Specificity and Versatility in Erythromycin Biosynthesis

Rembert Pieper, Camilla Kao and Chaitan Khosla*
Department of Chemical Engineering, Stanford University, Stanford, CA 94305–5025, USA
Guanglin Luo and David E. Cane*
Department of Chemistry, Brown University, Box H, Providence, RI 02912, USA

1 Introduction

More than 40 years ago, Woodward and Gerzon suggested that macrolide antibiotics such as erythromycin A (**1**) might be formed from simple propionate building blocks.[1,2] Since that time, not only has this prediction proved to be remarkably perceptive, but a wealth of information has been gained concerning the biosynthesis of this medicinally important class of natural products. Incorporation experiments with [^{14}C]-, [^{13}C]-, [^{18}O]-, and [^{2}H]-labelled substrates and intermediate analogues have confirmed the propionate origin of erythromycin and related metabolites and established that formation of the parent erythromycin macrolide, 6-deoxyerythronolide B (6-dEB, **2**), occurs by a processive mechanism in which the oxidation level and stereochemistry of the growing polyketide chain are adjusted immediately after each step of polyketide chain elongation.[3–5] Intriguingly, although more than 100 individual macrolides, made up of various combinations of acetate, propionate, and butyrate subunits, have been identified, all of these metabolites can be described by a general stereochemical model, illustrated in Figure 1, first proposed by Celmer.[6] Indeed, the existence of such striking regularities among a large number of metabolites, produced by a wide range of Actinomycete species, first suggested the possible modularity of the biosynthetic enzymes responsible for the formation of these polyketide natural products.[7] The central challenge has, therefore, been to unravel the mystery that shrouds the molecular genetic and biochemical basis for the intricate programming of the biosynthesis of complex polyketides.

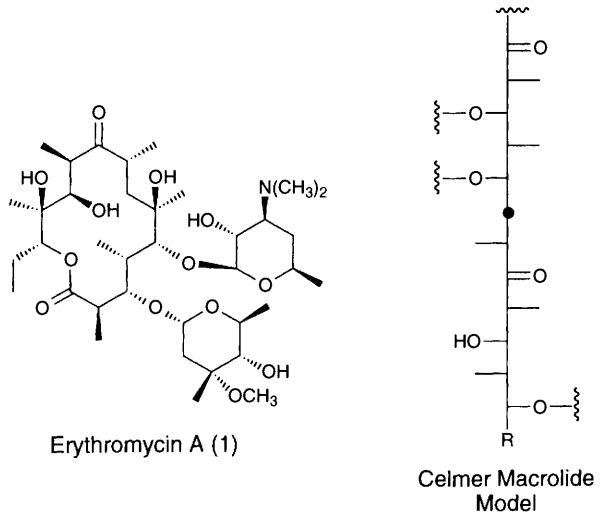
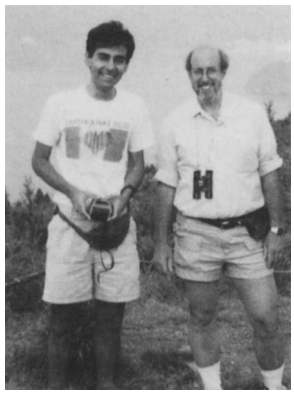


Figure 1 The broad spectrum antibiotic erythromycin A **1** and Celmer's macrolide stereochemical model.

2 Isolation and Characterization of the Genes for a Modular Polyketide Synthase

The first direct experimental evidence for the modular hypothesis came from the work of the groups of Peter Leadlay at the University of Cambridge and Leonard Katz at the Abbott Laboratories.[8,9] Working independently, these two groups of investigators demonstrated that the structural genes responsible for the formation of 6-dEB consist of three contiguous open reading frames of 10 kb each, encoding three large (*ca.* 3000 amino acid) multidomain proteins, designated deoxyerythronolide B synthase (DEBS) 1, 2, and 3 by Leadlay (Figure 2). Detailed sequence comparisons revealed that each of these proteins consists of eight to ten domains with considerable sequence similarity to enzymes responsible for each of the individual steps of fatty acid biosynthesis. Moreover, these domains



Chaitan Khosla (left)
David Cane (right)

Rembert Pieper was born in 1963 in Osnabruck, Germany, and graduated from Philipps Universität, Marburg, Germany, with a degree in pharmacy in 1989. This was followed by a PhD from the Technische Universität, Berlin, in 1993 on the isolation and characterization of enniatin synthetase, postdoctoral work (1994–1996) at Stanford University on the enzymology of polyketide synthases, and a current position as research associate at the National Institutes of Health, Bethesda.

Camilla Kao was born in 1970 in Midland, MI, and graduated in chemical engineering from Rice University, Houston, TX in May 1992. This was followed by graduate school at Stanford University in September 1992, also in chemical engineering, and a current position as Lieberman Fellow at Stanford University.

Chaitan Khosla (born 1964) is Assistant Professor of Chemical Engineering, of Chemistry (by courtesy), and of Biochemistry (by courtesy) at Stanford University. He received his PhD in 1990 at Caltech. After completing his postdoctoral studies in the laboratory of Professor Sir David Hopwood at the John Innes Centre in the UK, he joined Stanford in 1992. His current research interests focus at the interface of biological chemistry and biomolecular engineering.

Guanglin Luo (born 1967) graduated from Fudan University in Shanghai with a BSc in 1987. After receiving his MSc in organic chemistry in 1990 from the Shanghai Institute of Organic Chemistry, Academia Sinica, where he obtained the First Grade Award as one of the top four students, he entered the PhD programme in chemistry at Brown University. Upon receiving his PhD in 1996, he was awarded the Potter Prize for the outstanding thesis in chemistry. He is currently a postdoctoral researcher in the laboratory of Professor E. J. Corey at Harvard University.

David E. Cane (born 1944) is Vernon K. Krieble Professor of Chemistry and Professor of Biochemistry at Brown University. After receiving his PhD in 1971 from Harvard University under Prof. E. J. Corey, he carried out postdoctoral research with Professor Duilio Arigoni at the ETH in Zürich. He joined the faculty at Brown in 1973. His current research interests are centred on the biosynthesis of natural products, with special emphasis on the mechanistic enzymology of isoprenoid and polyketide biosynthesis.

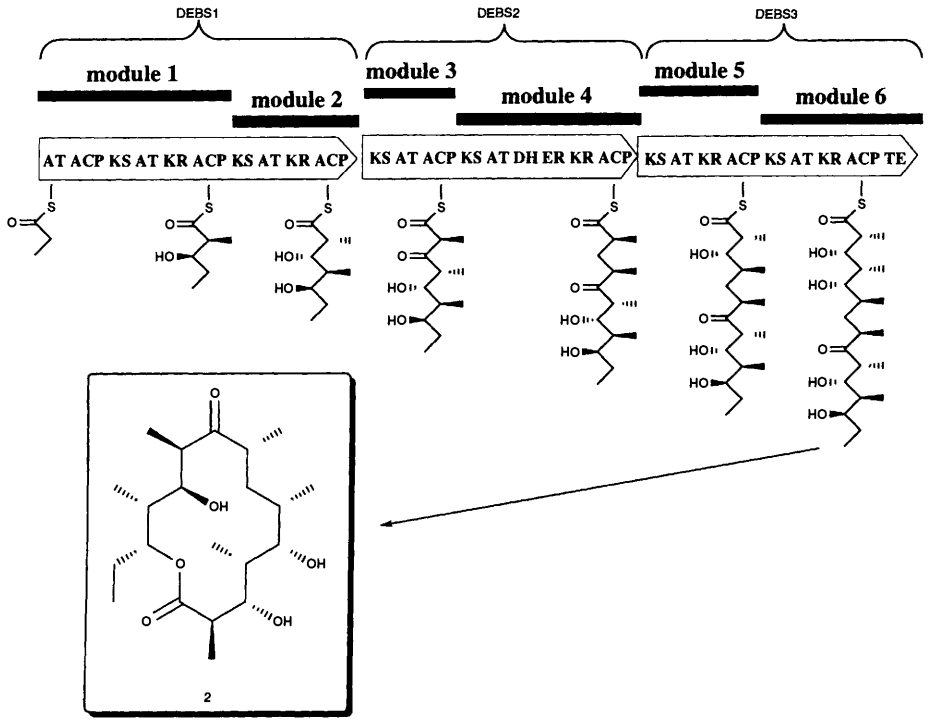


Figure 2 Model for the modular organization of 6 deoxyerythronolide B synthase (DEBS) and the biosynthesis of 6-deoxyerythronolide B (2) by DEBS1 + 2 + 3 Each DEBS subunit carries two complete modules and each of the six modules accounts for one cycle of polyketide chain extension and β-ketore duction, as appropriate The active sites are designated as follows acyltransferase (AT), β-ketoacyl ACP transferase (KS), acyl carrier protein (ACP), β-ketoreductase (KR), dehydratase (DH), enoylreductase (ER), and thioesterase (TE)

are arranged such that each protein contains two functional units or modules, each of which carries all the requisite catalytic activities for one of six cycles of polyketide chain elongation and reductive modification of the resultant β-ketoacyl thioester The significance of this discovery cannot be overstated Not only did the availability of the structural genes for 6-DEB synthase provide an invaluable tool for much of the subsequent experimental work in the polyketide synthase area, but the model that emerged from these studies reshaped the thinking about the programming of complex polyketide synthases and has provided the conceptual framework for the design of many of the most important experiments carried out over the last several years

According to the now widely accepted model (Figure 2), the acyltransferase (AT) domain at the N-terminus of DEBS1 initiates the polyketide chain-building process by transferring the propionyl-CoA primer unit, *via* the pantetheinyl residue of the first acyl carrier protein (ACP) domain, to the active site cysteine of the ketosynthase of module 1 (KS1) The acyltransferase in module 1 (AT1) loads methylmalonyl-CoA onto the thiol terminus of the ACP domain of module 1 KS1 then catalyses the first polyketide chain elongation reaction by decarboxylative acylation of the methylmalonyl residue by the propionyl starter unit, resulting in the formation of a 2-methyl-3-ketopentanoyl-ACP thioester The latter intermediate is then reduced by the ketoreductase of module 1 (KR1), giving rise to enzyme-bound (2*S*,3*R*)-2-methyl-3-hydroxypentanoyl-ACP At this point, module 1 has finished its task and the diketide product is transferred to the core cysteine of KS2, whereupon it undergoes another round of condensation and reduction, resulting in the formation of the corresponding triketide This process is repeated several times, with each module being responsible for a separate round of polyketide chain elongation and reduction, as appropriate, of the resulting β-ketoacyl thioester Finally, the thioesterase (TE) at the C-terminus of DEBS3 is thought to catalyse release of the finished polyketide chain by lactonization of the product generated by module 6

Following the characterization of the DEBS genes, Leadlay and coworkers succeeded in purifying the corresponding three DEBS proteins from the natural erythromycin producer, *Saccharopolyspora erythraea* The three proteins were, as predicted from the DNA sequence, of unusually large size – DEBS1 (M_r 370000), DEBS2 (M_r 380000) and DEBS3 (M_r 330000) [10] Partial proteolysis studies further established that propionyl-CoA specifically acylates the N-terminal domain of DEBS1, consistent with the proposed role of this region in loading the propionate starter on the polyketide synthase [11] In a very important set of experiments, the Cambridge group also established that (2*S*)-methylmalonyl-CoA is the exclusive substrate for polyketide chain elongation, based on the stereospecific acylation both of intact DEBS proteins and of selected partial proteolytic fragments [12] Unfortunately, neither the native protein preparations isolated from *Sac erythraea*, nor recombinant derivatives expressed in *Escherichia coli*, which lacked the requisite pantetheinyl moieties, were able to catalyse polyketide chain elongation In fact, until recently, there had been no reports of successful cell-free synthesis of macrolide-type polyketides, in spite of more than 30 years of intense efforts by numerous research groups

Very recently, the Cambridge group have reported the results of experiments which shed light on the subunit organization of the DEBS multienzyme system [13] Gel filtration, and ultracentrifugation experiments both indicate that the individual DEBS proteins (as well as entire modules derived therefrom) are associated as homodimers These conclusions were reinforced by crosslinking experiments in which purified module 5, obtained by partial elastase digestion of DEBS3, was crosslinked with 1,3-dibromopropanone, a reagent previously used to crosslink the sulfhydryl residues of the 4′-phosphopantetheine of the ACP domain and the active site cysteine of the ketosynthase in yeast and animal fatty acid synthases In a control experiment, an elastase fragment representing module 6 but lacking its ACP did not undergo dimerization upon addition of 1,3-dibromopropanone These results suggest that the ACP of one module interacts with the KS from its identical partner within each homodimeric unit

3 Genetic Manipulation of DEBS

The availability of the cloned DEBS structural genes opened the door to genetic modification of the DEBS proteins themselves In a

pioneering experiment, Katz and his coworkers generated a *Sac erythraea* mutant carrying a large, in-frame deletion in the ketoreductase domain of DEBS module 5 (KR5) and demonstrated that this mutant produced erythromycin analogues derived from **3** with a keto group at the predicted site, C-5[9] (Figure 3) The latter experiment provided not only direct experimental verification of the modular hypothesis suggested by the DEBS gene sequences, but established that the downstream domains in module 6 were capable of processing modified polyketide chain-elongation intermediates, thereby opening up the exciting possibility of rationally engineering the production of novel polyketide metabolites The Katz group also effected a similar reprogramming of polyketide synthesis by mutation of the presumed NADPH binding motif of the enoyl reductase domain of module 4 (ER4) The resulting mutant strain produced macrolides derived from **4** with the predicted $\Delta^{6,7}$-anhydroerythronolide skeleton [14]

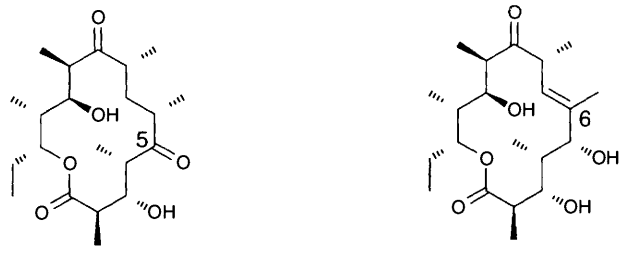
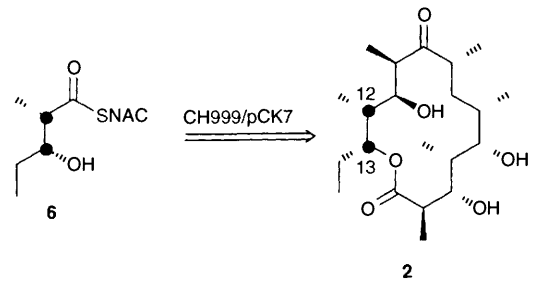


Figure 3 Novel analogues of 6-dEB produced by DEBS mutants carrying a deletion in KR5 (compound **3**) or a mutation in ER4 (compound **4**)

In 1994, using a specially engineered host–vector system for the expression of recombinant polyketide synthases, we succeeded in expressing the complete set of DEBS structural genes in an actinomycete host species, *Streptomyces coelicolor*, which normally produces neither erythromycin nor any other macrolides [15] The resultant strain, designated *S coelicolor* CH999/pCK7, produced substantial quantities (> 40 mg dm $^{-3}$) of 6-deoxyerythronolide B **2**, accompanied by a novel cometabolite, 8,8a-deoxyoleandolide **5** (> 10 mg dm $^{-3}$) (Figure 4a) Analysis of the protein constituents of *S coelicolor* CH999/pCK7 by sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS–PAGE) revealed the presence of three characteristically large proteins, DEBS1, 2 and 3 The production of 6-dEB demonstrates that DEBS1, 2, and 3 carry all the necessary biosynthetic activities to support generation of the full-length polyketide and cyclization to the aglycone **2** Furthermore, it is evident that ancillary activities required for the essential phosphopantetheinylation of the ACP domains are present in the host strain, and that the recombinant DEBS is fully functional in the heterologous host That 6-dEB is being formed by the normal biosynthetic pathway was confirmed by the incorporation of [1-^{13}C]propionate, giving rise to the expected labelling pattern in the ^{13}C NMR spectrum of the resultant sample of 6-dEB In an analogous experiment, the starter unit of 8,8a-deoxyoleandolide **5** was labelled by [1,2-$^{13}C_2$]acetate, thereby confirming that DEBS can tolerate an acetate starter in place of its normal propionyl-CoA substrate, and re-emphasizing a potential catalytic flexibility first suggested by the molecular genetic experiments of Katz The utilization of acetyl-CoA as a starter is presumably due to the lower intracellular concentration of propionyl-CoA in *S coelicolor* as compared to the native erythromycin producer, *Sac erythraea* Together, these results have raised intriguing possibilities for the use of PKS systems for the controlled formation of 'unnatural' natural products by rational control of the modular composition of the PKS, as well as the structures and relative amounts of the available substrates

In further experiments with *S coelicolor* CH999/pCK7, feeding of (2*S*,3*R*)-[2,3-$^{13}C_2$]-2-methyl-3-hydroxypentanoyl-*N*-acetylcysteamine (NAC) thioester **6** to the engineered organism resulted in the formation of 6-dEB **2** labelled with ^{13}C at C-12 and C-13, as evidenced by the appearance of the predicted set of enhanced and coupled doublets, consistent with the intact incorporation of the diketide chain elongation intermediate[16] (Scheme 1) The level of



Scheme 1 Incorporation of the intact chain elongation intermediate **6** into 6 dEB **2** by *S coelicolor* CH999/pCK7

enrichment (15—20 atom%) was especially noteworthy, being nearly 100 times more efficient than the levels usually observed for the incorporation of NAC thioesters into microbial polyketide metabolites Whatever the levels of incorporation, however, it is evident that the DEBS protein can recognize the relevant structural and stereochemical features of the exogenously administered NAC thioester and load the intermediate analogue on to the appropriate ketosynthase domain, presumably KS2, from where it will be processed in the normal manner These results, as well as the intact incorporation of advanced polyketide chain elongation intermediates into a wide variety of other polyketides, strongly suggest that superimposed on the purely organizational level of control over the programming of polyketide biosynthesis intrinsic to the sequential organization of the modular DEBS proteins, there is an additional level of substrate molecular recognition exercised by the various catalytic domains

In order to explore further the function of the modular DEBS proteins, we next constructed a plasmid, pCK9, carrying only the genes for the first open reading frame of the DEBS cluster[17] (Figure 4b) When this plasmid was used to transform *S coelicolor* CH999, protein extracts of the resultant recombinant strain were found to contain a protein shown to be identical with DEBS1 by a variety of methods Moreover, *S coelicolor* CH999/pCK9 produced 1—3 mg dm $^{-3}$ of a triketide lactone, (2*R*,3*S*,4*S*,5*R*)-2,4-dimethyl-3,5-dihydroxy-*n*-heptanoic acid δ-lactone **7**, the structure of which was unambiguously established by direct spectroscopic and chromatographic comparison with an authentic sample of **7** prepared by total synthesis Feeding of [1-^{13}C]propionate resulted in the formation of **7** enriched as expected at C-1, C-3 and C-5 These results established for the first time that the DEBS1 protein is fully competent to support the first two cycles of polyketide biosynthesis involved in the formation of 6-dEB, and does not require association with either DEBS2 or DEBS3 for activity Interestingly, **7**, which is generated by lactonization of the corresponding acyclic triketide intermediate attached to the ACP of module 2, has previously been reported by Katz to be an abortive chain-elongation product generated by the DEBS construct carrying the deletion in KR5 [18]

It is apparent that the triketide lactone can be released from DEBS1 alone without a requirement for the thioesterase (TE) domain In order to analyse the substrate specificity and function of the thioesterase, we next constructed yet another mutant CH999/pCK12 which expresses a PKS, DEBS1 + TE, in which the thioesterase, originally at the C-terminus of DEBS3, has now been fused to the C-terminus of DEBS1 [19] (Figure 4c) In fact, the latter strain not only produced substantially enhanced amounts of the triketide lactone **7** (> 20 mg dm $^{-3}$), but up to 10 mg dm $^{-3}$ of a cometabolite originating from incorporation of an acetate starter, (2*R*,3*S*,4*S*,5*R*)-2,4-dimethyl-3,5-dihydroxy-*n*-hexanoic acid δ-lactone **8** As with the formation of 8,8a-deoxyoleandolide **5** from an acetate starter, the production of **8** most likely reflects some combination of the limited availability of propionyl-CoA in the *S coelicolor* host strain, as well as the greater abundance of the alternative substrate acetyl-CoA

Meanwhile, the Cambridge group had reported analogous results

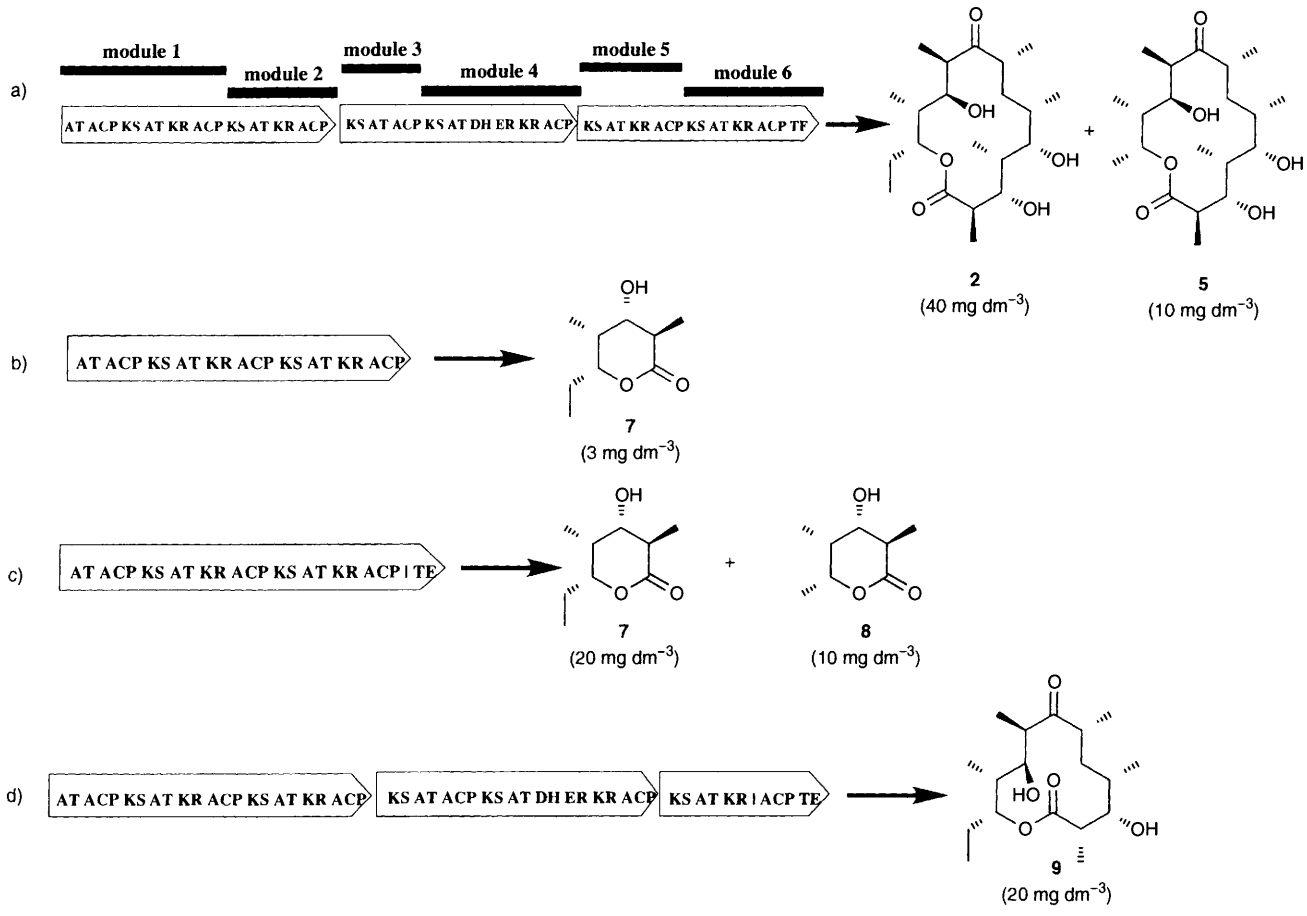


Figure 4 Polyketides produced by engineered *S coelicolor* CH999 strains (a) CH999/pCK7, (b) CH999/pCK9, (c) CH999/pCK12, (d) CH999/pCK15

on the formation of **7** [20] Using genetic recombination techniques in *Sac erythraea*, they fused the thioesterase domain from DEBS3 to the C-terminus of DEBS1 Through a second recombination they also deleted the structural genes for DEBS2 and DEBS3 Each of the resultant mutants produced the expected triketide lactone **7**, while neither produced the macrolactone 6-dEB, indicating that not only was the thioesterase capable of cyclizing the normal triketide product of DEBS1, but that this process completely suppresses transfer of the acyclic triketide intermediate to the downstream enzymes By contrast, a mutant carrying an inactive TE fused to the C-terminus of DEBS1 produced significantly lower levels of triketide lactone (< 0 1 mg dm ^{3}) These results not only confirmed that DEBS1 is fully competent to support the formation of triketide, but that the thioesterase can play an active role in catalysing the release of this substrate

In a further series of experiments, the Cambridge group also expressed a DEBS1 + TE construct in *S coelicolor* [21] Consistent with our own results with DEBS1 + TE in this organism, both the triketide lactone **7** and the C_8-lactone **8** were produced, but the natural tripropionate triketide **7** was found to be the minor constituent The reasons for these differences in product ratios is unclear, given the more than 30-fold preference of DEBS1 + TE for a propionyl-CoA over an acetyl-CoA primer (see below), but most likely reflects differences in timing and levels of protein expression as well as the varying sizes of intracellular pools of the two primers and the methylmalonyl extender under the growth conditions used for *S coelicolor* by the two research groups

Our group has constructed yet another deletion mutant of DEBS, expressed by *S coelicolor* CH999/pCK15, containing the first four DEBS modules, present as DEBS1 and DEBS2, plus a chimeric fifth module in which the ACP6 TE didomain region has now been fused just downstream of KR5[19] (Figure 4d) Assuming that KS5 and ACP6 could productively interact to catalyse a fifth condensation, it was expected that the mutant strain would be able to support five complete rounds of polyketide chain elongation and that the thioesterase/cyclase would catalyse the lactonization of the resultant polyketide Indeed, *S coelicolor* CH999/pCK15 produced 20 mg dm ^{3} of a completely new macrolactone, (8*R*,9*S*)-8,9-dihydro-8-methyl-9-hydroxy-10-deoxymethynolide **9** The structure of **9**, which is a novel analogue of 10-deoxymethynolide, the aglycone of the macrolide antibiotic methymycin, was unambiguously confirmed by extensive 2-D NMR and mass spectrometric analysis, supplemented by specific labelling by a variety of [^{13}C]enriched propionate precursors The latter results established that the TE activity, which naturally supports the formation of a 14-membered ring product, catalysed exclusive formation of the 12-membered ring lactone, by esterification of the acyl thioester with the C-11 hydroxy, in preference to the ordinarily kinetically favoured generation of a δ-lactone by esterification with the C-5 hydroxy These results further confirmed the structural and functional independence of individual modules of modular PKSs They also demonstrated the feasibility of constructing hybrid modules *via* genetic engineering

4 Cell-free Formation of Polyketides

Unlike the related peptide synthetases, for which well-developed enzymological methods have been available for the past several decades,[22] detailed mechanistic studies on PKSs have been seriously hampered by a lack of fully active cell-free systems Indeed, until quite recently, 6-methylsalicylic acid synthase and its closely related homologue, orsellinic acid synthase, were the only known examples of microbial PKSs the activities of which (including chain-elongation activity) had been reconstituted *in vitro* [23–25] The situation has changed dramatically over the past few years with several reports demonstrating enzymatic synthesis of polyketides using the PKS responsible for the polyketide component of cyclosporin,[26] the tetracenomycin synthase,[27] truncated forms of

modular PKSs,[28 29] and the complete DEBS assembly [28] In at least the last three cases, the development of fully active cell-free systems has benefited from the availability of high-level expression systems derived *via* genetic engineering

As described above, DEBS1, 2, and 3 are large multifunctional proteins carrying a total of at least 28 distinct active sites The DEBS proteins were present in a cell-free protein preparation from *S coelicolor* CH999/pCK7 at a level of *ca* 3—5% total cellular protein [30] In the presence of 150 mmol dm^{-3} sodium phosphate buffer, the multi-enzyme assembly was found to catalyse the formation of 6-dEB **2**, as well as the abortive chain elongation product **7**, upon addition of propionyl-CoA, (2*RS*)-methylmalonyl CoA, and NADPH[28] (Scheme 2a) A high phosphate concentration in the protein preparation and reaction buffers was found to be very important for the observed enzymatic activity, presumably by enhancing the assembly of the multi-enzyme complex *via* hydrophobic interactions [28 30] The DEBS-catalysed formation of 6-dEB, as well as the formation of the triketide lactone **7** described below, is inhibited by both cerulenin and *N*-ethylmaleimide, each a well-known inhibitor of the condensation reactions of fatty acid biosynthesis [28] The apparent k_{cat} parameters for the formation of 6-dEB and the triketide lactone **7** by DEBS1, 2 and 3 are 0 5 and 0 23 min^{-1}, respectively, pointing to the relative inefficiency of chain transfer from DEBS1 to DEBS2 *in vitro* [30] The complete DEBS system has been substantially purified (to > 50% homogeneity) without significant losses in specific activity The individual recombinant proteins behave, as expected, as homodimers upon gel filtration

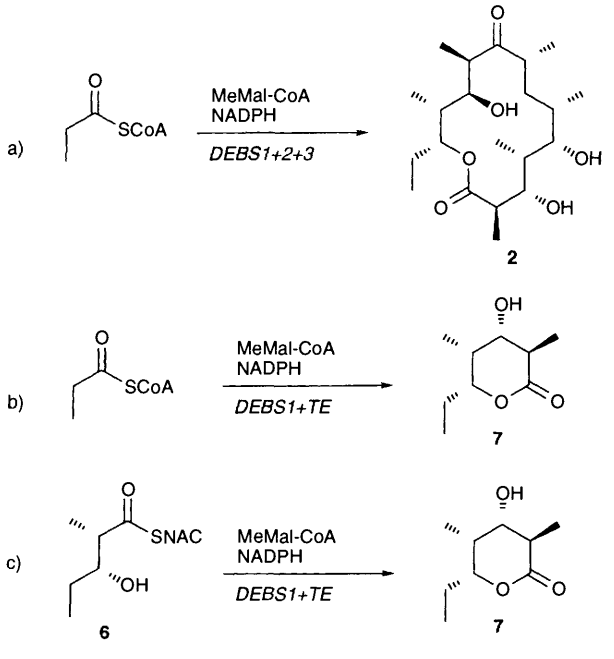
Analogous to the above studies with the complete DEBS system, similar (and more extensive) investigations have also been carried out on the formation of the triketide lactone **7** by the truncated DEBS1 + TE protein (Figure 2b) Despite the relative simplicity of the system, it harbours key molecular recognition features of the overall DEBS system the first two modules generate methyl-branched carbon centres as well as secondary alcohols with both D and L stereochemistry Furthermore, this mini-PKS has been shown to be highly active as judged by a k_{cat} value (3 4 min^{-1}) that is comparable to the estimated rate constant *in vivo*[30] and by the fact that products can be synthesized on scales that facilitate structural analysis *via* NMR spectroscopy [28] The apparent K_m for methylmalonyl-CoA in DEBS1 + TE catalysed synthesis of the tripropionate lactone **7** is 24 μmol dm^{-3} [30] In contrast, the K_m for propionyl-CoA is not easily measured, since the enzyme can readily decarboxylate methylmalonyl-CoA (or methylmalonyl-ACP) to generate a propionate primer which is turned over into **7** without any effect on the apparent k_{cat} [30] In the presence of (2*RS*)-methylmalonyl-CoA and NADPH, DEBS1 + TE could also convert the exogenously added diketide chain elongation intermediate, (2*S*,3*R*)-2-methyl-3-hydroxypentanoyl-NAC thioester **6**, to the triketide lactone **7**, as verified by both ^{14}C- and ^{13}C-labelling experiments[28] (Scheme 2c), completely consistent with the previously described experiments with intact cells which had demonstrated the incorporation of **6** into 6-deoxyerythronolide B **2**

DEBS1 + TE has a remarkably broad specificity towards alternative primer units In addition to propionyl-CoA, both acetyl- and butyryl-CoA can serve as surrogate chain initiators, giving rise to the corresponding triketide lactones **8** and **10**, respectively[29 31] (Scheme 3) Consistent with these observations, DEBS1 + TE can be acylated by radiolabelled acetyl-, propionyl-, and butyryl-CoA with comparable efficiency [31] Preincubation of DEBS1 + TE with iodoacetamide fails to inhibit acylation by propionyl-CoA,[28] while partial proteolysis of the labelled protein indicates exclusive acylation of the N-terminal portion of the protein containing the first AT domain [11] Notwithstanding this tolerance for different starter units, however, DEBS1 + TE exhibits a 32-fold and an eight-fold kinetic preference towards propionyl primers over acetyl and butyryl primers, respectively, suggesting the existence of one or more active sites with discriminating molecular recognition features [30]

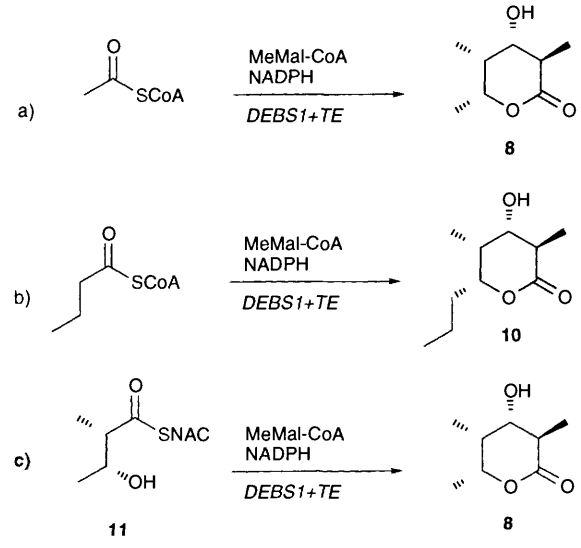
The broad substrate specificity of DEBS towards unnatural substrates is also illustrated by two other types of experiments First, DEBS1 + TE can also recognize the unnatural diketide **11**, processing it to the corresponding C_8 triketide lactone **8**[31] (Scheme 3c) Second, if NADPH is excluded from the reaction mixture, the enzyme can convert propionyl-CoA and methylmalonyl-CoA to the pyran-2-one **12**, presumably formed by lactonization of the unreduced diketoacylthioester product[31] (Scheme 4a) Alternatively, incubation of DEBS1 + TE with methylmalonyl-CoA and the NAC-diketide **6** in the absence of NADPH leads to formation of the ketolactone **13**[32] (Scheme 4b), emphasizing the ability of the DEBS protein to mediate the formation of polyketides in a variety of oxidation states

5 Programming and Reprogramming of Modular Polyketide Synthases

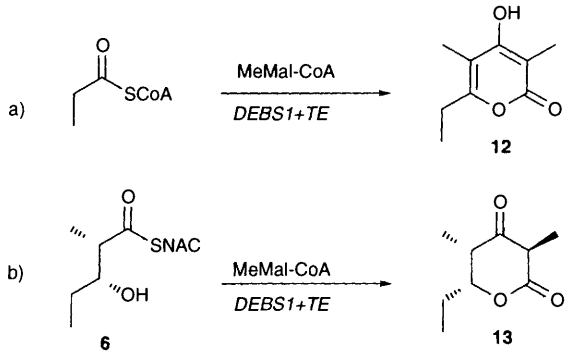
The example of the erythromycin PKS presents an elegant evolutionary solution to the problem of programming a complex sequence of biosynthetic reactions based on the repetitive utilization of a small repertoire of biochemical transformations Two levels of catalytic control are evident in this multi-enzyme assembly First, the modular



Scheme 2 Enzyme catalysed formation of (a) 6 dEB catalysed by DEBS1 + 2 + 3, (b) and (c) triketide lactone **7** by DEBS1 + TE



Scheme 3 Broad substrate specificity of the DEBS enzyme Processing of anomalous primer substrates (a) acetyl CoA, (b) butyryl CoA and (c) **11** by DEBS1 + TE



Scheme 4 Catalysis of polyketide chain elongation by DEBS1 + TE in the absence of NADPH using a) propionyl-CoA and b) **6** as substrates

structure of the proteins provides organizational control at the level of dictating the sequence of reactions to be employed in the overall catalytic cycle Secondly, the molecular recognition features of some or all individual domains introduce an additional level of selectivity into the multi-step transformation Importantly, neither of these two control mechanisms results in absolute specificity, as illustrated vividly by several examples reviewed here Thus, the intrinsic tolerance within modular PKSs towards reprogramming presents an exciting opportunity for the rational design of novel 'unnatural' natural products, and for the combinatorial generation of molecular diversity within this medicinally important family of molecules

Several diverse and complementary strategies for genetic and/or chemical reprogramming can be envisioned As discussed above, a few examples suggest the feasibility of genetically knocking out individual active sites without impairing the remainder of the catalytic cycle In an extreme case, it may even be possible to delete entire modules, as illustrated by experiments in which the terminal thioesterase is fused to various upstream modules Beyond inactivation, it will be interesting to explore the extent to which individual modules (or domains therein) can be substituted by heterologous modules/domains with unnatural molecular recognition features This has been successfully demonstrated in the cases of the structurally smaller aromatic PKSs as well as the modular peptide synthetases Finally, the potential for genetically engineering new catalytic functions into existing PKS pathways also remains to be evaluated

In addition to genetic reprogramming, the availability of fully active cell-free systems, in conjunction with facile mutagenesis tools, opens up new possibilities for chemically reprogramming modular PKSs to produce polyketides that might otherwise be inaccessible *via in vivo* engineered biosynthesis For example, as reviewed above, modular PKSs can turn over a variety of unnatural substrates into polyketide products They can also function under non-biological conditions, such as in the absence of reducing equivalents Given the structural complexity of most natural or engineered products of modular PKSs, 'one-pot' enzymatic synthetic methodologies could be an attractive complement to established chemical synthesis efforts aimed at elucidating the structure–activity relationships of lead molecules with potential human therapeutic, veterinary and agrochemical utility

Acknowledgments. This work was supported by a grant from the National Institutes of Health (GM22172) to D E C and in part by a grant from the National Institutes of Health (CA 66736-01), a National Science Foundation Young Investigator Award, and a David and Lucile Packard Fellowship for Science and Engineering 0to C K

6 References

1 R B Woodward, *Angew Chem* , 1957, **69**, 50
2 K Gerzon, E H Flynn, M V Sigal, P F Wiley, R Monahan and U C Quarck, *J Am Chem Soc* , 1956, **78**, 6398
3 D E Cane, H Hasler, P B Taylor and T -C Liang, *Tetrahedron*, 1983, **39**, 3449
4 D E Cane, T C Liang, P B Taylor, C Chang and C C Yang, *J Am Chem Soc* , 1986, **108**, 4957
5 D E Cane and C -C Yang, *J Am Chem Soc* , 1987, **109**, 1255
6 W D Celmer, *J Am Chem Soc* , 1965, **87**, 1801
7 D E Cane, W D Celmer and J W Westley, *J Am Chem Soc* , 1983, **105**, 3594
8 J Cortes, S F Haydock, G A Roberts, D J Bevitt and P F Leadlay, *Nature*, 1990, **348**, 176
9 S Donadio, M J Staver, J B McAlpine, S J Swanson and L Katz, *Science*, 1991, **252**, 675
10 P Caffrey, D J Bevitt, J Staunton and P F Leadlay, *FEBS Lett*, 1992, **304**, 225
11 J F Aparicio, P Caffrey, A F A Marsden, J Staunton and P F Leadlay, *J Biol Chem* , 1994, **269**, 8524
12 A F A Marsden, P Caffrey, J F Aparicio, M S Loughran, J Staunton and P F Leadlay, *Science*, 1994, **263**, 378
13 J Staunton, P Caffrey, J F Aparicio, G A Roberts, S S Bethell and P F Leadlay, *Nature Struct Biol* , 1996, **3**, 188
14 S Donadio, J B McAlpine, P J Sheldon, M Jackson and L Katz, *Proc Natl Acad Sci USA*, 1993, **90**, 7119
15 C M Kao, L Katz and C Khosla, *Science*, 1994, **265**, 509
16 D E Cane, G Luo, C Khosla, C M Kao and L Katz, *J Antibiot* , 1995, **48**, 647
17 C M Kao, G Luo, L Katz, D E Cane and C Khosla, *J Am Chem Soc* , 1994, **116**, 11612
18 S Donadio, M J Staver, J B McAlpine, S J Swanson and L Katz, *Gene*, 1992, **115**, 97
19 C M Kao, G Luo, L Katz, D E Cane and C Khosla, *J Am Chem Soc* , 1995, **117**, 9105
20 J Cortes, K E H Wiesmann, G A Roberts, M J B Brown, J Staunton and P F Leadlay, *Science*, 1995, **268**, 1487
21 M J B Brown, J Cortes, A L Cutter, P F Leadlay and J Staunton, *J Chem Soc Chem Commun* , 1995, 1517
22 H Kleinkauf and H von Dohren, *Eur J Biochem* , 1990, **192**, 1
23 P Dimroth, H Walter and F Lynen, *Eur J Biochem* , 1970, **13**, 98
24 J B Spencer and P M Jordan, *Biochemistry*, 1992, **31**, 9107
25 E R Wo, I Fujii, Y Ebizuka, U Sankawa, A Kawaguchi, S Huang, J M Beale, M Shibuya, U Mocek and H G Floss, *J Am Chem Soc* , 1989, **111**, 5498
26 M Offenzeller, Z Su, G Santer, H Moser, R Traber, K Memmert and E Schneider Scherzer, *J Biol Chem* , 1993, **268**, 26127
27 B Shen and C R Hutchinson, *Science*, 1993, **262**, 1535
28 R Pieper, G Luo, D E Cane and C Khosla, *Nature*, 1995, **378**, 263
29 K E H Wiesmann, J Cortes, M J B Brown, A L Cutter, J Staunton and P F Leadlay, *Chem Biol* , 1995, **2**, 583
30 R Pieper, S Ebert Khosla, D E Cane and C Khosla, *Biochemistry*, 1996, **35**, 2054
31 R Pieper, G Luo, D E Cane and C Khosla, *J Am Chem Soc* , 1995, **117**, 11373
32 G Luo, R Pieper, C Khosla and D E Cane, *Bioorg Med Chem* , 1996, **4**, 995