Published in final edited form as:

Synlett. 2015 ; 26(8): 1008–1025. doi:10.1055/s-0034-1380264.

***In Vitro* Reconstitution of Metabolic Pathways: Insights into Nature's Chemical Logic**

Brian Lowry^a, Christopher T. Walsh^b, and Chaitan Khosla^{a,b,c}

^aDepartment of Chemical Engineering, Stanford University, 443 Via Ortega, Stanford, CA 94305, USA; blowry@stanford.edu

^bStanford University Chemistry, Engineering, and Medicine for Human Health (ChEM-H), Stanford University, 443 Via Ortega, Stanford, CA 94305

^cDepartment of Chemistry, 333 Campus Drive Mudd Building, Stanford University, Stanford, CA 94305, USA; khosla@stanford.edu

Abstract

In vitro analysis of metabolic pathways is becoming a powerful method to gain a deeper understanding of Nature's core biochemical transformations. With astounding advancements in biotechnology, purification of a metabolic pathway's constitutive enzymatic components is becoming a tractable problem, and such *in vitro* studies allow scientists to capture the finer details of enzymatic reaction mechanisms, kinetics, and the identity of organic product molecules. In this review, we present eleven metabolic pathways that have been the subject of *in vitro* reconstitution studies in the literature in recent years. In addition, we have selected and analyzed subset of four case studies within these eleven examples that exemplify remarkable organic chemistry occurring within biology. These examples serves as tangible reminders that Nature's biochemical routes obey the fundamental principles of organic chemistry, and the chemical mechanisms are reminiscent of those featured in traditional synthetic organic routes. The illustrations of biosynthetic chemistry depicted in this review may inspire the development of biomimetic chemistries via abiotic chemical techniques.

Keywords

Biosynthesis; Enzymes; Antibiotics; Antifungal Agents; Antiviral Agents; Natural Products

1 Introduction

In the quest to better understand chemical reactions in the context of biology, scientists have sought to excise complete enzymatic pathways from their native cellular environments. The hope is that by removing these elements from cells and recapturing their catalytic activities *in vitro*, researchers can apply a variety of analytical tools to study the finer details of their chemical transformations. This concept, which we will define in this review as the *in vitro* reconstitution of metabolic pathways, has existed for over 100 years in the biochemistry

community. Perhaps the earliest example of this concept lies within the study of the fermentation pathways in yeast. In 1897, Eduard Buchner and his colleagues desired a better understanding of how yeast convert sugars into alcohol, for which the underlying chemical principles were a mystery.¹ Shrouded behind scientific contention for many years, it was thought that the fermentation transformations could only take place inside of cells. By methods that would certainly be viewed in this day and age as crude, Buchner used sand to grind up yeast cells and fashioned cheesecloth to serve as an effective membrane by which he could separate and collect an aqueous, enzyme-containing yeast extract. This extract, containing glucose, produced carbon dioxide and ethanol and validated that the cellular machinery within the yeast were responsible for fermentation. Cell-free reconstitution of yeast metabolism was instrumental in identifying the intermediates within glycolysis and the role of phosphate and ATP in driving the fermentation process.²

While these early studies of fermentation demonstrate the power of cell-free analysis of pathways, the means by which contemporary scientists study pathways *in vitro* has greatly changed as technology has advanced. With the advent of recombinant DNA technology and technologies for heterologous expression and purification of enzymes, researchers can now select seemingly any enzymes of interest, produce them in a host cell system, and obtain analytically pure samples of product molecules for testing and analysis. In the context of this review, we will consider *in vitro* reconstitution to be the study of a set of enzymes obtained as pure components through modern protein purification techniques. Furthermore, we will define a pathway as a series of enzyme-catalyzed chemical reactions in which the enzymes catalyze at least four chemical transformations. Throughout this review, we will discuss the *in vitro* reconstitution of 11 representative metabolic pathways originating from bacteria, plants, and fungi.

While the transformations in these pathways occur via biological machinery, it is important to remember that molecules in Nature obey the rules of organic chemistry and follow non-enzymatic mechanisms, and they may do so with exquisite regio- and stereospecificity. The architecture of natural product scaffolds have served for decades as evidence of how structural complexity can be assembled from simple, primary metabolic building blocks and have also served as an inspiration to synthetic chemists for “biomimetic” logic in devising abiotic synthetic routes. Thus, the goals of this review are two-fold. First, we will describe the story behind reconstitution of the selected pathways in a manner that emphasizes the basic chemistry of their enzymatic reactions. By recounting these examples from the literature, we convey the key results from each study and what chemical insights were gained in each case. Second, among these 11 pathways, we have selected four specific examples of remarkable chemical logic for in depth discussion. By bringing forth these particularly interesting cases, we illustrate bridges between biosynthesis and chemical synthesis, possibly inspiring the design of new synthetic chemical routes.

2 Bacterial Metabolites

2.1 Fatty Acids

Fatty acids are ubiquitous within the metabolism of nearly all forms of life. Chemically, these molecules are simple, aliphatic carbon chains possessing a terminal carboxylate group,

and they serve a multitude of essential roles including energy storage, structural components for cellular membranes, and recognition elements for cell-cell communication.³ While the fatty acid biosynthesis pathway serves a central role in primary metabolic processes, it is also heavily integrated into the secondary metabolism of many organisms. Here, we focus our attention on the synthesis of fatty acids within bacteria, which not only use fatty acids as essential precursors for interior cell membrane lipids but also as an anchor for molecules protruding from the exterior cell membrane surface. Specifically, fatty acyl groups derived from fatty acids serve as the hydrophobic moiety of the lipid A endotoxin within the membrane-bound lipopolysaccharide (LPS) of Gram-negative bacteria.⁴ Additional details on the corresponding components of LPS, their function, and details of their biosynthesis will be presented in Section 2.3.

Bacterial fatty acids are constructed in a repetitive fashion from simple metabolic building blocks derived from acetate.^{5,6} However, in the pathway they exist as the chemically activated coenzyme A (CoA) thioesters, acetyl- and malonyl-CoA, abundant throughout bacterial primary metabolism. A set of nine discrete enzymes and a carrier protein collectively form what is known as the bacterial fatty acid synthase (Figure 1A). We first focus on the reactions involving FabD and FabH on the periphery of Figure 1A. The malonyl-CoA: ACP transacylase, FabD, catalyzes the transfer of the malonyl functionality from malonyl-CoA to the acyl carrier protein, ACP. It is important to note that all ACPs are modified at a conserved serine residue with a phosphopantetheine post-translational modification. This so-called phosphopantetheine arm is derived from coenzyme A and contains a nucleophilic thiol which acts as the acyl carrier.⁷ The resulting malonyl-ACP species is ready to participate in a decarboxylative Claisen-like condensation reaction with acetyl-CoA (catalyzed by a ketosynthase, FabH), yielding acetoacetyl-ACP. This species undergoes an NADPH cofactor-dependent reduction at the β -carbon (catalyzed by the ketoreductase, FabG), followed sequentially by a dehydration to yield an α,β -unsaturated acyl-ACP (catalyzed by either one of two dehydratases, FabA or FabZ) and a second NADPH (or sometimes NADH)-dependent reduction to afford butyryl-ACP. Further elaboration of this C₄ intermediate is catalyzed by either one of two ketosynthases (FabB or FabF) via a second decarboxylative condensation reaction that utilizes another molecule of malonyl-ACP. The cycle of extension, reduction, dehydration, and reduction continues in an iterative manner until the fatty acyl-ACP reaches its full length (typically C₁₆, but can be C₁₄ or C₁₈). At that point, the acyl chain may either be transferred to glycerol-3-phosphate (in phospholipid synthesis) or a carbohydrate (in Lipid A synthesis), or it may be hydrolyzed to a free fatty acid (as shown in Figure 1A, by thioesterase TesA).

The decarboxylative Claisen condensation reaction within the fatty acid synthase represents one of the major biosynthetic manifolds for carbon-carbon bond formation in nature and a key chemical transformation that will reappear throughout this review. Figure 1B illustrates the finer catalytic details of FabH and FabB/F during chain extension. A key catalytic feature of the ketosynthase is an active site cysteine bearing a nucleophilic thiol, which can carry reactive intermediates as thioesters. These ketosynthases catalyze two reactions – they promote transfer of the growing acyl chain to their cysteine residues and also collaborate with a nucleophilic malonyl-ACP species to catalyze attack on the enzyme-bound electrophilic thioester, thereby elongating the acyl chain by two carbon atoms. In both

reactions, it is believed that the ketosynthase cradles the reactant(s) deep within its active site, providing protection from unwanted side reactions. We will revisit this mode of carbon-carbon bond formation when analyzing the biosynthesis of polyketides.

Notwithstanding the large body of work to define the detailed chemical mechanisms for fatty acid biosynthesis in bacteria, it was not until recently that the full system had been studied *in vitro*, yielding insights into its kinetics and regulation. Motivated by the desire for a deeper understanding into this prototypical bacterial metabolic pathway, Yu and colleagues reported *in vitro* reconstitution of the fatty acid synthase derived from *E. coli* by overexpressing all nine Fab enzymes and the ACP in the natural *E. coli* host and purifying the enzymes to homogeneity.⁸ Upon supplementing the ten protein species with acetyl-CoA, malonyl-CoA, and NADPH, C₁₄-C₁₈ fatty acid species were observed, initially using ¹⁴C-isotope incorporation experiments and subsequently via UV-spectrophotometry.⁹ Remarkably, under conditions of maximum turnover frequency, the dehydratase FabZ was identified as the principal rate-determining component of the *E. coli* fatty acid synthase, whereas the turnover rate of a cyanobacteria fatty acid synthase was limited by an entirely different component, FabH.¹⁰ The reconstituted multi-enzyme system has also highlighted how subtle changes in relative activities of individual components can substantially influence the partitioning between unsaturated and saturated fatty acid products.⁹ In addition to yielding new insights into the catalytic mechanisms of fatty acid synthases, the reconstituted system could also provide a cell-free platform for antibacterial discovery or for optimizing the production of fatty acid-derived biofuels.^{9,11}

2.2 Farnesene

Isoprenoids are versatile compounds found in prokaryotic and eukaryotic organisms, serving a wide variety of roles within the cell. These molecules include cholesterol, other steroids, defense agents, and cellular pigments.¹² It is estimated that at least 25,000 isoprenoids have been characterized, and society is continuing to appreciate their many beneficial properties.¹² For examples, plant isoprenoids serve as powerful therapeutic and nutritional agents (i.e. artemisinin, paclitaxel, and lycopene).¹³ It has recently become evident that bacteria produce a multitude of isoprenoids with fascinating chemistry and properties.^{14,15,16} Much like fatty acid biosynthesis, the pathways that build isoprenoids can be found in both primary and secondary metabolism. The well-studied mevalonate (MVA) and methylerythritol phosphate (MEP) pathways synthesize critical precursors for constructing complex isoprenoids. Here, we focus on the bacterial MVA pathway (Figure 2A) which has been rerouted to produce a useful biofuel secondary metabolite.

A hallmark of isoprenoid biosynthesis is the use of two 5-carbon precursors to construct larger, more complex molecules. These two simple precursors, known as isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP), are constructed by the MVA pathway using a series of seven enzymes. Featuring acetate as a basic building block, the MVA pathway begins with the condensation of three acetyl-CoA molecules to generate 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA). These condensation reactions are sequentially followed by an NADPH-dependent, four electron reduction to afford mevalonate (thioester to alcohol), two consecutive phosphorylations to give mevalonate-5-

pyrophosphate, and an ATP-dependent decarboxylation to arrive at IPP (or DMAPP if Idi-catalyzed isomerization occurs). At this point, the farnesyl pyrophosphate synthase, IspA, catalyzes two rounds of isoprenoid chain extension involving two molecules of IPP and one molecule of DMAPP generating farnesyl pyrophosphate (FPP). Incorporation of α -farnesene synthase at the end of the MVA pathway diverts FPP to farnesene, releasing free pyrophosphate. The chain extension chemistry and farnesene synthesis will be analyzed further in Section 5.1.

Seeking to develop a route to the jet fuel farnesene, Zhu and colleagues reconstituted the MVA pathway *in vitro*.¹⁷ Similar to fatty acid synthase reconstitution, these researchers expressed the eight MVA pathway enzymes and the α -farnesene synthase in an *E. coli* host. The purified enzymes worked in tandem with the requisite NADPH and ATP cofactors to produce farnesene, as confirmed by gas chromatography-mass spectrometry. Unexpectedly, Idi concentration was found to be most influential on the turnover rate of this pathway. This finding may come as a surprise because previous experiments have pointed to the HMG-CoA reductase catalyzed reaction as the rate-determining step in cholesterol biosynthesis.^{12,18} Furthermore, synergistic effects on the rate were observed for Idi and the HMG-CoA synthase, ERG13, giving an optimal turnover frequency for the whole system of 0.4 s^{-1} . Using the optimized *in vitro* protein ratios as a guide, the researchers developed a metabolically engineered *E. coli* strain that was capable of producing 5.5-fold higher farnesene titers than the unoptimized strain. This work further reinforces the utility in *in vitro* pathway reconstitution as a guide to programming improved *in vivo* biosynthesis of biofuels and related metabolites.

2.3 O-Polysaccharides

The bacterial cell surface can have numerous polysaccharides protruding into the extracellular environment to perform a variety of functions. Perhaps the most interesting of these functions is their critical role in bacterial virulence for which Raetz and Whitfield put forth long-term efforts in understanding the biosynthesis and pathogenesis of lipopolysaccharide (LPS) endotoxins.¹⁹ In Gram-negative bacteria such as *E. coli*, LPS is incorporated into the outer cell membrane and consists of a hydrophobic anchor known as lipid A, a small polysaccharide core, and a large repeating polysaccharide known as an O-polysaccharide antigen. In brief, the lipid A anchor is synthesized by a series of acyltransferase enzymes that catalyze the transfer of β -hydroxymyristoyl, myristoyl, and lauroyl fatty acyl groups onto a short polysaccharide chain before being exported to the periplasmic space.¹⁹ This short chain is further decorated by addition of a so-called polysaccharide core. The O-polysaccharide is synthesized in a separate series of steps and is subsequently ligated to the lipid A – core species. The completed LPS is exported to the outer cell membrane by a specialized transporter. While the detailed steps for construction of the entire LPS are beyond the scope of this review, we will focus on the steps involved in O-polysaccharide biosynthesis.

The O-polysaccharide is constructed in an assembly line fashion, for which the biosynthetic machinery exerts exquisite control over sugar regiochemistry.²⁰ Polysaccharide biosynthesis is initiated on a surrogate carrier unit, undecaprenyl phosphate (und-PP) that spans the inner

cell membrane, facing the interior cytoplasm of the bacterial cell (Figure 3). Here, the carrier is sequentially loaded with 5 sugars, the first of which is N-acetyl-D-galactosamine (GalNAc). An additional GalNAc unit is added via an α -1,3 glycosidic linkage catalyzed by WbnH. Next, WbnJ catalyzes the addition of galactose (Gal) via a β -1,3 linkage, followed by WbnK-catalyzed addition of fucose (Fuc) via an α -1,2 linkage to galactose. Finally, WbnI catalyzes the transfer of a second Gal by an α -1,3 linkage to the first Gal, giving a branched pentasaccharide (RU-PP-Und). A “flippase” known as Wzx translocates the chain to the periplasmic face of the membrane. Here, a polymerase, Wzy, is believed to catalyze repetitive condensation of pentasaccharides. Polysaccharide chain length is controlled by Wzz, a co-polymerase protein.

Motivated by the goal of gaining a better understanding of Wzy and Wzz from *E. coli*, Woodward and colleagues reconstituted *in vitro* the O-polysaccharide biosynthetic pathway by overcoming significant experimental hurdles. While the purification, isolation, and *in vitro* analysis of WbnJ, WbnK, and WbnI had been reported previously,²¹ isolation of the integral membrane protein Wzy had proven challenging because of the loss of integrity of this protein outside of its native membrane environment. Circumventing the problem by expressing a truncated version of this enzyme in *E. coli*, the researchers obtained pure and active Wzy. Another challenge was the lack of an *in vitro* assay to analyze Wzy and Wzz reaction products. To remedy this problem, these researchers devised an elegant ³H-incorporation assay that could be readily monitored by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). In this assay, ³H-RU-PP-Und was generated using the Wbn enzymes. Polymerized polysaccharides of various length could then be separated by SDS-PAGE and visualized by autoradiography, and the key insight gained was that Wzy alone produced a wide range of chain lengths indicating stochastic control over chain length. Upon addition of Wzz however, the chain length became more narrowly defined. Interestingly, Wzz proteins from different organisms produced drastically different SDS-PAGE band patterns. The *in vitro* system established here has laid the groundwork for answering outstanding question relating to protein interactions between Wzy and Wzz and the mechanistic basis for chain length control by Wzz. It has also been used to probe the mechanism of WaaL, a ligase protein that links the O-polysaccharide to the lipid A core.²²

3 Plant Metabolites

Plants produce an extraordinary array of secondary metabolites (flavonoids, alkaloids, polyketides, glucosides, to name a few).²³ Once thought to be waste products of plant primary metabolism, work over the last 50 years has demonstrated their utility as defense compounds, pigments, and signaling molecules.^{23,24} In plants, defense compounds are analogous to an innate immune system, as they provide protection from infectious agents. Plant-derived drugs have been used for many decades including well-known compounds such as the opiate pain reliever morphine and the antimalarial drug quinine.²⁵ Even within the last two decades, numerous plant-derived compounds have received FDA approval, including the alkaloids galantamine (used to treat early-onset Alzheimer’s disease) and varenicline (used to aid smoking cessation).²⁵ Here we will focus on two plant metabolites, dhurrin and camalexin which are examples of glucosides and alkaloids, respectively.

3.1 Dhurrin

Cyanogenic glucosides (CG) are produced in cyanogenic plants, and serve a variety of roles in plant physiology, including defense.²⁶ Perhaps the most studied CG in terms of its biosynthesis is dhurrin, produced by the plant *Sorghum bicolor*. Dhurrin is derived from L-tyrosine and glucose which are incorporated into the final product via a series of enzyme catalyzed oxidations. L-tyrosine is oxidized by the sequential action of two cytochrome P450 monooxygenases (Figure 4A). Cytochrome P450 enzymes represent a class of extremely versatile catalysts that perform a wide array of chemistries; they are a central theme in plant metabolite biosynthesis. The first enzyme, CYP79, performs an interesting multistep oxidation. Two successive hydroxylation reactions (consuming molecular oxygen and NADPH) and a dehydration give (Z)-p-hydroxyphenylacetaldoxime ((Z)-p-HAO). The second cytochrome P450, P450ox, is believed to catalyze a dehydration and an additional NADPH-dependent oxidation (a hydroxylation leading to the penultimate compound, phydroxymandelonitrile (p-HMAN)). The dehydration reaction is unusual in that it has been empirically shown to be NADPH-dependent. A glucosyltransferase is responsible for installing the sugar moiety in the final step. While the CYP79 enzyme had been studied previously, very little was known about the P450ox enzyme and how it bridges the gap between (Z)-p-HAO and dhurrin.

In order to probe more deeply into the action of the P450ox enzyme, Kahn and colleagues reconstituted *in vitro* the entire dhurrin biosynthetic pathway using enzymes from the natural host organism.²⁷ While isolation of CYP79 and glucosyltransferase were relatively straightforward, P450ox purification proved challenging as the protein is naturally associated with the cell membrane. Through painstaking efforts, the researchers were able to obtain P450ox in the microsomal fraction of the *S. bicolor* lysates. The microsomal environment afforded catalytically active CYP79 and P450ox, and dhurrin synthesis was observed through radioactive TLC analysis when combining the three enzymes, ¹⁴C-tyrosine, UDP-glucose, and NADPH. Aside from demonstrating that these components were necessary and sufficient for dhurrin production, the researchers showed that P450ox catalyzes two steps and identified the nitrile intermediate, p-hydroxyphenylacetone nitrile (p-HPAN), in the overall transformation. While the mechanistic details of P450ox catalysis remain unknown, this work illustrates the power of reconstituting a multi-step plant metabolic pathway *in vitro*. It also sets the stage for probing the potential role of complex formation between CYP79 and P450ox.²⁸

3.2 Camalexin

Alkaloids represent a very broad class of plant secondary metabolites and much like glycosides are best known for their defense properties against herbivores and pathogens.²⁹ Including aforementioned drugs such as morphine and quinine, as well as caffeine and nicotine, alkaloids are nitrogen-containing heterocyclic compounds which are derived from amino acid building blocks.²⁹ In this section, we will focus on the biosynthesis of the alkaloid camalexin, a secondary metabolite from *Arabidopsis thaliana*. Camalexin is also classified as a phytoalexin (a compound produced by plants in response to pathogenic stress), and is a key defense mechanism against fungal infections.^{30,31} Camalexin is constructed from L-tryptophan and L-cysteine. Its biosynthesis features a series of

cytochrome P450 oxidations akin to those observed in the dhurrin pathway (Figure 4B). In the initial step, CYP79B2 catalyzes decarboxylation and N- hydroxylation of tryptophan to indole-3- acetaldoxamine (IAOx). A second P450 enzyme is believed to catalyze an oxidative coupling of cysteine to IAOx, and the resulting cysteine-indole-3-acetonitrile (Cys-IAN) compound is decarboxylated and cyclized by the action of CYP71A15 to form the thiazole ring structure within camalexin. The hallmark of the camalexin pathway is the unusual C-S bond forming reaction prior to cyclization of the thiazole moiety, which makes identification of the second P450 enzyme of considerable interest.

The identity of the second enzyme in the pathway and its mechanism was unknown until recently. By using a combination of gene expression data and protein sequence analysis, Klein and coworkers were able to not only identify a P450 enzyme capable of performing the C-S coupling reaction but were also able to reconstitute the entire camalexin pathway *in vitro*.³² Much like the study involving the dhurrin pathway, all P450 enzymes were isolated from the microsomal fraction of the expression host (in this case, *Saccharomyces cerevisiae*). An enzyme known as CYP71A13 was identified as the missing second P450 in the pathway and the key player in facilitating C-S bond formation. Starting with the IAOx precursor, it was observed that CYP71A13 catalyzed the formation of Cys-IAN in the presence of L-cysteine and NADPH. However, it was also observed that glutathione could also function as a thiol donor when presented in comparable concentrations as cysteine and that rate of IAOx consumption was similar in both cases. The researchers reasoned that the P450 enzyme may not directly catalyze the C-S bond formation but rather act to generate a reactive electrophile that can serve as a partner for the nucleophilic thiol substrate. It was then shown that CYP79B2, CYP71A13, and CYP71A15 were necessary and sufficient for *in vitro* synthesis of camalexin. These results, however, do not rule out the possibility that other enzymes, for instance a glutathione transferase enzyme, may be needed to optimize the coupling reaction and protect reactive intermediates. The study sheds light on a long-standing mystery relating to the interesting C-S coupling chemistry in plant biosynthesis. In addition, this work highlights the power of modern genetic and bioinformatic tools for locating proverbial “needle in a hay stack” enzymes that are responsible for chemical transformations of interest.

4 Polyketides and Nonribosomal Peptides

The remainder of our case studies will focus on polyketides and nonribosomal peptides which are large classes of natural product metabolites with incredible structural diversity and remarkable medicinal utility. Both types of molecules can be found in numerous species of plants, bacteria, and fungi, but a common theme is that these molecules are used as biochemical weaponry to ward off predators and pathogens or to gain an edge over neighboring organisms for space and resources.³³ Humans have made use of polyketides for therapeutic purposes, including antibiotics (erythromycins, tetracyclines), antifungals (amphotericin), antitumor agents (geldnamycin), and cholesterol lowering agents (lovastatin).^{34,35,36} Clinically relevant nonribosomal peptides (and their nonribosomal peptide – polyketide hybrid counterparts) also serve as antibiotics (penicillins, cephalosporins), immunosuppressants (rapamycin, FK506), and anticancer agents

(epothilones).^{37,38} For this review, we will focus on a set of five polyketides and one nonribosomal peptide whose biosynthesis has been reconstituted *in vitro* (Figure 5A).

Polyketide and nonribosomal peptide biosynthesis occurs by the action of successive condensations between simple metabolic building blocks in a manner similar to that of fatty acid biosynthesis. Polyketides are assembled by enzymatic complexes known as polyketide synthases (PKSs). PKSs can vary widely in their overall architecture, as they can exist as discrete domains (Section 4.1), a single modular polypeptide (Section 4.2), and even multimodular assembly lines (Section 4.3). While their architectures are different, the same series of chemical transformations are conserved. For our purposes, PKSs accept simple metabolic building blocks derived from either acetate or propionate, which are accepted by the synthases as activated CoA thioesters, malonyl- or methylmalonyl-CoA. As in the fatty acid synthase, core catalytic components of the PKS are a ketosynthase (KS), an acyltransferase (AT), and an acyl carrier protein (ACP), which carry out three essential chemical reactions: acyl-CoA selection, acyl transfer, and a decarboxylative Claisen-like condensation reaction (Figure 5B). The AT selects the appropriate acyl-CoA extender unit (either a malonyl or methylmalonyl unit), and catalyzes transfer of this acyl functionality to a phosphopantetheine arm of the ACP via an enzyme-bound serine-ester intermediate. Meanwhile, the KS catalyzes transfer of its electrophilic substrate onto its active site cysteine via trans-thioesterification. This sets the stage for the KS and the ACP collaborate to catalyze C-C bond formation, thereby elongating the polyketide chain by one extender unit. Further chemical and stereochemical diversity at the β -carbonyl position can be generated by so-called tailoring reactions catalyzed by different combinations of ketoreductase (KR), dehydratase (DH), and enolreductase (ER) enzymes. The intermediate is then transferred back to the same KS or onto a different KS for additional rounds of chain processing until the chain is released. As will be seen in Sections 4 and 5, chain release often occurs by fascinating cyclization mechanisms, and can lead to even further diversification of the finished polyketide product.

Nonribosomal peptides are synthesized by enzymes with analogous functions but with strikingly different chemistry. Nonribosomal peptides are synthesized by enzyme complexes called nonribosomal peptide synthetases (NRPSs), which consist of enzymatic domains arranged in multimodular assembly lines. Figure 5C depicts the basic catalytic transformations of NRPSs, whose building blocks are amino acids. The core catalytic components of the NRPS are a condensation (C) domain, an adenylation (A) domain, and a peptidyl carrier protein (PCP). Analogous to a PKS's AT domain, the NRPS A domain is responsible for selecting the appropriate amino acid extender unit. However, rather than directly forming an acyl intermediate, the A domain first activates the carboxylate via adenylation, and then transfers the activated building block to the phosphopantetheine arm of the PCP (an identical posttranslational modification to that required for ACPs). The C domain then facilitates C-N bond formation between two aminoacyl-PCP moieties, in effect catalyzing an S $\rightarrow$ N shift of the growing peptide chain onto the amino acid building block. Following chain growth, tailoring and cyclization can also impart chemical diversity on the final peptide product, but greatest factor contributing to nonribosomal peptide diversity is

nature's spectacular pool of amino acid building blocks.³⁹ We will now explore the fascinating chemistry of PKSs and NRPSs in the context of six examples.

4.1 Aromatic Polyketides

4.1.1 Actinorhodin-derived Polyketides—Soil bacteria synthesize a variety of polyketide structures, including aromatic polyketides such as oxytetracycline, tetraneomycin, and actinorhodin, all of which serve as antibiotics.^{40,36} Aromatic PKSs operate by a mechanism nearly identical to that of the bacterial fatty acid synthase by using a set of discrete enzymes to catalyze decarboxylative Claisen condensations in a repetitive fashion.³⁴ Much work over the years by several laboratories has contributed to our present-day understanding of aromatic PKS biosynthesis by soil bacteria.^{41,42,43} Here, we will focus on the enzymatic machinery from the actinorhodin PKS (Figure 6). The core components of the actinorhodin PKS consist of a malonyl-CoA:ACP transacylase (MAT), an ACP, a KS, and a chain length factor (CLF).⁴⁴ The MAT, ACP, and KS collaborate to catalyze extender unit selection, transacylation, and chain extension as described above. Analogous to Wzz in O-antigen biosynthesis, the CLF specifies the number of chain extension cycles, thereby controlling polyketide chain length. While the actinorhodin PKS lacks tailoring domains, chemical diversity arises at the point of cyclization and release. Intramolecular aldol condensations and dehydrative aromatizations drive the cyclization pattern of the nascent polyketide intermediate, leading to the shunt metabolites, SEK4 and SEK4b.⁴⁵

While aromatic polyketides had been produced *in vivo*, *in vitro* reconstitution of their PKSs had not been observed until Carreras and colleagues reconstituted the actinorhodin PKS from its purified components.⁴⁶ Using ¹⁴C-labeled malonyl-CoA, SEK4 and SEK4b were observed by radio-TLC analysis. Notably, the MAT utilized by the actinorhodin PKS is shared by the fatty acid synthase in the natural host; this dual functionality has been observed for many aromatic PKSs. Perhaps the most interesting finding in this work was the discovery that the KS and CLF proteins formed a heterodimer. Its identification paved the way for further mechanistic analyses of how the CLF controls polyketide chain length⁴⁷ and also enabled *in vitro* reconstitution of the pathway to more advanced intermediates in actinorhodin synthesis.^{48,49}

4.1.2 Enterocin—Enterocin is a related but considerably more complex aromatic polyketide whose biosynthesis has been extensively investigated by Moore and coworkers. Unlike the previous example from a soil bacterium, enterocin is part of a vast collection of marine natural products derived from ocean-dwelling bacteria.^{50,51,52} The early stages of enterocin biosynthesis are nearly identical to the actinorhodin PKS with seven rounds of chain elongation catalyzed by the core PKS components. However, the enterocin PKS is primed with benzoic acid. Enterocin formation also requires a regiospecific ketoreductase that reduces the C-9 carbonyl on the full-length polyketide chain. The hallmark of the enterocin pathway is EncM, a unique flavoprotein, which catalyzes the cyclization and aromatization of the linear polyketide chain. The fascinating mechanism of EncM will be further described in depth in Section 5.3. Following cyclization, a sequential methyl transfer and hydroxylation lead to the final product.

Using nine enzymes from the enterocin pathway, Cheng and colleagues reconstituted enterocin synthesis *in vitro* and uncovered a remarkable chemical mechanism in the process.⁵³ HPLC product analysis determined that all nine enzymes were necessary and sufficient to produce enterocin. It was also observed that omission of EncM led to shunt products reminiscent of the SEK compounds produced by the actinorhodin PKS enzymes, which provided strong evidence the EncM plays a major role in cyclization. Previous *in vivo* work had hypothesized EncM's key role in enterocin cyclization, implicating EncM as a putative "Favorskiiase" enzyme catalyzing chain cyclization in a manner reminiscent of Alexei Favorskii's rearrangement mechanism within organic synthesis.^{54,55} The enzyme was observed to be flavin-dependent and likely acts upon the EncC-bound linear polyketide chain. In addition, EncK and EncR were shown to be the SAM-dependent methyltransferase and the cytochrome P450, respectively, which complete the enterocin pathway. This work opened the door for an elegant analysis of the detailed mechanism of EncM, which we will revisit in depth later.⁵⁶

4.2 Fungal Polyketides

4.2.1 Norsolorinic Acid—While it has not been emphasized thus far, eukaryotic organisms also produce polyketides by similar chemistries to that of their bacterial counterparts. Fungi produce polyketides that possess clinically relevant antifungal (griseofulvin), immunosuppressive (mycophenolic acid), and cholesterol lowering properties (lovastatin).⁵⁷ Others, such as the mycotoxin aflatoxin B1, are extremely potent environmental toxins.⁵⁷ For the purposes of this discussion, fungal PKSs can be classified into two varieties by their degree of reduction: non-reducing PKSs (NR-PKSs) or highly-reducing PKSs (HR-PKSs). Organizationally, fungal PKSs are reminiscent of aromatic bacterial PKSs in that their enzymatic domains perform iterative cycles of chain extension. However, fungal PKS domains collectively exist as a single polypeptide. This modular protein houses all of the core enzymatic domains (KS, MAT, and ACP) while also containing all tailoring, patterning, and release domains. The defining feature of the NR-PKSs versus HR-PKSs is that NR-PKS lack tailoring domains that reduce the polyketide β -carbonyl. This subtle difference in domain combinations imparts a profound impact on the final molecular structure. Our first fungal polyketide example is norsolorinic acid, a precursor to aflatoxin B1, which is synthesized by a NR-PKS known as PksA (Figure 7A). Extensive work by Townsend and coworkers has led to the definition of PksA as a prototypical NR-PKS.⁵⁸ In its native form, PksA exists as a single polypeptide containing six domains. Polyketide synthesis begins with selection of a hexanoyl priming unit derived from hexanoyl-CoA by the starter unit:ACP transacylase (SAT). The primer undergoes ten chain extension reactions with malonyl-CoA. Cyclization is followed by post-PKS reactions to convert norsolorinic acid into aflatoxin A1.

Crawford and coworkers reconstituted the activity of PksA *in vitro* after painstaking work to express and purify the functional enzyme.⁵⁹ While PksA exists as a single, six-domain polypeptide, the researchers were only able to recover soluble protein when the single polypeptide was genetically deconstructed into multiple proteins. Segregation at the MAT-PT junction gave two soluble proteins that produced norsolorinic acid in the presence of hexanoyl-CoA and methylmalonyl-CoA. The identity of the PT (product template) domain

was previously unknown, and bioinformatic analysis revealed that it bore no homology with any characterized PKS domains. Based on its proximity to the ACP, it was hypothesized that the PT controlled chain cyclization by stabilizing the poly- β -ketone intermediates in specific conformations. Indeed, experiments with deconstructed PksA variants lacking the PT domain were unable to yield norsolorinic acid, rather giving shunt products. Thus, the PT serves to protect the growing chain during chain elongation and to act as an aromatase enzyme by facilitating regiospecific cyclization. In addition to throwing light on the nature of the PT domain, *in vitro* reconstitution of this pathway also laid the groundwork for applying advanced proteomic approaches to interrogate the active site occupancy within PksA.⁶⁰

4.2.2 Dihydromonacolin L—The cholesterol-lowering drug lovastatin is derived from a compound, dihydromonacolin L, produced by a HR-PKS in *Aspergillus terreus*. Fungal HR-PKSs share the same modular architecture and iterative catalysis as NR-PKSs, but contain additional tailoring domains to further modify the β -carbon. Over the last 5 years, Tang and coworkers have made impressive strides toward understanding HR-PKS enzymology in the context of the lovastatin synthase. The lovastatin PKS, LovB, contains the same basic core domains as PksA, but its full complement of tailoring domains leads to a highly-reduced polyketide core that cyclizes into a remarkably different chemical structure (Figure 7B). Biosynthesis of the nonaketide dihydromonacolin L is accomplished by a multidomain polypeptide (LovB) which accepts ten malonyl-CoA building blocks. LovB also contains a set of four tailoring domains: a DH domain, a SAM-dependent methyltransferase (MT) domain, an inactive ER domain, and a KR domain. In addition to the tailoring domains within LovB, a stand-alone *trans*-ER domain (LovC) acts to reduce the enoyl intermediates after the third, fourth, and sixth rounds of chain elongation. A feature of the final polyketide that draws a chemist's interest is the decalin ring. Most remarkably, the hexaketide intermediate of dihydromonacolin L undergoes an intramolecular Diels-Alder reaction to form this moiety of the final product. The details of this interesting biochemical Diels-Alder chemistry will be discussed more in depth in Section 5.4.

Ma and colleagues described *in vitro* reconstitution of dihydromonacolin L biosynthesis using LovB, LovC, and a thioesterase domain and in the process were able to elucidate the critical role of LovC.⁶¹ This study was made possible only upon finding a suitable host for expression and purification of LovB. Since the natural host *A. terreus* was unable to yield useful amounts of enzyme, a modified strain of *S. cerevisiae* was employed to obtain increased yield of LovB with high activity. Upon obtaining LovB, LovC and a heterologous thioesterase were added to afford dihydromonacolin L. An interesting finding was that LovC required full activity of the *cis*-tailoring domains in order to produce the expected final product. In particular, it was noted that in the absence of SAM (required for MT activity), LovC activity failed, leading to derailment of truncated ACP-bound polyketides. Thus, it was hypothesized that LovC might act as a mechanistic gatekeeper, permitting only the correct dihydromonacolin L precursor to proceed through all rounds of chain processing. This study was instrumental in developing a platform for studying LovC the identity of other modifying enzymes and intermediates along the path to lovastatin.^{62,63}

4.3 Assembly Line Polyketides and Nonribosomal Peptides

4.3.1 6-deoxyerythronolide B—In addition to iterative PKSs, bacteria are known to harbor enormous, multimodular PKSs that synthesize their products in an assembly line fashion.^{35,64} Multimodular PKSs build their polyketides by sequential chemical transformations in a manner that is reminiscent of the automobile assembly line. Here we focus on DEBS, the most-well characterized multimodular PKS, which produces the aglycone core, 6-deoxyerythronolide B (6-dEB), to the erythromycin antibiotics.^{65,66} For over two decades, the laboratories of Khosla, Cane, and Leadlay have investigated this enzymatic assembly line. DEBS consists of three polypeptides, DEBS1, DEBS2, and DEBS3 and each polypeptide contains a set of two modules (Figure 8A). Each module possesses domains that catalyze a single round of chain elongation and processing. The final product is cyclized and released with the aid of a thioesterase. Multimodular PKSs accept acyl-CoAs as priming and extender units (the most common extender units being malonyl- and methylmalonyl-CoA) and also consume reducing equivalents in the form of NADPH. The assembly line architecture facilitates the so-called vectorial processing of the growing polyketide chain as it is channeled sequentially from one active site to the next. In this regard, a unique feature of multimodular PKSs are docking domains which are depicted as black tabs flanking the DEBS proteins (Figure 8A). These coiled-coil structural motifs are critical for correct matching of non-covalently interacting modules.^{67,68} Other types of protein-protein interactions have proven critical in nearly all catalytic steps within DEBS (acyl-CoA transacylation, chain elongation and chain translocation).^{69,70,71}

Despite two decades of research, *in vitro* reconstitution of the entire synthase from purified protein components was not demonstrated until recently. Seeking a better understanding of the kinetic parameters and substrate specificity of the native assembly, Lowry and colleagues were able to demonstrate *in vitro* production of 6-dEB.⁷² The primary hurdle in reconstituting DEBS was that DEBS1 expressed poorly in several heterologous hosts making purification of reasonable amounts of enzyme intractable. To circumvent this problem, DEBS1 was deconstructed into three proteins, the loading didomain, module 1, and module 2, which could all be expressed and purified. By monitoring the NADPH consumption of complete DEBS, it was shown that maximal 6-dEB turnover is approximately 1 min⁻¹. This measurement served as validation of protein activity, as this number is consistent with estimates of 6-dEB productivity in the natural host organism, *S. erythraea*. Furthermore, the researchers sought to probe the substrate specificity of the AT domains in the context of a full assembly line by presenting DEBS with both methyl- and ethylmalonyl-CoA. Surprisingly, it was confirmed that AT3 could accept an ethylmalonyl unit leading to a 8-ethyl-8-desmethyl 6-dEB analog. This observation was unprecedented, as isolated AT domains have been shown in several studies to have strict specificity for their native CoA substrates and to require protein engineering for specificity alteration.⁷³ The reconstitution of DEBS will open avenues to further study the kinetics and mechanism of an intact PKS assembly line that had not been possible. In addition, the reconstituted system forms the basis of a test ground for incorporation of non-natural CoA extender units into 6-dEB, generating novel polyketides.

4.3.2 Asperlicin—Nonribosomal peptides are another class of incredibly diverse and useful natural product compounds, and in particular, they are well known for their utility as antibiotics.³³ Nonribosomal peptide synthetases are similar in architecture to multimodular PKS in that they are assembly lines.^{37,74} However, nonribosomal peptide chemical diversity is rooted in the extender units, which are amino acids rather than acyl-CoAs. NRPSs contain additional features in common with multimodular PKSs in that they channel intermediates in a vectorial manner prior to chain release. For many years, the laboratories of Walsh and Marahiel have studied all aspects of nonribosomal peptide biosynthesis; *in vitro* reconstitution of NRPSs is elegantly outlined in a previous review.⁷⁴ Here, we will focus on a recently discovered fungal NRPS known as the asperlicin synthetase from *Aspergillus alliaceus* (Figure 8B). An assembly line made up of only two modules, the asperlicin NRPS (AspA) produces two structural forms of asperlicin by catalyzing three rounds of chain processing which will be discussed below. AspA accepts two molecules of anthranilate (Ant) and one molecule of L-tryptophan to build the peptide core. The PCP-bound tripeptide then undergoes cyclization and release by the action of the C_T domain to form asperlicin C and D.

The use of two modules to catalyze the addition of three amino acids may seem counterintuitive based on our discussion NRPS biochemistry. However, this anomaly within asperlicin biosynthesis was elucidated through the work of Gao and colleagues by reconstituting AspA *in vitro*.⁷⁵ The asperlicin scaffold would suggest that the synthetase processes two molecules of Ant followed by one molecule of L-tryptophan. Consequentially, assays of A domain-catalyzed acylation revealed that the preferred substrates of A1 and A2 (the A domains of modules 1 and 2, respectively) are Ant and L-tryptophan, respectively. Upon showing that purified AspA in combination with its amino acid substrates and ATP were necessary and sufficient to produce asperlicin C and D, the researchers employed precursor feeding studies to determine how exactly the two module AspA can conduct three chain extensions. When PCP2 (the PCP within module 2) was loaded with a L-Trp-Ant dipeptide substrate, addition of Ant, A1, and PCP1 led to the formation of asperlicin, which provided strong evidence that the C domain is capable of catalyzing a second amide bond between dipeptide's aniline NH₂ group and the second molecule of Ant. The cyclization, catalyzed through collaboration of C_T and PCP2, was hypothesized to proceed through a mechanism in which the terminal aniline group serves as the nucleophile for thioester displacement. Once cyclized, the cyclic tripeptide can then differentiate into asperlicin C or D by one of two distinct cyclodehydration patterns. This work elegantly outlines the how AspA behaves as a unique NRPS in performing a non-traditional iterative behavior in an enzymatic assembly line. Furthermore, this work showcases how nature has devised a remarkably short pathway to build a complex, nitrogen-containing heterocyclic scaffold. An analogous abiotic route would certainly be advantageous for synthetic chemistry.

5 Discussion of Chemical Insights

5.1 Allylic Carbocation Chemistry Drives Farnesene Synthesis

As mentioned in our discussion of the MVA pathway, the isoprenoid chain extension reaction involving IPP and DMAPP represents another canonical methodology by which nature orchestrates carbon-carbon bond formation. These enzymatic reactions take advantage of forming stabilized allylic carbocations, which are central to many reactions in organic synthesis such as Friedel-Crafts alkylation chemistry and carbocation-based Nazarov reactions.⁷⁶ Catalyzed by prenyltransferase enzymes (in this case, IspA, the farnesyl-pyrophosphate synthase), these head-to-tail alkylative chain extension reactions take advantage of an allylic carbocation as an electrophile, matching it with a corresponding nucleophilic pi bond. The reactions are best depicted mechanistically as an S_N1 process. Figure 9A illustrates the IPP / DMAPP condensation reaction in which an allylic carbocation is formed from DMAPP. IPP's pi bond attacks the cation leading to extension of the isoprenoid chain. It is believed that a general base within IspA acts to remove an α -proton to generate geranyl-pyrophosphate (GPP). By an analogous set of chemical transformations, IspA catalyzes a second round of chain extension between GPP and IPP to yield FPP. At this point, AFS catalyzes an interesting variation on allylic carbocation chemistry (Figure 9B) to generate farnesene. An interesting study by Faraldos *et al.* methodically builds a proposed mechanism by which AFS acts as an "eliminase" enzyme, taking advantage of the reactivity of an allylic farnesyl carbocation.⁷⁷ It was shown that pyrophosphate loss is not a concerted, one-step process. Rather, the pyrophosphate is lost (generating a carbocation) but then collapses back to the 3-position generating a nerolidyl pyrophosphate (NPP) intermediate. Pyrophosphate is eliminated by abstraction of an adjacent proton, leading to the final degree of unsaturation in farnesene. The final step in AFS catalysis is reminiscent (although not identical) to intramolecular Friedel-Crafts reactions reported in the literature.⁷⁸ Biosynthesis of farnesene serves as an excellent example of how enzymes use allylic carbocation chemistry to conduct different chemical transformations which are not unfamiliar in the world of synthetic organic chemistry.

5.2 Multifunctional Oxidation Chemistry of Cytochrome P450 Enzymes

As alluded to earlier, cytochrome P450 enzymes are one of nature's largest and most versatile classes of biological catalysts. Much like an organic chemist calling upon the Woodward reaction to perform a *cis*-hydroxylation or selecting a peroxy-acid to perform an olefin epoxidation, nature calls upon P450s to catalyze a range of reactions including hydroxylations, epoxidations, and dehydrogenations.⁷⁹ The mechanism of these enzymes has been analyzed in detail since the early 1970s when Minor Coon's laboratory successfully reconstituted P450 activity in microsomes obtained from cell lysates.⁸⁰ While a complete analysis of the P450 catalytic mechanism is beyond the scope of this review, all P450 enzymes share a catalytic core which contains an iron-heme cofactor, which is the epicenter of the catalytic cycle.⁸¹ In brief, the cycle can be summarized in several key steps: binding of the small molecule substrate, reduction of the heme's Fe^{III} atom to Fe^{II} by accepting electrons from a reductase partner, capture of molecular oxygen by Fe^{II} , reduction of oxygen by a second electron transfer leading to loss of water, and finally insertion of an activated oxygen atom into the small molecule substrate. Multiple comprehensive reviews

elegantly describe the finer details and experimental history of studying these enzymes.^{80,81,79} The dhurrin biosynthetic pathway's CYP79 is a good illustration of P450 hydroxylation chemistry (Figure 10). While the overall transformation takes L-tyrosine to (Z)-p- hydroxyphenylacetaldoxime ((Z)-p-HAO), three intermediates are generated via P450 along this path.⁸² By the canonical P450 mechanism, two successive N-hydroxylation reactions consume molecular oxygen and NADPH to produce N-hydroxytyrosine (N-HTY) and subsequently N,N-dihydroxytyrosine (N,N-DTY). N-DTY is then dehydrated and decarboxylated to give (E)-p-HAO, and an isomerization leads to (Z)-p-HAO. It is believed that the dehydration and decarboxylation proceed by non-enzymatic means although further investigation could show that CYP79 has even greater biocatalytic versatility.

5.3 A Remarkable “Favorskiiase” Enzyme with in Enterocin Biosynthesis

Before the turn of the 20th century, Alexei Favorskii discovered a reaction in which α -halo ketones could rearrange into carboxylic acids.⁵⁵ This reaction was later termed the Favorskii rearrangement in his honor and has become a convenient route for preparing branched carboxylic acids and ring contracted compounds. In a typical Favorskii rearrangement, a base is used to generate an enolate that can eliminate the α -halogen in an intramolecular fashion. The subsequent product is a ring-strained carbonyl that can react with a second molecule of base giving a carboxylic acid within the rearranged scaffold. This type of transformation is not confined to the chemical synthetic realm, as the enterocin biosynthetic pathway is believed to have a Favorskii-like rearrangement in its mechanism catalyzed by EncM. While the enterocin scaffold lacks halogens, a Favorskii rearrangement is entirely consistent with formation of the final product (Figure 11).⁵⁴ The full extended polyketide chain bound to EncC (carbon chain depicted in red) undergoes a dual oxidation across the C4-C5 positions to give intermediate **1**. It is likely at this point that EncM cradles the polyketide chain in the proper orientation to facilitate tautomerization of the C1 enol and generate a strained ring-strained ketone at C3 (intermediate **2**). The Favorskii rearrangement is completed when the C7-OH attacks the C3 carbonyl to give a rearranged cyclic moiety (intermediate **3**, a highly strained cyclopropanone). Proving itself to have great versatility, EncM also catalyzes two aldol condensations (C2-C9 and C5-C10) and a heterocyclization (C15-O11) within intermediate **4** to arrive at desmethyl-5-deoxyenterocin. This example illustrates how nature has beautifully mimicked a canonical reaction within organic synthesis along the pathway to building a highly complex natural product.

5.4 Diels-Alder Chemistry Appears in Biology

Our final case study is appearance of one of the most useful synthetic chemical reactions within biology. Otto Diels and Kurt Alder made their landmark discovery in 1928 and later were awarded the Nobel Prize in Chemistry in 1950, and since then, the Diels-Alder reaction has become a simple yet indispensable chemical route for [4+2] cycloadditions.⁸³ The well-known reaction mechanism is the coupling of a diene to a dienophile which leads to cycloaddition and C-C bond formation.⁸⁴ Diels-Alder chemistry increased greatly in popularity when Woodward and colleagues incorporated it into chemical synthesis routes to create biomolecules like the steroids cortisone and cholesterol.⁸³ Interestingly, nature has seemingly followed the same logic when designing reactions as Woodward, as examples biosynthetic Diels-Alder reactions have begun to emerge in the literature.⁸⁵ The lovastatin

PKS provides an excellent example of a Diels-Alder reaction playing a major role in fashioning the final polyketide. Recall that LovB catalyzes 8 rounds of chain elongation to build dihydromonacolin L. Upon priming with malonyl-CoA, the first two chain extensions are followed by ketoreduction and dehydration leading to a formation of a diene. The third and fourth chain extensions are followed by ketoreduction, dehydration, and enoyl reduction giving diene-containing acyl pentaketide (Figure 12). One more round of chain extension and dehydration generates the dieneophile which is uniquely positioned to pair with the corresponding diene, leading to cyclization of the decalin moiety. This example also illustrates the power of chain tailoring within polyketide biosynthesis, as the installation of additional functional groups allows for even greater structural diversity to arise by non-enzymatic means.

6 Concluding Remarks

Through visiting examples from the literature and specific molecular case studies, we illustrate that the importance and utility of *in vitro* reconstitution of metabolic pathways is becoming self-evident. Following in the footsteps of biochemists in the late 19th century, modern chemists, biochemists, and chemical engineers have gained an appreciation for the wealth of experiments and great insights that can be derived from such studies. Armed with modern technologies for chemical analysis, the researchers conducting the studies in our eleven examples have been able to probe deeply into the enzymology and chemical mechanisms involved in synthesizing complex natural products using relatively simple, repetitive biological pathways. In particular, each example has emphasized the major role that modern molecular biology techniques (in particular, recombinant DNA technology and heterologous protein expression) in studying enzymatic pathways. We can now reliably determine product identity and accurately measure kinetics by applying tools like HPLC, LC/MS, radioisotope analysis, and UV-Vis spectrometry to *in vitro* reactions. It is noteworthy that through rigorous searches of the literature, it is evident that very few pathways from primary metabolism have been reconstituted and studied in full. For example, the nucleotide biosynthetic pathways and tricarboxylic acid cycle could serve as excellent starting points for reconstitution studies. Such pathways are at the heart of metabolism and involve a wealth of interesting chemical transformations. *In vitro* studies could yield valuable insights into chemical mechanisms, enzyme kinetics, and possibly drug development and discovery. Above all, we hope that the final four case studies have reinforced the idea that nature obeys the same rules of organic chemistry that are obeyed by thousands of synthetic chemical routes used daily in organic synthesis laboratories. Nature should serve as a powerful inspiration for the advancement of organic synthesis in terms of designing new routes and merging biosynthetic and synthetic routes.

Acknowledgment

This research was supported by a grant from the NIH (R01 GM087934) to C.K.

References

1. Barnett J, Lichtenthaler F. Yeast. 2001; 18:363. [PubMed: 11223946]
2. Barnett JA. Yeast. 2003; 20:509. [PubMed: 12722184]

3. Berg, JM.; Tymoczko, JL.; Stryer, L. Biochemistry. 7th ed.. New York, NY, USA: W.H. Freeman and Company; 2012.
4. Raetz CRH, Reynolds CM, Trent MS, Bishop RE. Annu. Rev. Biochem. 2007; 76:295. [PubMed: 17362200]
5. Rock CO, Jackowski S. Biochem. Biophys. Res. Commun. 2002; 292:1155. [PubMed: 11969206]
6. Cronan, JE.; Thomas, J. Methods in Enzymology. Vol. 459. Elsevier Inc; 2009. p. 395-433.
7. Pfeifer BA, Khosla C. Microbiol. Mol. Biol. Rev. 2001; 65:106. [PubMed: 11238987]
8. Yu X, Liu T, Zhu F, Khosla C. Proc. Natl. Acad. Sci. U. S. A. 2011; 108:18643. [PubMed: 22042840]
9. Xiao X, Yu X, Khosla C. Biochemistry. 2013; 52:8304. [PubMed: 24147979]
10. Kuo J, Khosla C. Metab. Eng. 2014; 22:53. [PubMed: 24395007]
11. Liu R, Zhu F, Lu L, Fu A, Lu J, Deng Z, Liu T. Metab. Eng. 2014; 22:10. [PubMed: 24333607]
12. Voet, D.; Voet, JG. Biochemistry. 3rd ed.. Hoboken, NJ, USA: John Wiley & Sons, Inc; 2004. p. 942-950.
13. Kirby J, Keasling JD. Annu. Rev. Plant Biol. 2009; 60:335. [PubMed: 19575586]
14. Cane DE, Ikeda H. Acc. Chem. Res. 2012; 45:463. [PubMed: 22039990]
15. Citron CA, Gleitzmann J, Laurenzano G, Pukall R, Dickschat JS. Chembiochem. 2012; 13:202. [PubMed: 22213220]
16. Ajikumar P, Tyo K, Carlsen S. Mol. Pharm. 2008; 5:167. [PubMed: 18355030]
17. Zhu F, Zhong X, Hu M, Lu L, Deng Z, Liu T. Biotechnol. Bioeng. 2014; 9999:1.
18. Goldstein J, Brown M. Nature. 1990; 343:425. [PubMed: 1967820]
19. Raetz CRH, Whitfield C. Annu. Rev. Biochem. 2002; 71:635. [PubMed: 12045108]
20. Whitfield C, Larue K. Nat. Struct. Mol. Biol. 2008; 15:121. [PubMed: 18250628]
21. Yi W, Shao J, Zhu L, Li M, Singh M, Lu Y, Lin S, Li H, Ryu K, Shen J, Guo H, Yao Q, Bush CA, Wang PGJ. Am. Chem. Soc. 2005; 127:2040.
22. Han W, Wu B, Li L, Zhao G, Woodward R, Pettit N, Cai L, Thon V, Wang PGJ. Biol. Chem. 2012; 287:5357.
23. Wink M. Annu. Plant Rev. 2010; 40:1.
24. Hartmann T. Phytochemistry. 2007; 68:2831. [PubMed: 17980895]
25. Salim AA, Chin Y, Kinghorn AD.
26. Vetter J. Toxicon. 2000; 38:11. [PubMed: 10669009]
27. Kahn RA, Bak S, Svendsen I, Halkier BA, Maller BL. Plant Physiol. 1997; 115:1661. [PubMed: 9414567]
28. Møller B, Conn E. J. Biol. Chem. 1980; 255:3049. [PubMed: 7358727]
29. Facchini P. Annu. Rev. Plant Biol. 2001; 52:29.
30. Ahuja I, Kissen R, Bones AM. Trends Plant Sci. 2012; 17:73. [PubMed: 22209038]
31. Møldrup ME, Geu-Flores F, Halkier BA. Plant Cell. 2013; 25:360. [PubMed: 23449503]
32. Klein AP, Anarat-Cappillino G, Sattely ES. Angew. Chem. Int. Ed. Engl. 2013; 52:13625. [PubMed: 24151049]
33. Walsh, CT. Antibiotics: Actions, Origins, Resistance. Washington, D.C., USA: ASM Press; 2003.
34. Carreras C, Pieper R, Khosla C. Top. Curr. Chem. 1997; 188:85.
35. Staunton J, Weissman KJ. Nat. Prod. Rep. 2001; 18:380. [PubMed: 11548049]
36. Hopwood DA. Chem. Rev. 1997; 97:2465. [PubMed: 11851466]
37. Fischbach MA, Walsh CT. Chem. Rev. 2006; 106:3468. [PubMed: 16895337]
38. Walsh CT. Science. 2004; 303:1805. [PubMed: 15031493]
39. Walsh CT, O'Brien RV, Khosla C. Angew. Chem. Int. Ed. Engl. 2013; 52:7098. [PubMed: 23729217]
40. Katz L, Donadio S. Annu. Rev. Microbiol. 1993; 47:875. [PubMed: 8257119]
41. Hopwood D, Sherman D. Annu. Rev. Genet. 1990; 24:37. [PubMed: 2088174]

42. McDaniel R, Ebert-Khosla S, Hopwood DA, Khosla C. *Science*. 1993; 262:1546. [PubMed: 8248802]
43. Shen B, Hutchinson CR. *Science*. 1993; 262:1535. [PubMed: 8248801]
44. Das A, Khosla C. *Acc. Chem. Res.* 2009; 42:631. [PubMed: 19292437]
45. Fu H, Ebert-Khosla S, Hopwood DA, Khosla C. *J. Am. Chem. Soc.* 1994; 116:4166.
46. Carreras C, Khosla C. *Biochemistry*. 1998; 37:2084. [PubMed: 9518007]
47. Tang Y, Tsai S, Khosla C. *J. Am. Chem. Soc.* 2003; 125:12708. [PubMed: 14558809]
48. Zawada RJX, Khosla C. *J. Biol. Chem.* 1997; 272:16184. [PubMed: 9195917]
49. Zawada RJ, Khosla C. *Chem. Biol.* 1999; 6:607. [PubMed: 10467128]
50. Akey DL, Gehret JJ, Khare D, Smith JL. *Nat. Prod. Rep.* 2012; 29:1038. [PubMed: 22498975]
51. Moore BS, Kalaitzis JA, Xiang L, Antonie Van Leeuwenhoek. 2005; 87:49. [PubMed: 15726291]
52. Watts KR, Tenney K, Crews P. *Curr. Opin. Biotechnol.* 2010; 21:808. [PubMed: 20956079]
53. Cheng Q, Xiang L, Izumikawa M, Meluzzi D, Moore BS. *Nat. Chem. Biol.* 2007; 3:557. [PubMed: 17704772]
54. Xiang L, Kalaitzis JA, Moore BS. *Proc. Natl. Acad. Sci. U. S. A.* 2004; 101:15609. [PubMed: 15505225]
55. Akhrem AA, Ustynyuk TK, Titov YA. *Russ. Chem. Rev.* 1970; 39:732.
56. Teufel R, Miyana A, Michaudel Q, Stull F, Louie G, Noel JP, Baran PS, Palfey B, Moore BS. *Nature*. 2013; 503:552. [PubMed: 24162851]
57. Simpson TJ, Cox RJ. *Wiley Encycl. Chem. Biol.* 2008
58. Crawford JM, Townsend CA. *Nat. Rev. Microbiol.* 2010; 8:879. [PubMed: 21079635]
59. Crawford JM, Thomas PM, Scheerer JR, Vagstad AL, Kelleher NL, Townsend CA. *Science*. 2008; 320:243. [PubMed: 18403714]
60. Vagstad AL, Bumpus SB, Belecki K, Kelleher NL, Townsend CA. *J. Am. Chem. Soc.* 2012; 134:6865. [PubMed: 22452347]
61. Ma SM, Li JW-H, Choi JW, Zhou H, Lee KKM, Moorthie VA, Xie X, Kealey JT, Da Silva NA, Vederas JC, Tang Y. *Science*. 2009; 326:589. [PubMed: 19900898]
62. Xu W, Chooi Y-H, Choi JW, Li S, Vederas JC, Da Silva NA, Tang Y. *Angew. Chem. Int. Ed. Engl.* 2013; 52:6472. [PubMed: 23653178]
63. Meehan MJ, Xie X, Zhao X, Xu W, Tang Y, Dorrestein PC. *Biochemistry*. 2011; 50:287. [PubMed: 21069965]
64. Khosla C, Tang Y, Chen AY, Schnarr NA, Cane DE. *Annu. Rev. Biochem.* 2007; 76:195. [PubMed: 17328673]
65. Cortes J, Haydock S, Roberts G. *Nature*. 1990; 348:176. [PubMed: 2234082]
66. Donadio S, Staver M, McAlpine J. *Science*. 1991; 252:675. [PubMed: 2024119]
67. Gokhale RS, Tsuji SY, Cane DE, Khosla C. *Science*. 1999; 284:482. [PubMed: 10205055]
68. Broadhurst RW, Nietlispach D, Wheatcroft MP, Leadlay PF, Weissman KJ. *Chem. Biol.* 2003; 10:723. [PubMed: 12954331]
69. Wong FT, Chen AY, Cane DE, Khosla C. *Biochemistry*. 2010; 49:95. [PubMed: 19921859]
70. Kapur S, Chen A, Cane D, Khosla C. *Proc. Natl. Acad. Sci. U. S. A.* 2010; 107:22066. [PubMed: 21127271]
71. Kapur S, Lowry B, Yuzawa S, Kenthirapalan S, Chen AY, Cane DE, Khosla C. *Proc. Natl. Acad. Sci. U. S. A.* 2012; 109:4110. [PubMed: 22371562]
72. Lowry B, Robbins T, Weng C-H, O'Brien RV, Cane DE, Khosla C. *J. Am. Chem. Soc.* 2013; 135:16809. [PubMed: 24161212]
73. Dunn BJ, Khosla CJR. *Soc. Interface*. 2013; 10:20130297. [PubMed: 23720536]
74. Sattely ES, Fischbach MA, Walsh CT. *Nat. Prod. Rep.* 2008; 25:757. [PubMed: 18663394]
75. Gao X, Jiang W, Jiménez-Osés G, Choi MS, Houk KN, Tang Y, Walsh CT. *Chem. Biol.* 2013; 20:870. [PubMed: 23890005]
76. Naredla RR, Klumpp DA. *Chem. Rev.* 2013; 113:6905. [PubMed: 23819438]

77. Faraldos JA, Gonzalez V, Li A, Yu F, Koksall M, Christianson DW, Alleman RAJ. Am. Chem. Soc. 2012; 134:20844.
78. Nicolaou KC, Kang Q, Wu TR, Lim CS, Chen DY-K. J. Am. Chem. Soc. 2010; 132:7540. [PubMed: 20462209]
79. Meunier B, de Visser SP, Shaik S. Chem. Rev. 2004; 104:3947. [PubMed: 15352783]
80. Coon MJ. Annu. Rev. Pharmacol. Toxicol. 2005; 45:1. [PubMed: 15832443]
81. Denisov IG, Makris TM, Sligar SG, Schlichting I. Chem. Rev. 2005; 105:2253. [PubMed: 15941214]
82. Sibbesen O, Koch B, Halkier BA, Moller BL. J. Biol. Chem. 1995; 270:3506. [PubMed: 7876084]
83. Nicolaou KC, Snyder SA, Montagnon T, Vassilikogiannakis G. Angew. Chemie Int. Ed. 2002; 41:1668.
84. Woodward R, Katz T. Tetrahedron. 1959; 5:70.
85. Stocking EM, Williams RM. Angew. Chem. Int. Ed. Engl. 2003; 42:3078. [PubMed: 12866094]

Biographies

Brian Lowry was born in Indianapolis, Indiana in 1988. He earned a B.S. in Chemical Engineering with Honors from Purdue University in 2010, graduating with highest distinction. He is currently pursuing a Ph.D. in Chemical Engineering at Stanford University under the direction of Chaitan Khosla and aims to complete his degree in 2015. During his time as a graduate student at Stanford, he was awarded a National Science Foundation Graduate Research Fellowship in 2011. His doctoral work has focused on studying the *in vitro* biosynthesis of 6-deoxyerythronolide B, the core precursor to the erythromycin family of antibiotics.

Christopher T. Walsh, born in 1944, studied biology at Harvard and completed his PhD in biochemistry under Fritz Lipmann at The Rockefeller University. From 1972 to 1987 he held a faculty position at MIT and later moved on to continue his work at Harvard Medical School. He served as Chair of the Department of Chemistry at MIT from 1978 to 1982 and of the Department of Biological Chemistry & Molecular Pharmacology at Harvard Medical School from 1987 to 1995. In 2014, he was appointed a Consulting Professor of Chemistry at Stanford University and a member of the Stanford Institute for Chemistry, Engineering, and Medicine for Human Health (ChEM-H). His research interests focus on the molecular logic and enzymatic machinery of peptide-based natural product biosynthesis.

Chaitan Khosla, born in 1964, earned a B.S. in Chemical Engineering from the Indian Institute of Technology (Bombay) and later completed a Ph.D. in Chemical Engineering at the California Institute of Technology in 1990 under the guidance of James Bailey. Upon completing his postdoctoral studies under David Hopwood at the John Innes Centre, he joined the Stanford University in 1992, where he is currently a Professor of Chemistry, Chemical Engineering, and (by courtesy) Biochemistry, and the Wells H. Rauser and Harold M. Petiprin Professor in the School of Engineering. He is also the founding director of Stanford ChEM-H. In 1999, he was awarded the Alan T. Waterman Award by the National Science Foundation, and in 2007 and 2009, he was elected to the American Academy of Arts and Sciences and the National Academy of Engineering, respectively. His research interests include the biosynthesis and mechanism of multimodular, assembly line polyketide synthases and the pathogenesis and treatment of celiac disease.

Figure 1.

A *Reconstitution of the Fatty Acid Synthase from *E. coli*.* Fatty acid synthesis is performed using the Fab enzymes, metabolic precursors (malonyl-CoA and acetyl-CoA), and a cofactor, NADPH. ACP, acyl-carrier protein; FabD, malonyl-CoA: ACP transacylase; FabH and FabB/F, ketosynthase; FabG, ketoreductase; FabA/Z, dehydratase; FabI enolreductase; TesA, thioesterase. **B** *Fatty Acid Chain Extension Chemistry.* The mechanisms for the ketosynthase catalyzed reactions of FabH and FabB/F proceed via the same chemistry. (I) represents the initial thioester transfer reaction of an acyl chain onto the ketosynthase's active site cysteine, and (II) represents the decarboxylative Claisen condensation that leads to extension of the growing fatty acyl chain.

Figure 2.

Reconstitution of the Mevalonate Pathway from E. coli for Farnesene Production. In vitro biosynthesis of farnesene is accomplished by using nine purified enzymes. HMG-CoA, 3-hydroxy-3-methylglutaryl-CoA; IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; FPP, farnesyl pyrophosphate; AtoB, acetoacetyl-CoA thiolase; ERG13, HMG-CoA synthase; HMGR, HMG-CoA reductase; ERG12, mevalonate kinase; ERG8, phosphomevalonate kinase; MVD1, mevalonate pyrophosphate decarboxylase; Idi, IPP isomerase; IspA, FPP synthase; AFS, farnesene synthase.

Figure 3.

Reconstitution of the E. coli O-polysaccharide Pathway. Biosynthesis of O-polysaccharides of defined chain lengths is catalyzed by the sequential action of four glycosyl transferases (WbnH, WbnJ, WbnK, and WbnI), a sugar polymerase (Wzy), and a chain length modality protein (Wzz). A flippase (Wzx) catalyzes translocation of the growing chain from the inner membrane to the outer membrane prior to polymerization. UDP, uridine diphosphate; RU-PP-Und, repeating unit pentasaccharide-PP-Und GDP; guanosine diphosphate; red - N-acetyl-D-galactosamine; blue - galactose; green - fucose (figure adapted from Woodward et. al *Nat. Chem. Biol.* **2010** 6, 418).

Figure 4.

A *Reconstitution of Dhurrin Biosynthesis from Sorghum bicolor*. Biosynthesis of dhurrin is accomplished by two cytochrome P450 enzymes and a glycosyltransferase. Z-p-HAO, (Z)-p-hydroxyphenyl-acetaldoxamine; p-HPAN, p-hydroxyphenylacetonitrile; p-HMAN, p-hydroxymandelonitrile. **B** *Reconstitution of Camalexin Biosynthesis from Arabidopsis thaliana*. Biosynthesis of camalexin occurs through the sequential activity of three cytochrome P450 enzymes. The putative intermediates for the oxidation of IAOx catalyzed by CYP71A13 and nucleophilic cysteine addition are shown in braces. IAOx, Indole-3-acetaldoxamine; Cys-IAN, cysteine-indole-3-acetonitrile.

Figure 5.

A *Polyketide and Nonribosomal Peptide Molecules to be Featured in Section 4.* These chemical structures represent the products of six biosynthetic pathways which have all been reconstituted *in vitro*. **B** *The Basic Mechanism for Polyketide Biosynthesis.* The basic steps in the catalysis of the core PKS enzymes are acyl-extender unit selection, extender unit transacylation, and chain extension. KS, ketosynthase; AT, acyltransferase; ACP, acyl carrier protein. **C** *The Basic Mechanism for Nonribosomal Peptide Biosynthesis.* The basic steps in the catalysis of the core NRPS enzymes are amino acid selection, adenylation to activate the amino acid substrate, transacylation of the amino acid, and peptide bond formation. A, adenylation domain; PCP, peptidyl carrier protein; C, condensation domain.

Figure 6.

A *Reconstitution of Aromatic Polyketide Biosynthesis from *Streptomyces coelicolor*.*

Biosynthesis of actinorhodin-derived aromatic polyketides is catalyzed by 3 discrete enzymes to impart actinorhodin shunt products, SEK4 and SEK4b. MAT, malonyl-CoA:ACP transacylase; ACP, acyl carrier protein, KS/CLF, ketosynthase/chain-length factor. **B** *Reconstitution of Enterocin Biosynthesis from *Streptomyces maritimus*.*

Biosynthesis of enterocin features an aromatic PKS and three additional enzymes. EncA, ketosynthase; EncB, chain-length factor; EncC, holo-ACP; FabD, malonyl-CoA:ACP transacylase from *S. glaucescens*; EncN, benzoate:ACP ligase; EncD, ketoreductase; EncM, "favorskiiase" flavoprotein; EncK, methyltransferase; EncR, cytochrome P450 hydroxylase.

Figure 7.

A *Reconstitution of Norsolorinic Acid Biosynthesis using the Fungal Polyketide Synthase, PksA.* *In vitro* biosynthesis of norsolorinic acid is demonstrated by using deconstructed variants of PksA. PksA activity was successfully reconstituted by segregating the natural protein into two proteins at the MAT-PT junction. PksA is primed with a hexanoyl fatty acyl unit and catalyzes seven iterative rounds of chain elongation. SAT, starter unit acyl carrier protein transacylase; KS, ketosynthase; MAT, malonyl-CoA:ACP transacylase; PT, product template; ACP, acyl carrier protein; TE/CLC, thioesterase/cyclization domain. **B** *Reconstitution of Dihydromonacolin L Biosynthesis using the Lovastatin Fungal Polyketide Synthase.* The lovastatin precursor dihydromonacolin L is synthesized by the modular enzyme LovB along with requisite the LovC domain and a heterologous thioesterase. KS, ketosynthase; MAT, malonyl-CoA:ACP transacylase; DH, dehydratase; ER⁰, enoyl reductase (inactive); KR, ketoreductase; ACP, acyl carrier protein; CON, condensation domain homolog; ER, enoyl reductase (active); TE, thioesterase.

Figure 8.

A *In vitro* Reconstitution of the 6-Deoxyerythronolide B Synthase. Biosynthesis of 6-deoxyerythronolide B is conducted using five enzymes derived from the multimodular DEBS PKS. The domains corresponding to each PKS module are shown in a distinct color to emphasize the assembly line architecture. The native DEBS1 protein was split at the LDD-Module 1 and Module 1-Module 2 interfaces to afford soluble, active proteins. LDD, loading didomain; KS, ketosynthase; AT, acyltransferase; ACP, acyl carrier protein; KR, ketoreductase; DH, dehydratase; ER, enoyl reductase; TE, thioesterase. **B** *In vitro* Reconstitution of the Asperlicin Nonribosomal Peptide Synthetase. The asperlicin synthetase, AspA, was reconstituted using a single bimodular enzyme. Modules 1 and 2, shown as distinct colors, utilize Ant and L-trp extender units, respectively. Module 1 was discovered to use two units of Ant in an interative fashion before extension by module 2. Ant, anthranilate; L-trp, L-tryptophan; A, adenylation domain; PCP, peptidyl carrier protein; C, condensation domain; CT, terminal condensation domain.

Figure 9.

A *Allylic Carbocation Chemistry Involved in IspA Catalysis.* The prenyltransferase, IspA, catalyzes the condensation of DMAPP and IPP to extend the isoprenoid chain, generating GPP. **B** *Elimination of Pyrophosphate in AFS Catalysis.* AFS takes advantage of allylic carbocation stability to eliminate pyrophosphate and drive the formation of farnesene. IPP, isopentenyl pyrophosphate; DMAPP, dimethylallyl pyrophosphate; GPP, geranyl pyrophosphate; FPP, farnesyl pyrophosphate; IspA, FPP synthase; AFS, farnesene synthase.

Figure 10.

Cytochrome P450 Oxidation Chemistry - CYP79. CYP79 converts L-tyrosine to (Z)-p-HAO via three intermediates along a path of cytochrome P450 reactions. N-HTY, N-hydroxytyrosine; N,N-DTY, N,N-dihydroxytyrosine; (E)-p-HAO, (E)-p-hydroxyphenyl-acetaldoxime; (Z)-p-HAO, (Z)-p-hydroxyphenyl-acetaldoxime.

Figure 11.

EncM, A Biological “Favorskiiase”. *EncM* facilitates the Favorskii-like rearrangement in enterocin biosynthesis by performing a dual oxidation of the initial EncC-bound polyketide chain to give **1**. Intermediates **2** and **3** represent the hypothesized Favorskii rearrangement products, and **3** undergoes two aldol condensations and a heterocyclization to give desmethyl-5-deoxyenterocin. The initial polyketide’s backbone carbon chain is shown in red (figure adapted from Teufel et. al. *Nature* **2013** 503, 552).

Figure 12.

Diels-Alder Chemistry within the Lov PKS. Following chain extension and tailoring of the acyl-pentaketide, the acyl-hexaketide product generated during Lov PKS chain extension possesses a diene and a dienophile which react by a [4+2] cycloaddition to give a decalin moiety. ACP, acyl carrier protein; DH, dehydratase; ER, enoyl reductase.