The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases

David E Cane[1] and Christopher T Walsh[2]

Polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs) catalyze chain elongation from simple building blocks to create a diverse array of natural products. PKS and NRPS proteins share striking architectural and organizational similarities that can be exploited to generate entirely new natural products.

Addresses: [1]Department of Chemistry, Box H, Brown University, Providence, RI 02912-9108, USA. [2]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA.

E-mail: David_Cane@brown.edu
walsh@walsh.med.harvard.edu

Chemistry & Biology December 1999, **6**:R319–R325

1074-5521/99/$ – see front matter

Polyketides and nonribosomal peptides are extremely large classes of natural products that are assembled from simple acyl-coenzyme A or amino acid monomers. Among the vast number of known polyketides and nonribosomal peptides are numerous important pharmaceuticals, agrochemicals, and veterinary agents including antibiotics, immunosuppresants, anti-cholesterolemics, as well as antitumor, antifungal and antiparasitic agents. These agents include the widely used broad-spectrum macrolide antibiotic erythromycin A (**1**; Figure 1) and the immunosuppressant cyclosporin A (**2**; Figure 1).

Modular organization of PKS and NRPS enzymes

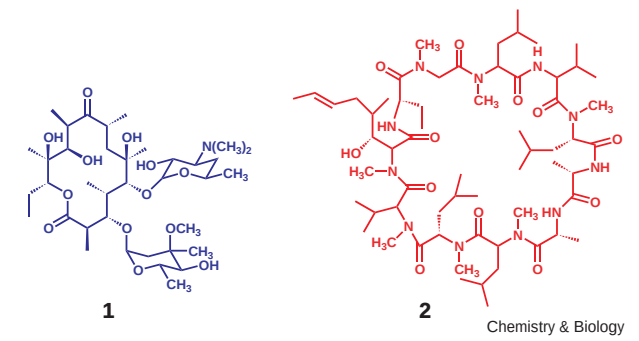
In spite of the enormous differences between polyketide and nonribosomal polypeptides, not only in overall structure, fundamental chemical building blocks and mechanisms of biosynthesis, numerous recent discoveries have highlighted striking similarities in the organization and function of the basic biosynthetic machinery by which these two families of metabolites are assembled [1–3]. Both are biosynthesized by polyfunctional megasynthases (200–2000 kDa) organized into repeated functional units known as modules, each of which is responsible for a discrete stage of polyketide or polypeptide chain elongation. Within each module are independently folded protein domains responsible for some combination of characteristic functional group modifications at each cycle of chain elongation. The modular megaproteins responsible for polyketide and polypeptide biosynthesis are known as polyketide synthases (PKSs) and nonribosomal polypeptide synthetases (NRPSs), respectively. (One group of polyketides, the aromatic polyketides, are biosynthesized by a distinct class of nonmodular proteins known as type II PKSs, in which the individual biochemical steps of polyketide biosynthesis are catalyzed by discrete subunits that are utilized repeatedly in the biosynthetic cycle [1].)

Modular PKS and NRPS proteins consist of one or more polyfunctional polypeptides, each of which is made up of as few as 1 or as many as 11 modules, with the amino-terminal to carboxy-terminal order of the individual modules paralleling the sequential order of polyketide or nonribosomal polypeptide chain elongation. In contrast, the corresponding structural genes for discrete PKS or NRPS open reading frames may be transcribed in an order that directly parallels the sequence of the eventual biosynthetic steps, as is the case for the erythromycin PKS (6-deoxyerythronolide B synthase, DEBS [4,5]), or in a partially convergent (the avermectin PKS [6]) or partially divergent manner (the rapamycin PKS [7]). Intriguingly, hybrid PKS/NRPS systems are also known, including rapamycin synthase (12 PKS modules and 1 NRPS module [7]) and yersiniabactin synthetase (3 NRPS modules and 1 PKS module [8]).

Core catalytic and carrier domains for iterated chain elongations

Each PKS or NRPS module is made up of a set of three mandatory or core domains, two of which are catalytic and one which acts as a carrier, that together are responsible for the central chain-building reactions of polyketide or polypeptide biosynthesis, as well as a variable set of auxiliary domains that mediate the modification (e.g. reduction, dehydration and methylation) of the newly extended polyketide or polypeptide chain. Both PKS and NRPS core domains utilize analogous acyl-chain elongation strategies in which the growing chain, tethered as an acyl-S-enzyme

Figure 1



Erythromycin A (1), a typical antibacterial polyketide produced by a modular PKS, and cyclosporin A (2), an immunosuppressant cyclic nonribosomal peptide produced by a NRPS.

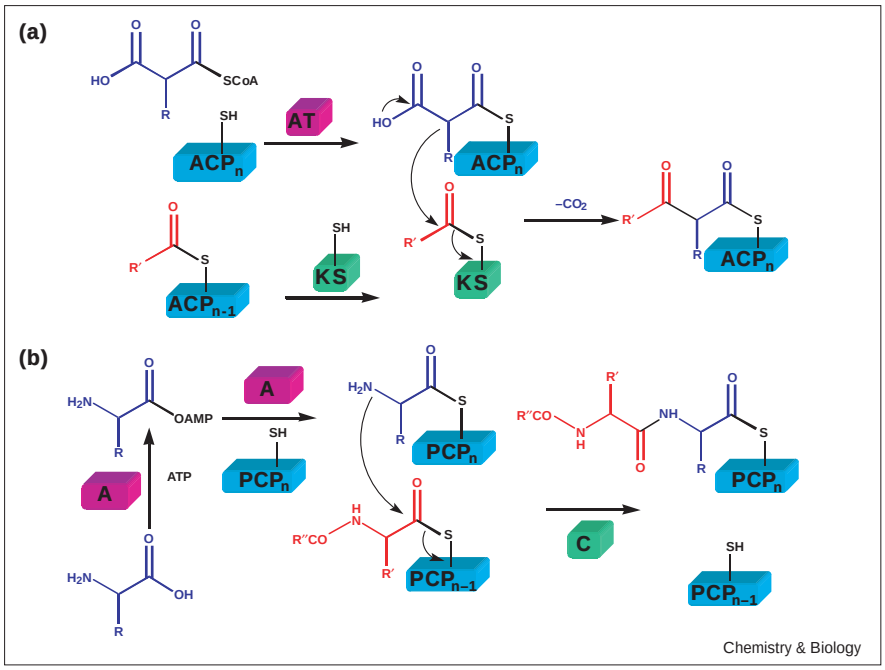
to the flexible 20 Å long phosphopantetheinyl arm of an (upstream) acyl carrier protein (ACP) or peptidyl carrier protein (PCP) domain acts as the electrophilic partner that undergoes attack by a nucleophilic chain-elongation unit, a malonyl- or aminoacyl-S-enzyme derivative, respectively, itself covalently bound to a downstream ACP/PCP domain (Figure 2a). In the case of a PKS, the fundamental chain-elongation reaction, a C–C bond-forming step, is mediated by a ketosynthase (KS) domain that catalyzes the transfer of the polyketide acyl chain to an active-site cysteine of the KS domain, followed by condensation with the methylmalonyl- or malonyl-S-ACP by a decarboxylative acylation of the malonyl donor unit. An additional essential component of the core PKS chain-elongation apparatus is an associated acyltransferase (AT) domain, which catalyzes the priming of the donor ACP sidearm with the appropriate monomer substrate, usually methylmalonyl- or malonyl-CoA.

The comparable core domains of an NRPS biosynthetic module function in a chemically distinct but architecturally and mechanistically analogous fashion. In the latter case, the key chain-building reaction, a C–N bond-forming reaction, involves the generation of the characteristic peptide bond by nucleophilic attack of the amino group of an amino acyl-S-PCP donor on the acyl group of an upstream electrophilic acyl- or peptidyl acyl-S-PCP chain, catalyzed by a condensation (C) domain (Figure 2b). It is not known whether this transfer, in mechanistic analogy to KS action, involves an acyl–enzyme intermediate between the electrophilic peptidyl chain and an active-site residue in the C domain. In functional analogy to the PKS AT domain, the core of the NRPS module utilizes an adenylation (A) domain to activate the donor amino-acid monomer as an amino acyl-AMP intermediate, which is then loaded onto the downstream PCP sidechain. Both the AT and A domains of the respective PKS and NRPS modules act as important gatekeepers for polyketide or polypeptide biosynthesis, exhibiting strict or at least high specificity for their cognate malonyl-CoA, methylmalonyl-CoA or amino acid substrates. For example the AT domains of all six DEBS modules show absolute specificity for (2*S*)-methylmalonyl-CoA [9].

Auxiliary PKS and NRPS domains

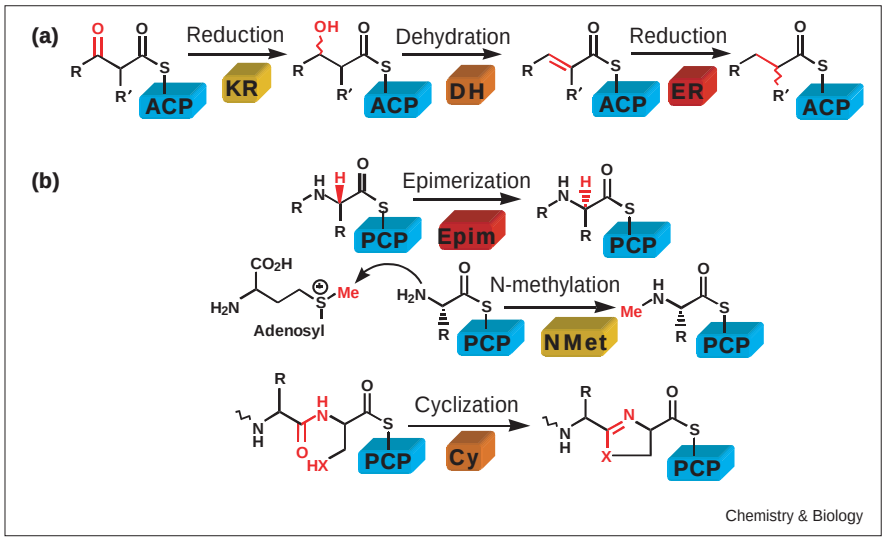
Although the core domains found in all PKS (KS–AT–ACP) or NRPS (C–A–PCP) modules are responsible for the chain-elongation steps, including the choice of chain-extension unit, supplementing these core chain-elongation domains are variable numbers of auxiliary domains that are responsible for modification of the growing polyketide or polypeptide acyl chain by a small set of iterated reactions including ketoreduction, dehydration and enoyl reduction (polyketides [1,2,4,5]) or epimerization, N-methylation, and heterocyclization (polypeptides [1–3]; Figure 3). The precise combination of these ancillary domains, acting locally on the acyl chains tethered at the proximal ACP/PCP waystation, contributes to the incredibly rich spectrum of structural complexity in the eventually formed mature polyketide or polypeptide product. Notably, rational modification of the number of modules and their constituent domains, in which conventional

Figure 2



Function and mechanism of PKS and NRPS core catalytic and carrier domains. **(a)** The acyltransferase (AT) domain of a PKS primes the flexible phosphopantetheinyl arm of the ACP domain with the malonyl (R = H) or methylmalonyl (R = Me) chain extender. The ketosynthase (KS) domain is acylated on an active-site cysteine residue by the acyl residue of a polyketide starter or chain elongation intermediate from an upstream ACP domain. The KS domain then catalyzes a decarboxylative acylation in which this acyl residue is transferred to the nucleophilic enol(ate) derived from the (methyl)malonyl-S-ACP partner. **(b)** In NRPS-catalyzed peptide bond formation, a dedicated adenylation (A) domain activates an amino-acid chain-extension unit (blue) as the aminoacyl adenylate, which is then transferred in a A-domain dependent reaction to the phosphopantetheinyl arm of the PCP domain. The condensation (C) domain then catalyzes nucleophilic attack of the amino nitrogen of the downstream (acceptor) aminoacyl-S-PCP (blue) on the electrophilic carbonyl carbon of the upstream (donor) peptidylacyl (or aminoacyl) residue (red).

Figure 3



Auxiliary (a) PKS and (b) NRPS domains mediating ketoreduction (KR), dehydration (DH), and enoylacyl reduction (ER) or epimerization (Epim), peptide N-methylation (NMet), and heterocyclization (Cy) of cysteine or serine residues. Each of these domains carries out its activities prior to transfer of the growing polyketide or polypeptide chain to the downstream module of release from the PKS/NRPS.

recombinant techniques have been used to delete, modify or swap individual domains with alternative domains taken from the same or foreign PKS/NRPS proteins, has allowed the engineered biosynthesis of numerous 'unnatural' natural products. For example, nearly 100 analogs of the erythromycin aglycone, 6-dEB, have already been prepared by such methods and many more can be expected [1,10].

Domains for chain initiation and termination

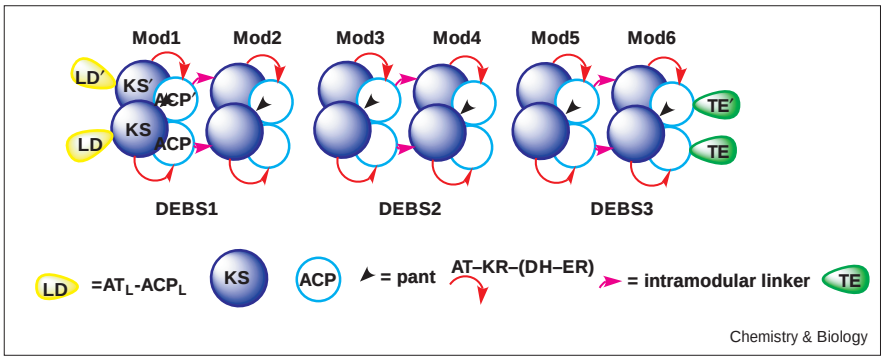
In addition to the basic subset of core and auxiliary domains found in each module, each PKS or NRPS also has a special set of dedicated domains responsible both for the initiation of acyl-chain assembly by loading of a starter unit onto the first, furthest upstream PKS/NRPS module, as well as a chain-terminating thioesterase (TE) domain, most often found fused to the carboxyl terminus of the last module, that is responsible for detachment of the most downstream covalent acyl enzyme intermediate and off-loading of the mature polyketide or polypeptide chain. In the case of the PKSs the TE domain normally catalyzes the release of the polyhydroxylated polyketide by macrolactonization as a consequence of intramolecular nucleophilic attack of a specific distal chain hydroxyl moiety on the acyl group of the polyketide acyl-S-ACP. (By contrast, the terminal TE of a modular polyether PKS should catalyze release of the free carboxylic acid by simple hydrolysis of the proximal polyketide acyl-S-ACP.) The analogous NRPS TE domains can likewise catalyze intramolecular macrocyclization (lactam or lactone formation) or simple hydrolysis, depending on the specific NRPS and its characteristic polypeptide product.

Models of quaternary structure

The higher order structure of the PKSs has been the subject of considerable experimentation as well as much ingenious speculation. Cross-linking, gel filtration, analytical ultrafiltration and partial proteolysis, as well as a series of mutant complementation experiments [11–13], have firmly established that each module of DEBS is a homodimer in which the active-site cysteine of the KS domain of each module lies close to the thiol residue of the phosphopantetheinyl arm of the ACP domain of the paired PKS polypeptide. A similar architecture has been deduced for the genetically, biochemically and mechanistically closely related vertebrate fatty acid synthases (FASs), in which the core KS cysteine residue and the thiol of the ACP of the paired strand have been shown to be separated by no more than 5 Å [14–17]. By contrast with unimodular FAS proteins, however, the polymodular PKSs must also transfer their products, the growing polyketide acyl chains, to the KS domains of an appropriate downstream module, which can lie either immediately downstream in the same multifunctional polypeptide *(in cis)* or in a distinct modular PKS subunit *(in trans)* altogether. Mutant complementation studies have shown that acyl transfer *in cis* takes place from the ACP of one module to the KS domain of the downstream module of the same polypeptide strand [12]. The factors controlling the timing and yield of transfer of growing polyketide acyl chains *in trans* to the target KS domain in the proper downstream PKS module are still obscure, however, but no doubt involve some balance of both protein–protein recognition and the intrinsic specificity of the downstream KS domain for its cognate polyketide substrate [18].

Although purified FAS proteins have been available for more than two decades, only recently have the three-dimensional structures of FAS components begun to be solved [19–21]. Pending direct information on the structure of any modular FAS or PKS, several schematic models

Figure 4



Tetrahedral model for modular PKS topology illustrated by the hexamodular 6-deoxyerythronolide B synthase (DEBS). The KS–ACP/KS′–ACP′ pairs of complementary peptide strands of a single module are organized as an $\alpha_2\beta_2$ tetrahedral homodimer. The intervening core AT and auxiliary KR–(DH–ER) domains are arranged along strands of variable length corresponding formally to edges of the tetrahedral array. The 20 Å long flexible phosphopantetheinyl arms of the ACP domains contact the KS of the complementary PKS strand within the same module and can also transfer polyketide acyl chains to the downstream KS domain for a successive round of chain elongation and processing. Short intramolecular linkers connect the ACP domains with the downstream KS domains on the same strand in polymodular PKS subunits, with the ACP_n–KS_{n+1} pairs also packed in a tetrahedral topology. The proposed topology also results in the pairing of the TE domains at the carboxyl terminus of module 6.

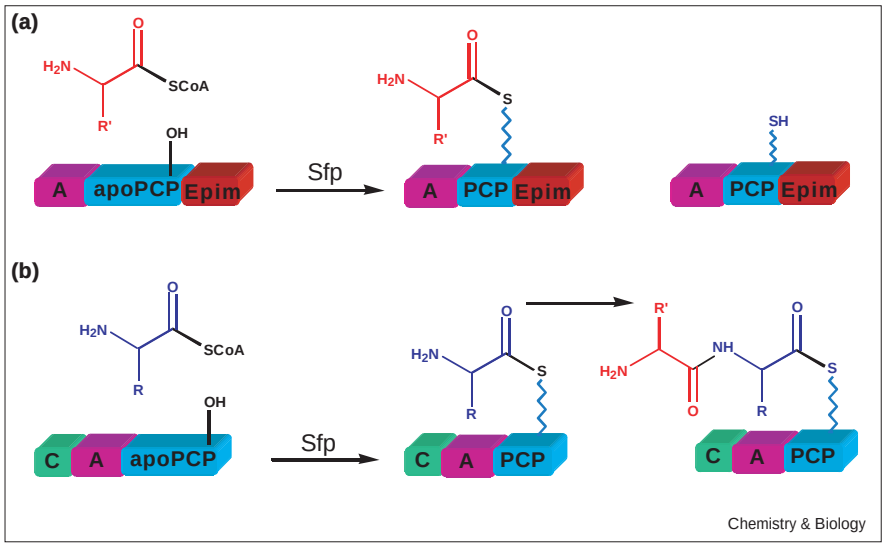
of modular PKS architecture have been suggested. Researchers at Cambridge University [11] have proposed an interesting model in which the homodimeric chains of the polymodular DEBS are intertwined in a double-helical array, keeping the paired KS, ACP and carboxy-terminal TE domains in close proximity, while allowing the required functional interaction between KS and ACP domains within the same module but on complementary strands. A somewhat different topological arrangement is proposed in Figure 4, in which the core KS and ACP pairs of each strand form an α′–β′ heterodimer, with each member of these two sets of domains situated at the corners of an distorted tetrahedron. The intervening AT domains, as well as any auxiliary KR, DH or ER domains would be strung along loops of variable length formally (but not actually) parallel to an edge of the tetrahedron. Modules can then readily interact by stacking individual tetrahedral arrays edge to edge, with the ACP_n–ACP_n' domains paired with the downstream KS_{n+1}–KS_{n+1}' domains in a complementary α′–β′ tetrahedral heterodimer linked either covalently or noncovalently by polypeptide linkers. In the latter model, the sequential stacking of tetrahedral modules could involve an overall intertwined (double-helical) topology or an equally plausible nontwisted connectivity. Both the tetrahedral model proposed here as well as the earlier double-helical model [11] not only are consistent with the apparent pairing of the carboxy-terminal TE domains, but can account for all the experimental data bearing on PKS structure and that of the closely related FASs. Of particular note is the recent observation by Smith and coworkers [17] that cross-linking of the active-site cysteine thiols of a fatty acid synthase with the thiol moiety of the ACP domains results not only in interstrand cross-linking, as previously reported [14], but in a substantial proportion of intrastrand cross-linking as well. The latter result suggests that the ACP domains of both strands are located close to the KS domains of both the parent and the paired FAS or PKS strand. Such a pairing is explicitly incorporated into the tetrahedral model and could, in principle, be reconciled with the double-helical model as well, although the current formulation of the latter model would seem to emphasize the pairing of the KS–AT/KS′–AT′ domains, with the coupled ACP domains diverging from the core. Definitive distinction between the two proposed topologies and refinement of the model will depend on direct studies of PKS module structure using X-ray crystallography, facilitated by the heterologous expression of individual PKS modules in *Escherichia coli*. By contrast, although the structure of a phenylalanine adenylation domain has been reported [22], no information is available yet on the three-dimensional organization of NRPS modules. Experiments analogous to those already carried out on FAS and PKS modules, such as cross-linking, mutant complementation, and native molecular weight determinations, should have high priority in investigations of NRPS enzymology in the near future. These studies will be materially facilitated by advances such as those summarized below involving the expression of active NRPS modules in *E. coli*. The existence of functional PKS–NRPS hybrids, however, strongly suggests that the NRPS proteins will turn out to be homodimeric as well.

Crosstalk: specificity and permissiveness in PKS and NRPS function

One of the greatest challenges in the study of polyketide and nonribosomal peptide biosynthesis continues to be the understanding of the relative roles played by modular organization and by molecular recognition in the control of the complex program of chemical events that results in the assembly of monomer substrates into oligomeric mature products. To address this issue at a fundamental biochemical level, it has been necessary to factor out the relative contributions to specificity of the domains responsible for

Figure 5

Investigation of selectivity of a C domain of the tyrocidine NRPS for (a) donor (electrophilic, red) or (b) acceptor (nucleophilic, blue) aminoacyl-S-PCP substrates. (a) Donor site selectivity was determined by using the *Bacillus* phosphopantetheine synthetase, Sfp, to directly load five different aminoacyl groups (R′, red) on the apo-PCP of the normal phenylalanine A–PCP–Epim module and mixing with the proline C–A–PCP acceptor module primed with [^{3}H]-Pro (blue), then determining the rate of formation of the resulting aminoacyl–prolyl peptide. (b) To measure acceptor site specificity, Sfp was used to load five different aminoacyl-S-CoAs (R, blue) on the proline C–A–PCP module, followed by mixing with the [^{3}H]-Phe-labeled A–PCP–Epim acceptor module and measuring the rate of product formation.

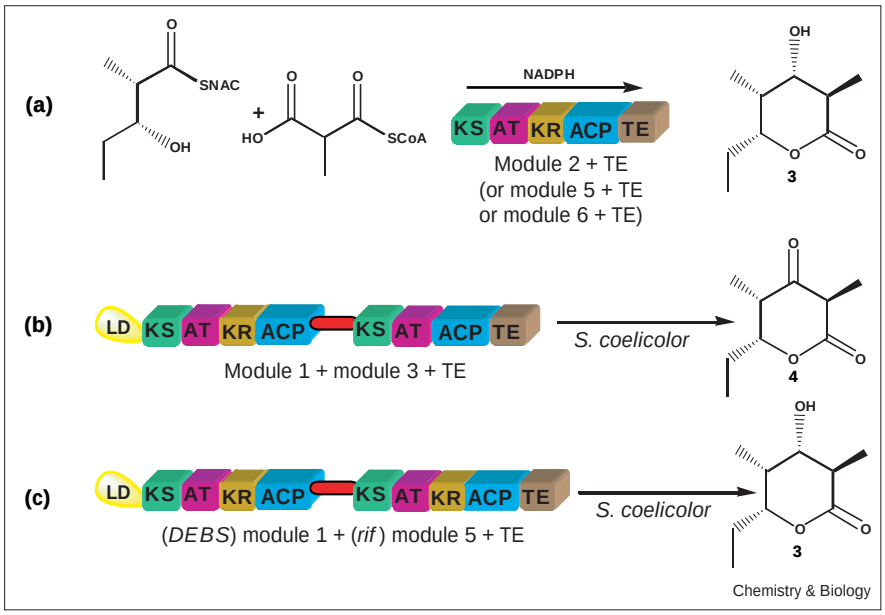


priming of the ACP/PCP thiol residues with chain-extension monomers (the AT and A domains of PKS and NRPSs, respectively) and the domains that mediate chain elongation (the KS and C domains, respectively). These questions cannot be addressed in more than a qualitative, or at best semi-quantitative, manner at the intact cell level. Even at the cell-free level, deconvolution of multiple enzymatic steps mediated by polymodular proteins is confounded by their arrays of covalently tethered acyl-enzyme intermediates and by the intrinsically narrow substrate specificity of the AT and A domains. To overcome the intrinsic specificity of monomer priming domains, one approach that has been used effectively has been to replace the native AT and A domains of a given module with heterologous AT or A domains with a different substrate specificity (e.g. malonyl-CoA for methylmalonyl-CoA transferase [23,24] or amino-acid adenylation swaps [25].) Very recently, Stachelhaus, Mootz and Marahiel [26] have shown that site-directed mutagenesis, informed by a detailed analysis of primary sequence and crystallographic data, can be used with impressive effectiveness to re-engineer the specificity of native amino-acid adenylation domains.

Independent technical advances, reported in two recent papers from our laboratories [27,28], have provided a powerful new set of tools to address these problems. Belshaw *et al.* [27] have now shown that it is possible to bypass the A domain of an NRPS module altogether by using a *Bacillus* phosphopantetheinyl transferase, known as Sfp [29], in combination with synthetic aminoacyl-CoA substrates [30] to effect stoichiometric transfer of intact aminoacyl-phosphopantetheines to inactive apo forms of isolated PCP (thiolation) domains of an NRPS (Figure 5). The resulting aminoacyl-S-PCPs could be used to establish that the first C domain of tyrocidine synthetase shows low selectivity at the upstream electrophilic residue (D-phenylalanine) and higher selectivity at the nucleophilic downstream residue (L-proline) in the formation of the chain-initiating D-Phe–L-Pro dipeptidyl-enzyme intermediate. Not only is the determination of C domain specificity of fundamental importance, but the same approach should be broadly applicable to the study of other NRPSs and PKSs as well, because of the broad tolerance of Sfp for acyl-CoAs as substrate. Sfp could in principle be used, therefore, to prepare a variety of 2-substituted malonyl-S-ACPs by loading the corresponding malonyl-CoA analogs onto the apo-ACP of a PKS, either in an intact module or as an isolated domain. Similarly, it should be possible to load aminoacyl residues onto PKS ACP domains and substituted malonyl or other acyl derivatives onto NRPS PCP domains in order to study the intrinsic specificity of KS and C domains for the nucleophilic component of the chain-building condensation reactions.

A contemporaneous advance from the Stanford–Brown team in the study of the erythromycin PKS system has shed important new light on the factors that are important for successful intermodular transfer of polyketide chain-elongation intermediates, and allowed study of the recognition properties of individual KS domains for the electrophilic component of the PKS condensation reaction. Gokhale *et al.* [28] have described the heterologous expression in *Escherichia coli* of single DEBS modules with the TE domain fused to the carboxyl terminus. Once again, Sfp was used to generate the holo-ACP form of the PKS, using CoASH as the donor, either *in vivo* or *in vitro*. Incubation of the natural diketide analog (2*S*,3*R*)-2-methyl-3-hydroxypentanoyl-S-(N-acetylcysteamine) (-NAC) with module

Figure 6



Formation of triketide lactones by recombinant modules. (a) DEBS module 2 + TE, expressed in *E. coli* with a suitably engineered amino-terminal peptide linker catalyzes the *in vitro* conversion of the diketide-SNAC thioester and methylmalonyl-CoA plus NADPH to the triketide lactone **3**. The same reaction is also catalyzed by recombinant module 5 + TE and module 6 + TE. (b) When expressed in *Streptomyces coelicolor* CH999, a chimeric PKS in which DEBS module 1, with its natural upstream loading domains, is fused to DEBS module 3 + TE by an engineered peptide linker generates the predicted triketide ketolactone **4**, resulting from transfer of the diketide product of module 1 to the KS domain of module 3. (c) A chimeric PKS in which DEBS module 1, with its natural upstream loading domains, is fused to rifamycin module 5 + DEBS TE by an engineered peptide linker generates **3**, resulting from transfer of the diketide product of module 1 to the KS domain of the heterologous *rif* module.

2 + TE, module 5 + TE, or module 6 + TE in the presence of methylmalonyl-CoA and NADPH led to formation of the lactonized triketide **3** (Figure 6a). Kinetic analysis showed nearly identical K_m values and similar k_{cat} values for all three modules with respect to the diketide substrate. The latter finding is particularly surprising in view of the fact that that the corresponding diketide-ACP intermediate is the natural substrate only for KS2 in the complete DEBS [18]. The key to the successful heterologous expression of individual DEBS modules in active form was the recognition that specific linkers at the amino terminus of each module were critical to obtaining active enzyme. This methodology promises to be widely applicable to the expression of a wide range of natural and hybrid PKS and NRPS modules in a convenient bacterial host such as *E. coli*. Excitingly, it was also found that use of appropriate intermodular linkers also allowed for the first time successful chain transfer between normally noncontiguous modules. Thus recombinant strains of *Streptomyces coelicolor* CH999 harboring fusions of DEBS module 1 + module 3 + TE produced the expected triketide ketolactone **4** (Figure 6b). Even more remarkably, a fusion of module 1 of DEBS with module 5 of the rifamycin PKS from *Amycolatopsis mediterranei* + TE also gave the predicted triketide lactone **3** when expressed in *S. coelicolor* (Figure 6c). It is expected that appropriate extensions of this technology should not only facilitate a wider variety of heterologous modular PKS fusions, but allow the engineering of hybrid PKS–NRPS fusions as well. In this manner, the factors controlling module–module interaction as well as the intrinsic specificity of both KS and C domains for both the electrophilic and nucleophilic components of the key polyketide and polypeptide chain building reactions can be evaluated on a systematic basis [31].

Prospects

Methodologies are now available to test and optimize replacement and repositioning of both core and auxiliary catalytic domains in PKSs and give direction to searches for similar rules for NRPS domain interchanges beyond the existing A domain swaps. For example, it should be possible by construction of appropriate chimeric modules and use of the Sfp enzyme, to test both qualitatively and quantitatively the ability of the KS domains of individual PKS modules to process amino acid substrates, or condensation (C) domains of NRPS systems to handle malonyl derivatives. The basic units of exchange in combinatorial biosynthesis could turn out to be libraries of intact modules, with optimized protein–protein and biochemical interactions among internal domains, which have been adapted for crosstalk by the use of appropriately designed linkers. The ability to reorganize protein assembly line domains would also facilitate creation of new hybrid PKS–NRPS assembly lines and lead to combinatorial variation in such therapeutically interesting hybrid molecules as rapamycin, the bleomycins and the epothilones. Particular KS domain and C domains that have shown useful selectivity/promiscuity ratios for the C–C and C–N bond-forming chain-elongation steps would be of particular interest in the reprogramming of the PKS and NRPS assembly lines.

Acknowledgements

Work from the authors' laboratories summarized here is supported by grants from the NIH to D.E.C. (GM22172) and C.T.W. (GM20011). We would also like to acknowledge the close collaboration of Professor Chaitan

Khosla of Stanford University in much of the work described, as well as for his comments on the manuscript.

References

1. Polyketide and nonribosomal polypeptide biosynthesis [Special issue]. (1997). *Chem. Rev.* **97**, 2463-2705, *(*Cane, D.E., guest ed.).
2. Cane, D.E., Walsh, C.T. & Khosla, C. (1998). Harnessing the biosynthetic code: combinations, permutations and mutations. *Science* **282**, 63-68.
3. Konz, D. & Marahiel, M.A. (1999). How do peptide synthetases generate structural diversity? *Chem. Biol.* **6**, R39-R48.
4. Donadio, S., Staver, M.J., McAlpine, J.B., Swanson, S.J. & Katz, L. (1991). Modular organization of genes required for complex polyketide biosynthesis. *Science* **252**, 675-679.
5. Cortes, J., Haydock, S.F., Roberts, G.A., Bevitt, D.J. & Leadlay, P.F. (1990). An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. *Nature* **348**, 176-178.
6. Ikeda, H., Nonomiya, T., Usami, M., Ohta, T. & Omura, S. (1999). Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*. *Proc. Natl Acad. Sci USA* **96**, 9509-9514.
7. Aparicio, J.F., *et al*., & Leadlay, P.F. (1996). Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymatic domains in the modular polyketide synthase. *Gene* **169**, 9-16.
8. Gehring, A.M., Mori, I., Perry, R.D. & Walsh, C.T. (1998). The nonribosomal peptide synthetase HMWP2 forms a thiazoline ring during biogenesis of yersiniabactin, an iron-chelating virulence factor of *Yersinia pestis*. *Biochemistry* **37**, 11637-11650.
9. Marsden, A.F.A., Caffrey, P., Aparicio, J.F., Loughran, M.S., Staunton, J. & Leadlay, P.F. (1994). Stereospecific acyl transfers on the erythromycin-producing polyketide synthase. *Science* **263**, 378-380.
10. McDaniel, R., Thamchaipenet, A., Gustafsson, C., Fu, H., Betlach, M. & Ashley, G. (1999). Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel 'unnatural' natural products. *Proc. Natl Acad. Sci. USA* **96**, 1846-1851.
11. Staunton, J., Caffrey, P., Aparicio, J.F., Roberts, G.A., Bethell, S.S. & Leadlay, P.F. (1996). Evidence for a double-helical structure for modular polyketide synthases. *Nat. Struct. Biol.* **3**, 188-192.
12. Kao, C.M., Pieper, R., Cane, D.E., & Khosla, C. (1996). Evidence for two catalytically independent clusters of active sites in a functional modular polyketide synthase. *Biochemistry* **35**, 12363-12368.
13. Pieper, R., Ghokale, R., Luo, G., Cane, D.E. & Khosla, C. (1997). Purification and characterization of bimodular and trimodular derivatives of the erythromycin polyketide synthase. *Biochemistry* **36**, 1846-1851.
14. Stoops, J.K. & Wakil, S.J. (1980). Yeast fatty acid synthetase: structure–function relationship and nature of the β-ketoacyl synthetase site. *Proc. Natl Acad. Sci. USA* **77**, 4544-4548.
15. Witkowski, A., Joshi, A. & Smith, S. (1996). Fatty acid synthase: *in vitro* complementation of inactive mutants. *Biochemistry* **35**, 10569-10575.
16. Joshi, A.K., Witkowski, A. & Smith, S. (1997). Mapping of functional interactions between domains of the animal fatty acid synthase by mutant complementation *in vitro*. *Biochemistry* **36**, 2316-2322.
17. Witkowski, A., Joshi, A.K., Rangan, V.S., Falick, A.M., Witkowska, H.E. & Smith, S. (1999). Dibromopropanone cross-linking of the phosphopantetheine and active-site cysteine thiols of the animal fatty acid synthase can occur both inter- and intrasubunit. Reevaluation of the side-by-side, antiparallel subunit model. *J. Biol. Chem.* **274**, 11557-11563.
18. Chuck, J.A., McPherson, M., Huang, H., Jacobsen, J.R., Khosla, C. & Cane, D.E. (1997). Molecular recognition of diketide substrates by a β-ketoacyl-acyl carrier protein synthase domain within a bimodular polyketide synthase. *Chem. Biol.* **4**, 757-766.
19. Holak, T.A., Nilges, M., Prestegard, J.H., Gronenborn, A.M. & Clore, G.M. (1988). Three-dimensional structure of acyl carrier protein in solution determined by nuclear magnetic resonance and the combined use of dynamical simulated annealing and distance geometry. *Eur. J. Biochem.* **175**, 9-15.
20. Huang, W., Jia, J., Edwards, P., Dehesh, K., Schneider, G. & Lindqvist, Y. (1998). Crystal structure of β-ketoacyl-acyl carrier protein synthase II from *E. coli* reveals the molecular architecture of condensing enzymes. *EMBO J.* **17**, 1183-1191.
21. Serre, L., Verbree, E.C., Dauter, Z., Stuitje, A.R. & Derewenda, Z.S. (1995). The *Escherichia coli* malonyl-CoA:acyl carrier protein transacylase at 1.5 Å resolution. Crystal structure of a fatty acid synthase component. *J. Biol. Chem.* **270**, 12961-12964.
22. Conti, E., Stachelhaus, T., Marahiel, M.A. & Brick, P. (1997). Structural basis for the activation of phenylalanine in the nonribosomal biosynthesis of gramicidin S. *EMBO J.* **16**, 4174-4183.
23. Oliynyk, M., *et al*., & Leadlay, P.F. (1996). A hybrid modular polyketide synthase obtained by domain swapping. *Chem. Biol.* **3**, 833-839.
24. Lau, J., Fu, H., Cane, D.E. & Khosla, C. (1999). Dissecting the role of acyltransferase domains of modular polyketide synthases in the choice and stereochemical fate of extender units. *Biochemistry* **38**, 1643-1651.
25. Stachelhaus, T., Schneider, A. & Marahiel, M.A. (1995). Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. *Science* **269**, 69-72.
26. Stachelhaus, T., Mootz, H.D. & Marahiel, M.A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* **6**, 493-505.
27. Belshaw, P.J., Walsh, C.T. & Stachelhaus, T. (1999). Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science* **284**, 486-489.
28. Gokhale, R.S., Tsuji, S.Y., Cane, D.E. & Khosla, C. (1999). Dissecting and exploiting intermodular communication in polyketide synthases. *Science* **284**, 482-485.
29. Lambalot, R.H., *et al*., & Walsh, C.T. (1996). A new enzyme superfamily – the phosphopantetheinyl transferases. *Chem. Biol.* **3**, 923-936.
30. Gehring, A.M., Lambalot, R.H., Vogel, K.W., Drueckhammer, D.G. & Walsh, C.T. (1997). Ability of *Streptomyces* spp. acyl carrier proteins and coenzyme A analogs to serve as substrates *in vitro* for *E. coli* holo-ACP synthase. *Chem. Biol.* **4**, 17-24.
31. Ranganathan, A., *et al*., & Leadlay, P.F. (1999). Knowlege-based design of bimodular and trimodular polyketide synthases based on domain and module swaps: a route to simple statin analogues. *Chem. Biol.* **6**, 731-741.