Spring 5-2015

Molecular Mechanism for the Biosynthesis of Antifungal HSAF and Antibacterial WAP-8294A2

Haotong Chen
University of Nebraska-Lincoln, haotong.chen@huskers.unl.edu

MOLECULAR MECHANISM FOR THE BIOSYNTHESIS OF ANTIFUNGAL HSAF AND ANTIBACTERIAL WAP-8294A2

by

Haotong Chen

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Chemistry

Under the Supervision of Professor Liangcheng Du

Lincoln, Nebraska

April, 2015

MOLECULAR MECHANISM FOR THE BIOSYNTHESIS OF ANTIFUNGAL HSAF AND ANTIBACTERIAL WAP-8294A2

Haotong Chen, Ph.D.

University of Nebraska, 2015

Advisor: Liangcheng Du

Bioactive natural products are a major source of anti-infectives. Many ubiquitous inhabitants of soil and water, such as the Gram-negative, prolific producers of natural products, *Lysobact*er, remain largely unexplored. This PhD thesis reports our studies of the biosynthetic mechanism for a novel antifungal natural product HSAF and a potent anti-MRSA natural product WAP-8294A2.

In the first chapter, we give a short review of the biosynthesis of polyketides, focusing on those which exhibit new biosynthetic features. In subsequent chapters, we present original research on the biosynthetic mechanisms of dihydromaltophilin, a heat stable antifungal factor (HSAF). HSAF is the main antifungal factor that the biocontrol agent *Lysobacter enzymogenes* produces to fight against fungal pathogens. Most interestingly, the genetic feature of HSAF biosynthesis suggests that the same polyketide synthase (PKS) module acts not only iteratively, but also separately. In the thesis, we demonstrated the *in vitro* and *in vivo* production of the polyene tetramate, providing direct evidence for this highly unusual iterative polyketide biosynthetic mechanism that is likely general for this type of hybrid polyketide-peptides. We also investigated four oxidoreductase (OX1-4) genes in the HSAF biosynthetic gene cluster, to define the minimal gene cluster required for HSAF biosynthesis. Together, the results support a new biosynthetic mechanism: a single set of domains of an iterative PKS-NRPS, in cooperating with a cascade of redox enzymes, to synthesize a complex and highly modified polycyclic tetramate macrolactam.

The final part of the thesis presents research on WAP-8294A, a complex of at least 20 cyclic lipodepsipeptides exhibiting remarkable activity against methicillin-resistant *Staphylococcus aureus* (MRSA). All WAP-8294A compounds contain a 3-hydroxy fatty acyl chain, varied only in the chain length and branching pattern. The mechanism for activating and introducing the 3-hydroxy fatty acid chain into the peptide is unclear. We have identified seven putative acyl CoA ligase (ACL) genes and generated gene-disruption mutants. We also expressed the genes in *E. coli* to obtain the pure enzymes. Both the *in vivo* and *in vitro* results showed that ACL-6 is the dominant enzyme responsible for the activation of 3-hydroxyl-7-methyloctanoic acid and the subsequent initiation of WAP core structure formation.

TABLE OF CONTENTS

Chapter 1: Iterative Events in the Biosynthesis of Natural Products by Bacterial Type I Polyketide Synthase

1. Introduction

Polyketides are a highly diverse group of natural products, which are biosynthesized from short acyl-CoA units by polyketide synthases (PKSs)[1-3]. They represent an important source of novel therapeutics. The research of polyketides has led to the discovery of an enormous variety of natural products with a wide range of biological activities, many of which have found use as new pharmaceuticals, such as antibiotics, immuno-suppressants, antiparasitics, hypolipidemics, and anti-tumoral agents[4].

Over the past two decades, the biosynthetic gene clusters for a broad range of PKSs have been identified and characterized, and the studies of the structures, mechanisms and activities of PKSs have expanded the pool of novel compounds with a great diversity of activities through engineered combinatorial biosynthesis[5, 6]. Thus the remarkable versatility and amenability of PKSs make them attractive targets for pathway engineering to produce natural product analogues[7-12]. Bacterial Type I PKSs, the best-characterized PKSs, are responsible for creating a significant percentage of bioactive natural products[13]. This group of synthases has a modular organization, and each module within the PKS carries out one cycle of polyketide chain elongation (thus "non-iterative" in function). It is possible to deduce the core structure of a polyketide product from the organization of a typical type I PKS, since there is a well-established co-linearity between the number of chain elongations and the number of modules[14-17]. This rule is helpful in the reprogramming and rational redesign of novel natural products, which has given natural products a renewed prominence in drug discovery programs[11, 17]. However, more and more counter examples to the

co-linearity rule have emerged[1, 18, 19]. In bacteria, the iterative use of one or more modules in a type I PKS has made it challenging to predict the structure of the product and to engineer the biosynthetic genes.

Iteratively used bacterial type I PKSs, which include module stuttering in modular PKSs[19], bacterial aromatic iterative PKSs[20], PUFA PKSs[21, 22], enediyne PKSs[23-25], and the PTM PKS[26, 27], give rise to an impressive structural diversity of polyketide biosynthetic products, which are hard to deduce from the domain/modular arrangement of the biosynthetic machinery[28]. Little is known about the mechanism of iterative modules at the molecular level[29, 30]. Elucidation of the assembly of iterative modules, how they recognize their substrates, and how they function are issues at the frontier of natural product research and provide new prospects and inspirations for novel compound discovery and engineering. The following is a brief review of examples of module iterative phenomena in bacterial type I PKSs and mechanistic studies of their processes.

1-1. Classification of PKSs

Nature supplies a large group of polyketides with highly diverse structures and biological activities. Many of these polyketides are medicines or leads for the development of anti-infective and anticancer drugs, such as erythromycin (antibacterial), ambruticin (antifungal), and daunorubicin (antitumor)[4]. Despite their tremendous diversity in structure and function, polyketides share a common biosynthetic mechanism, as the carbon backbones are biosynthesized from short acyl-CoA precursors by a sequential condensation catalyzed by polyketide synthase (PKS).[31] In general, polyketide synthase in charge of selecting, fusing, and processing the building blocks consists of individual catalytic domains. The chain initiation and

extension substrates are selected and activated by an acyl transferase (AT) domain, and transferred to the 4'-phosphopantetheinylated acyl carrier protein (ACP) domain. The ketosynthase (KS) domain catalyzes the decarboxylative Claisen condensation

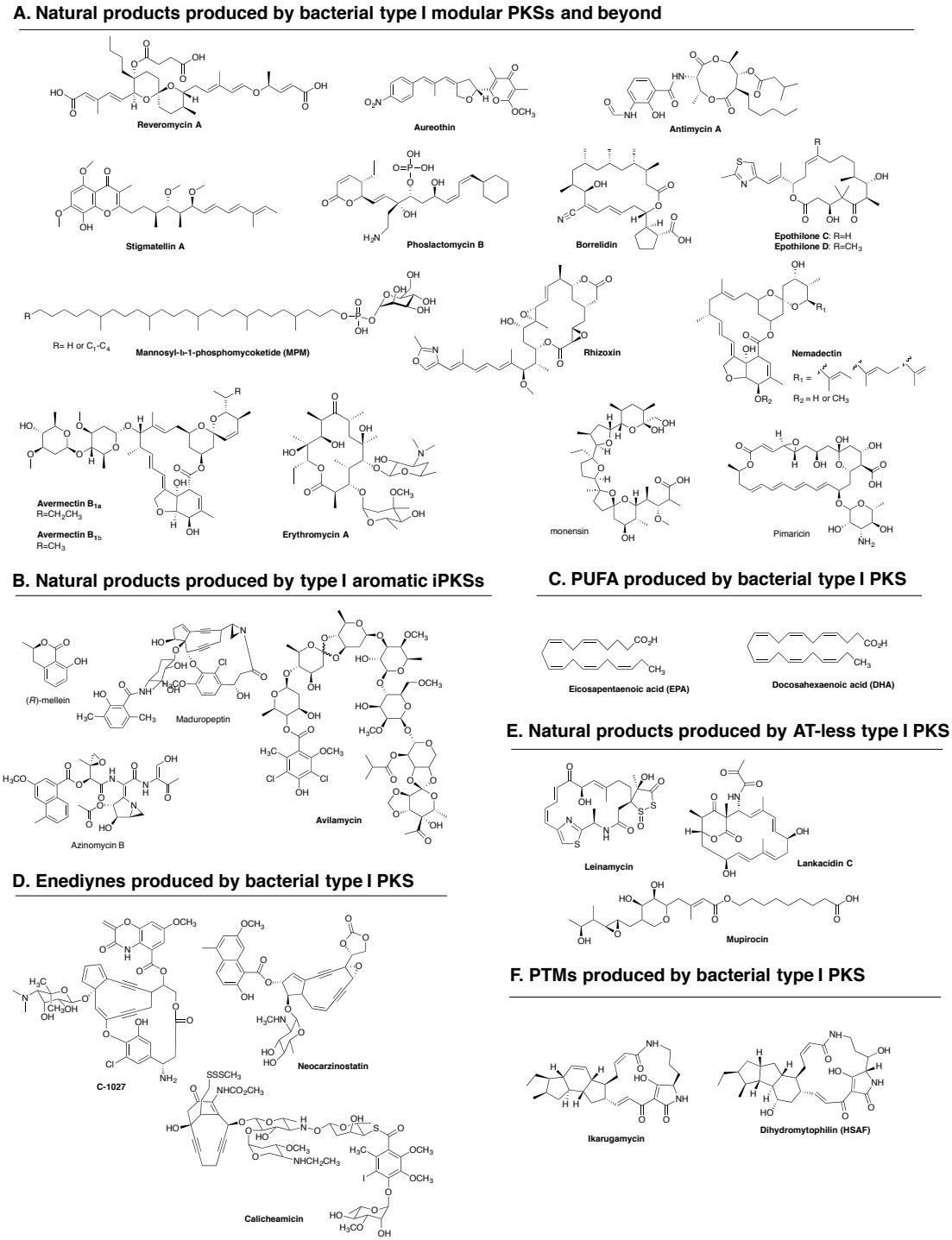


Figure 1-1. Examples of natural products synthesized by bacterial type I PKSs, which are described in this review.

between the substrate and the growing polyketide, forming a carbon-carbon covalent bond between the alpha carbon of the extender unit and the thioester carbonyl of the ACP-bound acyl chain. While a PKS requires a minimum of these three domains to function, other commonly present domains are the ketoreductase (KR), dehydrogenase (DH), and enolreductase (ER), which reduce the β-keto group into a fully saturated acyl chain. Most PKSs vary in the use of these reductive steps in the processing of the β-keto group, thus generating a greater diversity and complexity of polyketide products.

The classification of PKSs is based on the structural organization of PKS domains. In type I PKSs, a single large polypeptide chain houses multiple different active sites (domains) capable of catalyzing different reactions. Such PKSs are found in both bacteria and fungus. Type II PKS are usually found in bacterial systems, assign the different enzymatic activities of the polyketide synthesis to separate dissociable proteins. The majority of known type II PKS metabolites are aromatic products, although exceptions are known[32]. Type III PKSs utilize acyl-coenzyme A (CoA) thioesters directly, rather than depending on extender units tethered to an acyl carrier protein (ACP). These systems also produce aromatic metabolites, but are mainly distributed in the plant kingdom, while seldom occurring in bacteria[33]. While these classifications of PKSs are useful generalizations, nature does not limit itself to only three rigidly-defined methods for generating polyketide metabolites[34]. We will limit our discussion to examples in which type I modular PKSs synthesize polyketides behave in a non-canonical iterative pattern, thus combining two traits once thought to be mutually exclusive to fungal or to bacterial systems.

1-2. Classification of type I PKSs.

Type I PKSs are often classified into two major groups: the iterative (fungal) type I PKSs and modular (bacterial) type I PKSs. In iterative I PKSs (iPKSs), consist of one multienzyme protein, containing only a single module subdivided into various extending and tailoring domains, which is nevertheless capable of conducting multiple rounds of chain extension and β-keto processing. The varied reduction level of the β-keto during each round of elongation is the most intriguing feature of this PKS type[35-37]. These types of iPKSs are largely confined to fungal systems, thus will not be discussed in this review. Modular type I PKSs, in contrast, include multiple sequential modules, each of which contain the needed extending and tailoring domains, and operate in a modular fashion with each module being responsible for only one round of chain elongation and subsequent β-keto processing before passing the nascent polyketide is passed on to the downstream module, which carries out another round of chain extension and processing. Examples of such modular type I PKSs include those responsible for the biosynthesis of the macrolides such as erythromycin[38, 39] and avermectin[40]. These bacterial type I PKSs thus carry out each successive chain extension cycle by a different set of active sites housed in separate modules, and one or more such modules may exist on each multifunctional polypeptide. Importantly, in most modular type I PKSs, each active site is used only once. This linear mechanism used to be regarded as general for the biosynthesis of polyketides by type I PKSs in bacterial systems.

The repeated emergence of unusual type I PKSs has expanded the repertoire of PKSs. Although all of them have the basic modular structures, the unusual type I PKSs often have intriguing arrangements of the functional domains and exhibit a phylogenetic distance from either iterative or modular PKSs, making them fall into five new subgroups, including bacterial aromatic iPKS, polyunsaturated fatty acid (PUFA)

synthases, enediyne, polycyclic tetramate macrolactams (PTM) and AT-less type I PKSs[41, 42]. The structures of typical natural products of each group are showed in Figure 1-1.

1-3. Phylogenic study of type I PKSs.

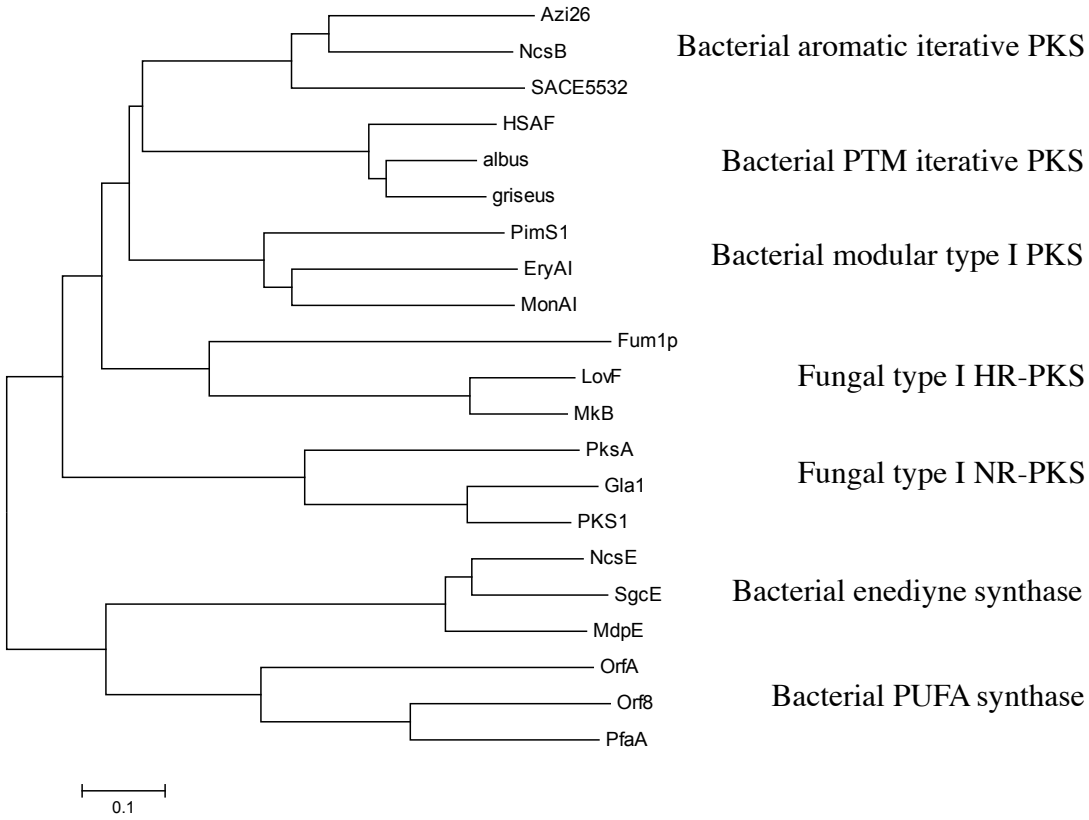


Figure 1-2. Phylogenetic analysis of different types of PKSs based on KS domains. Enzymes and their GenBank accession numbers are listed as follows: bacterial aromatic iPKS (3 sequences, including AziB for Azinomycin biosynthesis, ABY83164.1; NcsB for Neocarzinostatin biosynthesis, AAM77986.1; SACE_5532 for Mellein biosynthesis, YP_001107644.1). PTM type iPKSs (3 sequences, including PKS for HSAF biosynthesis, EF028635.2; PKS from *Streptomyces albus* J1074, ABYC01000481; PKS from *Streptomyces griseus,* AP009493); modular PKSs (3 sequences, including PimS1 module 1 for Pimaricin biosynthesis, CAC20931.1; MonAI: module 1 for Monensin biosynthesis, AAO65796.1; EryA1: module 2 for Erythromycin biosynthesis, YP_001102988.1); HR-PKSs (3 sequences, including

Fum1p for Fumonisin biosynthesis, AAD43562.2; LovF for Lovastatin biosynthesis, AAD34559.1; MkB for Monacolin biosynthesis, ABA02240.1), NR-PKS (3 sequences, including PksA for Aflatoxin biosynthesis from *Aspergillus. flavus*, AAS90093.1; PKS1 for Melanin biosynthesis from *Colletotrichum lagenaria*, BAA18956.1; PKS1 for Melanin biosynthesis from *Glarea lozoyensis*, AAN59953.1); enediyne synthases (3 sequences, including SgcE for C-1027 biosynthesis, ZP_11383500.1; MdpE for Maduropeptin biosynthesis, AAQ17110.2; NcsE for Neocarzinostatin biosynthesis, AAM78012.1) and PUFA synthases (3 sequences, including PfaA for EPA biosynthesis from *Photobacterium profundum*, AAL01060.1; Orf8 for DHA biosynthesis from *Moritella marina*, BAA89382.2; OrfA for EPA and DHA biosynthesis from *Schizochytrium* AAK72879.2). Numbers above branches indicate bootstrap values (1000 replicates) of the major clades. Similar sequences were aligned with ClustalW and the tree shown was generated using the MEGA 5.0).

Phylogenetic studies[20, 43] have investigated the evolutionary relation and catalytic diversity of PKSs. The bacterial aromatic iterative PKSs and PTM- type PKSs make up two independent clades that lie close to each other and sister to modular PKSs, but are phylogenetically distant from the clade containing two other types of bacterial iPKSs, that is, enediyne PKSs and PUFA synthases, and the clade containing fungal NR-PKSs and HR-PKSs (Figure 1-2). The phylogenetic differences between the subgroups in type I PKSs helps understand the relationships between these type I PKSs, and has given rise to the development of a rapid PCR approach to specifically access the genes encoding these enzymes[44, 45].

1-4. Non-canonical bacterial type I PKS

Bacterial type I PKSs were originally thought to be rigidly modular and co-linear with the core structure of final products, so that the core structure of an unknown compound produced by this type of PKS should be capable of deduction from

bioinformatics analysis of PKS genes. Due to the high level of programming, type I PKS systems have the potential to be engineered to make novel analogues of known metabolites. Through biosynthetic engineering, novel compounds have been accessed by re-programming type I PKS systems in a number of ways. Among other experiments, loading modules have been swapped, modules have been deleted or inserted, domains swapped or inactivated, and domains reprogrammed to impart new functions, among other experiments[46].

However, there are many deviations from the co-linearity in PKS systems have been discovered in recent years. Two major reasons contribute to the loss of co-linearity: non-canonical domains and non-canonical modules. The phenomenon of non-canonical domains includes additional domains, loss of specificity of AT domains, absence of domains for required functions, and novel domain functions and domain organization[47]. For example, the newly discovered branching domain acetylates an enzyme-bound enoyl moiety by means of a Michael-type conjugate addition, resulting in a β-branch instead of an elongated backbone in rhizoxin biosynthesis (Figure 1-3B)[48]. *Trans* domain activities refer to discrete enzymes which act only at specific points during the synthesis of nascent acyl chains, separate from the usual modifying activities. Examples are trans AT in AT-less PKS system such as leinamycin synthase (Figure 1-3B)[41, 49], a discrete KR domain (*antM*) in the biosynthesis of antimycins[50], and a discrete DH domain in the biosynthesis of phoslactomycin [51, 52]. Structures of the compounds discussed above are listed in Figure 1-1.

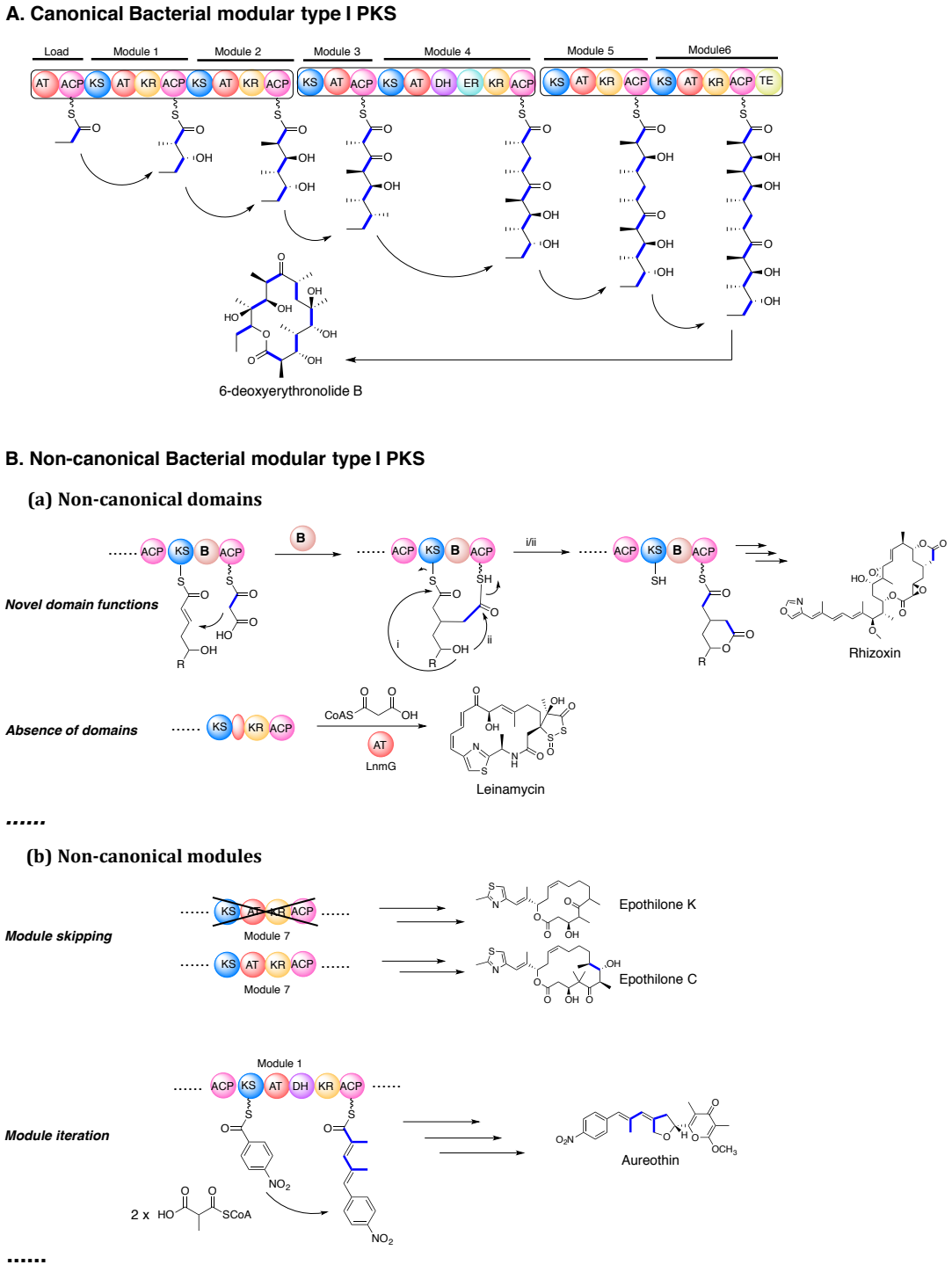


Figure 1-3. Examples of canonical bacterial type I PKSs and non-canonical bacterial type I PKS. KS, ketosynthase; AT, acyltransferase; DH, dehydratase; KR, ketoreductase; ACP, acyl carrier protein; TE, thiolesterase; C-MeT, C-methyltransferase; ER, enoyl reductase; B, branching domain.

The non-canonical domains affect the co-linearity mainly in the modification of the backbone, such as generating unpredictable reduction levels of the β-keto group. But the non-canonical modules sometimes result in changing the size of the backbone in the final products. The most frequent of such aberrations are the skipping of various modular PKS single domains, or sometimes of entire modules; conversely, individual modules may be used more than once, giving rise to the so-called "module stuttering", an example of a bacterial iterative PKS. Such events lead to inconsistency between the genetic organization of the biosynthetic gene cluster and its corresponding proteins, and its ultimate PKS metabolite.

A PKS which skips over modules usually does so in "programmed skips", passing over extraneous functional domains or modules in a PKS for which there is no obvious need in the biosynthesis of the corresponding polyketide; one such case is that of the similar KS-ACP bi-domain module in the biosynthesis of mupirocin[42]. There is also a trace amount of truncated derivatives produced during the biosynthesis of avermectin and epothilone (Figure 1-3B)[53, 54], suggesting that the truncated compounds are produced by the aberrant bypass of one or other module during biosynthesis. The reveromycin-producing PKS is one example where the genetic organization of the biosynthetic gene cluster and the corresponding proteins are not co-linear[55]. The analysis of the reveromycin PKS shows that it is possible that after translation, RevC protein (module 1-3), associates with RevA (module 4-6) instead of RevD, whose gene is located next to and downstream of RevC to bring the set of contiguous domains together to form functional modules. Therefore, even when the PKS genes are not present on the chromosome in a co-linear arrangement, the gene products form a specific head-to-tail complex through which the polyketide chain is processed in a programmed fashion and biosynthesis occurs in a co-linear fashion.

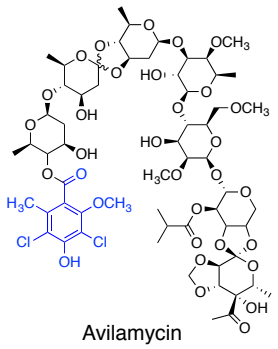
Iterative use of one or more modules will often result in longer polyketide chains than those predicted by the co-linear model (Figure 1-3B). This "module stuttering" composes a large group of non-canonical bacterial type I PKS. Beside those with a special domain organization, which already fall into novel subclasses of type I PKSs (bacterial aromatic iPKS, polyunsaturated fatty acid (PUFA), enediyne PKSs), there are still many extraordinary examples which lie beyond any existing classifications. In the following sections, I summarize the iteration events known to occur in bacterial type I PKSs and discuss the proposed mechanisms behind this phenomenon.

2. Iteratively used bacterial type I PKS

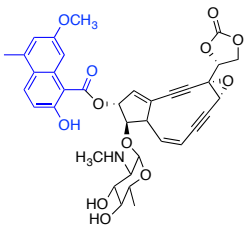
Iterative type I PKSs can be considered in two categories: entirely iterative PKSs ("fungal" type iPKS) and partially iterative PKSs (which contain stuttering modules).

2-1. Entirely iterative bacteria PKSs

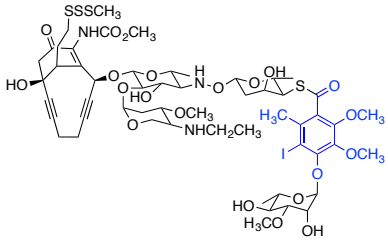
(a) Bacterial aromatic iPKS KS AT DH KR ACP



Avilamycin



Neocarzinostatin



Calicheamicin

(b) Bacterial PUFA synthase KS AT ACPn KR DH

Eicosapentaenoic acid (EPA)

Docosahexaenoic acid (DHA)

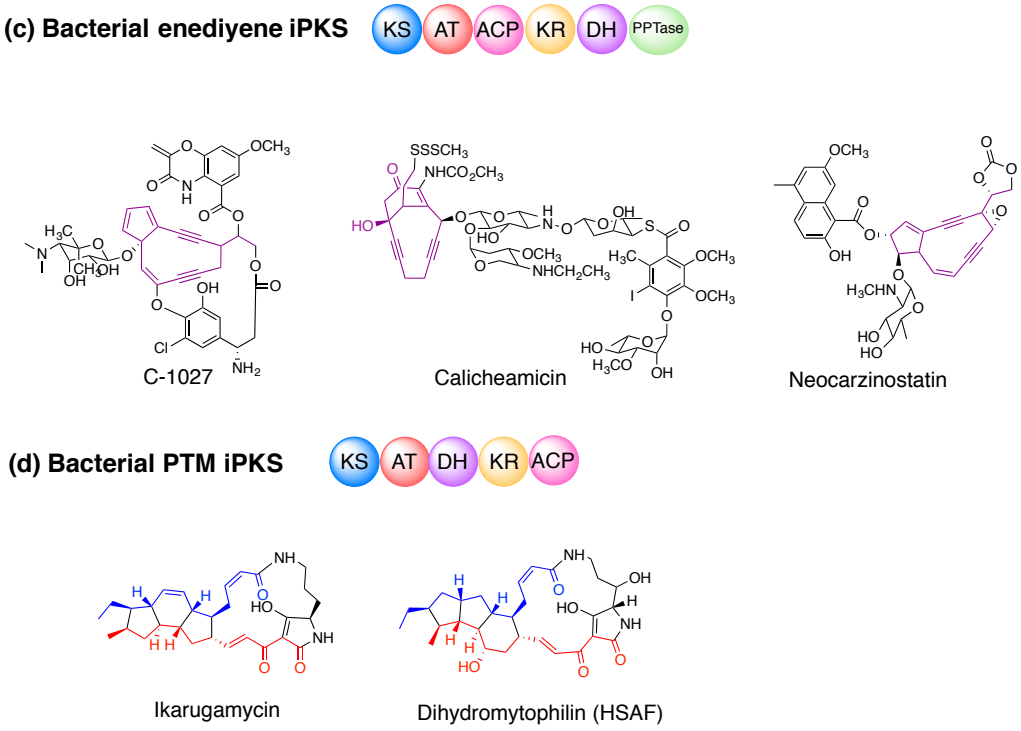


Figure 1-4. Domain organization of bacterial type I iPKSs and represented compounds produced by these synthases. Domain organizations of iterative type I PKSs, including aromatic iPKS, enediyne PKS, PUFA synthase and PTM PKS in bacteria. ACPn, multiple ACPs; PPTase, phosphopantetheinyl transferase.

2-1-1. Bacterial aromatic iPKSs

Most aromatic compounds produced by bacteria were believed to be synthesized by type II PKSs, until the first iterative type I PKS, AviM, was cloned in 1997 (Figure 1-4a)[56]. Heterologous expression of AviM and the biosynthesis of ordellinic acid in vitro confirmed that AviM is an iterative type I PKS that is able to catalyze aromatic polyketide synthesis. CalO5, which is responsible for biosynthesis of calicheamicin (Figure 1-4a), was successfully characterized to be another iterative type I PKS in bacteria[25, 57]. Besides the monocyclic aromatic polyketides, there are iterative type I

PKSs that also can generate higher-order aromatic polyketides, such as NcsB, which catalyze the formation of a naphthalinic acid moiety in neocarzinostatin biosynthesis[58]. So far, the known bacterial aromatic iPKSs are associated with the biosynthesis of chlorothricin[44], maduropeptin[59], polyketomycin[60], pactamycin[61], tiacumicin B[62], azinomycin B[63] and (R)-mellein[64].

Bacterial aromatic iPKSs are typically organized in the order of KS-AT-DH-(KR)-ACP (Figure 1-4a), and share high homology with each other, which is very similar to that of modular type I PKSs. Although the bacterial iPKSs are also highly homologous to fungal 6-MSA synthases[37], the absence of a *C*-methylation (CMeT) and enoylreductase (ER) domains is the most notable difference, suggesting a convergent evolution of PKSs to produce similar aromatic cores.

Bacterial aromatic iPKSs contributes to the diversity of polyketides generally through their unusual domain functions, different mechanisms for polyketide off-loading and post-synthesis modifications. For instance, most of the aromatic PKS KR domains contain a motif similar to the Leu-Asp-Asp motif of B-type KRs, suggesting that all the KR domains of bacterial aromatic iPKSs specifically generate *D*-hydroxyl group by reduction of the β-keto group (Figure 1-5B)[65, 66]. But different KR domains vary the hydroxylation pattern, generating different cyclized rings (5-methyl-1- NPA for AziB and 2-hydroxyl-5-methyl-1-NPA for NcsB) (Figure 1-5A). An unusual off-loading mechanism was discovered in bacterial aromatic iPKSs. In the biosynthesis of chlorothricin, pactamycin, avilamycin, tiacumicin B, and calicheamicin, a dedicated AT is utilized to release the polyketides products from PKSs (Figure 1-5C)[67-70].

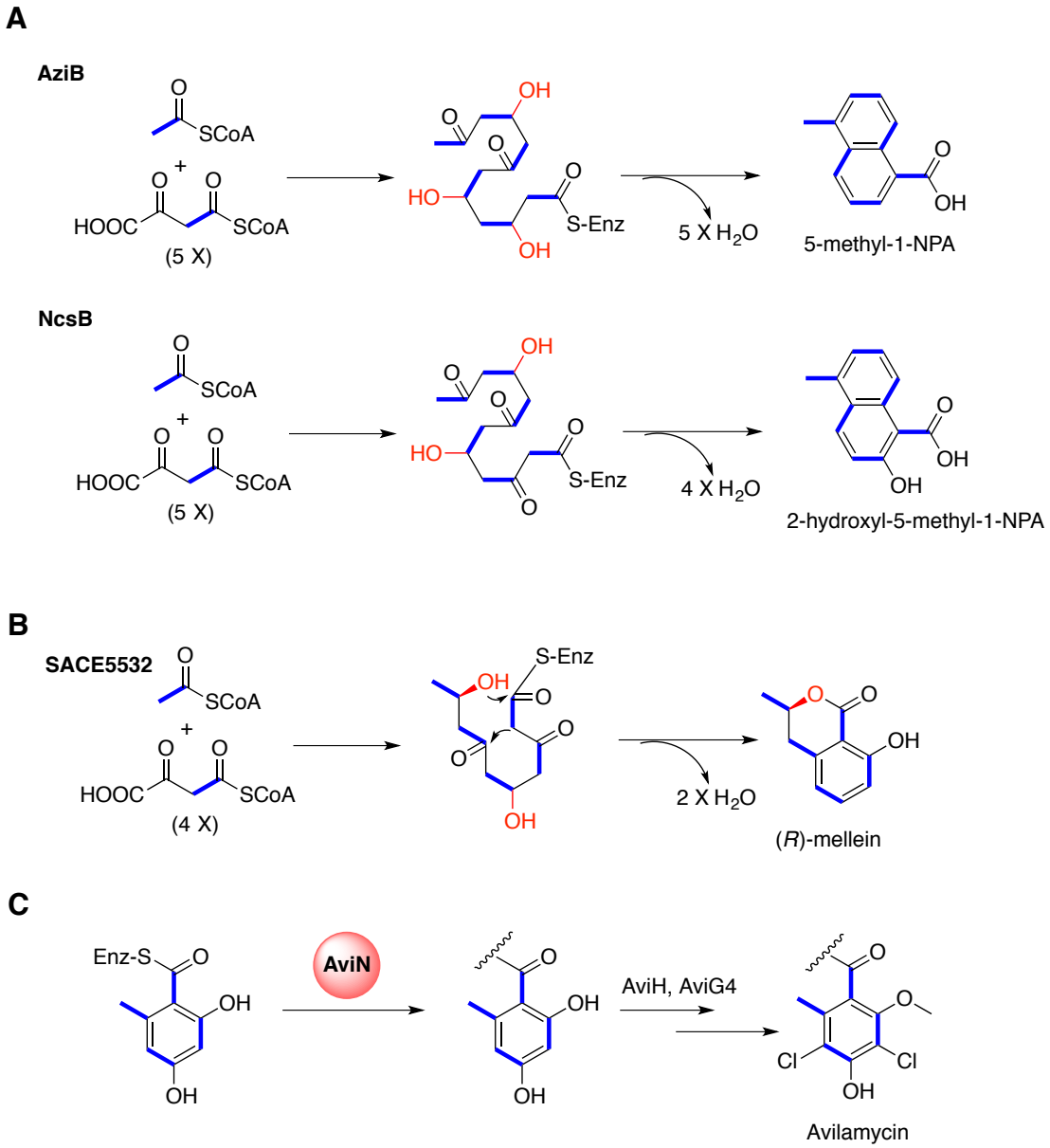


Figure 1-5. Bacterial aromatic iPKS-catalyzed reactions and the off-loading mechanism in the biosynthesis of avilamycin. The selective keto-reduction patterns are highlighted in red. ATs are represented by red rectangles.

Unlike the complex multicyclic aromatic scaffolds produced by type II PKSs, polyketides produced by bacterial aromatic iPKSs are relatively simple and consist of a set of mono- or bicyclic aromatic products. Thus bacterial aromatic iPKSs provide a

relatively simple system for understanding the rationale behind the catalysis achieved by these impressive enzymes. Further understanding of these domains in bacterial aromatic iPKSs will hopefully facilitate the alterations of the functionalities on the aromatic polyketides, opening the door to bioengineering studies and the production of novel natural products with improved biological activities.

2-1-2. PUFA

Polyunsaturated fatty acids (PUFAs) are essential to human health and nutrition. PUFAs contain multiple *cis* double bonds and have been identified outside of the eukaryotic realm in various marine bacteria. A gene cluster *pfa* was cloned into *Escherichia coli*, from a *Shewanella* sp., and was found to be responsible for eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) biosynthesis (Figure 1-4b)[22]. Five open reading frames were subsequently found to be the minimal gene set for PUFA biosynthesis[71]. Based on analysis of PUFA iPKSs, five to nine tandem ACP domains are present in the iPKSs, besides basic domain composition of KS, AT, KR, and DH (Figure 1-4b).

After heterologous expression the PUFA iPKS, (*pfa*ABCDE) from *S. japonica* in *E. coli,* together with a series of *in vitro* and *in vivo* characterizations of these iPKS proteins, Jiang and co-workers speculated that each of the bioinformatically predicted ACPs can be phosphopantetheinylated by the PPTase at the N-terminus both *in vitro* and *in vivo*[72]. Each of the tandem ACPs was functionally equivalent for PUFA biosynthesis, but the number of functional ACPs correlated directly to PUFA titers.

Further efforts will be made to rigorously dissect PUFA iPKSs from broader microbial PUFA producers.

2-1-3. Enediyne biosynthesis

The enediyne are a family of antibiotics containing a strained ring system (9 or 10 membered), which are extremely cytotoxic (Figure 1-4c)[73].

The biosynthetic gene clusters for two such metabolites, the chromoprotein antibiotic C-1027[23] and calicheamicin[25], have been identified and published, and the enediyne core of these metabolites was believed to be synthesized by a type I PKS. Ben Shen and co-workers cloned a gene cluster from *Streptomyces globisporus* containing the PKS gene *sgcE*,[23] part of the gene cluster responsible for the biosynthesis of the model nine-membered enediynes C-1027. The enediyne PKS *sgcE* contains six domains (Figure 1-4c). The region between AT and KR was proposed to be an ACP, which is a quite unusual location for an ACP domain. More interestingly, the C-terminus region of the PKS was predicted to be a phosphopantetheinyl transferase (PPTase) domain, loading the phosphopantetheinyl arm of CoA onto the ACP. This is an extremely rare event and probably unique to enediyne PKSs. Overall, the enediyne PKS shows the greatest homology to the PUFA PKSs, and like them they lack ER domains.

The involvement of SgcE in C-1027 biosynthesis was confirmed by gene inactivation and complementation: a Δ*sgcE* mutant eliminated the production of C-1027, and overexpression of SgcE in Δ*sgcE* restored C-1027 production[23]. This iPKS could iteratively catalyze the assembly of a linear polyunsaturated intermediate, which was further processed by a series of desaturations to furnish the two alkyne groups and cyclized to afford the enediyne core. Indeed, a group of five to ten genes is flanking the *sgcE* PKS gene, which are highly conserved and homologous to the oxidoreductases or proteins of unknown functions that are only associated with

enediyne biosynthesis, hinting at the possible novel chemistry involved in the synthesis of these natural products.

CalO5, which is also discussed above, catalyzes the biosynthesis of the orsellinic acid (OSA) core (Figure 1-4c). The gene was cloned together with another PKS encoded gene calE8 within the same cluster from *Micromonospora echinospora* sp. *calichensis*[25]. The resulting compound is calicheamicin, which is a model for the 10-membered enediyne antibiotics. The involvement of CalE8 in calicheamicin biosynthesis was also confirmed by gene inactivation. The domain organization of calE8 is homologous to sgcE. The observed similarity between SgcE and CalE8 clearly suggests a common polyketide pathway for the biosynthesis of both nine- and ten-membered enediynes. The biosynthetic mechanism of calicheamicin is very similar to that of C-1027; the CalE8 PKS catalyzes the biosynthesis of a nascent polyunsaturated intermediate in an iterative process. Modifications by the associated oxidoreductases within the gene cluster then form an enediyne core intermediate. SgcE and CalE8, therefore, represent a novel family of iterative type I PKSs, establishing a new paradigm for enediyne biosynthesis[58, 74].

The discovery of this type of biosynthetic mechanism led to a PCR-based approach to access the enediyne PKS loci and its associated redox genes[75]. The highly programmed biosynthetic mechanism suggested the bioengineering potential in generating new enediynes.

2-1-4. PTM iPKS

Polycyclic tetramate macrolactams are a group of compounds that have recently been shown to be produced by a variety of bacterial species (Figure 1-4d)[27]. A combination of small molecule chemistry, biosynthetic analysis, and genome mining in diverse

bacteria led to the recognition of this group of compounds as a new class of NPs with distinct structure and novel modes of action. The most striking feature of PTM biosynthesis is that there is only one single-module PKS-NRPS hybrid gene in the biosynthetic gene clusters [76]. This feature suggests that the single-module PKS responsible for the synthesis of [76]the PTM polyketide moieties is not only iterative, but also capable of catalyzing the biosynthesis of two separate polyketide chains and linking them to the amino acid activated by the NRPS module.

The hybrid PKS-NRPS for PTM biosynthesis has a typical modular domain organization, including KS-AT-DH- KR-ACP for the PKS portion and C-A-PCP-TE for the NRPS portion (Figure 1-6). There is no obvious remnants of an inactive enoylreductase (ER°) domain or a methyltransferase (CMeT) domain, as seen in several iterative fungal PKS-NRPS with the similar organization, or other types of batcerial iPKSs. Although this type of iterative PKS-NRPS is commonly seen in fungi, it was not thought to be common in bacteria until recent years. Phylogenetic analyses showed that there is a close relationship between PTM iPKSs and bacterial aromatic iPKSs, rather than fungal iPKSs (Figure 1-6).

The heat–stable antifungal factor (HSAF), which was isolated from the Gram-negative bacterium *Lysobacter enzymogenes*, exhibits potent inhibitory activities against a wide range of fungi, including the life-threatening human pathogen *Aspergillus fumigatus*[77-79]. Like other bacterial PTM biosynthetic gene clusters, the HSAF biosynthetic gene cluster contains only a single PKS-NRPS and the PKS contains only a single module, although the HSAF scaffold is apparently derived from two separate hexaketide chains along with an ornithine residue[76-79]. This suggests that the same PKS module would act not only iteratively, but also separately, in order to

construct and link the two different hexaketide chains with the NRPS-activated ornithine to form the characteristic PTM scaffold. We heterologously expressed the HSAF biosynthetic gene cluster in *Streptomyces* host, in which the native PKSs have been deleted, and the expected HSAF analogs have been detected and isolated. We also demonstrated the iterative polyketide biosynthetic mechanism *in vitro* using purified PKS and NRPS proteins[26].

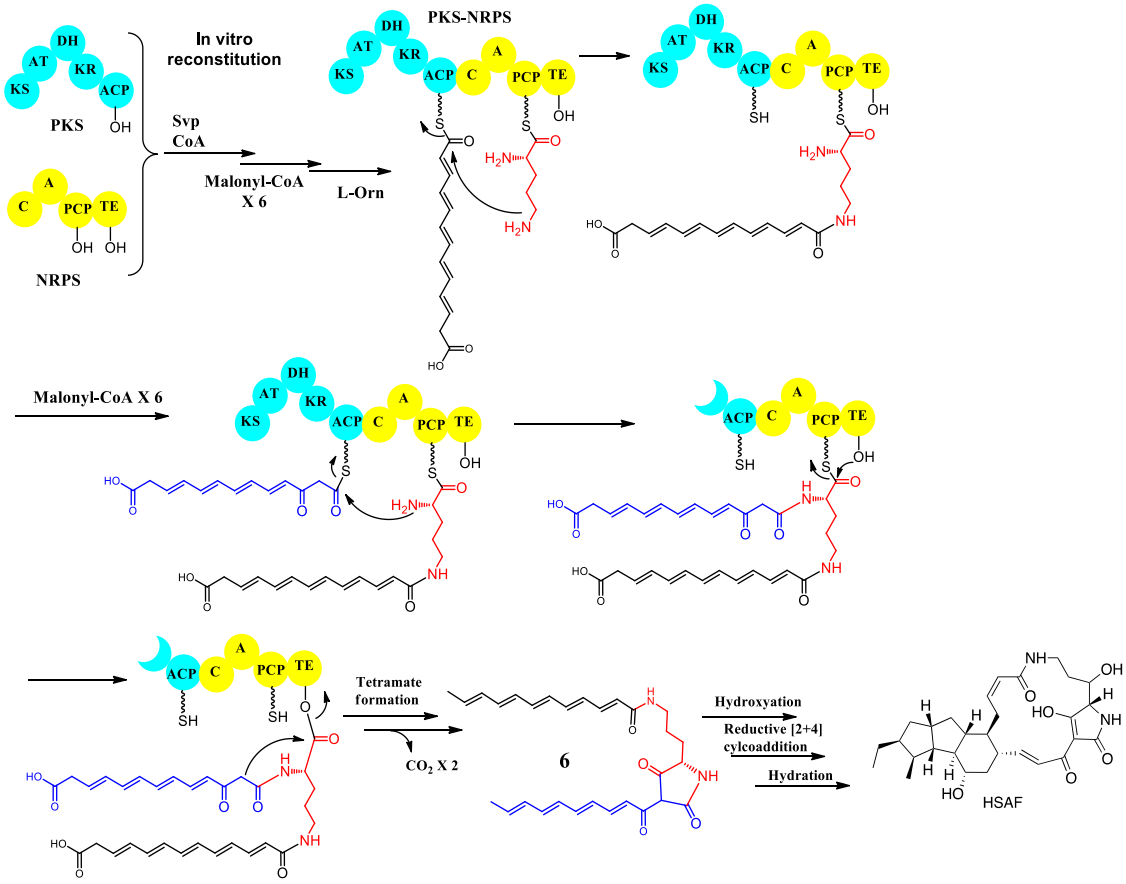


Figure 1-6. Proposed biosynthetic mechanism for HSAF.

In the *in vitro* reaction, the hexaketide polyene was the only detected intermediate. The results support that the single-module PKS is indeed an iterative PKS, which catalyzes five cycles of chain elongation in polyketide biosynthesis. The timing of the

second hexaketide chain biosynthesis and transfer is also interesting (Figure 1-6). The second hexaketide chain is synthesized and transferred right after the fifth cycle of polyketide chain elongation, in which the newly formed β-keto group of the second hexaketide chain has not been processed by the KR domain and DH domain. All of PTM natural products hitherto discovered retain this unprocessed β-keto group in the final structure (HSAF structure, Figure 1-6). The determining factor for this timing is not known, but is possibly related to the tailoring redox enzymes that are proposed to cooperate with the PKS-NRPS in elaborating the PTM scaffold.

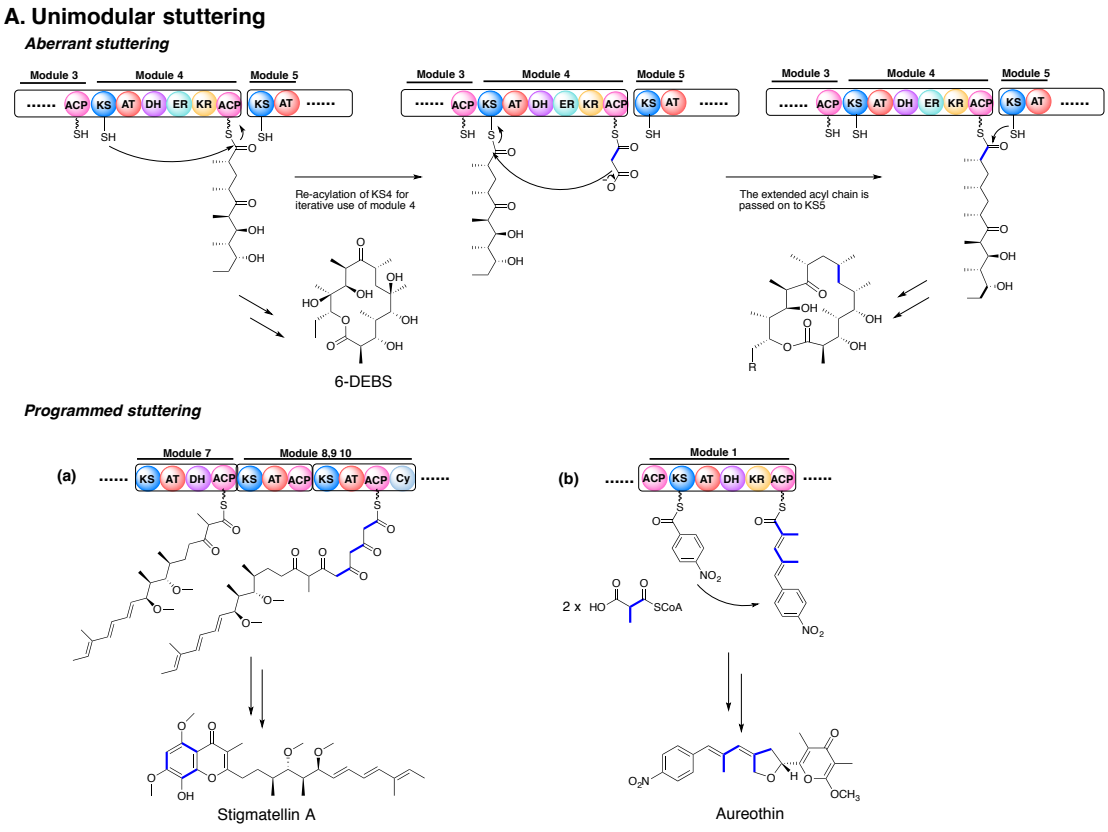
Recently, the Gulder group reported heterologous expression of the ikarugamycin biosynthetic gene cluster in *E. coli*[80], and the Zhang group reported the enzymatic mechanism for formation of the inner 5-memebered ring and demonstrated the polyketide origin of the ikarugamycin skeleton[81]. Ikarugamycin is a *Streptomyces*-derived PTM which has a 5,6,5-tricyclic system (Figure 1-4c). Both the Gulder and Zhang groups showed that a three-gene cluster is sufficient for ikarugamycin biosynthesis.

In light of the huge number of uninvestigated PTM-type gene clusters in databases, we believe our studies will facilitate research in a new area of iterative bacterial type I polyeketides. The understanding of the iterative mechanism could be used to guide biosynthetic engineering efforts. Knowledge of these biosynthetic mechanisms may also transfer to the exploration and discovery of new bioactive NPs from other *Lysobacter* species and allow optimization of the yield, structure, and activity of the *Lysobacter* anti-infectives.

2-2. Partially iterative bacteria PKSs.

The iteration pattern in entirely iterative bacterial type I PKSs is similar to that in

single-module fungal iPKSs, in which the iteration events occur in only one complete PKS module. But in bacterial modular PKSs, the iteration events can happen within only one or two modules out of many in the giant assembly line. “Module stuttering” is the term used to describe the iterative use of only one module in a multi-modular PKS, and is usually implicated when the number of modules present in a PKS is fewer than the number of condensation events required for complete biosynthesis of the corresponding polyketide[47]. In recent years, many examples of module stuttering have been reported in the biosynthesis of polyketides. Stuttering can be either aberrant or programmed.



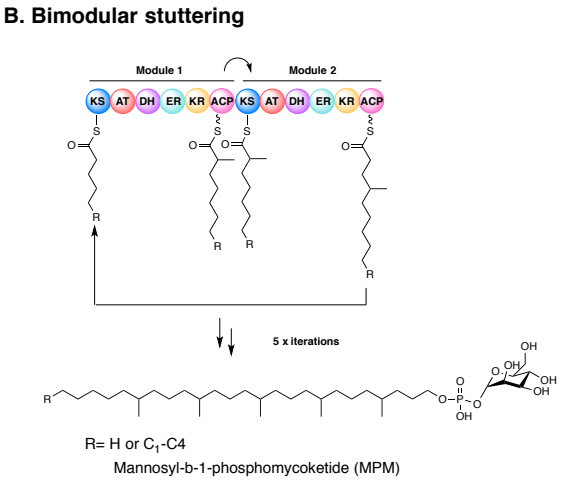


Figure 1-7. Mechanism of modular stuttering.

2-2-1. Unimodular iterations

In the biosynthesis of erythromycin A in *S. erythraea,* the detection and isolation of minor metabolites (Figure 1-7) in the fermentation broth of *S. erythraea* strain No. 5 revealed the first example of how a processive and co-linear modular PKS can operate in an iterative fashion in an aberrant manner[29]. This iteration is considered as aberrant because of the trace amounts of the iterative products, as native modular PKSs generally exert a rigorous control over the chain length.

Structural elucidation of the final product indicated that the iteration event happened during the fourth chain extension (Figure 1-7A). Such a homologation reaction would require the normal pentaketide product of module 4, attached to ACP4 (DEBS2), to be transferred back onto the active site cysteine residue of the KS4 domain of the same module, rather than being processed onto the downstream KS5 domain of DEBS3 (Figure 1-7A). A second, module 4-catalyzed chain extension cycle would

occur, producing a novel hexaketide chain attached to the ACP4 (DEBS2). This new hexaketide intermediate would be transferred to the KS5 (DEBS3) domain and the elongation process would then progress in the normal manner to give a ring-expanded octaketide product. The explanation for this aberrant stuttering is that the location of the module 4 allows the iteration event to occur before the transfer of the pentaketide intermediate to the next module. But this event only happened occasionally.

The same biosynthetic stuttering has also been found in epothilone PKS, which represents the 18-membered macrolactones[30]. Both of the module 5 and module 6 could be iteratively used, resulting in 6 different byproducts, based on different substrate selectivity of AT domains. The diversity of the iterative by-products indicated that the iteration events could happen to different modules within the same assembly line.

Apart from the occasional and unpredictable aberrant stuttering event in Type I bacterial PKS biosynthesis, most iterative PKS events are tightly controlled and apparently have developed during the evolution of the assembly lines. A number of programmed iterative events in Type I modular PKS that correspond to a series of natural products have been identified. The best studied and classic examples for the programmed iterative use of a type I PKS module are the stigmatellin, aureothin (two iterations each), and borellidin (three iterations) pathways (Figure 1-7A).

As an aromatic myxobacterial electron transport inhibitor, stigmatellin was isolated from the myxobacterium *Stigmatella aurantiaca*[32]. By analysis of the structure and gene cluster, Muller and co-workers believed that there must be 3 more malonyl-CoA units added to the backbone by two modules, StiH or StiJ (Figure 1-7A)[32, 82]. This stuttering is different from the ring expansion in erythromycin biosynthesis, since

stigmatellin is the only product observed in the culture broth. It strongly suggested a dedicated and programmed iterative processing in the PKS assembly.

The authors speculated thateither StiH or StiJ could be used iteratively, while the other worked classically. It's also possible that this iteration was actually from an *in trans* module outside the gene cluster. To exclude this possibility, the authors mutated every PKS-containing gene cluster in the genome of *S. aurantiaca*[83]. However, none of these mutations eliminated the production of stigmatellin, suggesting that its biosynthesis is solely the responsibility of the *sti* locus, apparently necessitating iterative use of the associated PKS.

The biosynthesis of lankacidin is a very unusual example of the iterative use of a PKS module. A module containing KS-DH-CM-ACP1-ACP1-KS-AT spread over four proteins is proposed to be iteratively used for the first five condensations[84, 85]. The presence of repeated, virtually identical, ACP domains is the most striking feature, suggesting that this module may be involved in iterative use, with twin ACP domains having different uses (Figure 1-7A).

The aureothin PKS also contains a programmed iteratively-used module, which represents a favorable model system to investigate the iteration process (Figure 1-7A). The iteration module, AurA, not only works iteratively, but also selects the *p*-nitrobenzoate starter unit[86]. Hertweck and co-workers proposed that the iteration happened by retrotransfer of the intermediate from one PKS strand to the opposite PKS strand[87]. In a recent published paper, Hertweck and co-workers confirmed that KS1 is in charge of priming the PKS, thus leaving a vacant position for the retrotransfer of the diketide. KS2 functions as a gatekeeper that controls the chain length of polyketide. The substitution of module 1 of AurA with the avemectin

module 1 resulted in abolition of iteration, indicating that iteration is completely due to aureothin module 1. The deletion of the TE domain enabled the authors to track the intermediates produced by *aur* PKS. The single AurA, which could catalyze up to 4 rounds of iteration, suggested a scenario for rational design of complex polyketides by pathway engineering.

2-2-2. Bimodular iterations

So far, module stuttering has usually been found to involve only one PKS module. However, a novel bimodular iteration was unraveled in 2008[88]. A protein involved in the biosynthesis of a phosphoglycolipid called mannosyl-β-1-phosphomycoketide (MPM) was proposed to contain two complete sets of modules and has been suggested to synthesize mycoketide by five alternating condensations of methylmalonyl and malonyl units, using an iterative mode of catalysis (Figure 1-7). The chemical structure of mycoketide contains branching at every alternate ketide unit, suggesting an alternative use of two modules wherein one module would condense a branched C3 ketide unit and the next would add a C2 unit. Site mutation and radiolabeled experiments confirmed that the AT1 prefers MMCoA while the AT2 recruits MCoA.

On the basis of biochemical, computational, mutagenic, analytical ultracentrifugation and atomic force microscopy studies, it was proposed that PKS12 protein is organized as a large supramolecular assembly mediated through specific interactions between the C-and N-terminus linkers. This interaction enables the intermolecular acyl chain transfer between two proteins. PKS12 protein thus forms a modular assembly to perform repetitive condensations analogous to iterative proteins. This novel intermolecular iterative biosynthetic mechanism provides new perspective on the

understanding of polyketide biosynthetic machinery and also suggests new ways to engineer polyketide metabolites.

3. Iterative mechanism

Although more and more non-colinear PKSs have been discovered, the mechanism that control the iteration processes in bacterial type I iPKSs has not been clearly elucidated[89]. A possibility could be that iteration is an intrinsic feature of the module like FAS and fungal PKSs. Moreover, a "force" coming from the downstream, for example, the KS domain which only accepts an intermediate with a particular chain length, restricts the upstream module to iterate. A combination of both is also possible. Finally, the timing of iteration and extender unit recruiting of the ketosynthase need to be coordinated to allow for an unobstructed metabolite flux.

3-1. Interchenar transfer vs. intrachenar transfer

The mechanism of fungal iPKS is summarized as: 1) KS domain catalyzes the Claisen condensation reaction between the starter and extender units, driven by decarboxylation of the extender unit. 2) The intermediate is passed back to the KS domain to be extended by another ketide unit, or 3) the finished polyketide chain will be released from the enzyme if the polyketide has reached its predetermined length, which is unique for each iPKS. The chain length is tightly controlled by a special product template (PT) domain in non-reducing fungal PKS[36].

In principle, the iteration might happen under the following scenarios: the intermediate is transferred from the ACP domain of one PKS chain to the (a) ACP domain on the same PKS chain ('intrachenar'); (b) ACP domain on the opposite PKS chain ('interchenar'); (c) KS domain upstream on the same PKS chain ('intrachenar');

or (d) KS domain upstream on the opposite chain (‘interchenar’) (Figure 1-8).

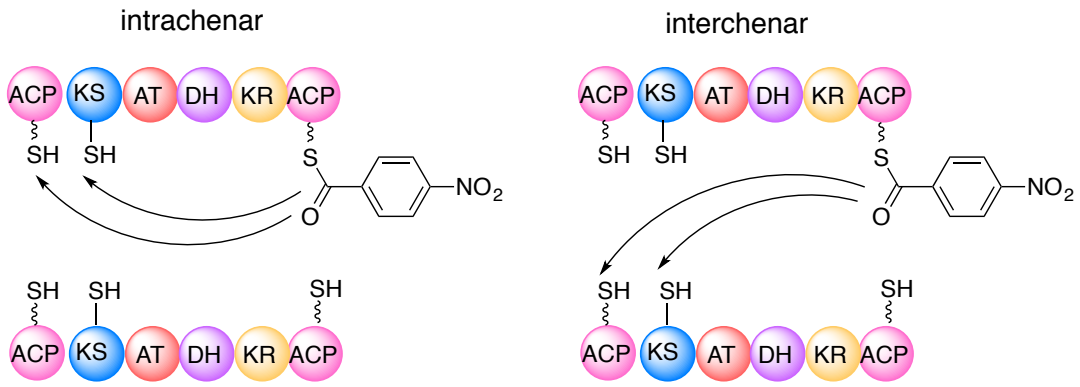


Figure 1-8. Mechanisms of intrachenar transfer and interchenar transfer.

Through targeted domain mutagenesis, cross-complementation experiments, and metabolic profiling carried out by Hertweck and co-workers, the mechanism of iterative events in a bacterial type I PKS has been addressed[87]. They revealed that the N-terminal ACP is not involved in the iteration process, ruling out an ACP−ACP shuttle (Figure 1-9). This is also consistent with the fact that most bacterial iPKSs don’t have an N-terminal ACP. Furthermore, an *aurA* (ΔKS, ΔACP) and *aurA* (ΔAT) heterodimer proved to be nonfunctional, whereas aureothin production was restored in a Δ*aurA* mutant complemented with *aur*A (ΔKS) − *aur*A (ΔACP). This finding supports that the most likely scenario is (d), a retrotransfer of the biosynthetic intermediate from the ACP onto the KS domain located on the opposite polypeptide strand.

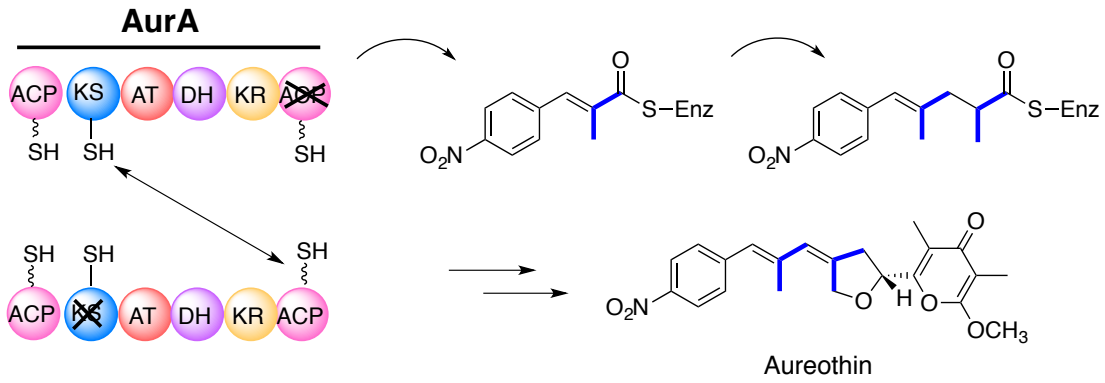


Figure 1-9. Interchenar retrotransfer of intermediate from ACP to KS in aureothin biocynthesis.

4. Summary

PKSs are generally classified into three canonical types according to their domain alignment and organizations. With the increasing number of PKSs falling into the category of "non-canonical" PKSs, the boundaries between the different types are no longer so well defined. These non-canonical type I PKS modules expand the diversity of natural product biosynthesis, making engineering and rational design in synthetic biology even more complicated. Iteration is likely the most unpredictable event that occurs in polyketide biosynthesis. Programmed iteration probably has great potential to be engineered to produce polyketides with designed length and modifications.

The iterations can happen with one module in the giant modular assembly line. The number of iterations can be controlled when using specific module, such as the case of the AurA module, which can be repetitively used for four times until selected by the next KS gatekeeper. The bi-modular iteration can be achieved by engineering

PKS12 in MPM biosynthesis. Upon switching the AT domain, one might change the substrate specificity of the 2 modules, creating the designed backbone for the target compounds. Wholly iterative PKSs, including bacterial iPKSs, PTM iPKS, PUFA and enediyne biosynthetic PKS, have the potential to provide the core structures of the target compounds, which could also be engineered to produce rationally-designed complex polyketides.

However, the mechanistic details for bacterial type I iterations are less understood. Although several groups have recently begun to investigate the mechanism of the iteration and the timing of retrotransfer, progress still needs to be made to understand these processes and to translate the knowledge into rational PKS engineering. At the same time, these iterative events reveal that nature recruits “unexpected” mechanisms to synthesize novel polyketide natural products. With the rapid growth of genome sequence data and in depth biochemical studies of the new PKSs, novel iteration mechanisms will continue to emerge.

Chapter 2: Iterative Assembly of Two Separate Polyketide Chains by the Same Single-module Bacterial Polyketide Synthase in the Biosynthesis of HSAF

Portions of this chapter were reported (Angewandte Chemie International Edition, 2014, 53 (29), 7524-7530) and are reproduced with the permission of the publisher[26].

2-1. Background and Significance

Fungal infections, including meningitis, thrush, ringworm, eye and nail infections, lung infections and those topical invasive and allergic diseases in humans and animals, represent an increasing threat to public health. According to the latest data, the mountain yellow-legged frog population has almost been wiped out in Kings Canyon National Park in California due to fungal infections, while 40% of the amphibian species have been lost in some areas in Central America, which has led to an estimated loss of at least $3.7 billion per year to U.S. agriculture[90].

In agriculture, fungal infections are a major threat to the food supply chain for billions of people. In many underdeveloped countries, contamination of crops such as maize and groundnuts has led to loss of crops and severe illness in both human and farm stock where contaminated food is ingested. For example, *Fusarium graminearum*, an important plant pathogen, is the major reason of *fusarium* head blight in wheat. This is one of the most serious wheat diseases in Midwest and central Canada. Its metabolite vomitoxin (also known as deoxynivalenol, DON) of *F. graminearum* belongs to a class of mycotoxins, tricothecenes, which are strong protein inhibitors. Contaminations of grains cause loss of billions of dollars every year.

Current therapeutic approaches for fungal infections rely mainly on antifungal drugs that target ergosterol of cell membranes and glucans of cell walls. However, the

over-reliance on limited targets has led to a rapid increase in drug resistance. Thus, it is important to continue the discovery and development of new antibiotics.

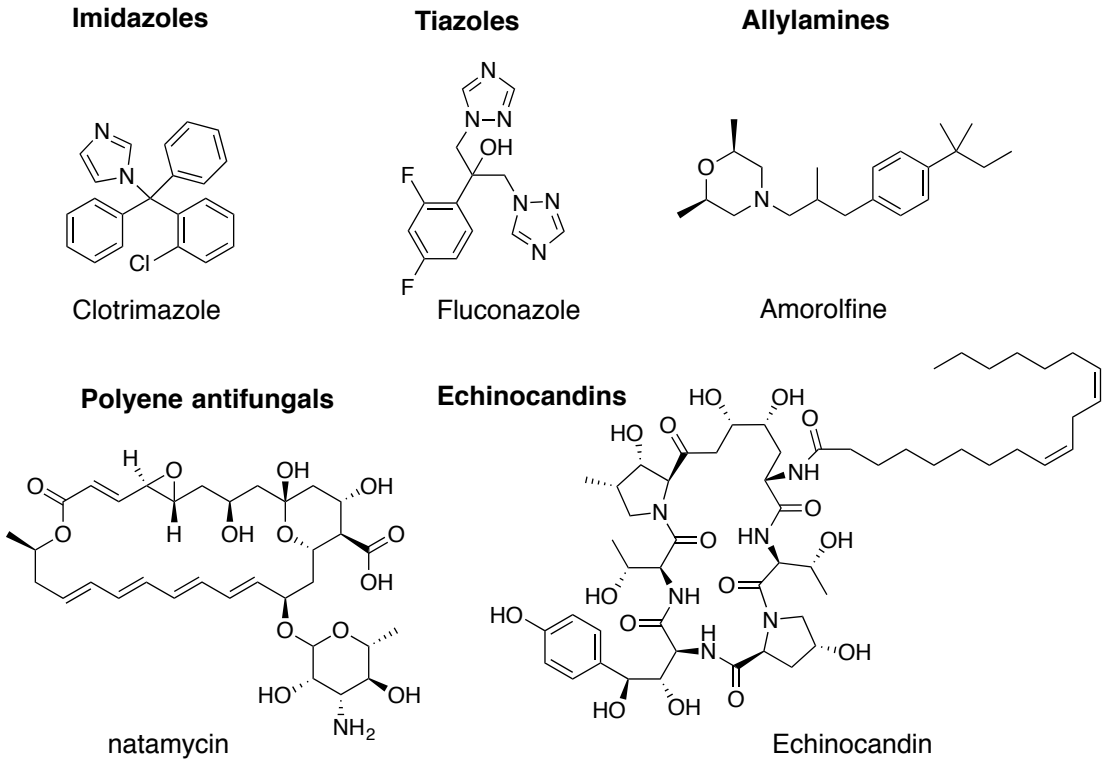


Figure 2-1. Classification of antifungal drug and typical examples.

Natural products are the most prolific source of anti-infectives and anticancer agents. About 70-80% of anti-infectives and anticancer drugs can be traced to or were inspired by natural products. Traditionally, soil bacteria, especially *Streptomyces*, have been the primary source for anti-infectives. More recently, marine organisms are also emerging as new sources for bioactive natural products with new structures and activities. However, the success has been limited partly due to difficulties with access to the marine organisms. In contrast, many "user-friendly", prolific natural product producers from terrestrial environments, other than the Gram-positive *Streptomyces,*

remain largely unexplored, such as the Gram-negative environmental bacteria *Lysobacter*[26, 78].

A number of *Lysobacter* species are now recognized as microbial predators with an impressive arsenal of bioactive small molecules, which gives them potential both as biocontrol agents and producers of promising drug leads.

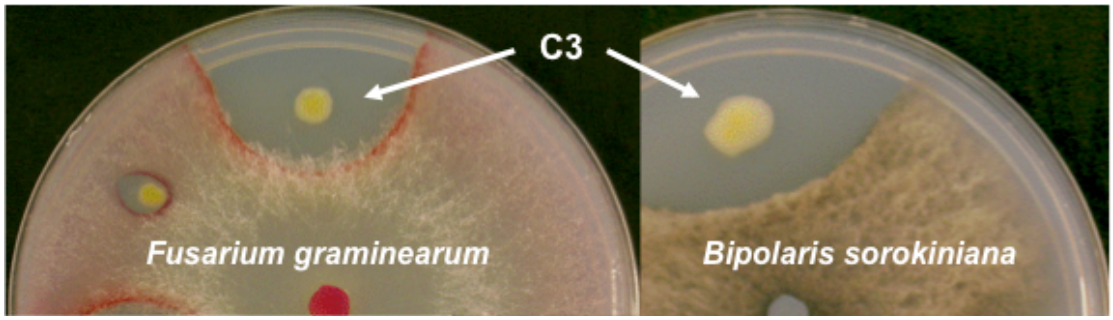


Figure 2-2. *Lysobacter enzymogenes* strain C3 and its antifungal activity. (These photos were taken by Prof. Gary Yuen at Department of Plant Pathology, UNL)

Our lab has been using *Lysobacter enzymogenes* strain C3 as a new source for bioactive natural products discovery. *Lysobacter enzymogenes* strain C3 (previously named *Stenotrophomonas maltophilia* strain C3) was originally isolated from grass leaves[77]. This strain is used as a biocontrol agent against fungal pathogens of plants. C3 exhibits wide antifungal activity; it can inhibit the growth of many fungi, such as *Bipolaris sorokiniana* and *Fusarium graminearum*. Our previous screening for antifungal compounds from *L. enzymogenes* strain C3 has led to the isolation of heat-stable antifungal factor (HSAF)[26, 78, 91, 92].

HSAF belongs to the polycyclic tetramate macrolactam (PTM), which is emerging as a new family of natural products with a distinct structure and new mode of action. HSAF contains a 5,5,6-tricyclic system fused with a 17-membered macrolactam that

includes a tetramate acid moiety. Tetramic acid is the key structural feature for many bioactive heterocycles that exhibit a wide range of biological activities including antibiotic and anticancer activities.

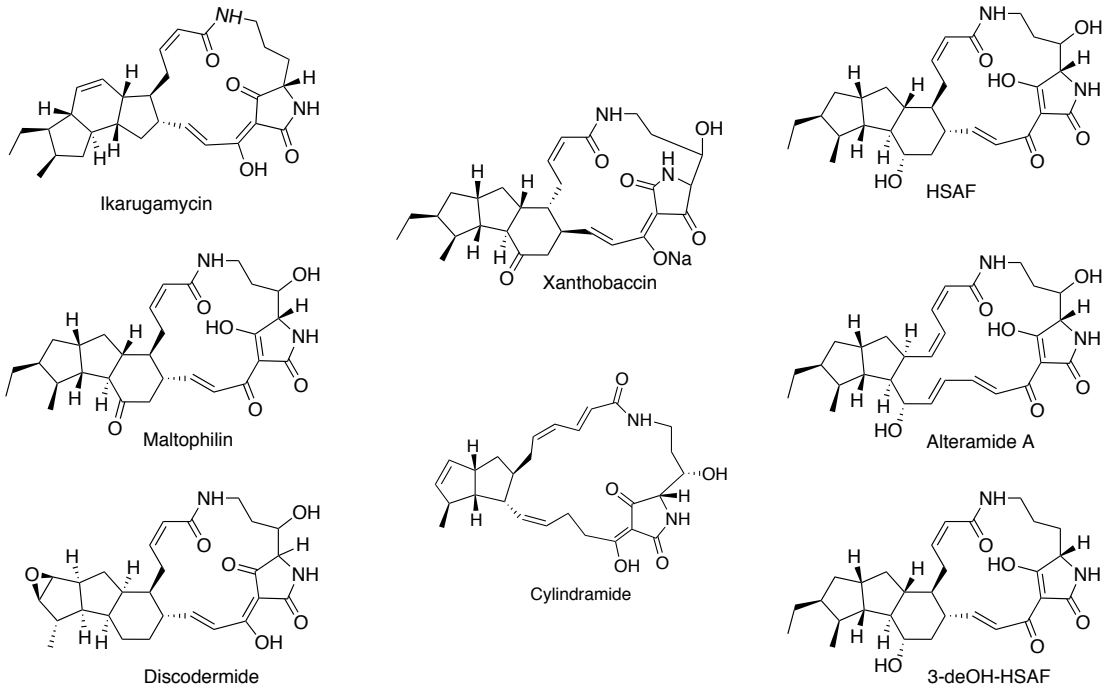
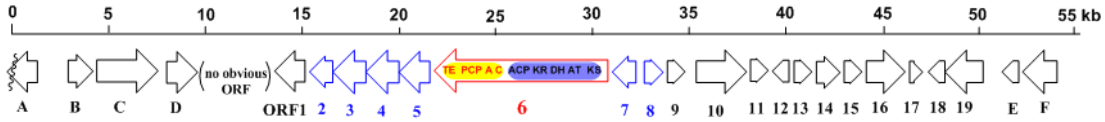


Figure 2-3. Chemical structure of PTMs.

The results of mode of action test indicated that HSAF specifically interferes with the biosynthesis of sphingolipids[78]. Sphingolipids represent an attractive new target for the development of novel antifungal drugs because their structure in fungal cells is distinct from that in mammalian cells. This novel mode of action together with the unusual chemical structure that is distinct from any existing antifungal drug on the market, make HSAF a potentially promising drug lead.

Through our previous study[76], the genetic locus responsible for HSAF biosynthesis in *L.e.* C3 and *L.e.* OH11 was identified. A series of disruption experiments defined 10

HSAF core genes, including a hybrid PKS-NRPS, making up a total of nine domains, and 6 flanking redox genes.



ORF	# aa	Blast Homolog
ORF1	496	Major facilitator superfamily (MSF) transporter
ORF2	351	OX4 (alcohol dehydrogenase zinc-binding protein)
ORF3	575	OX3 (FAD-dependent oxidoreductase)
ORF4	578	OX2 (FAD-dependent oxidoreductase)
ORF5	547	OX1 (FAD-dependent oxidoreductase)
ORF6	3132	PKS-NRPS (KS-AT-DH-KR-ACP-C-A-PCP-TE)
ORF7	353	FA hydroxylase/sterol desaturase
ORF8	257	Ferredoxin NADP reductase/FAD/NADP-binding protein
ORF9	318	Arginase
ORF10	840	TonB-dependent outer membrane receptor
ORF11	174	Competence-damaged protein
ORF12	313	MarR family transcriptional regulator
ORF13	220	LexA repressor
ORF14	350	Recombinase A
ORF15	291	Regulatory protein RecX
ORF16	877	Alanyl-tRNA synthetase
ORF17	67	Carbon-storage regulator, global regulator
ORF18	243	Two-component system regulatory protein
ORF19	906	Two-component system sensor protein
ORF-A	~873	DNA Mismatch repair protein (incomplete)
ORF-B	362	SGNH-hydrolase (lipase/esterase)
ORF-C	1017	Apolipoprotein *N*-acyltransferase
ORF-D	471	Amino acid permease
ORF-E	199	K^+-transporting ATPase subunit C
ORF-F	687	K^+-transporting ATPase subunit B

Figure 2-4. The dihydromaltophilin (HSAF) biosynthetic gene cluster. The "HSAF core" is highlighted in red for PKS-NRPS and blue for the redox enzymes[76].

We previously expressed and purified the NRPS protein[76]. The in vitro characterization demonstrated that the NRPS is responsible for forming the two amide bonds as well as the carbon-carbon bond between ornithine's carbonyl carbon and one hexaketide's α carbon, forming the expected tetramate product (Figure 2-5)[76]. The SD enzyme was able to convert a 3-deOH HSAF to HSAF in the presence of NADPH *in*

vitro[93]. The result demonstrated that the SD gene is the 3-hydorxylase whose activity could be significantly enhanced when co-expressed with the ferredoxin reductase gene (Figure 2-6).

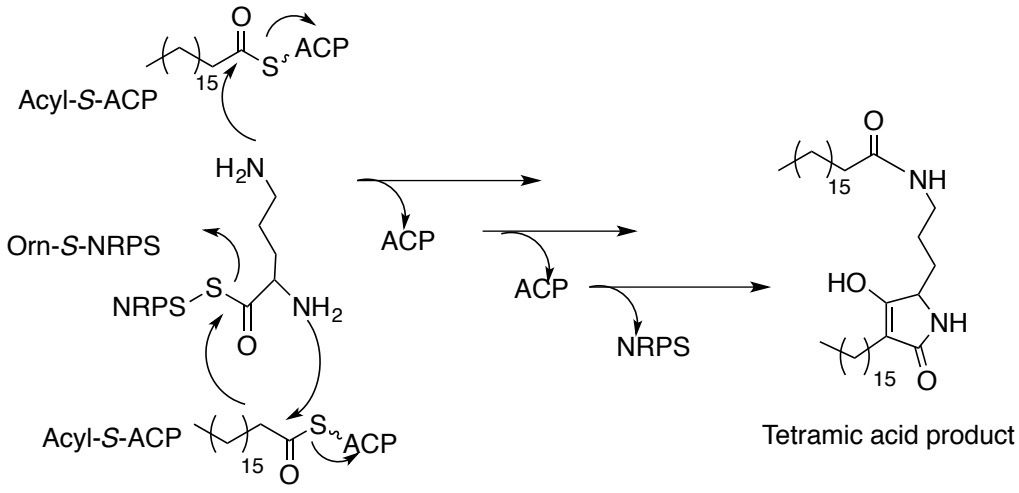


Figure 2-5. In vitro synthesis of a tetramate product using the purified dihydromaltophilin NRPS. The acyl-*S*-ACP substrate was prepared by incubating ACP and stearoyl-CoA with Svp, and Orn-*S*-NRPS was prepared by incubating NRPS and CoA with Svp, followed by addition of L-ornithine and ATP.

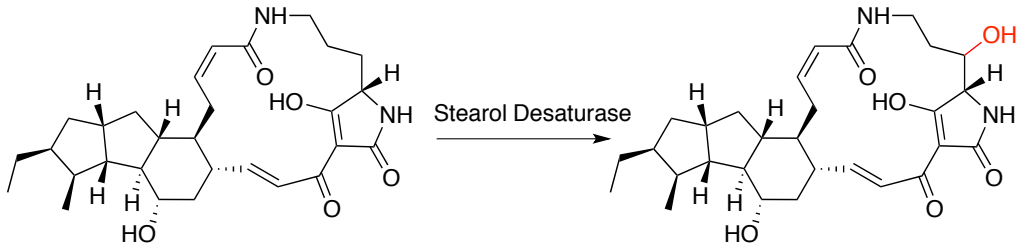


Figure 2-6. Hydroxylation of the carbon-3 of 3-deOH HSAF by Sterol Desaturase.

The most striking feature of the cluster is that only a single module PKS-NRPS is present, although the biosynthesis of HSAF apparently required two separate hexaketide chains that are linked together by one amino acid ornithine via two amide

bonds. This suggests that the same PKS module would act not only iteratively, but also separately. However, prior to our work, typical bacterial modular PKSs were known to be non-iterative, and there had been no report of a bacterial iterative PKS that was capable of making two separate polyketide chains.

Clearly, there is a novelty in the molecular logic of HSAF biosynthesis in which nature employs known "old" enzymes to form distinct "new" structures. The goal here is to prove that the single-module PKS-NRPS is sufficient for two separate polyketide chains and the tetramate macrolactam, which is fused to a 5,5,6-tricyclic system.

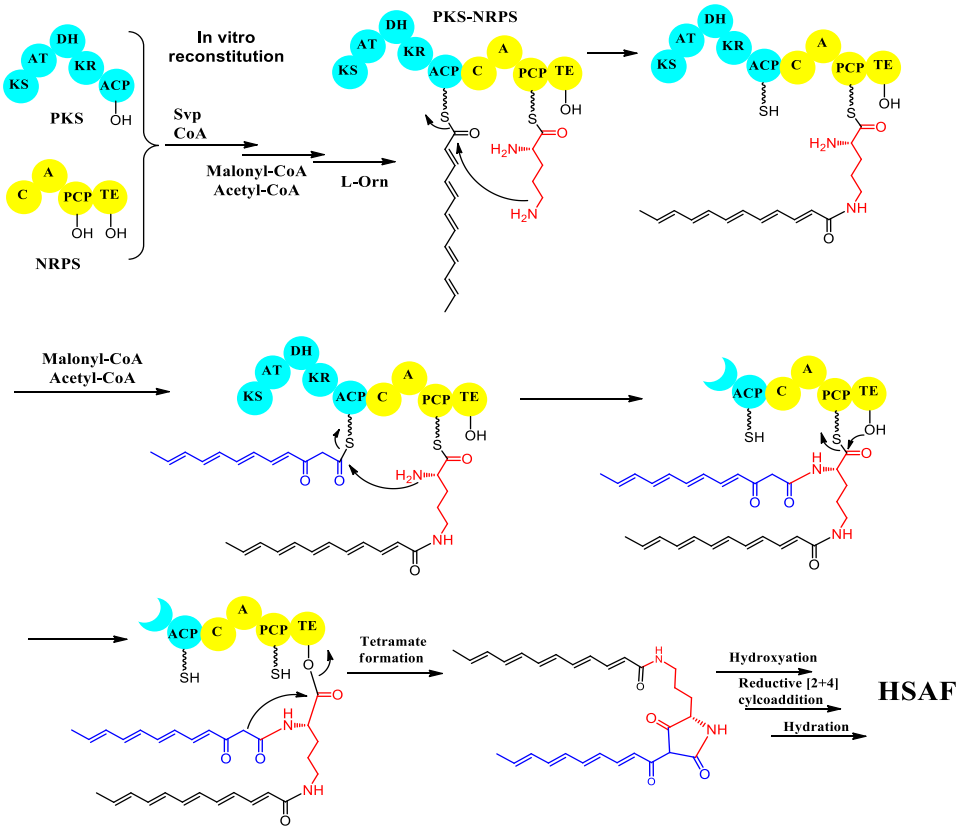


Figure 2-7. Hypothesis for formation of the key intermediate: polyene tetramate formation[26].

2-2 Materials and Methods

2-2-1. Strains, Vectors, Chemicals, and Molecular Biology Agents

Escherichia coli DH5α and EPI300 (Epicenter Biotechnologies, Madison, WI) were used as the hosts for general plasmid DNA propagation. *E. coli* SG13009 (pREP4) were used for protein expression. *E. coli* BW25113/pKD46 was used for Red recombination. *E. coli* ET12567/pUZ8002 was used as the conjugal strain. *Streptomyces sp.* LZ35 was isolated from the intertidal soil collected at Jimei, Xiamen, China[94]. Strain SR107 was derived from *Streptomyces sp*. LZ35 by deletion of four PKS gene clusters[95]. Strain ZM12 was derived from *S. coelicolor* by deletion of all ten native PKS and NRPS gene clusters present in its genome[96]. Strains SR107 and ZM12 were used as the hosts for the heterologous production of HSAF. Vector pANT841 was used for general cloning, vector pQE60 was used for protein expression in *E. coli* SG13009 (pREP4), and vector pSET152 was used for integration of the HSAF gene cluster into *Streptomyces* hosts. Chemicals were purchased from Sigma or Fisher Scientific. PCR primers were synthesized by Integrated DNA Technologies (IDT, Coralville, IA). Kits for plasmid preparation and DNA extraction were from Qiagen (Valencia, CA). Standard molecular biology methods were used for all other DNA manipulations.

2-2-2. HSAF polyketide synthase (PKS) expression in *E. coli*

The procedure for constructing the entire PKS gene (5526 bp) for expression in *E. coli* is illustrated in Figure 2-8. Two of the three pieces that make up the complete PKS gene were amplified by PCR using Cos4'-1 as the template. The following two pairs of primers were used for the PCR:

P1, KS-*Nco*I-Forward: 5'-GAG ACC ATG GAG GAC CGC ATC GCC A-3'

P2, KS-*Eco*RI-Reverse: 5'- CAT CAT CGT GCC GGT GGC GGT G-3'

P3, ACP-*Kpn*I-Forward: 5'- ACC GCG GCG ATG CGG TCG AAC-3'

P4, ACP-*Bam*HI-Reverse: 5'- GGG AAG GAT CCC AGC GCG TTC TGG T-3'

Amplification of the entire PKS gene (5526bp) failed after multiple trials. The construction of PKS expression vector is shown in Figure 1-12.

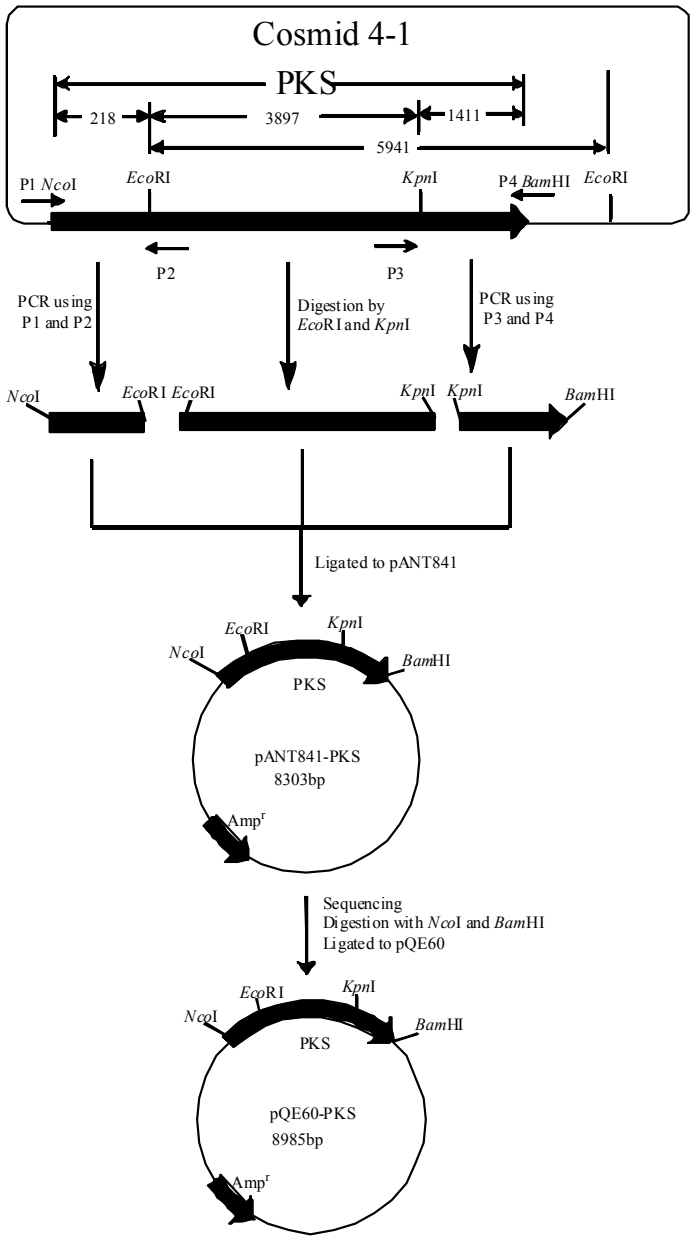


Figure 2-8. The construction of PKS expression vector.

The KS-fragment (218 bp) was digested with *Nco*I and *Eco*RI; the ACP-fragment (1411 bp) was digested with *Kpn*I and *Bam*HI. Both of the fragments were cloned into pANT841 at the same sites. The middle fragment was obtained by digestion of cosmid COS4-1 with *Eco*RI, followed by digestion of the resulting selected fragment (5941 bp) with *Kpn*I. The final targeted fragment (3897 bp) was cloned into pANT841 at the same sites. The construct was sequenced, and the result showed that the three fragments were ligated together and the entire PKS gene sequence was correct. The PKS gene was released from pANT841 as a *Bam*HI/*Nco*I fragment and cloned into expression vector pQE60.

2-2-3. Expression and purification of PKS protein.

To express the PKS, the pQE60 construct was introduced into *E.coli* SG13009. Single colonies were inoculated in 3 ml of liquid LB medium containing ampicillin (50μg/ml) and incubated in a shaker (250 rpm) at 37°C overnight. The overnight culture was added to 50 ml fresh LB medium and incubated in a shaker (250 rpm) at 37°C until the cell density (OD600) reached 0.6. To induce the expression of the PKS, IPTG (0.3mM) was added to the culture, and the cells were allowed to grow at the room temperature for another 12 hours. To prepare the soluble fraction, the cells were harvested and resuspended in 2 ml of PBS buffer (500 mM NaCl, 50 mM Na_2HPO_4, 50 mM NaH_2PO_4, pH 7.8). The cell suspension was sonicated on ice. The soluble fraction of protein extracts was loaded onto a Ni-NTA column which was previously calibrated with PBS buffer containing 10 mM imidazole. The column was washed three times with the buffer containing 20 mM imidazole, and the C-His6-tagged PKS

protein (199.8 kDa) was purified by using an imidazole step-gradient as described by the manufacturer's protocol. The purity of the protein was analyzed by SDS-PAGE, and the fractions containing purified protein were pooled, concentrated, and dialyzed against 50 mM NaCl, 50 mM Na_2HPO_4, 50 mM NaH_2PO_4, pH 7.8, containing 15% glycerol. Finally the protein solution was frozen in liquid nitrogen and stored at -80 °C until use.

2-2-4. *In vitro* assay for the activity of the PKS module

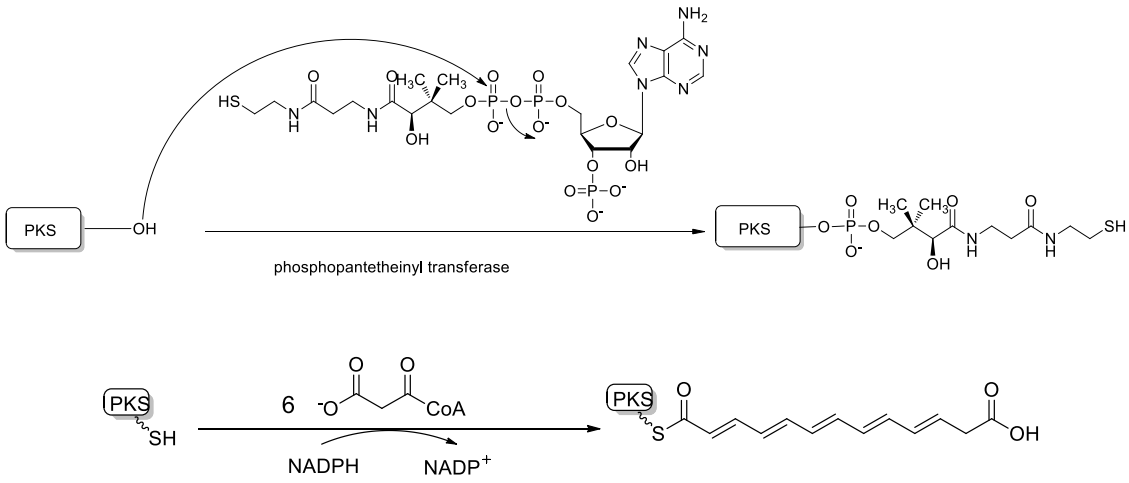


Figure 2-9. Reactions of PKS activation and polyketide chain assembly.

The first step was to convert the purified PKS to its *holo* form, through 4'-phospho-pantetheinylation at the active site serine residue of its ACP domain. This was carried by Svp[97], a promiscuous 4'- phosphopantetheinyl transferase (PPTase). The reaction contained the PKS protein (25 μM), CoA (2 mM) and Svp (10 μM), in a 45 μl solution of Tris-HCl (100 mM, pH 8.0), $MgCl_2$ (10 mM), and TCEP (0.5 mM). After incubation at 37 °C for 2 h, the reaction was stopped by boiling for 10 min. Due to the huge size (199.8 kDa) of the PKS protein, a trypsin digestion procedure was

employed prior to the ESI-Q-TOF analysis of the mass of the PKS. The protein was loaded to a SDS-PAGE (7.5%), and the PKS band was excised from the gel and subjected to trypsin digestion[98]. Briefly, gel pieces were digested by trypsin (no.V5111, Promega, Madison, WI) and digested peptides were extracted in 5% formic acid / 50% acetonitrile and separated using C18 reversed phase LC column (75 micron × 15 cm, BEH 130 , 1.7 micron Waters, Milford, MA). A Q-TOF Ultima tandem mass spectrometer couple with a Nanoaquity HPLC system (Waters) with electrospray ionization was used to analyze the eluting peptides. The system was user controlled employing MassLynx software (v 4.1, Waters) in data-dependent acquisition mode with the following parameters: 0.9-sec survey scan (380–1900 Da) followed by up to three 1.4-sec MS/MS acquisitions (60–1900 Da). The instrument was operated at a mass resolution of 8000. The instrument was calibrated using the fragment ion masses of doubly protonated Glu-fibrinopeptide. The peak lists of MS/MS data were generated using Distiller (Matrix Science, London, UK) using charge state recognition and de-isotoping with the other default parameters for Q-TOF data. Data base searches of the acquired MS/MS spectra were performed using Mascot (Matrix Science, v2.2.0, London, UK). The NCBI database (20100701) was used in the searches. Search parameters used were: no restrictions on protein molecular weight or pI, enzymatic specificity was set to trypsin with up to 3 missed cleavage sites, carbamidomethylation of C was selected as a fixed modification, and methionine oxidation was selected as a variable modification. Mass accuracy settings were 0.15 dalton for peptide mass and 0.12 dalton for fragment ion masses. The *apo*-PKS, *holo*-PKS, and the various forms of acylated PKS were analyzed by the MS method, which had been successfully used in other studies[99, 100].

To test the activity of the PKS, the reaction mixture (45 μl) containing the *holo*-PKS was added with malonyl-CoA (10 mM), NADPH (10 mM) and acetyl-CoA (2 mM), and the final volume was brought to 50 μl. The reactions were incubated at 37 °C for 20 h. A reaction containing boiled PKS protein served as the control. The reactions were stopped by boiling for 10 min, and a 10 μl aliquot of the reaction mixtures was loaded on SDS-PAGE (7.5%), and the PKS band was cut out for trypsin digestion. Since the acyl chain synthesized by the PKS could have various chain lengths and oxidation levels, the samples were subject thorough search for various potential acylated products.

2-2-5. Expression of PKS with a point-mutated KS domain in *E. coli* and activity assay of the synthase

The *in vitro* data obtained from the purified PKS suggest that the synthase may use malonyl-CoA as both the starter and the extender in the polyketide chain synthesis because a carboxylate appeared present in the intermediate (Figure 2-18). To probe the starter specificity, we generated another PKS expression construct, in which the active site cysteine in the KS domain of this PKS was mutated to alanine (C176A, with the protein accession number ABL86391). Primer extension PCR was performed to generate a 2151-bp fragment of PKS gene with C176A mutation. The following two pairs of primers were used to amplify a 359-bp fragment and a 1832-bp fragment (Figure 2-19):

KS (C/A) up-fw: 5'-CGC GGC GGC TAC ATC GAC GGC TTC GAC G-3'

KS (C/A) up-rv: 5'-GGC GAC CAG GGA GGA ACT AGC GGC GGT GTC GAT CGA CAG GCT-3'

KS (C/A) down-fw: 5'-AGC CTG TCG ATC GAC ACC GCC GCT AGT TCC TCC CTG GTC GCC-3'

KS (C/A) down-rv: 5'-GTA GAG CTC GTG GCC GGC GGC GAT GCC-3'

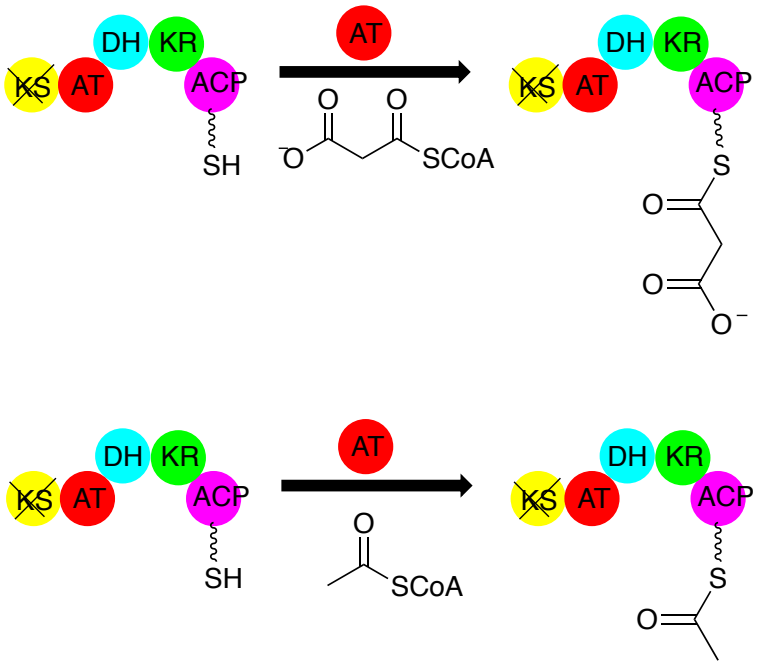


Figure 2-10. Mechanism of determination of the starter unit by active-site point-mutation in the KS domain.

After purification of the PCR fragments, an overlapping PCR was carried out using these 2 fragments as template. The resulted 2151-bp fragment was digested with *Eco*RI and *Sac*I, and cloned into previously constructed pANT841-PKS (Figure 2-12) at the same sites. The mutated PKS gene (mPKS) was released from pANT841 as a *Bam*HI/*Nco*I fragment and cloned into expression vector pQE60 (Figure 2-19). The subsequent expression and purification are the same as for the wild type PKS.

To assay the activity, mPKS (25 μM) was first converted into its *holo* form, by incubating with CoA (2 mM) and Svp (10 μM) in a 50 μl reaction containing Tris-HCl (100 mM, pH 8.0), $MgCl_2$ (10 mM), and TCEP (0.5 mM). After incubation at 37 °C for 2 h, the reaction mixture was added with 10 mM malonyl- CoA, 2 mM

acetyl-CoA, and 10 mM NADPH. A reaction containing the wild type PKS served as the control. After continual incubation at 37°C for 20 h, the reactions were stopped by boiling. A 10 μl aliquot of the reaction mixtures was loaded on SDS-PAGE for analysis. The mPKS band was cut from the gel and subject to trypsin digestion and Q-TOF MS analysis as described above.

2-2-6. *In vitro* reconstitution of the PKS module and the NRPS module

The PKS module was activated to its *holo*-form and incubated with proper substrates as described in the reactions above. The expression and purification of the NRPS module had been described previously[76]. The purified NRPS module was also converted to its *holo* form by incubating with CoA (0.83 mM) with Svp (5.6 μM) in a 60 μl reaction containing Tris-HCl (100 mM, pH 8.0), $MgCl_2$ (10 mM), and TCEP (0.5 mM). The reactions were incubated at 37 °C for 2 h. The NRPS reaction mixture and the PKS reaction mixture were then combined into a tube, which had a 40 μl solution containing L-Orn (1.5 mM), ATP (3 mM), Tris-HCl (100 mM, pH 8.0), $MgCl_2$ (10 mM), NaCl (50 mM), EDTA (0.1 mM), and TCEP (0.5 mM). For control, the reaction contained all the components of the reconstitution reaction except that NRPS was replaced with the total proteins of the *E. coli* host strain that was transformed only with the blank expression vector pQE60. After continual incubation for 5 h at 37 °C, the reactions were stopped by adding 150 μl of 0.2 mM TCA in methanol and were frozen at -20 °C for 30 min. The mixtures were centrifuged at 13,200 rpm for 20 min in a desktop Eppendorf centrifuge, and the supernatants were transferred to new tubes. The solutions were dried in a Speed-Vac, and the residues in the tubes were re-dissolved in 150 μl methanol. The methanol extracts were centrifuged, and supernatants were transferred to new tubes and dried again. Finally,

the residues in the tubes were re-dissolved in 300 μl 50% methanol and analyzed by LC-MS. The sample (50 μl) was loaded to a Waters BEH column (C18 1.7 μm, 1.0 × 100 mm), and eluted with the following solvent program: Solvent A was H_2O containing 0.1% formic acid; Solvent B was acetonitrile containing 0.1% formic acid. The flow rate was 50 μl/min with a binary gradient system (0-15 min, 5%-45% B gradient; 15-35 min, 45% B to 90% B gradient; 35-46 min, 90% B to 100% B gradient; 46-55 min, 100% B; 55-56 min, 100% B to 5% B gradient; 56-66 min, 5% B). The mass spectrometer was Mass Spec Waters Snapt G2-S Q-TOF, operated in positive ion high resolution mode with lock mass, mass range 100-1200 amu, 1sec scans. The LC-MS was set to scan for various masses of potential products that could be synthesized by the reconstituted PKS-NRPS and then released into the medium by the thioesterase (TE) domain in the NRPS module[76, 79, 92, 101]. One eminent peak at 15.48 min was detected in the reconstituted reaction, but no in the control reaction, when the scan was at m/z 475.2597 (Figure 2-23). This peak was subject to HR-ESI-MS analysis and gave m/z 475.2572 (Figure 2-23). This mass is coincident with the expected $[M+H]^+$ for the polyene tetramate intermediate.

2-2-7. Construction of HSAF expression vectors and heterologous production in *Streptomyces*

To subclone the HSAF gene cluster into the integrative vector pSET152, two restriction sites, *Acl*I and *Bam*HI, were introduced into the vector by insertion of two DNA fragments (Figure 2-11). The two fragments (1.4 and 1.7 kb) were amplified from the cosmid clone, Cos4'-1, which contains the entire HSAF gene cluster isolated from *Lysobacter enzymogenes* C3[78]. The following two pairs of primers were used for the PCR:

HSAF-LF: 5'-GAA GAT CTC GGG AAA CTC CGA GGC AGA AAT-3'

HSAF-LR: 5'-AGG AAT TCC GGT GAT GCA AGG GCT GTC GA-3'

HSAF-RF: 5'-AGG AAT TCA ACA CCT GGA GTT CTT CAG CAT CA-3'

HSAF-RR: 5'-GAA GAT CTA AAT CAG CGC CGG TGG CGT TG-3'

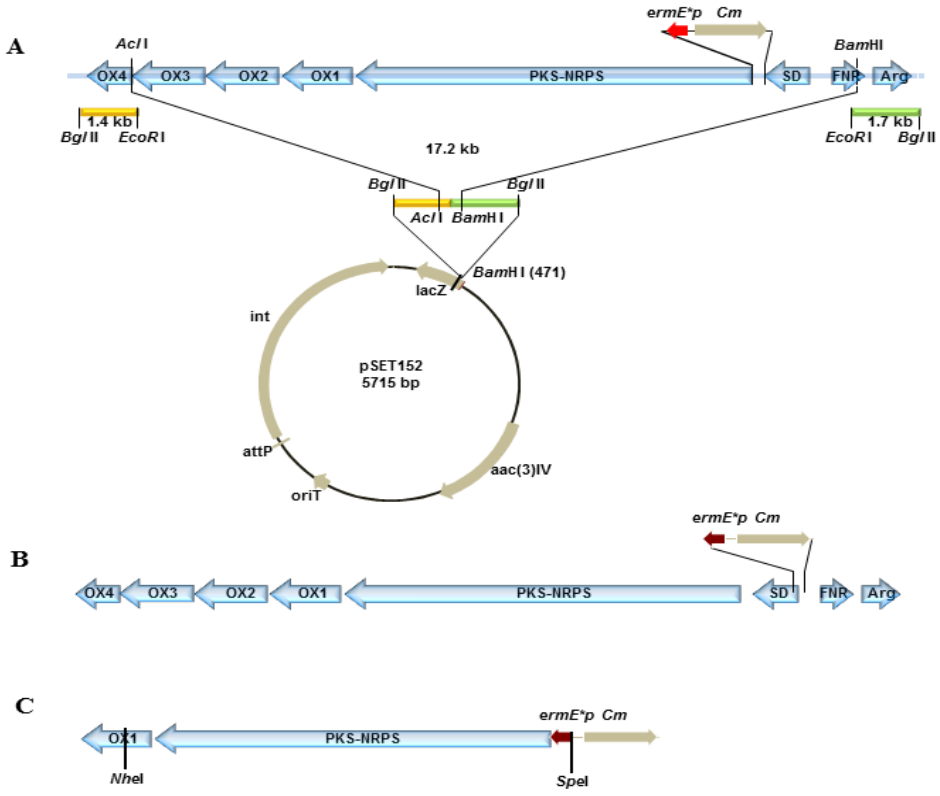


Figure 2-11. Strategy for constructing the HSAF expression vectors for the production of HSAF analogs in *Streptomyces* hosts. A, construct pSETHSAF3 that resulted in the recombinant *Streptomyces* strains SR107HSAF1 and ZM12HSAF1; B, construct pSETHSAF4 that resulted in the recombinant *Streptomyces* strain SR107HSAF2; C, construct pSETHSAF5 resulting in the recombinant *Streptomyces* strain SR107PKS/NRPS.

The resulting PCR products were digested with *Bgl*II and *Eco*RI, respectively, and ligated to the *Bam*HI- digested pSET152 to create pSETHSAF1. Then, a 17.2 kb

fragment harboring the bulk of the HSAF gene cluster was cut off from Cos4'-1 by digestion with *Acl*I and *Bam*HI and inserted into the same sites of pSETHSAF1 to yield pSETHSAF2 (Figure 2-11).

The λ Red mediated recombination in *E. coli* BW25113/pKD46[102] was used to replace the native promoter of the PKS-NRPS gene or of the SD gene with *Streptomyces ermE** promoter[103, 104]. About 100 ng of pSETHSAF2 was transformed into *E. coli* BW25113/pKD46 by electroporation at 2500 V, 25 μF, and 200 Ohms. Clones were selected on solid LB agar with 100 μg/ml ampicillin and 30 μg/ml apramycin. The transformants were cultured in LB medium at 30 °C and induced by addition of 10 mM L-arabinose for the expression of the red genes. Two fragments including the chloramphenicol-resistance cassette (Cm) and the *ermE** promoter were amplified from pUC19-Cm-PermER* (unpublished data) using the following two pairs of primers, respectively:

ER2HSAFF: 5'-gaa cca gtt tgt tcc aga tca cgg ccg gaa cgc gtt ccg ACA GCT TAT CAT CGA ATT TC-3'

ER2HSAFR: 5'-gcg gca gcc gat gcc gat gat ggc gat gcg gtc ctc cat ATG TCC GCC TCC TTT GGT CG-3'

CmER2SDF: 5'-cgc cgc ggc cta gcc gcg aat att cgc cgg cat tat ctg ACA GCT TAT CAT CGA ATT TC-3'

CmER2SDR: 5'-agc acg cgt ttc caa agt gtg ctc tga taa tcc gtt cat ATG TCC GCC TCC TTT GGT CG-3'

(note: the bases in lower case are the sequence for Red recombination, and the bases in upper case are for amplification of the target genes) Each of the fragments (~ 200 ng) was electroporated into *E. coli* BW25113/pKD46/pSETHSAF2 to generate pSETHSAF3 and pSETHSAF4, respectively. Transformants were selected for resistance to chloramphenicol at 25 μg/ml. In addition, the 10 kb *Spe*I /*Nhe*I fragment from pSETHSAF3 was inserted into the *Xba*I site of pSET152 to create pSETHSAF5,

which only contained the PKS/NRPS gene with the *ermE** promoter. The constructs pSETHSAF3-5 were individually transformed into *E. coli* ET12567/pUZ8002, which were used to mate with *Streptomyces sp.* SR107 and ZM12. This intergenetic conjugation resulted in the transfer of three different HSAF expression constructs into *Streptomyces* hosts, construct pSETHSAF3 leading to the recombinant strains SR107HSAF1 and ZM12HSAF1, construct pSETHSAF4 to strain SR107HSAF2, and construct pSETHSAF5 to strain SR107PKS/NRPS.

2-2-8. Production and analysis of the metabolites in *Streptomyces* strains

For production of secondary metabolites, all *Streptomyces* strains were inoculated on petri dishes containing the Fermentation Medium (soluble starch 25 g, soybean 15 g, yeast extract 2 g, agar 20 g, pH 7.2, 40 ml) and cultivated at 28 °C for 11 days. The whole solid cultures were diced and extracted with AcOEt/MeOH/AcOH (80:15:5, v/v/v) at room temperature, and the crude solution was concentrated under reduced pressure. The dried extract was dissolved in 2 ml methanol, and a 20 μl aliquot of each extracts was analyzed by HPLC (Agilent 1200, ZORBAX Eclipse XDB-C18, 4.6×250 mm, 5 μ). For SR107HSAF1-2 and ZM12HSAF1, chromatographic conditions were as follows: solvents-A, water-0.04% trifluoroacetic acid (TFA); solvent-B, acetonitrile-0.04% TFA; solvent gradient from 20% B to 35% B in the first 3 min, increased to 45% at 10 min, to 90% B at 17 min, followed by 3 min with 90% B, to 100% B at 22 min, followed by 4 min with 100% B; flow rate 1 ml/min and UV detection at 380 nm or 318 nm (Figure 2-24). For SR107PKS/NRPS, chromatographic conditions were as follows: solvents-A, water-0.1% formic acid (FA); solvent-B, acetonitrile-0.1% FA; solvent gradient from 20% B to 55% B in the first 5 min, increased to 70% at 8 min, followed by 8 min with 70% B, to 100% B at

17 min, followed by 3 min with 100% B, to 20% B at 21 min, followed by 4 min with 20% B; flow rate 1 ml/min and UV detection at 380 nm (Figure 2-26).

2-2-9. Isolation and structure determination of compounds 2 and 3

SR107HSAF culture (10 liters) was allowed to grow on the Fermentation Medium at 28 °C for 11 days. The whole solid cultures were diced and extracted three times with AcOEt/MeOH/AcOH (80 : 15 : 5, v/v/v) at room temperature, and the crude extract solution was concentrated under reduced pressure, and the concentrated extract was sequentially solvent partitioned into petroleum ether-soluble extract and MeOH-soluble extract. The MeOH extract was loaded to a Sephadex LH-20 column for separation. The column was eluted with MeOH to obtain six fractions, Fr.1 - 6. Fr.3 was subjected to medium-pressure liquid chromatography (MPLC; 80 g RP-18 silica gel; acetonitrile/H2O 30%, 50%, 70%, 80%, and 100%, 200 ml each) to afford five subfractions, Fr.3a – 3e. Fr.3c (60 mg) was subjected to preparative HPLC (Agilent 1200, ZORBAX Eclipse XDB-C18, 9.4×250 mm, 5 μ), using an isocratic solvent of 55% acetonitrile-0.04% TFA, at flow rate of 4 ml/min, detected at UV 360 nm. This afforded compound 2 (10 mg) and compound 3 (4 mg). To determine the structures of the compounds, 1D- and 2D-NMR were performed on a Bruker DRX-600 spectrometer. MS was carried out on an LTQ-Orbitrap-XL mass spectrometer (Thermo Scientific).

2-2-10. LC-MS-MS analysis of the metabolites in strain SR107PKS/NRPS

Strain SR107PKS/NRPS containing the construct pSETHSAF5 (PKS-NRPS alone under the control of *ermE** promoter) generated three new peaks as shown in Figure 2-24. To identify the metabolites, the peaks were individually collected from HPLC, and each of the peaks was further analyzed by UV-Vis spectrometry and

HPLC-HRMS using a Prominence Modular HPLC (Shimadzu Corporation, YMC-Pack Pro C18, 5 μm, 4.6 × 250 mm) coupled to a LTQ Velos Pro HRMS instrument (Thermo Scientific). For peak1, chromatographic condition was 60% acetonitrile/H_2O (containing 0.2% FA) for 10 min; for peaks 2 and 3, chromatographic condition was 85% acetonitrile/H_2O (containing 0.2% FA) for 10 min.

2-3 Result and discussion

The HSAF biosynthetic gene cluster contains only a single-module polyketide synthase-nonribosomal peptide synthetase (PKS-NRPS), although the PTM scaffold is apparently derived from two separate hexaketide chains and an ornithine residue. This suggests that the same PKS module would act not only iteratively, but also separately, in order to link the two hexaketide chains with the NRPS-activated ornithine to form the characteristic PTM scaffold. In the following sections, I describe the results on formation of tetramate polyenes in vitro using the individually purified PKS and NRPS, and on production of the Gram-negative *Lysobacter*-originated HSAF in Gram-positive *Streptomyces* heterologous hosts. The results provide direct evidence for this highly unusual biosynthetic mechanism that is likely general for the PTM- type hybrid polyketide-peptides.

2-3-1. HSAF polyketide synthase (PKS) expression in *E. coli*

To obtain direct evidence for this iterative single-module PKS, we expressed the PKS in *E. coli* and purified the 199.8 kDa protein (Figure 2-13). To test its activity, we converted the PKS to its holo form by incubating it with CoA and Svp, a promiscuous

4'-phosphopantetheinyl transferase (PPTase)[105]. Due to the huge size of this protein, we treated the PKS with trypsin after the reaction and followed the mass change of the tryptic fragment within the ACP domain of the PKS, to which the PPT moiety and biosynthetic intermediates are covalently linked. Specifically, the trypsin digestion is predicted to release a 26-residue fragment, VKPEQIDADASLNALGLDSLLAMELR (the active site serine residue underlined), within the ACP domain.

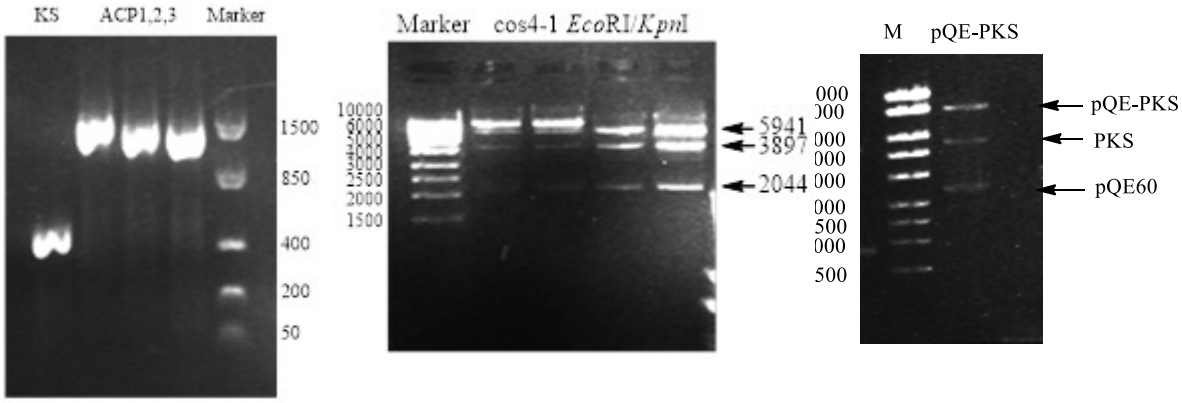


Figure 2-12. Construction of expression vector pQE60-PKS. A: PCR results of KS-fragment and ACP-fragment. B: Middle fragment digested from Cosmid4-1by *Eco*RI/*Kpn*I (after digestion of Cosmid4-1 by *Eco*RI, the 5941bp band was purified and digested by *Kpn*I into 2 fragments (3897bp and 2044bp)). C: Double digestion check of pQE60-PKS construction by *Nco*I/*Bam*HI.

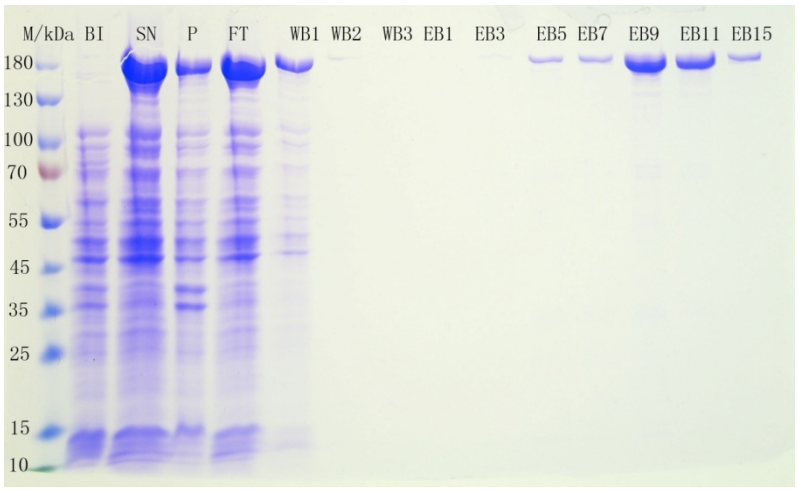


Figure 2-13. SDS-PAGE of the HSAF PKS expressed in *E. coli*. M, markers; BI, total proteins before IPTG induction; SN, total soluble proteins after sonication; P, total precipitated proteins after sonication; FT, flow-through fraction of the Ni-NTA column; WB1-3, wash-off fractions of the Ni-NTA column; EB1-15, imidazole-eluted fractions of the Ni-NTA column. The purified PKS is indicated with an arrow.

2-3-2. *In vitro* assay for the activity of the PKS module

The trypsin digestion of this *apo*-PKS is expected to release a 26-residue fragment, VKPEQIDADASLNALGLDSLLAMELR, which contains the active site serine residue (underlined) of the ACP domain of the PKS (Table 2-1, shown the mass of a number of selected ions, both the predicted and the observed). Q-TOF MS data showed that the tryptic fragment from the apo-PKS isolated from *E. coli* had m/z 928.1374 for $[M+3H]^{3+}$ (calculated 928.1604) (Figure 2-15). This fragment was subject to tandem MS-MS analysis (Figure 2-14). The result showed that this tryptic fragment had the expected amino acid sequence, confirming that the 199.8 kDa protein purified from *E. coli* was HSAF PKS.

After the correct tryptic PKS fragment was identified, we analyzed the holo-PKS and detected a tryptic fragment of m/z 1041.5487 for $[M+3H]^{3+}$ (calculated 1041.5221) and m/z 781.4270 for $[M+4H]^{4+}$ (calculated 781.3936) (Figure 2-16, Table 2-1). The holo-PKS was also confirmed using phosphopantetheine ejection assay[106, 107], which detected m/z= 357.2021 (calculated 357.0880) for the predicted product 5 (Figure 2-17). The data showed that the heterologously produced PKS was as expected and the enzyme was active.

VKPEQIDADASLNALGLDSLLAMELR

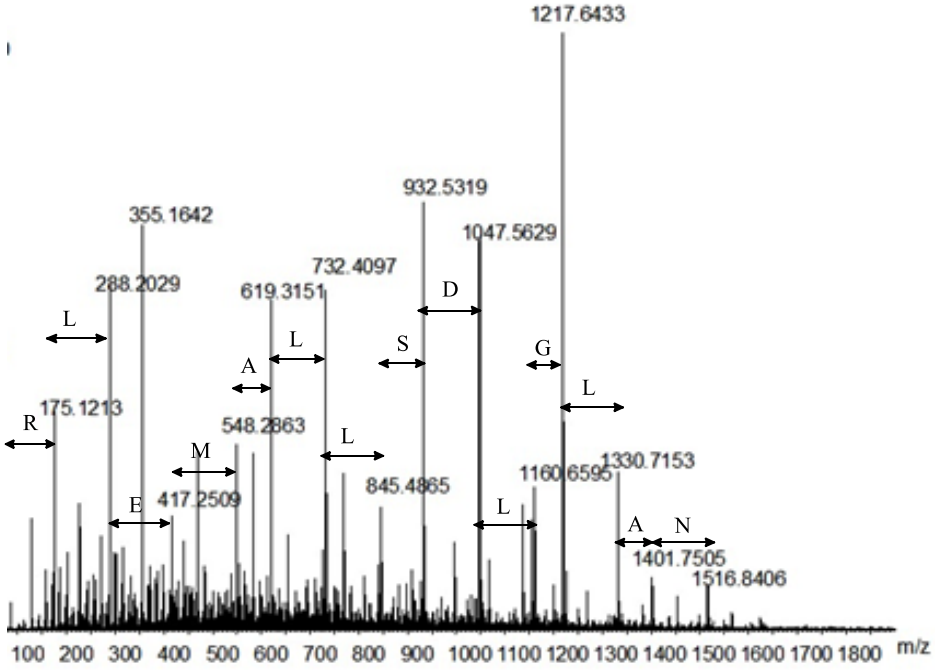


Figure 2-14. Tandem MS-MS data for the tryptic PKS fragment that harbors the serine active site of the ACP domain, VKPEQIDADASLNALGLDSLLAMELR.

VKPEQIDADASLNALGLDSLLAMELR

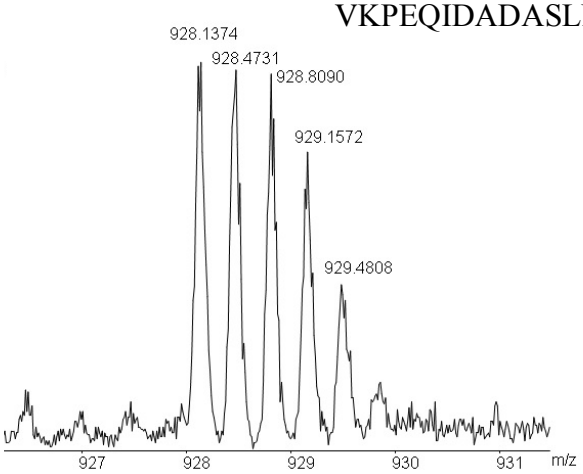


Figure 2-15. Q-TOF MS-MS data for the tryptic PKS fragment that harbors the serine active site. (VKPEQIDADASLNALGLDSLLAMELR).

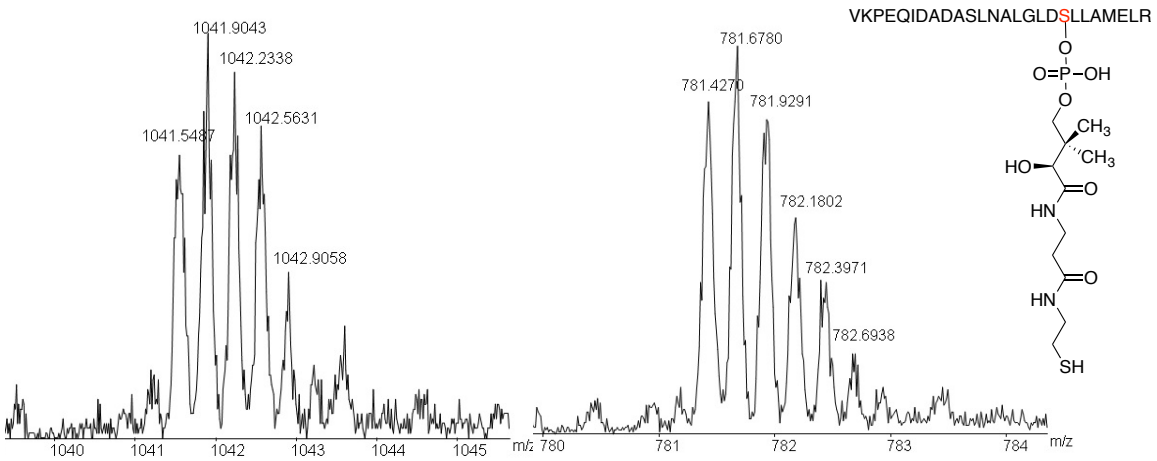


Figure 2-16. Q-TOF MS data for the tryptic fragment of PKS, after phosphopantetheinylation by PPTase, Svp, using CoA as substrate. The MS at left side showed $[M+3H]^{3+}$ *m/z* = 1041.5487 (calculated mass: 1041.52), and the MS at right side showed $[M+4H]^{4+}$ *m/z* = 781.4270 (calculated mass: 781.39).

Upon 4'-phosphopantetheinylation, the tryptic fragment from the holo-PKS is expected to have a 340.1 mass increase. Q-TOF MS data showed m/z 1041.5487 for $[M+3H]^{3+}$ (calculated 1041.5221) (Figure 2-16) and m/z 781.4270 for $[M+4H]^{4+}$ (calculated 781.3936). In addition, the 4'-phosphopantetheinylation of the PKS was examined using PPT ejection assay[106]. The expected compound was detected from the holo-PKS (Figure 2-17). The results demonstrated that the apo-form of the PKS expressed in *E. coli* had been converted into its holo-form *in vitro*.

Note: during vibrational activation of gas-phase carrier domains, a facile elimination occurs in bench-top mass spectrometers. The phosphopantetheine ejection pathways have been observed. The actual MS spectrum is shown below.

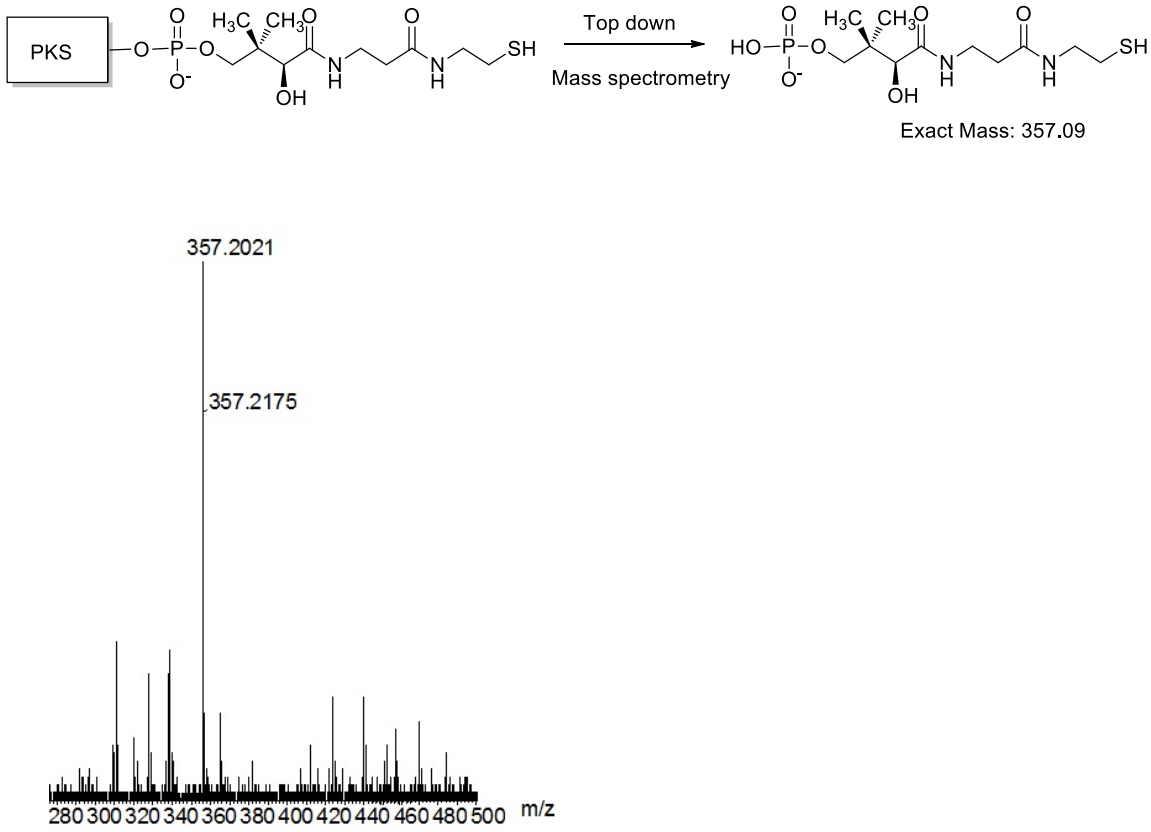


Figure 2-17. MS detection of the compound resulted from a PPT ejection, with *m*/*z* 357.2021 for $[M]^-$ (calculated 357.0891).

Next, we sought for potential biosynthetic intermediates that were covalently attached to the ACP domain of the PKS. Upon incubating the holo-PKS with acyl-CoA and NADPH, the tryptic fragment released from the PKS showed *m*/*z* 1113.4847 for $[M+3H]^{3+}$ (calculated 1113.5533) and *m*/*z* 835.4001 for $[M+4H]^{4+}$ (calculated 835.4170) (Figure 2-18). This mass change of the ACP fragment is coincident with a hexaketide polyene intermediate attached to the PPT of the PKS (Figure 2-18). We

also varied reaction conditions and searched for other potential biosynthetic intermediates with a varied carbon chain.

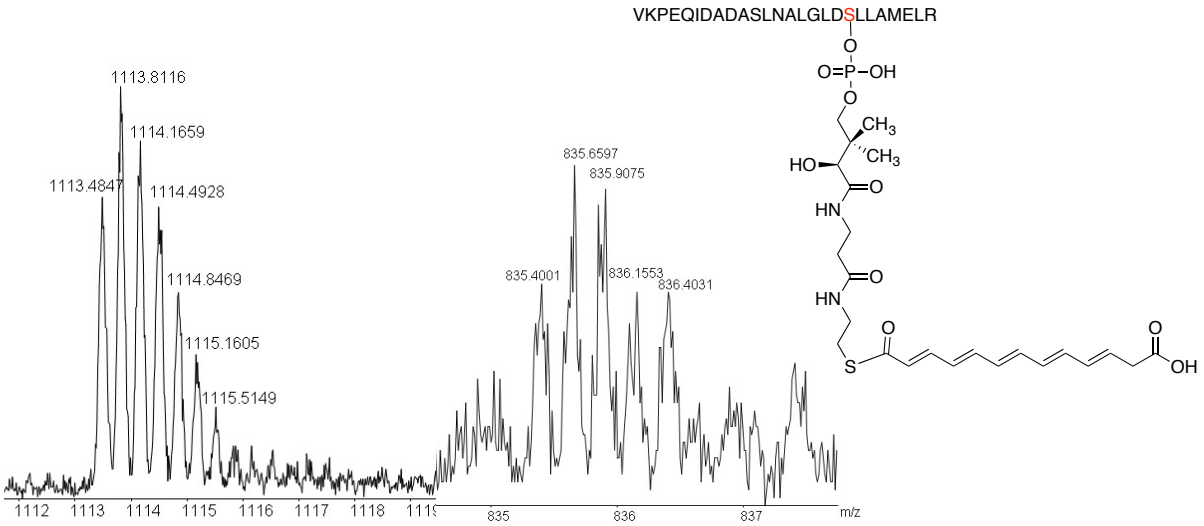


Figure 2-18. Polyene production in the in vitro reaction of PKS. The PKS protein was digested by trypsin, and the tryptic fragments were analyzed by Q-TOF-MS. The MS at left side showed $[M+3H]^{3+}$ m/z = 1113.4847 (calculated mass: 1113.55), and the MS at right side showed $[M+4H]^{4+}$ m/z = 835.4001 (calculated mass: 835.40).

Substrate	Predicted product	Predicted Mass (Da)		Observed Mass (Da)		Error (Da)
		M+3H	M+4H	M+3H	M+4H	M+1H
-	apo-PKS	928.1604	696.3722	928.1347	NA	0.0613
CoA	holo-PKS	1041.5221	781.3936	1041.5487	781.4270	0.0798
Mal-CoA	PKS-polyene	1113.5533	835.4170	1113.4847	835.4001	0.2058

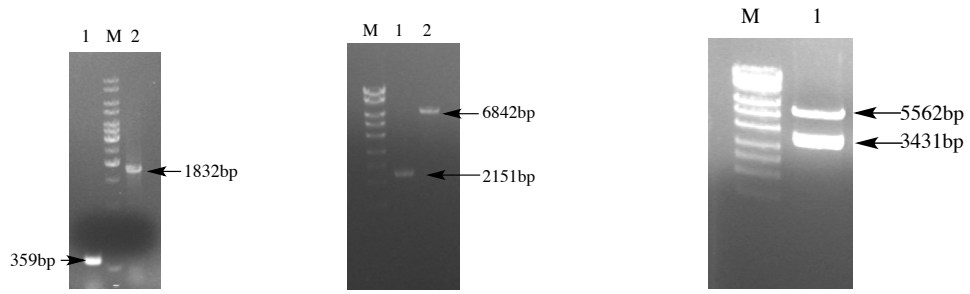
Table 2-1. The predicted mass and observed mass of the tryptic PKS fragment containing the ACP domain.

Hexaketide polyene was the only one detected. The results support that the single-module PKS found in the HSAF biosynthetic cluster is indeed an iterative PKS, preferably to catalyze five cycles of chain elongation in polyketide biosynthesis.

Interestingly, the data suggest that the PKS used malonyl-CoA as both the starter and the extender in the polyketide chain synthesis because a carboxylate appeared present in the intermediate (Figure 2-18).

2-3-3. Expression of PKS with a point-mutated KS domain in *E. coli* and activity assay of the synthase

To verify this hypothesis, we generated another PKS expression construct, in which the active site cysteine in the KS domain of this PKS was mutated to alanine. Since all the domains, except the KS domain, are still active, the synthase is expected to transfer the starter unit to the ACP but unable to elongate the polyketide chain. So the acyl group from the starter unit will be stuck on the ACP domain. Indeed we detected the mass for malonyl-*S*-ACP, instead of acetyl-*S*-ACP, and we also detected the mass of the mutated KS fragment and the active ACP domain. The result demonstrated that PKS prefers malonyl-CoA to acetyl-CoA as the starter.



(A) (B) (C)

Figure 2-19. Gel electrophoresis of A: PCR amplification of the KS upstream and downstream fragments (359 bp and 1832 bp). These two fragments were used as templates in overlapping PCR. The marker is the 1 kb ladder from Fermentas. B: Lane 1 is the overlapping PCR fragment (2151 bp), which harbors the C/A mutation

site. Lane 2 is the pANT841-PKS (original) vector digested by *Eco*RI and *Sac*I. Marker is the high range ladder No.393 from Fermentas. C: Lane 1 is *Nco*I and *Bam*HI double digested pQE60-mPKS, resulted in 2 bands that have the same lengths (5562 bp and 3431 bp, which are the lengths of mutated PKS and pQE60).

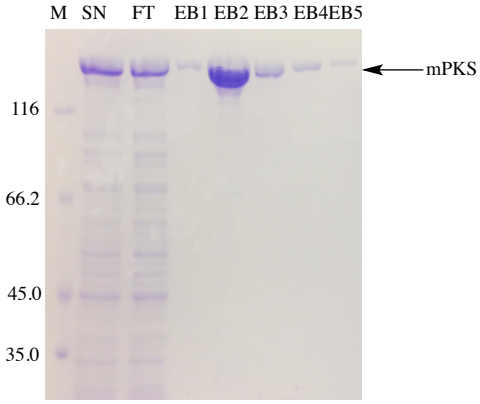


Figure 2-20. SDS-PAGE of the mutated PKS expressed in *E. coli*. M, markers; SN, total soluble proteins after sonication; FT, flow-through fraction of the Ni-NTA column; EB1-5, imidazole-eluted fractions of the Ni-NTA column. The purified PKS is indicated with an arrow.

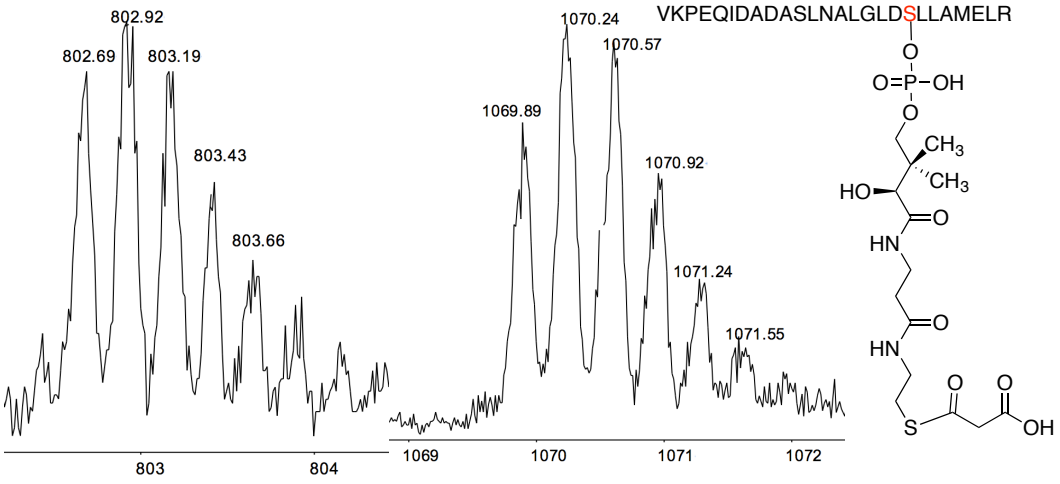
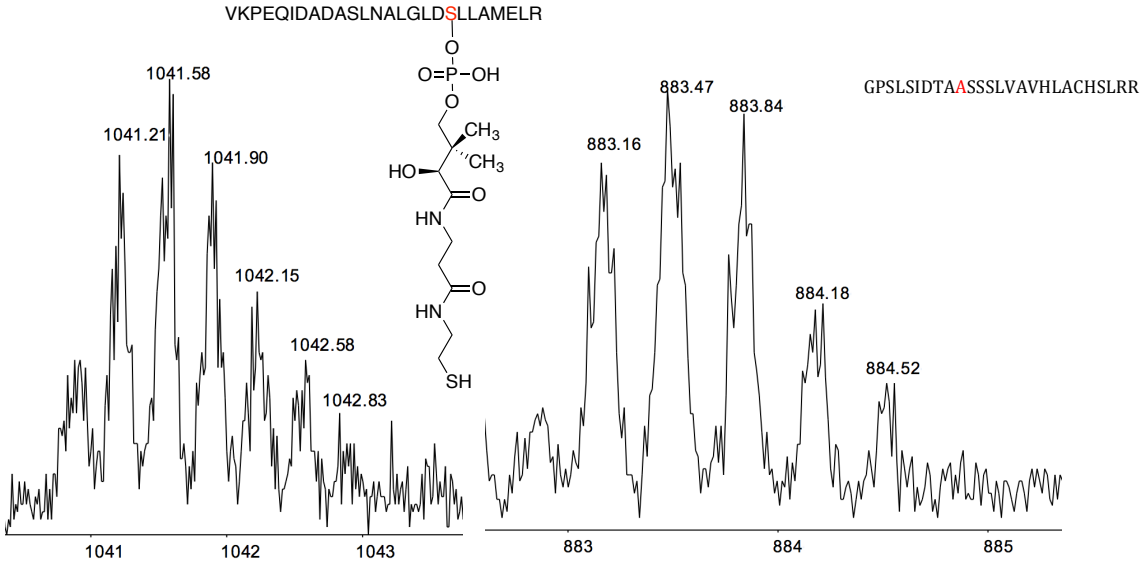


Figure 2-21. Q-TOF MS data for the tryptic fragment of holo-mutated PKS, after incubating with acyl-CoA, malonyl-CoA and NADPH. The MS at left side showed

$[M+3H]^{3+}$ *m/z* = 1069.89 (calculated mass: 1070.19), and the MS at right side showed$[M+4H]^{4+}$ *m/z* = 802.69 (calculated mass: 802.87).



(A) (B)

Figure 2-22. Q-TOF MS data for the tryptic fragment of PKS, after phosphopantetheinylation by the PPTase, Svp, using CoA as substrate. The MS showed $[M+3H]^{3+}$ *m/z* = 1041.21 (calculated mass: 1041.52). B: Q-TOF MS data for the tryptic fragment of PKS (incomplete digestion), which harbored the Cys to Ala mutation site. The MS showed $[M+3H]^{3+}$ *m/z* = 883.16 (calculated mass: 883.45)

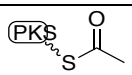
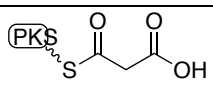
Starter unit	Predicted product	Predicted Mass (Da)		Observed Mass (Da)	
		M+3H	M+4H	M+3H	M+4H
Acetyl-CoA		1055.52	791.86	NA	NA
Malonyl-CoA		1070.19	802.87	1069.89	802.69

Table 2-2. The predicted mass and observed mass of the tryptic mutated PKS fragment containing the ACP domain.

Substrate	Predicted product	M+3H	
		Predicted	Observed
-	Tryptic fragment of holo-PKS	1041.52	1041.21
-	Tryptic fragment with C-A mutation	883.45	883.16

Table 2-3. The predicted mass and observed mass of the tryptic fragment of holo-mPKS and the tryptic fragment of PKS, which harbored Cys to Ala mutation.

2-3-4. *In vitro* reconstitution of the PKS module and the NRPS module

Finally, we incubated the polyene-*S*-PKS with L-ornithine-*S*-NRPS to reconstitute the activity of the hybrid PKS-NRPS *in vitro*. The NRPS module (C-A-PCP-TE, 148.6 kDa) was separately expressed and purified. Prior to reconstituting the activity, the purified NRPS was converted to the holo form by incubating with CoA and Svp, followed by incubating with L-ornithine and ATP to form L-ornithine-*S*-NRPS. Since a thioesterase (TE) domain is present in this module, we expect the product(s) be released into the reaction medium[76, 79, 94]. We used LC-MS to search for released product(s) of the reconstitution reaction (Figure 2-23). Two closely associated LC peaks, one at 21.4 min and one at 23.2 min, were detected when the MS scan was set at m/z 475.07-476.07. The peak at 21.4 min gave m/z 475.54, and the peak at 23.2

min gave m/z 475.60 (Figure 2-23). These peaks were not found in the control reaction. The mass of the species in these peaks is in agreement with the production of a tetramate-containing polyene-ornithine-polyene product (calculated 475.26 for $[M+H]^+$). Different from the hexaketide polyene produced in the PKS-alone reaction, the products from the reconstitution of both PKS and NRPS did not contain the carboxylate from the starter malonyl-CoA. Together, the data from the *in vitro* PKS-NRPS reconstitution and from the analysis of the PKS tryptic fragments support that the PKS module acts iteratively and, together the NPRS module, are sufficient for the synthesis of the framework of PTM.

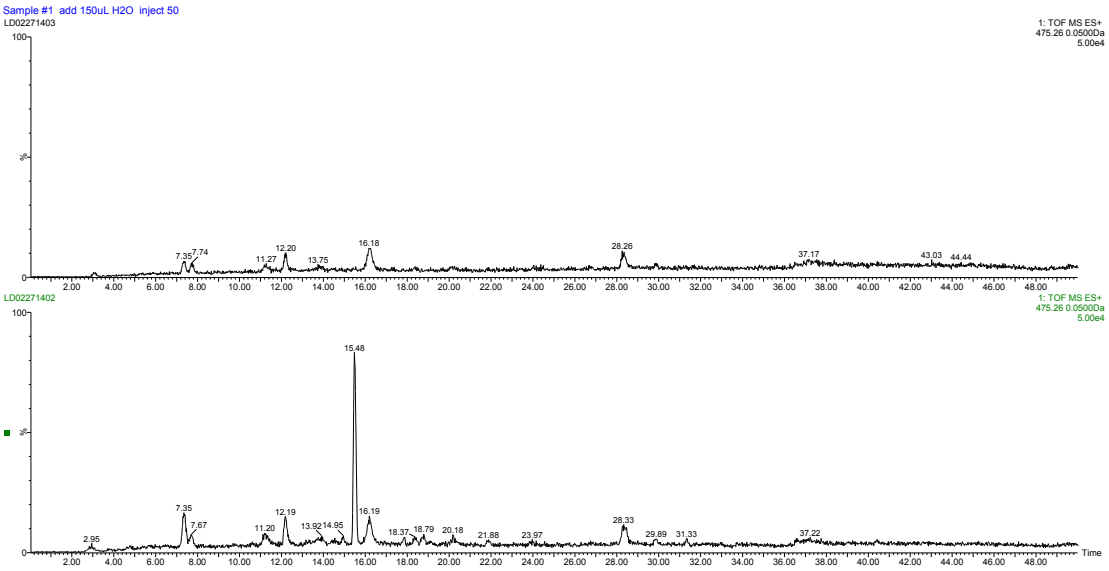
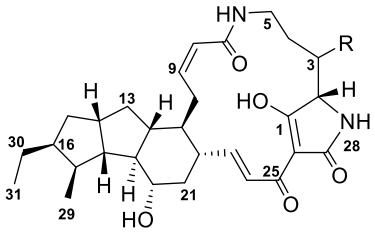


Figure 2-23. Analysis of the in vitro reconstituted PKS-NRPS. Total ion chromatography of the selected mass (475.2597) for compound **6** is shown for the control reaction (top panel) and the PKS-NRPS reaction (bottom). The peak at 15.48 min in the PKS-NRPS reaction gave *m*/*z* 475.2572 in HR-ESI-MS.

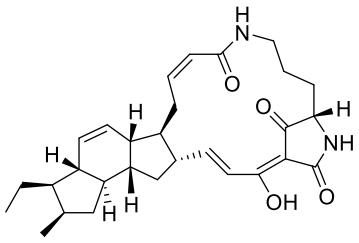
2-3-5. Construction of HSAF expression vectors and heterologous production in *Streptomyces*

(I acknowledge Yaoyao Li for construction of HSAF expression vectors and heterologous production in *Streptomyces*.)

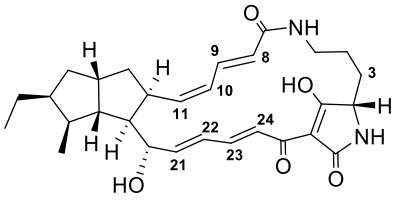
We isolated a cosmid clone, Cos4'-1, from the genomic library of *L. enzymogenes* C3, which contains the entire HSAF biosynthetic gene cluster[78]. The gene cluster was then transferred into vectors for expression in *Streptomyces sp*. However, the transformants failed to produce any detectable HSAF or analogs. Subsequently, we made two modifications in the experiments. First, we replaced the putative promoter at the 5'-nontranslated region of the PKS-NRPS gene with the *ermE** promoter, generating pSET152HSAF3 (Figure 2-11). Second, we chose to use an engineered host, strain SR107 derived from *Streptomyces sp*. LZ35 through deleting its four native PKS gene clusters. This host is expected to provide a relatively "clean" background for the heterologous production of HSAF. We introduced pSET152HSAF3 into strain SR107 and analyzed the metabolites in the transformant using HPLC. The transformant strain produced approximately seven eminent peaks that were absent in the control strain SR107 (Figure 2-24). We first focused our attention on the main peak 2 at 18.8 min because it falls in the region that HSAF and analogs would appear (16-20 min).



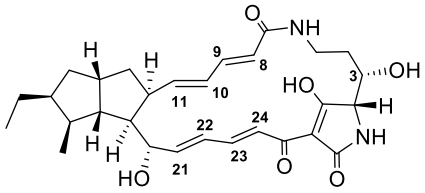
HSAF (**1**), R = OH
3-deOH-HSAF (**3**), R = H



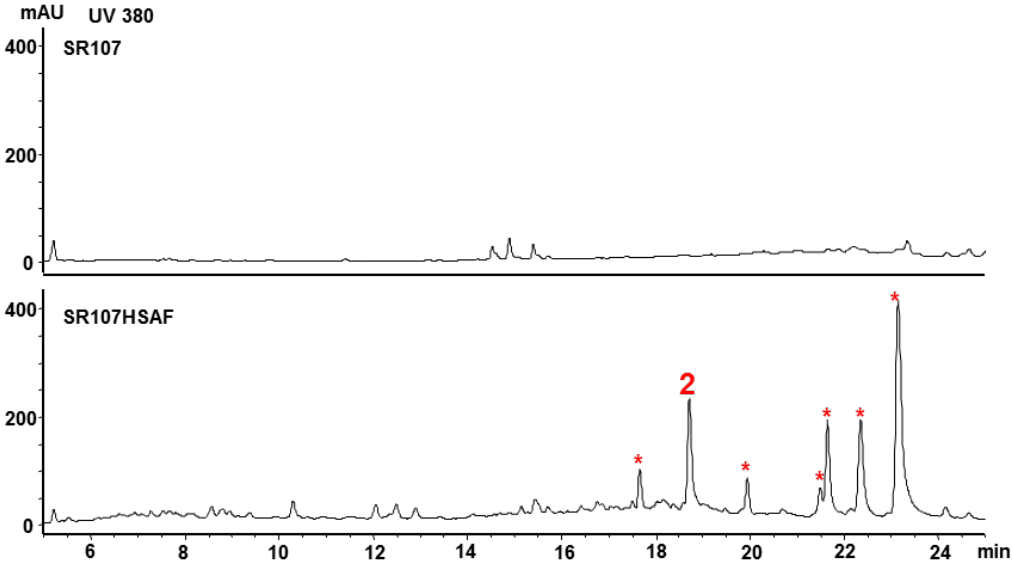
Ikarugamycin (**4**)



2



Alteramide A (**5**)



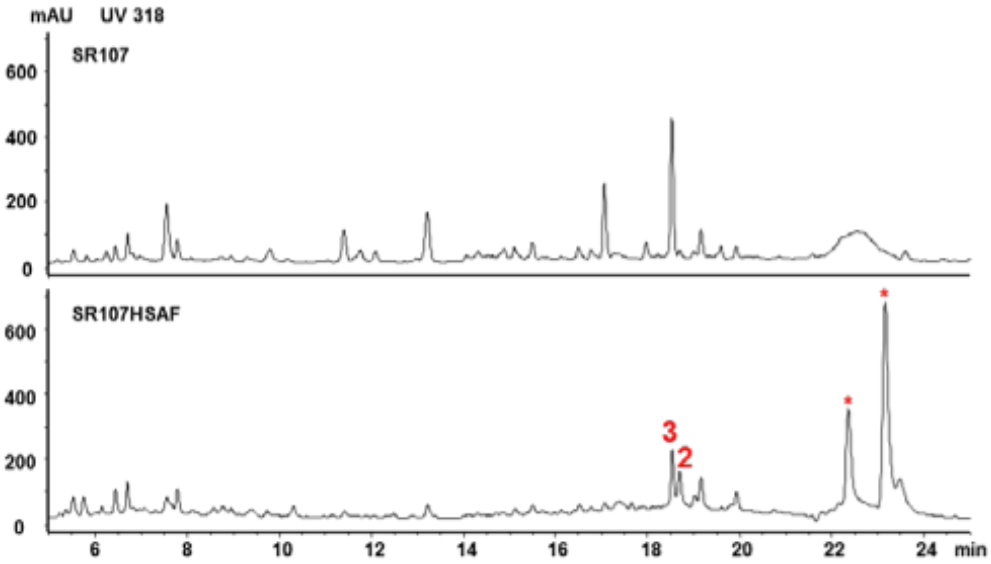


Figure 2-24. Chemical structures of HSAF and analogs and production of HSAF analogs in *Streptomyces*. SR107, HPLC analysis of metabolites from *Streptomyces sp*. LZ35 strain SR107, in which four native PKS gene clusters had been deleted; SR107HSAF, HPLC analysis of metabolites from strain SR107 containing the HSAF biosynthetic gene cluster under the control of *ermE** promoter. The first two HPLC profiles show the metabolites detected at 380 nm, and the bottom two show the metabolites detected at 318 nm. The asterisks indicate peaks that were absent in the control.

Compound **2** was isolated as yellow powder. HR-ESI-MS gave a quasi molecular ion at *m*/*z* 495.2837 for $[M + H]^+$ (calculated 495.2853 for $C_{29}H_{38}N_2O_5$). The complete NMR assignments were carried out by the analysis of 1D and 2D NMR data (HSQC, HMBC and 1H-1H COSY). The NMR comparison of **2** with alteramide A (**4**) (Figure 2-3) indicated that the compounds are structurally similar[80], except for the absence of the hydroxyl group at C3 and the Z-geometry for C10,C11 double bond in **2** (Figure 2-24). The relative configuration of **2** was established by proton couplings and NOE correlations. The large coupling constants (~ 15.0 Hz) between olefinic protons (H8/H9, H21/H22, and H23/H24) led to assignment of the *E*-configuration for the three double bonds, whereas the *Z*-configuration of C10,C11 double bond was deduced from the small coupling constant (11.2 Hz) between H10 and H11. The relative stereochemistry of the bicycle unit was determined from the NOESY experiment, which is identical to alteramide A (**4**).

The rest of the eminent peaks (indicated by asterisks in Figure 2-24) detected at 380 nm were fairly labile. The compounds eventually disappeared over the process of purification. To see if any other isolable HSAF analog was produced in the

transformant, we checked the metabolites under other wavelengths. At 318 nm, two peaks were detected at the HSAF region, one at 18.8 min (compound **2**) and the other at 18.6 min (compound **3**) (Figure 2-24). These two compounds were not produced by the control strain SR107 (at this wavelength, a main peak at 18.6 min was also detected in the control, but showed a different UV-Vis spectrum than compound **3**). Compound **3** was then isolated for structural determination. It appeared as white powder, with a quasi molecular ion at m/z 497.3021 for $[M + H]^+$ (calculated 497.3015 for $C_{29}H_{40}N_2O_5$) as determined by HR-ESI-MS. Comparison of the ^{1}H-NMR spectrum of **3** to that of 3-deOH-HSAF readily established the structure of **3** as 3-deOH-HSAF.

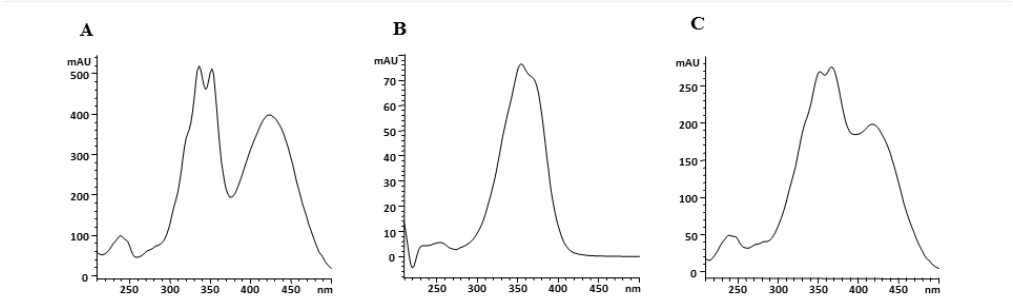
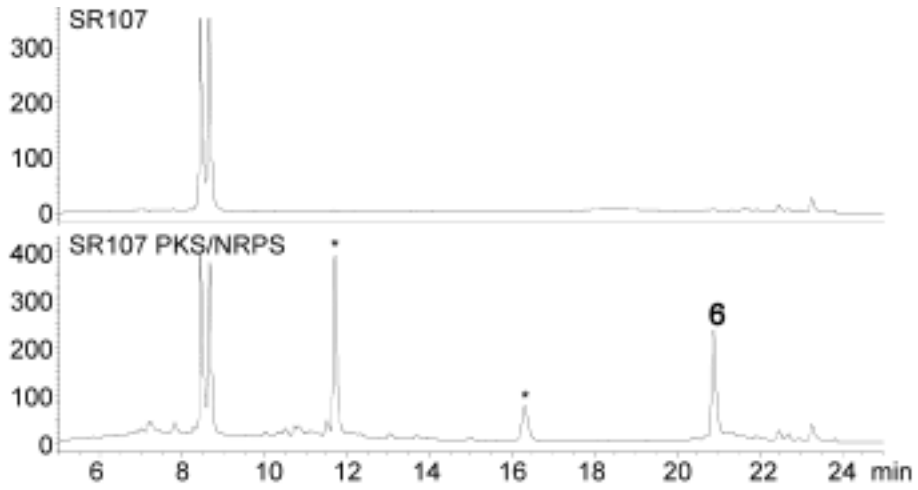


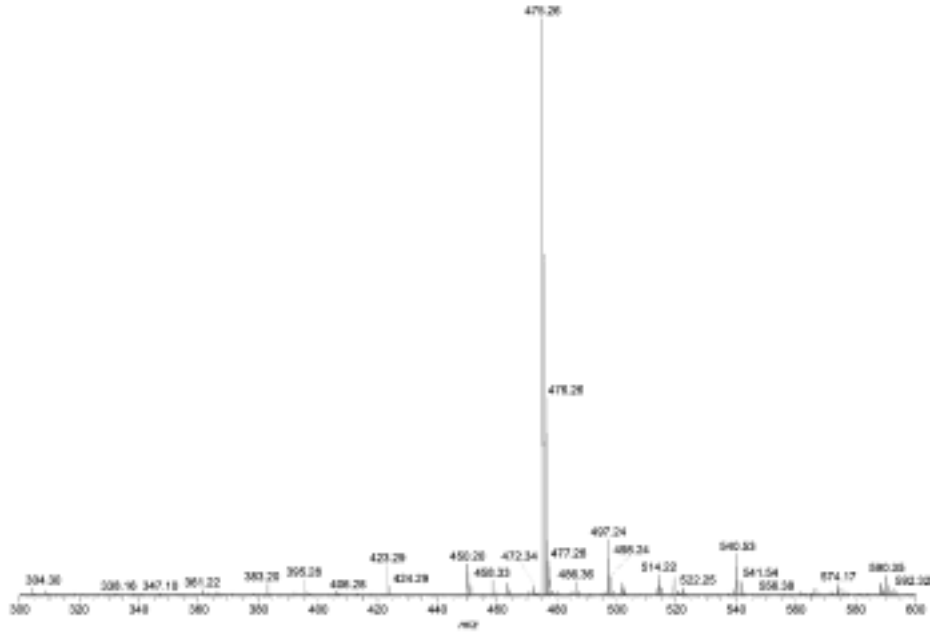
Figure 2-25. UV-Vis spectra of the three peaks observed in HPLC of the metabolites from *Streptomyces* strain SR107PKS/NRPS that was transformed with PKS-NRPS alone. A, peak at 11.7 min; B, peak at 16.3 min; C, peak at 20.8 min (polyene tetramate intermediate).

The production of compounds **2** and **3** in a *Streptomyces* strain supports the notion that the single-module hybrid PKS-NRPS in pSETHSAF3 is probably sufficient for the assembly of the PTM scaffold. To further prove this point, we transferred pSETHSAF3 into a second *Streptomyces* host that has a completely "clean" background. Strain ZM12 was derived from *S. coelicolor*, probably the best-studied

model *Streptomyces*, through deletion of all ten native PKS and NRPS gene clusters present in its genome[96]. HPLC showed that this strain produced very few metabolites. However, upon introduction of pSETHSAF3, both compounds **2** and **3** were produced as the main metabolites in strain ZM12, in addition to a number of minor peaks. No other PKS-NRPS is present in strain ZM12, so our data clearly demonstrated that the biosynthesis of HSAF only requires a single-module hybrid PKS-NRPS. The results also imply that the five domains (ketosynthase (KS)/acyltransferase (AT)/dehydratase (DH)/ketoreductase (KR)/acyl carrier protein (ACP)) of this PKS module act two separate times to assemble two separate hexaketide chains.

Compounds **2** and **3** lack the 3-hydroxy group of HSAF (**1**), which suggests that the sterol desaturase (SD) gene (see Figure 2-4 for the cluster) was not functional in the expression construct pSETHSAF3, in which the *ermE** promoter was placed in front of the PKS-NRPS gene (Figure 2-11). The SD gene encodes the 3-hydroxylase converting 3-deOH-HSAF (**3**) into HSAF[93]. We subsequently generated a second expression construct, pSETHSAF4, in which the *ermE** promoter was placed in front of the SD gene (Figure 2-11). The construct was introduced into *Streptomyces sp.* SR107 to generate the transformant strain SR107HSAF2. However, a careful search of the metabolites in this strain did not find any HSAF-like compound (Figure 2-26). The reason for this is unclear at this moment; one possibility is that the insertion of the *ermE** cassette in front of the SD gene might not lead to transcription of the PKS-NRPS gene and downstream tailoring genes because the SD gene and the other genes in the cluster do not appear to share the same promoter (Figure 2-11).


SR107
300
200
100
0
SR107 PKS/NRPS
400
300
200
100
0
*
*
6
6
8
10
12
14
16
18
20
22
24 min


476.26
476.28
407.24
486.24
514.22
540.33
541.04
558.38
574.17
580.39
592.32
423.29
434.29
450.28
456.33
472.34
477.28
466.36
394.36
386.16
347.18
361.22
383.28
395.28
486.26
320
340
360
380
400
420
440
460
480
500
520
540
560
580
600
m/z

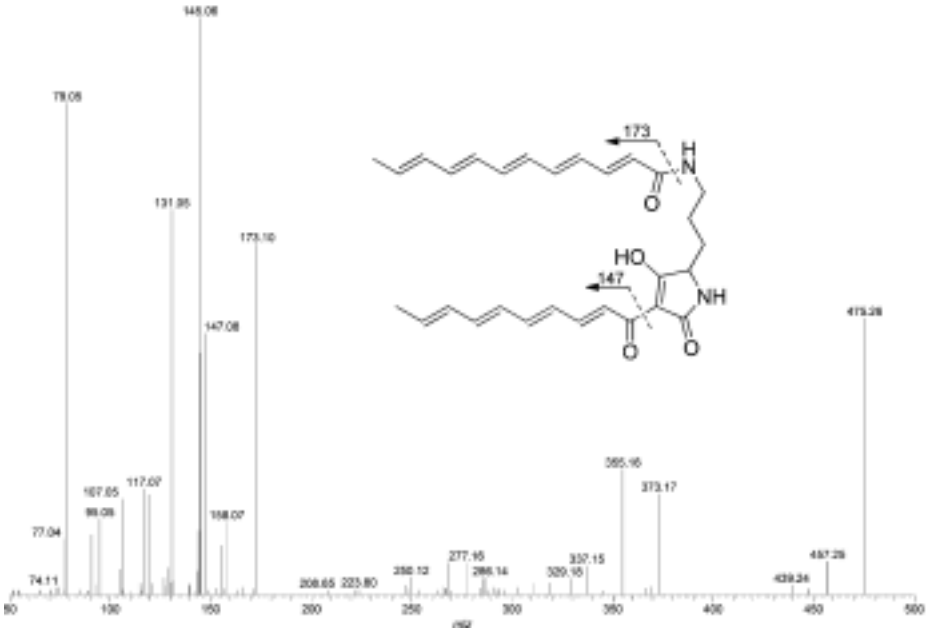

148.08
79.05
131.05
173.10
147.08
107.05
117.07
99.08
168.07
77.04
74.11
208.65
223.80
258.13
277.15
286.14
329.15
337.15
385.16
373.17
439.24
457.28
479.28
173
147
HO
NH
O
O
O
O
H
N
50
100
150
200
250
300
350
400
450
500
m/z

Figure 2-26. Production of the polyene tetramate **6** in *Streptomyces* transformed with the PKS-NRPS only. HPLC analysis (top); full ESI-MS spectrum of **6** (bottom left) and MS/MS analysis of **6** (bottom right). SR107: metabolites from nontransformed *Streptomyces sp*. LZ35 strain SR107; SR107PKS/NRPS: metabolites from the strain transformed with only the PKS-NRPS gene under control of the *ermE** promoter. The metabolites were detected at 380 nm. The asterisks indicate peaks that were absent in the control.

Based on the data, we proposed a mechanism for the formation of PTM scaffold (Figure 1-6). First, the five-domain PKS module iteratively catalyzes the formation of a polyene hexaketide. This hexaketide is transferred to the four-domain NRPS module, which activates L-orinithine and catalyzes the first amide bond formation, probably between the delta-amino group of L-ornithine and thioester carbonyl group of the hexaketide. This leads to a NRPS-bound polyene-ornithine intermediate. Meantime, the PKS module continues to assemble the second hexaketide, which is subsequently transferred to the NRPS through forming the second amide bond, probably between the alpha-amino group of L-ornithine and the second hexaketide chain. The timing of this transfer is right after the fifth cycle of polyketide chain elongation, in which the newly formed β-keto group of the second hexaketide chain has not been processed by the KR domain and DH domain. All PTM natural products contain this keto group in the final structure (C-25 in HSAF structure, Figure 2-24).[27, 76, 108] The determining factor for this timing is not known, but probably related to the redox enzymes that are proposed to cooperate with the PKS-NRPS in morphing the PTM scaffold[81]. The transfer of the second hexaketide chain leads to a NRPS-bound

polyene-ornithine-polyene intermediate. Finally, the intermediate is released from the NRPS by formation of the tetramate unite, which is through the attack of the nucleophilic alpha carbon of the second hexaketide at the thioester carbonyl carbon of L-ornithine. The presence of the β-keto group on the second hexaketide makes the α-carbon a good nucleophile, which promotes the tetramate formation and product release. This perhaps contributes to the timing of the transfer of this hexaketide to NRPS. In case the β-keto was fully processed, the second polyene chain would not undergo tetramate formation, which results in an "unproductive process". *In vivo*, the tetramate-containing polyketide-ornithine- polyketide is ultimately morphed into the PTM scaffold by tailoring enzymes, probably when the intermediate is still NRPS-bound. Our previous study showed that this NRPS prefers a C12 chain when forming tetramate acyl-ornithine-acyl products *in vitro*. Thus, the specificity of the domain (most likely the C domain) responsible for transferring the polyketide chain from the PKS to the NRPS is the key determinant for the macrolactam ring size observed in the characteristic PTM framework.

2-4. Summary and final remarks

HSAF is the main antifungal factor that the biocontrol agent *Lysobacter enzymogenes* uses to fight against fungal diseases and thus to protect crop plants. *Lysobacter* are Gram-negative gliding bacteria; they are ubiquitous in the environment, but largely remain untapped for bioactive natural products[109]. HSAF has a potent activity against a broad spectrum of fungi, using a new mode of action. Its chemical structure is distinct from any existing fungicide or antifungal drug. Most interestingly, its biosynthesis involves an unprecedented mechanism – using a single module

PKS-NRPS to assemble two separate hexaketide chains and link them to ornithine residue, eventually resulting in a polycyclic tetramate macrolactam. Furthermore, PTM-type gene clusters have been identified from numerous bacterial genomes[27, 108], and yet none of them has been biochemically investigated.

In over 20 years, bacterial modular PKSs have been found to synthesize polyketides using a non-iterative mechanism — one module catalyzes one cycle of elongation of the same polyketide chain. A single-module PKS to catalyze multi-cycle elongations of two separate polyketide chains has only been found in the PTM-type PKS. Our studies of HSAF provide direct evidence for a previously uncharacterized iterative biosynthetic mechanism for bacterial polyketide-peptide natural products. The evidence comes from both the *in vivo* approach - heterologous production of HSAF in the Gram-positive bacterial hosts, *Streptomyces* strains, and the *in vitro* approach - expression of the ~200 kDa PKS and the ~150 kDa NRPS separately in *E. coli*, purification of the giant enzymes, and reconstitution of the biosynthetic activity using these two enzymes. In light of the huge number of uninvestigated PTM-type gene clusters in databases, our studies present here will facilitate the research of a new area of hybrid polyketide-peptide biosynthesis.

Chapter 3: Defining the Minimal Gene Cluster Required for PTM Biosynthesis

Portions of this chapter were reported (Journal of American Chemical Society. 2011, 144:643-645) and are reproduced with the permission of the publisher[76].

3-1. Background and Significance

Among the natural products, polycyclic tetramate macrolactam (PTM) are emerging as a new class of natural products with a distinct structure and a new mode of action. HSAF was discovered from *Lysobacter enzymogenes* strain C3 that has powerful antifungal activity[78, 110]. HSAF contains a 5,5,6-tricyclic system fused with a 17-membered macrolactam that includes a tetramate acid moiety. Tetramic acid is the key structural feature for many bioactive heterocycles that exhibit a wide range of biological activities including antibiotic and anticancer activities.

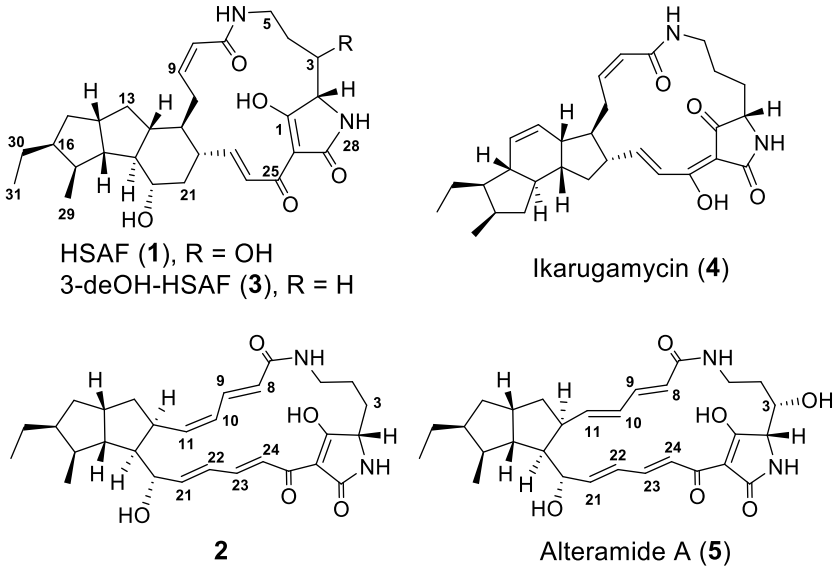
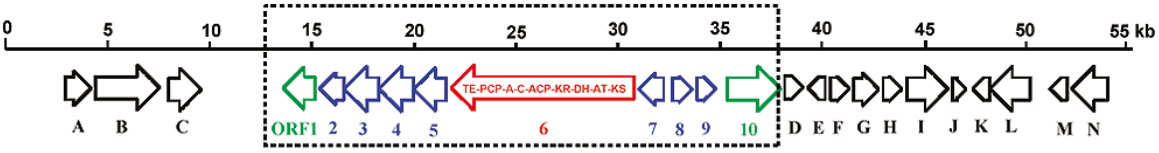


Figure 3-1. Chemical structure of selected PTMs.

Previous study of the mode of action indicated that HSAF specifically interferes with the biosynthesis of sphingolipids. Sphingolipids represent an attractive new target for the development of novel antifungal drugs because their structure in fungal cells is distinct from that in mammalian cells. This novel mode of action together with the unusual chemical structure that is distinct from any existing antifungal drug on the market, make HSAF a potentially promising drug lead[78].

Through our previous study, the genetic locus responsible for HSAF biosynthesis in *L.e.* C3 was identified[76]. Since the genes for secondary metabolite biosynthesis often cluster together, the complete biosynthetic gene cluster of HSAF may very likely be sitting within the region between ORF1 and ORF10, in which there is only one PKS-NRPS hybrid (Fig. 3-2). Similar gene clusters were also found by Clardy and co-workers from the genomes of a number of *Streptomyces* strains, all of which contain one single modular PKS/NRPS[27].

(A)



(B)

ORF	# nucleotides	# aa	Blastp Homolog	% identity/Similarity	E-value
ORF1	13746-15236	496	Major facilitator superfamily (MSF) transporter	36/53	2e-63
ORF2	15354-16409	351	OX4 (alcohol dehydrogenase zinc-binding protein)	70/81	2e-143
ORF3	16406-18133	575	OX3 (FAD-dependent oxidoreductase)	78/89	0.0
ORF4	18147-19883	578	OX2 (FAD-dependent oxidoreductase)	69/82	0.0
ORF5	19964-21619	547	OX1 (FAD-dependent oxidoreductase)	68/79	0.0
ORF6	21720-31091	3132	PKS/NRPS (KS-AT-DH-KR-ACP-C-A-PCP-TE)	63/74	0.0
ORF7	31409-32470	353	FA hydroxylase/sterol desaturase	59/71	3e-87
ORF8	33000-33773	257	Ferredoxin NADP reductase/FAD/NADP-binding protein	68/83	1e-104
ORF9	33993-34949	318	Arginase	73/84	6e-127
ORF10	35902-38424	840	TonB-dependent outer membrane receptor	64/76	0.0

Figure 3-2. (A) Map of HSAF biosynthetic gene cluster. ORFs within the rectangle are the putative minimal gene cluster for HSAF biosynthesis. (B) Annotation of the genes within the HSAF biosynthetic gene cluster.

In the previous work, 3-deOH HSAF was isolated from the disruption mutant of Sterol Desaturase (SD) gene within HSAF gene cluster and was observed to be transformed into HSAF in the presence of Sterol Desaturase[93]. This showed that the SD gene is also required in modification after the assembly of the HSAF core structure. The disruption of the ferredoxin reductase gene led to a lower level of dihydromaltophilin and proportional increase of a putative precursor/shunt product, suggesting that this gene is related to HSAF production but its role may be compensated by other similar genes in the genome. Finally, the disruption of the arginase gene (ORF9) had no obvious effect. Arginase catalyzes the conversion of arginine to urea and ornithine and is one of the enzymes in the Urea Cycle. This arginase may provide ornithine needed in dihydromaltophilin biosynthesis, but it is not essential as other arginases in primary metabolism could compensate its function. In the meantime, we constructed a series of heterologous expression constructs, as illustrated in Chapter 2, and introduced the constructs into *S. lividans* and analyzing the metabolites in the transformants[26]. The successful isolation of the polyene-tetramate-polyene intermediate and 3-deOH HSAF indicated that a single PKS-NRPS is sufficient for synthesizing the basic structure of HSAF and that the introduction of these core genes into a proper heterologous host will lead to basal production of HSAF (or its precursors/analogs).

As shown in Figure 3-2, this PKS-NRPS is flanked by six redox genes, four NADP/FAD-dependent oxidoreductase (OX) genes on the left and a hydroxylase/sterol desaturase gene and a NADP/FAD-ferredoxin reductase gene on the right. The genes for PKS-NRPS, OX1, 2, 3, and 4 appear to form an operon,

because they are either translationally coupled or only have a small intergenic region in between. Based on sequence alignment, it is not that all of the 4 redox genes are required for biosynthesis of PTMs with different ring systems. For example, in ikarugamycin biosynthetic gene cluster, there are only 2 redox genes next to the PKS-NRPS hybrid gene, which means that two redox enzymes are enough to make the tri-cyclic ring system. If this is applicable for all PTM compounds, it would suggest that there is one or two redox genes in HSAF biosynthetic gene cluster is/are functionally redundant. A characterization of the exact function of the OXs is required.

The objective of this project was to determine the minimal gene cluster required for assembling the hexaketide-ornithine-hexaketide skeleton, forming the tetramic acid unit, and synthesizing the 5,5,6-tricylic system of HSAF. My contribution in this work is the study of two of the oxidoreductase genes, using mutagenesis approach.

3-2. Materials and Methods

All oligonucleotides for PCR reactions were purchased from Integrated DNA Technologies (Coralville, IA). Molecular cloning reagents, including DNA polymerases, restriction enzymes and T4 DNA ligase, were purchased from New England BioLabs (Ipswich, MA). Chemicals for preparing solid and liquid medium were purchased from Fisher Scientific (Pittsburgh, PA) or Sigma-Aldrich (St. Louis, MO). Chemicals for metabolite extraction and HPLC solvents were purchased from Sigma-Aldrich. Plasmid was extracted using Qiagen kits (Valencia, CA). Other DNA manipulations were performed according to the procedure described in molecular cloning manuals. Plasmid DNA was propagated in *Escherichia coli* strain XLBLUE1. *Escherichia coli* strain S17-1 served as the host for the conjugal transfer of

replacement cassettes cloned in conjugation vector pJQ200SK. All PCR targeting gene transfers were performed using *Escherichia coli* strain BW25113.

3-2-1. Generation of four OX gene disruption mutants

(note: I acknowledge Yaoyao Li for preparing the four oxidoreductase gene disruption mutants.)

To construct the gene disruption vectors, an internal fragment was amplified from each of the four OX genes using the primer pairs listed below.

OX1For 5'-GCC GCT GGA GAA CTG GAA GG-3'

OX1Rev 5'-GCTGAA GTA GGT GAC GTG GAT GA-3'

OX2For 5'-GCT GTA CCC GTT CCT GAC GC-3'

OX2Rev 5'-TTC TTC CTC GCT GAG CAG ATA AGT-3'

OX3For 5'-TCC CGG TGC TGC CGT TCT A-3'

OX3Rev 5'-TGC GGT TCG GCG ATG GT-3'

OX4For 5'-AAG CTC GAC CTG CGC CAC T-3'

OX4Rev 5'-TTC TGC TCG GTG AAC CAC TCC-3'

The sizes of the PCR fragments for OX1, 2, 3, and 4 were 712, 622, 328 and 539 bp, respectively. The fragments were cloned into pGEM-T vector and their identity was confirmed by DNA sequencing. Each of the fragments was then transferred to the conjugal vector pJQ200SK as an *Apa*I/*Spe*I fragment to produce pJQ200SK-OX1-4. Each of the pJQ200SK-OX was transformed into *E. coli* S17-1, which was mated

with *Lysobacter enzymogenes* C3 for conjugal transfer of the constructs. The positive colonies grown on LB plates containing gentamicin (20μg/ml) and kanamycin (25μg/ml) were picked up and inoculated into liquid cultures containing these two antibiotics with same concentration.

Genomic DNA was prepared from each of the cultures, and diagnostic PCR reactions were performed to identify mutants that resulted from a homologous recombination at the respective OX genes. The following primer pairs were used for these experiments.

MOX1For 5'-GGG CAA CAA CGC CAA CC-3'

MOX1Rev 5'-CCC TCA TCA GTG CCA ACA TAG-3'

MOX2For 5'-CTG CAT CGA ATG GCT GAT C-3'

MOX2Rev 5'-TCT TTA TAG TCC TGT CGG GTT T-3'

MOX3For 5'-CCC CTA CTT CAA CGC CAT CC-3'

MOX3Rev 5'-CAC CCT CAT CAG TGC CAA C-3'

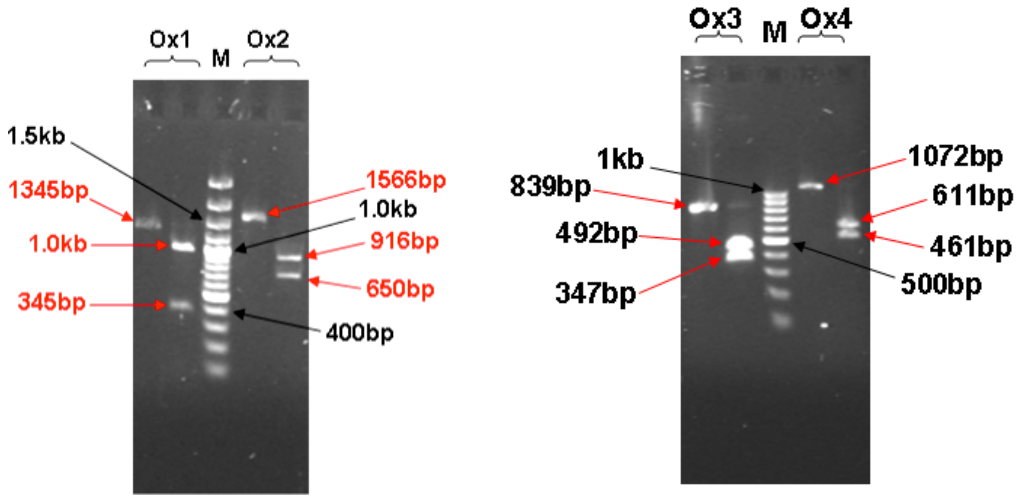
MOX4For 5'-ACC TTC CCG ATG GTG TTCC C-3'

MOX4Rev 5'-ACG ACC GAG CGT AGC GAG T-3'

In each of the primer pairs, the forward primer will anneal to the corresponding OX gene at a location that is outside the PCR fragment used in the gene disruption constructs, and the reverse primer will anneal to a location within pJQ200SK vector. The expected size of the PCR product from OX1 mutants was 1345 bp, which will release two fragments of 1000 and 345 bp upon *Apa*I digestion. The expected size of the PCR product from OX2 mutants was 1566 bp, which will release two fragments

of 916 and 650 bp upon ApaI digestion. The expected size of the PCR product from OX3 mutants was 839 bp, which will release two fragments of 492 and 347 bp upon *Apa*I digestion. The expected size of the PCR product from OX4 mutants was 1072 bp, which will release two fragments of 611 and 461 bp upon *Apa*I digestion. As shown in Figure 3-3, the diagnostic PCR and *Apa*I digestion produced the expected results and all of mutants were real OX gene disrupted mutants. (Fig. 3-3A)

(A)



(B)

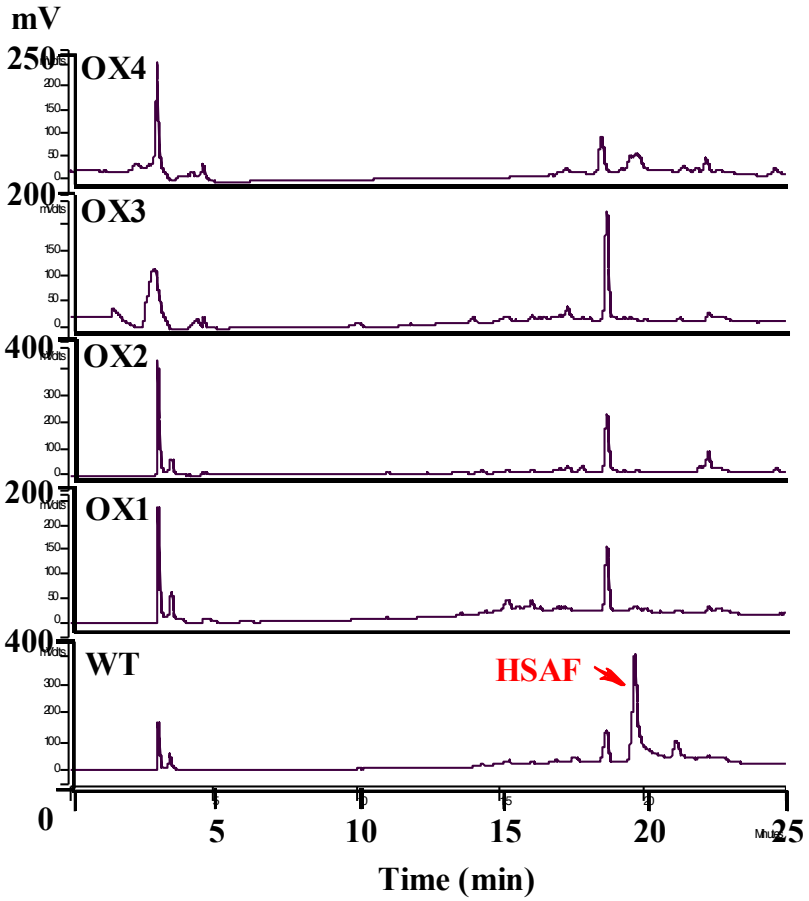


Figure 3-3. (A) Diagnostic PCR and restriction enzyme digestion of the PCR products to verify the OX gene disruption mutants. (B) HPLC analysis of the metabolites produced by the OX gene disrupted mutants. The HSAF peak produced by the wild type strain is indicated with an arrow.

To analyze the metabolites, the mutants were grown in 1/10 TSB medium containing gentamicin (20 μg/ml). The procedure for metabolite extraction and analysis was the same as described previously[78]. The results from HPLC showed that none of the OX mutants produced HSAF, while the wild type C3 clearly produced HSAF (Fig. 3-3B).

3-2-2. Generation of OX1, 2 gene deletion mutants

To construct the OX1 in-frame deletion vector, a 543-bp fragment of OX1 gene with an *Eco*RI site introduced to replace a 27-bp sequence within the gene. This would replace 9 amino acid residues (LVEDGRAVG) within a sequence motif that is highly conserved among the FAD-dependent oxidoreductases. The following primers were used for these experiments.

P1-OX1-Forw: 5'-AGA CTC GAG AAG CTA GCC ACG CTG C-3'

P2-OX1-Revs-fusion: 5'-CGT CTC GAA GAC GAA TTC CAG CTC TAC CAC CCG GTC GAC GCG GGT-3'

P1'-OX1-forw-fusion: 5'-GTA GAG CTG GAA TTC GTC TTC GAG ACG GTA CGG CTC AAG AAC GGC-3'

P2'-OX1- Rev: 5'- ACT GGA TCC CGA CTC GCC GGC CGG CA-3'

The following primers that flanked the left arm for homologous recombination were used to screen for OX1 single crossover mutants. A 1083-bp fragment was produced from the positive colonies.

P1-OX1 5'-TAA GGA TCC GCC GGC ATG GCC ACC G -3'
P2-OX1 5'-ACC ATG ATT ACG CCA AGC-3'
P3-OX1 5'-TCA CTC GAG TCA GGC GGC CGG AAG C-3'

Genomic DNA of *Lysobacter enzymogenes* C3 was used as template to amplify the complete OX1 in-frame deletion cassette. The upstream homologous region (left arm, LA) was amplified with P1-OX1-For/P2-OX1-Rv-Fusion and the downstream homologous region (right arm, RA) was amplified with P1'-OX1-For-Fusion/P2'-OX1-Rv. The PCR program for the amplification of both arms was as follows: denaturation at 98 ºC for 5 min, followed by 35 cycles of 98 ºC for 30 sec, 65 ºC for 30 sec and 72 ºC for 40 sec. Purified PCR products from the first PCR reaction served both as template and the primers for extension PCR that runs 30 cycles of amplification under the same condition in a 20 μl PCR system. Primers P1-OX1-For/P2'-OX1-Rv were then added to further amplify the fusion PCR products for 20 cycles. The fusion product was subjected to 1% agarose gel analysis. The product with the desired size was purified from the gel and served as the template to amplify the fusion PCR product for 35 cycles in a 40 μl PCR system. The fragment was then cloned into the conjugal vector pJQ200SK as an *Xho*I/*Bam*HI fragment to produce pJQ200SK-OX1 (Fig. 3-4B).

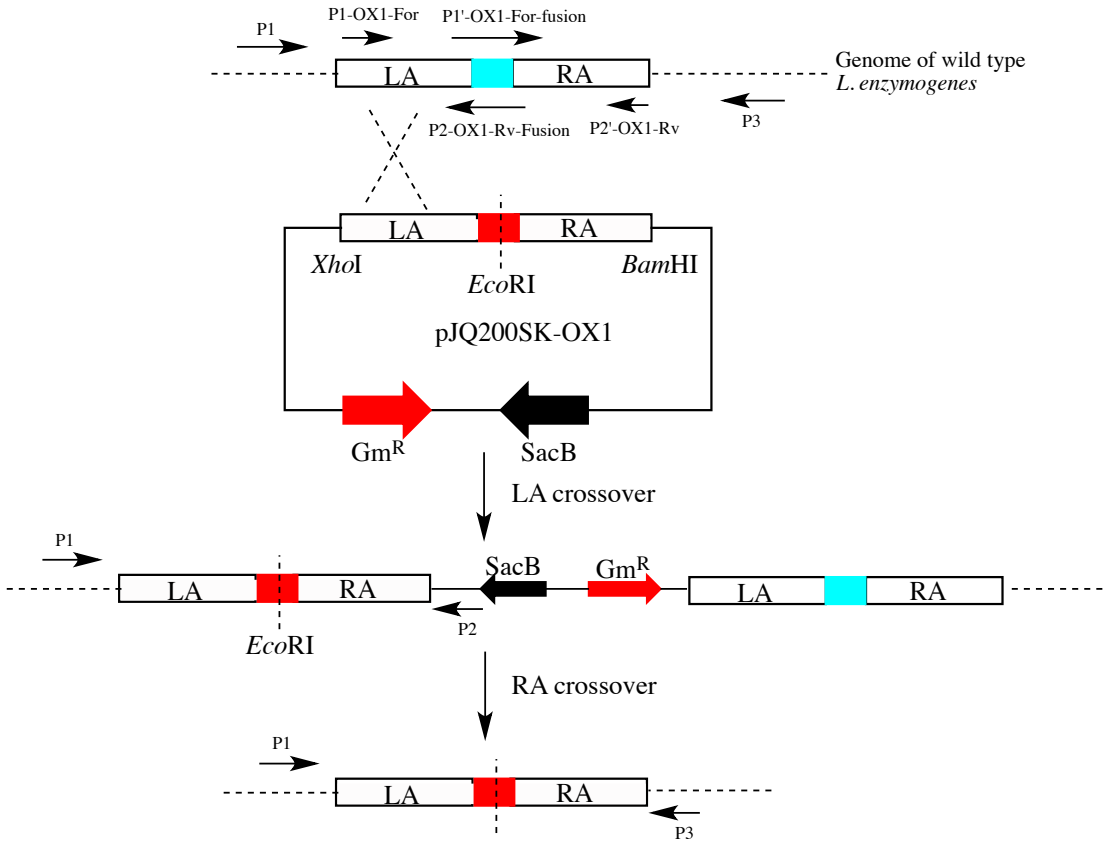


Figure 3-4. Scheme for OX1 gene deletion. The replacement cassette was cloned into the conjugation vector pJQ200SK through *Xho*I/*Bam*HI. Red bar represents the *Eco*RI site introduced to the host genome to replace a 27 bp sequence of the original gene. Primers P1/P2 were used for single crossover mutant screening. Primers P1/P3 were used for double crossover mutant screening.

To construct the OX2 gene in-frame deletion vector, a 536-bp fragment was amplified from OX2 gene, with an *Eco*RI site introduced to replace an 18-bp sequence within the gene. This would replace 6 amino acid residues (GGGVGG) within a sequence motif that is highly conserved among the FAD-dependent oxidoreductases. The procedure for generating OX2 in-frame deletion mutant was identical to that for OX1. The following primers pairs were used for these experiments.

P1-OX2- Forw: 5’-CCA CTC GAG CCC GAC CAT GCC GCA C-3’

P2-OX2-Revs-fusion: 5’-TGC GCC GAA TTC GCC AGC GAT GAC GAT CAC CCG CGG CGT-3’

P1’-OX2- Forw-fusion: 5’-GCT GGC GAA TTC GGC GCA CTG TCC ACC GGC GTC TAC GGC-3’

P2’-OX2- Revs: 5’-CTT GGA TCC GCG GCG CAG GTC GGC-3’

The following primers were used to screen for OX2 single crossover mutants. A 657-bp fragment was produced from the positive colonies.

N1-OX2 5’-ACC ACG AGG ACC CGA CCA TGC-3’

N2-OX2 5’-ACC ATG ATT ACG CCA AGC-3’

N3-OX2 5’- CCACTCGAGCCCGACCATGCCGCAC -3’

The OX2 double crossover mutants were screened using primers N1-OX2 and N3-OX2, which amplified a 536-bp fragment from the positive mutants. An *Eco*RI digestion of this fragment yielded two products of 161-bp and 375-bp, as expected.

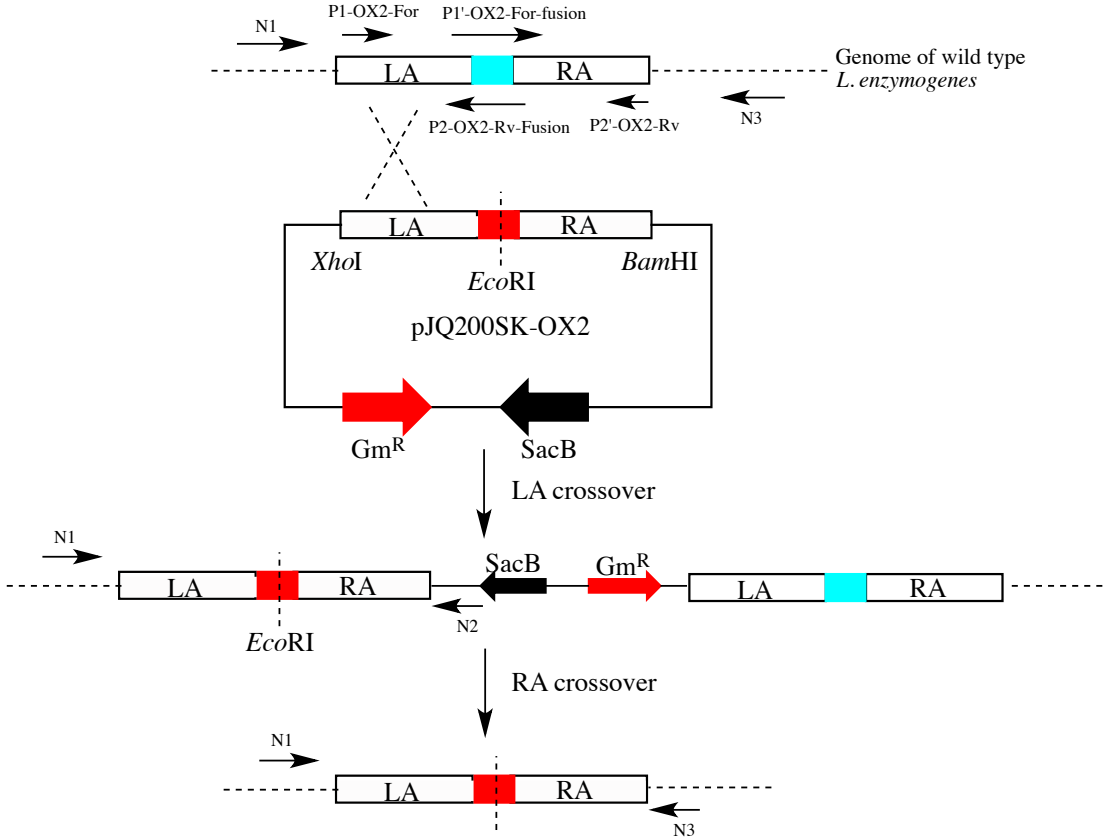


Figure 3-5. OX2 gene deletion scheme, replacement cassette was cloned into the conjugation vector pJQ200SK through *Xho*I/*Bam*HI. Red bar represents the *Eco*RI site introduced to the host genome to replace a 18 bp sequence of the original gene. Primers N1/N2 were used for single crossover mutant screening. Primers N1/N3 were used for double crossover mutant screening.

To conjugally transfer pJQ200SK-OX1 and pJQ200SK-OX2 into *Lysobacter enzymogenes* strain C3, the procedure was identical to the previously described. To screen the single crossover mutants, the genomic DNA was prepared from 1/10 TSB cultures of the individual colonies to serve as templates for PCR reaction. For the OX1- deletion mutants, primers P1 and P2 were used to screen the left arm single crossover mutants (Fig. 3-4). A 1083-bp fragment was expected from the positive colonies. The single crossover mutants of OX2-deletion mutants were identified similarly using primers N1 and N2 (Fig. 3-5). A 657-bp fragment was expected from the positive colonies. For both PCR reactions, the genomic DNA of wild type *Lysobacter enzymogenes* C3 served as the template for a negative control. No product with expected size was observed in the negative control.

The single colonies identified from single crossover were subjected to liquid cultures containing 5% (W/V) sucrose to select for the loss of the vector through a second homologous recombination. The single crossover mutants were grown in 1/10 tryptic soy broth (TSB) medium for 14 hours. Then the individual cultures were re-inoculated into 1/10 TSB (with 1:100 ratio) containing 5% sucrose and 25μg/ml kanamycin medium. Aliquots (50 μl) were taken out every 3 hours from this liquid culture and spread onto 1/10TSA plates (5% sucrose, 25 μg/ml kanamycin). Single

colonies were picked up from the plates and continued to grow in 1/10TSB (5% sucrose, 25 μg/ml kanamycin) liquid medium for another 2-5 days. To confirm the double crossover mutants, the genomic DNA was prepared from 1/10TSB cultures of the individual single colonies to serve as templates for PCR reaction. OX1 double crossover mutants were screened using primers P1 and P3, which will amplify a 1198-bp fragment from the positive mutants. An *Eco*RI digestion of this fragment will yield two products of 736-bp and 446-bp. OX2 double crossover mutants were screened using primers N1 and N3, which will amplify a 536-bp fragment from the positive mutants. An *Eco*RI digestion of this fragment will yield two products of 375-bp and 161-bp.

3-2-3. Generation of OX3, 4 gene deletion mutants

(note: I acknowledge Yunxuan Xie for preparing the OX3, 4 gene deletion mutants.)

To construct the OX3 in-frame deletion vector, primer extension PCR reactions were performed to generate a 593-bp fragment of OX3 gene with an *Eco*RI site introduced to replace a 27-bp sequence within the gene. This would delete 9 amino acid residues (LVEDDRAVG) within a sequence motif that is highly conserved among the FAD-dependent oxidoreductases. The procedure for generating OX3 in-frame deletion mutant was identical to that for OX1. The following primers were used for these experiments.

P1-OX3-For 5′-ATTCTAGATTCCTCCCCTACTTCAACGCCATCCG-3′

P2-OX3-Rv-Fusion
5′-GCCCGTCGCTCAGGCGGAGAATTCGATCTCCTCGACTTTCGC-3′

P1'-OX3-For-Fusion

5′-GCGAAAGTCGAGGAGATCGAATTCTCCGCCTGAGCGACGGGC-3′

P2'-OX3-Rv 5′-atctcgagactgcacgttgatgctcgggtggcg-3′

To construct the OX4 gene in-frame deletion vector, a 525-bp fragment was amplified from OX4 gene, with an *Eco*RI site introduced to replace a 33-bp sequence within the gene. This would delete 11 amino acid residues (PFRPGDLVQGF) within a sequence motif that is highly conserved among the Zn-dependent alcohol dehydrogenases. The following primers pairs were used for these experiments.

P1-OX4-For 5′-ATACTAGTGACCGCAAGC ACATGAAGACCGAACG-3′

P2-OX4-Rv-Fusion

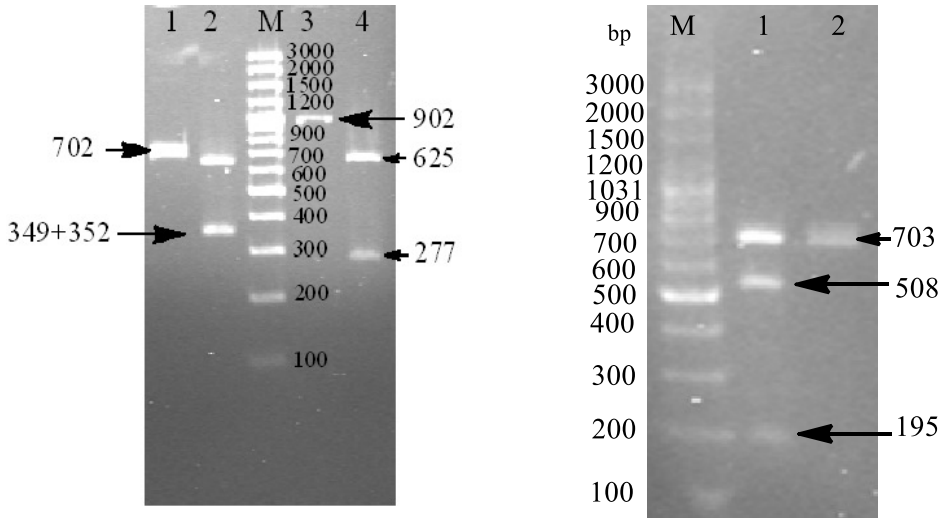
5′-AGGTGCGTGCGCCAGCCGCCGAATTCCGCGTTCGGGCCGGCATCGA-3′

P1'-OX4-For-Fusion

5′-TCGATGCCGGCCCGAACGCGGAATTCGGCGGCTGGCGCACGC ACCT-3′

P2'-OX4-Rv 5′-ATCTCGAGAACCAGGGTGCCGATCATGCCCGAGG-3′

To conjugally transfer pJQ200SK-OX4 and pJQ200SK-OX3 into Lysobacter enzymogenes strain C3, the procedure was identical to OX1 and OX2.



(A) (B)

Figure 3-6. PCR screening of the OX3 and OX4 in-frame deletion mutants. (A) OX3 and OX4 deletion double crossover mutants were identified through a PCR reaction using primers P1/P3 and N1/N3. The expected size of the PCR product from OX3 deletion mutants is 702 bp (Lane 1) and yields 349+352 bp fragments upon *Eco*RI single digestion (Lane 2). The expected size of PCR product from OX4 deletion mutants is 902 bp (Lane 3) and yields 625+277 bp fragments upon *Eco*RI single digestion (Lane 4). (B) SmaI digestion of OX3 deletion PCR product (Lane 2). The expected size of the digestion product is 508+195 bp (Lane 1). 100 bp DNA ladder plus is used as the marker.

3-2-4. Metabolites analysis of OX1-4 in-frame deletion mutants

The verified OX1-4 in-frame deletion mutants were cultured 1/10TSB for metabolites analysis. Metabolites were extracted and analyzed according to the method as previously described[78]. A peak with similar retention time of HSAF was obtained from OX4 deletion mutant and was further identified using high resolution mass spectrometry on Finnigan LCQ.

3-2-5. Structure identification of the new compounds produced by OX1-4 deletion mutant

(note: This part was accomplished by collaboration with Dr. Shen Yuemao at School of Life Sciences, Shandong University, Jinan, China.)

A 10-liter solid LB culture with 1.5% maltose was used for culturing OX4 mutant. The solid medium was chopped and extracted with ethyl acetate. Crude extract (5.8 g) was fractionated on normal pressure reverse phase C18 column (80 g) and eluted with

methanol/water mobile phase starting from 1 liter 60% methanol in water, followed by 2 liters 80% methanol in water and 1 liter 100% methanol. Each fraction was analyzed on HPLC (Agilent, ZORBAX SB-C18, 4.6×250 mm, 1 ml/min, 318 nm) using gradient elution program with mobile phase A (water with 0.005 M TFA) and mobile phase B (acetonitrile with 0.005 M TFA) starting from 5% phase B in A in the first five minutes, 5% to 40% phase B in A from 5 to 10 minutes, 40% to 60% phase B in A from 10 to 15 minutes, 60% phase B in A from 15 to 20 minutes, 60% to 100% phase B in A from 20 to 22 minutes, 100% phase B in A from 22 to 24 minutes and 100% to 5% phase B in A from 24 to 26 minutes. The 80% methanol/water eluent (3.4 g, designated as component A) was found to contain the desired compound. Component A was then further fractionated on Sephadex LH-20 (140 g, 3 ml/tube, 8 s/d), and each fraction was analyzed by HPLC. Desired compound was obtained as 50% methanol (0.005 M TFA) eluent by HPLC (Agilent, ZORBAX SB-C18, 9.4×250 mm, 4 ml/min, 318 nm).

3-3. Results and Discussion

3-3-1. Construction of OX1 and OX2 in-frame deletion mutants

The in-frame deletion mutants were constructed through two homologous recombination events. To identify OX1-deletion single crossover mutant at left arm, primers N1/N2 were used (Fig. 3-4). A 1198-bp fragment was obtained from the positive colonies as expected (Fig. 3-7A). The OX1-deletion single crossover mutants were identified similarly using primers N1/N2 (Fig. 3-5). A 536-bp fragment was obtained from the positive colonies as expected (Fig. 3-7B). To identify both double

crossover mutants, diagnostic PCR reactions were performed using primer pair P1/P3 and N1/N3. Expected size of each PCR product was obtained (Fig. 3-7C-D).

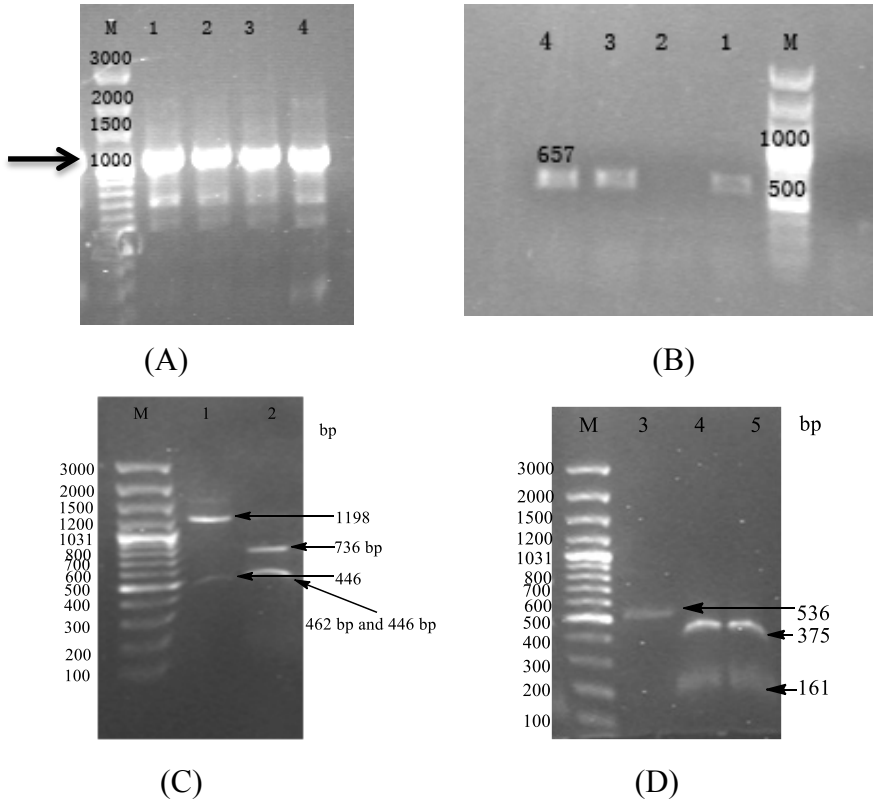


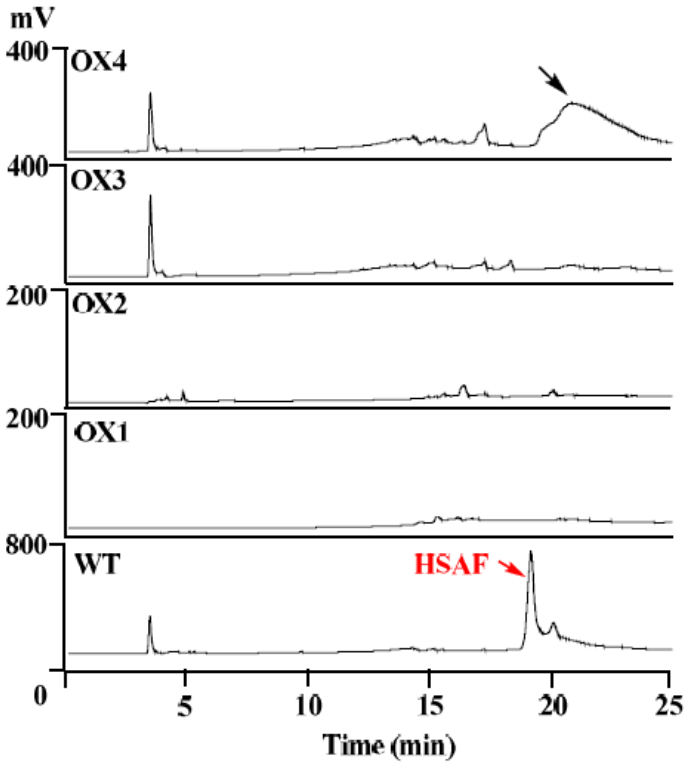
Figure 3-7. PCR screening of the OX1 and OX2 in-frame deletion single and double crossover mutants. (A) OX1 deletion single crossover mutants were identified through a PCR reaction using primers P1/P2. The expected size of amplicon is 1083 bp. (B) OX2 deletion single crossover mutants were identified through a PCR reaction using primers N1/N3. The expected size of amplicon is 657bp. (C) OX1 deletion double crossover mutants were identified through a PCR reaction using primers P1/P3. The expected size of the PCR product from OX1 deletion mutants is 1198 bp (Lane 1) and yields 736 and 446 bp fragments upon *Eco*RI single digestion (Lane 2) (D) OX2 deletion double crossover mutants were identified through a PCR reaction using primers N1/N3. The expected size of the PCR product from OX2 deletion mutants is

536 bp (Lane 3) and yields 375 and 161 bp fragments upon *Eco*RI single digestion (Lane 4,5).

3-3-2. Metabolites analysis of OX1-4 in-frame deletion mutants

Verified OX1-4 deletion mutants were cultured in 1/10 tryptic soy broth in which wild type C3 produces large quantities of HSAF. Metabolites of the wild type C3 and the four mutants were analyzed by HPLC. OX1-3 deletion mutants did not produce HSAF and its intermediate compounds, while OX4 deletion mutant produced a compound with similar retention time to HSAF (Figure 3-8A). High resolution mass spectrometry analysis showed that the compound had two mass units smaller than HSAF (Figure 3-8B). The structure of this compound was determined by ^{1}H-NMR and ^{13}C-NMR to be alteramide A (**5**) (Figure 3-1). These results suggested that the OX genes are involved in HSAF biosynthesis and are part of the HSAF gene cluster.

(A)



(B)

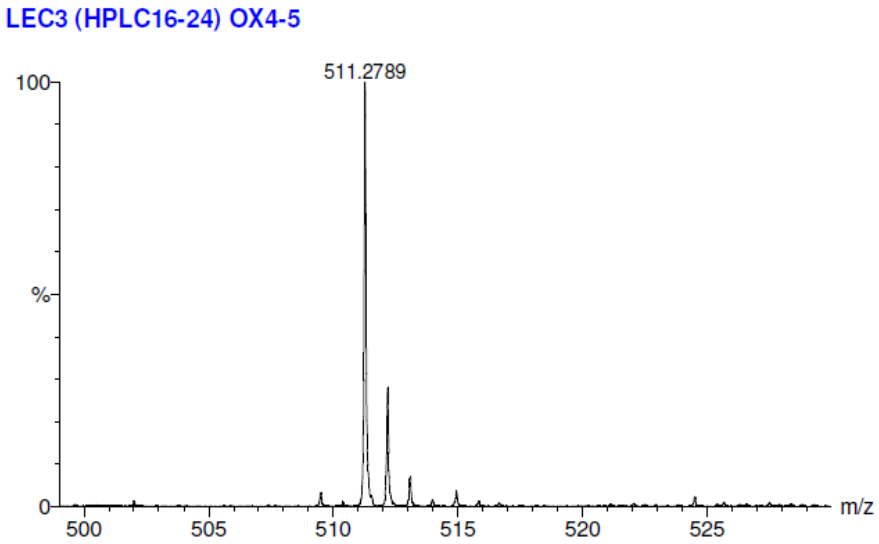


Figure 3-8. (A) HPLC analysis of the metabolites produced by the four OX gene in-frame deletion mutants. The new compound produced by the OX4 mutant and the HSAF peak produced by the wild type strain are indicated with an arrow. (B) MS analysis of the peak at ~22 minute in the OX4 mutant in (A).

3-4. Summary and final remarks

HSAF represents a new family of antifungal compounds with a novel mechanism of action and distinct structural features, including a 5,5,6-tricyclic system fused with a 17-membered macrolactam that includes the tetramic acid moiety. These features make HSAF an interesting target to study its biosynthetic mechanism. Based on sequence annotation of the HSAF gene cluster, we first proposed that a single modular PKS could synthesize two polyenes that joint together by the amino acid ornithine via two amide bonds. Through co-operation with the 6 redox enzymes, the polyene-tetramate-polyene intermediate will finally be cyclized and tailored into

HSAF or analogs. In this work, we defined the boundary of HSAF biosynthetic gene cluster, which suggested that all OXs may be involved in HSAF production, although only 2 redox genes are present in the ikaguramycin biosynthetic gene cluster.

Besides these 6 redox genes, ORF 1 and 10 are also considered to be part of the HSAF biosynthetic core genes. ORF 1 encodes a putative major facilitator superfamily transporter, whereas ORF 10 encodes a putative TonB-dependent outer membrane receptor (Figure 3-2). These two genes might not be essential for biosynthesis of HSAF core structure, but are likely to be involved in HSAF transport and regulation. Together, these results indicate that the region between ORF1 and ORF10 makes the minimal gene cluster of HSAF biosynthesis.

Chapter 4: *In vitro* Characterization of Redox Enzymes in HSAF Biosynthetic Pathway

Portions of this chapter were reported (RSC Advances. 2015, 5, 11644-11648) and are reproduced with the permission of the publisher[111].

4-1. Background and significance

Polycyclic tetramate macrolactams are identified by using a combination of small molecule chemistry, biosynthetic analysis, and genome mining in diverse bacteria. The PTM group emerges as a new class of natural products with a distinct structure and a new mode of action. The most striking feature of the PTM in biosynthesis, where only a single hybrid PKS-NRPS is sufficient for the synthesis of the characteristic PTM scaffold[26, 76]. The results presented in chapters 2 and 3 showed that a single-module PKS is responsible for the biosynthesis of two separate polyketide chains. The hybrid PKS-NRPS for HSAF biosynthesis has a typical modular organization, including KS-AT-DH-KR-ACP for the PKS portion and C-A-PCP-TE for the NRPS portion. There is no obvious remnant of an inactive enoylreductase (ER°) domain or a methyltransferase (CMeT) domain, as seen in several iterative fungal PKS-NRPS with the similar organization, such as PKS for tenellin, lovastatin and compactin[112, 113]. Although this type of iterative PKS-NRPS is commonly seen in fungi, it is not known in bacteria until recent years. The phylogenic study also suggested that the PTM iPKS evolves not far from aromatic iPKS (Figure 1-2).

As shown in Figure 3-2, the HSAF PKS-NRPS is flanked by six redox genes, four NADP/FAD-dependent oxidoreductase (OX) genes on the left and a

hydroxylase/sterol desaturase gene and a NADP/FAD-ferredoxin reductase gene on the right. The disruption or in-frame deletion of the PKS-NRPS, OX1, OX2, or OX3 eliminated HSAF production, whereas the OX4 mutant and hydroxylase/desaturase mutant produced putative precursor/shunt products [77,78].

The objective here is to express and purify the redox enzymes, characterizing the enzymatic reactions, and reconstitute HSAF biosynthesis using the enzymes and substrates.

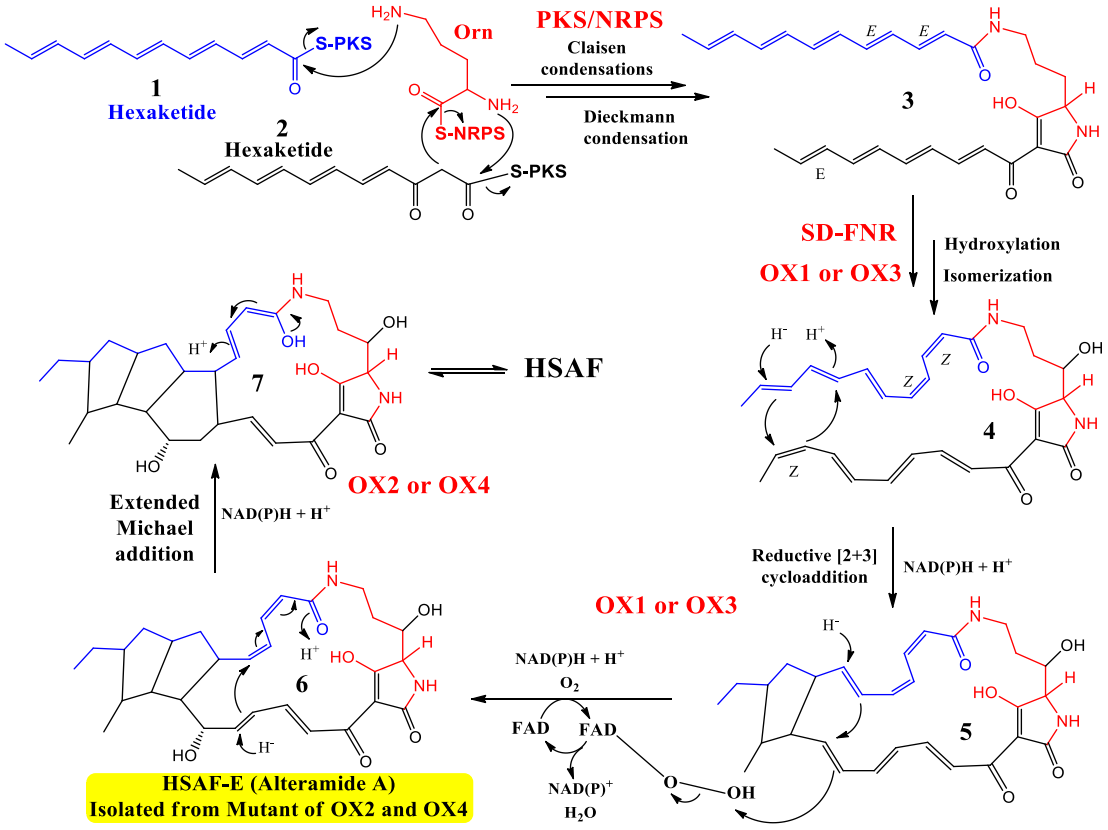


Figure 4-1. Proposed mechanism of redox enzymes participate in HSAF tailoring[78].

We have already shown that the NRPS module catalyzes the formation of two amide bonds between acyl-SNAC and the two amino groups of ornithine, as well as the formation of the tetramic acid unit, to produce a product (see Fig. 2-5 in Chapter 2) similar to polyene-tetramate-polyene intermediate **3** in Figure 4-1. In the previous

experiments, we used saturated acyl chains for the ease of substrate preparation[78]. However, the nascent products of the PKS are polyenes (**1** and **2** in Figure 4-1) and the NRPS uses the polyene acyl-*S*-PKS as substrates to condense with ornithine (Figure 4-1). We propose the PKS-NRPS cooperates with the redox enzymes to tailor the polyene chains of **3** into the tricyclic system of HSAF. The polyene-tetramate-polyene feature may help form a relatively rigid scaffold and bring together the alkene functionalities in close proximity to allow intramolecular cyclizations to take place. One issue here is the configuration of the alkenes. The structure of HSAF suggests that three of the alkenes need to be in *cis* configuration (Figure 4-2). PKS DH domains usually generate a trans double bond, although some can make a *cis* double bond [65, 114]. As only a single DH domain is present, we propose all the double bonds in **1** and **2** are in the *trans* form and three of them are subsequently isomerized to the *cis* form in **4**. No alkene isomerase-like enzyme is present in the cluster. However, flavin-dependent redox enzymes could catalyze the isomerizations [115], perhaps via a reversible hydrogenation- dehydrogenation. One or more of the OX enzymes in the cluster may be involved in the isomerizations.

Next, through two reductive intramolecular cycloaddition reactions, one [2+4] and one [2+3], **4** is converted to **5** and then **6**. "Diels-Alderase" activity has been observed in the biosynthesis for several fungal polyketides/peptides, such as lovastatin, solanapyrone A, macrophomic acid, and spinosyn A[116-122]. In the HSAF type of compounds, because a reduction is also needed in order to form the unusual ring system, we propose both the PKS-NRPS and the redox enzymes cooperate in the cyclizations (Figure 4-1). The substrate specificity of the redox enzymes may also play a key role in determining the chain length and reduction level of the initial PKS products **1** and **2**. The cycloadditions are apparently stereospecific, resulting in nine

chiral centers. It should be noted that the cycloadditions could precede the tetramic acid formation, while the intermediate is still bound on NRPS.

In addition to cycloaddition, a possible route to ring formation is via epoxide opening, followed by reductive coupling. The HSAF gene cluster does not contain a candidate epoxidase or epoxide hydrolase, so this possibility may be less likely. In summary, there are many very intriguing aspects in the proposed biosynthetic mechanism.

We have established the methods for PKS and NRPS *in vitro* reactions. The *in vitro* reactions for polyketide-peptide biosynthesis combine ornithine, acetyl-CoA, malonyl-CoA, ATP and NADPH, with various combinations of HSAF enzymes. To test the activity of the redox enzymes, we will include the redox enzymes, individually or in various combinations, in above reaction mixtures. The expectation is that the PKS-NRPS-synthesized product like **3** will serve as *in situ* substrate for the redox enzymes. By including various combinations of enzymes, we investigated the activity of individual enzymes. Additionally, chemically synthesized polyene acyl-SNAC, mimics of **1** and **2**, were used in the tests. By using acyl-SNAC (Figure 4-2) with various chain lengths and reduction levels, the specificity of the enzymes will be revealed.

4-2. Materials and Methods

4-2-1. Strains, Vectors, Chemicals and Molecular Biology Agents

Escherichia coli DH5α was used as the host for general plasmid DNA propagation. *E. coli* BL21 was used for protein expression. Vector pANT841 was used for general cloning, vector pET28a was used for protein expression in *E. coli* BL21, and vector

pET32b, pBAD-myc-His, pQE60, pMAL, pGEX-6p-1, pYES were tested to increase the protein solubility. Chemicals were purchased from Sigma or Fisher Scientific. PCR primers were synthesized by Integrated DNA Technologies (IDT, Coralville, IA). Kits for plasmid preparation and DNA extraction were from Qiagen (Valencia, CA). Standard molecular biology methods were used for all other DNA manipulations.

4-2-2. OX1-4 expression in *E. coli*

To construct the expression vector, OX1-4 genes were amplified by PCR using Taq DNA polymerase and the genome extracted from *L. enzymogenes* OH11 as template. The primers are listed in table.

Primer design:

OX1-Rev-*Xho*I: 5'- TCACTCGAGTCAGGCGGCCGGAAGC -3'

OX1-Forw-*Bam*HI: 5'-TAAGGATCCGCCGGCATGGCCACCG-3'

OX2-Rev-*Hin*dIII: 5'- TCAAAGCTTTCAGACGATGGTGGCG-3'

OX2-Forw-*Bam*HI: 5'-TAAGGATCCGCGCGCGATAACGGCC-3'

OX3-Rev-*Xho*I: 5'- TCACTCGAGTCATGTGCTTGCGGTC -3'

OX3-Forw-*Bam*HI: 5'-TAAGGATCCATGCCGAGCGGCAAGC-3'

OX4-Rev-*Xho*I: 5'- TCACTCGAGTCACACCCGCACCAGCAC -3'

OX4-Forw-*Bam*HI: 5'-TAAGGATCCATGAAGACCGAACGTTG-3'

The PCR fragment was digested and cloned into pANT841 at the same sites. The constructs were sequenced, and the result showed that all the OX1-4 sequences were

correct. The OX genes were released from pANT841 and cloned into expression vector pET28a. To express the OXs, each of the pET28a constructs was introduced into *E. coli* BL21 (DE3). Single colonies were inoculated in 3 ml of liquid LB medium containing kanamycin (50 μg/ml) and incubated in a shaker (250 rpm) at 37 °C overnight. The overnight culture was added to 50 ml fresh LB medium and incubated in a shaker (250 rpm) at 37 °C until the cell density (OD_{600}) reached 0.6. To induce the expression of the OXs, IPTG (0.1 mM) was added to the culture, and the cells were allowed to grow at the same conditions for another 3 h. To prepare the soluble fraction, the cells were harvested and resuspended in 2 ml of Tris-HCl buffer (250 mM NaCl, 100 mM LiCl, 50 mM Tris). The cell suspension was treated with lysozyme (1 mg/ml) and sonicated five times on ice. The soluble fraction of protein extracts was loaded onto a Ni-NTA column, which was previously equilibrated with Tris buffer containing 10 mM imidazole. The column was washed three times with the buffer containing 20 mM imidazole, and the His_6-tagged protein was eluted twice with 200 μl of Tris buffer containing 250 mM imidazole.

Other expression vectors were also used to increase the protein solubility, including pBAD-myc-His, pQE60, pMAL, and pGEX-6p-1. The expression methods are according to the standard protocols.

4-2-3. Redox enzymes expression in yeast

The four redox genes OX1-4 were amplified by PCR using *Taq* DNA polymerase and the cosmid COS4-1 as template, which contains part of the HSAF gene cluster. The OX1, OX3 and OX4 fragment were digested with *Xho*I and *Bam*HI; the OX2 fragment was digested with *Hin*dIII and *Bam*HI. The fragments were cloned into

pET28a vector at the same sites. The four redox genes were released from pET28a and cloned into expression vector pYES2.

To express the redox enzymes, the pYES2 constructs were introduced in *S. cerevisiae* strain INVSc1. The transformation and expression were followed the instruction from Invitrogen (Cat. no.V825–20). The cells were pelleted and resuspended in microsome extraction buffer (50 mM Tris, 1 mM EGTA, 1 mM β-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride, 2 μg/mL pepstatin A, pH 7.5) (2 mL/g). After adding glass beads, the cells were vortexed 60 seconds for 6 cycles. The glass beads were filtered out, followed by centrifuge at 5000 × g for 10 min. The supernatants were centrifuged at 100,000 × g for 1 hour using Optima^{TM} MAX-XP Ultracentrifuge (Beckman Coulter, CA). Microsomes were suspended at 1 mL/g in the microsome extraction buffer with 33% glycerol and stored at -80 °C.

4-2-4. *In vitro* assay of OX1-OX4 microsomes using alteramide A as substrate

The *in vitro* activity assay comprises 400 μM alteramide A, 1 mM FAD, 2 mM NADH, 50 mM NaCl, and 10 μg OX1-4 microsomes, in 50 mM Tris·HCl buffer (pH 8.0). The control reactions used micorsomes extracted from yeast cells which host pYES2 blank vectors. After continual incubation for 2 h at 37 °C, the reactions were stopped by adding an equal volume of ethyl acetate. The mixtures were centrifuged at 13,200 rpm for 5 min, and the top layers were transferred to new tubes. The solutions were air dried, and the residues in the tubes were re-dissolved in 150 μl methanol. The methanol solutions were centrifuged, transferred to new tubes, and dried again.

In the *in vitro* HSAF biosynthetic reaction, the system described in Chapter 2 was incubated with the microsomes of the four OXs, 10 μg each, together with 1 mM

FAD, 2 mM NADH and 0.1 mM BSA. After incubation for another 2 h at 37 °C, the reactions were stopped and extracted using the same method described above.

Finally, the products of the reactions were re-dissolved in 20 µl methanol and analyzed by LC-MS (HPLC: Agilent Technologies, 1220 Infinity LC; MS: Finnigan mat, LCQ). LC-MS column was Alltima C18LL (5 µ, 1.0 mm × 250 mm). Solvent A was H_2O containing 0.1% formic acid; Solvent B was acetonitrile containing 0.1% formic acid. The flow rate was 50µl/min, with a binary gradient system (0-15 min, 5%-45% B gradient; 15-35 min, 45% B to 90% B gradient; 35-46 min, 90% B to 100% B gradient; 46-55 min, 100% B; 55-56 min, 100% B to 5% B gradient; 56-66 min, 5% B).

4-2-5. Synthesis of a 2,4,6,8,10-dodecapentanoic acid thioester

The synthesis of a 2,4,6,8,10-dodecapentanoic acid thioester was published on RSC Advances by Andrew Olson, Haotong Chen, Liangcheng Du, and Patrick Dussault[111].

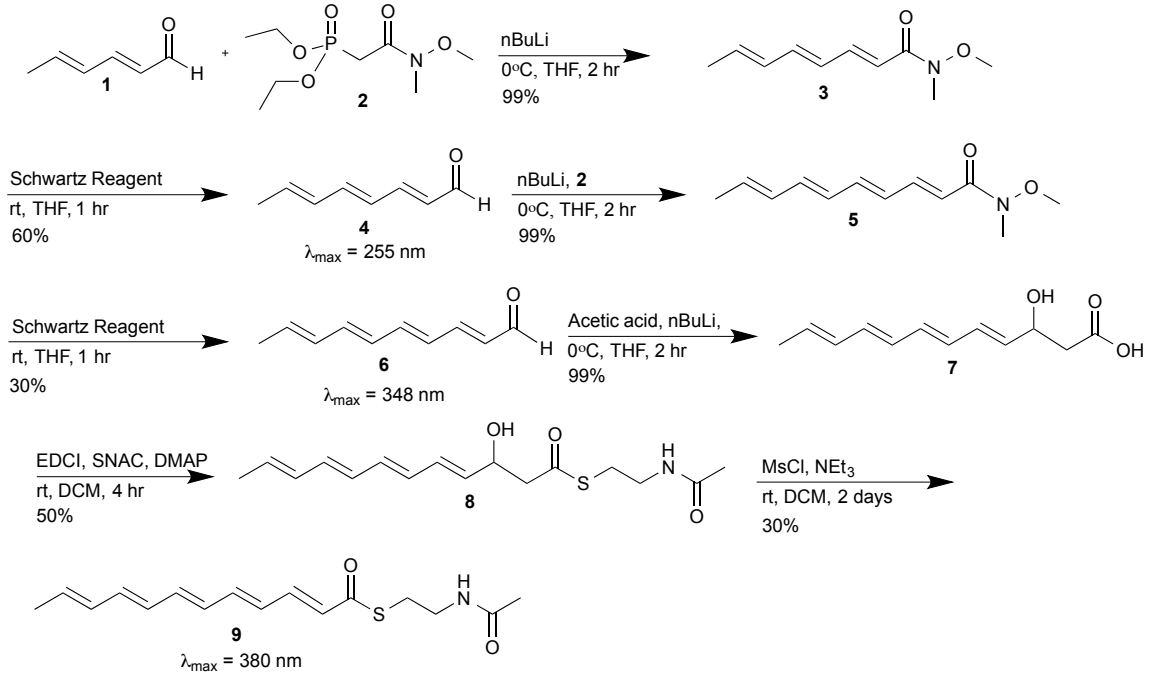


Figure 4-2. Synthesis of all-E-2,4,6,8,10-dodecapentenoate SNAC thioester.

4-2-6. *In vitro* assay for the activity of NRPS using the pentaenoic acid thioester substrate

NRPS needs to be converted to its holo-form by a promiscuous 4'-phosphopantetheinyl transferase (PPTase), Svp, by incubating with coenzyme A. The holo-NRPS was obtained by incubating protein (3 μM) and CoA (0.83 mM) with Svp (5.6 μM) in a 60 μl reaction containing Tris-HCl (100 mM, pH 8.0), $MgCl_2$ (10 mM), and TCEP (0.5 mM). Reaction was incubated at 37 °C for 2 h. Finally, the reaction mixture was added a 40 μl solution containing L-Orn (1.5 mM), ATP (3 mM), Tris-HCl (100 mM, pH 8.0), $MgCl_2$ (10 mM), NaCl (50 mM), EDTA (0.1 mM), and TCEP (0.5 mM). A reaction without NRPS served as the control. The cell free extract (CFE) extracted from ΔPKS mutant[78] was considered to provide crude redox enzymes, which catalyse the 5,5,6-tricyclic ring system formation in HSAF, coupled with 0.5 mM FAD/NADH. The pentaenoate thioester was added right away in a dark room, to make a final concentration of 0.5 mM. After continual incubation overnight at 30 °C, the reactions were stopped by adding 150 μl of 0.2 mM TCA in methanol and were frozen at -20 °C for 30 min. The mixtures were centrifuged at 13,200 rpm for 20 min in a desktop Eppendorf centrifuge, and the supernatants were transferred to new tubes. The solutions were dried in a Speed-Vac, and the residues in the tubes were re-dissolved in 20 μl methanol, and analyzed by Agilent LC-1200 (Santa Clara, CA) connected to a 2.1 x 100 mm Symmetry ODS column from Waters (Milford, MA) and a Triple Quadrupole Mass Spectrometer model 4000 QTrap from ABSciex (Framingham, MA) operating in either single quadrupole (Q1), enhanced mass spectrum (EMS), MS/MS or multiple reaction monitoring (MRM) modes. The samples were injected onto the column and a gradient from 98% mobile phase A (0.1% formic acid in water, J.T. Baker) to 60% B (0.1% formic acid in acetonitrile, Acros

Organics) was run over 15 minutes, followed by 5 minutes of 98% B and 5 min of 98 % A, all at a flow rate of 0.25 mL/min.

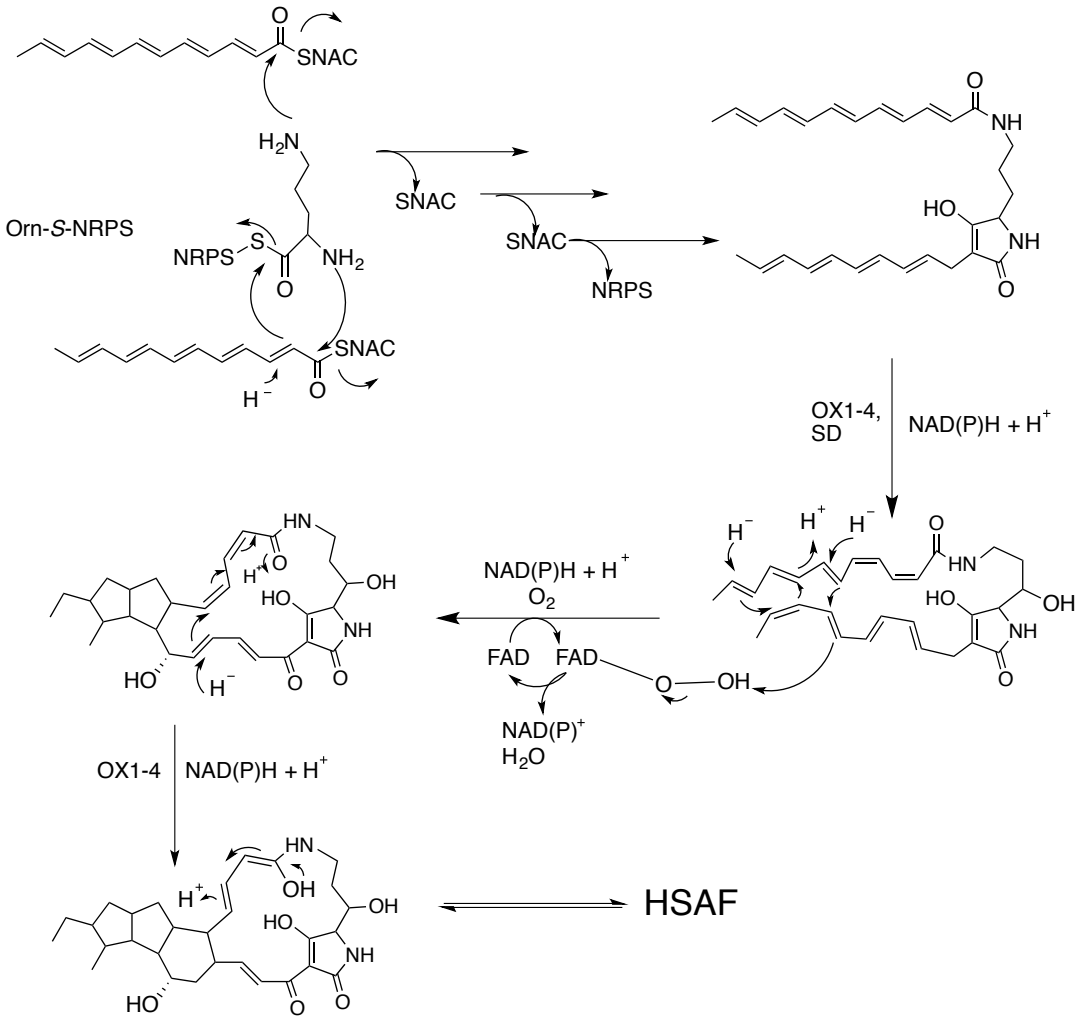


Figure 4-3. A proposed pathway of in vitro biosynthesis of HSAF from *E*-2,4,6,8,10-dodecapentenoate SNAC thioester (**9**).

4.3. Results and discussion

4-3-1. Redox enzymes expression

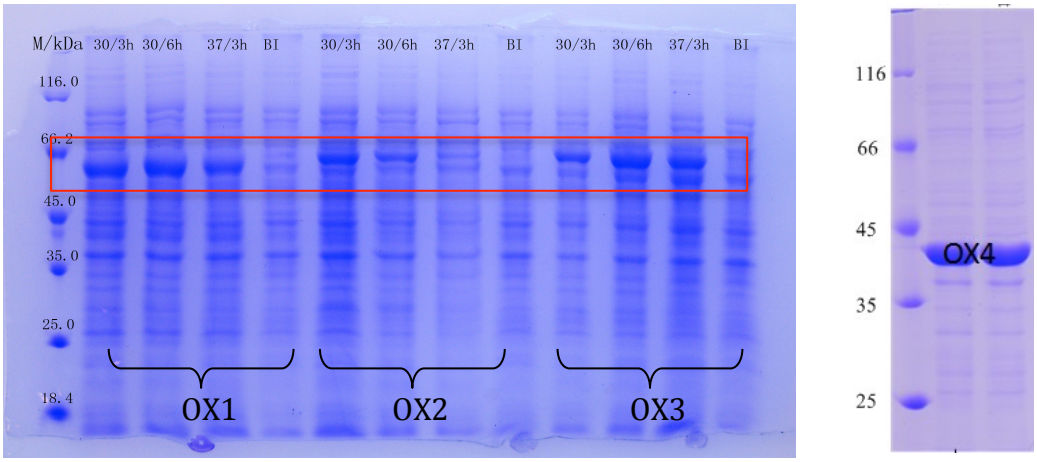
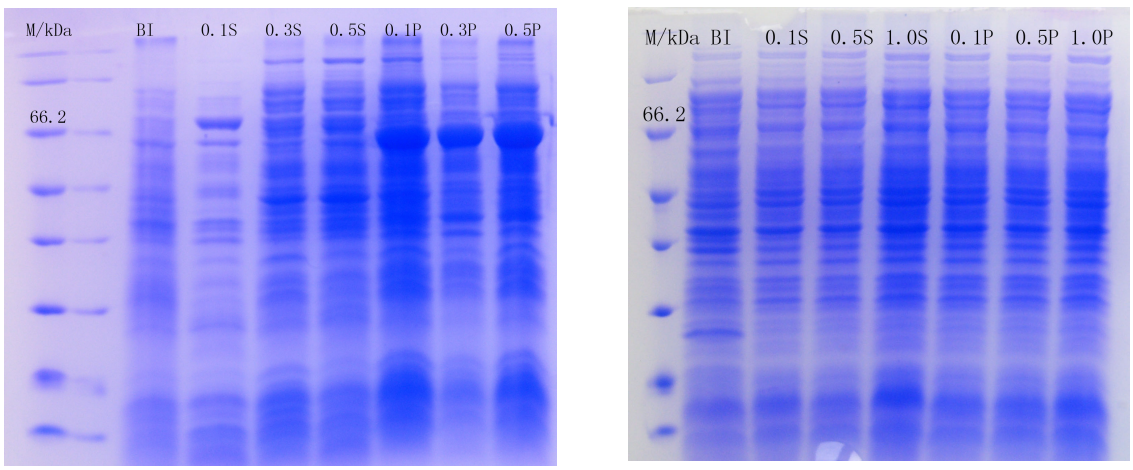


Figure 4-4. SDS-PAGE of OX1-4 expressed in *E. coli* under different expression conditions. M, markers; BI: cell lysate before induction as control. 30, 37: induction temperature is 30 or 37 °C. 3h, 6h: the induction time is 3 or 6 hours.

Since the expressed proteins all ended up in the inclusion bodies, the conditions were optimized, including temperature and induction time, together with switching the IPTG concentration. Other plasmids were used to improve solubility.



(A) (B)

Figure 4-5. (A) SDS-PAGE of the OX1 expressed in *E. coli* under induction of different concentration of IPTG. M, markers; BI: cell lysate before induction as

control. 0.1, 0.3, 0.5: concentration of IPTG is 0.1 mM, 0.3 mM and 0.5 mM. S: total soluble proteins after sonication; P: precipitate portion after sonication.

(B) SDS-PAGE of the OX1-3 in pGEX-6p-1 vector expressed in *E. coli* under induction of different concentration of IPTG. M, markers; BI: cell lysate before induction as control. 0.1, 0.5, 1: concentration of IPTG is 0.1 mM, 0.5 mM and 1 mM. S: total soluble proteins after sonication; P: precipitate portion after sonication.

The results showed that OX1-4 expressed in *E. coli* were present predominantly in the inclusion bodies. To circumvent this problem, OX1-4 was expressed in yeast, and microsomes were extracted and used in *in vitro* assay.

4-3-2. *In vitro* reaction of OX4 using alteramide A as substrate

Recently, the Zhang group reported the enzymatic mechanism for the formation of the inner 5-membered ring and demonstrated the polyketide origin of the ikarugamycin skeleton[81]. A three-gene-cassette ikaABC from the marine-derived *Streptomyces* sp. ZJ306 is sufficient for conferring ikarugamycin production in a heterologous host. IkaC catalyzes a reductive cyclization reaction to form the inner five-membered ring by a Michael addition-like reaction.

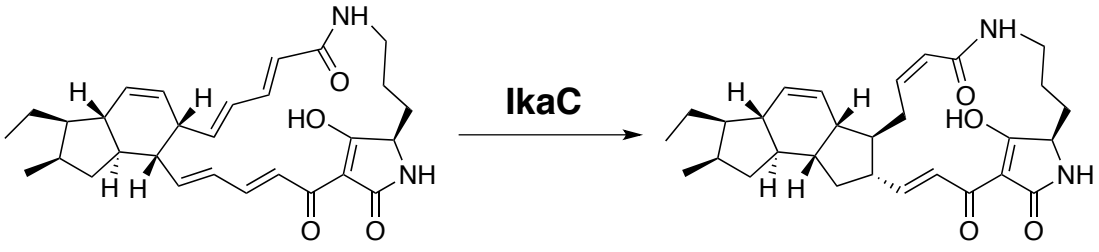


Figure 4-6. The proposed mechanism of IkaC, a redox enzyme from ikarugamycin gene cluster[81].

In the HSAF gene cluster, OX4 gene has been predicted to encoding an alcohol dehydrogenase zinc-binding protein, which has the greatest similarity to *ika*C, thus becoming a candidate for catalyzing the last ring closure. Alteramide A was isolated from the wild type and OX4 disruption mutants, which has the bi-cyclic ring system (Figure 3-1). Although the configuration of H on 12-C is different from HSAF, it could still possible to serve as an alternative substrate for OX4. The molecular modeling using SWISS MODEL indicated a good docking of alteramide A with OX4 (Identity: 32%, Score: -146) using NADPH-dependent oxidoreductase as template (PDB accession number: 2J3I). Thus this compound was used in *in vitro* activity assay for the 4 OX enzymes.

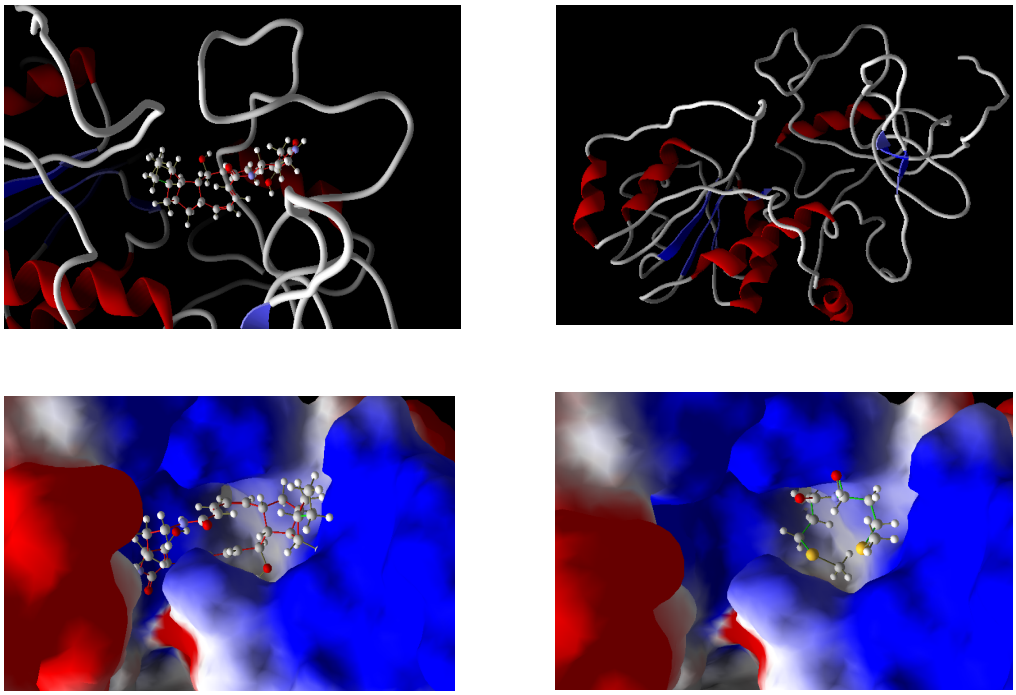


Figure 4-7. Molecular modeling of OX4 with alteramide A as ligand. Template: NADPH-dependent oxidoreductase[123].

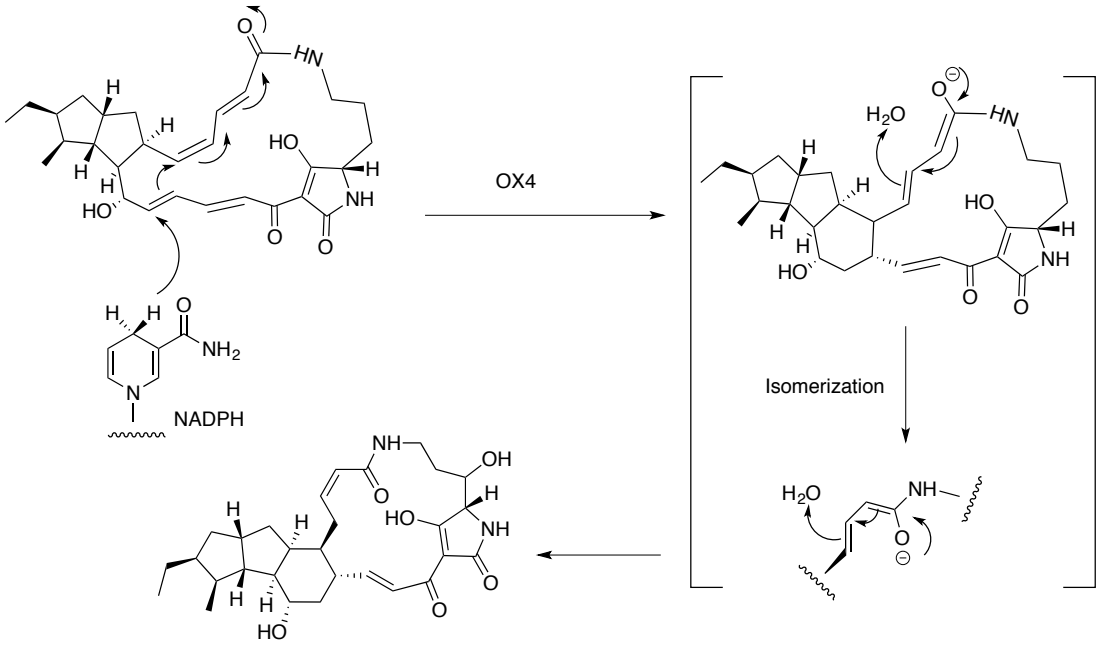


Figure 4-8. The proposed mechanism of last ring closure in HSAF biosynthesis.

In the *in vitro* activity assay, 400 μM alteramide A was incubated with 10 μg OX4 microsomes, together with 1 mM FAD, 2 mM NADH, 50 mM NaCl, in 50 mM Tris·HCl buffer (pH 8.0). The control reaction used micorsomes extracted from yeast cells which host pYES2 blank vectors. After the reactions were stopped, the extracts were analyzed by HPLC and MS analysis. Although the HPLC chromatogram showed no obvious difference between the control and the real reaction, with the time growing, the HPLC peaks revealed several reproducible differences in real reaction group.

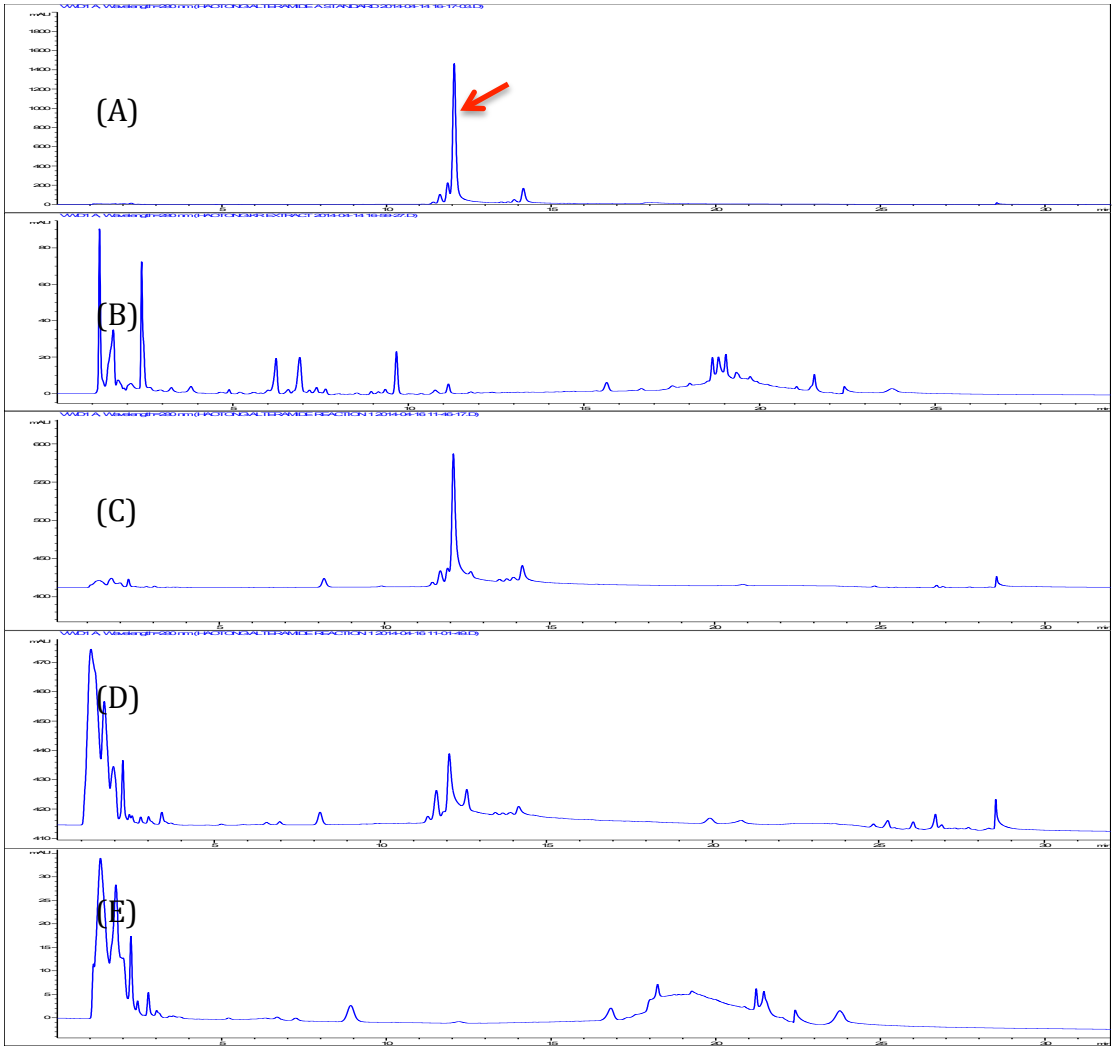


Figure 4-9. HPLC analysis of *in vitro* assays. (A): Alteramide A standard. (B): ΔKR mutant which has no HSAF or analog produced. (C): Control reaction w/o OX4 added. (D): Real reaction with OX4 after 10 hours incubation. (E): Real reaction with OX4 after 20 hours incubation.

The HPLC analysis showed that alteramide A is being consumed during the incubation, but unfortunately, new HSAF-like peaks are not observed by HPLC.

4-3-3. *In vitro* assay for the activity of NRPS using the pentaenoic acid thioester substrate

Although it was proved that the NRPS module catalyzes the formation of two amide bonds between acyl-SNAC and the two amino groups of ornithine, as well as the formation of the tetramic acid unit, to produce a product (see Figure 2-5 in Chapter 2) very similar to intermediate **3** (Figure 4-1) [82], the nascent products of the PKS (**1** and **2**) and the nascent substrates of NRPS, acyl-*S*-PKS, has not been added and evaluated in the *in vitro* reaction system. To examine the activity and selectivity of NRPS toward different substrates, the native substrate all-*E*-2,4,6,8,10-dodecapentenoate SNAC thioester was chemically synthesized and added into the *in vitro* reaction.

The expected final product is the polyene-tetramate-polyene shown in Figure 4-1, and this polyene-tetramate-polyene intermediate was previously obtained by incubating the hybrid PKS-NRPS with a series of co-enzyme and substrates[26]. But this intermediate is very unstable and difficult to be isolated under normal condition. The OXs, which are supposed to catalyze the ring closure reactions to form a rigid 5,5,6-tricyclic ring system, were added into the reaction, so that the resulting products with the ring system should be more stable and easier to be detected.

Pentaenoate thioester **9** was incubated with the NRPS protein, ornithine, phosphopantetheinyl transferase, CoA, crude redox enzymes, ATP, and FAD/NADH at 30 °C, and the products were compared with those resulting from an incubation conducted in the absence of the NRPS protein. LC-MS analysis found a variety of products having a mass consistent with the predicted HSAF precursor or analogues (Figure 4-12), and the compounds with spontaneously formed amide bond (Figure 4-14). However, no HSAF was observed.

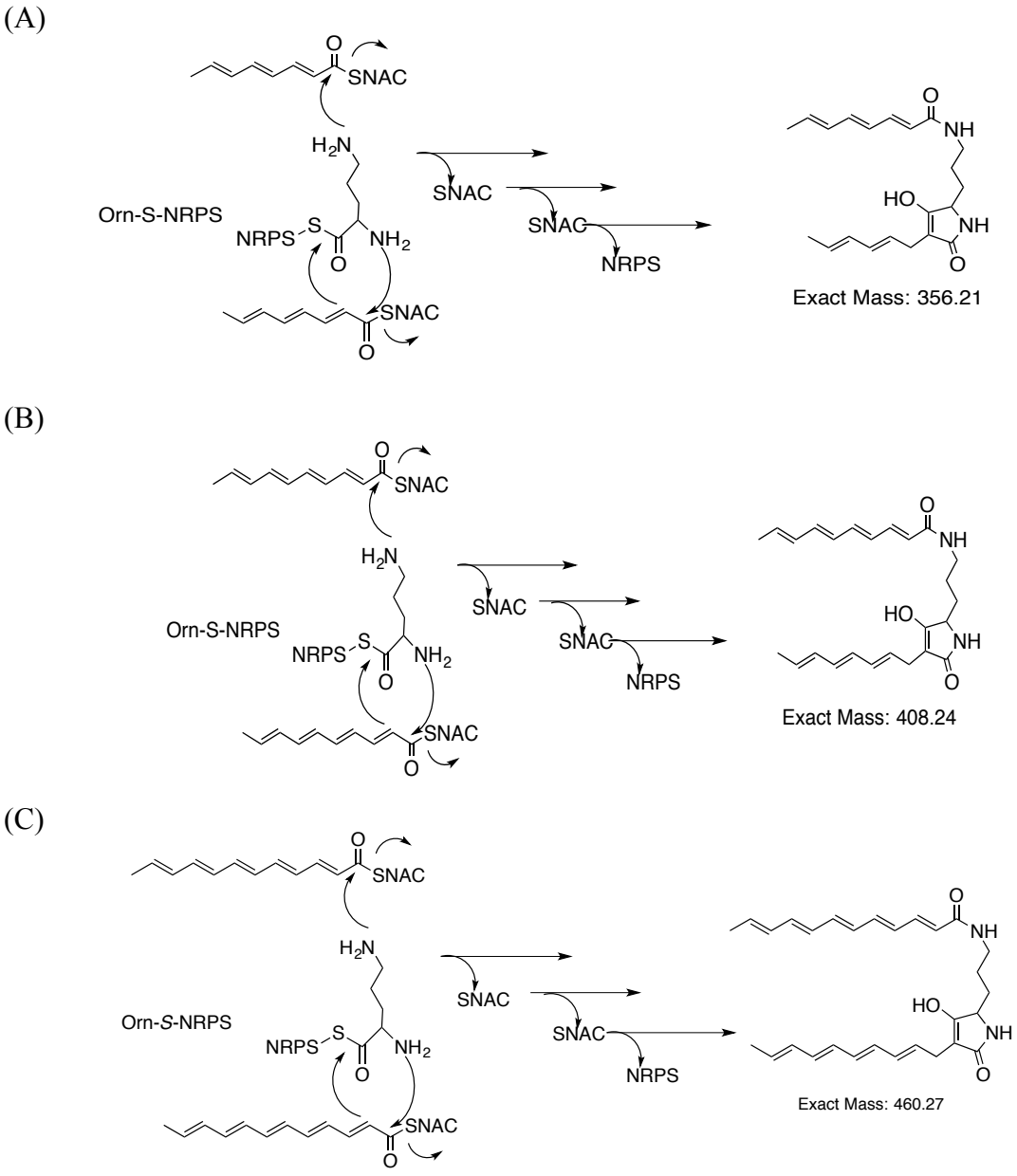


Figure 4-10. Proposed NRPS-catalzyed biosynthesis of (A) tetraketide, (B) pentaketide, and (C) hexaketide polyene tetramate intermediates using the corresponding chemically synthesized SNAC substrates.

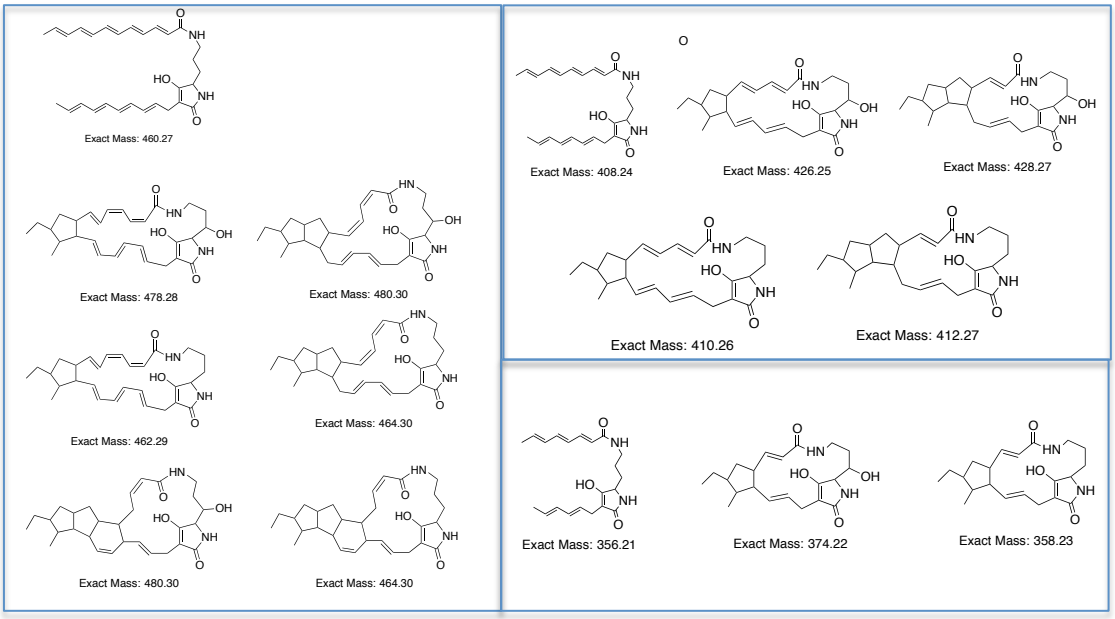


Figure 4-11. Structure of possible intermediates produced in the *in vitro* reactions after cyclization by redox enzymes.

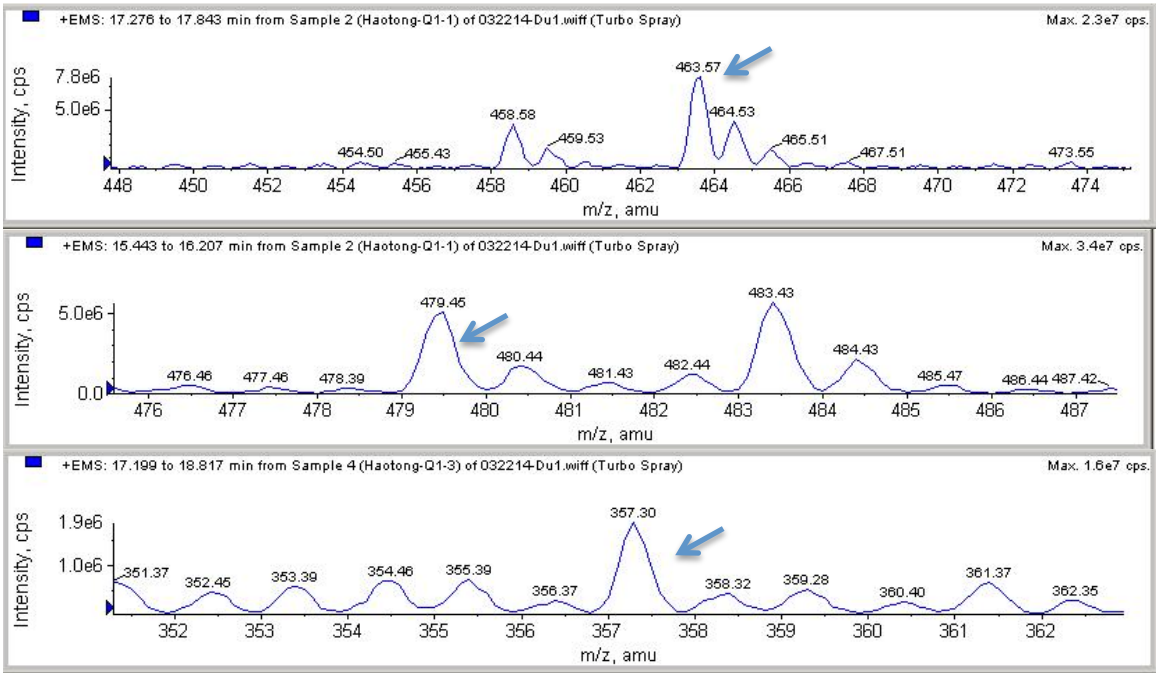


Figure 4-12. LC-MS result of the *in vitro* reactions. The pointed peaks showed the corresponding masses of the predicted possible structures.

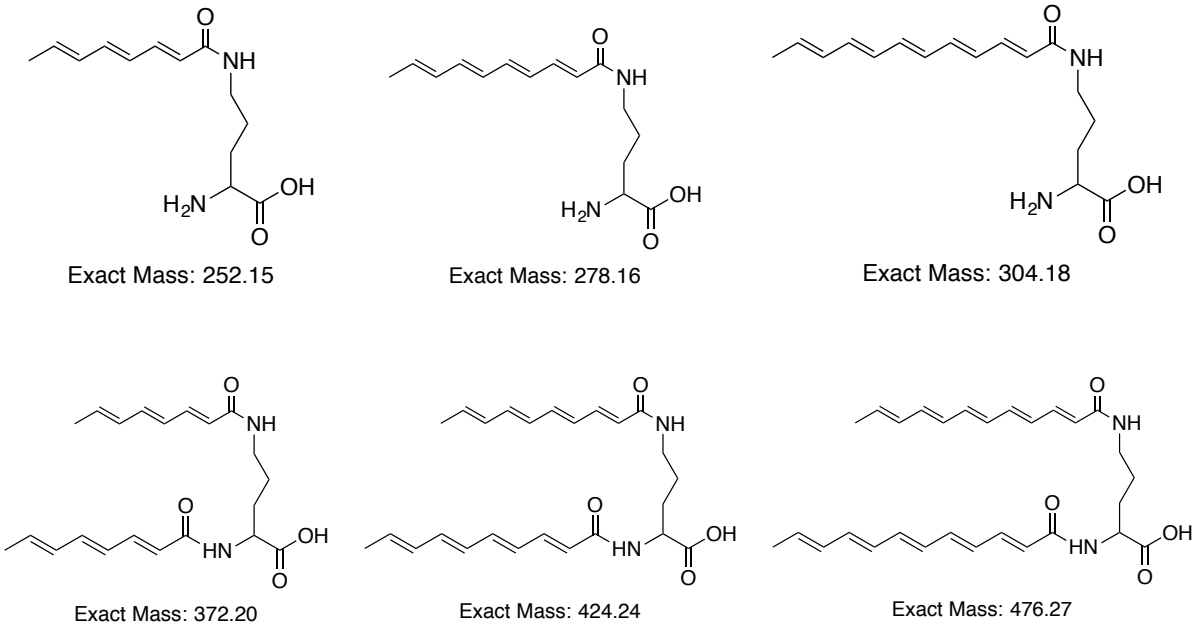


Figure 4-13. Possible structures and expected masses of the compounds with spontaneously formed amide bond.

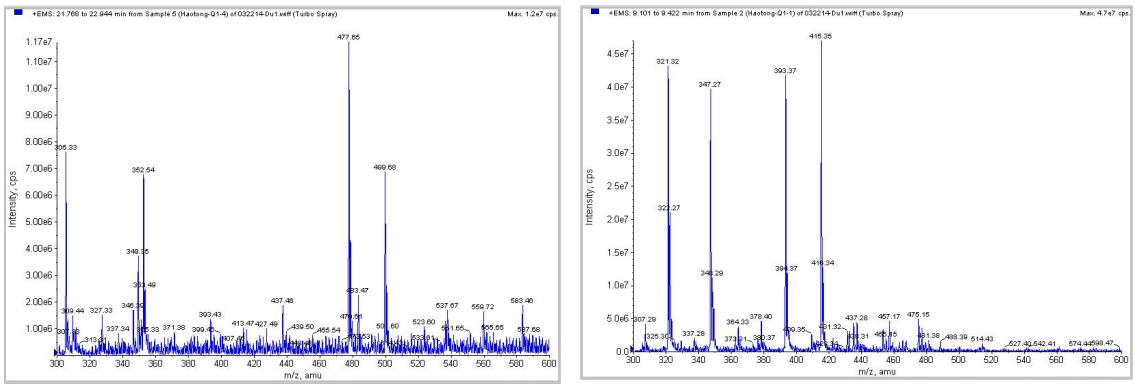


Figure 4-14. Mass spectrum of the detected compounds with automatically formed amide bond.

4-4. Summary and final remarks

The experiments described in this chapter support a new mechanism for the biosynthesis of bacterial hybrid polyketide-peptide, i.e. a single set of domains of an

iterative PKS-NRPS, in cooperating with a cascade of redox enzymes, to synthesize a complex and highly modified polycyclic tetramate macrolactam. OXs in CFE do have activities. They are able to catalyze the ring closure. NRPS also shows a broad specificity toward the substrates with different chain length, and it's able to catalyze the two amide bonds formation and tetramate acid ring formation. The novelty is, very few Diels-Alder-like redox enzymes have been recognized in fungi or bacteria, representing a new way by which nature assembles a complex and distinct molecular scaffold. The mechanistic understanding of the biosynthetic process potentially expand approaches to other bioactive PTM. A big problem here is that it's hard to obtain soluble OX enzymes. Further efforts will be needed to improve the solubility of the proteins. Another problem is that the polyene-tetramate-polyene intermediate is very unstable. Without the ring structure formed by OXs enzymes, this intermediate is difficult to isolate and characterize.

Chapter 5: Fatty Acids Activation and Selection in the Biosynthesis of WAP-8294A, a Group of Anti-MRSA Cyclic Lipodepsipeptides

5-1. Background and Significance

The over-reliance on limited drug targets has led to the rapid emergence of anti-infectious drug resistance. For instance, the recent outbreak of "NDM-1 symptoms" (short for New Delhi Metallo-beta-lactamase) in 2009 is the result of a superbug that are resistant to many antibiotics, including those drugs that are reserved as the last resorts for treatment of infectious disease defense[124]. This is a life-threatening concern because few new antibiotics have been developed in the past four decades. For antibacterial drugs, since the introduction of fluoroquinolones in the 1960s, there was an almost four-decade innovation gap in new antibiotic discovery until linezolid was introduced in 2000 and daptomycin (Cubicin®) in 2003. Serious threats, such as the multi-drug resistant NDM-1, demand new antibacterial agents with novel structures and activities. Especially, nosocomial infections caused by methicillin-resistant *Staphylococcus aureus* (MRSA) in hospitals have become a serious clinical problem[125]. Thus, it is imperative to continue the discovery and development of new antibiotics[126].

Natural products (NPs) are the most prolific source of bioactive compounds, especially anti-infective and anti-cancer agents. Our lab has been using *Lysobacter enzymogenes* strain C3 and OH11 as a new source for bioactive NP discovery[26, 76]. Figure 5-1 are NPs isolated from various *Lysobacter* species. Both lysobactin and tripropeptins exhibit strong anti-MRSA and anti-VRE (vancomycin resistant *Enterococcus*) activities[92], while dihydromaltophilin (HSAF) discussed in the previous chapters exhibits strong inhibitory activities against a wide range of fungi.

WAP- 8294A is a family of cyclic lipodepsipeptides originally isolated from *Lysobacter staphylocidin,* that display anti-MRSA activity[125, 127-129]. This group of compounds exhibits up to 14-fold greater antibacterial activity than vancomycin, appear to exert their cytotoxic effect by selective interactions with phospholipids in the target cell membrane resulting in membrane damage[129]. Among the member of this family of NPs, , WAP-8294A2 had reached Phase-I/II clinical studies.

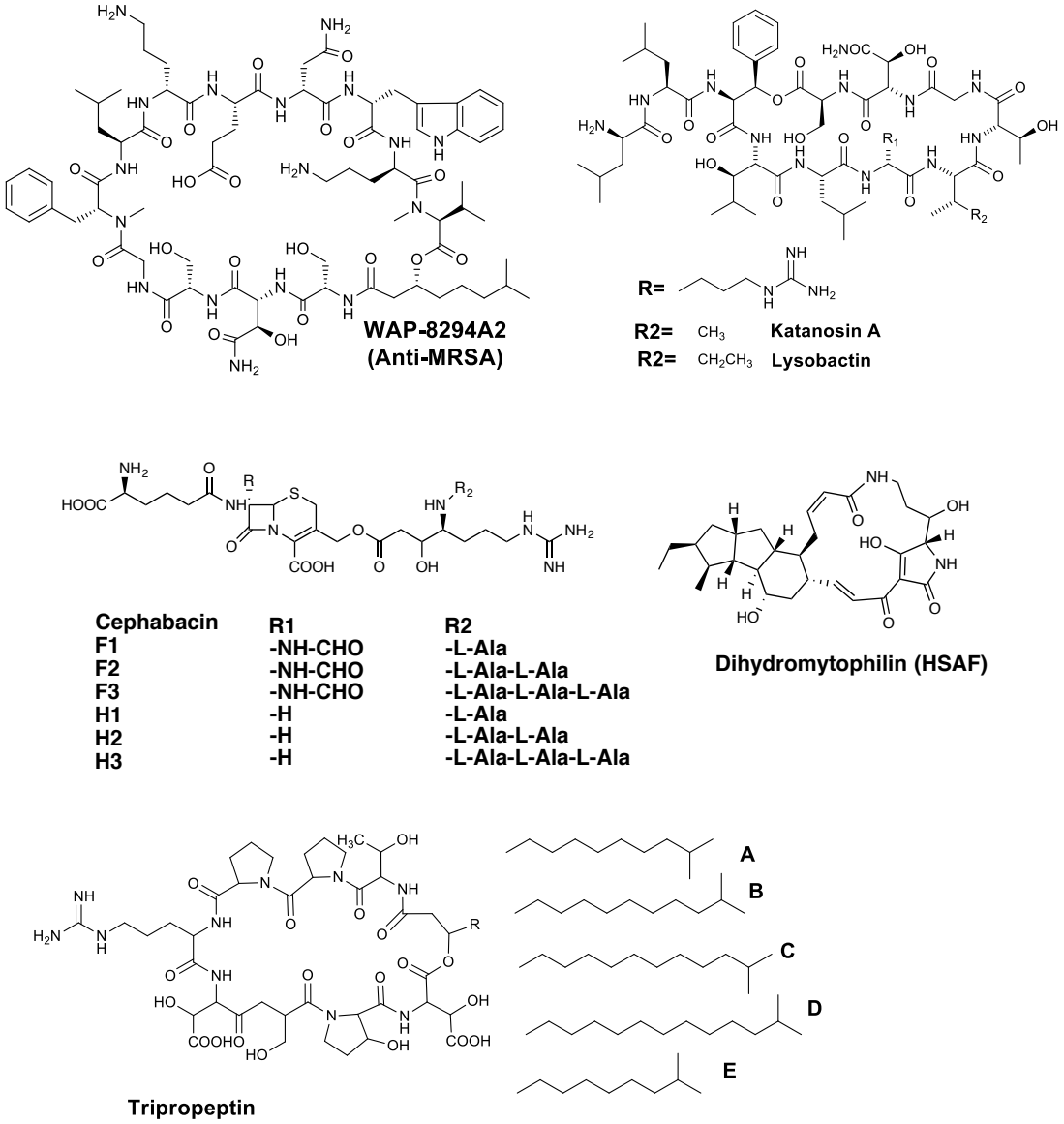
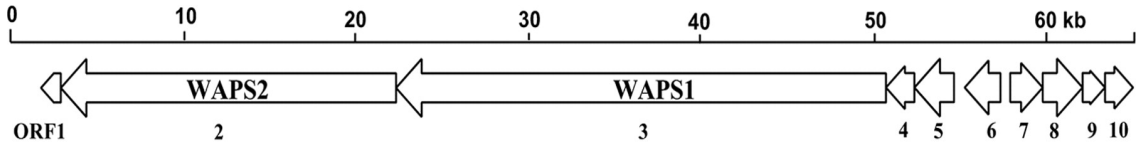


Figure 5-1. Natural products isolated from various *Lysobacter* species[110]. The figure is reproduced with the permission of American Society for Microbiology

We reported the identification and characterization of the WAP- 8294A2 biosynthetic gene cluster from *L. enzymogens* strain OH11[110]. After genome sequencing of *L. enzymogenes* OH11, we found at least nine clusters which contain genes for polyketide synthase and nonribosomal peptide synthetase (PKS-NRPS). Through genome annotation and a series of gene disruption experiments, a total of 45 domains that make up 12 NRPS modules were identified to encode enzymes involved in WAP-8294As core structure biosynthesis (Figure 5-2)[110].



ORF	bp	aa	Blastx Homolog
ORF1	219	72	MbtH-like protein
ORF2	22191	7396	WAPS2 (5 modules) ($C\text{-}A_8\text{-}T\text{-}C\text{-}A_9\text{-}T\text{-}E\text{-}C\text{-}A_{10}\text{-}T\text{-}E\text{-}C\text{-}A_{11}\text{-}T\text{-}E\text{-}C\text{-}A_{12}\text{-}MT\text{-}T\text{-}TE$)
ORF3	27897	9298	WAPS1 (7 modules) ($C\text{-}A_1\text{-}T\text{-}C\text{-}A_2\text{-}T\text{-}E\text{-}C\text{-}A_3\text{-}T\text{-}C\text{-}A_4\text{-}T\text{-}C\text{-}A_5\text{-}MT\text{-}T\text{-}E\text{-}C\text{-}A_6\text{-}T\text{-}C\text{-}A_7\text{-}T\text{-}E$)
ORF4	972	323	Taurine catabolism dioxygenase
ORF5	1821	606	Efflux transporter
ORF6	1698	565	Pectin lyase fold/virulence factor
ORF7	1227	408	Lytic endopeptidase preproenzyme
ORF8	3096	1031	TonB-dependent receptor
ORF9	774	257	SapC-related protein
ORF10	1011	336	Pass1-related protein

Figure 5-2. Putative biosynthetic gene cluster of WAP-8294As[110]. The figure is reproduced with the permission of American Society for Microbiology.

Disruption mutants of the NRPS genes lost the inhibitory activity against *B. subtilis,* suggesting that the NRPS cluster is likely to be a major contributor to the antibacterial activity of *L. enzymogenes*. Total metabolites extracted from the mutants revealed the absence of a major peak, which presents in the wild-type strain; this peak was then

determined to be WAP-8294A2 by mass spectrometry and NMR analysis. The results demonstrated that the NRPS cluster is responsible for the biosynthesis of WAP-8294A2, as well as other members in the WAP family that differ only in the length of the acyl side chain.

Compounds in the WAP-8294 family differ mainly in the nature of the 3-hydroxy fatty acyl moiety, which all compounds in the family possess. In most cases, the free fatty acids are activated as CoA thioesters, necessitating the existence of an enzyme to activate and introduce the 3-hydroxy fatty acid chain into the peptide. Such acyl-CoA ligase enzymes are often located in the vicinity of the NRPS cluster[130]. While the amphi-enterobactin[130], daptomycin[131] and calcium-dependent antibiotic (CDA)[132] gene clusters all contain a dedicated acyl-CoA ligase and an ACP loading domain for the activation and incorporation of the fatty acyl side chain, the surfactin[133] and the WAPS cluster do not contain neither an ACL, nor an ACP domain in front of the WAPS assembly line, suggesting that the WAP-8492A biosynthetic pathway likely recruits an ACL outside the gene cluster (in *trans*). Generally, this type of in *trans* ACLs exhibits a broad specificity and is able to function redundantly[134]. This could be the reason that WAP-8492A compounds share a common amino acid composition and only differ in the length of the fatty acid side chain. These features could turn ACLs into interesting targets for production of novel WAP analogs by genetic engineering[135]. The hypothesized pathway starts with the activation of the 3-hydroxy fatty acid into acyl-*S*-CoA by the action of an in *trans* ACL[110]. A type III C domain is present at the *N*- terminus of the WAPs NRPS, which is a common feature found in NRPSs dedicated to cyclic lipopeptides, such as surfactin, daptomycin and CDA. In the biosynthesis of such lipopeptides, the activated acyl-*S*-CoA can directly serve as substrate to be incorporated into the peptide[136] or, alternatively, can be transferred to a

dedicated acyl carrier protein (ACP) as acyl-*S*-ACP, which is then incorporated into the peptide.

The aim of this work is to determine the activation and mechanism for incorporation of fatty acid side chain in the biosynthesis of WAP-8294A. Here, we show that several ACLs play an important role in the initiation process. We used chemically synthesized fatty acids as substrates, providing evidence both *in vitro* and *in vivo* that the ACL-6 is the dominate enzyme that activates 3-hydroxy-7-methyloctanoic acid and initiates WAP8294A2 formation.

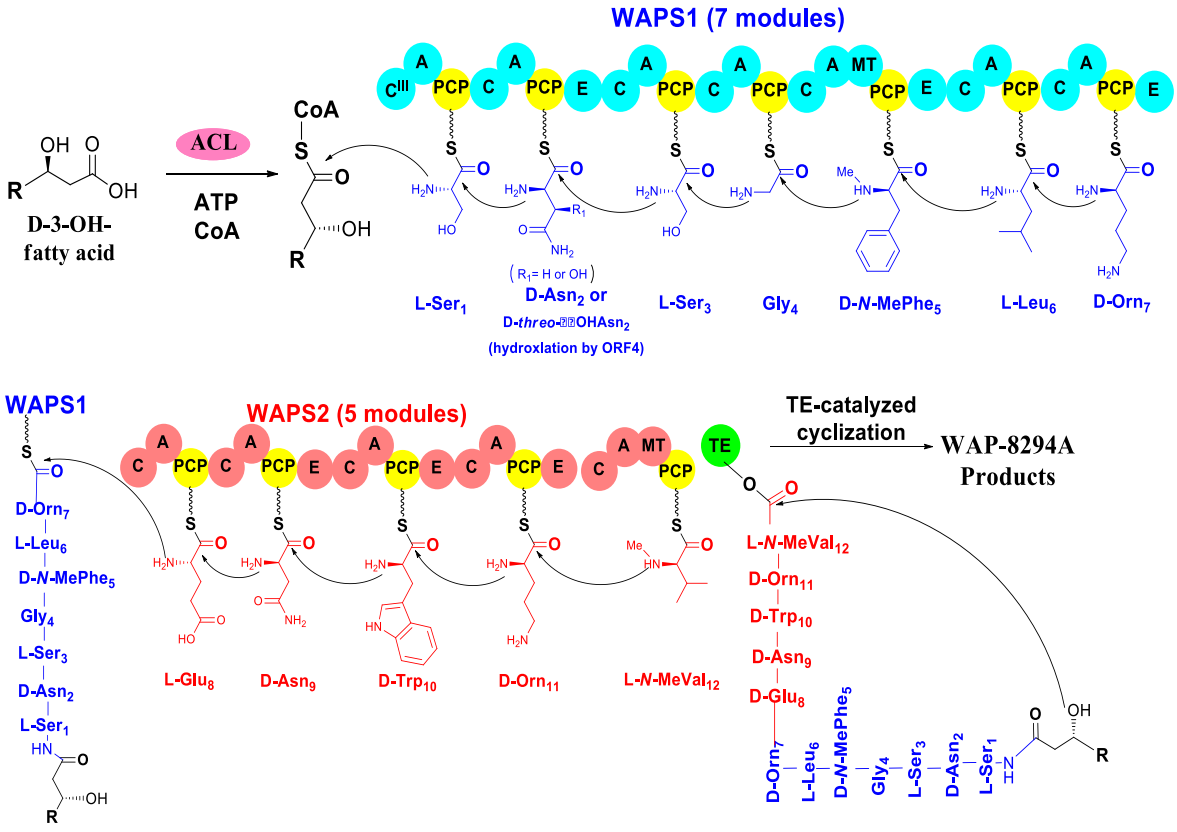


Figure 5-3. Proposed biosynthetic pathway of WAP-8294As[110]. The figure is reproduced with the permission of American Society for Microbiology.

5-2. Materials and methods

5-2-1. Bacterial strains, plasmids, and general DNA manipulations

Chemicals used in this study were purchased from Fisher Scientific or Sigma. Oligonucleotide primers for PCR were synthesized by Fisher. Plasmid preparation and DNA extraction were carried out with Qiagen kits (Valencia, CA), and all other DNA manipulations were carried out according to standard methods. Ni-NTA agarose was purchased from Qiagen (Valencia, CA). *Escherichia coli* strain XL Blue was used as the host for general plasmid DNA propagation, and the cloning vector was pANT841. Vector pET28a was used for protein expression in *E. coli* BL21 DE3. *E. coli* S17-1 was used as the conjugal strain. *Lysobacter enzymogenes* and other bacterial strains were grown in Luria- Bertani (LB) broth medium, 1/10-strength tryptic soy broth (1/10 TSB, Sigma) or NYGB/A medium. EnzChek Pyrophosphate Assay Kit was purchased from Sigma.

5-2-2. Generation of gene disruption mutants

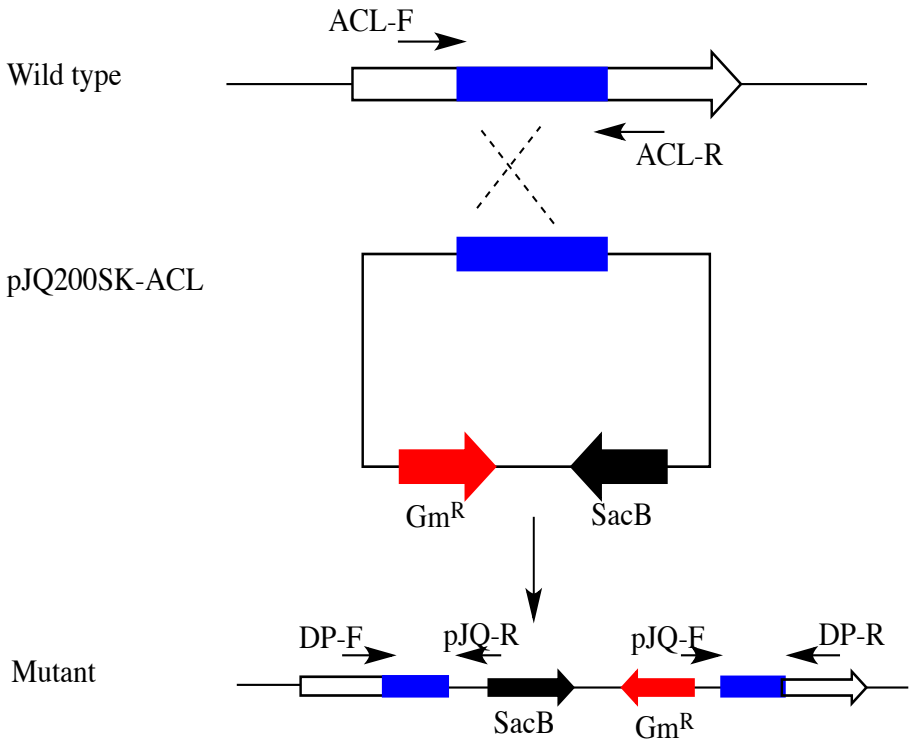


Figure 5-4. Strategy of gene disruption by single crossover leading to plasmid insertion.

To construct the plasmid vectors for gene disruption, an internal fragment was amplified from each of the ORFs using the primer pairs described below. Genomic DNA from the wild type *Lysobacter enzymogenes* OH11 served as the PCR template. The homologous regions of these ACLs were amplified by PCR and introduced into the conjugation vector pJQ200SK to produce seven versions of pJQ200SK-ACLs. Each of the pJQ200SK-ACL constructs was transformed into *E. coli* S17-1, which was mated with *L. enzymogenes* OH11 for conjugal transfer of the vectors. The positive colonies grown on LB plates containing gentamicin (20 μg/ml) were picked up and inoculated into liquid cultures containing gentamicin. Genomic DNA was prepared from each of the cultures, and diagnostic PCR were performed to identify mutants that resulted from a homologous recombination. To screen the ACL gene disruption mutants, PCR were performed using the diagnostic primers listed in Table 5-1, which would amplify correct fragments when mutants resulted from a homologous recombination, but not from the wild type or mutants resulted from a random insertion of the construct (Figure 5-4).

Table 5-1. Primers used in this paper

Primers for 7 ACL genes disruption			
Name	No.	Fragment Length	Primers
ACL-1	LysEGL003270	496bp	Forward: TTACTCGAGTCAATCCCAAGCTCAAGC Reverse: TATGGATCCTAGTTGTGCAGCACCAGC

ACL-2	LysEGL000201	488bp	Forward: TAACTCGAGATTCCAGCCACGAA
			Reverse: TATGGATCCATGCTTGAGGATGCGTTC
ACL-3	LysEGL000048	579bp	Forward: TAACTCGAGAGCTCGAACTCAAGAAGG
			Reverse: AATGGATCCTGACGATGATTTCCTGGC
ACL-4	LysEGL000406	555bp	Forward: ATCCTCGAGTTCTTCCTGCTGATGTTC
			Reverse: TATGGATCCTAGCGCATCGTGCAACT
ACL-5	LysEGL003038	604bp	Forward: TATCTCGAGTCAACACCAATCCGATGT
			Reverse: TAG GGATCCTTTTCCTGCAGTTCCTTC
ACL-6	LysEGL001572	630bp	Forward: TAACTCGAGCACTCCGAAAACCATCTG
			Reverse: AATGGATCCTCGTCGTGGAAGATCG
ACL-7	LysEGL003969	700bp	Forward: TAACTCGAGGATCTCGCATCGCAAC
			Reverse: TATGGATCCGGTCGTGAGGTAATAGC

Primers for 7 ACL genes expression

Name	Fragment Length	Protein size (Da)	Primers
ACL-1	1677bp	58677.4	Forward: ATAGAATTCAGCCCGACCGTCAGCC
			Reverse: AGGAAGCTTAGCACAACAGGACCC
ACL-2	1635bp	58767.3	Forward: AACGAATTCATGCTCTCCCGACTGA
			Reverse: TATCTCGAGCGGAGCGAAAGGGAAG
ACL-3	1737bp	60454.5	Forward: TTAGAATTCTGGGCGCTCGTATCC
			Reverse: ATACTCGAGCGCTTTCACCGCAATG
ACL-4	1695bp	59778.6	Forward: TATGAATTCCCACCCGCTCGGCTGA
			Reverse: TAAAAGCTTCAGTGCCGCCTCTCAA
ACL-5	1677bp	61084.6	Forward: TATGAATTCAGTTTGAACCGTCCGT
			Reverse: TATAAGCTTACCCATGCTCACGCGT
ACL-6	1395bp	47157.1	Forward: TAAGAATTCCAACAACGAATCGCTG
			Reverse: ATGAAGCTTATGACCCGGGATTTAT
ACL-7	1713bp	58643	Forward:TAAGAATTCTCAAGGACCGATCGAA

Reverse: ATTAAGCTTCCATCTGCAGGGTCAT

Primers for construction of ACL-6 deletion mutant			
ACL-6	upstream	412bp	ATC GTC GAC CTT CTT CCT GCA A
			CTC GGATCC TTC GTT GTT GAG G
ACL-6	downstream	280bp	CTA GGATCC CCT GAA GCA GAT C
			GCG TCT AGA TGA GCG GTT CCA T

Diagnostic Primers		
No.	Forw-diagnostic primer sequence	Expected length
ACL-1	TGTACCACGACAAATCCATCC	632bp
ACL-2	TGATGATCGTGCACACCTCC	903bp
ACL-3	AGATAGATCCCGACGCGCTG	796bp
ACL-4	TGTGCAACATCGCCGC	840bp
ACL-5	AGCTCAAGCTCAAGAAGGGC	770bp
ACL-6	ATCGACATGCTGGCGTTG	906bp
ACL-7	TGTTCGAAGCCAAGGCC	903bp
ACL-6 deletion	GAA TAC CTG GCC CTG CAA	690bp
Primer on vector	ACCATGATTACGCCAAGC	

5-2-3. Production and analysis of the metabolites in mutants

L. enzymogenes OH11 and its ACL mutants were grown in 1/10 TSB for 1 day, and an aliquot of 200 µl was spread on the solid NYGA fermentation medium (Bacteriological Peptone, 5 g/L; yeast extract, 3 g/L; glycerol, 20 g/L, and agar 15 g/L). The plates were incubated at 28 °C for 2 days. To extract the metabolites, the solid culture broth was collected and extracted with methanol. The organic phase was dried with a rotavapor (Buchi, Rotavapor R-200) to afford the crude extract. The extract was dissolved in 2 ml methanol containing 0.05% TFA. A 20 µl aliquot of each extracts was analyzed by HPLC (1220 Infinity LC, Agilent Technologies) using

a reversed-phase column (Cosmosil 5C18-AR-II, 4.6 ID × 250 mm). Water/0.01M TFA (solvent A) and acetonitrile/methanol = 1:1 (solvent B) were used as the mobile phases with a flow rate of 1.0 mL/min. The HPLC program was as follows: 57% B in A in the first 5 min, 57–100% B in 5–32 min, 100% B in 32–40 min, back to 57% B at 41 min, and maintained to 48 min. The metabolites were detected at 280 nm on a UV detector. LCQ-MS was used to verify the mass of the peak of WAP-8294A2 and analogues.

5-2-4. Generation of gene deletion mutants

To construct the ACL-6 in-frame deletion vector, primer extension PCR reactions were performed to generate a 705-bp fragment of ACL-6 upstream and downstream regions. This deleted the whole ACL-6 gene. The primers used for these experiments are listed in Table 5-1. The PCR amplified fragment was then cloned into the conjugal vector pJQ200SK as an *Xho*I/*Bam*HI fragment to produce pJQ200SK-ACL-6. To conjugally transfer pJQ200SK-ACL-6 into *L. enzymogenes* strain OH11, the procedure was identical to the previously described. The correct single crossover mutant was confirmed by diagnostic PCR. The single colonies identified from single crossover were cultured in the presence of 5% (W/V) sucrose to select for the loss of the vector through a second homologous recombination. The single crossover mutants were grown in 1/10 tryptic soy broth (TSB) medium for 14 hours. Then the individual cultures were re-inoculated into 1/10 TSB (with 1:100 ratio) containing 5% sucrose and 25 μg/ml kanamycin medium. Aliquots (50 μl) were taken out every 3 hours from this liquid culture and spread onto 1/10 TSA plates (5% sucrose, 25 μg/ml kanamycin). Single colonies were picked up from the plates and continued to grow in 1/10 TSB (5% sucrose, 25 μg/ml kanamycin) liquid medium for another 2-5 days. To

confirm the double crossover mutant, the diagnostic PCR was performed using genomic DNA extracted from the single colonies as template.

5-2-5. Construction of ACL-6 overexpression and complementation mutants

The coding region of ACL-6 was amplified, using primers ACL-6-up and ACL-6-down (Table 5-1), from the genomic DNA of the wild type strain. The PCR product was digested with *Bam*HI/*Xho*I and cloned into the same sites of the vectors pHmgA to generate pHmgA-ACL-6[137]. This vector included one homologous region of *hmgA* gene, and the plasmid containing the ACL-6 gene is therefore expected to integrate into the *hmgA* gene in OH11 genome by homologous recombination. The *hmgA* gene is predicted to encode a homogentisate 1,2-dioxygenase, which catalyzes the oxidative cleavage of the aromatic ring of tyrosine/phenylalanine, which is a key step in aromatic amino acid degradation pathway. The disruption of *hmgA* gene blocks the oxidative cleavage reaction and leads to the accumulation of homogentisate, exhibiting black color in the mutant organisms and allowing for easy selection of the single-crossover mutants into whose genomes the construct has been integrated[141, 142]. In each the constructs, the coding region of ACL-6 was placed downstream from the respective promoter. The constructs were validated and transferred into *L. enzymogenes* OH11 and ACL-6 deletion mutant by conjugation. The black colored colonies were selected and verified by diagnostic PCR[138, 139].

5-2-6. Supplementation of ACL-6 mutant with 3-hydroxy-7-methyloctanoic acyl SNAC

ACL-6 mutant was grown in 1/10 TSB for 1 day, and an aliquot of 1 mL was transferred to each of two 1 L NYGB liquid cultures. The culture was fed with the fatty acyl-SNAC dissolved in DMSO under pulse feeding on 2 following days (0.12

mmol/L and 0.06 mmol/L)[140]. Metabolites were extracted from the supernatant with ethyl acetate: methanol: acetic acid = 80: 15: 5. The organic phase was dried with a rotavapor (Buchi, Rotavapor R-200) to afford the crude extract. The extract was dissolved in 2 ml methanol containing 0.05% TFA. A 20 μl aliquot of each extracts was analyzed by HPLC and MS as described above.

5-2-7. Expression of ACL1-7 genes and purification of the proteins

To construct the expression vectors for the ACLs, genes fragments from all seven ACL genes were amplified by PCR using *Taq* DNA polymerase with genomic DNA extracted from *L. enzymogenes* OH11 as template. The primers are listed in the table 5-1. The PCR fragment was digested and cloned into pANT841 at the same sites. The constructs were sequenced, and the result showed that all the ACL sequences were correct. The ACLs were released from pANT841 and cloned into expression vector pET28a. To express the ACLs, the pET28a construct was introduced into *E. coli* BL21 (DE3). Single colonies were inoculated in 3 ml of liquid LB medium containing kanamycin (50 μg/ml) and incubated in a shaker (250 rpm) at 37 °C overnight. The overnight culture was added to 50 ml fresh LB medium and incubated in a shaker (250 rpm) at 37 °C until the cell density (OD_{600}) reached 0.6. To induce the expression of the ACLs, IPTG (0.1 mM) was added to the culture, and the cells were allowed to grow at the same conditions for another 3 h. To prepare the soluble fraction, the cells were harvested and resuspended in 2 ml of Tris-HCl buffer (250 mM NaCl, 100 mM LiCl, 50mM Tris). The cell suspension was supplemented with lysozyme (1 mg/ml) and sonicated five times on ice. The soluble fraction of protein extracts was loaded onto a Ni-NTA column, which was previously calibrated with Tris buffer containing 10 mM imidazole. The column was washed three times with

the buffer containing 20 mM imidazole, and the His_6-tagged protein was eluted twice with 200 μl of Tris buffer containing 250 mM imidazole (Figure 5-11). The fractions containing pure ACLs were collected and dialyzed against Tris buffer containing 15 % glycerol.

5-2-8. *In vitro* assay of acyl-CoA ligase activity

Enzymatic reactions to generate fatty acyl-CoA were set up using the following conditions: 10 μM ACLs, 0.2 mM CoA, 10 mM ATP, 20 μM fatty acid in DMSO, 5 mM $MgCl_2$, 0.2% (w/v) Triton, 200 μM DTT, and 100 mM Tris-HCl pH 7.8. All reactions were incubated at 30 °C overnight. Samples were analyzed by HPLC (1220 Infinity LC, Agilent Technologies) using a reversed-phase column (Phenomen-ex, 4.6 ID × 150 mm). Water/25 mM ammonium acetate/0.5% acetic acid (solvent A) and acetonitrile/0.5% acetic acid (solvent B) were used as the mobile phases with a flow rate of 1.0 mL/min. The HPLC program was as follows: 10% B in A in the first 3 min, 10-90% B in 3–15 min, 90% B in 15-25 min, back to 10% B at 26 min, and maintained to 30 min[141]. The metabolites were detected at 260 nm on a UV detector. MS (Finnigan mat, LCQ) was used to verify the mass of the peak of products. To further elucidate the structure of the product, Agilent LC-1200 (Santa Clara, CA) connected to a 2.1× 100 mm Symmetry ODS column from Waters (Milford, MA) and a Triple Quadrupole Mass Spectrometer model 4000 QTrap from ABSciex (Framingham, MA) operating in either single quadrupole (Q1), enhanced mass spectrum (EMS), MS/MS or multiple reaction monitoring (MRM) modes were used. The samples were injected onto the column and a gradient from 98% mobile phase A (0.1% formic acid in water, J.T. Baker) to 60% B (0.1% formic acid in

acetonitrile, Acros Organics) was run over 15 minutes, followed by 5 minutes of 98% B and 5 min of 98 % A, all at a flow rate of 0.25 mL/min.

5-2-9. Kinetic analysis of acyl-CoA ligase activity

The EnzChek Pyrophosphate Assay Kit was used to measure the enzymatic activities of ACLs.

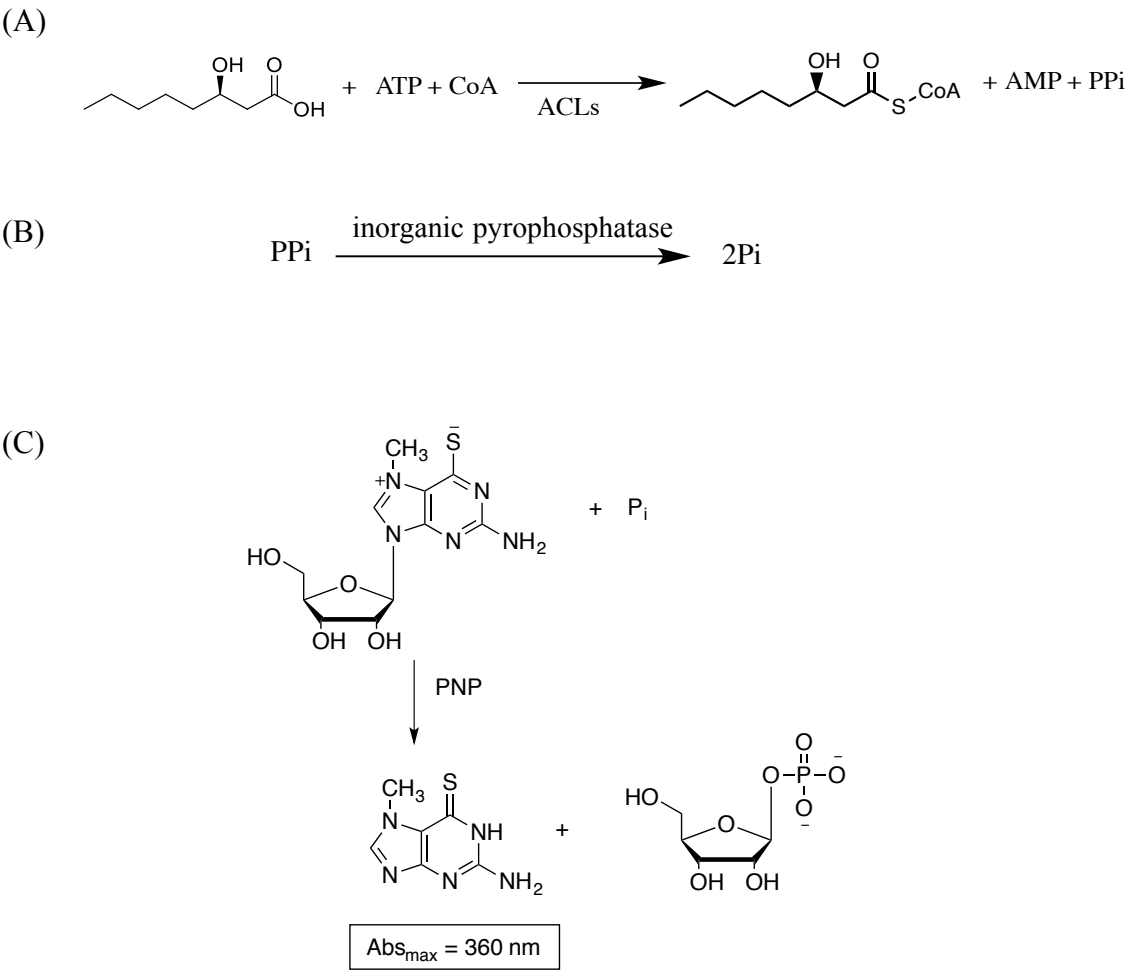


Figure 5-5. Enzymatic conversion of A: fatty acids to fatty acyl-*S*-CoA by ACLs, releasing a pyrophosphate group. B: A pyrophosphate group to 2 phosphate groups via the hydrolysis of inorganic pyrophosphatase. C: 2-amino-6-mercapto-7-methyl-purine ribonucleoside (MESG) to ribose 1-phosphate and 2-amino-6-mercapto-7-methylpurine by purine ribonucleoside phosphorylase (PNP).

A standard curve for the pyrophosphate assay was generated using the pyrophosphate standard as a source of PPi. The linear range of the assay for PPi extends from 1 μM to about 75 μM. The standard 1 mL reaction mixture contains: 730 μL - x μL dH_2O (to create a final volume of 1.0 mL), 50 μL 20 × reaction buffer, 200 μL MESG substrate solution, various μL-volumes of sample to be analyzed for pyrophosphate content, 10 μL purine nucleoside phosphorylase (1 U), 10 μL inorganic pyrophosphatase (0.03 U). In this reaction, the analytical sample volumes were 0, 10, 20, 40, 80, 120 and 150 μL, respectively, giving concentrations of 0, 5, 10, 20, 40, 60 and 75 μM, respectively. The reaction solutions were mixed well and incubated for 30-60 minutes at 22 °C. The absorbance at 360 nm was recorded over this time period. The background absorbance determined for the no-pyrophosphate control was subtracted from each sample before plotting the curve. A standard curve ($y = 0.0156x + 0.0004$, $R^2 = 0.99603$) for the pyrophosphate assay was generated using the pyrophosphate standard as a source of PPi.

The proper ACL 1-7 protein concentrations in the reaction were determined spectrophotometrically by varying protein concentration between 1 μM to 15 μM.

Initial kinetic parameters (k_{cat} and K_M) were obtained by saturating with two substrates and varying the concentration of a third substrate[142]. For fatty acid kinetic constant determination, both CoA (400 μM) and ATP (400 μM) were kept at fixed saturating concentrations, and fatty acid (1 μM–2 mM) was kept at variable concentrations. Individual substrate saturation kinetic data were fitted to Equation 1,

$$v = \frac{V_{\max}[S]}{K_m + [S]} \qquad \text{（Equation 1）}$$

where v is the reaction rate, V_{max} is the maximal velocity, [S] is the substrate concentration, and K_M is the Michaelis-Menten constant.

To measure the activity of acyl-CoA ligase, the following reagents were combined in 0.1 mL reaction volumes: 0.4 mM ATP, 0.4 mM CoA, 0.4 mM free fatty acids in DMSO, 0.2 mM MESG, 1 U purine nucleoside phosphorylase, 0.01 U pyrophosphatase, 1× reaction buffer and various ACLs with final concentration of 15 μM. After incubating at 22°C for 30 minutes, the absorbance at 360 nm was measured and corrected for absorbance at 360 nm of a control reaction which the ACL protein was replaced by the purified protein extracted from the pET28a blank vector hosted *E. coli* cell. The substrates being tested are chemically synthesized 3-hydroxyl-7-methyloctanoic acid (which corresponded to WAP-8294A2) and 3-hydroxyl-octanoic acid (which corresponded to WAP-8294A1). The kinetics of ACL 1-7 were determined by initial velocity experiments using the coupled assay. These data were globally fitted to Michaelis-Menten Equation, yielding *Km* values for the substrates.

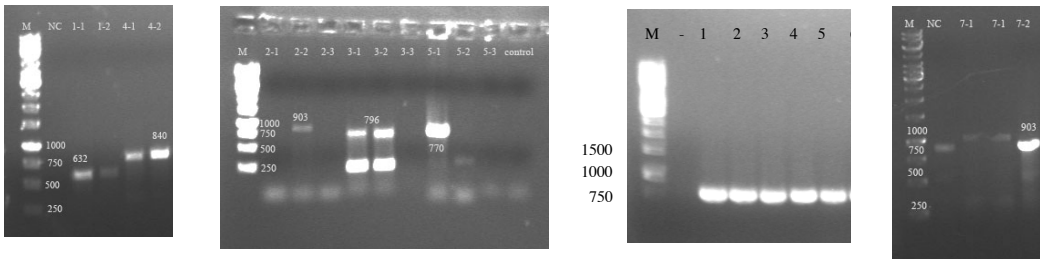
5-3. Results and discussion

In the previous study, an ACL-coding gene (accession number JN596953) present in one of the uncharacterized PKS/NRPS clusters within the genome of *L.e.* OH11 was disrupted to test if the WAP assembly could use an ACL in other metabolic pathways[110]. This mutant showed a partial loss of the antibacterial activity and produced approximately 50% of the WAP-8492A yield of the wild type. This result suggests that the WAP-8492A biosynthetic pathway is likely to recruit an ACL from other pathways. This type of in *trans* ACLs generally exhibit a broad specificity and

functional redundancy, which is consistent with the fact that all WAP-8492A compounds have the same amino acid composition but vary in the length of the fatty acid side chain.

5-3-1. Generation of gene disruption mutants

We have identified 14 putative ACL genes in the genome, among which 7 are associated with secondary metabolism. To test if the WAP assembly could use an ACL in other metabolic pathways, we disrupted the seven ACL encoding ORFs. The HPLC and MS analysis showed that the individual disruptions of the seven ACLs failed to eliminate the production of WAPs. Although the mutants showed varied yields of WAPs, the ACL-6 disruption mutant gave the lowest yield (Figure 5-7), making this putative ACL the best candidate for WAP's "normal" ACL. In our following experiments we focused our attention on this putative ACL enzyme.



(a) (b) (c) (d)

Figure 5-6. ACL1-6 disruption mutants diagnostic PCR:

a: Diagnostic PCR result of ACL-1 and ACL-4 mutants.

b: Diagnostic PCR result of ACL-2, ACL-3, and ACL-5 mutants.

c: Diagnostic PCR result of ACL-6 mutants.

d: Diagnostic PCR result of ACL-7 mutants.

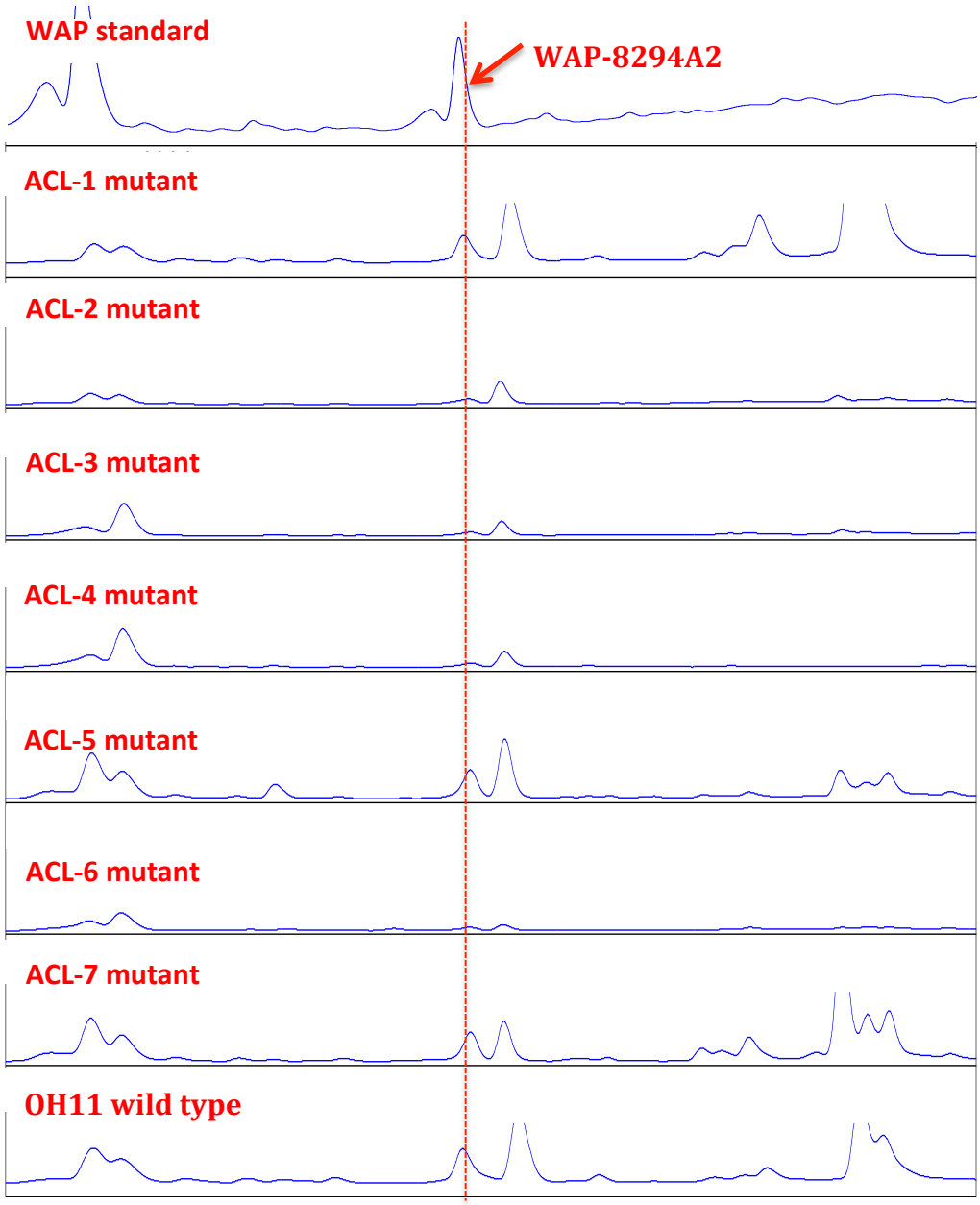


Figure 5-7. WAP standard and metabolite profiles of *Lysobacter enzymogenes* OH11 wild type, ΔACL 1-7: ACL disruption mutants 1-7.

5-3-2. Synthesis of 3-hydroxy fatty acids and 3-hydroxy fatty acyl SNAC

(note: I acknowledge Andrew Olson for preparing the 3-hydroxy fatty acids and 3-hydroxy fatty acyl SNAC)

To generate the initial carbon backbone, we generated the dianion of methyl acetoacetate with LDA; subsequent alkylation with the aliphatic halide provide the initial carbon skeleton in roughly 20-50% yields (1a-c)[143]. A subsequent asymmetric reduction of the β-keto ester with borane dimethyl sulfide complex in the presence of CBS-oxazaborolidine provides the β-hydroxy ester in 20-40% yield (2a-c) with a 90-92% enantiomeric excess of *R* isomer[144]. Hydrolysis to the carboxylic acid from the methyl ester was accomplished in almost quantitative yield (3a-c) by the addition of LiOH. Selective formation of a thioester in the presence of the free alcohol is accomplished through carbodiimide-mediated coupling with *N*-acetylcysteamine (SNAC) by way of DCC with catalytic amount of DMAP to provide the 3-hydroxy thioester in 70-80% yields (4a-c).

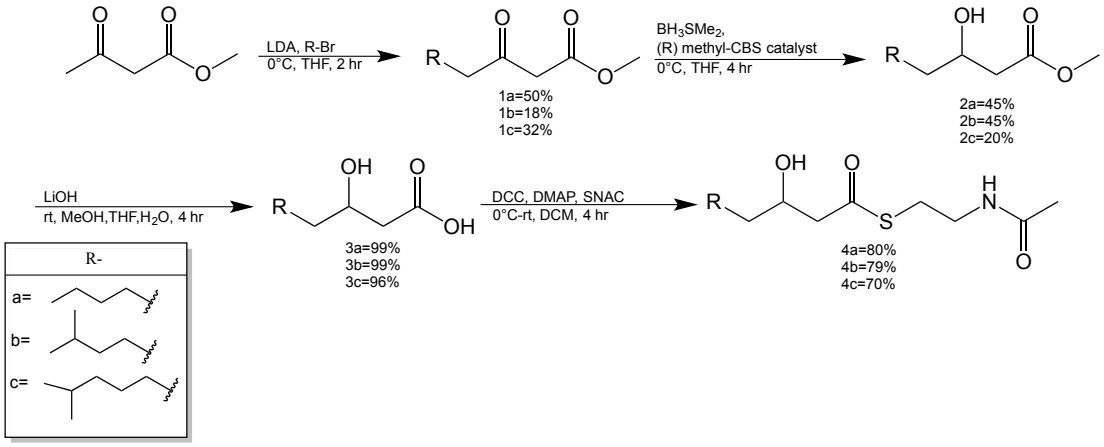


Figure 5-8. Chemical synthesis of 3-hydroxy fatty acids and 3-hydroxy fatty acyl-SNAC.

5-3-3. Construction of ACL-6 in-frame deletion mutants and complementation by feeding experiments

To eliminate any possibility of polar effects caused by insertion of the conjugal vector into the ACL gene, we also generated in-frame deletion mutants for ACL-6 gene (Figure 5-9). WAP production of this mutant strain was monitored by HPLC-MS. Similar to the ACL-6 disruption mutant, the WAP A2 production in the in-frame deletion mutant was diminished to less than 10% compared with the wild type, thus confirming the gene disruption results shown above.

Next, we attempted to chemically restore WAP-8294A2 biosynthesis in the ACL-6 mutant by adding the CoA surrogate, the *N*-acetylcysteamine thioester (SNAC) of 3-hydroxy-7-methyloctanoic acid. After feeding the synthetic 3-hydroxy 7-methyloctanoic acyl-SNAC into the culture by pulse feeding on two following days, the WAP production was successfully restored (Figure 5-10). This result clearly showed that ACL-6 is involved in starter unit activation, but does not participate in NRPS priming; the NRPS could probably be primed by any acyl thioester carrying the 3-hydroxy functional group. Interestingly, an ACL-6 overexpression mutant didn't increase the yield to any significant extent compared with the wild type.

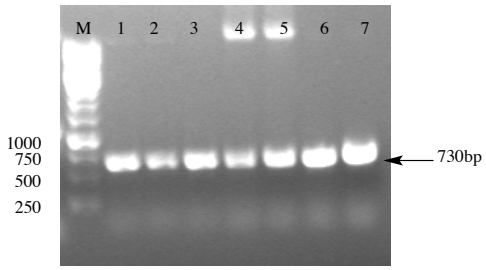


Figure 5-9. Diagnostic PCR result of ACL-6 double-crossover mutants (ACL-6 deletion).

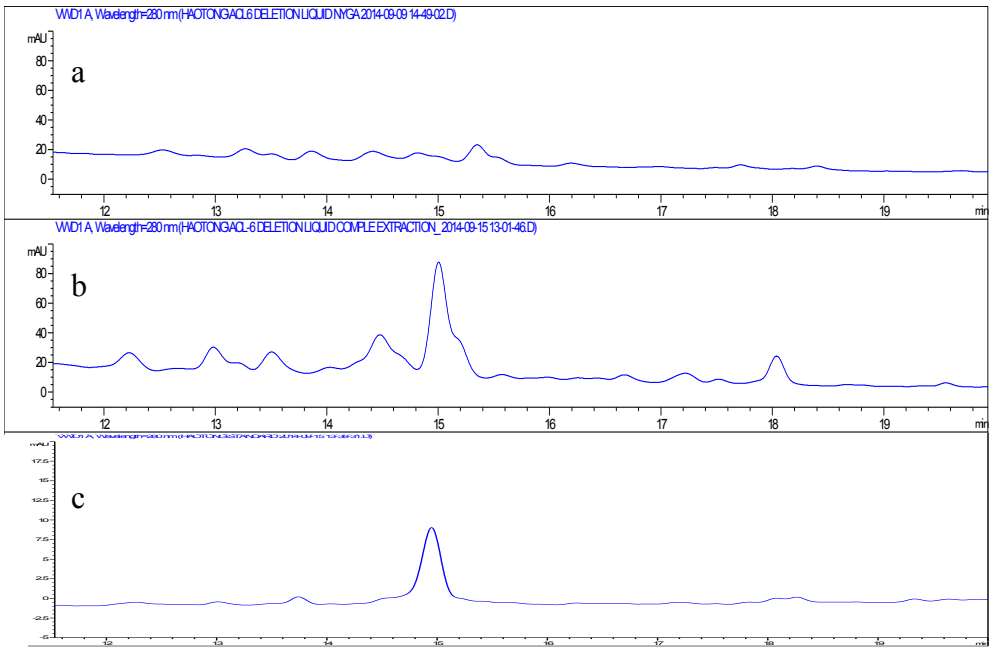


Figure 5-10. Metabolic profiles of a) *Lysobacter enzymogenes* OH11 ACL-6 in-frame deletion mutant (ΔACL-6), b) ΔACL-6 mutant supplemented with 3-hydroxy-7-methyloctanoic acyl-SNAC, c) WAP-8294 A2 standard.

5-3-4. ACLs expression and purification

To identify the relevant ACL gene(s) and reveal the molecular basis for acyl chain diversity, we also expressed the seven putative ACL genes heterologously and tested the substrate specificity of the enzymes. Seven ACL genes were amplified and cloned into expression vector pET28a. After DNA verification by sequencing, heterologous expression in *E. coli* BL21 yielded satisfactory quantities of soluble ACL 1-7 proteins. As judged by SDS-PAGE, the observed ACL 1-7 molecular masses were consistent with the predicted molecular weights (58.7, 58.8, 60.4, 59.8, 61.0, 47.2, 58.6 kDa) calculated from the amino acid sequence, including the His_6tag (Figure 5-11).

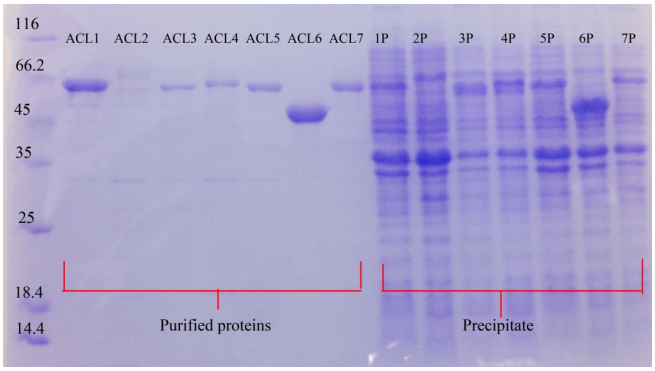


Figure 5-11. Purified ACL1-7 proteins with the expected size shown on SDS-PAGE. 1P-7P: precipitate portion of ACL 1-7 protein extracts.

5-3-5. *In vitro* assay of acyl-CoA ligase activity

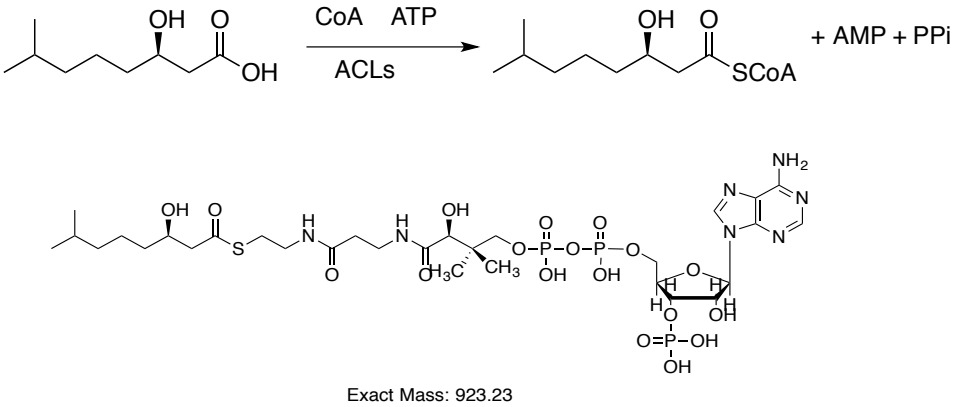
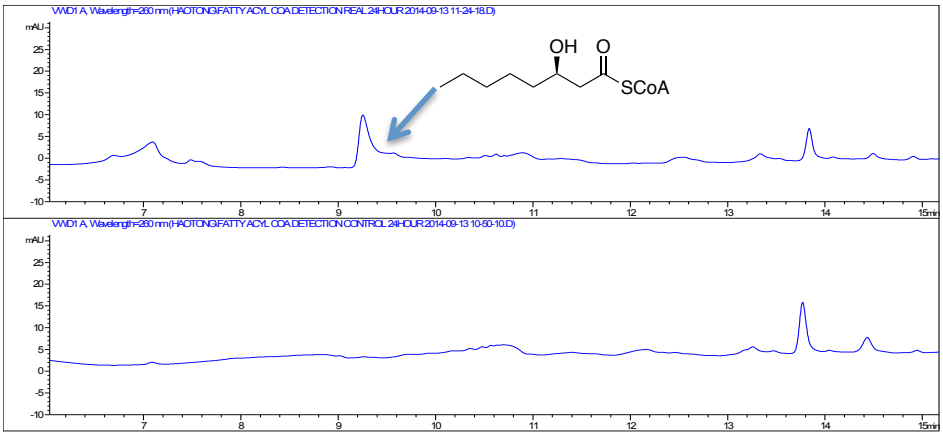


Figure 5-12. Reaction scheme of the activation of 3-hydroxy-7-methyloctanoic acid by ACLs and the structure of 3-hydroxy-7-methyloctanoic acyl-CoA.

We incubated ACL1-7 with 3-hydroxy fatty acids, ATP, and CoA as described above. The assays were analyzed by HPLC. A distinct peak was found around 9 min in the

ACL-6 reaction system. This peak was collected and subjected to MS, giving a mass identical to that of 3-hydroxy-7-methyloctanoic acyl-CoA (*m*/*z* =924.23) (Figure 5-14). Tandem mass spectrometer fragmented the compound into product ions, and both the precursor ion (m/z = 924.90) and product ion (m/z = 417.20) were detected, while the control reaction (w/o enzyme) only exhibited the precursor and product ions of CoA (*m*/*z*=768.12 and *m*/*z* =261.20) (Figure 15-18). The result shows that ACL-6 catalyzes the thioester formation with CoA. ACL-3 and ACL-5 also showed some activity, but the product was so minute that the peak could not be detected by HPLC. The target mass could only be detected by MS, suggesting that the substrate selectivity and activity of these ACLs *in vitro* are weak toward the 3-hydroxy acyl thioester substrate. Additionally, activity of enzymes ACL1-7 against 3-hydroxy-octanoic acid, which is supposed to be the fatty acid substrate of WAP-8294A1, was tested. HPLC and MS data showed that the fatty acid substrate of WAP-8294A1 is favored by ACL-5 instead of ACL-6, indicating that ACL-6 has some specificity toward 3-hydroxy 7-methyloctanoic acid (Figure 5-13).

Both *in vitro* and *in vivo* assay verify that ACL-6 is the most likely enzyme that activates 3-hydroxy 7-methyloctanoic acid into WAP assembly line.



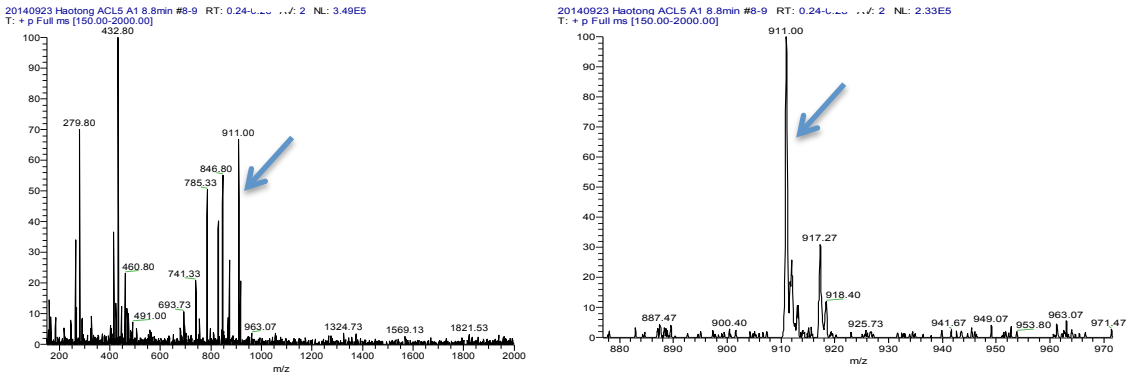
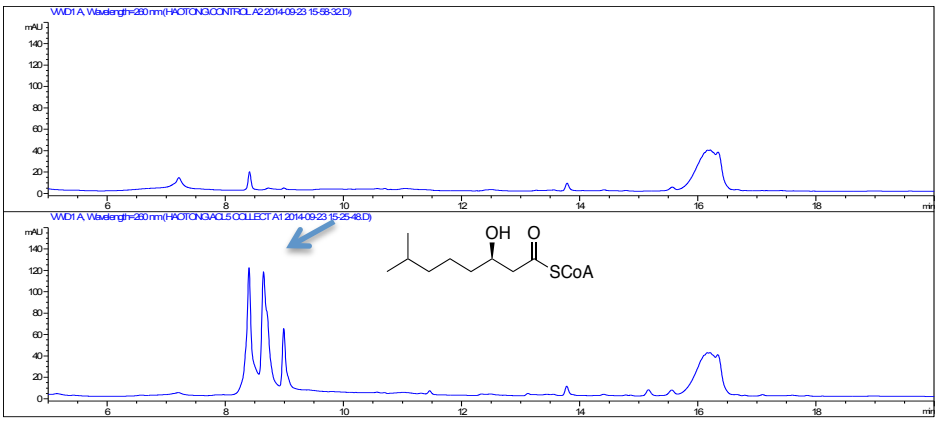


Figure 5-13. HPLC and MS analysis of the activation of 3-hydroxyl-octanoic acid by ACL-5 *in vitro*. The ESI-MS signals were measured in positive single ion mode.



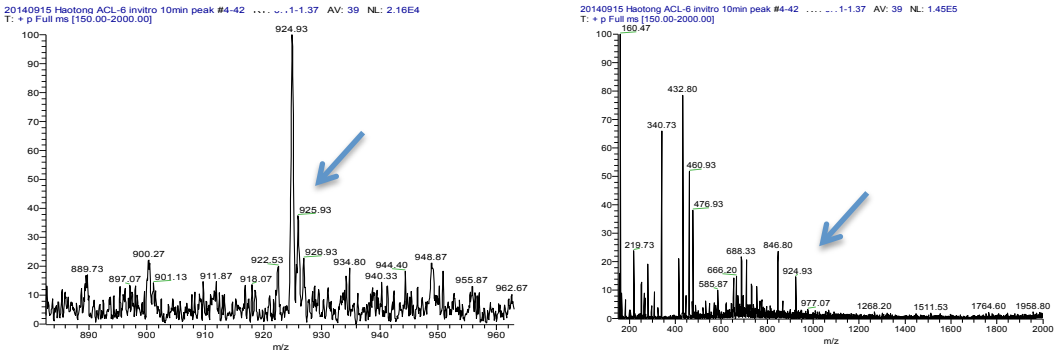


Figure 5-14. HPLC and MS analysis of the activation of 3-hydroxyl-7-methyl acid by ACL6 *in vitro*. The ESI-MS signals were measured in positive single ion mode.

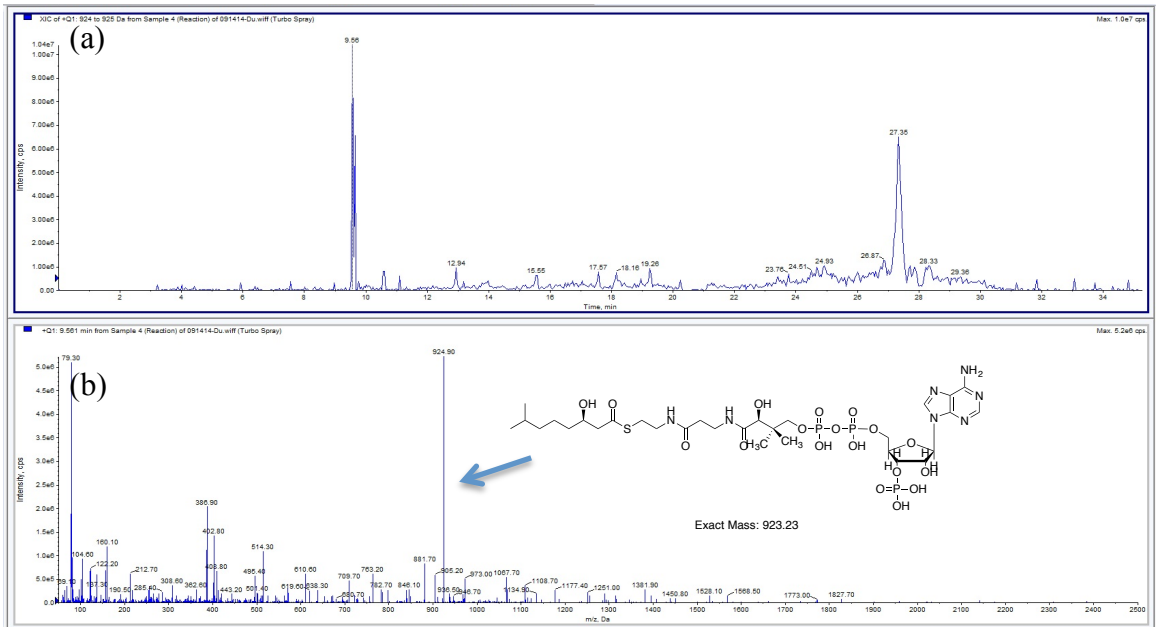


Figure 5-15. LC-MS (+) of the enzymatic reaction. a: Extracted Ion Chromatogram (XIC) of the enzymatic reaction (3-hydroxyl-7-methyloctanoic acyl-CoA with m/z=923.23). b: Mass Spectrum of the selected product peak.

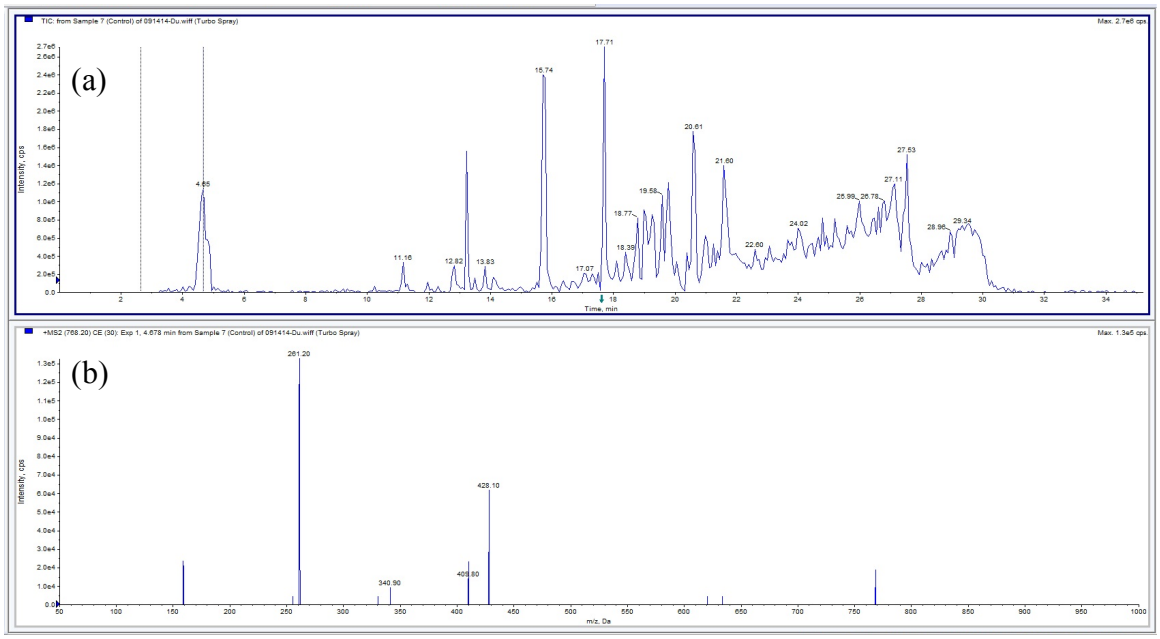
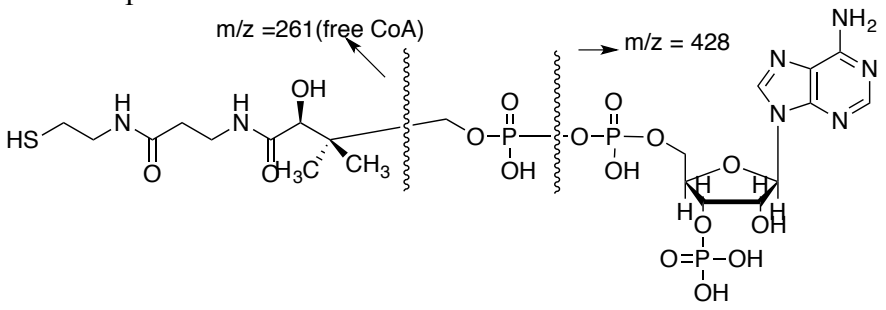


Figure 5-16. MS/MS (+) of the control reaction a: Total Ion Chromatogram (TIC) of the control reaction b: MS-MS Spectrum of the selected peak showing expected fragments for CoA.

Fragmentation pattern:



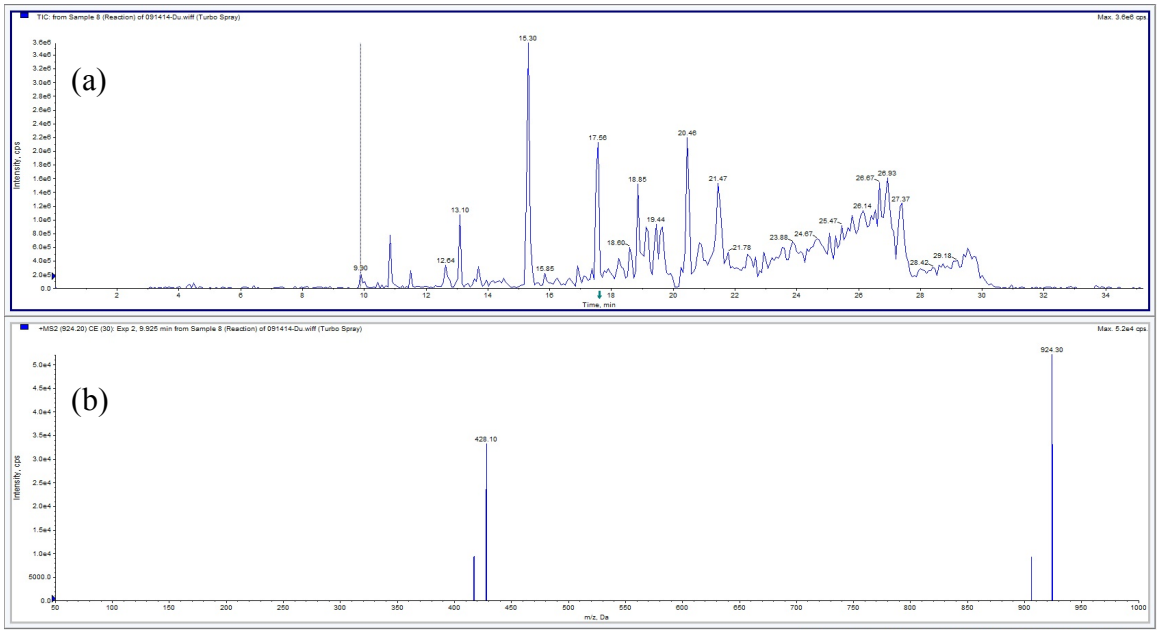
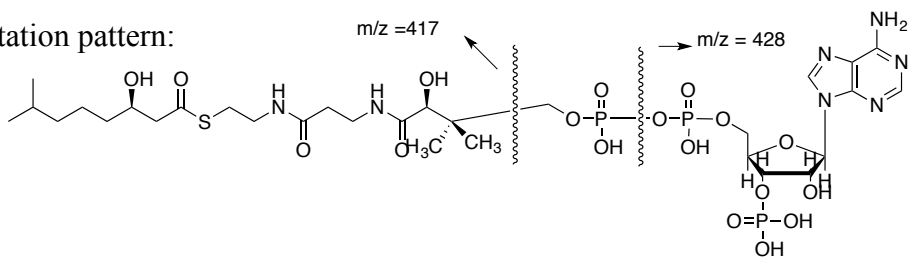


Figure 5-17. MS/MS (+) of the enzymatic reaction. a: Total Ion Chromatogram of the enzyme reaction. b: MS-MS Spectrum of the selected peak showing that expected fragments for product.

Fragmentation pattern:



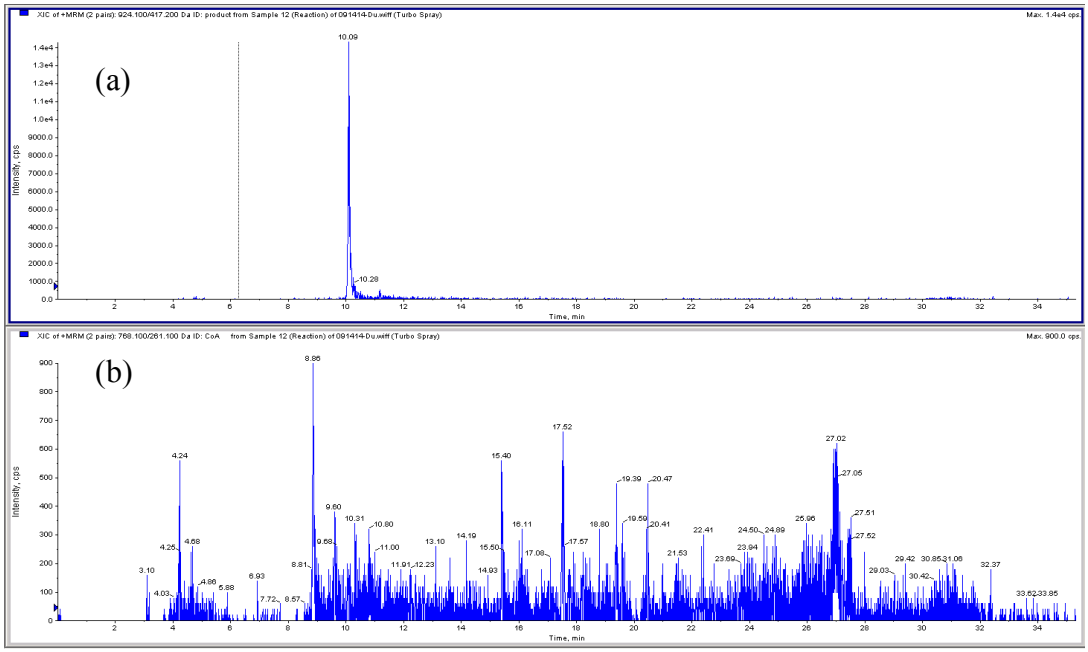


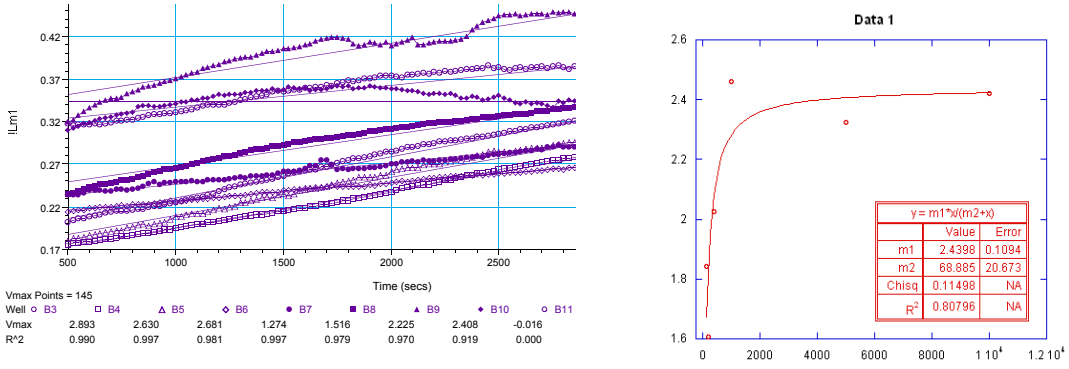
Figure 5-18. Multiple Reaction Monitoring (MRM) Chromatogram of Enzyme Reaction. a: MRM Chromatogram of monitoring Q1(+) (the first mass analyzer) *m*/*z*=924.23 and Q3 (+) (the third mass analyzer) *m*/*z* =417.23, confirming the identity of the product formed. b: MRM Chromatogram of monitoring Q1 (+)*m*/*z* =768.2 and Q3 (+) *m*/*z* =261.0, showing that there is no peak for CoA, indicating that the substrates are consumed by the reaction.

5-3-6. Kinetic analysis of acyl-CoA ligase activity

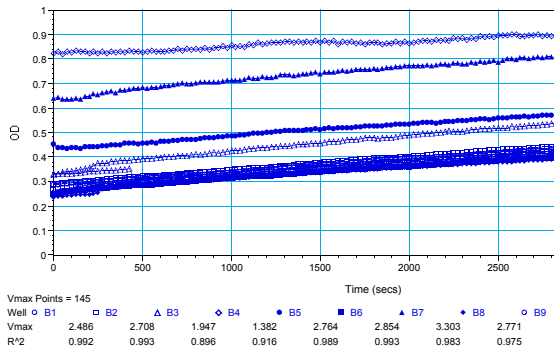
To quantify the enzymatic activity of ACL-6, we performed the initial velocity studies [142]. The substrate selectivity was also evaluated by the kinetics assay using EnzChek Pyrophosphate Assay Kit. The inorganic pyrophosphatase in this kit catalyzes conversion of PPi into two equivalents of Pi. The Pi is then consumed by the 2-amino-6-mercapto-7-methyl-purine ribonucleoside (MESG) by purine ribonucleoside phosphorylase (PNP) and the product 2-amino-6-mercapto-7-methyl purine is detected by an increase in absorbance at 360 nm (Figure 5-5). Thus the

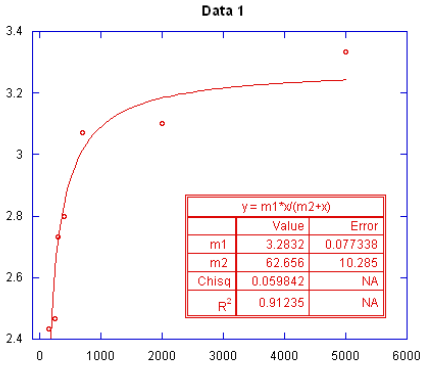
EnzChek Pyrophosphate Assay Kit can be used for the quantitation of PPi in solution or for the continuous determination of PPi released in enzymatic reactions[145, 146]. To measure the activity of ACLs, the enzymes were incubated with free fatty acid, ATP, CoA and the coupled reaction system provided by the kit as described above. The kinetic mechanisms of ACL 1-7 were determined by initial velocity experiments using the coupled assay. These data were globally fitted to Michaelis-Menten Equation, yielding K_M values for 3-hydroxyloctanoic acid. The data and graphs suggested that substrate 3-hydroxyloctanoic acid is favored by ACL-5 (K_m=15.001 ± 1.894 μM, k_{cat}=0.0198 ± 0.0003 min^{-1}), while 3-hydroxyl-7-methyloctanoic acid is favored by ACL-6 (K_m= 4.900 ± 0.0810 μM, k_{cat}= 0.326 ± 0.005 min^{-1}), but not completely ignored by the other ACLs. This is consistent with the idea that *in trans* ACLs generally exhibit a broad specificity and are functionally redundant (Table 5-2 a-b). Saturation kinetics was not achievable for ACL-2 at concentrations less than 10 μM. No steady signal was detected in the enzyme-free control reactions, showing that the data obtained resulted from the enzymatic activities of ACLs. This is consistent with the *in vitro* and *in vivo* data that the ACL-6 is the corresponding ligase of 3-hydroxy 7-methyloctanoic acid and CoA, while ACL-5 selectively catalyzes the ligation of 3-hydroxy octanoic acid with CoA. Although these ACLs differ in K_M to a large extent, both k_{cat} and v_{max} were relatively small. The slow reaction rates might explain the very low yield of WAPs in the microorganism.

(A)

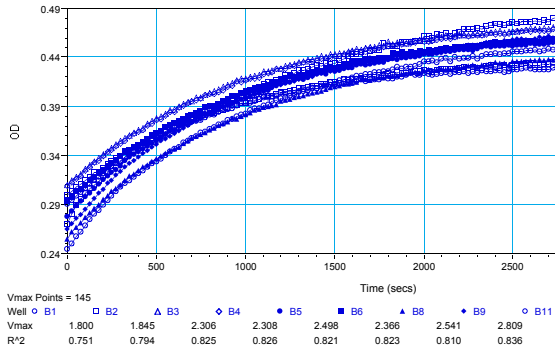


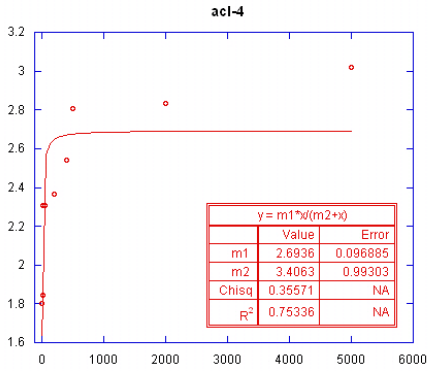
(B)


OD
Time (secs)
Vmax Points = 145
Well B1 B2 B3 B4 B5 B6 B7 B8 B9
Vmax 2.486 2.708 1.947 1.382 2.764 2.854 3.303 2.771
R^2 0.992 0.993 0.896 0.916 0.989 0.993 0.983 0.975

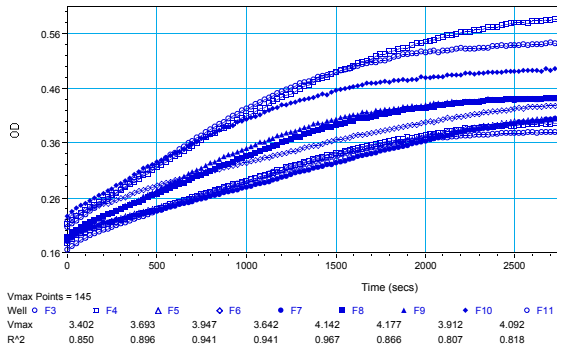

Data 1
y = m1*x/(m2+x)
Value Error
m1 3.2832 0.077338
m2 62.656 10.285
Chisq 0.059842 NA
R² 0.91235 NA

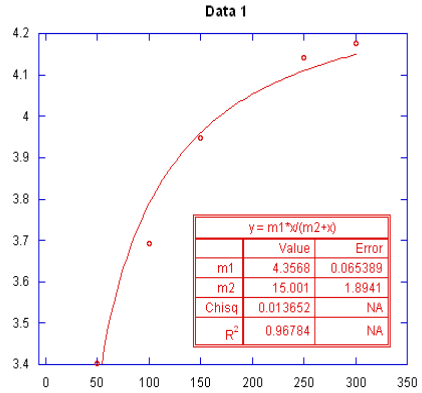
(C)


OD
Time (secs)
Vmax Points = 145
Well B1 B2 B3 B4 B5 B6 B8 B9 B11
Vmax 1.800 1.845 2.306 2.308 2.498 2.366 2.541 2.809
R^2 0.751 0.794 0.825 0.826 0.821 0.823 0.810 0.836

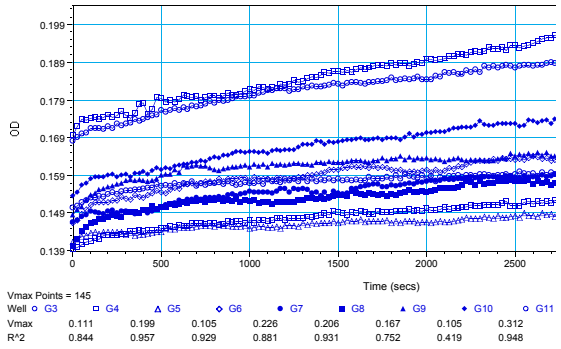

acl-4
y = m1*x/(m2+x)
Value Error
m1 2.6936 0.096885
m2 3.4063 0.99303
Chisq 0.35571 NA
R² 0.75336 NA

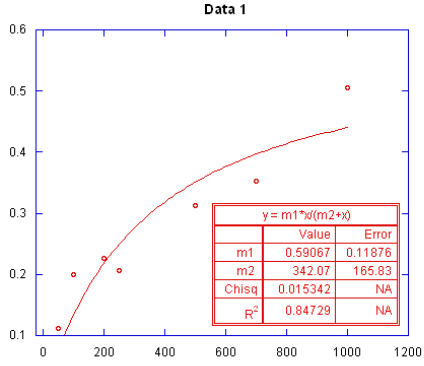
(D)


OD
Time (secs)
Vmax Points = 145
Well F3 F4 F5 F6 F7 F8 F9 F10 F11
Vmax 3.402 3.693 3.947 3.642 4.142 4.177 3.912 4.092
R^2 0.850 0.896 0.941 0.941 0.967 0.866 0.807 0.818


Data 1
y = m1*x/(m2+x)
Value Error
m1 4.3568 0.065389
m2 15.001 1.8941
Chisq 0.013652 NA
R² 0.96784 NA

(E)


OD
Time (secs)
Vmax Points = 145
Well G3 G4 G5 G6 G7 G8 G9 G10 G11
Vmax 0.111 0.199 0.105 0.226 0.206 0.167 0.105 0.312
R^2 0.844 0.957 0.929 0.881 0.931 0.752 0.419 0.948


Data 1
y = m1*x/(m2+x)
Value Error
m1 0.59067 0.11876
m2 342.07 165.83
Chisq 0.015342 NA
R² 0.84729 NA

(F)

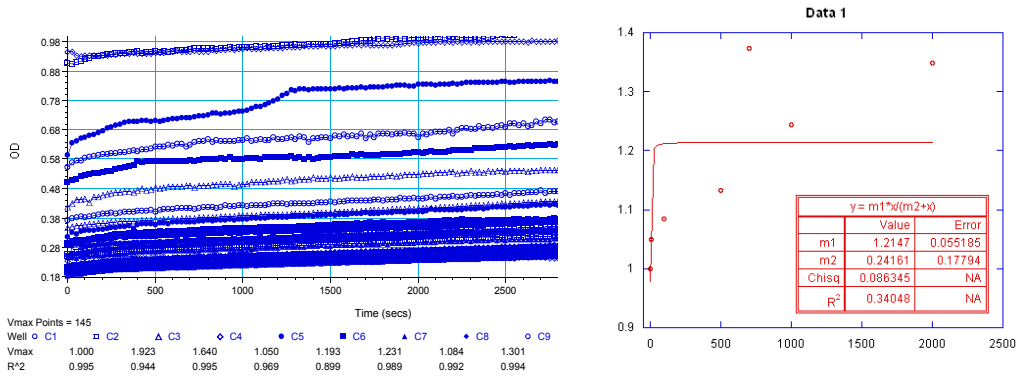


Figure 5-19. Kinetic mechanisms of ACL 1-7 (except ACL-2) using 3-hydroxyl-7-methyloctanoic acid as substrate. A-F: ACL1, 3, 4, 5, 6, 7. Left: initial rate curve. Right: Michealis-Menten equation fitted curve.

Table 5-2. a: Kinetic parameters for ACL 1-7 against 3-hydroxy-octanoic acid

	K_M (μM)	v_{max} (μM/min)	k_{cat} (min-1)
ACL-1	68.88 ± 20.67	0.1671 ± 0.0075	0.0111 ± 0.0005
ACL-3	62.65 ± 1.29	0.2248 ± 0.0053	0.0149 ± 0.0003
ACL-4	3.41 ± 0.99	0.1845 ± 0.0066	0.0123 ± 0.0004
ACL-5	15.00 ± 1.89	0.2983 ± 0.0044	0.0198 ± 0.0003
ACL-6	342.07 ± 16.58	0.0405 ± 0.0081	0.0027 ± 0.0005
ACL-7	15.49 ± 1.05	0.0864 ± 0.0039	0.0057 ± 0.0002

b: Kinetic parameters for ACL 1-7 against 3-hydroxy 7-methyloctanoic acid

	K_M (uM)	v_{max} (μM/min)	k_{cat} (min-1)
ACL-1	70.24 ± 12.43	0.2420 ± 0.0065	0.0161 ± 0.0004
ACL-3	6.12 ± 1.67	0.3148 ± 0.0073	0.0209 ± 0.0004
ACL-4	511.76 ± 379.39	0.6342 ± 0.0859	0.0422 ± 0.0057
ACL-5	49.98 ± 11.87	0.9535 ± 0.0397	0.0635 ± 0.0026
ACL-6	5.84 ± 1.28	4.9006 ± 0.0810	0.3267 ± 0.0054
ACL-7*			

*: ACL7 does not show distinct activity toward this substrate, thus the kinetics parameters could not be obtained.

*: Saturation kinetics was not achievable for ACL-2 at concentrations less than 10 μM.

5-4. Summary and final remarks

WAP-8294As are cyclic lipodepsipeptides with potent antibacterial activity. Although many lipopeptide antibiotics are known, including daptomycin (Cubicin), WAP-8294As exhibit exceptionally potent activities against MRSA and VRSA (ED_{50} is 14 times higher than that of vancomycin). The fact that WAP-8294A2 reached the Phase-I/II clinical studies testifies to its promise. Thus, study and engineering of WAPs biosynthetic gene cluster to produce various WAPs analogues represent a

valuable new opportunity to generate new antibiotics. In the previous study,[110] we identified a total of 45 domains from the two NRPS genes within the WAPS cluster, representing one of the largest NRPS systems.

The N-terminus of WAPs is acylated with a fatty acid side chain, which is a common feature found in other cyclic lipopeptides, such as surfactin, amphi-enterobactin[130], daptomycin, fengycin[147] and CDA[148]. When analyzing the C domain phylogeny of this type of NRPS, the N-terminus C domains present in the biosynthetic gene cluster of the lipopeptides listed above did not cluster with the known C domain subtypes, not even with the C domains downstream in the same assembly line. This N-terminus C domain is supposed to serve as an acceptor for a fatty acid, which is activated by ACLs[149-151]. In the case of surfactin biosynthesis, there are four putative ACL genes in the *B. subtilis* genome, two of which have the ability to catalyze the ligation of 3-hydroxyfatty acids and CoA[152]. Disruption of each of the four ACL could not eliminate the production of surfactin. Similarly, a disruption of an ACL gene (accession number JN596953) found in the genome of *L. enzymogenes* OH11 lowered the antibacterial activity and the yield of WAP-8294A2, suggesting a broad specificity and redundancy of these ACLs[110]. This is consistent with the fact that all WAP-8492A compounds have the same amino acid composition but vary in the length of the fatty acid side chain.

To further prove this feature and discover the most essential ACLs for fatty acyl priming in WAP biosynthesis, we constructed seven ACL disruption mutants. None of these mutants eliminates WAP production completely, suggesting that the 3-hydroxy fatty acid substrates can be activated by various ACLs. Among the 7

mutants, the ACL-6 mutant had the lowest yield of WAPs, thus became an important candidate in our investigation.

The ACL-6 deletion mutant was constructed to eliminate any possibility of polar effects caused by gene disruption, which inserts a large plasmid into the genome. We also constructed complementation and overexpression vectors to verify the ACL-6 activity, but were not able to obtain the ACL-6 complementary mutant. Instead, we used a chemically synthesized CoA surrogate, 3-hydroxy-7-methyloctanoic acid SNAC, in a feeding experiment to verify the activity. The subsequent restoration WAPs production supported that ACL-6 is required in WAP fatty acid substrate activation, but not NRPS priming. It also suggested that feeding the ΔACL-6 mutant with different side chain substrates to generate novel WAP analogues is possible.

We were able to express the seven ACL genes in *E. coli* and purify the enzymes. The kinetic data from *in vitro* studies revealed that ACL-6 is the most likely enzyme involved in WAP-8294A2 synthesis, while ACL-5 is the most likely enzyme involved in activating the acyl substrate of WAP-8294A1. Based on the results presented in this study, we propose the mechanism shown in Figure 5-20 for the initiation of the biosynthesis of WAPs.

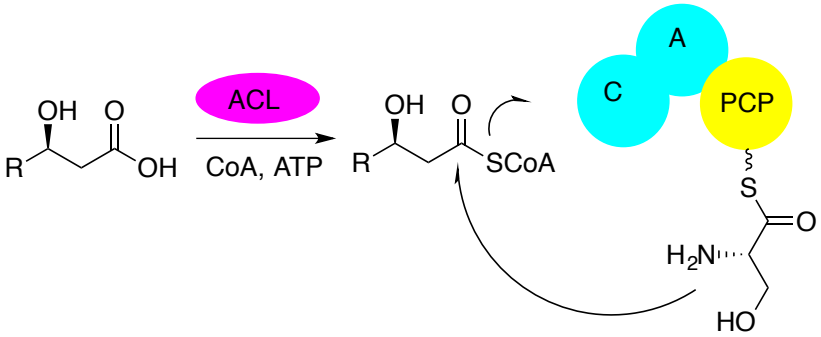


Figure 5-20. Proposed lipidation reaction during WAP biosynthesis. The fatty acid is activated by ACLs and then recognized by the donor site of C domain, which catalyzes the nucleophilic attack of the amino group of the PCP-bound serine on the carbonyl of fatty acyl-CoA.

The understanding of the activation mechanisms of the fatty acyl portion of the WAPs obtained in this study can be used to guide future biosynthetic engineering efforts. Fatty acid moieties have strong influence on the activity and properties of natural products. The final goal of this study is to generate a small family of WAP-8294A2 analogues with optimal structure-activity features (such as solubility verses activity by varying the fatty acid chain length and replacing certain amino acid residues) by feeding the wild type as well as the mutants with fatty acids with different chain lengths and various amino acids to direct the biosynthesis toward an optimized compound. It should also allow transfer of such logic to the discovery of new bioactive NPs from other *Lysobacter* species and allow optimization of the yield, structure, and activity of the *Lysobacter* anti-infectives.

An important task for the future is to understand the NRPS initiation in WAPs biosynthesis by investigating the activity of the first NRPS module (C-A_{ser}-PCP) *in vitro*. Although most C domain exhibits a greater selectivity towards the amino acid activated by the A domain within the same module compared with the intermediate in upstream modules[153, 154], it's incorrect to draw the conclusion that the fatty acyl substrate selectivity is only determined by ACLs. If the first C domain catalyzes the transfer of activated 3-hydroxy fatty acid at its donor site to the peptidyl carrier

protein (PCP) bound amino acid serine, the mechanism of lipoinitiation of WAPs could be explored using the purified C-A_{ser}-PCP.

Reference

1. Hertweck, C. (2009) The Biosynthetic Logic of Polyketide Diversity, *Angew. Chem. Int. Edit. 48*, 4688-4716.
2. Staunton, J., and Weissman, K. J. (2001) Polyketide biosynthesis: a millennium review, *Nat. Prod. Rep. 18*, 380-416.
3. Hopwood, D. A. (1997) Genetic contributions to understanding polyketide synthases, *Chem. Rev. 97*, 2465-2497.
4. Fischbach, M. A., and Walsh, C. T. (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms, *Chem. Rev. 106*, 3468-3496.
5. Khosla, C. (2000) Natural product biosynthesis: A new interface between enzymology and medicine, *J. Org. Chem. 65*, 8127-8133.
6. Wilkinson, B., Kendrew, S. G., Sheridan, R. M., and Leadlay, P. F. (2003) Biosynthetic engineering of polyketide synthases, *Expert. Opin. Ther. Pat. 13*, 1579-1606.
7. Zabala, A. O., Cacho, R. A., and Tang, Y. (2012) Protein engineering towards natural product synthesis and diversification, *J. Ind. Microbiol. Biot. 39*, 227-241.
8. Shen, B., and Thorson, J. S. (2012) Expanding nature's chemical repertoire through metabolic engineering and biocatalysis, *Curr. Opin. Chem. Biol. 16*, 99-100.
9. Watanabe, K. (2008) Exploring the Biosynthesis of Natural Products and Their Inherent Suitability for the Rational Design of Desirable Compounds through Genetic Engineering, *Biosci. Biotech. Bioch. 72*, 2491-2506.
10. Winter, J. M., and Tang, Y. (2012) Synthetic biological approaches to natural

product biosynthesis, *Curr. Opin. Biotech. 23*, 736-743.

11. Weissman, K. J., and Leadlay, P. F. (2005) Combinatorial biosynthesis of reduced polyketides, *Nat. Rev. Microbiol 3*, 925-936.
12. Williams, G. J. (2013) Engineering polyketide synthases and nonribosomal peptide synthetases, *Curr. Opin. Struc. Biol. 23*, 603-612.
13. Chan, Y. A., Podevels, A. M., Kevany, B. M., and Thomas, M. G. (2009) Biosynthesis of polyketide synthase extender units, *Nat. Prod. Rep. 26*, 90-114.
14. Sherman, D. H. (2005) The Lego-ization of polyketide biosynthesis, *Nat. Biotechnol. 23*, 1083-1084.
15. Khosla, C., Kapur, S., and Cane, D. E. (2009) Revisiting the modularity of modular polyketide synthases, *Curr. Opin. Chem. Biol. 13*, 135-143.
16. Challis, G. L. (2008) Genome mining for novel natural product discovery, *J. Med. Chem. 51*, 2618-2628.
17. Winter, J. M., Behnken, S., and Hertweck, C. (2011) Genomics-inspired discovery of natural products, *Curr. Opin. Chem. Biol. 15*, 22-31.
18. Shen, B., Cheng, Y. Q., Christenson, S. D., Jiangi, H., Ju, J. H., Kwon, H. J., Lim, S. K., Liu, W., Nonaka, K., Seo, J. W., Smith, W. C., Standage, S., Tang, G. L., Van Lanen, S., and Zhang, J. (2007) Polyketide Biosynthesis beyond the Type I, II, and III Polyketide Synthase Paradigms: A Progress Report, *Acs. Sym. Ser. 955*, 154-166.
19. Moss, S. J., Martin, C. J., and Wilkinson, B. (2004) Loss of co-linearity by modular polyketide synthases: a mechanism for the evolution of chemical diversity, *Nat. Prod. Rep. 21*, 575-593.
20. Zhang, O., Pang, B., Ding, W., and Liu, W. (2013) Aromatic Polyketides

Produced by Bacterial Iterative Type I Polyketide Synthases, *Acs. Catal. 3*, 1439-1447.

21. Kaulmann, U., and Hertweck, C. (2002) Biosynthesis of polyunsaturated fatty acids by polyketide synthases, *Angew. Chem. Int. Edit. 41*, 1866-+.

22. Metz, J. G., Roessler, P., Facciotti, D., Levering, C., Dittrich, F., Lassner, M., Valentine, R., Lardizabal, K., Domergue, F., Yamada, A., Yazawa, K., Knauf, V., and Browse, J. (2001) Production of polyunsaturated fatty acids by polyketide synthases in both prokaryotes and eukaryotes, *Science 293*, 290-293.

23. Liu, W., Christenson, S. D., Standage, S., and Shen, B. (2002) Biosynthesis of the enediyne antitumor antibiotic C-1027, *Science 297*, 1170-1173.

24. Van Lanen, S. G., Lin, S. J., and Shen, B. (2008) Biosynthesis of the enediyne antitumor antibiotic C-1027 involves a new branching point in chorismate metabolism, *P. Natl. Acad. Sci. USA 105*, 494-499.

25. Ahlert, J., Shepard, E., Lomovskaya, N., Zazopoulos, E., Staffa, A., Bachmann, B. O., Huang, K. X., Fonstein, L., Czisny, A., Whitwam, R. E., Farnet, C. M., and Thorson, J. S. (2002) The calicheamicin gene cluster and its iterative type I enediyne PKS, *Science 297*, 1173-1176.

26. Li, Y. Y., Chen, H. T., Ding, Y. J., Xie, Y. X., Wang, H. X., Cerny, R. L., Shen, Y. M., and Du, L. C. (2014) Iterative Assembly of Two Separate Polyketide Chains by the Same Single-Module Bacterial Polyketide Synthase in the Biosynthesis of HSAF, *Angew. Chem. Int. Edit. 53*, 7524-7530.

27. Blodgett, J. A. V., Oh, D. C., Cao, S. G., Currie, C. R., Kolter, R., and Clardy, J. (2010) Common biosynthetic origins for polycyclic tetramate macrolactams from phylogenetically diverse bacteria, *P. Natl. Acad. Sci. USA 107*,

11692-11697.

28. Fisch, K. M. (2013) Biosynthesis of natural products by microbial iterative hybrid PKS-NRPS, *Rsc. Adv. 3*, 18228-18247.

29. Wilkinson, B., Foster, G., Rudd, B. A. M., Taylor, N. L., Blackaby, A. P., Sidebottom, P. J., Cooper, D. J., Dawson, M. J., Buss, A. D., Gaisser, S., Bohm, I. U., Rowe, C. J., Cortes, J., Leadlay, P. F., and Staunton, J. (2000) Novel octaketide macrolides related to 6-deoxyerythronolide B provide evidence for iterative operation of the erythromycin polyketide synthase, *Chem. Biol. 7*, 111-117.

30. Hardt, I. H., Steinmetz, H., Gerth, K., Sasse, F., Reichenbach, H., and Hofle, G. (2001) New natural epothilones from Sorangium cellulosum, strains So ce90/B2 and So ce90/D13: Isolation, structure elucidation, and SAR studies, *J. Nat. Prod. 64*, 847-856.

31. Zhu, G., LaGier, M. J., Stejskal, F., Millership, J. J., Cai, X. M., and Keithly, J. S. (2002) Cryptosporidium parvum: the first protist known to encode a putative polyketide synthase, *Gene 298*, 79-89.

32. Gaitatzis, N., Silakowski, B., Kunze, B., Nordsiek, G., Blocker, H., Hofle, G., and Muller, R. (2002) The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase, *J. Biol. Chem. 277*, 13082-13090.

33. Moore, B. S., and Hopke, J. N. (2001) Discovery of a new bacterial polyketide biosynthetic pathway, *Chembiochem. 2*, 35-38.

34. Muller, R. (2004) Don't classify polyketide synthases, *Chem. Biol. 11*, 4-6.

35. Kennedy, J., Auclair, K., Kendrew, S. G., Park, C., Vederas, J. C., and Hutchinson, C. R. (1999) Modulation of polyketide synthase activity by

accessory proteins during lovastatin biosynthesis, *Science 284*, 1368-1372.

36. Crawford, J. M., and Townsend, C. A. (2010) New insights into the formation of fungal aromatic polyketides, *Nat. Rev. Microbiol. 8*, 879-889.

37. Fujii, I. (2010) Functional analysis of fungal polyketide biosynthesis genes, *J. Antibiot 63*, 207-218.

38. Cortes, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J., and Leadlay, P. F. (1990) An Unusually Large Multifunctional Polypeptide in the Erythromycin-Producing Polyketide Synthase of Saccharopolyspora-Erythraea, *Nature 348*, 176-178.

39. Smith, S., and Tsai, S. C. (2007) The type I fatty acid and polyketide synthases: a tale of two megasynthases, *Nat. Prod. Rep. 24*, 1041-1072.

40. Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., Kikuchi, H., Shiba, T., Sakaki, Y., and Hattori, M. (2001) Genome sequence of an industrial microorganism Streptomyces avermitilis: Deducing the ability of producing secondary metabolites, *P. Natl. Acad. Sci. USA 98*, 12215-12220.

41. Cheng, Y. Q., Tang, G. L., and Shen, B. (2003) Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis, *P. Natl. Acad. Sci. USA 100*, 3149-3154.

42. El-Sayed, A. K., Hothersall, J., Cooper, S. M., Stephens, E., Simpson, T. J., and Thomas, C. M. (2003) Characterization of the mupirocin biosynthesis gene cluster from Pseudomonas fluorescens NCIMB 10586, *Chem. Biol. 10*, 419-430.

43. Jenke-Kodama, H., Sandmann, A., Muller, R., and Dittmann, E. (2005) Evolutionary implications of bacterial polyketide synthases, *Mol. Biol. Evol*.

22, 2027-2039.

44. Jia, X. Y., Tian, Z. H., Shao, L., Qu, X. D., Zhao, Q. F., Tang, J., Tang, G. L., and Liu, W. (2006) Genetic characterization of the chlorothricin gene cluster as a model for spirotetronate antibiotic biosynthesis, *Chem. Biol. 13*, 575-585.

45. Zhao, Q. F., He, Q. L., Ding, W., Tang, M. C., Kang, Q. J., Yu, Y., Deng, W., Zhang, Q., Fang, J., Tang, G. L., and Liu, W. (2008) Characterization of the azinomycin B biosynthetic gene cluster revealing a different iterative type I polyketide synthase for naphthoate biosynthesis, *Chem. Biol. 15*, 693-705.

46. Xu, W., Qiao, K. J., and Tang, Y. (2013) Structural analysis of protein-protein interactions in type I polyketide synthases, *Crit. Rev. Biochem. Mol. 48*, 98-122.

47. Katz, L. (2009) The Debs Paradigm for Type I Modular Polyketide Synthases and Beyond, *Methods Enzymol. 459*, 113-142.

48. Bretschneider, T., Heim, J. B., Heine, D., Winkler, R., Busch, B., Kusebauch, B., Stehle, T., Zocher, G., and Hertweck, C. (2013) Vinylogous chain branching catalysed by a dedicated polyketide synthase module, *Nature 502*, 124.

49. Cheng, Y. Q., Tang, G. L., and Shen, B. (2002) Identification and localization of the gene cluster encoding biosynthesis of the antitumor macrolactam leinamycin in Streptomyces atroolivaceus S-140, *J. Bacteriol. 184*, 7013-7024.

50. Liu, J., Zhu, X., Seipke, R. F., and Zhang, W. (2014) Biosynthesis of Antimycins with a Reconstituted 3-Formamidosalicylate Pharmacophore in Escherichia coli, *Acs Synth. Biol*.

51. Palaniappan, N., Alhamadsheh, M. M., and Reynolds, K. A. (2008)

cis-Delta(2,3)-double bond of phoslactomycins is generated by a post-PKS tailoring enzyme*, J. Am. Chem. Soc. 130*, 12236.

52. Chen, Y. L., Zhao, J., Liu, W., Gao, J. F., Tao, L. M., Pan, H. X., and Tang, G. L. (2012) Identification of phoslactomycin biosynthetic gene clusters from Streptomyces platensis SAM-0654 and characterization of PnR1 and PnR2 as positive transcriptional regulators, *Gene 509*, 195-200.

53. Molnar, I., Schupp, T., Ono, M., Zirkle, R., Milnamow, M., Nowak-Thompson, B., Engel, N., Toupet, C., Stratmann, A., Cyr, D. D., Gorlach, J., Mayo, J. M., Hu, A., Goff, S., Schmid, J., and Ligon, J. M. (2000) The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from Sorangium cellulosum So ce90, *Chem. Biol. 7*, 97-109.

54. Julien, B., Shah, S., Ziermann, R., Goldman, R., Katz, L., and Khosla, C. (2000) Isolation and characterization of the epothilone biosynthetic gene cluster from Sorangium cellulosum, *Gene 249*, 153-160.

55. Takahashi, S., Toyoda, A., Sekiyama, Y., Takagi, H., Nogawa, T., Uramoto, M., Suzuki, R., Koshino, H., Kumano, T., Panthee, S., Dairi, T., Ishikawa, J., Ikeda, H., Sakaki, Y., and Osada, H. (2011) Reveromycin A biosynthesis uses RevG and RevJ for stereospecific spiroacetal formation, *Nat. Chem. Biol. 7*, 461-468.

56. Gaisser, S., Trefzer, A., Stockert, S., Kirschning, A., and Bechthold, A. (1997) Cloning of an avilamycin biosynthetic gene cluster from Streptomyces viridochromogenes Tu57, *J. Bacteriol. 179*, 6271-6278.

57. Whitwam, R. E., Ahlert, J., Holman, T. R., Ruppen, M., and Thorson, J. S. (2000) The gene calC encodes for a non-heme iron metalloprotein responsible

for calicheamicin self-resistance in Micromonospora, *J. Am. Chem. Soc. 122*, 1556-1557.

58. Zazopoulos, E., Huang, K. X., Staffa, A., Liu, W., Bachmann, B. O., Nonaka, K., Ahlert, J., Thorson, J. S., Shen, B., and Farnet, C. M. (2003) A genomics-guided approach for discovering and expressing cryptic metabolic pathways, *Nat. Biotechnol. 21*, 187-190.

59. Van Lanen, S. G., Oh, T. J., Liu, W., Wendt-Pienkowski, E., and Shen, B. (2007) Characterization of the maduropeptin biosynthetic gene cluster from Actinomadura madurae ATCC 39144 supporting a unifying paradigm for enediyne biosynthesis, *J. Am. Chem. Soc. 129*, 13082-13094.

60. Daum, M., Peintner, I., Linnenbrink, A., Frerich, A., Weber, M., Paululat, T., and Bechthold, A. (2009) Organisation of the Biosynthetic Gene Cluster and Tailoring Enzymes in the Biosynthesis of the Tetracyclic Quinone Glycoside Antibiotic Polyketomycin, *Chembiochem. 10*, 1073-1083.

61. Ito, T., Roongsawang, N., Shirasaka, N., Lu, W. L., Flatt, P. M., Kasanah, N., Miranda, C., and Mahmud, T. (2009) Deciphering Pactamycin Biosynthesis and Engineered Production of New Pactamycin Analogues, *Chembiochem. 10*, 2253-2265.

62. Xiao, Y., Li, S. M., Niu, S. W., Ma, L. A., Zhang, G. T., Zhang, H. B., Zhang, G. Y., Ju, J. H., and Zhang, C. S. (2011) Characterization of Tiacumicin B Biosynthetic Gene Cluster Affording Diversified Tiacumicin Analogues and Revealing a Tailoring Dihalogenase, *J. Am. Chem. Soc. 133*, 1092-1105.

63. Ding, W., Deng, W., Tang, M. C., Zhang, Q., Tang, G. L., Bi, Y. R., and Liu, W. (2010) Biosynthesis of 3-methoxy-5-methyl naphthoic acid and its incorporation into the antitumor antibiotic azinomycin B, *Mol. Biosyst. 6*,

1071-1081.

64. Sun, H. H., Ho, C. L., Ding, F. Q., Soehano, I., Liu, X. W., and Liang, Z. X. (2012) Synthesis of (R)-Mellein by a Partially Reducing Iterative Polyketide Synthase, *J. Am. Chem. Soc. 134*, 11924-11927.

65. Wu, J. Q., Zaleski, T. J., Valenzano, C., Khosla, C., and Cane, D. E. (2005) Polyketide double bond biosynthesis. Mechanistic analysis of the dehydratase-containing module 2 of the picromycin/methymycin polyketide synthase, *J. Am. Chem. Soc. 127*, 17393-17404.

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

67. He, Q. L., Jia, X. Y., Tang, M. C., Tian, Z. H., Tang, G. L., and Liu, W. (2009) Dissection of Two Acyl-Transfer Reactions Centered on Acyl-S-Carrier Protein Intermediates for Incorporating 5-Chloro-6-methyl-O-methylsalicyclic Acid into Chlorothricin, *Chembiochem. 10*, 813-819.

68. Castillo, Y. P., and Perez, M. A. (2008) Bacterial beta-ketoacyl-acyl carrier protein synthase III (FabH): an attractive target for the design of new broad-spectrum antimicrobial agents, *Mini-Rev. Med. Chem. 8*, 36-45.

69. Weitnauer, G., Muhlenweg, A., Trefzer, A., Hoffmeister, D., Sussmuth, R. D., Jung, G., Welzel, K., Vente, A., Girreser, U., and Bechthold, A. (2001) Biosynthesis of the orthosomycin antibiotic avilamycin A: deductions from the molecular analysis of the avi biosynthetic gene cluster of Streptomyces viridochromogenes Tu57 and production of new antibiotics, *Chem. Biol. 8*, 569-581.

70. Weitnauer, G., Hauser, G., Hofmann, C., Linder, U., Boll, R., Pelz, K., Glaser, S. J., and Bechthold, A. (2004) Novel avilamycin derivatives with improved

polarity generated by targeted gene disruption, *Chem. Biol. 11*, 1403-1411.

71. Okuyama, H., Orikasa, Y., Nishida, T., Watanabe, K., and Morita, N. (2007) Bacterial genes responsible for the biosynthesis of eicosapentaenoic and docosahexaenoic acids and their heterologous expression, *Appl. Environ. Microb. 73*, 665-670.

72. Jiang, H., Zirkle, R., Metz, J. G., Braun, L., Richter, L., Van Lanen, S. G., and Shen, B. (2008) The role of tandem acyl carrier protein domains in polyunsaturated fatty acid biosynthesis, *J. Am. Chem. Soc. 130*, 6336.

73. Lam, K. S., Veitch, J. A., Golik, J., Krishnan, B., Klohr, S. E., Volk, K. J., Forenza, S., and Doyle, T. W. (1993) Biosynthesis of Esperamicin-a(1), an Enediyne Antitumor Antibiotic, *J. Am. Chem. Soc. 115*, 12340-12345.

74. Belecki, K., Crawford, J. M., and Townsend, C. A. (2009) Production of Octaketide Polyenes by the Calicheamicin Polyketide Synthase CalE8: Implications for the Biosynthesis of Enediyne Core Structures, *J. Am. Chem. Soc. 131*, 12564.

75. Liu, W., Ahlert, J., Gao, Q. J., Wendt-Pienkowski, E., Shen, B., and Thorson, J. S. (2003) Rapid PCR amplification of minimal enediyne polyketide synthase cassettes leads to a predictive familial classification model, *P. Natl. Acad. Sci. USA 100*, 11959-11963.

76. Lou, L. L., Qian, G. L., Xie, Y. X., Hang, J. L., Chen, H. T., Zaleta-Riyera, K., Li, Y. Y., Shen, Y. M., Dussault, P. H., Liu, F. Q., and Du, L. C. (2011) Biosynthesis of HSAF, a Tetramic Acid-Containing Macrolactam from Lysobacter enzymogenes, *J. Am. Chem. Soc. 133*, 643-645.

77. Li, S. J., Du, L. C., Yuen, G., and Harris, S. D. (2006) Distinct ceramide synthases regulate polarized growth in the filamentous fungus Aspergillus

nidulans, *Mol. Biol. Cell. 17*, 1218-1227.

78. Yu, F. G., Zaleta-Rivera, K., Zhu, X. C., Huffman, J., Millet, J. C., Harris, S. D., Yuen, G., Li, X. C., and Du, L. C. (2007) Structure and biosynthesis of heat-stable antifungal factor (HSAF), a broad-spectrum antimycotic with a novel mode of action, *Antimicrob. Agents Chemother. 51*, 64-72.

79. Lou, L. L., Chen, H. T., Cerny, R. L., Li, Y. Y., Shen, Y. M., and Du, L. C. (2012) Unusual Activities of the Thioesterase Domain for the Biosynthesis of the Polycyclic Tetramate Macrolactam HSAF in Lysobacter enzymogenes C3, *Biochemistry 51*, 4-6.

80. Antosch, J., Schaefers, F., and Gulder, T. A. (2014) Heterologous reconstitution of ikarugamycin biosynthesis in E. coli, *Angew. Chem. Int. Edit. 53*, 3011-3014.

81. Zhang, G. T., Zhang, W. J., Zhang, Q. B., Shi, T., Ma, L., Zhu, Y. G., Li, S. M., Zhang, H. B., Zhao, Y. L., Shi, R., and Zhang, C. S. (2014) Mechanistic Insights into Polycycle Formation by Reductive Cyclization in Ikarugamycin Biosynthesis, *Angew. Chem. Int. Edit. 53*, 4840-4844.

82. Beyer, S., Kunze, B., Silakowski, B., and Muller, R. (1999) Metabolic diversity in myxobacteria: identification of the myxalamid and the stigmatellin biosynthetic gene cluster of Stigmatella aurantiaca Sg a15 and a combined polyketide-(poly)peptide gene cluster from the epothilone producing strain Sorangium cellulosum So ce90, *Bba-Gene Struct. Expr. 1445*, 185-195.

83. Broadhurst, R. W., Nietlispach, D., Wheatcroft, M. P., Leadlay, P. F., and Weissman, K. J. (2003) The structure of docking domains in modular polyketide synthases, *Chem. Biol. 10*, 723-731.

84. Mochizuki, S., Hiratsu, K., Suwa, M., Ishii, T., Sugino, F., Yamada, K., and

Kinashi, H. (2003) The large linear plasmid pSLA2-L of Streptomyces rochei has an unusually condensed gene organization for secondary metabolism, *Mol. Microbiol. 48*, 1501-1510.

85. Nowak-Thompson, B., Gould, S. J., and Loper, J. E. (1997) Identification and sequence analysis of the genes encoding a polyketide synthase required for pyoluteorin biosynthesis in Pseudomonas fluorescens Pf-5, *Gene 204*, 17-24.

86. Busch, B., Ueberschaar, N., Behnken, S., Sugimoto, Y., Werneburg, M., Traitcheva, N., He, J., and Hertweck, C. (2013) Multifactorial Control of Iteration Events in a Modular Polyketide Assembly Line, *Angew. Chem. Int. Edit. 52*, 5285-5289.

87. Busch, B., Ueberschaar, N., Sugimoto, Y., and Hertweck, C. (2012) Interchenar Retrotransfer of Aureothin Intermediates in an Iterative Polyketide Synthase Module, *J. Am. Chem. Soc. 134*, 12382-12385.

88. Chopra, T., Banerjee, S., Gupta, S., Yadav, G., Anand, S., Surolia, A., Roy, R. P., Mohanty, D., and Gokhale, R. S. (2008) Novel intermolecular iterative mechanism for biosynthesis of mycoketide synthase by a bimodular polyketide synthase, *Plos. Biol. 6*, 1584-1598.

89. Busch, B., and Hertweck, C. (2009) Evolution of metabolic diversity in polyketide-derived pyrones: Using the non-colinear aureothin assembly line as a model system, *Phytochemistry 70*, 1833-1840.

90. Kupferschmidt, K. (2012) Attack of the Clones, *Science 337*, 636-638.

91. Lou, L., Qian, G., Xie, Y., Hang, J., Chen, H., Zaleta-Rivera, K., Li, Y., Shen, Y., Dussault, P. H., Liu, F., and Du, L. (2011) Biosynthesis of HSAF, a tetramic acid-containing macrolactam from Lysobacter enzymogenes, *J. Am. Chem. Soc. 133*, 643-645.

92. Xie, Y. X., Wright, S., Shen, Y. M., and Du, L. C. (2012) Bioactive natural products from Lysobacter, *Nat. Prod. Rep. 29*, 1277-1287.

93. Li, Y. Y., Huffman, J., Li, Y., Du, L. C., and Shen, Y. M. (2012) 3-Hydroxylation of the polycyclic tetramate macrolactam in the biosynthesis of antifungal HSAF from Lysobacter enzymogenes C3, *Medchemcomm 3*, 982-986.

94. Jiang, Y., Wang, H., Lu, C., Ding, Y., Li, Y., and Shen, Y. (2013) Identification and characterization of the cuevaene A biosynthetic gene cluster in streptomyces sp. LZ35, *Chembiochem. 14*, 1468-1475.

95. Zhao, G. S., Li, S. R., Wang, Y. Y., Hao, H. L., Shen, Y. M., and Lu, C. H. (2013) 16,17-dihydroxycyclooctatin, a new diterpene from Streptomyces sp. LZ35, *Drug discoveries & therapeutics 7*, 185-188.

96. Zhou, M., Jing, X., Xie, P., Chen, W., Wang, T., Xia, H., and Qin, Z. (2012) Sequential deletion of all the polyketide synthase and nonribosomal peptide synthetase biosynthetic gene clusters and a 900-kb subtelomeric sequence of the linear chromosome of Streptomyces coelicolor, *Fems. Microbiol. Lett. 333*, 169-179.

97. Sanchez, C., Du, L., Edwards, D. J., Toney, M. D., and Shen, B. (2001) Cloning and characterization of a phosphopantetheinyl transferase from Streptomyces verticillus ATCC15003, the producer of the hybrid peptide-polyketide antitumor drug bleomycin, *Chem. Biol. 8*, 725-738.

98. Kayser, J. P., Vallet, J. L., and Cerny, R. L. (2004) Defining parameters for homology-tolerant database searching, *J. Biomol. Tech. 15*, 285-295.

99. Hitchman, T. S., Crosby, J., Byrom, K. J., Cox, R. J., and Simpson, T. J. (1998) Catalytic self-acylation of type II polyketide synthase acyl carrier proteins,

Chem. Biol. 5, 35-47.

100. Crosby, J., Byrom, K. J., Hitchman, T. S., Cox, R. J., Crump, M. P., Findlow, I. S., Bibb, M. J., and Simpson, T. J. (1998) Acylation of Streptomyces type II polyketide synthase acyl carrier proteins, *FEBS letters 433*, 132-138.

101. Du, L., and Lou, L. (2010) PKS and NRPS release mechanisms, *Nat. Prod. Rep. 27*, 255-278.

102. Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products, *Proc. Natl. Acad. Sci. U S A 97*, 6640-6645.

103. Wilkinson, C. J., Hughes-Thomas, Z. A., Martin, C. J., Bohm, I., Mironenko, T., Deacon, M., Wheatcroft, M., Wirtz, G., Staunton, J., and Leadlay, P. F. (2002) Increasing the efficiency of heterologous promoters in actinomycetes, *J. Mol. Microbiol. Biotechnol. 4*, 417-426.

104. Bibb, M. J., Janssen, G. R., and Ward, J. M. (1985) Cloning and analysis of the promoter region of the erythromycin resistance gene (ermE) of Streptomyces erythraeus, *Gene 38*, 215-226.

105. Sanchez, C., Du, L., Edwards, D. J., Toney, M. D., and Shen, B. (2001) Cloning and characterization of a phosphopantetheinyl transferase from *Streptomyces verticillus* ATCC15003, the producer of the hybrid peptide-polyketide antitumor drug bleomycin, *Chem. Biol. 8*, 725-738.

106. Dorrestein, P. C., Bumpus, S. B., Calderone, C. T., Garneau-Tsodikova, S., Aron, Z. D., Straight, P. D., Kolter, R., Walsh, C. T., and Kelleher, N. L. (2006) Facile detection of acyl and peptidyl intermediates on thiotemplate carrier domains via phosphopantetheinyl elimination reactions during tandem mass spectrometry, *Biochemistry 45*, 12756-12766.

107. Gerber, R., Lou, L., and Du, L. (2009) A PLP-dependent polyketide chain releasing mechanism in the biosynthesis of mycotoxin fumonisins in Fusarium verticillioides, *J. Am. Chem. Soc. 131*, 3148-3149.

108. Cao, S., Blodgett, J. A., and Clardy, J. (2010) Targeted discovery of polycyclic tetramate macrolactams from an environmental Streptomyces strain, *Org. Lett. 12*, 4652-4654.

109. Shigemori, H., Bae, M. A., Yazawa, K., Sasaki, T., and Kobayashi, J. (1992) Alteramide-a, a New Tetracyclic Alkaloid from a Bacterium-Alteromonas Sp Associated with the Marine Sponge Halichondria-Okadai, *J. Org. Chem. 57*, 4317-4320.

110. Zhang, W., Li, Y. Y., Qian, G. L., Wang, Y., Chen, H. T., Li, Y. Z., Liu, F. Q., Shen, Y. M., and Du, L. C. (2011) Identification and Characterization of the Anti-Methicillin-Resistant Staphylococcus aureus WAP-8294A2 Biosynthetic Gene Cluster from Lysobacter enzymogenes OH11, *Antimicrob. Agents Chemother. 55*, 5581-5589.

111. Olson, A. S., Chen, H. T., Du, L. C., and Dussault, P. H. (2015) Synthesis of a 2,4,6,8,10-dodecapentanoic acid thioester as a substrate for biosynthesis of heat stable antifungal factor (HSAF), *Rsc. Adv. 5*, 11644-11648.

112. Ma, S. M., Li, J. W. H., Choi, J. W., Zhou, H., Lee, K. K. M., Moorthie, V. A., Xie, X. K., Kealey, J. T., Da Silva, N. A., Vederas, J. C., and Tang, Y. (2009) Complete Reconstitution of a Highly Reducing Iterative Polyketide Synthase, *Science 326*, 589-592.

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

114. Alhamadsheh, M. M., Palaniappan, N., Daschouduri, S., and Reynolds, K. A. (2007) Modular polyketide synthases and cis double bond formation: establishment of activated cis-3-cyclohexylpropenoic acid as the diketide intermediate in phoslactomycin biosynthesis, *J. Am. Chem. Soc. 129*, 1910-1911.

115. Liavonchanka, A., and Feussner, I. (2008) Biochemistry of PUFA double bond isomerases producing conjugated linoleic acid, *Chembiochem 9*, 1867-1872.

116. Auclair, K., Sutherland, A., Kennedy, J., Witter, D. J., Van den Heever, J. P., Hutchinson, C. R., and Vederas, J. C. (2000) Lovastatin nonaketide synthase catalyzes an intramolecular Diels-Alder reaction of a substrate analogue, *J. Am. Chem. Soc. 122*, 11519-11520.

117. Katayama, K., Kobayashi, T., Chijimatsu, M., Ichihara, A., and Oikawa, H. (2008) Purification and N-terminal amino acid sequence of solanapyrone synthase, a natural Diels-Alderase from Alternaria solani, *Biosci. Biotechnol. Biochem. 72*, 604-607.

118. Watanabe, K., Mie, T., Ichihara, A., Oikawa, H., and Honma, M. (2000) Detailed reaction mechanism of macrophomate synthase. Extraordinary enzyme catalyzing five-step transformation from 2-pyrones to benzoates, *J. Biol. Chem. 275*, 38393-38401.

119. Ose, T., Watanabe, K., Mie, T., Honma, M., Watanabe, H., Yao, M., Oikawa, H., and Tanaka, I. (2003) Insight into a natural Diels-Alder reaction from the structure of macrophomate synthase, *Nature 422*, 185-189.

120. Stocking, E. M., and Williams, R. M. (2003) Chemistry and biology of biosynthetic Diels-Alder reactions, *Angew. Chem. Int. Ed. Engl. 42*,

3078-3115.

121. Kelly, W. L. (2008) Intramolecular cyclizations of polyketide biosynthesis: mining for a "Diels-Alderase"?, *Org. Biomol. Chem. 6*, 4483-4493.

122. Kim, H. J., Ruszczycky, M. W., Choi, S. H., Liu, Y. N., and Liu, H. W. (2011) Enzyme-catalysed [4+2] cycloaddition is a key step in the biosynthesis of spinosyn A, *Nature 473*, 109-112.

123. Youn, B., Kim, S. J., Moinuddin, S. G. A., Lee, C., Bedgar, D. L., Harper, A. R., Davin, L. B., Lewis, N. G., and Kang, C. (2006) Mechanistic and structural studies of apoform, binary, and ternary complexes of the Arabidopsis alkenal double bond reductase At5g16970, *J. Biol. Chem. 281*, 40076-40088.

124. Kumarasamy, K. K., Toleman, M. A., Walsh, T. R., Bagaria, J., Butt, F., Balakrishnan, R., Chaudhary, U., Doumith, M., Giske, C. G., Irfan, S., Krishnan, P., Kumar, A. V., Maharjan, S., Mushtaq, S., Noorie, T., Paterson, D. L., Pearson, A., Perry, C., Pike, R., Rao, B., Ray, U., Sarma, J. B., Sharma, M., Sheridan, E., Thirunarayan, M. A., Turton, J., Upadhyay, S., Warner, M., Welfare, W., Livermore, D. M., and Woodford, N. Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study, *Lancet. Infect. Dis. 10*, 597-602.

125. Harad, K. I., Suzuki, M., Kato, A., Fujii, K., Oka, H., and Ito, Y. (2001) Separation of WAP-8294A components, a novel anti-methicillin-resistant staphylococcus aureus antibiotic, using high-speed counter-current chromatography, *J. Chromatogr. A 932*, 75-81.

126. Walsh, C. (2003) Where will new antibiotics come from?, *Nat. Rev. Microbiol. 1*, 65-70.

127. Kato, A., Hirata, H., Ohashi, Y., Fujii, K., Mori, K., and Harada, K. (2011) A

new anti-MRSA antibiotic complex, WAP-8294A II. Structure characterization of minor components by ESI LCMS and MS/MS, *J. Antibiot. 64*, 373-379.

128. Kato, A., Nakaya, S., Kokubo, N., Aiba, Y., Ohashi, Y., Hirata, H., Fujii, K., and Harada, K. (1998) A new anti-MRSA antibiotic complex, WAP-8294A. I. Taxonomy, isolation and biological activities, *J. Antibiot. 51*, 929-935.

129. Kato, A., Nakaya, S., Ohashi, Y., and Hirata, H. (1997) WAP-8294A(2), a novel anti-MRSA antibiotic produced by Lysobacter sp., *J. Am. Chem. Soc. 119*, 6680-6681.

130. Zane, H. K., Naka, H., Rosconi, F., Sandy, M., Haygood, M. G., and Butler, A. (2014) Biosynthesis of Amphi-enterobactin Siderophores by Vibrio harveyi BAA-1116: Identification of a Bifunctional Nonribosomal Peptide Synthetase Condensation Domain, *J. Am. Chem. Soc. 136*, 5615-5618.

131. Baltz, R. H. (2009) Daptomycin: mechanisms of action and resistance, and biosynthetic engineering, *Curr. Opin. Chem. Biol. 13*, 144-151.

132. Chong, P. P., Podmore, S. M., Kieser, H. M., Redenbach, M., Turgay, K., Marahiel, M., Hopwood, D. A., and Smith, C. P. (1998) Physical identification of a chromosomal locus encoding biosynthetic genes for the lipopeptide calcium-dependent antibiotic (CDA) of Streptomyces coelicolor A3(2), *Microbiol-Uk 144*, 193-199.

133. Kraas, F. I., Helmetag, V., Wittmann, M., Strieker, M., and Marahiel, M. A. (2010) Functional Dissection of Surfactin Synthetase Initiation Module Reveals Insights into the Mechanism of Lipoinitiation, *Chem. Biol. 17*, 872-880.

134. Arora, P., Vats, A., Saxena, P., Mohanty, D., and Gokhale, R. S. (2005)

Promiscuous fatty acyl CoA ligases produce acyl-CoA and acyl-SNAC precursors for polyketide biosynthesis, *J. Am. Chem. Soc. 127*, 9388-9389.

135. Miao, V., Coeffet-Le Gal, M. F., Nguyen, K., Brian, P., Penn, J., Whiting, A., Steele, J., Kau, D., Martin, S., Ford, R., Gibson, T., Bouchard, M., Wrigley, S. K., and Baltz, R. H. (2006) Genetic engineering in Streptomyces roseosporus to produce hybrid lipopeptide antibiotics, *Chem. Biol. 13*, 269-276.

136. Imker, H. J., Krahn, D., Clerc, J., Kaiser, M., and Walsh, C. T. (2010) N-Acylation during Glidobactin Biosynthesis by the Tridomain Nonribosomal Peptide Synthetase Module GlbF, *Chem. Biol. 17*, 1077-1083.

137. Wang, Y., Qian, G. L., Liu, F. Q., Li, Y. Z., Shen, Y. M., and Du, L. C. (2013) Facile Method for Site-specific Gene Integration in Lysobacter enzymogenes for Yield Improvement of the Anti-MRSA Antibiotics WAP-8294A and the Antifungal Antibiotic HSAF, *Acs Synth. Biol. 2*, 670-678.

138. Milcamps, A., and de Bruijn, F. J. (1999) Identification of a novel nutrient-deprivation-induced Sinorhizobium meliloti gene (hmgA) involved in the degradation of tyrosine, *Microbiol-Uk 145*, 935-947.

139. Arias-Barrau, E., Olivera, E. R., Luengo, J. M., Fernandez, C., Galan, B., Garcia, J. L., Diaz, E., and Minambres, B. (2004) The homogentisate pathway: A central catabolic pathway involved in the degradation of L-phenylalanine, L-tyrosine, and 3-hydroxyphenylacetate in Pseudomonas putida, *J. Bacteriol. 186*, 5062-5077.

140. Werneburg, M., Busch, B., He, J., Richter, M. E. A., Xiang, L. K., Moore, B. S., Roth, M., Dahse, H. M., and Hertweck, C. (2010) Exploiting Enzymatic Promiscuity to Engineer a Focused Library of Highly Selective Antifungal and Antiproliferative Aureothin Analogues, *J. Am. Chem. Soc. 132*, 10407-10413.

141. Dalluge, J. J., Gort, S., Hobson, R., Selifonova, O., Amore, F., and Gokarn, R. (2002) Separation and identification of organic acid-coenzyme A thioesters using liquid chromatography/electrospray ionization-mass spectrometry, *Anal. Bioanal. Chem. 374*, 835-840.

142. Vergnolle, O., Xu, H., and Blanchard, J. S. (2013) Mechanism and regulation of mycobactin fatty acyl-AMP ligase FadD33, *J. Biol. Chem. 288*, 28116-28125.

143. Huckin, S. N., and Weiler, L. (1974) Alkylation of Dianions of Beta-Keto-Esters, *J. Am. Chem. Soc. 96*, 1082-1087.

144. Hasdemir, B., Onar, H. C., and Yusufoglu, A. (2012) Asymmetric synthesis of long chain beta-hydroxy fatty acid methyl esters as new elastase inhibitors, *Tetrahedron-Asymmetr 23*, 1100-1105.

145. Webb, M. R. (1992) A Continuous Spectrophotometric Assay for Inorganic-Phosphate and for Measuring Phosphate Release Kinetics in Biological-Systems, *P. Natl. Acad. Sci. USA 89*, 4884-4887.

146. Lloyd, A. J., Thomann, H. U., Ibba, M., and Soll, D. (1995) A Broadly Applicable Continuous Spectrophotometric Assay for Measuring Aminoacyl-Transfer-Rna Synthetase-Activity, *Nucleic Acids Res. 23*, 2886-2892.

147. Tosato, V., Albertini, A. M., Zotti, M., Sonda, S., and Bruschi, C. V. (1997) Sequence completion, identification and definition of the fengycin operon in Bacillus subtilis 168, *Microbiology 143 (Pt 11)*, 3443-3450.

148. Baltz, R. H., Miao, V., and Wrigley, S. K. (2005) Natural products to drugs: daptomycin and related lipopeptide antibiotics, *Nat. Prod. Rep. 22*, 717-741.

149. Konz, D., Doekel, S., and Marahiel, M. A. (1999) Molecular and biochemical

characterization of the protein template controlling biosynthesis of the lipopeptide lichenysin, *J. Bacteriol. 181*, 133-140.

150. Rausch, C., Hoof, I., Weber, T., Wohlleben, W., and Huson, D. H. (2007) Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution, *Bmc. Evol. Biol. 7*.

151. Roongsawang, N., Lim, S. P., Washio, K., Takano, K., Kanaya, S., and Morikawa, M. (2005) Phylogenetic analysis of condensation domains in the nonribosomal peptide synthetases, *Fems. Microbiol. Lett. 252*, 143-151.

152. Steller, S., Sokoll, A., Wilde, C., Bernhard, F., Franke, P., and Vater, J. (2004) Initiation of surfactin biosynthesis and the role of the SrfD-thioesterase protein, *Biochemistry 43*, 11331-11343.

153. Belshaw, P. J., Walsh, C. T., and Stachelhaus, T. (1999) Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis, *Science 284*, 486-489.

154. Lautru, S., and Challis, G. L. (2004) Substrate recognition by nonribosomal peptide synthetase multi-enzymes, *Microbiology 150*, 1629-1636.