JOC Perspective

pubs.acs.org/joc

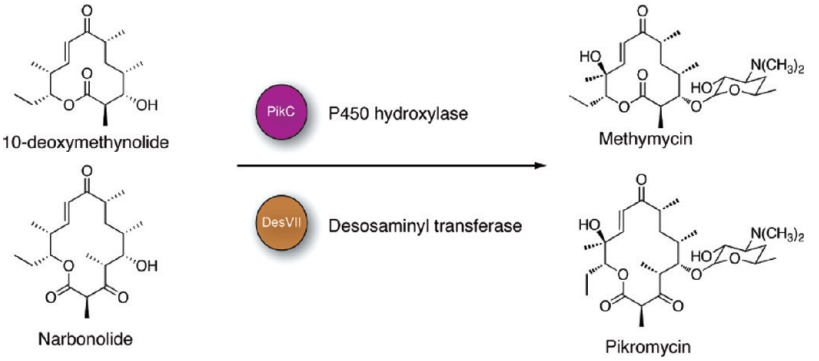
Frontiers and Opportunities in Chemoenzymatic Synthesis

Jonathan D. Mortison and David H. Sherman*

Life Sciences Institute, Department of Chemistry and Department of Medicinal Chemistry, 210 Washtenaw Avenue, University of Michigan, Ann Arbor, Michigan 48109

davidhs@umich.edu

Received June 24, 2010



Natural product biosynthetic pathways have evolved enzymes with myriad activities that represent an expansive array of chemical transformations for constructing secondary metabolites. Recently, harnessing the biosynthetic potential of these enzymes through chemoenzymatic synthesis has provided a powerful tool that often rivals the most sophisticated methodologies in modern synthetic chemistry and provides new opportunities for accessing chemical diversity. Herein, we describe our research efforts with enzymes from a broad collection of biosynthetic systems, highlighting recent progress in this exciting field.

Introduction

Polyketides, nonribosomal peptides, and their hybrids constitute large classes of structurally diverse natural products that possess a wealth of pharmacological activities including antimicrobial, antimycotic, antiparasitic, antitumor, and immunosuppressive, making them extremely valuable lead compounds in drug discovery and development. In fact, over two-thirds of newly introduced drugs worldwide in the past two decades have been natural products or derivatives thereof.[1] However, clinical development of many of these promising candidate drugs is challenged by difficulties in large-scale compound production, as natural product isolates are often low-yielding and their structural complexities typically make total synthesis a limited option. Furthermore, difficulty in generating analogue libraries of a parent structure limits the effectiveness of SAR studies to modulate the desired properties of the lead compound. Thus, in order to efficiently identify next-generation drugs from natural product scaffolds, it has become imperative to explore new methods for rapid generation of structurally diverse compound libraries.

To address this issue, modern efforts to expand access to chemical diversity have increasingly employed multidisciplinary tools and strategies. From these tools, promising chemoenzymatic approaches to generate natural products have emerged, particularly through the manipulation of enzymes from biosynthetic pathways of secondary metabolites. These strategies leverage the unique selectivity and catalytic power of these enzymes, which are responsible for producing countless natural products in marine and terrestrial microorganisms. These methods may utilize a large assembly line of enzymes to produce a full natural product scaffold such as 6-deoxyerythronolide B (6-DEB) (Figure 1a) or a single enzyme to effect a difficult chemical transformation, such as the selective P450-catalyzed oxidation of a C−H bond (Figure 1b). In this Perspective, we present some of the most recent advances in chemoenzymatic synthesis and identify current trends and opportunities in this growing field. Particular attention will be centered on chemoenzymatic oxidation (e.g., hydroxylation and epoxidation) and chain termination (e.g., macrocyclization) strategies.

Polyketide Synthases and Nonribosomal Peptide Synthetases

Many enzymes utilized in chemoenzymatic synthesis come from natural product biosynthetic pathways. The building of polyketide and nonribosomal peptide natural products occurs on large megaenzyme complexes called polyketide synthases[2] (PKS) and nonribosomal peptide synthetases[3]

DOI: 10.1021/jo101124n Published on Web 09/30/2010

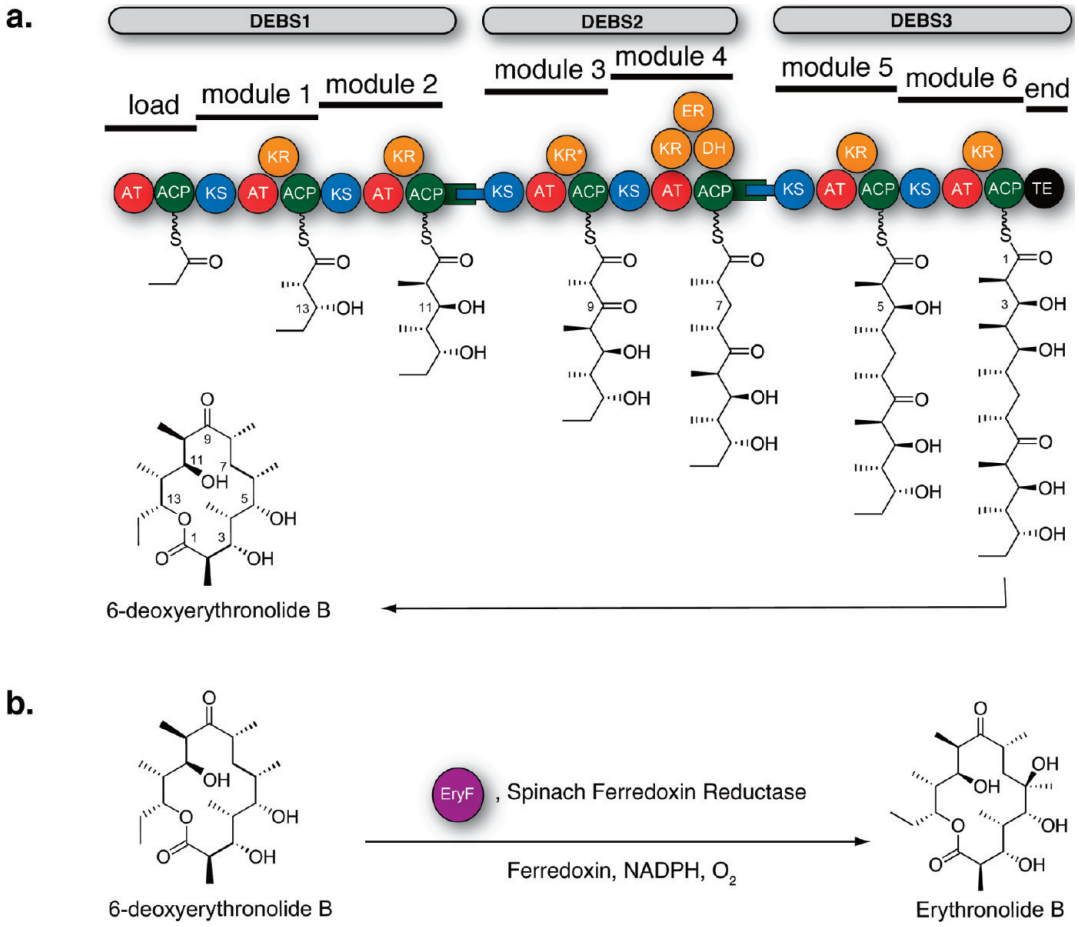


FIGURE 1. Illustration of (a) the DEBS PKS responsible for biosynthesis of the erythromcycin aglycon 6-deoxyerythronolide B (6-DEB) and (b) in vitro chemoenzymatic oxidation of 6-DEB to erythronolide B by the EryF P450 hydroxylase: KS = ketosynthase, AT = acyltransferase, ACP = acyl carrier protein, KR = ketoreductase, DH = dehydratase, ER = enoyl reductase, TE = thioesterase. * denotes an inactive domain.

(NRPS), respectively. In type I PKS systems, a modular assembly of enzymatic domains links simple malonyl-CoA derivatives via sequential decarboxylative Claisen condensations to build complex polyketide chains (Figure 2a). Modules from a type I PKS are responsible for a single elongation step in building the growing polyketide chain, and a minimal module consists of three core domains: a ketosynthase (KS), an acyltransferase (AT), and an acyl carrier protein (ACP). During elongation, the AT domain selects an appropriate acyl-CoA extender unit and loads it onto the ACP. The KS domain then catalyzes decarboxylation of the extender unit to generate a transient enolate nucleophile that reacts in a Claisen condensation with the growing polyketide chain to give an extended β-keto acyl-ACP intermediate. Following extension, the intermediate can then undergo various levels of processing by reductive enzyme domains (i.e., ketoreductases, dehydratases, and enoyl reductases) to introduce functional and stereochemical diversity to the polyketide backbone. The architecture of these PKS systems is such that the collinear arrangement of the modules encodes the chemical structure of the corresponding polyketide product. As a result, rational engineering of these PKS enzymes provides an exciting avenue for design of new polyketide compounds.

In NRPS systems, enzymes are again organized into modules containing a similar trio of core domains: an adenylation domain (A), a condensation domain (C), and a peptidyl carrier protein (PCP) (Figure 2b). The A domain selects an amino acid monomer, activates the carboxyl group as an *O*-AMP ester, and loads it onto the PCP. The C domain then catalyzes formation of the new peptide bond between the PCP-loaded monomer and the growing peptidyl intermediate. Similar to PKS systems, the collinear nature of the NRPS modules leads to products with predictable amino acid sequences. While many of these biosynthetic systems are composed exclusively of PKS or NRPS modules, some contain hybrid PKS/NRPS modules, generating products that contain both polyketide and peptidyl units in their structures.

Following PKS and NRPS extension and processing, the intermediates are often converted to macrocyclic ring scaffolds via a terminal cyclization domain or off-loaded as linear structures through alternative chain termination events.[4,5] The final bioactive compounds are then obtained by further tailoring of these scaffolds by enzymes such as P450 hydroxylases and glycosyl transferases.

Macrocyclization

In their biologically relevant form, many PKS and NRPS-derived natural products contain macrocyclic rings that equal or exceed eight atoms, making them challenging targets for synthetic chemists. In generating macrocyclic lactones[6] and lactams by synthetic methods, common hurdles

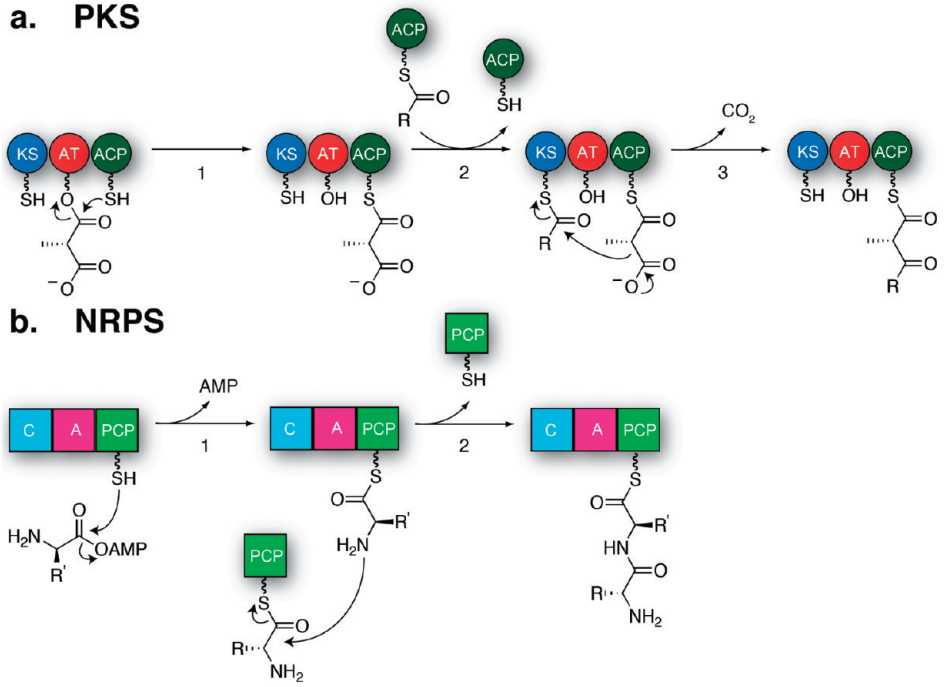


FIGURE 2. Illustration of (a) PKS mechanism: (1) AT loading of the ACP with a methylmalonate extender unit, (2) loading of the KS with the pendant polyketide intermediate from the upstream ACP domain, and (3) KS-catalyzed decarboxylative Claisen condensation of the extender unit with the polyketide intermediate. Illustration of (b) NRPS mechanism: (1) loading of the AT activated *O*-AMP amino acid onto the PCP, (2) C domain-catalyzed condensation of the PCP-loaded amino acid with the upstream PCP-bound peptidyl intermediate.

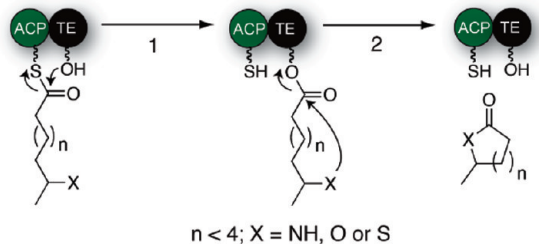


FIGURE 3. Illustration of TE-mediated macrocyclization: (1) transfer of final intermediate to TE active site serine and (2) regioselective intramolecular cleavage of the acyl enzyme intermediate by an internal nucleophile.

include the need for orthogonal protecting group strategies, complex conformational effects (entropic and enthalpic factors), regioselectivity, and competing intermolecular dimerizations/oligomerizations. These factors often lead to extended synthetic routes and decreased yields of target compounds.

In PKS and NRPS systems, cyclization is often catalyzed by a discrete terminal thioesterase (TE) domain, containing a characteristic serine, histidine, aspartate catalytic triad found in the homologous serine hydrolases. During this process, the final biosynthetic intermediate from the last PKS or NRPS elongation module is passed to the TE active site serine residue to form an *O*-acyl enzyme intermediate, followed by cleavage of the acyl enzyme by regiospecific nucleophilic attack of an internal hydroxyl or amino group to produce the macrocyclic lactone or lactam (Figure 3). Due to the unique ability of these TE enzymes to efficiently catalyze macrocyclization, many efforts have explored their potential as biocatalysts when excised from their native PKS or NRPS multifunctional protein context.

P450 Hydroxylation and Epoxidation

Another set of exceptional challenges for synthetic organic chemists are the selective hydroxylation and epoxidation of unactivated C−H bonds and olefins, respectively, by traditional methodologies. Mild oxidation of organic compounds, however, can be accomplished biocatalytically by the P450 superfamily of enzymes, which are heme-containing proteins that couple with a reductase partner protein and ferredoxin cofactor to activate molecular oxygen using NADPH/NADH.[7] Specifically, P450 enzymes from natural product biosynthetic pathways are capable of effecting many difficult oxidative transformations on natural product scaffolds in both a regio- and stereospecific manner, making them attractive biocatalysts for chemoenzymatic synthesis.

Cryptophycin

The cryptophycins are a class of hybrid PKS/NRPS natural products isolated from a pair of related lichen-derived cyanobacterial symbionts, *Nostoc* sp. ATCC 53789 and GSV 224.[8,9] These compounds are exceptionally potent tubulin-depolymerizing agents and are not active substrates of P-glycoprotein (P-gp) or multiple drug resistance-associated protein (MRP),[10] making them viable chemotherapeutic alternatives against cancers that overexpress both of these transporters and are resistant to the vinca alkaloids and paclitaxel. Recently, synthetic crytophycin analogue LY355703 (cryptophycin-52) was discontinued as a candidate in phase II clinical trials due to dose-limiting toxicity resulting in peripheral neuropathy.[11] Nevertheless, these natural products still garner significant interest in their development as anticancer drugs,[12] with current efforts aimed toward generating cryptophycins that eliminate these undesired effects.

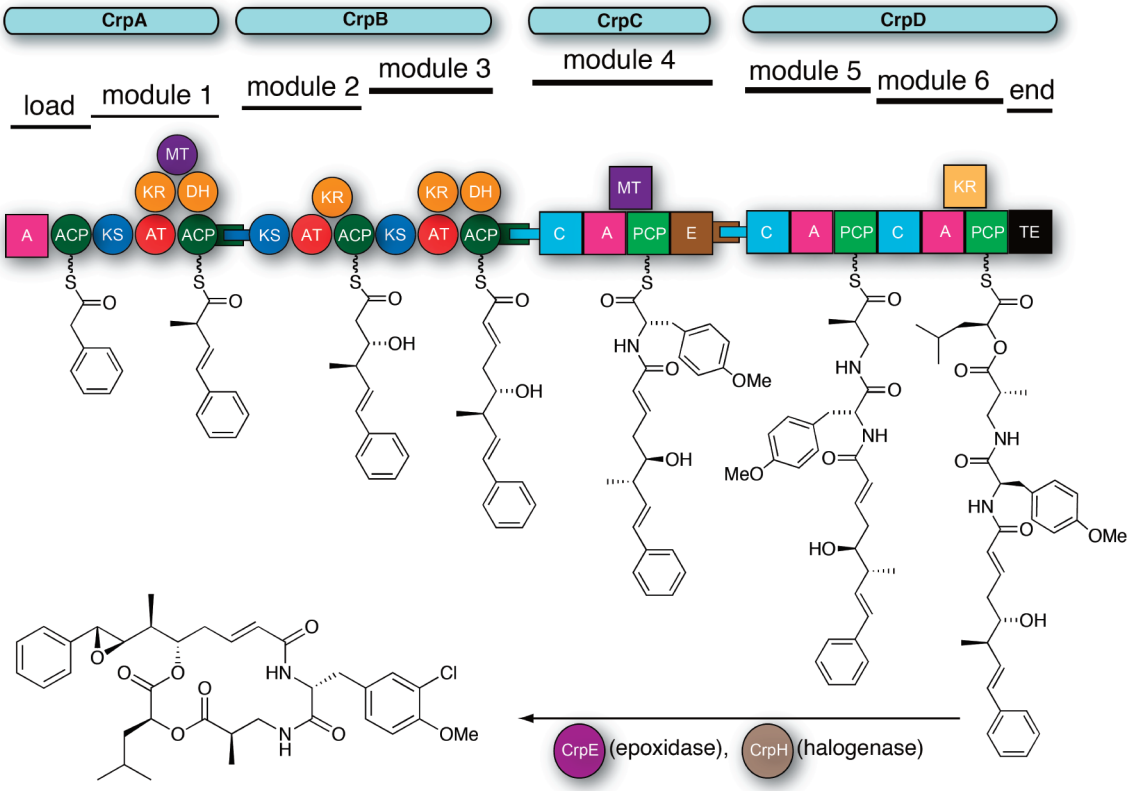


FIGURE 4. Illustration of the Crp PKS/NRPS responsible for biosynthesis of the cryptophycin peptolide core. Epoxidation by CrpE and chlorination by CrpH generates cryptophycin 1: C = condensation domain, A = adenylation domain, PCP = peptidyl carrier protein, MT = methyltransferase, E = epimerase.

Recent access to the cryptophycin (Crp) biosynthetic gene cluster has spurred development of specific enzymes from this pathway as biocatalysts in the generation of new cryptophycins with improved pharmacological profiles. The peptolide core of the most abundant cryptophycin analogue, cryptophycin 1, consists of a PKS-derived phenyloctenoic acid (unit A) and three NRPS-derived amino acids: 3-chloro-*O*-methyl-D-tyrosine (unit B), methyl β-alanine (unit C), and L-leucic acid (unit D) (Figure 4). However, there is significant diversity within this class of compounds, which contains more than 25 analogues that incorporate many substitutional variations on the core scaffold. This diversity suggests a high degree of flexibility in the enzymes of the Crp biosynthetic machinery, a hypothesis that was further borne out in precursor-directed biosynthesis studies, where a significant number of unnatural subunits were incorporated into new cryptophycin analogues by the Crp enzymes.[13]

In generating synthetic Crp analogues, efficient macrocyclization and epoxidation represent two challenging issues that can be addressed by chemoenzymatic methods. In a recently described approach, the Crp TE was excised and heterologously overexpressed as a recombinant enzyme for cyclization of linear Crp precursors.[14] NAC thioester-activated *seco*-cryptophycins were chemically synthesized, incorporating three different Unit C moieties, and utilized to interrogate the in vitro activity and substrate specificity of Crp TE (Figure 5a). Robust macrocylization was observed in each case, suggesting changes to the β-alanine site are well tolerated by the thioesterase. Signficantly, a linear cryptophycin precursor lacking the phenyl ring on the styryl moiety of unit A was poorly cyclized by Crp

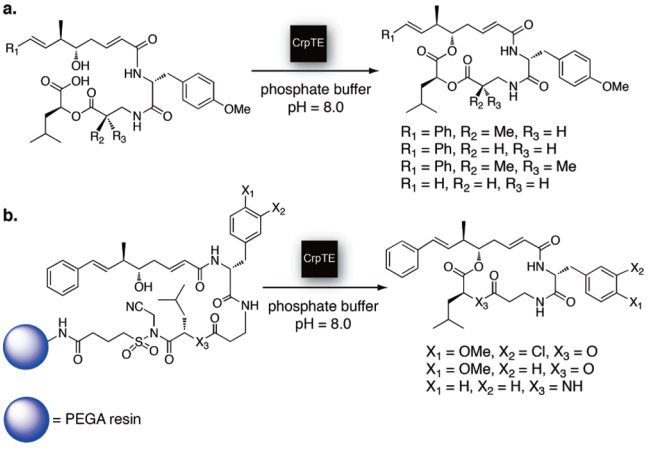


FIGURE 5. Chemoenzymatic macrocyclization of (a) solution-phase and (b) solid-phase linear cryptophycin intermediates by Crp TE in vitro.

TE and instead primarily hydrolyzed to the *seco*-acid. This indicated that a terminal aryl group is critical for efficient macrocyclization.

In a subsequent study, it was shown that this methodology could be expanded to solid-phase bound cryptophycin precursors.[15] Several *seco*-cryptophycin analogues were synthesized as acyl sulfonamides on safety-catch PEGA resin, which were activated with iodoacetonitrile and subsequently subjected to macrocyclization reactions with Crp TE (Figure 5b). Here, Crp TE was not only capable of catalyzing cyclization of these substrates directly from the activated solid support but also

tolerated changes to the unit B tyrosyl ring as well as a switch from an ester to an amide linkage between units C and D. These results clearly demonstrate the value of the cryptophycin thioesterase as a versatile biocatalyst for the synthesis of novel cryptophycins from linear and resin-bound precursors.

The cryptophycin β-epoxide group confers a 100-fold increase in compound potency, thus representing a key functional group in this class of natural products.[16] Introduction of the functional group can only be accomplished with modest diastereoselectivity by chemical methods, and separation of the resulting mixture of α/β diastereomers is cumbersome.[17,18] Recently, the engineered P450 epoxidase CrpE identified from the cryptophycin biosynthetic gene cluster has been shown to be a viable alternative to synthetic strategies. In those studies, CrpE was expressed as a recombinant protein containing an N-terminal maltose-binding protein (MBP) tag and incubated with a small library of desepoxy cryptophycin substrates in vitro[19] (Figure 6). The resulting β-epoxy cryptophycins were generated efficiently as single diastereomers, though linear cryptophycins were not epoxidized, indicating that cyclization is a prerequisite for CrpE activity. At the same time, the epoxidase also demonstrated tolerance toward functional changes on units B and C. Finally, it was also shown in vitro that Crp TE and CrpE can be used in tandem to cyclize and epoxidize linear cryptophycins in a single reaction.[13] This has the potential to greatly streamline the chemoenzymatic synthesis of new cryptophycins by accomplishing the two most difficult steps in an efficient and economical manner.

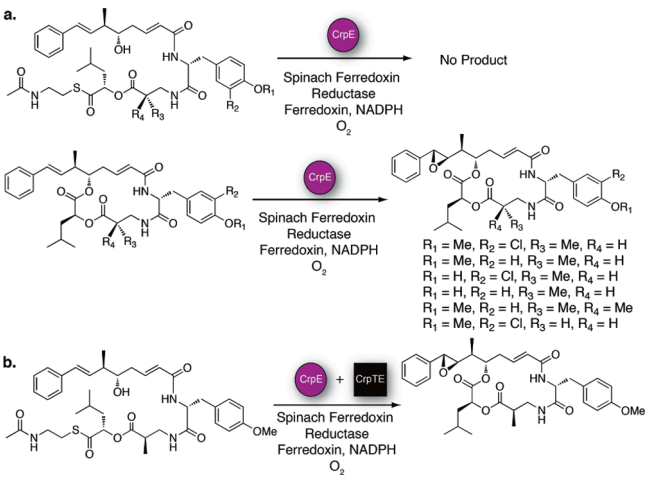


FIGURE 6. In vitro chemoenzymatic (a) epoxidation of cryptophycin intermediates by the CrpE epoxidase and (b) tandem epoxdation and macrocyclizaiton of linear Crp intermediates by CrpE and Crp TE.

Pikromycin and Erythromycin

Macrolides are a large family of antibiotics that bind to the 50S ribosomal subunit of pathogenic bacteria and block peptidyl transferase activity for amino acid chain elongation.[20] The representative member of this family, erythromycin, has been in clinical use for nearly 60 years to combat a variety of common infections and has served as a template for the production of new antibiotics (e.g., ketolides) in this class.[21,22] Despite the success of the macrolides, however, there is a critical need for rapid access to new antibiotics due to the evolution of macrolide-resistant bacteria, pathogens that pose a growing threat to human health.[23] Currently, the complex structural and functional features of macrolides make them particularly difficult targets for synthetic efforts, and semisynthesis has been the only

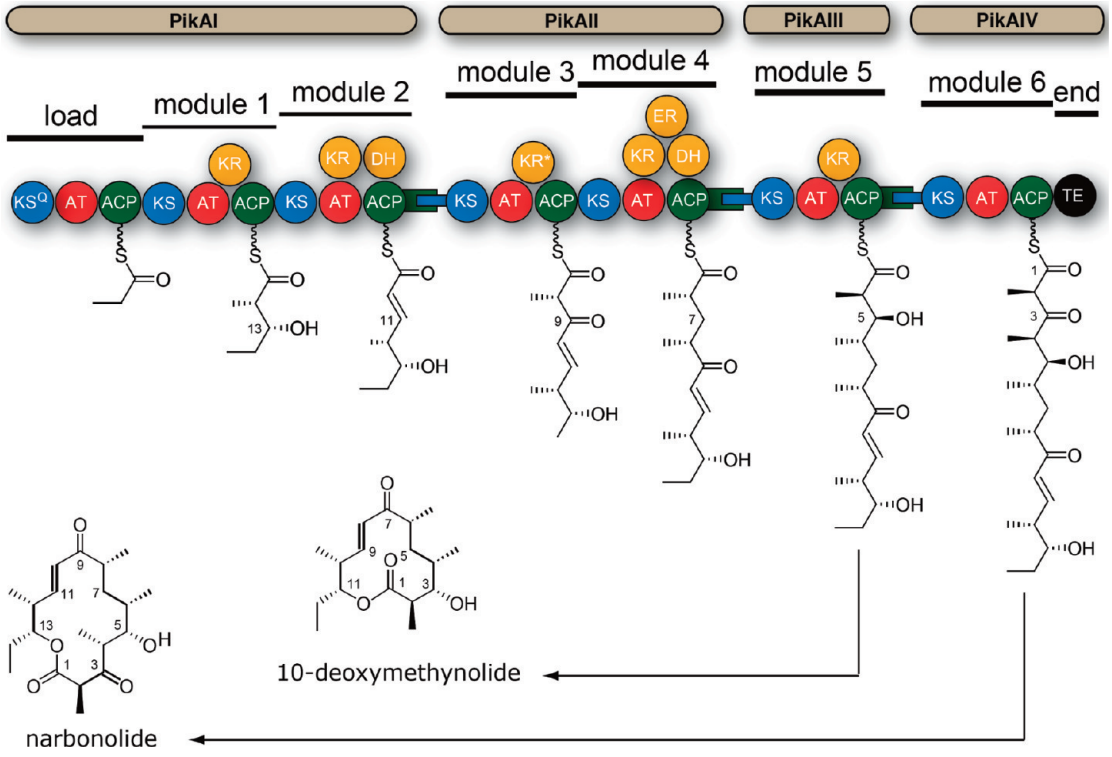


FIGURE 7. Illustration of the Pik PKS responsible for the biosynthesis of the methymycin and pikromycin aglycons 10-deoxymethynolide (10-Dml) and narbonolide (Nbl), respectively.

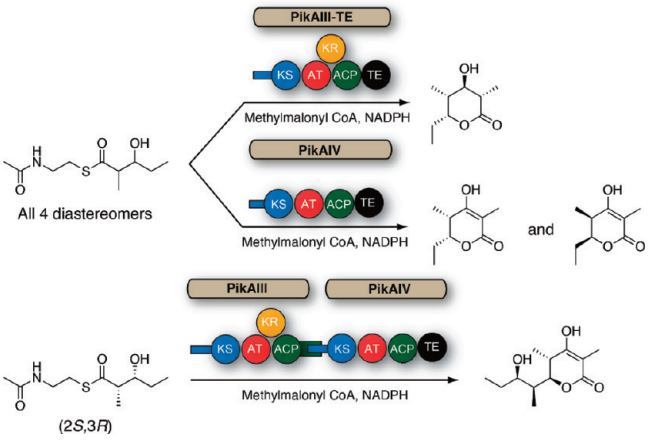


FIGURE 8. Probing molecular specificity in PikAIII and PikAIV modules with diastereomeric diketide substrates to give triketide lactones and tetraketides. Only *syn* diketides were accepted by either module with a strong preference for the (2*S*,3*R*) diketide. PikAIII only accepted the (2*S*,3*R*) diastereomer.

reliable but limited avenue for next generation macrolides such as the azalides[24] and ketolides.[21] As a result, focused study has been centered on engineered biosystems for the chemoenzymatic synthesis of new macrolide scaffolds.

The erythromycin[25,26] (DEBS, Figure 1a, vide supra) and pikromycin[27] (Pik, Figure 7) PKSs are hallmarks for polyketide natural product biosynthetic systems, producing the antibiotic aglycons 6-deoxyerythronolide B and narbonolide, respectively. The modular architecture of these two enzyme systems makes them prime candidates for rational pathway engineering, through which specific and predictable changes can be introduced to the core structures of their putative polyketide products.[28−32]A wealth of studies have been performed on both the Pik and DEBS PKSs, exploiting engineered whole modules as well as individual enzymes such as isolated thioesterase domains and P450 monooxygenases for chemoenzymatic synthesis.

Utilization of complete PKS modules for chemoenzymatic synthesis presents a significant challenge due to the size (often greater than 150 kDa) and complexity of the enzymes involved. However, the reward for overcoming this challenge is great, as efficient generation of hybrid PKSs through combinatorial reorganization of PKS modules would lend access to extraordinary chemical diversity. Toward this end, modules 5 (PikAIII) and 6 (PikAIV) from the Pik PKS present an excellent experimental system, since both modules are natively expressed as monomodular enzymes. Initial studies in this system exploited the facile functional expression and purification of recombinant PikAIII, PikAIII-TE (module 5 with a C-terminal thioesterase fusion), and PikAIV, which were then incubated with short diketide model substrates to generate tri- and tetraketide lactone products in vitro[33,34] (Figure 8). These studies not only demonstrated the versatility of recombinant PKS modules but also laid a foundational understanding of their substrate specificity and catalytic efficiency. During subsequent work, synthetic full-length native penta- and hexaketide chain elongation intermediates for PikAIII and PikAIV were used for the in vitro chemoenzymatic synthesis of the complete 12- and 14-membered ring aglycons 10-deoxymethynolide (10-Dml) and narbonolide (Nbl)[35] (Figure 9). Interestingly, comparative steady-state kinetic analysis of the loading, extension, processing, and

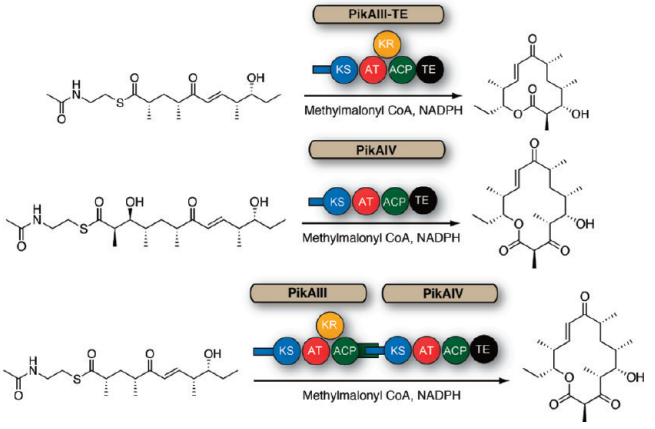


FIGURE 9. Chemoenzymatic synthesis of macrolactones with PikAIII and PikAIV modules, utilizing native chain-elongation intermediates.

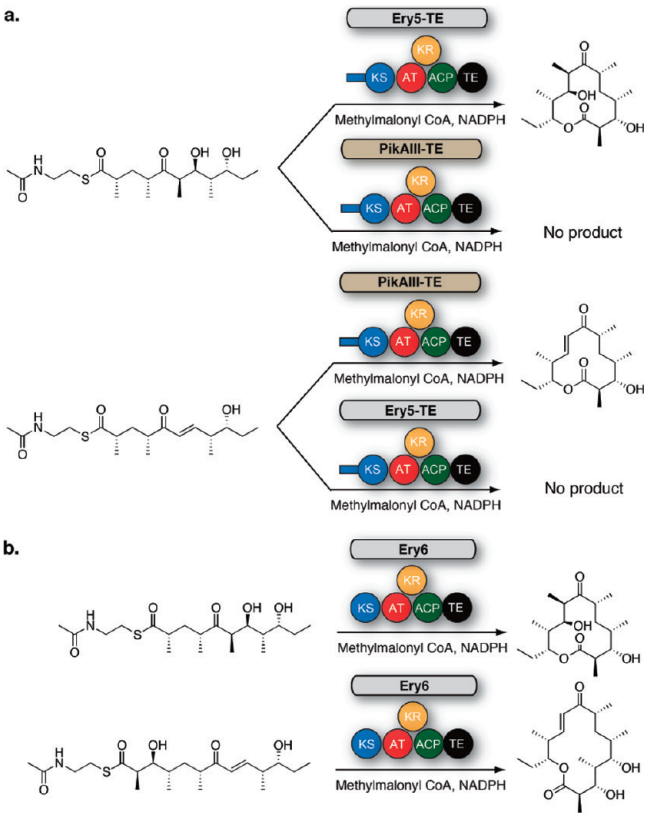


FIGURE 10. Comparative analysis of molecular specificity in the Pik and DEBS PKS: (a) Ery5-TE and PikAIII-TE accept, elongate, and cyclize their native pentaketide substrates but do not tolerate substrates from the reciprocal system; (b) Ery6 shows remarkable flexibility in processing the non-native DEBS pentaketide and Pik hexaketide substrates.

cyclization events in these modules revealed a significant difference between rates when diketide substrates were used versus native chain elongation intermediates. An inherent preference for native pentaketide and hexaketide substrates (presumably controlled through key molecular recognition elements) has been demonstrated by the ability of PikAIII and PikAIV modules to process them 2−3 orders of magnitude more efficiently than model diketides.

Similar in vitro studies have been conducted in the DEBS system, where individual engineered DEBS modules 5 (Ery5)

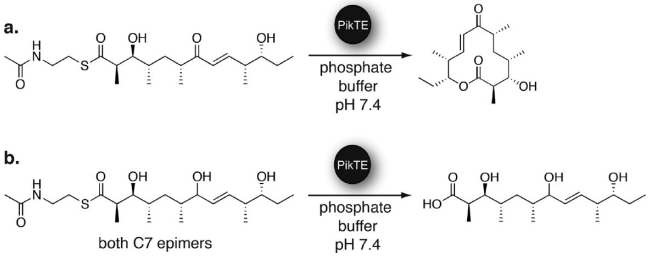


FIGURE 11. (a) In vitro chemoenzymatic synthesis of 10-deoxymethynolide with Pik TE. (b) Reduction at C7 of the hexaketide substrate abolishes TE-mediated cyclization.

and 6 (Ery6) were incubated with short-chain diketide substrates to give triketide lactone products. As in the Pik system, it was also revealed that each of the final DEBS modules has an inherent substrate specificity profile.[36−38] At the same time, though, it was demonstrated that the DEBS modules were nearly 3 orders of magnitude more efficient than their Pik counterparts in processing these diketides. Recent experiments in the DEBS PKS with synthetic native-chain elongation intermediates confirmed these findings and laid out a more detailed understanding of their molecular specificity parameters, as well as a comparative analysis with the analogous Pik modules. In these studies, both Pik module 5 (PikAIII) and DEBS module 5 (Ery5) were shown to exhibit a fairly high specificity, as evidenced by the fact that the synthetic pentaketide substrate from each system was not tolerated by the corresponding reciprocal module[39] (Figure 10a). This was surprising as the Pik and DEBS pentaketide substrates are structurally similar, only differing at the C6−C7 positions. Nevertheless, these changes were enough to preclude efficient enzymatic processing of the pentaketides by the noncognate PKS modules. On the other hand, Pik module 6 (PikAIV) and DEBS module 6 (Ery6) displayed a remarkable flexibility toward noncognate substrates, with Ery6 showing a particularly relaxed specificity. Ery6 was capable of accepting, extending, processing, and cyclizing both the non-native synthetic DEBS pentaketide substrate and the Pik hexaketide substrate (Figure 10b), which not only have differing chain lengths of 10 and 12 carbons, respectively, but contain variant proximal and distal functional group substitution patterns. This relaxed specificity profile indicates that module 6 of both Pik and DEBS may be excellent candidates for future PKS engineering efforts.

Due to the pivotal role thioesterase domains play in natural product biosynthesis, understanding the specificity of the Pik and DEBS TEs is important for their applications both in vitro and in vivo. As a result, both TEs have been the subject of experiments to explore their chemoenzymatic potential. When the Pik TE was excised from its native context, it was shown to catalyze macrocyclization of the synthetic pikromycin hexaketide SNAC substrate, exclusively forming 10-Dml[40] (Figure 11). However, when the α,β-unsaturated ketone of the substrate was reduced at C7 to the corresponding allylic alcohol, cyclization by Pik TE was abrogated. Based on this result, we reasoned that the rigid enone structure of the Pik hexaketide provides a favorable entropic contribution necessary for TE-mediated cyclization. This was further supported by structural biology studies where the Pik TE was cocrystallized with covalently bound affinity labels mimicking a portion of the pikromycin heptaketide chain elongation intermediate (PDB ID 1HFJ).[41]

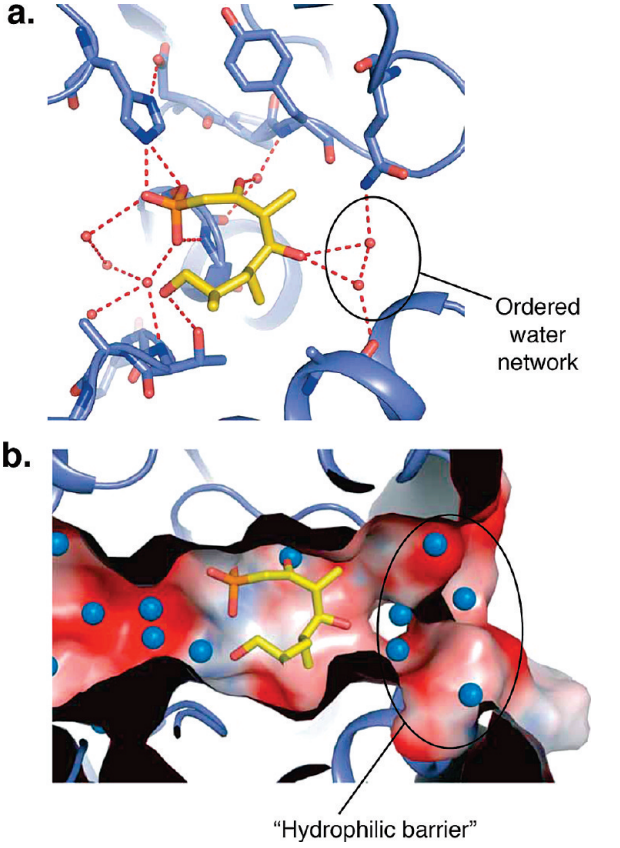


FIGURE 12. Pik TE crystal structure (1.8 Å) (a) active site with a bound affinity label mimicking the Pik cyclization intermediate. An ordered water network on right side of the channel forms a "hydrophilic barrier", inducing a curl into the cyclization intermediate; (b) electrostatic surface representation of the Pik TE substrate channel.

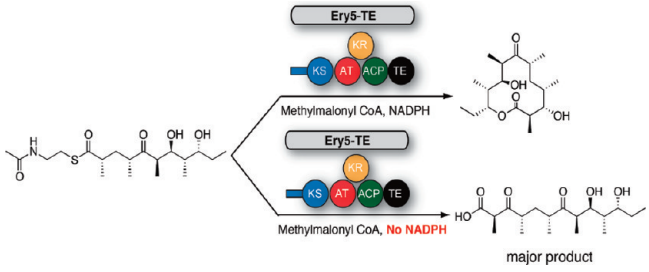


FIGURE 13. Probing specificity of the DEBS TE from in vitro reactions of Ery5-TE with synthetic DEBS pentaketide SNAC. Formation of 12-membered ring macrolactone is attenuated upon exclusion of NADPH, giving C3-unreduced hexaketide *seco*-acid as the major product.

Here, it was shown that cyclization appears to be primarily a structure-driven process with Pik TE, as minimal contacts were seen between the affinity label and the TE active site residues. A water network on one side of the thioesterase channel appeared to form a "hydrophilic barrier" that directed curling of the linear intermediate toward cyclization (Figure 12). Also, molecular modeling of the enone structure for the Pik hexa- and heptaketide cyclization intermediates showed a favorable disposition of the appropriate internal hydroxyl group as a nucleophile for macrocyclization.

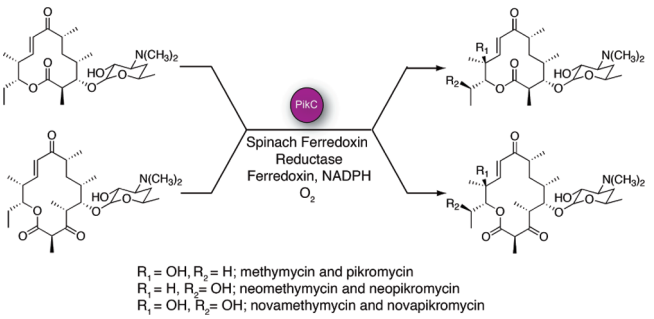


FIGURE 14. Hydroxylation patterns of YC-17 and narbomycin by the PikC P450.

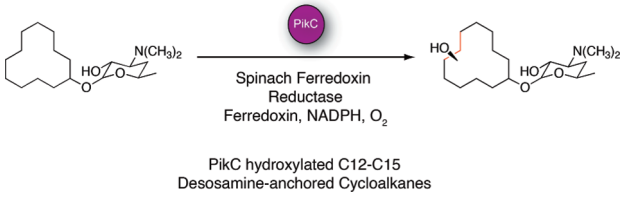


FIGURE 15. Demonstration of substrate flexibility in the PikC P450 hydroxylase on desosamine-anchored unsubstituted cycloalkanes. Carbons distal to the desosamine were hydroxylated (shown in red).

Until recently, despite the availability of solved crystal structures for DEBS TE,[42,43] little information had been gained about its ability to catalyze macrocyclization. Recent in vitro work, however, has provided fresh insights into key mechanistic requirements for macrolactonization with DEBS TE. When Ery5-TE was incubated with its native DEBS pentaketide substrate, an unnatural 12-membered ring macrolactone was formed, demonstrating a similar flexibility to Pik TE for generating varying ring sizes.[39] However, reactions with DEBS pentaketide substrate that excluded NADPH cofactor (thus eliminating β-keto reduction by the KR domain) resulted in drastically reduced levels of cyclized product and instead DEBS TE-catalyzed hydrolysis to a linear hexaketide intermediate (Figure 13). Here, the oxidation state at the β-position of the pendant hexaketide cyclization intermediate was critical, as a change from a β-hydroxy to a β-keto group resulted in attenuation of TE-mediated macrocyclization, thus indicating a key enzyme−substrate active site hydrogen bond interaction at this position. This hypothesis is consistent with a previous study involving modeling of 6-DEB into the active site of the DEBS TE crystal structure, which suggested that the 3-position hydroxyl does participate in a hydrogen-bond with Asn-180 and the backbone carbonyl of Tyr-171.[42] This contrasts with the above-described substrate binding mode for the Pik TE, which is capable of efficiently cyclizing both a hexaketide containing a β-hydroxy and a heptaketide containing a β-keto group to give 10-Dml and Nbl, respectively.[35]

Finally, the PikC P450 hydroxylase from pikromycin biosynthesis represents an attractive candidate for chemoenzymatic synthesis since it exhibits broad substrate selectivity. In its native form, PikC installs the difficult quaternary C10 and C12 hydroxyls in the methymycin and pikromycin metabolites, respectively.[44] In a series of in vitro studies with recombinant PikC, the P450 was shown to be extraordinarily flexible, hydroxylating both the C10 and C12 positions of the

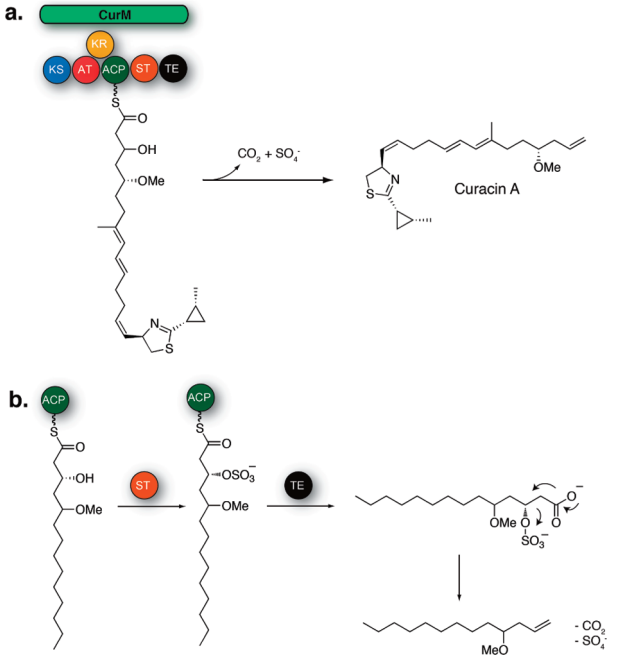


FIGURE 16. Illustration of (a) chain termination in the final CurM module, generating the terminal olefin in curacin A. (b) The sulfotransferase (ST) in CurM preferentially adds a sulfate to the (*R*)-hydroxyl of a synthetic curacin mimic, followed by sequential hydrolysis and decarboxylation by the Cur TE to give the terminal alkene.

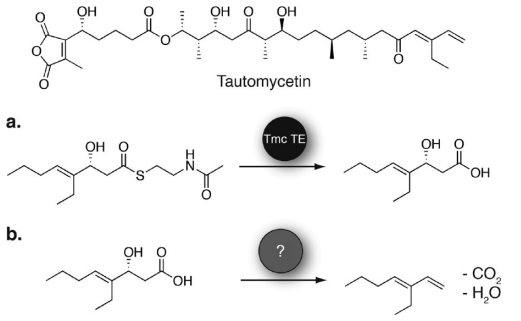


FIGURE 17. In vitro assay of tautomycetin (Tmc) TE activity. (a) TMC TE preferentially hydrolyzes synthetic Tmc SNAC substrate mimics containing an (*R*)-hydroxyl at the β-position. (b) Decarboxylation and dehydration of the resulting *seco*-acid are catalyzed by unidentified enzymes to give the terminal olefin.

12-membered ring macrolide YC-17 (10-Dml glycoslyated with desosamine on the C-3 hydroxyl) as well as both the C12 and C14 positions of the 14-membered ring macrolide narbomycin (narbonolide glycosylated with desosamine on the C-5 hydroxyl)[44−46] (Figure 14). Here, it was determined that substrate anchoring of the macrolactones by the desosamine sugar in the PikC active site was vital for activity and selectivity.[47] Also, the alternative hydroxylation patterns seen in these in vitro studies were identified natively in vivo as minor Pik metabolites neomethymycin/novamethymycin and neopikromycin/novapikromycin.[46,48] The only required substitution for hydroxylation by PikC was the desosamine sugar appendage on the macrocycle. Also, the cocrystal structures between PikC and substrate revealed that due to the desosamine anchor the most distal carbons on the

macrocycles are positioned close to the enzyme Fe center, and are subsequently hydroxylated (Figure 15). This detailed understanding of PikC function has recently motivated an effort to explore its ability to catalyze remote C−H bond activation for hydroxylation of a series of desosaminylated hydrocarbon rings of varying size. Remarkably, PikC showed a facile ability to selectively hydroxylate these compounds that also exhibited surprising antibiotic activity.[49]

Curacin and Tautomycetin

With the ongoing discovery of new natural products and their putative biosynthetic pathways, the biochemical toolbox for chemoenzymatic synthesis continues to expand. This has resulted in increasing levels of chemical diversity and the uncovering of enzymes that can perform unique chemical transformations. As has been discussed earlier, macrocyclization by a PKS or NRPS terminal thioesterase domain is often a final step in forming bioactive natural products; however, some natural product pathways have evolved to terminate in unusual ways. This is the case with the natural products curacin A[50] (Cur) and tautomycetin[51] (Tmc), where distinctive chain termination events give rise to linear products with rare terminal olefin groups. In both cases, a series of hydrolysis, decarboxylation, and dehydration reactions appears to be required for formation of this functionality, though the mechanisms for accomplishing this are divergent in the two systems.

For the Cur and Tmc systems, terminal thioesterases hydrolyze their respective intermediate metabolites from upstream ACPs to give linear *seco*-acids. In curacin biosynthesis, this TE-mediated cleavage is preceded by an unusual embedded sulfotransferase (ST) enzyme in the final CurM module that transfers a sulfonate to the β-hydroxy group of the ACP-tethered intermediate[4] (Figure 16). Following sulfonation, the TE domain catalyzes hydrolysis of the intermediate phosphopantetheinyl thioester with concomitant decarboxylation and elimination of sulfate to give the terminal olefin. By contrast, in tautomycetin biosynthesis the linear *seco*-acid formed by TE hydrolysis of the final intermediate is unactivated for elimination at the β-hydroxy position. Here, hydrolysis occurs in a stereospecific manner favoring the (*R*) hydroxyl group at the β-position. At the same time, unlike in Cur biosynthesis, decarboxylation/elimination is uncoupled from the thioesterase activity.[5] Rather, it appears that committed decarboxylase/dehydratase enzymes are responsible for decarboxylation and elimination of water to give the final product (Figure 17). Currently, two putative decarboxylases and a dehydratase, encoded by *tmcJ*, *tmcK*, and *tmcM* are candidate enzymes for catalyzing these reactions.

Future Outlook

Realization of the promise of chemoenzymatic synthesis is rooted in a thorough understanding of the fundamental mechanistic underpinnings of the enzymes of interest. Through a synergistic merging of synthetic chemistry with enzymology and structural biology, detailed knowledge can be obtained for natural product biosynthetic systems, thus paving the way for their rational application toward generating new compounds of biological interest.

The ability of recombinant thioesterases to chemoenzymatically generate macrocyclic natural product scaffolds holds outstanding promise for medicinal chemistry, providing a robust and environmentally friendly alternative to difficult chemical macrocyclization strategies. The examples presented in this Perspective are representative of the broad utility of TEs as biocatalysts. Salient examples have also been demonstrated with thioesterases in other PKS/NRPS systems such as tyrocidine[52−56] and epothilone,[57,58] further supporting this general premise. In order to leverage the catalytic power and specificity of recombinant TEs, however, future studies must first aim at detailed understanding of their many underlying structural details and catalytic limitations. It can be envisioned that through structure-based protein engineering efforts, rational development of TEs with expanded substrate tolerance and increased catalytic efficiencies will be within reach.

At the same time, P450 hydroxylases and epoxidases are well positioned to address the challenging task of selectively oxidizing C−H bonds and alkenes in a stereo- and regiospecific manner. Both the CrpE epoxidase and PikC hydroxylase are powerful examples of the utility of P450s for post-PKS/NRPS tailoring of natural product scaffolds. Nonetheless, development of P450 enzymes for chemoenzymatic synthesis includes significant challenges. For example, P450s require reductase enzyme partners and ferredoxin cofactors for activity and may not operate efficiently with in vivo partners during heterologous expression. For in vitro work, this requires the use of an exogenous reductase/ferredoxin partner, such as commercially available spinach ferredoxin reductase.[19,44] This strategy, however, is not economical for large-scale chemoenzymatic synthesis and does not necessarily reflect general reductase compatibility of other natural product pathway P450s. In the case of PikC, this hurdle was overcome by making a PikC-RhFRED fusion to generate a self-sufficient reductase-coupled P450.[45] General applicability of this strategy was demonstrated in the same study, where a similar fusion was made with one of the erythromycin P450s, EryF, to provide a self-sufficient enzyme. Finally, substrate specificity requirements of P450 enzymes, such as the desosamine anchoring mechanism of PikC,[47,49] must be considered in order to successfully apply them for biocatalytic conversion of new substrates. Nevertheless, despite these challenges, the opportunity afforded by biosynthetic P450s remains an exciting prospect for future work.

Working in larger systems, the manipulation of complete PKS modules with multiple enzymatic domains represents a distinctly challenging goal in chemoenzymatic synthesis but one that can pay extraordinary dividends. By harnessing the biosynthetic potential of full modules, it can be envisioned that an array of polyketide scaffolds containing myriad stereochemical and functional permutations could be accomplished in combinatorial fashion. In this Perspective, the examples from the Pik and DEBS PKSs are a testament to the utility of intact modules for chemoenzymatic synthesis both in vitro and in vivo. At the same time, current work in other systems such as epothilone[57,58] and rapamycin[59] continues to expand our access to new PKS modules with unique substrate specificities and catalytic capabilities. Thus, while current understanding of PKS modules in many systems is still insufficient or incomplete for practical metabolic engineering efforts, the most recent studies using native chain elongation intermediates from Pik and DEBS[35,39] have been particularly

crucial in guiding detailed biochemical evaluations of underlying mechanisms in those two systems. This suggests that future studies must continue to employ native or near-native substrates and their analogues to rigorously probe inherent molecular specificities of modular PKS domains for accepting, elongating, processing, and cyclizing substrates.

Finally, the unique and divergent strategies for rendering a terminal alkene product from the curacin A and tautomycetin biosynthetic pathways highlight the malleability of PKS and NRPS systems toward the evolution of new enzyme activities and chemical functionality. Indeed, among biosynthetic chain termination strategies, new patterns continue to emerge that depart from more conventional macrocyclization and hydrolysis strategies. In addition to terminal olefin formation, reductase-catalyzed off-loading of aldehyes[60] and TE-mediated intermolecular cyclooligomerizations[61−63] have also been reported. Still, these enzyme activities only represent a small fraction of all marine and terrestrial metabolic diversity. Mechanisms for constructing other novel structures such as the cyclopropane ring and vinyl chloride moieties in curacin A and the jamaicamides[64] and the exocyclic enones on the bryostatin core[65] have also been elucidated recently. Thus, with the ongoing discovery and addition of new enzymes to the biochemical toolbox, the prospects for accessing valuable chemical diversity through chemoenzymatic synthesis continue to burn bright.

Acknowledgment. The generous support of NIH (NIGMS and NCI) through grants to DHS and the Hans W. Vahlteich Professorship at the University of Michigan has made this work possible. The cover art photography was graciously provided by Pam Schultz.

References

(1) Newman, D. J.; Cragg, G. M.; Snader, K. M. *J. Nat. Prod.* **2003**, *66*, 1022–1037.
(2) Staunton, J.; Weissman, K. J. *Nat. Prod. Rep.* **2001**, *18*, 380–416.
(3) Schwarzer, D.; Finking, R.; Marahiel, M. A. *Nat. Prod. Rep.* **2003**, *20*, 275–287.
(4) Gu, L.; Wang, B.; Kulkarni, A.; Gehret, J. J.; Lloyd, K. R.; Gerwick, L.; Gerwick, W. H.; Wipf, P.; Hakansson, K.; Smith, J. L.; Sherman, D. H. *J. Am. Chem. Soc.* **2009**, *131*, 16033–16035.
(5) Scaglione, J. B.; Akey, D. L.; Smith, J. L.; Sherman, D. H. *Angew. Chem., Int. Ed.* **2010**, *49*, 5726−5730.
(6) Parenty, A.; Moreau, X.; Campagne, J. M. *Chem. Rev.* **2006**, *106*, 911–939.
(7) Ortiz de Montellano, P. R. *Cytochrome P450: Structure, Mechanism and Biochemistry*, 2nd ed.; Plenum Press: New York, 1995.
(8) Schwartz, R. E.; Hirsch, C. F.; Sesin, D. F.; Flor, J. E.; Chartrain, M.; Fromtling, R. E.; Harris, G. H.; Salvatore, M. J.; Liesch, J. M.; Yudin, K. *J. Ind. Microbiol. Biotechnol.* **1990**, *5*, 113–123.
(9) Golakoti, T.; Yoshida, W. Y.; Chaganty, S.; Moore, R. E. *J. Nat. Prod.* **2001**, *64*, 54–59.
(10) Smith, C. D.; Zhang, X.; Mooberry, S. L.; Patterson, G. M.; Moore, R. E. *Cancer Res.* **1994**, *54*, 3779–3784.
(11) Edelman, M. J.; Gandara, D. R.; Hausner, P.; Israel, V.; Thornton, D.; DeSanto, J.; Doyle, L. A. *Lung Cancer* **2003**, *39*, 197–199.
(12) D'Agostino, G.; del Campo, J.; Mellado, B.; Izquierdo, M. A.; Minarik, T.; Cirri, L.; Marini, L.; Perez-Gracia, J. L.; Scambia, G. *Int. J. Gynecol. Cancer* **2006**, *16*, 71–76.
(13) Magarvey, N. A.; Beck, Z. Q.; Golakoti, T.; Ding, Y.; Huber, U.; Hemscheidt, T. K.; Abelson, D.; Moore, R. E.; Sherman, D. H. *ACS Chem. Biol.* **2006**, *1*, 766–779.
(14) Beck, Z. Q.; Aldrich, C. C.; Magarvey, N. A.; Georg, G. I.; Sherman, D. H. *Biochemistry* **2005**, *44*, 13457–66.
(15) Seufert, W.; Beck, Z. Q.; Sherman, D. H. *Angew. Chem., Int. Ed.* **2007**, *46*, 9298–9300.
(16) Eggen, M.; Georg, G. I. *Med. Res. Rev.* **2002**, *22*, 85–101.
(17) Barrow, R. A.; Hemscheidt, T.; Liang, J.; Paik, S.; Moore, R. E.; Tius, M. A. *J. Am. Chem. Soc.* **1995**, *117*, 2479–2490.
(18) Hoard, D. W.; Moher, E. D.; Martinelli, M. J.; Norman, B. H. *Org. Lett.* **2002**, *4*, 1813–1815.
(19) Ding, Y.; Seufert, W. H.; Beck, Z. Q.; Sherman, D. H. *J. Am. Chem. Soc.* **2008**, *130*, 5492–5498.
(20) Schlunzen, F.; Zarivach, R.; Harms, J.; Bashan, A.; Tocilj, A.; Albrecht, R.; Yonath, A.; Franceschi, F. *Nature* **2001**, *413*, 814–821.
(21) Or, Y. S.; Clark, R. F.; Wang, S. Y.; Chu, D. T. W.; Nilius, A. M.; Flamm, R. K.; Mitten, M.; Ewing, P.; Alder, J.; Ma, Z. K. *J. Med. Chem.* **2000**, *43*, 1045–1049.
(22) Denis, A.; Agouridas, C.; Auger, J. M.; Benedetti, Y.; Bonnefoy, A.; Bretin, F.; Chantot, J. F.; Dussarat, A.; Fromentin, C.; D'Ambrieres, S. G.; Lachaud, S.; Laurin, P.; Le Martret, O.; Loyau, V.; Tessot, N.; Pejac, J. M.; Perron, S. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 3075–3080.
(23) Moellering, R. C., Jr.; Craig, W.; Edmond, M.; Farrell, D. J.; Ferraro, M. J.; File, T. M., Jr.; Klein, J.; Lonks, J.; Metlay, J. P.; Sahm, D.; Talbot, G. H. *J. Chemother.* **2002**, *14*, 42–56.
(24) Mutak, S. *J. Antibiot.* **2007**, *60*, 85–122.
(25) Cortes, J.; Haydock, S. F.; Roberts, G. A.; Bevitt, D. J.; Leadlay, P. F. *Nature* **1990**, *348*, 176–178.
(26) Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L. *Science* **1991**, *252*, 675–679.
(27) Xue, Y.; Zhao, L.; Liu, H. W.; Sherman, D. H. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12111–6.
(28) Donadio, S.; McAlpine, J. B.; Sheldon, P. J.; Jackson, M.; Katz, L. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 7119–7123.
(29) McDaniel, R.; Kao, C. M.; Fu, H.; Hevezi, P.; Gustafsson, C.; Betlach, M.; Ashley, G.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **1997**, *119*, 4309–4310.
(30) Tang, L.; Fu, H.; McDaniel, R. *Chem. Biol.* **2000**, *7*, 77–84.
(31) Yoon, Y. J.; Beck, B. J.; Kim, B. S.; Kang, H. Y.; Reynolds, K. A.; Sherman, D. H. *Chem. Biol.* **2002**, *9*, 203–214.
(32) Menzella, H. G.; Reid, R.; Carney, J. R.; Chandran, S. S.; Reisinger, S. J.; Patel, K. G.; Hopwood, D. A.; Santi, D. V. *Nat. Biotechnol.* **2005**, *23*, 1171–1176.
(33) Beck, B. J.; Aldrich, C. C.; Fecik, R. A.; Reynolds, K. A.; Sherman, D. H. *J. Am. Chem. Soc.* **2003**, *125*, 12551–12557.
(34) Yin, Y.; Lu, H.; Khosla, C.; Cane, D. E. *J. Am. Chem. Soc.* **2003**, *125*, 5671–5676.
(35) Aldrich, C. C.; Beck, B. J.; Fecik, R. A.; Sherman, D. H. *J. Am. Chem. Soc.* **2005**, *127*, 8441–8452.
(36) Gokhale, R. S.; Tsuji, S. Y.; Cane, D. E.; Khosla, C. *Science* **1999**, *284*, 482–485.
(37) Wu, N.; Kudo, F.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **2000**, *122*, 4847–4852.
(38) Wu, N.; Tsuji, S. Y.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **2001**, *123*, 6465–6474.
(39) Mortison, J. D.; Kittendorf, J. D.; Sherman, D. H. *J. Am. Chem. Soc.* **2009**, *131*, 15784–15793.
(40) Aldrich, C. C.; Venkatraman, L.; Sherman, D. H.; Fecik, R. A. *J. Am. Chem. Soc.* **2005**, *127*, 8910–8911.
(41) Akey, D. L.; Kittendorf, J. D.; Giraldes, J. W.; Fecik, R. A.; Sherman, D. H.; Smith, J. L. *Nat. Chem. Biol.* **2006**, *2*, 537–542.
(42) Tsai, S. C.; Miercke, L. J. W.; Krucinski, J.; Gokhale, R.; Chen, J. C. H.; Foster, P. G.; Cane, D. E.; Khosla, C.; Stroud, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14808–14813.
(43) Tsai, S. C.; Lu, H. X.; Cane, D. E.; Khosla, C.; Stroud, R. M. *Biochemistry* **2002**, *41*, 12598–12606.
(44) Xue, Y.; Wilson, D.; Zhao, L.; Liu, H.; Sherman, D. H. *Chem. Biol.* **1998**, *5*, 661–7.
(45) Li, S.; Podust, L. M.; Sherman, D. H. *J. Am. Chem. Soc.* **2007**, *129*, 12940–12941.
(46) Lee, S. K.; Park, J. W.; Kim, J. W.; Jung, W. S.; Park, S. R.; Choi, C. Y.; Kim, E. S.; Kim, B. S.; Ahn, J. S.; Sherman, D. H.; Yoon, Y. J. *J. Nat. Prod.* **2006**, *69*, 847–9.
(47) Sherman, D. H.; Li, S.; Yermalitskaya, L. V.; Kim, Y.; Smith, J. A.; Waterman, M. R.; Podust, L. M. *J. Biol. Chem.* **2006**, *281*, 26289–97.
(48) Zhang, Q.; Sherman, D. H. *J. Nat. Prod.* **2001**, *64*, 1447–50.
(49) Li, S. Y.; Chaulagain, M. R.; Knauff, A. R.; Podust, L. M.; Montgomery, J.; Sherman, D. H. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 18463–18468.
(50) Chang, Z.; Sitachitta, N.; Rossi, J. V.; Roberts, M. A.; Flatt, P. M.; Jia, J.; Sherman, D. H.; Gerwick, W. H. *J. Nat. Prod.* **2004**, *67*, 1356–67.
(51) Choi, S. S.; Hur, Y. A.; Sherman, D. H.; Kim, E. S. *Microbiology* **2007**, *153*, 1095–1102.
(52) Trauger, J. W.; Kohli, R. M.; Mootz, H. D.; Marahiel, M. A.; Walsh, C. T. *Nature* **2000**, *407*, 215–218.
(53) Kohli, R. M.; Walsh, C. T.; Burkart, M. D. *Nature* **2002**, *418*, 658–661.
(54) Trauger, J. W.; Kohli, R. M.; Walsh, C. T. *Biochemistry* **2001**, *40*, 7092–7098.
(55) Kohli, R. M.; Burke, M. D.; Tao, J. H.; Walsh, C. T. *J. Am. Chem. Soc.* **2003**, *125*, 7160–7161.
(56) Lin, H. N.; Walsh, C. T. *J. Am. Chem. Soc.* **2004**, *126*, 13998–14003.
(57) Boddy, C. N.; Schneider, T. L.; Hotta, K.; Walsh, C. T.; Khosla, C. *J. Am. Chem. Soc.* **2003**, *125*, 3428–3429.
(58) Boddy, C. N.; Hotta, K.; Tse, M. L.; Watts, R. E.; Khosla, C. *J. Am. Chem. Soc.* **2004**, *126*, 7436–7437.

(59) Gregory, M. A.; Petkovic, H.; Lill, R. E.; Moss, S. J.; Wilkinson, B.; Gaisser, S.; Leadlay, P. F.; Sheridan, R. M. *Angew. Chem., Int. Ed.* **2005**, *44*, 4757–4760.
(60) Gaitatzis, N.; Kunze, B.; Müller, R. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 11136–11141.
(61) Gehring, A. M.; Mori, I.; Walsh, C. T. *Biochemistry* **1998**, *37*, 2648–2659.
(62) Wu, X. M.; Bu, X. Z.; Wong, K. M.; Yan, W. L.; Guo, Z. H. *Org. Lett.* **2003**, *5*, 1749–1752.
(63) Lombo, F.; Velasco, A.; Castro, A.; de la Calle, F.; Brana, A. F.; Sanchez-Puelles, J. M.; Mendez, C.; Salas, J. A. *ChemBioChem* **2006**, *7*, 366–376.
(64) Gu, L. C.; Wang, B.; Kulkarni, A.; Geders, T. W.; Grindberg, R. V.; Gerwick, L.; Hakansson, K.; Wipf, P.; Smith, J. L.; Gerwick, W. H.; Sherman, D. H. *Nature* **2009**, *459*, 731–735.
(65) Buchholz, T. J.; Rath, C. M.; Lopanik, N. B.; Gardner, N. P.; Hakansson, K.; Sherman, D. H. *Chem. Biol.*, in press.