Polyketide and Nonribosomal Peptide Antibiotics: Modularity and Versatility

Christopher T. Walsh, *et al.*

Science **303**, 1805 (2004);
DOI: 10.1126/science.1094318

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.

48. G. Shapiro *et al.*, *Proc. Am. Soc. Clin. Oncol.* **22**, 199 (2003).
49. L. Meijer, E. Raymond, *Acc. Chem. Res.* **36**, 417 (2003).
50. T. Schindler *et al.*, *Science* **289**, 1938 (2000).
51. B. Nagar *et al.*, *Cancer Res.* **62**, 4236 (2002).
52. N. Schiering *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 12654 (2003).
53. C. Breitenlechner *et al.*, *Structure (Cambridge)* **11**, 1595 (2003).
54. W. F. De Azevedo, Jr. *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2735 (1996).
55. T. G. Davies *et al.*, *Nature Struct. Biol.* **9**, 745 (2002).
56. L. Tong *et al.*, *Nature Struct. Biol.* **4**, 311 (1997).
57. C. E. Fitzgerald *et al.*, *Nature Struct. Biol.* **10**, 764 (2003).
58. J. Stamos, M. X. Sliwkowski, C. Eigenbrot, *J. Biol. Chem.* **277**, 46265 (2002).
59. L. Shewchuk *et al.*, *J. Med. Chem.* **43**, 133 (2000).
60. B. Zhao *et al.*, *J. Biol. Chem.* **277**, 46609 (2002).
61. L. N. Johnson *et al.*, *Pharmacol. Ther.* **93**, 113 (2002).
62. E. De Moliner, E. D. Lowe, K.-Y. Cheng, L. N. Johnson, unpublished observations; Protein Data Bank ID 1PKD.
63. D. Komander *et al.*, *Biochem. J.* **375**, 255 (2003).
64. R. A. Engh, D. Bossemeyer, *Pharmacol. Ther.* **93**, 99 (2002).
65. H. S. Cho *et al.*, *Nature* **421**, 756 (2003).
66. T. A. Young, B. Delagoutte, J. A. Endrizzi, A. M. Falick, T. Alber, *Nature Struct. Biol.* **10**, 168 (2003).
67. S. Holton *et al.*, *Structure (Cambridge)* **11**, 1329 (2003).
68. G. Kontopidis *et al.*, *Structure (Cambridge)* **11**, 1537 (2003).
69. A. S. Corbin, P. La Rosee, E. P. Stoffregen, B. J. Druker, M. W. Deininger, *Blood* **101**, 4611 (2003).
70. M. E. Gorre *et al.*, *Science* **293**, 876 (2001).
71. M. Azam, R. R. Latek, G. Q. Daley, *Cell* **112**, 831 (2003).
72. L. N. Johnson, M. E. Noble, D. J. Owen, *Cell* **85**, 149 (1996).

REVIEW

Polyketide and Nonribosomal Peptide Antibiotics: Modularity and Versatility

Christopher T. Walsh

Polyketide (PK) and nonribosomal peptides (NRP), constructed on multimodular enzymatic assembly lines, often attain the conformations that establish biological activity by cyclization constraints introduced by tailoring enzymes. The dedicated tailoring enzymes are encoded by genes clustered with the assembly line genes for coordinated regulation. NRP heterocyclizations to thiazoles and oxazoles can occur on the elongating framework of acyl-*S* enzyme intermediates, whereas tandem cyclic PK polyether formation of furans and pyrans can be initiated by post–assembly line epoxidases. Macrocyclizations of NRP, PK, and hybrid NRP-PK scaffolds occur in assembly line chain termination steps. Post–assembly line cascades of enzymatic oxidations also create cross-linked and cyclized architectures that generate the mature scaffolds of natural product antibiotics. The modularity of the natural product assembly lines and permissivity of tailoring enzymes offer prospects for reprogramming to create novel antibiotics with optimized properties.

Polyketides and nonribosomal peptides comprise two families of natural products biosynthesized with comparable logic by multimodular enzymes acting in assembly line arrays. The monomeric building blocks are organic acids or amino acids, respectively. Copolymerization via mixed modules of enzymatic machinery can be used to assemble hybrid polyketide-nonribosomal peptide molecules of useful structural complexity and therapeutic activity.

Since the topic of harnessing the biosynthetic code for natural products was reviewed in this journal a half-dozen years ago (*1*), knowledge about gene clusters that encode enzymes for the biosynthesis of many therapeutically useful PK, NRP, and hybrid PK-NRP natural products has increased dramatically, with hundreds of identified or predicted clusters available in public databases. In just one example, the genome of the avermectin producer, *Streptomyces avermitilis* (*2*), revealed 24 additional polyketide synthase (PKS) or nonribosomal peptide synthase (NRPS) clusters for unidentified secondary metabolites, suggesting natural product biosynthetic capacity is tremendously underestimated by the products detected in fermentations. The systematic variation of culture conditions to up-regulate transcriptomes selectively, and thereby the encoded proteomes, may begin to coax microbes to reveal their full secondary metabolome capacity (*3*). The need for new antibiotics to combat the widening circle of resistance that ensues each time such a drug is launched into widespread human use puts a premium on the discovery of therapeutically useful molecules. Deciphering the structure of heretofore cryptic microbial secondary metabolites may well hasten the discovery of new antibiotic scaffolds useful as novel therapeutics or as reprogrammable scaffolding elements for semisynthetic modifications.

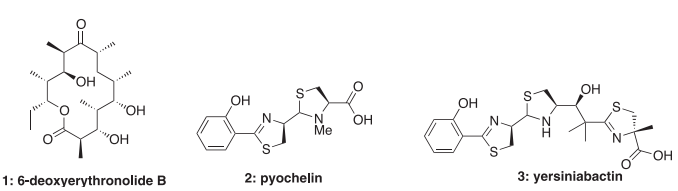


Fig. 1. Polyketide and nonribosomal peptide natural products from reconstituted assembly lines.

The molecular logic used by PKS and NRPS multimodular assembly lines in acyl and peptidyl chain initiation, elongation, and termination reactions has largely been deconvoluted and put to use for mechanistic and synthetic efforts, including making new analogs of the estrogen receptor antagonist R1128 (*4*–*6*). The role of domains and subunits that tailor the growing chains both on the assembly lines and after release have been inventoried (*7*–*9*). Reconstitution of assembly lines from purified assembly line proteins in vitro has been reported for deoxyerythronolide B **1** (DEB) (*10*), the NRP siderophore pyochelin **2** (*11*), and the hybrid PK-NRP iron chelator yersiniabactin **3** (Ybt) (*12*) (Fig. 1). PKS and PKS-NRPS gene clusters encoding DEB (*13*) and Ybt (*14*), respectively, have been moved from producing streptomycetes to *E. coli* and reconstituted in vivo, opening up the genetic toolbox that has been developed over decades for gene and protein studies in *E. coli*. The detection of intersubunit linkers (*15*, *16*) and their portability between PKS and NRPS subunits (*17*) has enabled initiation of mix-and-match module strategies to create new molecular skeletons.

Heterocyclization and Macrocyclization in Assembly Line Enzymology

Among the features that endow polyketide and nonribosomal peptide natural products with high affinity for biological targets are

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA. E-mail: christopher_walsh@hms.harvard.edu

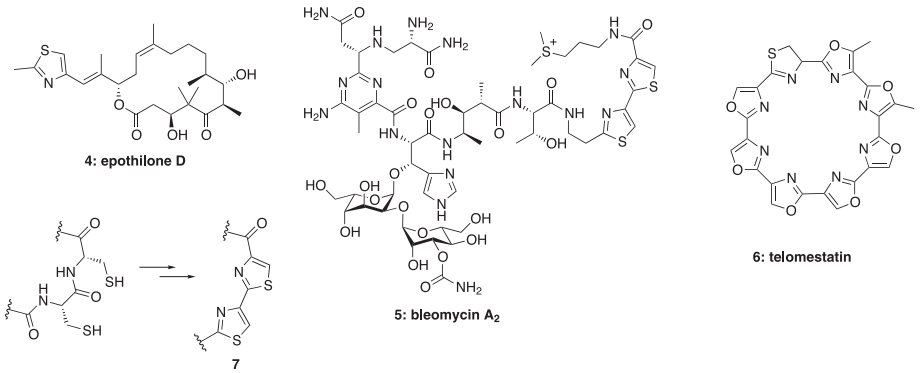


Fig. 2. Heterocyclization in NRP and PK-NRP natural products.

the constraints imposed by cyclization. This includes heterocyclization of cysteines and serines or threonines to thiazolines (pyochelin) (*11*) and oxazolines (vibriobactin) (*18*). Subsequent oxidation produces thiazoles (epothilone **4** and bleomycin **5**) (*19*) and oxazoles (telomestatin **6** and disorazoles) (*20*, *21*), respectively, whereas thiazoline reduction yields thiazolidines (pyochelin **2** and yersiniabactin **3**) (Fig. 2). The five-atom heterocycles can interact with proteins (GE2270), RNA (thiostrepton), and DNA targets (bleomycin) (*22*). The tandem conversion of two cysteines to a bithiazole moiety **7** during bleomycin assembly (*23*) generates a DNA intercalating moiety. The cyclization domains in NRPS assembly lines responsible for the cyclodehydration of peptide bonds to heterocycles have been characterized (*24*, *25*) and, if generally portable, might allow generation of tandem bithiazole peptide libraries. Among the most remarkable of the natural products in this class is telomestatin, a nanonomolar inhibitor of telomerase (*20*), with all eight amino acid constituents heterocyclized: five serines to oxazoles, two threonines to methyloxazoles, and the single cysteine to thiazoline. This highly constrained octamer is also macrocyclized, presumably by enzymatic machinery.

The second cyclization constraint is macrocyclization in polyketides, such as the antibiotic erythromycin and the antitumor epothilones, as well as NRPs including the recently approved antibiotic daptomycin **8**; ramoplanin **9**, currently in clinical trials; and the PK-NRP hybrid pristinamycin components of the antibiotic synercid **10** (Fig. 3). Macrocyclization occurs at the most downstream way station of PKS and NRPS assembly lines during chain release by thioesterase (TE) domains (*9*, *26*). In competition with intermolecular hydrolytic chain release, an intrachain nucleophile can capture the full-length acyl-*O*-TE intermediate. In PKS assembly lines, the internal nucleophiles are alcohol oxygens, leading to lactones, e.g., 14-membered in erythronolide and 16-membered in epothilone (Scheme 1, A and B). The size of polyketide macrolactones varies from 22 to 40 atoms in the polyether ionophore antibiotics. The gene cluster for the 38-membered nystatin **11** has been reported (*27*), and the regioselective NysK macrocyclizing thioesterase domain has been identified.

In NRPS assembly lines, the intramolecular chain capture by amines leads to peptidolactams: the amino terminus reacts in "head-to-tail" cyclic amide formation in tyrocidine **12** (Scheme 1C), whereas the ε-NH_2 of Lys^7 attacks Asn^{12} to create the "lariat" structure of the antibiotic bacitracin **13** (Fig. 4). Capture by side chain hydroxyl groups of threonyl or tyrosyl residues yields peptidolactones in the lipopeptide daptomycin **8** and fengycin **14**. Lipopeptides can also be cyclized through the β-OH of a β-OH fatty acyl moiety as in surfactin **15**, creating a range of regiospecific connections, based on substrate and TE combinations (*9*, *27*).

The isolated TE domains of some PKS and NRPS assembly lines retain autonomous capacity to catalyze such regio- and stereospecific cyclizations, exemplified by epothilone synthase TE (Scheme 1B) (*28*) and the tyrocidine synthetase TE (*29*, *30*) (Scheme 1C). Tyrocidine synthetase TE shows remarkable promiscuity, cyclizing peptidyl thioesters containing 6 to 14 residues, D-phenyllactyl[1] analogs to peptidolactones (*31*) instead of peptidolactams, and showing high promiscuity for substitution at positions 5 to 8 of the linear decapeptidyl substrates for cyclization. Peptidyl-*O*-esters on solid phase beads are substrates, permitting library approaches to tyrocidine synthetase TE-mediated cyclization with optimization of antibiotic selectivity (*32*). This same TE is also able to macrocyclize PK-NRP thioesters to produce novel hybrid macrocycles (*33*). These approaches are amenable to library strategies both for substrates and TE domains from various PKS and NRPS assembly lines. Titration of the thermodynamic activation of peptidyl thioesters for excised TE domains has been

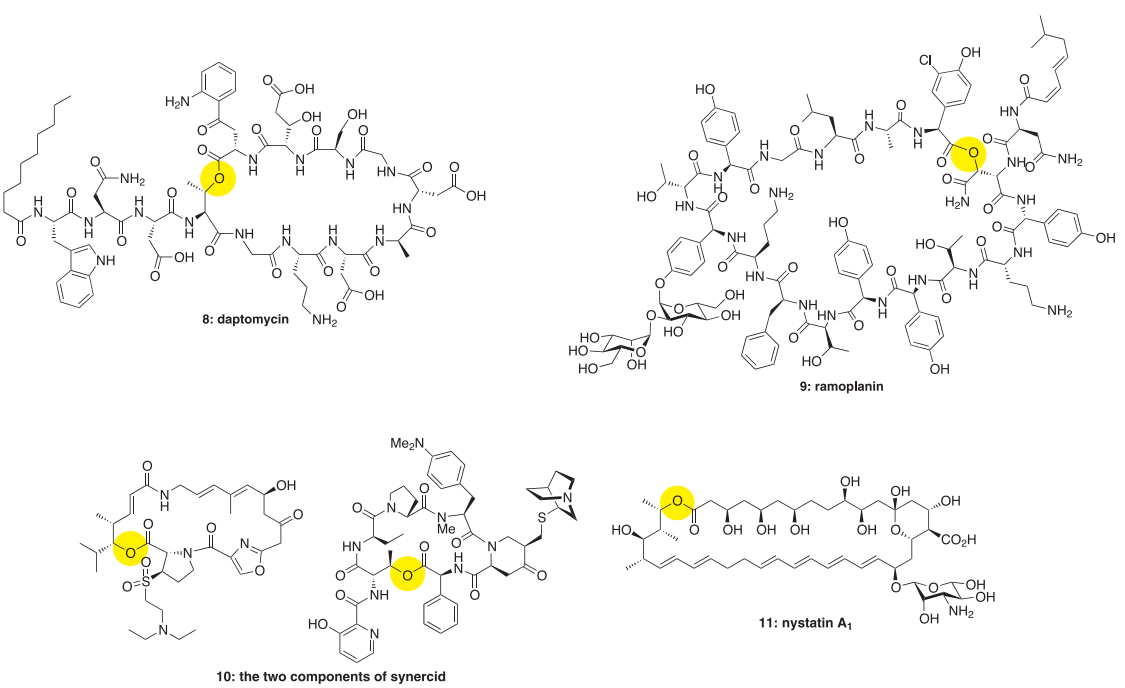
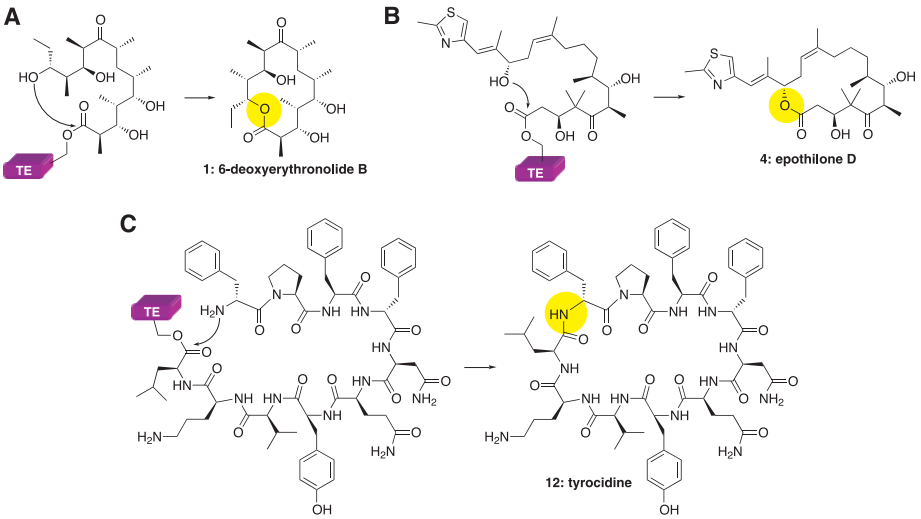


Fig. 3. Macrocyclized PK and NRP antibiotics.



Scheme 1. Macrocyclizations mediated by the C-terminal thioesterase domains of PKS and NRPS multimodular assembly lines: (**A**) 14-membered PK macrolide; (**B**) 16-membered PK macrolide; and (**C**) head-to-tail cyclization to the cyclic decapeptide tyrocidine.

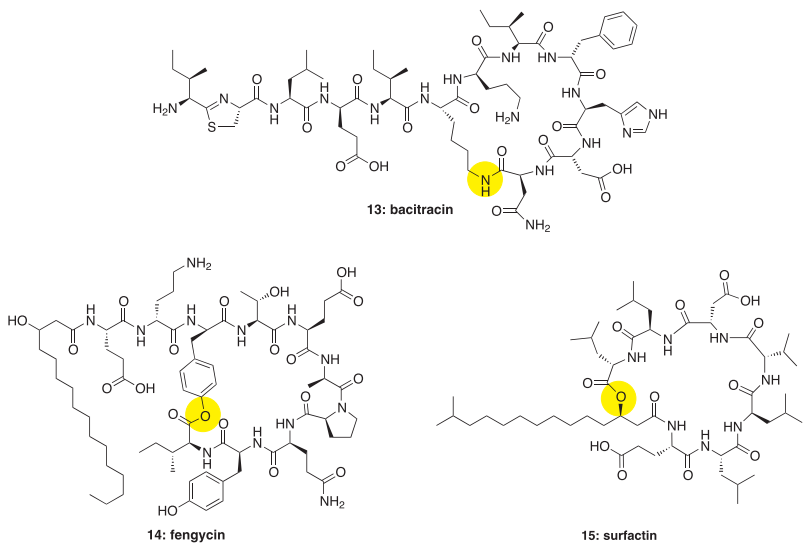
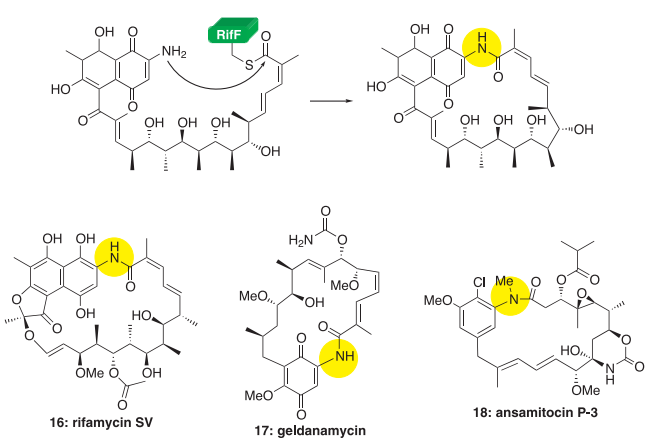


Fig. 4. Macrolactam and macrolactone nonribosomal peptides.



Scheme 2. Macrolactam formation in the biosynthesis of ansa-bridged rifamycin and geldanamycin antibiotics by discrete amide synthase subunits.

demonstrated by the observation that the TE domain from fengycin synthetase will not macrocyclize lipopeptidyl *N*-acetylcysstamine thioesters but will proceed when more activated *S*-aryl leaving groups are employed, with faithful regio- and stereospecificity of macrocyclization on unprotected substrates (*34*).

In vivo swaps of modules and deletions have been reported in the surfactin system (*35*). Directed incorporation of fluorinated aromatic amino acids in both calcium-dependent antibiotic and balhimycin fermentations further augments the possibilities for novel variants (*36*, *37*). Control of macrolactone ring size in vivo in polyketide assembly of erythromycin and pikromycin scaffolds has been investigated (*38*, *39*), and mechanisms have been determined for module skipping and stuttering to cyclize 12-, 14-, and 16-membered rings. The ability to mix and match monomers and vary macrocyclic ring sizes and points of connectivity in vitro and/or in vivo by such approaches along with evolution of TE domain specificities might be of special interest in PK-NRP hybrids such as the pristinamycin antibiotics and the antitumor agents leinamycin (*40*) and cryptophycin (*41*).

A variation of the molecular logic for macrocyclization in natural product assembly lines occurs in the ansa antibiotics, including the antitubercular drugs of the rifamycin group **16**, the Hsp90 inhibitor geldanamycin **17,** and the maytansine family antitumor agent ansamitocin **18.** The biosynthetic genes are clustered as typical for PKS assembly lines (*42*–*44*), but the final amide bond formation in macrocyclization occurs by action of separate amide synthase subunits rather than embedded TE domains (Scheme 2). These amide synthases are homologous to amine acyl transferases and may permit macrocyclization of variant linear precursors to novel ansa bridged macrocycles.

Alternate Biosynthetic Strategies to Constrain and Modify Antibiotic Scaffolds

Two additional enzymatic strategies to modify and thus constrain the conformations and architecture are exemplified by: (i) the glycopeptide antibiotics of the vancomycin **19** and teicoplanin **20** class (*45*, *46*) and (ii) the polyether subfamily of polyketides, exemplified by monensin **21**, nanchangmycin **22**, and nonactin **23**, all of whose biosynthetic gene clusters have been sequenced in the past few years (*47*–*49*) (Fig. 5).

Aryl crosslinks. The distinguishing characteristic of vancomycin, teicoplanin, balhimycin, and dozens of related glycopeptides is the oxidative crosslinks that occur between electron-rich aryl side chains. These convert what would be acyclic, floppy heptapeptides into rigid, cup-shaped scaffolds that are the pharmacophores for binding the D-Ala-D-Ala

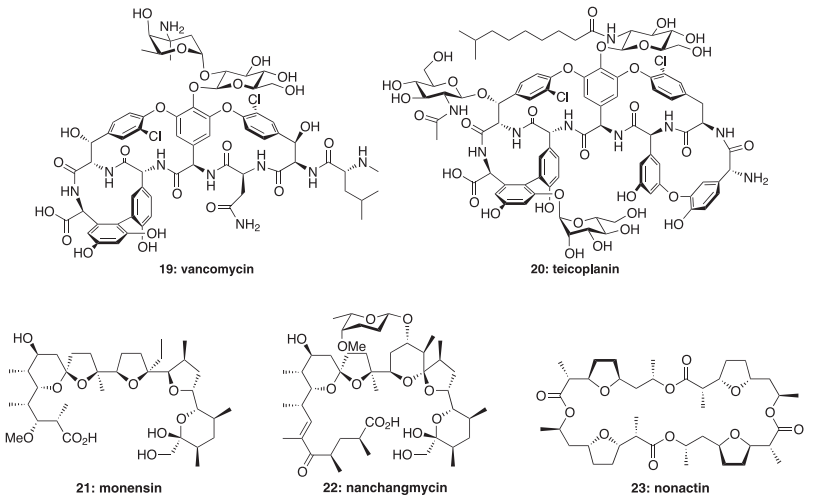
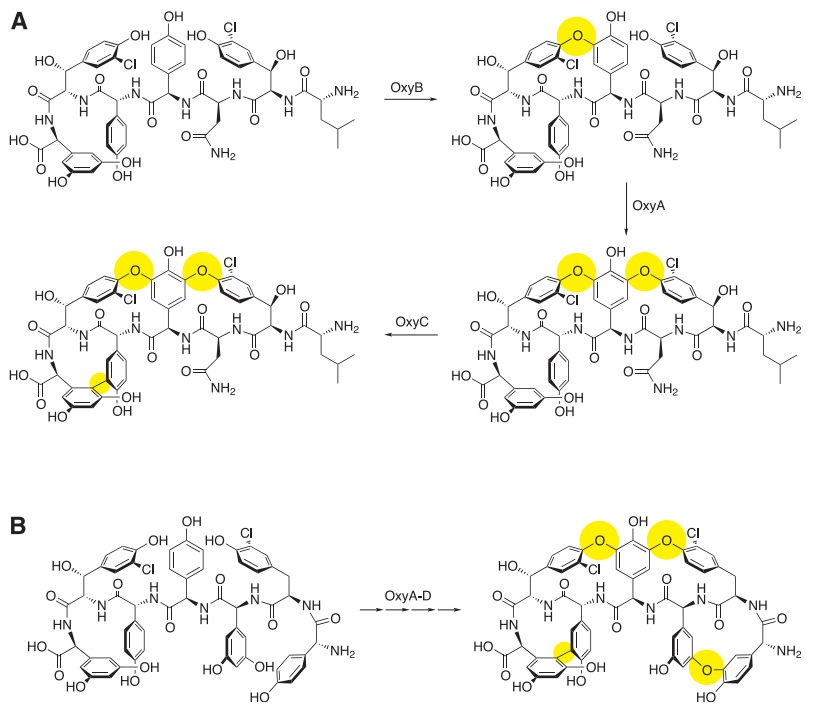
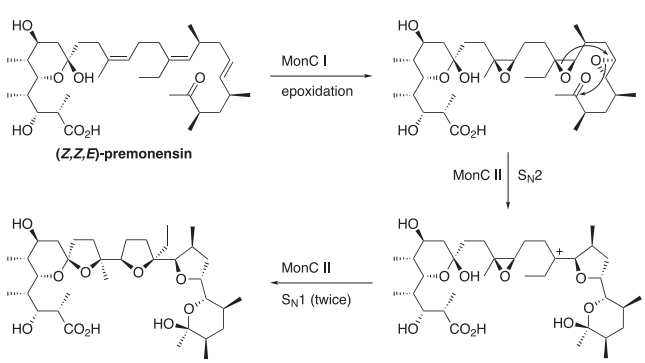


Fig. 5. Glycopeptide (NRP) and polyether (PK) antibiotics.



Scheme 3. Oxidative crosslinking of the aryl side chains in glycopeptide antibiotics by hemeprotein oxidases: (**A**) three crosslinks (2-4, 4-6, 5-7) introduced to the heptapeptide aglycone scaffold of vancomycin and (**B**) a fourth crosslink (1-3) introduced in the heptapeptide scaffold of teicoplanin subfamily antibiotics.



Scheme 4. Proposal for the enzymatic formation of tetrahydrofuran and tetrahydropyran rings in the ionophore monensin via enzymatic epoxidation of olefins, followed by intramolecular capture [adapted from (*48*)].

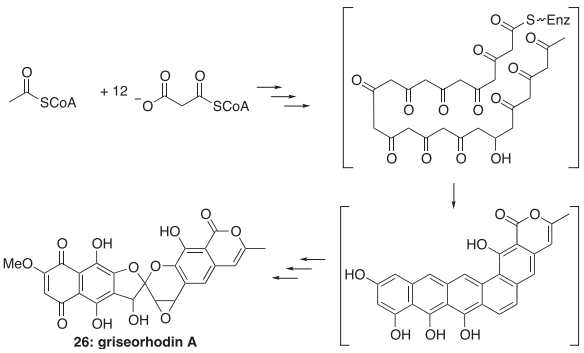
termini of bacterial peptidoglycan strands. The gene clusters for chloroeremomycin, balhimycin, teicoplanin, and two teicoplanin analogs, A47934 and A40926 (*50*–*54*), reveal three to four conserved tandem hemeproteins, OxyA to OxyD, that have the spectral and architectural features of cytochrome P450s. Knockouts in the balhimycin cluster (*55*) reveal that OxyB first catalyzes the 4-6 phenolic crosslink, OxyA the 2-4 phenolic crosslink, and OxyC the 5-7 C-C crosslink (Scheme 3A). Presumably this is one-electron oxidative chemistry on prefolded peptide substrates in the hemeprotein active sites, with both regiospecificity and atropisomer stereospecificity control (*45*) during crosslink formation. The teicoplanin subfamily gene clusters encode the fourth hemeprotein, consistent with the observation that, in teicoplanin, residues 1 and 3 are now electron-rich aromatics and undergo a fourth oxidative crosslink, such that all seven residues (1-3, 2-4, 4-6, 5-7) are covalently constrained (Scheme 3B). If these aromatic residue crosslinking oxygenases could be coaxed to work on other aryl peptides, electronically and sterically tuned, then libraries of highly constrained peptide architectures would be possible that might select high-affinity biological targets, e.g., variants of arylomycins (*56*).

Polyether formation and macrocyclization. The biosynthesis of polyether antibiotics proceeds via multimodular PKS assembly lines, 60 domains in the proteins conducting the 12 elongation cycles for nanchangmycin synthase (*48*), with redox control to generate polyolefinic chains that are oxygenated and converted to tetrahydropyran and tetrahydrofuran rings. These cyclic ether groups, together with the C-terminal carboxylate, chelate alkali metal cations, sit in plasma membranes and depolarize them via ion flow. Insight into the logic of polyether ionophore assembly has come with the sequencing of the monensin A cluster from *Streptomyces cinnamonensis* and that from *Streptomyces nanchangensis* for nanchangmycin biogenesis (*47*, *48*). The pattern of ether rings in monensin is 6-5-5-5-6 whereas the rings in nanchangmycin **22** are 6-5-6-5-6, representing distinct olefin placement in the acyclic precursor chains. The key chemical steps appear to be catalyzed by epoxidase-creating and -cleaving enzymes, which convert the olefins to epoxides that then spring open, via cationic mechanisms (Scheme 4), to set the furan and pyran ether ring systems. The double-bond geometry of the olefin precursors is probably a key determinant for epoxidase recognition (*57*).

A variant use of furan ether moieties occurs in the enzymatic cyclotetramerization of enantiomers of C_{10} furan-containing acyl CoAs by nonactin synthase (*49*), yielding the 32-membered tetralactone ring of **23**. Remarkably, two ketosynthase domains of nonactin synthase catalyze C-O lactone bond formation rather than the prototypic C-C bond

Claisen-type condensation of PKS modules (*49*). If portable to other contexts, this versatility of ketosynthase domains would be useful in combinatorial biosynthesis.

The enediyne family of antibiotics (*6*) targets DNA for oxidative strand scission after cycloaromatization of the enediyne warhead to a reactive arene diradical species. Both 9-membered and 10-membered enediyne cores are known, as in neocarzinostatin **24** and calicheamicin **25** (Fig. 6). The PKS cassettes encoding the two variants of enediyne core scaffold have been classified into two groups by polymerase chain reaction (PCR) amplification (*58*) and may provide starting points for evolution of the related but distinct scaffold-building enzymes.

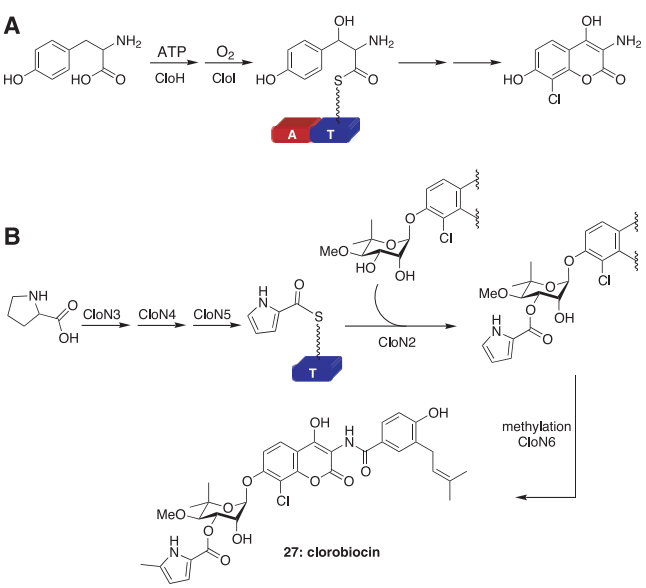


Scheme 5. Proposal for assembly of the griseorhodin skeleton via tridecaketidyl-*S*-enzyme formation, multiple cyclizations, and extensive oxidative tailoring reactions [adapted from (*64*)].

Tailoring Reactions During and After Assembly Line Operation

The reaction types of the preceding section, the oxidative crosslinking of the aryl rings in the heptapeptide scaffold of the glycopeptide antibiotics, the conversion of trienes to triene epoxides to five cyclic ethers in ionophores, and the construction of enediyne cores, represent tailoring reactions of enzymes dedicated to working on the NRP and PK chains, respectively, during or after release from the assembly lines. These and other tailoring reactions add to the structural versatility of antibiotic natural products and may be required for biological activity.

Glycosylations of the erythronolide macrolactone are essential for gain of antibiotic activity. The desosamine and cladinose sugars in erythromycin orient the aglycone at the 50S bacterial ribosome subunit as evidenced by co-crystallization of carbomycin with the 50S subunit where the disaccharide points back to the peptidyl transferase center (*59*). Variation of the antibiotic sugars that decorate PK and NRP aglycone scaffolds is possible in vivo (*60*–*63*) and in vitro (*64*–*66*). In the vancomycin system, it has been possible to vary the aglycone platform and both sugars in the disaccharide (D-glucosyl-1,2-L-vancosamine) chain to create new molecular entities.



Scheme 6. Role for two freestanding modules in clorobiocin assembly: (**A**) tyrosine activation, hydroxylation, and subsequent oxidative cyclization to the aminocoumarin core and (**B**) proline activation, oxidation to pyrrole, and transfer in the last step of antibiotic acylation.

Griseorhodin, **26**, a member of the rubromycin family, is a potent inhibitor of telomerase enzyme activity. It arises by action of aromatic polyketide synthase catalysis through 12 elongation cycles of malonyl units to generate a presumed tridecaketidyl intermediate that can cyclize to an angular hexacyclic intermediate and then undergo up to 11 redox transformations to create the napthoquinone and spiroketal moieties in griseorhodin (*67*), which are required for telomerase inhibition (Scheme 5). This is the largest number of redox tailoring enzymes yet seen in a PKS or NRPS cluster and should provide insights into the remarkable chemistry when the chemical function of each of the 11 enzymes is deconvoluted.

Nontraditional Assembly Lines: Domain and Module Shuffling in Nature

A search of genomic databases reveals some single NRPS modules, chiefly adenylation/thiolation (A/T) didomains responsible for amino acid selection and activation (*68*), in the absence of any other assembly line components. These reflect modules dedicated to the processing and transfer of amino acids, shunting them from primary to secondary metabolism, including β-hydroxylations, desaturations, and cyclopropanations. The biosynthetic gene cluster for the aminocoumarin antibiotic clorobiocin **27** (*69*) contains an A/T subunit for activating tyrosine to a β-hydroxytyrosyl thioester, which is then further oxidized (*70*) and cyclized to the aminocoumarin core of the antibiotic (Scheme 6A). Later in the pathway, proline is diverted to a pyrrolyl-*S*-enzyme intermediate by unusual, freestanding A and T domains. Desaturation of the prolyl-*S*-T domain (*71*) enables pyrrolyl transfer to the 3′-hydroxyl of the noviosyl sugar (Scheme 6B) to create the pharmacophore that blocks the adenosine triphosphate (ATP) site of the GyrB subunit of the target enzyme DNA gyrase. Transfer of evolved A/T subunits (as well as discrete A and T domains) into antibiotic producers may assist in diversity creation and give impetus to the combinatorial biosynthetic shuffling operations being carried out naturally in antibiotic-producing bacteria and fungi. Analogous separation of PKS assembly line catalytic and carrier protein domains is at work in the leinamycin (*40*) and other clusters (*72*), which are presumably evolving new catalytic functions and new protein partners.

A wide collection of tools, genes, enzymes, and variant monomers are becoming available for in vitro and in vivo manipulations of natural product structures to explore gain of function of novel biological activities.

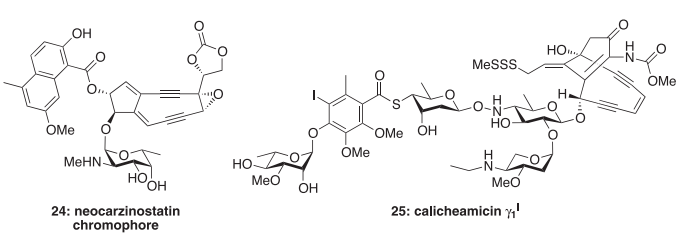


Fig. 6. Enediyne antibiotics from polyketide synthases.

References and Notes

1. D. E. Cane, C. T. Walsh, C. Khosla, *Science* **282**, 63 (1998).
2. S. Omura *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12215 (2001).
3. H. B. Bode, B. Bethe, R. Hofs, A. Zeeck, *ChemBioChem* **3**, 619 (2002).
4. Y. Tang, T. S. Lee, C. Khosla, *PLoS Biol.*, **2**, 227 (2004).
5. H. D. Mootz, D. Schwarzer, M. A. Marahiel, *ChemBioChem* **3**, 490 (2002).
6. B. Shen, *Curr. Opin. Chem. Biol.* **7**, 285 (2003).
7. C. T. Walsh *et al.*, *Curr. Opin. Chem. Biol.* **5**, 525 (2001).
8. T. A. Keating *et al.*, *ChemBioChem* **2**, 99 (2001).
9. R. M. Kohli, C. T. Walsh, *Chem. Commun.* **2003**, 293, (2003).
10. R. Pieper, G. Luo, D. E. Cane, C. Khosla, *Nature* **378**, 263 (1995).
11. H. M. Patel, C. T. Walsh, *Biochemistry* **40**, 9023 (2001).
12. D. A. Miller, L. Luo, N. Hillson, T. A. Keating, C. T. Walsh, *Chem. Biol.* **9**, 333 (2002).
13. B. A. Pfeifer, S. J. Admiraal, H. Gramajo, D. E. Cane, C. Khosla, *Science* **291**, 1790 (2001).
14. B. A. Pfieifer, C. C. Wang, C. T. Walsh, C. Khosla, *Appl. Environ. Microbiol.* **69**, 6698 (2003).
15. C. Khosla, P. B. Harbury, *Nature* **409**, 247 (2001).
16. N. Wu, S. Y. Tsuji, D. E. Cane, C. Khosla, *J. Am. Chem. Soc.* **123**, 6465 (2001).
17. S. E. O'Connor, C. T. Walsh, F. Liu, *Angew. Chem. Int. Ed. Engl.* **42**, 3917 (2003).
18. T. A. Keating, C. G. Marshall, C. T. Walsh, *Biochemistry* **39**, 15522 (2000).
19. T. L. Schneider, B. Shen, C. T. Walsh, *Biochemistry* **42**, 9722 (2003).
20. K. Shin-ya *et al.*, *J. Am. Chem. Soc.* **123**, 1262 (2001).
21. R. Jansen, H. Irschik, H. Reichenbach, V. Wray, G. Höfle, *Liebigs Ann. Chem.* 759 (1994).
22. R. S. Roy, A. M. Gehring, J. C. Milne, P. J. Belshaw, C. T. Walsh, *Nat. Prod. Rep.* **16**, 249 (1999).
23. L. Du, C. Sanchez, M. Chen, D. J. Edwards, B. Shen, *Chem. Biol.* **7**, 623 (2000).
24. D. Konz, A. Klens, K. Schorgendorfer, M. A. Marahiel, *Chem. Biol.* **4**, 927 (1997).
25. N. J. Hillson, C. T. Walsh, *Biochemistry* **42**, 766 (2003).
26. S. A. Sieber, M. A. Marahiel, *J. Bacteriol.* **185**, 7036 (2003).
27. T. Brautaset *et al.*, *Chem. Biol.* **7**, 395 (2000).
28. C. N. Boddy, T. L. Schneider, K. Hotta, C. T. Walsh, C. Khosla, *J. Am. Chem. Soc.* **125**, 3428 (2003).
29. J. W. Trauger, R. M. Kohli, H. D. Mootz, M. A. Marahiel, C. T. Walsh, *Nature* **407**, 215 (2000).
30. J. W. Trauger, R. M. Kohli, C. T. Walsh, *Biochemistry* **40**, 7092 (2001).
31. R. M. Kohli, J. W. Trauger, D. Scwarzer, M. A. Marahiel, C. T. Walsh, *Biochemistry* **40**, 7099 (2001).
32. R. M. Kohli, C. T. Walsh, M. D. Burkart, *Nature* **418**, 658 (2002).
33. R. M. Kohli, M. D. Burke, J. Tao, C. T. Walsh, *J. Am. Chem. Soc.* **125**, 7160 (2003).
34. S. A. Sieber, J. Tao, C. T. Walsh, M. A. Marahiel, *Angew. Chem. Int. Ed. Engl.* **43**, 493 (2004).
35. H. D. Mootz *et al.*, *J. Am. Chem. Soc.* **124**, 10980 (2002).
36. Z. Hojati *et al.*, *Chem. Biol.* **9**, 1175 (2002).
37. S. Weist *et al.*, *Angew. Chem. Int. Ed. Engl.* **41**, 3383 (2002).
38. B. Wilkinson *et al.*, *Chem. Biol.* **7**, 111 (2000).
39. Y. J. Yoon *et al.*, *Chem. Biol.* **9**, 203 (2002).
40. Y.-Q. Cheng, G.-L. Tang, B. Shen, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3149 (2003).
41. D. Panda, K. DeLuca, D. Williams, M. A. Jordan, L. Wilson, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9313 (1998).
42. P. R. August *et al.*, *Chem. Biol.* **5**, 69 (1998).
43. T.-W. Yu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7968 (2002).
44. A. Rascher *et al.*, *FEMS Microbiol. Lett.* **218**, 223 (2003).
45. B. K. Hubbard, C. T. Walsh, *Angew. Chem. Int. Ed. Engl.* **42**, 730 (2003).
46. K. C. Nicolaou, C. N. Boddy, S. Brase, N. Winssinger, *Angew. Chem. Int. Ed. Engl.* **38**, 2096 (1999).
47. M. Oliynyk *et al.*, *Mol. Microbiol.* **49**, 1179 (2003).
48. Y. Sun *et al.*, *Chem. Biol.* **10**, 431 (2003).
49. H. J. Kwon *et al.*, *Science* **297**, 1327 (2002).
50. A. M. van Wageningen *et al.*, *Chem. Biol.* **5**, 155 (1998).
51. J. Recktenwald *et al.*, *Microbiology* **148**, 1105 (2002).
52. J. Pootoolal *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8962 (2002).
53. M. Sosio, S. Stinchi, F. Beltrametti, A. Lazzarini, S. Donadio, *Chem. Biol.* **10**, 541 (2003).
54. M. Sosio *et al.*, *Microbiology* **150**, 95 (2004).
55. D. Bischoff *et al.*, *Angew. Chem. Int. Ed. Engl.* **40**, 4688 (2001).
56. A. Holtzel *et al.*, *J. Antibiot.* **55**, 571 (2002).
57. Z. A. Hughes-Thomas, C. B. W. Stark, I. U. Bohm, J. Staunton, P. F. Leadlay, *Angew. Chem. Int. Ed. Engl.* **42**, 4475 (2003).
58. W. Liu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 11959 (2003).
59. J. L. Hansen *et al.*, *Mol. Cell* **10**, 117 (2002).
60. D. Hoffmeister *et al.*, *Chem. Biol.* **7**, 821 (2000).
61. C. Mendez, J. A. Salas, *Trends Biotechnol.* **19**, 449 (2001).
62. L. L. Remsing *et al.*, *J. Am. Chem. Soc.* **124**, 1606 (2002).
63. D. Hoffmeister, G. Dräger, K. Ichinose, J. Rohr, A. Bechthold, *J. Am. Chem. Soc.* **125**, 4678 (2003).
64. H. C. Losey *et al.*, *Biochemistry* **40**, 4745 (2001).
65. H. C. Losey *et al.*, *Chem. Biol.* **9**, 1305 (2002).
66. C. Walsh, C. L. Freel Meyers, H. C. Losey, *J. Med. Chem.* **46**, 3425 (2003).
67. A. Li, J. Piel, *Chem. Biol.* **9**, 1017 (2002).
68. H. Chen *et al.*, *Biochemistry* **40**, 11651 (2001).
69. F. Pojer, S.-M. Li, L. Heide, *Microbiology* **148**, 3901 (2002).
70. H. Chen, C. T. Walsh, *Chem. Biol.* **8**, 301 (2001).
71. M. G. Thomas, M. D. Burkart, C. T. Walsh, *Chem. Biol.* **9**, 171 (2002).
72. C. R. Hutchinson, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 3010 (2003).
73. Supported by NIH grants GM 20011, GM 49338, AI 42738. Special thanks to D. Vosburg for preparation of artwork.

VIEWPOINT

Organic Chemistry in Drug Discovery

Malcolm MacCoss[1]* and Thomas A. Baillie[2]

The role played by organic chemistry in the pharmaceutical industry continues to be one of the main drivers in the drug discovery process. However, the precise nature of that role is undergoing a visible change, not only because of the new synthetic methods and technologies now available to the synthetic and medicinal chemist, but also in several key areas, particularly in drug metabolism and chemical toxicology, as chemists deal with the ever more rapid turnaround of testing data that influences their day-to-day decisions.

Numerous changes are now occurring in the pharmaceutical industry, not just in the way that the industry is perceived, but also in the rapid expansion of biomedical and scientific knowledge, which affects the way science is practiced in the industry. The recent changes in the way that synthetic chemistry is practiced in this environment center around new scientific advances in synthetic techniques and new technologies for rational drug design, combinatorial chemistry, automated synthesis, and compound purification and identification. In addition, with the advent of high-throughput screening (HTS), we are now faced with many targets being screened and many hits being evaluated. However, success in this arena still requires skilled medicinal chemists making the correct choices, often with insight gleaned from interactions with computational chemists and structural biologists, about which "hits" (*1*) are likely to play out as true "lead" (*1*) structures that will meet the plethora of hurdles that any drug candidate must surmount.

In the recent past, the usual flow of information that was generated regarding any new compound prepared in the laboratory of a drug discovery company followed a paradigm similar to that shown in Fig. 1. This scheme was driven by the need to get the initial information on a compound first, before deciding whether its properties met appropriate criteria before moving onto the next evaluation step. Such a linear sequence of events, although sparing of the number of compounds taken down the pathway, often meant that a considerable amount of time passed (several weeks) before it was known whether a particular change in a molecule was in fact a useful transformation, or whether it was a potency-enhancing change in the primary in vitro assay but was perhaps a liability in a downstream evaluation. Thus, the delay in getting appropriate feedback to the synthetic chemist meant that decisions about which molecules to prepare in the next round of synthesis were not guided by input from downstream data. With the advent of faster synthetic technologies, including advances in nuclear magnetic resonance (NMR) methods,

[1]Department of Basic Chemistry, Merck Research Laboratories, 126 East Lincoln Avenue, Rahway, NJ 07065, USA. [2]Department of Drug Metabolism, Merck Research Laboratories, Sumneytown Pike, West Point, PA 19486, USA.

*To whom correspondence should be addressed. E-mail: malcolm_maccoss@merck.com