## Title

Mechanism Studies on Fungal Type I Highly-Reducing Polyketide Synthases and Polyketide Synthase-Nonribosomal Peptide Synthetase Hybrids



## Author

Xu, Wei

## Publication Date

2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

# Mechanism Studies on Fungal Type I Highly-Reducing Polyketide Synthases and Polyketide Synthase-Nonribosomal Peptide Synthetase Hybrids

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Chemical Engineering

by

Wei Xu

2013

## ABSTRACT OF THE DISSERTATION

Mechanism Studies on Fungal Type I Highly-Reducing Polyketide Synthases and Polyketide Synthase-Nonribosomal Peptide Synthetase Hybrids

by

Wei Xu

Doctor of Philosophy in Chemical Engineering

University of California, Los Angeles, 2013

Professor Yi Tang, Chair

Filamentous fungi are known as promising sources for bio-active natural products, some of which are blockbuster drugs, such as penicillin and lovastatin. Fungal polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs) are large highly complex multidomain megasynthetases with complex programming rules. Different from their well-studied bacterial counterparts, the mechanisms of these megasynthases are not well understood to date. The thesis focuses on mechanism study of two fungal highly reducing PKSs involved in lovastatin biosynthesis and a PKS-NRPS hybrid from *Apsergillus sp*.

Using the developed expression system in *S. cerevisiae*, we were able to probe the property of these enzymes under different conditions with *in vivo* experiment in the heterologous host and *in vitro* assays.

In the lovastatin project, we observed and studied interaction between highly-reducing polyketide synthase (HR-PKS) and its product releasing partner. These PKS/releasing-enzyme pairs could widely exist in the HR-PKS systems. It also shows applications to drug industry because the efficiency of product releasing from PKSs could directly determine the yield of the natural products. In the study related to PKS-NRPS hybrid, we were able to be the first group to reconstitute full function of this type of megasynthetase *in vitro* and probe the programming on both PKS and NRPS module in this hybrid. More importantly, we also observed interaction between these two modules.

The dissertation of Wei Xu is approved.

Michael Jung

James Liao

Tatiana Segura

Yi Tang, Committee Chair

University of California, Los Angeles

2013

TABLE OF CONTENTS

## TABLE OF FIGURES

## LIST OF TABLES

## ACKNOWLEDGMENTS

1. Section 1.1.3 and 1.3.1.1 contain materials from Xu, W.; Qiao, K.; Tang, Y. *Critical reviews in biochemistry and molecular biology* **2013**, *48*, 98.
2. Section 2.1 contains materials from Xie, X.; Meehan, M. J.; Xu, W.; Dorrestein, P. C.; Tang, Y. *J Am Chem Soc* **2009**, *131*, 8388. and Xu, W.; Chooi, Y. H.; Choi, J. W.; Li, S.; Vederas, J. C.; Da Silva, N. A.; Tang, Y. *Angew. Chem. Int. Ed. Engl.* **2013**, DOI: 10.1002/anie.201302406.
3. Section 2.2 contains material from Xu, W.; Cai, X. L.; Jung, M. E.; Tang, Y. *J Am Chem Soc* **2010**, *132*, 13604.

The work described in this dissertation is supported by NIH grants 1R01GM085128 and 1R01GM092217 to Yi Tang, in part by Natural Sciences & Engineering Research Council of Canada (NSERC) and Canada Research Chair in Bioorganic & Medicinal Chemistry to our collaborator John C. Vederas.

I would like to thank Prof. D. K. Ro for providing the expression plasmid of LovA in the LovG work and thank Prof. David Tirrell for providing all the unnatural phenylalanine analogs as substrates for the ApdA *in vitro* assays.

Additionally, I would like to thank the many individuals who have helped me with my research or my personal and professional development during the path I was pursuing my PhD.

First and foremost, I would like to give my sincere appreciations to my advisor Professor Yi Tang, not only for his constant support and encouragement, but also for his guidance, advices and inspirations to me. The many skills I've learnt from Dr. Tang have been extremely valuable to me and I will treasure this experience for the rest of my life.

I would also like to thank my committee members Professor Michael Jung, Professor James Liao and Professor Tatiana Segura for their support of my doctoral candidacy and insightful questions and discussions along the way.

I sincerely thank Dr. Xinkai Xie and Dr. Wenjun Zhang their help and training at the beginning of my research in UCLA. I am very grateful to Dr. Yit-Heng Chooi for the insightful discussion and suggestions. Also special thanks to Dr. Yanran Li, Dr. Kangjian Qiao, Angelica Zabala, Dr. Hsiao-Ching Lin and Dr. Shuang Li for the cooperation and discussion on some projects.

I would like to thank Dr. Robert Peterson, Prof. Robert T. Clubb and his graduate students Albert Chan, Megan Sjodt and Brendan Amer for their help with protein NMR experiments.

Also many thanks to all my colleagues in the lab, Dr. Woncheol Kim, Dr. Zhen Gu, Dr. Hui Zhou, Dr. Lauren Pickens, Dr. Jaclyn Winter, Dr. Yingtong Di, Anuradha Biswas, Peng Wang, Xue Gao, Muxun Zhao, Ralph Cacho, Jingjing Wang and Bo Wang, for their company and support.

Finally I would like to thank my parents, my sister and her family for their continual encouragement and support over all these years.

## VITA

| | |
|---|---|
| 2002 | B. S., Applied Chemistry<br>China University of Petroleum<br>Dongying, China |
| 2005 | M. S., Inorganic Chemistry<br>Nanjing University<br>Nanjing, China |
| 2009-2012 | Teaching Assistant<br>Department of Chemical and Biomolecular Engineering<br>University of California, Los Angeles, USA |
| 2009-2013 | Research Assistant<br>Department of Chemical and Biomolecular Engineering<br>University of California, Los Angeles, USA |

## PUBLICATIONS

1) Xie, X., Meehan, M. J. **Xu, W.**, Dorrestein, P. C., Tang, Y. "Acyltransferase Mediated Polyketide Release from Fungal Megasynthase." *J. Am. Chem. Soc*. **2009**, 131, 8388-8389.

2) Li, Y., **Xu, W.**, Tang, Y. "Classification, Prediction and Verification of the Regioselectivity of Fungal Polyketide Synthase Product Template Domains." *J. Biol. Chem.* 2010, 285, 22762-22771.

3) **Xu, W.**, Cai, X., Jung, M. E., Tang, Y. "Analysis of Intact and Dissected Fungal Polyketide Synthase-Nonribosomal Peptide Synthetase in vitro and in *Saccharomyces cerevisiae*" *J. Am. Chem. Soc.* **2010**, 132, 13604-13607.

4) Meehan, M. J., Xie, X., Zhao, X., **Xu, W.**, Tang, Y., Dorrestein, P. C. "FT-ICR-MS characterization of intermediates in the biosynthesis of the α-methylbutyrate side chain of lovastatin by the 277 kDa polyketide synthase LovF." *Biochemistry* **2011**, 50, 287–299.

5) Qiao, K., Zhou, H., **Xu, W.**, Zhang, W., Garg, N., Tang, Y. "A Fungal Nonribosomal Peptide Synthetase Module that can Synthesize Thiopyrazines." *Org. Lett.* **2011**, 13, 1758-1761.

6) Liu, Q., Yao, F., Chooi, Y., Kang, Q., **Xu, W.**, Li, Y., Shao, Y., Shi, Y., Deng, Z., Tang, Y.., You, D. "Elucidation of Piericidin A1 Biosynthetic Locus Revealed an Thioesterase-dependent Mechanism of α-Pyridone Ring Formation." *Chem. Biol.* **2012**, 19, 243-253.

7) Zabala, A. O., **Xu, W.**, Chooi, Y., Tang, Y. "Discovery and Characterization of a Silent Gene Cluster that Produces Azaphilones from *Aspergillus niger* ATCC 1015 Reveal a Hydroxylation-Mediated Pyran-Ring Formation." *Chem. Biol.* **2012**, 19, 1049–1059.

8) Ames, B. D., Nguyen, C., Bruegger, J., Smith, P., **Xu, W.**, Ma, S. M., Wong, E., Wong, S., Xie, X., Li, J. W., Vederas, J. C., Tang, Y., Tsai, S. C. **"**Crystal structure and biochemical studies of the trans-acting polyketide enoyl reductase LovC from lovastatin biosynthesis" *Proc. Natl. Acad. Sci. U. S. A*. **2012**, 109, 11144-11149.

9) **Xu, W.**, Qiao, K., Tang, Y. "Structural Analysis of Protein-Protein Interactions in Type I Polyketide Synthases." *Crit. Rev. Biochem. Mol. Biol*. **2012**, 48, 98-122.

10) Lin, H-S., Chooi, Y., Dhingra, S., **Xu, W.**, Calvo, A. M., Tang, Y. "The Fumagillin Biosynthetic Gene Cluster in *Aspergillus fumigatus* Encodes a Cryptic Terpene Cyclase Involved in the Formation of β-*trans*-Bergamotene." *J. Am. Chem. Soc.* **2013**, 135, 4616-4619.

11) **Xu, W.**, Chooi, Y., Choi, J. W., Li, S., Vederas, J. C., Da Silva, N. A., Tang, Y. "LovG: The Thioesterase Required for Dihydromonacolin L Release and Lovastatin Nonaketide Synthase Turnover in Lovastatin Biosynthesis." *Angew. Chemie. Intl. Ed.* **2013**, in press.

## 1. Background and Motivation

### 1.1 Natural Products and Natural Product Biosynthesis

In order to live, grow, and reproduce, all organisms have developed ways to interconvert a large number of organic compounds. All these compounds, or metabolites, can be called natural products. Although the organisms can be extremely different in their characteristics, their primary metabolites, including intermediates for the synthesis and modification of carbohydrates, proteins and nucleic acids, are essentially the same. The term "natural products" often refers to the secondary metabolites, which are not necessarily produced under all conditions and are specific to the organisms. In most cases, the functions of secondary metabolites to the organism are not yet known. Some toxic natural products are believed to provide defenses against predators or competitors.

Natural products have played important roles throughout the world in treating human diseases. From ancient times, mankind knew to search natural sources for curing diseases from herb mixture. Later, people started to realize that some of these medicinal functions originate from certain compounds that could be purified from natural organisms. For example, morphine, a single molecule that isolated from opium, was widely used to relieve severe pain. Penicillin, the first effective drug against many previously lethal diseases, is a natural product isolated from certain species of *Penicillium.* Nowadays, natural products still play significant roles in drug discovery and development [1-4]. In modern drug market, natural products are still the major sources for developing innovative pharmaceuticals. From 1981 to 2010, the percentage of natural products and natural product derived compounds in the approved small-molecule drugs has always been at ~30% [1].

According to their biosynthetic origins, natural products are usually classified into several major catalogs: alkaloids, terpenoids, polyketides and nonribosomal peptides.

### 1.1.1 Alkaloids and Alkaloid Biosynthesis

Alkaloids are a diverse group of natural products that contain one or multiple nitrogen atoms in their structure. They are generally derived from nitrogen containing precursors such as amino acids. In 1806, morphine was isolated as the first example of alkaloid but its structure was not elucidated until 1952[5]. Among the approximately 12,000 known alkaloids, many have been exploited as pharmaceuticals and are currently in clinical use[6], including the analgesics morphine and codeine,[7] the muscle relaxant papaverine and tubocurarine[8], the gout suppressant colchicine[9], the anticancer agents vinblastine[10], the antibiotic sanguinarine[11]. Other important alkaloids include caffeine, nicotine, quinine and cocaine (**Figure 1**).

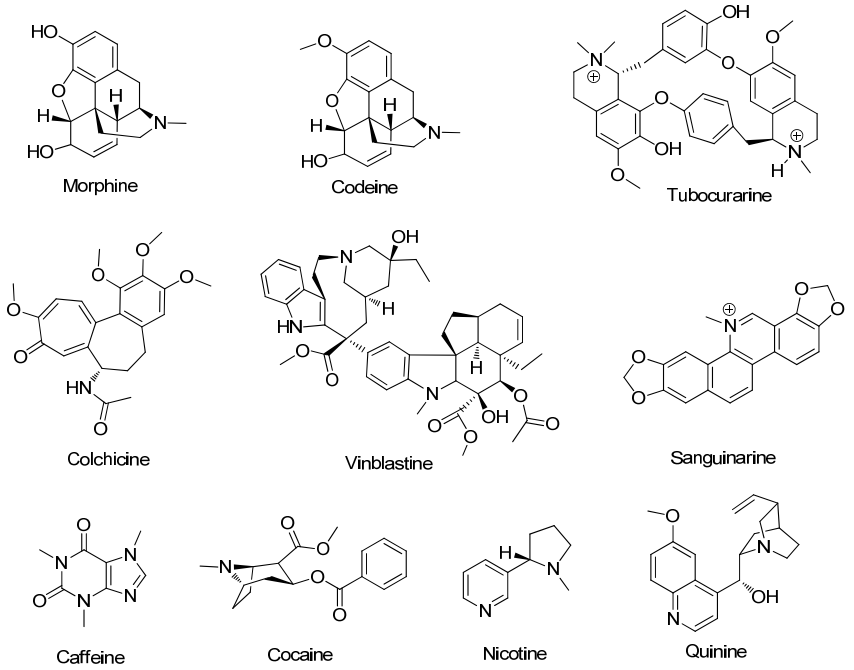



**Figure 1.** Chemical structures of some important alkaloids

Unlike other types of secondary metabolites, alkaloids usually don't share a common biosynthetic origin. Each class of alkaloids are biosynthesized through unique strategy[6]. For

example, monoterpenoid indole alkaloids are condensational products of an indole moiety derived from tryptamine and a monoterpenoid component; benzylisoquinoline alkaloids biosynthesis involves the condensation of two tyramine or dopamine moieties, which are derived from decarboxylation of tyrosine or dihydroxyphenylalanine by tyrosine decarboxylase; tropane alkaloids and nicotine biosynthesis begin with the methylation of putrescine to N-methylputrescine; purine alkaloids biosynthesis begin with the N-methylation of xanthosine [6] (**Figure 2**).

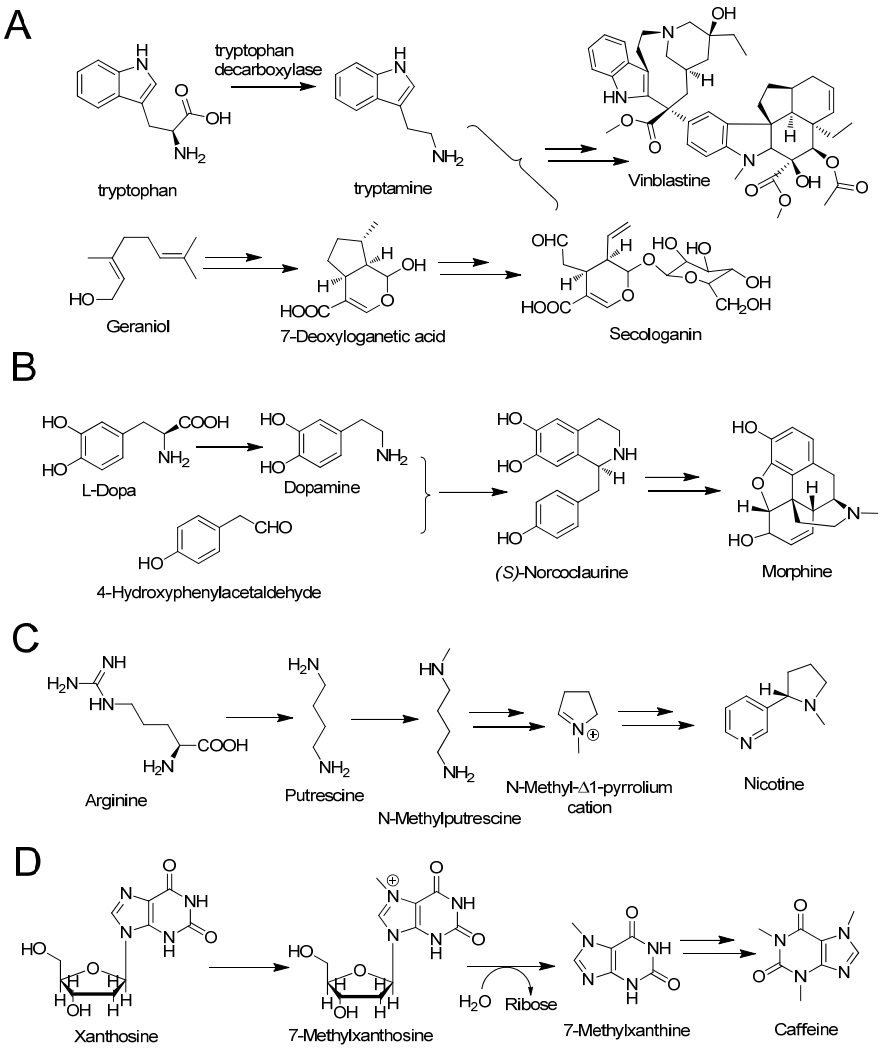



**Figure 2.** Biosynthesis of different classes of alkaloids (**A**) monoterpenoid indole alkaloids; (**B**) benzylisoquinoline alkaloids; (**C**) tropane alkaloids; (**D**) purine alkaloids

### 1.1.2 Terpenoids and Terpenoid Biosynthesis

Terpenoids are derived from five-carbon isoprene building blocks, hence, sometimes termed isoprenoids[12]. The building blocks for terpenoids are isopentenyl pyrophosphate (IPP) and its carbon–carbon double bond isomer dimethylallyl pyrophosphate (DMAPP). According to the numbers of isoprene unit, they are classified as hemiterpenoids (C5), monoterpenoids (C10), sesquiterpenoids (C15), diterpenoids (C20), etc. Terpenoids form a big family of natural products that show diverse bioactivities. For example, menthol is a monoterpenoid that is used to relieve minor throat irritation; another monoterpenoid, thymol, is well known as a biocide; artemisinin is a sesquiterpenoid lactone whose derivative are now standard treatment for *P. falciparum* malaria; forskolin is a diterpenoid commonly used to raise levels of cyclic AMP (cAMP) in cell physiology research; the diterpenoid taxol is a widely used drug in cancer chemotherapy because of its promising mitotic inhibitory activity[13] (**Figure 3**).

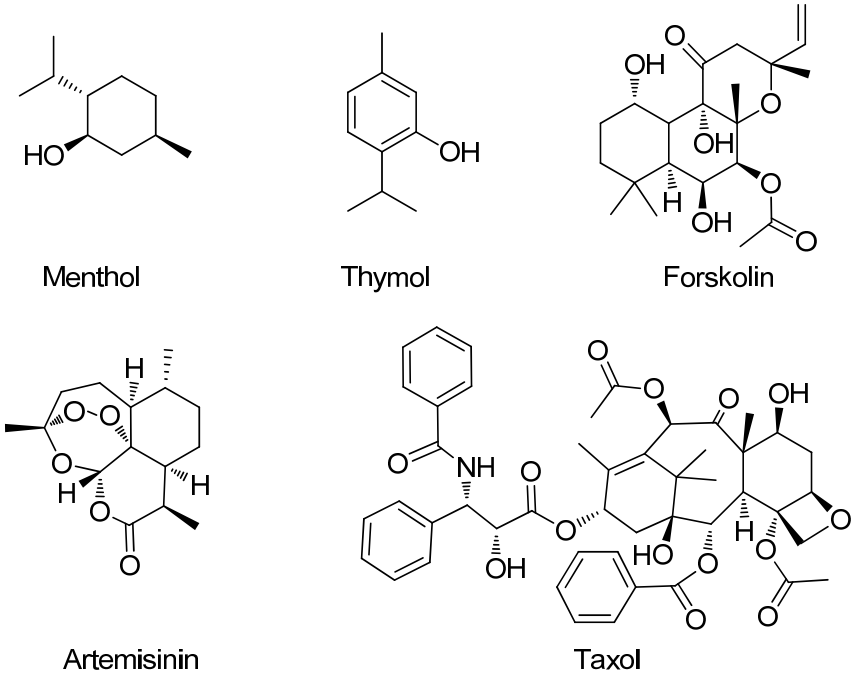



**Figure 3.** Chemical structures of some important terpenoids

Surprisingly, although over 55, 000 terpenoids have been characterized to date, they are derived from a handful of linear isoprenoid precursors, which are geranyl diphosphate (C10), farnesyl diphosphate (C15), geranylgeranyl diphosphate (C20), and geranylfarnesyl diphosphate (C25). This diversity is appreciated by the manifold ways in which the constituent isoprene units are cyclized and rearranged **(Figure 4)**. Terpenoid cyclases are the major enzymes that control the folding processes and generate a unique hydrocarbon product with remarkable structural and stereochemical precision[14].

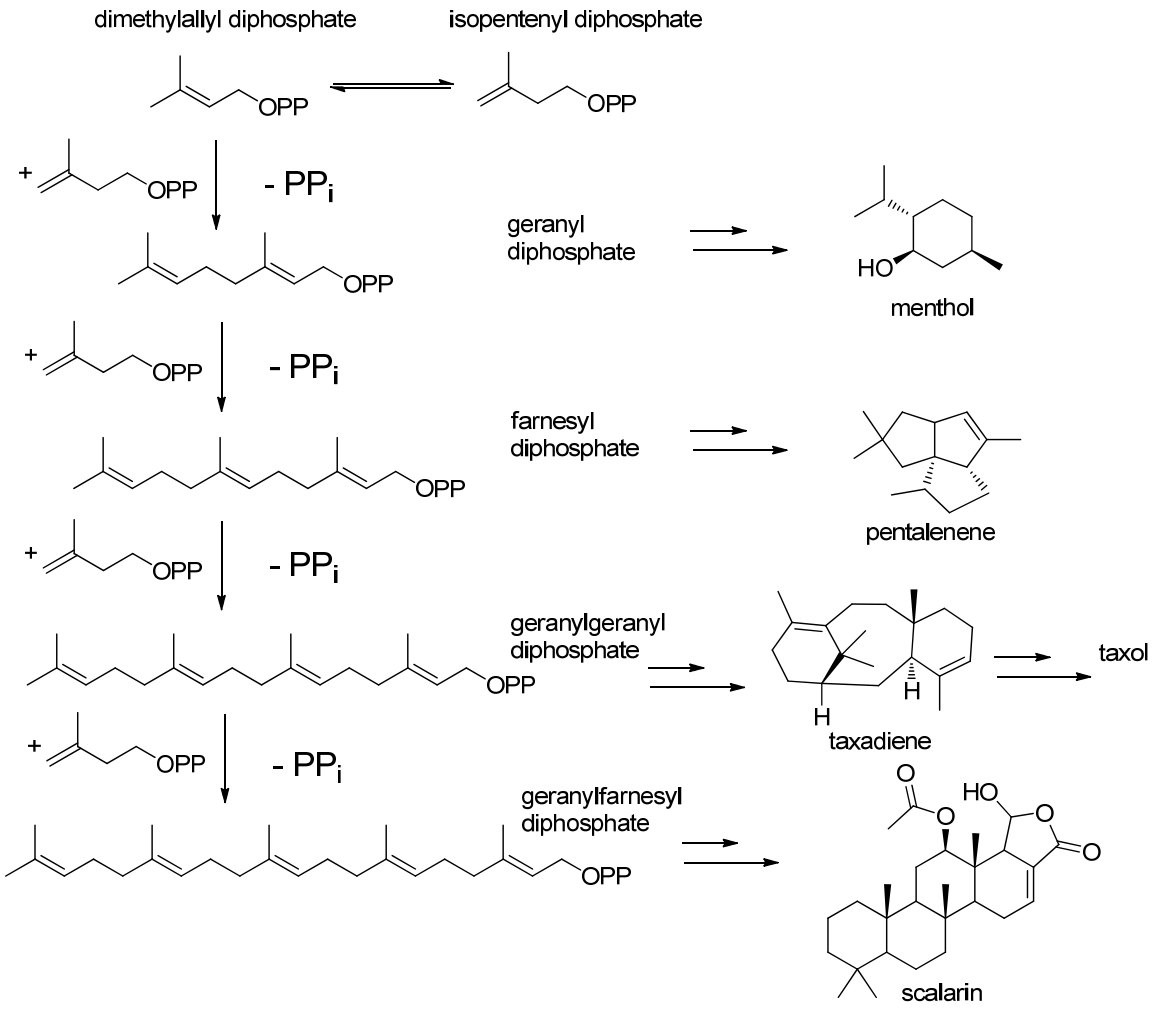



**Figure 4.** Linear isoprenoid precursors in terpenoid biosynthesis

### 1.1.3 Polyketides and Polyketide Biosynthesis

Polyketides constitute another major family of secondary metabolites that have diverse structures and broad biological activities [15, 16], and they have been ample sources for the development of frontline therapeutics towards different diseases [1]. Many polyketide natural products have become important drugs, including human antibiotics such as erythromycin A and tetracycline; veterinary antibiotic tylosin; anti-cancer agent doxorubicin; immunosuppressant rapamycin and cholesterol-lowering drug lovastatin (**Figure 5**).

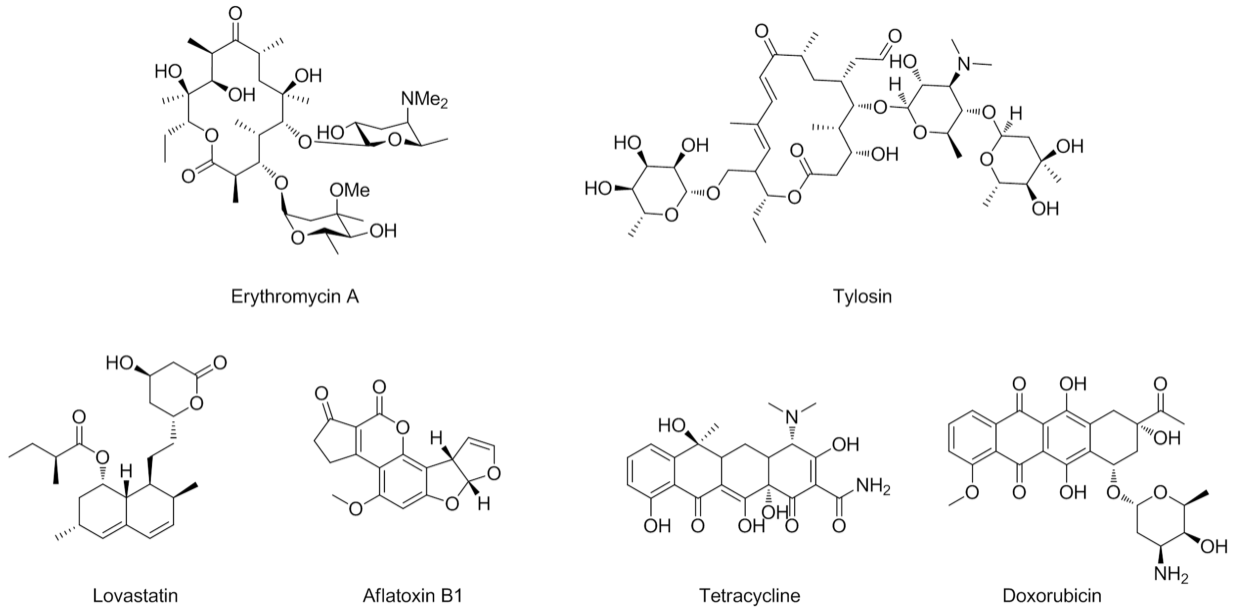



**Figure 5.** Examples of polyketides synthesized by different types of polyketide synthases

Notwithstanding the diversity displayed by the structures, all polyketides are assembled from simple cellular building blocks, such as acetate, by a group of enzymes called polyketide synthases (PKSs) [17]. The polyketide biosynthesis resembles the biosynthesis of fatty acids. Both fatty acid synthases (FASs) and PKSs are composed of catalytic domains that work in coordination to polymerize the simple building blocks (mostly in the form of acyl-CoAs). The catalytic processes on both FASs and PKSs can be divided into three steps: initiation, elongation

and termination. In the initiation step, the acyl carrier protein (ACP) must first be posttranslationally phosphopantetheinylated at the active site serine within the conserved DSL motif by a phosphopantetheinyl transferase (PPTase). The PPTase is either encoded in the PKS gene cluster as a dedicated enzyme, or is shared with FASs or other PKSs in the same host [18]. The acyl-CoA:ACP transacylase (AT) is responsible for transferring the monomeric acyl groups, such as malonyl or methylmalonyl, from the coenzyme A carrier to the free thiol of the phosphopantetheine (Ppant) arm on the modified (*holo*-) ACP. The ketosynthase (KS) catalyzes the successive decarboxylative Claisen condensations between the growing polyketide chain and the monomeric/extender unit supplied by ACP to elongate the nascent polyketide backbone. AT, KS and ACP are the minimal domains required for the PKS synthesis hence are called minimal PKSs. Following each chain extension step in which the polyketide backbone is increased in size by one ketide, tailoring domains such as β-ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) can selectively modify the newly formed polyketide. Whereas in FASs, the β-keto positions are completely reduced to methylenes and yield a saturated fatty acyl product, the tailoring domains in PKS can be used in different combinations at each β-position to yield products of different oxidation states. Also unique to PKSs, a methyl group may be introduced at the α position of the growing chain by *C*-methyltransferase (MT) domain [19]. Finally, when the polyketide chain reaches the desired size, the product can be offloaded from the PKSs via different release mechanisms, such as hydrolysis or macrocyclization catalyzed by a thioesterase (TE) [20] (**Figure 6**).

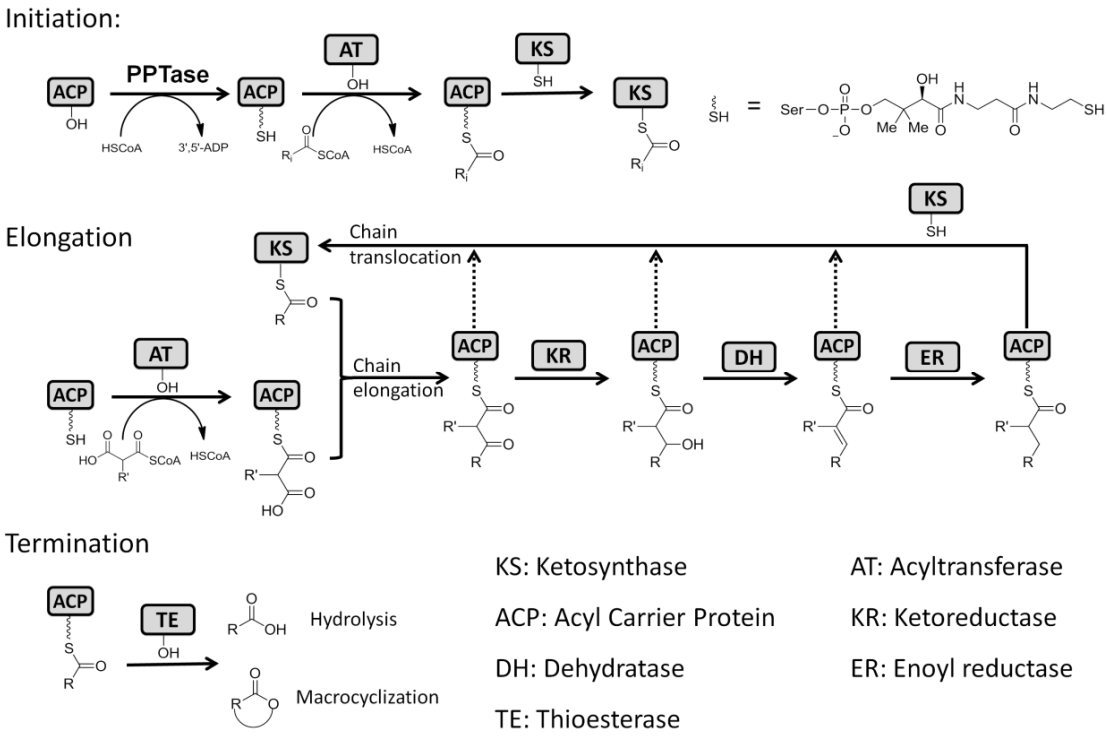



**Figure 6.** General chemical reactions catalyzed by PKSs

Based on the enzyme primary sequence, catalytic domain organization and product structure, PKSs can be classified into three distinct types [19, 21]: type I PKSs are large multidomain enzymes; type II PKSs consist of dissociated enzymes that function iteratively, and are typically associated with aromatic polyketides from bacteria; Types III PKSs contain single KS domains that catalyze condensations between various acyl-CoAs directly without the use of ACP domains. This classification is also supported by the phylogenetic analysis based on the amino acid sequence of KS domains [22, 23].

Among the three classes of PKSs, type I PKSs have the highest complexity in their protein structures and product diversities. These megasynthases are featured by having multiple catalytic domains juxtaposed on a single polypeptide chain. These covalently joined domains result in increased reaction rates and protection of labile intermediates. Depending on whether the polyketide backbone extension is performed using distinct sets of domains (each set contains

minimal PKS domains and is referred as a module), or repeatedly on a single set of domains, type I PKSs can be divided into two subcategories: modular (non-iterative) and iterative PKSs [19, 21, 24].

Modular type I PKSs, which are usually found in bacteria, synthesize polyketide in an assembly-line fashion: the growing polyketide chain is delivered from one module to the next and each catalytic domain in each module is typically used only once [25]. The number of the modules present in a PKS defines the number of chain extension steps, while the presence of KR, DH and ER domains in each module determine the degree of β-keto reduction after each elongation [26, 27]. Therefore, the modular nature of the PKSs and the associated structures of the polyketide products are colinear, which can be used to predict product from PKS sequence and vice versa [27]. The best studied modular PKS is the 6-deoxyerythronide B synthase (DEBS) involved in the synthesis of 6-deoxyerythronide B (6-dEB), which is the aglycon of the antibiotic erythromycin A. In DEBS, one loading module and six extending modules are required to process one propionate and six methylmalonyl-CoA units to produce 6-dEB (**Figure 7**). The colinear feature of type I PKSs has greatly enabled the rational manipulation of the domains and modules towards the generation of structurally altered products. In the case of DEBS, over fifty different macrolides have been synthesized through the engineering of AT and β-processing domains [28].

Iterative type I PKSs (IPKS) is also multidomain enzyme, the characteristic property of this type of PKSs are repeating usage their functional domains for the polyketide assembling. With a few exceptions, most fungal PKSs belong to iterative type I PKSs. Detailed discussion about this type of PKSs can be found in section 1.3.

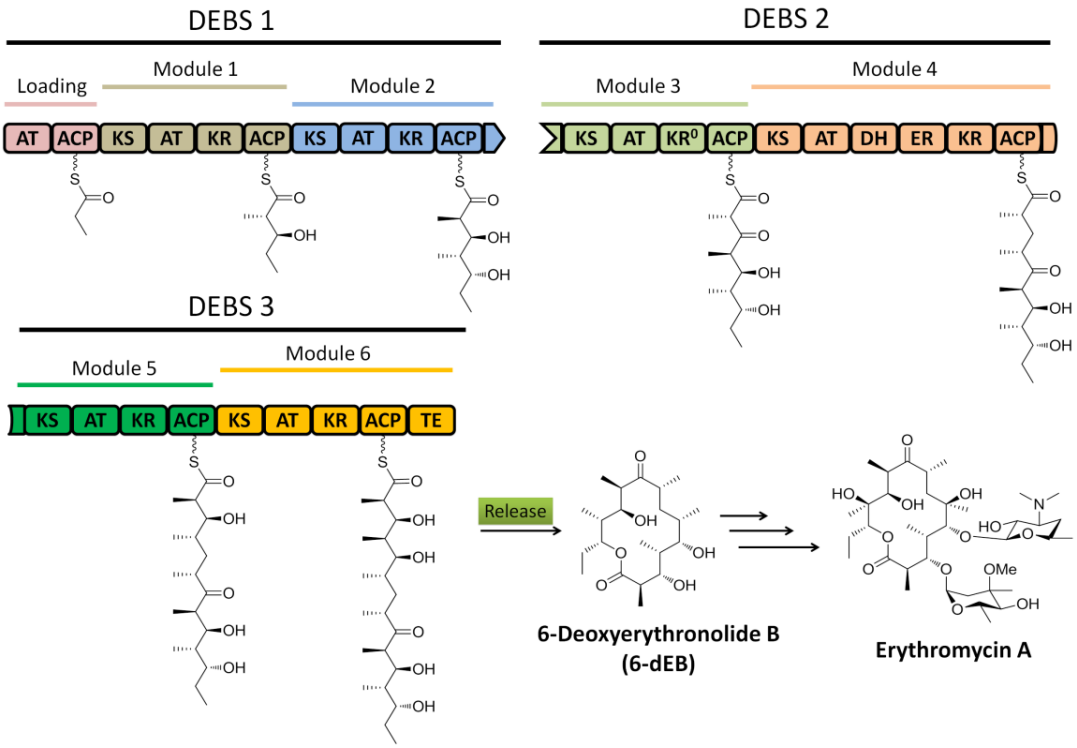



**Figure 7.** Biosynthesis of 6-deoxyerythronolide B on type-I modular PKS DEBSs (The domain with a "$^{0}$" notation indicate nonreactive domain which misses important catalytic/binding site)

### 1.1.4 Nonribosomal Peptides and Nonribosomal Peptides Biosynthesis

Nonribosomal peptides are peptides that are not synthesized by ribosomes. They represent a family of nitrogen containing secondary metabolites that are usually produced by microorganisms like bacteria and fungi. Nonribosomal peptides exhibit broad range of biological activities and many of them are used as important drugs. Important examples are: penicillin[29], the very first general antibiotics; vancomycin[30, 31], a "drug of last resort" antibiotic used for treatment of infections; enterobactin[32, 33], the high affinity siderophore that acquires iron and immunosuppressive agent cyclosporine A[34] (**Figure 8**).

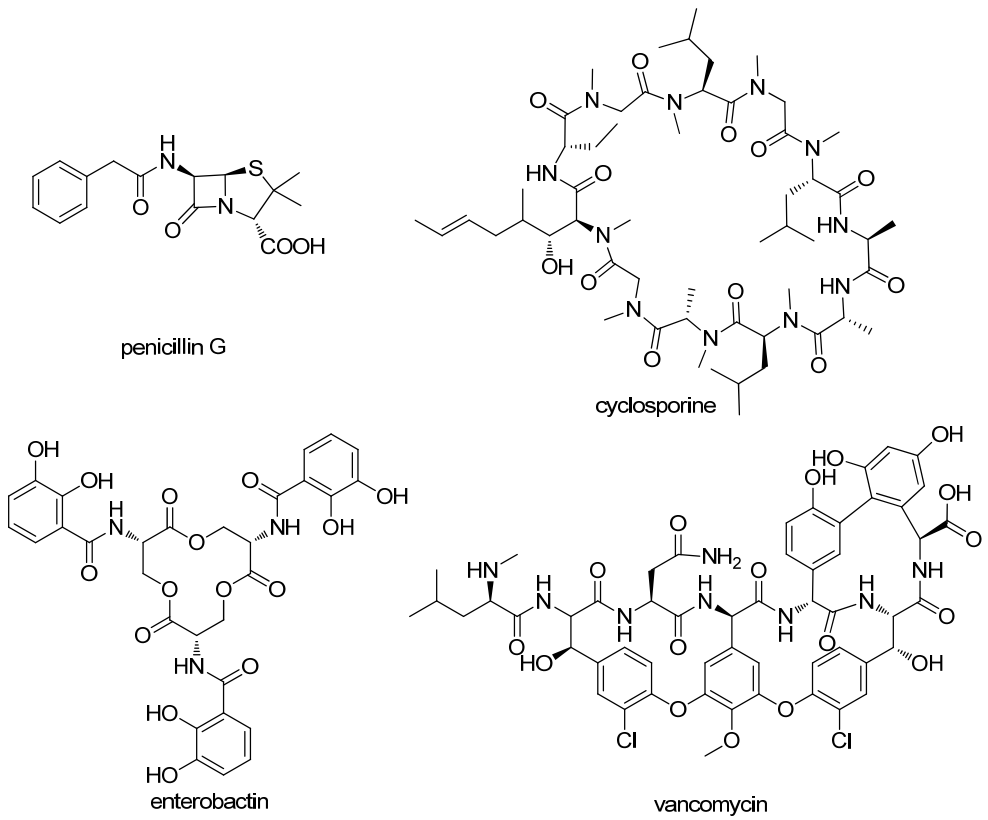



**Figure 8.** Examples of important nonribosomal peptides

Nonribosomal peptides are also biosynthesized in an assembly-line pattern, which is similar to polyketide biosynthesis from modular PKSs. The megasynthetases that assemble nonribosomal peptides are called nonribosomal peptide synthetases (NRPSs)[35]. Different to PKSs, NRPSs utilize amino acids (sometimes are non-proteinogenic amino acids) for the synthesis, which leads to a more diverse set of building blocks [36]. NPRSs are usually modular, means they are composed of a series of modules, in which unique domain organization can be found and one round of peptide elongation is catalyzed [37]. C-A-T is the typical module organization in a NRPS. "A" stands for adenylation domain that catalyzes conversion from selected amino acid to aminoacyl-O-AMP through consumption of ATP. "T" is the thiolation domain (sometimes it is also called Peptide Carrier Protein, PCP) that is similar to the ACP domain; it transfers the aminoacyl group from the activated aminoacyl-O-AMP to the thiol end

of its phosphapantethiol arm. "C" is the condensation domain that makes the new peptide bonds between the two intermediates on the two adjacent PCPs. The synthesis can then continue along the assembly line until the final product is released from the NPRS (**Figure 9)**. Tailoring domains can also been found in some of the modules, which increase the diversity of NRPs. For example, Epimerase (E) domain alters the stereochemistry of α-carbon of aminoacyl intermediate when it's in the module; N-methylation domain methylates the α-amino on the intermediate. Detailed reviews of NPRSs are available in numerous review papers [35, 38].

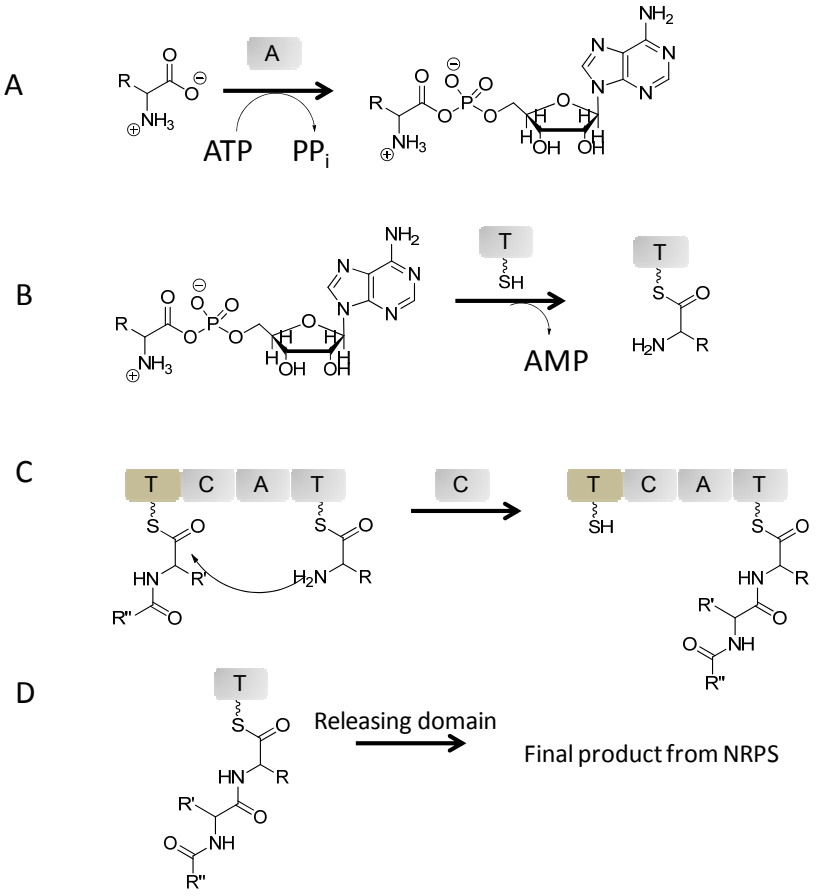



**Figure 9.** Catalytic steps during nonribosomal peptides biosynthesis (**A**) Adenylation; (**B**) Thiolation; (**C**) Condensation; (**D**) Product releasing

## 1.2 Strategies for Investigating Natural Product Biosynthesis

As the major sources for new drug discovery, natural products are still of crucial importance to human society. Gathering the natural products that are biosynthesized from certain host is still an

important route to producing natural product. Many of the natural products are structurally complex, containing multiple chiral centers and labile connectivities. They are difficult to be chemically synthesized. Fermentative approaches are therefore important tools in the production and development of these compounds for pharmaceutical, agricultural and related applications [39]. The key motivation for studying the natural product biosynthetic pathways is to expand our knowledge of how these complex molecules are made. By means of genetic and biochemical characterization, we gain knowledge of function and property of each single gene in the pathway. The society can be benefit by increasing the efficiency of production through metabolic engineering, which involves altering the metabolic profile or improving biosynthetic capabilities by introduction of rational changes in the pathway [40]. More importantly, people can also can produce new "non-natural" natural products and increase the diversity of natural products with proper engineering processes (protein engineering or combinatorial biosynthesis).

Understanding of the biosynthesis of natural products is also necessary for discovery of new natural products. Many natural product pathways have been evolved to be active only under appropriate conditions, such as presence or lack of specific nutrients or being challenged by other co-growing strains [41]. As a result, some natural product pathways are silent under the standard lab culture conditions [42]. These silent pathways can be activated under certain culture conditions or when provided with specific inducers [43]. Over-expression of the regulatory gene that is specific to the target biosynthetic gene cluster has also been proved to be an effective way to activate the pathway [44]. All this requires certain information about the biosynthesis pathways. This section will discuss tools and strategies used towards understanding natural product biosynthesis.

### 1.2.1 Bioinformatic Analysis of Biosynthetic Genes

The biosynthesis of a secondary metabolite requires involvement of many genes. In some organism, the whole set of genes responsible for the biosynthesis of the secondary metabolite are clustered on the genome in a region spanning 10–100 kb. For the organism that doesn't show gene clusters for secondary metabolites, the biosynthetic studies are hard to approach.

Genes for production of polyketides or non-ribosomal peptides are usually clustered: genes encoding tailoring enzymes, transcriptional regulators and transporters are often located adjacent to PKS or NRPS genes. As a result, biosynthesis studies of secondary metabolite have mainly focused on polyketides and non-ribosomal peptides. Nowadays, precise identification of functional domains or even assignment of substrate specificity became approachable based on bioinformatics analysis of biosynthetic genes, this would definitely aid in determining the biosynthetic mechanism. Bioinformatic analysis would also assist to rule out unrelated cluster, when one is looking for a candidate clusters for a known natural product.

### 1.2.2 Characterization of Biosynthetic Genes in Native Hosts

Once a candidate gene cluster for target natural product is identified, the next step is to perform gene deletion or inactivation experiments to the important genes in the cluster to confirm the responsible gene cluster for the target compound. This experiment is also used to interrupt the normal biosynthetic procedure and identify the function or timing of each enzyme in the pathway.

In order to perform the gene deletion or inactivation experiment, a wide variety of genetic technologies has been developed [45]. Usually, a typical fungal transformation method consists of two processes: preparation of the protoplast and introduction of the heterologous DNA into the protoplast. The former process is mainly achieved by digesting the fungal cell wall using

lytic enzymes to give protoplasts with permeable cell membranes, while the heterologous DNAs were introduced via polyethylene glycol (PEG)-mediated method in the presence of calcium chloride. The selection of transformants was accomplished by the use of selective markers. *Agrobacterium*-mediated transformation has been demonstrated to introduce DNA into both plant and fungi [46-48].

RNA-silencing is another strategy to inactivate a target gene [49]. Inactivation of the PKS-NRPS gene responsible for biosynthesis of chaetoglobosin A in *Penicillium expansum* was achieved through this strategy[50].

### 1.2.3 Characterization of Biosynthetic Genes in Heterologous Hosts

Due to the difficulty in transformation for some fungi, expression of target gene or genes in heterologous host that is easy to manipulate became a reasonable substitution.

*Escherichia coli* is one of the most powerful heterologous platforms with established genetics. Many PKSs and NRPSs have been successfully reconstituted in *E. coli*.[51-55]

Beside *E. coli*, the unicellular yeast *Saccharomyces cerevisiae* is another workhorse for heterologous expression of PKSs and NRPSs. Recently, a phosphopantetheinyl transferase gene *npgA* was integrated into a vacuolar protease deficient strain BJ5464 and the resulting strain[56] is reported as a very powerful expression system for fungal megasynthases. In both *E. coli* and *Saccharomyces cerevisiae*, the transformed DNA sequences have to be intronless.

Some well studied *Aspergillus* strains, like *Aspergillus oryzae* and *Aspergillus nidulans,* have been used to express other fungal gene[57] or even the whole gene cluster[58]. In these cases, the similarity of the promoter and intron splicing system between the native hosts the *Aspergillus*

hosts was able to allow direct expression of genes in the cluster through usage of gDNA information.

## 1.3 Fungal Polyketide Synthases and Polyketide Synthase-Nonribosomal Peptide Synthetase Hybrid

### 1.3.1 Fungal Polyketide Synthases

Most fungal PKSs belong to iterative type I PKSs (IPKSs) It is now well-accepted that the fungal IPKSs can be grouped into three major classes: the nonreducing (NR), the partial-reducing (PR) and the highly reducing (HR) PKSs [23, 24]. Compared to their bacterial type I PKS counters, the fungal IPKSs are evolutionarily more related to the mammalian FASs (mFASs) and share significant protein sequence similarity and domain architecture.

Iterative type I PKSs, such as the lovastatin nonaketide synthase LNKS (or LovB) from *Aspergillus terreus*, are mostly found in filamentous fungi, with some recent discoveries of bacterial examples, including the PKSs that synthesize the polyketide core of enediynes [59]. Each iterative type I PKS consists of a single set of catalytic domains, which can vary and define different extends of reductive modification in the product. While the linear organization and the iterative use of the domains resemble those of mammalian fatty acid synthase (mFAS), the programming rules dictating the permutative use of catalytic domains in iterative type I PKSs are significantly more complicated and mysterious. For example, different combinations of MT-KR-DH-ER usage are involved in the eight cycles of chain elongation catalyzed by LovB during production of dihydromonacolin L (**Figure 10B**). This means the generation of different β-oxidation states of different chain length intermediates by the same set of domains. The necessity of such complex programming rule is evident, because a precisely prepared triene at the

hexaketide stage decalin core is required for formation of decalin core through the intramolecular Diels-Alder cyclization (**Figure 10B**)[60-62]. In sharp contrast, LovF, a different type I PKS in the same lovastatin pathway which has nearly the same domain architecture as LovB (**Figure 10A**), performs only a single iteration of chain extension and tailoring to yield the 2-methylbutyrate side chain in lovastatin (**Figure 10B**) [61, 63]. The contrast in programming rules between LovB and LovF highlights the complexity and perhaps engineering potential of this subclass of type I PKSs.

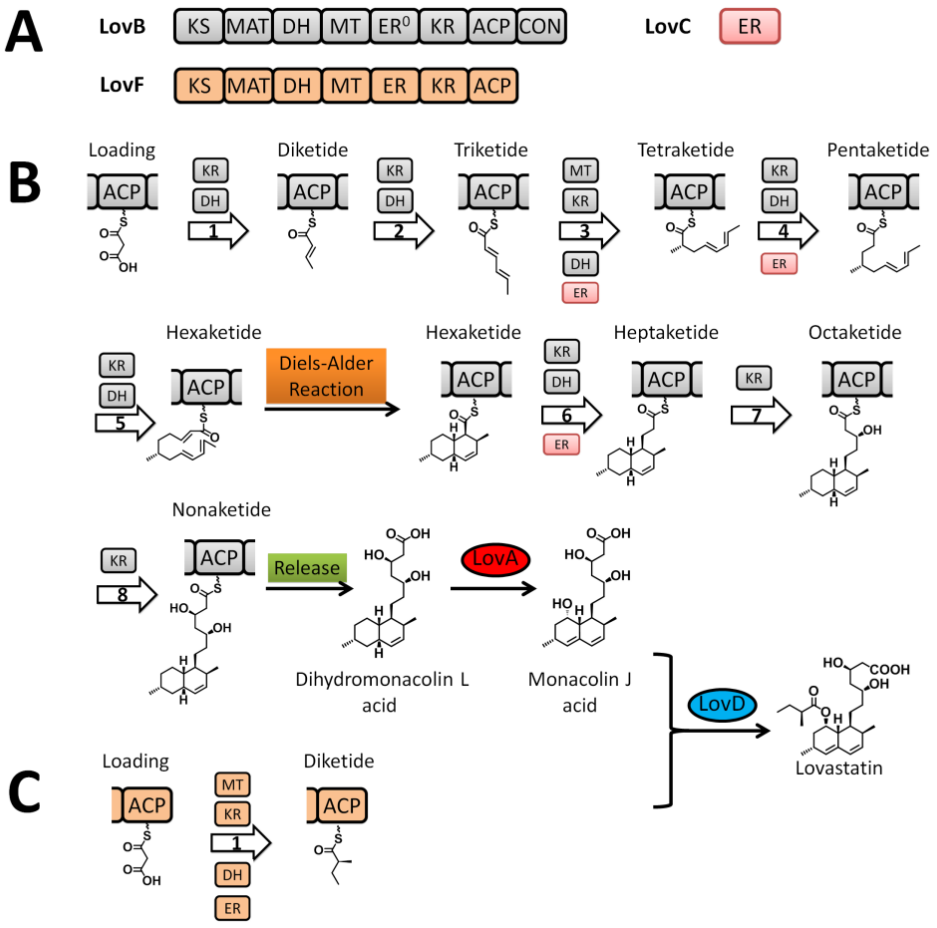



**Figure 10.** IPKSs in the biosynthesis of lovastatin (**A**) Domain organization of type I iterative PKSs LovB and LovF; (**B**) Biosynthesis of dihydromonacolin L on LovB in presence of LovC; (**C**) Biosynthesis of 2-methylbutyrate side chain on LovF

#### 1.3.1.1 Iterative Polyketide Synthases and mammalian FAS

The mammalian FASs are evolutionarily closely related to the type I iterative and modular PKSs [64]. In addition to the highly similar sequences of catalytic domains and linear positioning of the domains (**Figure 11A**), the parallel between mFAS and PKSs can be extent to the linker regions and noncatalytic domains ψ-MT and ψ-KR. In particular, the presence of the ψ-MT domain in mFAS and the catalytic MT domain in the iterative fungal PKSs indicate shared evolutionary ancestor. Despite the domain similarities, the mFAS and type I PKSs are functionally very different. In the case of modular PKSs, each module only performs one chain elongation and one defined set of β-processing, and the chain is passed unidirectionally down the assembly line to downstream modules. The different substrate specificities of the domains, different combinations of KR-DH-ER cassette, and the juxtaposition of different mFAS-like modules, all contribute to the diverse structures that can be generated from type I PKSs. So far, no X-ray structure of an intact PKS module has been solved, although the structures of nearly all the individual domains have been obtained. From a structural point of view, the fascinating aspects of module PKSs include the spatial arrangement of different modules with respect to each other, and how protein-protein interactions between the modules are facilitated through ACPs and linker regions to achieve the assembly-line like biosynthetic logic.

While the modular PKSs is certainly nature's example of "more is better" when it comes to generating chemical diversity, the iterative type I PKSs can be considered a more compact version in which modules of different functions are condensed into a single set of domains. Although iterative PKSs are functionally and structurally more closely related to mFAS, there are significant differences that make this subclass of PKS fascinating from a biochemical and structural perspective. Depending on the type of product, different tailoring domains are present

in the PKS. For nonreduced polyketides in which β-processing is absent, additional domains such as product template (PT) and Claisen-like cyclase (TE/CLC) are present to fix the regioselectivity of intramolecular cyclization reactions. For reduced polyketides, the more familiar KR, DH, ER and MT domains can be found in the PKS, although it is common that some domains lost their functions during evolution. The current structural understanding of the domains from iterative type I PKSs is not as extensive as in the module PKS, mostly due to the difficulties in isolating functional and soluble domains. In recent years, however, numerous structures, including that of AT, PT, TE/CLC have been solved, which have provided insights into how some of these distinct domains can fit into the overall dimeric structure of the PKSs. Another distinction between mFAS and iterative PKS is the recruitment of smaller, *in trans* enzymes for product tailoring and product release. This can be exemplified in the LovB example, in which a dissociated ER enzyme (LovC) is required for enoylreduction of three intermediates, and a dissociated TE or acyltranferase is required for product release. From a limited number of experiments, the interaction between the iterative PKS and these dissociated enzymes appear to be specific. Hence understanding this unique protein-protein interaction is important for fully decoding the programming rules of iterative PKSs.

Finally, the most intriguing differences between mFAS and iterative PKSs are undoubtedly the ability of the tailoring domains in the reducing iterative PKSs to only function on selected intermediates during the different cycles of chain extension, and how overall chain length is determined. Currently, nearly no information is known regarding how the selectivity is so precisely maintained. As a result of this cryptic selectivity, highly similar iterative PKSs can produce drastically different products. Recent genome sequences of different fungi species have revealed that each contains dozens of such PKSs with unknown products, yet unlike the modular

PKSs, no product can be predicted from these enzymes due to the utter lack of understanding of their sequence-function relationships. It is expected that crystal structures of the iterative PKS domains, including KS, KR, DH, ER, MT and ACP, will provide significant breakthroughs needed for decoding these machineries.

#### 1.3.1.2 Structures of Iterative Polyketide Synthases and mammalian FAS

While the crystal structure of an intact module (from modular PKS) or an entire iterative PKS is not yet available, significant structural information of individual PKS domains have been obtained in recent years. These include the X-ray structures of excised KS-AT [65, 66], KR [67-70], DH [71, 72], KR-ER [73] and TE [72, 74-76] domains from type I modular PKSs. Structures of ACPs from different families of PKSs have been obtained via high-resolution, liquid-state NMR analysis. These high resolution structures have greatly increased our understanding of the structural biology and enzymology of the PKSs components and domains. A number of publications have reviewed the various aspects of structural enzymology of FASs and PKSs [77-80].

Because of the similarity in the domain organization between mFAS and type I PKSs (**Figure 11A**), much of the finding in mFASs can be extended to that of type I PKSs, although significant differences are expected. The X-ray crystal structures of mFAS have provided the first look into the organization of the catalytic units in the multidomain type I PKS megasynthases [81, 82].

The head-to-head model of FASs was proposed through sets of mutation complementation experiments [83-85]. Inactivation mutations can be introduced to different catalytic domains without disrupting the folding of the domains, thereby allowing the combination of different mutant monomers to generate homodimers and heterodimers [83]. The ACP domains were

shown to crosstalk with the KS domain from either subunits to accomplish chain elongations in FAS [85], but could only interact with the DH domain from the same subunit [84, 86]. Heterodimers containing one subunit with either an inactive KS or AT domain, when combined with the other subunit inactivated in any one of the other five domains, including ACP, KR, DH, ER and TE domain, retained the ability to synthesize fatty acids[86]. Additionally, a null mutant FAS with all seven active site-compromised domains was capable of synthesizing fatty acid product in vitro after pairing with a wild-type FAS subunit [87]. In the head-to-head model, the ACP is able to interact with KS-AT from both subunits and can only interact with DH-ER-KR-TE domains within the same subunit [88]. The model was further supported by the cross-linking of KS and ACP via 1, 3-dibromopropanone (DBP), in which intra-subunit linked heterodimers were recovered [89]. The KS domains were proposed to be positioned at the interface of the two subunits [90]. This is motivated by biochemical evidences that KS domains from type I and type II FASs are always found as dimmers [90-92], and that the two subunits can be covalently crosslinked at low concentrations [90]. Therefore the KS domains must play a key role in maintaining the dimeric structures of FAS, and by inference, PKSs.

The 4.5 Å resolution X-ray structure of porcine FAS, which was subsequently refined to 3.2 Å resolution, showed the overall structure of mFAS adopts an X-shaped organization that is consistent with the head-to-head model [81, 82, 93]. The individual catalytic domains were mapped out by fitting the density diagram with previously-solved type II FAS components, including FabB (KS) [94], FabA (DH) [95], FabZ (DH) [96], and FabG (KR) [97] from *E. coli*; FabD (MAT, malonyl-CoA:ACP transacylase) from *S. coelicolor* [98]; and a zinc-free quinine reductase (ER) from *Thermus thermophilus* [99]. Starting from the N-terminus of the mFAS, the dimeric KS domains are located at the center of the lower body where each is connected to the

monomeric 'leg' MAT domain (**Figure 11B and Figure 11C**). Dimeric DH, ER and monomeric KR together form the upper body of the FAS, in which KR domains are the 'arms'. The 'chest' DH adopts a 'double-hot-dog' fold containing portion from core domain of FAS. KR domain plays a central role in connecting the 'shoulder' ER and DH domains as well as two newly identified non-enzymatic domains consisting of the rest of the core domain: 'pseudo-methyltransferase' (ψ-MT) and 'pseudo-ketoreductase' (ψ-KR) due to their structural similarity toward methyltransferase and truncated KR. In the lower body, the linker region between KS and MAT adopts a α2β3-fold. The upper body and lower body of the mFAS are barely contact with each other with a contact surface area of 230 Å. In contrast, the KS, DH and the ER domains contribute to ~5000 $Å^2$ dimer interface of the mFAS. The linker region between the MAT and the DH domains also contribute to dimer formation.

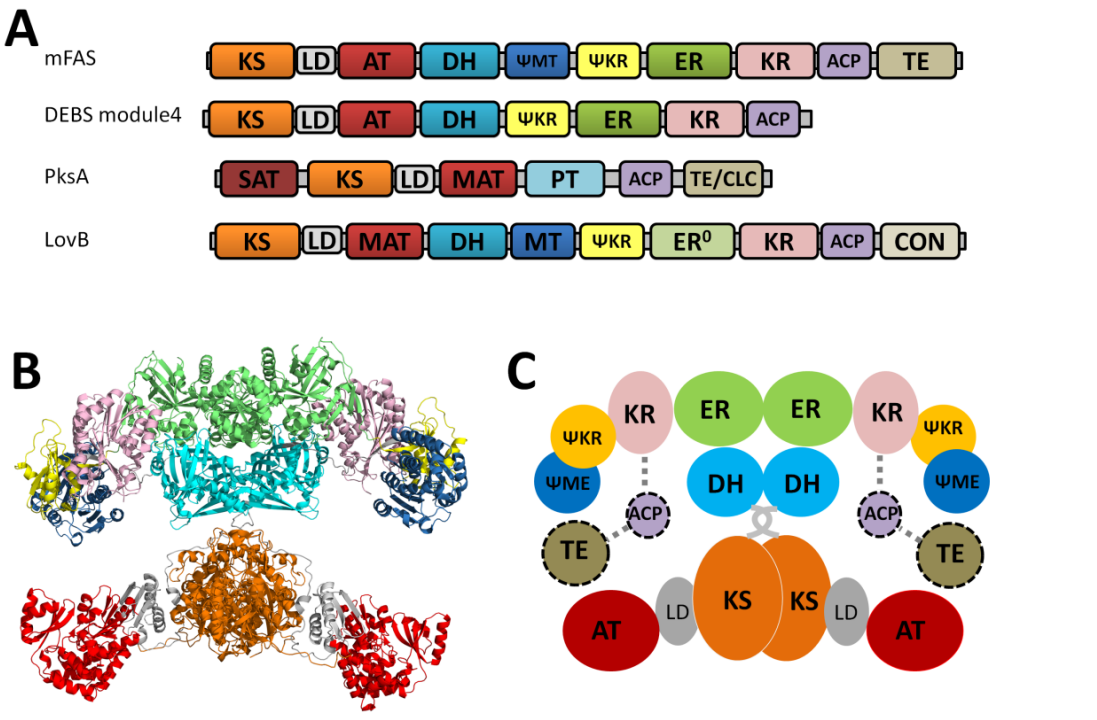



**Figure 11.** Structural overview of mammalian FAS (**A**) Comparison of the linear organizations of mFAS and type I PKSs. (**B**) The crystal structure of mFAS with a resolution of 3.2 Å. PDB code: 2VZ9 (**C**) Cartoon representation of the mFAS organization with predicted ACP-TE domains

Collectively, the dimeric mFAS clearly contains two reaction chambers, in which fatty acid assembly could take place independently. Interestingly, the two reaction chambers appeared to be asymmetrical: the distance between the active sites of MAT and KR domains is 15 Å shorter in the narrower chamber comparing to that in the wider chamber, suggesting that the fatty acid synthesis in each of reaction chamber might not proceed at the same pace and conformation changes. The anchor point of ACP was revealed at the center of the upper portion of the reaction chambers. The ACP and TE domains were unsolved in the reported mFAS crystal due to the inherent domain flexibility, but structures of stand-alone human TE [100] and rat FAS ACP [101] are available. The ψ-MT domain is located at vicinity of ACP anchor point, confining the movement of ACP within the region of the active sites of the catalytic domains. These structure features enable the *holo*-ACP to dock at different locations and shuttle the substrate to the specific active site from one chamber but not active sites from the other. Considering that the upper and lower bodies of the FAS is connected by the flexible MAT-DH linker and the limited protein-protein interactions between the two halves, it is highly possible that the lower body of one subunit could swivel more than 90 degree with respects to the upper body to interact with the ACP from the other subunit (**Figure 12A**). This hypothesis is verified by the electron microscopic analysis, in which an 80-100 degree rotation of the lower body with respects to the upper body was clearly present in reconstructed structures. The combined X-ray and EM structures explain the biochemical findings in mutant complementation and crosslinking experiments. This swiveling mechanism also injects FASs with additional catalytic route that contributes to approximately 20 % of fatty acid synthesis rate, as estimated from *in vitro* mutant complementation [88].

To examine the dynamics of the flexible FAS that permit the functional interactions of the ACP with all the catalytic domains, single-particle electron microscopy (EM) were utilized to continuously capture a series of configurations of FAS dimer from *Rattus norvegicus* [102]. The FAS mutant (Δ22) was constructed by the removal of the 22-residue ACP-TE linker to restrict the mobility of the TE domain. In the absence of substrates, both ACP and TE domains were visible in EM reconstructions. In the presence of substrates, asymmetrical arrangement of the upper β-processing domains was observed. Specifically, a swinging motion (up to 25 degrees) of the lower body relative to the upper body resulted in the opening of one of the reaction chambers and closing of the other (**Figure 12B**). While the open configuration of one of the chamber is required by the β-process domains to catalyze the reduction steps, the closure of the reaction chamber is proposed to facilitate the substrate loading and condensation reactions by the KS and AT domains. Additionally, since the TE domain is completely 'locked' outside upon the closure of the chamber, the chain releasing step likely happens in the open chamber.

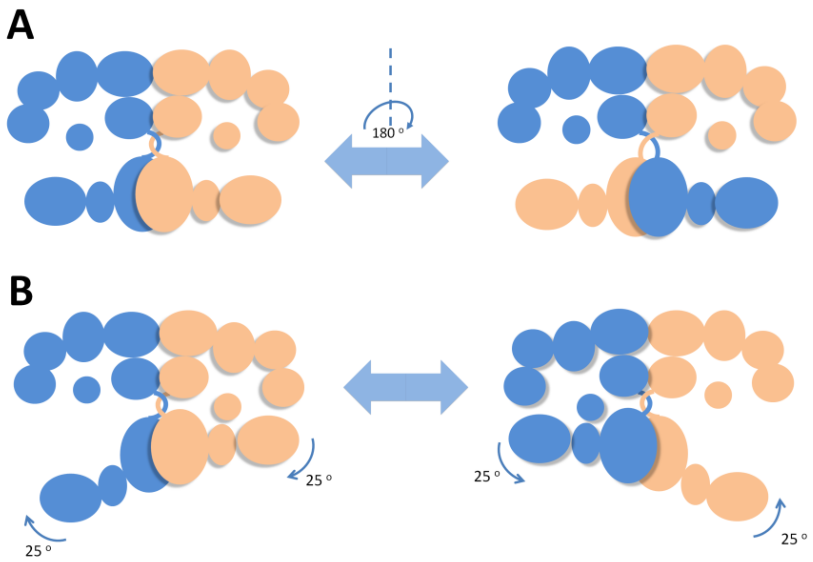



**Figure 12.** Schematic configurations adopted by mFAS to position catalytic domains into the proximity of ACP domain for domain-domain interactions (**A**) A full 180° swiveling of lower body relative to the upper body permits the ACP to crosstalk with KS-AT domains from both subunits. (**B**) A 25° rotation of the lower body against the upper body leads to the two states of the reaction chambers-open and closed.

#### 1.3.1.3 Non-reducing Polyketide Synthases (NR-PKSs)

NR-PKSs are IPKSs that don't have any reductive tailing domains. Besides the minimal PKS domains (KS, AT, and ACP), the NR-PKSs could have additional functional domains including the starter unit:ACP transacylase (SAT) domain, product template (PT) domain, and a releasing domain like a TE (an example is PksA in **Figure 11A**). Due to available structural information, NR-PKS has become the best understood type of IPKSs.

The SAT domain is located at the N-terminus of the protein. It belongs to the family of acyltransferases and exhibits similarity to MATs domains. The molecular basis for the incorporation of starter unit has been discovered by the Townsend group in the context of the PksA [103]: it transfers a starter acyl unit to the ACP, which then primes the KS domain for initial cycle of extension.

The poly-β-keto chains produced by NR-IPKS will be cyclized into mono- or polycyclic compounds. The three common first-ring cyclization patterns of fungal aromatic polyketides are C2−C7, C4−C9, and C6−C11. Recent structure study of PT domain in PksA explains the ability of NR-PKSs to stabilize the highly reactive nascent poly-β-keto chains and control the regioselectivity of the first ring cyclization [104]. Phylogenetic analysis [23] suggests that iterative nonreducing PKSs are evolved from ancient reducing PKSs. Combining this with the fact that the PksA PT was found to be dimeric both in the crystal and in solution, and has a double "hot-dog" fold like the DH domain, it is reasonable to infer that PTs are evolved from the DH and are located at the positions of DH domains in the three dimensional structure (**Figure 13C**). PTs catalyze regio-specific aldol cyclization of the unreduced poly-β-ketoacyl:ACP intermediate synthesized by the iterative PKS. The PT domain of PksA catalyzes stepwise first- and second-ring cyclizations to yield the ACP-tethered bicyclic intermediate [105]. A deep

pocket that can accommodate the phosphopantetheinyl tethered intermediate chain is found in PksA PT crystal. The pocket can be divided into three regions: a Ppant-binding region close to the surface, a cyclization chamber that can accommodate the bicyclic product core and a hexyl moiety binding site sitting in the innermost part of the pocket that is specific for the hexanoyl starter unit (**Figure 13B**). Notably, a 'wet' side of the cyclization chamber, which contains a network of crystallographic water molecules and nearby hydrophilic residues, is proposed to keep C9 in the substrate in the electrophilic keto tautomeric form for the intramolecular aldol cyclizations. The catalytic dyad, located at the surface of the cyclization chamber, is also composed of a histidine and an aspartate from two different "hot-dog" folds like in the DH domain. The catalytic histidine acts as base to deprotonate the methylene and generates the nucleophile. The aspartate is involved in polarizing the catalytic histidine, instead of helping removing a hydroxyl as in DH domain.

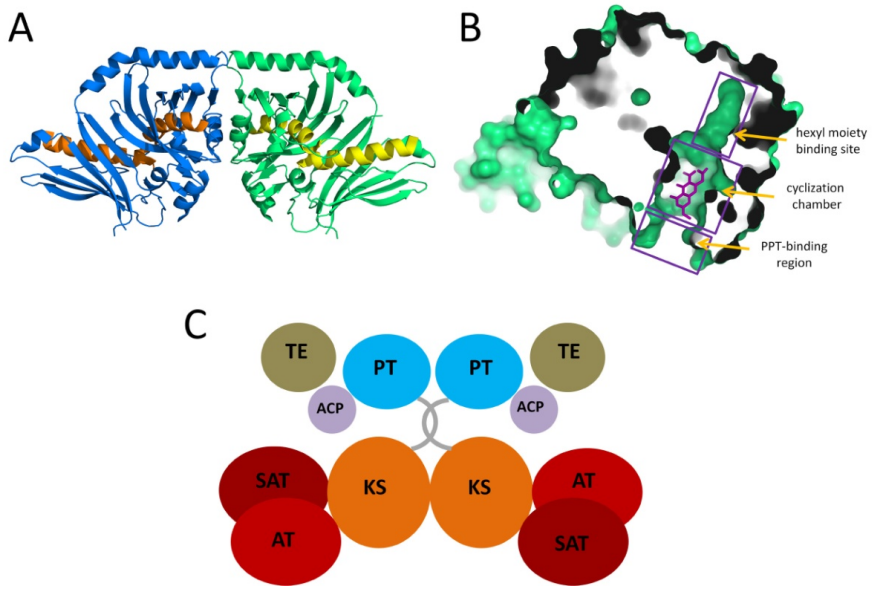



**Figure 13.** The crystal structure of the PT domain from PksA (**A**) Ribbon representation of the PT dimer showing the double hot dog (DHD) fold. PDB code: 3HRR. (**B**) Surface representation of the PT monomer substrate pocket. The product analogue HC8 is co-crystallized and is shown as sticks (**C**) Predicted domain organization of PksA.

PT domains of different cyclization regioselectivity have been functionally characterized and can be grouped phylogenetically [106]. It is expected that the active site geometry will dictate the positioning of the unreduced polyketide chain with respect to the histidine residue, which can then lead to different cyclization patterns.

The monomeric TE/CLC domain from the norsolorinic acid PKS, PksA, shares similar α/β-hydrolase fold with the typical modular PKS TEs [107]. This type of C-terminal TE domains of nonreducing iterative PKSs from fungi catalyze the regioselective Claisen condensation using thioester carbonyl as the electrophile and release of the aromatic polyketide, and hence are also referred to as TE/CLC domain [108]. In the absence of the TE/CLC domain, the polyketide intermediate can be released in as α-pyrone shunt products by nucleophilic attack from an enolate oxygen [105, 108, 109]. The substrate-binding region is similarly located between the core and the lid, and contains hydrophobic residues for binding of the aromatic substrate. A loop region in the lid is highly flexible, and is proposed to be the entrance to the binding pocket. Once delivered, the polyketide acyl intermediate is transesterified from the pPant arm of the ACP to the active site serine. After release of the freed ACP, the lid can close and the diketo portion of the substrate can rotate into a conformation that favors enolate formation and subsequent attack on the carbonyl of the oxyester. Therefore, ACP-TE/CLC protein-protein interaction and reorganization of the TE/CLC pocket are prerequisites for substrate delivery and product formation. In support to this proposed interaction, NMR titration experiments have shown interaction between analogous region of the NRPS TE domain and the peptidyl carrier protein partner [110, 111].

#### 1.3.1.4 Partially-reducing Polyketide Synthases (PR-PKSs)

Partially-reduing polyketide synthases lack ER domain in their linear domain prediction. The knowledge about PR-PKSs is far more limited than NR and HR PKSs, it has so far only been linked to production of 6-methylsalicylic acid (6-MSAS), which requires a single reduction of a specific keto group in a tetraketide backbone[112]. Phylogenetic analyses have shown that the 6-MSASs and other uncharacterized PR-PKSs are less related to the NR- and HR-PKSs but more related to the bacterial modular type I PKSs, suggesting an origin from actinobacteria by horizontal gene transfer.

#### 1.3.1.5 Highly-reducing Polyketide Synthases (HR-PKSs)

The HR-PKSs exhibit higher degree of complexity in their programming compared to that of the NR-PKSs. The HR-PKSs selectively catalyze β-keto reductions (ketoreduction, dehydration and enoylreduction) during each extension cycles. In addition, a C-methyl transferase (CMeT) domain could be found in many HR-PKSs. It is responsible for methylation of α-carbon on the newly extended polyketide chain using S-adenosylmethionine (SAM) as the methyl donor. All these bring high complexity in the programming. As discussed above, LNKS (LovB) requires 35 or more highly orchestrated sequence of catalytic steps to produce dihydromonacolin L.

Most HR-PKSs lack a built-in releasing domain, which is a unique feature among the HR-PKSs. This means it requires an additional releasing enzyme for product releasing from HR-PKSs, but the releasing enzyme is often not identified. For example, LovB isolated from heterologous host was reported to being able to assemble dihydromonacolin L but not able to turn over efficiently. This incomplete understanding of HR-PKS releasing mechanism has also complicated efforts in the heterologous reconstitution of HR-PKS systems. Recent advances in

multigene expression tools and *in vitro* reconstitution methods have improved our understanding of the HR-PKS systems.

### 1.3.2 Fungal Polyketide Synthase-Nonribosomal Peptide Synthetase Hybrids

Some HR-PKSs are also fused with a complete C-terminal nonribosomal peptide synthetase (NRPS) module, forming a PKS-NRPS hybrid megasynthetase that can incorporate an amino acid into the polyketide product. This increases the size of HR-PKSs and also complicates the programming on the HR-PKS. The NRPS module in PKS-NRPS hybrid usually shares a domain organization as (C−A−T−R) including the condensation (C), adenylation (A), thiolation or peptidyl carrier protein (T), and the reduction/Dieckmann cyclase (R) domains.

## 2. Results and Discussion

### 2.1 Mechanisms Studies on Polyketide Synthases in Lovastatin Biosynthesis

### 2.1.1 Introduction

Lovastatin, a polyketide produced by the fungus *Aspergillus terreus*[113], is one of the most important natural products discovered to date. Both lovastatin and the semi-synthetic derivative simvastatin are widely prescribed hypercholesterolemia drugs because of their inhibitory activities towards *3S*-hydroxy-3-methyl-glutaryl-CoA reductase (HMGR).[114] The fermentative production of lovastatin has therefore been one of highest grossing processes involving a natural product.

The biosynthetic gene cluster for lovastatin acid (**1**) in *A. terreus* was first identified in a pioneering work by Kennedy et al.[61] Through genetic[61, 115] and biochemical[62, 116, 117] characterizations, it is now known that two highly reducing iterative type I polyketide synthases

(HR-PKSs) play central roles in biosynthesis of **1**. The lovastatin nonaketide synthase (LNKS or LovB), together with the dissociated enoylreductase LovC, are responsible for the programmed assembly of dihydromonacolin L acid (**2**).[61] **2** is then modified by the cytochrome P450 monooxygenase LovA via hydroxylation and dehydration to afford monacolin L acid (**4**). LovA also catalyzes further hydroxylation at C-8 to produce **3**.[117] Lastly, the α-methylbutyryl side chain is synthesized by lovastatin diketide synthase (LDKS or LovF) and is transferred to the C-8 hydroxyl group of **3** by the acyl transferase LovD to yield **1** (**Figure 14**).[116] It is the function of LovD that was exploited for the enzymatic synthesis of simvastatin.[118]

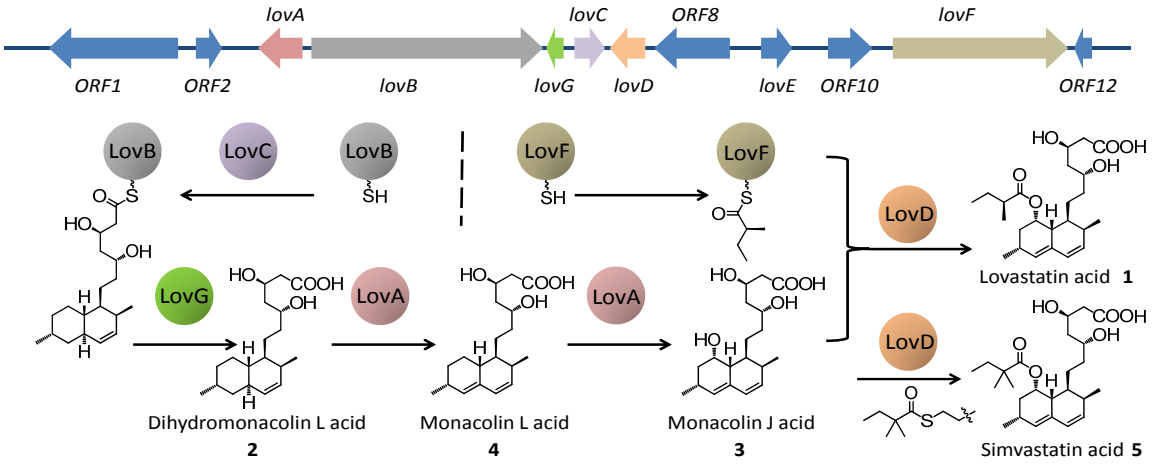



**Figure 14.** Lovastatin biosynthetic gene cluster and pathway

Both LovB and LovF belong to a large family of megasynthases, highly reducing iterative polyketide synthases (HR-IPKSs) from filamentous fungi.[119, 120] This family of megasynthases produce important natural products such as lovastatin **1**,[61] squalestatin,[121] and fumonisin.[122] Unlike other families of PKSs, the biosynthetic programming rules of fungal HR-IPKSs remain largely unexplored. One particularly unique feature of HR-IPKSs is the lack of a built-in offloading domain that facilitates the release of completed products. In sharp contrast to bacterial type I or fungal nonreducing PKSs, in which a dedicated thioesterase (TE) domain is appended at the end of the megasynthase and catalyzes the release of polyketides via

macrocyclization[74, 123, 124], Claisen cyclization[105, 125, 126], or hydrolysis[64, 127, 128]. Understanding the mechanisms of product release is therefore an important goal in demystifying the functions of HR-IPKSs.

### 2.1.2 Acyltransferase Mediated Polyketide Release from Lovastatin Diketide Synthase (LDKS or LovF)

A version of this section was published as: Xie, X.; Meehan, M. J.; Xu, W.; Dorrestein, P. C.; Tang, Y. *J. Am. Chem. Soc.* **2009**, *131*, 8388.

The diketide synthase LovF consists of seven linearly arranged domains and has been proposed to biosynthesize the α-S-methylbutyrate side chain of **1** (**Figure 14**). Transfer of the diketide side chain from LovF to monacolin J (**3)** is hypothesized to be catalyzed by a dissociated acyltransferase LovD.[61] The proposed acyltransferase-mediated product release mechanism is unprecedented among known PKSs. We demonstrate LovF-LovD protein−protein interactions play an important role in LovF product release and lovastatin biosynthesis.

#### 2.1.2.1 Materials and Methods

**Materials**

Monacolin J acid was prepared as described previously. [129] All other chemicals were purchased from standard sources. *E. coli* XL1-Blue and *E. coli* TOPO10 were used for construction of recombinant DNA molecules. *E. coli* BL21 (DE3) was used for protein expression of LovF ACP domain. *Saccharomyces cerevisiae* BJ5464-NpgA, which is constructed from *S. cerevisiae* BJ5464 (*MATα ura3-52 trp1 leu2Δ1 his3Δ200 pep4::HIS3 prb1Δ1.6R can1 GAL*) and contains a chromosomal copy of the phosphopantetheinyl (Ppant) transferase *npgA*.[56] was used for protein expression of the entire LovF.

**Expression and purification of holo-LovF from *S. cerevisiae***

For 1 L of liquid culture, the *S. cerevisiae* cells were grown at 28 °C in YPD media with 1% dextrose for 72 hours. The cells were harvested by centrifugation (3500 rpm, 15 minutes, 4 °C), resuspended in 20 mL lysis buffer (50 mM PBS, pH 8.0, 0.5M NaCl, 10 mM imidazole) and lysed by sonication. Cellular debris was removed by centrifugation (15,000 g, 1 hour, 4 °C). Ni-NTA agarose resin was added to the supernatant (2 mL/L of culture) and the solution was gently mixed at 4°C for overnight. The protein-resin mixture was loaded into a gravity flow column and LovF was purified with increasing concentration of imidazole in Buffer A (50 mM Tris-HCl, pH 7.9, 2 mM EDTA, 2 mM DTT, 500 mM NaCl). The collected protein solution was buffer exchanged to Tris buffer (50 mM, pH 7.9, 2 mM EDTA, 2 mM DTT) to remove NaCl before anion exchange chromatography. A HiTrap CM FF anion exchange column (GE Scientific) was used in conjunction with Buffer A and Buffer B (50 mM Tris-HCl, pH 7.9, 2 mM EDTA, 2 mM DTT, 1 M NaCl) to purify the LovF protein. LovF was eluted at approximately 400 mM NaCl. Purified proteins were concentrated and buffered exchanged into Buffer A with Centriprep filter devices (Amicon Inc.) and stored in 10% glycerol solution at -80°C. Protein concentrations were determined by Bradford assay using BSA as a standard. The final yield of C-terminal hexahistidine tagged LovF is ~3.5 mg/L.

**Purification protein from *E. coli* BL21 (DE3)**

The ACP domain of LovF (last 459 bp of LovF c-DNA) was amplified by PCR and inserted into the expression vector pET28a to give pXK64. pXK64 was transformed to *E. coli* BL21 (DE3) by electroporation. The resulting transformant was grown in LB medium at 37°C to OD 600 of 0.5, at which time 1 mM IPTG was added to the culture and protein expression was

performed at 16°C C overnight. The purification steps are similar to those used to purify LovF. A final yield of ~1 mg/L LovF ACP was obtained. LovD protein in this study was also purified in a similar way as reported.

**LovF activity assays**

50 μM LovF was incubated with 50 nM LovD, 100 μM monacolin J, 2 mM malonyl-CoA, 2 mM S-(5'-adenosyl)-L-methionine chloride (SAM), 2 mM NADPH in 100 mM PBS, pH 7.4. At 1, 2, 3, and 4 hours time points, 10 μL solution was removed to a microcentrifuge tube, extracted with 200 μL 1% acetic acid (AcOH)/99% ethyl acetate (EA), dried by speedvac, and re-dissolved in 0.1 M NaOH to yield the ring open form lovastatin acid. The solution was then analyzed by LC-MS. LC-MS was conducted with a Shimadzu 2010 EV Liquid Chromatography Mass spectrometer by using both positive and negative electrospray ionization and a Phenomenex Luna 5μ 2.0 x 100 mm C18 reverse-phase column. Samples were separated on a linear gradient of 5 to 95% $CH_3CN$ (30 mins) and 95% $CH_3CN$ (15 mins) at a flow rate of 0.1 ml/min. Both $H_2O$ and $CH_3CN$ were supplemented with 0.05% (vol/vol) formic acid.

**Synthesis of α-S-methylbutyryl-CoA (MB-CoA)**

41 mg coenzyme A trilithium salt (0.05 mmol), 42.5 mg PyBOP (0.08 mmol), 20 mg $KCO_3$ (0.2 mmol), and 13.3 μL α-S-methylbutyric acid were dissolved in 4 ml 50% tetrahydrofuran/50% $H_2O$ and stirred at room temperature for 1 hour. The reaction mixture was lyophilized to remove all the solvent and re-dissolved in 200 μL DI water. The solution was then purified with a preparative reverse-phase HPLC column (SunFire Prep C18 OBD column, 5μ, 19 x 50 mm). A 5–70% CH3CN in water (0.1% TFA) gradient (30 mins) with a flow rate of 2 ml/min was used. MB-CoA eluted with a retention time of 28 min. The fractions were combined, lyophilized, and dissolved in 50 mM PBS buffer (pH 6.5). The final yield was ~70%. The identity of MB-CoA

was confirmed by MALDI-TOF (Applied Biosystems Voyager-DE-STR MALDI-TOF) to give m/z of 852.18 (calculated 852.18).

**Kinetic study of intact LovF protein towards lovastatin biosynthesis**

150 μM, 100 μM, 50 μM, 25 μM, 10 μM, 5 μM LovF were each incubated with 2 mM malonyl-CoA, 2 mM SAM, 2 mM NADPH in 100 mM PBS, pH 7.4 for one hour for the preloading of α-S-methylbutyrate to the ACP domain of LovF, 100 μM MJ acid and 50 nM LovD were then added to the solution to start the reaction. At each time point, 10 μL aliquots of each reaction sample were removed, quenched by addition of EA/1% AcOH. The organic phase was dried by speedvac, and re-dissolved in 0.1M NaOH for LC-MS analysis. Initial rates of the acyl transfer reactions were measured by fitting the linear portions of the time course analysis. The experiments were repeated three times.

**Kinetic study of MB-CoA towards lovastatin biosynthesis**

4 mM, 2 mM, 1 mM, and 0.5 mM MB-CoA were incubated with 100 μM MJ acid and 10 μM LovD in 50 mM HEPES buffer (pH 7.9). At each time point, 50 μL solutions of all samples were removed, quenched by addition of EA/1% AcOH. The organic phase was dried by speedvac, and re-dissolved in 0.1M NaOH for LC-MS analysis. Initial rates of the acylation reactions were measured by fitting the linear portions of the time course analysis. The experiments were repeated three times.

**Phosphopantetheinylation of stand-alone ACP domain of LovF**

The stand-alone *apo* LovF ACP was converted to α-S-methylbutyryl-LovF_ACP (MB-ACP) by using Sfp and MB-CoA. 60 μM LovF ACP was incubated with 10 mM $MgCl_2$, 250 μM MB-CoA, 5 μM Sfp in 100 mM MES, pH 6.5. The reaction progress was monitored by LC/MS (Applied Biosystems Q-STAR Elite Quad-TOF Hybrid LC/MS/MS System) with a Jupiter 5μ

C4 300A 50 x 2.0 mm reverse phase Column. A step gradient from 5% $CH_3CN$ (15 mins) to 75% $CH_3CN$ (15 mins) at a flow rate of 0.2 ml/min was used. Both $H_2O$ and $CH_3CN$ were supplemented with 0.1% (vol/vol) formic acid. MagTran v1.03 (Amgen) was used for deconvolution. After deconvolution, the mass of *apo* LovF ACP and MB-ACP were determined to be 18953 (calculated 18952) and 19376 (calculated 19376) respectively. The reaction mixture was then buffer exchanged to 50 mM PBS buffer (pH 7.9).

**Kinetic study of LovF ACP towards lovastatin biosynthesis**

The MB-ACP derived above was used directly for kinetic study. 0.5 μM, 1.0 μM, 2.5 μM, 5.0 μM, 10 μM MB-ACP was incubated with 100 μM MJ acid and 100 nM LovD in 50 mM HEPES buffer (pH 7.9). At different time points, 20 μL aliquots of each sample were removed, quenched by addition of EA/1% AcOH, and subjected to LC-MS analysis. Initial rates of the acylation reactions were measured by fitting the linear portions of the time course analysis. The experiments were repeated three times.

### 2.1.2.2 Results and Discussions

***Holo*-LovF megasynthase can be expressed from *Saccharomyces cerevisiae* BJ5464-NpgA**

To obtain soluble LovF megasynthase for *in vitro* studies, the *lovF* gene (7.6 kb) from *Aspergillus terreus* was inserted into a yeast 2μ expression vector under the control of the *ADH2* promoter. The resulting vector was transformed into *Saccharomyces cerevisiae* BJ5464-NpgA, which contains a chromosomal copy of the phosphopantetheinyl (Ppant) transferase NpgA. Recombinant LovF (277 kDa) was purified to homogeneity by affinity and anion exchange chromatography steps (**Figure 15**). The ACP domain of LovF was determined to be phosphopantetheinylated after tryptic digestion, offline HPLC purification, and Ppant ejection analysis by FTMS [130].

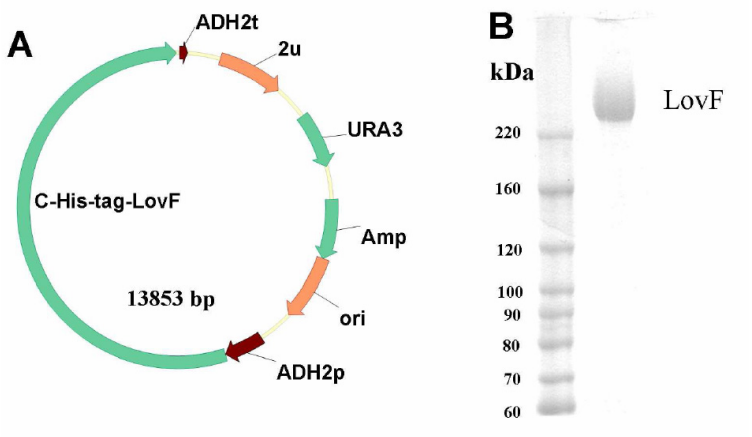



**Figure 15.** Expression and purification of LovF from *Saccharomyces cerevisiae* BJ5464-NpgA (**A**) Map of the LovF expression vector; (**B**) SDS-PAGE of the purified LovF

**LovF is functional in its purified form but not able to offload product**

The purified LovF was incubated with malonyl-CoA, SAM, and NADPH to test its activity. No products were observed directly by LC-MS analysis of the organic extract after reaction. Careful Ppant ejection analysis on the digested LovF after reaction shows the presence of intermediates on LovF ACP domain. When malonyl-CoA was supplied at substoichiometric concentrations of LovF, a new species with a molecular weight of 84.0576 Da attached to the Ppant arm of the ACP was detected using FTMS (**Figure 16**). The Ppant ejection ion that resulted from source fragmentation of the fraction containing the ACP active site peptide was determined to have a mass of 345.1843 Da. This mass is consistent with that of Ppant covalently bound to the expected diketide product α-methylbutyrate. Furthermore, an additional round of fragmentation confirmed that this ion was indeed the α-methylbutyrate loaded phosphopantetheinyl ejected ion [131]. No released diketide products were detected in the reaction mixture, confirming the inability of LovF alone to offload the product.

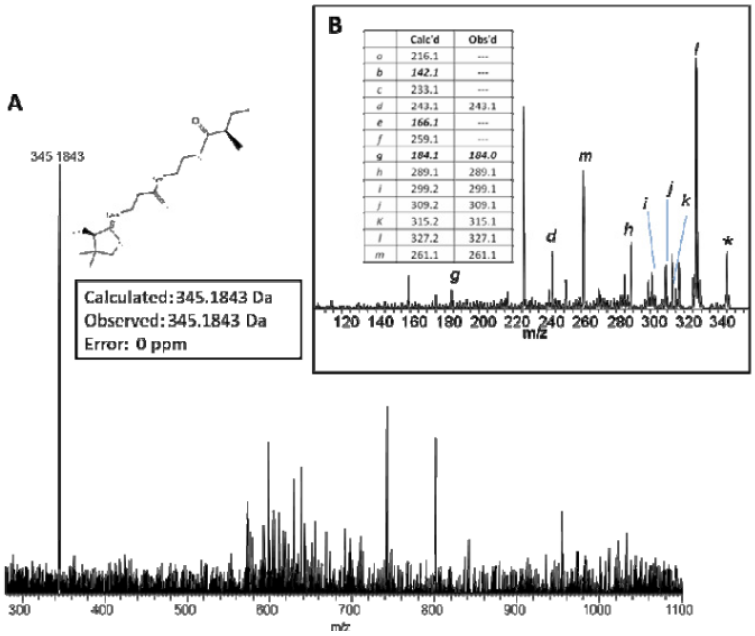



**Figure 16.** Source fragmentation and FT-ICR-MS detection of α-methylbutyrate-bound Ppant ejection

**LovD transfers α-methylbutyl from LovF to monacolin J acid *in vitro***

Because LovF lacks the ability to offload products, experiments were set up to verify that LovD offloads α-methylbutyl from LovF and transfers it to the acyl acceptor monacolin J acid **3**. We add LovD and monacolin J to the previous reaction and test the organic extract. LovD was supplied at low concentrations to minimize the reverse hydrolysis reaction.[129] LC-MS analysis of the organic Extract revealed the formation of a single product and the accompanying decrease of **3** (**Figure 17**). The product eluted with the same retention time and has the same mass fragmentation pattern as the standard **1**. In contrast, reactions either without LovD or with an S76A active site mutant of LovD did not afford any **1**, demonstrating the essential catalytic role of LovD in facilitating the acyl transfer. Importantly, synthesis of **1** confirmed that the entire range of LovF functions can be reconstituted *in vitro*, including activities of all six catalytic domains, as well as the proposed interactions with LovD.

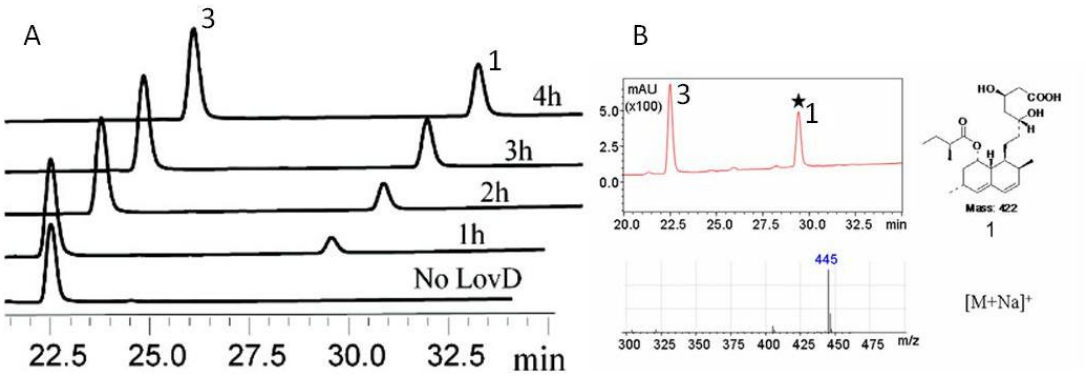



**Figure 17.** *In vitro* reconstitution of LovF and LovD activity

**Kinetics study shows that LovD strongly interacts with LovF**

To quantify the rate of synthesis of **1** using α-*S*-methylbutyryl-LovF (MB-LovF) as an acyl donor, different concentrations of LovF were first treated with malonyl-CoA and all the cofactors for 1 h to preload the ACP domain with α-S-methylbutyrate. LovD and **3** were then added, and the initial velocities of the reactions were measured by HPLC. The reaction displayed Michaelis−Menten kinetics, and the kinetic parameters were determined as shown in **Table 1**. The catalytic efficiency (kcat/Km) of LovD toward MB-LovF was 1.29 $min^{-1}\mu M^{-1}$. We also measured the reaction kinetics of LovD toward α-S-methylbutyryl-CoA (MB-CoA,) and α-S-methylbutyryl-S-N-acetylcysteamine (MB-SNAC). Compared to that of MB-LovF, LovD displayed a 20 000-fold and 800 000-fold attenuation in kinetic efficiency toward MB-CoA and MB-SNAC, respectively. This means that both $K_m$ and $k_{cat}$ were significantly improved when the methylbutyrate side chain is attached to LovF, which demonstrates that protein−protein interactions between LovF and LovD play a key role in facilitating rapid offloading of the diketide substrate from LovF to LovD and ensure efficient biosynthesis of **1**. This is analogous to the intermodular transfer of polyketide intermediates in bacterial type I PKSs, in which protein−protein interactions significantly lowers the $K_m$ of the acyl transfer reaction.[132] The

lower $K_m$ is critical for the ping-pong bibi reaction, as LovD can be inhibited by **3** when the acyl donor substrate binds with a high $K_m$, such as in the case of MB-CoA and MB-SNAC.

**Table 1.** Kinetic Parameters of LovD towards MB-SNAC, MB-CoA, MB-ACP, and MB-LovF

| | $k_{cat}$ ($min^{-1}$) | $K_m$ (mM) | $k_{cat}/K_m$ ($min^{-1}\mu M^{-1}$) |
|---|---|---|---|
| MB-LovF | 48.4 ± 3.6 | 0.039 ± 0.008 | 1.29 |
| MB-ACP | ND | ND | 0.32 |
| MB-CoA | 0.17 ± 0.01 | 2.5 ± 0.31 | $6.9 \times 10^{-5}$ |
| MB-SNAC | 0.024 ± 0.001 | 9.2 ± 0.9 | $1.6 \times 10^{-6}$ |

**LovD interacts with LovF through the ACP domain**

ACP domain of LovF was cloned and expressed the from *Escherichia coli* expression strain BL21 (DE3) and the *apo* form of the stand-alone ACP was confirmed by ESI to be 100% *apo* (**Figure 18**). The broad spectrum Ppant transferase Sfp was used to prepare α-S-methylbutyryl-LovF-ACP (MB-ACP) using *apo*-LovF ACP and MB-CoA.[133] The conversion from *apo* to *holo* was monitored by mass spectrometry. We observed that MB-ACP made from *apo*-ACP was rapidly converted into *holo*-ACP; this means that MB-ACP was prone to spontaneous hydrolysis of the thioester after preparation. In the full length LovF, we did not see the same phenomena. This may because of the lack of protection to the bound substrate in a stand-alone ACP. As a result, the individual $k_{cat}$ and $K_m$ values were not determined and only the catalytic efficiency was measure at low MB-ACP concentrations. As shown in **Table 1**, LovD displayed comparable efficiency toward MB-ACP when compared to that of MB-LovF, indicating the protein−protein interaction between the LovF ACP domain and LovD is sufficient to afford high catalytic efficiency. There remains the possibility that LovD makes additional contacts with the LovF megasynthase, which may account for the 4-fold difference in kcat/Km between MB-LovF and MB-ACP.

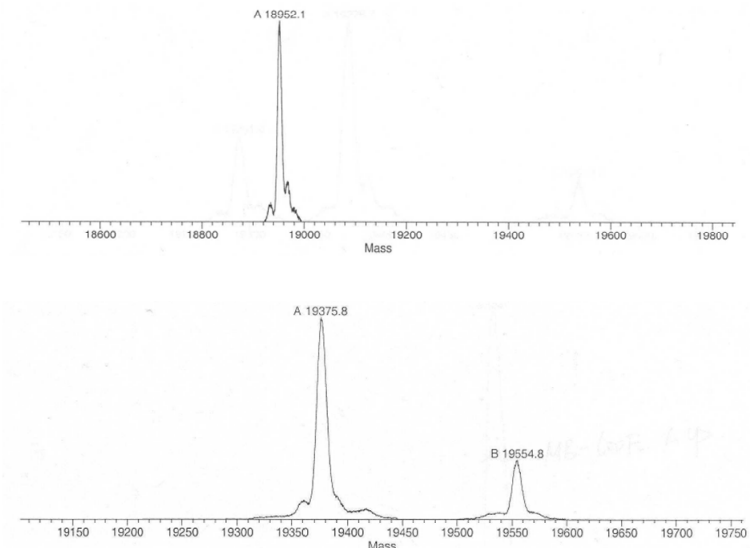



**Figure 18.** Characterization of *apo* LovF ACP and MB-ACP by ESI-TOF

**LovD is highly specific towards LovF ACP.**

When the methylbutyryl modified versions of heterologous ACPs were presented to LovD and **3**, including OxyC (type II PKS), DEBS ACP3 (bacterial type I), and $ACP_p$ (*E. coli* FAS), we were unable to detect any trace of **1** in the reaction mixture. This suggests that LovD interact specifically to LovF ACP.

**Discussion**

Our experiment showed that LovF was able to communicate with LovD and transfer the assembled product onto LovD then to monacolin J. This involves central interaction in PKS based synthesis: the interaction between ACP and other catalytic domains.

The acyl carrier protein (ACP) is a small (~80-100 residues) non-catalytic domain that tethers the growing polyketide and building block on the Ppant arm. The posttranslational modification that transforms *apo* ACP domains into its *holo* form is catalyzed by a phosphopantetheinyl transferase (PPTase), such as *sfp* from *Bacillus subtillis* [134]. Using the

20Å Ppant arm, the ACP delivers substrates to the different catalytic domains during PKS catalysis. To facilitate the assortment of protein-protein interactions, ACP binds to the specific enzyme partners through different structural features on the small protein. Therefore, obtaining the structures of ACPs and decoding the residues that are responsible for the different (and orthogonal) interactions is the key in understanding PKS function. This is especially important in type I PKSs, in which fine-tuned KS:ACP interactions play key roles in conferring the unidirectional substrate transfer of modular PKSs [135, 136], and controlling the permutation of tailoring domain functions in iterative PKSs.

Due to the inherent flexibility of the ACP domains and the adjacent linker regions, ACPs are typically absent in the crystallographic snapshots of intact megasynthases. However, the solution structures of ACPs have been obtained via solution-phase NMR spectroscopy. To date, several solution structures of ACPs have been characterized, including: DEBS ACP2 [137], *holo* and *apo* DEBS ACP6 [138], calicheamicin ACP [139], curacin $ACP_I$ [140], *apo*- [141], *holo*- [142], derivatized forms [143, 144] of actinorhodin ACP, frenolicin *holo*-ACP [145], oxytetracycline *apo*-ACP [146] and norsolorinic PKS *holo*-ACP [147].

Despite the overall low sequence similarity among ACPs from different PKSs, nearly all ACPs exhibit a similar right-handed twisted bundle consisting of three α-helices (helix I, helix II and helix IV) connected by two loops. In some structures an additional short helix (helix III) is found in the second loop depends on its helix-loop equilibrium conformations [148, 149]. The longer helix-I is positioned anti-parallel to helix II and helix IV, while helix I and helix IV are arranged at a small angle relative to each other. The overall helical bundle is stabilized via inter-helical hydrophobic interactions. The well-conserved “DSL” motif containing the active serine is located at the N-terminus of helix II, which is identified as the universal protein-protein

'recognition helix' [78, 150, 151]. This recognition helix is featured by a large number of negatively charged residues and is demonstrated by mutagenesis studies to play a crucial role for ACPs to communicate with other catalytic domains [152].

ACP is involved in all of the catalytic steps during each cycle for polyketide extension, including extender-unit loading by AT, chain elongation and translocation by the KS, reductive modifications by KR, DH, ER and/or other domains. For the multidomain type I PKSs, the mobility and flexibility of the ACP domain facilitates the shuttling of substrates to and between all the different active sites. The binding of ACP with other functional domains are either modulated by the recognition of the acyl substrates or a specific region on the ACP or both. However, these dynamic domain-domain interactions have to be weak and transient to ensure efficient product synthesis. Studying the determinants for efficient transport or delivery of carrier protein-tethered substrates is essential for understanding the complex catalytic process on PKSs.

Because of its transient nature, detailed insights into ACP based interactions are less likely to come from single crystallographic snapshot of intact multidomain enzymes. Currently, several approaches are involved to study the ACP-centered protein-protein interactions, including computer simulation of docking models based on reported structures, mutagenesis followed by biochemical assays and solution-phase NMR spectroscopy to probe the dynamic behaviors.

We have initiated the structure determination of LovF ACP domain using solution-phase NMR, in collaboration with Dr. Robert Clubb lab in UCLA. We purified $^{15}N$ and $^{13}C$ labeled LovF ACP from *E. coli* and the resulting sample is analyzed on the NMR instrument. Experimental results showed that the protein is folded and we were able to work on structure determination. Detailed data are not shown in this dissertation due to incomplete structure determination at current stage.

### 2.1.2 LovG Releases Polyketide from Lovastatin Nonaketide Synthase (LNKS or LovB) and Increases Turnover

A version of this section was published as: Xu, W.; Chooi, Y. H.; Choi, J. W.; Li, S.; Vederas, J. C.; Da Silva, N. A.; Tang, Y. *Angew. Chem. Int. Ed. Engl.* **2013**, DOI:10.1002/anie.201302406.

One unsolved biochemical step in the lovastatin biosynthetic pathway is the release of dihydromonacolin L acid (**2)** from LovB that allows multiple turnovers by this enzyme. In the previous *in vitro* reconstitution work from our lab [62], we showed that purified LovB and LovC were sufficient to assemble **2** tethered to LovB from malonyl-CoA, but were not able to release the product. In this work, we demonstrate that LovG, a thioesterase found in the lovastatin biosynthetic pathway, is the missing link in lovastatin biosynthesis.

#### 2.1.2.1 Materials and Methods

**Materials**

*E. coli* TOP10 (Invitrogen) and *E. coli* XL1-Blue (Stratagene) were used for cloning, following standard recombinant DNA techniques. DNA restriction enzymes were used as recommended by the manufacturer (New England Biolabs). PCR was performed using AccuPrime™ Pfx DNA Polymerase (Invitrogen). The constructs of pCR-Blunt vector (Invitrogen) containing desired PCR products were confirmed by DNA sequencing. *E. coli* BL21 (DE3) (Novagen) was used for protein expression. *Saccharomyces cerevisiae* strain BJ5464-NpgA was used as the yeast expression host. BJ5464-NpgA was created through chromosomally integrating the *A. nidulans* phosphopantetheinyl transferase gene *npgA* to BJ5464 (*MATα ura3-52 his3-Δ200 leu2-Δ1 trp1 pep4::HIS3 prb1 Δ1.6R can1 GAL*).[153]

**Cloning of N-His$_6$ LovG expression vector in *E. coli.***

DNA sequences of *lovG* were duplicated from the genomic DNA of *Aspergillus terreus* NIH2624, using the primers LovG Nhis-NdeI-F and LovG Nhis-NotI-R listed in **Table 2**. The pCR product was ligated into pCR-Blunt vectors and subsequently digested with *Nde*I and *Not*I and then inserted into pET-23a (+) to yield pXW157.

**Table 2.** Primers used for LovG study

| Primer | Sequence (5'-3') | Description |
|---|---|---|
| LovG Nhis-*Nde*I-F | AA**CATATG**CACCATCATCACCACCACATGCGCT-ACCAAGCATCTCCAG-3' | Forward cloning primer for construction of N-His tagged LovG in *E. coli* |
| LovG Nhis-*Not*I-R | AA**GCGGCCGC**TCACTCCAATGTCTGGACCGTC | Reverse cloning primer for construction of N-His tagged LovG in *E. coli* |
| LovG *Nde*I-F | AA**CATATG**CGCTACCAAGCATCTCCAG | Forward cloning primer for construction of C-His tagged LovG in *S. cerevisiae* |
| LovG Chis-*Pme*I-R | AA**GTTTAAAC**TCAGTGGTGGTGATGATGGTGCT-CCAATGTCTGGACCGTC | Reverse cloning primer for construction of C-His tagged LovG in *S. cerevisiae* |
| LovG_KO_screen-f | CAGATGGACTCTCCGCGAAG | external primers for PCR screening of double homologous recombination |
| LovG_KO_P2-F | TCACCGTGGGAACCTGGTTCTC | fusion PCR of *lovG* deletion cassette |
| LovG_KO_P3-R | CGTCCGTCCGTCTCTCCGCATGGTGTGACGGCC-CAGACATTGGAG | fusion PCR of *lovG* deletion cassette |
| GpdA-F | CATGCGGAGAGACGGACGGAC | Amplification of *gpdA-ble* from pAN8-1 |
| BLE-R | TCAGTCCTGCTCCTCGGCCAC | Amplification of *gpdA-ble* from pAN8-1 |
| LovG_KO_P4-F | CGTGGCCGAGGAGCAGGACTGAGGAGATGCTT-GGTAACGCATG | fusion PCR of *lovG* deletion cassette |
| LovG_KO_P5-R | CTGTAGAAGACAATGAGCTTG | fusion PCR of *lovG* deletion cassette |
| LovG_KO-Screen-r | TAATCCGCATAGCAGCCATGC | external primers for PCR screening of double homologous recombination |
| for in BLE | GGCCTGGACGAGCTGTACGC | internal primers for PCR screening of double homologous recombination |
| rev in BLE | CTGATGAACAGGGTCACGTCG | internal primers for PCR screening of double homologous recombination |
| LovG-F-Hind | TCAAGG**AAGCTT**GCAAAACGTAGGGGCAAACA AA | Forward cloning primer for *ADH2* promoter-LovG-*ADH2* terminator |
| LovG-R-Hind | CGAACT**AAGCTT**GGGAGCAAAAAGTAGAATAT TATC | Reverse cloning primer for *ADH2* promoter-LovG-*ADH2* terminator |

* Restriction recognition sites are underlined.

**Cloning of C-His$_6$ LovG cotransform vector in *Saccharomyces cerevisiae.***

The primer pair LovG *Nde*I-F and LovG Chis-*Pme*I-R was used to amplify *lov*G sequence (**Table 2**). The resulting PCR product was ligated into pCR-Blunt vector and subsequently digested with *Nde*I and *Pme*I and then inserted into the digested pXW02 (an expression vector with a *LEU2* marker) to create pXW161 (**Table 3**).

**Cloning of LovC-LovG cotransform vector in *Saccharomyces cerevisiae.***

The primer pair LovG *Nde*I-F and LovG Chis-*Pme*I-R was used to amplify *ADH2p-lov*G-*ADH2t* casseette. The resulting PCR product was ligated into pCR-Blunt vector and subsequently digested with *Hind*III and then inserted into the digested pXW06 (the LovC expression plasmid with a *TRP1* marker) to create pSL05 (**Table 3**).

**Table 3.** Expression plasmids used in LovG study

| Plasmid | Description |
|---|---|
| pXW157 | N-terminus hexahistidine tagged LovG in *E. coli*, ampicillin resistance |
| pXW161 | C-terminus hexahistidine tagged LovG in *S. cerevisiae*, *LEU2* marker |
| pXW06 | C-terminus hexahistidine tagged LovC in *S. cerevisiae*, *TRP1* marker |
| YepLovB-6His | C-terminus hexahistidine tagged LovB in *S. cerevisiae*, *URA3* marker |
| pSL05 | C-terminus hexahistidine tagged LovC and LovG in *S. cerevisiae*, *TRP1* marker |
| *S-LovA/CPR::pESC-Leu2d* | Synthetic LovA and CPR in *S. cerevisiae*, *LEU2* marker |

***lovG* gene knockout in *Aspergillus terreus***

Inactivation cassette containing the bleomycin gene (*ble*) was constructed through fusion PCR[154], two 2 kb homologous regions in *Aspergillus terreus* NIH2624 genome were used to flank the resistant marker that is under the *gpdA* promoter.[155] Fusion PCR product was gel purified, ligated into pCR-Blunt II vector and sequenced before using for transformation. Inactivation cassette was linearized by PCR and transformed into *Aspergillus terreus* NIH2624 through a polyethylene glycol–mediated protocol as described previously.[156] Transformants were grown on glucose minimal agar medium[157] supplemented with 1.2 M sorbitol and 250 μg/mL zeocin. To confirm correct integration of the inactivation cassette into the genome, gDNA was extracted from the transformants and used as template for PCR using primer pairs (**Figure 22**).

**Expression and Purification of LovG from *E. coli***

A chaperone plasmid pG-KJE8 (Takara) was co-transformed with pXW157 into *E. coli* BL21 (DE3) strain through electroporation. The transformant was incubated at 37 °C in 1L LB

medium containing 35 μg/mL ampicillin, 25 μg/mL chloramphenicol, 0.5 mg/mL *L*-arabinose and 5 ng/mL tetracycline till an OD600 of 0.4-0.6. The cells were incubated on ice for 10 minutes, and then induced with isopropylthio-β-D-galactoside (IPTG) (0.1 mM final concentration). The induced culture was incubated at 16°C for 16 hours before the cells were harvested by centrifugation (2500 g, 15 minutes, 4°C). The cells were re-suspended in 30 mL lysis buffer (50 mM Tris-HCl, 2 mM EDTA, 2 mM DTT, 500 mM NaCl, 5 mM imidazole, pH=7.9) and lysed through sonication on ice. Cellular debris was removed by centrifugation (30,000 g, 30 min, 4°C). Ni-NTA agarose resin was added to the supernatant and the mixture was binding at 4°C for at least 2 hours. LovG is elute together with chaperone protein GroEL with increasing concentration of imidazole in buffer A (50 mM Tris-HCl, 500 mM NaCl, pH=7.9) on a gravity flow column. The eluate was concentrated and exchanged into buffer E (50 mM Tris-HCl, 100 mM NaCl, pH=7.9). The resulting solution was loaded onto a sephacryl-S100 size-exclusion column (GE Healthcare Life Sciences) under isocratic condition (0.5 mL/min flow rate with buffer E) for isolation of LovG from GroEL.

**Expression and Purification of proteins in *S. cerevisiae***

The expression plasmids were transformed into *S. cerevisiae* strain BJ5464-NpgA for expression of PKS proteins, by using S. c. EasyComp$^{TM}$ Transformation Kit (Invitrogen). For 1 L of yeast culture, the cells were grown at 28°C in YPD media with 1% dextrose for 72 hours. The cells were harvested by centrifugation (2500 g, 20 minutes, 4°C), resuspended in 20 mL lysis buffer (50 mM $NaH_2PO_4$ pH 8.0, 0.15 M NaCl, 10 mM imidazole) and lysed through sonication on ice. Cellular debris was removed by centrifugation (35,000 g, 1 hour, 4°C). Ni-NTA agarose resin was added to the supernatant and the solution was stirred at 4°C overnight.

The protein/resin mixture was loaded into a gravity flow column and proteins were purified with increasing concentration of imidazole in Buffer A (50 mM Tris-HCl, pH=7.9, 2 mM EDTA, 2 mM DTT).

**Storage of purified proteins.**

Purified proteins were concentrated and exchanged into buffer E (50 mM Tris-HCl, 100 mM NaCl, pH=7.9) + 10% glycerol. The concentrated enzyme solutions were aliquoted and flash frozen. Protein concentrations were determined with the Bradford (Biorad) assay using BSA as a standard.

***In vitro* turnover assays of LovG**

For *in vitro* assays in this study, a typical reaction volumn is 100 μL and the final concentrations of enzymes were 10 μM. Cofactor concentrations are 2 mM NADPH, 1 mM SAM. The assays were carried out in 100 mM $NaH_2PO_4$ pH 7.4 buffer, at room temperature. MatB[158] (20 μM) was used to replenish malonyl-CoA from 10 mM CoA and 100 mM malonate or [2-$^{13}$C]-malonate in presence of 20 mM ATP and 7 mM $MgCl_2$.. The reaction mixtures were incubated at room temperature for 12 hrs, then quenched and extracted twice with equal volume of 99% ethyl acetate (EA)/1% acetic acid (AcOH). The resultant organic extracts were evaporated to dryness, redissolved in 0.05M NaOH in methanol, and then analyzed on LC-MS. LC-MS was conducted with a Shimadzu 2010 EV Liquid Chromatography Mass Spectrometer by using both positive and negative electrospray ionization, and a Phenomenex Luna 5μ 2.0×100 mm C18 reverse-phase column. Samples were separated on a linear gradient of 5 to 95% $CH_3CN$ (v/v) in $H_2O$ supplemented with 0.05% (v/v) formic acid at a flow rate of 0.1 mL/min.

***In vivo* coexpression assays**

For *in vivo* assays in this study, *S. cerevisiae* strain BJ5464-NpgA harboring the corresponding expression plasmids was inoculated to the corresponding Yeast Synthetic Drop-Out medium without the nutruant like uracil, tryptophan or leucin. The cells were grown for 72 hours with constant shaking at 28°C. The seed culture was inoculated to 50 mLYPD (1% dextrose) to an initial $OD_{600}$ 0.1. The culture is incubated with constant shaking at 28°C. For induction of LovA and CPR, total 1% galactose was added to the culture after 24 hours of shaking. The sample culture was extracted twice with equal volume of 99% ethyl acetate (EA)/1% acetic acid (AcOH). The resultant organic extracts were evaporated to dryness, redissolved in 0.05M NaOH in methanol, and then analyzed on LC-MS. LC-MS analysis was conducted with a Shimadzu 2010 EV Liquid Chromatography Mass Spectrometer by using both positive and negative electrospray ionization, and a Phenomenex Luna 5μ 2.0×100 mm C18 reverse-phase column. Samples were separated on a linear gradient of 5 to 95% $CH_3CN$ (v/v) in $H_2O$ supplemented with 0.05% (v/v) formic acid at a flow rate of 0.1 mL/min.

**2.1.2.1 Results and Discussions**

**A natural TE is necessary for the efficient product releasing from LovB**

Known that LovB and LovC cannot produce **2** *in vitro*, we tested whether this is also the case in *S. cerevisiae*. LovB and LovC were coexpressed in *S. cerevisiae* BJ5464-NpgA, which is a vacuolar protease-deficient yeast strain harbouring an *A. nidulans* phosphopantethienyl transferase *npgA*.[56] Following extraction of the 3-day yeast culture, no trace of **2** was found by selective ion monitoring (**Figure 19**), confirming the lack of a releasing function with LovB and LovC only.

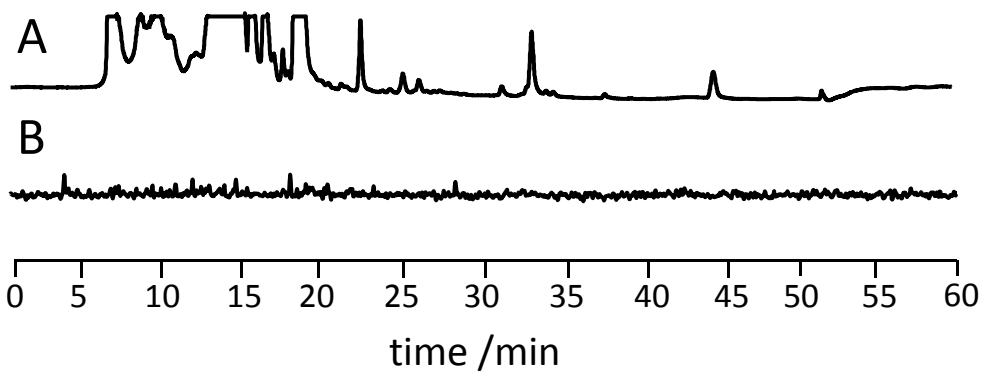



**Figure 19.** Coexpression of LovB and LovC alone in BJ5464-NpgA cannot produce dehydromonocolin L. (**A**) HPLC trace at 200 nm. (**B**) Extracted ion chromatogram for 2, m/z $[M-H]^-$ 323.

Serendipitously in our *in vitro* work, we identified non-native thioesterase (TE) domains from unrelated fungal PKSs that can offload **2**.[62] Co-expression of fungal TE domains, such as the PKS13 TE,[159] Hpm3 TE,[160] etc, with LovB and LovC resulted in detectable levels of **2** from BJ5464-NpgA, albeit with very low titers (< 400 μg/L, **Figure 20**). Given the known expression levels of LovB (~4 mg/L) [62], this low titer corresponds to an average turnover of less than 90 for LovB. These studies strongly indicate that 1) the low titer of **2** from *S. cerevisiae* must be overcome by increasing the turnover rate of LovB; and 2) there must exist a natural TE that partners with LovB in the release of **2.**

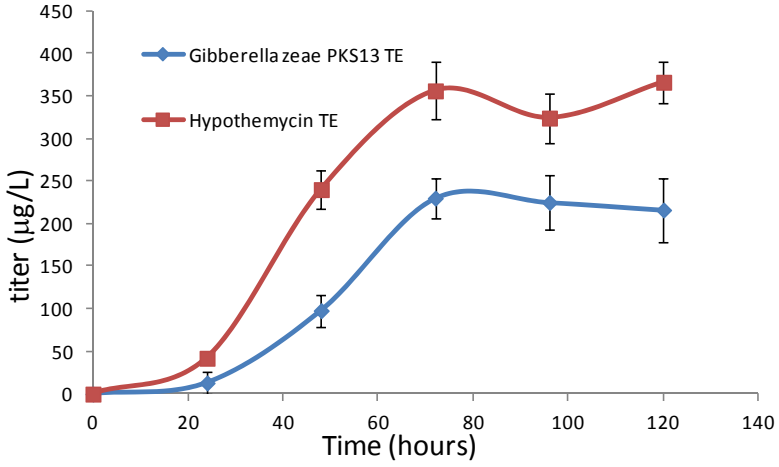



**Figure 20.** Time course of the DML acid production in *S. cerevisiae* strain BJ5464-NpgA that co-transformed with *lovB*, *lovC* and heterologous TE genes.

### *lovG* Gene might encode the missing thioesterase

To identify the TE that is involved in the biosynthesis of **2**, we revisited the *A. terreus lov* gene cluster and searched for a likely candidate. The only known TE-like enzyme in the gene cluster, LovD, is highly specific towards LovF and does not function with LovB. A thorough bioinformatic analysis of genes of unassigned function suggests that a gene (ATEG_09962) that is located between *lovB* and *lovC* (**Figure 21**) warrants further scrutiny. This gene was initially assigned as *orf5*[61] and later as *lovG*[161], with an annotation of the protein product as either a hypothetical protein or an oxidoreductase. Conserved domain analysis of LovG indicates that LovG in fact belongs to the esterase-lipase family of serine hydrolases (**Figure 21**).

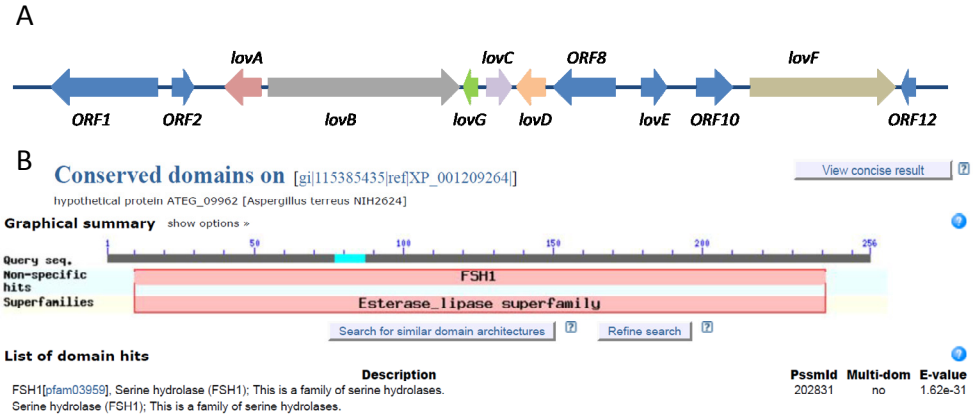



**Figure 21.** The hypothetical thioesterase gene *lovG* (**A**) Location of LovG in the lovastatin gene cluster; (**B**) NCBI Conserved Domain Search result for LovG LovG belongs to serine hydrolase family

Notably, close homologs of LovG are also found in related biosynthetic pathways, including MlcF in the compactin pathway in *Penicillium citrinum*[162] (58% identity, 71% similarity); and MokD in the monacolin K (lovastatin) pathway in *Monascus pilosus*[163] (67% identity, 76% similarity). As is the case for *lovG,* both *mlcF* and *mokD* are located between the genes encoding nonaketide PKS and the *trans*-acting enoylreductase. RT-PCR analysis showed that *mlcF* and *mokD* are moderately co-transcribed with the other biosynthetic genes in their corresponding clusters, which suggests that these enzymes should be involved in the biosynthetic

pathways.[162, 163] Therefore, we hypothesized that LovG is the missing link that is required for LovB turnover and release of **2**.

***lovG* is involved in the lovastatin biosynthesis**

We first determined the role of *lovG* in biosynthesis of **1** in *A. terreus* through deletion of *lovG*. The genetic disruption of *lovG* was performed using a double-crossover recombination with the zeocin-resistant marker, followed by identification of desired *ΔlovG* mutants via diagnostic PCR (**Figure 22B**). Metabolite analysis of the *ΔlovG* mutants showed dramatically reduced level of **1** (< 5%) compared to the wild type (**Figure 22C**). No intermediate such as **2**, **4** or **3** was accumulated in the culture, indicating LovG is involved in the early part of the pathway.

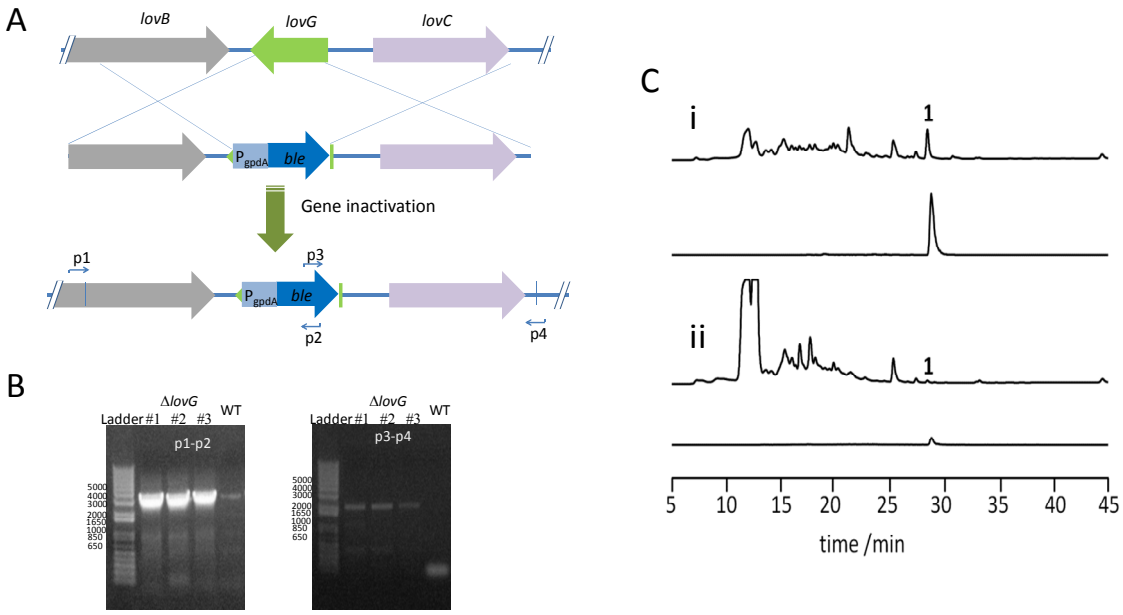



**Figure 22.** Functional deletion of *lovG* (**A**) Knockout strategy employed for deletion of *lovG* from the *lov* cluster; (**B**) Diagnostic PCR analysis of Δ*lovG* mutant using primer pairs described above. (Primers: p1—LovG_KO_screen-f; p2—rev in BLE; p3—for in BLE; p4—LovG_KO-Screen-r); (**C**) Metabolic profiles of fermentation extracts from cultures of different *A. terreus* strains. i) HPLC chromatograms at 238 nm and the extracted ion chromatograms of observed masses *m/z* $[M+Na]^+$ 445 corresponding to **1** in wild type *A. terreus*; **ii**) HPLC chromatograms at 238 nm and the extracted ion chromatograms of observed masses *m/z* $[M+Na]^+$ 445 corresponding to **1** in *ΔlovG A. terreus* mutant. The traces are drawn to the same scale.

The significant attenuation of **1** biosynthesis indicates that *lovG* is intimately involved in the *lov* pathway, but not absolutely essential for the production of **1**. This is not unexpected given

our previous findings: an endogenous TE in *A. terreus* may be recruited for the hydrolysis of **2**, albeit at a much lower efficiency. In a recent study, the region encoding *lovG* (between *lovB* and *lovC*) was replaced by sequences encoding the nonribosomal peptide synthetase (NRPS) module from the CheA PKS-NRPS to yield LovB fused with a NRPS module.[164] The authors observed a similar reduction in the level of **1,** and attributed the drop in yield to the reduced efficiency of the chimeric LovB.[164] We believe the decrease titer of **1** in that study can also be from the unintended deletion of *lovG*.

**LovG is functional in releasing dihydromonacolin L from LovB**

To further confirm the hydrolytic role of LovG, we probed the direct involvement of this enzyme in turnover of **2** by *in vitro* and *in vivo* reconstitution assays. First, LovG was expressed and purified as *N*-terminal His-tagged protein from *E. coli* BL21 (DE3). Chaperone proteins GroES/EL, DnaK/J and GrpE were co-expressed[165] to assist the folding of LovG and to improve the solubility (**Figure 24A**). By incubating equimolar amounts (10 μM) of purified LovB, LovC and LovG with malonyl-CoA and the necessary cofactors (NADPH, SAM) for 12 hours at 25°C, we observed production of **2** in the assay without the need to add either base (1M KOH) or heterologous TE (**Figure 23B**). With LovG included, the turnover rate of **2** was estimated to be ~ 0.12 $hour^{-1}$.(**Figure 24B**) The slow, albeit observable turnover may be due to the large fraction of recombinant LovB being inactive in the *in vitro* assay as previously noted.[62] Furthermore, *lovG* was cloned under *S. cerevisiae ADH2* promoter and coexpressed with *lovB* and *lovC*. This resulted in the production of **2** from *S. cerevisiae* BJ5464-NpgA at an elevated titer of ~35 mg/L after ~50 hours of culturing (**Figure 23A, Figure 24C**). Collectively, these results confirmed LovG is the natural releasing partner of LovB during lovastatin biosynthesis in *A. terreus*.

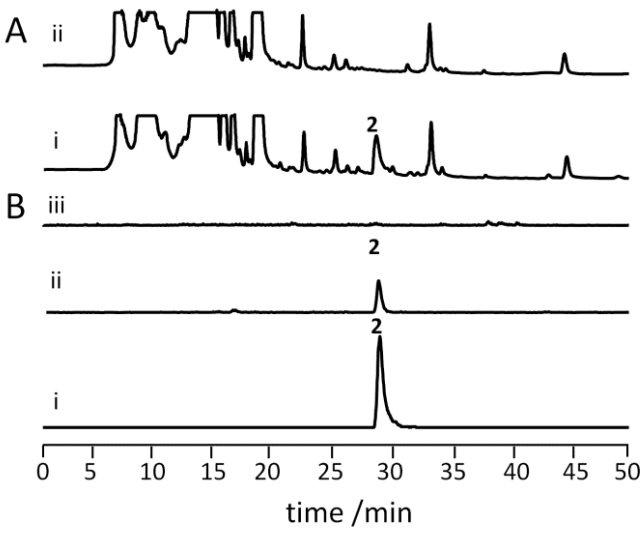



**Figure 23.** LovG releases dihydromonacolin L from LovB. (**A**) HPLC chromatogram at 200 nm of fermentation extracts from 3-day cultures of BJ5464-NpgA expressing **i**) LovB, LovC and LovG; **ii**) LovB and LovC; and **B**) Extracted ion chromatograms of *m/z* $[M-H]^-$ 323 corresponding to 2 in the *in vitro* experiments containing **i**) LovB, LovC and LovG; **ii**) LovB and LovC, treated with 1M KOH; **iii**) LovB and LovC with no 1M KOH treatment. All enzymes are added to a final concentration of 10 μM. The reactions were performed at 25°C for 12 hours.

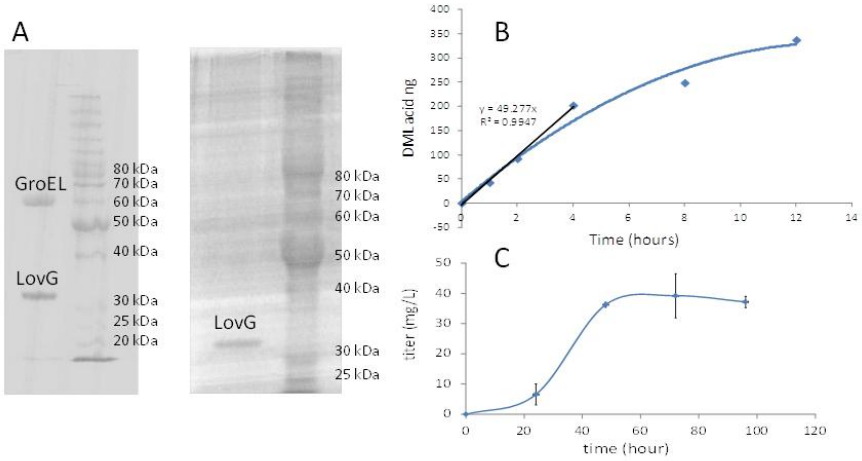



**Figure 24.** Heterologously expressed LovG is functional. (**A**) SDS PAGE gel images of the heterogeneously expressed LovG from *E. coli*, before and after size-exclusion chromatography. (**B**) *In vitro* turnover assay for production of **2** by LovB, LovC and LovG. Parallel reactions were taken place in 100 μL solution buffered at pH 7.4. Equimolar amounts (10 μM) LovG, LovB and LovC were incubated with necessary cofactors: 2 mM NADPH and 1 mM SAM. 20 μM MatB was used to replenish malonyl-CoA from 10 mM CoA and 100 mM malonate in presence of 20 mM ATP and 7 mM $MgCl_2$. The produced 2 is quantified on LC-MS by comparing the MS peak area to the standard. The initial turnover rate was calculated using the first four data points that are in the linear region of the curve. (**C**) Time course of the DML acid (**2**) production in *S. cerevisiae* strain BJ5464-NpgA that co-transformed with *lovB*, *lovC* and *lovG* genes. The sample culture was extracted twice with equal volume of 99% ethyl acetate (EA)/1% acetic acid (AcOH). The resultant organic extracts were evaporated to dryness, redissolved in 0.05M NaOH in methanol, and then analyzed on LC-MS. The produced **2** is quantified on LC-MS by comparing the MS peak area to the standard.

**Production of monacolin J from yeast**

The successful production of **2** following LovG coexpression prompted us to examine the ability of the yeast host to produce more advanced intermediates such as the desired **3.** The 2μ vector encoding both *lovA* and the endogenous *A. terreus* cytochrome P450 oxidoreductase (CPR) from *A. terreus* [117] was introduced into BJ5464-NpgA and coexpressed with LovB, LovC and LovG. In the absence of expression of LovA and CPR, which are controlled by the divergent *GAL1-GAL10* promoter, we were only able to observe the accumulation of **2**. When galactose was added to the culture media at day 2, nearly all of **2** was oxidized to **4** and **3.** After 48 hours, both **3** and **4** can be detected from the yeast culture at ~20 mg/L each (**Figure 25**). It is important to note here that this initial demonstration of biosynthesis of **3** has not been optimized. With additional metabolic engineering approaches, we expect the conversion from **4** to **3**, as well as the titer of **3** can be significantly improved.

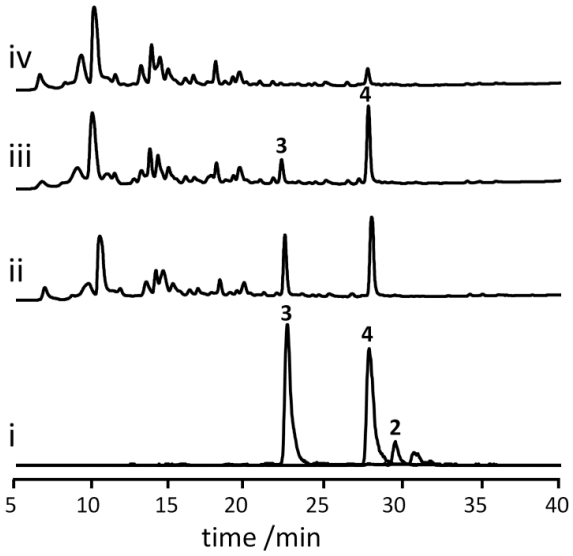



**Figure 25.** Production of monacolin J from *S. cerevisiae* BJ5464-NpgA co-expressing LovB, LovC, LovG, LovA and *A. terreus* cytochrome P450 oxidoreductase (CPR) **i**) Extracted ion chromatograms showing detection of **2**, **3** and **4** in culture 48 hours after induction of LovA and CPR expression. *m/z* $[M-H]^-$ for **2**: 323, **3**: 337 and **4**: 321. **ii-iv**) HPLC Chromatograms at 238 nm of fermentation extracts. **ii**) 48 hours after induction of LovA and CPR expression; **iii**) 24 hours after induction of LovA and CPR expression; **iv**) Before induction of LovA and CPR expression.

LovG joins a growing list of fungal TEs that are involved in product release via either hydrolysis or acyl transfer to an on-pathway intermediate (such as LovD [116]). All of the TEs associated with HR-PKSs characterized to date are stand-alone enzymes, in contrast to the fused TE domains in mammalian fatty acid synthases (FAS)[81] or some of the nonreducing PKSs (NR-PKS), in which the fused TE can act as Claisen-like Cyclases (TE/CLC).[108] Recently, TE/CLCs from NR-PKS was also shown to possess editing functions during the PKS function and hydrolyze stalled products.[166] Given that LovB is noted to be highly accurate in the assembly of **2** as a sole product both *in vitro* and here, *in vivo*, we hypothesized that LovG may exert proofreading functions during the iterative LovB functions to remove aberrant tailored products. This function can ensure the timely offloading of incorrect, stalled intermediates and free LovB for more efficient turnover of **2**.

**LovG also releases aberrant product from LovB**

To assay the possible editing function of LovG, we performed *in vitro* reconstitution experiments in the absence of LovC, of which the earliest function in the pathway is to reduce the α-β enoyl intermediate at the tetraketide stage. We previously observed small amounts of offloaded products from LovB in the forms of methylated, conjugated α-pyrones such as **6** and **7** (**Figure 26A**) via continued chain elongation and spontaneous esterification); and ketones (not visible at the scale drawn in **Figure 26A**) via hydrolysis of β-keto thioesters and decarboxylation), when LovC is excluded.[61, 62] Upon addition of LovG at equimolar concentration to LovB, the *in vitro* assay mixture turned significantly more yellow and the extract contained high amounts of offloaded products **8-10**, in addition to **6** and **7** (**Figure 26B).** The compounds **8-10** were not stable and precluded the precise structural determination by NMR. However, through comparison of UV absorption spectra and mass increase upon using [2-$^{13}$C]

malonate, as well as previously deciphered programming rules of LovB[62], we were able to deduce the structures of these compounds as shown in **Figure 26**. **8** is a hydrolyzed -keto acid, while **9** and **10** are ketones resulting from the decarboxylation of longer polyene β-keto acids. Interestingly, while the levels of **8-10** are high in the LovG-containing assay, the relative amounts of pyrones **6** and **7** remained nearly the same as in assay with LovB alone, indicating that LovG can readily hydrolyze the incorrectly tailored, β-keto thioesters from LovB to give **8-10**. In contrast, the spontaneous hydrolysis of β-keto intermediates is much slower without LovG, resulting in LovB having to offload the products through an additional round of chain elongation, which can then be cyclically released as **6** and **7** (**Figure 27**)**.** A recent study showed that NR-PKS TEs can also promote the formation of pyrones.[167] Therefore, LovG may also assist in the formation of **6** and **7**, albeit likely at a much lower level as evident in Figure 26.

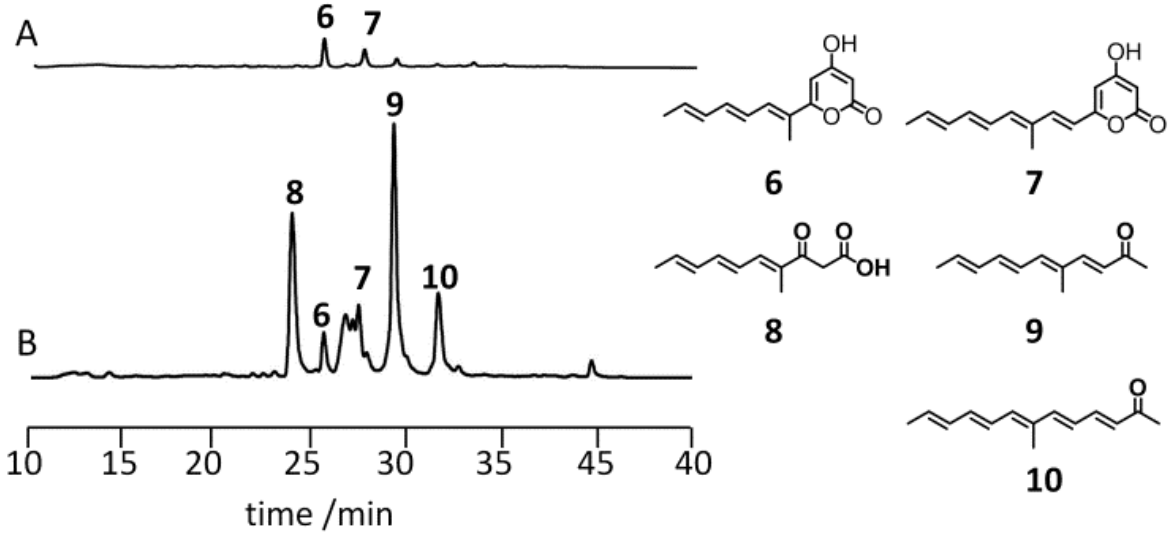



**Figure 26**. Aberrant products released from LovB by LovG (HPLC Chromatograms are extracted at 330 nm.) (**A**) *in vitro* assay with LovB; (**B**) *in vitro* assay with LovB + LovG.

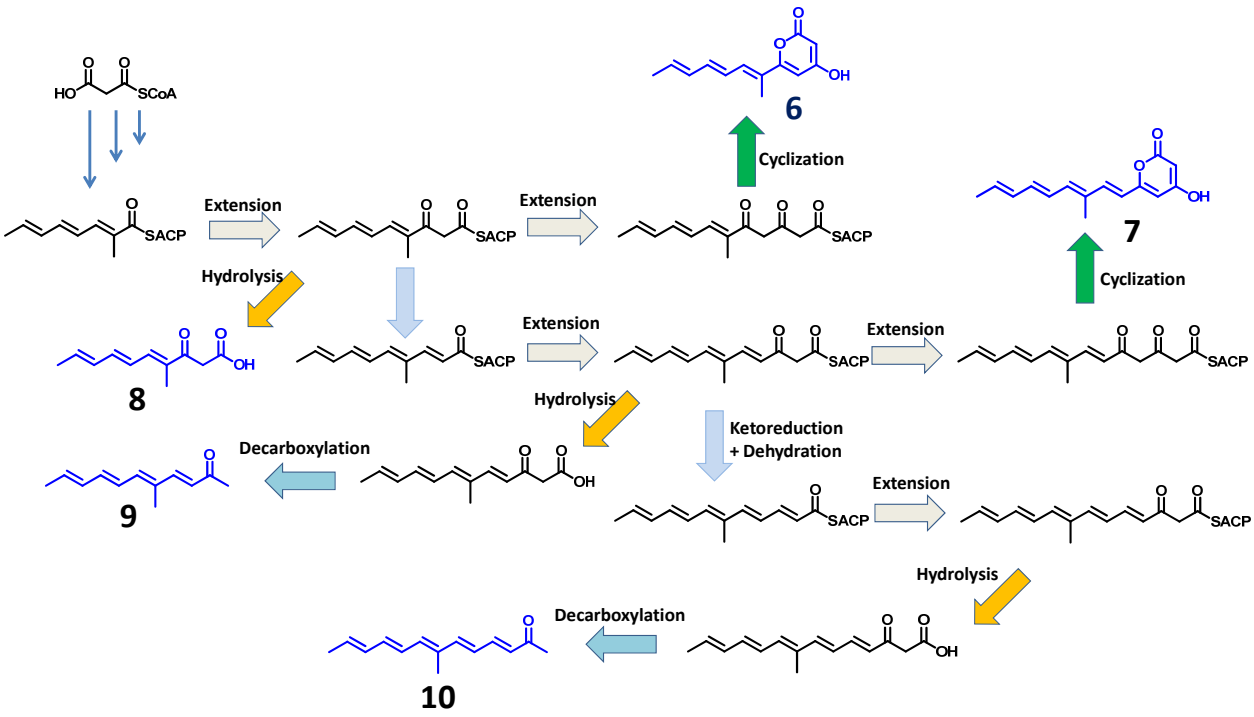



**Figure 27.** Editing function of LovG in absence of LovC

## Discussion

The finding of *lovG* confirmed the conclusion that HR-PKS lacks of product releasing mechanism and indicates that stand-alone TE-like gene in the cluster encodes the PKSs may be the candidates to release the product from HR-PKS.

Also, as a multifunctional esterase, LovG not only involved in the release of the correct product **2** from LovB, but is also shown to play a role in the clearance of aberrant intermediates from LovB. Construction of the LovG-containing pathway capable of *de novo* synthesis of **3** also opens up new metabolic engineering opportunities for statin production from yeast.

## 2.2 Mechanisms Studies on Polyketide Synthase-Nonribosomal Peptide Synthetase Hybrid in Aspyridone Biosynthesis[1]

### 2.2.1 Introduction

Filamentous fungi produce a diverse array of bioactive secondary metabolites. Among the small molecule natural products, polyketides and nonribosomal peptides are synthesized by multidomain enzymes such as polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs), respectively. The enzymology and biochemical properties of these mega-enzymes are highly complex and are different from the well-studied bacterial PKSs and NRPSs.[119, 168] In particular, the highly-reducing PKSs (HR-PKSs), in which individual domains are programmed to function in different permutations during the iterative process of chain elongation, are particularly enigmatic as exemplified by the lovastatin nonaketide synthase LovB.[62] Notwithstanding the complexity of HR-PKSs, an even more impressive biosynthetic machinery that is found in nearly all filamentous fungi is the PKS-NRPS hybrid, in which a single module of NRPS is translationally fused to the *C*-terminus of a HR-PKS. A typical PKS-NRPS hybrid contains ~ 10 catalytic domains; exceeds 400 kDa as a stand-alone protein; and synthesizes acyltetramic acid (3-acyl-pyrrolidine-2,4-dione) as the product. The tetramic acids are further tailored into complex natural products, such as fusarin C,[169] equisetin,[170] aspyridone,[171] pseurotin A,[172] cyclopiazonic acid,[173, 174] chaetoglobosin A,[175] tenellin,[176] etc (**Figure 28A**). The widespread presence of fungal PKS-NRPS machineries,[177, 178] along with their unique biosynthetic capabilities and highly evolved programming rules, make them attractive targets for biochemical investigations *in vitro* and in genetically tractable hosts.

[1] Compounds are numbered independently in this section

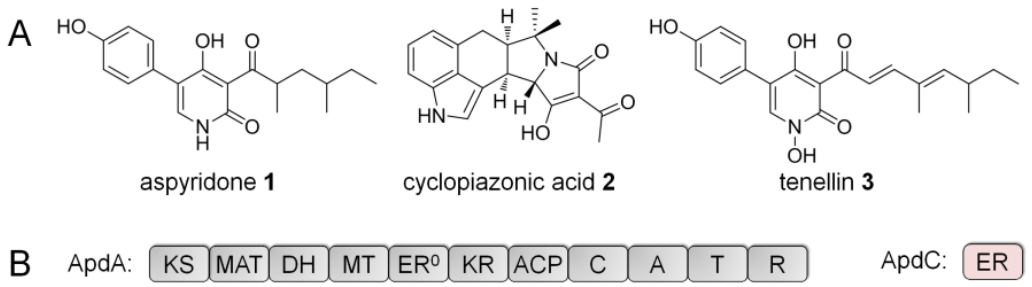



**Figure 28**. Tetramic acid synthesized from fungal PKS-NRPS (**A**) Selected tetramic acid-derived natural products. (**B**) Domain organization of ApdA and ApdC

### 2.2.2 Reconstitution of PKS-NRPS Hybrid in Aspyridone Biosynthesis

ApdA is the PKS-NRPS that is involved in the synthesis of aspyridone A (**1)** (**Figure 28A**). The role of ApdA was discovered via artificial transcription activation of the cryptic *apd* gene cluster in *Aspergillus nidulans*.[171] The HR-PKS module of the enzyme, which consists of ketosynthase (KS), malonyl–CoA:ACP transacylase (MAT), dehydratase (DH), methyltransferase (MT), ketoreductase (KR), and acyl carrier protein (ACP) (**Figure 28B**) domains, is proposed to synthesize the ACP-bound 4,6-dimethyl-3-oxooctanyl thioester **5**. A dissociated enoylreductase (ER) ApdC is proposed to function *in trans* with ApdA, analogous to the role played by LovC during LovB-catalyzed synthesis of dihydromonacolin L.[61] The downstream NRPS module contains the condensation (C), adenylation (A) and thiolation (T) domains, and catalyzes the formation of the L-tyrosinyl-thioester **4** and the amide linkage between **4** and **5** to yield **6** tethered to the T domain. The bimodular assembly line is terminated with a putative reductase (R) domain that facilitates formation and release of the tetramic acid product. Combined with ApdC, ApdA is proposed to synthesize a precursor of **1** via ~25 enzymatic steps. Although the PKS-NRPSs responsible for tenellin (TenS) and cyclopiazonic acid (CpaS) have both been reconstituted in the heterologous host *Aspergillus oryzae,* [179, 180] *in vitro* reconstitution of fungal PKS-NRPS activities has not been accomplished. Reconstitution

of purified biosynthetic enzymes *in vitro* enables the most direct analysis of the functions and products of the megasynthases.[181] Here we show that using the *Saccharomyces cerevisiae* strain BJ5464-NpgA as the expression host,[56, 62] the complete activities of the ApdA and ApdC can be reconstituted, which led to the identification of an acyltetramic acid **7** as the product.

### 2.2.2.1 Materials and Methods

#### Materials

*E. coli* TOP10 (Invitrogen) and *E. coli* XL1-Blue were used for cloning, following standard recombinant DNA techniques. DNA restriction enzymes were used as recommended by the manufacturer (New England Biolabs). PCR was performed using Platinum Pfx DNA polymerase (Invitrogen). The constructs of pCR-Blunt vector (Invitrogen) containing desired PCR products were confirmed by DNA sequencing. *E. coli* BL21 (DE3) (Novagen) and BAP1 were used for protein expression. *Saccharomyces cerevisiae* strain BJ5464-NpgA (*MATα ura3-52 his3-Δ200 leu2- Δ1 trp1 pep4::HIS3 prb1 Δ1.6R can1 GAL*) was used as the yeast expression host.[153],[56]

#### Re-annotation of *apdA* and *apdC* gene

Both *apdA* and *apdC* were found to be mis-annotated. An extra base pair was found in *apdA* around 127 bps before the only intron that was defined by AN8412.3 (**Figure 29**). Insertion of this base into the sequence eliminates the intron and makes the *apdA* a single ORF. ClustalW2 protein sequence alignment (http://www.ebi.ac.uk/Tools/clustalw2) with other PKS-NRPS genes shows that the inserted region (extra 48 residues) is conserved in the other reported PKS-NRPS genes (**Figure 29**). The protein coded by AN8409.3 (ApdC) was aligned with other trans-ERs. The result showed that ApdC also misses some conserved residues. Later experiment showed that the protein based on AN8409.3 was not solublly expressed from *E. coli*. An alternative

annotation of *apdC*, which has 21 extra residues, can recover the missing region (**Figure 29**). So the ApdA and ApdC cloning were performed based on the new annotations.

```
ApdA-New   6651    cttctcttccttggcgagtatcattggcaatgctggccaatccaactacc   6700
AN8412.3   6651    cttctcttccttggcgagtatcattggcaatgcag-ccaatccaactacc   6699

ApdA-New   6701    acgctgcaaacatgttcctctctggtcttgccgccaaccgccgcaagaag   6750
AN8412.3   6700    acgctgcaaacatgttcctctctggtcttgccgccaaccgccgcaagaag   6749

ApdA-New   6751    ggtctggcagcgtccgtcttacacatagggctagtcaccgatgtaggtta   6800
AN8412.3   6750    ggtctggcagc-----------------------------------------   6760

ApdA-New   6801    tgttgctcgcaggggtcaagcaatggaggaacgactccgcaggctgttct   6850
AN8412.3   6760    ----------------------------------------------------   6760

ApdA-New   6851    ttttgccattgtcggaatcagatatccatcatgcgatggcacaaactatt   6900
AN8412.3   6760    ----------------------------------------------------   6760

ApdA-New   6901    ctagctagtccgcataacagctcgcgtcagcacgaaattatcttgggttt   6950
AN8412.3   6761    ----ctagtccgcataacagctcgcgtcagcacgaaattatcttgggttt   6806
```

```
AN8412.3 (2223-2260)    SIIGNAAN----------PTTTLQTCSSLVLPP----------------------------------TAARRVWQPS-PHNS
ApdA-New (2223-2308)    SIIGNAGQSNYHAANMFLSGLAANRRKKGLAASVLHIGLVTDVGYVAR--RGQAMEERLRRLFFLPLSESDIHHAMAQTTLAS-PHNS
TenS     (2343-2431)    AILNNTGQSNYHCANLYMDSLVTNRRSRGLAASIIHVGHVCDTGYVARLVDDTKVQMSLGTTRVMSVSETDVHHAFAEAVRGGQPDSR
CheA     (2240-2326)    MLIGNPGQANYHAANLYMASLAANRRTRGLAASIIHLGYIADVGYVTR--QDQSMRERLEKLLFRPLCESDIHNAFAEAIRASRPNSR
CpaS     (2198-2284)    SVIGTPKQVAYHAPSLFMTDLIQRRRMRGLVGSVMALGMVVDAGYFSR--QGKEVIQRMMHHGYAPLSESDLHYAFGEAVAAGAPEAE
FusA     (2045-2130)    SVVGNRGQANYAASNLFMSAIAEQRRAKGLAASVMHIGMVLGVGYVSS---TGAYEATLRSSNYMAISETDLLNMFSQAILVGQPSST
EqiS     (2212-2298)    ACLGNSGQSNYAAANMFMTTLAFQRRARGVPGSVIDLSSLMGIGHVGR--SDVFNADYFRSLGATSVSEVDLHQMFAEAIHVGRPDSK
```

**Figure 29.** Reanotation of *apdA*. PCR products of *apdA* from *Aspergillus nidulans* strain FGSC A4 were sequenced and an extra base pair was inserted at 6686 bp position compare to the sequence in the database. Insertion of this base into the sequence makes the *apdA* a single ORF. ClustalW2 protein sequence alignment with other PKS-NRPS genes shows that the inserted region (extra 48 residues) is conserved in the other reported PKS-NRPS genes.

```
AN8409.3 (152-165)      QAGKSP------------------------------SGYAAITT
ApdC-New (152-186)      QAGKSPY-------VLVYGGSTATGTLAIQILTRSGYAAITT
CheB     (153-188)      APAKKGFP------VLVSGAGTATGALATQILTLSGLQPIVT
TenS3    (176-127)      QPHDGAFDSDANVFILVYGGGTSTGAIAIQILKAAGFHPITC
MlcG     (157-195)      YSETKP---SKKTYVLIYGGSTATATIAMQFMRLSGYTPIAT
LovC     (154-195)      PSADQPPTHSKPVYVLVYGGSTATATVTMQMLRLSGYIPIAT
```

**Figure 30.** Reanotation of *apdC*. The protein coded by AN8409.3 (ApdC) was aligned with other *in trans*-ERs. The result showed that ApdC also misses some conserved residues. Later experiment showed that the protein based on AN8409.3 was not soluble from *E. coli*. An alternative annotation of *apdC*, which has 21 extra residues, can recover the missing region.

## Plasmid construction:

The re-annotated *apdA* and *apdC* genes were amplified from the genomic DNA of *A. nidulans* (FGSC *A4*) using the primers listed in **Table 4**. The *cpaS* gene was cloned from the genomic DNA of *Aspergillus flavus* NRRL3357. The 2μ yeast-E.coli shuttle plasmids with different auxotrophic markers (*URA3*, *TRP1* and *LEU2*) are used for construction of the yeast expression vectors.[182]

**Table 4**: Primers used for ApdA study

| Primer | Sequence |
|---|---|
| ApdA-*Nde*I-F | 5'-ACATATGCAGGACTTGATCGCCATTGTTG-3' |
| ApdA-*Pml*I-R | 5'-ACCCACGTGAACGGAGGTCTTGGAAC-3' |
| ApdA-*Pml*I-F | 5'-CGTTCACGTGGGTGTGATGAACTCAGA -3' |
| ApdA-*Pml*I-mutation-R | 5'-GGCACGTGAATGGATATAAACCCAGTCACTG -3' |
| ApdA-*Sma*I-F | 5'-TTCACTGCCCGGGTGATCGACGCG-3' |
| ApdA-*Nhe*I-R | 5'-AATGTCCTCTGGGCTAGCGAAG-3' |
| ApdA-*Sma*I-mutation-F | 5'-AACCCGGGGAAATTCTTCGTCATCAAGCAG-3' |
| ApdA-His-Stop-*Swa*I-R | 5'-CCATTTAAATCAATGATGATGATGATGATGAAGAGACCCCGCTTCCCCCATA-3' |
| ApdA-*Spe*I-start-F | 5'-AACTAGTCAGGACTTGATCGCCATTGTTG-3' |
| ApdC-*Nhe*I-start-F | 5'-AAGCTAGCATGATCCCTCCCAAGCAACAAACC-3' |
| ApdC-SOE2-F | 5'-GATCCTTACGCGCTCCGGCTATGCCGCGATTAC-3' |
| ApdC-SOE2-R | 5'-CATAGCCGGAGCGCGTAAGGATCTGGATCGC-3' |
| ApdC-chistag-NotI-R | 5'- TTGCGGCCGCGTGACTCGCGACCTCATACACAAG-3' |
| ApdC-*Nde*I-mutation-F | 5'-CTGCTTAGGATTCCTGAGCACATGGACTACAATGAAGCTGC-3' |
| ApdC-*Nde*I-mutation-R | 5'-GCAGCTTCATTGTAGTCCATGTGCTCAGGAATCCTAAGCAG-3' |
| ApdC-*Nde*I-F | 5'-AACATATGATCCCTCCCAAGCAAC-3' |
| ApdC-Histag-stop-*Swa*I-R | 5'-AAATTTAAATTCAGTGGTGGTGGTGGTGGTGACTCGCGACCTCATACAC-3' |
| ApdA_NRPS-*Nde*I-start-F | 5'-TTCATATGACGGAGAAGCCTGAGACA-3' |
| ApdA_NRPS-*Hind*III-R | 5'-GTAAGCTTAAGAGACCCCGCTTCCCCCA-3' |
| ApdA_PKS-His-Stop-*Swa*I-R | 5'-ACATTTAAATCTAGTGATGATGATGATGGTGGTGGTTGGGCGCAACCGTGTCTTGC-3' |
| CpaS-CATR-*Nde*I-f | 5'-AACATATGGTCAAAGATCTCGATTTCATAG-3' |
| CpaS-*Pml*I-r | 5'-TGCACGTGATTACCCAAGTAGGGATATCGC-3' |

**Table 5.** Expression plasmids used in ApdA study

| Plasmid | Description |
|---|---|
| pXW45 | C-terminus hexahistidine tagged ApdA in *S. cerevisiae*, uracil marker |
| pXW58 | N-terminus FLAG tagged, C-terminus hexahistidine tagged ApdA in *S. cerevisiae*, uracil marker |
| pXW44 | C-terminus hexahistidine tagged ApdC in *E. coli*, ampicillin resistance |
| pXW51 | C-terminus hexahistidine tagged ApdC in *S. cerevisiae*, tryptophan marker |
| pXW71 | C-terminus hexahistidine tagged ApdA PKS in *S. cerevisiae*, uracil marker |
| pXW68 | C-terminus hexahistidine tagged ApdA NRPS in *E. coli*, ampicillin resistance |
| pXW87 | C-terminus hexahistidine tagged CpaS NRPS in *S. cerevisiae*, uracil marker |
| pXW88 | C-terminus hexahistidine tagged CpaS NRPS in *S. cerevisiae*, leucine marker |

## Expression and Purification of proteins in *E. coli*

The expression plasmids were transformed into *E. coli* BL21 (DE3) or *E. coli* BAP1 for protein expression through electroporation. The cells were grown at 37 °C in 1L LB medium with 100 μg/mL ampicillin till an $OD_{600}$ of 0.4-0.6. At which time the cells were incubated on ice for 10 minutes, and then induced with 0.1 mM isopropylthio-β-D-galactoside (IPTG) for 16 hours at 16°C. The cells were harvested by centrifugation (2500 g, 15 minutes, 4°C), re-suspended in 30 mL lysis buffer (50 mM Tris-HCl, 2 mM EDTA, 2 mM DTT, 500 mM NaCl, 5 mM imidazole, pH=7.9) and lysed through sonication on ice. Cellular debris was removed by centrifugation (30,000 g, 30 min, and 4°C). Ni-NTA agarose resin was added to the supernatant and the solution was stirred at 4°C for at least 2 hours. The mixture was loaded into a gravity

flow column and proteins were purified with increasing concentration of imidazole in buffer A (50 mM Tris-HCl, 500 mM NaCl, pH=7.9).

For the expression of ApdC, a chaperone plasmid pG-KJE8 was co-transformed with pXW44 into *E. coli* BL21 (DE3) strain through electroporation. The transformants were incubated at 37 °C in 1L LB medium containing 100 μg/mL ampicillin, 25 μg/mL chloramphenicol and 5 ng/mL tetracycline till an OD600 of 0.4-0.6. The following steps are the same as in a regular *E. coli* protein purification.

**Expression and Purification of proteins in *S. cerevisiae***

The expression plasmids were transformed into *S. cerevisiae* strain BJ5464-NpgA for expression of PKS proteins, by using S. c. EasyCompTM Transformation Kit (Invitrogen). For 1 L of yeast culture, the cells were grown at 25°C in YPD media with 1% dextrose for 72 hours. The cells were harvested by centrifugation (2500 g, 20 minutes, 4°C), resuspended in 20 mL lysis buffer (50 mM $NaH_2PO_4$ pH 8.0, 0.15 M NaCl, 10 mM imidazole) and lysed through sonication on ice. Cellular debris was removed by centrifugation (35,000 g, 1 hour, 4°C). Ni-NTA agarose resin was added to the supernatant (2 mL per liter of culture) and the solution was stirred at 4°C overnight. The protein/resin mixture was loaded into a gravity flow column and proteins were purified with increasing concentration of imidazole in Buffer A (50 mM Tris-HCl, pH=7.9, 2 mM EDTA, 2 mM DTT).

**Purification of FLAG-tagged Protein**

FLAG-tagged protein was purified by using ANTI-FLAG® M2 Affinity Gel (Sigma-Aldrich), following the standard protocols. The clear cell lysate was applied onto a column which was packed with ANTI-FLAG affinity gel. After proper washing processes, the protein was eluted with FLAG peptide.

**Storage of Purified Proteins**

Purified proteins were concentrated and exchanged into buffer E (50 mM Tris-HCl, 100 mM NaCl, pH=7.9) + 10% glycerol. The concentrated enzyme solutions were aliquoted and flash frozen. Protein concentrations were determined with the Bradford (Biorad) assay using BSA as a standard.

***In vitro* Turnover Assays**

For *in vitro* assays in this study, the final concentrations of enzymes were 25 μM, with cofactor concentrations as 2 mM malonyl-CoA, 2 mM NADPH, 1 mM SAM, 0.5 mM amino acid, 25 mM ATP and 10 mM $MgCl_2$. The assays were carried out in 100 mM $NaH_2PO_4$ pH 7.4 buffer, with 2 mM DTT at room temperature. In the assays for $^{13}C$ labeling, 25 μM MatB[158] was used to replenish malonyl-CoA from 20 mM CoA and 100 mM malonate or [2-$^{13}C$]-malonate. The reaction mixtures were usually incubated at room temperature for 12 hrs, then quenched and extracted twice with equal volume of 99% ethyl acetate (EA)/1% acetic acid (AcOH). The resultant organic extracts were evaporated to dryness, redissolved in methanol, and then analyzed on LC-MS. LC-MS was conducted with a Shimadzu 2010 EV Liquid Chromatography Mass Spectrometer by using both positive and negative electrospray ionization, and a Phenomenex Luna 5μ 2.0 x 100 mm C18 reverse-phase column. Samples were separated on a linear gradient of 5 to 95% $CH_3CN$ (v/v) in $H_2O$ supplemented with 0.05% (v/v) formic acid at a flow rate of 0.1 mL/min.

**Co-transformation and Fermentation**

ApdA and ApdC were observed to produce compound **7** in vitro by incubating these two enzymes with needed building blocks and cofactors. However, large-scale purification of compound **7** was performed *in vivo*. pXW45 and pXW51 were encoded for ApdA and ApdC

under the *S. cerevisiae ADH2* promoter, with uracil and tryptophan marker respectively. Both plasmids were co-transformed into *S. cerevisiae* BJ5464-NpgA by using S. c. EasyCompTM Transformation Kit (Invitrogen). The cells were first grown under a minimal medium lack of both uracil and tryptophan for 3 days. Then the seed culture was inoculated into YPD media with 1% dextrose and the cells were grown at 25°C for 5 days before extraction.

**Purification of compound 7**

The culture was centrifuged (2500g, 20 minutes) and the supernatant was acidified to pH 4 with HCl then extracted with equal volume of ethyl acetate twice. The organic extract was combined and concentrated. The residue was defatted by partitioning between n-hexane and 10% aqueous methanol. The methanol fraction was concentrated and the resulting aqueous residue was extracted with ethyl acetate and evaporated to dryness. The defatted extract was separated on a Sephadex LH-20 column with 50%:50% methanol-chloroform as mobile phase. The fractions that contain compound **7** ($R_f \approx 0.3$ on silica TCL under EA) were combined and dried. The residue was dissolved in HPLC grade methanol, then subjected to reverse phase-HPLC preparative purification using an XTerra Prep MS C18 5μm, 19 mm X 50 mm column. A gradient of water (+0.1% TFA) and acetonitrile (+0.1% TFA) solvent system was used. The collected fraction was evaporated and the residual aqueous solution was extracted with ethyl acetate. The EA phase was dried under vacuum and yield the pure compound as yellowish oil.

**Characterization of compound 7**

1D and 2D NMR of **7** were performed on Bruker DRX-500 spectrometer using $CDCl_3$ or $CD_3OD$ as the solvent.

#### 2.2.2.2 Results and Discussions

**Reconstitution of ApdA and ApdC *in vitro***

The uninterrupted *apdA* and *apdC* genes were cloned from the sequenced *A. nidulans* strain FGSC A4 after re-annotation of introns (**Figures 29-30**). A gene encoding *N*-terminal FLAG, *C*-terminal hexahistidine tagged ApdA was placed under the control of the *ADH2* promoter [183] in the 2μ episomal vector pXW58. Full length ApdA (439 kDa) was then solubly expressed from BJ5464-NpgA and purified to near homogeneity with an anti-FLAG antibody affinity resin to a final of yield of ~1.2 mg/L (**Figure 31B**). ApdC was solubly expressed from *Escherichia coli* BL21 (DE3) when coexpressed with chaperone proteins GroEL and GroES (**Figure 35**). To test the activities of the two enzymes and identify possible tetramic acid products, equimolar amounts (25 μM) of purified ApdA and ApdC were incubated at room temperature for 12 hours with all the cofactors and building blocks (2 mM NADPH, 1 mM SAM, 1 mM L-Tyr, 25 mM ATP, 10 mM $MgCl_2$ and 2 mM malonyl-CoA). LC-MS analysis of the organic extract showed the production of a predominant compound **7** with a UV absorption maximum ($\lambda_{max}$) at 279 nm with *m/z* $[M+H]^+$ = 332 (see **Figure 33B**, trace **i**). Exclusion of any of the cofactors abolished the synthesis of **7**. To assess the length of the polyketide portion of **7**, the *in vitro* assay was performed in the presence of [2-$^{13}$C] malonate (100 mM) and MatB (25 μM), which can generate [2-$^{13}$C] malonyl-CoA in the presence of ATP, $Mg^{2+}$ and CoA.[158] The assay yielded the same product profile and the mass of **7** was increased to *m/z* $[M+H]^+$ = 336, suggesting the presence of a tetraketide chain which is consistent with the proposed function of AdpA as shown in **Figure 31A**.

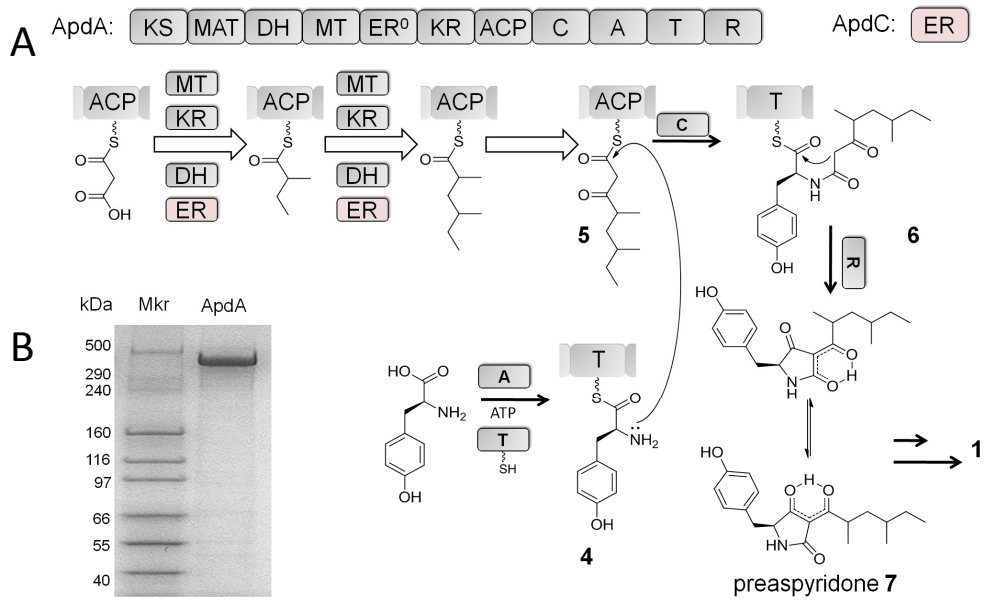



**Figure 31.** PKS-NRPS hybrid ApdA from BJ5464-NpgA (**A**) Proposed mechanism for the biosynthesis of preaspyridone 7. The phosphopantetheinyl arms of the ACP and T domains are shown as squiggle lines. The white arrow represents one round of chain elongation. **7** can be converted to the final natural product **1** after extra enzymatic oxidation and ring rearrangement.[171] (**B**) The SDS-PAGE of purified ApdA (439 kDa, tags included). The enzyme is purified from *S. cerevisiae* BJ5464-NpgA by an anti-FLAG antibody affinity resin.

**Structure elucidation of preaspyridone (7)**

To obtain sufficient amounts of **7** for structural elucidation, BJ5464-NpgA was cotransformed with expression plasmids (pXW58 and pXW51) for both ApdA and ApdC and cultured as previously described.[160],[184] The retention time, UV and mass spectra of **7** biosynthesized from yeast was identical to that of the in vitro sample **(Figure 33B**, trace ii). After 4 days, the culture was extracted and **7** was purified (final titer ~ 4 mg/L) and fully characterized by $^{1}H$, $^{13}C$ and correlation NMR spectroscopy.

Interestingly, both proton and carbon NMR spectra in $CDCl_3$ showed two sets of signals with a relative ratio of 3:1 and these two sets collapsed into one single set when the NMR solvent was switched to $CD_3OD$, which is characteristic for 3-acyltetramic acids. The two sets of signals in $CDCl_3$ arise from the external tautomers, which slowly interconvert through the rotation around the acyl-tetramic acid linkage (**Figure 32**).[185, 186] The collapse can be explained by the

inclination of external tautomeric equilibrium to one side as a result of solvent effect.[187] The proton spectrum showed characteristics of a typical AA'BB' spin system, which strongly indicated the presence of a *para*-substituted phenyl ring. COSY coupling between high field protons suggests a spin system of a branched aliphatic chain. Combining all the NMR information and in comparison to that of pretenellin-A[188] and **1**, **7** was determined as the 3-acyltetramic acid shown in **Figure 31A** and is named preaspyridone.

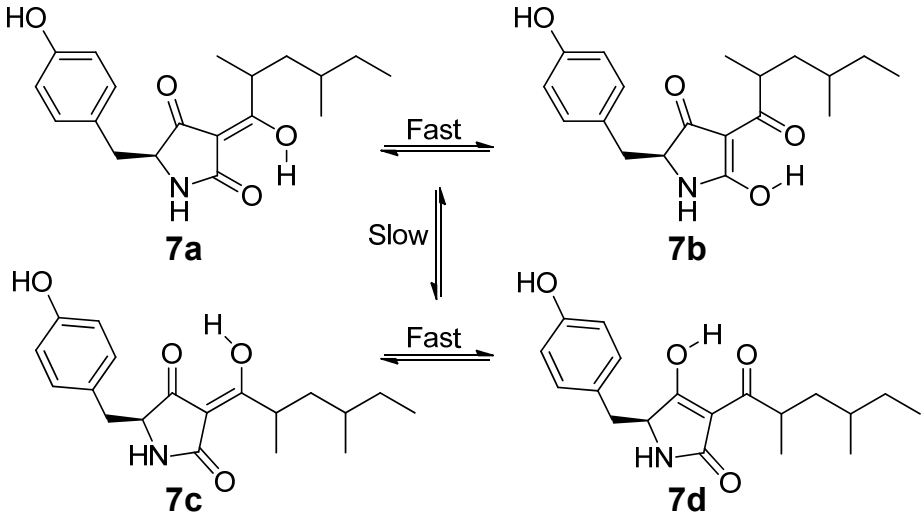



**Figure 32.** Major tautomers of 3-acyltetramic acid preaspyridone. Two sets of rapidly interchanging internal tautomers 7a-7b and 7c-7d are involved, where each set arises through proton transfer along the intramolecular hydrogen bond. The interconvertions of two pairs of external tautomers 7a-7c and 7b-7d, which is arisen from the rotation of the acyl side chain, are relatively slower. The internal tautomerizations are too rapid to be detected on the time scale of an NMR experiment but the external tautomerism can be measured.[185] Research showed that in $CDCl_3$, the tetramic acid tautomeric forms c-d lead over a-b through a ratio as around 77% to 23%, but a-b are the dominating forms in $CD_3OD$. [187]

Synthesis of **7** using purified ApdA and ApdC confirms the complete reconstitution of the activities of the recombinant PKS-NRPS megasynthetase. Under in vitro conditions, the assembly of **7** can be initiated by loading of malonyl-CoA onto ApdA, followed by decarboxylation to yield the acetyl starter unit. Under *in vivo* conditions, priming may also be initiated directly with an acetyl-CoA unit. The growing polyketide chain then elongates into the tetraketide **5** as shown in **Figure 31A** through differential tailoring of the α and β positions.

None of the tailoring domains were used in the last PKS elongation step, which affords the β-keto thioester **5** required for formation of the acyltetramic acid. The acyl chain of **5** is then amidated with the α-$NH_2$ of **4** to afford **6.** No other tetramic acid or amidated polyketide product was observed in the in vitro reaction in detectable amounts, indicating the programming rules of AdpA are precise and only the correctly tailored polyketide acyl intermediates are transferred from the PKS to the NRPS. The *in vitro* synthesis of **7** in the absence of additional oxidative enzymes also confirms the release of **6** via a R-domain catalyzed Dieckmann cyclization instead of the originally proposed NADPH-dependent reductive release followed by cyclization and enzymatic oxidation.[171] The Dieckmann cyclase property of ApdA R domain is therefore analogous to those of dissected R domains from the EqiS and CpaS fungal PKS-NRPSs.[189, 190]

**Incorporation of unnatural amino acid moiety to synthesize preaspyridone analog**

To probe the flexibility of the C and A domains towards different aromatic amino acids, we performed the *in vitro* reactions with various *para*-substituted phenylalanines and the synthesis of analogs of **7** was monitored by LC-MS (**Figure 33A**, trace iii, iv and v). ApdA was able to synthesize the corresponding analogs **8** and **9** in the presence of l-phenylalanine and l-4-fluorophenylalanine, respectively, at ~50% efficiency compared to **7**. Increasing the sizes of the *para* substituent (-$NH_2$, -Cl, -Br) led to decreases in product turnover. Surprisingly, adding l-tryptophan resulted in the synthesis of the indole-containing analog **10** (**Figure 33A**, trace v) as revealed by selected ion monitoring, albeit at significantly lower levels. Therefore, these results demonstrate while the NRPS module clearly prefers l-tyrosine in the *in vitro* assay, both the activation (A) and the condensing (C) domains have some flexibility in utilizing other aromatic amino acids and lead to formation of analogs of **7**.

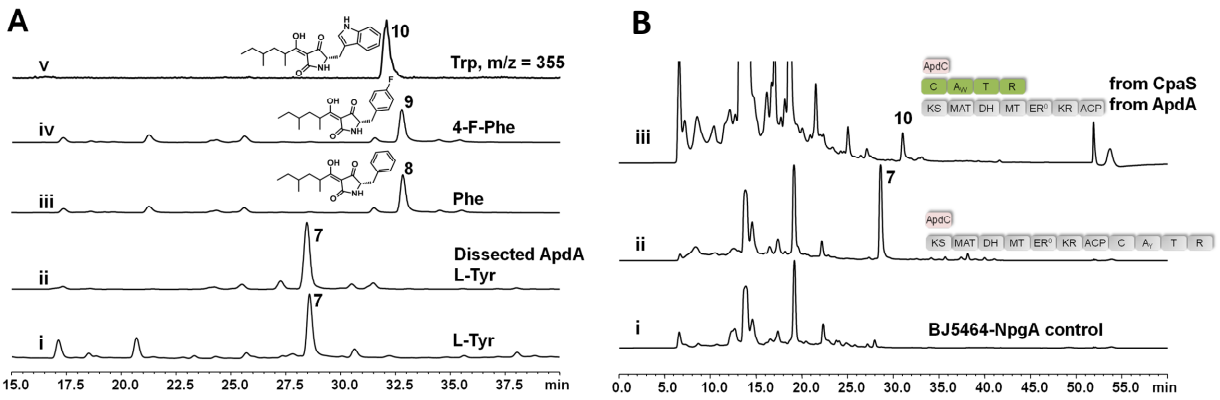
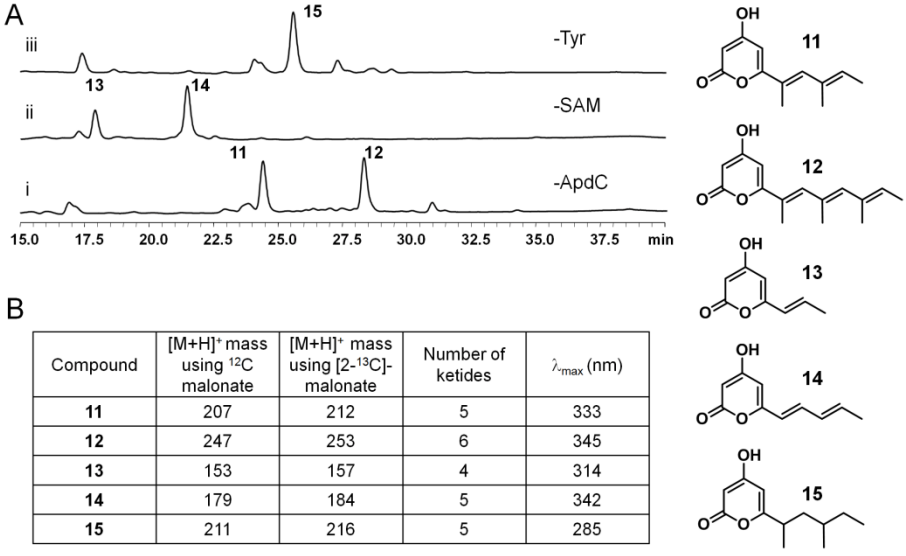



**Figure 33**. Reconstitution of ApdA *in vitro* and in *S. cerevisiae* (**A**) In vitro reconstitution of ApdA using purified enzymes. In each trace, the supplied L-amino acid (1 mM) is shown. All experiments were performed with intact ApdA except trace ii in which dissected PKS and NRPS modules of ApdA were reassembled. Trace **i-iv** are HPLC traces monitored at UV λ=280 nm; trace **v** is a selected ion monitoring trace $[M+H]^+$=355, which is produced at lower titers. (**B**) HPLC traces (280 nm) that shows the *in vivo* reconstitution of ApdA in BJ5464-NpgA. The cultures were extracted with ethyl acetate on the 5$^{th}$ day. Trace **i**: untransformed BJ5464 control; trace **ii**: BJ5464-NpgA transformed with ApdA. This culture was used to purify **7** for structural determination; trace **iii**: BJ5464-NpgA transformed with expression plasmids that separately encode the ApdA PKS and the CpaS NRPS modules. The synthesis of the expected analog **10** was observed in the extract (Trace **iii** is scaled differently to clearly show **10**).




B

| Compound | $[M+H]^+$ mass using $^{12}C$ malonate | $[M+H]^+$ mass using [2-$^{13}C$]-malonate | Number of ketides | $\lambda_{max}$ (nm) |
|---|---|---|---|---|
| **11** | 207 | 212 | 5 | 333 |
| **12** | 247 | 253 | 6 | 345 |
| **13** | 153 | 157 | 4 | 314 |
| **14** | 179 | 184 | 5 | 342 |
| **15** | 211 | 216 | 5 | 285 |

**Figure 34**. *In vitro* assays probing of the programming rules of the PKS-NRPS system. The proposed structures of these polyketides are listed on the right. (**A**) HPLC UV traces of that show the production of different polyketides through removal of the indicated components from the assay that synthesizes 7. Trace i and ii are monitored at λ=330 nm; trace iii is monitored at λ=280 nm. The relative retention times of compounds are in agreement with their relative polarities based on the proposed structures. (**B**) The masses of the mentioned polyketides. The increment in the mass when fed with [2-$^{13}$C]-malonate indicates the number of ketides for the corresponding compound. Conjugation onto the pyrone ring will cause the $\lambda_{max}$ to shift a longer wavelength.

## Probing the programming rule of HR-PKS portion of ApdA

The *in vitro* assay developed for ApdA similarly allows probing of the programming rules of the HR-PKS portion of the megasynthetase. Towards this end, we first examined whether removal of SAM in the *in vitro* reaction can lead to unmethylated versions of **7**. While no tetramic acid was observed, two unsaturated α-pyrones (**13** and **14**) were produced (**Figure 34A**, trace ii). **13** was previously recovered as a derailment product of CalE8[191] and **14** was isolated from an *in vitro* LovB assay in the absence of SAM, in which the ER LovC failed to function without *C*-methylation of a tetraketide intermediate.[62] Therefore, the synthesis of **13** and **14** here indicates *C*-methylation is also a prerequisite modification for enoylreduction by the ER partner ApdC. The incorrectly tailored polyketide intermediates were then offloaded by ApdA as **13** and **14**, analogous to that observed for LovB upon derailment in the tailoring steps. These convergent observations between ApdA/ApdC and LovB/LovC thus point to a conserved programming rule, in which a subtle α-methylation can modulate the interaction between HR-PKSs and their dissociated ER partners.

Similarly, removal of ApdC from the assay produced two α-pyrones **11** and **12** (**Figure 33A**, trace i). Repeating the assay with 2-$^{13}$C malonate resulted in increase of 5 and 6 mu for **11** and **12,** respectively (**Figure 34B**). We also observed increase in the $\lambda_{max}$ of the α-pyrones, which is indicative of increased levels of conjugation. These observations, combined with inference from biosynthetic logic and comparison to the LovB system[62],[61, 192], led us to propose the structures of **11** and **12** as shown in Figure 34, of which the two compounds contain two and three α-methyl modifications, respectively. The failure of ApdA to synthesize tetramic acids containing unsaturated polyketide portions is evident of the substrate specificity of the NRPS module towards incorrectly tailored polyketide acyl substrates. These stall products were instead offloaded as pyrones to afford **11** and **12.** Intriguingly, this result is different than that of TenS

reconstituted in *A. oryzae* in which the conjugated prototenellins were synthesized in the absence of the ER TenC.[188] This is most likely due to the more relaxed substrate tolerance of the downstream C domain of TenS.

**Interaction between PKS and NRPS module**

The formation of **11, 12** and **14** from the above assays points to an unexpected feature of ApdA, which is its ability to synthesize polyketide products longer than the tetraketide observed in **1** and **7**. To probe the chain length specificity of the HR-PKS, we sought to assay the PKS portion independent of the activities of the NRPS module. To do so, we i) excluded L-tyrosine from the assay that produced **7**; and ii) cloned and expressed the PKS portion as a stand-alone enzyme (**Figure 34B**) from BJ5464-NpgA (5 mg/L) and assayed with ApdC in the presence of malonyl-CoA, SAM and NADPH. In both assays, the same α-pyrone **15** was recovered from the reaction mixtures (**Figure 34A**, trace iii;). The size of **15** was confirmed to be a pentaketide and $\lambda_{max}$ of 285 suggests that the linear portion of the molecule is not conjugated and fully reduced, which is expected if all the tailoring domains functioned properly. Therefore the HRPKS module can indeed function correctly without the downstream NRPS module and synthesize a pentaketide. However, during the synthesis of **7,** in order to prevent continual elongation of the polyketide beyond the β-keto tetraketide and eventual offloading as **15**, the C domain of the NRPS must interact with the PKS accurately to capture **5** and condense with **4**.

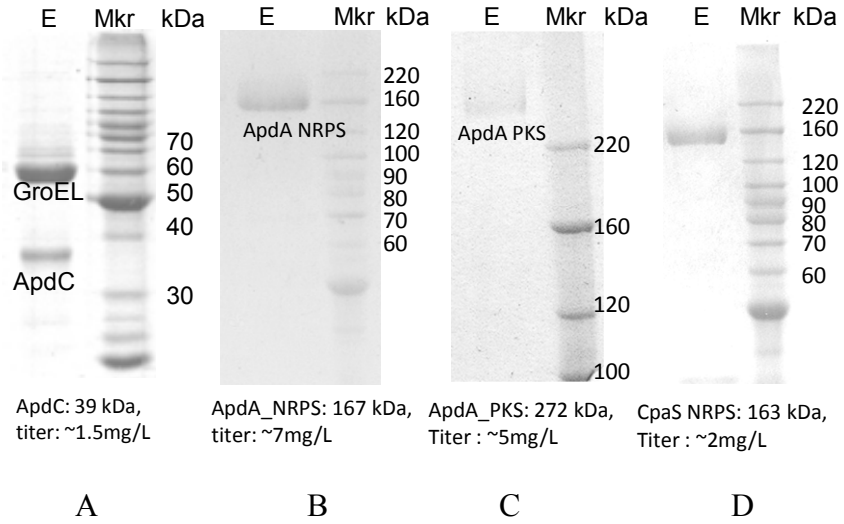



**Figure 35.** SDS PAGE of the heterogeneously expressed proteins in this study. All these enzymes are with C-terminal His-Tag and purified by Ni-NTA agarose affinity resin. (**A**) ApdC (39 kDa) that expressed from *Escherichia coli* BL21 (DE3). ApdC is coexpressed with chaperone proteins GroEL and GroES. The GroEL stays with ApdC but experiment results show that GroEL does not affect ApdC function. (**B**) The ApdA PKS module (272 kDa) from BJ5464-NpgA. (**C**) The ApdA NRPS module (167 kDa) from *E. coli* BAP1. D) The *A. flavus* CpaS NRPS module (163 kDa) from BJ5464-NpgA.

To examine if this critical interaction between the HRPKS and NRPS can be established *in trans*, we expressed and purified the NRPS module as a *C*-terminus hexahistidine tagged stand-alone enzyme from *E. coli* BAP1[193] (5 mg/L) (Figure 35C). Incubation of the dissected ApdA PKS and NRPS modules in equimolar amounts (25 μM), along with ApdC, cofactors and building blocks, produced **7** in comparable yield as the intact ApdA (**Figure 33A**, trace ii). This is the first demonstration of physically separating a fungal PKS-NRPS megasynthetase into the two independent, functional modules. Encouraged by the *in trans* complementation findings, we explored whether heterologous PKS and NRPS modules can be functionally matched, which was previously engineered for some bacterial PKS-NRPS hybrids.[194] We cloned and expressed the NRPS module of *Aspergillus flavus* CpaS (**Figure 34D**), which incorporates l-tryptophan into the tetramic acid intermediate of **2**.[180] When BJ5464-NpgA was cotransformed with expression plasmids that expressed ApdA PKS, CpaS NRPS and ApdC, small amounts of **10** (~0.1 mg/L) was observed from the extracts of the yeast culture (**Figure 33B**, trace iii), indicating successful communication between the two modules. The lower titer of **10** may be attributed to the substrate specificity of the CpaS C domain donor site [195] towards the noncognate ApdA ACP, as well as the bulkier **5** compared to the natural acetoacetyl thioester acyl donor. The lower turnover of **10** can also arise from the inability of the CpaS R domain to correctly process an unnatural substrate. Nevertheless, the result from the straightforward mix and match experiment shows the potential to use this dissociated platform to biosynthesize new tetramic acid compounds.

**Discussion:**

We reconstituted the complete activities of the PKS-NRPS ApdA through the synthesis of **7.** To the best of our knowledge, this represents the largest biosynthetic megasynthases

reconstituted *in vitro*. Using both *in vitro* assays and the yeast host, we were able to dissect the programming rules of ApdA, including the protein-protein interactions and key acyl transfer step between the PKS and NRPS modules.

**3. Conclusions**

A series of biochemical studies on the core megasynthetases from fungal secondary metabolism were performed for investigating the programming rules in fungal PKSs and PKS-NRPS hybrids.

First of all, we realized the importance of the product releasing enzyme to fungal HR-PKS. Specifically, LovD is responsible for the efficient acyl transfer on to monacolin J to complete the last step of lovastatin synthesis. Kinetics data shows that LovD is very specific to LovF protein. LovG, the thioesterase that releases product form LovB, increases the turnover rate for this megasynthase dramatically.

Secondly, we were able to reconstitute PKS-NRPS megasynthetase from yeast and probe their property or programming rules on the biosynthesis.

Moreover, with complete understanding of the gene functions in the pathway, we were able to use our expression system to manipulator multiple genes and reconstitute efficient biosynthesis of important natural product in the yeast host.

Lastly, our study also shows that protein-protein interaction in the polyketide or nonribosomal peptide biosynthesis is important and determines the rate of biosynthesis. In most case this interaction are specific, which suggest that detailed study on the factors for this recognition is meaning for bring more engineering.

## 4. References


1. Newman, D.J. and G.M. Cragg, *Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010.* Journal of Natural Products, 2012. **75**(3): p. 311-335.

2. Butler, M.S., *Natural products to drugs: natural product-derived compounds in clinical trials.* Natural Product Reports, 2008. **25**(3): p. 475-516.

3. Butler, M.S., *Natural products to drugs: natural product derived compounds in clinical trials.* Natural Product Reports, 2005. **22**(2): p. 162-95.

4. Newman, D.J. and G.M. Cragg, *Natural products as sources of new drugs over the last 25 years.* Journal of Natural Products, 2007. **70**(3): p. 461-77.

5. Facchini, P.J., *ALKALOID BIOSYNTHESIS IN PLANTS: Biochemistry, Cell Biology, Molecular Regulation, and Metabolic Engineering Applications.* Annual Review of Plant Physiology and Plant Molecular Biology, 2001. **52**: p. 29-66.

6. Ziegler, J. and P.J. Facchini, *Alkaloid biosynthesis: metabolism and trafficking.* Annual Review of Plant Biology, 2008. **59**: p. 735-69.

7. Furlan, A.D., et al., *Opioids for chronic noncancer pain: a meta-analysis of effectiveness and side effects.* Canadian Medical Association Journal, 2006. **174**(11): p. 1589-1594.

8. Rossokhin, A., et al., *Interaction of d-tubocurarine with potassium channels: Molecular modeling and ligand binding.* Molecular Pharmacology, 2006. **69**(4): p. 1356-1365.

9. Borisy, G.G. and E.W. Taylor, *Mechanism of Action of Colchicine - Binding of Colchincine-3h to Cellular Protein.* Journal of Cell Biology, 1967. **34**(2): p. 525-&.

10. Rowinsky, E.K. and R.C. Donehower, *The Clinical-Pharmacology and Use of Antimicrotubule Agents in Cancer Chemotherapeutics.* Pharmacology & Therapeutics, 1991. **52**(1): p. 35-84.

11. Schmeller, T., B. LatzBruning, and M. Wink, *Biochemical activities of berberine, palmatine and sanguinarine mediating chemical defence against microorganisms and herbivores.* Phytochemistry, 1997. **44**(2): p. 257-266.

12. Gao, Y., R.B. Honzatko, and R.J. Peters, *Terpenoid synthase structures: a so far incomplete view of complex catalysis.* Natural Product Reports, 2012. **29**(10): p. 1153-1175.

13. Rowinsky, E.K., L.A. Cazenave, and R.C. Donehower, *Taxol - a Novel Investigational Antimicrotubule Agent.* Journal of the National Cancer Institute, 1990. **82**(15): p. 1247-1259.

14. Christianson, D.W., *Structural biology and chemistry of the terpenoid cyclases.* Chemical Reviews, 2006. **106**(8): p. 3412-3442.

15. Walsh, C.T., *The chemical versatility of natural-product assembly lines.* Acc Chem Res, 2008. **41**(1): p. 4-10.

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

17. Khosla, C., *Structures and mechanisms of polyketide synthases.* J Org Chem, 2009. **74**(17): p. 6416-20.

18. Lambalot, R.H., et al., *A new enzyme superfamily - the phosphopantetheinyl transferases.* Chem Biol, 1996. **3**(11): p. 923-36.

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

20. Du, L. and L. Lou, *PKS and NRPS release mechanisms.* Nat Prod Rep, 2010. **27**(2): p. 255-78.

21. Weissman, K.J., *Introduction to Polyketide Biosynthesis.* Complex Enzymes in Microbial Natural Product Biosynthesis, Part B: Polyketides, Aminocoumarins and Carbohydrates, 2009. **459**: p. 3-16.

22. Jenke-Kodama, H., et al., *Evolutionary implications of bacterial polyketide synthases.* Mol Biol Evol, 2005. **22**(10): p. 2027-39.

23. Kroken, S., et al., *Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(26): p. 15670-5.

24. Cox, R.J., *Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets.* Organic and Biomolecular Chemistry, 2007. **5**(13): p. 2010-26.

25. Meier, J.L. and M.D. Burkart, *The chemical biology of modular biosynthetic enzymes.* Chem Soc Rev, 2009. **38**(7): p. 2012-45.

26. Cane, D.E. and C.T. Walsh, *The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases.* Chem Biol, 1999. **6**(12): p. R319-25.

27. 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. 3468-96.

28. McDaniel, R., et al., *Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(5): p. 1846-1851.

29. Smith, D.J., A.J. Earl, and G. Turner, *The multifunctional peptide synthetase performing the first step of penicillin biosynthesis in Penicillium chrysogenum is a 421,073 dalton protein similar to Bacillus brevis peptide antibiotic synthetases.* EMBO Journal, 1990. **9**(9): p. 2743-50.

30. Merkel, A.B., et al., *The position of a key tyrosine in dTDP-4-Keto-6-deoxy-D-glucose-5-epimerase (EvaD) alters the substrate profile for this RmlC-like enzyme.* Journal of Biological Chemistry, 2004. **279**(31): p. 32684-32691.

31. Oberthur, M., et al., *A systematic investigation of the synthetic utility of glycopeptide glycosyltransferases.* Journal of the American Chemical Society, 2005. **127**(30): p. 10747-10752.

32. Gehring, A.M., K.A. Bradley, and C.T. Walsh, *Enterobactin biosynthesis in Escherichia coli: isochorismate lyase (EntB) is a bifunctional enzyme that is phosphopantetheinylated by EntD and then acylated by EntE using ATP and 2,3-dihydroxybenzoate.* Biochemistry, 1997. **36**(28): p. 8495-503.

33. Gehring, A.M., I. Mori, and C.T. Walsh, *Reconstitution and characterization of the Escherichia coli enterobactin synthetase from EntB, EntE, and EntF.* Biochemistry, 1998. **37**(8): p. 2648-59.

34. Weber, G., et al., *The Peptide Synthetase Catalyzing Cyclosporine Production in Tolypocladium-Niveum Is Encoded by a Giant 45.8-Kilobase Open Reading Frame.* Current Genetics, 1994. **26**(2): p. 120-125.

35. Schwarzer, D., R. Finking, and M.A. Marahiel, *Nonribosomal peptides: from genes to products.* Natural Product Reports, 2003. **20**(3): p. 275-287.

36. Pavela-Vrancic, M., et al., *Editing of non-cognate aminoacyl adenylates by peptide synthetases.* Biochemical Journal, 1999. **342**: p. 715-719.

37. Doekel, S. and M.A. Marahiel, *Biosynthesis of natural products on modular peptide synthetases.* Metabolic Engineering, 2001. **3**(1): p. 64-77.

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

39. Demain, A.L. and J.L. Adrio, *Contributions of microorganisms to industrial biology.* Mol. Biotechnol., 2008. **38**: p. 41-55.

40. Stephanopoulos, G., J. Nielsen, and A. Aristodou, *Metabolic Engineering*. 1998, San Diego: Academic Press.

41. Challis, G.L. and D.A. Hopwood, *Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by Streptomyces species.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**: p. 14555-14561.

42. Sanchez, J.F., et al., *Advances in Aspergillus secondary metabolite research in the post-genomic era.* Natural Product Reports, 2012. **29**(3): p. 351-371.

43. Brakhage, A.A. and V. Schroeckh, *Fungal secondary metabolites - Strategies to activate silent gene clusters.* Fungal Genetics and Biology, 2011. **48**(1): p. 15-22.

44. Chiang, Y.M., et al., *A Gene Cluster Containing Two Fungal Polyketide Synthases Encodes the Biosynthetic Pathway for a Polyketide, Asperfuranone, in Aspergillus nidulans.* Journal of the American Chemical Society, 2009. **131**(8): p. 2965-2970.

45. Kuck, U. and B. Hoff, *New tools for the genetic manipulation of filamentous fungi.* Appl Microbiol Biotechnol, 2010. **86**(1): p. 51-62.

46. Gelvin, S.B., *Agrobacterium-mediated plant transformation: the biology behind the "gene-jockeying" tool.* Microbiol Mol Biol Rev, 2003. **67**(1): p. 16-37, table of contents.

47. Mata, M.M., et al., *Agrobacterium-mediated insertional mutagenesis of the ochratoxigenic fungus Aspergillus westerdijkiae.* Can J Microbiol, 2007. **53**(1): p. 148-51.

48. Sugui, J.A., Y.C. Chang, and K.J. Kwon-Chung, *Agrobacterium tumefaciens-mediated transformation of Aspergillus fumigatus: an efficient tool for insertional mutagenesis and targeted gene disruption.* Appl Environ Microbiol, 2005. **71**(4): p. 1798-802.

49. Nakayashiki, H., et al., *RNA silencing as a tool for exploring gene function in ascomycete fungi.* Fungal Genetics and Biology, 2005. **42**(4): p. 275-283.

50. Schumann, J. and C. Hertweck, *Molecular basis of cytochalasan biosynthesis in fungi: Gene cluster analysis and evidence for the involvement of a PKS-NRPS hybrid synthase by RNA silencing.* Journal of the American Chemical Society, 2007. **129**(31): p. 9564-+.

51. Pfeifer, B.A., et al., *Biosynthesis of complex polyketides in a metabolically engineered strain of E-coli.* Science, 2001. **291**(5509): p. 1790-1792.

52. Watanabe, K., et al., *Engineered biosynthesis of an ansamycin polyketide precursor in Escherichia coli.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(17): p. 9774-9778.

53. Peiru, S., et al., *Production of the potent antibacterial polyketide erythromycin C in Escherichia coli.* Appl Environ Microbiol, 2005. **71**(5): p. 2539-2547.

54. Gruenewald, S., et al., *In vivo production of artificial nonribosomal peptide products in the heterologous host Escherichia coli.* Appl Environ Microbiol, 2004. **70**(6): p. 3282-3291.

55. Watanabe, K., et al., *Total biosynthesis of antitumor nonribosomal peptides in Escherichia coli.* Nature Chemical Biology, 2006. **2**(8): p. 423-428.

56. Lee, K.K., N.A. Da Silva, and J.T. Kealey, *Determination of the extent of phosphopantetheinylation of polyketide synthases expressed in Escherichia coli and Saccharomyces cerevisiae.* Anal Biochem, 2009. **394**(1): p. 75-80.

57. Bailey, A.M., et al., *Characterisation of 3-methylorcinaldehyde synthase (MOS) in Acremonium strictum: first observation of a reductive release mechanism during polyketide biosynthesis.* Chemical Communications, 2007(39): p. 4053-4055.

58. Sakai, K., et al., *Construction of a Citrinin Gene Cluster Expression System in Heterologous Aspergillus oryzae.* Journal of Bioscience and Bioengineering, 2008. **106**(5): p. 466-472.

59. Shen, B., *Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms.* Current Opinion in Chemical Biology, 2003. **7**(2): p. 285-295.

60. Auclair, K., et al., *Lovastatin nonaketide synthase catalyzes an intramolecular Diels-Alder reaction of a substrate analogue.* Journal of the American Chemical Society, 2000. **122**(46): p. 11519-11520.

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

62. Ma, S.M., et al., *Complete reconstitution of a highly reducing iterative polyketide synthase.* Science, 2009. **326**(5952): p. 589-92.

63. Xie, X., et al., *Acyltransferase mediated polyketide release from a fungal megasynthase.* Journal of the American Chemical Society, 2009. **131**(24): p. 8388-9.

64. Smith, L., et al., *Analysis of specific mutants in the lasalocid gene cluster: evidence for enzymatic catalysis of a disfavoured polyether ring closure.* ChemBioChem, 2008. **9**(18): p. 2967-75.

65. Tang, Y., et al., *The 2.7-Angstrom crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(30): p. 11124-9.

66. Tang, Y., et al., *Structural and mechanistic analysis of protein interactions in module 3 of the 6-deoxyerythronolide B synthase.* Chem Biol, 2007. **14**(8): p. 931-43.

67. Keatinge-Clay, A.T. and R.M. Stroud, *The structure of a ketoreductase determines the organization of the beta-carbon processing enzymes of modular polyketide synthases.* Structure, 2006. **14**(4): p. 737-48.

68. Keatinge-Clay, A.T., *A tylosin ketoreductase reveals how chirality is determined in polyketides.* Chem Biol, 2007. **14**(8): p. 898-908.

69. Zheng, J., et al., *Structural and functional analysis of A-type ketoreductases from the amphotericin modular polyketide synthase.* Structure, 2010. **18**(8): p. 913-22.

70. Zheng, J. and A.T. Keatinge-Clay, *Structural and functional analysis of C2-type ketoreductases from modular polyketide synthases.* J Mol Biol, 2011. **410**(1): p. 105-17.

71. Keatinge-Clay, A., *Crystal structure of the erythromycin polyketide synthase dehydratase.* J Mol Biol, 2008. **384**(4): p. 941-53.

72. Akey, D.L., et al., *Crystal Structures of Dehydratase Domains from the Curacin Polyketide Biosynthetic Pathway.* Structure, 2010. **18**(1): p. 94-105.

73. Zheng, J., et al., *Divergence of multimodular polyketide synthases revealed by a didomain structure.* Nature Chemical Biology, 2012. **8**: p. 615-621.

74. Tsai, S.C., et al., *Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: versatility from a unique substrate channel.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(26): p. 14808-13.

75. Tsai, S.C., et al., *Insights into channel architecture and substrate specificity from crystal structures of two macrocycle-forming thioesterases of modular polyketide synthases.* Biochemistry, 2002. **41**(42): p. 12598-12606.

76. Akey, D.L., et al., *Structural basis for macrolactonization by the pikromycin thioesterase.* Nat Chem Biol, 2006. **2**(10): p. 537-42.

77. Weissman, K.J. and R. Muller, *Protein-protein interactions in multienzyme megasynthetases.* Chembiochem, 2008. **9**(6): p. 826-48.

78. Tsai, S.C. and B.D. Ames, *Structural enzymology of polyketide synthases.* Methods Enzymol, 2009. **459**: p. 17-47.

79. Smith, S. and S.C. Tsai, *The type I fatty acid and polyketide synthases: a tale of two megasynthases.* Nat Prod Rep, 2007. **24**(5): p. 1041-72.

80. Keatinge-Clay, A.T., *The structures of type I polyketide synthases.* Natural Product Reports, 2012: p. DOI: 10.1039/C2NP20019H

81. Maier, T., M. Leibundgut, and N. Ban, *The crystal structure of a mammalian fatty acid synthase.* Science, 2008. **321**(5894): p. 1315-22.

82. Maier, T., S. Jenni, and N. Ban, *Architecture of mammalian fatty acid synthase at 4.5 A resolution.* Science, 2006. **311**(5765): p. 1258-62.

83. Witkowski, A., A. Joshi, and S. Smith, *Fatty acid synthase: In vitro complementation of inactive mutants.* Biochemistry, 1996. **35**(32): p. 10569-10575.

84. Joshi, A.K., A. Witkowski, and S. Smith, *Mapping of functional interactions between domains of the animal fatty acid synthase by mutant complementation in vitro.* Biochemistry, 1997. **36**(8): p. 2316-2322.

85. Joshi, A.K., A. Witkowski, and S. Smith, *The malonyl/acetyltransferase and beta-ketoacyl synthase domains of the animal fatty acid synthase can cooperate with the acyl carrier protein domain of either subunit.* Biochemistry, 1998. **37**(8): p. 2515-2523.

86. Joshi, A.K., V.S. Rangan, and S. Smith, *Differential affinity labeling of the two subunits of the homodimeric animal fatty acid synthase allows isolation of heterodimers consisting of subunits that have been independently modified.* J Biol Chem, 1998. **273**(9): p. 4937-43.

87. Joshi, A.K., et al., *Engineering of an active animal fatty acid synthase dimer with only one competent subunit.* Chemistry & Biology, 2003. **10**(2): p. 169-173.

88. Rangan, V.S., A.K. Joshi, and S. Smith, *Mapping the functional topology of the animal fatty acid synthase by mutant complementation in vitro.* Biochemistry, 2001. **40**(36): p. 10792-10799.

89. Witkowski, A., et al., *Dibromopropanone cross-linking of the phosphopantetheine and active-site cysteine thiols of the animal fatty acid synthase can occur both inter- and intrasubunit - Reevaluation of the side-by-side, antiparallel subunit model.* Journal of Biological Chemistry, 1999. **274**(17): p. 11557-11563.

90. Witkowski, A., et al., *Head-to-head coiled arrangement of the subunits of the animal fatty acid synthase.* Chemistry and Biology, 2004. **11**(12): p. 1667-76.

91. Moche, M., et al., *Structure of the complex between the antibiotic cerulenin and its target, beta-ketoacyl-acyl carrier protein synthase.* Journal of Biological Chemistry, 1999. **274**(10): p. 6031-6034.

92. Price, A.C., C.O. Rock, and S.W. White, *The 1.3-Angstrom-resolution crystal structure of beta-ketoacyl-acyl carrier protein synthase II from Streptococcus pneumoniae.* Journal of Bacteriology, 2003. **185**(14): p. 4136-4143.

93. Asturias, F.J., et al., *Structure and molecular organization of mammalian fatty acid synthase.* Nature Structural & Molecular Biology, 2005. **12**(3): p. 225-232.

94. Olsen, J.G., et al., *Structures of beta-ketoacyl-acyl carrier protein synthase I complexed with fatty acids elucidate its catalytic machinery.* Structure, 2001. **9**(3): p. 233-43.

95. Leesong, M., et al., *Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: two catalytic activities in one active site.* Structure, 1996. **4**(3): p. 253-64.

96. Kimber, M.S., et al., *The structure of (3R)-hydroxyacyl-acyl carrier protein dehydratase (FabZ) from Pseudomonas aeruginosa.* J Biol Chem, 2004. **279**(50): p. 52593-602.

97. Price, A.C., et al., *Structure of beta-ketoacyl-[acyl carrier protein] reductase from Escherichia coli: negative cooperativity and its structural basis.* Biochemistry, 2001. **40**(43): p. 12772-81.

98. Keatinge-Clay, A.T., et al., *Catalysis, specificity, and ACP docking site of Streptomyces coelicolor malonyl-CoA:ACP transacylase.* Structure, 2003. **11**(2): p. 147-54.

99. Shimomura, Y., Y. Kakuta, and K. Fukuyama, *Crystal structures of the quinone oxidoreductase from Thermus thermophilus HB8 and its complex with NADPH: implication for NADPH and substrate recognition.* J Bacteriol, 2003. **185**(14): p. 4211-8.

100. Chakravarty, B., et al., *Human fatty acid synthase: structure and substrate selectivity of the thioesterase domain.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(44): p. 15567-72.

101. Reed, M.A.C., et al., *The type I rat fatty acid synthase ACP shows structural homology and analogous biochemical properties to type II ACPs.* Organic & Biomolecular Chemistry, 2003. **1**(3): p. 463-471.

102. Brignole, E.J., S. Smith, and F.J. Asturias, *Conformational flexibility of metazoan fatty acid synthase enables catalysis.* Nat Struct Mol Biol, 2009. **16**(2): p. 190-7.

103. Crawford, J.M., et al., *Identification of a starter unit acyl-carrier protein transacylase domain in an iterative type I polyketide synthase.* Proc Natl Acad Sci U S A, 2006. **103**(45): p. 16728-33.

104. Crawford, J.M., et al., *Structural basis for biosynthetic programming of fungal aromatic polyketide cyclization.* Nature, 2009. **461**(7267): p. 1139-43.

105. Crawford, J.M., et al., *Deconstruction of iterative multidomain polyketide synthase function.* Science, 2008. **320**(5873): p. 243-6.

106. Li, Y., W. Xu, and Y. Tang, *Classification, Prediction, and Verification of the Regioselectivity of Fungal Polyketide Synthase Product Template Domains.* Journal of Biological Chemistry, 2010. **285**(30): p. 22762-22771.

107. Korman, T.P., et al., *Structure and function of an iterative polyketide synthase thioesterase domain catalyzing Claisen cyclization in aflatoxin biosynthesis.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(14): p. 6246-51.

108. Fujii, I., et al., *Identification of Claisen cyclase domain in fungal polyketide synthase WA, a naphthopyrone synthase of Aspergillus nidulans.* Chem Biol, 2001. **8**(2): p. 189-97.

109. Ma, S.M., et al., *Redirecting the cyclization steps of fungal polyketide synthase.* J Am Chem Soc, 2008. **130**(1): p. 38-9.

110. Koglin, A., et al., *Structural basis for the selectivity of the external thioesterase of the surfactin synthetase.* Nature, 2008. **454**(7206): p. 907-11.

111. Frueh, D.P., et al., *Dynamic thiolation-thioesterase structure of a non-ribosomal peptide synthetase.* Nature, 2008. **454**(7206): p. 903-6.

112. Beck, J., et al., *The multifunctional 6-methylsalicylic acid synthase gene of Penicillium patulum. Its gene structure relative to that of other polyketide synthases.* Eur J Biochem, 1990. **192**(2): p. 487-98.

113. Alberts, A.W., et al., *Mevinolin: a highly potent competitive inhibitor of hydroxymethylglutaryl-coenzyme A reductase and a cholesterol-lowering agent.* Proc Natl Acad Sci U S A, 1980. **77**(7): p. 3957-61.

114. Istvan, E.S. and J. Deisenhofer, *Structural mechanism for statin inhibition of HMG-CoA reductase.* Science, 2001. **292**(5519): p. 1160-4.

115. Hendrickson, L., et al., *Lovastatin biosynthesis in Aspergillus terreus: characterization of blocked mutants, enzyme activities and a multifunctional polyketide synthase gene.* Chem Biol, 1999. **6**(7): p. 429-39.

116. Xie, X., et al., *Acyltransferase mediated polyketide release from a fungal megasynthase.* J Am Chem Soc, 2009. **131**(24): p. 8388-9.

117. Barriuso, J., et al., *Double oxidation of the cyclic nonaketide dihydromonacolin L to monacolin J by a single cytochrome P450 monooxygenase, LovA.* J Am Chem Soc, 2011. **133**(21): p. 8078-81.

118. Xie, X., et al., *Biosynthesis of lovastatin analogs with a broadly specific acyltransferase.* Chem Biol, 2006. **13**(11): p. 1161-9.

119. Cox, R.J., *Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets.* Organic & Biomolecular Chemistry, 2007. **5**(13): p. 2010-2026.

120. Chooi, Y.H. and Y. Tang, *Navigating the fungal polyketide chemical space: from genes to molecules.* Journal of Organic Chemistry, 2012. **77**(22): p. 9933-53.

121. Cox, R.J., et al., *Rapid cloning and expression of a fungal polyketide synthase gene involved in squalestatin biosynthesis.* Chemical Communications, 2004(20): p. 2260-1.

122. Gerber, R., L. Lou, and L. Du, *A PLP-dependent polyketide chain releasing mechanism in the biosynthesis of mycotoxin fumonisins in Fusarium verticillioides.* Journal of the American Chemical Society, 2009. **131**(9): p. 3148-9.

123. Akey, D.L., et al., *Structural basis for macrolactonization by the pikromycin thioesterase.* Nature Chemical Biology, 2006. **2**(10): p. 537-42.

124. Boddy, C.N., et al., *Epothilone C macrolactonization and hydrolysis are catalyzed by the isolated thioesterase domain of epothilone polyketide synthase.* Journal of the American Chemical Society, 2003. **125**(12): p. 3428-9.

125. Fujii, I., et al., *Identification of Claisen cyclase domain in fungal polyketide synthase WA, a naphthopyrone synthase of Aspergillus nidulans.* Chemistry & Biology, 2001. **8**(2): p. 189-97.

126. Ma, S.M., et al., *Redirecting the cyclization steps of fungal polyketide synthase.* Journal of the American Chemical Society, 2008. **130**(1): p. 38-9.

127. Moldenhauer, J., et al., *Biosynthesis of the antibiotic bacillaene, the product of a giant polyketide synthase complex of the trans-AT family.* Angewandte Chemie. International Edition in English, 2007. **46**(43): p. 8195-7.

128. Strieter, E.R., et al., *Cascade reactions during coronafacic acid biosynthesis: elongation, cyclization, and functionalization during Cfa7-catalyzed condensation.* Journal of the American Chemical Society, 2009. **131**(6): p. 2113-5.

129. Xie, X., et al., *Biosynthesis of lovastatin analogs with a broadly specific acyltransferase.* Chemistry & Biology, 2006. **13**(11): p. 1161-9.

130. Meehan, M.J., et al., *FT-ICR-MS Characterization of Intermediates in the Biosynthesis of the alpha-Methylbutyrate Side Chain of Lovastatin by the 277 kDa Polyketide Synthase LovF.* Biochemistry, 2011. **50**(2): p. 287-299.

131. Meluzzi, D., et al., *Top-down mass spectrometry on low-resolution instruments: Characterization of phosphopantetheinylated carrier domains in polyketide and non-ribosomal biosynthetic pathways.* Bioorganic & Medicinal Chemistry Letters, 2008. **18**(10): p. 3107-3111.

132. Wu, N., et al., *Assessing the balance between protein-protein interactions and enzyme-substrate interactions in the channeling of intermediates between polyketide synthase modules.* Journal of the American Chemical Society, 2001. **123**(27): p. 6465-6474.

133. Quadri, L.E.N., et al., *Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases.* Biochemistry, 1998. **37**(6): p. 1585-1595.

134. Quadri, L.E., et al., *Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases.* Biochemistry, 1998. **37**(6): p. 1585-95.

135. Kapur, S., et al., *Molecular recognition between ketosynthase and acyl carrier protein domains of the 6-deoxyerythronolide B synthase.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(51): p. 22066-22071.

136. Kapur, S., et al., *Reprogramming a module of the 6-deoxyerythronolide B synthase for iterative chain elongation.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(11): p. 4110-5.

137. Alekseyev, V.Y., et al., *Solution structure and proposed domain domain recognition interface of an acyl carrier protein domain from a modular polyketide synthase.* Protein Sci, 2007. **16**(10): p. 2093-107.

138. Tran, L., et al., *Insights into protein-protein and enzyme-substrate interactions in modular polyketide synthases.* Chem Biol, 2010. **17**(7): p. 705-16.

139. Lim, J., et al., *Solution structures of the acyl carrier protein domain from the highly reducing type I iterative polyketide synthase CalE8.* PLoS One, 2011. **6**(6): p. e20549.

140. Busche, A., et al., *Characterization of molecular interactions between ACP and halogenase domains in the Curacin A polyketide synthase.* ACS Chem Biol, 2012. **7**(2): p. 378-86.

141. Crump, M.P., et al., *Solution structure of the actinorhodin polyketide synthase acyl carrier protein from Streptomyces coelicolor A3(2).* Biochemistry, 1997. **36**(20): p. 6000-8.

142. Evans, S.E., et al., *An ACP structural switch: conformational differences between the apo and holo forms of the actinorhodin polyketide synthase acyl carrier protein.* Chembiochem, 2008. **9**(15): p. 2424-32.

143. Evans, S.E., et al., *Probing the Interactions of early polyketide intermediates with the Actinorhodin ACP from S. coelicolor A3(2).* J Mol Biol, 2009. **389**(3): p. 511-28.

144. Haushalter, R.W., et al., *Binding and pK(a) Modulation of a Polycyclic Substrate Analogue in a Type II Polyketide Acyl Carrier Protein.* Acs Chemical Biology, 2011. **6**(5): p. 413-418.

145. Li, Q., et al., *Solution structure and backbone dynamics of the holo form of the frenolicin acyl carrier protein.* Biochemistry, 2003. **42**(16): p. 4648-57.

146. Findlow, S.C., et al., *Solution structure and dynamics of oxytetracycline polyketide synthase acyl carrier protein from Streptomyces rimosus.* Biochemistry, 2003. **42**(28): p. 8423-33.

147. Wattana-amorn, P., et al., *Solution structure of an acyl carrier protein domain from a fungal type I polyketide synthase.* Biochemistry, 2010. **49**(10): p. 2186-93.

148. Sharma, A.K., et al., *Solution structures of conformationally equilibrium forms of holo-acyl carrier protein (PfACP) from Plasmodium falciparum provides insight into the mechanism of activation of ACPs.* Biochemistry, 2006. **45**(22): p. 6904-16.

149. Cantu, D.C., et al., *Acyl carrier protein structural classification and normal mode analysis.* Protein Science, 2012. **21**(5): p. 655-66.

150. Zhang, Y.M., et al., *Key residues responsible for acyl carrier protein and beta-ketoacyl-acyl carrier protein reductase (FabG) interaction.* J Biol Chem, 2003. **278**(52): p. 52935-43.

151. Weissman, K.J., et al., *Evidence for a protein-protein interaction motif on an acyl carrier protein domain from a modular polyketide synthase.* Chem Biol, 2006. **13**(6): p. 625-36.

152. Chen, A.Y., et al., *Extender unit and acyl carrier protein specificity of ketosynthase domains of the 6-deoxyerythronolide B synthase.* J Am Chem Soc, 2006. **128**(9): p. 3067-74.

153. Jones, E.W., *Tackling the protease problem in Saccharomyces cerevisiae.* Methods in Enzymology, 1991. **194**: p. 428-53.

154. Szewczyk, E., et al., *Fusion PCR and gene targeting in Aspergillus nidulans.* Nat Protoc, 2006. **1**(6): p. 3111-20.

155. Punt, P.J., I.E. Mattern, and C.A.M.J.J. van den Hondel, *A vector for Aspergillus transformation conferring phleomycin resistance.* Fungal Genet. Newsl., 1988. **35**: p. 25-30.

156. Chooi, Y.H., R. Cacho, and Y. Tang, *Identification of the viridicatumtoxin and griseofulvin gene clusters from Penicillium aethiopicum.* Chem Biol, 2010. **17**(5): p. 483-94.

157. Cove, D.J., *The induction and repression of nitrate reductase in the fungus Aspergillus nidulans.* Biochim Biophys Acta, 1966. **113**(1): p. 51-6.

158. An, J.H. and Y.S. Kim, *A gene cluster encoding malonyl-CoA decarboxylase (MatA), malonyl-CoA synthetase (MatB) and a putative dicarboxylate carrier protein (MatC) in Rhizobium trifolii--cloning, sequencing, and expression of the enzymes in Escherichia coli.* Eur J Biochem, 1998. **257**(2): p. 395-402.

159. Wang, M., et al., *A Thioesterase from an Iterative Fungal Polyketide Synthase Shows Macrocyclization and Cross Coupling Activity and May Play a Role in Controlling Iterative Cycling through Product Offloading.* Biochemistry, 2009. **48**(27): p. 6288-6290.

160. Zhou, H., et al., *Enzymatic synthesis of resorcylic acid lactones by cooperation of fungal iterative polyketide synthases involved in hypothemycin biosynthesis.* Journal of the American Chemical Society, 2010. **132**(13): p. 4530-1.

161. Auclair, K., et al., *Conversion of cyclic nonaketides to lovastatin and compactin by a lovC deficient mutant of Aspergillus terreus.* Bioorg Med Chem Lett, 2001. **11**(12): p. 1527-31.

162. Abe, Y., et al., *Molecular cloning and characterization of an ML-236B (compactin) biosynthetic gene cluster in Penicillium citrinum.* Mol Genet Genomics, 2002. **267**(5): p. 636-46.

163. Chen, Y.P., et al., *Identification of the mokH gene encoding transcription factor for the upregulation of monacolin K biosynthesis in Monascus pilosus.* J Agric Food Chem, 2010. **58**(1): p. 287-93.

164. Boettger, D., et al., *Evolutionary Imprint of Catalytic Domains in Fungal PKS-NRPS Hybrids.* Chembiochem, 2012. **13**(16): p. 2363-73.

165. Nishihara, K., et al., *Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in Escherichia coli.* Appl Environ Microbiol, 1998. **64**(5): p. 1694-9.

166. Vagstad, A.L., et al., *Interrogation of global active site occupancy of a fungal iterative polyketide synthase reveals strategies for maintaining biosynthetic fidelity.* J Am Chem Soc, 2012. **134**(15): p. 6865-77.

167. Newman, A.G., et al., *Analysis of the cercosporin polyketide synthase CTB1 reveals a new fungal thioesterase function.* Chem Commun (Camb), 2012. **48**(96): p. 11772-4.

168. Staunton, J. and K.J. Weissman, *Polyketide biosynthesis: a millennium review.* Natural Product Reports, 2001. **18**(4): p. 380-416.

169. Song, Z., et al., *Fusarin C biosynthesis in Fusarium moniliforme and Fusarium venenatum.* ChemBioChem, 2004. **5**(9): p. 1196-1203.

170. Sims, J.W., et al., *Equisetin biosynthesis in Fusarium heterosporum.* Chem Commun (Camb), 2005(2): p. 186-8.

171. Bergmann, S., et al., *Genomics-driven discovery of PKS-NRPS hybrid metabolites from Aspergillus nidulans.* Nat Chem Biol, 2007. **3**(4): p. 213-7.

172. Maiya, S., et al., *Identification of a hybrid PKS/NRPS required for pseurotin A biosynthesis in the human pathogen Aspergillus fumigatus.* ChemBioChem, 2007. **8**(14): p. 1736-43.

173. Tokuoka, M., et al., *Identification of a novel polyketide synthase-nonribosomal peptide synthetase (PKS-NRPS) gene required for the biosynthesis of cyclopiazonic acid in Aspergillus oryzae.* Fungal Genetics and Biology, 2008. **45**(12): p. 1608-15.

174. Chang, P.K., B.W. Horn, and J.W. Dorner, *Clustered genes involved in cyclopiazonic acid production are next to the aflatoxin biosynthesis gene cluster in Aspergillus flavus.* Fungal Genetics and Biology, 2009. **46**(2): p. 176-82.

175. Schümann, J. and C. Hertweck, *Molecular basis of cytochalasan biosynthesis in fungi: gene cluster analysis and evidence for the involvement of a PKS-NRPS hybrid synthase by RNA silencing.* J Am Chem Soc, 2007. **129**(31): p. 9564-5.

176. Eley, K.L., et al., *Biosynthesis of the 2-pyridone tenellin in the insect pathogenic fungus Beauveria bassiana.* Chembiochem, 2007. **8**(3): p. 289-97.

177. von Dohren, H., *A survey of nonribosomal peptide synthetase (NRPS) genes in Aspergillus nidulans.* Fungal Genetics and Biology, 2009. **46 Suppl 1**: p. S45-52.

178. Collemare, J., et al., *Biosynthesis of secondary metabolites in the rice blast fungus Magnaporthe grisea: the role of hybrid PKS-NRPS in pathogenicity.* Mycol Res, 2008. **112**(Pt 2): p. 207-15.

179. Heneghan, M.N., et al., *First Heterologous Reconstruction of a Complete Functional Fungal Biosynthetic Multigene Cluster.* ChemBioChem, 2010.

180. Seshime, Y., et al., *Functional expression of the Aspergillus flavus PKS-NRPS hybrid CpaA involved in the biosynthesis of cyclopiazonic acid.* Bioorg Med Chem Lett, 2009. **19**(12): p. 3288-92.

181. Sattely, E.S., M.A. Fischbach, and C.T. Walsh, *Total biosynthesis: in vitro reconstitution of polyketide and nonribosomal peptide pathways.* Natural Product Reports, 2008. **25**(4): p. 757-93.

182. Mutka, S.C., et al., *Metabolic pathway engineering for complex polyketide biosynthesis in Saccharomyces cerevisiae.* FEMS Yeast Res, 2006. **6**(1): p. 40-7.

183. Lee, K.M. and N.A. DaSilva, *Evaluation of the Saccharomyces cerevisiae ADH2 promoter for protein synthesis.* Yeast, 2005. **22**(6): p. 431-40.

184. Reeves, C.D., et al., *Genes for the biosynthesis of the fungal polyketides hypothemycin from Hypomyces subiculosus and radicicol from Pochonia chlamydosporia.* Applied and Environment Microbiology, 2008. **74**(16): p. 5121-9.

185. Royles, B.J.L., *Naturally-Occurring Tetramic Acids - Structure, Isolation, and Synthesis.* Chemical Reviews, 1995. **95**(6): p. 1981-2001.

186. Schobert, R. and A. Schlenk, *Tetramic and tetronic acids: an update on new derivatives and biological aspects.* Bioorg Med Chem, 2008. **16**(8): p. 4203-21.

187. Saito, K. and T. Yamaguchi, *Nmr Spectroscopic Studies of Tautomerism in Tetramic Acid Analogs and Their Anilides .3. Polar-Solvent Effects on Tautomeric Populations.* Bulletin of the Chemical Society of Japan, 1978. **51**(2): p. 651-652.

188. Halo, L.M., et al., *Authentic heterologous expression of the tenellin iterative polyketide synthase nonribosomal peptide synthetase requires coexpression with an enoyl reductase.* ChemBioChem, 2008. **9**(4): p. 585-94.

189. Sims, J.W. and E.W. Schmidt, *Thioesterase-like role for fungal PKS-NRPS hybrid reductive domains.* J Am Chem Soc, 2008. **130**(33): p. 11149-55.

190. Liu, X. and C.T. Walsh, *Cyclopiazonic acid biosynthesis in Aspergillus sp.: characterization of a reductase-like R* domain in cyclopiazonate synthetase that forms and releases cyclo-acetoacetyl-L-tryptophan.* Biochemistry, 2009. **48**(36): p. 8746-57.

191. Belecki, K., J.M. Crawford, and C.A. Townsend, *Production of octaketide polyenes by the calicheamicin polyketide synthase CalE8: implications for the biosynthesis of enediyne core structures.* J Am Chem Soc, 2009. **131**(35): p. 12564-6.

192. Burr, D.A., X.B. Chen, and J.C. Vederas, *Syntheses of conjugated pyrones for the enzymatic assay of lovastatin nonaketide synthase, an iterative polyketide synthase.* Organic Letters, 2007. **9**(1): p. 161-4.

193. Pfeifer, B.A., et al., *Biosynthesis of complex polyketides in a metabolically engineered strain of E. coli.* Science, 2001. **291**(5509): p. 1790-2.

194. Liu, F., S. Garneau, and C.T. Walsh, *Hybrid nonribosomal peptide-polyketide interfaces in epothilone biosynthesis: minimal requirements at N and C termini of EpoB for elongation.* Chem Biol, 2004. **11**(11): p. 1533-42.

195. Samel, S.A., et al., *Structural and functional insights into a peptide bond-forming bidomain from a nonribosomal peptide synthetase.* Structure, 2007. **15**(7): p. 781-92.