Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms

Ben Shen

Recent literature on polyketide biosynthesis suggests that polyketide synthases have much greater diversity in both mechanism and structure than the current type I, II and III paradigms. These examples serve as an inspiration for searching novel polyketide synthases to give new insights into polyketide biosynthesis and to provide new opportunities for combinatorial biosynthesis.

Addresses
Division of Pharmaceutical Sciences and Department of Chemistry, University of Wisconsin, Madison, WI 53705, USA
e-mail: bshen@pharmacy.wisc.edu

Current Opinion in Chemical Biology 2003, **7**:285–295

This review comes from a themed section on Biocatalysis and biotransformation
Edited by Tadhg Begley and Ming-Daw Tsai

1367-5931/03/$ – see front matter
© 2003 Elsevier Science Ltd. All rights reserved.

DOI 10.1016/S1367-5931(03)00020-6

Abbreviations

ACP	acyl carrier protein
AT	acyl transferase
DEBS	6-deoxyerythromycin B synthase
DH	dehydratase
KR	ketoreductase
KS	ketoacyl synthase
NRPS	nonribosomal peptide synthetase
PKS	polyketide synthase
TD	terminal domain

Introduction

Polyketides are a large family of natural products found in bacteria, fungi and plants, and include many clinically important drugs such as tetracycline, daunorubicin, erythromycin, rapamycin and lovastatin. They are biosynthesized from acyl CoA precursors by polyketide synthases (PKSs). Much of the current research on polyketide biosynthesis is driven by: first, the unparalleled biological activities and enormous commercial value of these natural products, which remain the most successful candidates for new drug discovery; second, the extraordinary structure, mechanism and catalytic reactivity of PKSs that provide an unprecedented opportunity to investigate the molecular mechanisms of enzyme catalysis, molecular recognition and protein–protein interaction; and third, the remarkable versatility and amenability of PKSs that allow the generation of novel compounds, difficult to access by other means, by combinatorial biosynthesis methods.

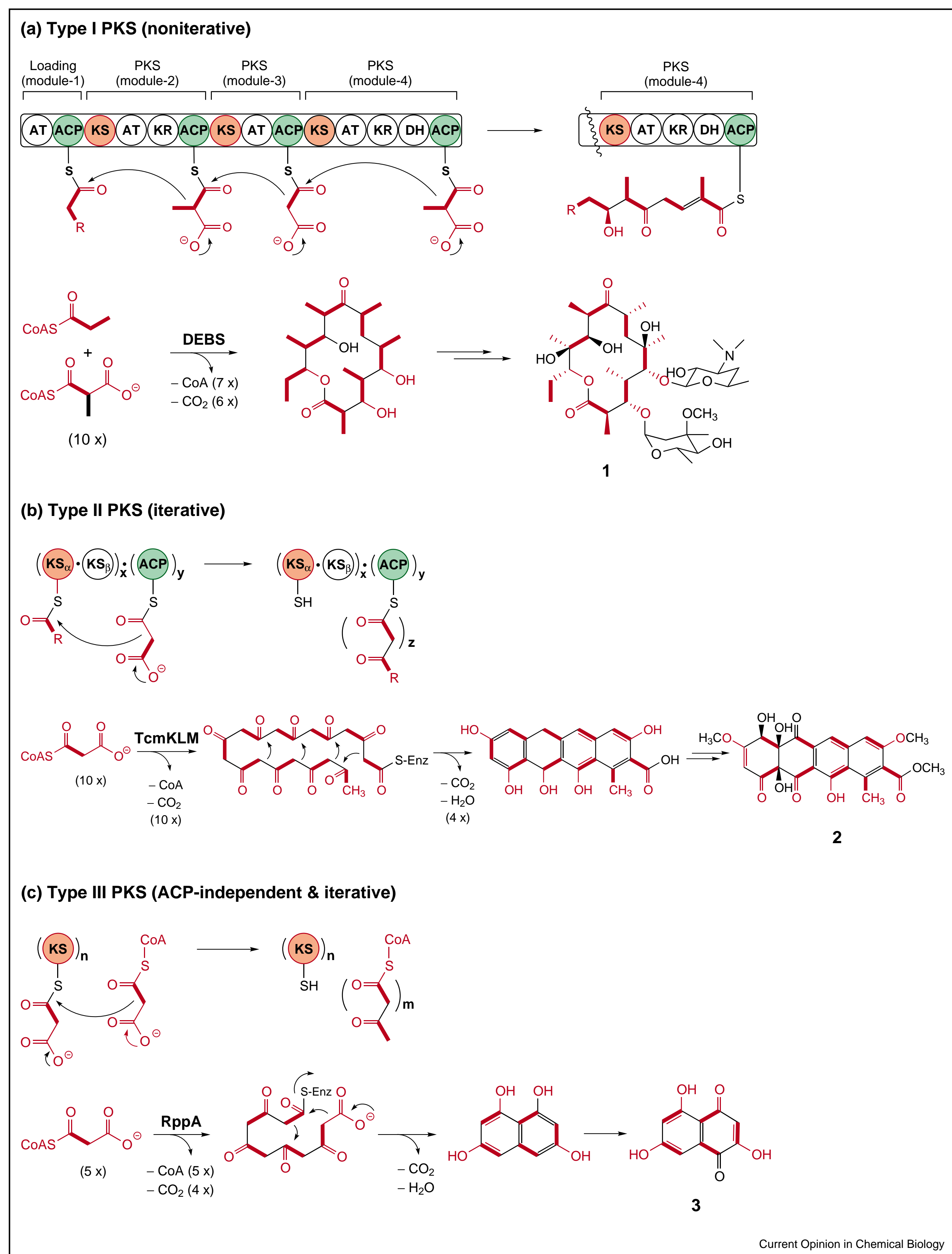
Three types of bacterial PKSs are known to date. First, type I PKSs are multifunctional enzymes that are organized into modules, each of which harbors a set of distinct, non-iteratively acting activities responsible for the catalysis of one cycle of polyketide chain elongation, as exemplified by the 6-deoxyerythromycin B synthase (DEBS) for the biosynthesis of reduced polyketides (i.e. macrolides, polyethers and polyene) such as erythromycin A (**1**) (Figure 1a) [1]. Second, type II PKSs are multienzyme complexes that carry a single set of iteratively acting activities, as exemplified by the tetracenomycin PKS for the biosynthesis of aromatic polyketides (often polycyclic) such as tetracenomycin C (**2**) (Figure 1b) [2]. Third, type III PKSs, also known as chalcone synthase-like PKSs, are homodimeric enzymes that essentially are iteratively acting condensing enzymes, as exemplified by the RppA synthase for the biosynthesis of aromatic polyketides (often monocyclic or bicyclic), such as flavolin (**3**) (Figure 1c) [3]. Type I and II PKSs use acyl carrier protein (ACP) to activate the acyl CoA substrates and to channel the growing polyketide intermediates, whereas type III PKSs, independent of ACP, act directly on the acyl CoA substrates. Despite structural and mechanistic differences, all types of PKSs biosynthesize polyketides by sequential decarboxylative condensation of the acyl CoA precursors, and the ketoacyl synthase (KS) domain (for type I PKSs) or subunit (for type II and III PKSs) catalyzes the C–C bond-forming step.

Since the first reports of bacterial type I PKS in 1990 [4,5], type II PKS in 1984 [6,7], and type III PKS in 1999 [8], the PKS paradigms have served the scientific community beyond the call of duty, providing the molecular basis to explain the vast structural diversity observed with polyketide natural products, and the biotechnological platform to produce 'unnatural' natural products by combinatorial biosynthesis methods with engineered PKSs. As the field stumbles into its adolescence, how much do we really know about polyketide biosynthesis? Here, selected examples from recent literature are presented to argue that PKSs have much greater diversity in both mechanism and structure than the currently well appreciated type I, II and III paradigms. These examples serve as an inspiration in searching for novel PKSs, both to give new insights into polyketide biosynthesis and to provide new opportunities for combinatorial biosynthesis.

Iterative type I PKSs for aromatic polyketide biosynthesis

Although aromatic polyketide biosyntheses in fungi are catalyzed by iterative type I PKSs, as exemplified by the

Figure 1



6-methylsalicyclic acid synthase, the paradigm for aromatic polyketide biosyntheses in bacteria is the iterative type II PKS (Figure 1b) [2]. However, because most of the type II PKSs studied so far were cloned according to the type II PKS paradigm, caution has to be taken to generalize type II PKSs for aromatic polyketide biosynthesis in bacteria. Bechthold and co-workers cloned the first iterative type I PKS, AviM, for aromatic polyketide biosynthesis in bacteria in 1997 [9]. The *aviM* gene was discovered from the avilamycin (**4**; Figure 2a) biosynthetic gene cluster that was cloned from *Streptomyces viridochromogens* Tü57 using a deoxysugar biosynthetic gene as a probe. AviM has the characteristic type I PKS domains of KS, acyl transferase (AT), dehydratatse (DH), and ACP (Figure 2b). Because the orsellinic acid moiety is the only structural element in **4** that could be of polyketide origin (Figure 2a), Bechthold and co-workers expressed *aviM* in *Streptomyces lividans* TK24 and *Streptomyces coelicolour* CH999 to verify if AviM can catalyzes orsellinic acid (**5**) biosynthesis *in vivo*. Production of **5** was observed in both hosts, confirming AviM as an orsellinic acid synthase that catalyzes aromatic polyketide biosynthesis from the acyl CoA precursors in an iterative process (Figure 2c). Ironically, AviM was treated more as an exception to the type II PKS paradigm than as an indication that aromatic polyketide biosynthesis in bacteria could be catalyzed by iterative type I PKS.

Thorson and co-workers [10] recently cloned the calicheamicin (**6**) biosynthetic gene cluster from *Micromonospora echinospora* ssp. *calichensis* by screening for genes conferring **6** resistance and subsequently sequenced and characterized the *cal* cluster [11••]. Two PKS genes, *calE8* and *calO5*, were identified within the cluster, each of which is characteristic of type I PKS. There are two structural elements within **6** that are of polyketide origin: the enediyne core and the orsellinic acid moiety (Figure 2a). Because CalE8 was established as the enediyne core synthase (see discussion below), CalO5 was proposed to be responsible for the biosynthesis of the orsellinic acid moiety of **6**. Strikingly, CalO5, consisting of KS, AT, DH and ACP domains, exhibits head-to-tail sequence homology and has an identical domain organization to AviM (Figure 2b). These results support the functional assignment of CalO5 as an orsellinic synthase, revealing the second example of iterative type I PKSs for aromatic polyketide biosynthesis in bacteria (Figure 2c).

While AviM and CalO5 clearly demonstrated that the biosynthesis of monocyclic aromatic polyketides (tetraketides) could be catalyzed by iterative type I PKSs, the finding of NcsB from the neocarzinostatin (**7**) biosynthetic gene cluster suggests that bacterial iterative type I PKSs are not limited to monocyclic aromatic polyketide biosynthesis and could catalyse the biosynthesis of higher-order aromatic polyketides, such as the naphthalinic acid moiety of **7** (hexaketide) (Figure 2a). In fact, early attempts to clone the *ncs* biosynthetic gene cluster for **7** on the assumption that its naphthalinic acid moiety is biosynthesized by a type II PKS failed, and, retrospectively, this approach cannot be successful because the *ncs* cluster harbors no type II PKS gene. We succeeded in cloning the *ncs* gene cluster by chromosomal walking from the *ncsA* gene that encodes the neocarzinostatin apoprotein [12••]. Two PKS genes were identified within the *ncs* gene cluster. One of them, *ncsB*, encodes a type I PKS consisting of the characteristic KS, AT, DH, ketoreductase (KR) and ACP domains (Figure 2d). In a mechanistic analogy to AviM and CalO5, NcsB could be envisaged as catalysing the biosynthesis of naphthalinic acid (**8**) from the acyl CoA precursors in an iterative process, with an exception of regiospecific reduction at C-6 by the KR domain (Figure 2e). This hypothesis is consistent with the remarkable homology in both amino acid sequence and domain organization among NcsB, AviM and CalO5. NcsB, therefore, represents the third example of bacterial iterative type I PKSs for aromatic polyketide biosynthesis.

Iterative type I PKSs for enediyne biosynthesis

The enediyne family of antibiotics is structurally characterized by the enediyne core, a unit consisting of two acetylenic groups conjugated to a double bond or incipient double bond within a nine-membered ring (e.g. **7** and C-1027 (**9**)) or ten-membered ring (e.g. **6**) (Figure 2a). Although feeding experiments with ^{13}C-labeled precursors unambiguously established that both the nine- and ten-membered enediyne cores were derived (minimally) from eight head-to-tail acetate units [13–16], it remained, until very recently [11••,12••,17••], controversial whether the enediyne cores are assembled by *de novo* polyketide biosynthesis or degradation from a fatty acid precursor. The enediyne cores bear no structural resemblance to any characterized polyketides, revealing little clue to what type of PKS is responsible for their biosynthesis.

As a model system for the nine-membered enediynes, we cloned the biosynthetic gene cluster for **9** from *Streptomyces globisporus* using both a deoxysugar biosynthesis gene and the *cagA* gene that encodes the C-1027 apoprotein as probes [17••,18]. Of the genes identified within

(Figure 1 Legend) Structures and mechanisms of bacterial PKSs. **(a)** Type I PKS consisting of non-iteratively acting domains as exemplified by DEBS for erythromycin (**1**) biosynthesis. **(b)** Type II PKS consisting of iteratively acting subunits as exemplified by TcmKLM for tetracenomycin (**2**) biosynthesis. **(c)** Type III PKS consisting of an iteratively acting single subunit as exemplified by RppA for flavolin (**3**) biosynthesis. KS and ACP domain or subunits are shown in orange and green, respectively. Atoms that were incorporated intact from the acyl CoA precursors to resultant polyketides are shown in bold.

Figure 2

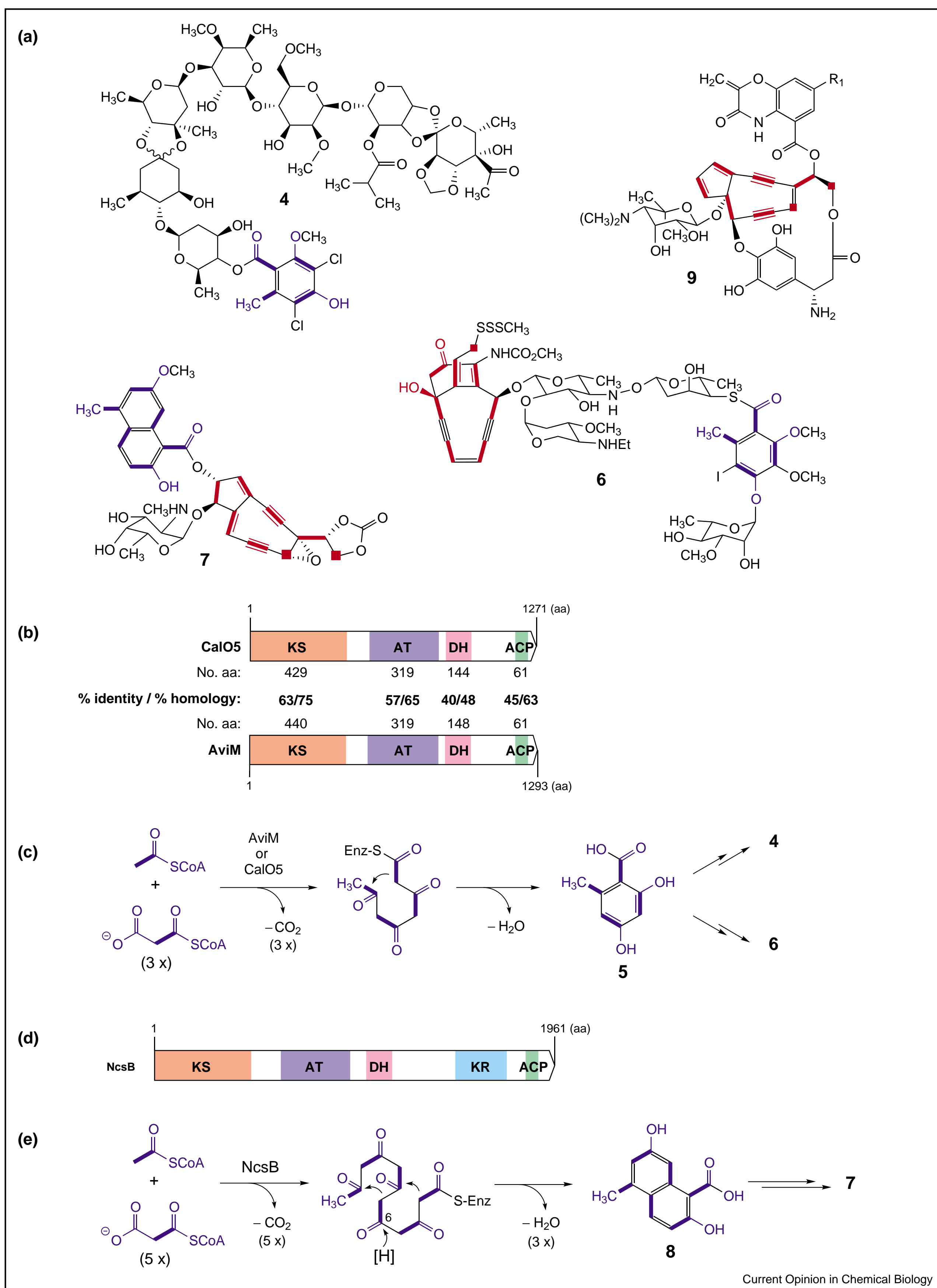
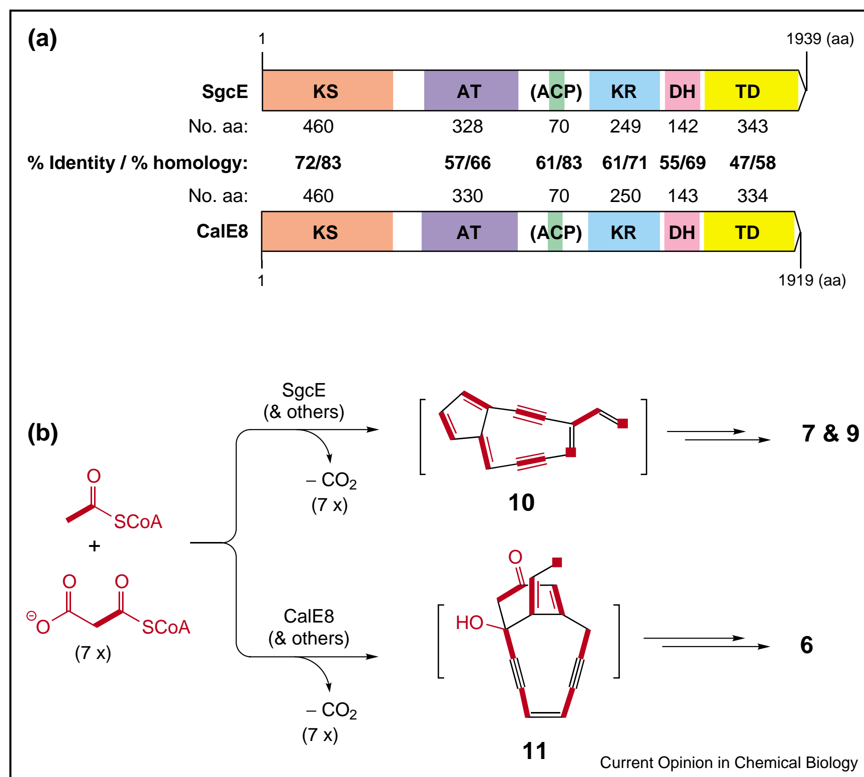


Figure 3



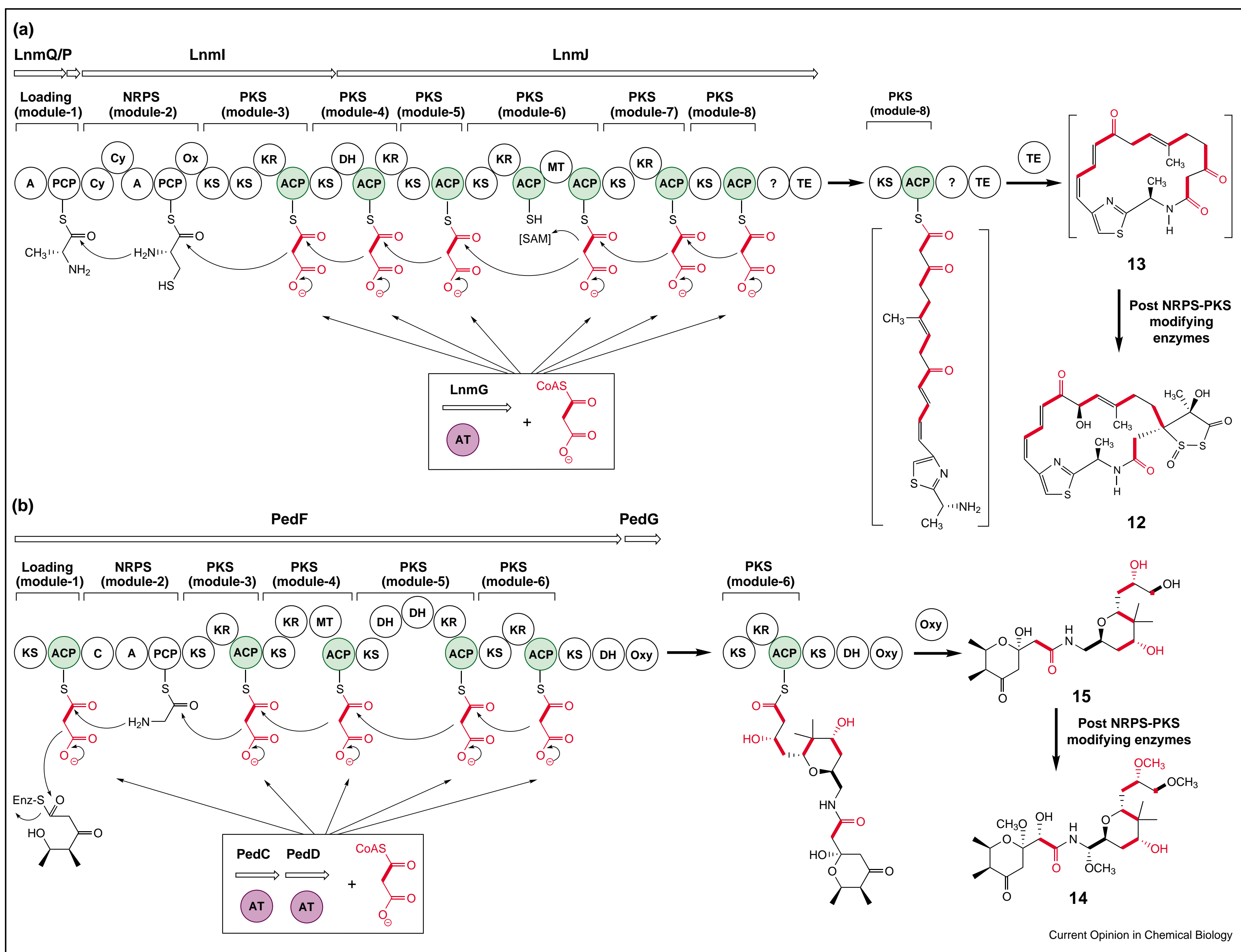
Mechanism of iterative type I PKSs for enediyne biosynthesis. **(a)** Domain organization of SgcE and CalE8 and amino acid sequence comparison between the two enediyne PKSs. **(b)** Biosynthetic hypothesis for SgcE- or CalE8-catalyzed biosynthesis of undefined polyketide intermediates from the acyl CoA precursors in an iterative process and their subsequent modifications by enediyne PKS associated enzymes into putative nine- or ten-membered enediyne cores such as **10** or **11**. Atoms that were incorporated intact from the acyl CoA precursors to the enediyne cores are shown in bold.

the C-1027 cluster, there is only one PKS gene, *sgcE*, whose deduced product consists of five domains. Four of the domains, KS, AT, KR and DH, are characteristic of known type I PKSs, and the fifth domain, residing at the C-terminus (terminal domain, TD), is unique to enediyne PKSs (Figure 3a). (It has been speculated on the basis of secondary structure predictions and solvent potential information that the region between AT and KR might contain an ACP domain and the TD domain might be a 4′-phosphopantetheinyl transferase [12••].) The involvement of SgcE in **9** biosynthesis was confirmed by gene inactivation and complementation: a *ΔsgcE* mutant lost its ability to synthesize **9**, and overexpression of *sgcE* in *ΔsgcE* restored **9** production [17••]. SgcE could be envisaged catalysing the assembly of a linear polyunsaturated intermediate from the acyl CoA precursor in an iterative process. The nascent intermediate, upon action of other enzyme activities, is subsequently desaturated to furnish the two yne groups and cyclized to afford the enediyne core (**10**) (Figure 3b) [17••]. The latter hypothesis is consistent with the findings that a group of five to ten genes, flanking the *sgcE* enediyne PKS gene, are highly conserved among all enediyne gene clusters characterized [11••,12••,17••]. These genes encode various oxidoreductases or proteins of unknown functions that are only associated with enediyne biosynthesis, serving as candidates for processing the nascent linear polyketide intermediate into an enediyne intermediate such as **10** or **11** (Figure 3b).

Complementary to **9**, Thorson and co-workers cloned the biosynthetic gene cluster for **6** from *M. calichensis* as a

(Figure 2 Legend) Structures and mechanism of iterative type I PKSs for aromatic polyketide biosynthesis. **(a)** Structures of avilamycin (**4**), calicheamicin (**6**), neocarzinostatin (**7**) and C-1027 (**9**). **(b)** Domain organization of AviM and CalO5 and amino acid sequence comparison between the two orsellinic synthases. **(c)** Biosynthetic hypothesis for AviM- or CalO5-catalyzed biosynthesis of orsellinic acid (**5**) from the acyl CoA precursors in an iterative process. **(d)** Domain organization of the NcsB PKS. **(e)** Biosynthetic hypothesis for NcsB-catalyzed biosynthesis of naphthalinic acid (**8**) from the acyl CoA precursors in an iterative process. Aromatic polyketide moiety and enediyne cores are shown in blue and red, respectively. Atoms that were incorporated intact from the acyl CoA precursors to resultant polyketide moieties are shown in bold.

Figure 4



model system for the 10-membered enediyne antibiotics by screening for genes conferring resistance to **6** [10,11••, 16]. Two PKS genes, *calE8* and *calO5*, were identified within the *cal* cluster, and *calO5*, as discussed above, encodes the orsellinic acid synthase (Figure 2b). The involvement of *calE8* in **6** biosynthesis was established by gene inactivation, and the resultant *ΔcalE8* mutant completely lost its ability to produce **6** [11••]. Remarkably, CalE8 exhibits head-to-tail sequence homology to and has the same domain organization of KS, AT, KR, DH and TD as SgcE (Figure 3a). The observed similarity between SgcE and CalE8 clearly suggests a common polyketide pathway for the biosynthesis of both nine- and ten-membered enediynes, despite the fact that their incorporation patterns by ^{13}C-labeled acetate feeding experiments were distinct — the two triple bonds of **7** and **9** were derived from intact acetate units whereas those of **6** were derived from adjacent acetate units (Figure 2a,3b) [13–16]. Like SgcE, the CalE8 PKS could be similarly envisaged catalysing the biosynthesis of a nascent polyunsaturated intermediate from the acyl CoA precursors in an iterative process. Modifications of the nascent polyketide intermediate by the 10-member enediyne cluster associated enzymes could then afford an enediyne core intermediate such as **11** (Figure 3b). SgcE and CalE8, therefore, represent a novel family of iterative type I PKSs, establishing a new paradigm for enediyne biosynthesis.

The SgcE and CalE8 enediyne PKSs have inspired further interest in searching for enediyne biosynthetic gene clusters. Using high-throughput genome-scanning methods to detect and analyse gene clusters involved in natural product biosynthesis, Farnet and co-workers identified multiple genetic loci, homologous to those of *sgcE* and *calE8*, from organisms including those that were not known as enediyne producers [12••]. Guided by the genomic information, they further demonstrated that enediyne production in these organisms could be induced under optimized growth conditions [12••]. We developed a PCR-based approach to access the enediyne PKS and its associated accessory genes directly and verified its effectiveness by cloning multiple enediyne PKS loci from organisms known to produce both the nine- and ten-membered enediyne antibiotics (Liu W, Ahlert J, Wendt-Pienkowski E, Thorson JS, Shen B, unpublished data). Taken together, these results not only further supported the iterative type I PKS paradigm for enediyne biosynthesis but also suggested a much greater diversity of enediyne natural products than previously appreciated.

Type I PKSs that lack the cognate AT domain and require a discrete AT enzyme acting iteratively *in trans* for reduced polyketide biosynthesis

Type I PKSs are characterized with non-iteratively acting modules, each of which minimally contains three domains of KS, AT and ACP that select, activate and catalyse a decarboxylative Claisen condensation between the extender unit and the growing polyketide chain, generating a β-ketoacyl-*S*-ACP intermediate. Optional domains are found between AT and ACP that carry out the variable set of reductive modifications of the β-keto group before the next round of chain extension. The order of modules in the PKS enzymes dictates the sequence of biosynthetic events, and the variation of domains within the modules affords the structural diversity observed in the resultant polyketide products (Figure 1a) [1].

We recently cloned the leinamycin (**12**; Figure 4) biosynthetic gene cluster from *Streptomyces atroolivaceus* S-140 that is characterized by a hybrid nonribosomal peptide synthetase (NRPS)–PKS system [19•,20••]. Surprisingly, the six PKS modules, encoded by the *lnmIJ* PKS genes, completely lack the cognate AT domain, whose missing activity instead was provided *in trans* by a discrete protein, LnmG (Figure 4a). Inactivation of *lnmGIJ* abolished **12** production, confirming that they are essential for **12** biosynthesis. LnmG was biochemically characterized *in vitro* as an AT enzyme, showing that it efficiently and specifically loaded malonyl CoA *in trans* to ACPs from all six LnmIJ PKS modules but not to the negative control of LnmP peptidyl carrier protein. These findings led us to propose that LnmG, acting iteratively, loads malonyl CoA *in trans* to the LnmIJ PKSs to initiate **12** biosynthesis; subsequently, decarboxylative condensations yield the nascent polyketide intermediate such as **13** that is finally converted into **12** (Figure 4a) [20••]. The biosynthesis of **12**, therefore, represented an unprecedented PKS architecture that is characterized by a discrete, iteratively acting AT protein that loads the extender units *in trans* to 'AT-less' type I PKS proteins for polyketide biosynthesis. Limited structural studies suggested that DEBS, the archetype of type I PKS, formed a parallel homodimer, possibly a helical structure [21,22]. At the core of the helix is a tetrahedron formed by the KS and AT domains of each PKS module with the ACP domain brought close to the KS domain of the opposite subunit [1,21,22]. This model is further supported by the recently solved crystal structure of the homodimer of the TE domain of DEBS [23]. The 'AT-less' PKS apparently deviates from the current type

(Figure 4 Legend) Structures and mechanism of type I PKSs that lack the cognate AT domain and require a discrete AT enzyme acting *in trans* for reduced polyketide biosynthesis. **(a)** Biosynthetic hypothesis for leinamycin (**12**) and modular organization of the 'AT-less' LnmIJ PKSs with the discrete LnmG AT enzyme loading malonyl CoA to all six PKS modules. **(b)** Biosynthetic hypothesis for pederin (**14**) and modular organization of the 'AT-less' PedFG PKSs with the discrete PedCD AT enzymes loading malonyl CoA to all five PKS modules. Atoms that were incorporated intact from the acyl CoA precursors to the polyketide moieties are shown in bold.

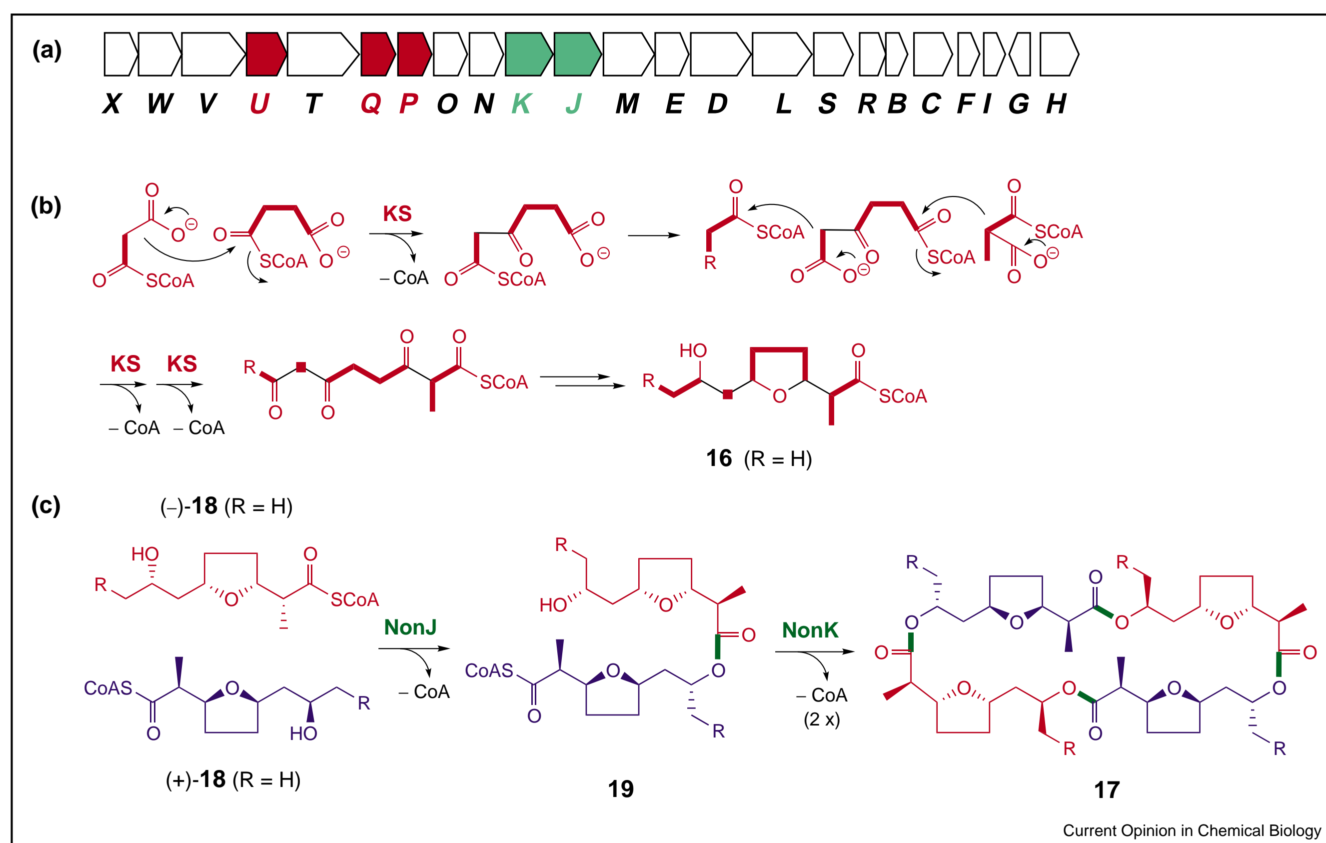
I PKS paradigm, suggesting an alternative model for type I PKS in which the KS and ACP domain of each module could minimally constitute the core structure.

Although LnmIJ remains to be the only 'AT-less' type I PKS that has been experimentally characterized, 'AT-less' type I PKS might be more abundant than we currently appreciate. Individual modules that lack the cognate AT domain have been noted from several type I PKS or hybrid NRPS–PKS systems [24–28,29•]. Piel [30•] recently reported the putative pederin (**14**) biosynthetic gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. The *pedF*, *pedG* and *pedH* genes together encode two NRPS modules and 10 PKS modules. The PKS modules consist of characteristic domains of type I PKS but completely lack the cognate AT domain. By contrast, two genes, *pedC* and *pedD*, encoding discrete AT enzymes, were identified within the *ped* cluster. In a mechanistic analogy to LnmG, it could be envisaged that PedC, PedD, or both, acting iteratively, load malonyl CoA *in trans* to the 'AT-less' PedF PKS to initiate **14** biosynthesis; subsequent decarboxylative condensations yield a nascent intermediate such as **15** that could be finally converted into **14** (Figure 4b).

Type II PKSs that act non-iteratively, lack ACP, utilize acyl CoAs as substrates directly and catalyse both C–C and C–O bond formation for macrotetrolide biosynthesis

Known PKSs use ACP to activate the acyl CoA substrates and channel the polyketide intermediates, except for type III PKSs that utilize acyl CoAs as substrates directly. Type III PKSs are structurally distinct from both type I and II PKSs. Although they possess the highly conserved Cys residue that is essential for PKS activity, the amino acid sequences of this Cys motif have no apparent similarity to those of the KSs of both type I and II PKSs. Intrigued by the unusual incorporation pattern of nonactic acid (**16**) by ^{13}C-labeled precursors (Figure 5b), we cloned and characterized the macrotetrolide (**17**) biosynthetic gene cluster from *Streptomyces griseus* to investigate its biosynthesis [31,32••,33••,34•].

Figure 5



Structures and mechanism of the macrotetrolide type II PKS that act non-iteratively, lack ACP, utilize acyl CoAs as substrate directly and catalyse both C–C and C–O bond formation. **(a)** The macrotetrolide biosynthetic gene cluster with the genes encoding C–C or C–O bond-forming type II KS shown in red and green, respectively. **(b)** Biosynthetic hypothesis for nonactic acid (**16**) from the acyl CoA precursors by three KSs (NonPQU) in a non-iterative process. Atoms that were incorporated intact from acyl CoA precursors to **16** are shown in bold. **(c)** The NonJK KS catalyzed stereospecific cyclotetramerization of nonactyl CoA (**18**) into macrotetrolides (**17**) via C–O bond formation. (–)-**18** is in red, (+)-**18** is in blue, and the newly formed C–O bonds are in green.

Our original hypothesis, based on the structural resemblance between **17** and other macrolide and polyether antibiotics, was biased towards a type I PKS for its biosynthesis. By contrast, the cloned *non* cluster was characterized with type II PKS genes, including *nonJKPQU* that encode five discrete KS enzymes, but no gene for ACP (Figure 5a) [32••]. Because sequence analysis clearly indicated that the five KS proteins are highly homologous to KSs of type I or II PKS, both of which are ACP dependent (Figure 1a and 1b), the lack of ACP from the macrotetrolide gene cluster was striking. To demonstrate that **17** biosynthesis is ACP-independent, the cloned gene cluster was expressed in *S. lividans* 1326, resulting in the production of **17** in the heterologous host [32••]. This result excluded the participation of *S. griseus* genes residing outside the cloned gene cluster in **17** biosynthesis. Gene inactivation, followed by fermenting the resultant mutants in the presence of exogenously added **16**, subsequently demonstrated that NonPQU were involved in the assembly of **16** from the acyl CoA precursors while NonJK were responsible for cyclotetramerization of (+)- and (−)-**16** into **17** [33••,34•]. Because the pathway proposed on the basis of feeding experiments suggested minimally three decarboxylative condensation steps, we proposed that NonPQU each would be responsible for one of the three steps for **16** biosynthesis from the four acyl CoA precursors (Figure 5b) [34•]. NonPQU together, therefore, represent a novel type II PKS that acts non-iteratively, lacks ACP and utilizes acyl CoA as substrates directly for polyketide biosynthesis (Figure 5b).

Because all known PKSs catalyse C–C bond formation [1–3], the finding that the NonJK KSs were responsible for the cyclotetramerization, via four C–O linkages was totally unexpected [33••]. A series of gene expression cassettes harboring *nonJK* in combination with various other genes within the cloned cluster were used to identify the minimal genes required for the C–O bond-forming cyclotetramerization steps, establishing that NonJKL were sufficient to biotransform **16** into **17** *in vivo*. Because NonL is a CoA ligase catalysing the formation of nonactyl CoA (**18**) from **16**, it was concluded that NonJK were responsible for the C–O bond-forming steps, acting directly on the CoA substrates of (+)- and (−)-**18** [33••]. All C–C bond-forming KSs are characterized with the Cys–His–His (Asn) catalytic triad, the His–His (Asn) of which are essential for malonyl-ACP or malonyl CoA decarboxylation to generate the corresponding carbon anion [35–37]. By contrast, NonJK are characterized by a mutated catalytic triad: Cys–Gly/Tyr–His, suggesting NonJK lack the decarboxylation activity. This is consistent with the proposal that NonJK catalyse C–O bond formation by using the –OH as the nucleophile directly (Figure 5c). Site-directed mutagenesis confirmed that the conserved Cys residue indeed played a catalytic role in the C–O bond-forming step [33••]. Finally, *in vitro* studies using cell-free preparations with overproduced NonJ or NonK revealed that NonJ catalyzes the stereospecific dimerization between (−)- and (+)-**18** to form (−)-nonactyl-(+)-nonactyl CoA (**19**) and NonK catalyzes the stereospecific cyclodimerization of **19** to afford **17** (Figure 5c) (Kwon H-J, Shen B: unpublished data). Thus, the NonJK KSs catalyse C–O bond-forming steps in **17** biosynthesis, acting non-iteratively, utilizing acyl CoA as substrates directly, and employing the same active site residue Cys as in KS catalysis for C–C bond formation in polyketide biosynthesis. NonJK, therefore, represent the first example of PKSs that could catalyse sequential condensation of acyl CoA substrates by forming new bonds other than C–C.

Conclusions and perspectives

The past decade has witnessed an exponential growth of genetic information on polyketide biosynthesis (and natural product biosynthesis in general). The type I, II and III PKS paradigms have been and will continue to be the cornerstones for polyketide biosynthesis. Innovations in methodologies for cloning biosynthetic gene clusters and advance in technologies for DNA sequencing and bioinformatics, however, have opened up many new opportunities to search for unprecedented biosynthetic machinery. It is in the spirit of this promise, one could not help but wonder how much do we know about polyketide biosynthesis? The examples presented here are by no means comprehensive [38••], but they do demonstrate the rich chemistry and architecture of PKSs beyond the type I, II and III paradigms. Structurally, PKSs could also be iterative type I, such as AviM, CalO5, NcsB, SgcE and CalE8; hybrid of non-iterative type I and iterative type II, such as LnmIJ/LnmG and PedFG/PedCD; and non-iterative type II, such as NonJKPQU. Mechanistically, PKSs could be ACP-dependent or ACP-independent, such as NonJKPQU. Synthetically, PKSs could make C–C bonds but also C–O bonds, such as NonJK. It is not a question if there are more novel PKSs, but rather how many there are, how to discover them, what are the evolutionary relationships among them, and what are the driving forces governing and advantages of having the myriad of PKS variants. One could remain optimistic that these novel systems will continue to give new insights into polyketide biosynthesis and provide new opportunities for combinatorial biosynthesis.

Acknowledgements

I am grateful to my co-workers who have made dreams into reality through their diligence and dedication. Studies described from the Shen laboratory were supported in part by the University of California BioSTAR Programme (Bio99-10045) and Kosan Biosciences, Inc., Hayward, CA; NIH grant CA78747, an NSF CAREER Award (MCB9733938); and an NIH Independent Scientist Award (AI51689).

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Staunton J, Weissman KJ: **Polyketide biosynthesis: a millennium review**. *Nat Prod Rep* 2001, **18**:380-416.

2. Shen B: **Biosynthesis of aromatic polyketides**. *Curr Top Chem* 2000, **209**:1-51.

3. Moore BS, Hopke JN: **Discovery of a new bacterial polyketide biosynthetic pathway**. *Chem Bio Chem* 2001, **2**:35-38.

4. Cortes J, Haydock SF, Roberts GA, Bevitt DJ, Leadlay PF: **An unusually large multifunctional polypeptide in the erythromycin producing polyketide synthase of *Saccharopolyspora erythraea***. *Nature* 1990, **348**:176-178.

5. Donadio S, Staver MJ, McAlpine JB, Swanson SJ, Katz L: **Modular organization of genes required for complex polyketide biosynthesis**. *Science* 1991, **252**:675-679.

6. Malpartida F, Hopwood DA: **Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host**. *Nature* 1984, **309**:462-464.

7. Motamedi H, Hutchinson CR: **Cloning and heterologous expression of a gene cluster for the biosynthesis of tetracenomycin C, the anthracycline antitumor antibiotic of *Streptomyces glaucescens***. *Proc Natl Acad Sci USA* 1987, **84**:4445-4449.

8. Funa N, Ohnishi Y, Fujii I, Shibuya M, Ebizuka Y, Horinouchi S: **A new pathway for polyketide synthesis in microorganisms**. *Nature* 1999, **400**:897-899.

9. Gaisser S, Trefzer A, Stockert S, Kirschning A, Bechthold A: **Cloning of an avilamycin biosynthetic gene cluster from *Streptomyces viridochromogenes* Tü57**. *J Bacteriol* 1997, **179**:6271-6278.

10. Whitwam RE, Ahlert J, Holman TR, Ruppen M, Thorson JS: **The gene *calC* encodes for a non-heme iron metalloprotein responsible for calicheamicin self-resistance in *Micromonospora***. *J Am Chem Soc* 2000, **122**:1556-1557.

11. •• Ahlert J, Shepard E, Lomovskaya N, Zazopoulos E, Staffa A, Bachmann BO, Huang K, Fonstein L, Czisny A, Whitwam RE *et al.*: **The calicheamicin gene cluster and its iterative type I enediyne PKS**. *Science* 2002, **297**:1173-1176.

Cloning, sequencing, and confirmation of the calicheamicin biosynthetic gene cluster are described. CalO5 was assigned as the orsellinic acid synthase, and CalE8 was assigned as the enediyne PKS, establishing a new paradigm for enediyne biosynthesis.

12. •• Zazopoulos E, Huang K, Staffa A, Liu W, Bachmann BO, Nonaka K, Ahlert J, Thorson JS, Shen B, Farnet CM: **A genomics-guided approach for discovering and expressing cryptic metabolic pathways**. *Nat Biotechnol* 2003, e-pub ahead of print, doi: 10.1038/nbt784.

A high-throughput genome scanning method to detect and analyse gene clusters involved in natural product biosynthesis was applied to the discovery of novel enediyne natural products, leading to the discovery of enediyne biosynthetic gene clusters from various organisms including those that were previously unknown as enediyne producers. Enediyne production by these organisms was confirmed by fermentation under the selective growth conditions.

13. Hensens OD, Giner J-L, Goldberg IH: **Biosynthesis of NCS Chrom A, the chromophore of the antitumor antibiotic neocarzinostatin**. *J Am Chem Soc* 1989, **111**:3295-3299.

14. Tokiwa Y, Miyoshi-Saitoh M, Kobayashi H, Sunaga R, Konishi M, Oki T, Iwasaki S: **Biosynthesis of dynemicin A, a 3-ene-1,5-diyne antitumor antibiotic**. *J Am Chem Soc* 1992, **114**:4107-4110.

15. Lam KS, Veitch JA, Golik J, Krishnan B, Klohr SE, Volk KJ, Forenza S, Doyle TW: **Biosynthesis of esperamicin A_1, and enediyne antitumor antibiotic**. *J Am Chem Soc* 1993, **115**:12340-12345.

16. Thorson JS, Shen B, Whitwam RE, Liu W, Li Y, Ahlert J: **Enediyne biosynthesis and self-resistance: a progress report**. *Bioorg Chem* 1999, **27**:172-188.

17. •• Liu W, Christenson SD, Standage S, Shen B: **Biosynthesis of the enediyne antitumor antibiotic C-1027**. *Science* 2002, **297**:1170-1173.

Cloning, sequencing, and confirmation of the C-1027 biosynthetic gene cluster are described. SgcE was assigned as the enediyne PKS. The high sequence homology and conserved domain organization between SgcE and CalE8 suggested a common polyketide pathway for both the nine- and ten-membered enediyne core, establishing a new paradigm for enediyne biosynthesis.

18. Liu W, Shen B: **Genes for production of the enediyne antitumor antibiotic C-1027 in *Streptomyces globisporus* are clustered with the *cagA* gene that encodes the C-1027 apoprotein**. *Antimicrob Agent Chemother* 2000, **44**:382-392.

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

A PCR method to clone a thiazole-containing natural product biosynthetic gene cluster was developed and applied to the cloning of the leinamycin biosynthetic gene cluster.

20. •• Cheng Y-Q, Tang G-L, Shen B: **Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis**. *Proc Natl Acad Sci USA* 2003, **100**:in press.

Sequencing and characterization of the leinamycin biosynthetic gene cluster revealed the ‘AT-less’ LnmIJ PKSs and the discrete LnmG AT enzyme. Biochemical characterization *in vitro* established that LnmG, acting iteratively, loads malonyl CoA *in trans* to all six PKS modules of LnmIJ, establishing a novel architecture of PKS. In a mechanistic analogy to LnmG, other ‘AT-less’ PKSs and their discrete AT counterparts were also discussed.

21. Staunton J, Caffrey P, Aparicio JF, Roberts G, Bethell S, Leadlay PF: **Evidence for a double-helical structure for modular polyketide synthase**. *Nat Struct Biol* 1996, **3**:188-192.

22. Gokhale RS, Lau J, Cane DE, Khosla C: **Functional orientation of the acyltransferase domain in a module of the erythromycin polyketide synthase**. *Biochemistry* 1998, **37**:2524-2528.

23. Tsai S-C, Miercke LJW, Krucinski J, Gokhale R, Chen JC-H, Foster PG, Cane DE, Khosla C, Stroud RM: **Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: versatility from a unique substrate channel**. *Proc Natl Acad Sci USA* 2001, **99**:14002-14007.

24. Duitman EH, Hamoen LW, Rembold M, Venema G, Seitz H, Saenger W, Bernhard F, Reinhardt R, Schmidt M, Ullrich C *et al.*: **The mycosubtilin synthetase of *Bacillus subtilis* ATCC6633: a multifunctional hybrid between a peptide synthetase, an amino transferase and a fatty acid synthase**. *Proc Natl Acad Sci USA* 1999, **96**:13294-13299.

25. Paitan Y, Alon G, Orr E, Ron EZ, Rosenberg E: **The first gene in the biosynthesis of the polyketide antibiotic TA of *Myxococcus xanthus* codes for a unique PKS module coupled to a peptide synthetase**. *J Mol Biol* 1999, **286**:465-474.

26. Paitan Y, Orr E, Ron EZ, Rosenberg E: **An unusual β-ketoacyl:acyl carrier protein synthase and acyltransferase motifs in TaK, a putative protein required for biosynthesis of the antibiotic TA in *Myxococcus xanthus***. *Gene* 2001, **203**:191-197.

27. Albertini AM, Caramori T, Scoffone F, Scotti C, Galizzi A: **Sequence around the 159° region of the *Bacillus subtilis* genome: the *pksX* locus spans 33.6kb**. *Microbiology* 1995, **141**:299-309.

28. Huang G, Zhang L, Birch RG: **A multifunctional polyketide-peptide synthetase essential for albicidin biosynthesis in *Xanthomonase albilineans***. *Microbiology* 2001, **147**:631-642.

29. • Zhu G, LaGier MJ, Stejskal F, Millership JJ, Cai X, Keithly JS: ***Cryptosporidium parvum*: the first protist known to encode a putative polyketide synthase**. *Gene* 2002, **298**:79-89.

Cloning and sequence analysis of a putative ‘AT-less’ PKS from the parasitic protist *C. parvum* is described. Neither the role of this cluster nor the functions of the gene product were known.

30. • Piel J: **A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles**. *Proc Natl Acad Sci USA* 2002, **98**:14808-14813.

Cloning and sequence analysis of the putative pederin biosynthetic gene cluster from an uncultured bacterial symbiont of *Paederus* beetles was described, revealing the *pedFGH* ‘AT-less’ PKS genes and *pedCD* discrete AT genes. Its involvement in pederin biosynthesis was based on sequence prediction only.

31. Smith WC, Xiang L, Shen B: **Genetic localization and molecular characterization of the *nonS* gene required for macrotetrolide biosynthesis in *Streptomyces griseus* DSM40695**. *Antimicrob Agent Chemother* 2000, **44**:1809-1817.

32. Kwon H-J, Smith WC, Xiang L, Shen B: **Cloning and heterologous expression of the macrotetrolide biosynthetic gene cluster revealed a novel polyketide synthase that lacks an acyl carrier protein**. *J Am Chem Soc* 2001, **123**:3385-3386.
•• Heterologous expression of the 23-gene *non* cluster is described, and resulted in the production of macrotetrolides. These results confirmed that macrotetrolide biosynthesis involves a type II PKS lacking ACP.

33. Kwon H-J, Smith WC, Scharon AJ, Hwang SH, Kurth MJ, Shen B: **C–O bond formation by polyketide synthases**. *Science* 2002, **297**:1327-1330.
•• *In vivo* genetic, mutagenetic and *in vitro* biochemical studies of macrotetrolide biosynthesis are described. These results established that NonJK catalyse the C–O bond-forming cyclotetramerization steps and utilize acyl CoAs as substrates directly, revealing an unprecedented enzyme activity for PKS.

34. Shen B, Kwon H-J: **Macrotetrolide biosynthesis: a novel type II polyketide synthase**. *Chem Rec* 2002, **2**:389-396.
• A critical review of macrotetrolide biosynthesis emphasizes how early studies inspired effort that eventually led to the discovery of a novel PKS.

35. He M, Varoglu M, Sherman DH: **Structural modeling and site-directed mutagenesis of the actinorhodin β-ketoacyl-acyl carrier protein synthase**. *J Bacteriol* 2000, **182**:2619.

36. Olsen JG, Kadziola A, von Wettstein-Knowles P, Siggaard-Andersen M, Lindquist Y, Larsen S: **The X-ray crystal structure of β-ketoacyl [acyl carrier protein] synthase I**. *FEBS Lett* 1999, **460**:46-52.

37. Jez JM, Ferrer J-L, Bowman ME, Dixon RA, Noel JP: **Dissection of malonyl-CoA decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase**. *Biochemistry* 2000, **39**:890-902.

38. Metz JG, Roessler P, Facciotti D, Levering C, Dittrich F, Lassner M, Valentine R, Lardizabal K, Domergue F, Yamada A *et al.*: **Production of polyunsaturated fatty acids by polyketide synthase in both prokaryotes and eukaryotes**. *Science* 2001, **293**:290-293.
•• Polyunsaturated fatty acid biosynthetic pathways were cloned from a marine bacteria *Shewanella* and a marine protist *Schizochytrium*, respectively, revealing that they were biosynthesized by PKSs with distinct mechanism and structure. Their role in polyunsaturated fatty acid biosynthesis was confirmed by heterologous expression. In addition to the examples discussed in the text, these PKSs demonstrated yet another diversity in polyketide biosynthesis.