Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of *E. coli*

Blaine A. Pfeifer,[1] Suzanne J. Admiraal,[1] Hugo Gramajo,[1]* David E. Cane,[4] Chaitan Khosla[1,2,3]†

The macrocyclic core of the antibiotic erythromycin, 6-deoxyerythronolide B (6dEB), is a complex natural product synthesized by the soil bacterium *Saccharopolyspora erythraea* through the action of a multifunctional polyketide synthase (PKS). The engineering potential of modular PKSs is hampered by the limited capabilities for molecular biological manipulation of organisms (principally actinomycetes) in which complex polyketides have thus far been produced. To address this problem, a derivative of *Escherichia coli* has been genetically engineered. The resulting cellular catalyst converts exogenous propionate into 6dEB with a specific productivity that compares well with a high-producing mutant of *S. erythraea* that has been incrementally enhanced over decades for the industrial production of erythromycin.

Erythromycin, a broad spectrum antibiotic synthesized by the bacterium *S. erythraea*, is a prototype of a class of complex natural products called polyketides (*1*). These biomolecules are synthesized from simple building blocks such as acetyl coenzyme A (acetyl-CoA), propionyl-CoA, malonyl-CoA, and methylmalonyl-CoA through the action of large modular megasynthases called PKSs (*2*). Polyketide structural complexity often precludes the development of practical laboratory synthetic routes, leaving fermentation as the only viable source for the commercial production of these pharmaceutically and agriculturally useful agents. But the challenges associated with developing scalable and economically feasible fermentation processes for polyketide production from natural biological sources (principally the *Actinomyces* family of bacteria) are enormous and represent the most serious bottleneck during polyketide preclinical and clinical development. Analogous to the pivotal role of *E. coli* in the production of therapeutic proteins, the ability to produce complex polyketides in this exceptionally well-developed heterologous host can be expected to benefit natural product drug and agrochemical development.

Notwithstanding the importance of this objective, three challenges are encountered in the production of complex polyketides in *E. coli*. The first is the ability to express correctly folded and posttranslationally modified modular PKSs in this heterologous host. For example, deoxyerythronolide B synthase (DEBS), which synthesizes the macrocyclic core of erythromycin, is a 2-megadalton $\alpha_2\beta_2\gamma_2$ complex composed of two equivalent sets of 28 distinct active sites, seven of which must be posttranslationally pantetheinylated. Once coexpressed in *E. coli* in an active form, the three DEBS subunits would represent the largest and perhaps the most highly elaborated polypeptides in this bacterium. Second, the building blocks for polyketide biosynthesis must be available in *E. coli*. For example, 6-deoxyerythronolide B (6dEB, **1**) is synthesized from one propionyl-CoA–derived primer unit and six (2*S*)-methylmalonyl-CoA–derived extender units (Fig. 1). Although *E. coli* synthesizes propionyl-CoA, this substrate is formed under conditions that are incompatible with typical high cell density fermentation conditions (*3*). Moreover, (2*S*)-methylmalonyl-CoA has not been observed as a metabolite in *E. coli* (*4*). Therefore production of 6dEB in *E. coli* would require re-engineering its metabolism to synthesize these two precursors. Lastly, in order to design an effective cellular catalyst for 6dEB biosynthesis, intracellular DEBS activity needs to be well-synchronized with precursor biosynthesis to maximize the relevant stoichiometric and kinetic parameters. Here we address each of these three issues.

Although the expression of DEBS genes has been achieved in the model actinomycete *Streptomyces coelicolor* (*5*), to our knowledge genes of this size have never been functionally expressed in *E. coli*. Therefore, we first attempted to express the genes encoding each of the three DEBS proteins, followed by in vitro assays of protein activity. DEBS3, DEBS2, and a variant of DEBS1, DEBS1+TE (*6*, *7*), were cloned individually into the pET21c expression vector (*8*) and introduced via transformation into *E. coli* BL21(DE3) harboring the *sfp* phosphopantetheinyl transferase gene (*9*) on pRSG56 (*10*). The expression levels of the three DEBS genes (*11*) were found to be comparable to those reported earlier from *S. erythraea* (*12*) or *S. coelicolor* (*13*). However, although DEBS3 was found to be active in these lysates, DEBS1+TE and DEBS2 lacked any detectable activity (*14*). Consistent with these results, recombinant DEBS3 could be purified from these lysates using procedures described earlier (*15*), but neither DEBS1+TE nor DEBS2 could be purified in detectable quantities. The key parameter that facilitated detection of in vitro activity and subsequent purification of DEBS1+TE and DEBS2 was the incubation temperature after IPTG (isopropyl-β-D-thiogalactopyranoside) induction. Upon lowering the expression temperature from 30° to 22°C, active DEBS1+TE, DEBS2, and DEBS3 proteins could be detected in recombinant *E. coli* lysates. Low temperature induction conditions were subsequently maintained throughout the course of this study.

Although plasmid-based coexpression of the *sfp* gene facilitated posttranslational modification of the DEBS proteins in *E. coli*, integration of this gene into the chromosome was desirable before attempting coexpression of all the DEBS genes. Using homologous recombination, a single copy of the *sfp* gene under control of the T7 RNA polymerase promoter was integrated in the *prp* operon of BL21(DE3), yielding *E. coli* BAP1 (Fig. 2) (*16*). This site was chosen for *sfp* gene insertion for two reasons. First, the *prp* operon is putatively responsible for propionate catabolism in *E. coli* (*17*). Because propionate was intended to be the sole source of carbon building blocks for 6dEB biosynthesis, concurrent propionate catabolism and anabolism were deemed undesirable. By deleting *prpRBCD* in the process of *sfp* integration, the ability of BAP1 to utilize propionate as a carbon and energy source was eliminated. Second, together with the *sfp* gene, the *prpE* gene in BAP1 was also placed under control of an IPTG-inducible T7 promoter. PrpE is thought to convert propionate into propionyl-CoA (*18*); therefore, in the presence of exogenous propionate, propionyl-CoA can be expected to accumulate inside the cell at the same time as DEBS is expressed in an active form.

As a final step toward engineering *E. coli* for 6dEB production, a suitable pathway for (2*S*)-methylmalonyl-CoA biosynthesis needed to be engineered. Recently, the propionyl CoA carboxylase (*pcc*) genes from *S. coelicolor* have been isolated and characterized (*19*). Preliminary experiments confirmed that the genes encoding the two subunits, *pccA* and *pccB*, could be expressed in *E. coli* in an active form and that the activity of the biotinylated subunit (*pccA*) could be enhanced upon coexpression of the *E. coli birA* biotin

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

*Permanent address: Facultad de Ciencias Bioquimicas y Farmaceuticas, Departamento de Microbiologia, Suipacha 531, Rosario 2000, Argentina.

†To whom correspondence should be addressed. E-mail: ck@chemeng.stanford.edu

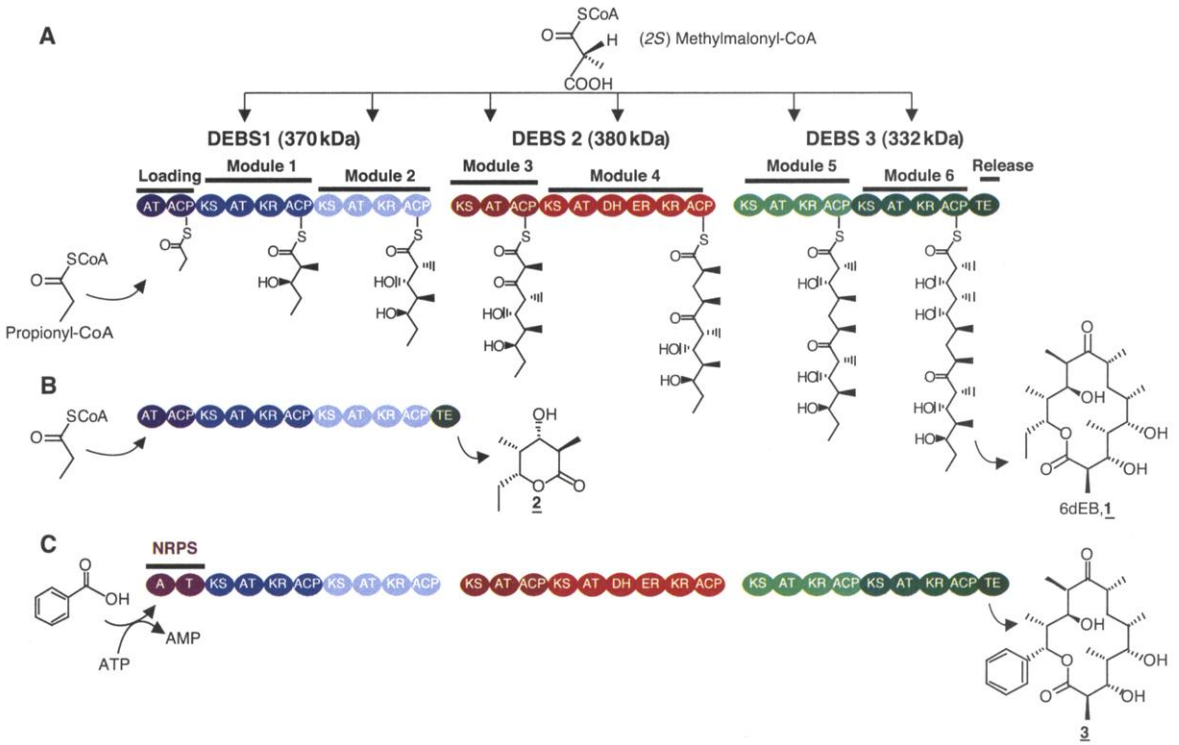

Fig. 1. (**A**) The 6-deoxyerythronolide B synthase. Catalytic domains: KS, ketosynthase; AT, acyl transferase; ACP, acyl carrier protein; KR, ketoreductase; ER, enoyl reductase; DH, dehydratase, TE, thioesterase. DEBS utilizes 1 mole of propionyl-CoA and 6 moles of (2*S*)-methylmalonyl-CoA to synthesize 1 mole of 6-deoxyerythronolide B (6dEB, **1**). (**B**) Truncated DEBS1+TE produces the triketide lactone **2**. (**C**) The rifamycin synthetase is a polyketide synthase that is naturally primed by a nonribosomal PKS loading module, comprised of two domains: an ATP-dependent adenylation domain (A) and a thiolation domain (T). Substitution of this A-T didomain in placeof the loading didomain of DEBS yields an engineered "hybrid" synthase that utilizes exogenous acids such as benzoic acid to synthesize substituted macrocycles such as **3** in an engineered strain of *E. coli*.

ligase gene (*20*). To test whether propionyl-CoA carboxylase can synthesize the correct isomer of methylmalonyl-CoA in vivo, BAP1 was transformed with three plasmids: pRSG32 carrying the DEBS1+TE gene under control of the pT7 promoter (*8*), pTR132 carrying the *pccAB* genes (*19*), and pCY216 carrying the *birA* gene (*21*). M9 minimal media cultures of transformed cells were grown at 37°C to mid-log phase, followed by induction at 22°C with 0.5 liter mM IPTG, arabinose (2.5 g/liter), and [1-^{14}C]– or [1-^{13}C]–propionate (26 mg/liter or 250 mg/liter), respectively (*22*). After 12 to 48 hours, the culture supernatant was extracted and analyzed for formation of the expected triketide lactone (**2**) product of DEBS1+TE. Formation of triketide lactone under both feeding conditions confirmed the ability of BAP1 to produce polyketides and encouraged us to attempt 6dEB production in the same host.

To coexpress all three DEBS genes and the *pcc* genes, vectors pET21c and pET28a (Novagen, Milwaukee, Wisconsin) were modified to express two and three genes, respectively (*23*). When tested individually, protein production was observed from each gene located on both plasmids. BAP1 was transformed with these plasmids together with the *birA* plasmid. Individual transformants were cultured, induced, and analyzed similar to the experiment for DEBS1+TE using [1-^{13}C]–propionate. Nuclear magnetic resonance (NMR) analysis of the crude organic extract revealed 6dEB as the major propionate-derived metabolite of these

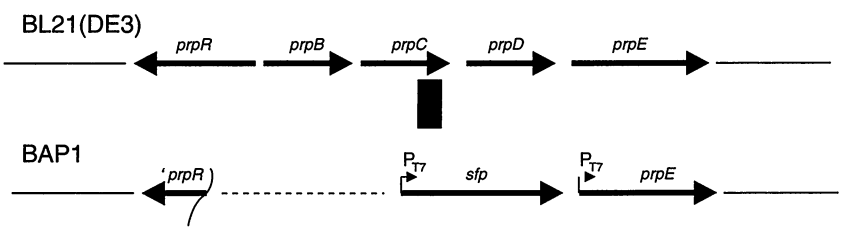

Fig. 2. Genetic design of *E. coli* BAP1.

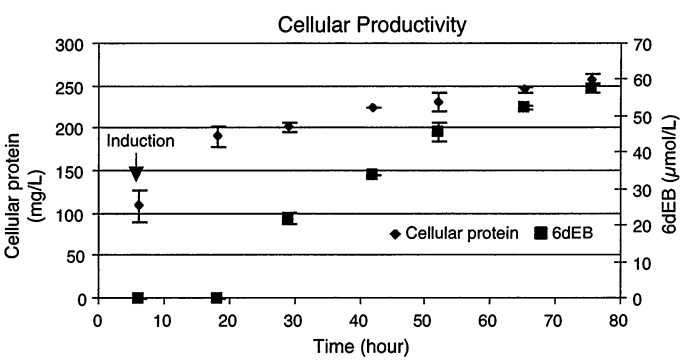

Fig. 3. Production of 6dEB in *E. coli*. Cellular protein content and 6dEB concentration are plotted versus time.

recombinant cells. The product was later purified by high-performance liquid chromatography (HPLC) and subjected to mass spectrometry yielding a major peak of the expected mass (*24*). To quantify the productivity of this newly characterized polyketide cellular system, culture samples were taken periodically, and the concentration of 6dEB was measured (Fig. 3). From this data it can be calculated that the specific productivity of this cellular catalyst is 0.1 mmol of 6dEB per gram of cellular protein per day. This is superior to wild-type *S. erythraea* and compares well to an industrially relevant strain that overproduces erythromycin (0.2 mmol of erythromycin per gram of cellular protein per day) (*25*) as a result of a decades-long program of directed-strain improvement based on random mutagenesis.

To demonstrate the power of engineering modular PKSs in this previously unknown het-

erologous system, we attempted to construct a derivative of DEBS in which a PKS module was fused to a nonribosomal peptide synthetase (NRPS)–like (*26*) module. The first module of the rifamycin synthetase has recently been shown to be an NRPS-like module comprised of two domains: an adenylation (A) and a thiolation (T) domain (*27*). The A domain activates 3-amino-5-hydroxybenzoate (as well as benzoate and several benzoate derivatives) (*27*) in an adenosine triphosphate (ATP)–dependent reaction and transfers the aryl adenylate onto the phosphopantetheine arm of the T domain (Fig. 1). This NRPS-like module was fused upstream of the first condensation module of DEBS in lieu of the loading didomain of DEBS (*28*). In the presence of exogenous propionate and benzoate, the resulting strain of *E. coli* produced the expected 6dEB analog (**3**), as confirmed by NMR and mass spectrometry (Fig. 1) (*29*).

In summary, we have demonstrated the feasibility of engineering *E. coli* to produce complex polyketide natural products. Multiple changes were made to the *E. coli* genome for relevant 6dEB production, including introduction of the three DEBS genes from *S. erythraea*, introduction of the *sfp* phosphopantetheinyl transferase gene from *Bacillus subtilis*, introduction of genes encoding a heterodimeric propionyl-CoA carboxylase from *S. coelicolor*, deletion of the endogenous *prpRBCD* genes, and overexpression of the endogenous *prpE* and *birA* genes. When gene expression was coordinately induced at low temperature, propionate could be converted into 6dEB by this metabolically engineered cellular catalyst with excellent kinetic parameters. Given the availability of well-established scalable protocols for fermenting *E. coli* to overproduce bioproducts, the ability to synthesize complex polyketides in this heterologous host bodes well for the practical production of these bioactive natural products. Equally important, as indicated by the hybrid PKS-NRPS described here, it opens the door for harnessing the enormous power of molecular biology in *E. coli* to engineer modular polyketide synthases using directed and random approaches.

References and Notes

1. D. O'Hagan, *The Polyketide Metabolites* (Horwood, Chichester, UK, 1991).
2. D. E. Cane, C. T. Walsh, C. Khosla, *Science* **282**, 63 (1998).
3. A. W. Tsang, A. R. Horswill, J. C. Escalante-Semerena, *J. Bacteriol.* **180**, 6511 (1998).
4. T. Haller, T. Buckel, J. Retey, J. A. Gerlt, *Biochemistry* **39**, 4620 (2000).
5. C. M. Kao, L. Katz, C. Khosla, *Science* **265**, 509 (1994).
6. C. M. Kao, G. Luo, L. Katz, D. E. Cane, C. Khosla, *J. Am. Chem. Soc.* **117**, 9105 (1995).
7. J. Cortes *et al.*, *Science* **268**, 1487 (1995).
8. See Supplemental Material for details regarding construction of plasmids encoding DEBS1+TE (pRSG32), DEBS2 (pBP49), and DEBS3 (pRSG50) genes. Text and Web figures 1 and 2 are available at *Science* Online at www.sciencemag.org/cgi/content/full/291/5509/1790/DC1.
9. R. H. Lambalot, *et al.*, *Chem. Biol.* **3**, 923 (1996).
10. R. S. Gokhale, S. Y. Tsuji, D. E. Cane, C. Khosla, *Science* **284**, 482 (1999).
11. Individual transformants were used to start 25 ml Luria Broth (LB) seed cultures containing carbenicillin (100 μg/ml) and kanamycin (50 μg/ml) at 250 rpm and 37°C. These cultures were used to inoculate 1 liter of LB medium, and the culture was grown under the same conditions. At mid-log phase (OD_{600} = 0.4 to 0.8) cells were induced with 1 mM IPTG and transferred to a 30°C incubator. Cells were harvested after 4 to 6 hours and their protein content was analyzed via 7.5% SDS–polyacrylamide gel electrophoresis (SDS-PAGE). The three DEBS proteins were expressed at about 1% total cellular protein [see Supplemental Materials for representative gel (*8*)].
12. P. Caffrey, D. J. Bevitt, J. Staunton, P. F. Leadlay, *FEBS Lett.* **304**, 225 (1992).
13. R. Pieper, G. Luo, D. E. Cane, C. Khosla, *Nature* **378**, 263 (1995).
14. DEBS1+TE (*13*) and DEBS3 were assayed as previously described [J. R. Jacobsen, D. E. Cane, C. Khosla, *Biochemistry* **37**, 4928 (1998)]. Although an assay for the entire DEBS2 has not yet been developed, the activity of module 3 on this protein can be assayed as previously described (*10*).
15. N. L. Pohl, R. S. Gokhale, D. E. Cane, C. Khosla, *J. Am. Chem. Soc.* **120**, 11206 (1998).
16. The method of Hamilton *et al.* [*J. Bacteriol.* **171**, 4617 (1989)] was used for chromosomal gene replacement. Specifically, ~1 kb of *prpR* and *prpE* were amplified via polymerase chain reaction (PCR) using the following oligonucleotides: (*prpE* 5′ oligonucleotide) TTTCATATGTCTTTTAGCGAATTTTATCAGCGTTC; (*prpE* 3′ oligonucleotide) TGAATTCTTAATTAACGGCGCTGAGAAACATCCGGCT; (*prpR* 5′ oligonucleotide) TTTATGCATGGCGTAATGCAGCAGAAAATGG; (*prpR* 3′ oligonucleotide) TTTCTAGACCGGAAAACTCACCTCTTCTAT. The *prpE* fragment was cloned as an Nde I–Eco RI cassette in pET21c and subsequently was mobilized as an Sph I–Pac I fragment into the temperature-sensitive plasmid, pMAK705. [This fragment, which contained the pT7 promoter from pET21c upstream of *prpE*, served as the right-hand homologous fragment for gene replacement (see Fig. 2).] The Nsi I–Xba I cassette encoding the 5′ end of *prpR* was used as the left-hand homologous fragment for gene replacement (Fig. 2). Between these homologous fragments, the *sfp* gene fused to the pT7 promoter (from pRSG56) was cloned in the orientation shown in Fig. 2. The resulting temperature-sensitive plasmid was used for gene replacement in the chromosome of BL21(DE3). Double crossover events resulted in generation of the BAP1 cell line (Fig. 2), which was subsequently verified by PCR amplification and by its ability to produce pantetheinylated, active DEBS proteins.
17. A. R. Horswill, J. C. Escalante-Semerena, *J. Bacteriol.* **181**, 5615 (1999).
18. ———, *Microbiology* **145**, 1381 (1999).
19. E. Rodriguez, H. Gramajo, *Microbiology* **145**, 3109 (1999).
20. H. Gramajo, unpublished results.
21. A. Chapman-Smith, D. L. Turner, J. E. Cronan Jr., T. W. Morris, J. C. Wallace, *Biochem. J.* **302**, 881 (1994).
22. ^{14}C-1-propionate feeding was performed as follows: Individual transformants were inoculated into M9 minimal media cultures with glucose [T. Maniatis *et al.*, Eds., *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 1982), p. 440] in the presence of carbenicillin (50 μg/ml), kanamycin (25 μg/ml), and chloramphenicol (17 μg/ml) at 37°C and 250 rpm. Cultures were grown to mid-log phase (OD_{600} = 0.6 to 0.8), cooled at 22°C for 5 min, and then centrifuged. The cell pellets were resuspended in 1 ml of the remaining supernatant and induced with 1 mM IPTG and 0.25% arabinose (for pCY216). In addition, ^{14}C-1-propionate (at 56 mCi/mmol) was added at final concentration of 0.27 mM. The culture was then stirred for an additional 12 to 15 hours at 22°C. At this point the culture was centrifuged and 100 μl of the supernatant was extracted (2×) with ethyl acetate (300 μl each time). The extract was dried in vacuo and subjected to thin-layer chromatography (TLC) analysis. Negative controls included cultures of BAP1/pRSG32/pCY216 and BL21(DE3)/pRSG32/pTR132/pCY216. ^{13}C-1-propionate feeding was performed as follows: A single transformant of BAP1/pRSG32/pTR132/pCY216 was used to start a 3-ml LB culture with carbenicillin (100 μg/ml), kanamycin (50 μg/ml), and chloramphenicol (34 μg/ml) at 37°C and 250 rpm. The starter culture was used to inoculate 100-ml M9 minimal media with glucose at the same antibiotic concentrations as above. These cultures were grown at 250 rpm and 37°C to mid-log phase (OD_{600} = 0.5 to 0.7), cooled for 15 min in a 22°C bath, and induced with 500 μM IPTG and 0.25% arabinose. ^{13}C-1-propionate was added to a final concentration of 100 mg/liter, and the cultures were incubated at 22°C for 12 to 15 hours. The sample was then centrifuged and the supernatant extracted twice with 300-ml ethyl acetate. The sample was dried in vacuo, resuspended in $CDCl_3$, and analyzed via ^{13}C-NMR. A negative control was performed with BL21(DE3)/pRSG32/pTR132/pCY216.
23. The construction of plasmids pBP130 and pBP144 is described in the Supplemental Materials (*8*).
24. Plasmids pBP130, pBP144, and pCY216 were transformed into BAP1 as previously described. Culture conditions were identical to those described for ^{13}C-1-propionate fed at 250 mg/liter (*22*). Cultures were sampled regularly over 3 days. Samples were centrifuged and the supernatant (either 2 or 20 μl) loaded onto a Hewlett-Packard 1090 HPLC using an initial 4.6 × 10 mm column (Inertsil, C18 ODS3, 5 μm; Metachem Technologies, Torrance, CA), was washed with water (1 ml/min for 2 min), and then was loaded onto a main 4.6 × 50 mm column with the same stationary phase and flow rate. A 6-min gradient was then applied starting with 100% water and finishing with 100% acetonitrile maintained for an additional 1 min. The samples were analyzed with an Alltech evaporative light scattering detection system (ELSD500; Deerfield, IL), and a peak at 6.4-min retention time was confirmed as hepta–^{13}C-labeled 6dEB by mass spectrometry (MW_{obs} = 393). Product concentrations were measured in comparison to standard 6dEB samples using the same detection scheme. At the end of the incubation period, the entire culture supernatant was extracted as before with ethyl acetate, dried, and analyzed by ^{13}C-NMR. Additionally, the final cell pellet was analyzed via SDS-PAGE to confirm the presence of the three DEBS proteins and the PCC. No differences were observed between the expression levels of the proteins at 12 and 48 hours post-induction. The stability of each plasmid in BAP1/pBP130/pBP144/pCY216 was also tested at 12 and 36 hours post-induction. No loss of pBP144 was observed at either time point, whereas pBP130 and pCY216 were maintained in 50 and 35% of the colonies at 12 and 36 hours, respectively. No rearrangement of any plasmid was detected at either time point, based on restriction analysis of multiple re-transformed colonies. Negative controls for the ^{13}C-NMR experiments included BAP1/pBP130/pCY216, BAP1/pBP144/pCY216, and BAP1/pBP130/pBP160/pCY216. {Plasmid pBP160 carries a C to A null mutation at the active site of the KS domain in module 1 [C. Kao, *et al.*, *Biochemistry* **35**, 12363 (1996)].}
25. W. Minas, P. Brunker, P. T. Kallio, J. E. Bailey, *Biotechnol. Prog.* **14**, 561 (1998).
26. H. D. Mootz, M. A. Marahiel, *Curr. Opin. Chem. Biol.* **1**, 543 (1997).
27. S. J. Admiraal, C. T. Walsh, C. Khosla, in preparation.
28. The construction of plasmid pBP165, carrying the rifamycin loading didomain fused to DEBS1 as well as the *pccAB* genes, is described in the Supplemental Materials (*8*).
29. ^{13}C-NMR ($CDCl_3$, 500 MHz) δ: 213.76, 177.43, 79.70, 76.60, 71.24, 37.72 (enriched carbon atoms only). Mass Spectrometry (AP-CI) for expected $^{12}C_{19}$-$^{13}C_6H_{38}O_6Na$, 463.2757; observed, 463.2847.

30. Supported by grants from the NSF (BES 9806774 to C.K.), the NIH (CA 66736 to C.K. and GM 22172 to D.E.C.), and grant ANPCyT (01-00078-01686 to H.G.). B.A.P. is a recipient of a predoctoral fellowship from the Stanford-NIH training program in biotechnology. S.J.A. is a recipient of a fellowship from the Cancer Research Fund of the Damon Runyon–Walter Winchell Foundation (DRG-1573). The assistance of Q. Li (^{13}C-NMR), T. Leaf (HPLC), and J. Carney (mass spectrometry) is gratefully acknowledged.

7 December 2000; accepted 23 January 2001