impede the GFAP promoter activation by the two cytokines. Expression of any of the p300 deletion mutants used in this experiment reduced the GFAP promoter activation by LIF and BMP2 either alone or in combination (*6, 11*) (Fig. 4C), possibly by competing with the endogenous p300 for the binding to STAT3 and Smad1. The p300(1–1736) fragment, which interacted with neither STAT3 nor Smad1, also behaved as an inhibitor, possibly because of the seclusion of downstream targets of endogenous p300.

It is interesting that the neuroepithelial cell cultures (*6*) over 3 days with LIF or BMP2 alone produced GFAP-positive cells, although their extent was smaller than that observed in the 2-day culture with a combination of LIF and BMP2 (*8*). This is consistent with previous observations in which LIF (*5*) or BMP2 (*22*) alone induced astrocyte differentiation in cultures for a relatively longer period. This may be due to the formation of a STAT3-Smad1-p300 complex induced by an exogenously added cytokine and the endogenous expression and accumulation of its counterpart.

We have proposed a mechanism by which p300 coordinates the interaction of STAT3 and Smad1, leading to synergistic astrocyte differentiation. Similar interactions between transcriptional coactivators and different kinds of transcription factors may explain synergistic actions of distinct types of cytokines in other biological signaling pathways.

References and Notes

1. T. Taga and T. Kishimoto, *Annu. Rev. Immunol.* **15**, 797 (1997).
2. C.-H. Heldin, K. Miyazono, P. ten Dijke, *Nature* **390**, 465 (1997); R. Derynck, Y. Zhang, X.-H. Feng, *Cell* **95**, 737 (1998).
3. K. K. Johe, T. G. Hazel, T. Muller, M. M. Dugich-Djordjevic, R. D. G. McKay, *Genes Dev.* **10**, 3129 (1996); P. Rajan and R. D. G. McKay, *J. Neurosci.* **18**, 3620 (1998).
4. R. McKay, *Science* **276**, 66 (1997).
5. A. Bonni *et al.*, *ibid.* **278**, 477 (1997).
6. Neuroepithelial cells were prepared from embryonic day 14 mouse telencephalons as described (*3, 5*). Initially, plated cells were expanded for 4 days in N2-supplemented Dulbecco's modified Eagle's medium (DMEM)/F12 containing 10 ng/ml of basic fibroblast growth factor on dishes precoated with poly-L-ornithine and fibronectin. Cells were then detached and replated on Chamber Slides (Nunc; 8×10^4 cells per well) or 24-well plates (Nunc; 6×10^5 cells per well) for immunofluorescent staining after 2-day culture with cytokines (*7*) (Fig. 1) and luciferase assay (*11*) (Figs. 2 to 4), respectively.
7. Cells cultured on Chamber Slides were fixed with 4% paraformaldehyde in phosphate-buffered saline and stained with antibody to GFAP (Dako) and rhodamine-conjugated second antibody (Chemicon). The cells were counterstained with Hoechst 33258 to identify nuclei. Images were obtained using fluorescent microscopy (AX70, Olympus).
8. K. Nakashima and T. Taga, unpublished data.
9. T. Imamura *et al.*, *Nature* **389**, 622 (1997).
10. A. Hata, G. Lagna, J. Massagué, A. Hemmati-Brivanlou, *Genes Dev.* **12**, 186 (1998).
11. Neuroepithelial cells were transfected with a luciferase reporter gene fused with the 2.5-kb GFAP promoter (GF1L-pGL3) (Figs. 2 to 4) or its modified versions (GF1L-S-pGL3, SBSW-GF1L-SB-pGL3, or GF1L-K-pGL3) (Fig. 2) using Trans-IT LT-1 (Mirus) according to the manufacturer's protocol. As an internal control, a plasmid containing sea pansy luciferase expression construct was cotransfected. The following expression constructs were also used: DN-STAT3 (*23*) and Smad6 (*9*) in pEF-BOS (*24*) (Fig. 2), full-length p300 in CMV-β (18) (Fig. 3), and fragments of p300 (1–682, 1–1030, 1–1736, and 1737–2414) in pcDEF3 (*18*) (Fig. 4). On the following day, cells were stimulated with each cytokine (80 ng/ml) for 8 hours and solubilized. Luciferase activity was measured according to the recommended procedures for the Pikkagene Dual Luciferase Assay System (Toyo Ink Inc.). Luminous CT-9000D (Dia-Iatron) was used for quantitation.
12. M. Miura, T. Tamura, K. Mikoshiba, *J. Neurochem.* **55**, 1180 (1990).
13. M. A. Kahn *et al.*, *ibid.* **68**, 1413 (1997).
14. S. Dennler *et al.*, *EMBO J.* **17**, 3091 (1998).
15. L. Zawel *et al.*, *Mol. Cell* **1**, 611 (1998).
16. Y. Shi *et al.*, *Cell* **94**, 585 (1998).
17. Y. Kamei *et al.*, *ibid.* **85**, 403 (1996); P. Dai *et al.*, *Genes Dev.* **10**, 528 (1996); R. Kurokawa *et al.*, *Science* **279**, 700 (1998); J. C. Chrivia *et al.*, *Nature* **365**, 855 (1993).
18. A. Nishihara *et al.*, *Genes Cells* **3**, 613 (1998).
19. R. Janknecht, N. J. Wells, T. Hunter, *Genes Dev.* **12**, 2114 (1998); X.-H. Feng, Y. Zhang, R.-Y. Wu, R. Derynck, *ibid.*, p. 2153; C. Pouponnot, L. Jayaraman, J. Massagué, *J. Biol. Chem.* **273**, 22865 (1998); J. N. Topper *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9506 (1998).
20. J. J. Zhang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 15092 (1996); S. Bhattacharya *et al.*, *Nature* **383**, 344 (1996); E. Pfitzner, R. Jähne, M. Wissler, E. Stoecklin, B. Groner, *Mol. Endocrinol.* **12**, 1582 (1998).
21. COS-7 cells were maintained in DMEM containing 10% fetal bovine serum and transfected with expression constructs using Trans-IT LT-1. After 24 to 36 hours of culture, cells were solubilized as described (*25*). Lysates were immunoprecipitated with antibodies to FLAG (Sigma) or to hemagglutinin (HA) (Santa Cruz). Precipitates and, in some cases, cell lysates were subjected to SDS–polyacrylamide gel electrophoresis (PAGE) followed by immunoblotting with anti-FLAG, anti-HA, or anti-Myc (Genosys).
22. R. E. Gross *et al.*, *Neuron* **17**, 595 (1996).
23. M. Minami *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 3963 (1996).
24. S. Mizushima and S. Nagata, *Nucleic Acids Res.* **18**, 5322 (1990).
25. K. Nakashima, M. Narazaki, T. Taga, *FEBS Lett.* **401**, 49 (1997).

26. We thank R. Eckner for HA-p300, M. Minami and S. Akira for DN-STAT3, K. Shimozaki and S. Nagata for pEF-BOS and R-Luc, K. Ikenaka for GFAP gene promoter, K. Yasukawa for sIL-6R, Genetics Institute for BMP2 and BMP4, A. Nishihara for p300 deletion constructs, J. Akai and R. Hata for help and discussion on luciferase assays, and T. Imamura for valuable discussion. We also thank Y. Nakamura for secretarial assistance and K. Saito for technical help. Supported by a grant-in-aid from the Ministry of Education, Science, Sports and Culture of Japan and by the Human Frontier Science Program, Cell Science Research Foundation, and Kowa Life Science Foundation.

15 December 1998; accepted 4 March 1999

Dissecting and Exploiting Intermodular Communication in Polyketide Synthases

Rajesh S. Gokhale,[1] Stuart Y. Tsuji,[1] David E. Cane,[4] Chaitan Khosla[1,2,3]*

Modular polyketide synthases catalyze the biosynthesis of medicinally important natural products through an assembly-line mechanism. Although these megasynthases display very precise overall selectivity, we show that their constituent modules are remarkably tolerant toward diverse incoming acyl chains. By appropriate engineering of linkers, which exist within and between polypeptides, it is possible to exploit this tolerance to facilitate the transfer of biosynthetic intermediates between unnaturally linked modules. This protein engineering strategy also provides insights into the evolution of modular polyketide synthases.

Since the discovery of the modular architecture of certain polyketide synthases (PKSs), several reports have highlighted the functional versatility of these multienzyme assemblies by experiments involving domain inactivation, substitution, or addition (*1*). Although these empirical gene fusion approaches have led to the biosynthesis of diverse "unnatural" natural products, they have usually resulted in decreased in vivo productivity (*2*). The reasons for the lower productivity are poorly understood but could include structural instability of the engineered protein, suboptimal chemistry within the altered module, or inefficient processing of the nonnatural polyketide intermediates by downstream modules.

An alternative strategy for combinatorial biosynthesis would be to recombine intact modules from the vast natural repertoire of PKSs. Such an approach would benefit from the use of highly evolved modules as intact catalytic units, thereby eliminating unwanted perturbations in module structure or chemistry. Along with the diverse chemistry observed in polyketide biosynthetic pathways, greater degrees of freedom

[1]Department of Chemical Engineering, [2]Department of Chemistry, [3]Department of Biochemistry, Stanford University, Stanford CA 94305–5025, USA. [4]Department of Chemistry, Brown University, Box H, Providence RI 02912–9108, USA.

*To whom correspondence should be addressed.

could be combinatorially accessed through the interweaving of PKS modules with those of nonribosomal peptide synthetases, as found in nature (*2*). However, productive chain transfer between heterologous PKS modules has not been reported, perhaps because of the lack of understanding of the molecular basis for intermodular communication. Furthermore, although many bacterial genomes encode more than one PKS, each composed of multiple modules, there is no evidence for crosstalk between modules belonging to different dedicated PKS assemblies. A variety of experiments and observations have reinforced the view that individual modules show considerable selectivity toward their cognate substrates and that this intrinsic selectivity places serious constraints on harnessing the combinatorial potential of PKSs (*1*, *2*). Here we present several lines of evidence that challenge this viewpoint. Our studies on four individually expressed, catalytically functional modules of the erythromycin PKS reveal that these modules have very similar kinetic parameters for extending a given diketide substrate into the corresponding triketide. Surprisingly, short intermodular segments of variable amino acid sequence, referred to as linkers, were found to play a crucial role in the assembly of functional modules as well as in the intermodular polyketide chain transfer. By appropriate engineering of these linkers, we show that it is possible to facilitate communication between heterologous modules. Our results suggest that whereas the chemistry within PKS modules may be substrate- and stereo-selective, chain transfer between intact modules is permissive as long as the evolutionarily optimized linkers can provide the connectivity between adjacent modules.

The remarkable overall selectivity exhibited by naturally occurring multimodular PKSs suggests that each module possesses significant substrate specificity. As a direct test of this hypothesis, we sought to functionally express and kinetically characterize representative modules of 6-deoxyerythronolide B synthase (DEBS) (Fig. 1), which synthesizes the macrocyclic core **1** of the antibiotic erythromycin (*3*). For technical convenience, *Escherichia coli* was chosen as the expression host. The thioesterase (TE) domain, which ordinarily occurs at the COOH terminus of module 6 of DEBS, was fused to the COOH-terminal end of each module to facilitate substrate turnover (*4*). A key barrier to the functional expression of PKS modules in *E. coli* is the inability of the host's phosphopantetheinyl transferases to posttranslationally modify the acyl carrier protein (ACP) domains of PKS modules. To overcome this problem, we coexpressed the *sfp* phosphopantetheinyl transferase from *Bacillus subtilis* in *E. coli* BL21(DE3) cells (*5*). Coexpression of the *sfp* gene was both necessary and sufficient for functional expression of NH_2-terminal modules such as module 3+TE (M3+TE) and M5+TE (*6*), as assayed by their ability to convert the diketide thioester **2** into the expected triketide products **3** and **4**, respectively (Fig. 2). However, no activity could be detected from similarly expressed COOH-terminal modules such as M2+TE and M6+TE (*6*), even though these proteins were homodimeric and chromatographically similar to their NH_2-terminal counterparts. This lack of activity was unexpected because M2+TE and M6+TE are known to accept diketide thioester **2** when presented in their natural bimodular contexts (*7*). Detailed analysis of each domain of recombinant M2+TE indicated that whereas its acyltransferase (AT) domain was catalytically competent (as assessed by selective labeling of methylmalonyl-CoA) and its ACP domain was pantetheinylated, the ketosynthase (KS) domain could not be acylated with radiolabeled diketide thioester **2** (*8*). Because the KS domain is present at the NH_2-terminal ends of these re-

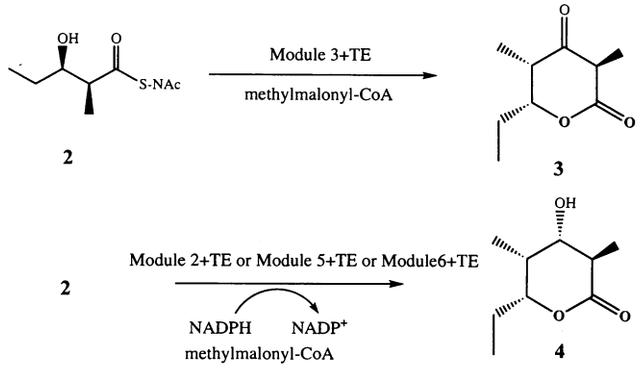

Fig. 2. Cell-free synthesis of triketides catalyzed by individual modules. All of the protein assays were done with varying concentrations of diketide thioester **2** (0.5 to 10 mM), 2.5 mM ^{14}C-methylmalonyl-CoA, and 100 pmol of purified protein in a 100-ml reaction. Because the KR domain is inactive in M3+TE, 1 mM of NADPH was only used in the assay mixtures for M2+TE, M5+TE, and M6+TE proteins. The reaction mixtures were quenched and extracted by ethyl acetate and separated by means of thin-layer chromatography (TLC). The products were confirmed by simultaneously running standards on TLC plates. Time courses for the formation of triketide lactone **4** and triketide ketolactone **3** were performed for 30 min. Quantitative measurements were done on a Packard InstantImager. All of these experiments were performed in triplicate.

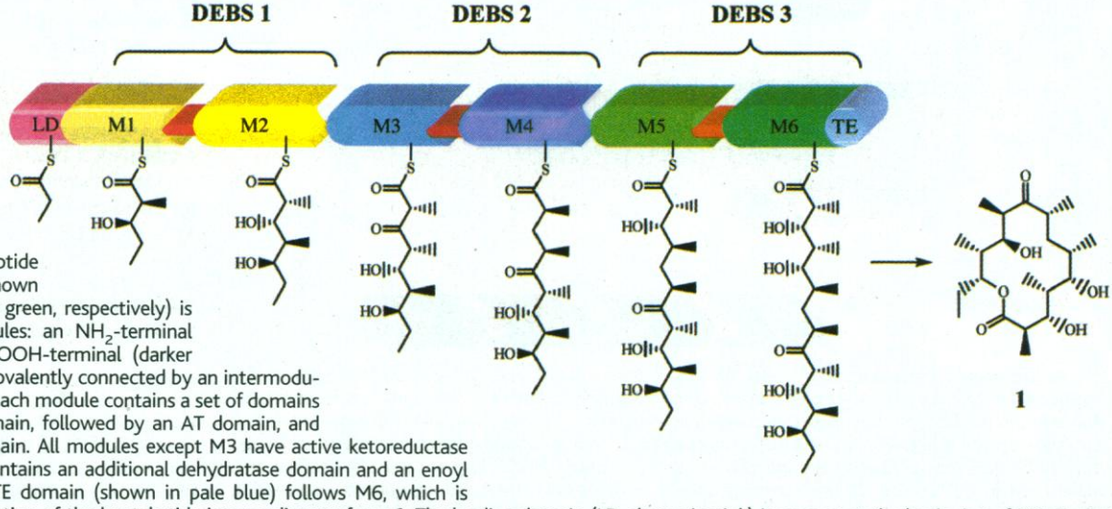

Fig. 1. Schematic representation of the modular organization of DEBS, which catalyzes the synthesis of the macrolide aglycone of erythromycin, 6-deoxyerythronolide B (6-dEB, **1**). DEBS is a hexameric protein complex ($\alpha 2\beta 2\gamma 2$). Each constituent polypeptide chain (DEBS 1, 2, and 3; shown here in yellow, blue, and green, respectively) is composed of two modules: an NH_2-terminal (lighter shade) and a COOH-terminal (darker shade) module that are covalently connected by an intermodular linker (shown in red). Each module contains a set of domains beginning with a KS domain, followed by an AT domain, and ending with an ACP domain. All modules except M3 have active ketoreductase (KR) domains, and M4 contains an additional dehydratase domain and an enoyl reductase domain. The TE domain (shown in pale blue) follows M6, which is responsible for the cyclization of the heptaketide intermediate to form **1**. The loading domain (LD, shown in pink) is present at the beginning of M1. During the course of reaction, all the intermediates are covalently sequestered as acylthioesters of the corresponding ACP and KS domains. The growing chain gets passed on from one module to next, like a baton in a relay race.

combinant proteins, we suspected a problem at this end of the protein for these COOH-terminal modules. Amino acid sequence comparisons of the six DEBS modules had led to the identification of linker regions of variable sequence that spanned the boundaries between the highly conserved ACP domain at the COOH-terminal end of an upstream module and the KS domain at the NH_2 terminus of the following module (*9*). Our analysis of these linker regions between DEBS modules, and those of other PKSs, revealed a marked contrast between the sequences separating covalently connected modules (for example, sequences between M1 and M2, M3 and M4, or M5 and M6 of DEBS) and those separating noncovalently connected modules (for example, sequences between M2 and M3 or M4 and M5 of DEBS). Specifically, the sequences separating covalently connected modules (hereafter referred to as intermodular linkers) are short and are characterized by the presence of a relatively conserved proline residue, whereas those upstream of NH_2-terminal modules and downstream of COOH-terminal modules (hereafter referred to as interpolypeptide linkers) are longer and relatively hydrophilic (Fig. 3). The NH_2-terminal ends of all KS domains contain a highly conserved stretch of four amino acids, PIAI (*10*), whose coding sequence often encompasses a Bsa BI restriction site (4 of 6 modules in DEBS and 8 of 10 modules in the rifamycin PKS possess this restriction site). Thus, construction of gene fusions at this site is a straightforward matter. Using this strategy, we replaced the DNA encoding the first 18 amino acids at the 5′ end of the inactive M2+TE construct described above with DNA encoding the first 39 amino acids from M5+TE (*11*). The resulting M2+TE protein could now accept diketide thioester **2** and convert it into the expected triketide lactone (Fig. 2). Likewise, the activity of the above-mentioned M6+TE gene was restored by a similar replacement. These results indicated that intermodular linkers play a crucial role in the assembly of functional PKS modules. Moreover, the development of this strategy for heterologous expression of single modules provided us with convenient access to the reagents needed for evaluating module specificity.

A Intermodular linker

```
M2ery: GGATGAEQAAPATT..APVD
M4ery: VGDAD..QAA.VRVVGAA.DES
M6ery: VGAAEAEQA.PALVREVPKDAD
M2rif: FGSA.A.NR.PAEIGTAAAE
M3rif: LG..ER.PAAPAPVTRDVSD
M5rif: GETVAGAPATPVTTVADAG
M3rap: .ELFTGENPAPVRGPVSAVGQD
M4rap: .ELFTGENPAPVRGPVSVVGQD
M7rap: .ELFTGENPAPVRGPVSA.GQD
```

B NH_2-terminal interpolypeptide linker

```
M3ery:  ......VTD SE KVAEYLRR .ATLDLRAAR QRIRE..LES
M5ery:  MSGDNGM.TE E.KLRRYLKR TVT.ELDSVT ARLRE..VEH RAG
M4rif:  ......MSAPNE QIVDAL.R ASLKE....N VRLQQENSAL AAAAA
M7rif:  ......VSASYE KVVEAL.R KSLEE....V GTLKKRNRQL ADAAG
M8rif:  ......V.AD EGQLRDYLKR .AIADARDAR TRLRE..VEE QAR
M9rif:  ......MATD E.KLLKYLKR .VTAELHS.. ..LRKQGARH .AD
M5rap:  ......MR.. EDQLLDAL.R KSVKE....N ARLRKANTSL RAAMD
M11rap: ......M.PEQD KVVEYL.R WATAELHTTR AKL.....EA LAAANT
```

Fig. 3. Amino acid sequence comparisons (*10*) of representative intermodular (**A**) and NH_2-terminal interpolypeptide (**B**) linkers from erythromycin (ery), rifamycin (rif), and rapamycin (rap) PKSs. These linker sequences have low sequence similarity. However, interpolypeptide linker residues have a higher occurrence of charged residues [bold residues in (B)], whereas nearly all of the intermodular linkers have a conserved proline residue [bold residues in (A)].

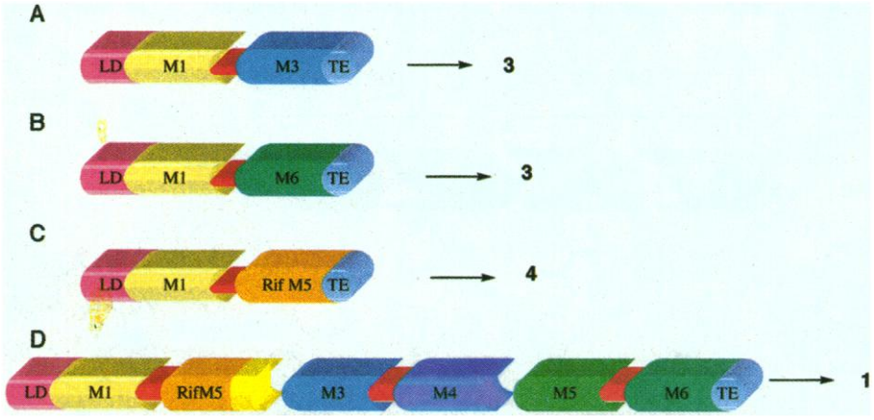

Fig. 4. Engineered heterologous modular fusions. Constructs (**A**) and (**B**) are bimodular constructs engineered from the erythromycin PKS, where chain transfer is observed between M1 and an unnatural downstream module (M3 and M6, respectively), as long as the intermodular linker (shown in red) that is naturally present between M1 and M2 is maintained. Communication between two modules derived from different PKSs (erythromycin and rifamycin) is illustrated in (**C**). The ery M1-M2 linker is also maintained in this construct. The construct shown in (**D**) demonstrates successful interpolypeptide chain transfer between heterologous modules. In addition to retaining the ery M1-M2 linker (shown in red), the intermodular linkers that are naturally present between M2 and M3 are also preserved (shown in yellow).

In addition to highlighting the important role of linkers in module connectivity, the above experiments showed that at least four distinct modules of DEBS were able to accept and extend diketide thioester **2**, which is the natural substrate of M2 only. To quantify this unexpected substrate tolerance, the steady-state kinetic parameters of the purified individual modules were measured. As seen in Table 1, the apparent catalytic rate constant (k_{cat}) and Michaelis constant (K_M) for **2** are very similar for every module except M3+TE. The value of k_{cat} for M3+TE is five times lower than that for the other three modules. A possible explanation for this difference is that the TE domain, which is present at the COOH-terminal ends of all four modules, discriminates between the ACP-bound β-hydroxythioester substrate generated by M2+TE, M5+TE, and M6+TE, and the β-ketothioester substrate generated by M3+TE. Recent studies with the isolated TE domain of DEBS have shown that it possesses significant specificity toward acylthioester substrates that have a (2*R*, 3*S*)-2-methyl-3-hydroxy stereochemistry (*12*), as is the case for the normal products of M2, M5, and M6. To assess the contribution of the TE domain to the apparent k_{cat} and K_M of M6+TE, these parameters were measured for **2** in the absence of NADPH (the reduced form of nicotinamide adenine dinucleotide phosphate), which eliminates the ketoreductase-catalyzed reaction from the catalytic cycle. Although the apparent K_M remained virtually unchanged (4.1 mM), the apparent k_{cat} decreased by half. We therefore conclude that individual modules of multimodular PKSs have an unexpectedly broad tolerance toward incoming ketide chains and that these modules faithfully process incoming substrates.

Because the acyl moiety of **2** is the product of the first module of DEBS, the above results suggest that it should be possible to synthesize novel polyketides by fusing M3, M5, or M6 downstream of M1. Although the previous attempts in our lab to construct a nonnative module fusion were unsuccessful (*13*), the results from heterologous expression of individual PKS modules in *E. coli* suggested that linkers might play a crucial role in facilitating chain transfer between adjacent modules. To assess the significance of the linkers, the fusions M1-M3+TE and M1-M6+TE were reconstructed

Table 1. Kinetic parameters of individual modules for the synthesis of triketide.

Proteins	$k_{cat} \times 100$ (min^{-1})	K_M (mM)
M2+TE	8 ± 0.6	4.6 ± 0.4
M3+TE	1.5 ± 0.3	4.4 ± 0.4
M5+TE	7.5 ± 0.7	4.7 ± 0.4
M6+TE	9.5 ± 0.6	4.3 ± 0.4
M6+TE (−NADPH)	4.5 ± 0.7	4.1 ± 0.4

in a manner that preserved the natural intermodular linker between M1 and M2 through use of the conserved Bsa BI restriction site described above (Fig. 4, A and B) (*14*). Remarkably, the recombinant strain of *Streptomyces coelicolor* CH999 harboring M1-M3+TE synthesized the expected triketide ketolactone **3**, as judged by nuclear magnetic resonance (NMR) spectroscopy, in quantities similar to the triketide synthesized by strains harboring DEBS1+TE (15 mg/liter). Likewise, M1-M6+TE also produced comparable quantities of the ketolactone **3**, as judged by NMR spectroscopy (*15*). As a more demanding test of the role of linkers in polyketide chain transfer, we replaced M2 of DEBS with M5 from the rifamycin (*rif*) PKS from *Amycolatopsis mediterranei* (*16*). The α-methyl and β-hydroxyl stereocenters generated by M5 of the *rif* PKS are the same as those generated by M2 of DEBS, although the incoming acyl chain differs substantially in the two cases. The bimodular PKS generated by linking *rif* M5 downstream of DEBS M1 produced the expected triketide lactone **4** in vivo, as judged by NMR spectroscopy (Fig. 4C) (*17*). Again, the productivity of the recombinant strain of *S. coelicolor* expressing this PKS is comparable to that for the triketide lactone synthesized by DEBS1+TE (15 mg/liter). The successful engineering of a DEBS M1–*rif* M5 junction prompted us to test whether the interpolypeptide linker between M2 and M3 of DEBS could also facilitate chain transfer from *rif* M5 to DEBS M3. For this experiment, the COOH-terminal linker that ordinarily occurs downstream of ACP2 in DEBS1 was fused in place of the TE domain in the bimodular construct described in Fig. 4C (*18*). Upon coexpression of the resulting bimodular protein with DEBS2 and DEBS3, the recombinant strain of *S. coelicolor* produced 15 mg/liter of **1**, as judged by NMR spectroscopy (Fig. 4D). These studies demonstrate the feasibility of functionally rewiring PKS modules. Thus, the linker hypothesis, which was developed using an intermodular linker between M1 and M2 of DEBS, is also applicable to interpolypeptide linkers (as illustrated by the linker between M2 and M3 of DEBS). The linker hypothesis has two important implications. First, it suggests a simple model for the evolution of modular PKSs and possibly for nonribosomal peptide synthetases (*2*, *19*), because gene duplication is both necessary and sufficient for the evolution of multimodular systems as long as linkers provide suitable module connectivity. Second, it provides a fundamentally new strategy for combinatorial biosynthesis, in which modules, rather than individual enzymatic domains, are the building blocks for genetic manipulation.

References and Notes

1. D. A. Hopwood, *Chem. Rev* **97**, 2465 (1997); L. Katz, *ibid.*, p. 2557; J. Staunton and B. Wilkinson, *ibid.*, p. 2611; C. Khosla, *ibid.*, p. 2577.
2. D. E. Cane, C. T. Walsh, C. Khosla, *Science* **282**, 63 (1998); R. McDaniel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 1846 (1999).
3. J. Cortes *et al.*, *Nature* **348**, 176 (1990); S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson, L. Katz, *Science* **252**, 675 (1991).
4. J. Cortes *et al.*, *Science* **268**, 1487 (1995); C. M. Kao, G. Luo, L. Katz, D. E. Cane, C. Khosla, *J. Am. Chem. Soc.* **117**, 9105 (1995); *ibid.* **118**, 9184 (1996).
5. R. H. Lambalot *et al.*, *Chem. Biol.* **3**, 923 (1996).
6. For this study, the gene encoding M3+TE was derived from an engineered trimodular construct pKAO318 (*13*), where an Nhe I site was engineered at the start codon of *eryII* (DEBS 2). Fusion of the gene encoding TE at the end of ACP3 has been described earlier in the construction of pCK13 (*4*). The Nhe I–Eco RI fragment was cloned in pET21c (Novagen) to construct the expression plasmid pRSG34. By using the Eco RI site, the natural stop codon at the end of the TE domain was deleted so that this protein could be overproduced as a COOH-terminal hexahistidine (His_6)–tagged fusion protein. Expression of M5+TE was achieved by combining the engineered Nde I site from pJRJ10 [J. R. Jacobsen *et al.*, *Biochemistry* **37**, 4928 (1998)] with the Eco RI site from pCK15 (*4*). The Nde I–Eco RI fragment was cloned in pET21c to obtain the expression plasmid pRSG46. These plasmids were introduced through transformation into *E. coli* BL21(DE3) cells for protein expression. Similar expression constructs were engineered for M2+TE and M6+TE using an engineered Nhe I site immediately upstream of the corresponding ketosynthases (at position 7570, 5′-GCTAGCGAGCCGATC-3′; at position 28710, 5′-GCTAGCGACCCGATC-3′). In order to coexpress the *sfp* gene, the Nde I–Hind III fragment derived from the pUC8-*sfp* [M. M. Nakano *et al.*, *Mol. Gen. Genet.* **232**, 313 (1992)] was cloned into pET, which has a kanamycin resistance gene. The resultant plasmid, pRSG56, was cotransformed into BL21(DE3) containing other single-module plasmids.
7. C. M. Kao, R. Pieper, D. E. Cane, C. Khosla, *Biochemistry* **35**, 12363 (1996); N. L. Pohl, R. S. Gokhale, D. E. Cane, C. Khosla, *J. Am. Chem. Soc.* **120**, 11206 (1998).
8. R. S. Gokhale, S. Y. Tsuji, D. E. Cane, C. Khosla, data not shown.
9. S. Donadio and L. Katz, *Gene* **111**, 51 (1992).
10. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
11. Expression plasmids for M2+TE, pRSG64, and M6+TE, pRSG54, were constructed by replacing the Bsa BI–Eco RI fragment in pRSG46 (*6*) by the corresponding fragment from pCK12 (*4*) (for M2+TE) and pJRJ10 (for M6+TE). The fusion of the TE gene at the end of M2 has been described earlier (*4*). Proteins were expressed in BL21(DE3) cells and induced with 1 mM isopropyl-β-D-thiogalactopyranoside (optical density at 600 nm = 0.6). Cells were harvested by centrifugation after 10 hours and resuspended in disruption buffer [200 mM sodium phosphate (pH 7.2), 200 mM sodium chloride, 2.5 mM dithiothreitol (DTT), 2.5 mM EDTA, 1.5 mM benzamidine, pepstatin and leupeptin (2 mg/liter), and 30% v/v glycerol]. The cells were lysed using a French press, and the lysate was collected after centrifugation. Nucleic acids were precipitated with polyethylenimine (0.15%) and removed with centrifugation. The supernatant was made 50% (w/v) saturated with ammonium sulfate and precipitated overnight. After centrifugation, the pellet containing protein was redissolved in buffer A [100 mM sodium phosphate (pH 7.2), 2.5 mM DTT, 2 mM EDTA, and 20% glycerol (v/v)] and stored at −80°C. For chromatography, the buffer was exchanged to buffer A + 1M ammonium sulfate using a gel filtration PD10 (Pharmacia) column. The resulting sample was loaded on a Butyl Sepharose (Pharmacia) column. Fractions containing DEBS proteins were pooled and applied on an anion exchange column (6 ml; Resource Q, Pharmacia). Purified protein fractions were pooled and concentrated with Amicon Centriprep30. Typical purified protein yields were ~3 to 4 mg per liter of culture. All these individual modules eluted as dimers on the gel filtration column. More than 90% of the proteins were phosphopantetheinylated in vivo as a result of the overexpression of *sfp* phosphopantetheinyl transferase. Although these proteins were expressed as His_6-tagged proteins, they did not bind to a Ni-column under experimental conditions. It is unclear whether this inability to bind to a Ni-agarose column is due to steric effects or whether the His_6 peptide was lost during purification.
12. R. S. Gokhale, D. Hunziker, D. E. Cane, C. Khosla, *Chem. Biol.* **6**, 117 (1999).
13. Previous attempts to generate heterologous module fusions, such as M1-M6+TE, M1-M3+TE, and M1-M2-M6+TE, have been unsuccessful [R. McDaniel *et al.*, *ibid.* **4**, 667 (1997); C. M. Kao, J. Lau, R. S. Gokhale, C. Khosla, unpublished results]. Although none of these recombinant PKSs, constructed by engineering restriction sites at arbitrarily chosen positions in the linkers between ACP and KS domains, yielded any polyketides in vivo, the fusion proteins were stable and could be purified as homodimers. In vitro assays carried out with radiolabeled propionyl-CoA and unlabeled methylmalonyl-CoA did not produce any radiolabeled triketide product. Instead, when diketide thioester **2** and radiolabeled methylmalonyl-CoA were used as substrates, these bimodular constructs produced radiolabeled triketide, probably by exogenously priming at the downstream module. It has been previously shown that **2** does not acylate at the KS domain of M1 [N. Tsukamoto *et al.*, *Biochemistry* **35** 15244 (1996)].
14. Bsa BI–Eco RI fragments containing *ery* M3 and *ery* M6 were cloned behind *ery* M1. The M1-M3+TE and M1-M6+TE genes were excised as Pac I–Eco RI fragments and inserted into pCK12 (*4*). The resultant plasmids pST97 and pST96, respectively, were transformed in *S. coelicolor* strain CH999.
15. The reason for the inability of KR6 to reduce the triketide keto ester is unclear, because incubation of M5-M6+TE with diketide **2** and NADPH can synthesize **4**.
16. P. R. August *et al.*, *Chem. Biol.* **4**, 69 (1998); T. Schupp, C. Toupet, N. Engel, S. Goff, *FEMS Microbiol. Lett.* **159**, 201 (1998).
17. The construction of *ery* M1–*rif* M5+TE was done as follows: Based on an alignment with *ery* ACP sequences, the natural sequence at 28024 5′-CGCGAC-3′ in *rif* ACP5 was replaced with the Spe I recognition sequence of 5′-ACTAGT-3′. Standard polymerase chain reaction techniques were used to carry out mutagenesis. The Bsa BI–Spe I fragment containing *rif* M5 was excised and used to replace the corresponding *ery* M2 fragment in pCK12. The resultant expression plasmid, pST110, was transformed into *S. coelicolor* strain CH999.
18. To graft the interpolypeptide linker from the COOH terminus of *ery* M2 onto the *rif* M5, the Pac I–Spe I fragment of pST110 was inserted into a derivative of pCK7 [C. M. Kao, *et al. Science* **265**, 509 (1994)]. The derivative had a Spe I site engineered in the homologous region at the end of ACP2. The resulting pST113 construct contained *ery* M1 linked to *rif* M5 (through the natural cis linker between *ery* M1 and M2), *rif* M5 noncovalently linked to *ery* M3 (through the natural interpolypeptide linker between *ery* M2 and M3), and the subsequent natural links to *ery* M4, M5, and M6.
19. P. J. Belshaw, C. T. Walsh, T. Stachelhaus, *Science* **284**, 486 (1999).

20. Supported by grants from NIH (CA66736 to C.K. and GM22172 to D.E.C.) and NSF (BES9806774 to C.K.). S.Y.T. is a recipient of a predoctoral fellowship from NSF and a NIH Biotechnology Graduate Training Fellowship. We thank N. Wu for his assistance with NMR.

19 January 1999; accepted 12 March 1999