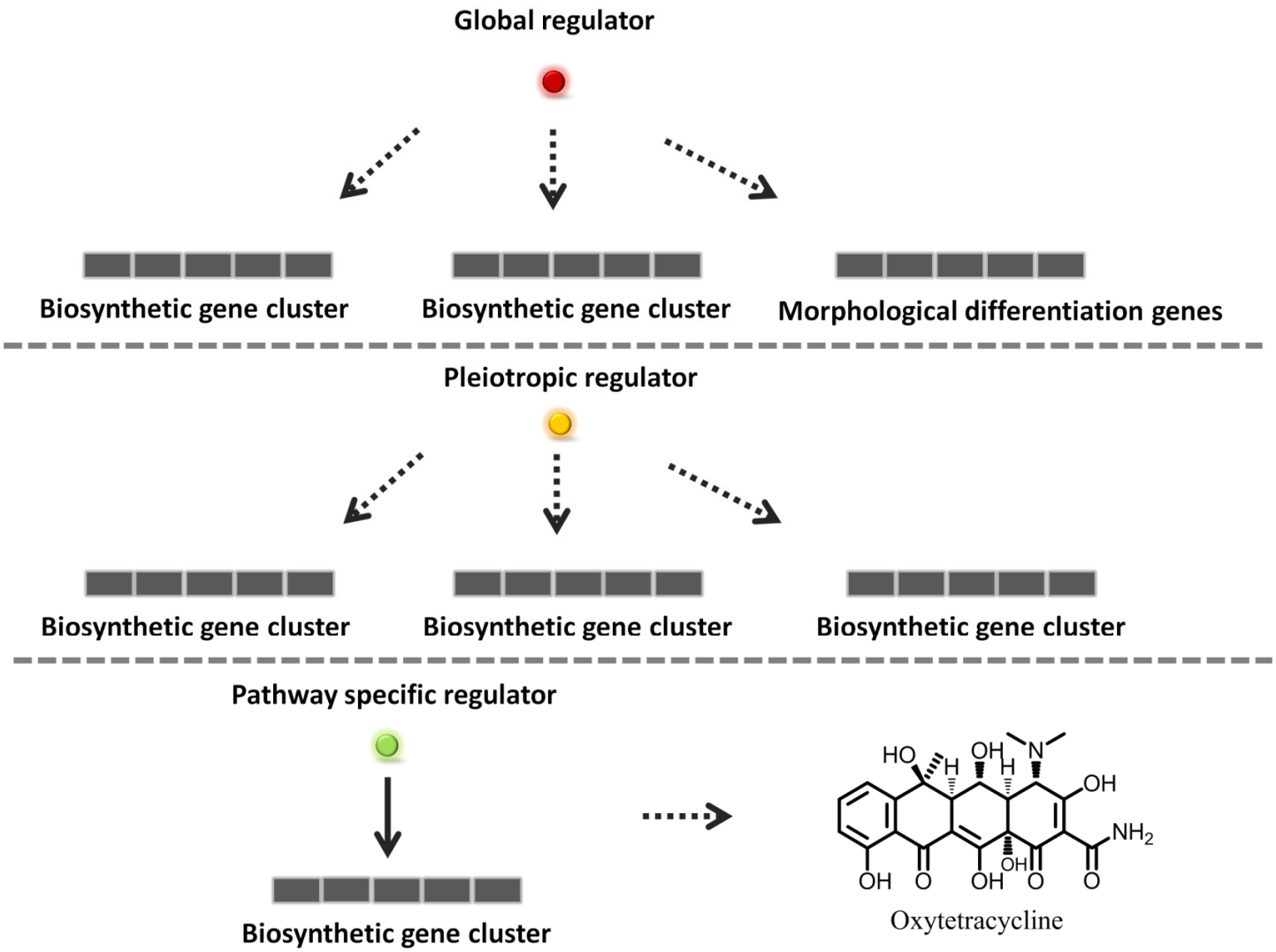


Figure 6. Three layers of secondary metabolite production regulations.

1.2. Redox tailoring enzymes involved in polyketide structural complexity generation

This section contains material which was published as: Wang, P., Gao, X., Tang, Y."Complexity Generation during Natural Product Biosynthesis using Redox Enzymes." *Curr. Opin. Chem. Biol.* **2012**, 16, 362-369.

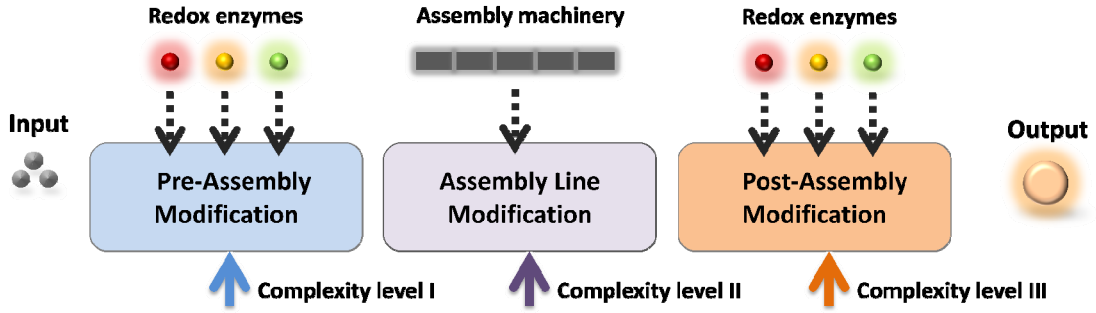


Figure 7. Complexity generation during natural product biosynthesis can take place at three different stages.

To expand chemical spaces of natural products, a variety of tailoring enzymes are adopted by microorganism, such as methyltransferase, glycosyltransferase, acyltransferase, aminotransferase, and redox enzymes [13]. Among them, redox enzymes play important roles in morphing natural product precursors into structurally diverse final products through cascade reactions at different stages. We selectively discuss a number of redox enzymes involved in polyketide natural product biosynthesis with a special focus on pre-assembly and post-assembly stages.

1.2.1 Pre-assembly modification in polyketide biosynthesis

Polyketides are a group of natural products synthesized by PKSs using successive Claisen-like condensation reactions [41]. Prior to entering the polyketide assembly line,

recruitment of starter and extender acyl thioesters is required. A variety of starter units are incorporated for polyketides produced by different families of PKSs [42]. Extender units for polyketide biosynthesis are selected by acyltransferases in PKSs in the form of CoA thioesters of dicarboxylic acids. The most common malonyl-CoA and methylmalonyl-CoA are siphoned from primary metabolism pool of the cell. In recent years, a growing numbers of polyketide natural products were discovered to contain acyl-malonyl building blocks synthesized by crotonyl-CoA carboxylase/reductase (CCR) homologs (Figure 8a) [43]. Through reductive carboxylation of the β carbons of a variety of α-β unsaturated fatty acyl-CoAs, CCRs can synthesize a number of dicarboxylic acids, including ethyl- [44], chloroethyl- [45], butyryl-, pentyl-, isopentyl-, hexylmalonyl-CoA [46].

The first dedicated CCR homolog in natural product biosynthetic pathways was SalG found in the salinosporamide A gene cluster from the marine bacterium *Salinispora tropica* [45]. SalG synthesizes chloroethylmalonyl-CoA, which is selected as a building block by the PKS SalA. The function of SalG was confirmed by gene inactivation and in vitro assay using 4-chlorocrotonyl-CoA as the substrate [45]. SalG was subsequently shown to also generate 4-bromoethylmalonyl-CoA and 4-fluoroethylmalonyl-CoA from the corresponding 4-halocrotonyl-CoAs, which was utilized by downstream enzymes to produce bromosalinosporamide [47] and fluorosalinosporamide [48], respectively. Since then a growing number of CCR homologs have been discovered in PKS pathways, including those of FK506 [46], sanglifehrins [49], divergolides [50], and ansalactams [51] *etc*.

The detailed mechanism of CCR-catalyzed carboxylative reduction of (*E*)-crotonyl-CoA to (2*S*)-ethylmalonyl-CoA was revealed by Alber and coworkers [52]. The reaction is initiated

by a hydride transfer from NADPH onto the β carbon followed by the *anti* carboxylation with CO_2. The structure of hexylmalonyl-CoA synthase CinF involved in cinnabaramides biosynthesis was recently reported [53]. The structure of CinF in complex with 2-octenoyl-CoA provided important insights into the mechanism and substrate specificity of the CCR homolog. Two smaller amino acid residues Gly362 and Ala163 in the CinF binding pocket are important for accommodating the larger 2-octenoyl chain. The corresponding residues in CCR from *S. coelicolor* are Phe370 and Ile171, which sterically prevent the binding of substrates larger than crotonyl-CoA. CinF small to large point-mutants Gly362Phe and Ala163IIe each lost activities toward 2-octenoyl-CoA, but remained active towards crotonyl-CoA [53]. The accumulation of CCR homologs and the newly gained understanding of CCR substrate specificity will enable dramatic expansion of polyketide extender unit pools. With concerted engineering of acyltransferase specificity and downstream PKS flexibility, these new extender units may lead to significantly more structure complexity in the final products.

1.2.2 Post-assembly modification in polyketide biosynthesis

Following the assembly-line like synthesis of the polyketide backbone, post-assembly modifications by different tailoring enzymes are required to convert the compound into its bioactive form. Among them, redox enzymes play crucial roles in rearranging the scaffolds into highly complex, sometimes unrecognizable final products. The biochemical mechanisms of two elusive oxidative cascades were recently established and are highlighted below.

Polyketides containing spiroacetal cores are widely found in nature, including reveromycin A [46], griseorhodin A [54] *etc*. Compounds in this family exhibit a broad spectrum of biological activities, such as the potent induction of apoptosis in osteoclasts by

reveromycin A. While the stereospecific spiroacetal cores are crucial for the observed activities, the enzymatic basis of the transformation from linear precursors has been unresolved until recently for the cyclization step in reveromycin A [46]. Two enzymes, a dihydroxy ketone synthase (RevG) and a spiroacetal synthase (RevJ), are involved in stereoselective spiroacetal formation (Figure 8b). Knockout of RevG led to accumulation of the acyclic precursor RM-A1a **1** that contains the 11, 15, 19-triol. In vitro assay with RevG and **1** led to formation of an unstable intermediate that contains the dihydroxy ketone moiety, which can spontaneous form the spiroacetal without stereocontrol. Addition of the previously unknown enzyme RevJ resulted in dehydrative cyclization of the dihydroxy ketone into exclusively the *15S* product **2** that is found in the final reveromycin A structure [46]. Interestingly, homologs of RevG and RevJ are not found in the known gene clusters of other spiroacetal-containing polyketides, suggesting different mechanisms of stereocontrol may be required to generate different cyclic scaffolds.

The gilvocarcins are a group of antitumor natural products bearing a unique benzo[*d*]naphtho[1,2-*b*]pyran-6-one chromophore linked with a variety of *C*-linked deoxysugar moieties [21]. The unique chromophore is rearranged from an angucycline polyketide via multiple oxidative modifications. Rohr and coworkers performed one-pot total enzymatic synthesis of the aglycon defucogilvocarcin M **7** using purified enzymes and various redox partners/cofactors (Figure 8c) [55]. In the "reaction pot", the ensemble of PKS enzymes assembled acetyl-CoA and malonyl-CoA into the angucycline scaffold **3**, which was tailored into **7** by a collection of four redox enzymes (GilOI, GilOII, JadF, GilR) and two methyltransferases (GilM, GilMT) in the presence of cofactors FAD, NADPH, *S*-adenosyl methionine (SAM), and cofactor regeneration enzymes. The precise control of reaction components through removal of individual tailoring enzymes in the pot, as well as the elimination of nonspecific reactions from

endogenous cellular enzymes under in vivo conditions, enabled recovery of true intermediates in the reaction cascades. Interestingly, the inability to process the angucycline UWM6 under any combinations of the tailoring enzymes indicated that UWM6 is in fact a shunt product, instead of the long-standing assignment as an intermediate. This finding also led to the assignment of the oxygenase JadF (the soluble replacement of GilOIV) in the release, decarboxylation and dehydration of the ACP-tethered polyketide product **3** to yield **4**. In addition, GilOI was identified as a bifunctional enzyme catalyzing both C12 oxidation and 4a,12b-dehydration of **4** to yield **5,** which undergoes oxidative cleavage catalyzed by GilOII to afford the dibenzochromen-6-one core. The last step in the synthesis of **7** from the hemiacetal **6** is catalyzed by the oxidoreductase GilR, which can also catalyze the conversion of pregilvocarcin V to the gilvocarcins V [56]. Crystal structure of GilR in complex with pregilvocarcin V was recently solved, which showed the cofactor FAD is covalently attached to the enzyme through His65 and Cys125 [57]. Two amino acids Tyr445 and Tyr448 located in the catalytic pocket were proposed and confirmed by mutagenesis to play crucial roles in catalysis. With the roles of the redox enzymes assigned in this cascade, one gains more appreciation of the nature's remarkable ability to control reactive intermediates en route to highly complex structures.

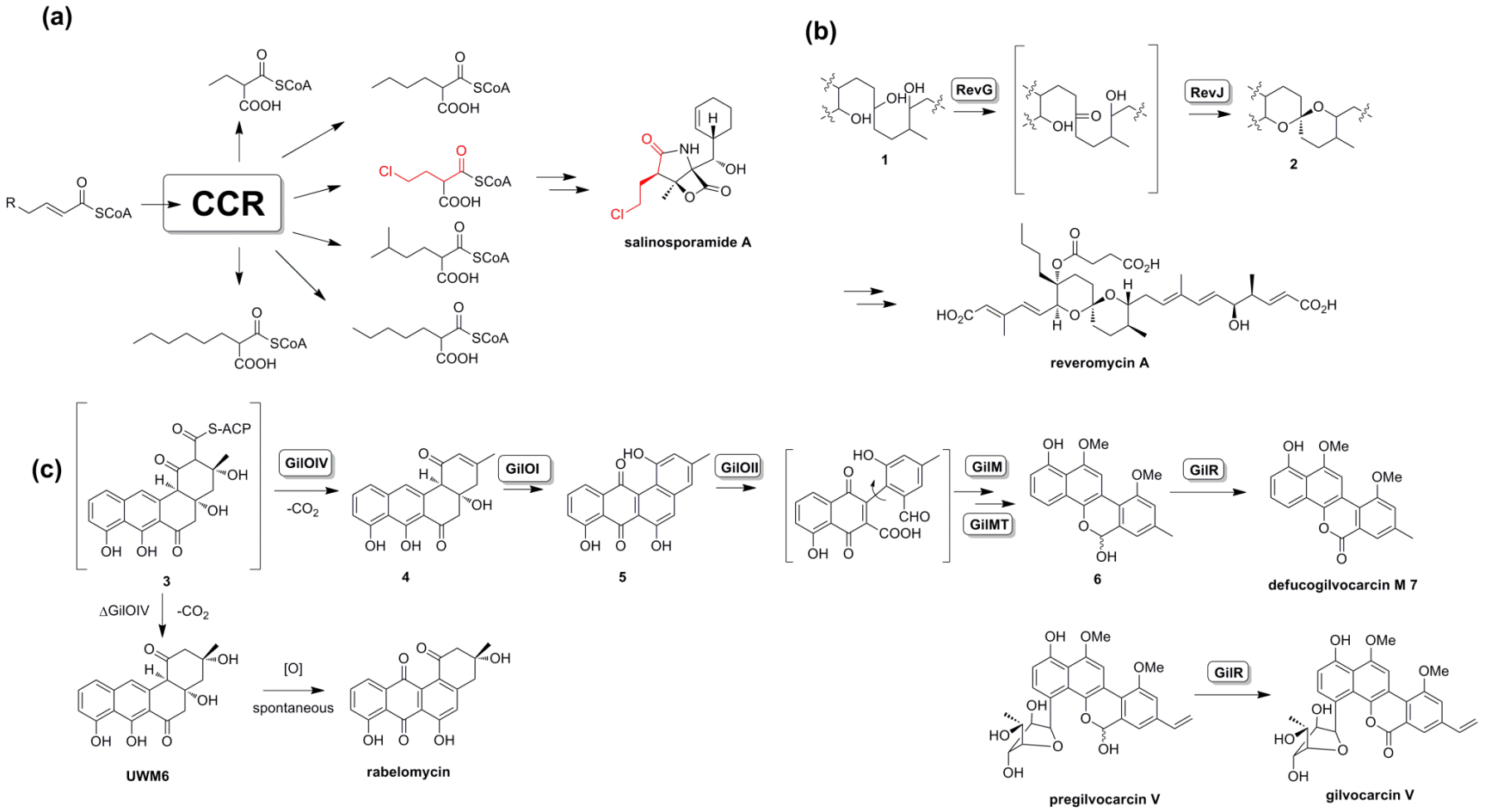


Figure 8. Representative pathways to generate natural complexity in polyketide biosynthesis.

1.3. Previous studies on tetracycline biosynthesis

Since the first discovery of chlortetracycline in 1948, only a number of natural tetracyclines were isolated during the following six decades including oxytetracycline [58], SF2575 [59], and dactylocyclines [60]. Attempts to understand the biosynthetic logic of tetracyclines have been intensively made by different research groups [61]. Fairly elucidated partial biosynthetic pathway of tetracycline was gradually unfolded based on metabolite analysis of blocked-pathway mutants in 1960s [62-64]; however, the detailed enzymatic basis for tailoring steps remains largely elusive. Thanks to recent advances of DNA sequencing technology, the investigation progress has been dramatically accelerated. Three natural tetracycline gene clusters have been documented by our group and others including which of chlortetracycline [65], oxytetracycline [66], and SF2575 [67].

By using both in vivo and in vitro analysis, the major portion of oxytetracycline biosynthetic pathway, starting from starter unit malonamoyl-CoA to key intermediate anhydrotetracycline, has been unravelled [68]. Established biosynthetic pathway is showed in Figure 10. Extended minimal polyketide synthase unit OxyABCD synthesizes an amidated poly-ketone backbone by using malonyl-CoAs as building blocks. After OxyJ catalyzes C9* ketone reduction, a number of dedicated cyclases catalyze aldol condensation like four ring cyclizations in a specific manner [69] (C7-C12, C5-C14, C3-C16, and C1-C18) to afford an aromatic four ring intermediate pretetramide. OxyF transfers a methyl group from *S*-adenosyl methionine (SAM) to a reactive C6 position, forming 6-methyl-pretetramide. By using 6-methyl-pretetramide as substrate, OxyLQT catalyzes sequential transformations to generate anhydrotetracycline through C4, C12a oxidations, C4 ketone reductive transamination, and C4

amino methylations [70]. In addition to essential biosynthetic enzymes, the functions of assistant enzymes were also discovered. During the starter unit priming steps, OxyP works as a policing machinery by hydrolyzing acetyl-ACP to free ACP. The suppression of acetyl-ACP favours correct starter unit malonamoyl-ACP to initiate the biosynthesis of oxytetracycline over a competing product 2-acetyl-2-decarboxyamido-oxytetracycline (ADOTC) [71]. Moreover, OxyE, a putative FAD dependent monooxygenase, was found to catalyze C4 hydroxylation on 6-methyl-pretetramide. The functional overlap between OxyE and OxyL at C4 position indicates an assistant mechanism has been developed by microorganism to accelerate lagging transformations by adding extra enzyme at the same position [72].

The enzymatic basis of the final transformations from anhydrotetracyline to oxytetracycline is a longstanding question since the discovery of oxytetracycline in 1950. Based on blocked-pathway mutant metabolite studies, a proposed three-step biosynthetic route was gradually composed [73]. It was proposed that two oxygenases catalyze two hydroxylations at C6 and C5 to form an unstable intermediate 5a-11a dehydrooxytetracycline; a cosynthetic factor I (F_{420}) dependent tetracycline dehydrogenase (reductase) converts this intermediate to final product oxytetracycline (Figure 9) [74]. Only OxyS (otcC) has been documented for initiating the final transformations by hydroxylating C6, while the other two enzymes remain unknown [75].

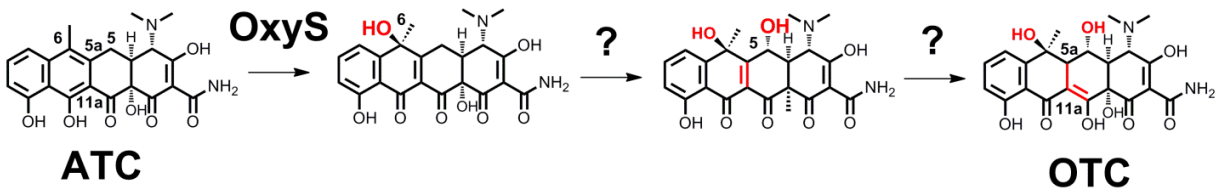


Figure 9. Final transformations from ATC to OTC remain elusive.

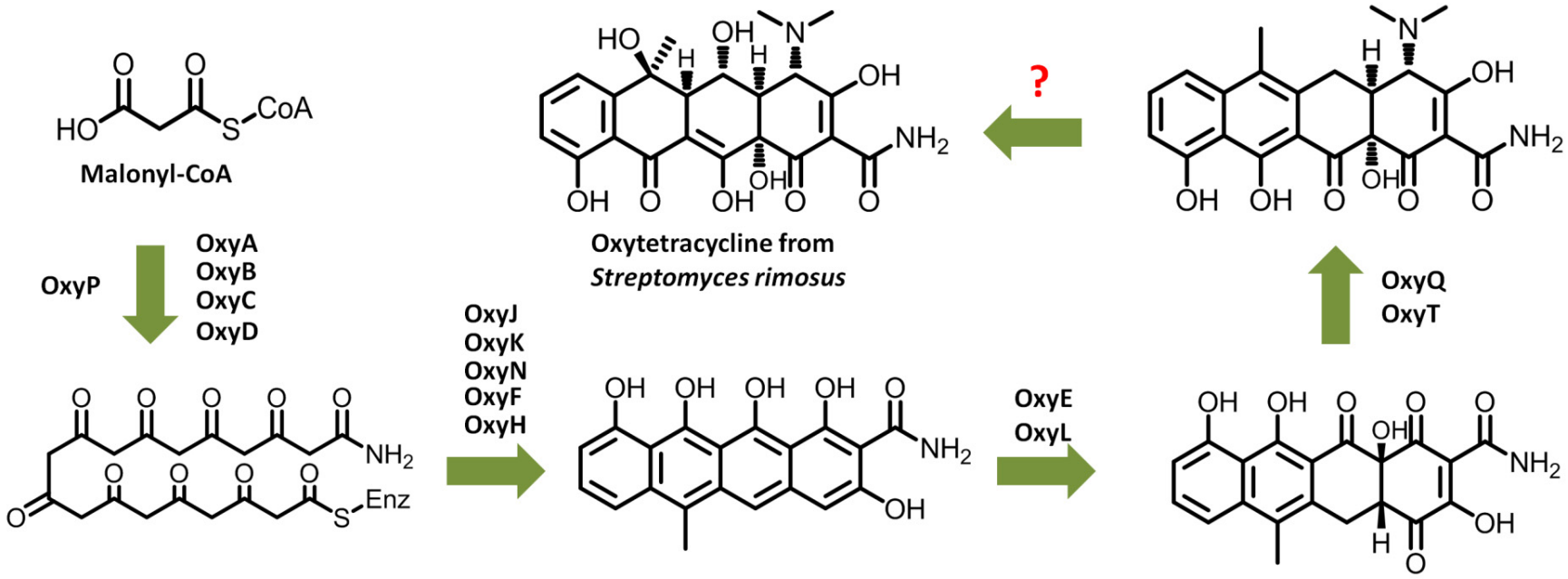


Figure 10. Established oxytetracycline biosynthetic pathway in previous studies.

2. Results and Discussion

2.1 Uncovering the enzymes that catalyze the final steps in oxytetracycline biosynthesis.

This section contains material which was published as: Wang, P., Bashiri, G., Gao, X., Sawaya, M. R. Tang, Y. "Uncovering the Enzymes that Catalyze the Final Steps in Oxytetracycline Biosynthesis." *J. Am. Chem. Soc.* **2013**, 135 (19), 7138–7141

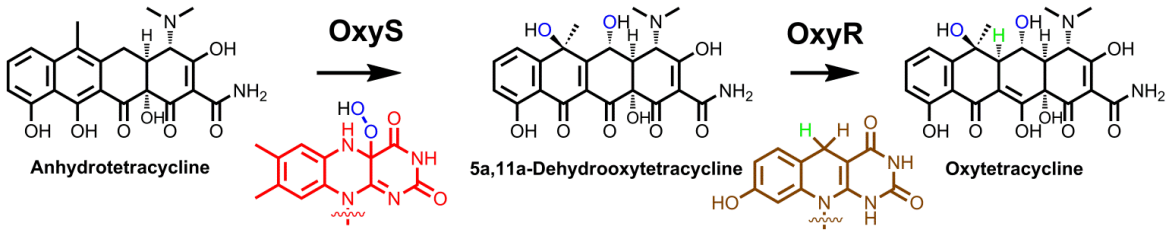


Figure 11. Redox enzymes OxyS and OxyR catalyze the final transformations involved in oxytetracycline biosynthesis.

2.1.1 Introduction

Tetracyclines are a group of aromatic polyketide natural products produced by soil-borne bacteria [8, 76]. Their broad-spectrum antibiotic activities led to the wide spread usage as the front line antibiotics. The emergence of resistant mechanisms reduced the effectiveness of tetracyclines, leading to the continual demand for semisynthetic analogs that can overcome resistance mechanisms [77-78]. As biosynthetic targets, the oxytetracycline **8** (oxy) and tetracycline **9**/chlorotetracycline **10** (ctc) pathways have been investigated using genetic and biochemical approaches [66, 70, 79-82]. The biosynthesis of both **8** from *S. rimosus* and **9** from *S. aureofaciens* are known to go through the common, late intermediate anhydrotetracycline **11**.

Although it is known that the anhydrotetracycline hydroxylase present in each pathway (OxyS and Cts8) is responsible for the (*S*)-hydroxylation of C6 in **11** [75, 83], the enzymatic basis of the remaining steps, including the additional hydroxylation of C5 in **8**, and the common C5a-C11a reduction that carves out the tetracycline structures are unclear (Figure 10). These steps are essential in sculpting the keto-enol containing lower periphery of tetracyclines, which is a critical feature in chelation of Mg^{2+} and binding to 30S ribosomal subunit [84]. Understanding the enzymology behind these transformations is therefore needed to complete our knowledge of the tetracycline biosynthetic pathway, as well as towards the generation of new analogs based on this privileged scaffold using synthetic biology approaches.

The final C5a-C11a reduction step has been proposed to involve a F_{420}-dependent tetracycline dehydrogenase, in which intermediates **12** or **13** is reduced to **9** or **8**, respectively [73-74, 85]. An earlier study proposed that the enzyme TchA encoded outside of the *ctc* gene cluster in *S. aureofaciens* is involved in this final reduction step of chlorotetracycline **10** (as well as **9**) biosynthesis [86]. However, recent evidences suggest that this enzyme may not play a direct role in C5a-C11a reduction, but may be a coenzyme F_{420}:L-glutamate ligase that is essential in the biosynthesis of the F_{420} cofactor itself (Figure 14) [87]. Initial bioinformatic analysis revealed no F_{420}-binding enzyme is encoded in either the *oxy* or the *ctc* cluster [66, 82]. However, recent heterologous expression of the *oxy* cluster (spanning otrA to otrB), which resulted in the production of **8** in *Myxococcus xanthus* [88] and *S. lividans* K4-114 [83], strongly suggests that all the enzymes required to convert **11** to **8** are encoded within the cluster. This therefore prompted us to examine the uncharacterized oxy genes.

2.1.2 Results and Discussion

OxyR is a small protein (16 kDa) encoded immediately adjacent, and in opposite polarity to *oxyS*, and the two are cotranscribed from the bidirectional promoter region (Figure 12) [89]. A similar genetic arrangement of homologs of OxyS and OxyR is also observed in the *ctc* gene cluster. OxyR shares 63% and 64% similarities to CtcR from *ctc* cluster and DacO4 from the recently sequenced dactylocycline (*dac*) gene cluster, respectively [83] [65]. A new phylogenetic profiling method was recently used to uncover previously unassigned, F_{420}-containing enzymes in bacteria [90]. Intriguingly, bioinformatic analysis using these newly reannotated enzymes now shows OxyR, CtcR, and DacO4 are predicted to use F_{420} as a cofactor. The up-to-date *in silico* analysis and the *oxyR-oxyS* genetic arrangement led us to examine the role of OxyR.

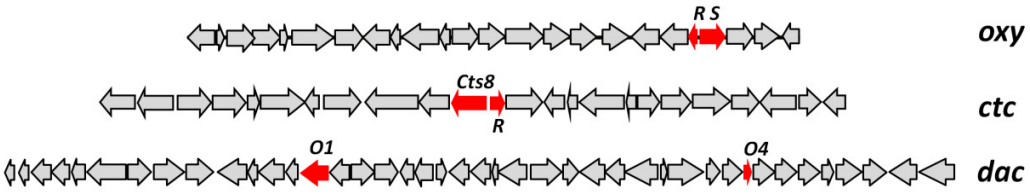


Figure 12. The organizations of the tetracycline biosynthetic gene clusters. Two essential genes *oxyS* and *oxyR* and the homologs are highlighted in red.

We inactivated *oxyR* gene in the oxy pathway that is heterologously transplanted in *S. lividans* K4-114, which resulted in elimination of biosynthesis of **8** (Figure 13). No known intermediate leading to and including **11** was accumulated from this strain. This hinted that OxyR participates in the transformation of **11** to **8**, most likely following the C6 hydroxylation of **11** to **12** catalyzed by OxyS. The 1,3,5-triketo-containing **12** is known to be highly unstable,

which may account for the lack of intermediates in the Δ*oxyR* strain. To probe the timing of OxyR, we constructed a recombinant *S. lividans* strain that overexpresses both OxyS and OxyR using the *ermE** promoter (Methods). **11** was then supplemented to a liquid culture of this strain as a substrate for biotransformation. Following 3 days of culturing, the production of **8** was detected (Figure 13), indicating that only OxyS and OxyR are sufficient to convert **11** to **8** in vivo. To investigate the role of OxyR in the pathway, and whether it is indeed the F_{420}-dependent tetracycline dehydrogenase, we then turned to in vitro experiments using purified enzymes.

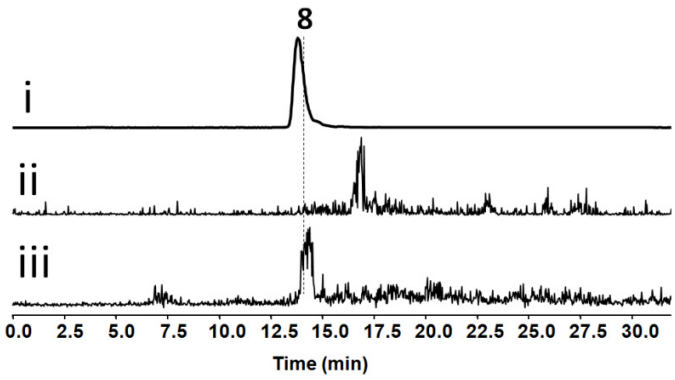


Figure 13 In vivo experiments showed that OxyS and OxyR are able to convert **11** to **8** in *S. lividans*. Mass filter (*m/z* $[M+H]^+$ 461) of i) **8** *S. lividans* producer ii) inactivation of oxyR in **8** *S. lividans* producer; iii) **11** feeding experiment with OxyS and OxyR in *S. lividans*.

The polyhistidine-tagged OxyR was expressed and purified from *S. lividans* TK64 (Figure 21). OxyR purifies with a light yellow color indicating that it is bound to a cofactor. Purified *holo*-OxyR was denatured and the resulting yellow supernatant was subjected to LC-MS analysis. The analysis showed released compounds with UV absorption spectrum, retention time, and mass consistent with coenzyme F_{420} purified from *Mycobacterium smegmatis* cells (Figure

14) [91]. This result hints that the small OxyR is the elusive F_{420}-dependent C5a-C11a reductase. To generate the reduced form of F_{420} cofactor (F_{420}-H_2) in situ (Figure 16, inset ii), the F_{420}-dependent glucose-6-phosphate dehydrogenase from *M. tuberculosis* (Mtb-FGD1) was expressed and purified from *M. smegmatis* [92-93].

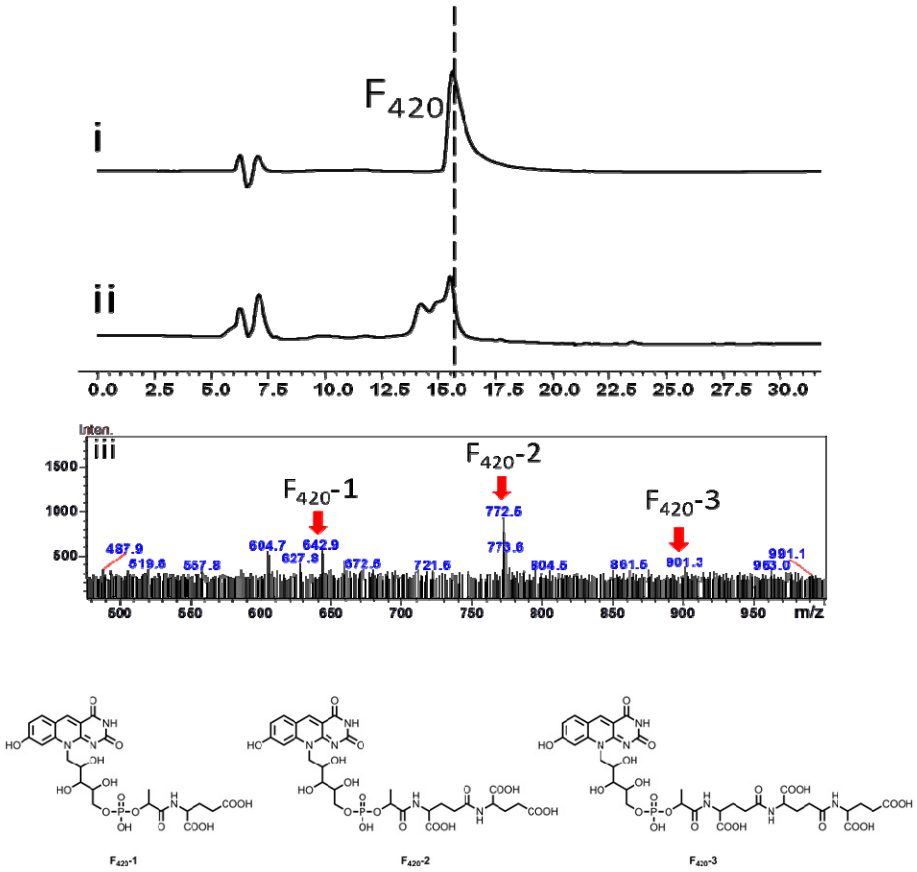


Figure 14. HPLC analysis (395 nm) of OxyR cofactor. i) oxidized F_{420} standard; ii) supernatant extract from denatured *holo*-OxyR; iii) mass analysis (*m/z* $[M-H]^-$) of supernatant extract from denatured *holo*-OxyR.

When *holo*-OxyS purified from *E. coli* (Figure 21) and *holo*-OxyR were added to **11** in the presence of NADPH and the F_{420} regeneration system, we detected the complete conversion

of **11** to **8** and **9** (Figure 15e). This is therefore consistent with our in vivo result that OxyS and OxyR are sufficient to produce the mature tetracycline scaffold. The unexpected production of **8**, instead of **9**, as the major product in this assay indicates that the C5 hydroxylation step that converts **12** to **13** is also catalyzed by OxyS (Figure 16 and see discussion below). Recombinant OxyR purified from *E. coli* does not exhibit any yellow color and does not release F_{420} upon denaturation, as would be expected from the lack of F_{420} in *E. coli* metabolism [94]. This *apo*-OxyR, when reconstituted with F_{420}, along with *holo*-OxyS, similarly catalyzed the conversion of **11** to **8** and **9** with a product ratio of 7:1 (Figure 15d).

Removal of OxyS from the reaction led to no conversion of **11** (Figure 15f), as expected since formation of the C5a-C11a double bond requires C6 hydroxylation and dearomatization of the C ring of **11** by OxyS. Exclusion of OxyR from the reaction led to complete consumption of **11**, but no stable product can be isolated for characterization (Figure 15g). At very early time points (~1 min), uncharacterized products with masses (*m/z* 443 $[M+H]^+$, *m/z* 459 $[M+H]^+$) consistent with those of singly and doubly hydroxylated products can be detected in the OxyS-only reaction mixture (Figure 20). However these products rapidly degraded, in line with our observation that the Δ*oxyR S. lividans* strain does not accumulate any isolatable intermediates.

The mechanisms of the OxyS and OxyR reactions are shown in Figure 16. OxyS catalyzes the stereospecific hydroxylation of **11** at C6 via the oxidized OxyS-Fl-4a-OOH form using the monooxygenase mechanism [95]. We propose the resulting **12** can be released by OxyS, and can be captured by OxyR to reduce the C5a-C11a double bond using the low-potential hydride provided by the F_{420}-H2 cofactor and yield **9**. However, OxyS can recapture **12**

and perform an additional hydroxylation with opposite stereochemistry at C5 to yield **13**. During their total synthesis of tetracycline, Myers and coworkers reported a C6 hydroperoxide analog of **12** that can equilibrate between ketone and enol forms at C11-C11a-C5a-C5, which confirms that vinylogous C5 proton in **12** is sufficiently acidic [96]. We propose that C5 hydroxylation of enol-**12** can also be initiated through the base-catalyzed proton extraction of the C11 hydroxyl. However, the C6-substituted and dearomatized **12** must bind in a different conformation in the active site of OxyS compared to that of **11**. This may account for the C5 carbanion attacking Fl-4a-OOH from the opposite face to give the flipped stereochemistry of the hydroxyl group at C5. The resulting **13** is then released and reduced by OxyR to yield **8**. The 8-hydroxy-7-desmethyl-5-deazariboflavin part of the oxidized F_{420} (F_{420}-Ox) is then reduced by Mbt-FGD1 using G6P as hydride donor (Figure 16, inset ii).

Our reconstitution of the final step of **8** biosynthesis raises questions with regard to the lack of C5 hydroxylation in the *ctc* or *dac* pathways (Figure 16, inset i). We examined the effect of substituting OxyR homologs from these pathways on product distribution from the in vitro assays. Both CtcR and DacO4 enzymes were purified in *apo* forms from *E. coli* (Figure 21) and reconstituted with F_{420} into the *holo* forms. Analysis of products showed that although **8** is still the dominant product, using CtcR or DacO4 decreased the ratio of **8** to **9** (Figure 15h and 15i). Notably when CtcR is used, the ratio is decreased to below 2:1. These mix and match assays showed that 1) both CtcR and DacO4 are the F_{420}-dependent C5a-C11a reductase in the respective pathways; and 2) compared to OxyR, CtcR displays a stronger affinity for **12** in competition with OxyS for the second hydroxylation step, albeit the latter still dominates to afford **8** as the major product.

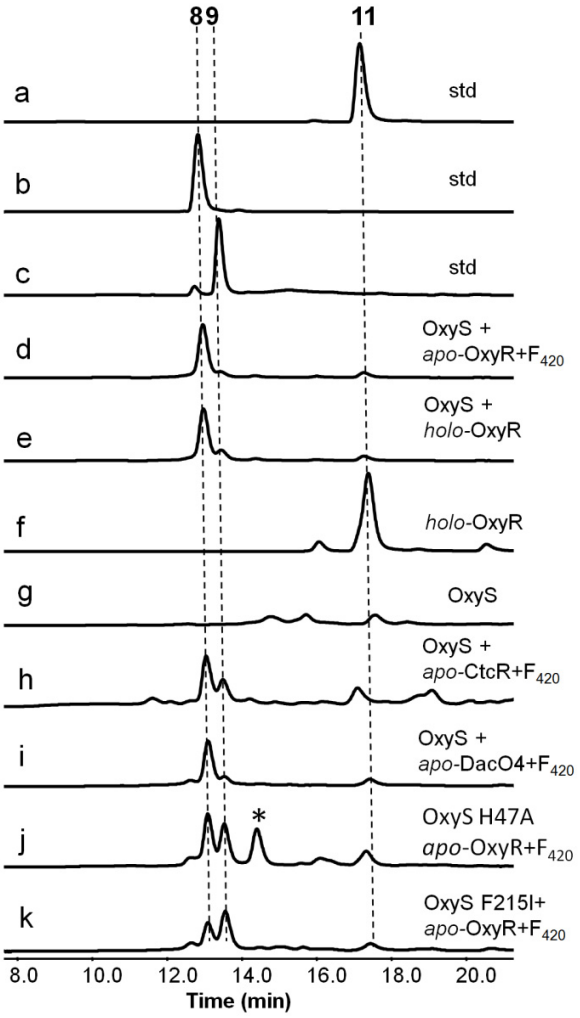


Figure 15. Analysis of OxyR and OxyS functions. The HPLC traces (270 nm) of the authentic compounds (a) **11**, (b) **8** and (c) **9**; and the enzymatic reaction products by (d) OxyS, *apo*-OxyR and F_{420}; (e) OxyS and *holo*-OxyR; (f) *holo*-OxyR only; (g) OxyS only; (h) OxyS, *apo*-CtcR and F_{420}; (i) OxyS, *apo*-DacO4 and F_{420}; (i) OxyS-H47A, *apo*-OxyR and F_{420}. (j) OxyS-F215I, *apo*-OxyR and F_{420}. Final concentrations are 20 μM OxyS, 20 μM OxyR, 100 μM 3 and 2 mM NADPH. To regenerate the reduced form of F_{420}, 2 μM FGD1, 20 μM F_{420}, and 2 mM G6P are added. The asterisk indicates an uncharacterized compound with the same mass as **9**.

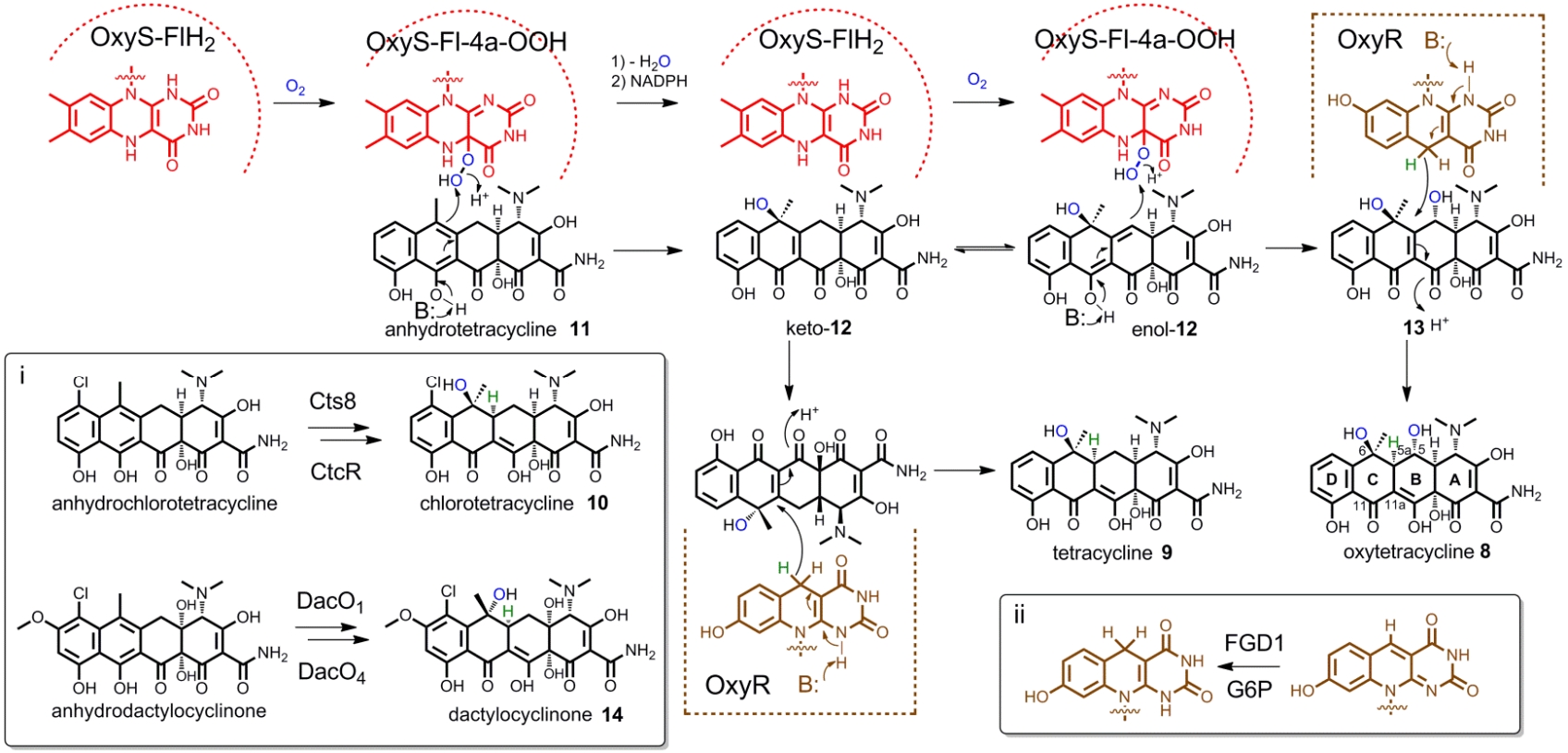


Figure 16. The proposed mechanisms of the transformations of **11** to **8** and **9**. The Fl-4a-OOH bound OxyS catalyzes the first hydroxylation at C6 of **11** to form **12**, which is in equilibrium between keto and enol forms. The latter is hydroxylated at C5 by the reoxidized OxyS-Fl-4a-OOH to form the intermediate **13**. *Holo*-OxyR with reduced F_{420} catalyzes the C5a-C11a reduction to afford the final product **8**. If C5 hydroxylation does not take place, OxyR reduces **12** into 9; Insets: i) The parallel oxidation and reduction steps that produce **10** and **14**; ii) The regeneration of F_{420}-Ox to F_{420}-H_2 catalyzed by F_{420}-dependent glucose-6-phosphate dehydrogenase (*Mbt*-FGD1) using glucose-6-phosphate (G6P) as substrate.

We attempted to examine if OxyS homologs have differential abilities to catalyze the C5 hydroxylation step. Unfortunately, we were not able to obtain soluble forms of either Cts8 or DacO1 from Streptomyces and *E. coli*. Therefore, to understand the structural basis of the tandem hydroxylation reactions and to guide mutational approaches, we determined the crystal structure of OxyS in complex with oxidized flavin (Flox) to 2.6 Å resolution (see Table 1 for statistics). The monomeric OxyS is comprised of three structural domains, including the FAD-binding domain (residues 1-175 and 271-389), the middle domain (residues 176-270) and the C-terminal thioredoxin-like domain (residues 390-503) (Figure 18). The overall structure of OxyS is similar to other FAD dependent monooxygenases found in the pathways of aromatic polyketides, including aklavinone-11 hydroxylase RdmE (PDB ID: 3ihg) from the rhodomycin pathway [97], as well as PgaE (PDB: 2qa1) and CabE (PDB: 2qa2) from angucycline pathways (Figure 17) [98]. Coordinates of these three molecules superimpose on OxyS with RMSDs of 1.6 Å, 1.5 Å, and 1.4 Å over 454, 458, and 458 alpha carbon pairs, respectively. Even the dimer interface is preserved in all three structures, involving the penultimate helix of the FAD-binding domain. The most notable differences among these structures involve 10˚ offsets in the orientation of the middle domain with respect to the FAD-binding domain, and an insertion of 8 residues (191-199) in proximity to the substrate binding pocket of OxyS.

Attempts to obtain OxyS structures bound to **11** were not successful. To model likely interactions between **11** and the substrate binding pocket of OxyS, we performed structural alignment with RdmE in complex with the cocrystallized, aklavinone **15** (Figures 18 and 19). Using the superimposed structures of OxyS and RdmE (the RMSD is 1.51 Å for 282 matching alpha carbons) and the coordinates of **15**, we can propose a putative binding pocket and binding orientation for **11**. The tetracycline substrate is anchored in a narrow hydrophobic cleft at the

interface between the FAD-biding domain and the middle domain, as indicated for **15** in Figure 18. The middle domain provides a hydrophobic and aromatic patch (Trp211-Leu217, OxyS numbering, for alignment see Figure 17) against the D and C rings of **15**. On the other side of the pocket, the well conserved patch PAGG helps to position the substrate with respect to the FAD isoalloxazine ring. As shown in Figure 18 for **15**, and can be envisioned for **11**, the planar C-D rings are perpendicular to the isoalloxazine ring of Flox, with the C6 carbon placed 4.7 Å away from the bridgehead C4a of Flox. Similarly, the C5 carbon is also located within striking distance (4.4 Å) of C4a, which may explain how OxyS can hydroxylate both positions. However, in order to achieve the opposite stereochemical outcomes, the substrate **11** and **12** must be tilted with respect to the FAD ring to facilitate attack of C6 and C5 on Fl-C4a-OOH from the opposite faces. Hence it is likely that slight perturbation in the geometry of the binding pocket can alter the ability and stereoselectivity of OxyS in catalyzing the hydroxylation steps.

To test this hypothesis, we further generated a modelled configuration of **11** in the active site of OxyS based on the D- and C-ring coordinates of **15**. This is to account for the nonplanar nature of the A and B rings in **11** (Methods and Figure 19). Comparison of the modelled positions of 8 and 3 led us to mutate His47 and Phe215, which appear to be in close contact with the substrates (Figure 18 and 19). The N1 of His47 imidazole side chain is located 3.7 Å away from the C3 oxygen atom in **15**, suggesting it might hydrogen bond to the same hydroxyl position in the A ring of **11** and **12**. When the assay was performed in the presence of a His47Ala mutant of OxyS (Figure 21), a decreasing ratio of **8** to **9** was observed. This implies that disrupting the possible hydrogen bonding interactions can indeed lead to a repositioning of the substrate and decrease the efficiency of the C5 hydroxylation step. Interestingly, we also observed the emergence of a new product from the assay that has similar UV and identical mass

(*m/z* 445 [M+H]+) as **9** (Figure 15j and 22). Although the structure of the compound cannot be determined due to low amounts produced from the in vitro assay, it is possible that this is the 6*R*-OH analog of **9** formed from the altered binding of **11** in the mutant active site.

On the other hand, Phe215 is located near the entrance of the active site and the phenyl side chain is positioned close to the C ring. The proximity of the bulky side chain may also play a role in orienting the position of **11** and **12** in the active site with respect to Fl-4a-OOH. To investigate the role of this residue, we made the mutation of Phe215Ile. Surprisingly, the slight decrease in the volume of the side chain at position 215 led to dramatic changes in the product ratio of 1:2 (Figure 15k), in which the amount of **9** now exceeded that of **8**. Although these mutagenesis studies were based on a structural alignment using the different substrate **15**, and the exact roles of these residues may differ during catalysis, it is clear that the ability of OxyS to hydroxylate both C6 and C5 has been intricately tuned during evolution. In Cts8 and DacO1, variations in the active-site configuration can lead to the exclusive production of the C6 hydroxylated products, as well as the opposite stereochemistry at C6 observed in dactylocyclinone (Figure 16, inset i). Similarly, the additional substitutions in D-rings of anhydrochlorotetracycline and anhydrodactylocyclinone, such as the C7 chlorine and C8 methoxy, can further lead to differences in substrate orientation in the corresponding active sites.

Table 1. Data collection and refinement statistics of OxyS

	OxyS
Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	76.1, 115.1, 121.0
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	2.55 (2.62-2.55)
R_{merge}	0.081 (0.507)
$I/\sigma I$	13.5 (3.4)
$CC_{1/2}$	99.8 (90.1)
Completeness (%)	89.0 (85.6)
Redundancy	5.0 (4.9)
Refinement	
Resolution (Å)	2.55
No. reflections	31454
R_{work}/ R_{free}	0.187/0.229
No. atoms	
Protein	7,440
Ligand/ion	106
Water	48
B-factors ($Å^2$)	
Protein	48.5
Ligand/ion	34.0
Water	36.8
R.m.s deviations	
Bond lengths (Å)	0.010
Bond angles (°)	1.2

*Highest resolution shell is shown in parenthesis.

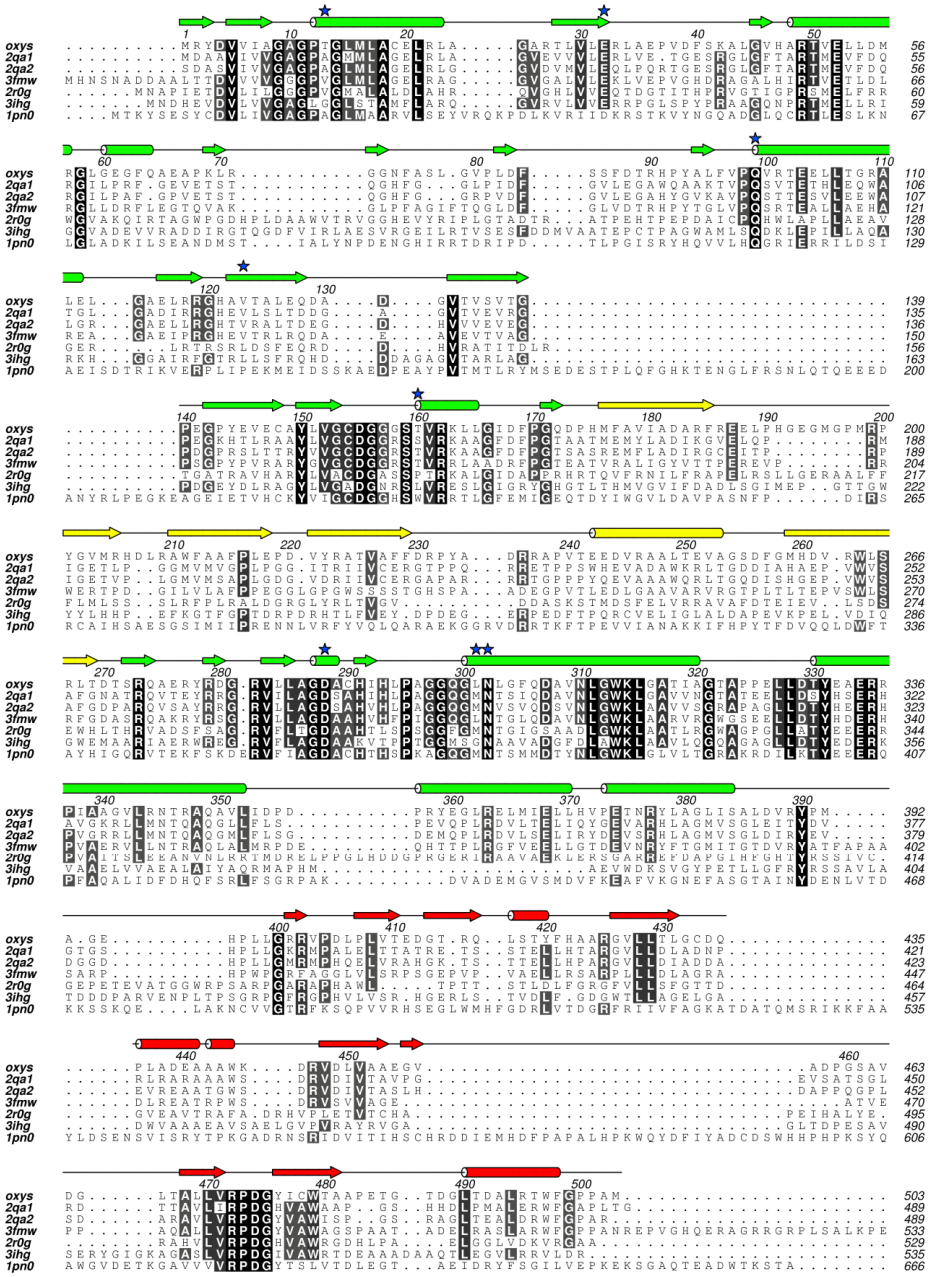


Figure 17. Structure-based sequence alignment of OxyS and RdmE (PDB:3ihg), PgaE (PDB: 2qa1), CabE (PDB: 2qa2), MtmOIV (PDB: 3fmw), RebC (PDB: 2r0g) and phenol hydroxylase (PDB: 1pn0) . The blue asterisks indicate residues that participate in hydrogen bonding with FAD.

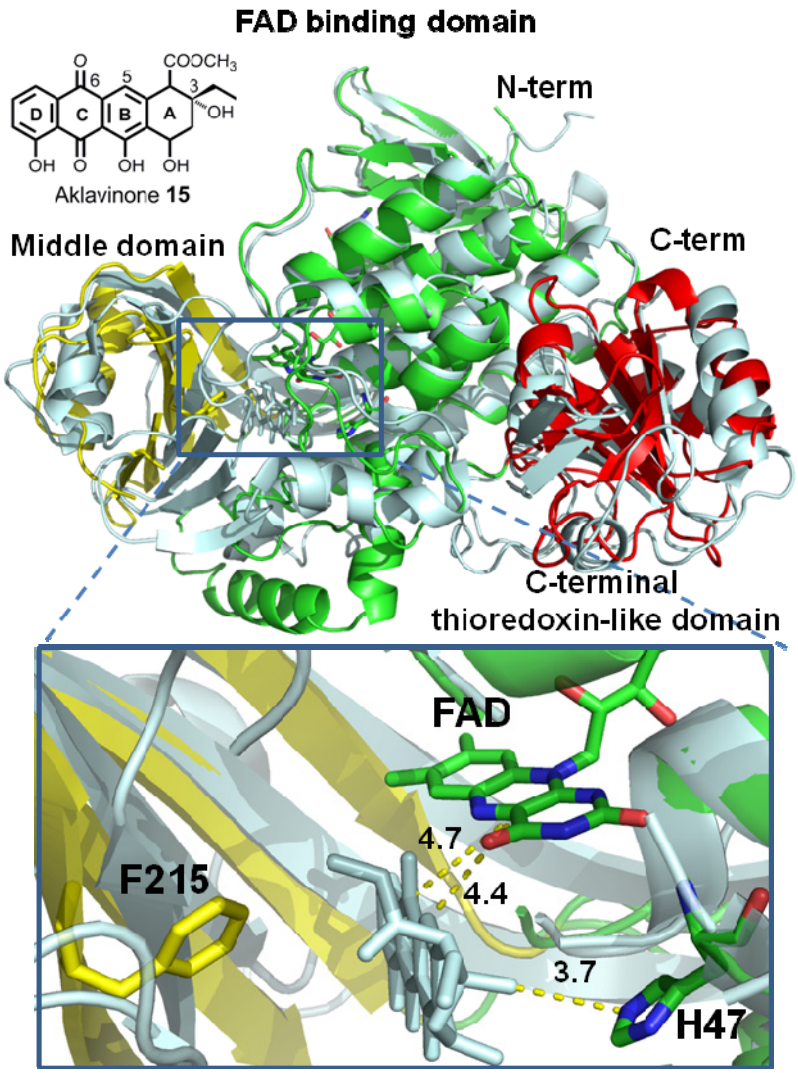


Figure 18. Overall structural alignment of OxyS with the complex of RdmE cocrystalized with aklavinone **15** (palecyan, PDB ID: 3ihg). RdmE is shown in pale cyan while OxyS is colored according to domains. The RMSD is 1.51 Å for 282 matching alpha carbons. The magnified region of the catalytic pocket indicates possible orientation of the substrate **11** or **12** using the cocrystalized **15** (shown in pale cyan). Mutation of H47A and F215I increased the amount of **9** to **8**.

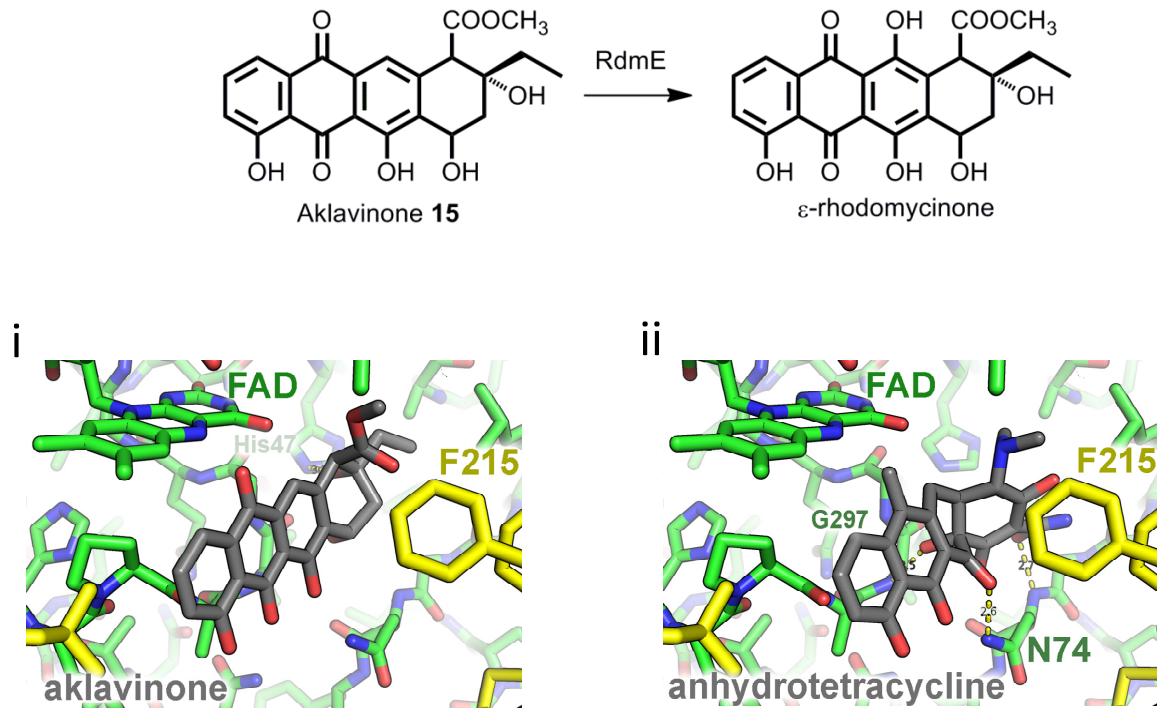


Figure 19. (Top) Reaction catalyzed by RdmE with Akalavinone **15**. (Bottom) (i) alternative view of **15** modeled in the active site of OxyS and (ii) Modeled view of **11** in the active site of OxyS based on the position of **15**.

In conclusion, we showed that OxyR, CtcR and DacO4 are F_{420}-dependent reductases catalyzing the key C5a-C11a reduction step in respective tetracycline biosynthetic pathways. Our in vitro results demonstrate that secondary metabolic pathways in actinobacteria can use F_{420} in natural product biosynthesis, which expands the utility of this cofactor beyond a nicotinamide equivalent in methogenic archaea [99] and mycobacteria [90]. This is in line with the isolation of F_{420} from the fermentation broths of numerous actinomycetes [100-101]. Other OxyR-like enzymes in secondary metabolism exist in the database, including ActVA2 (64% similarity), an enzyme of unknown function from the actinorhodin biosynthetic pathway [102]. Using structural-guided mutagenesis, we showed that OxyS is responsible for both C6 and C5 hydroxylation steps during

the conversion of **11** to **8**. While substrate specificities of OxyS and OxyR are important for the observed relative ratio of **8** to **9,** it is the unexpected ability of OxyS to catalyze the C5 hydroxylation that results in formation of **8** in the *oxy* pathway.

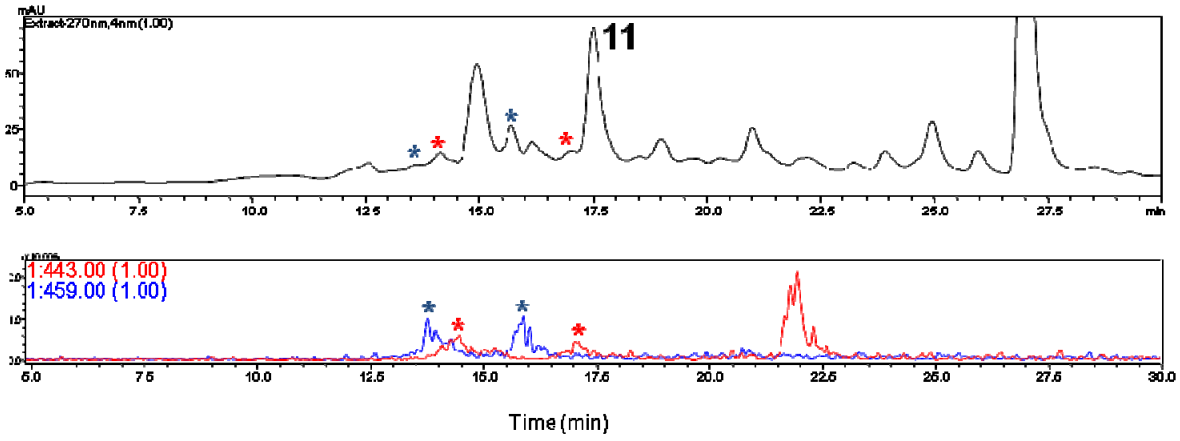


Figure 20. OxyS-only in vitro assay quenched and analyzed after ~1 min showed metabolites with mass corresponding to singly and doubly hydroxylated products. Mass filter *m/z* $[M+H]^+$ 443 (red) and 459 (blue). The red asterisks indicate the metabolites with *m/z* $[M+H]^+$ 443, the blue asterisks indicate the metabolites with *m/z* $[M+H]^+$ 459.

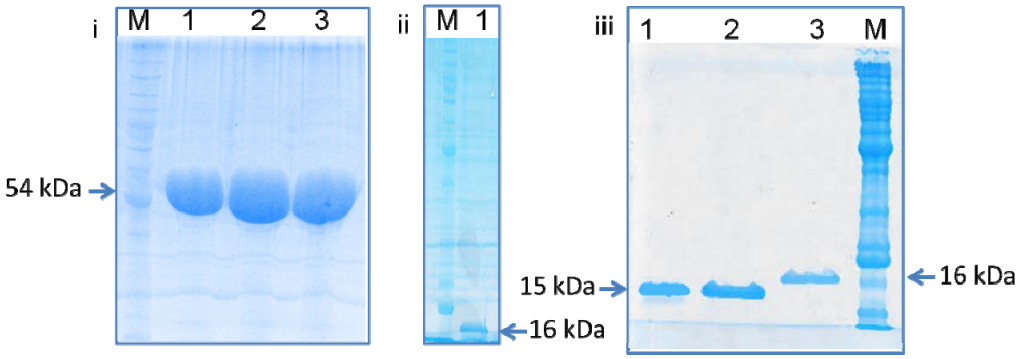


Figure 21. SDS-PAGE analysis of the recombinant proteins in this study. i) 1. OxyS-H47A; 2. OxyS-F215I, and 3. OxyS; ii) 1. *holo*-OxyR purified from *S. lividans* TK64, iii) 1. *apo*-DacO4, 2. *apo*-CtcR, 3. *apo*-OxyR. M stands for benchmark protein ladder.

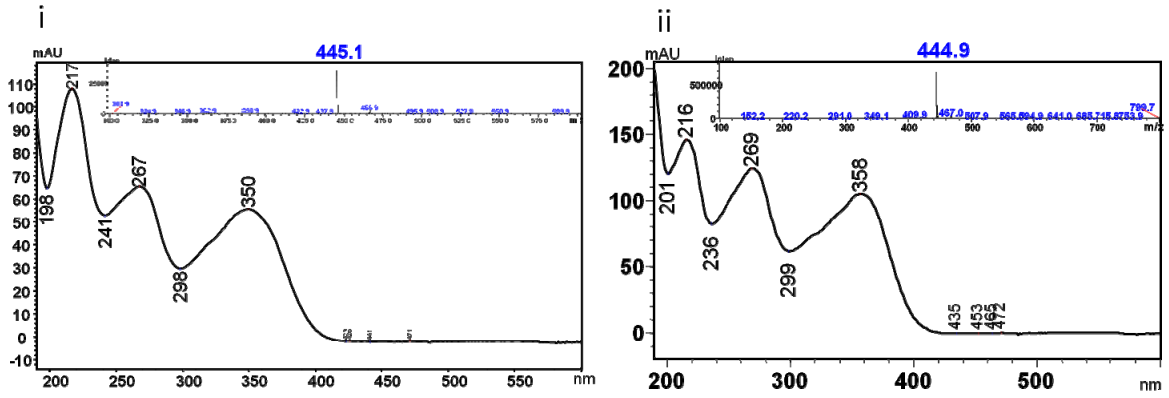


Figure 22. (i) UV and mass analysis of the new compound produced by OxyS His47Ala mutant assay and (ii) of **9** with *m/z* $[M + H]^+$ 445.

2.1.3 Materials and Methods

Expression Plasmid Construction.

OxyS encoding gene was cloned into pET28a (Invitrogen) using NdeI and HindIII sites. OxyR, CtcR, and DacO4 encoding genes were cloned into pET28a using NdeI and EcoRI sites. All the resulting plasmids were transformed into *E.coli* BL21(DE3) for protein expression. The *Streptomyces* expression vector pPWW50-Gen Poly [103] contain *oxyR* cloned using NdeI and BamHI sites. The resulting plasmid was transformed into *Streptomyces lividans* TK64 for further protein expression.

In vivo Gene Inactivation and Feeding Experiments.

The *oxyR* gene was deleted in-frame from pOTC [83], harboring the *oxy* gene cluster, by using standard lambda-RED-mediated gene deletion protocol [104]. The resulting plasmid pRKO was then co-transformed with pCtc11 into *S. lividans* K4-114 to generate OxyR-inactivated K4 strain K4/RKO. The K4/RKO strain was grown on R5 agar with 25 μg/mL

thiostrepton and 70 μg/mL kanamycin at 30 °C for 7 days and subjected to organic extraction and LC-MS analysis.

For the 3 feeding experiment, *oxyS* and *oxyR* were cloned *in cis* under the *ermE** promoter to generate pSR. The pSR plasmid was then transformed into *S. lividans* K4 to generate OxyS and OxyR coexpressed K4 strain K4/SR. The K4/SR was inoculated into 50 mL R5 liquid with 25 μg/mL thiostrepton and grown at 30 °C, 250 rpm for ~3 days. Total 4 mg **11** was then added into this 50 mL R5 culture. After 3 additional days of culturing, the R5 broth was subjected to organic extraction and LC-MS analysis.

Heterologous Protein Expression and Purification.

OxyS, OxyR, CtcR, and DacO4 were overexpressed in *E.coli* BL21(DE3) with an N-terminal hexahistidine tag. Seed cultures of the four proteins in 5 ml LB medium with 35 μg/liter kanamycin were inoculated into 500 ml LB medium with 35 μg/liter kanamycin and grown at 37 ºC until the A600 at 0.4 to 1.0. Isopropyl 1-thio-β-d-galactopyranoside was then added at the final concentration of 120 μM to induce protein expression at 16 ºC overnight. The overnight induced 500 ml cultures were subjected to centrifugation to remove supernatant. The cell pellets were resuspended in Buffer A (50 mM Tris-HCl, pH 7.9, 10 mM imidazole, and 50mM NaCl). The resuspended cells were subjected to sonication to disrupt cell membranes followed by centrifugation at 14,000 rpm 30 minutes. The soluble supernatants were incubated with nickel-nitrilotriacetic acid resin (GE Healthcare) for at least 2 hours. The mixtures of protein and resin were then loaded on to a gravity flow column. Buffer A with increasing concentrations of imidazole and 2 mM DTT was used to wash the column stepwise. Buffer A with 250 mM imidazole was used to release the target proteins from nickel-nitrilotriacetic acid resin. Amicon filtration columns (Millipore) were used to concentrate proteins and exchange protein buffer to

Buffer B (50 mM Tris-HCl, pH 7.9, 2 mM EDTA, 2 mM DTT). The concentrated protein buffers were stored with 10% glycerol at -80 ºC. Bradford assay was used to determine the concentration of target proteins [105]. The protein purification process was carried out at 4 ºC. For the expression of OxyR in S. lividans TK64, the transformants harboring OxyR expression plasmid were inoculated into 500 ml R5 medium with 25 μg/liter thiostrepton and grown at 30 ºC 4-7 days. The cell pellets were harvest by centrifugation at 3,750 rpm. Buffer D (3 mM sodium chloride, 250 mM sodium phosphate pH 7.1, 2 mM EDTA, 2 mM DTT, 1 mM benzamidine, 3 mg/l leupeptin, 3 mg/l pepstatin, and 30% glycerol) was used to resuspend cell pellets. The subsequent process is the same as described above. Mycobacterium tuberculosis FGD1 (Mtb-FGD1) was expressed and purified from *M. smegmatis* cells as previously described [93]. The coenzyme F_{420} was prepared from large-scale *M. smegmatis* cultures [91], the identity of which was confirmed using mass spectrometry (data not shown).

LC-MS analysis of in vitro assays.

General in vitro assay was carried out under the following condition with specified modification by removing different components: total 100 μL reaction contains 20 μM OxyS (OxyS H47A, OxyS F215I), 20 μM OxyR (CtcR or DacO4), 2 μM FGD1, 20 μM F420, 2 mM NADPH, 2 mM G6P, 100 μM **11** in 100 mM Tris-HCl pH 7.5 buffer at room temperature for 1 hour. The reactions were quenched by 1 mL ethyl acetate (1% acetic acid) and then the organic extracts were dried in vacuo. Samples were redissolved in 50 μL methanol for LC-MS analysis. LC-MS analyses of organic extracts of in vitro and in vivo assays were performed on a Shimadzu 2010 EV Liquid Chromatography Mass Spectrometer using both positive and negative electrospray ionization and a Phenomenex Luna C18 reverse-phase column (5μ 2.0 x 100 mm).

A linear gradient of 5 to 95% CH3CN (v/v) in H_2O (0.1% formic acid) program was used to separate samples at a flow rate of 0.1 mL/min for 30 minutes.

Crystallization and Structure Determination

For crystallization of OxyS, dialysis was used to remove imidazole from protein buffer instead of Amicon filtration column method. The hanging drop vapor diffusion method was used to grow OxyS protein crystals at room temperature. OxyS crystals grew from the reservoir condition: 16% w/v PEG-8000, 0.1 M Na/K phosphate pH 6.0, and 0.2 M NaCl. The crystals were then soaked with saturated FAD and ATC in the same buffer as the reservoir.

OxyS crystals belonged to space group P212121 with two OxyS molecules in the asymmetric unit. X-ray diffraction data were collected at the Advanced Photon Source (Argonne National Laboratory), beamline 24-ID-C, using an ADSC Quantum 315 CCD detector. Crystals were cryo-protected by a quick dip in a solution consisting of 6.5 µL reservoir and 3.5 µL glycerol. Crystals were cryo-cooled to 100 K during the data collection. One-hundred-twenty 1.0° oscillation frames were collected at a wavelength of 0.9794 Å. Data reduction and scaling were performed using XDS [106]. Diffraction to 2.55 Å resolution was observed.

The crystal structure was determined by the molecular replacement method using the program PHASER [107]. The search model for this procedure was a homology model built by SWISS-MODEL [108] using as a template MtmOIV [109], a monooxygenase from the mithramycin biosynthetic pathway, PDB entry 3FMW chain C with 47% identity to OxyS from *S. rimosus*. PHENIX AutoBuild wizard was used to improve the model [110]. Further refinement was performed using REFMAC5 [111] and Buster/TNT [112] with TLS parameterization of domain disorder [113]. After each refinement step, the models were visually inspected in COOT

[114], using both 2Fo-Fc and Fo-Fc difference maps. The models were validated with the following structure validation tools: PROCHECK [115], ERRAT [116], and VERIFY3D [117]. The structure contains 90.2% of the residues within the most favored region of the Ramachandran plot, 9.1% were in additional allowed regions, and 0.7% were in generously allowed regions. There were no residues in disallowed regions. The Errat score, 94.2%, indicates that this percentage of residues fall below the 95% confidence limit of being erroneously modeled. Verify3D reports 99.4% of the residues have an averaged 3D-1D score greater than 0.2. Data collection and refinement statistics are reported in Table 1. The coordinates of the final models and structure factors have been deposited in the Protein Data Bank with PDB code 4k2x. The structures were illustrated using Pymol (PyMOL). Coordinates were superimposed using the secondary structure matching algorithm, SSM, implemented in COOT [118]. The structure based sequence alignment in Figure 17 was performed with the programs MAPS [119] and, STACCATO [120] and displayed with the program ALINE [121].

2.2. Heterologous expression and manipulation of three tetracycline biosynthetic pathways.

This section contains material which was published as: Wang, P., Kim. W., Pickens, L. B., Gao, X., Tang, Y. "Heterologous Expression and Manipulation of Three Tetracycline Biosynthetic Pathways." *Angew. Intl. Chem. Ed*. **2012**, 51, 11136–11140.

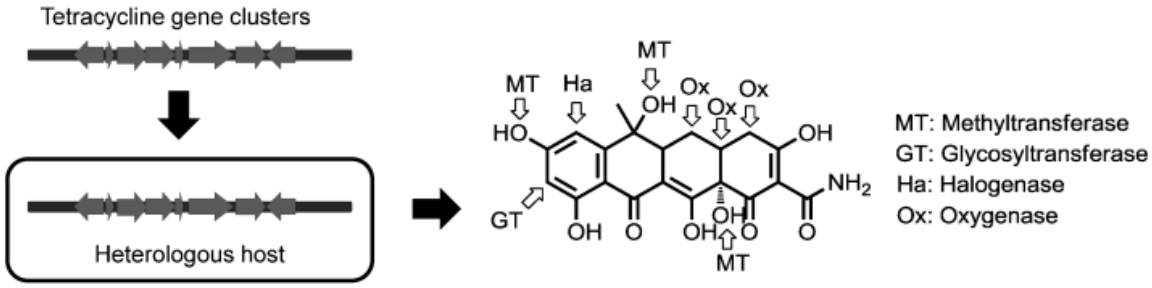


Figure 23. Three tetracycline biosynthetic pathways were overexpressed and manipulated in the heterologous host *S. lividans* K4-114. Through the inactivation of various genes and characterization of the resulting biosynthetic intermediates, new tetracycline-modifying enzymes were identified.

2.2.1 Heterologous expression of oxytetracycline biosynthesis

Our first objective is to establish a heterologous host for reconstitution of the *oxy* cluster and biosynthesis of **8**, which will circumvent genetic manipulation in its native *S. rimosus* [122]. While systematic build-up of the oxy genes in *S. coelicolor* CH999 has provided key insights into the earlier steps of the pathway [66, 70], biosynthesis of **8** using this approach has not been accomplished due to the incomplete understanding of the last few steps of the pathway, namely from anhydrotetracycline **11** to **8**. Despite previous reports of production of **8** from *S. lividans* [80] and *M. xanthus* [88], we did not detect the biosynthesis of **8** in the model heterologous host *S. lividans* K4-114 [35] upon insertion of the oxy cluster (oxyA-T, 25 kb) into its chromosome

using the integrative vector pOTC. To enhance the transcription of the oxy pathway, a *Streptomyces* antibiotic regulatory protein (SARP), encoded by *ctc11* from the chlorotetracycline gene cluster [65], was placed under the constitutive *ermE** promoter [123-124] in pCTC11 and co-transformed into K4-114 with pOTC, resulting in the strain K4/pOTC-Ctc11 (Table S1). We chose *ctc11* because no corresponding SARP is found in the vicinity of the oxy cluster. HPLC-MS analysis showed the production of **8** as a predominant product at ~20 mg/L (Figure 25C). Using pOTC as the template, the λ-Red-mediated recombination method [104, 125] was employed to delete *oxyS* encoding the anhydrotetracycline oxygenase (Scheme 1A). Cotransformation of the resulting plasmid with pCTC11 into K4-114 led to the accumulation of **11** and disappearance of **8** in K4/pSKO-Ctc11 (Figure 25D). While OxyS has been linked to the C6 hydroxylation of **11** to yield 5a, 11a-dehydrotetracycline **12** [79, 81], previous knockout studies in *S. rimosus* did not lead to the isolation of **11**, presumably due to off-pathway modifications by endogenous enzymes [75]. Therefore, the availability of the "clean" heterologous host will enable the complete understanding of the enzymology between **11** and **8**, a critical requirement for generating tetracycline-like analogs.

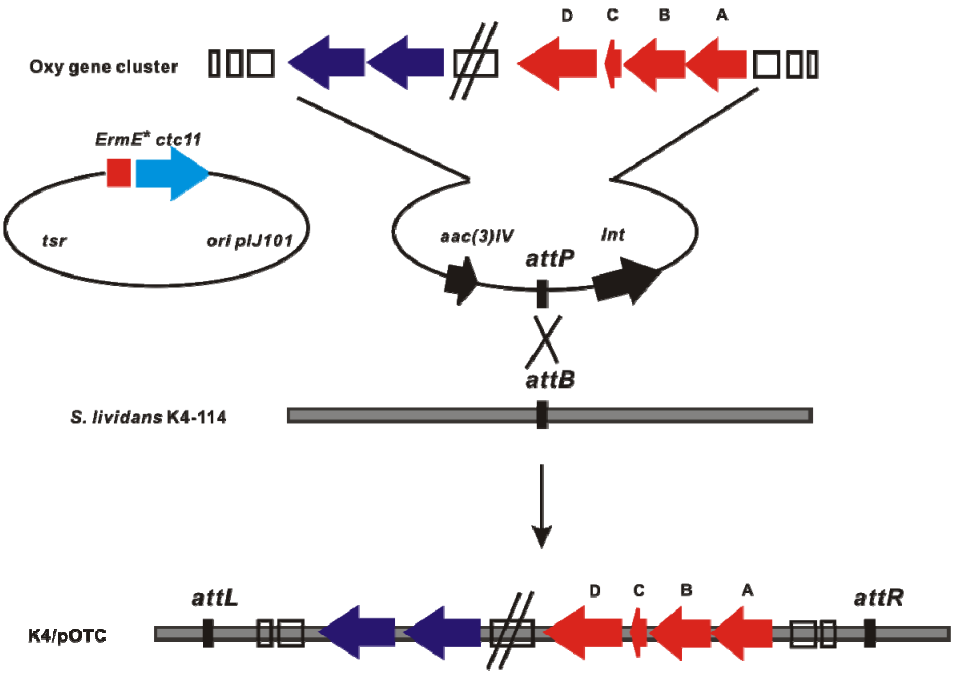


Figure 24. Scheme representation of heterologous expression of *oxy* biosynthetic pathway in *S. lividans* K4

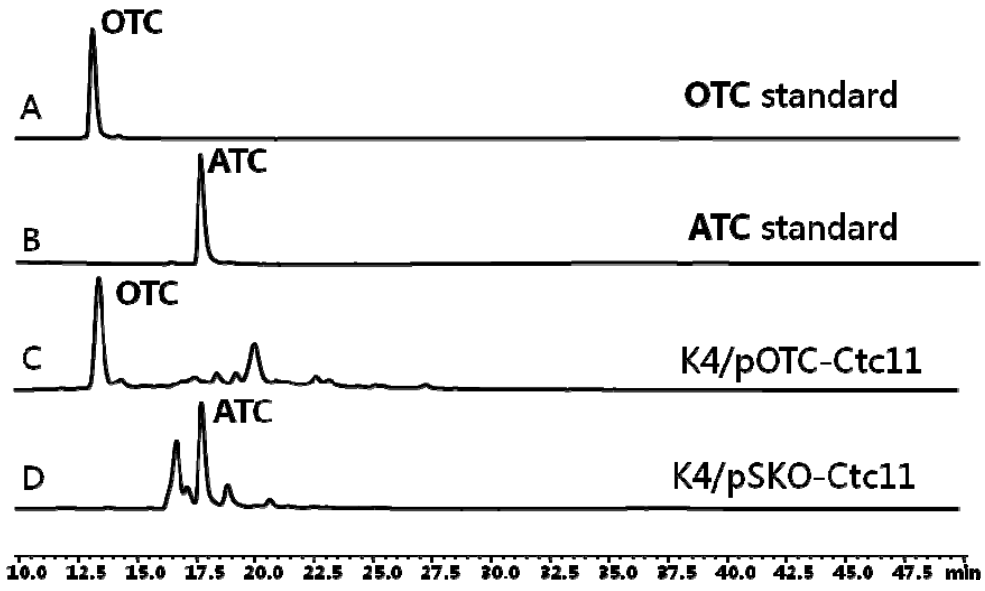


Figure 25. HPLC analysis of *oxy* heterologous expression in *S. lividans* K4. A, C) λ = 358 nm; B, D) λ = 430 nm;

2.2.2 Identification of tailoring enzymes involved in SF2575 biosynthesis

2.2.2.1 Introduction

The tetracycline family is a group of polyketide natural products produced by soil-borne bacteria [82]. Prime examples include chlortetracycline **10** and oxytetracycline **8** produced by *S. aureofaciens* [126] and *S. rimosus* [58], respectively. They were widely used as first-line antibiotics due to their broad-spectrum antibiotic properties and low toxicity in animals [127]. However, widespread production and usage in human and animals led to the rapid development drug-resistance mechanism and decreased the effectiveness of tetracyclines as clinical antibiotic medicines [128-129]. To combat with these tetracycline resistance pathogens, new generation tetracyclines that have been developed, include doxycycline [130], minocycline [131], and tigecycline [132-134] by modifications of tetracycline scaffold using a semi-synthetic strategy. Moreover total synthesis of tetracyclines was developed by Myers and coworkers recently [135-136]. On the other hand, however, only a few natural tetracyclines that have been isolated, since **10** and **8** were discovered decades ago, include SF2575 **16** [59, 137] and dactylocycline [138] (Figure 31 and 35). Among these, SF2575, showing exceptionally potent anticancer rather than antibiotic activities [59, 137], was isolated from *S.* sp. SF2575. The closely SF2575 relative compounds TAN-1518A and TAN-1518B were reported as DNA topoisomerase I inhibiting reagents [139] indicating SF2575 might act on the same molecular target. This novel action mechanism likely attributes to the unique modifications on tetracycline in SF2575 structure including two methoxy groups at C-6 and C-12a, a 4-(*R*)-salicylate, a C-9 C-glycoside of 2-6-dideoxy-arabino-hexopyranose (D-olivose), and acylation of the O-4' of D-olivose with (Z)-2-methyl-2-butenoate (angelate). These promising novel biological activities inspire the structure-

activity relationship (SAR) studies on tetracycline and imply the tetracycline is an "evolutionarily privileged" scaffold for further drug discovery.

In order to understand the biosynthetic logic and provide enzymatic tools to generate new tetracycline analogs, biosynthetic machinery of tetracycline was intensively investigated [68]. Tetracycline biosynthesis is initiated by an extended minimal polyketide synthase (PKS) consisting of a ketosynthase (KS), a chain-length factor (CLF), an acyl carrier protein (ACP), an amidotransferase and a proof editing thiolase to afford amidated decaketide backbone [66, 71]. Dedicated cyclases and C-9* reductase were employed to buckle the highly reactive poly-beta keto chain in a specific fashion to form the linearly fused four ring compound pretetramide **18** [140-141], followed by C-6 methylation and double hydroxylation to reach the last common intermediate 4-keto-anhydrotetracycline **19** [70, 72]. Previously studied tetracycline biosynthesis indicated oxytetracycline and SF2575 share very similar biosynthetic pathway, differing from 1**9** [67]. After this point, a significant number of steps remain to be resolved, including two methylations at C-6-OH and C-12a-OH, 4-(*R*) reduction, C-9 glycosylation, C-6 hydroxylation and C5a-C11a reduction. Fully understanding of the mechanism behind these tailoring reactions will not only provide an access to generate analogs with improved pharmaceutical properties but also reveal the natural strategy to transform the rigid four ring compound **18** to more dynamic penultimate product **20**. For instance, tigecycline demonstrates the potential benefits of modifying at C-9 position, therefore, identification and characterization of C-9 glycosyltransferase might provide an enzymatic tool to access C-9 position. On the other hand, for example, doxycycline, a 6-deoxy analogue of tetracycline bears enhanced stability toward both acid and basic degradation [142] implying C-6-OH methylation might offer similar stabilizing effects on tetracycline scaffold during the serial chemical transformations. Although

the genes encoding the corresponding tailoring enzymes were assigned in *ssf* gene cluster based on bioinformatic analyses [67], only functions of SsfX3 and SsfL were identified via in vitro experiments followed by protein structural characterization of SsfX3 [143]. Therefore, the lack of experimental evidences of tailoring enzymes and the ambiguousness of reaction sequence impedes the attempts to further elucidate the mechanism of biochemical reactions and the generation of new tetracycline analogs through combinatorial biosynthesis and metabolite engineering strategies.

To this end, here we report the characterization of serial reactions, diverging SF2575 biosynthesis from tetracycline biosynthesis from **19,** by identifying new tetracycline tailoring enzymes via combination of biochemical and genetic investigation in heterologous host *S. lividans* K4-114. Structural elucidations of tetracycline analogs obtained from gene inactivation mutants not only verify the functions of individual enzymes but also reveal serial dedicated reactions which transform a rigid structure **18** to more dynamic tetracycline product **20**. The concerted actions of 4-ketone reductase SsfF and methyltransferase SsfM1 at A ring as well as the reactions catalyzed by a flavin-dependent oxygenase SsfO1, and a methyltransferase SsfM2 at C ring were assessed via the combination of biochemical and genetic investigation, providing insights into the natural strategy to stabilize A ring and to convert an aromatic C-ring to more dynamic structure with different specific (*R*) stereochemistry at C-6. Most interestingly, the redundancy of SsfM3 demonstrates an evolutionary event in which combinatorial biosynthesis was employed by nature to morph the tetracycline structure.

2.2.2.2 Results

Expression of *ssf* gene cluster in heterologous host *S. lividans* K4-114

Heterologous expressions of second metabolite gene cluster have been used for a number of natural products including fredericamycin[123], novobiocin [144], and polyoxin [145]. Although we established genetic manipulation of *S*. sp SF2575 to perform gene inactivation in native strain, the low frequency of homologous recombination and consequently time-consuming process hindered the progress of the investigation. In order to circumvent genetic hindrance of *S.* sp SF2575, a heterologous platform for SF2575 biosynthesis was set up in *S. lividans* K4-114 via constructing a shuttle vector pSF2575 that contains the entire *ssf* gene cluster. The whole *ssf* gene cluster was introduced and integrated into the chromosome of *S. lividans* K4-114 to generate K4/pSF2575. HPLC-MS analysis of K4/pSF2575 revealed that K4/pSF2575 produced **16** and two major intermediates compound **20** and compound **21** (Figure 26c).

To improve the titer of SF2575 in *S. lividans* K4-114, a *Streptomyces* antibiotic regulatory protein (SARP) family regulator [146-147] encoding gene *ssfT1*, was constructed into a high-copy number shuttle plasmid to generate pT1, under the control of the constitutive promoter *ermE** [148]. The pT1 and pSF2575 was co-transformed into *S. lividans* K4-114 to afford K4/pSF2575-T1. Subsequent HPLC-MS analysis of K4/ pSF2575-T1 showed the titers of SF2575 and the other two intermediates increased (Figure 26d). To verify the role of SsfT1 in the titer improvement of SF2575 and intermediates, *ssfT1* was removed, using in-frame deletion, from pSF2575 via RED mediated gene inactivation procedure to generate pT1KO. The pT1KO was transformed into *S. lividans* K4-114 to afford K4/T1KO. As expected, the production of SF2575 and intermediates was abolished (Figure 26b). To maintain the high yield of SF2575

and its intermediates, we performed gene inactivations in the background of SsfT1 overexpression.

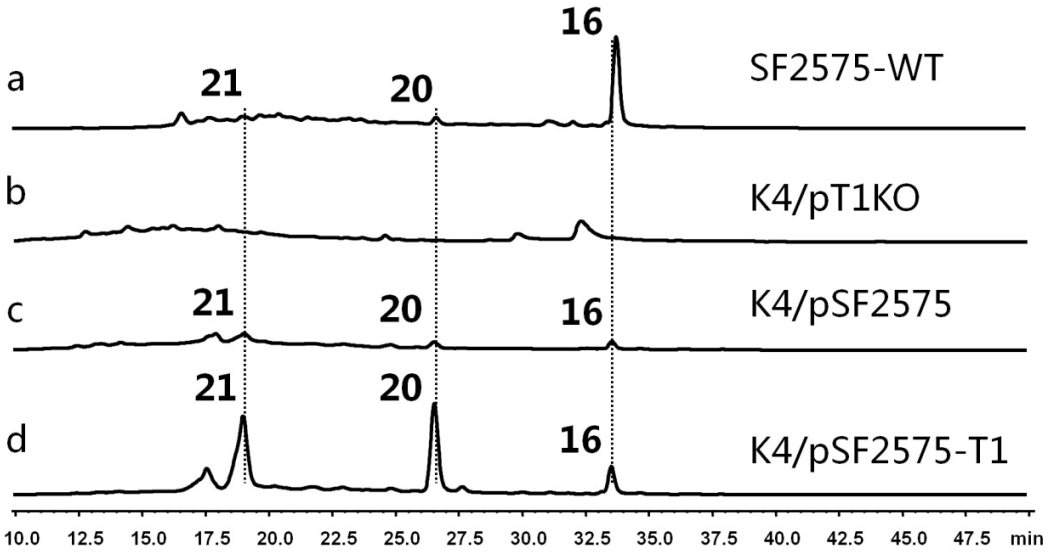


Figure 26. Heterologous expression of *ssf* gene cluster in *S. lividans* K4.

Functional identifications of four methyltransferases

In the biosynthetic gene cluster of SF2575, four methyltransferase genes *ssfM1-ssfM4* were identified through protein amino acid sequence comparison [67]. On the tetracycline aglycone, three positions require methyl addition activities: C-6, C-12a, and C-6-OH groups. SsfM4 was confirmed as C-6 methyltransferase in previous studies. Based on bioinformatic analyses, SsfM1 and SsfM2 bear strong assembly towards O-methyltransferase, while SsfM3 shows similarity with both C-methyl and O-methyltransferase. In mithramycin biosynthesis, a SsfM1 homologue MtmMI (48% identity) was reported to catalyze methylation at C-4-OH of 4-demethyl-premithramycinone A ring to afford premithramycinone [149]. Since C-4-OH of SF2575 is occupied by salicylic acid, the only open position at A ring is C-12a-OH. Therefore

we proposed the SsfM1 is more likely functions as C-12a-OH methyltransferase. A SsfM2 homologue NcsB1 (43% identity) was characterized to convert 2,7-dihydroxy-5-methyl naphthoic acid to 7-methoxy-2-hydroxy-5-methyl naphthoic acid indicating ssfM2 is an O-methyltransferase, moreover, ssfM2 is located right downstream of SsfO1, implying they are co-transcribed in a same operon. Therefore SsfM2 was proposed as C-6-OH methyltransferase. Interestingly, SsfM3 shows highly homology (39% identity) with MtmMII in mithramycin biosynthesis [149]. MtmMII was characterized as C-9 methyltransferase to transform premithramycin A3' to premithramycin A3 [149]. However, SF2575 structure shows C-9 position is occupied by D-olivose implying the redundancy of SsfM3 in SF2575 biosynthesis. We proposed SsfM3 was inactivated during the evolutionary process. To support our hypothesis, we performed sequence alignment of SsfM3 with other five homologues in which NcsB1 was structural characterized [150] (Figure 33). As expected, we found two amino acids mutants occurs in SsfM2 in highly conserved region which are Phe249 and Leu250 compared to the other methyltransferases. The counterpart of these two amino acids are His and Asp/His which are confirmed as catalytic residues to deprotonate phenol hydroxyl residue and initiate a nucleophilic attack to *S*-adenosyl methionine and sequentially form a O-C bond [151]. Based on the bioinformatic analyses, we proposed that SsfM3 used to work as C-9 methyltransferase and was inactivated during the evolutionary process. Therefore, as for SF2575 biosynthesis, SsfM3 is redundant. In order to prove the hypothesis of SsfM3, we performed gene inactivation of ssfM3 on pSF2575 and generate mutant strain K4/pM3KO-T1. As expected, K4/pM3KO-T1 did not abolish the production of SF2575 confirming the redundancy of SsfM3 in SF2575 biosynthesis (Figure 30b).

To confirm the involvements of the other three methyltransferases, we performed gene inactivations of *ssfM1*, *ssfM2*, and *ssfM4* on pSF2575 and generate mutant strains K4/pM1KO-T1, K4/pM2KO-T1, and K4/pM4KO-T1 respectively as described in method section. HPLC-MS analysis of organic extract from K4/pM1KO-T1 R5 agar revealed the production of three major products **22**, **23**, **24** with mass (*m/z* 764, $[M + H]^+$; *m/z* 682, $[M + H]^+$; and *m/z* 562, $[M + H]^+$) (Figure 30c). **22** was isolated and structurally elucidated via full NMR (^1H, ^{13}C, COSY, HMQC, and HMBC) analyses (Table S3). Comparing with **16**, signals of C-12a sp3 methylene carbon [δ_H 3.51 s, δ_C 54.4] were missing; consistent with this, the C-12a was found to shift 6.0 ppm upfield (Table S3 and Figure S3), indicating the replacement of methoxyl group by hydroxyl group. Therefore, **22** was established as TAN-1518A. Consequently, products **23** and **24** were identified as C-18-OH-deangelate-TAN-1518A and C-4-OH-desalicyclic-C-18-OH-deangelate-TAN-1518A respectively based on Mass spectrum. HPLC-MS analyses of metabolites from K4/pM2KO-T1 R5 agar showed 3 major intermediates **25**, **26**, and **27** with 14 lower mass (*m/z* 764, $[M + H]^+$; *m/z* 682, $[M + H]^+$; and *m/z* 562, $[M + H]^+$) compared to **16**, **20** and **21** respectively (Figure 30d). The compound **25** was isolated and analyzed by NMR (^1H, ^{13}C, COSY, HMQC, and HMBC) (Table S4 and Figure S4). Comparing with **16**, signals of C6-OH sp3 methylene carbon [δ_H 3.81 s, δ_C 50.5] were missing, indicating no methylation occurred at the C-6 hydroxyl position. Therefore the compound **25** was established as C-6-OH-demethyl-SF2575. Consequently, compound **26** and **27** were identified as C-6-OH-demethyl-18-OH-deangelate-SF2575 and C-6-OH-demethyl-C-4-OH-desalicyclic-C-18-OH-deangelate-SF2575 respectively based on mass spectrum. The characterization of compound **25** structure confirms ssfM2 servers as C-6-OH O-methyltransferase in *ssf* biosynthesis. In addition, complementation of *ssfM1* and *ssfM2* in K4/pM1KO-T1 and K4/pM2KO-T1 restored the production of final

SF2575 respectively, further confirming the roles of SsfM1 and SsfM2 in *ssf* biosynthesis (Figure 32). Therefore, gene inactivation and complementation results clearly showed that SsfM1 is responsible for C-12a-OH O-methylation while SsfM2 serves as C-6-OH O-methyltransferase involved in *ssf* biosynthesis. Furthermore, the 6 intermediates compound **22**, **23**, **24**, **25**, **26** and **27** demonstrated the tolerance of post tailoring enzymes toward the demethyl-substrates either at C-6-OH or C-12-OH positions.

The first methyl transfer reaction in SF2575 pathway is the methylation of C-6 position of pretetramid to produce 6-methyl-pretetramid. Previously, SsfM4 was identified to catalyze this transformation by heterologous reconstitution [67]. To generate C-6-desmethyl analogs, *ssfM4* gene was inactivated on pSF2575 to afford mutant K4/pM4KO-T1. HPLC-MS analysis of organic extraction from K4/pM4KO-T1 R5 agar revealed two major compounds **28** and **29** were accumulated with mass (*m/z* 530, $[M + H]^+$, *m/z* 560, $[M + H]^+$) respectively (Figure 30e and Figure 31). Both compounds bear similar UV pattern with anhydrotetracycline (ATC) with λ_{max} at 430. The two compounds were then isolated and subjected to full NMR analyses (Tables S5, S6 and Figures S5, S6). Compared to **16**, signals of D-olivose 8 protons and 6 carbons are present in both compounds and further confirmed by COSY, HMQC, and HMBC (Tables S5, S6 and Figures S5, S6). Signals of C-14 sp3 methylene carbon [δ_H 0.96 s, δ_C 24.5] and C6-O-CH_3 sp3 methylene carbon [δ_H 3.13 s, δ_C 114.5] were missing; consistent with this, the emergence of one aromatic carbon [δ_H 7.14 s, δ_C 117.6] in compound **28** indicates that C6 methylation was eliminated due to the *ssfM4* inactivation. Surprisingly, as for compound **29**, compared to compound **28**, signals of a methoxy group emerged and the C-6 carbon was found to shift 66.1 ppm downfield. HMBC shows correlations between the methylene proton with an aromatic carbon with chemical shift 143.9 ppm (Table S6 and Figure S6) confirming the methoxy group

located at C-6 position. The structure of compound **29** indicates the unexpected methoxy group might be generated by downstream enzymes C-6-hydroxylase or air oxidation and the corresponding C-6-OH methyltransferase SsfM2.

Glycosyltransferase SsfS6 catalyzes C-9 glycosylation providing an anchor for further oxidation

Only one glycosyltransferase homolog, SsfS6, was found encoded in *ssf* gene cluster and shows highly homology to a C-glycosyltransferase, HedJ (40% identity), involved in the hedamycin biosynthesis [152] and UrdGT2 (33% identity), involved in urdamycin biosynthesis [153]. Although angucyclines are natural substrates of both of them, UrdGT2 showed relaxed substrate specificities towards both aglycone and activated sugar donors [154-155]. Especially, UrdGT2 is able to attach D-olivose to a tetracycline like aglycone premithramycinone at the C-9 position indicating SsfS6 might bear broad substrate specificities as same as UrdGT2. Previous studies implied that typically glycosyl transfer reactions occur towards the end of the biosynthetic pathway [156]. Therefore, following the biosynthetic logic of oxytetracycline, we proposed that the D-oliviose was transferred after the post modification steps of C-ring including C-6 methoxylation and C5a-C11a reduction [67]. Alternatively, as exampled by mithramycin biosynthesis, aglycone premithramycinone is fully sugar loaded before final oxidation and reduction to afford mithramycin [109]. To confirm the involvement of SsfS6 and find out the natural aglycone of SF2575, gene inactivation of *ssfS6* was performed as described above to generate mutant K4/S6KO-T1. A single new metabolite **30** was accumulated from the *ssfS6* mutant K4/S6KO-T1 along with the abolishment of SF2575 production. HPLC-MS analysis of organic extract from K4/S6KO-T1 R5 agar revealed the production of the compound **30** with

mass (*m/z* 414, $[M + H]^+$), and shares the same UV pattern with **11** (Figure 30g). The product of K4/S6KO-T1 was isolated, and characterized by full NMR analyses. Compared with **21** [137], signals of D-olivose, 8 protons and 6 carbons were missing along with emergence of one aromatic proton at C-9 position [δ_H 6.80 d, δ_C 110.8], indicating the lost of D-olivose at C-9 position (Table S7 and Figure S7). In addition, signals of C-6-OH sp3 methylene carbon [δ_H 3.13 s, δ_C 114.5] were missing, consistently the C-6 was found to shift 43.6 ppm downfield, indicating the lost of methoxyl group at C-6-OH position. Taken together with data of COSY, HMQC, and HMBC, the structure of compound **30** was established as C-12a-methoxy, C-4-hydroxyl-ATC. Therefore, these results confirmed the function of SsfS6 is C-9 glycosyltransferase with the aglycon C-12a-methoxy, C4-hydroxyl-ATC as substrate. This result was further confirmed by complementary experiment (Figure 32). The structure of **30** implies that the D-olivose attachment is essential for the successive oxidation steps, consistent with the biosynthesis of mithramycin. The D-olivose is possible to act as an anchor to orientate substrate approaching the activate site of downstream oxygenase.

Identification of redox enzymes involved in SF2575 biosynthesis

The first required step diverging from oxytetracycline biosynthesis is reduction of C-4-ketone to produce **31**. Three reductases are encoded in *ssf* gene cluster including SsfU, SsfK, and SsfF. Among them, only SsfU was proved as C-9* reductase during polyketide backbone biosynthesis via in vivo experiments as described in previous studies [67]. SsfK shows 59% similarity with SAV_2786 (ACM60257) with putative role as 3-ketoacyl-(ACP) reductase, while SsfF belongs to a NADPH-dependent oxidoreductase family with specificity towards aromatic and nonpolar substrates (PF00248). Gene inactivations of *ssfK* and *ssfF* were performed and

generate mutants K4/pKKO-T1 and K4/pFKO-T1 respectively [157]. The inactivation of *ssfK* largely decreased but not abolished the production of SF2575 ruling out the involvement of SsfK in the C-4-ketone reduction. However, HPLC-MS analysis of K4/pFKO-T1 organic extraction showed the abolishment of SF2575 production along with the accumulation of WJ135 **32**. Compound **32** is a shunt product degraded from the final common intermediate **19** between biosynthetic pathways of oxytetracycline and SF2575 (Figure 30h). This result demonstrates biosynthesis of SF2575 was blocked after SsfO2, a dioxygenase which catalyzes the C-4 and C-12a hydroxylation steps on 6-methyl-pretetramide to generate degraded shunt product **32**. Complementation of *ssfF* in K4/pFKO-T1 restored the production of SF2575 (Figure 32), confirming the involvement of SsfF. The reduction of **19** at C-4-ketone quenches the enolization between C-4 and C-4a of **19** at A-ring consequently enhance the stability of 4-hydroxyl-ATC **31** by avoiding retro-claisen like cleavage steps proposed in previous studies [70]. Therefore, we conclude that SsfF is responsible for C-4-ketone reduction, the key diverging step from oxytetracycline biosynthesis in SF2575 biosynthesis.

Two flavin-dependent oxygenases were found encoded in *ssf* gene cluster. One is SsfO2, a dioxygenase, was proved to generate **19** [67]. Another flavin-dependent oxygenase, SsfO1 shows 54% similarity to OxyS indicating that SsfO1 catalyzes C-6 hydroxylation. Compared to oxytetracycline, SF2575 bears (*R*) configuration at C-6 position. We proposed the intrinsic difference between SsfO1 and OxyS leads to different stereochemistry at C-6 of SF2575 and oxytetracycline. We performed the *ssfO1* inactivation on pSF2575 and generate mutant K4/pO1KO-T1. HPLC-MS analysis of organic extraction of K4/pO1KO-T1 showed accumulation of **33** at the titer of ~33 mg/L with mass (*m/z* 543, $[M + H]^+$) along with abolishment of SF2575 production (Figure 30i). The product of **33** was isolated, and confirmed

by full NMR analyses (Table S8 and Figure S8). Compared with **30**, signals of D-olivose, 8 protons and 6 carbons were found in spectrum of **33**, indicating the existence of D-olivose at C-9 position. As same as **30**, the lack of C6-OH sp3 methylene carbon [δ_H 3.13 s, δ_C 114.5] signals indicates the loss of methoxyl group at C-6 position. Taken together, the structure of **33** was established as C-4-hydroxyl-C-12a-O-methyl-C9-D-olivose-ATC. Thus, we identified the function of SsfO1 is C-6 hydroxylase with **33** as substrate. These results further confirmed that deoxysugar substituent is added prior to oxidation by SsfO1 indicating existence of the anchor effect [158]. Furthermore this result was confirmed by complementary experiment (Figure 32).

To further confirm the C-6-(*R*)-hydroxylation activity of SsfO1 in converting **33** to **34**, soluble and *holo*-SsfO1 was overexpressed from *E. coli* and purified. The bright yellow colour of the recombinant SsfO1 indicates the binding of co-enzyme FAD. The reaction containing SsfO1 (20 μM), **33** (0.8M) and NADPH (2 mM) was incubated at room 25 °C for 16 h. After this, the reaction mixture was extracted with 1 mL of organic solvent (99% ethyl acetate, 1% acetic acid). The organic phase was evaporated *in vacuo* and resuspended in 30 μL methanol followed by HPLC-MS analysis. New compound **34** was found with mass (*m/z* 560, $[M + H]^+$) and λ_{max} at 400 nm (Figure 27) sharing the similar UV pattern with DHTC [159]. Therefore, the structure of **34** was proposed as shown in scheme 2 (W. Kim and L. B. Pickens, unpublished data) [157]. Due to the lack of C5a-C11a reductase in the reaction, **33** was converted to **34** by SsfO1 at (*R*) configuration. By swapping SsfO1 with OxyS, we might be able to generate tetracycline analogs with different stereochemistry at C-6. Therefore, we further incubated **33** with OxyS in the same reaction as describe above, but no new compound was detected in HPLC-MS trace (Figure 27) (W. Kim and L. B. Pickens, unpublished data) [157]. These results showed that OxyS in oxytetracycline is not able to replace SsfO1.

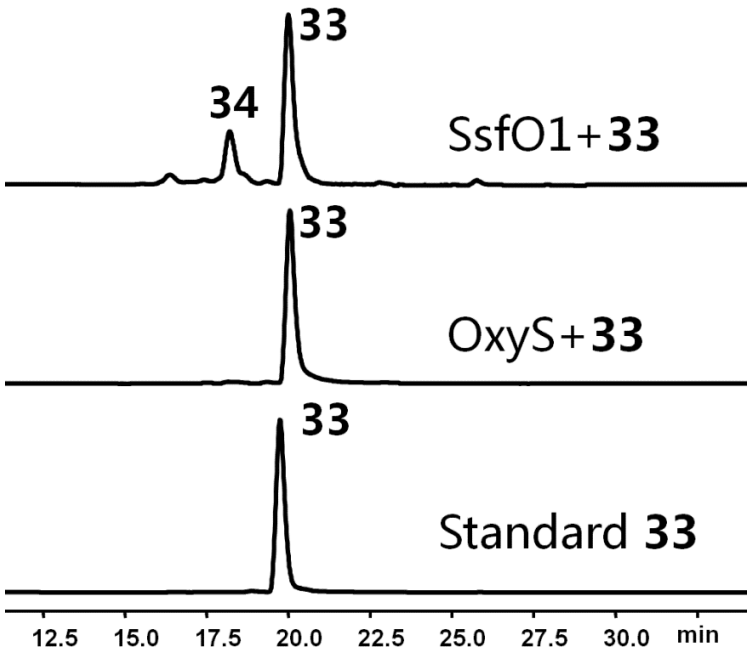


Figure 27. SsfO1 (L. B. Pickens) [157] and OxyS in vitro assay with substrate **33**.

The C5a-C11a reduction step involved in tetracycline family biosynthesis has been a longstanding missing link due to the lack of candidate in sequenced gene clusters including which of chlortetracycline and oxytetracycline. A flavin-dependent oxidoreductase TchA was proposed to catalyze this step [160]. However, because of the lack of direct experimental evidence, authors were not able to conclude that TchA catalyzes the final conversion. Recently, we uncovered the enzymatic basis of the final transformations in oxytetracycline biosynthesis (See discussion in Section 2) [161] . Unlike chlortetracycline and oxytetracycline biosynthetic gene cluster, a putative dehydrogenase encoding gene *ssfP* was found in *ssf* gene cluster [67]. Interestingly, SsfP shows no similarity with TchA, but shares 54% sequence resemblance to a putative F_{420}-dependent methylenetetrahydromethanopterin reductase. To confirm the involvement of SsfP, gene inactivation of *ssfP* was performed on pSF2575 to generate mutant strain K4/pPKO-T1 (L. B. Pickens) [157]. Surprisingly, the HPLC-MS analyses of K4/pPKO-

T1 revealed the presence of compound **36** with mass (m/z 574, $[M + H]^+$) and λ_{max} at ~400 nm (Figure 30j). The structure of compound **36** was proposed as shown in Figure 31 indicating the inactivation of SsfP stalls C5a-C11a reduction and the SsfM2 catalyzes C6-OH methylation right after SsfO1 hydroxylation to eventually generate **36**. To eliminate the activity of SsfM2, we performed three-gene inactivation by replacement of *ssfM2*, *ssfO1* and *ssfP* on pSF2575 (L. B. Pickens) [157]. As expected, HPLC-MS analyses of K4/pM2O1PKO-T1 showed the same profile of which in K4/pO1KO-T1. Sequentially we complemented SsfO1 into K4/pM2O1KO-T1 to restore the production of compound **34** which emerged in SsfO1 in vitro assay. The emergence of **34** in the complemented strain verified the function of SsfP as C5a-11a reductase. Further purification of compound **34** and **36** failed due to the compound degradation (L. B. Pickens) [157]. .

To understand the protein structural basis of SsfO1, we determined crystal structure of SsfO1 at resolution 2.8 Å with oxidized FAD cofactor (See Table 2 for statistics). Compared to OxyS structure, SsfO1 bears the same domains as OxyS (Figure 28) including FAD binding domain (residues 1-193 and 278-399), middle domain (residues 194-277), and C-terminal thioredoxin-like domain (residues 400-514). Structural alignment SsfO1 with OxyS shows RMSD of 0.974 Å through 352 alpha carbon pairs (Figure 29).

Table 2. Data collection and refinement statistics of SsfO1

	SsfO1
Data collection	
Space group	$P4_32_12$
Cell dimensions	
a, *b*, *c* (Å)	90.3, 90.3 195.0
α, β, γ (°)	90.0, 90.0, 90.0
protomers/asymmetric unit	1
Resolution (Å)	2.8
R_{merge}	0.072 (0.507)
$I/\ \ I$	32.7 (3.6)
Completeness (%)	99.7 (99.8)
Redundancy	17.7 (7.5)
Refinement	
Resolution (Å)	2.8
No. reflections	20,578
$R_{work/}\ R_{free}$	0.196/0.240
No. atoms	
Protein	3746
Ligand/ion	53
Water	40
B-factors (Å^2)	
Protein	96.1
Ligand/ion	63.5
Water	68.4
R.m.s deviations	
Bond lengths (Å)	0.010
Bond angles (°)	1.2

*Highest resolution shell is shown in parenthesis.

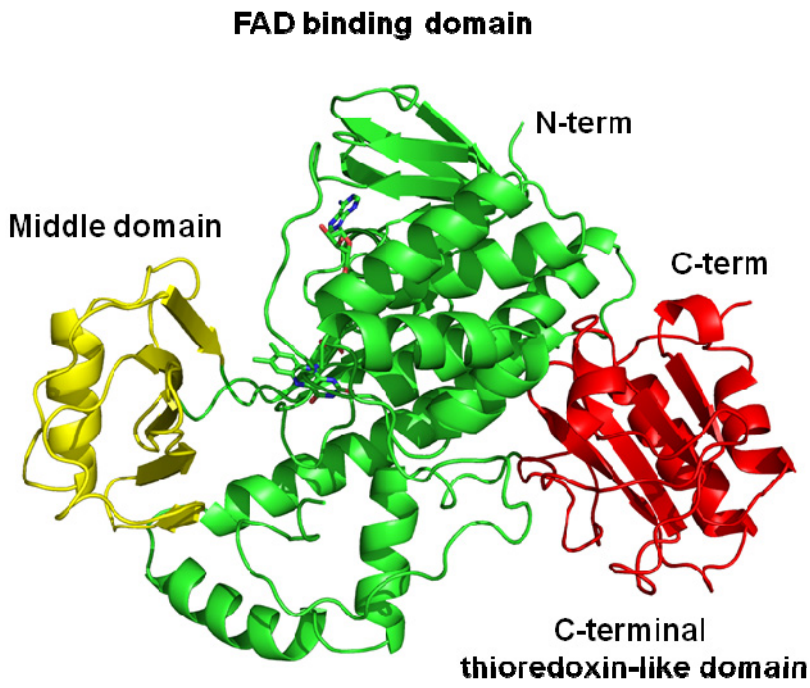


Figure 28. Overall structure of SsfO1 shows three domains including FAD binding domain, Middle domain, and C-terminal thioredoxin-like domain.

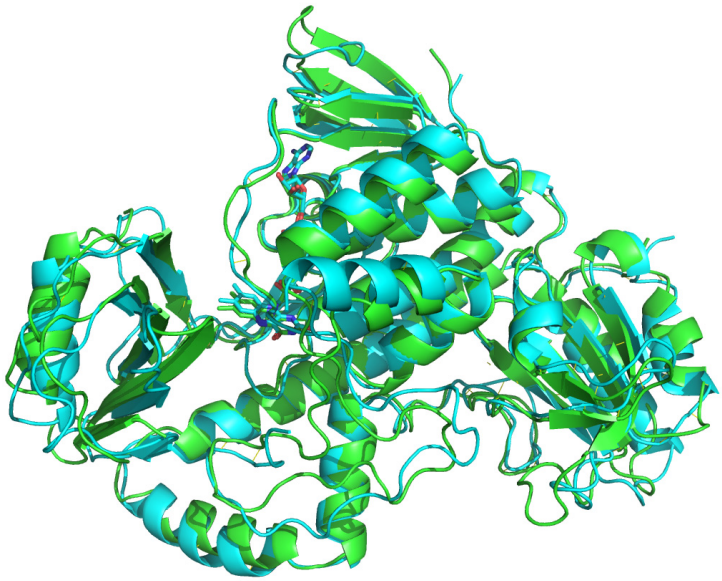


Figure 29. Structural alignment of SsfO1 (blue) with OxyS (green).

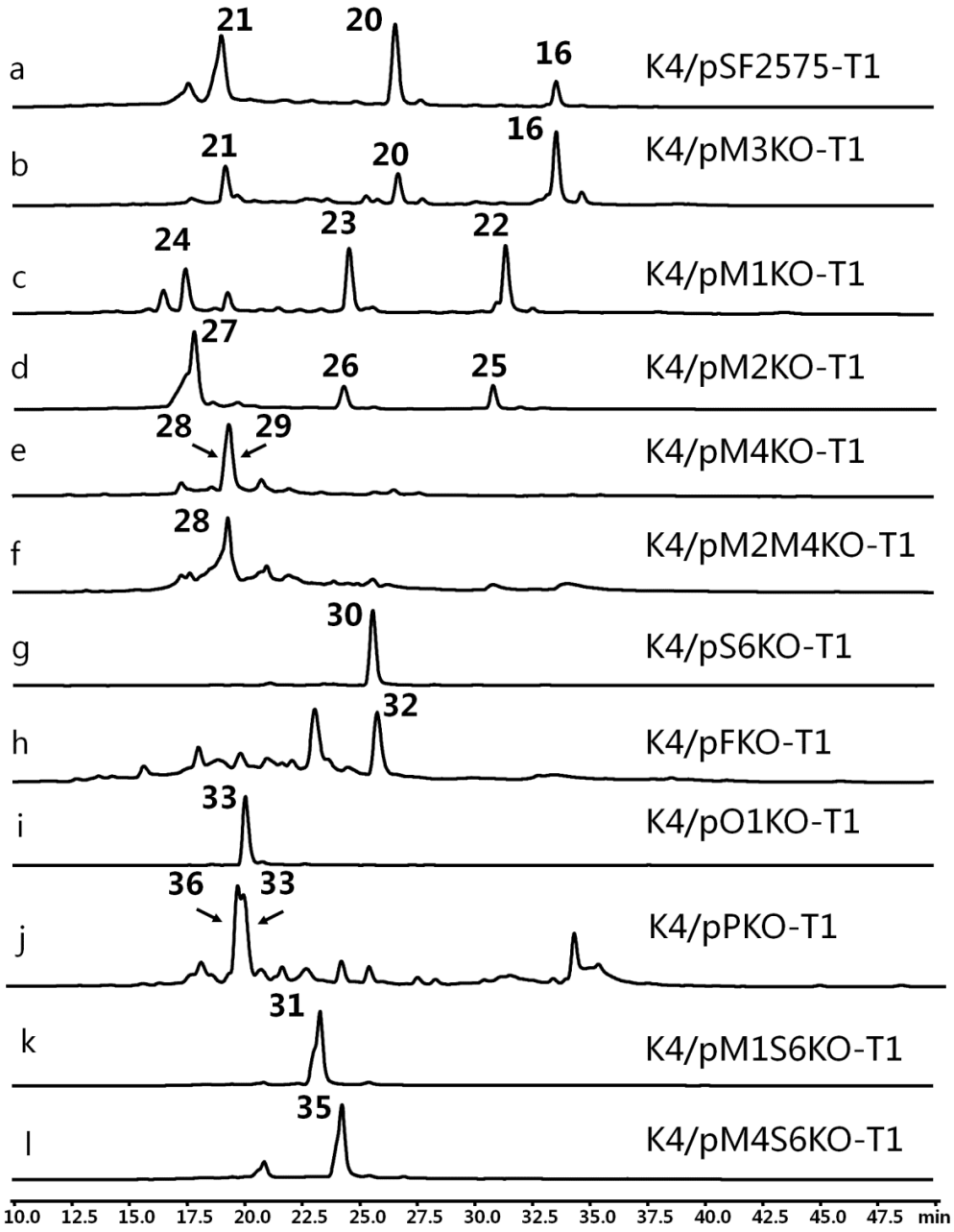


Figure 30. HPLC profiles of extracts from cultures of K4-114 integrated with different knockouts of the *ssf* pathways. a-d) λ=358 nm; e-g, i, k, l) λ=430 nm; h, j) λ=400 nm.

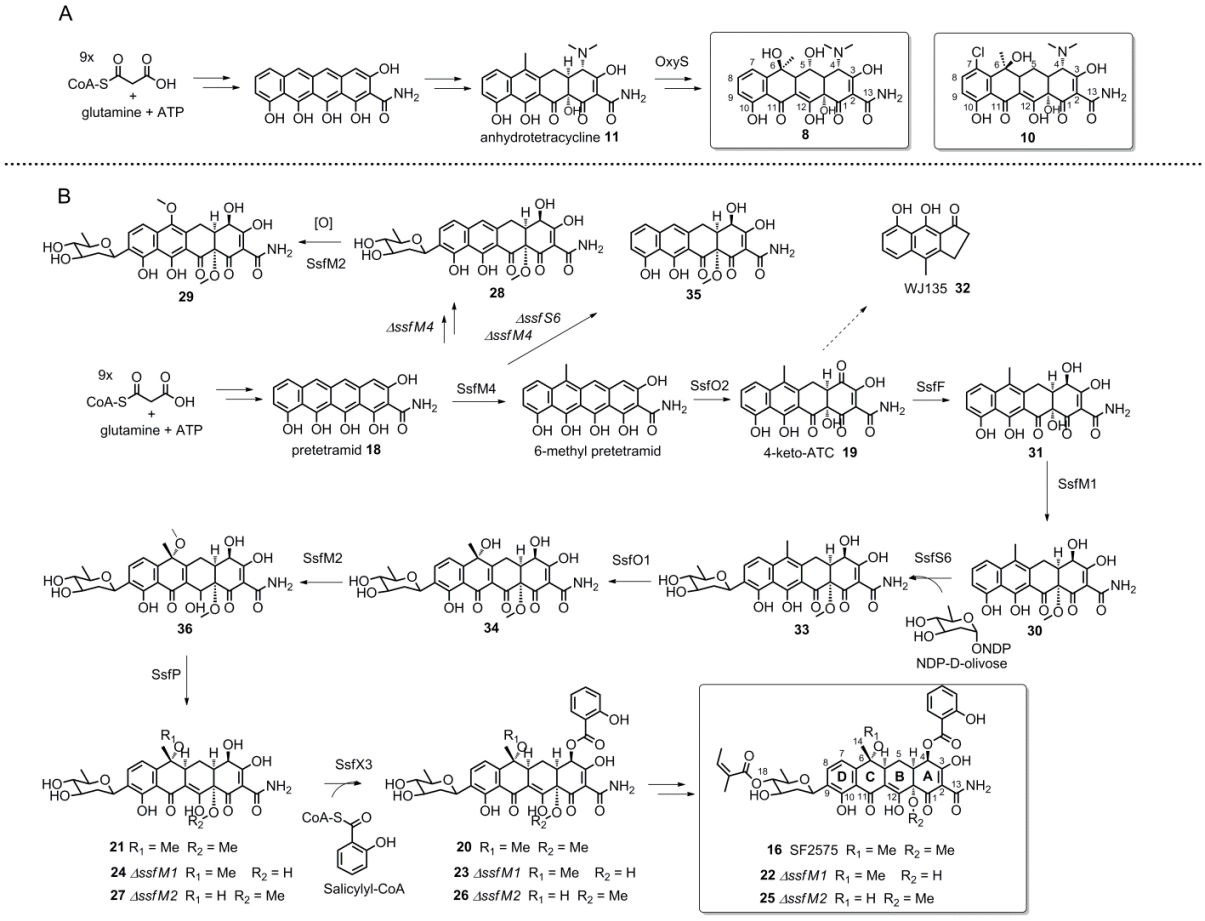


Figure 31. Biosynthetic pathways of A) oxytetracycline; B) SF2575. ATC = anhydrotetracycline, NDP = nucleotide diphosphate

Generation of tetracycline analogs by using double gene inactivation

The single gene inactivation experiments established the biosynthetic sequence and the functions of enzymes which are responsible for the transformations from 4-keto-ATC **19** to the penultimate product **20**. We proposed the 4-ketone reductase SsfF stabilizes the **19** by dearomatizing A ring to generate relatively stable intermediate **31**. To prove this hypothesis, we performed *ssfM1* inactivation on *ssfS6* mutant plasmid pS6KO and generate K4/pS6M1KO-T1. The structure of compound **31** in organic extraction of K4/pS6M1KO-T1 was confirmed by full NMR analyses (1H, 13C, COSY, HMQC, and HMBC) (Table S9 and Figure S9). The emergence of **31** supports the hypothesis that reduce C-4-ketone of **19** is a natural strategy to maintain the carbon skeleton of SF2575.

SsfM4 works as C-6 methyltransferase to convert pretetramide **18** to 6-methyl **18** by adding a methyl group at a reactive C-6 site. In viridicatumtoxin biosynthesis, C-6 position is modified by a geranyl group [162]. To open this position, we performed *ssfM4* inactivation on pS6KO to generate K4/pS6M4KO-T1, the compound **35** was purified and confirmed by full NMR analyses (1H, 13C, COSY, HMQC, and HMBC). The further modification at C-6 position to generate new tetracycline like compounds has been reported [163].

Bioactivity and TetR induced luminescence assays of tetracycline analogs

Identification of a wealth of new tailoring enzymes provides an access to the biosynthesis of novel tetracycline compounds. The relationship of compounds and bioactivities can shed light on the features which are important to potent bioactivities of either as antibiotic or anticancer agents. To evaluate antibiotic activities of these tetracycline analogues, we measured the minimum inhibitory concentration (MIC) values against four strains, including Gram-positive

strain: *Bacillus subtilis*; Gram-negative strains: *Salmonella enterica*, *Pseudomonas fluorescens Pf0-1*, and *E. coli* DH10B (Table 1). MIC values showed that ATC-like compounds: **31**, **30**, and **35** preserved most of antimicrobial activities, while only sugar attached compounds dramatically increased the MIC values towards both Gram-positive and Gram-negative strains, implying the D-olivose sugar moiety strongly depletes the antibiotic activities of ATC scaffold. Interestingly, salicylic acid group and angelic acid group partially restored the antimicrobial activities towards *Bacillus subtilis*, indicating the further modifications at C-4 and C-9 are essential for antibiotic activities. TetR induced luminescence assays were performed with tetracycline as positive control. Surprisingly all of the tetracycline analogs did not show any induction in TetR induced luminescence assays indicating intrinsic difference between TetR and SsfR.

Table 3. Minimum inhibitory concentration (MIC) values of compounds in this study

μg/ml	*Salmonella enterica*	*Bacillus subtilis*	*Pseudomonas fluorescens* Pf0-1	*E. coli* DH10B
31	16	1	8	8
30	32	0.25	32	16
33	>128	64	>128	>128
25	>128	1	>128	>128
22	>128	4	>128	>128
23	>128	16	>128	>128
28	>128	>128	>128	>128
29	>128	>128	>128	>128
35	8	0.5	8	16
11	8	0.5	8	8
8	0.5	0.5	0.5	1

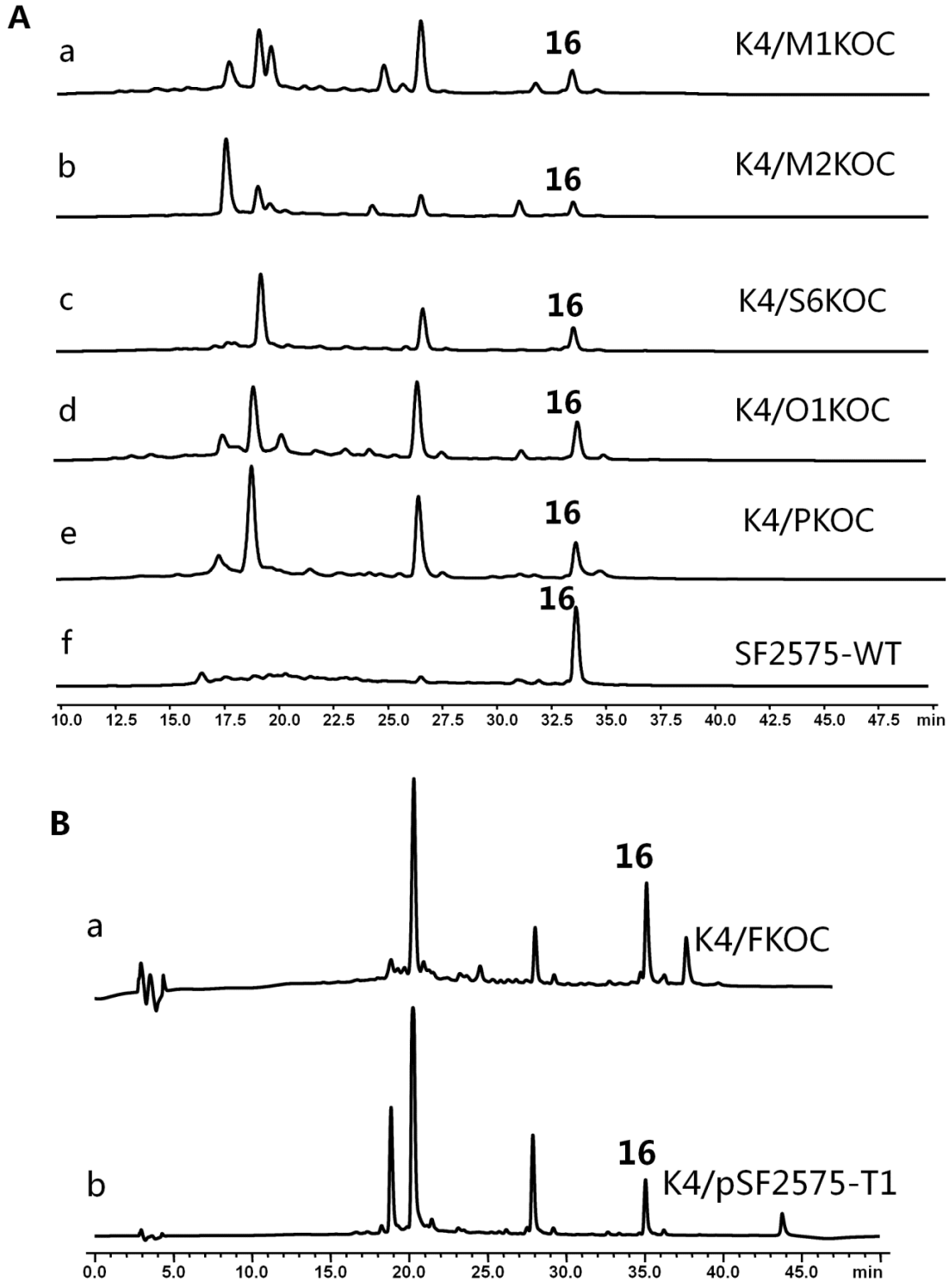


Figure 32. HPLC analyses of *ssf* gene complementary mutants.

2.2.2.3 Discussion

SF2575 belongs to naturally occurring polyketides, which comprise a class of natural products with diverse biological activities [10, 164-165]. Tetracyclines, a group of polyketides, were recognized as broad-spectrum antibiotics against both Gram-positive and Gram-negative microorganisms [76]. In contrast to antibiotic activities, the unexpected anticancer activities of SF2575 demonstrate that the structural diversity of tetracyclines is able to lead to emerge novel biological activities. In SF2575 biosynthetic pathway, a number of post-tailoring enzymes catalyze the transformations from a structurally rigid precursor pretetramid **18** to a structurally dynamic final product SF2575. The cascade transformations delineate an efficient biosynthetic route to expand chemical space of tetracycline scaffold. Therefore, identifications of tetracycline tailoring enzymes will provide a set of enzymatic tools to access the multiple positions around tetracycline four-ring structure. To this end, we elucidated seven key tailoring biosynthetic steps and established SF2575 biosynthetic pathway starting from 4-keto-ATC **19**, the last intermediate diverge from oxytetracycline biosynthesis, to the penultimate product **20** in the *ssf* heterologously transplanted *S. lividans* K4-114 strain.

Due to the difficulties and time consuming of genetic manipulation in SF2575 native producer *S.* sp 2575, the heterologous host *S. lividans* K4-114 was employed to overcome this barrier. Co-expression *ssf* gene cluster with pathway specific regulator SsfT1 in *S. lividans* K4-114 led to successful production of **16** and two key intermediates with a combined titer of ~127 mg/L, indicating the SsfT1 improved the transcription of *ssf* gene cluster. Given the good titer of SF2575 production in *S. lividans* K4-114 and the ease of genetic manipulation in *E. coli*, we performed gene inactivation and complementation using this heterologous system.

Methylations of natural products have significant effects on either compound biological activities or biochemical transformations during biosynthesis [166-168]. By gene inactivation of SsfM1 in heterologous host, production of **22** unequivocally confirmed SsfM1 as C-12a-O-methyltransferase. Interestingly, the appearance of **23** and **24** demonstrates the timing of C-12a-O-methlation acted right after C-4-ketone reduction, the following tailoring enzymes are able to act on the demethylated substrates all the way to **22**. We proposed that the C-12a-OH methylation might enhance the stability of **19** by replacing hydrogen with methyl group to extinguish the rearrangement in A ring. However, the existence of **31** accumulated in double knock out strain K4/pSFS6M1 indicated the step might not be essential for this chemical transformation. Alternatively C-12a-OH methylation might work as a protection step to avoid redundant modification in native host *S.* sp SF2575 such as glucuronic acid modification in *S. rimosus* [72]. On the other hand, SsfM2 inactivation led to accumulate **25**, **26**, and **27**, which are consistent with SsfM1 inactivation results. The timing of C-6-O-methylation is right after C-6-hydroxylation also supported by the emergence of compound **36** in *ssfP* knock out strain. The methylation at C-6-OH might stabilize **36** to avoid the rearrangement in C-ring which might lead to rapid degradation of **34** when **34** is released from SsfO1 catalytic pocket. However, the production of **25**, **26**, and **27** shows that the absence of SsfM2 is not critical for C5a-C11a reduction indicating the possible protection function of SsfM2 the same as SsfM1. Previous heterologous expression results confirmed the function of SsfM4 as C-6 methyltransferase. Gene *ssfM4* inactivation in *S. lividan*s K4-114 led to accumulate compound **28** and compound **29**. The structure of **28** shows glycosyltransferase is able to recognize C6-demethyl-4-hydroxy-ATC as a substrate. However, differing to O-methyltransferase inactivations, the following tailoring steps were stalled at **28** mostly due to the incompatibility between SsfO1 and **28**. In addition, the

unexpected **29** presents an unusual C-6-O-methylation resulting from C-6 oxidization followed by SsfM2 methylation. Due to the highly reactive nature of C6 at C ring, C-6 oxidization might result from air oxidation or catalyzed by SsfO1. The lack of methyl group at C-6 position maintains the aromaticity of C-ring sequentially abolishes the recognition of successive tailoring enzymes. Since three methylation steps were confirmed to be catalyzed by SsfM1, SsfM2 and SsfM4 respectively, leaving SsfM3 as a redundant enzyme. Interestingly SsfM3 shows 39% identity with MtmMII which catalyzes methyl transfer from SAM to premithramycin A3' at C-9 position [149] indicating SsfM3 catalyzes methylation at C-9 of SF2575 intermediate, the same spot which the glycosyltransferase SsfS6 acts on. We proposed the competition between SsfM3 and SsfS6 forces bacteria to abandon SsfM3 by inactivating SsfM3 during the evolutionary process. To support this hypothesis, we performed alignment SsfM3 with other five methyltransferases [150] (Figure 33). Two amino acids mutants were found in highly conserved region. Compared to NcsB1, catalytic diad His246 and Asp247 were mutated to Phe249 and Leu250. These mutations might inactivate SsfM3. As expected, the gene inactivation of SsfM3 did not affect the production of final product **16** and the other two intermediates, indirectly supporting this hypothesis. Glycosyltransferases are commonly found in natural product biosynthesis and used to generate novel compounds [169-170]. Compared to O-glycosidic bond, the C-glycosidic bond occurs less frequently but more stable against glycosidase degradation. Up to date, a few C-glycosyltransferases were characterized in natural product biosynthesis [171]. Among them, UrdGT2, a glycosyltransferse involved in urdamycin biosynthesis was investigated intensively [154-155]. The structure of UrdGT2 was reported and the mechanism of the C-C bond formation was proposed based on structure analysis [172]. Authors proposed that an Asp137 in UrdGT2 stabilizes an 8-phenolate ion using hydrogen bond with 7-hydroxyl and 8-

hydroxyl group via s water. Consequently, the 8-phenolate ion activates the C-9 of UWM6 to undergo nucleophilic attack to the anomeric carbon of dTDP-D-olivose [172]. Interestingly, UrdGT2 showed broad substrate specificities and is able to attach D-olivose to premithramycinone at C-9 position [173]. Compared to premithramycinone, tetracycline scaffold bears similar phenol-phenol-ketone structure which can be used to form hydrogen bond with a putative Asp137 in SsfS6 and successively improve nucleophilicity of C9 through phenol resonance mechanism. The activated C9 is possibly undergoing nucleophilic attack on anomeric carbon of dTDP-D-olivose via a direct Friedel-Crafts alkylation mechanism [172]. It is reflected to the results that the inactivation of SsfS6 led to accumulation of compound **30**. The structure of SsfS6 aglycone **30** is different with the previous proposal [67]. The glycosylation occurs before C-6 methoxylation indicating the phenol-phenol-ketone configuration is required for activation of C9 to undergo Friedel-Crafts like reaction catalyzed by SsfS6. Moreover, the recently approved semi-synthesis drug tigecycline demonstrated the potential value of modification at C-9 position [133]. The C9-glycosyltransferase SsfS6 opens C-9 position of tetracycline aglycone. Therefore, elucidation of SsfS6 aglycone is of interest from both combinatorial biosynthesis and biochemistry angles. In addition, Bechthold and coworkers reported that aryl-C-glycoside catalyst can be generated from a natural product O-glycosyltransferase via rational mutagenesis [174], indicating there is a large space to train SsfS6 as a broad glycosyltransferase towards structurally diverse substrates. The semi-synthesis tigecycline might be afforded by using SsfS6 variants.

Redox enzymes are vital for tetracycline biosynthesis. The insertions of oxygen atoms by oxidation and hydrogen atoms by reduction convert a structurally rigid precursor **18** to more structurally dynamic structures of tetracycline family compounds. The dynamic chemical

behaviours of these compounds exhibit enhanced diverse biological activities [175]. In SF2575 biosynthesis, the C-4-ketoreductase SsfF branches SF2575 biosynthesis away from oxytetracycline biosynthesis by converting **19** to **31**. Unlike oxytetracyline, SF2575 represents 4-(*R*)-OH rather than 4-(*S*)-dimethylamine indicating **31** bears *R* configuration at C-4. By interrogating the mechanism of **19** degradation in previous study [70], we proposed the reduction of C-4-ketone reduces reactivity of A ring consequently quenches the degradation process of **31**. The existence of **31** produced by *ssfS6*, *ssfM1* double knock strain supports this hypothesis. Therefore, SsfF catalyzes C-4-ketone reduction providing a solution to avoid degradation caused by SsfO2 catalyzed dioxidation. Furthermore, the nature of hydroxyl group provides a nucleophilic group at C-4 position, further modification at this position will be investigated. Differing to OxyS in oxytetracycline biosynthesis, SsfO1 catalyses hydroxylation step at C ring and converts **33** an aromatic structure to **34**. Instead of (*S*) configuration at C-6 of oxytetracycline, **16** bears *R* configuration at C-6. The different stereochemistry might results from the attachment of D-olivose or the intrinsic differences between SsfO1 and OxyS. Aminosugar attachment in natural products has been reported as anchor part to orientate substrate approaching catalytic pocket [158]. It can explain that the inactivation of SsfS6 stalls the oxidation at C-6 position. The scenario of C-6 hydroxylation might be initiated by fitting D-olivose into a substrate pocket of SsfO1, orientating C6 to face the activated FAD to eventually form C-O bond at *R* configuration. At this point, the C-6 methylation might provide an inductive effect to facilitate the nucleophilic attack and force the process of dearomatization of C-ring. In addition, the configuration of hydroxyl group is essential for the biological activities of tetracyclines. Exampled as dactylocycline, a naturally occurring tetracycline isolated from *Dactylosporangium* that shows novel antibiotic activities towards tetracycline resistant pathogen

Staphylococcus aureus [138]. The C-6-(*R*)-OH glycosylation provides the key feature that avoids the tetracycline resistant mechanism [138]. Therefore, SsfO1, the only enzyme was confirmed to work as C-6-(*R*)-hydroxylation, will provide new insights in the relationship between *S*-*R* configurations and novel tetracycline activities. The instability of **36** requires successive reactions to maintain this chemical structure. Although the O-methylation enhances the electron density at C-6, the possible electron rearrangement across C5a-C11a double bond increases the risk of degradation. In order to eliminate the double bond, SsfP, a C5a-C11a reductase, was employed to convert **36** to a more stable compound **21** by inserting two hydrogen atoms. In chlortetracycline biosynthesis, a putative F_{420} dependent oxidoreductase TchA was reported to catalyze this step [86]. However, up-to-date evidence suggests that TchA may work as a coenzyme F_{420}:L-glutamate ligase involved in F_{420} biosynthesis [87].

The MIC values of these new tetracycline analogs indicate the relationships between diverse structures and biological activities also provide starting materials for further semi-synthesis. ATC-like compounds show more active than the other intermediates. The attachment of D-olivose on **33** strongly hinders the antibiotic activities, implying the D-olivose moiety might affect the invasion of tetracycline analogs passing through cell membrane. Interestingly, the modification of C-4-*R* salicylic acid and O-4' on D-olivose angelic acid group improves the activities toward Gram-positive strain *Bacillus subtilis* only. This might result from the acid functionalities improved the uptake of compounds into cells. Surprisingly, the C-12a-OH methylation compound **30** does not eliminate antimicrobial activities compared to compound **31**. It has been reported that the C-12a formyl esterified tetracycline analogues bear comparable antibacterial activities with that of tetracycline [176]. Conclusively, elucidation of SF2575 tailoring enzymes including four methyltransferases, one glycosyltransferase, three redox

enzymes provide great opportunities to generate new tetracycline analogs via genetic manipulations. We have updated our previous SF2575 pathway, outlined in Figure 31. Biosynthetic intermediates were isolated and characterized, followed by antimicrobial assay. The bioassay results imply the relationships between tetracycline structures and diverse biological activities (Table 3).

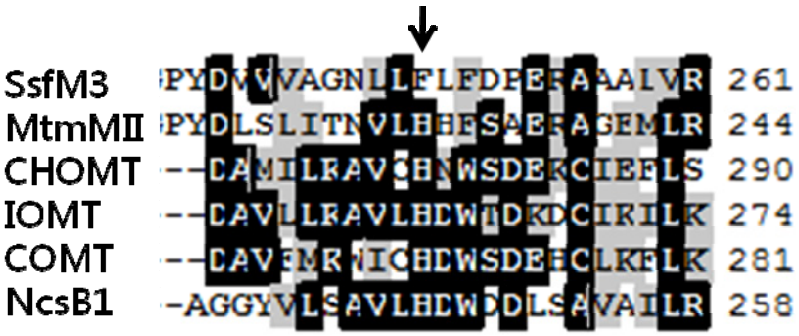


Figure 33. Sequence alignment of SsfM3 with five different methyltransferases. The black arrow indicates amino acid mutation found in SsfM3. MtmMII, accession NO.:109154789; CHOMT, accession NO.:44887779; IOMT, accession NO.:197304942; COMT, accession NO.:23200298; Ncsb1, accession NO.:257097665.

2.2.2.4 Materials and Methods

Bacterial strains, Plasmids and Reagents.

Strains and plasmids used in this study are listed in Table S1. *S.* sp. SF2575 was obtained from Meiji Co. Engineered host *S. lividans* K4-114 [35] was used as heterologous host for production of SF2575 and intermediates. *E. coli* XL-1 Blue (Stratagene) and *E. coli* Topo10 (Invitrogen) were used for the plasmid DNA manipulation. Restriction enzymes were purchased from New England Biolabs. T4 DNA ligase was purchased from Fisher. High fidelity pfx DNA polymerase were purchased from Invitrogen. Primers were synthesized in Integrated DNA Technologies.

DNA manipulation techniques.

Standard DNA manipulation techniques were used for DNA isolation and manipulations. *S. lividans* K4-114 protoplast preparation and PEG mediated DNA transformation techniques were performed according to Hopwood *et al* [177].

Heterologous expression of SF2575 biosynthesis in *S. lividans* K4-114.

Fosmid 5-F15 was digested with *Hind*III to generate 25kb DNA fragment and ligated into fosmid 11A12 *Hind*III site to afford pDP16 containing the whole *ssf* gene cluster. The *cat* gene in pDP16 was replaced by *int-acc(3)IV* cassette to generate pSF2575 using lambda-RED-mediated gene replacement procedures [178-179] SsfT1 was ligated into the *Xba*I site of pPW189 to generate pT1 under the control of the promoter *ermE**. pSF2575 and pT1 were co-transformed into *S. lividans* K4-114 using PEG mediated DNA transformation techniques (Figure S1). [177]. The transformants were streak out on R5 agar plus 25 μg/mL thiostrepton

and 50 μg/mL apramycin, and were well grown at 30 °C for 7-10 days. Quarter of R5 agar was then chopped and extracted with 10 mL of oraganic solvent (89% ethyl acetate, 10% methanol, 1% acetic acid). The solvent was evaporated by speedvap, and were resuspended in 50 μL methanol followed by HPLC-MS analyses.

Gene inactivation in *S. lividans* K4-114 heterologous host.

The gene replacements in *S. lividans* K4-114 host were performed using the lambda-RED-mediated protocol referring to Gust *et al* [104]. The details of individual gene inactivation were described in Supporting Information. (Figure S2)

Production and Isolation of SF2575 analogues.

Mutants in *S. lividans* K4-114 host were inoculated in 50 mL R5 broth in baffled flask for 4 days at 30℃ at 250 rpm. 0.5 mL of seed cultures were inoculated on the R5 agar plates and incubated at 30℃ for 8 days. After that, cultured agar plates were extracted with equal volume of ethyl acetate. Crude extracts were prepared by drying the ethyl acetate extract *in vacuo*. The crude extracts were partitioned with H_2O, $CHCl_3$, and methanol. The $CHCl_3$ extracts were subjected to Sephadex LH-20 using a methanol/chloroform (1:1). Desired fractions were collected and purified by preparative HPLC using an HPLC Alltech Alltima reverse-phase column (5 μm, 10 mm × 250 mm) with a linear gradient of 20-75% CH3CN in water (0.1% TFA) in 30 min at a flow rate of 2.5 mL/min.

Spectroscopic Analyses.

^{1}H- and ^{13}C NMR and 2D NMR spectra of compounds were recorded on a Bruker ARX500 NMR spectrometer at the University of California Department of Chemistry and

Biochemistry NMR facility. HPLC-MS spectra were obtained on a Shimadzu 2010 EV liquid chromatography mass spectrometer using positive and negative electrospray ionization. The samples were separated using a Phenomenex Luna reverse-phase column (5 μm, 2.0 mm × 100 mm C18) on a linear gradient of 5 to 95% CH3CN in water (0.1% formic acid) in 30 min at a flow rate of 0.1 mL/min.

Expression and Purification of Recombinant SsfO1.

SsfO1 was cloned into a pET28a vector using *Nde*I and *Eoc*RI sites to generate expression vector pLP35. *E. coli* BL21 (DE3) strain carrying pLP35 was inoculated into 5 mL LB broth containing 35 mg/L kanamycin for overnight seed culture. This overnight culture was inoculated a 500 mL flask of LB containing 35 mg/L kanamycin and grown at 37 °C to an OD_{600} of around 0.5 at which point isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 120 μM to induce protein expression for overnight shake at 16 °C. Cell membranes were disrupted by sonication in Buffer A (50 mM Tris-HCl, pH 7.9, 10 mM imidazole and 50 mM NaCl). Cellular lysate was centrifuged at 16,000 rpm for 0.5 h. The soluble fraction was incubated with Ni-NTA resin (Qiagen) for 2 h. Protein-resin mixture was loaded to a gravity flow column, and eluting with increasing step gradients of imidazole in Buffers A. The target protein was eluted in Buffer A with 250 mM imidazole. Buffer exchange and concentration of the protein solution were subjected to a 30 kDa MWCO Amicon filtration column (Millipore). Purified SsfO1 was stored in Buffer B (50 mM Tris-HCl pH 7.9, 2 mM EDTA, 2 mM DTT, 10% glycerol) at -80 °C.

SsfO1 in vitro assay.

A standard SsfO1 assay was performed with the following condition: 50 mM Tris pH 8.0, 0.5 mM NADPH, 0.1 mM FAD^+, 0.5 mM substrate, and 50 μM SsfO1. Reaction mixture was incubated at 30°C for 2 h. The assay mixtures were extracted with 2 vol of ethyl acetate (1% acetic acid) and were evaporated *in vacuo*. Samples were then dissolved in 30 μL methanol and were subjected to HPLC-MS analysis as described in Spectroscopic Analyses.

2.3 Cloning and identification of dactylocycline biosynthetic gene cluster.

2.3.1 Introduction

To further utilize the K4-114 strain as a heterologous host for tetracycline biosynthesis, we targeted the biosynthesis of **14**, which is the aglycone of dactylocyclines (Figure 35). Dactylocyclines are heavily modified tetracyclines isolated from *Dactylosporangium* sp. SC 14051 in a pool in New Jersey [60] and is active against *Staphylococcus aureus* strains that are resistant towards tetracycline [138]. The additional structural features from tetracycline include C8 methoxyl, C7 chloride, and the C6 deoxysugar [138]. Co-purified dactylocycline aglycon dactylocyclinone shows cross-resistance activity with tetracycline resistant microorganism indicating the essentialness of deoxysugar entities [138]. However, the acid liable of the ether bond between C6 *R* hydroxyl group of dactylocyclinone and the hydroxylamino sugar hampered the development of the dactylocycline as clinical usage antibiotics [138]. Therefore, understanding the biosynthesis of dactylocyclines may lead to the development of new tetracycline analogs. However, this is particularly challenging in the native host due to its exceedingly slow growth rate and complete genetic incalcitrance. Reconstituting the biosynthesis of **14** (and ultimately dactylocyclines) in K4-114 may serve as a promising alternative.

Table 4. Putative functions of ORFs in the *dac* biosynthetic gene cluster

name	closest homologue	putative role	size (kDa)	identity (%)/similarity (%)	accession number	*oxy* homologue	*ssf* homologue
DacA	SsfA	ketosynthase	45	77/88	ADE34518.1	OxyA	SsfA
DacB	SsfB	chain length factor	41	70/83	ADE34519.1	OxyB	SsfB
DacC	Acyl carrier protein	acyl carrier protein	10	59/72	AEM44307.1	OxyC	SsfC
DacD	OxyD	amidotransferase	68	71/79	AAZ78328.1	OxyD	SsfD
DacK	StfQ	aromatase	35	53/64	CAJ42327.1	OxyK	SsfY1
DacN	OxyN	cyclase	28	70/80	AAZ78337.1	OxyN	SsfY2
DacG	MtmX	cyclase	17	62/75	CAA61988.1	OxyI	SsfY4
DacH	Acyl CoA ligase	acyl-CoA ligase	56	58/67	AEY86471.1	OxyH	SsfL2
DacM1	O-methyltransferase	O-methyltransferase	38	65/77	ZP_06908842.1	OxyF	SsfM4
DacM2	O-methyltransferase	O-methyltransferase	38	54/66	CAE17534.1	OxyT	-
DacM3	StfMI	O-methyltransferase	25	60/70	CAJ42328.1	-	-
DacO1	OxyS	oxygenase	54	55/67	AAZ78342.1	OxyS	SsfO1
DacO2	SsfO2	oxygenase	61	65/74	ADE34483.1	OxyL	SsfO2
DacO3	PokO4	oxygenase	53	54/64	ACN64854.1	-	-
DacO4	OxyR	oxidase	15	50/65	AAZ78341.1	OxyR	-
DacO5	OxyG-oxygenase	oxygenase	12	55/75	AEY86472.1	OxyG	-
DacJ	Flavin reductase	flavin reductase	21	52/67	ABO15845.1	-	-
DacP	ZhuC	acyltransferase	33	60/67	AAG30190.1	OxyP	SsfV
DacQ	OxyQ	aminotransferase	40	65/76	AAZ78340.1	OxyQ	-
DacE	Cts4	halogenase	60	71/81	BAA07389.1	-	-
DacT1	RubRg3	regulator	29	69/78	CAI94728.1	-	SsfT1
DacT2	Response regulator	regulator	21	40/55	ZP_09539699.1	TA1	SsfT2
DacT3	regulator	regulator	24	40/56	YP_003765346.1	OtcG	
DacR1	OtrA	tetracycline resistance protein	70	56/67	CAA37477.1	OtrA	-
DacR2	Snas_5185	drug resistance transporter	55	45/62	YP_003513913.1	OtrB	SsfR
DacR3	Na^+/H^+ exchanger	Na^+/H^+ exchanger	44	62/77	YP_003339806		
DacS1	TDP-glucose synthase	TDP-glucose synthase	39	55/69	CAI94732.1	-	SsfS1
DacS2	Aminotransferase	aminotransferase	40	76/84	AAP85347.1	-	-
DacS3	Hypothetical protein	sugar 3-C-methyltransferase	46	73/81	CAI94697.1	-	-
DacS4	Epimerase	dTDP-4-keto-6-deoxyglucose 3, 5-epimerase	22	63/74	CCD33158.1	-	-
DacS5	Dehydratase	NDP-hexose 2, 3-dehydratase	52	69/79	AAD13549.1	-	SsfS3
DacS6	Hypothetical protein	nitrososynthase	44	67/74	CAI94696.1	-	-
DacS7	Hypothetical protein	methyltransferase	35	62/75	CAI94729.1	-	-
DacS8	Glycosyl transferase	glycosyl transferase	43	44/57	AEM44284.1	-	-
DacS9	4-ketoreductase	sugar 4-ketoreductase	37	45/55	BAC55215.1	-	-
DacP1	PdmP1	biotin carboxylase	47	69/79	ABM21735.1	-	-
DacP2	Biotin attachment	biotin carboxyl carrier protein	32	44/54	YP_482014.1	-	-
DacP3	Unnamed protein	carboxyl transferase	61	59/72	YP_004584141.1	-	-

2.3.2 Results and Discussion

To overcome the limitation of dactylocycline as antibiotic agent, we isolated the genomic DNA of *Dactylosporagium* sp. and sequenced the genome by using 454 sequencing method [180]. To fish out the biosynthetic gene cluster which is responsible to synthesize dactylocycline molecules, halogenase Cts4 [65] and ketosynthase OxyA [66] were used as probes to detect homologous sequence in the library of draft genome sequence of *Dactylosporagium*. Contig299 was identified to harbor the *dac* gene cluster spanning continuous 71 kb DNA fragment. By using annotation online software (http://www0.nih.go.jp/~jun/cgi-bin/frameplot.pl) and BLAST (http://blast.ncbi.nlm.nih.gov/), 38 putative ORFs were identified with their proposed functions listed in Table 4. The *dac* gene cluster was deposited in GenBank with access number JX262387.1.

2.3.2.1 Proposed genes involved in backbone synthesis

Dactylocycline, SF2575, and oxytetracycline share the linear four ring scaffold implying that *dac* gene cluster bears similar genotype of SF2575 or oxytetracycline. To synthesize the 20 carbon polyketide backbone, minimal polyketide synthase unit, consisting of a ketosynthase (KS), a chain length factor (CLF), and an acyl carrier protein (ACP), is required in *dac* gene cluster. As expected, DacA (KS), DacB (CLF), and DacC (ACP) are encoded in *dac* gene cluster. In addition, an amidotransferase DacD, sharing 71% identities with OxyD, is encoded downstream of *dacC* indicating a co-transcriptional event. As described in oxytetracycline biosynthesis, DacD works as an amidotransferase to transfer an amine group into malonyl-CoA or malonyl-ACP to form malonamate starter unit [66]. DacABC complex catalyzes iterative Claisen-like condensations by using the malonamate starter unit and malonyl-CoA as extender

unit [66]. During the biosynthesis process, an acyltransferase DacP, sharing 60% identities with ZhuC [181], might work as an editing enzyme the same as OxyP which was reported to monitoring the starter unit steps by hydrolysing acetyl-ACP to free ACP [71]. The set of extended minimal PKS unit comprising of DacABCDP is encoded in *dac* gene cluster to biosynthesize the dactylocycline polyketide backbone.

2.3.2.2 Proposed genes involved in post modification steps.

The reactive nature of poly-beta-ketone backbone requires dedicated cyclases to sculpture the linearly fused four ring structure. Unlike *oxy* gene cluster, C-9* ketoreductase encoding gene is absent in *dac* gene cluster which stands in line with the remaining C-8 hydroxyl group on dactylocycline structure. We proposed that the first C7-C12 cyclization of C9 unreduced nascent backbone is catalyzed by DacK, which sharing 53% identities with StfQ, an aromatase involved in steffimycin biosynthesis [182]. The putative second ring cyclase DacN shows 60% identities with OxyN which was confirmed to catalyze the second ring cyclization in *oxy* biosynthesis [66]. DacG and DacH, sharing 62% and 58% identities with MtmX [183] and SsfL2 respectively [67] (L. B. Pickens) [157], were proposed to promote the third and forth ring cyclizations to generate linearly fused four aromatic ring system. A OxyF [140] homolog DacM1 was assigned as a C6 methyltransferase to modify C6 position by installing a methyl group to afford C8 hydroxyl-C6 methyl-pretetramide intermediate. To dearomatize A ring, DacO2 was assigned to double hydroxylate C4 and C12a positions followed by DacQ catalyzed reductive transamination to generate C8-hydroxyl version of 4-amino anhydrotetracycline. DacM2 works as methyltransferase to modify the 4-amino group to dimethylamino group. In addition to C8-hydroxyl group, an extra C4a hydroxyl group is present on dactylocycline structure. A

monooxygenase DacO, sharing 54% identities with PokO4 involved in polyketomycin biosynthesis was assigned to catalyze the C4a hydroxylation [184]. The timing of the C4a hydroxylation remains unknown. We proposed two routes for the A ring transformations: 1) the C4a hydroxylation follows C4 and C12a hydroxylations; 2) the DacO3 catalyzes the C4a hydroxylation using C8-hydroxyl anhydrotetracycline as substrate. Halogenase DacE shows 71% identities with Cts4 involved in chlortetracycline biosynthesis[65] and is responsible for the C7 chlorination. DacM3, a methyltransferase, transfers a methyl group onto C8 hydroxyl group to eventually form the key intermediate anhydrodactylocyclinone.

The transformations from anhydrodactylocyclinone to dactylocyclinone are proposed to bear similar mechanism with of anhydrotetracycline to tetracycline. OxyS homolog DacO1 is proposed to convert anhydrodactylocyclinone to 5a, 11a-dehydrodactylocyclinone followed by C5a-C11a reduction catalyzed by OxyR homolog DacO4. Compared to OxyS, DacO1 only adds one hydroxyl group from the opposite side of anhydrodactylocyclinone to give hydroxyl group *R* configuration in final product.

2.3.2.3 Proposed genes involved in hydroxylaminosugar biosynthesis

The structure-activity relationship study of dactylocycline and its aglycone showed the unique hydroxylaminosugar on C6 hydroxyl group leads to improved antibiotic activities against tetracycline resistant microorganisms [138]. Therefore, understanding the biosynthetic logic of the hydroxylaminosugar will offer promising enzymatic tools to morph tetracycline structures to hydroxylaminosugar armed novel analogs with enhanced biological activities.

Total nine sugar biosynthetic enzymes are encoded in *dac* gene cluster. We proposed a biosynthetic pathway based on bioinformatics analysis (Figure 36). Starting from glucose-1-

phosphate, TDP-glucose synthase DacS1 activates glucose by loading glucose onto NDP at anomeric carbon. The 4, 6 dehydration of activated glucose is catalyzed by a NDP-4, 6 dehydratase. However, the dedicated enzyme encoding gene was not found in the *dac* gene cluster. It might be encoded outside the *dac* gene cluster. DacS4, a dTDP-4-keto-6-deoxyglucose 3, 5-epimerase is responsible to convert C4 hydroxyl group and C5 methyl group to the opposite side followed by 2, 3-dehydration catalyzed by DacS5. The C5 ketone group is then replaced by an amino group. DacS2 shows 76% identities with an aminotransferase [152] and was proposed to catalyze this step. C4 reduction and C3 methylation are carried out by DacS9 and DacS3 respectively. DacS6 shows 67% identities with a nitrososynthase RubN8 involved in rubradirin pathway [185] and was proposed to catalyze the oxidation of amino group at C3. DacM4 is proposed to load a methyl group at C4 hydroxyl group to complete the unique hydroxylaminosugar biosynthesis. DacS8, a glycosyltransferase, is responsible to transfer the hydroxylaminosugar onto C6 hydroxyl group of dactylocyclinone to form the final product dactylocycline with novel activities against tetracycline resistant pathogens.

2.3.2.4 Proposed genes involved in regulation, resistance, and precursor supplementation.

Three regulators are encoded in *dac* gene cluster including DacT1, DacT2, and DacT3. DacT1 shares 69% identities with regulation activator RubRg3 [185] indicating the similar role as SsfT1 in SF2575 biosynthesis. DacT2 shows similarity with OtcG which was reported as a positive regulator [186]. DacT3 is proposed as a response regulator to control the transcription of DacR2 [187]. DacR2 and DacR1 are encoded to establish the resistance mechanisms working as transporter and ribosomal modification protein respectively [188-189].

To enhance the precursor supplementation, DacP1, DacP2, and DacP3 are encoded in *dac* gene cluster to generate malonyl-CoA from acetyl-CoA [190]. DacJ shows 52% identities with a flavin reductase [191] indicating the supporting role to regenerate reduced FAD like cofactor for flavin-dependent proteins such as DacO1 and DacO2.

2.3.2.5 Heterologous expression of dactylocyclinone.

To transplant the *dac* pathway into K4-114, two plasmids encoding all the necessary genes proposed for the formation of **14** and overexpression of the SARP DacT1 were constructed and transformed into K4-114 to generate K4/pDac-T1O3E. HPLC-MS analysis showed the emergence of two major compounds with same mass and UV absorption spectrum as **14** at a titer of ~1 mg/L (Figure 34). Proton NMR analysis confirmed one of the compounds to be **14**, while the other compound was assigned to be the C4-epimer of **14** [192]. The isolation of **14** confirms the role of the *dac* gene cluster in the biosynthesis of dactylocyclines, and affords a third set of enzymatic tools that can modify the tetracycline scaffold. This also represents the first report of reconstituting natural products originating from *Dactylosporangium* in *Streptomyces.*

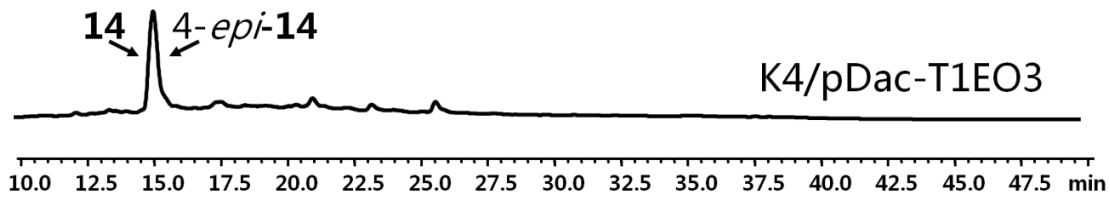


Figure 34. HPLC analysis of *dac* heterologous expression in *S. lividans* K4.

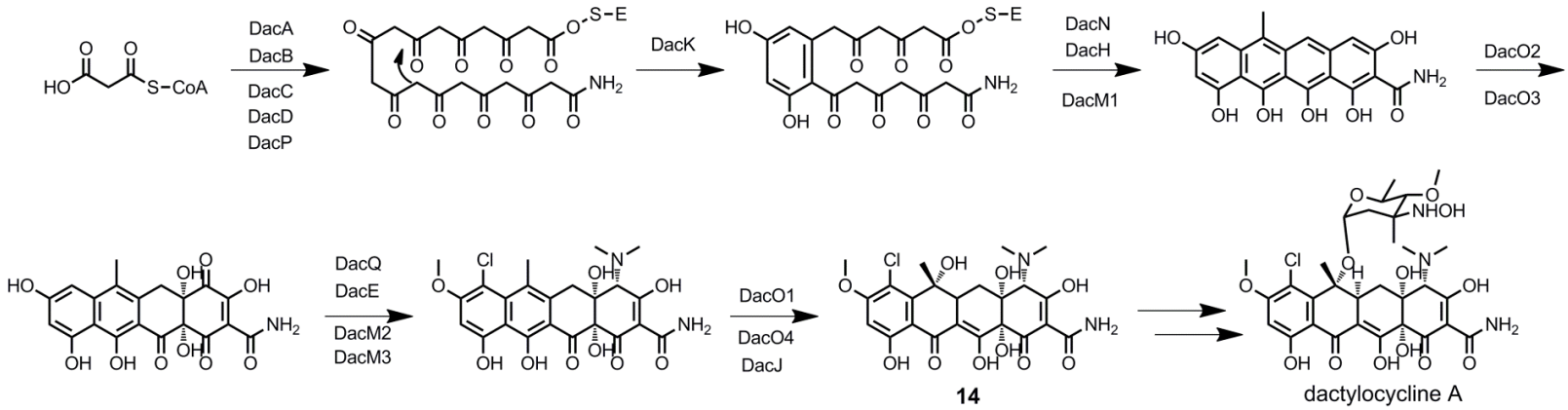


Figure 35. Proposed biosynthetic pathway of dactylocycline A.

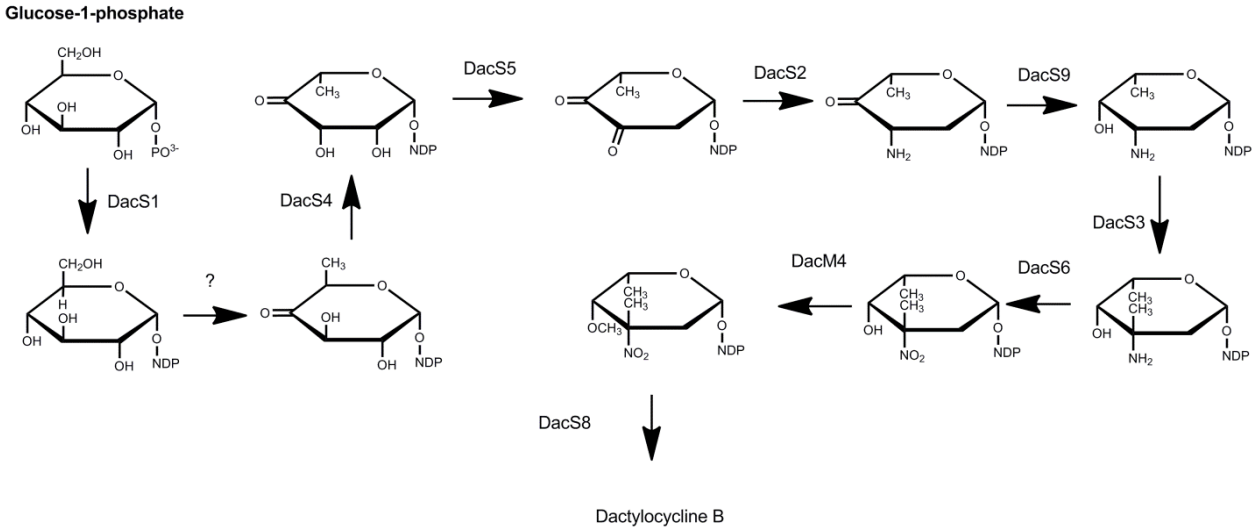


Figure 36. Proposed biosynthetic pathway of hydroxylaminosugar in dactlyocycline biosynthesis.

3. Conclusions

Three natural tetracycline biosynthetic pathways have been characterized by using genetic, biochemical, and protein structure analysis. The linearly fused four-ring scaffold is shared by tetracycline family and essential for their diverse biological activities, especially for the broad-spectrum antibiotic activities. However, the emergence of tetracycline resistance mechanisms in microbial communities reduces the effectiveness of tetracyclines, resulting in high demand of new generation tetracycline antibiotics. Additionally, the use of non-antibiotic tetracycline analogs [78] and the exceptional antitumor activities of SF2575 intrigue us to explore chemical spaces of tetracyclines for development of both antibiotic and non-antibiotic tetracycline medical agents by using biosynthesis engineering approach.

First, we completed our understanding of oxytetracycline biosynthetic pathway by uncovering the enzymatic basis of the final transformations from anhydrotetracycline to oxytetracycline. Two redox enzymes OxyS and OxyR catalyze three tandem reactions including C6, C5 hydroxylation and C5a-C11a reduction respectively. The unexpected OxyS catalyzed C5 hydroxylation led us to determine the crystal structure of OxyS. The structural complex of OxyS with oxidized FAD revealed three-domain architecture consisting of FAD binding domain, middle domain, and C-terminal thioredoxin-like domain. Structural alignment of OxyS-FAD complex with RdmE-aklavinone structural complex led us to simulate the possible interactions between OxyS and anhydrotetracycline. The structural analysis based site-directed mutants of OxyS evaluated the possible functions of His47 and Phe215 in OxyS catalytic pocket. In addition, a novel F_{420} dependent enzyme OxyR was revealed to transform 5a-11a dehydrotetracycline and 5a-11a dehydrooxytetracycline to tetracycline and oxytetracycline by

reducing the C5a-C11a double bond. To our best knowledge, it is the first F_{420} dependent enzyme was characterized in secondary metabolite biosynthesis.

Second, we identified a set of tetracycline modifying enzymes in SF2575 biosynthetic pathway. To circumvent genetic barriers of SF2575 native producer, a heterologous expression host in *S. lividans* K4-114 was established to produce SF2575. (The same strategy was also employed in oxytetracycline biosynthesis investigation.) By using the surrogate host, we were able to identify the functions of SF2575 tailoring enzymes by using gene inactivations and mutant metabolite analysis. Four methyltransferases, three redox enzymes, and one glycosyltransferase were identified and a number of tetracycline analogs were generated by using mutagenesis. SsfO1, a homolog of OxyS in SF2575 biosynthetic pathway, catalyzes C6 hydroxylation but gives *R* configuration compared to OxyS catalyzed reaction in oxytetracycline biosynthesis. To understand the intrinsic differences between SsfO1 and OxyS, we determined the structure of SsfO1. As expected, SsfO1 bears a similar three domain structure with OxyS but shows bigger binding pocket to accommodate D-olivose attached tetracycline scaffold.

Third, we identified dactylocycline biosynthetic gene cluster by using 454 genome sequencing approach. The third natural tetracycline biosynthesis offers us more enzymatic tools to modify tetracycline scaffold. In addition to oxytetracycline and SF2575, a number of new functionalities are found in dactylocycline structure including C4a hydroxyl, C8 methoxy, C7 chloride, and C6-OH hydroxylamino sugar groups. The heterologous expression of dactylocyclinone in *S. lividans* K4 validated the dactylocycline biosynthetic gene cluster and demonstrated the first successful heterologous expression of gene cluster from *Dactylosporangium* in *Streptomyces*.

Appendices

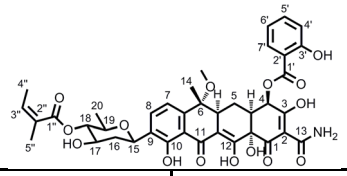
Table S1. Strains and plasmids in this study

Strain or plasmid	Characteristics	Reference or source
Strains		
E. coli		
TOP 10	*recA1*	Invitrogen
XL-1 blue	*recA1*	Stratagene
BW25113	*ΔaraBAD ΔrhaBAD*	[125]
Streptomyces lividans		
K4-114	Heterologous expression host strain	[35]
K4/pOTC-Ctc11	K4-114 harboring pOTC and pCtc11	This study
K4/pSKO-Ctc11	K4-114 harboring pOxySKO and pCtc11	This study
K4/pSF2575-T1	K4-114 harboring pSF2575 and pT1	This study
K4/pM1KO-T1	K4-114 harboring pM1KO and pT1	This study
K4/pM2KO-T1	K4-114 harboring pM2KO and pT1	This study
K4/pM3KO-T1	K4-114 harboring pM3KO and pT1	This study
K4/pM4KO-T1	K4-114 harboring pM4KO and pT1	This study
K4/pM2M4KO-T1	K4-114 harboring pM2M4KO and pT1	This study
K4/pS6KO-T1	K4-114 harboring pS6KO and pT1	This study
K4/pO1KO-T1	K4-114 harboring pO1KO and pT1	This study
K4/pKKO-T1	K4-114 harboring pKKO and pT1	This study
K4/pFKO-T1	K4-114 harboring pFKO and pT1	This study
K4/pM1S6KO-T1	K4-114 harboring pM1S6KO and pT1	This study
K4/pM4S6KO-T1	K4-114 harboring pM4S6KO and pT1	This study
K4/M1KOC	K4-114 harboring pM1KO and pM1KOC	This study
K4/M2KOC	K4-114 harboring pM2KO and pM2KO	This study
K4/S6KOC	K4-114 harboring pS6KO and pS6KOC	This study
K4/O1KOC	K4-114 harboring pO1KO and pO1KOC	This study
K4/FKOC	K4-114 harboring pFKO and pFKOC	This study
K4/pDac-T1O3E	K4-114 harboring pDac and pDacT1O3E	This study
Plasmid		
pKD46	Ap^r, λ-RED for PCR-targeting	[125]
pIJ778	*aadA* gene source	[178]
pIJ787	*tet-int* cassette source	[179]
pDP16	Fosmid harboring entire *ssf* gene cluster	This study
pDP24	*int- aac(3)IV* cassette source	This study
pDP82	Km^r gene source	This study
pYT264	Cosmid harboring entire *oxy* gene cluster	This study
pCtc11	pJTU870 derivative in which *ctc11* under the control of *ermE**	This study
pOTC	pYT264 derivative in which *bla* was replaced by *tet-int* cassette	This study
pOxySKO	pOTC derivative in which *oxyS* was in-frame deleted.	This study
pT1	pJTU870 derivative in which *ssfT1* under the control of *ermE**	This study
pSF2575	pDP16 derivative in which *cat* was replaced by *int- aac(3)IV*	This study
pM1KO	pSF2575 derivative in which *ssfM1* was disrupted by *aadA+oriT*	This study
pM2KO	pSF2575 derivative in which *ssfM2* was disrupted by *aadA+oriT*	This study
pM3KO	pSF2575 derivative in which *ssfM3* was disrupted by Km^r gene	This study
pM4KO	pSF2575 derivative in which *ssfM4* was disrupted by Km^r gene	This study
pM2M4KO	pSFM2KO derivative in which *ssfM4* was disrupted by Km^r gene	This study
pS6KO	pSF2575 derivative in which *ssfS6* was disrupted by *aadA+oriT*	This study
pO1KO	pSF2575 derivative in which *ssfO1* was disrupted by *aadA+oriT*	This study
pKKO	pSF2575 derivative in which *ssfK* was disrupted by *aadA+oriT*	This study
pFKO	pSF2575 derivative in which *ssfF* was disrupted by *aadA+oriT*	This study
pM1S6KO	pSFS6KO derivative in which *ssfM1* was disrupted by Km^r gene	This study
pM4S6KO	pSFS6KO derivative in which *ssfM4* was disrupted by Km^r gene	This study
pM1KOC	pT1 derivative in which *ssfM1* was constructed after *ssfT1*	This study
pM2KOC	pT1 derivative in which *ssfM2* was constructed after *ssfT1*	This study
pS6KOC	pT1 derivative in which *ssfS6* was constructed after *ssfT1*	This study
pO1KOC	pT1 derivative in which *ssfO1* was constructed after *ssfT1*	This study
pFKOC	pT1 derivative in which *ssfF* was constructed after *ssfT1*	This study
pDac	Fosmid harboring part *dac* gene cluster	This study
pDacT1O3E	pJTU870 derivative in which *dacT1, dacO3, and dacE* under the control of *ermE**	This study

Table S2. Primers used in this paper

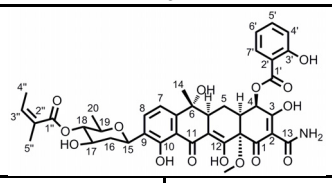
Primer	Sequence
OxySKO_5	5'-CTGCGCCTGGCCGGTGCCAGGACCCTGGTGCTGGAACGCATTCCGGGGATCCGTCGACC-3'
OxySKO_3	5'-CTCCGGAGCCGCGGTCCAGCAGATGTAGCCGTCGGGGCGTGTAGGCTGGAGCTGCTTC-3'
M1KO_5	5'-GACAGGCTGTCCTGAGCGACGAAAGGAAGGGTCACAGTGATTCCGGGGATCCGTCGACC-3'
M1KO_3	5'-GGCGCCCGCGCTGCGCGCCGTCCTGCCGCCCGCCGGTCATGTAGGCTGGAGCTGCTTC-3'
M2KO_5	5'-TCACCACCGACGATCCGAACCGCGAAAGGAAGCGACATGATTCCGGGGATCCGTCGACC-3'
M2KO_3	5'-GTCCGCGTCCCGGCCGGCGCTTCGGAGTCCGGGGGATCATGTAGGCTGGAGCTGCTTC-3'
M3KO_5	5'-GAGACCGCGATGCTGCGCACCGCGCTCGAACTCCCTAGGCACGTAGAAAGCCAGTCCGC-3'
M3KO_3	5'-GCGCAGCACGCGGATACCGGTCAGGCCGGCCGCGCTAGCTCAGAAGAACTCGTCAAGAA-3'
M4KO_5	5'- GCCCGGGCCGCCGCGCTGCGGGCGGCGACCCAGCCTAGGCACGTAGAAAGCCAGTCCGC-3'
M4KO_3	5'- CAGGCCGCCCTTCTCGATCAGCGTGCGGAACTCGCTAGCTCAGAAGAACTCGTCAAGAA-3'
S6KO_5	5'-GAGCATCGGCCGAGCCGTGCGGGCTTGGATGCCGCCATGATTCCGGGGATCCGTCGACC-3'
S6KO_3	5'-AGCCCGGGTCACGCCGTCCCGTTCAGACCGGCAGGCTCATGTAGGCTGGAGCTGCTTC-3'
O1KO_5	5'-GCCGCCCCGGTAGCAAGCGTGGAGGAGGACGACCGCATGATTCCGGGGATCCGTCGACC-3'
O1KO_3	5'- GTCGCTTCCTTTCGCGGTTCGGATCGTCGGTGGTGATCATGTAGGCTGGAGCTGCTTC-3'
KKO_5	5'-AGGAATCCCATGCGTCTCAAGGACAAGGTCGCCGTGATCATTCCGGGGATCCGTCGACC-3'
KKO_3	5'-GTTGACTTCGAGCACATGGCCGTTGACGTAGCTGCTGTCTGTAGGCTGGAGCTGCTTC-3'
FKO_5	5'-TGCTCGGGCCATCAGCCCGACCGGATGGAGCACACGATGATTCCGGGGATCCGTCGACC-3'
FKO_3	5'-TCCTCGTCGCCGGTGTGACTGCTCGGACTGGTGCGCTCATGTAGGCTGGAGCTGCTTC-3'
OxySKOT_5	5'-AACGGCATGAAGCGTCCTCG-3'
OxySKOT_3	5'-TCAGCTGCAGCAGCGAGCGG-3'
M1KOT_5	5'-TCACGATGCGCGACGACAGG-3'
M1KOT_3	5'- CTGGTGAACGGTTCGGACCC-3'
M2KOT_5	5'-CCGACCACCAGTGATCACCACC-3'
M2KOT_3	5'-AGCACCGTTCGAGGCCCGTT-3'
M3KOT_5	5'-AAAACTAGTGGAGGAGCACATATGCCGGACACGGCAGGTGC-3'
M3KOT_3	5'-AAAGCTAGCTCATGACGGAAGTCCCTCCTTC-3'
M4KOT_5	5'-AAGGACTAGTGGAGGAAGCCATATGACCAGCACCGACACCGA-3'
M4KOT_3	5'-AAAGCTAGCTCAGCTCGCCACCGTGCACT-3'
S6KOT_5	5'-GTGGACTGCGGCTCGAGCATCG-3'
S6KOT_3	5'-ACATGACCCGGGCTGCCAGAGC-3'
O1KOT_5	5'-ATGTCGTCTCCCGCGGTCGGCAATGACACACCTATGT -3'
O1KOT_3	5'-ATCCCGTCCACGTGCGTGGT-3'
KKOT_5	5'-AAGGTTTCATTTAGCGCCTC-3'
KKOT_3	5'-GCGAGCTTTGATCAACGACC-3'
FKOT_5	5'-ACATCGAATCGAGCGCCGCC-3'
FKOT_3	5'-CCATGGGTGCGAACTCCTCG-3'
FRT_5	5'-ATTCCGGGGATCCGTCGACC-3'
FRT_3	5'-TGTAGGCTGGAGCTGCTTC-3'
Ctc11_5	5'-AAACTAGTATGGTGAAGTTCAACATTCTGGGTC-3'
Ctc11_3	5'-AAGCTAGCTCAGGCCGAACGGGCCCGCAGG-3'
OxyS_5	5'-ATCTAGAGGAGGAGCCCATATGCGGTACGACGTGGTGAT-3'
OxyS_3	5'-AGCTAGCTCACATGGCGGGCGGCCCG-3'
SsfT1_5	5'-AAGGACTAGTGGAGGAAGCCATATGATGTTTAAGATCCTTGGC-3'
SsfT1_3	5'-AAAGCTAGCCTAGGGGGCCAGGAGGACGGG-3'
M1_5	5'-AAGGACTAGTGGAGGAGCACCGATGACCGACGCCGCCGCTATCTC-3'
M1_3	5'-AAGGGCTAGCTCAGGCCTTGTGGGCGATGG-3'
M2_5	5'-AAGGACTAGTGGAGGAAGCCATATGGCTGACGCCCCC-3'
M2_3	5'-AAGGGCTAGCTCAGCCGGCGGTGAACTCCA-3'
S6_5	5'-AAGGACTAGTGGAGGAAGCCATATGAGAATCCTGGTCATCGC -3'
S6_3	5'-AAAGCTAGCTCACGCCTGCTCGATCAGCC-3'
O1_5	5'-AAAACTAGTGGAGGAGCACATATGGAACCAGAGGTGGTGGT-3'
O1_3	5'-AAGGGCTAGCTCACTGGTGGTCGGGAGCGC-3'
F_5	5'-AAAACTAGTGGAGGAGCACATATGGAGTACCGAAAGCTGGG-3'
F_3	5'-AAAGCTAGCTCAGCGGCGCACCGGGACCCGCG-3'
DacT1_5	5'-AAACTAGTGGAGGAGCCCATATGATTTTTCGGGTACTTGG-3'
DacT1_3	5'-AACCTAGGTCAGGCGAGCCGGGCGTGCCGG-3'
DacE_5	5'-AAACTAGTGGAGGAGCCCATATGCCCACCGACCGCGACTACGACGTCA-3'
DacE_3	5'-AACCTAGGTCAGTGGGTGGGCGCGGCGG-3'
DacO3_5	5'-AAACTAGTGGAGGAGCCCATATGAACACGACGGACATCGCGGTGGTCGG-3'
DacO3_3	5'-AACCTAGGTCAGCCCGCGCCGGCGAAGTACGTC-3'

Table S3 NMR Data of **22** in DMSO Measured on Bruker NMR Spectrometer.

22				
				
NO.	^{13}C δ(ppm)	^{1}H δ(ppm) (m, J_{HH}(Hz))	HMBC	HMQC
1	189.2[a]	-	-	-
2	95.1	-	13 (NH_2)	-
3 (OH)	188.6[a]	-	-	-
4	69.8	6.28 (d, 4.8)	-	H4
4a	40.4	2.83 (m)	-	H4a
5	19.1	1.55 (m), 2.15 (m)	-	H5-2
5a	34.9	3.48 (m)	H14	H5a
6-OCH_3	50.3	3.18 (s)	-	6-OCH_3
6	77.9	-	6-OCH_3, H7, H14	-
6a	147.4	-	H8, H14	-
7	114.9	6.97 (d, 7.9)	-	H7
8	133.8	7.70 (d, 7.9)	-	H8
9	128.6	-	H7, 10(OH)	-
10 (OH)	158.1	12.15 (s)	H8, 10(OH)	-
10a	114.5	-	H7, 10(OH)	-
11	192.3	-	-	-
11a	106.1	-	-	-
12(OH)	176.0	14.91 (s)	12(OH)	-
12a	75.1	-	12(OH)	-
13 (NH_2)	173.2	9.37 (s)	-	-
14	24.4	1.03 (s)	-	H14
15	70.3	4.81 (d, 11.4)	H8	H15
16	40.9	1.49 (m), 2.21 (m)	-	H16-2
17	69.1	3.79 (m)	H18	H17
17(OH)	-	5.13 (b)	-	-
18	77.6	4.58 (t, 9.2)	H20	H18
19	73.7	3.60 (m)	H18, H20	H19
20	18.0	1.11 (d, 6.1)	-	H20
1'	167.2	-	H7'	-
2'	113.3	-	3'(OH), H4', H6'	-
3'(OH)	159.9	10.34 (s)	3'(OH), H5', H7'	-
4'	117.7	7.03 (d, 8.7)	3'(OH), H6'	H4'
5'	135.9	7.55 (t, 7.8)	H7'	H5'
6'	119.5	7.00 (m)	H4'	H6'
7'	130.3	7.84 (d, 7.9)	H5'	H7'
1''	166.9	-	H18, H5''	-
2''	127.7	-	H4'', H5''	-
3''	136.9	6.10 (m)	H4'', H5''	H3''
4''	15.5	1.92 (d, 7.2)	-	H4''
5''	20.3	1.86 (s)	H3''	H5''

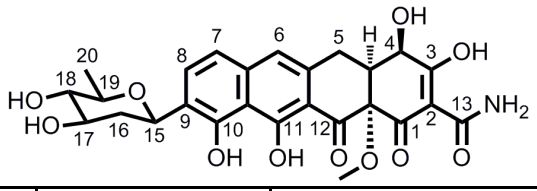
[a] Signal is weak.

Table S4 NMR Data of **25** in DMSO Measured on Bruker NMR Spectrometer.

25				
				
NO.	^{13}C δ(ppm)	^{1}H δ(ppm) (m, J_{HH}(Hz))	HMBC	HMQC
1	189.8	-	-	-
2	95.6	-	-	-
3 (OH)	189.6[a]	-	-	-
4	69.8	6.09 (m)	-	H4
4a	41.6	2.99 (m)	H14,	H4a
5	19.5	1.56 (m), 2.41 (m)	-	H5-2
5a	36.5	3.15 (m)	-	H5a
6	71.3	-	H7, H14,	-
6 (OH)		5.47 (s)		
6a	152.3	-	H8, H14,	-
7	114.6	7.22 (d, 8.0)	-	H7
8	133.8	7.67 (d, 8.0)	-	H8
9	128.1	-	H7, H10(OH),	-
10 (OH)	157.3	11.98 (s)	H8, H10(OH),	-
10a	113.0	-	H7, H10(OH),	-
11 (OH)	192.9	-	-	-
11a	116.2	-	-	-
12(OH)	174.5	14.89 (s)	-	-
12a	81.0	-	12a-OCH_3	-
12a-OCH_3	54.1	3.49 (s)	-	12a-OCH_3
13 (NH_2)	173.7	9.40 (s), 9.48 (s)	-	-
14	24.3	0.98 (s)	-	H14
15	70.3	4.80 (d, 11.2)	H8,	H15
16	40.4	1.46 (m), 2.18 (m)	-	-
17	69.1	3.78 (b)	-	H17
17(OH)		5.13 (d, 5.6)	-	-
18	77.5	4.57 (t, 9.3)	H20	H18
19	73.7	3.60 (m)	H20	H19
20	18.0	1.10 (d, 6.1)	-	H20
1'	167.1	-	H7'	-
2'	113.1	-	3'(OH), H4', H6'	-
3'(OH)	160.1	10.37 (s)	3'(OH), H5', H7'	-
4'	117.7	7.03 (d, 8.4)	3'(OH), H6'	H4'
5'	136.1	7.57 (m)	H7'	H5'
6'	119.5	7.00 (m)	H4'	H6'
7'	130.2	7.87 (dd, 7.9, 1.6)	H5'	H7'
1''	166.9	-	H18, H5''	-
2''	127.7	-	H4'', H5''	-
3''	136.9	6.10 (m)	H4'', H5''	H3''
4''	15.5	1.92 (dd, 7.2, 1.4)	-	H4''
5''	20.3	1.86 (s)	-	H5''

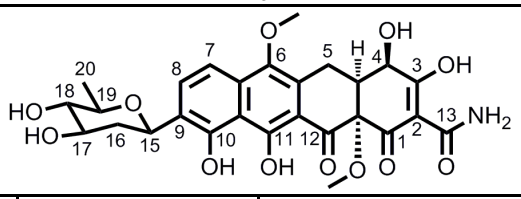
[a] Signal is weak.

Table S5 NMR Data of **28** in DMSO Measured on Bruker NMR Spectrometer.

28				
				
NO.	^{13}C δ(ppm)	^{1}H δ(ppm) (m, J_{HH}(Hz))	HMBC	HMQC
1	196.9[a]	-	-	-
2	98.0	-	-	-
3 (OH)	188.2[a]	-	-	-
4	69.8	4.61 (s)	-	H4
4a	39.6	3.28 (m)	-	H4a
5	26.3	2.83 (m) 3.26 (m)	H6	H5-1, H5-2
5a	134.7	-	-	-
6	117.6	7.14 (s)	H7	H6
6a	138.0	-	H6, H7, H8	-
7	118.2	7.25 (d, 8.3)	H6	H7
8	130.4	7.64 (d, 8.4)	-	H8
9	123.8	-	H7, H15	-
10 (OH)	152.1	9.80 (s)	H8	-
10a	111.4	-	H7	-
11 (OH)	164.1	15.38 (s)	-	-
11a	110.2	-	H6	-
12	200.0	-		-
12a	84.2	-	12a-OCH_3	-
12a-OCH_3	54.6	3.49 (s)	-	12a-OCH_3
13 (NH_2)	173.0	9.14 (s), 9.20 (s)	-	-
15	70.5	4.83 (d, 10.8)	H8, H16-1	H15
16	40.4	1.37 (m), 2.10 (m)	-	H16-1, H16-2
17	71.9	3.51 (m)	H18	H17
18	77.2	2.88 (m)	H16-2	H18
19	76.1	3.32 (m)	H18, H20	H19
20	18.5	1.23 (d, 6.1)	H16-1, H18	H20

[a] Signal is weak.

Table S6 NMR Data of **29** in DMSO Measured on Bruker NMR Spectrometer.

29				
				
NO.	^{13}C δ(ppm)	^{1}H δ(ppm) (m, J_{HH}(Hz))	HMBC	HMQC
1	197.0[a]	-	H4a	-
2	98.1	-	-	-
3 (OH)	188.2[a]	17.78 (s)	H4a	-
4	67.4	4.63 (s)	H4a	H4
4a	40.1	3.23 (m)	H5-1	H4a
5	20.7	2.63 (m), 3.33 (m)	-	-
5a	115.6	-	6-OCH_3, H8	-
6	143.9	-	6-OCH_3, H7	-
6a	132.9	-	H8	-
7	112.6	7.44 (d, 8.7)	-	H7
8	130.6	7.73 (d, 8.4)	H7, H15	H8
9	124.6	-	H7, H8, 10(OH), H15, H16-1	-
10 (OH)	152.8	9.95 (s)	H7, H8, H15	-
10a	112.2	-	H7	-
11 (OH)	160.9	15.34 (s)	-	-
11a	109.5	-	H7, 12a-OCH_3	-
12	200.0	-		-
12a	84.0	-	H4a, 12a-OCH_3	-
12a-OCH_3	54.6	3.48 (s)	-	12a-OCH_3
13 (NH_2)	173.0	9.15 (s), 9.20 (s)	-	-
6-OCH_3	60.6	3.72 (s)	-	6-OCH_3
15	70.5	4.85 (d, 12.1)	H8	H15
16	40.4	1.38 (m), 2.12 (m)	-	H16-1, H16-2
17	71.9	3.52 (m)	H15, H16-1, H16-2, H18	H17
18	77.2	2.89 (t, 8.9)	H16-1, H16-2, H19, H20	H18
19	76.1	3.32 (m)	H15, H16-2, H18, H20	H19
20	18.4	1.24 (d, 6.1)	H18	H20

[a] Signal is weak.

Table S7 NMR Data of **30** in DMSO Measured on Bruker NMR Spectrometer.

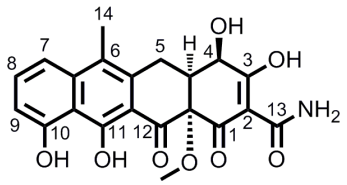
30				
				
NO.	^{13}C δ(ppm)	^{1}H δ(ppm) (m, J_{HH}(Hz))	HMBC	HMQC
1	197.1	-	-	-
2	98.1	-	-	-
3 (OH)	188.4	-	-	-
4	67.7	4.68 (s)	-	H4
4a	39.6	3.39 (m)	H5-1	H4a
5	25.2	2.64 (m) 3.21 (m)	H4a	H5-1,H5-2
5a	132.6	-	H5-1	-
6	121.4	-	H4a, H5-1, H7, H14	-
6a	138.5	-	H7, H8, H14	-
7	114.8	7.30 (d, 8.5)	H8, H9	H7
8	131.3	7.50 (t, 8.1)	H4a, H7, H14	H8
9	110.8	6.79 (d, 7.8)	H7, H8	H9
10 (OH)	157.7	-	H7, H8, H9	-
10a	111.8	-	H9	-
11 (OH)	163.0	15.47 (s)	-	-
11a	109.8	-	H4a	-
12	200.0	-	-	-
12a	83.3	-	H4a, H5-2,12a-OCH_3	-
12a-OCH_3	54.6	3.52 (s)	H4a	12a-OCH_3
13 (NH_2)	173.0	9.21 (s), 9.90 (s)	-	-
14	13.9	2.21 (s)	-	H14

Table S8 NMR Data of **33** in DMSO Measured on Bruker NMR Spectrometer.

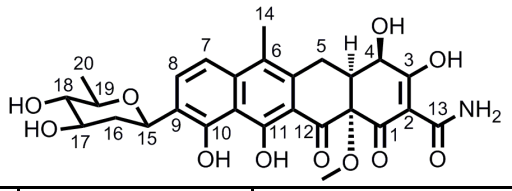
33				
				
NO.	^{13}C δ(ppm)	^{1}H δ(ppm) (m, J_{HH}(Hz))	HMBC	HMQC
1	197.2	-	-	-
2	98.2	-	-	-
3 (OH)	188.4	-	-	-
4	67.6	4.66 (s)	-	H4
4(OH)	-	6.05 (s)	-	-
4a	39.7	3.20 (m)	-	H4a
5	25.0	2.67 (m), 3.36 (m)	-	H5-1, H5-2
5a	131.2	-	H14	-
6	121.7	-	H7,H14	-
6a	137.5	-	H7, H8, H14	-
7	115.0	7.43 (d, 8.6)	-	H7
8	130.3	7.68 (d, 8.6)	H15	H8
9	123.8	-	H7, H15	-
10 (OH)	152.8	10.13 (s)	H8, H15	-
10a	111.2	-	H7	-
11 (OH)	163.0	15.76 (s)	-	-
11a	109.9	-	H5-2	-
12(OH)	200.4	-	-	-
12a	83.4	-	12a-OCH_3, H5-2	-
12a-OCH_3	54.7	3.51 (s)	-	12a-OCH_3
13 (NH_2)	173.0	9.14 (s), 9.20 (s)	-	-
14	13.8	2.28 (s)	-	H14
15	70.5	4.83 (d, 11.2)	H8, H16-1,	H15
16	40.4	1.37 (m), 2.13 (m)	-	H16-1, H16-2
17	72.0	3.51 (m)	H16-1, H16-2, H18	H17
18	77.3	2.89 (t, 8.7)	H16-2	H18
19	76.1	3.34 (m)	H18, H20	H19
20	18.5	1.24 (d, 5.9)	H16-1, H18	H20

Table S9 NMR Data of **31** in DMSO Measured on Bruker NMR Spectrometer.

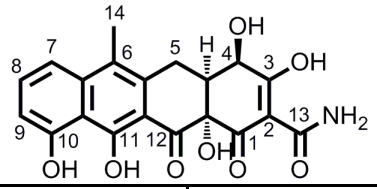
31				
				
NO.	^{13}C δ(ppm)	^{1}H δ(ppm) (m, J_{HH}(Hz))	HMBC	HMQC
1	196.7	-	-	-
2	97.4	-	-	-
3 (OH)	190.4	18.02 (s)	-	-
4	68.0	4.77 (s)	-	H4
4a	42.7	2.91 (m)	H5-1, H5-2	H4a
5	25.0	2.62 (m) 3.37 (m)	-	H5-1, H5-2
5a	132.7	-	H5-1	-
6	114.9	-	H7	-
6a	138.7	-	H7, H8, H14	-
7	121.5	7.41 (d, 8.4)	H5-2, H8, H9, H14	H7
8	131.9	7.57 (t, 8.1)	H7, H14	H8
9	110.8	6.85 (d, 7.7)	H7	H9
10 (OH)	157.7	9.95 (s)	H7, H8, H9	-
10a	112.0	-	H7, H9	-
11 (OH)	162.7	15.37 (s)	H7	-
11a	109.0	-	H5-2	-
12	201.2	-	-	-
12a	78.2	-	H5-2	-
13 (NH_2)	173.2	9.09 (s)	-	-
14	14.0	2.31 (s)	-	H14

Figure S1. Gene organization of oxytetracycline [66], SF2575 [67], and dactylocycline biosynthsis.

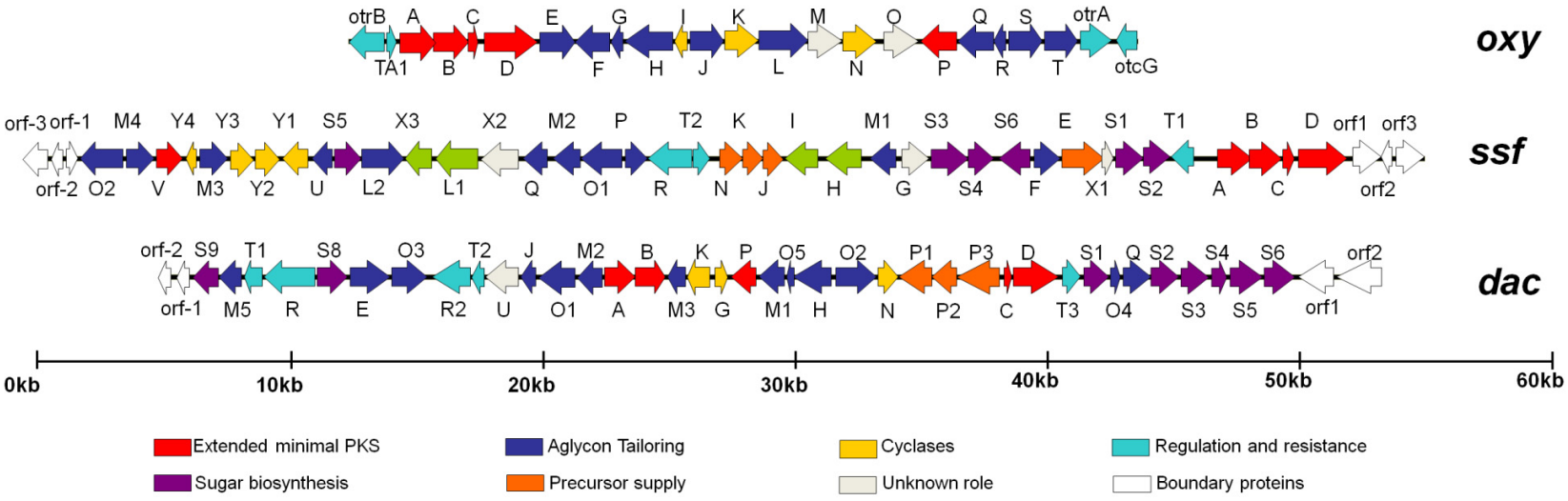
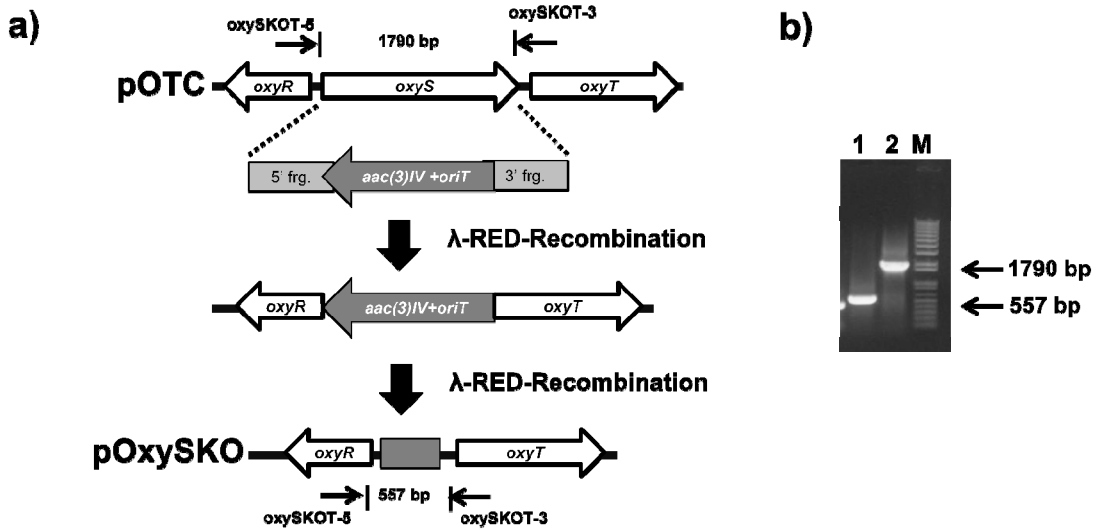
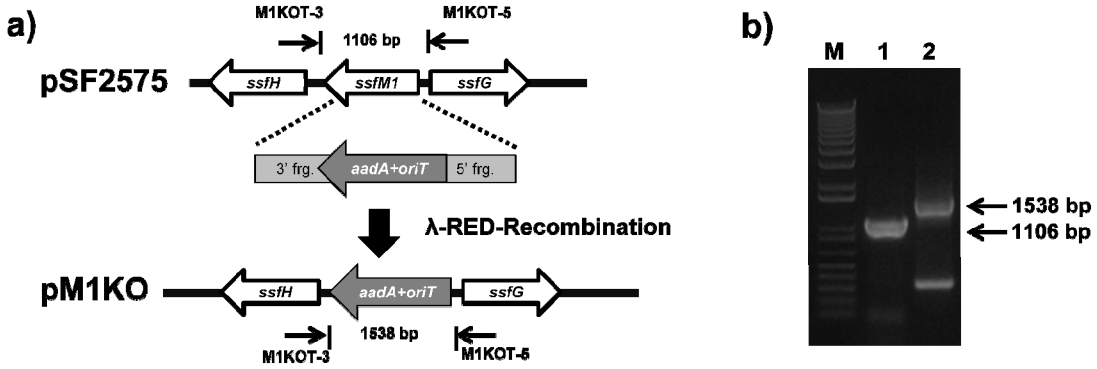


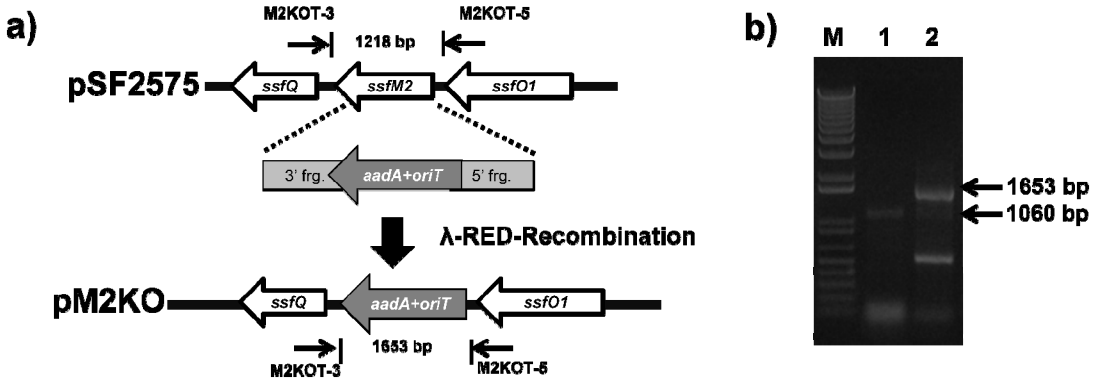
Figure S2. Depictions of construction of gene inactivation plasmids which were confirmed by PCR.



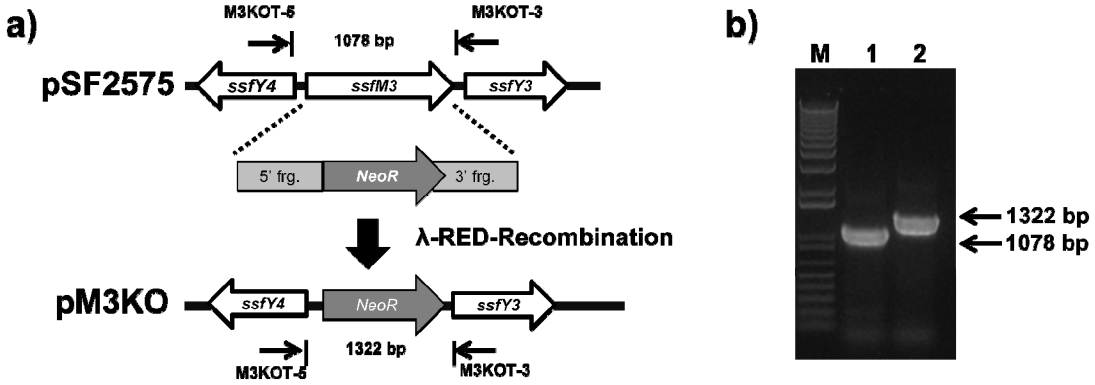
A. Construction of *oxyS*-inactivation plasmid pSKO. (a) Scheme of *oxyS*-inactivation. pSKO was generated by replacing a 1314 bp *oxyS* fragment with a 1382 bp *aac(3)IV* and *oriT* DNA cassette and then removed it via lambda-RED-mediated recombination. (b) PCR confirmation of mutant plasmid. M, 1 kb plus DNA ladder ; lane 1) pSKO; lane 2) pOTC



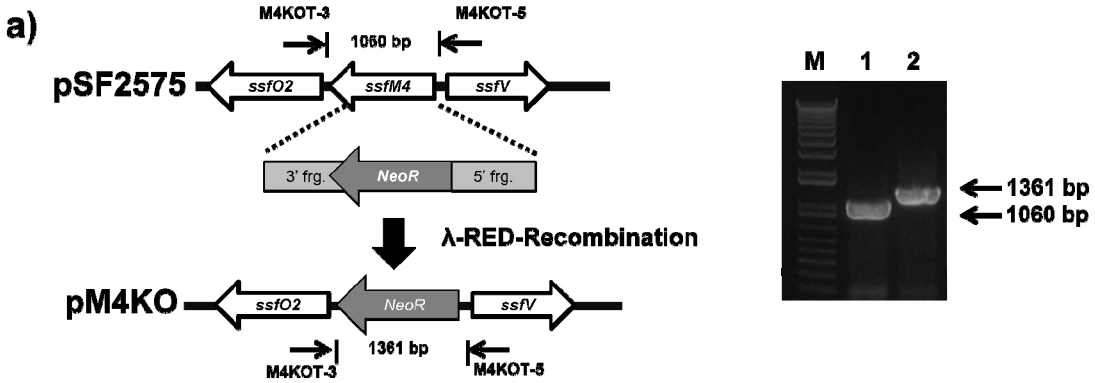
B. Construction of *ssfM1*-inactivation plasmid pM1KO. (a) Scheme of *ssfM1*-inactivation. pM1KO was generated by replacing a 993 bp *ssfM1* fragment with a 1425 bp *aadA* and *oriT* DNA cassette via lambda-RED-mediated recombination. (b) PCR confirmation of mutant plasmid. M, 1 kb plus DNA ladder ; lane 1) pSF2575; lane 2) pM1KO



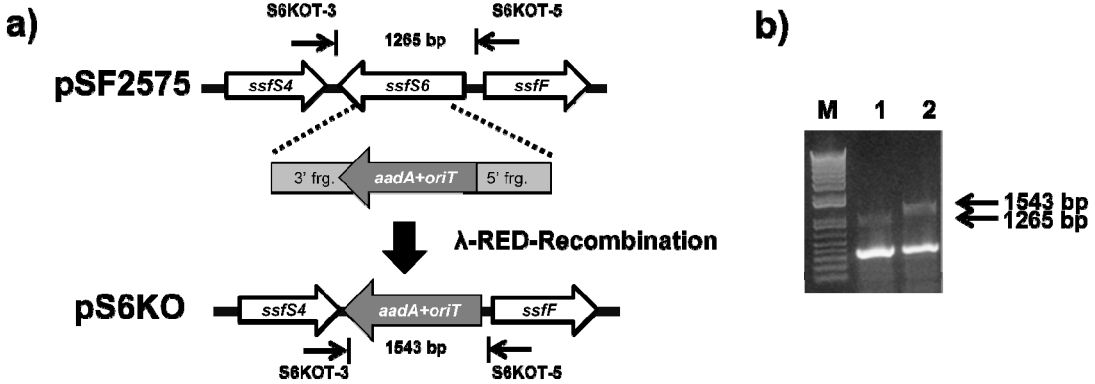
C. Construction of *ssfM2*-inactivation plasmid pM2KO. (a) Scheme of *ssfM2*-inactivation. pM2KO was generated by replacing a 990 bp *ssfM2* fragment with a 1425 bp *aadA* and *oriT* DNA cassette via lambda-RED-mediated recombination. (b) PCR confirmation of mutant plasmid. M, 1 kb plus DNA ladder ; lane 1) pSF2575; lane 2) pM2KO



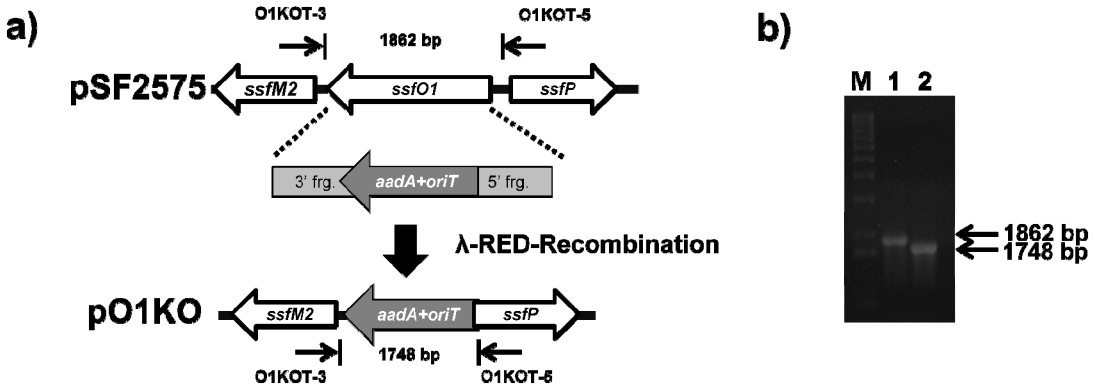
D. Construction of *ssfM3*-inactivation plasmid pM3KO. (a) Scheme of *ssfM3*-inactivation. pM3KO was generated by replacing a 852 bp *ssfM3* fragment with a 1096 bp *NeoR* DNA cassette via lambda-RED-mediated recombination. (b) PCR confirmation of mutant plasmid. M, 1 kb plus DNA ladder ; lane 1) pSF2575; lane 2) pM3KO



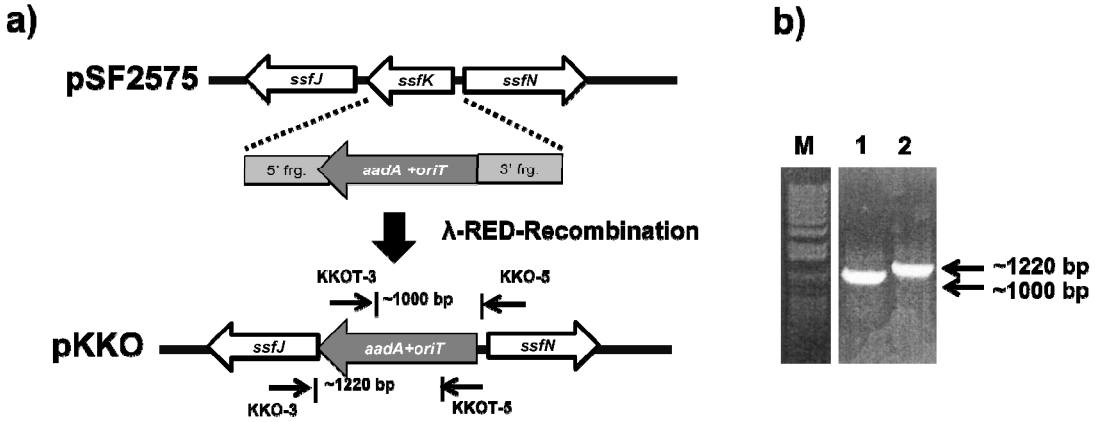
E. Construction of *ssfM4*-inactivation plasmid pM4KO. (a) Scheme of *ssfM4*-inactivation. pM4KO was generated by replacing a 795 bp *ssfM4* fragment with a 1096 bp *NeoR* DNA cassette via lambda-RED-mediated recombination. (b) PCR confirmation of mutant plasmid. M, 1 kb plus DNA ladder ; lane 1) pSF2575; lane 2) pM4KO



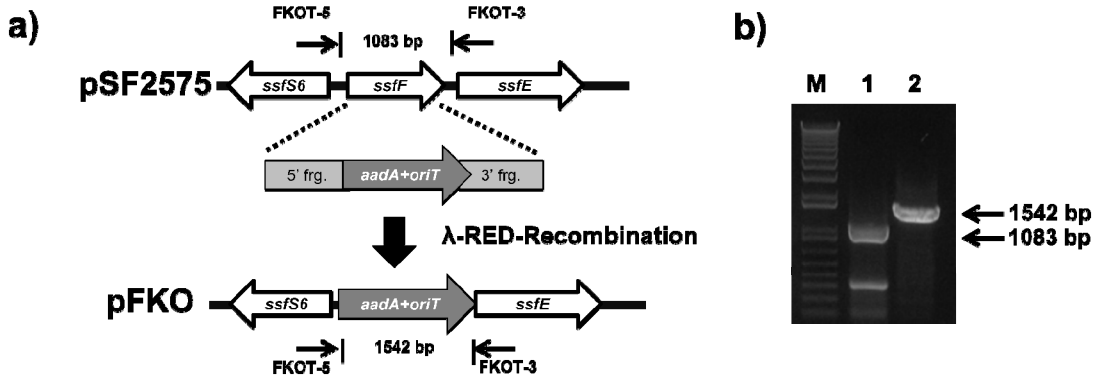
F. Construction of *ssfS6*-inactivation plasmid pS6KO. (a) Scheme of *ssfS6*-inactivation. pS6KO was generated by replacing a 1146 bp *ssfS6* fragment with a 1424 bp *aadA* and *oriT* DNA cassette via lamda-RED-mediated recombination. (b) PCR confirmation of mutant plasmid. M, 1 kb plus DNA ladder ; lane 1) pSF2575; lane 2) pS6KO



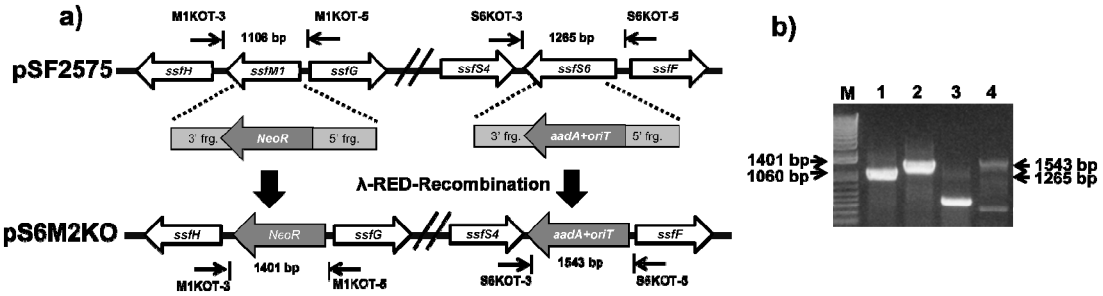
G. Construction of *ssfO1*-inactivation plasmid pO1KO. (a) Scheme of *ssfO1*-inactivation. pO1KO was generated by replacing a 1539 bp *ssfO1* fragment with a 1425 bp *aadA* and *oriT* DNA cassette via lamda-RED-mediated recombination. (b) PCR confirmation of mutant plasmid. M, 1 kb plus DNA ladder lane 1) pSF2575; lane 2) pO1KO



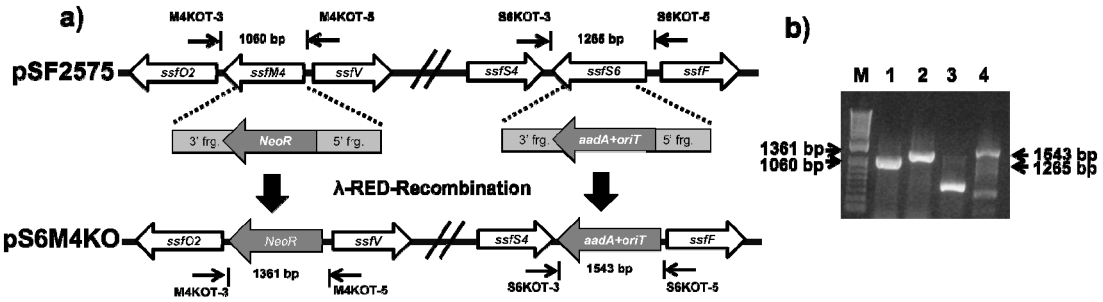
H. Construction of *ssfK*-inactivation plasmid pKKO. (a) Scheme of *ssfK*-inactivation. pKKO was generated by replacing a 651 bp *ssfK* fragment with a 1425 bp *aadA* and *oriT* DNA cassette via lambda-RED-mediated recombination. (b) PCR confirmation of mutant plasmid. M, 1 kb plus DNA ladder ; lane 1) pKKOL; lane 2) pKKOR.



I. Construction of *ssfF*-inactivation plasmid pFKO. (a) Scheme of *ssfF*-inactivation. pFKO was generated by replacing a 966 bp *ssfF* fragment with a 1425 bp *aadA* and *oriT* DNA cassette via lambda-RED-mediated recombination. (b) PCR confirmation of mutant plasmid. M, 1 kb plus DNA ladder ; lane 1) pSF2575; lane 2) pFKO



J. Construction of *ssfS6M1*-inactivation plasmid pS6M1KO. (a) Scheme of *ssfS6M1*-inactivation. pM1KO was generated by replacing a 801 bp *ssfM1* fragment with a 1096 bp *aadA* and *oriT* DNA cassette on pS6KO via lambda-RED-mediated recombination. (b) PCR confirmation of mutant plasmid. M, 1 kb plus DNA ladder ; lane 1) pSF2575 with M1KOT primers; lane 2) pS6M1KO with M1KOT primers; 3) pSF2575 with S6KOT primers; 4) pS6M1KO with S6KOT primers.



K. Construction of *ssfS6M4*-inactivation plasmid pM4KO. (a) Scheme of *ssfS6M4*-inactivation. pS6M4KO was generated by replacing a 795 bp *ssfM4* fragment with a 1096 bp *NeoR* DNA cassette on pS6KO via lambda-RED-mediated recombination. (b) PCR confirmation of mutant plasmid. M, 1 kb plus DNA ladder ; lane 1) pSF2575 with M4KOT primers; lane 2) pS6M4KO with M4KOT primers; 3) pSF2575 with S6KOT primers; 4) pS6M4KO with S6KOT primers.

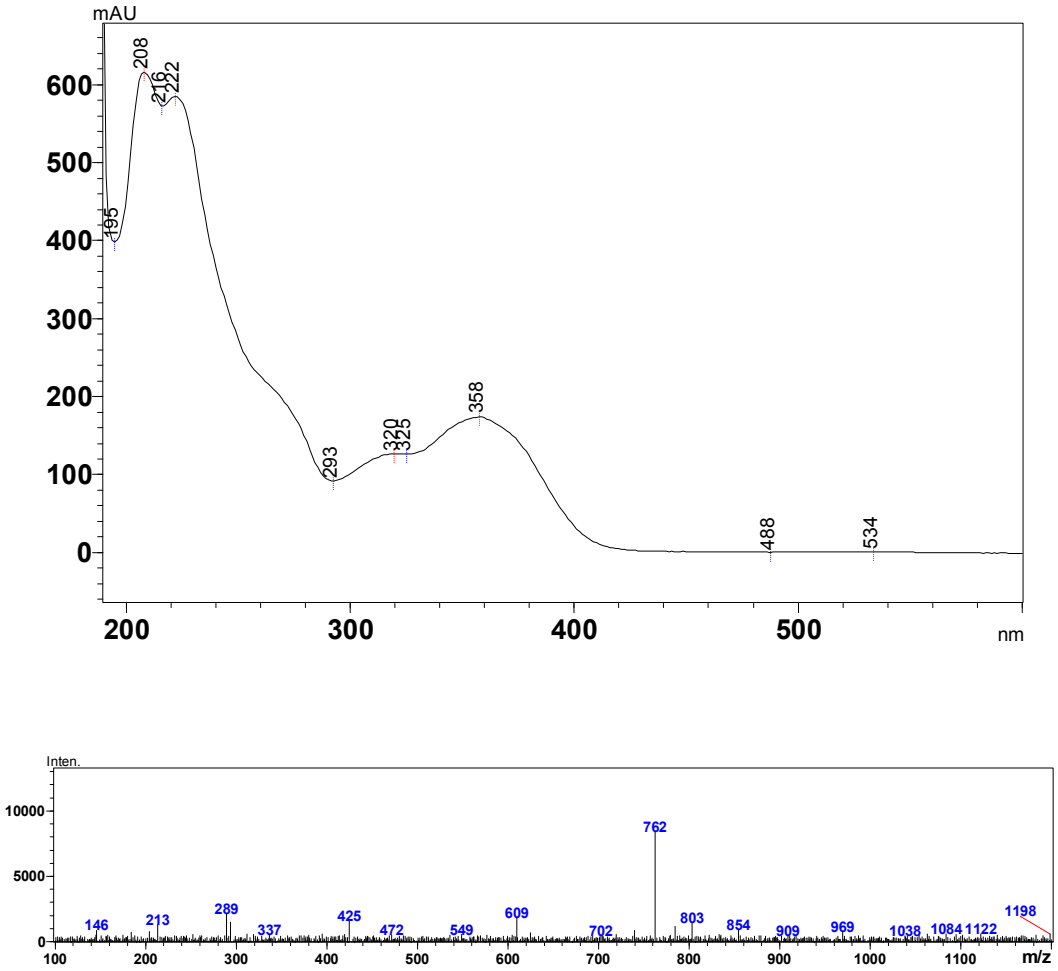


Figure S3.1. UV spectrum and MS Measured during LC-MS for **22**.

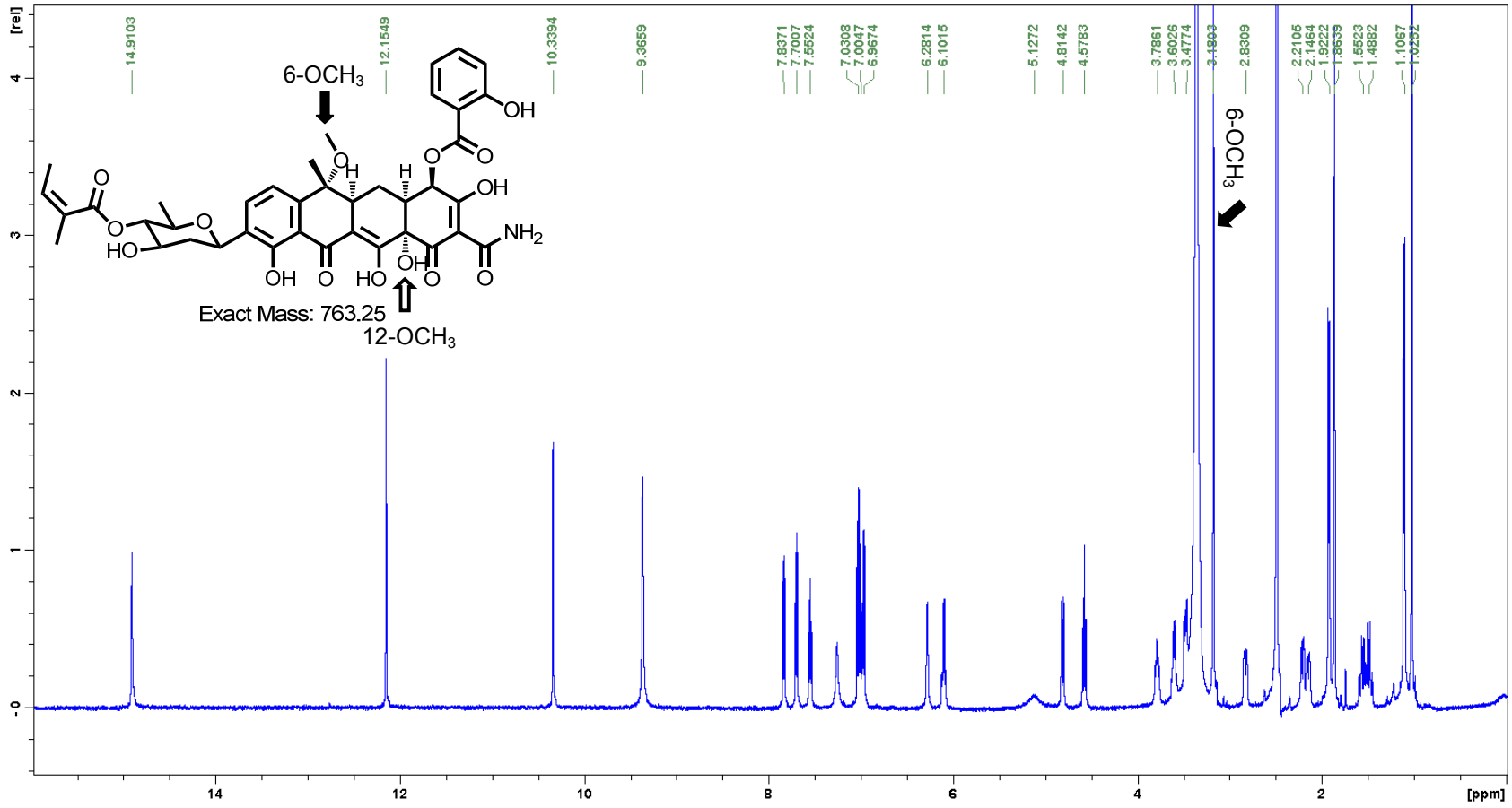


Figure S3.2 ^{1}H NMR spectrum of **22**. Measured in DMSO-d_6 at 500MHz

12-OCH_3 signal is missing in ^{1}H NMR spectrum of **22** (empty arrow), while 6-OCH_3 signal is clearly observed (black arrow).

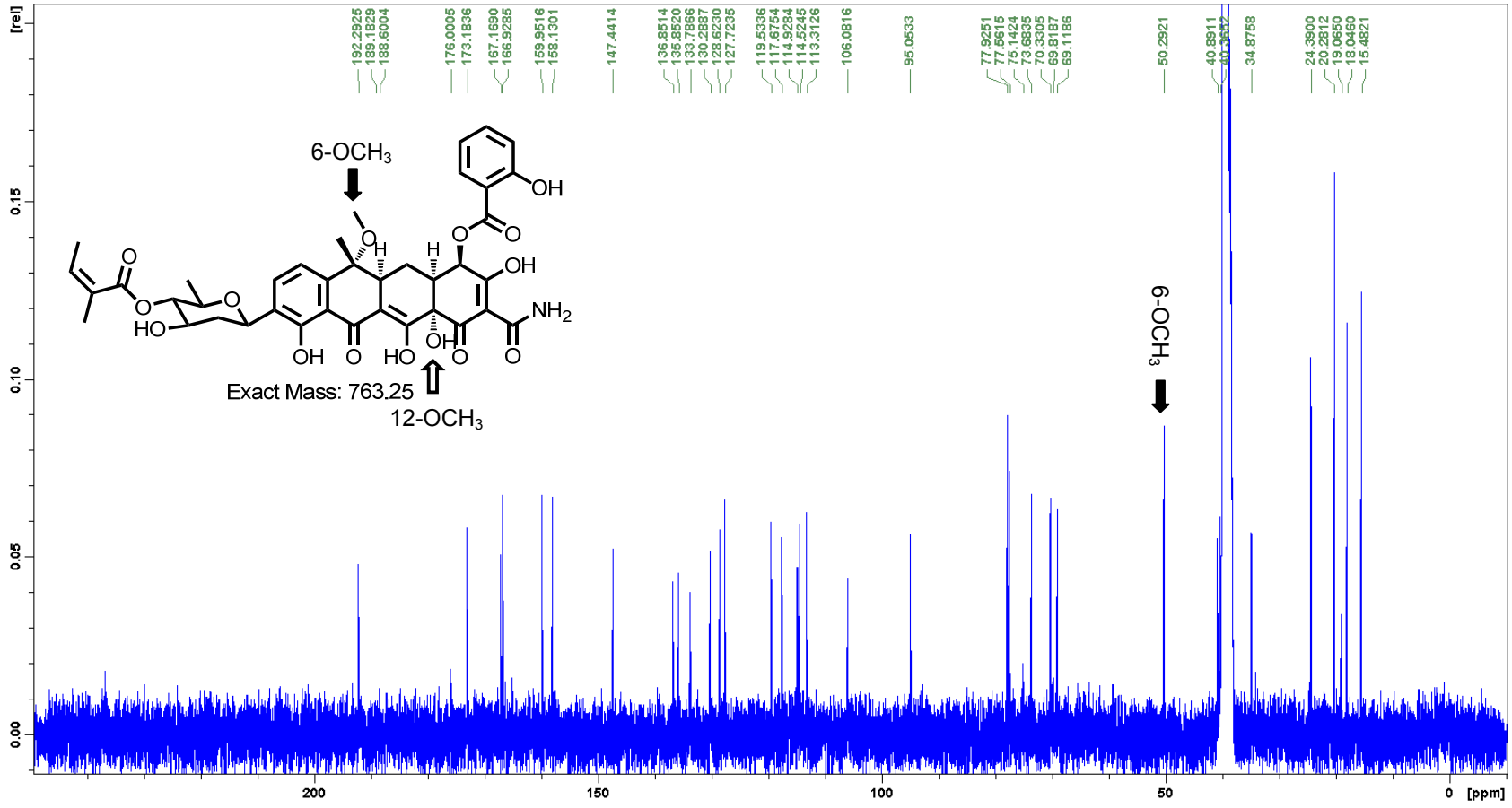


Figure S3.3. ^{13}C NMR spectrum of **22**. Measured in DMSO-d_6 at 125MHz

12-OCH_3 signal is missing in ^{13}C NMR spectrum of **22** (empty arrow), while 6-OCH_3 signal is clearly observed (black arrow).

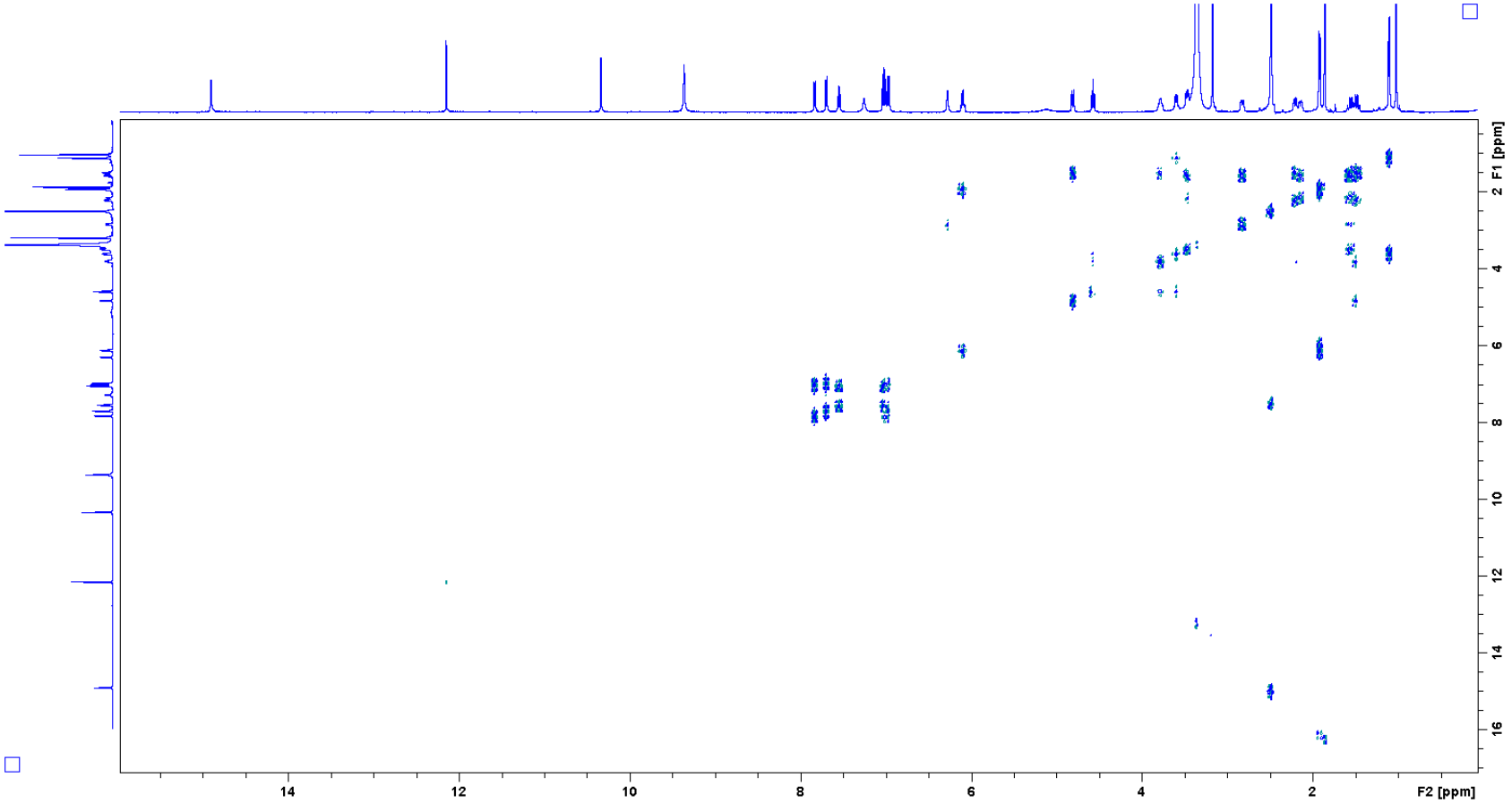


Figure S3.4. ^{1}H-^{1}H COSY NMR spectrum of **22**.

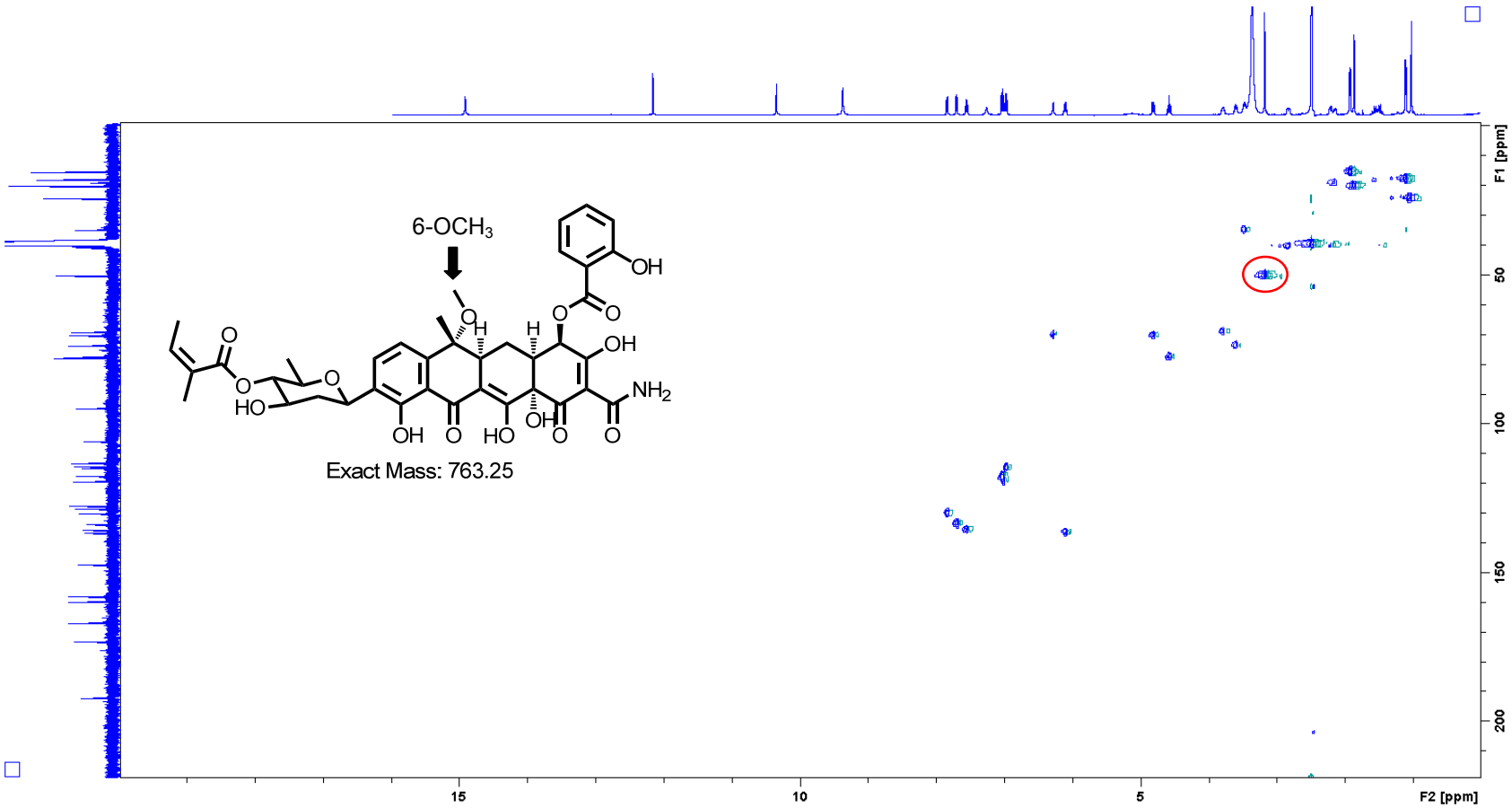


Figure S3.5. ^{1}H-^{13}C HMQC NMR spectrum of **22**.

Correlation between 6-OCH_3 and its proton is showed in red circle.

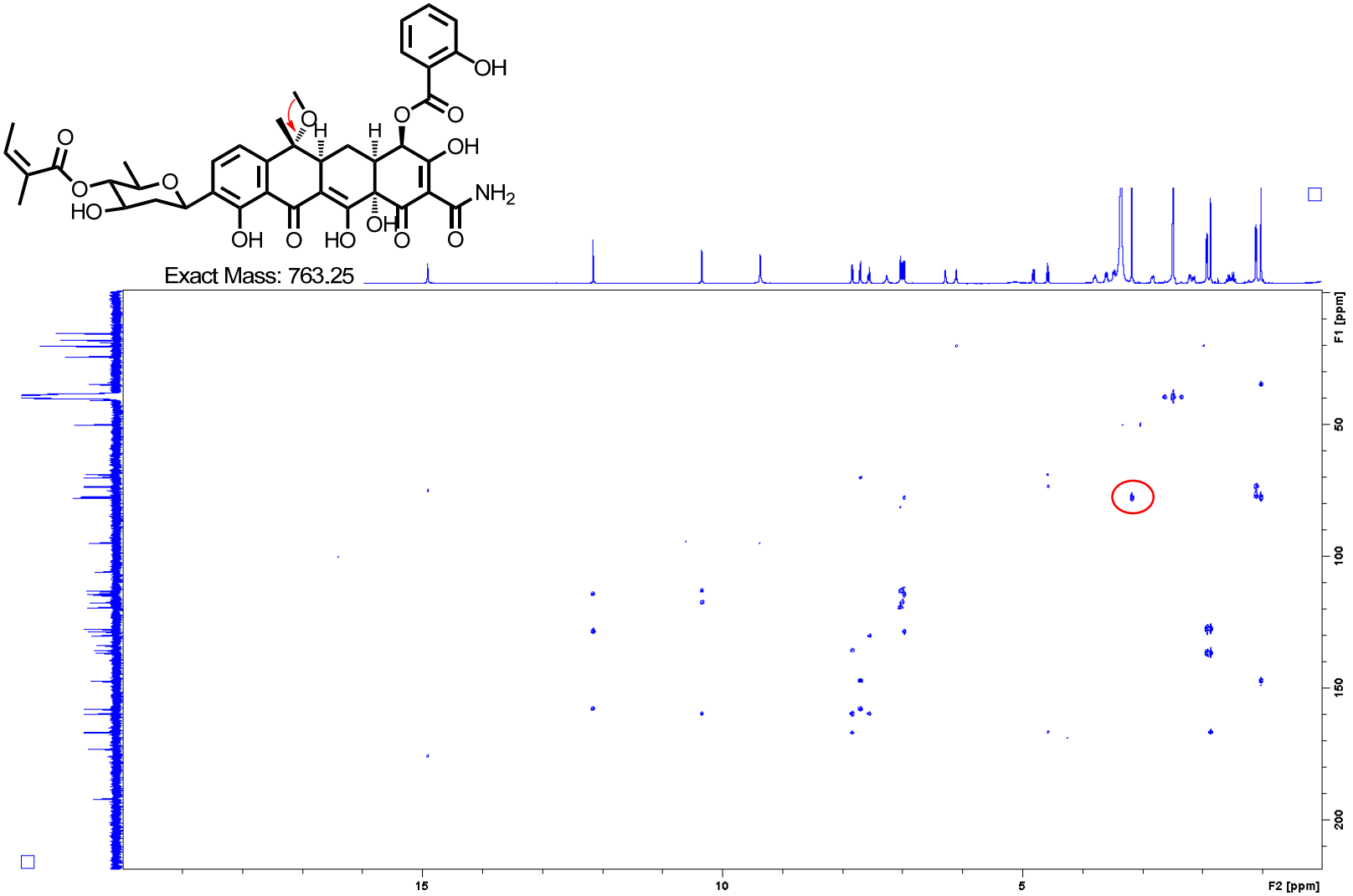


Figure S3.6. ^{1}H-^{13}C HMBC N.MR spectrum of **22**

Correlation between 6-OCH_3 and C6 (in red circle) establishes the position of the methyl group.

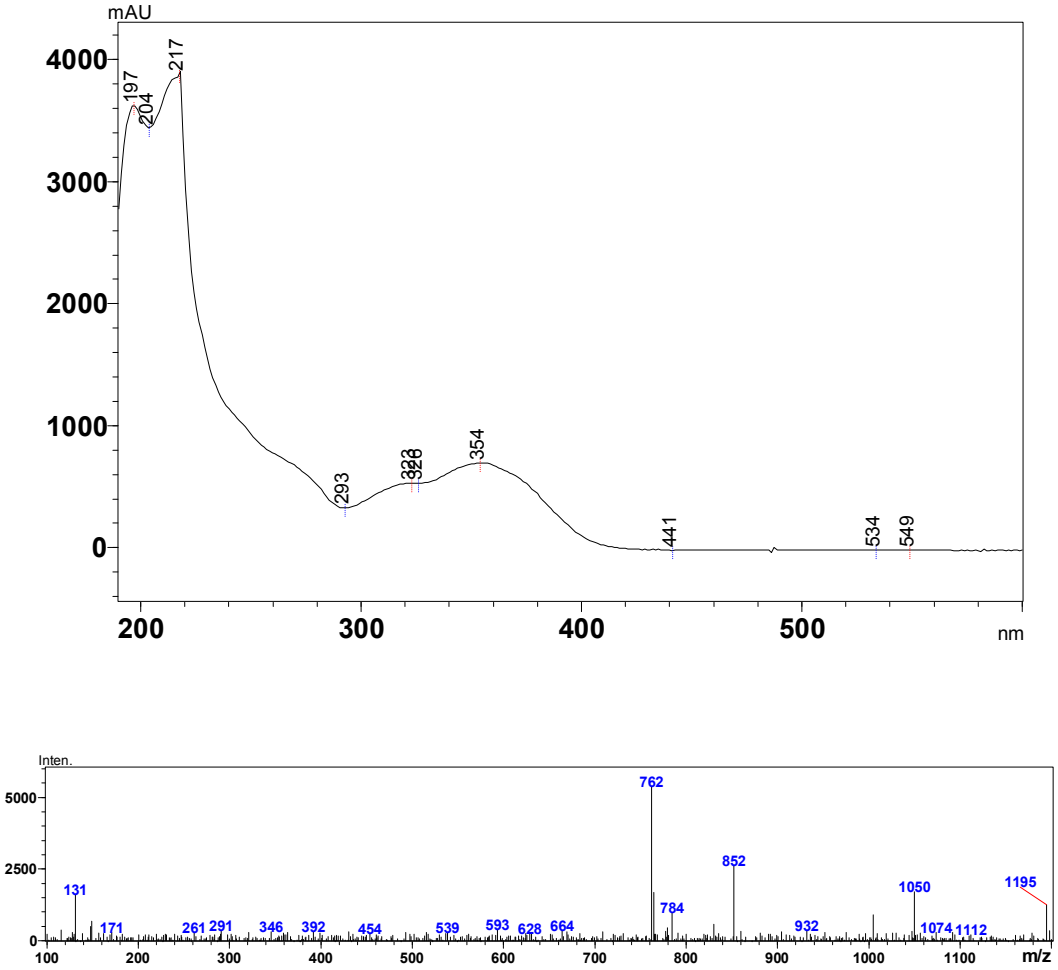


Figure S4.1. UV spectrum and MS Measured during LC-MS for **25**.

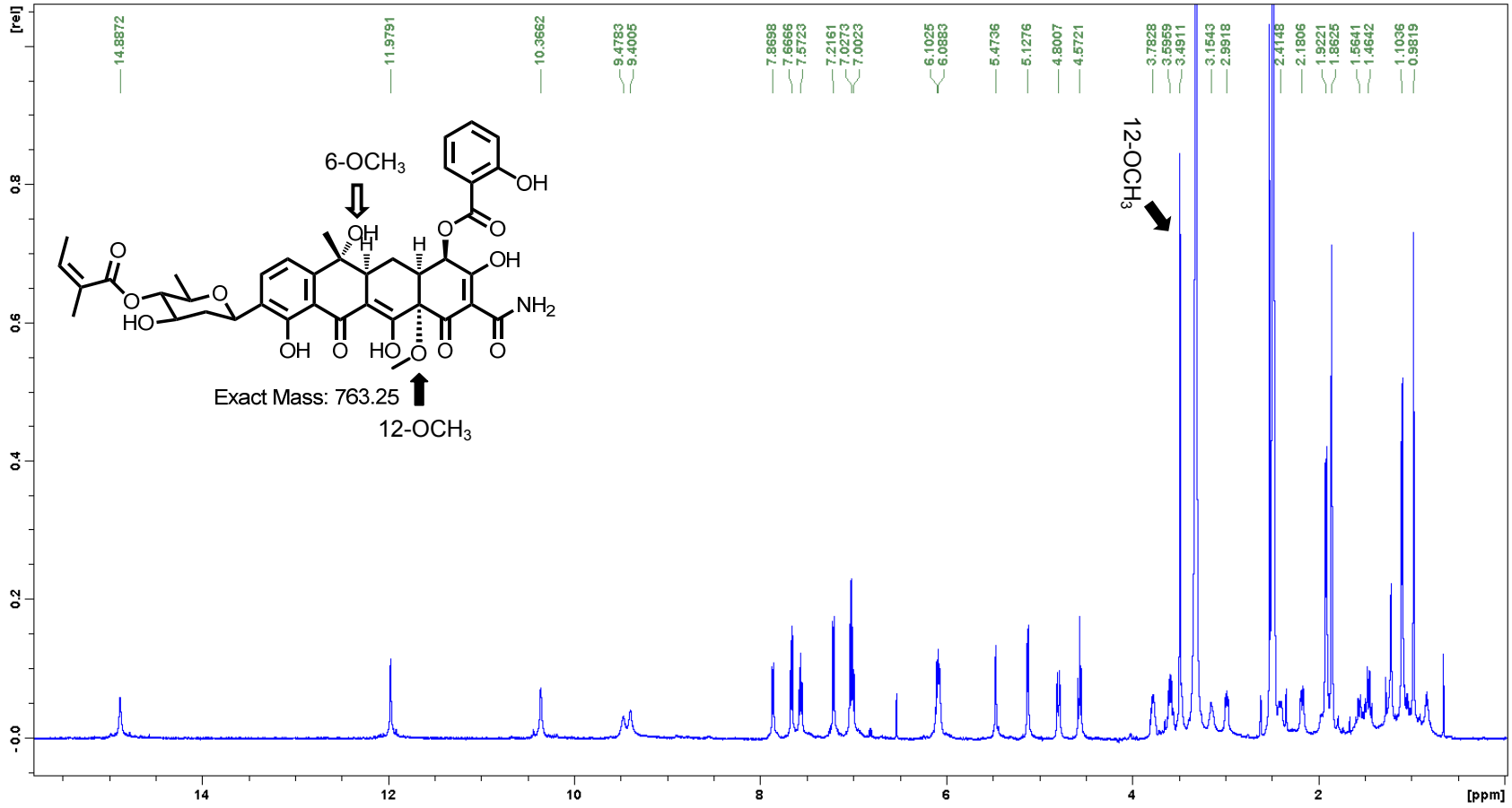


Figure S4.2. ^{1}H NMR spectrum of **25**. Measured in DMSO-d_6 at 500MHz

6-OCH_3 signal is missing in ^{1}H NMR spectrum of **25** (empty arrow), while 12-OCH_3 signal is clearly observed (black arrow).

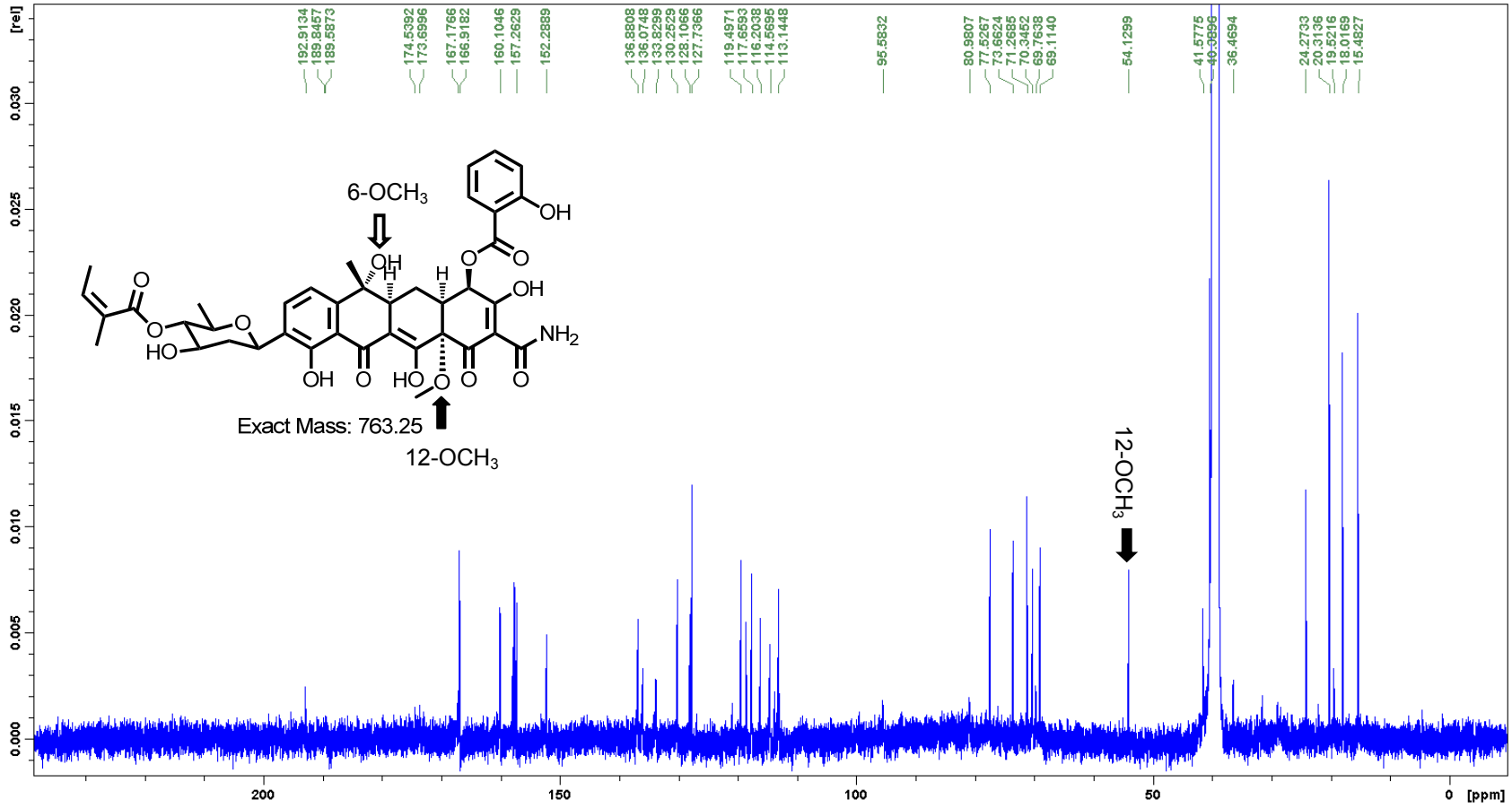


Figure S4.3. ^{13}C NMR spectrum of **25**. Measured in DMSO-d_6 at 125MHz

6-OCH_3 signal is missing in ^{13}C NMR spectrum of **25** (empty arrow), while 12-OCH_3 signal is clearly observed (black arrow).

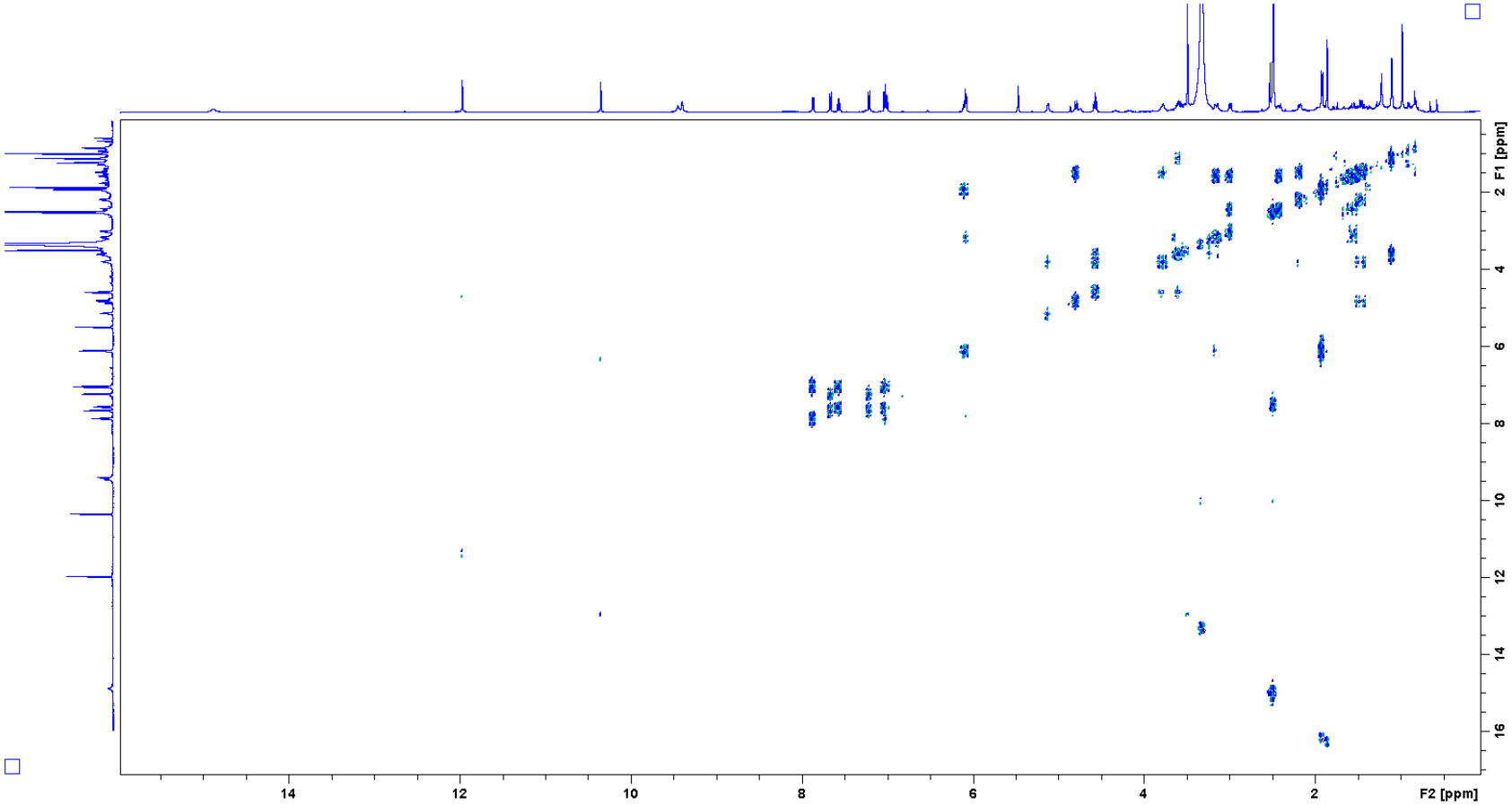


Figure S4.4. ^{1}H-^{1}H COSY NMR spectrum of **25**.

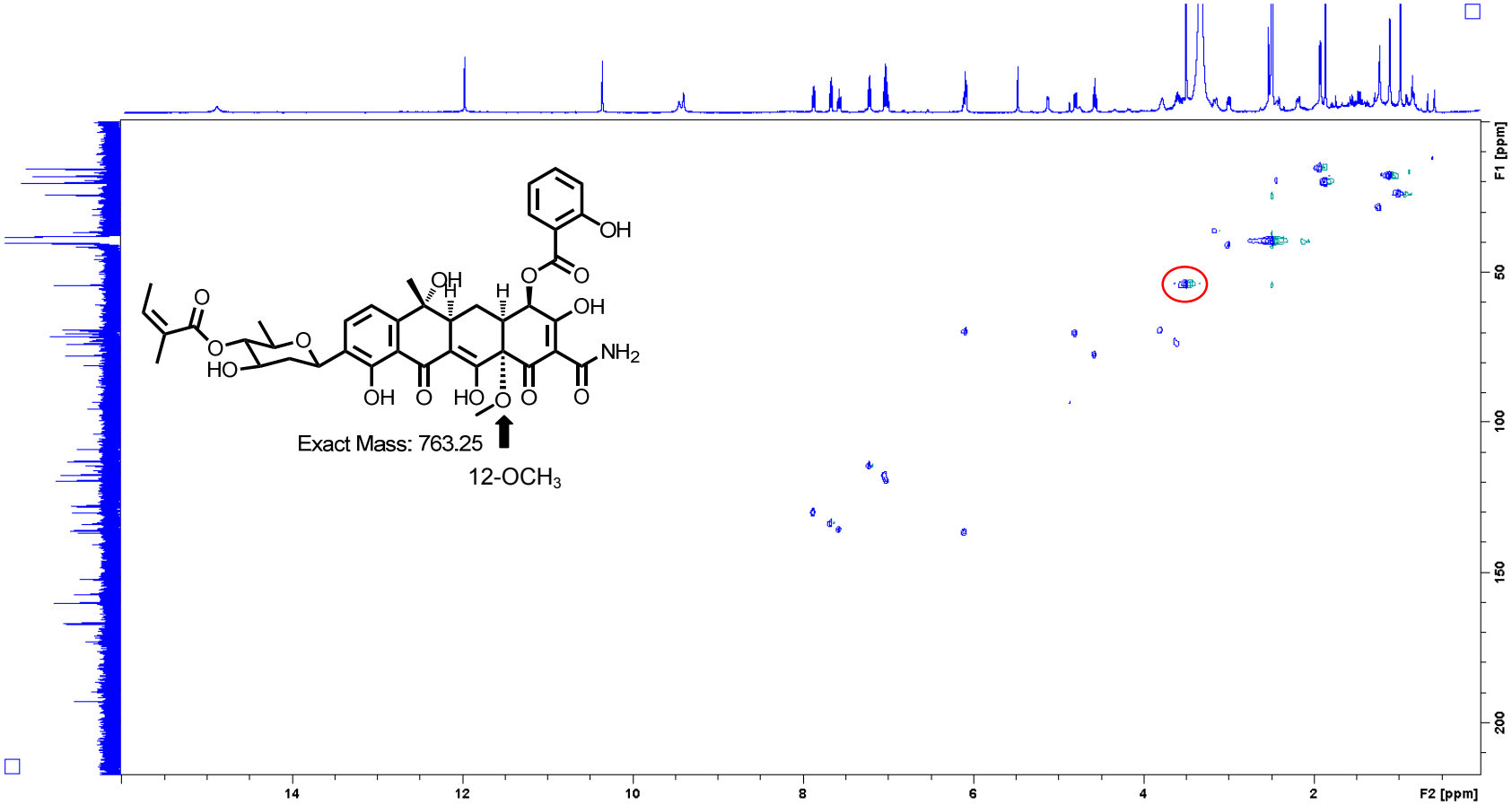


Figure S4.5. ^{1}H-^{13}C HMQC NMR spectrum of **25**.

Correlation between 12-OCH_3 and its proton is showed in red circle.

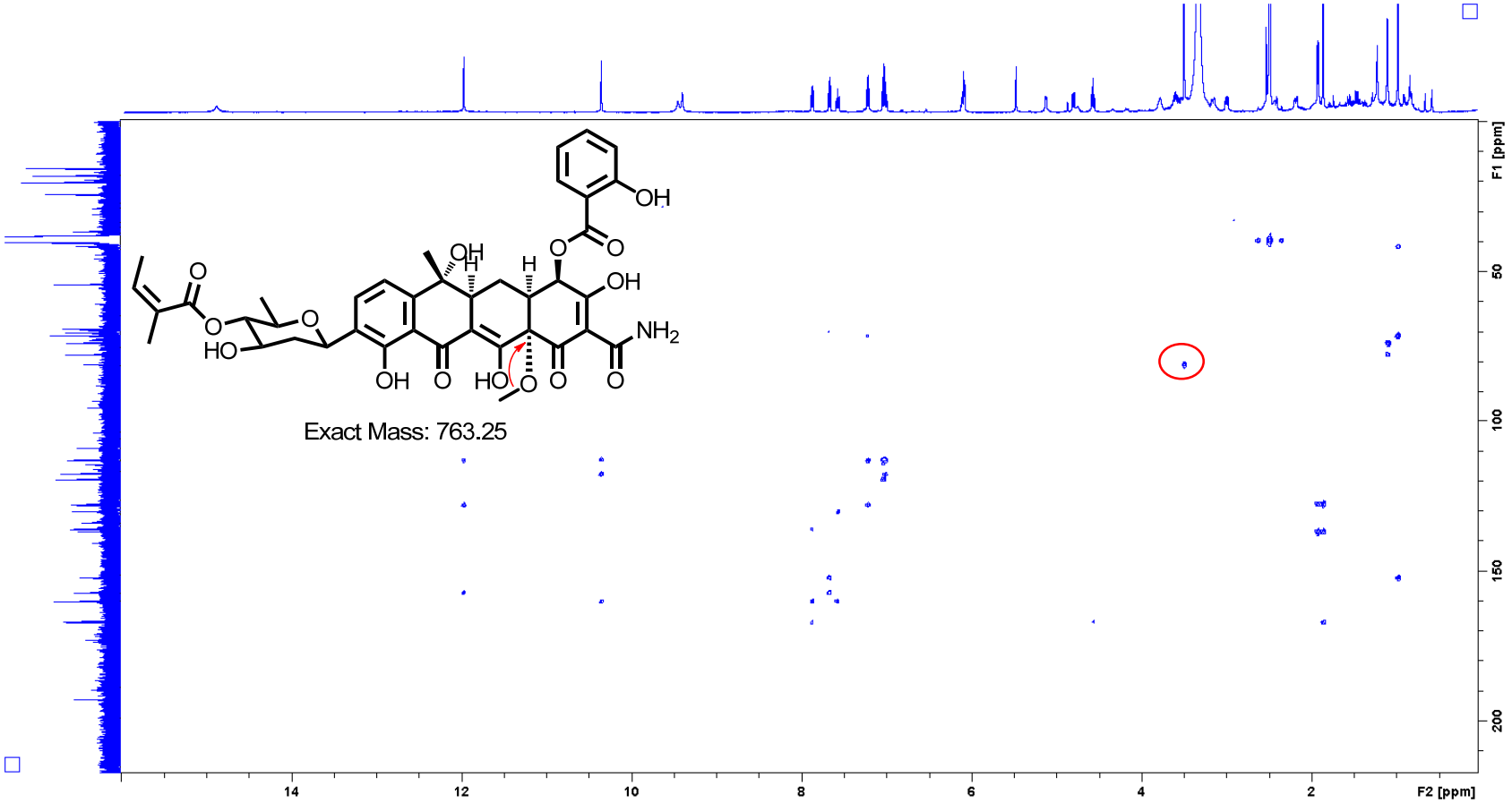


Figure S4.6. ^{1}H-^{13}C HMBC NMR spectrum of **25**.

Correlation between 12-OCH_3 and C12a (in red circle) establishes the position of the methyl group.

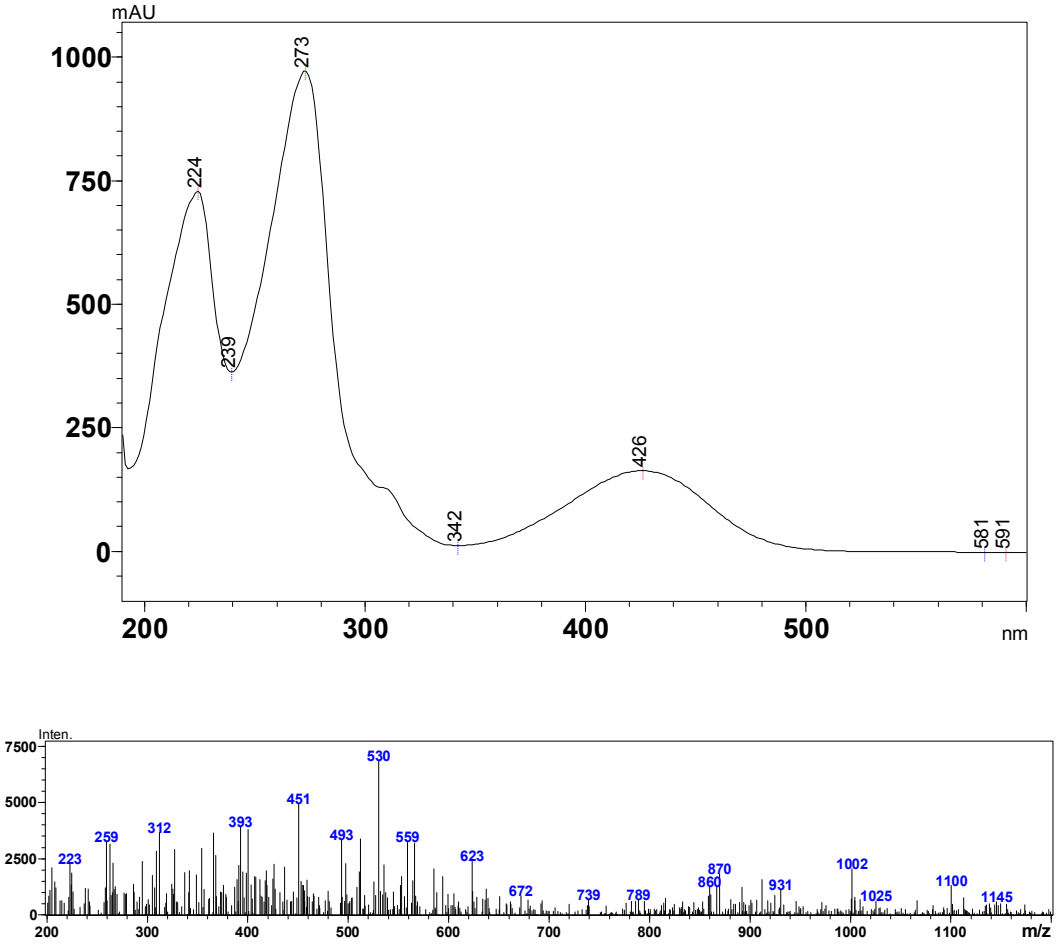


Figure S5.1. UV spectrum and MS Measured during LC-MS for **28**.

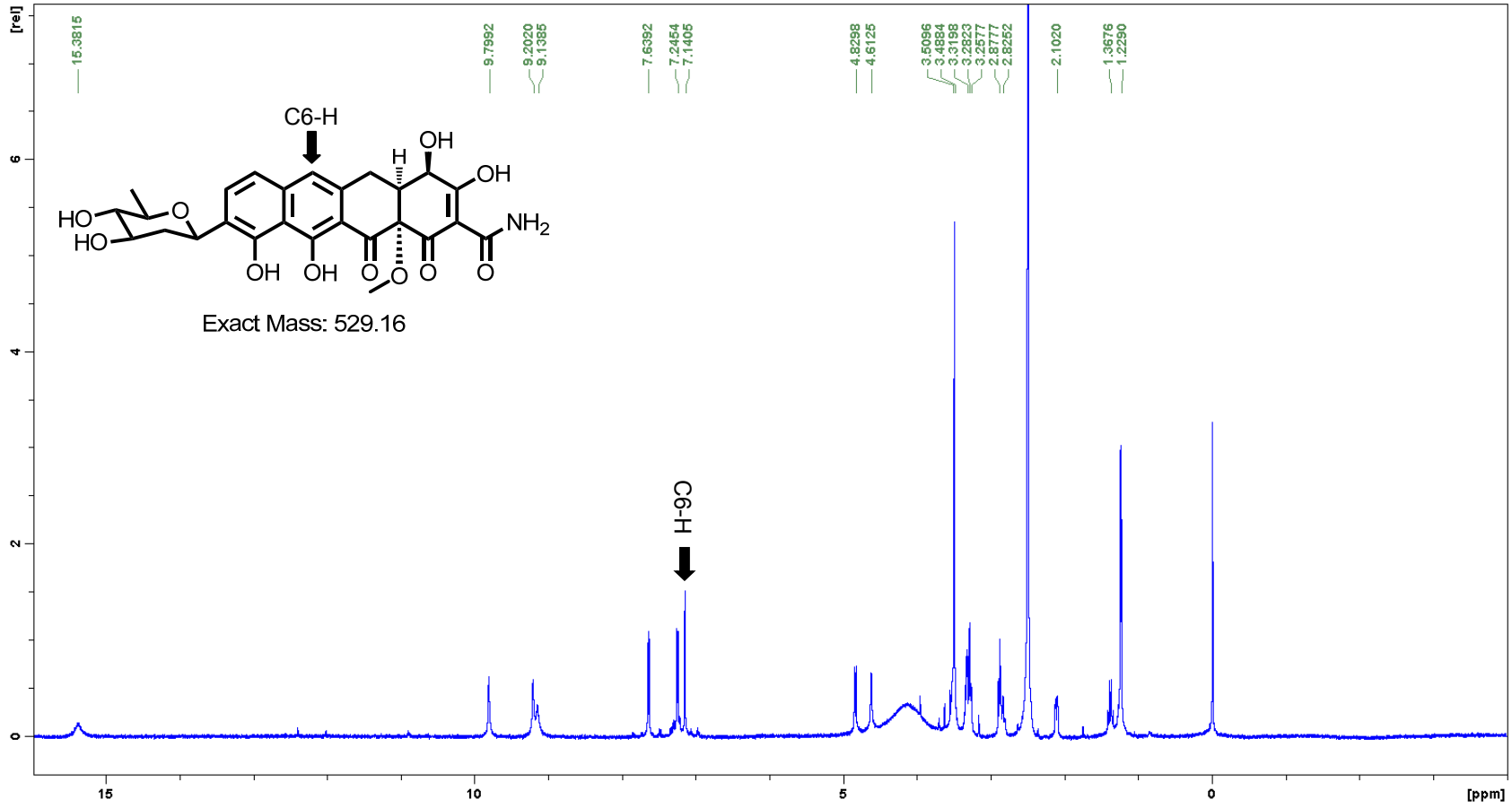


Figure S5.2. ^{1}H NMR spectrum of **28**. Measured in DMSO-d_6 at 500MHz

The proton signal at C14 is missing, while C6-H signal is clearly observed (black arrow).

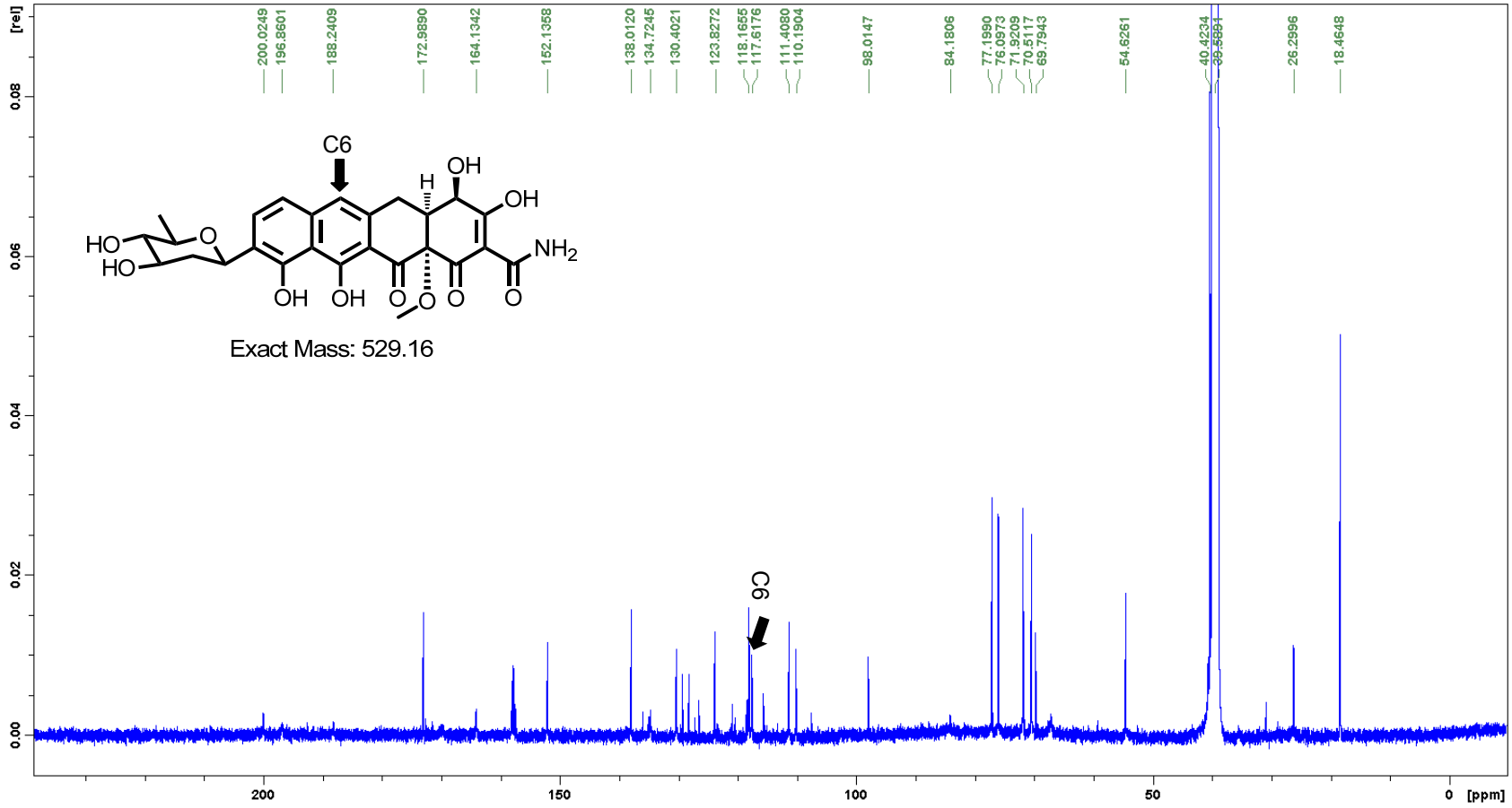


Figure S5.3. ^{13}C NMR spectrum of **28**. Measured in DMSO-d_6 at 125MHz

C6 carbon signal is clearly observed (black arrow).

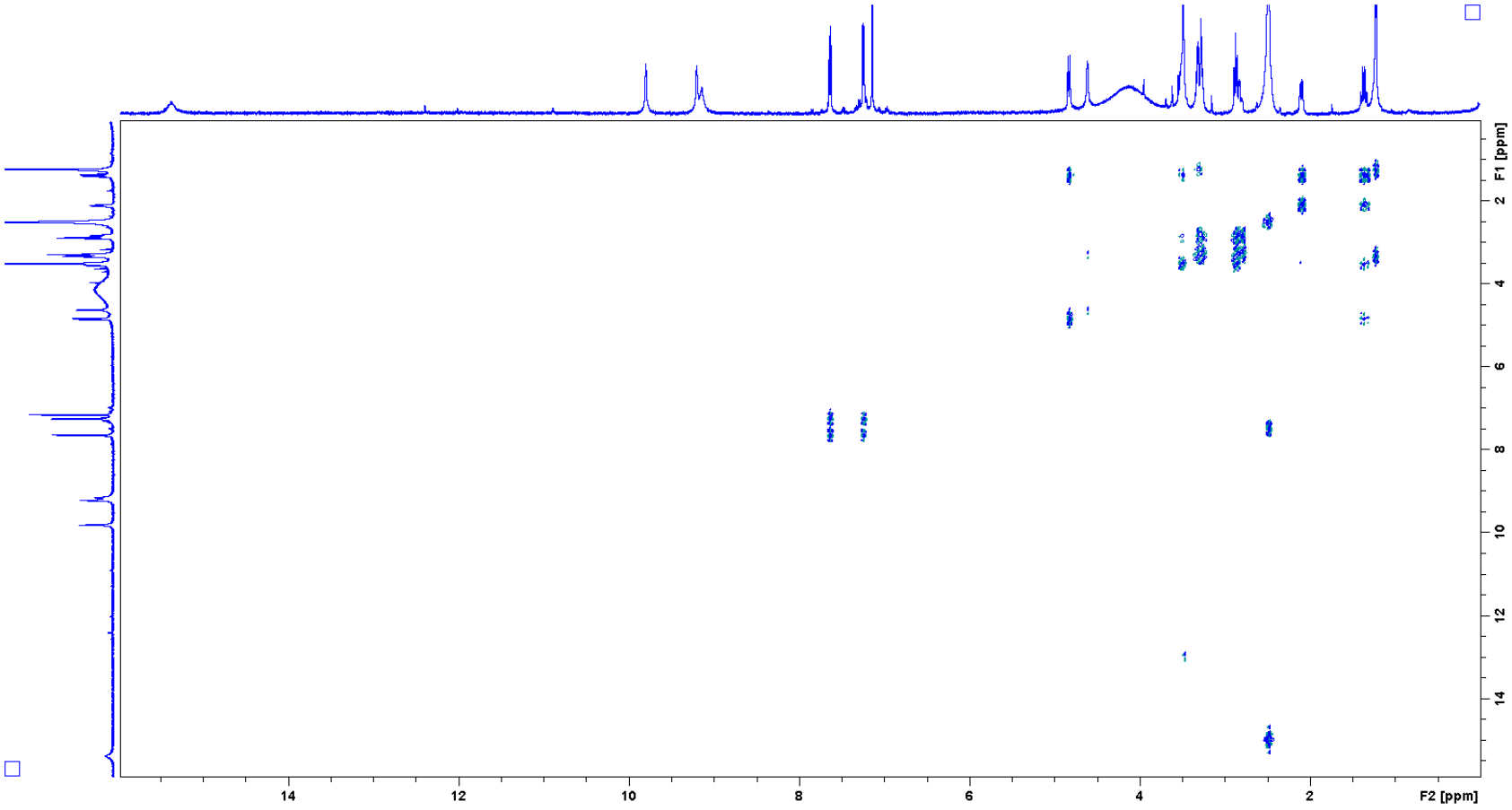


Figure S5.4. ^{1}H-^{1}H COSY spectrum of **28**.

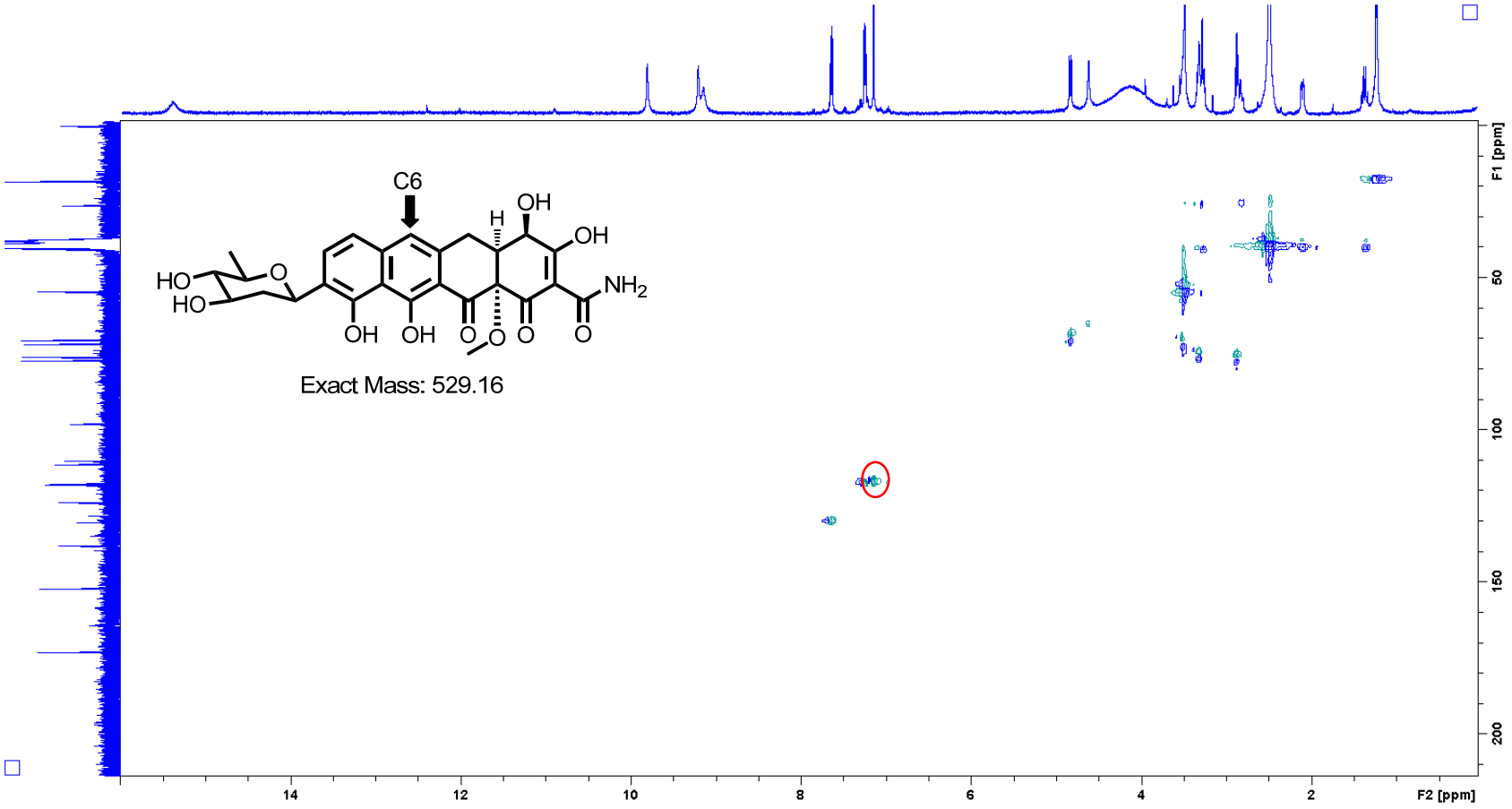


Figure S5.5. HMQC spectrum of **28**.

Correlation between C6 and its proton is showed in red circle.

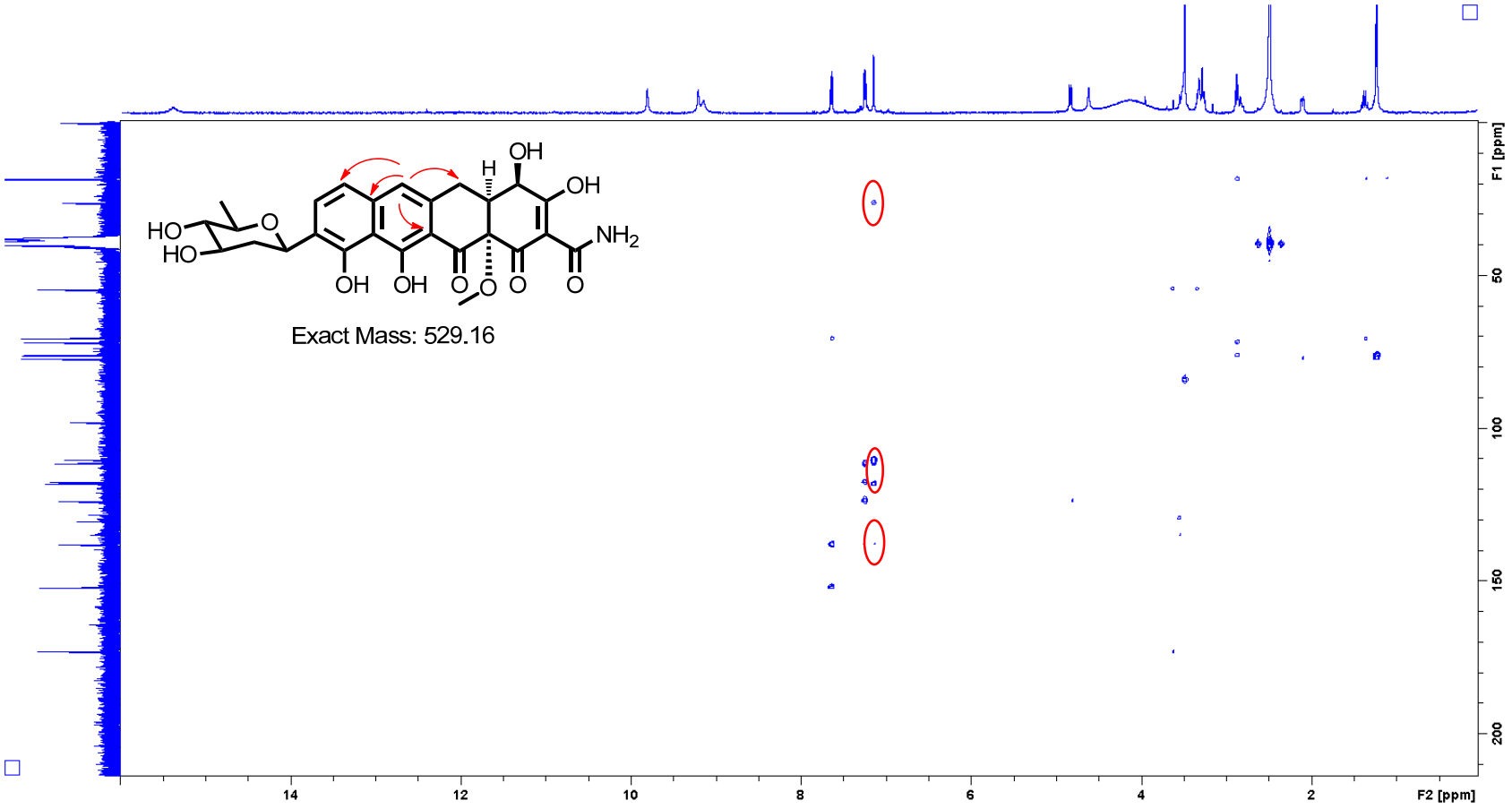


Figure S5.6. HMBC spectrum of **28**.

Correlations between C6-H and C5, C6a, C7, C11a establishes the position of C6-H.

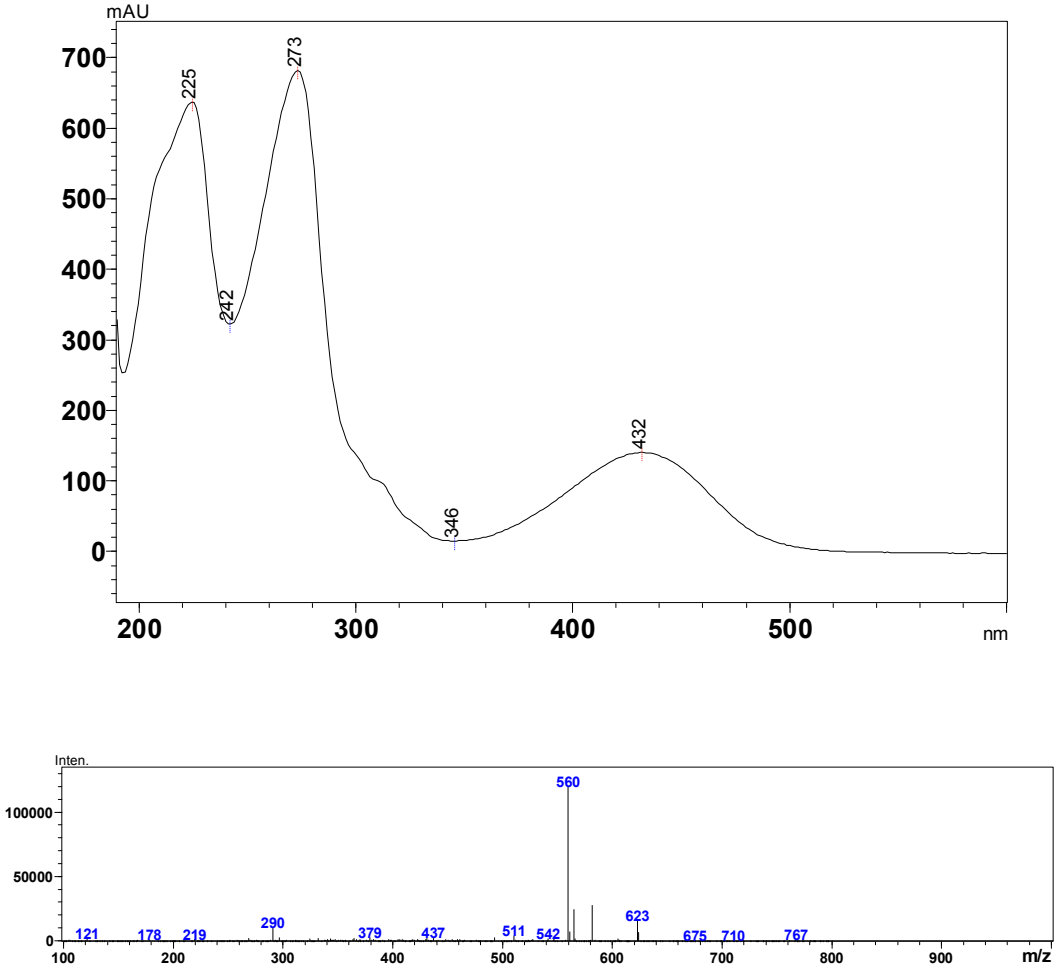


Figure S6.1. UV spectrum and MS Measured during LC-MS for **29**.

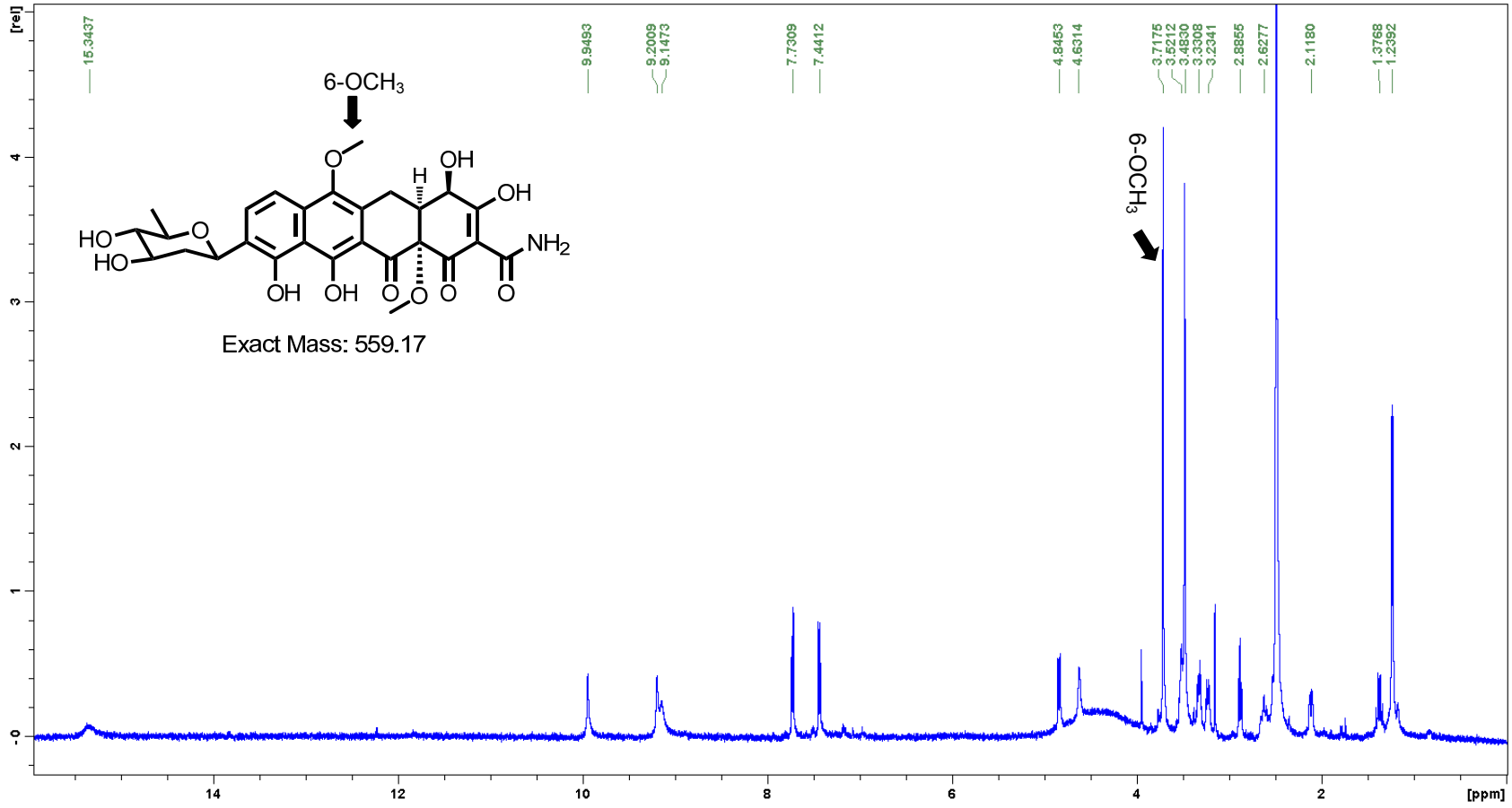


Figure S6.2. ^{1}H NMR spectrum of **29**. Measured in DMSO-d_6 at 500MHz

6-OCH_3 signal is clearly observed (black arrow).

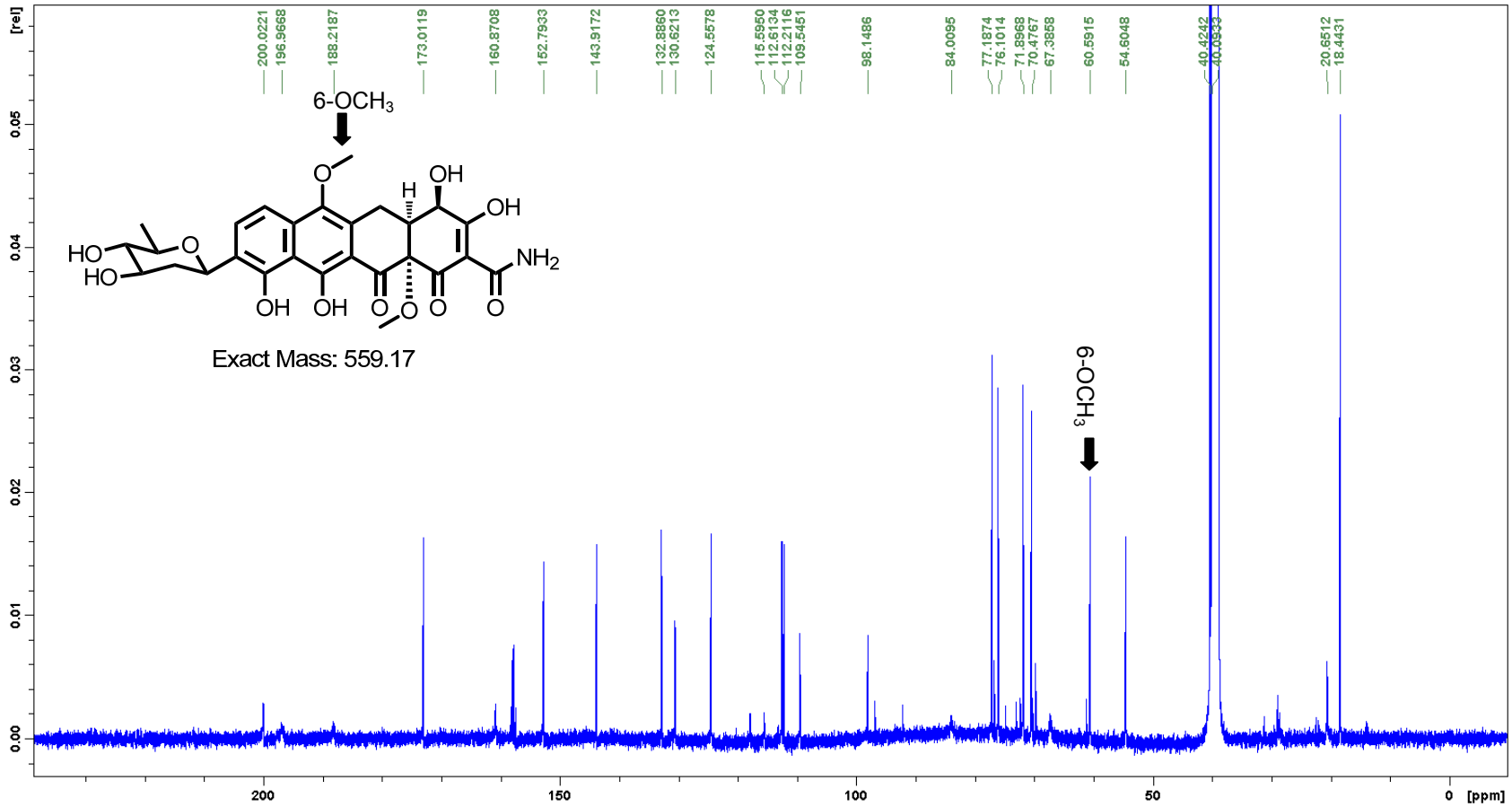


Figure S6.3. ^{13}C NMR spectrum of **29**. Measured in DMSO-d_6 at 125MHz
6-OCH_3 signal is clearly observed (black arrow).

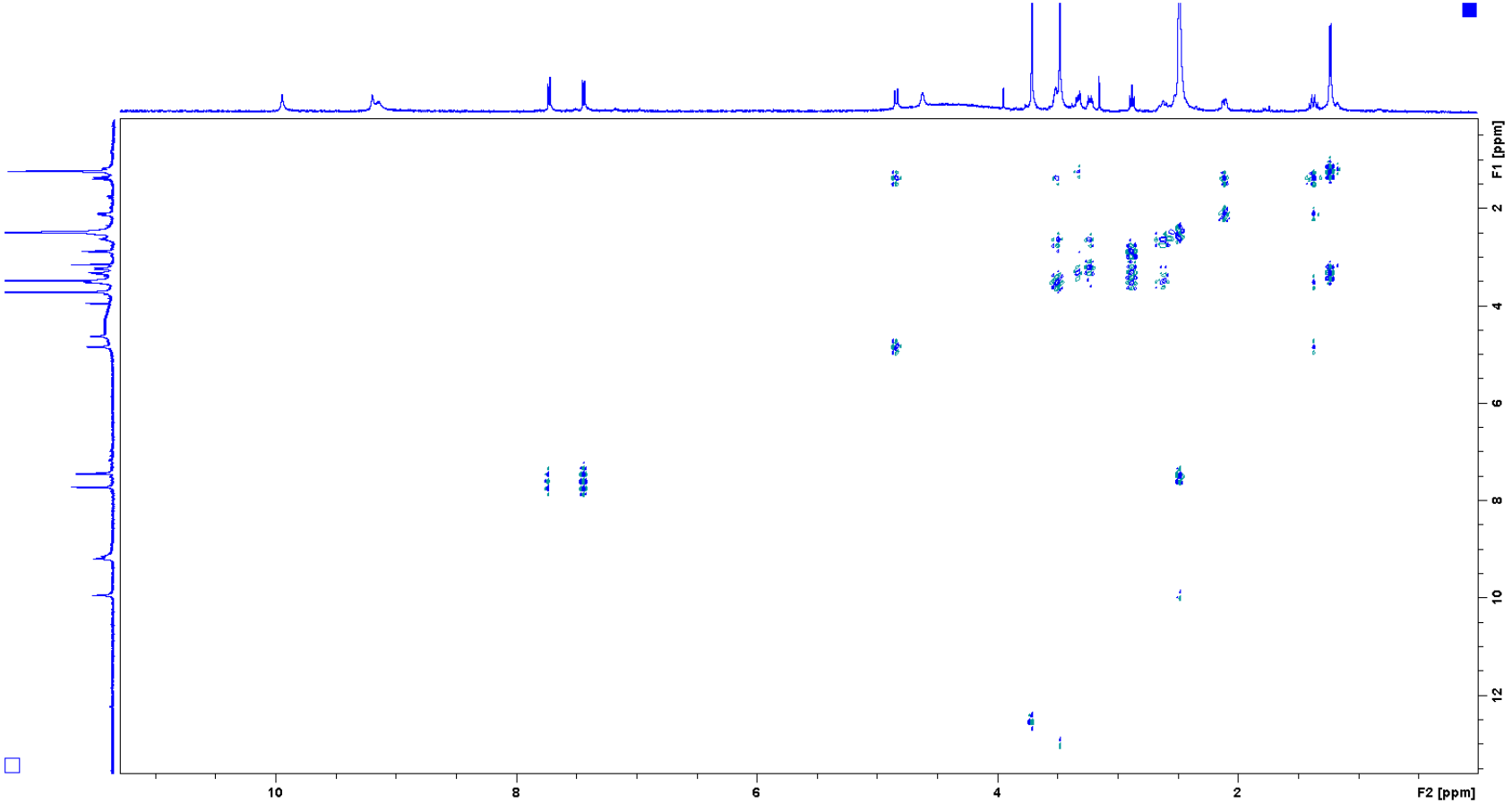


Figure S6.4. ^{1}H-^{1}H COSY spectrum of **29**.

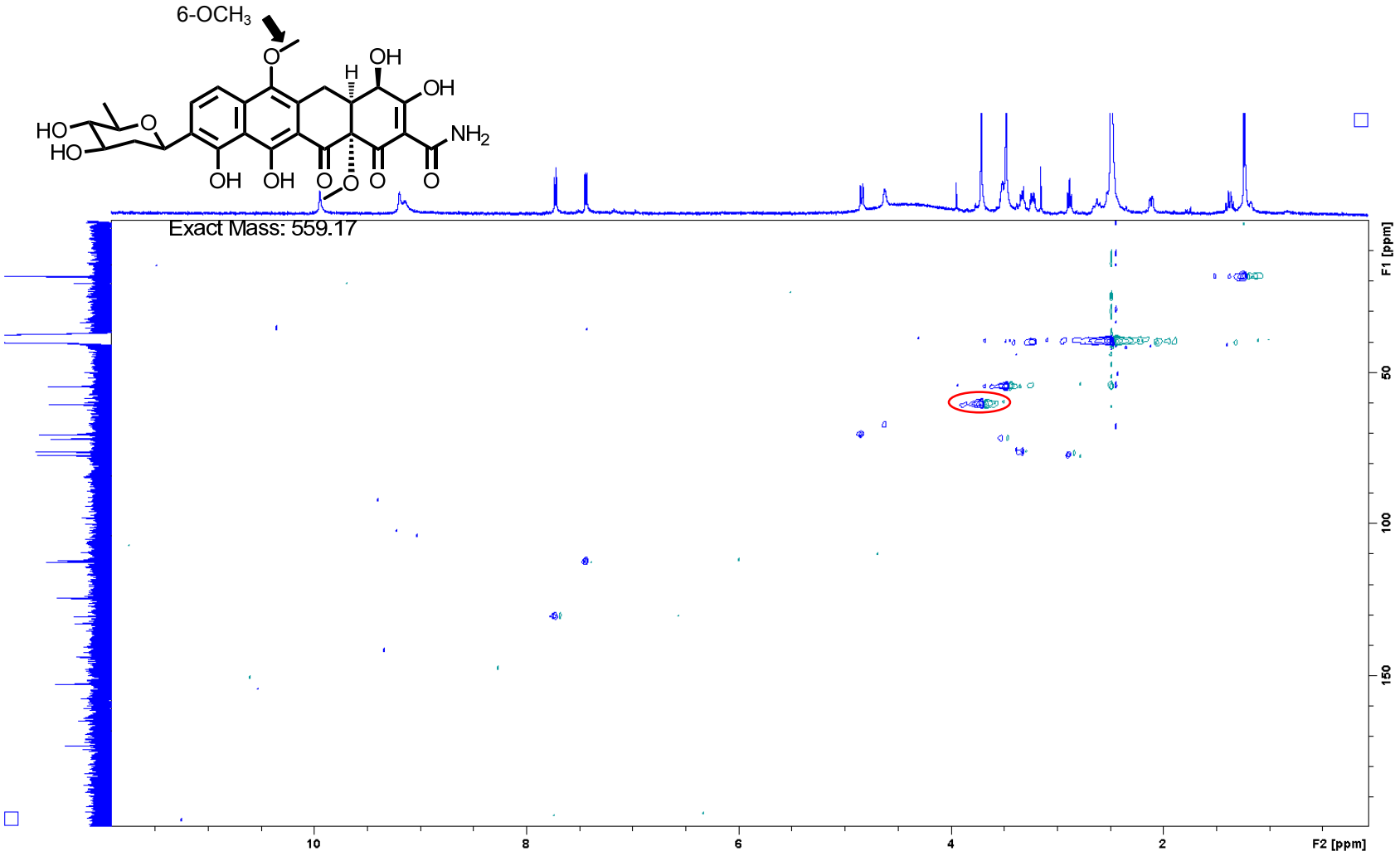


Figure S6.5. HMQC spectrum of **29**

Correlation between 6-OCH_3 and its proton is showed in red circle.

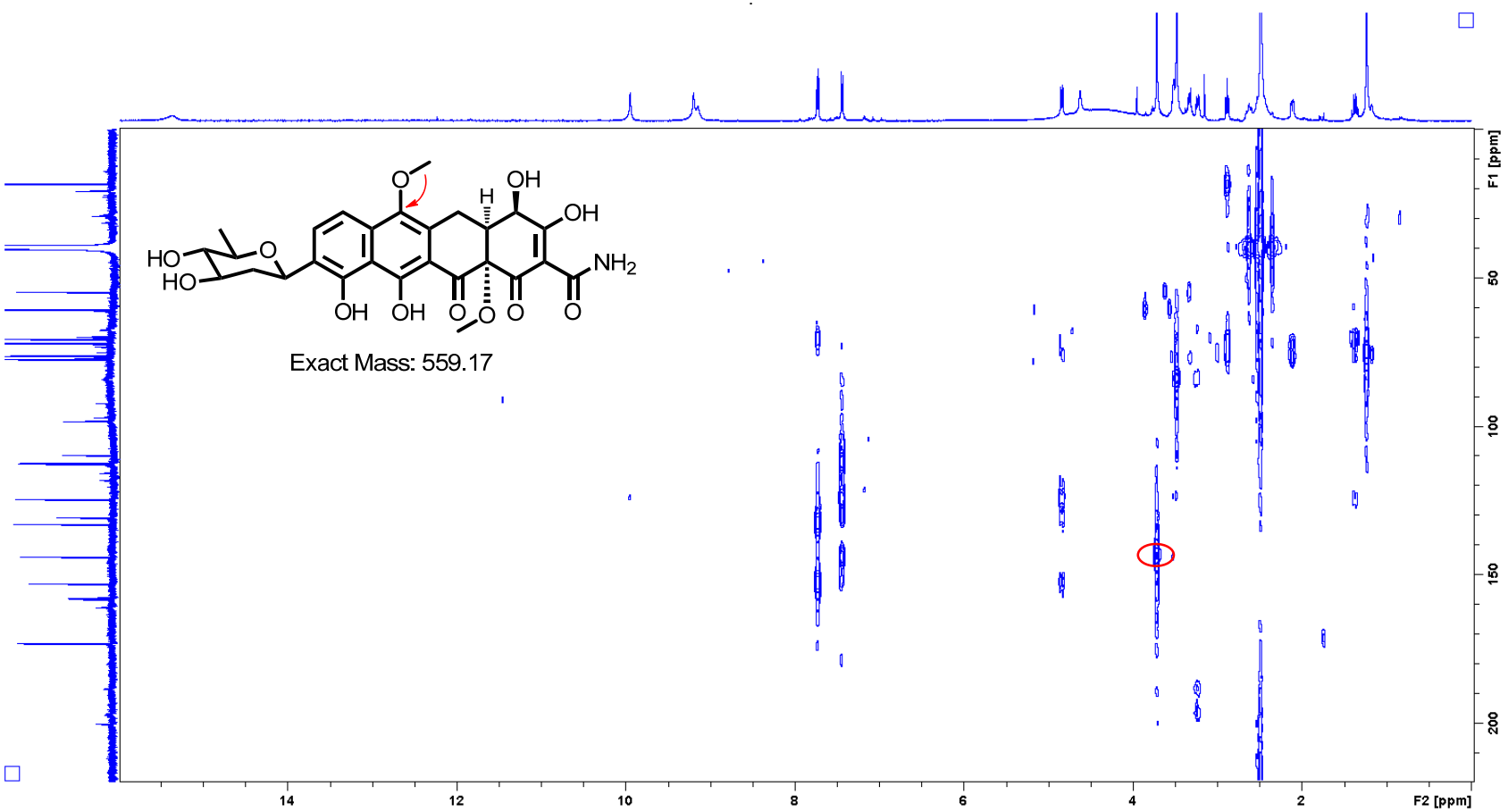


Figure S6.6. HMBC spectrum of **29**.

Correlations between 6-OCH_3 and C6 establishes the position of 6-OCH_3.

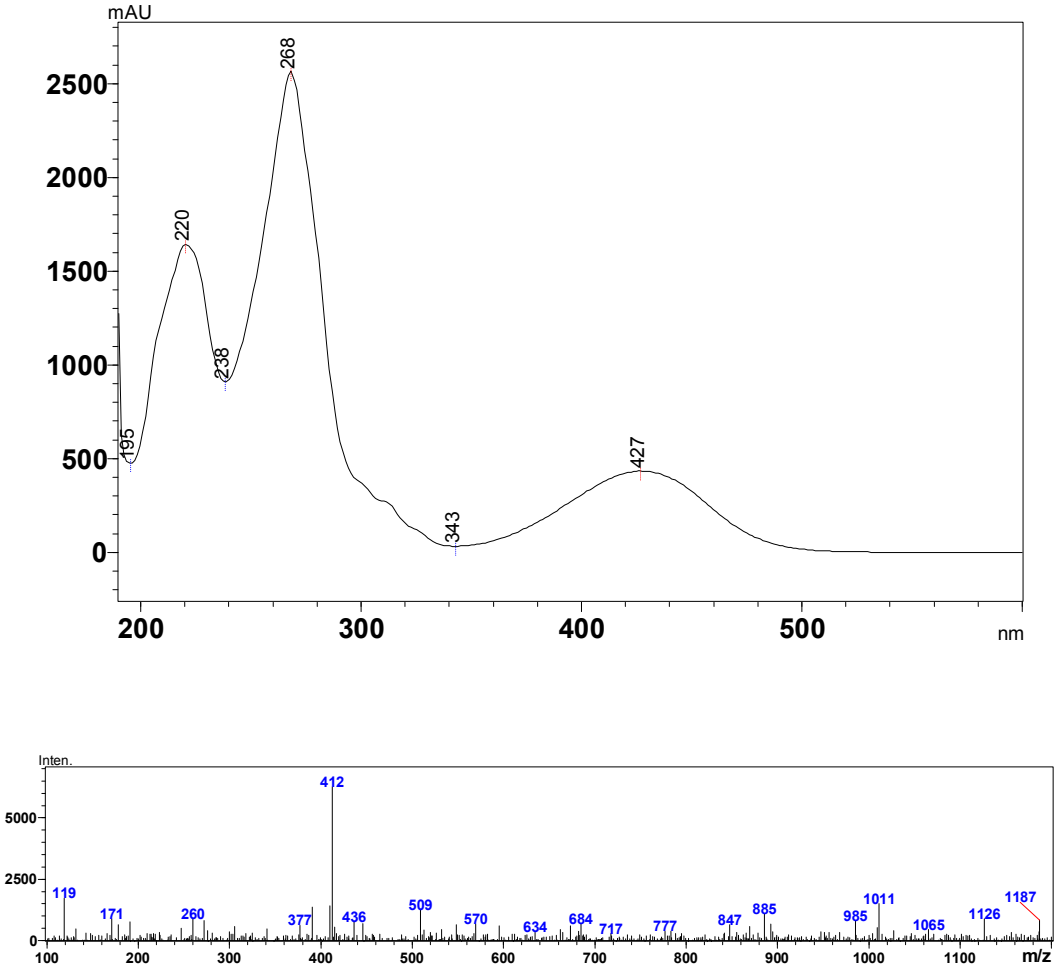


Figure S7.1. UV spectrum and MS Measured during LC-MS for **30**.

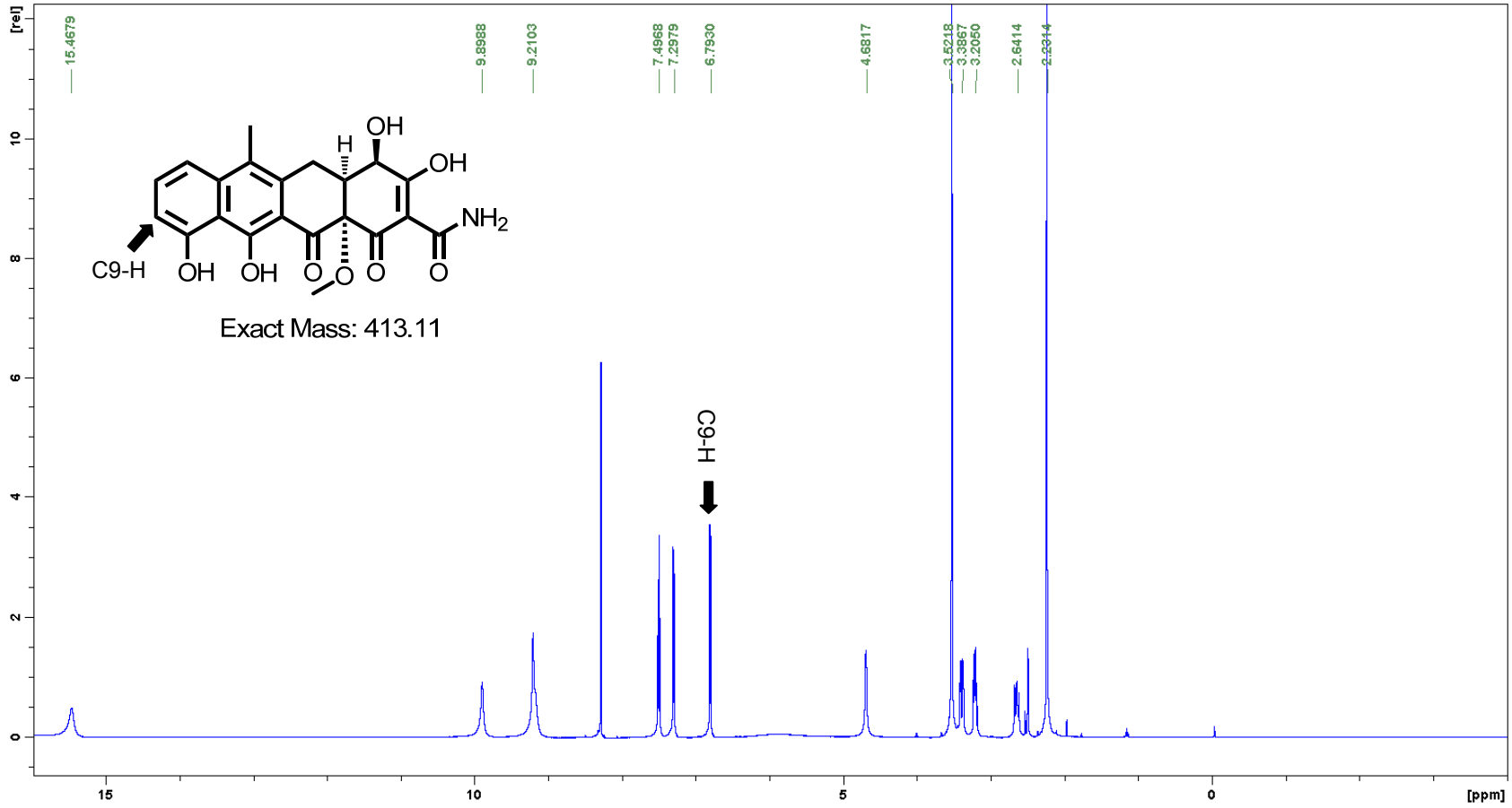


Figure S7.2. ^{1}H NMR spectrum of **30**. Measured in DMSO-d_6 at 500MHz

C9-H signal is clearly observed (black arrow), while whole set of D-olivose signals are missing.

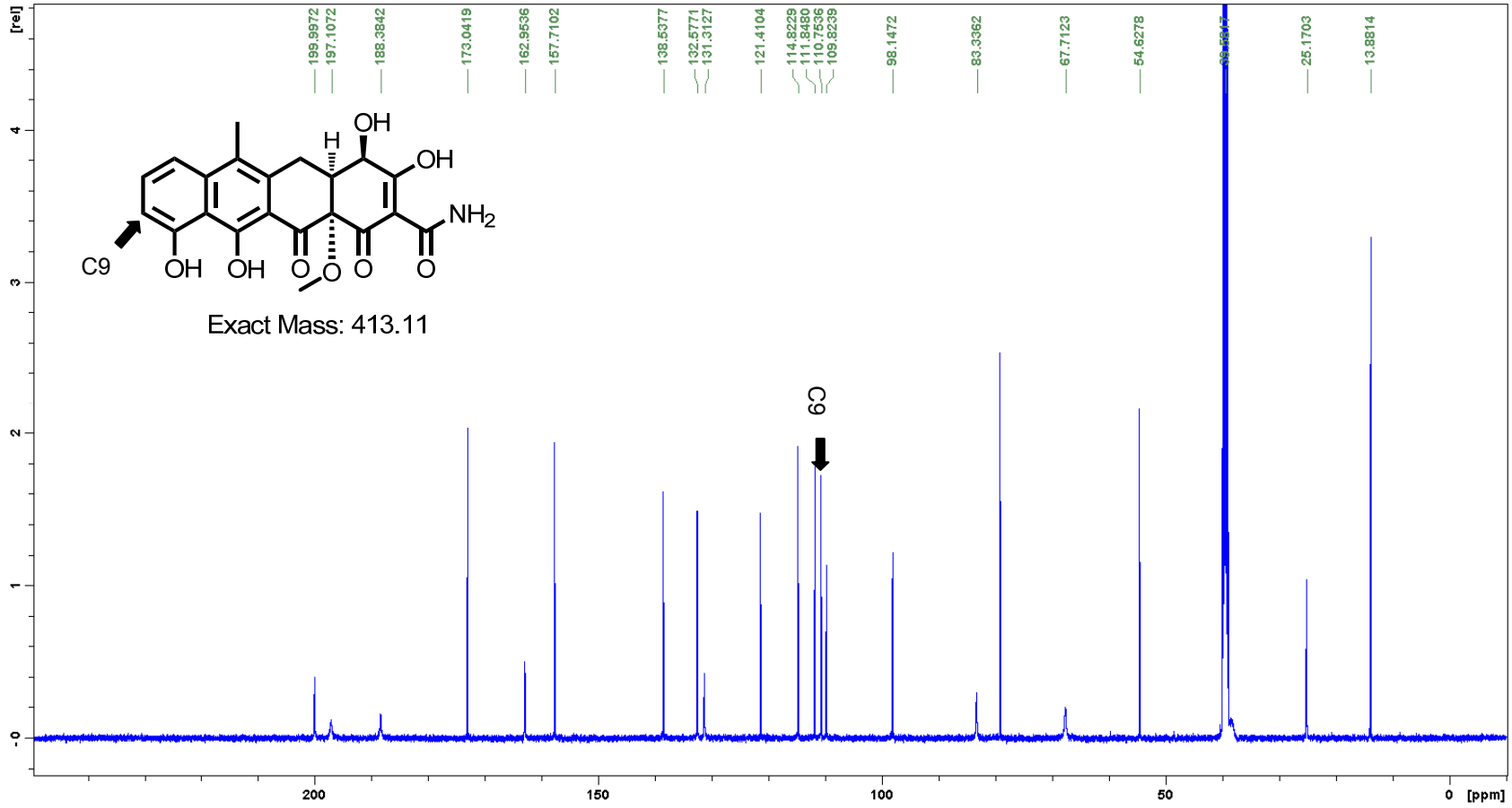


Figure S7.3. ^{13}C NMR spectrum of **30**. Measured in DMSO-d_6 at 125MHz

C9 carbon signal shifts upfield about 13 ppm compared to compound **33** (see below), when D-olivose is removed by SsfS6 inactivation.

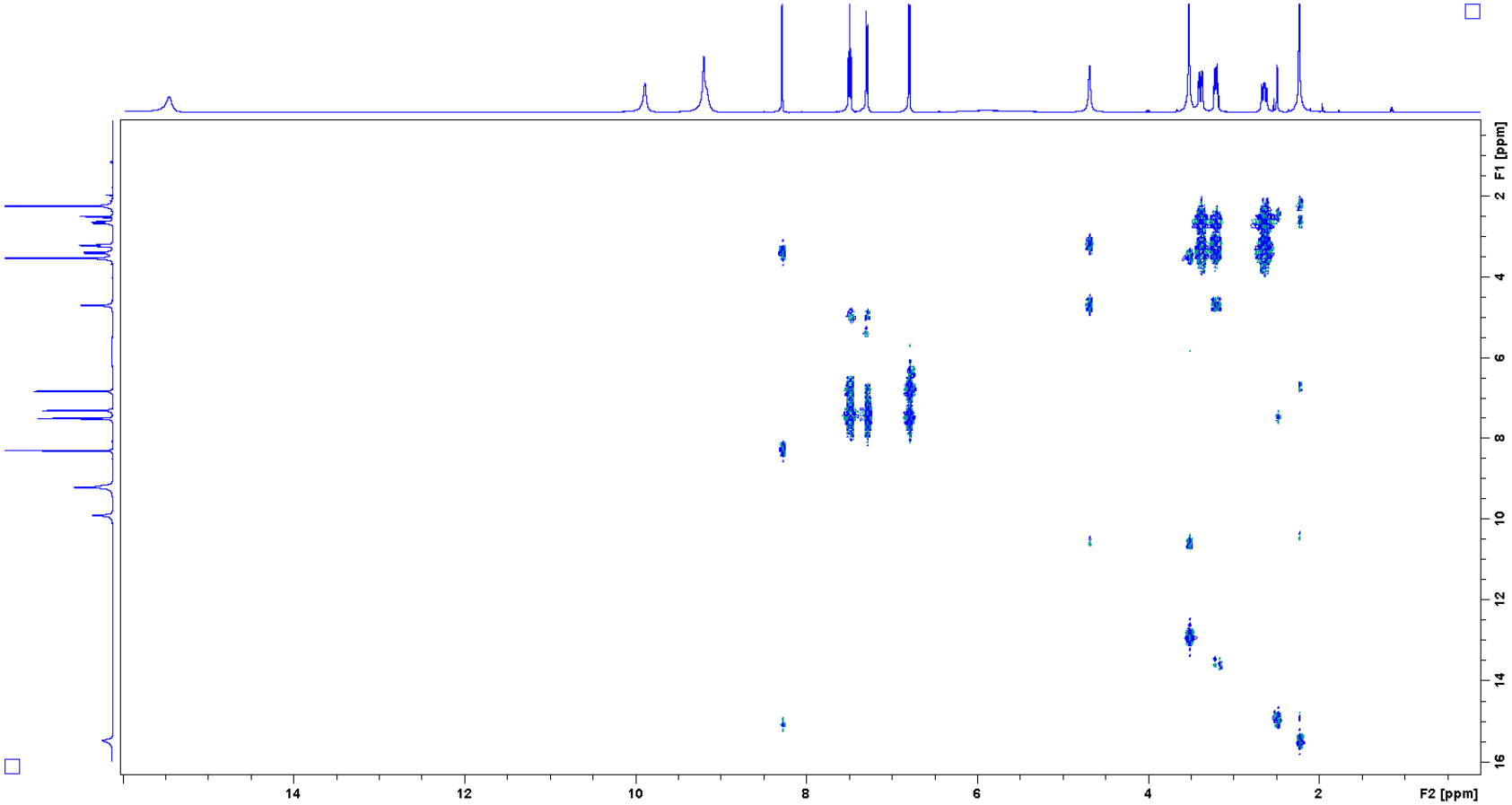


Figure S7.4. ^{1}H-^{1}H COSY NMR spectrum of **30**.

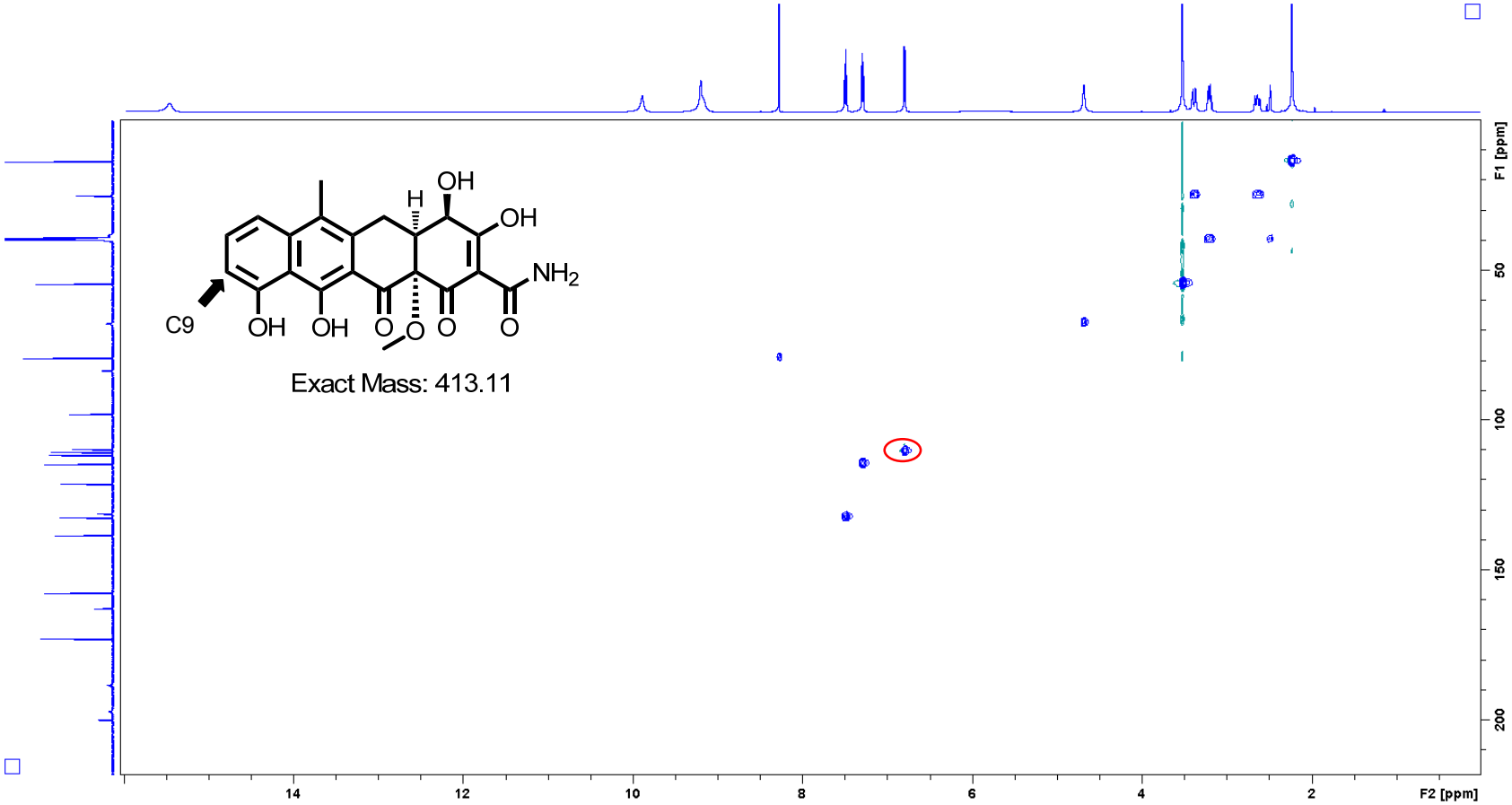


Figure S7.5. ^{1}H-^{13}C HMQC NMR spectrum of **30**.

Correlation between C9 and its proton is showed in red circle.

.

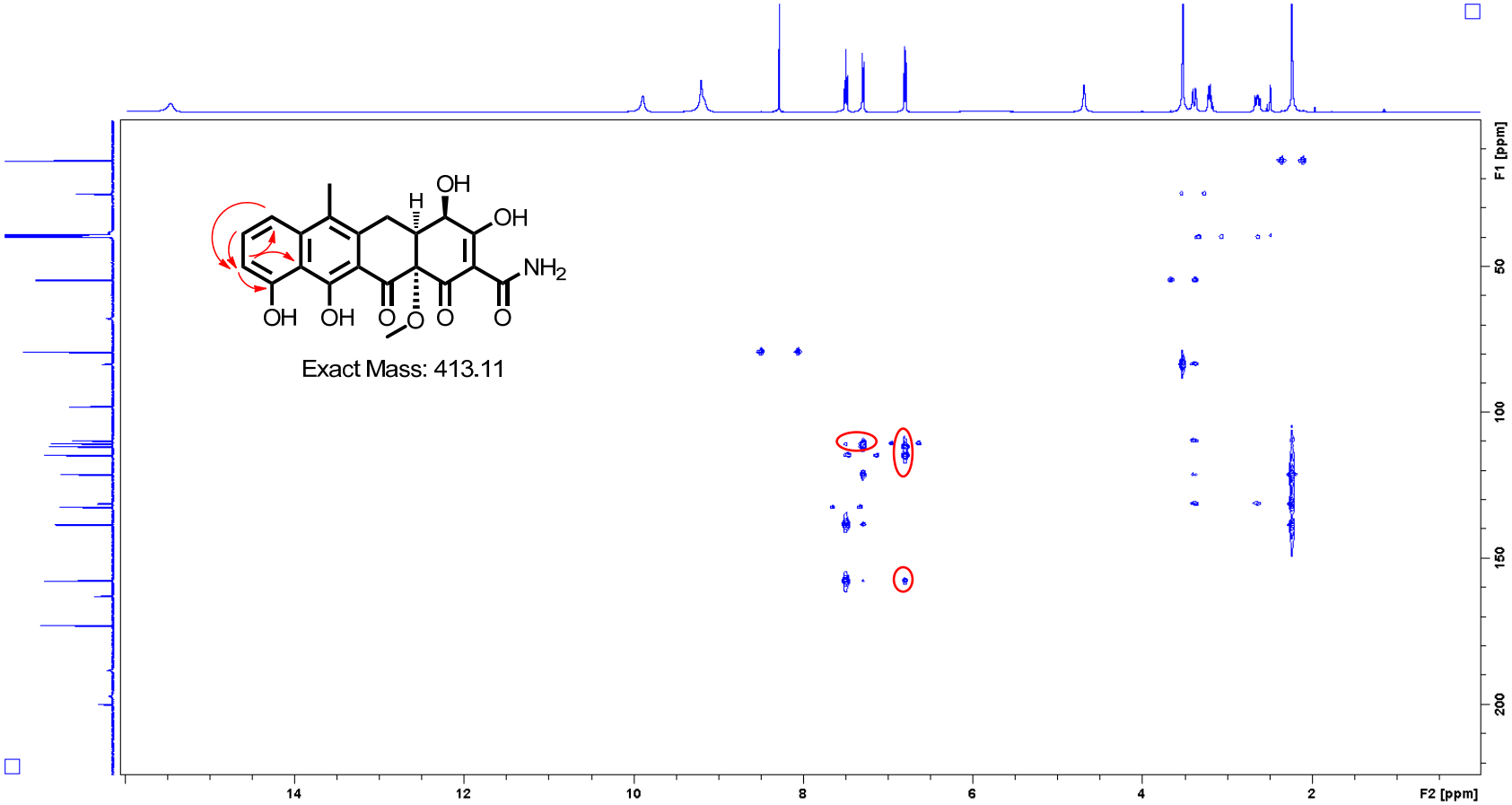


Figure S7.6. ^{1}H-^{13}C HMBC NMR spectrum of **30**.

Correlations between C9-H and C7, C8, C10, C10a establish the position of C9-H, confirming the loss of D-olivose at C9.

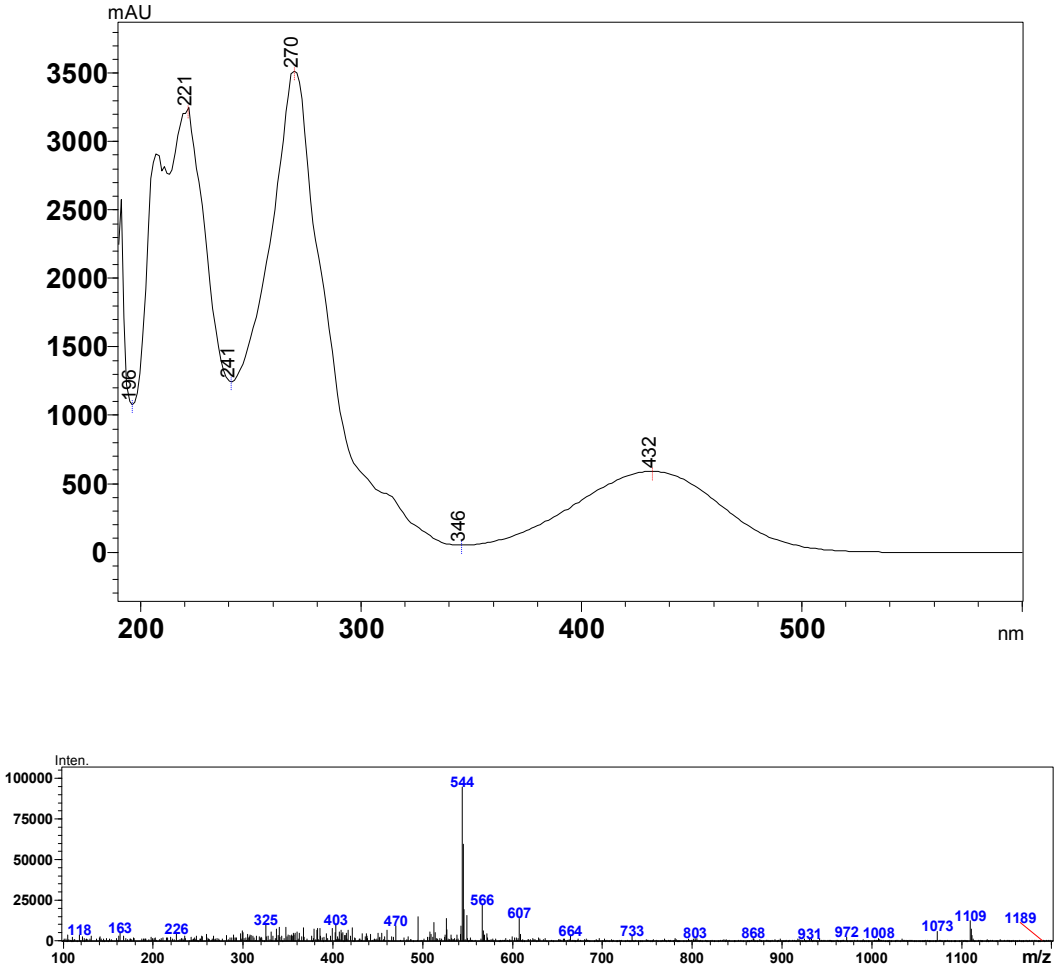


Figure S8.1. UV spectrum and MS Measured during LC-MS for **33**.

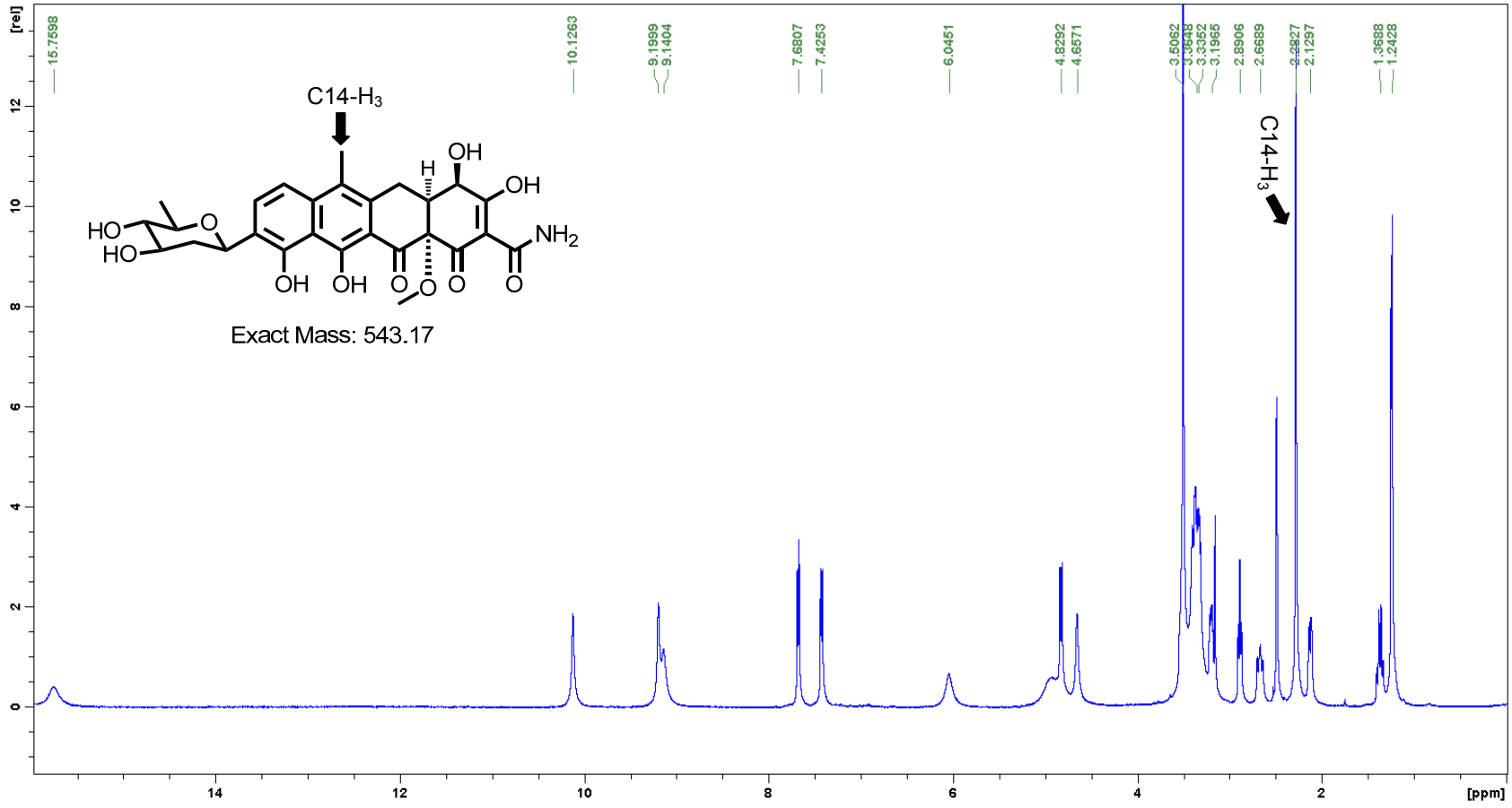


Figure S8.2. ^{1}H NMR spectrum of **33**. Measured in DMSO-d_6 at 500MHz

C14-H_3 carbon signal shifts downfield about 1.3 ppm and C5a-H is missing compared to compound **25** confirming the loss of a hydroxyl group at C6 by SsfO1 inactivation.

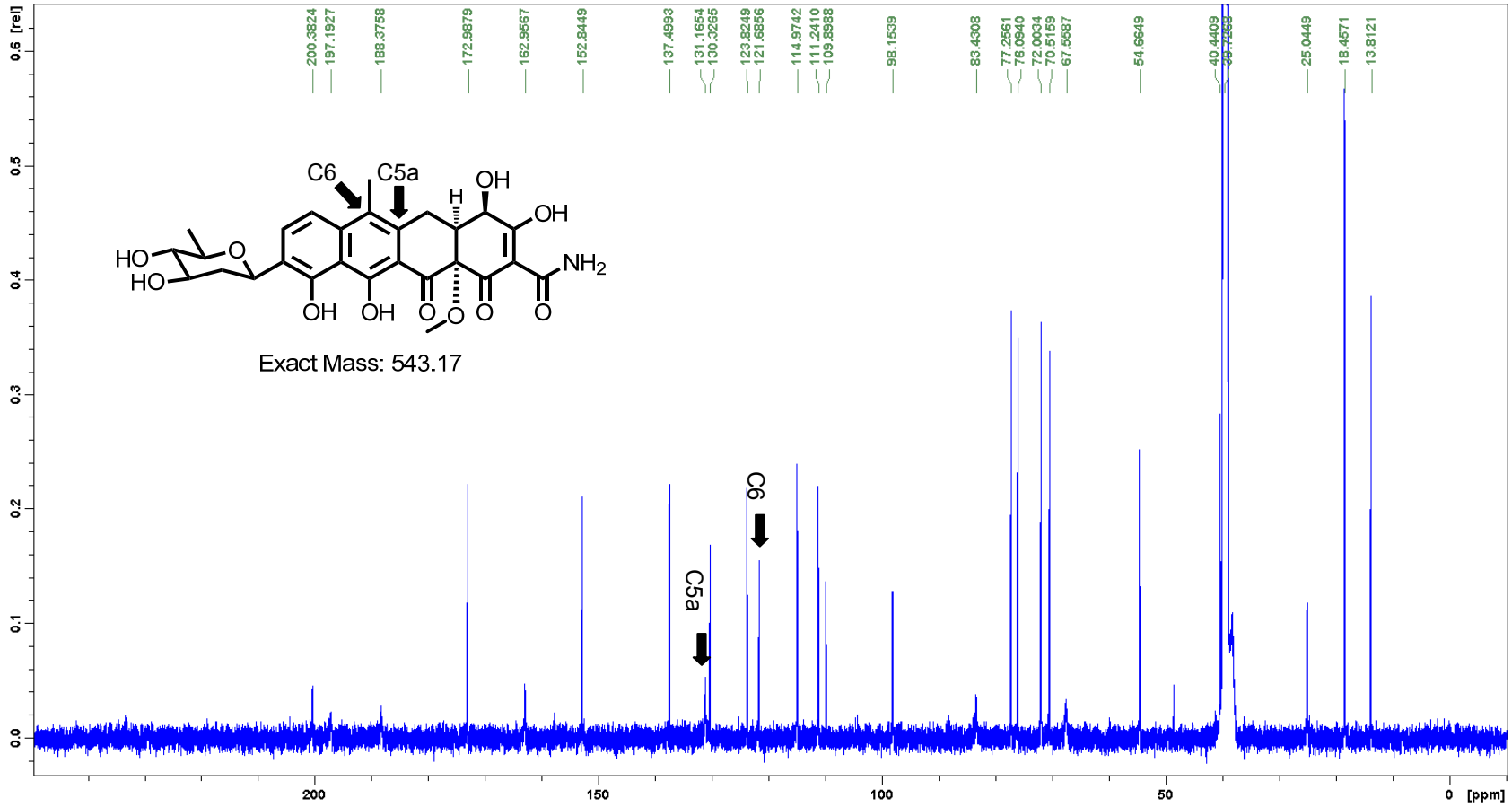


Figure S8.3. ^{13}C NMR spectrum of **33**. Measured in DMSO-d_6 at 125MHz

C6 and C5a shifts downfield about 50 ppm and 95 ppm compared to compound **25** confirming the loss of a hydroxyl group at C6 by SsfO1 inactivation.

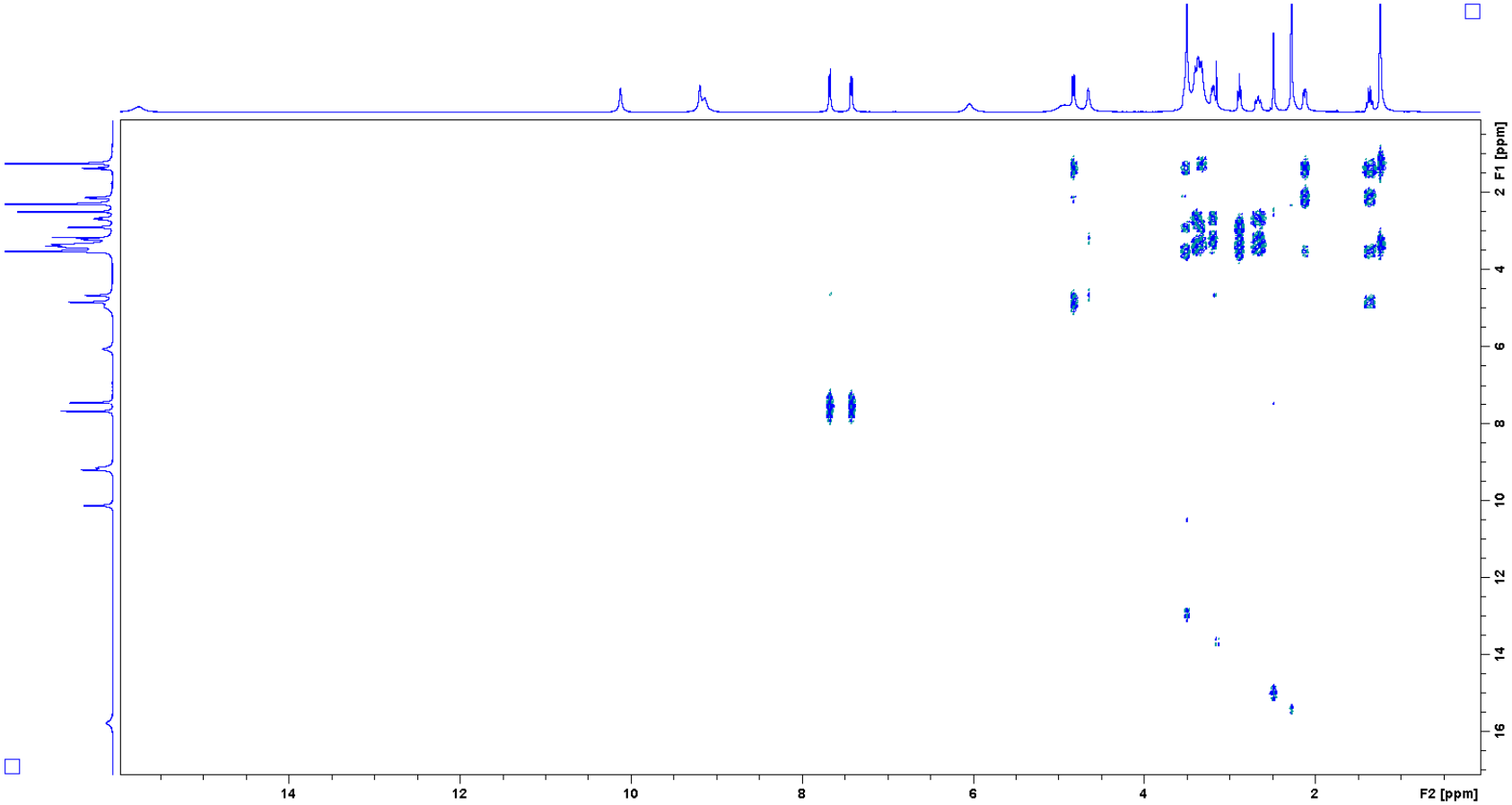


Figure S8.4. ^{1}H-^{1}H COSY NMR spectrum of **33**.

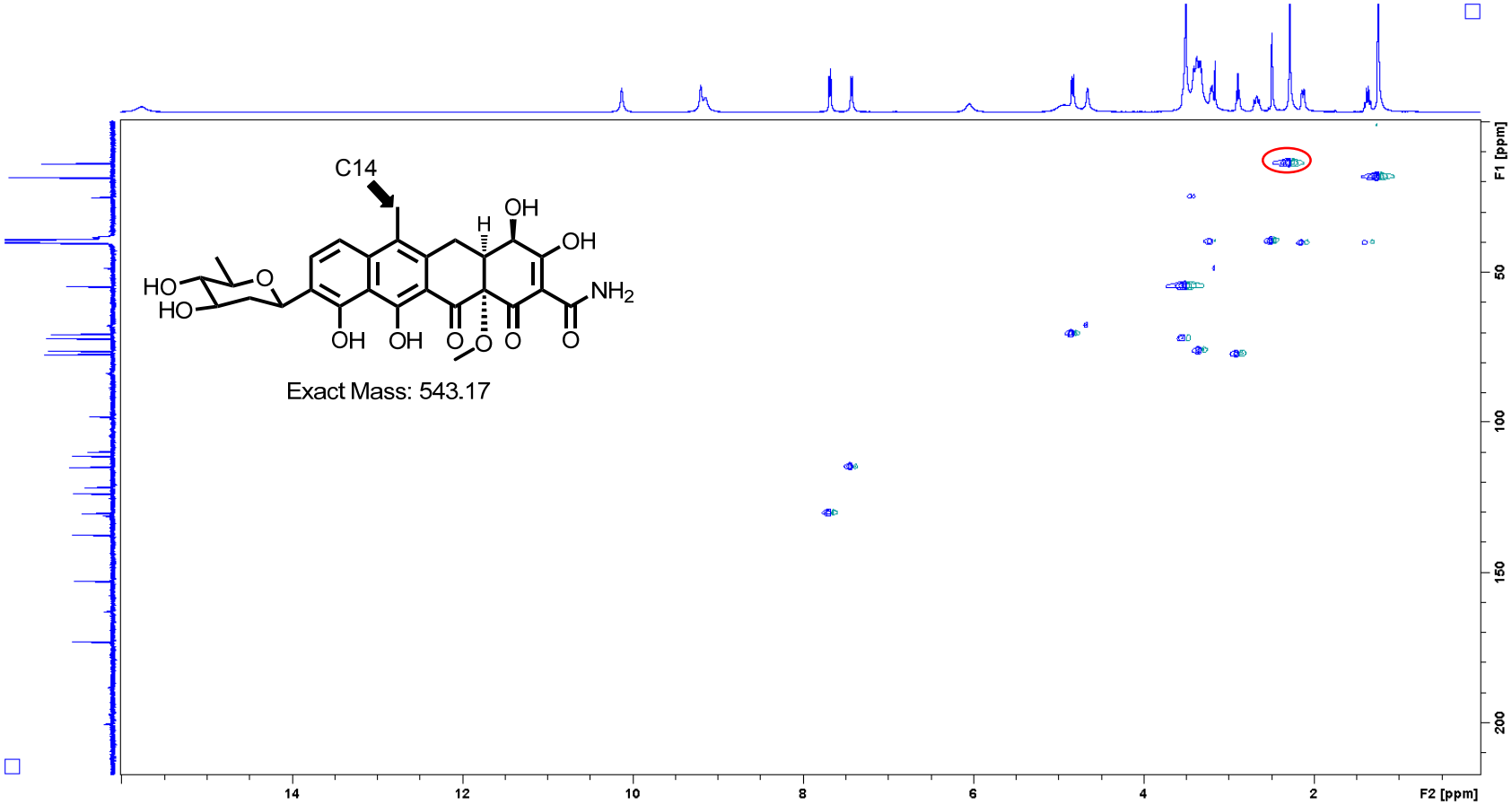


Figure S8.5. ^{1}H-^{13}C HMQC NMR spectrum of **10**.

Correlation between C14 and its proton is showed in red circle.

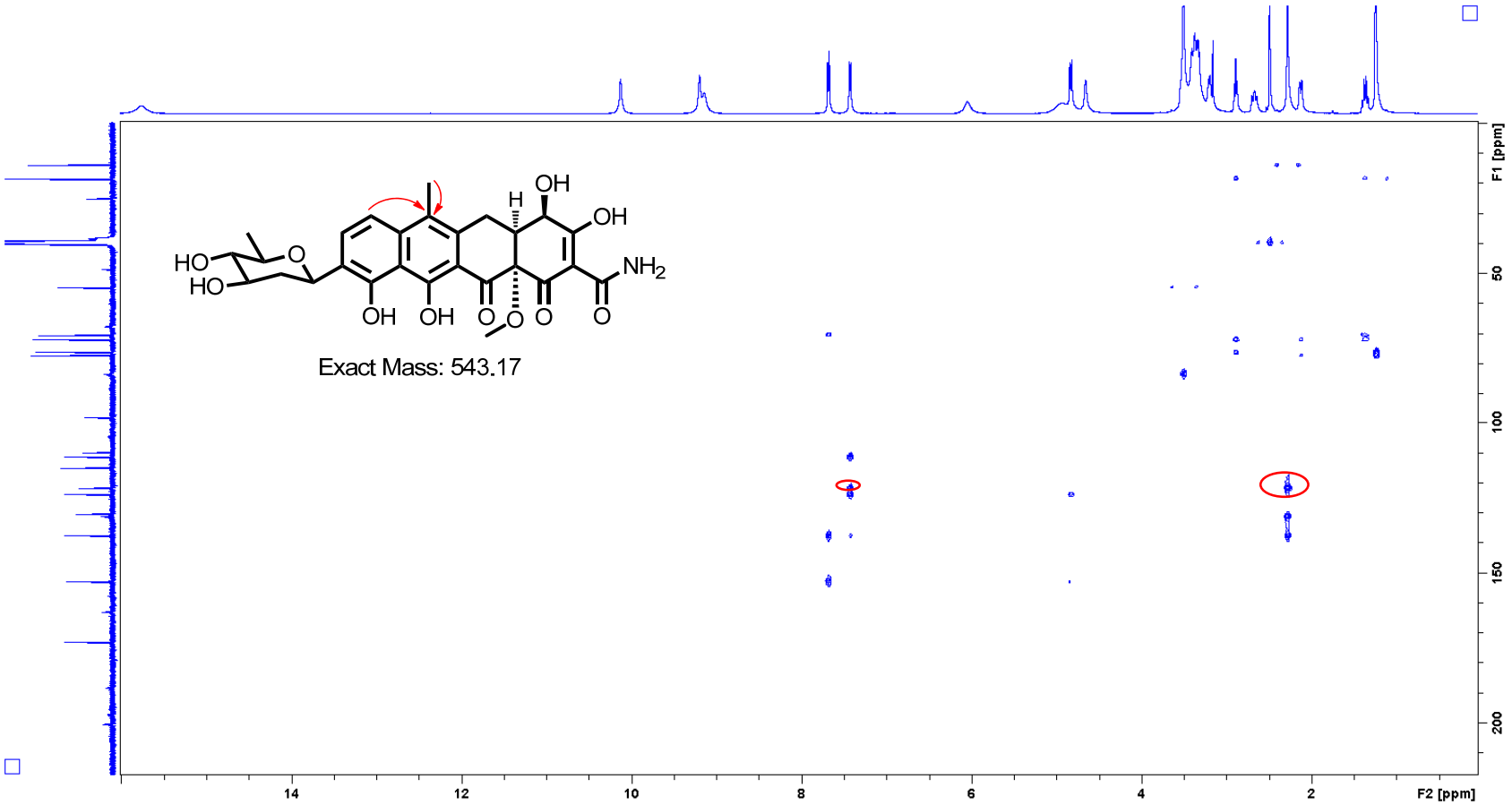


Figure S8.6. ^{1}H-^{13}C HMBC NMR spectrum of **33**.

Correlations between C9-H, C14-H_3 and C6 establish the position of C6, confirming the loss of a hydroxyl group at C6.

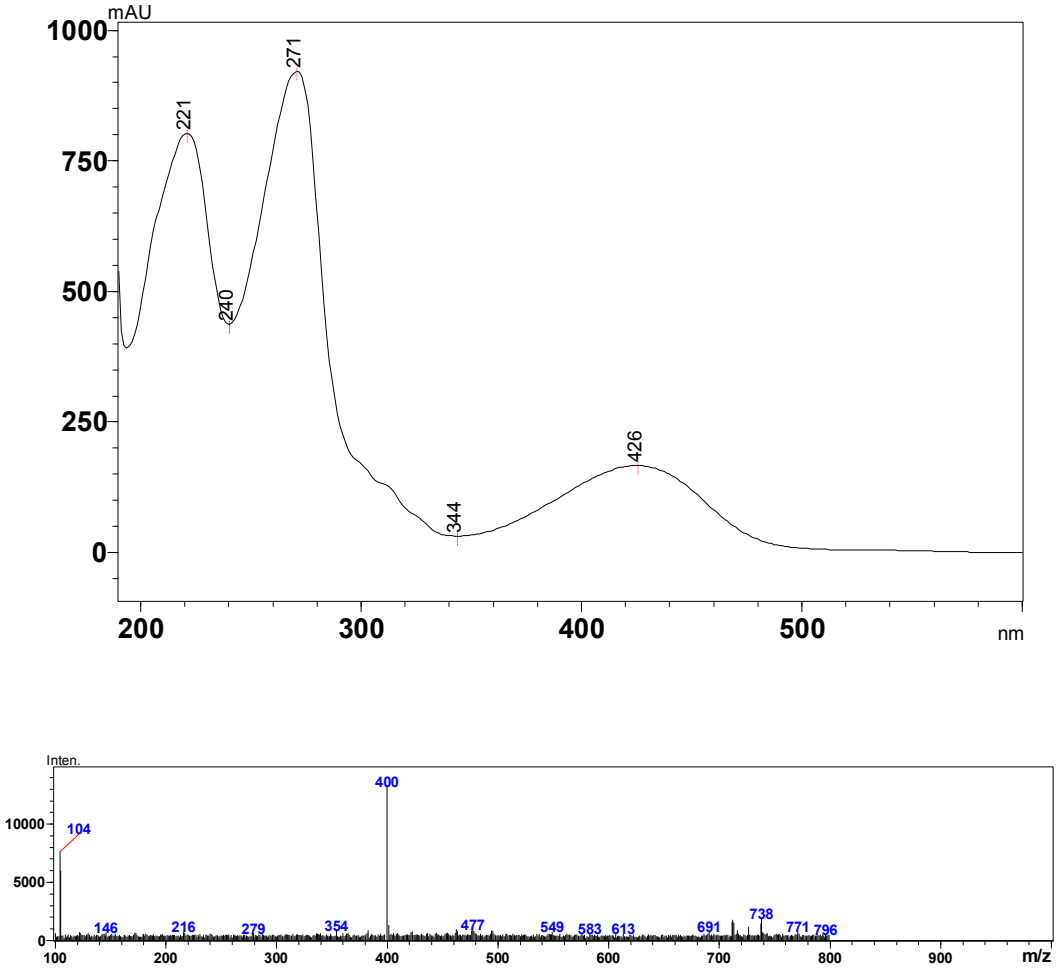


Figure S9.1. UV spectrum and MS Measured during LC-MS for **31**.

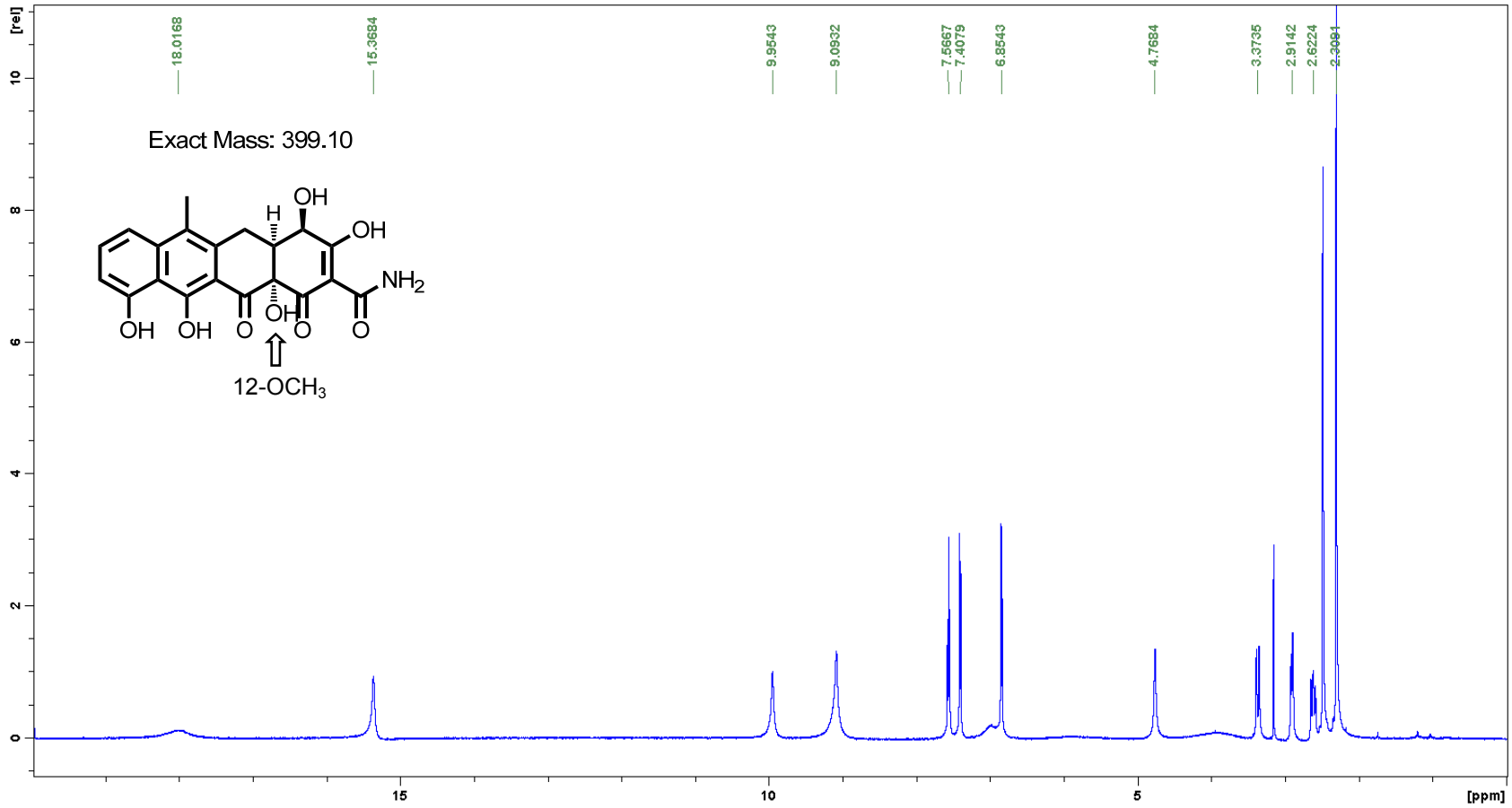


Figure S9.2. ^{1}H NMR spectrum of **31**. Measured in DMSO-d_6 at 500MHz

The proton signal of 12-OCH_3 (empty arrow) is missing, confirming the loss of the methyl group.

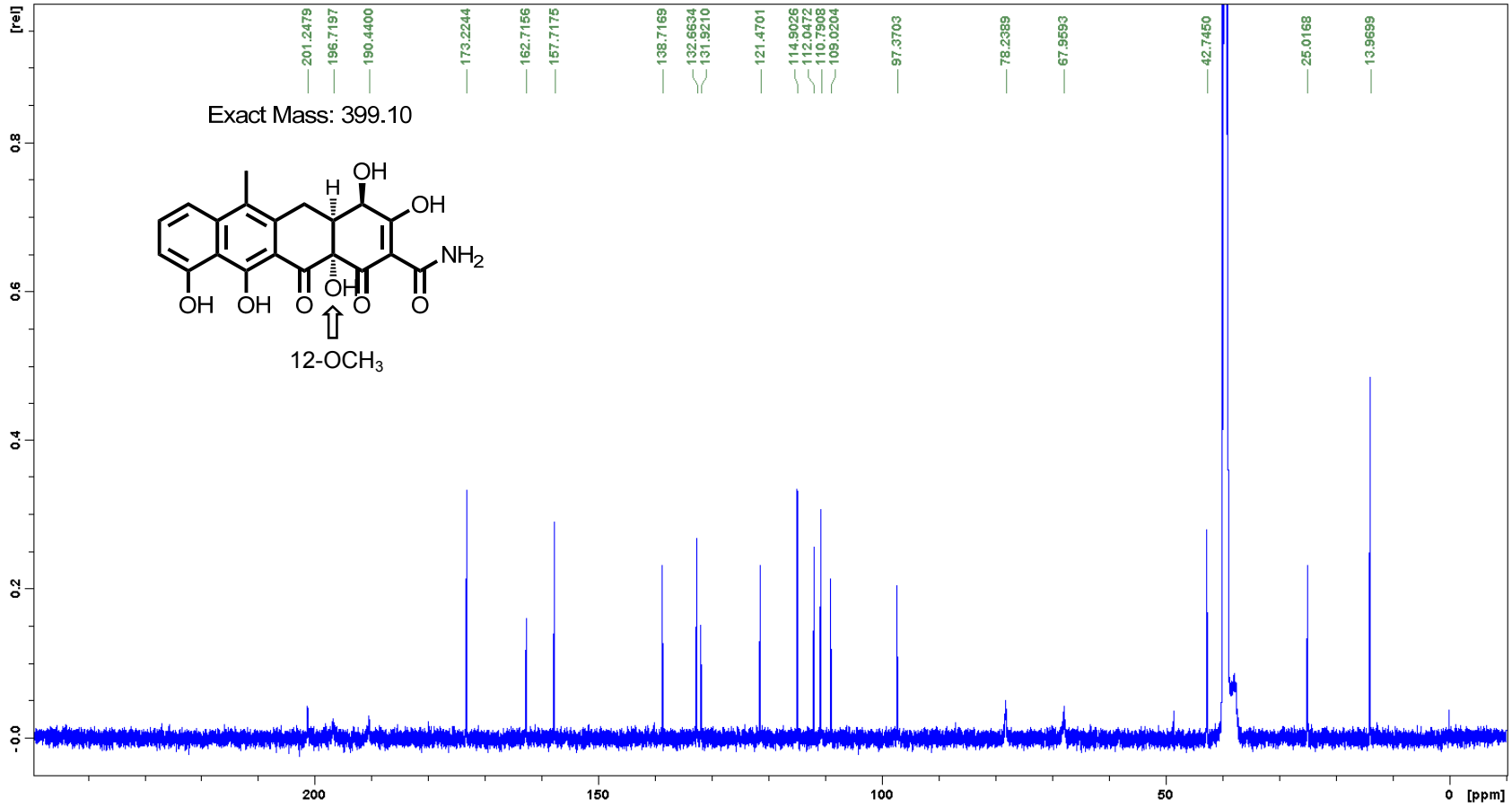


Figure 9.3. ^{13}C NMR spectrum of **31**. Measured in DMSO-d_6 at 125MHz

The carbon signal of 12-OCH_3 (empty arrow) is missing, confirming the loss of the methyl group.

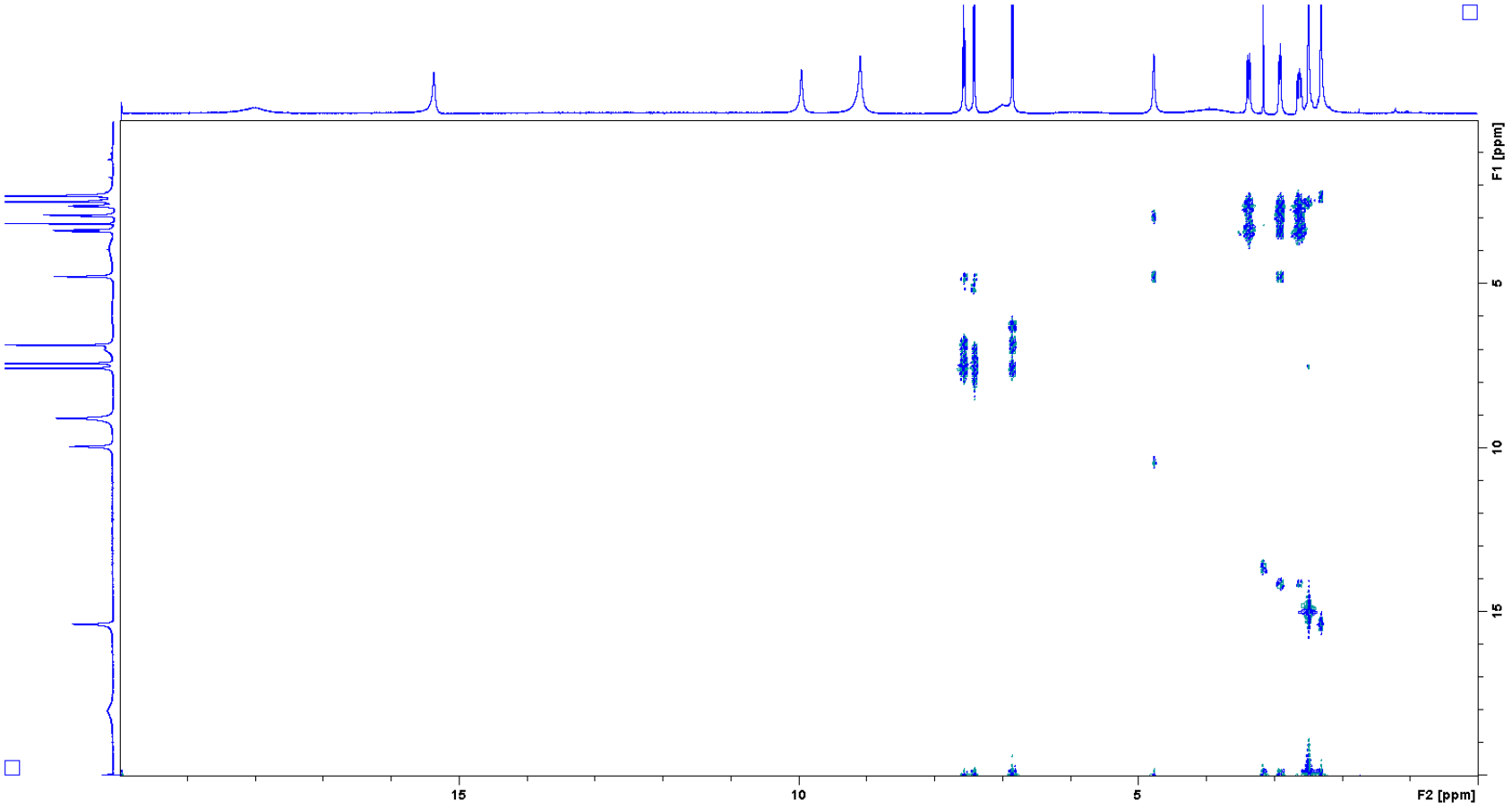


Figure S9.4. ^{1}H-^{1}H COSY spectrum of **31**.

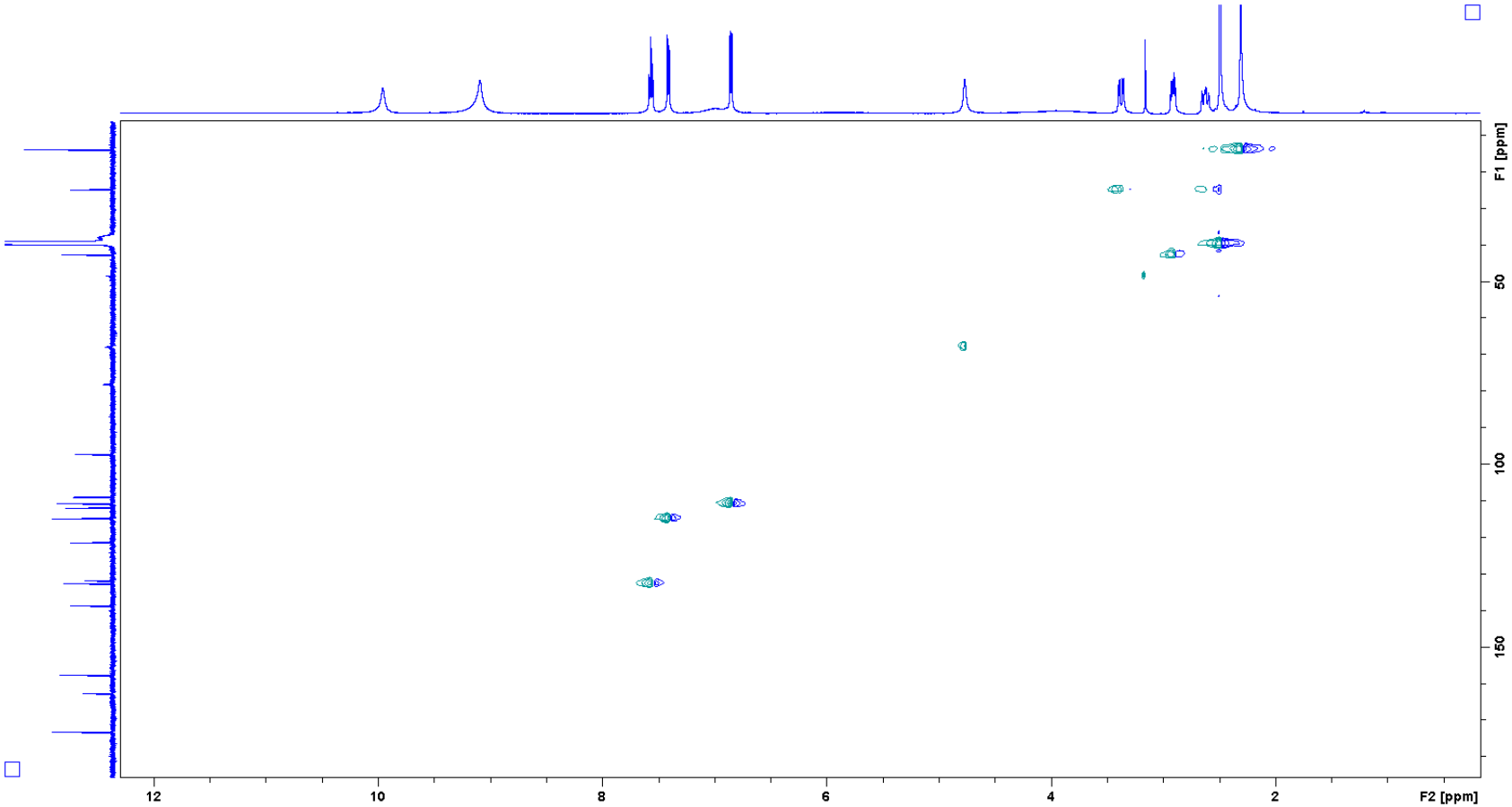


Figure S9.5. HMQC spectrum of **31**.

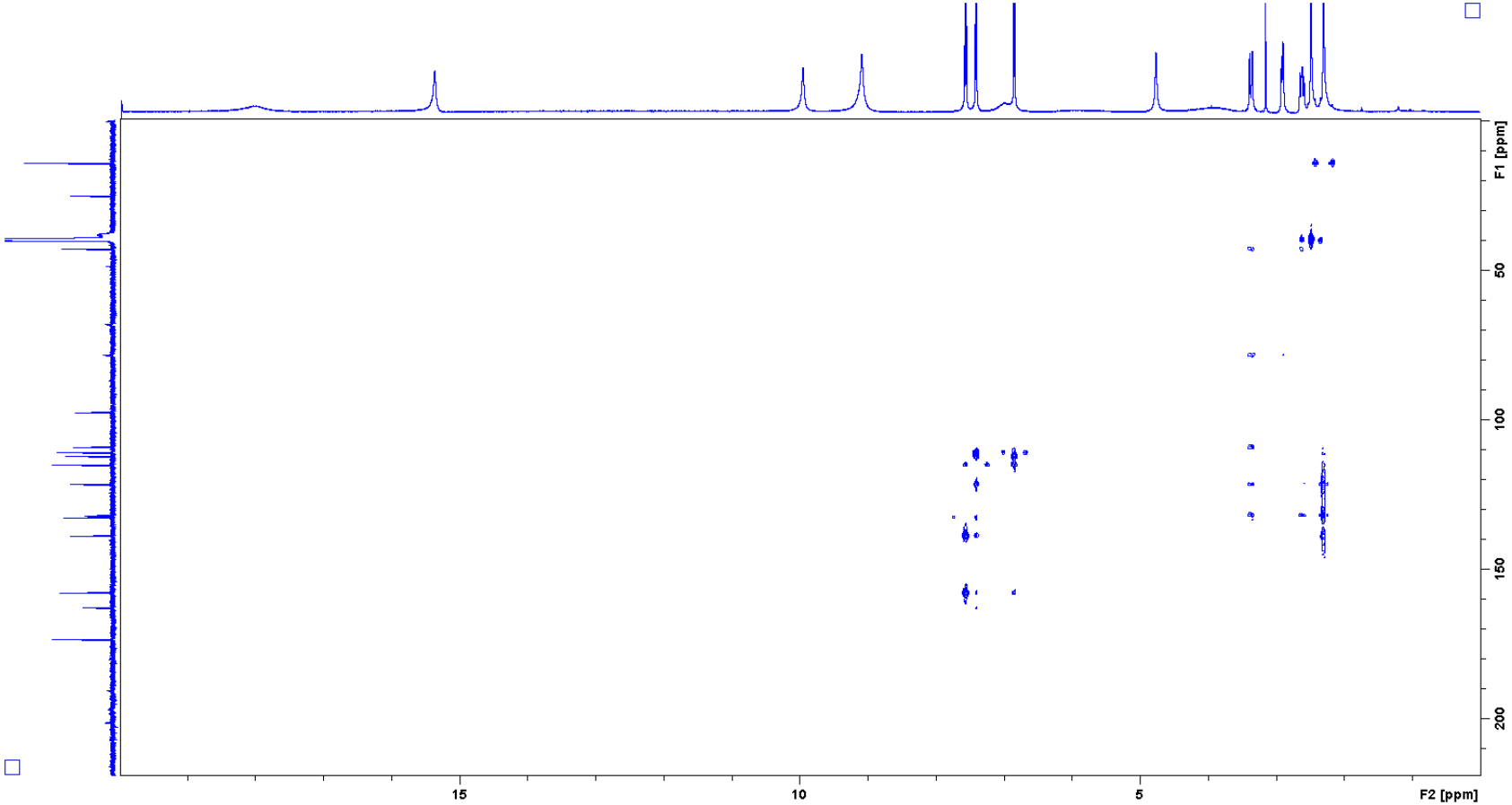


Figure S9.6. HMBC spectrum of **31**.

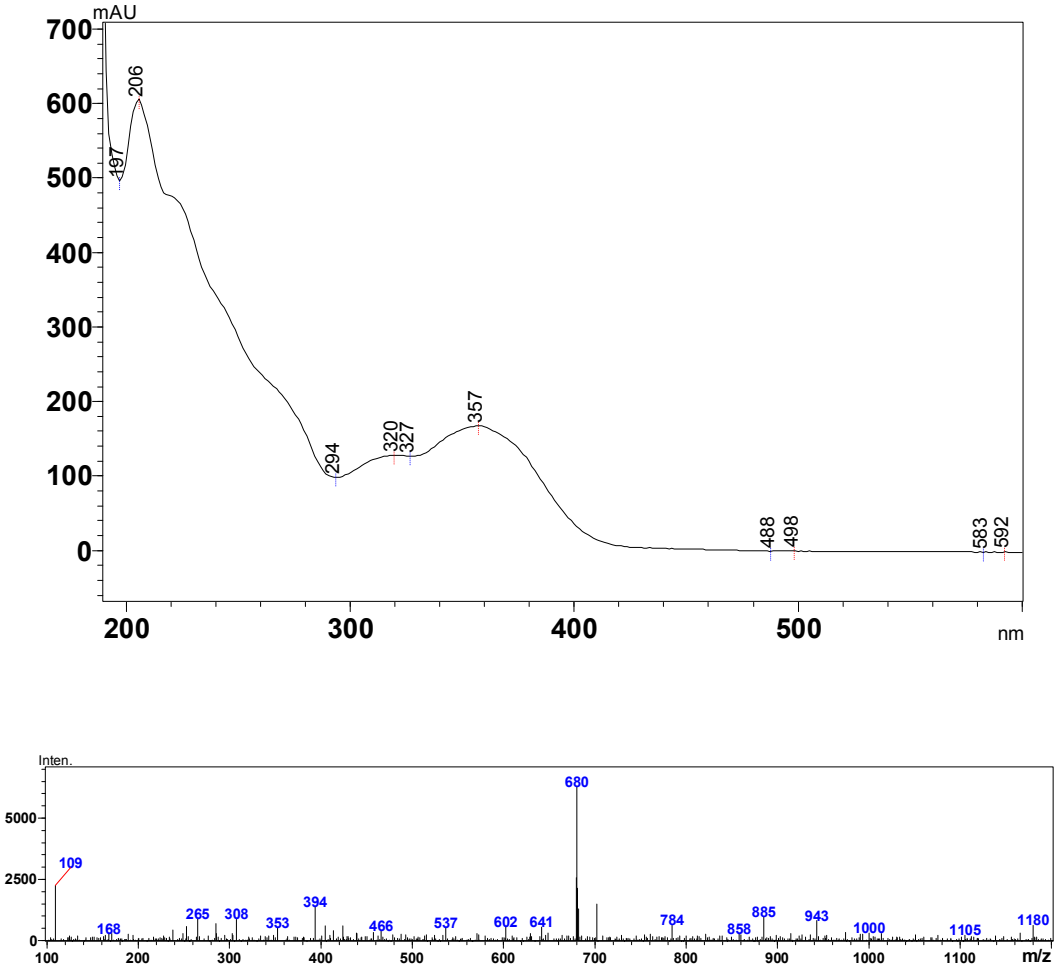


Figure S10. UV spectrum and MS Measured during LC-MS for **23**.

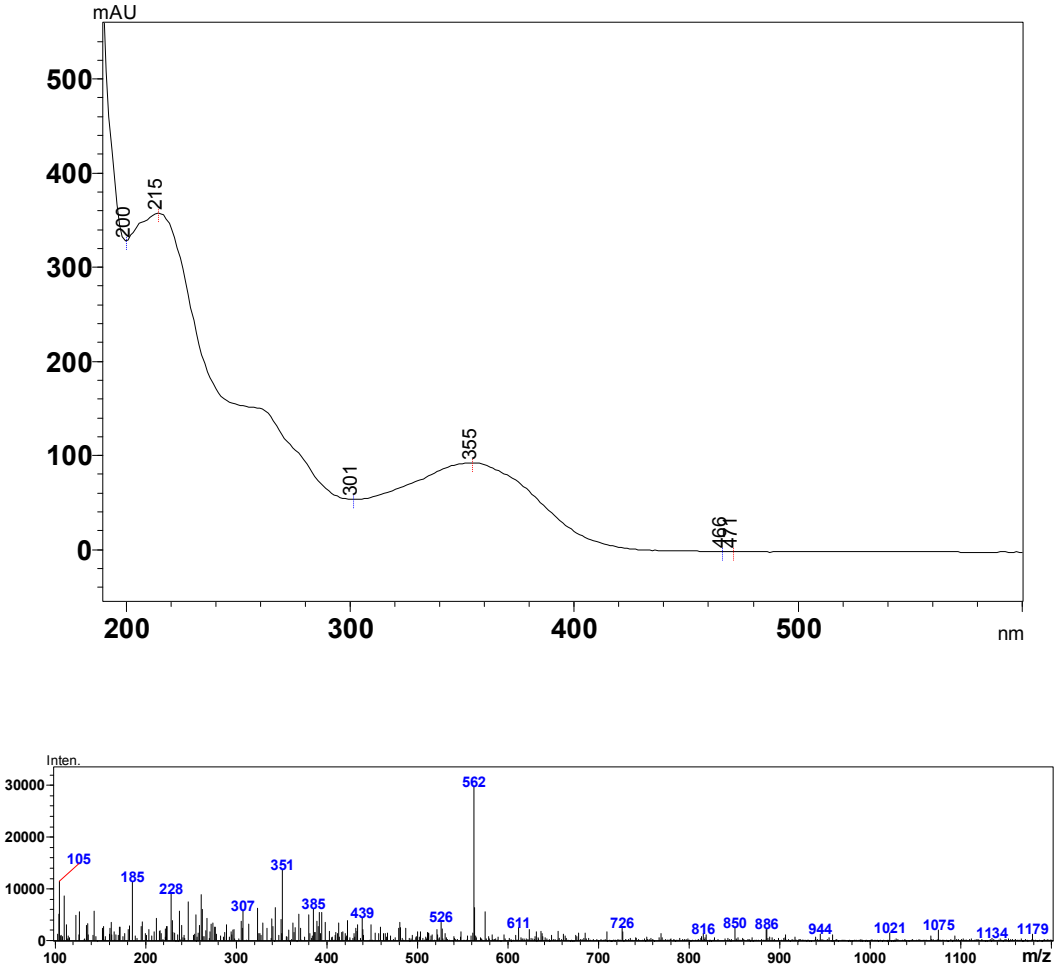


Figure S11. UV spectrum and MS Measured during LC-MS for **24**.

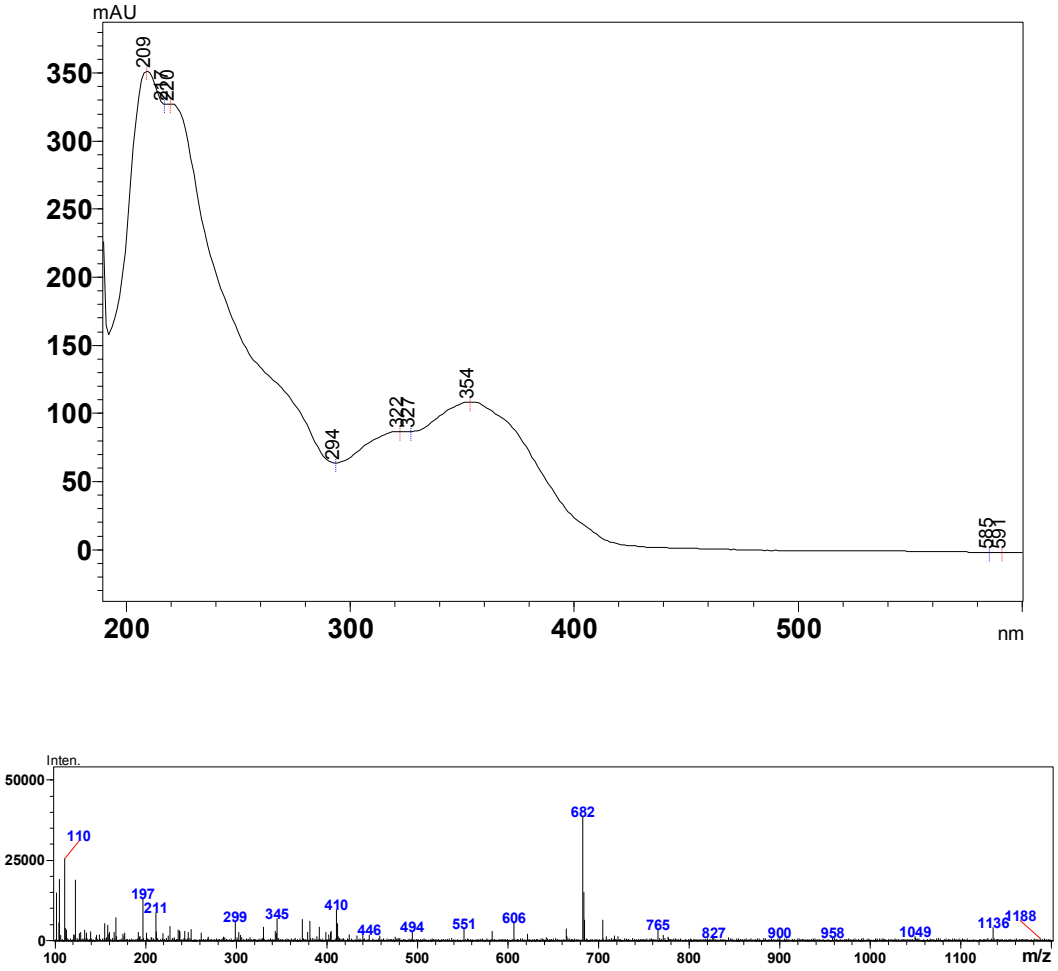


Figure S12. UV spectrum and MS Measured during LC-MS for **26**.

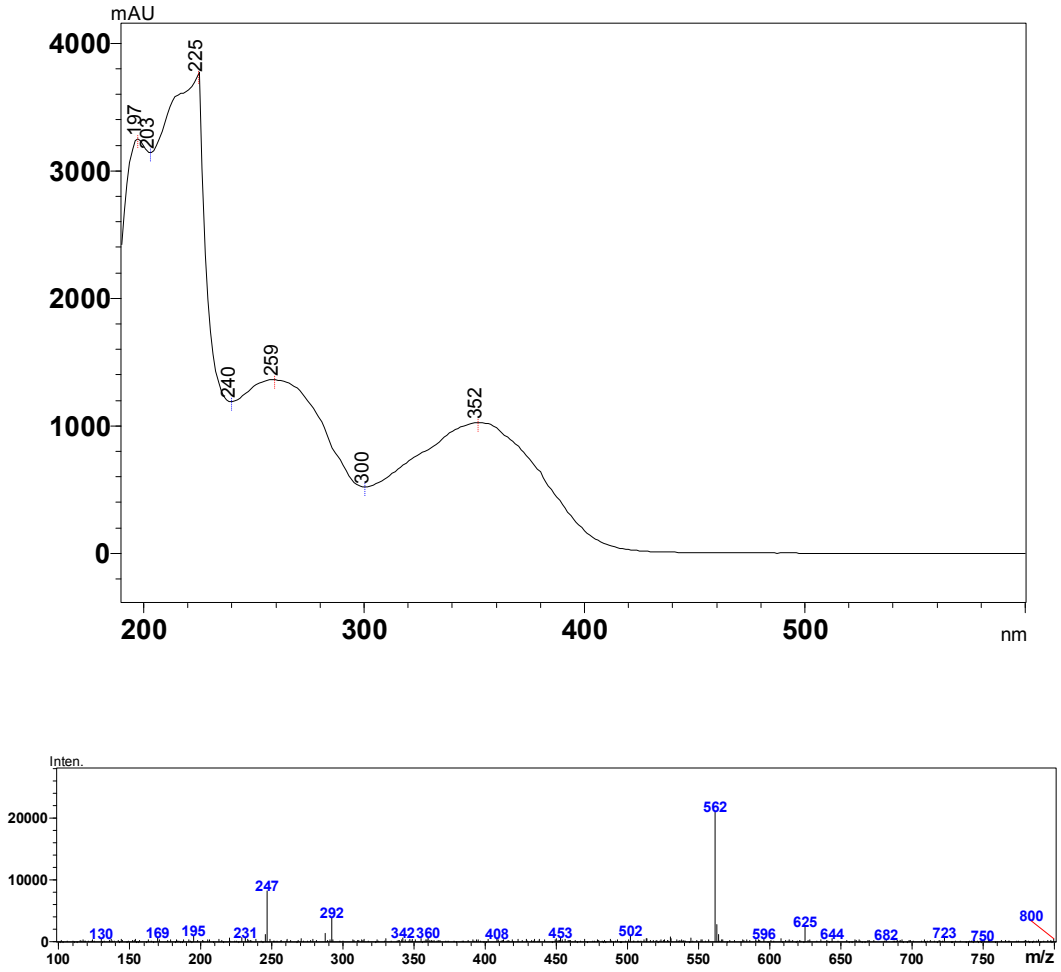


Figure S13. UV spectrum and MS Measured during LC-MS for **27**.

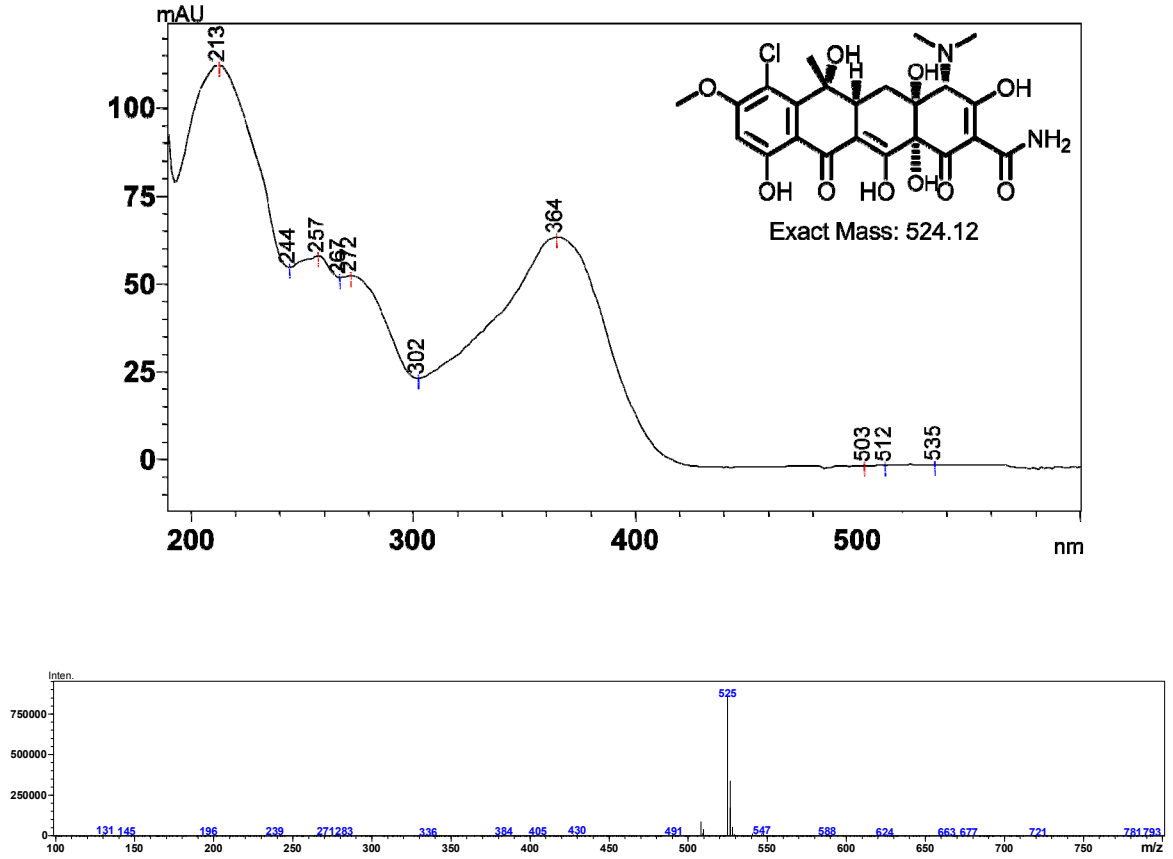


Figure S14. UV spectrum and MS Measured during LC-MS for **14**

References

1. Baltz, R.H., *Molecular engineering approaches to peptide, polyketide and other antibiotics.* Nat Biotechnol, 2006. **24**(12): p. 1533-1540.

2. Weissman, K.J. and P.F. Leadlay, *Combinatorial biosynthesis of reduced polyketides.* Nat Rev Microbiol, 2005. **3**(12): p. 925-936.

3. Olano, C., C. Mendez, and J.A. Salas, *Antitumor compounds from actinomycetes: from gene clusters to new derivatives by combinatorial biosynthesis.* Nat Prod Rep, 2009. **26**(5): p. 628-660.

4. Staunton, J. and K.J. Weissman, *Polyketide biosynthesis: a millennium review.* Nat Prod Rep, 2001. **18**(4): p. 380-416.

5. Donadio, S. and L. Katz, *Organization of the enzymatic domains in the multifunctional polyketide synthase involved in erythromycin formation in Saccharopolyspora erythraea.* Gene, 1992. **111**(1): p. 51-60.

6. Cane, D.E., C.T. Walsh, and C. Khosla, *Harnessing the biosynthetic code: combinations, permutations, and mutations.* Science, 1998. **282**(5386): p. 63-8.

7. Kennedy, J., et al., *Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis.* Science, 1999. **284**(5418): p. 1368-72.

8. Hertweck, C., et al., *Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork.* Nat Prod Rep, 2007. **24**(1): p. 162-190.

9. Mcdaniel, R., et al., *Engineered Biosynthesis of Novel Polyketides.* Science, 1993. **262**(5139): p. 1546-1550.

10. Hutchinson, C.R., *Biosynthetic studies of daunorubicin and tetracenomycin C.* Chem Rev, 1997. **97**(7): p. 2525-2535.

11. Henkel, T., et al., *Landomycins, new angucycline antibiotics from Streptomyces sp. I. Structural studies on landomycins A-D.* J Antibiot (Tokyo), 1990. **43**(5): p. 492-503.

12. Austin, M.B. and A.J.P. Noel, *The chalcone synthase superfamily of type III polyketide synthases.* Nat Prod Rep, 2003. **20**(1): p. 79-110.

13. Olano, C., C. Méndez, and J.A. Salas, *Post-PKS tailoring steps in natural product-producing actinomycetes from the perspective of combinatorial biosynthesis.* Nat. Prod. Rep., 2010. **27**: p. 571-616.

14. Khosla, C., et al., *Structure and mechanism of the 6-deoxyerythronolide B synthase.* Annu Rev Biochem, 2007. **76**: p. 195-221.

15. Cortes, J., et al., *An unusually large multifunctional polypeptide in the erythromycin producing polyketide synthase of Saccharopolyspora erythraea.* Nature, 1990. **348**(6297): p. 176-178.

16. Donadio, S., et al., *Modular organization of genes required for complex polyketide biosynthesis.* Science, 1991. **252**(5006): p. 675-679.

17. Cane, D.E., *Programming of erythromycin biosynthesis by a modular polyketide synthase.* J Biol Chem, 2010.

18. Marsden, A.F.A., et al., *Engineering broader specificity into an antibiotic-producing polyketide synthase.* Science, 1998. **279**(5348): p. 199-202.

19. Kao, C.M., L. Katz, and C. Khosla, *Engineered biosynthesis of a complete macrolactone in a heterologous host.* Science, 1994. **265**(5171): p. 509-12.

20. Xue, Q., et al., *A multiplasmid approach to preparing large libraries of polyketides.* Proc Natl Acad Sci U S A, 1999. **96**(21): p. 11740-5.

21. Kharel, M.K., et al., *Angucyclines: Biosynthesis, mode-of-action, new natural products, and synthesis.* Nat Prod Rep, 2012. **29**(2): p. 264-325.

22. Weber, S., et al., *Investigations of the Biosynthesis and Structural Revision of Landomycin A.* J. Org. Chem., 1994. **59**(15): p. 4211-4214.

23. Westrich, L., et al., *Cloning and characterization of a gene cluster from Streptomyces cyanogenus S136 probably involved in landomycin biosynthesis.* FEMS Microbiol Lett, 1999. **170**(2): p. 381-7.

24. Weisshaar, B. and G.I. Jenkins, *Phenylpropanoid biosynthesis and its regulation.* Curr Opin Plant Biol, 1998. **1**(3): p. 251-257.

25. Austin, M.B., et al., *An aldol switch discovered in stilbene synthases mediates cyclization specificity of type III polyketide synthases.* Chem Biol, 2004. **11**(9): p. 1179-1194.

26. Abe, I. and H. Morita, *Structure and function of the chalcone synthase superfamily of plant type III polyketide synthases.* Nat Prod Rep, 2010. **27**(6): p. 809-838.

27. Hwang, E.I., et al., *Production of plant-specific flavanones by Escherichia coli containing an artificial gene cluster.* Appl Environ Microbiol, 2003. **69**(5): p. 2699-2706.

28. Miyahisa, I., et al., *Efficient production of (2S)-flavanones by Escherichia coli containing an artificial biosynthetic gene cluster.* Appl Environ Microbiol, 2005. **68**(4): p. 498-504.

29. Katsuyama, Y., et al., *Synthesis of unnatural flavonoids and stilbenes by exploiting the plant biosynthetic pathway in Escherichia coli.* Chem Biol, 2007. **14**(6): p. 613-621.

30. Katsuyama, Y., et al., *Production of curcuminoids by Escherichia coli carrying an artificial biosynthesis pathway.* Microbiol-Sgm, 2008. **154**: p. 2620-2628.

31. Pfeifer, B.A. and C. Khosla, *Biosynthesis of polyketides in heterologous hosts.* Microbiol Mol Biol Rev, 2001. **65**(1): p. 106-18.

32. Baltz, R.H., *Streptomyces and Saccharopolyspora hosts for heterologous expression of secondary metabolite gene clusters.* J. Ind. Microbiol. Biotechnol., 2010. **37**(8): p. 759-72.

33. Bentley, S.D., et al., *Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2).* Nature, 2002. **417**(6885): p. 141-7.

34. Chouayekh, H. and M.J. Virolle, *The polyphosphate kinase plays a negative role in the control of antibiotic production in Streptomyces lividans.* Mol. Microbiol., 2002. **43**(4): p. 919-930.

35. Ziermann, R. and M.C. Betlach, *Recombinant polyketide synthesis in Streptomyces: engineering of improved host strains.* Biotechniques, 1999. **26**(1): p. 106-110.

36. Elliot, M.A. and N.J. Talbot, *Building filaments in the air: aerial morphogenesis in bacteria and fungi.* Curr Opin Microbiol, 2004. **7**(6): p. 594-601.

37. Chater, K.F. and S. Horinouchi, *Signalling early developmental events in two highly diverged Streptomyces species.* Mol. Microbiol., 2003. **48**(1): p. 9-15.

38. Gramajo, H.C., E. Takano, and M.J. Bibb, *Stationary-phase production of the antibiotic actinorhodin in Streptomyces coelicolor A3(2) is transcriptionally regulated.* Mol. Microbiol., 1993. **7**(6): p. 837-45.

39. Floriano, B. and M. Bibb, *afsR is a pleiotropic but conditionally required regulatory gene for antibiotic production in Streptomyces coelicolor A3(2).* Mol. Microbiol., 1996. **21**(2): p. 385-96.

40. Sheldon, P.J., S.B. Busarow, and C.R. Hutchinson, *Mapping the DNA-binding domain and target sequences of the Streptomyces peucetius daunorubicin biosynthesis regulatory protein, DnrI.* Mol. Microbiol., 2002. **44**(2): p. 449-60.

41. Fischbach, M.A. and C.T. Walsh, *Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms.* Chem Rev, 2006. **106**(8): p. 3468-3496.

42. Hertweck, C., *The biosynthetic logic of polyketide diversity.* Angewandte Chemie. International Edition in English, 2009. **48**(26): p. 4688-716.

43. Wilson, M.C. and B.S. Moore, *Beyond ethylmalonyl-CoA: the functional role of crotonyl-CoA carboxylase/reductase homologs in expanding polyketide diversity.* Nat Prod Rep, 2012. **29**(1): p. 72-86.

44. Erb, T.J., et al., *Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway.* Proc Natl Acad Sci USA, 2007. **104**(25): p. 10631-6.

45. Eustaquio, A.S., et al., *Biosynthesis of the salinosporamide A polyketide synthase substrate chloroethylmalonyl-coenzyme A from S-adenosyl-L-methionine.* Proc Natl Acad Sci USA, 2009. **106**(30): p. 12295-12300.

46. Takahashi, S., et al., *Reveromycin A biosynthesis uses RevG and RevJ for stereospecific spiroacetal formation.* Nat Chem Biol, 2011. **7**(7): p. 461-468.

47. Liu, Y., et al., *Biosynthesis of salinosporamides from alpha,beta-unsaturated fatty acids: implications for extending polyketide synthase diversity.* J Am Chem Soc, 2009. **131**(30): p. 10376-10377.

48. Eustaquio, A.S. and B.S. Moore, *Mutasynthesis of fluorosalinosporamide, a potent and reversible inhibitor of the proteasome.* Angewandte Chemie. International Edition in English, 2008. **47**(21): p. 3936-8.

49. Qu, X.D., et al., *Cloning, sequencing and characterization of the biosynthetic gene cluster of sanglifehrin A, a potent cyclophilin inhibitor.* Mol. Biosyst., 2011. **7**(3): p. 852-861.

50. Xu, Z.L., L. Ding, and C. Hertweck, *A branched extender unit shared between two orthogonal polyketide pathways in an endophyte.* Angewandte Chemie. International Edition in English, 2011. **50**(20): p. 4667-4670.

51. Wilson, M.C., et al., *Structure and biosynthesis of the marine streptomycete ansamycin ansalactam a and its distinctive branched chain polyketide extender unit.* J Am Chem Soc, 2011. **133**(6): p. 1971-1977.

52. Erb, T.J., et al., *Carboxylation mechanism and stereochemistry of crotonyl-CoA carboxylase/reductase, a carboxylating enoyl-thioester reductase.* Proc Natl Acad Sci USA, 2009. **106**(22): p. 8871-6.

53. Quade, N., et al., *Unusual carbon fixation gives rise to diverse polyketide extender units.* Nature Chemical Biology, 2012. **8**(1): p. 117-124.

54. Yunt, Z., et al., *Cleavage of four carbon-carbon bonds during biosynthesis of the griseorhodin a spiroketal pharmacophore.* J Am Chem Soc, 2009. **131**(6): p. 2297-2305.

55. Pahari, P., et al., *Enzymatic total synthesis of defucogilvocarcin M and its implications for gilvocarcin biosynthesis.* Angew. Chem. Int. Ed. Engl., 2012. **51**(5): p. 1216-1220.

56. Kharel, M.K., et al., *GilR, an unusual lactone-forming enzyme involved in gilvocarcin biosynthesis.* Chembiochem, 2009. **10**(8): p. 1305-8.

57. Noinaj, N., et al., *The crystal structure and mechanism of an unusual oxidoreductase, GilR, involved in gilvocarcin V biosynthesis.* J Biol Chem, 2011. **286**(26): p. 23533-23543.

58. Finlay, A.C., et al., *Terramycin, a new antibiotic.* Science, 1950. **111**: p. 85.

59. Hatsu, M., et al., *A new tetracycline antibiotic with antitumor activity. I. Taxonomy and fermentation of the producing strain, isolation and characterization of SF2575.* J. Antibiot., 1992. **45**: p. 320-324.

60. Wells, J.S., et al., *Dactylocyclines, novel tetracycline derivatives produced by a Dactylosporangium sp. I. Taxonomy, production, isolation and biological activity.* J. Antibiot., 1992. **45**(12): p. 1892-8.

61. Pickens, L.B. and Y. Tang, *Decoding and engineering tetracycline biosynthesis.* Metab. Eng., 2009. **11**(2): p. 69.

62. McCormick, J.R.D., et al., *Biosynthesis of the tetracyclines. IV. Biological rehydration of the 5a,6-anhydrotetracyclines.* J. Am. Chem. Soc., 1962. **84**(15): p. 3023-3025.

63. McCormick, J.R.D., S. Johnson, and N.O. Sjolander, *Biosynthesis of the tetracyclines. V. Naphthacenic precursors.* J. Am. Chem. Soc., 1963. **85**(11): p. 1692-1694.

64. McCormick, J.R.D., et al., *Biosynthesis of the tetracyclines. VI. Total synthesis of a naphthacenic precursor: 1,3,10,11,12-pentahydroxynaphthacene-2-carboxamide.* J. Am. Chem. Soc., 1963. **85**(11): p. 1694-1695.

65. Ryan, M.J., *Cloning of the biosynthetic pathway for chlortetracycline and tetracycline formation and cosmids useful therein*, U.S. Patent, Editor. 1999: United States.

66. Zhang, W., et al., *Engineered Biosynthesis of a novel amidated polyketide, using the malonamyl-specific initiation module from the oxytetracycline polyketide synthase.* Appl. Environ. Microbiol., 2006. **72**: p. 2573-2580.

67. Pickens, L.B., et al., *Biochemical analysis of the biosynthetic pathway of an anticancer tetracycline SF2575.* J. Am. Chem. Soc., 2009. **131**(48): p. 17677.

68. Pickens, L.B. and Y. Tang, *Oxytetracycline biosynthesis.* J Biol Chem, 2010. **285**(36): p. 27509-15.

69. Zhou, H., Y. Li, and Y. Tang, *Cyclization of aromatic polyketides from bacteria and fungi.* Nat Prod Rep, 2010. **27**(6): p. 839-68.

70. Zhang, W., et al., *Identifying the Minimal Enzymes Required for Anhydrotetracycline Biosynthesis.* J. Am. Chem. Soc., 2008. **130**(19): p. 6068-6069.

71. Wang, P., et al., *Genetic characterization of enzymes involved in the priming steps of oxytetracycline biosynthesis in Streptomyces rimosus.* Microbiology, 2011. **157**(Pt 8): p. 2401-9.

72. Wang, P., et al., *Identification of OxyE as an Ancillary Oxygenase during Tetracycline Biosynthesis.* ChemBioChem, 2009. **10**(9): p. 1544-1550.

73. Rhodes, P.M., et al., *Biochemical and genetic characterization of Streptomyces rimosus mutants impaired in oxytetracycline biosynthesis.* J. Gen. Microbiol., 1981. **124**(Jun): p. 329-338.

74. Walsh, C., *Naturally occurring 5-deazaflavin coenzymes: biological redox roles.* Acc. Chem. Res., 1986. **19**(7): p. 216-221.

75. Peric-Concha, N., et al., *Ablation of the otcC gene encoding a post-polyketide hydroxylase from the oxytetracyline biosynthetic pathway in Streptomyces rimosus results in novel polyketides with altered chain length.* J. Biol. Chem., 2005. **280**(45): p. 37455-37460.

76. Chopra, I. and M. Roberts, *Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance.* Microbiol Mol Biol Rev, 2001. **65**(2): p. 232-60 ; second page, table of contents.

77. Thaker, M., P. Spanogiannopoulos, and G.D. Wright, *The tetracycline resistome.* Cell Mol Life Sci, 2010. **67**(3): p. 419-31.

78. Nelson, M.L. and S.B. Levy, *The history of the tetracyclines.* Ann N Y Acad Sci, 2011. **1241**: p. 17-32.

79. Béhal, V., Z. Hošťálek , and Z. Vaněk, *Anhydrotetracycline oxygenase activity and biosynthesis of tetracyclines in streptomyces aureofaciens.* Biotechnol. Lett., 1979. **1**(4): p. 177-182.

80. Binnie, C., M. Warren, and M.J. Butler, *Cloning and heterologous expression in Streptomyces lividans of Streptomyces rimosus genes involved in oxytetracycline biosynthesis.* J. Bacteriol., 1989. **171**(2): p. 887-95.

81. Butler, M.J. and B.N. Gedge, *Purification of anhydrotetracycline oxygenase from Streptomyces rimosus using fast protein liquid chromatography.* Biotechnol. Tech., 1989. **4**: p. 235-238.

82. Hunter, I.S. and R.A. Hill, *Tetracyclines*, in *Biotechnology of Antibiotics*, W.R. Strohl, Editor. 1997, Marcel Dekker: New York. p. 659-682.

83. Wang, P., et al., *Heterologous expression and manipulation of three tetracycline biosynthetic pathways.* Angew Chem Int Ed Engl, 2012. **51**(44): p. 11136-40.

84. Brodersen, D.E., et al., *The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit.* Cell, 2000. **103**(7): p. 1143-1154.

85. McCormick, J.R.D., et al., *Cosynthesis of tetracyclines by pairs of Streptomyces aureofaciens mutants.* J. Am. Chem. Soc., 1960. **82**(18): p. 5006-5007.

86. Nakano, T., et al., *Identification and cloning of the gene involved in the final step of chlortetracycline biosynthesis in Streptomyces aureofaciens.* Biosci. Biotechnol. Biochem., 2004. **68**(6): p. 1345-1352.

87. Nocek, B., et al., *Structure of an amide bond forming F(420):gamma-glutamyl ligase from Archaeoglobus fulgidus -- a member of a new family of non-ribosomal peptide synthases.* J Mol Biol, 2007. **372**(2): p. 456-69.

88. Stevens, D.C., et al., *Heterologous expression of the oxytetracycline biosynthetic pathway in Myxococcus xanthus.* Appl Environ Microbiol, 2010. **76**(8): p. 2681-3.

89. McDowall, K.J., A. Thamchaipenet, and I.S. Hunter, *Phosphate control of oxytetracycline production by Streptomyces rimosus is at the level of transcription from promoters overlapped by tandem repeats similar to those of the DNA-binding sites of the OmpR family.* J Bacteriol, 1999. **181**(10): p. 3025-32.

90. Selengut, J.D. and D.H. Haft, *Unexpected abundance of coenzyme F_{420}-dependent enzymes in Mycobacterium tuberculosis and other actinobacteria.* J. Bacteriol., 2010. **192**(21): p. 5788-5798.

91. Bashiri, G., et al., *Metabolic engineering of cofactor F_{420} production in Mycobacterium smegmatis.* PLoS One, 2010. **5**(12): p. e15803.

92. Purwantini, E. and L. Daniels, *Purification of a novel coenzyme F_{420}-dependent glucose-6-phosphate dehydrogenase from Mycobacterium smegmatis.* J Bacteriol, 1996. **178**(10): p. 2861-6.

93. Bashiri, G., et al., *Crystal structures of F_{420}-dependent glucose-6-phosphate dehydrogenase FGD1 involved in the activation of the anti-tuberculosis drug candidate PA-824 reveal the basis of coenzyme and substrate binding.* J Biol Chem, 2008. **283**(25): p. 17531-41.

94. Eirich, L.D., G.D. Vogels, and R.S. Wolfe, *Distribution of coenzyme F_{420} and properties of its hydrolytic fragments.* J Bacteriol, 1979. **140**(1): p. 20-7.

95. Walsh, C.T. and T.A. Wencewicz, *Flavoenzymes: versatile catalysts in biosynthetic pathways.* Nat Prod Rep, 2013. **30**(1): p. 175-200.

96. Charest, M.G., D.R. Siegel, and A.G. Myers, *Synthesis of (-)-tetracycline.* J Am Chem Soc, 2005. **127**(23): p. 8292-3.

97. Lindqvist, Y., et al., *Structural basis for substrate recognition and specificity in aklavinone-11-hydroxylase from rhodomycin biosynthesis.* J Mol Biol, 2009. **393**(4): p. 966-77.

98. Koskiniemi, H., et al., *Crystal structures of two aromatic hydroxylases involved in the early tailoring steps of angucycline biosynthesis.* J Mol Biol, 2007. **372**(3): p. 633-48.

99. Graham, D.E. and R.H. White, *Elucidation of methanogenic coenzyme biosyntheses: from spectroscopy to genomics.* Nat Prod Rep, 2002. **19**(2): p. 133-47.

100. Coats, J.H., et al., *Discovery, production, and biological assay of an unusual flavenoid cofactor involved in lincomycin biosynthesis.* J Antibiot (Tokyo), 1989. **42**(3): p. 472-4.

101. Kuo, M.S., et al., *Isolation and identification of 7,8-didemethyl-8-hydroxy-5-deazariboflavin, an unusual cosynthetic factor in streptomycetes, from Streptomyces lincolnensis.* J Antibiot (Tokyo), 1989. **42**(3): p. 475-8.

102. Caballero, J., et al., *Organization and functions of the actVA region of the actinorhodin biosynthetic gene cluster of Streptomyces coelicolor.* Mol. Gen. Genet., 1991. **230**: p. 401-412.

103. Doumith, M., et al., *Analysis of genes involved in 6-deoxyhexose biosynthesis and transfer in Saccharopolyspora erythraea.* Mol. Gen. Genet., 2000. **264**(4): p. 477-485.

104. Gust, B., et al., *PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin.* Proc. Natl. Acad. Sci. U. S. A., 2003. **100**(4): p. 1541-6.

105. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal. Biochem., 1976. **72**(1-2): p. 248-254.

106. Kabsch, W., *Xds.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 125-32.

107. McCoy, A.J., et al., *Phaser crystallographic software.* J. Appl. Crystallogr., 2007. **40**(Pt 4): p. 658-674.

108. Schwede, T., et al., *SWISS-MODEL: an automated protein homology-modeling server.* Nucleic Acids Res., 2003. **31**(13): p. 3381-3385.

109. Beam, M.P., et al., *Crystal structure of Baeyer-Villiger monooxygenase MtmOIV, the key enzyme of the mithramycin biosynthetic pathway.* Biochemistry, 2009. **48**(21): p. 4476-87.

110. Terwilliger, T.C., et al., *Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard.* Acta Crystallogr D Biol Crystallogr, 2008. **64**(Pt 1): p. 61-9.

111. Murshudov, G.N., A.A. Vagin, and E.J. Dodson, *Refinement of macromolecular structures by the maximum-likelihood method.* Acta Crystallogr. D Biol. Crystallogr., 1997. **53**: p. 240-255.

112. Blanc, E., et al., *Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT.* Acta Crystallogr. D Biol. Crystallogr., 2004. **60**: p. 2210-2221.

113. Winn, M.D., G.N. Murshudov, and M.Z. Papiz, *Macromolecular TLS refinement in REFMAC at moderate resolutions.* Methods Enzymol., 2003. **374**: p. 300-321.

114. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr. D Biol. Crystallogr., 2004. **60**: p. 2126-2132.

115. Laskowski, R.A., et al., *Procheck - a Program to Check the Stereochemical Quality of Protein Structures.* J. Appl. Crystallogr., 1993. **26**: p. 283-291.

116. Colovos, C. and T.O. Yeates, *Verification of protein structures: Patterns of nonbonded atomic interactions.* Protein Sci., 1993. **2**(9): p. 1511-1519.

117. Luthy, R., J.U. Bowie, and D. Eisenberg, *Assessment of protein models with three-dimensional profiles.* Nature, 1992. **356**(6364): p. 83-5.

118. Krissinel, E. and K. Henrick, *Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2256-68.

119. Lu, G.G., *TOP: a new method for protein structure comparisons and similarity searches.* J. Appl. Crystallogr., 2000. **33**: p. 176-183.

120. Shatsky, M., et al., *BioInfo3D: a suite of tools for structural bioinformatics.* Nucleic Acids Res., 2004. **32**(Web Server issue): p. W503-7.

121. Bond, C.S. and A.W. Schuttelkopf, *ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments.* Acta Crystallogr D Biol Crystallogr, 2009. **65**(Pt 5): p. 510-2.

122. Petkovic, H., et al., *Genetics of Streptomyces rimosus, the oxytetracycline producer.* Microbiol. Mol. Biol. Rev., 2006. **70**(3): p. 704-728.

123. Chen, Y., E. Wendt-Pienkowski, and B. Shen, *Identification and utility of FdmR1 as a Streptomyces antibiotic regulatory protein activator for fredericamycin production in Streptomyces griseus ATCC 49344 and heterologous Hosts.* J. Bacteriol., 2008. **190**(16): p. 5587-5596.

124. Jung, W.S., et al., *Enhanced heterologous production of desosaminyl macrolides and their hydroxylated derivatives by overexpression of the pikD regulatory gene in Streptomyces venezuelae.* Appl. Environ. Microbiol., 2008. **74**(7): p. 1972-9.

125. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products.* Proc. Natl. Acad. Sci. U. S. A., 2000. **97**(12): p. 6640-5.

126. Duggar, B.M., *Aureomycin; a product of the continuing search for new antibiotics.* Ann. NY Acad. Sci., 1948. **51**(2): p. 177-181.

127. Chopra, I., P.M. Hawkey, and M. Hinton, *Tetracyclines, molecular and clinical aspects.* J Antimicrob Chemother, 1992. **29**(3): p. 245-77.

128. Levy, S.B., et al., *Nomenclature for tetracycline resistance determinants.* Antimicrob. Agents Chemother., 1989. **33**(8): p. 1373-1374.

129. Levy, S.B., et al., *Nomenclature for new tetracycline resistance determinants.* Antimicrob. Agents Chemother., 1999. **43**(6): p. 1523-1524.

130. Martell, M.J. and J.H. Boothe, *The 6-deoxytetracyclines. VII. alkylated aminotetracyclines possessing unique antibacterial activity.* J. Med. Chem., 1967. **10**(1): p. 44-46.

131. Church, R.F.R., R.E. Schaub, and M.J. Weiss, *Synthesis of 7-Dimethylamino-6-demethyl-6-deoxytetracycline (Minocycline) via 9-Nitro-6-demethyl-6-deoxytetracycline.* J. Org. Chem., 1971. **36**(5): p. 723-725.

132. Sum, P.-E. and P. Petersen, *Synthesis and structure-activity relationship of novel glycylcycline derivatives leading to the discovery of GAR-936.* Bioorg. Med. Chem. Lett., 1999. **9**(10): p. 1459-1462.

133. Sum, P.-E., et al., *Synthesis and antibacterial activity of 9-substituted minocycline derivatives.* Bioorg. Med. Chem. Lett., 2006. **16**(2): p. 400-403.

134. Agwuh, K.N. and A. MacGowan, *Pharmacokinetics and pharmacodynamics of the tetracyclines including glycylcyclines.* J. Antimicrob. Chemother., 2006. **58**(2): p. 256-265.

135. Charest, M.G., et al., *A convergent enantioselective route to structurally diverse 6-deoxytetracycline antibiotics.* Science, 2005. **308**(5720): p. 395-8.

136. Sun, C., et al., *A robust platform for the synthesis of new tetracycline antibiotics.* J Am Chem Soc, 2008. **130**(52): p. 17913-27.

137. Hatsu, M., et al., *A new tetracycline antibiotic with antitumor activity. II. The structural elucidation of SF2575.* J. Antibiot., 1992. **45**: p. 325-330.

138. Tymiak, A.A., et al., *Dactylocyclines: novel tetracycline glycosides active against tetracycline-resistant bacteria.* J. Org. Chem., 1993. **58**(3): p. 535-537.

139. Horiguchi, T., et al., *New naphthacenecarboxamide antibiotics, TAN-1518A and B, have inhibitory activity against mammalian DNA topoisomerase I.* J. Antibiot., 1994. **47**(5): p. 545-556.

140. Zhang, W., et al., *Investigation of early tailoring reactions in the oxytetracycline biosynthetic pathway.* J. Biol. Chem., 2007. **282**: p. 25717-25725.

141. Zhang, W., et al., *A New Mechanism for Benzopyrone Formation in Aromatic Polyketide Biosynthesis.* J. Am. Chem. Soc., 2007. **129**(30): p. 9304-9305.

142. Brown, J.R. and D.S. Ireland, *Structural Requirements for Tetracycline Activity*, in *Advances in Pharmacology and Chemotherapy*, S. Garattini, et al., Editors. 1978, Academic Press: New York. p. 161-202.

143. Pickens, L.B., et al., *Structural and biochemical characterizations of the salicylyl-acyltranferase SSFX3 from a tetracycline biosynthetic pathway.* J. Biol. Chem., 2011.

144. Eustaquio, A.S., et al., *Production of 8'-halogenated and 8'-unsubstituted novobiocin derivatives in genetically engineered streptomyces coelicolor strains.* Chem Biol, 2004. **11**(11): p. 1561-72.

145. Chen, W., et al., *Characterization of the polyoxin biosynthetic gene cluster from Streptomyces cacaoi and engineered production of polyoxin H.* J Biol Chem, 2009. **284**(16): p. 10627-38.

146. Wietzorrek, A. and M. Bibb, *A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold.* Mol Microbiol, 1997. **25**(6): p. 1181-4.

147. Bruheim, P., et al., *High-yield actinorhodin production in fed-batch culture by a Streptomyces lividans strain overexpressing the pathway-specific activator gene actII-ORF4.* J Ind Microbiol Biotechnol, 2002. **28**(2): p. 103-11.

148. Bibb, M.J., et al., *The mRNA for the 23S rRNA methylase encoded by the ermE gene of Saccharopolyspora erythraea is translated in the absence of a conventional ribosome-binding site.* Mol Microbiol, 1994. **14**(3): p. 533-45.

149. Lozano, M.J., et al., *Characterization of two polyketide methyltransferases involved in the biosynthesis of the antitumor drug mithramycin by Streptomyces argillaceus.* J. Biol. Chem., 2000. **275**(5): p. 3065-74.

150. Cooke, H.A., et al., *Molecular Basis of Substrate Promiscuity for the SAM-Dependent O-Methyltransferase NcsB1, Involved in the Biosynthesis of the Enediyne Antitumor Antibiotic Neocarzinostatin.* Biochemistry, 2009. **48**(40): p. 9590-9598.

151. Cooke, H.A., et al., *Molecular basis of substrate promiscuity for the SAM-dependent O-methyltransferase NcsB1, involved in the biosynthesis of the enediyne antitumor antibiotic neocarzinostatin.* Biochemistry, 2009. **48**(40): p. 9590-8.

152. Bililign, T., et al., *The hedamycin locus implicates a novel aromatic PKS priming mechanism.* Chem Biol, 2004. **11**(7): p. 959-69.

153. Kunzel, E., et al., *Inactivation of the urdGT2 gene, which encodes a glycosyltransferase responsible for the C-glycosyltransfer of activated D-olivose, leads to formation of the novel urdamycins I, J, and K.* Journal of the American Chemical Society, 1999. **121**(48): p. 11058-11062.

154. Durr, C., et al., *The glycosyltransferase UrdGT2 catalyzes both C- and O-glycosidic sugar transfers.* Angew. Chem. Int. Ed. Engl., 2004. **43**(22): p. 2962-2965.

155. Hoffmeister, D., et al., *The C-glycosyltransferase UrdGT2 is unselective toward D- and L-configured nucleotide-bound rhodinoses.* J. Am. Chem. Soc., 2003. **125**(16): p. 4678-4679.

156. Zhang, H., et al., *Complete biosynthesis of erythromycin A and designed analogs using E. coli as a heterologous host.* Chem Biol, 2010. **17**(11): p. 1232-40.

157. Pickens, L.B., *Characterization of the biosynthetic pathway of anticancer tetracycline SF2575*, in *Department of Chemical and Biomolecular Engineering* 2011, University of California, Los Angeles.: Los Angeles.

158. Li, S., et al., *Selective oxidation of carbolide C-H bonds by an engineered macrolide P450 mono-oxygenase.* Proc Natl Acad Sci U S A, 2009. **106**(44): p. 18463-8.

159. Wasserman, H.H., T.J. Lu, and A.I. Scott, *On the total synthesis of tetracycline.* J.Am. Chem. Soc., 1986. **108**(14): p. 4237-4238.

160. Nakano, T., et al., *Identification and cloning of the gene involved in the final step of chlortetracycline biosynthesis in Streptomyces aureofaciens.* Bioscience Biotechnology and Biochemistry, 2004. **68**(6): p. 1345-1352.

161. Wang, P., et al., *Uncovering the enzymes that catalyze the final steps in oxytetracycline biosynthesis.* J Am Chem Soc, 2013. **135**(19): p. 7138-41.

162. Chooi, Y.-H., R. Cacho, and Y. Tang, *Identification of the Viridicatumtoxin and Griseofulvin Gene Clusters from Penicillium aethiopicum.* Chemistry & Biology, 2010. **17**(5): p. 483-494.

163. Chooi, Y.H., et al., *Discovery and characterization of a group of fungal polycyclic polyketide prenyltransferases.* J Am Chem Soc, 2012. **134**(22): p. 9428-37.

164. Hertweck, C., *The biosynthetic logic of polyketide diversity.* Angew. Chem. Int. Ed. Engl., 2009. **48**(26): p. 4688-716.

165. Lombó, F., et al., *The aureolic acid family of antitumor compounds: structure, mode of action, biosynthesis, and novel derivatives.* Appl. Environ. Microbiol., 2006. **73**(1): p. 1-14.

166. Martin, W., et al., *Syntheses in tetracycline series .1. total synthesis of d,1-7-chloro-6-desoxytetracyclines and d,1-7-chloro-6-desmethyl-6-desoxytetracyclines of natural, 5a-epi and 6-epi series.* Tetrahedron Letters, 1973(36): p. 3513-3516.

167. Madduri, K. and C.R. Hutchinson, *Functional characterization and transcriptional analysis of a gene cluster governing early and late steps in daunorubicin biosynthesis in Streptomyces peucetius.* J Bacteriol, 1995. **177**(13): p. 3879-84.

168. Wu, Y., et al., *N-methylation of the amide bond by methyltransferase asm10 in ansamitocin biosynthesis.* ChemBioChem, 2011. **12**(11): p. 1759-66.

169. Salas, J.A. and C. Mendez, *Engineering the glycosylation of natural products in actinomycetes.* Trends Microbiol, 2007. **15**(5): p. 219-32.

170. Mendez, C., et al., *Deoxysugars in bioactive natural products: development of novel derivatives by altering the sugar pattern.* Curr Top Med Chem, 2008. **8**(8): p. 710-24.

171. Bililign, T., B.R. Griffith, and J.S. Thorson, *Structure, activity, synthesis and biosynthesis of aryl-C-glycosides.* Nat Prod Rep, 2005. **22**(6): p. 742-60.

172. Mittler, M., A. Bechthold, and G.E. Schulz, *Structure and action of the C-C bond-forming glycosyltransferase biosynthesis of the UrdGT2 involved in the antibiotic urdamycin.* J. Mol. Biol., 2007. **372**(1): p. 67-76.

173. Trefzer, A., et al., *Rationally designed glycosylated premithramycins: Hybrid aromatic polyketides using genes from three different biosynthetic pathways.* J. Am. Chem. Soc., 2002. **124**(21): p. 6056-6062.

174. Harle, J., et al., *Rational design of an aryl-C-glycoside catalyst from a natural product O-glycosyltransferase.* Chem Biol, 2011. **18**(4): p. 520-30.

175. Duarte, H.A., et al., *Importance of tautomers in the chemical behavior of tetracyclines.* Journal of Pharmaceutical Sciences, 1999. **88**(1): p. 111-120.

176. Blackwood, R.K. and A.R. English, *Structure–Activity relationships in the tetracycline series*, in *Adv. Appl. Microbiol.*, D. Perlman, Editor. 1970, Academic Press. p. 237-266.

177. Kieser, T., et al., *Practical Streptomyces Genetics*. 2000, Norwich, UK: John Innes Foundation.

178. Gust, B., et al., *PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin.* Proc Natl Acad Sci U S A, 2003. **100**(4): p. 1541-6.

179. Gust, B., et al., *Lambda red-mediated genetic manipulation of antibiotic-producing Streptomyces.* Adv Appl Microbiol, 2004. **54**: p. 107-28.

180. Margulies, M., et al., *Genome sequencing in microfabricated high-density picolitre reactors.* Nature, 2005. **437**(7057): p. 376-80.

181. Tang, Y., A.T. Koppisch, and C. Khosla, *The acyltransferase homologue from the initiation module of the R1128 polyketide synthase is an acyl-ACP thioesterase that edits acetyl primer units.* Biochemistry, 2004. **43**(29): p. 9546-55.

182. Gullon, S., et al., *Isolation, characterization, and heterologous expression of the biosynthesis gene cluster for the antitumor anthracycline steffimycin.* Appl Environ Microbiol, 2006. **72**(6): p. 4172-83.

183. Kunzel, E., et al., *Tetracenomycin M, a novel genetically engineered tetracenomycin resulting from a combination of mithramycin and tetracenomycin biosynthetic genes.* Chem. Eur. J., 1997. **3**(10): p. 1675-1678.

184. Daum, M., et al., *Organisation of the biosynthetic gene cluster and tailoring enzymes in the biosynthesis of the tetracyclic quinone glycoside antibiotic polyketomycin.* ChemBioChem, 2009. **10**(6): p. 1073-1083.

185. Sohng, J.K., et al., *Identification of a gene cluster of biosynthetic genes of rubradirin substructures in S. achromogenes var. rubradiris NRRL3061.* Mol Cells, 1997. **7**(5): p. 674-81.

186. Lesnik, U., et al., *Regulatory Elements in Tetracycline-Encoding Gene Clusters: the otcG Gene Positively Regulates the Production of Oxytetracycline in Streptomyces rimosus.* Food Technol. Biotech., 2009. **47**(3): p. 323-330.

187. Ramos, J.L., et al., *The TetR family of transcriptional repressors.* Microbiol Mol Biol Rev, 2005. **69**(2): p. 326-56.

188. Ohnuki, T., et al., *Molecular cloning of tetracycline resistance genes from Streptomyces rimosus in Streptomyces griseus and characterization of the cloned genes.* J. Bacteriol., 1985. **161**(3): p. 1010-1016.

189. Burdett, V., *Tet(M)-promoted release of tetracycline from ribosomes is GTP dependent.* J. Bacteriol., 1996. **178**(11): p. 3246-3251.

190. Kim, B.C., et al., *Cloning, sequencing, and characterization of the pradimicin biosynthetic gene cluster of Actinomadura hibisca P157-2.* J Microbiol Biotechnol, 2007. **17**(5): p. 830-9.

191. Zhang, X. and R.J. Parry, *Cloning and characterization of the pyrrolomycin biosynthetic gene clusters from Actinosporangium vitaminophilum ATCC 31673 and Streptomyces sp. strain UC 11065.* Antimicrob Agents Chemother, 2007. **51**(3): p. 946-57.

192. Tymiak, A.A., et al., *Dactylocyclines, novel tetracycline derivatives produced by a Dactylosporangium sp. II. Structure elucidation.* J. Antibiot., 1992. **45**(12): p. 1899-906.