# Molecular recognition of diketide substrates by a β-ketoacyl-acyl carrier protein synthase domain within a bimodular polyketide synthase

Jo-Anne Chuck[1], Michael McPherson[1], Hai Huang[1], John R Jacobsen[2], Chaitan Khosla[2,3,4] and David E Cane[1]

**Background:** Modular polyketide synthases (PKSs) are large multifunctional proteins that catalyze the biosynthesis of structurally complex bioactive products. The modular organization of PKSs has allowed the application of a combinatorial approach to the synthesis of novel polyketides via the manipulation of these biocatalysts at the genetic level. The inherent specificity of PKSs for their natural substrates, however, may place limits on the spectrum of molecular diversity that can be achieved in polyketide products. With the aim of further understanding PKS specificity, as a route to exploiting PKSs in combinatorial synthesis, we chose to examine the substrate specificity of a single intact domain within a bimodular PKS to investigate its capacity to utilize unnatural substrates.

**Results:** We used a blocked mutant of a bimodular PKS in which formation of the triketide product could occur only via uptake and processing of a synthetic diketide intermediate. By introducing systematic changes in the native diketide structure, by means of the synthesis of unnatural diketide analogs, we have shown that the ketosynthase domain of module 2 (KS2 domain) in 6-deoxyerythronolide B synthase (DEBS) tolerates a broad range of variations in substrate structure, but it strongly discriminates against some others.

**Conclusions:** Defining the boundaries of substrate recognition within PKS domains is crucial to the rationally engineered biosynthesis of novel polyketide products, many of which could be prepared only with great difficulty, if at all, by direct chemical synthesis or semi-synthesis. Our results suggest that the KS2 domain of DEBS1 has a relatively relaxed specificity that can be exploited for the design and synthesis of medicinally important polyketide products.

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

Correspondence: David E Cane and Chaitan Khosla
E-mail: david_cane@brown.edu and ck@chemeng.stanford.edu

**Key words:** β-ketoacyl-acyl carrier protein synthase, erythromycin biosynthesis, polyketide synthase, substrate analog

Received: **21 August 1997**
Revisions requested: **8 September 1997**
Revisions received: **12 September 1997**
Accepted: **15 September 1997**

**Chemistry & Biology** October 1997, **4**:757–766
http://biomednet.com/elecref/1074552100400757

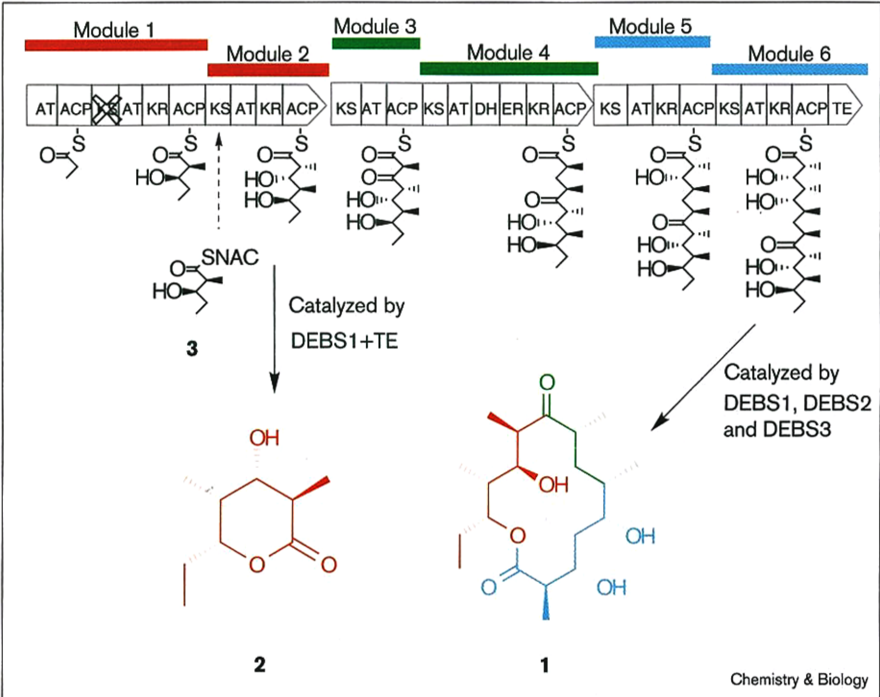
© Current Biology Ltd ISSN 1074-5521

## Introduction

Polyketides represent a growing family of natural products that exhibit a wide array of biological properties, including antibiotic, antiparasitic, antifungal, immunosuppressive and antitumor activities [1]. Each polyketide is synthesized through the repetitive condensation of simple monomers in a manner analogous to fatty acid biosynthesis [2,3]. Polyketide synthases (PKSs) introduce structural diversity in their enzymic products by using a variety of organic-acid extender units and controlling the degree of β-carbonyl reduction after each successive chain extension. The PKS also exhibits high stereochemical control over the new chiral centers formed during polyketide-chain elongation, while maintaining regioselectivity in the final lactone ring cyclization [4].

The factors which control the ultimate structure of a PKS product have been the object of intensive study, most notably through the manipulation of 6-deoxyerythronolide B synthase (DEBS), the PKS which catalyzes the biosynthesis of 6-deoxyerythronolide B (6-dEB; **1**; Figure 1), the parent macrolide aglycone of erythromycin [5]. The key to much of this effort has been the discovery that the structural genes responsible for 6-dEB synthesis consist of three contiguous open reading frames (of ~10 kb each), encoding three large multidomain proteins (~3000 amino acids each) designated DEBS1, DEBS2 and DEBS3 [6,7]. Each DEBS protein contains two functional units known as modules and each module in turn contains all the active sites required for one cycle of condensation and β-carbonyl processing. Unlike a typical fatty acid synthase (FAS) that uses the same set of active sites for repetitive rounds of chain elongation and reductive processing, the PKS multienzyme assembly carries a distinct active site for each individual catalytic step in carbon-chain formation and modification. The modular PKS can also control the level of β-ketoreduction through the presence of all, some, or none of the ketoreductase (KR), enoyl reductase (ER) and

**Figure 1**




Modular organization of the multidomain PKS, 6-deoxyerythronolide B synthase (DEBS). The complete DEBS system is composed of three large multifunctional polypeptides (each with a molecular weight > 300 kDa) designated DEBS1, DEBS2 and DEBS3. Each DEBS subunit consists of two functional modules, each consisting of the active sites required for one round of polyketide-chain elongation and any necessary reduction and modification of the β-keto function. DEBS1+TE is an engineered protein which consists of DEBS1 fused to the thioesterase domain of DEBS3. The biosynthesis of 6-deoxyerythronolide B (**1**) from propionyl-CoA, methylmalonyl-CoA, and NADPH is catalyzed by DEBS1, DEBS2, and DEBS3. DEBS1+TE catalyzes the formation of the $C_9$-triketide lactone (2*R*,3*S*,4*S*,5*R*)-2,4-dimethyl-3,5-dihydroxy-*n*-heptanoic acid δ lactone (**2**) from the same three substrates. Abbreviations are as follows: ACP, acyl carrier protein, AT, acyltransferase; DH, dehydratase, ER, enoyl reductase; KR, β-ketoreductase; KS, β-ketoacyl-ACP synthase; SNAC, N-acetyl cysteamine thioester; TE, thioesterase.

dehydratase (DH) domains, all of which are present in a vertebrate FAS.

Previous results have shown that the modular arrangement of the PKS can be exploited to create a wealth of novel products through domain inactivation [7–9], module deletion [10–13], domain replacement [14,15], and domain insertion ([16,17]), whilst maintaining the catalytic integrity of the enzyme. Although this powerful technique has paved the way towards the rational design of new compounds by combinatorial biosynthesis [18,19], proper exploitation of the biosynthetic power of modular PKSs to create novel metabolites with wide structural diversity requires an understanding of the balance between the intrinsic specificity and the structural permissiveness of the individual catalytic domains.

Our aim was to probe the structural features that are important in molecular recognition and processing of chain-elongation intermediates by a modular PKS. With this in mind, we decided to explore the limits of substrate specificity exhibited by a single PKS domain. DEBS1+TE is an engineered bimodular PKS in which the thioesterase (TE) domain of DEBS3 has been fused to the carboxyl terminus of DEBS1. DEBS1+TE catalyzes the formation of triketide lactone **2** (Figure 1) from one equivalent of propionyl-CoA and two equivalents of methylmalonyl-CoA, plus two equivalents of NADPH [20,21]. In the absence of propionyl-CoA, DEBS1+TE can supply its own primer via the enzyme-catalyzed decarboxylation of methylmalonyl-CoA [22,23]. The DEBS1+TE mini-PKS is also capable of processing **3** (Figure 1), the N-acetyl cysteamine (NAC) thioester of an advanced diketide intermediate to the corresponding triketide lactone **2** [22,23]. Radiolabeling studies have identified the specific site of transacylation by the diketide **3** as the active-site cysteine of KS2, the β-ketoacyl-acyl carrier protein synthase (β-ketoacyl-ACP) domain in module 2 of DEBS1 [24].

We have prepared and kinetically tested a series of diketide analogs in order to establish the structural basis of substrate recognition and the intrinsic limits of molecular discrimination by the KS2 domain. By using a blocked mutant of DEBS1+TE — DEBS1+TE($KS1^0$) — in which the β-ketoacyl-ACP synthase domain of module 1 (KS1) has been inactivated by mutation of the active-site cysteine to alanine [24,25], we could effectively isolate the KS2 domain for study by abolishing competition with the natural biosynthetic route.

## Results and discussion

### Design and synthesis of target diketide analogs

In order to identify specific regions of molecular recognition by the KS2 domain, we systematically examined features of the native diketide structure that we thought would be crucial for recognition and processing by its cognate active site, namely chain length, stereochemistry, and the presence of individual substituents or functional groups. We chose five diketide analogs, **4**–**8** (Figure 2), which fulfilled these requirements. Diketides **4** and **5**

**Figure 2**

Both DEBS1+TE and DEBS1+TE(KS1$^0$), an engineered mutant of DEBS1+TE containing an inactive KS1 domain, can accept the N-acetyl cysteamine (NAC) thioester of the diketide **3** and incorporate it into the triketide lactone **2**. Compounds **4–13** represent the various diketide NAC thioester analogs used in this study and the corresponding triketide products expected from successful processing by the DEBS1+TE(KS1$^0$) enzyme. Dashed arrows indicate lack of conversion of the diketide substrates **7** and **8** to the expected products **12** and **13**.

**Figure 3**

Synthesis of diketides **6** and **8** and triketides **11** and **13**. For experimental details, see the Materials and methods section. Conditions: *i*, LiOH·$H_2O$, $H_2O_2$, THF:$H_2O$ (2:1), room temperature, 1.5 h; *ii*, N-acetylcysteamine, $Et_3N$, $Ph_2P(O)N_3$, DMF, 0°C → room temperature, 24 h; *iii*, lithium diisopropylamide, MeI, THF:hexamethylphosphoramide (4:1), 0°C, 15 min then −60°C → room temperature, 12 h; *iv*, *t*-butyldimethylsilyl trifluoromethanesulfonate, $Et_3N$, $CH_2Cl_2$, 0°C → room temperature, 2.5 h; *v*, diisobutylaluminium, $CH_2Cl_2$, −78°C, 45 min; *vi*, di-*n*-butylboron triflate, *i*-$Pr_2EtN$, (4*R*)-3-propionyl-4-benzyl-2-oxazolidinone, $CH_2Cl_2$, 0°C, 15 min then −78°C, 2 h and 0°C, 1.5 h; *vii*, *t*-butyldimethylsilyl trifluoromethylsulfonate, *i*-$Pr_2EtN$, $CH_2Cl_2$, 0°C → room temperature, 12 h; *viii*, LiOH·$H_2O$, $H_2O_2$, THF:$H_2O$ (3:1) 0°C → room temperature, 12 h; *ix*, 49% aqueous HF, $CH_3CN$:$H_2O$ (3:1), room temperature, 12 h.

retain the stereochemistry of the natural diketide but differ in the length of the carbon chain by one fewer or one additional methylene unit, respectively. Diketides **6** and **7** are analogs of the natural diketide substrate lacking either the 2-methyl group or the 3-hydroxyl substituent. Diketide **8** is a diastereomer of the natural diketide **3** in which the stereochemistry of the C-2 methyl group has been inverted.

Diketide **6** and an authentic sample of the expected product, 4-normethyltriketide lactone **11**, were synthesized from commercially available (-)-methyl (*R*)-3-hydroxyvalerate (**14**; Figure 3). The required sample of **8** was prepared by stereoselective methylation [26] of **14**, to give (2*R*,3*R*)-**15** (Figure 3) in 84% diastereomeric excess over the (2*S*,3*R*)-isomer, followed by routine conversion to the NAC thioester **8**. The corresponding 4-*epi*-triketide lactone **13** (see Figure 3) was prepared from **15** by the chiral oxazolidinone method of Evans *et al.* [27], as described previously [10,23, 28]. All other diketide substrates and the corresponding triketide lactone derivatives were synthesized using methods described previously [10,23,28].

### *In vitro* formation of polyketides catalyzed by DEBS1+TE(KS1$^0$)

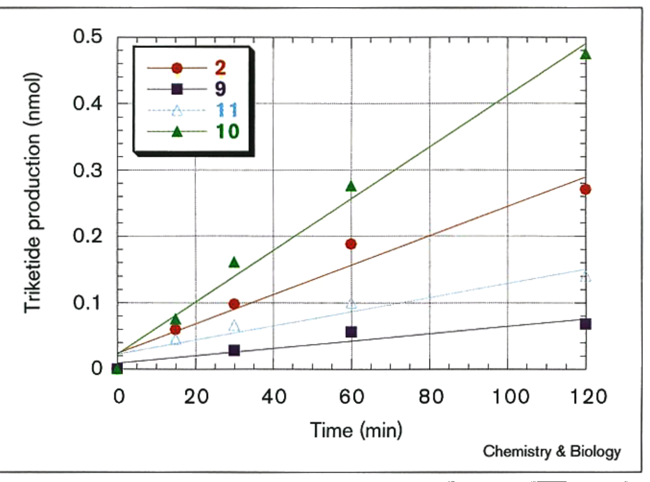
The ability of the DEBS1+TE(KS1$^0$) mutant enzyme to use both natural and aberrant diketide substrates was analyzed *in vitro* using a radiochemical thin layer chromatography (TLC) assay [22]. The conversion of NAC-diketides **4** and **5** to the corresponding $C_8$- and $C_{10}$-triketide lactones has been previously reported [22,23]. Incubation of the 2-normethyl diketide **6** with the KS1$^0$ mutant,

[2-*methyl*-$^{14}$C]methylmalonyl-CoA, and NADPH, gave the corresponding 4-normethyl triketide lactone **11**. The identity of **11** was confirmed by the addition of synthetic unlabeled **11** and recrystallization of the mixture to constant radioactivity. Interestingly, **11** is also biosynthesized by a mutant of DEBS1+TE in which the acyltransferase 1 (AT1) domain has been replaced by a malonyl transferase from the rapamycin PKS [14]. Incubation of the deoxydiketide NAC thioester **7** with partially purified DEBS1+TE(KS1$^0$), [2-*methyl*-$^{14}$C]methylmalonyl-CoA, and NADPH, resulted in the formation of trace amounts of a polar radioactive product detected by TLC/phosphoimaging that was not present in control experiments. Although formation of this compound could be blocked by pre-incubation of DEBS1+TE(KS1$^0$) with cerulenin, a known inhibitor of polyketide biosynthesis, the new band did not exhibit the same chromatographic mobility as an authentic sample of the predicted 5-deoxytriketide product, the hydroxyacid **12**, and was therefore not examined further. The diastereomeric diketide, (2*R*,3*R*)-2-methyl-3-hydroxylpentanoyl NAC thioester (**8**) was incubated with DEBS1+TE(KS1$^0$) and the usual co-substrates under the conditions described above. Gas chromatography/mass spectrometry (GC/MS) analysis using synthetic 4-*epi*-methyl triketide lactone **13** as a standard indicated the complete absence of **13** in the incubation product. (A small amount of **2** was generated as a result of the presence of 8% of the natural (2*S*,3*R*)-**3** in the synthetic sample of **8**.)

Having identified the products of the incubations of the various diketide analogs, we characterized the kinetic behavior of these unnatural substrates. In one series of experiments, fixed amount of diketides **3**, **4**, **5**, **6** and **8** were separately incubated with DEBS1+TE(KS1$^0$) in the presence of saturating concentrations of [2-*methyl*-$^{14}$C] methylmalonyl-CoA and NADPH and the incubation mixtures were assayed at a series of time intervals up to 2 h, during which time the rate of formation of all products remained linear. The results of these experiments are illustrated in Figure 4. Interestingly, the homologous (2*S*,3*R*)-2-methyl-3-hydroxyhexanoyl NAC thioester (**5**) was actually processed at a rate 75% faster than that of the natural diketide intermediate **3**, while the 2-normethyl diketide (**6**) was processed at a rate 50% that of **3** (Table 1). The shorter chain homolog, (2*S*,3*R*)-2-methyl-3-hydroxybutyryl NAC thioester (**4**), was slower still, but was nevertheless turned over at a rate 25% of the natural diketide. Finally, as expected from the results above, the epimeric (2*R*,3*R*)-2-methyl-3-hydroxypentanoyl NAC thioester (**8**) was completely inactive as a substrate.

To quantify more precisely the level of selectivity for each intermediate analog, a series of competition experiments was carried out in which the individual diketides **4**, **5**, and **6** were incubated with DEBS1+TE(KS1$^0$) in the presence of the natural diketide **3** and the usual co-substrates

**Figure 4**




Time-course for triketide synthesis from the various diketide NAC thioester analogs catalyzed by DEBS1+TE(KS1$^0$). Incubation mixtures (100 μl) contained the following concentration of substrates: diketide NAC thioester (5 mM), [2-*methyl*-$^{14}$C]-methylmalonyl-CoA (specific activity 5mCi/mmol, 100 μM) and NADPH (1 mM). A typical incubation contained a DEBS1+TE(KS1$^0$) protein concentration of ~0.1 μM, determined as described previously [21].

(Figure 5). The relative $V_{max}/K_m$ (where $V_{max}$ is maximum rate and $K_m$ is the Michaelis constant) for each analog could be calculated from the relative rates of formation of the corresponding triketide lactone products, after correction for the initial concentration of each substrate (Table 1). Once again, homolog **5** was found to be modestly preferred, with a $V_{max}/K_m$ 1.55 times that of the native diketide, while DEBS1+TE(KS1$^0$) showed a 12-fold preference for **3** over the truncated diketide analog **4** (Figure 5a). Interestingly, earlier experiments showed that DEBS1+TE has a 8:1 preference for the natural starter propionyl-CoA over butyryl-CoA, the precursor of the enzyme-bound diketide intermediate corresponding to **5**, whereas a 35:1 preference in $V_{max}/K_m$ has been measured for propionyl-CoA over acetyl-CoA, the precursor of the enzyme-bound equivalent of **4** [22,23]. In contrast, when **6** was incubated with the KS1$^0$ mutant the absence of the 2-methyl substituent resulted in only a sixfold reduction in $V_{max}/K_m$ compared to the value for the incubation with **3**.

In a third set of experiments, the steady-state kinetic parameters were determined for the two most active substrates of DEBS1+TE(KS1$^0$), **3** and **5** (Table 1). The native diketide **3** had a $K_m$ of 2.7 ± 0.6 mM and a $k_{cat}$ of 0.54 ± 0.07 min$^{-1}$, while the higher homolog **5** had a $K_m$ of 2.2 ± 0.2 mM and a slightly enhanced $k_{cat}$ of 0.85 ± 0.03 min$^{-1}$, consistent with the results of the above-described competition experiments. The $K_m$ for **3** with DEBS1+TE itself was also estimated to be in the millimolar range, although more precise determination of the steady state parameters

**Table 1**

**Steady state kinetic parameters for the turnover of diketide-SNAC analogs into their corresponding triketide lactones.**

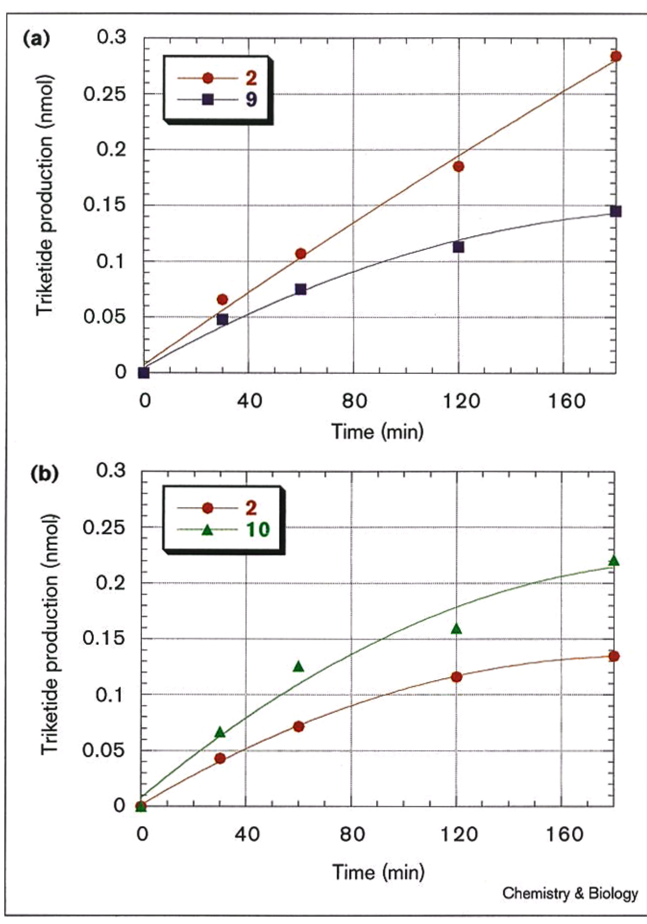
| Diketide | $K_m$ (mM) | $k_{cat}$ ($min^{-1}$) | $V_0$* (pmol/min) | $V_0^{(rel)}$ | $V_{max}/K_m$ (rel)† |
|---|---|---|---|---|---|
| 3 | 2.7 | 0.54 | 2.2 | 100 | 100 |
| 4 | n.d. | n.d. | 0.6 | 25 | 8 |
| 5 | 2.2 | 0.85 | 3.9 | 175 | 155 |
| 6 | n.d. | n.d. | 1.1 | 50 | 15 |

*Initial rates were determined using diketide concentrations of 5 μM (Figure 4); †$V_{max}/K_m$ values were determined from competition assays (Figure 5); n.d., not determined.

for diketide substrates of DEBS1+TE was precluded by competing triketide synthesis as a result of the endogenous generation of the propionyl-CoA starter by decarboxylation of methylmalonyl-CoA. In conjunction with our earlier findings, the current results suggest that an active site or sites other than the loading AT domain are responsible for the starter-substrate selectivity exhibited by DEBS1 and its derivatives. It is also evident that while module 1 of DEBS1 may discriminate against substrates other than propionyl-CoA, this discrimination is attenuated when the resulting intermediates are processed by the KS2 of module 2. Indeed, the butyryl-derived homolog **5** is actually modestly preferred over the natural diketide intermediate, both in terms of a somewhat lower $K_m$ and a slightly higher $k_{cat}$.

The fact that the turnover of the natural diketide **3** to give **2** is ~10-fold slower than the formation of the same triketide lactone from propionyl-CoA catalyzed by DEBS1+TE may reflect a combination of several factors, including the following: the diketide-chain elongation intermediate is normally generated as an enzyme-bound thioester covalently attached to the pantetheinyl residue of ACP1. This intermediate is thought to undergo direct transfer to the active-site cysteine of the KS2 domain. This internal transfer is likely to be more rapid than acylation of the same active-site cysteine by exogenously added diketide thioesters. In addition, N-acetylcysteamine thioesters are often processed less efficiently by fatty acid synthase component enzymes than are the natural ACP or CoA substrates [29]. Finally, and least likely, the DEBS1+TE(KS1$^0$) mutant may generate triketide lactone at an inherently slower rate than does the wild-type DEBS1+TE.

**Figure 5**




Competition of diketide NAC thioesters for DEBS1+TE(KS1$^0$). **(a)** Competition of diketide NAC thioester **4** and 'natural' diketide NAC thioester **3** as substrates for DEBS1+TE(KS1$^0$). Incubation mixtures contained **4** (5 mM) and **3** (0.5 mM), [2-*methyl*-$^{14}$C]methylmalonyl-CoA (specific activity 5 mCi/mmol, 100 μM) and NADPH (1 mM). The concentrations of $C_9$-triketide lactone **2** and $C_8$-triketide lactone **9** synthesized in the same reaction are plotted versus time. Rates were calculated from the linear range of the data. **(b)** Competition of diketide NAC thioester **5** and NAC thioester **3** as substrates for DEBS1+TE(KS1$^0$). Incubation mixtures contained **5** (1 mM) and **3** (0.85 mM), [2-*methyl*-$^{14}$C]-methylmalonyl CoA (specific activity 5 mCi/mmol, 100 μM) and NADPH (1 mM). The concentrations of **2** and $C_{10}$-triketide lactone **10** synthesized in the same reaction are plotted versus time. Rates were calculated from the linear range of the data.

For purposes of comparison, we also measured the $K_m$ of the co-substrate methylmalonyl-CoA using DEBS1+TE(KS1$^0$). In the presence of saturating concentrations of diketide **3**, (2*RS*)-methylmalonyl-CoA was found to have a $K_m$ of 19.1 ± 3.9 μM. Because DEBS has been shown to have a strict specificity for (2*S*)methylmalonyl-CoA [21,30], the actual $K_m$ for the natural 2*S*-enantiomer is calculated to be 9.5 μM, assuming no competition from the inactive (2*R*)-methylmalonyl-CoA. These values are in excellent agreement with the $K_m$ of 24 μM previously determined for (2*RS*)-methylmalonyl-CoA and DEBS1+TE, suggesting

that the KS1$^0$ mutation has not significantly disturbed the interaction of the PKS with the methylmalonyl-CoA substrate. The fact that the diketide-NAC thioesters exhibit $K_m$ values that are two orders of magnitude higher than those for methylmalonyl-CoA (and probably propionyl-CoA) is more than likely to be a reflection of the significantly higher effective molarity of the covalently bound diketide–ACP1 intermediate, as well as a preference for the pantetheinyl thioester over the NAC analog.

The two inactive diketide analogs — **7**, lacking the 3-hydroxy substituent of the natural chain-elongation intermediate, and the 2-*epi*-methyl diketide **8** — were each tested as inhibitors of triketide lactone formation in the presence of the natural diketide **3**. Incubation of DEBS1+TE(KS1$^0$) with diketide **3** in the presence of 10-fold higher concentrations of either **7** or **8**, along with [2-*methyl*-$^{14}$C]methylmalonyl-CoA and NADPH, had no detectable effect on the observed rate of formation of **2**. These results indicate that both the presence of the hydroxyl moiety and the stereochemistry at C-2 of the diketide are essential for molecular recognition of the native diketide intermediate by the KS2 domain.

## Molecular recognition and specificity of the KS2 domain

The data reported here, in combination with the results of earlier studies, have provided a detailed picture of the selectivity of DEBS module 2 for a wide spectrum of diketide substrates. Of the three catalytic domains — KS2, KR2, and the appended TE domain — only KS2, the β-ketoacyl-ACP synthase domain of DEBS module 2, is likely to exercise the degree of structural selectivity that has been observed. Previous experiments have shown that **3**, the NAC-thioester analog of the natural diketide-chain elongation intermediate, covalently acylates the active-site cysteine of the KS2 domain [24]. Were this acylation to be non-selective, with either the downstream KR2 or TE domains exercising the observed selectivity, then we would expect that the two inactive diketide analogs, **7** and **8**, would also acylate the active site of KS2 and therefore act as irreversible inactivators of DEBS1+TE(KS1$^0$). In fact neither **7** nor **8** is an inhibitor of polyketide chain elongation from **3**. Based on the available data, we cannot as yet distinguish whether the observed selectivity is primarily due to KS2-catalyzed self-acylation by the diketide, or to the subsequent condensation step, or to some combination of the two events. In addition, we cannot exclude the possibility that either the KR2 or the TE domain, or both, influence the observed selectivity, either by positively amplifying or negatively modulating the intrinsic specificity of the KS2 domain. The factors described below affect the molecular recognition of diketide substrates by KS2.

### *Chain length*

KS2 will process diketides with a variety of chain lengths. Earlier experiments using acetyl-CoA and butyryl-CoA starting units had established that DEBS can tolerate a range of substrate analogs [20–23]. The behavior of the derived diketides **4** and **5** has now confirmed this observation and indicated that in some cases diketide chain length homologs can actually be processed at a somewhat faster rate than are the natural diketides. We have also recently reported *in vivo* experiments, with a mutant of the complete DEBS harboring a defective KS1 domain, which established the ability of the PKS to process diketide analogs with even longer starter chains (**16**) or carrying a bulky benzyl substituent (**17**; [31]; Figure 6).

**Figure 6**



Additional diketide and triketide substrates of DEBS.

### *Substitution at C-2*

The natural diketide-chain elongation intermediate is substituted at C-2 with a methyl group in the *2S* configuration. In the ketosynthase reaction catalyzed by the KS domains of DEBS, the introduction of the methylmalonyl extender unit is controlled by the corresponding AT domain, which has been shown to have a strict requirement for (2*S*)methylmalonyl-CoA [21,30]. The condensation reaction itself is a decarboxylative acylation with inversion of configuration [32]. Replacement of the native DEBS AT domains with malonyl-AT domains from a variety of heterologous PKS gene clusters [14,33] allows the generation of chain-elongation intermediates and derived polyketides lacking the usual 2*S*-methyl substituent. We have now shown that the analog **6**, in which the 2*S*-methyl group has been replaced by a hydrogen atom, is processed at a significant rate compared to that of the natural intermediate. On the other hand, the diastereomeric diketide **8**, in which the methyl group stereochemistry at C-2 has been inverted, is completely inactive as a substrate for the KS2 domain.

### *C-3 oxygen substituent*

We and others have shown previously that unreduced β-ketoacyl thioesters, generated by inactivation of the corresponding KR domains or by the simple omission of NADPH from incubation mixtures, are viable substrates for downstream condensing (KS) domains [7,22,23]. On the other hand, the fact that the 3-deoxydiketide **7** is inactive as a chain-elongation substrate establishes that an oxygen atom at C-3 is normally required for processing by the KS2 domain. Interestingly, the unsaturated triketide **18** has been shown to be processed by the KS2 domain to

give, ultimately, a 16 membered ring unsaturated octaketide lactone [31]. Thus, in the absence of an oxygen atom at C-3, the KS2 domain may be able to recognize the alternate hydroxyl substituent at C-5.

## Significance

Polyketide synthases (PKSs) are multifunctional proteins that catalyze the biosynthesis of a number of medicinally important natural products. Their modular arrangement allows a combinatorial approach to the synthesis of novel polyketides, but the specificity of PKSs for their natural substrates may limit the molecular diversity of mutabiosynthetic polyketide products. We have analyzed the molecular recognition features of a β-ketoacyl-acyl carrier protein synthase (KS) domain within a modular PKS and identified specific structural features of the natural substrate that are required by its cognate active site. By using a blocked mutant of a bimodular PKS, we were able to explore the intrinsic specificity limits of the KS2 domain as a result of its ability to process unnatural substrate analogs into their corresponding polyketide products. These findings pave the way for the production of novel polyketides through a mutabiosynthetic process that complements the current methods employed in the rational design of new polyketide products. This approach creates a new avenue for obtaining the increased molecular diversity currently required in screening for novel compounds of medicinal importance.

## Materials and methods

### *Materials*

All reactions were run under nitrogen atmosphere using oven-dried syringes and flame-dried glassware when appropriate. THF and $Et_2O$ were distilled from Na/benzophenone ketyl. $CH_2Cl_2$ was distilled from $CaH_2$. Flash column chromatography was performed with EM reagent silica gel 60 (230-240 mesh) and the indicated solvent systems. DL-[2-*methyl*-$^{14}$C]Methylmalonyl-CoA (56.4 mCi/mmol) was obtained from ARC, Inc. Methyl (3R)-3-hydroxy valerate was purchased from Fluka. All other reagents and solvents were obtained from Aldrich.

### *Protein isolation*

*Streptomyces coelicolor* CH999/pCK16, which expresses DEBS1+TE(KS1$^{0}$), a null mutant of the DEBS1+TE in which the KS1 domain has been inactivated by site-directed mutagenesis [25], was used as a source of protein. DEBS1+TE(KS1$^{0}$) was partially purified by ammonium sulfate precipitation and gel filtration [34]. The concentration of DEBS1+TE(KS1$^{0}$) was determined by densitometric scanning of stained SDS-PAGE gels using myosin as a standard. Adobe Photoshop 2.5.1, NIH image 1.5.2, and Molecular Analyst (BioRad) software packages were used for image analysis and quantification. Typically preparations contained 0.2 μM DEBS1+TE(KS1$^{0}$).

### *Assays*

Assays were carried out essentially by the method reported [22]. Unless stated otherwise, the various diketide-SNAC (5 mM) were incubated with (2*RS*)-[2-*methyl*-$^{14}$C]methylmalonyl-CoA (100 μM, 5 mCi/mmol), 1 mM NADPH and DEBS1+TE(KS1$^{0}$) (10 pmol) in a final volume of 100 μl. After incubation for various time periods (15–180 min) at 30°C, the assay mixtures were extracted and products separated by TLC [22]. Quantification of radioactive compounds was carried out using a Molecular Imager GS-363 (BioRad) and authentic [$^{14}$C]-labeled triketide lactone **2** as radioactive standard. Initial velocities were determined by linear regression. Kinetic values were determined by varying diketide concentrations from 0.5–15 mM and [$^{14}$C]-methylmalonyl-CoA concentrations from 1–100 μM. Initial velocities were fitted directly to the Michaelis–Menten equation by non-linear least-squares regression using standard kinetic software packages and used to calculate $K_m$ and $V_{max}$. $V_{max}/K_m$ values were determined from competition assays where diketide-SNAC **3** was incubated in the presence of a second diketide analog. Except where indicated, each diketide was at a concentration of 1 mM. The ratio of the rate of formation of the two triketide lactones being synthesized simultaneously in the one assay gave the relative $V_{max}/K_m$, after adjustment for differences in the two initial diketide substrate concentrations.

### *Product analysis and characterization*

$^{1}$H and $^{13}$C nuclear magnetic resonance (NMR) spectra were recorded on Bruker WM-250 and AM-400 NMR spectrometers. Chemical shifts are reported in parts per million relative to tetramethylsilane. Infra-red (IR) spectra were recorded on a Perkin-Elmer 1600 series Fourier transform IR (FTIR) spectrophotometer. Optical rotations were obtained using a Perkin-Elmer 241 Polarimeter using the Na D line at 25°C and are reported as follows: $[\alpha]_D$ (concentration, *c*, g/100 ml, solvent). Mass spectra were obtained by chemical ionization or electron ionization using a Kratos MS80RFA Mass Spectrometer and are reported as *m/z*. GC/MS was performed on a Hewlett-Packard 5988A mass spectrometer interfaced to an HP 5890 II capillary gas chromatograph. Optical rotations were measured on a Perkin-Elmer 241 polarimeter using the Na D line at 25°C. The formation of triketide lactones **2**, **9**, **10** and **11** in incubations with DEBS1+TE(KS1$^{0}$) was confirmed in each case by recrystallization to constant specific activity with their synthetic counterparts [24]. Diketides **2** and **13** had identical $R_f$ values on TLC and therefore could not be resolved by this method. GC/MS analysis was therefore used to analyze the product of incubation of 2-*epi*-methyl diketide **8** with DEBS1+TE(KS1$^{0}$) and co-substrates. The separation was carried out on an HP cross-linked silicone gum capillary column (12.5 m × 0.22 mm × 0.33 μm) with positive chemical ionization detection (methane) and selected ion monitoring at 173 amu. Separation of the triketide lactone epimers was achieved with an oven temperature gradient of 80°C to 250°C at 25°C/min and a solvent delay of 2 min. Under these conditions triketide lactones **2** and **13** had retention times of 5.67 min and 5.54 min, respectively. Only a small quantity of **2** was detected in the incubation product, formed as a result of the presence of < 8% **3** in the sample of **8**.

### *(3R)-3-Hydroxypentanoic acid*

Methyl (3*R*)-3-hydroxyvalerate (**14**) (105 mg, 0.80 mmol, 1.0 eq.) was dissolved in 4 ml of THF and 2 ml of $H_2O$. $LiOH{\cdot}H_2O$ (67 mg, 1.60 mmol, 2.0 eq.) was added and the reaction mixture was stirred at room temperature for 1.5 h. THF was evaporated, the aqueous solution was acidified with HCl to pH 2, and then extracted with EtOAc. After drying over $Na_2SO_4$, evaporation of EtOAc yielded 93 mg of (3*R*)-3-hydroxypentanoic acid (99%). $^{1}$H NMR ($CDCl_3$): δ 3.95-3.88 (m, 1H, C(3)-H), 2.62-2.55 (m, 1H, one of C(2)-$H_2$), 2.55-2.42 (m, 1H, one of C(2)-$H_2$), 1.65-1.49 (m, 2H,C(4)-$H_2$), 1.01 (t, 3H, C(5)-$H_3$).

### *(3R)-3-Hydroxypentanoic acid N-acetyl cysteamine thioester **6***

To a dry 10 ml three-neck flask was added 60 mg (0.51 mmol, 1.0 eq.) of (3*R*)-3-hydroxypentanoic acid and 5 ml of DMF. The solution was cooled to 0°C and 0.28 ml (2.03 mmol, 4 eq.) triethylamine and 0.33 ml (1.53 mmol, 3.0 eq.) diphenylphosphoryl azide were added. The reaction mixture was stirred at 0 °C for 3 h. N-acetylcysteamine (400 mg, 3.57 mmol, 7 eq.) was added and the reaction mixture was warmed to room temperature and stirred for 24 h. The reaction was then quenched with 3 ml $H_2O$, extracted with EtOAc, and dried overnight ($Na_2SO_4$). After evaporation of the solvent, the crude product was purified by $SiO_2$ column chromatography with 4% MeOH in $CH_2Cl_2$ to afford 47 mg of **6** as a colorless oil (42%). $R_f$ = 0.08 (4% MeOH in $CH_2Cl_2$). $^{1}$H NMR ($CDCl_3$): δ 5.85-5.75 (br, 1H, N-H), 4.05-3.93 (m, 1H, C(3)-H), 3.50-3.43 (m, 2H, N-$CH^2$) 3.05-2.99 (m, 2H, S-$CH_2$), 2.78-2.74 (m, 2H,C(2)-$H_2$), 1.97 (s, 3H, $COCH_3$), 1.58-1.48 (m, 2H, C(4)-$H_2$), 0.97 (t, 3H, C(5)-$H_3$).

*Methyl (3R)-3-(t-butyldimethylsiloxy)-valerate*
A solution of 415 mg (3.14 mmol, 1.0 eq.) methyl-(3*R*)-3-hydroxyvalerate **14** in 15 ml $CH_2Cl_2$ was cooled to 0°C under $N_2$. *t*-Butyldimethylsilyl trifluoromethanesulfonate (TBDMSOTf; 1.08 ml, 4.70 mmol, 1.5 eq.) was added and the mixture was stirred for 10 min, after which 1.10 ml $NEt_3$ (7.85 mmol, 2.5 eq.) was added. The reaction was allowed to warm to room temperature in 2.5 h before being quenched with brine. The aqueous layer was extracted with $CH_2Cl_2$ and the combined organics were washed with pH 11 NaOH and dried over $Na_2SO_4$. The crude product was purified by $SiO_2$ column chromatography with 1:9 EtOAc: hexanes to afford methyl (3*R*)-3-(*t*-butyldimethylsiloxy)-valerate as a colorless oil (696 mg, 90%). $^1$H NMR ($CDCl_3$): δ 4.18-4.05 (m, 1H, C(3)-H), 3.68 (s, 3H, $OCH_3$), 2.42-2.40 (m, 2H, C(2)-$H_2$), 1.58-1.48 (m, 2H, C(4)-$H_2$), 0.91-0.86 (m, 12H, *t*-Bu in TBS and C(5)-$H_3$), 0.05-0.02 (m, 6H, dimethyl in TBS).

*(3R)-3-(t-Butyldimethylsiloxy)-pentanal*
To a flame-dried 25 ml two-neck flask was added 0.307 g (1.35 mmol, 1.0 eq.) of methyl (3*R*)-3-(*t*-butyldimethylsiloxy)-valerate in 6 ml $CH_2Cl_2$. The solution was cooled to −78°C and 1.25 ml of DIBAL-H (1.0 M in $CH_2Cl_2$, 1.0 eq.) was added. The reaction mixture was stirred at −78°C for 1 h, then quenched with 3 ml saturated Rochelle salt solution. The resulting mixture was then warmed to room temperature and stirred for 1 h. The aqueous layer was extracted with $CH_2Cl_2$. The combined organic extracts were dried with $Na_2SO_4$ and the concentrated extract was purified by flash $SiO_2$ column chromatography with 1:9 EtOAc: hexanes to afford 256 mg (89%) of (3*R*)-3-(*t*-butyldimethylsiloxy)-pentanal as a colorless oil. $R_f$ = 0.40 (10% EtOAc/hexanes). $^1$H NMR ($CDCl_3$): δ 9.82-9.80 (m, 1H, C(1)-H), 4.20-4.10 (m, 1H, C(3)-H), 2.52-2.49 (m, 2H, C(2)-$H_2$), 1.60-1.49 (m, 2H, C(4)-$H_2$), 0.92-0.85 (m, 12H, t-Bu in TBS and C(5)-$H_3$), 0.06-0.01 (m, 6H, dimethyl in TBS). IR (neat; $cm^{-1}$): 2956, 2885, 2858, 2720, 1728, 1472, 1464.

*(4R)-3-[(2′R,3′S,5′R)-2′-Methyl-3′-hydroxy-5′-t-butyldimethylsiloxy-heptanoyl]-4-benzyl-2-oxazolidinone*
To a solution of 358 mg (1.54 mmol, 1.75 eq.) of (4*R*)-3-propionyl-4-benzyl-2-oxazolidinone in 12 ml $CH_2Cl_2$ at 0°C under nitrogen was added 1.49 ml di-*n*-butylboron triflate ($Bu_2BOTf$) (1.0 M in $CH_2Cl_2$, 1.7 eq.) followed by the addition of 0.29 ml diisopropylethylamine (1.67 mmol, 1.9 eq.). The solution was stirred at 0°C for 50 min, then cooled to −78°C and 0.190 g (0.88 mmol, 1.0 eq.) of (3*R*)-3-(*t*-butyldimethylsiloxy)-pentanal was added. The resulting solution was stirred for 2 h at −78°C, then for 1.5 h at 0°C before quenching with 12 ml 1M NaOAc in 80% methanol/water. After 5 min, 3 ml of 30% aqueous $H_2O_2$ was added dropwise. After being stirred at ambient temperature for 30 min, the mixture was partitioned between 90 ml water and 120 ml hexanes. The organic layer was washed with aqueous $NaHCO_3$ and brine, dried ($Na_2SO_4$) and concentrated. Purification by flash $SiO_2$ column chromatography with 20% EtOAc in hexanes afforded 0.316 g (80%) of (4*R*)-3-[(2′*R*,3′*S*,5′*R*)-2′-methyl-3′-hydroxy-5′-*t*-butyldimethylsiloxy-heptanoyl]-4-benzyl-2-oxazolidinone as a colorless oil. $^1$H NMR ($CDCl_3$): δ 7.38-7.20 (m, 5H, Ph-H), 4.75-4.65 (m, 1H, N-CH), 4.29-4.18 (m, 2H, O-$CH_2$), 4.18-4.07 (m, 1H, C(3′)-H), 3.95-3.90 (m, 1H, C(2′)-H), 3.82-3.72 (m, 1H, C(5′)-H), 3.32-3.26 (m, 1H, one of Ph-$CH_2$), 2.82-2.73 (m, 1H, one of Ph-$CH_2$) 1.70-1.57 (m, 2H, C(4′)-$H_2$), 1.32-1.20 (m, 5H, C(6′)-$H_2$ and C(2′)-$CH_3$), 0.91-0.85 (m, 12H, *t*-Bu in TBS and C(7′)-$H_3$), 0.09-0.07 (m, 6H, dimethyl in TBS).

*(4R)-3-[(2′R,3′S,5′R)-2′-Methyl-3′,5′-bis(t-butyldimethylsiloxy)-heptanoyl]-4-benzyl-2-oxazolidinone*
To a stirred solution of 0.316 g of (4*R*)-3-[(2′*R*,3′*S*,5′*R*)-2′-methyl-3′-hydroxy-5′-*t*-butyldimethylsiloxy-heptanoyl]-4-benzyl-2-oxazolidinone (0.70 mmol, 1.0 eq.) in 10 ml $CH_2Cl_2$ was added 0.36 ml (1.6 mmol, 2.3 eq.) of TBDMSOTf at 0°C followed by 0.36 ml (2.0 mmol, 2.8 eq.) diisopropylethylamine. The mixture was allowed to warm to room temperature overnight and was then quenched with brine. The separated organic layer was washed with aqueous $NaHCO_3$ and brine, dried ($Na_2SO_4$) and concentrated. Purification by flash $SiO_4$ column chromatography with 10% EtOAc in hexanes afforded 0.342 g (85%) of (4*R*)-3-[(2′*R*,3′*S*,5′*R*)-2′-methyl-3′,5′-bis(*t*-butyldimethylsiloxy)-heptanoyl]-4-benzyl-2-oxazolidinone as white crystals. $^1$H NMR ($CDCl_3$): δ 7.32-7.18 (m, 5H, Ph-H), 4.65-4.52 (m, 1H, N-C-H), 4.18-4.09 (m, 3H, O-$CH_2$ and C(3′)-H), 4.02-3.95 (m, 1H, C(2′)-H), 3.65-3.60 (m, 1H, C(5′)-H), 3.30-3.26 (m, 1H, one of Ph-$CH_2$), 1.67-1.54 (m, 2H, C(4′)-$H_2$), 1.31-1.19 (m, 5H, C(6′)-$H_2$ and C(2′)-$CH_2$), 0.94-0.83 (m, 21H, two *t*-Bu in TBS and C(7′)-$H_3$), 0.10-0.02 (m, 12H, two dimethyl in TBS).

*(2R,3S,5R)-2-Methyl-3,5-bis(t-butyldimethylsiloxy)-heptanoic acid.*
To a stirred solution of 0.342 g (0.61 mmol, 1.0 eq.) (4*R*)-3-[(2′*R*,3′*S*,5′*R*)-2′-methyl-3′,5′-bis(*t*-butyldimethylsiloxy)-heptanoyl]-4-benzyl-2-oxazolidinone in 11.3 ml THF and 3.8 ml $H_2O$ at 0°C was added 0.50 ml (4.89 mmol, 8 eq.) of $H_2O_2$ followed by 51 mg (1.22 mmol, 2 eq.) of LiOH·$H_2O$. The resulting mixture was warmed to room temperature overnight before quenching with 1.7 ml 20% $Na_2SO_3$. THF was stripped off and the aqueous solution was acidified with 1N HCl and extracted with $CH_2Cl_2$. The organic solution was dried ($Na_2SO_4$) and concentrated to afford 0.185 g (76%) of (2*R*,3*S*,5*R*)-2-methyl-3,5-bis(*t*-butyldimethylsiloxy)-heptanoic-acid as a colorless oil. $^1$H NMR ($CDCl_3$): δ 4.22-4.14 (m, 1H, C(3)-H), 3.80-3.70 (m, 1H, C(5)-H), 2.74-2.67 (m, 1H, C(2)-H), 1.80-1.70 (m, 2H, C(4)-$H_2$), 1.59-1.47 (m, 2H, C(6)-$H_2$), 1.20-1.12 (d, 3H, C(2)-$CH_3$), 0.92-0.84 (m, 21H, two *t*-Bu in TBS and C(7)-$H_3$), 0.10-0.02 (m, 12H, two dimethyl in TBS).

*(2R,3S,5R)-2-Methyl-3,5-dihydroxy-n-heptanoic acid δ lactone* ***11***
To a solution of 0.185 g of (2*R*,3*S*,5*R*)-2-methyl-3,5-bis(*t*-butyldimethylsiloxy)-heptanoic-acid in 12 ml $CH_3CN$ and 2.4 ml $H_2O$ was added 1.67 ml of 49% aqueous HF at room temperature. The reaction was stirred overnight before quenching with sat. $NaHCO_3$. The mixture was extracted with $CH_2Cl_2$ and the organic solution was dried ($Na_2SO_4$) and concentrated. Purification by $SiO_2$ column chromatography with 50% EtOAc in hexanes afforded 32 mg (55%) of **11** as white crystals. $R_f$ = 0.22 (50% EtOAc/hexanes), $^1$H NMR ($CDCl_3$): δ 4.20-4.13 (m, 1H, C(5)-H), 3.80-3.71 (m, 1H, C(3)-H), 2.41-2.32 (m, 1H, C(2)-H), 2.25-2.17 (m, 1H, one of C(4)-$H_2$), 1.82-1.59 (m, 3H, C(6)-$H_2$ and one of C(4)-$H_2$), 1.41 (d, J = 6.95 Hz, 3H, C(2)-$CH_3$), 1.00 (t, J = 7.5 Hz, 3H, C(7)-$H_3$). $^{13}$C NMR ($CDCl_3$): δ 173.4 (C-1), 77.7 (C-5), 70.3 (C-3), 45.1 (C-2), 37.7 (C-4), 28.8 (C-6), 13.5 (C(2)-$CH_3$), 9.1 (C-7), $[\alpha]_D$ = +132.0 (c 1.13, $CHCl_3$).

*Methyl (2R,3R)-2-methyl-3-hydroxyvalerate* ***15***
To a flame dried 10 ml two-neck flask containing 0.31 ml (2.23 mmol, 2.9 eq.) of diisopropylamine at 0°C was added with stirring 1.14 ml *n*-butyllithium (1.6 M in hexanes, 2.4 eq.). The reaction mixture was stirred at 0°C for 20 min before cooling to −60°C. After 15 min, 0.097 g (0.73 mmol, 1.0 eq.) methyl (3*R*)-3-hydroxyvalerate **14** was added. The mixture was held at −60°C for 15 min before being warmed to 0°C and stirred for 1 h. $CH_3I$ (0.37 ml, 5.8 mmol, 8.0 eq.) and 0.25 ml HMPA were then added and the reaction was stirred at room temperature overnight. After addition of saturated aqueous $NH_4Cl$, the mixture was extracted with 1:1 hexanes:ethyl ether. The organic extract was dried over $Na_2SO_4$ and purified by flash column $SiO_2$ chromatography with 1:1 hexanes:ethyl ether to afford 48 mg (46%) of methyl (2*R*,3*R*)-2-methyl-3-hydroxyvalerate **15** as a colorless oil, in 84% diastereomeric excess over the (2*S*,3*R*)-diastereomer, as determined by $^1$H NMR on the derived NAC thioester **8**. $^1$H NMR ($CDCl_3$): δ 3.71 (s, 3H, O-$CH_3$), 3.64-3.54 (m, 1H,C(3)-H), 2.61-2.50 (m, 2H, OH and C(2)-H), 1.64-1.53 (m, 1H, one of C(4)-H), 1.50-1.38 (m, 1H, one of C(4)-H), 1.23-1.20 (d, 3H, C(2)-$CH_3$), 0.98-0.95 (m, 3H, C(5)-$H_3$). $^{13}$C NMR ($CDCl_3$): 176.8, 74.7, 52.1, 44.7, 28.8, 14.8, 9.9. IR (neat; $cm^{-1}$): 3447, 2968, 2881, 1744, 1460.

*(2R,3R)-2-Methyl-3-hydroxypentanoic acid*
Methyl (2*R*,3*R*)-2-methyl-3-hydroxyvalerate **15** (54 mg, 0.38 mmol, 1.0 eq.) was dissolved in 3 ml THF and 1.5 ml $H_2O$. LiOH·$H_2O$ (32 mg, 0.76 mmol, 2.0 eq.) was added and the reaction mixture was stirred at room temperature for 1.5 h. The THF was evaporated and the aqueous solution was acidified with 1N HCl to pH 2, then extracted with EtOAc. After drying over $Na_2SO_4$, evaporation of EtOAc yielded 40 mg (92%)

of (2*R*,3*R*)-2-methyl-3-hydroxypentanoic acid. $^1$H NMR ($CDCl_3$): δ 3.68-3.62 (m, 1H, C(3)-H), 2.62-2.44 (m, 1H, C(2)-H), 1.68-1.45 (m, 2H, C(4)-$H_2$), 1.27-1.24, (d, 3H, C(2)-$CH_3$), 1.03-0.97 (t, 3H, C(5)-$H_3$).

### *(2*R*,3*R*)-2-Methyl-3-hydroxypentanoic acid NAC-thioester* 8

To a dry 10 ml three-neck flask was added in 40 mg (0.29 mmol, 1.0 eq.) of (2*R*,3*R*)-2-methyl-3-hydroxypentanoic acid and 5 ml DMF. The solution was cooled to 0°C and 0.16 ml (1.16 mmol, 4 eq.) of triethylamine and 0.19 ml (0.87 mmol, 3 eq.) of diphenylphosphoryl azide were added. The reaction mixture was stirred at 0°C for 3 h. N-acetylcysteamine (0.24 g, 0.20 mmol, 7 eq.) was then added and the reaction mixture was warmed to room temperature and stirred for 24 h. The reaction was then quenched with 3 ml $H_2O$, extracted with EtOAc, and the extract was dried over $Na_2SO_4$. The crude product was purified by $SiO_2$ column chromatography with 4% MeOH in $CH_2Cl_2$ to afford 41 mg (54%) of **8** as a colorless oil. $R_f$ = 0.14 (4% MeOH in $CH_2Cl_2$). $^{25°C}[\alpha]_D$ = −21.0° (c 0.29 $CHCl_3$). $^1$H NMR ($CDCl_3$): δ 5.83 (br, 1H, N-H), 3.69-3.62 (m, 1H, C(3)-H), 3.54-3.38 (m, 2H, N-$CH_2$), 3.07-3.02 (m, 2H, S-$CH_2$), 2.83-2.69 (m, 1H, C(2)-H), 2.10 (br, 1H, O-H), 1.97 (s, 3H, C=O-$CH_3$), 1.66-1.55 (m, 1H, one of C(4)-$H_2$), 1.55-1.38 (m, 1H, one of C(4)-$H_2$), 1.24-1.21 (d, 3H, C(2)-$CH_3$), 1.00-0.94 (t, 3H, C(5)-$H_3$). $^{13}$C NMR $CDCl_3$: δ 204.1 (C-1), 170.3 (NCO), 75.1 (C-3), 53.8 (C-2), 39.4 ($CH_2$N), 28.6 ($CH_2$S), 27.7 $CH_3$CO), 23.2 (C-4), 15.1 (C-2-$CH_3$), 9.7 (C-5). IR (neat; $cm^{-1}$): 3318, 3088, 2975, 2935, 2878, 1652, 1558, 1456, 1375.

### *Methyl (2*R*,3*R*)-2-methyl-3-(*t*-butyldimethylsiloxy)valerate*

A solution of 0.20 g (1.39 mmol, 1.0 eq.) methyl (2*R*,3*R*)-2-methyl-3-hydroxyvalerate **15** in 15 ml $CH_2Cl_2$ was cooled to 0°C under $N^2$ and 0.48 ml of TBDMSOTf (2.09 mmol, 1.5 eq.) was added. The mixture was stirred for 10 min, followed by addition of 0.49 ml $NEt_3$ (3.48 mmol, 2.5 eq.). The reaction was allowed to warm to room temperature over 2.5 h before quenching with brine. The aqueous layer was extracted with $CH_2Cl_2$ and the combined organic extracts were washed with pH 11 aqueous NaOH and dried over $Na_2SO_4$. The crude product was purified by $SiO_2$ column chromatography with 1:9 EtOAc: hexanes to afford 0.356 g (99%) of methyl (2*R*,3*R*)-2-methyl-3-(*t*-butyldimethylsiloxy)valerate as a colorless oil. $^1$H NMR ($CDCl_3$): δ 3.95-3.88 (m, 1H, C(3)-H), 3.66 (s, 3H, O-$CH_3$), 2.67-2.62 (m, 1H, C(2)-H), 1.55-1.44 (m, 2H, C(4)-$H_2$), 1.09-1.06 (d, 3H, C(2)-$CH_3$), 0.92-0.84 (m, 9H, *t*-Bu in TBS and C(5)-$H_3$), 0.07-0.02 (m, 6H, dimethyl in TBS). $^{13}$C NMR $CDCl_3$: δ 175.6, 74.3, 51.4, 44.9, 25.7, 18.0, 12.2, 8.3, -2.9, -4.4, -5.0. IR (neat; $cm^{-1}$): 2960, 2886, 2858, 1744, 1464, ,1382, 1351.

### *(2*R*,3*R*)-2-Methyl-3-(*t*-butyldimethylsiloxy)pentanal*

To a flame-dried 25 ml two-neck flask was added 1.613 g (6.25 mmol, 1.0 eq.) of methyl (2*R*,3*R*)-2-methyl-3-(*t*-butyldimethylsiloxy)valerate in 30 ml $CH_2Cl_2$. The solution was cooled to −78°C and 6.25 ml DIBAL-H (1.0 M in $CH_2Cl_2$, 1.0 eq.) was added. The reaction mixture was stirred at −78°C for 45 min, then quenched with 15 ml saturated Rochelle salt solution. The reaction mixture was then warmed to room temperature and stirred for 1 h. The aqueous layer was extracted with $CH_2Cl_2$ and the combined organic layers were dried over $Na_2SO_4$ and purified by flash $SiO_2$ column chromatography with 5% ethyl ether in hexanes to afford 0.66 g (45%) of (2*R*,3*R*)-2-methyl-3-(*t*-butyldimethylsiloxy)pentanal as a colorless oil. $R_f$ = 0.3 (5% ethyl ether in hexanes). $^{25°C}[\alpha]_D$ = −40.0 (c 0.72, $CHCl_3$): $^1$H NMR ($CDCl_3$) δ 9.76-9.75 (s, 1H, C(1)-H), 3.88-3.84 (m, 1H, C(3)-H), 2.58-2.46 (m, 1H, C(2)-H), 1.63-1.50 (m, 2H, C(4)-$H_2$), 1.08-1.05 (d, 3H, C(2)-$CH_3$), 0.98-0.86 (m, 12H, *t*-Bu in TBS and C(5)-$H_3$), 0.06-0.01 (m, 6H, dimethyl in TBS). IR (neat; $cm^{-1}$): 2984, 2961, 2860, 1728.

### *(4*R*)-3-[(2′*R*,3′*S*,4′*R*,5′*R*)-2′,4′-dimethyl-3′,5′-di(*t*-butyldimethylsiloxy)-heptanoyl]-4-benzyl-2-oxazolidinone*

To a stirred solution of (4*R*)-3-propionyl-4-benzyl-2-oxazolidinone (0.373 g, 1.6 mmol, 1.75 eq.) in dry $CH_2Cl_2$ (9 ml) at 0°C under nitrogen was added 1.55 ml of $Bu_2$BOTf (1.0 M in $CH_2Cl_2$, 1.55 mmol, 1.7 eq.) followed by the addition of 0.35 ml of diisopropylethylamine (1.73 mmol, 1.9 eq.). The solution was kept stirring at 0°C for 50 min. After the solution was cooled to −78°C for 30 min, (2*R*,3*R*)-2-methyl-3-(*t*-butyldimethylsiloxy)-pentanal (209 mg, 0.91 mmol, 1 eq.) was added neat (plus a 1 ml $CH_2Cl_2$ wash). The resulting solution was stirred for 2 h at −78°C and 1.5 h at 0°C and then quenched with 1M NaOAc in 90% methanol/$H_2O$ (8 ml). After 5 min, 1.6 ml of 30% $H_2O_2$ was added dropwise and the mixture was left to stir at ambient temperature for 30 min. The mixture was partitioned between 48 ml $H_2O$ and 64 ml hexanes and the organic layer was washed (aqueous $NaHCO_3$, brine), dried ($Na_2SO_4$) and concentrated. Purification by $SiO_2$ flash chromatography (5→15% EtOAc/hexanes) gave the product, (4*R*)-3-[(2′*R*,3′*S*,4′*R*,5′*R*)-2′-4′-dimethyl-3′-hydroxy-5′-(*t*-butyldimethylsiloxy)-heptanoyl]-4-benzyl-2-oxazolidinone, as a clear oil (280 mg, 94%) plus 62 mg of recovered aldehyde. This material was used immediately in the next protection step.

To a stirred solution of (4*R*)-3-[(2′*R*,3′*S*,4′*R*,5′*R*)-2′-4′-dimethyl-3′-hydroxy-5′-(*t*-butyldimethylsiloxy)-heptanoyl]-4-benzyl-2-oxazolidinone (280 mg, 0.61 mmol, 1 eq.) in dry $CH_2Cl_2$ (8 ml) was added 2,6-lutidine (0.115 ml, 0.98 mmol, 1.6 eq.) at 0°C. TBDMSOTf (0.21 ml, 0.915 mmol, 1.5 eq.) was added in one portion and the reaction mixture was left to stir at ambient temperature for 12 h. The reaction was quenched with brine and the separated organic layer was washed with aqueous $NaHCO_3$, brine, dried ($Na_2SO_4$) and concentrated. Purification by $SiO_2$ flash chromatography (5→10% EtOAc/hexanes) gave the title compound as a colorless oil (260 mg, 74%): $R_f$ = 0.42 (15% EtOAc/hexanes, stains blue with vanillin) $^1$H NMR ($CDCl_3$): δ 7.34-7.21 (m, 5H, Ph-$H_5$), 4.64-4.58 (m, 1H, N-CH), 4.34 (dd, 1H, *J* = 1.4 and 6.24 Hz, C(3′)-H), 4.14-4.08 (m, 2H, $OCH_2$), 4.03-3.99 (m, 1H, C(2′)-H), 3.55-3.51 (m, 1H, C(5′)-H), 3.48-3.44 (dd, 1H, *J* = 3.27 and 13.11 Hz, one of the Ph-$CH_2$), 2.62-2.55 (dd, 1H, *J* = 10.93 and 13.07 Hz, one of Ph-$CH_2$), 1.84-1.80 (m, 1H, C(4′)-H), 1.48-1.42 (m, 2H, C(6′)-$H_2$), 1.16 (d, 3H, *J* = 7.03 Hz, C(2′)-$CH_3$), 0.91-0.80 (m, 24H, 2Si(C($CH_3$)$_3$, C(7′)-$H_3$ and C(4′)-$CH_3$), 0.10-0.0 (ss, 12H, 2Si($CH_3$)$_2$).

### *(2*R*,3*S*,4*R*,5*R*)-2,4-Dimethyl-3,5-dihydroxy-*n*-heptanoic acid S lactone* 13

To a stirring solution of the diprotected imide (158 mg, 0.274 mmol, 1 eq.) in THF (4.6 ml) and $H_2O$ (1.5 ml) at 0°C was added 30% aqueous $H_2O_2$ (0.26 ml, 2.19 mmol, 8 eq.) followed by the addition of LiOH·$H_2O$ (23 mg, 0.548 mmol, 2 eq.). The resulting solution was warmed to room temperature overnight (12–13 h). The reaction was quenched by the addition of 20% aqueous $Na_2SO_3$ (2.4 ml). The THF was removed *in vacuo* and the aqueous layer was acidified with 1N HCl and extracted with $CH_2Cl_2$ (4 × 25 ml). The organic layers were combined, dried ($Na_2SO_4$) and concentrated. The crude acid was purified by $SiO_2$ flash chromatography (20% EtOAc/hexanes) to give the product (105 mg, 92%). This was immediately subjected to deprotection.

The free acid (105 mg, 0.25 mmol, 1 eq.) was dissolved in a mixture of acetonitrile (6.1 ml) and $H_2O$ (2.3 ml). Aqueous HF solution (49%, 1.6 ml) was added dropwise to the stirred solution. The reaction was kept at ambient temperature overnight (12 h) and then quenched by dropwise addition of saturated $NaHCO_3$ solution. The mixture was extracted with $CH_2Cl_2$ (3 × 10 ml). The combined organic layers were dried ($Na_2SO_4$) and concentrated. $SiO_2$ flash chromatography (50% EtOAc/hexanes) afforded a colorless oil (40 mg, 93%): $R_f$ = 0.28 (50% EtOAc/hexanes, stains blue with vanillin). $^1$H NMR ($CDCl_3$): δ 1.05 (t, *J* = 7.4 Hz, 3H, H-7), 1.07 (d, *J* = 7.4 Hz, 3H, C(4)-$CH_3$), 1.27 (d, *J* = 7.4 Hz, 3H, C(2)-$CH_3$), 1.59-1.68 (m, 1H, one of C(6)-$H_2$), 1.76-1.88 (m, 2H, C(4)-H, one of C(6)-$H_2$), 3.74-3.81 (m, 2H, C(5)-H and C(3)-H). $^{13}$C NMR $CDCl_3$: δ 9.4 (C-7), 11.2 (C-4-$CH_3$), 14.2 (C-2-$CH_3$), 26.2 (C-6), 39.0 (C-4), 41.8 (C-2), 75.9 (C-3), 82.6 (C-5), 174.2 (C-1). $^{25°C}[\alpha]_D$ = +82.2 (c 1.19, $CHCl_3$), HRMS (CI, isobutane) $[M+H]^+$ calc'd for $C_9H_{17}O_3$: 173.1177, found: 173.1178.

## Acknowledgements

This work was funded by a grant from the National Institutes of Health (GM22172) to D.E.C. and by a National Science Foundation grant (NSFMCB9417) and an NSF Young Investigator Award, as well as a David and Lucile Packard Fellowship for Science and Engineering to C.K.

## References


1. O'Hagan, D. (1995). *The Polyketide Metabolites*. Ellis Horwood, Chichester, U.K.
2. Wakil, S.J. (1989). Fatty acid synthase, a proficient multifunctional enzyme. *Biochemistry* **28**, 4523-4530.
3. Hopwood, D.A. & Sherman, D.H. (1990). Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu. Rev. Genet.* **24**, 37-66.
4. O'Hagan, D. (1995). Biosynthesis of fatty acid and polyketide metabolites. *Nat. Prod. Rep.* **12**, 1-32.
5. Pieper, R., Kao, C., Khosla, C., Luo, G. & Cane, D.E. (1996). Specificity and versatility in erythromycin biosynthesis. *Chem. Soc. Rev.* **25**, 297-302.
6. Cortes, J., Haydock, S.F., Roberts, G.A., Bevitt, D.J. & Leadlay, P.F. (1990). An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of Saccharopolyspora erythrea. *Nature* **348**, 176-178.
7. Donadio, S., Staver, M.J., McAlpine, J.B., Swanson, S.J. & Katz, L. (1991). Modular organization of genes required for complex polyketide biosynthesis. *Science* **252**, 675-679.
8. Donadio, S., McAlpine, J.B., Sheldon, P.J., Jackson, M. & Katz, L. (1993). An erythromycin analog produced by reprogramming of polyketide synthesis. *Proc. Natl Acad. Sci. USA* **90**, 7119-7123.
9. Bedford, D., Jacobsen, J.R., Luo, G., Cane, D.E. & Khosla, C. (1996). A functional chimeric modular polyketide synthase generated via domain replacement. *Chem. Biol.* **3**, 827-831.
10. Kao, C.M., Luo, G., Katz, L., Cane, D.E. & Khosla, C. (1994). Engineered biosynthesis of a triketide lactone from an incomplete modular polyketide synthase. *J. Am. Chem. Soc.* **116**, 11612-11613.
11. Cortes, J., Weismann, K.E.H., Roberts, G.A., Brown, M.J.B., Staunton, J. & Leadlay, P.F. (1995). Repositioning of a domain in a modular polyketide synthase to promote specific chain cleavage. *Science* **268**, 1487-1489.
12. Kao, C.M., Luo, G., Katz, L., Cane, D.E. & Khosla, C. (1995). Manipulation of macrolide ring size by directed mutagenesis of a modular polyketide synthase. *J. Am. Chem. Soc.* **117**, 9105-9106.
13. Kao, C.M., Luo, G., Katz, L., Cane, D.E. & Khosla, C. (1996). Engineered biosynthesis of structurally diverse tetraketides by a trimodular polyketide synthase. *J. Am. Chem. Soc.* **118**, 9184-9185.
14. Oliynyk, M., Brown, M.J.B., Cortes, J., Staunton, J. & Leadlay, P.F. (1996). A hybrid polyketide synthase obtained by domain swapping. *Chem. Biol.* **3**, 833-839.
15. Kuhstoss S., Huber, M., Turner, J.R., Paschal, J.W. & Rao, R.N. (1996). Production of a novel polyketide through the construction of a hybrid polyketide synthase. *Gene* **183**, 231-236.
16. McDaniel, R., *et al.*, & Khosla, C. (1997). Gain-of-function mutagenesis of a modular polyketide synthase. *J. Am. Chem. Soc.* **119**, 4309-4310.
17. Kao, C.M., McPherson, M., McDaniel, R.N., Fu, H., Cane, D.E. & Khosla, C. (1997). Gain of function mutagenesis of the erythromycin polyketide synthase. 11. Engineered biosynthesis of an eight-membered ring tetraketide lactone. *J. Am. Chem. Soc.*, in press.
18. Hutchinson, C.R. & Fujii, I. (1995). Polyketide synthase gene manipulation: a structure-function approach in engineering novel antibiotics. *Annu. Rev. Microbiol.* **49**, 201-238.
19. Tsoi, C.J. & Khosla, C. (1995). Combinatorial biosynthesis of 'unnatural' natural products: the polyketide example. *Chem. Biol.* **2**, 355-362.
20. Pieper, R., Luo, G., Cane, D.E. & Khosla, C. (1995). Cell-free biosynthesis of polyketides by recombinant erythromycin polyketide synthases. *Nature* **378**, 263-266.
21. Weismann, K.E.H., *et al.*, & Leadlay, P.F. (1995). Polyketide synthesis *in vitro* on a modular polyketide synthase. *Chem. Biol.* **2**, 583-589.
22. Pieper, R., Ebert-Khosla, S., Cane, D. & Khosla, C. (1996). Erythromycin biosynthesis: kinetic studies on a fully active modular polyketide synthase using natural and unnatural substrates. *Biochemistry* **35**, 2054-2060.
23. Pieper, R., Luo, G., Cane, D.E. & Khosla, C. (1995). Remarkable broad substrate specificity of a modular polyketide synthase in a cell-free system. *J. Am. Chem. Soc.* **117**, 11373-11374.
24. Tsukamoto, N., *et al.*, & Cane, D.E. (1996). 6-Deoxyerythronolide B synthase 1 is specifically acylated by a diketide intermediate at the β-ketoacyl-acyl carrier protein synthase domain of module 2. *Biochemistry* **35**, 15244-15248.
25. Kao, C.M., Pieper, R., Cane, D.E. & Khosla, C. (1996). Evidence for two catalytically independent clusters of active sites in a functional modular polyketide synthase. *Biochemistry* **35**, 12363-12368.
26. Fráter, G. (1979). Über die stereospezifität der alkylierung von β-hydroxycarbonsaüreestern. (Translation: The stereospecificity of alkylation of β-hydroxycarboxylic acid esters.) *Helv. Chim. Acta* **62**, 2825-2828.
27. Evans, D.A., Bartroli, J. & Shih, T.L. (1981). Enantioselective aldol condensations. 2. Erythro-selective chiral aldol condensations via boron enolates. *J. Am. Chem. Soc.* **103**, 2127-2129.
28. Cane, D.E., Lambalot, R.H., Prabhakaran, P.C. & Ott, W.R. (1993). Macrolide biosynthesis. 7. Incorporation of chain elongation intermediates into methymycin. *J. Am. Chem. Soc.* **115**, 522-526.
29. Wong, H., Mattick, J.S. & Wakil, S.J. (1983). The architecture of the animal fatty acid synthetase. III. Isolation and characterization of β-ketoacyl reductase. *J. Biol. Chem.* **258**, 15305-15311.
30. Marsden, A.F.A., Caffrey, P., Aparicio, J.F., Loughran, M.S., Staunton, J. & Leadlay, P.F. (1994). Stereospecific acyl transfers on the erythromycin-producing polyketide synthase. *Science* **263**, 378-380.
31. Jacobsen, J.R., Hutchinson, C.R., Cane, D.E. & Khosla, C. (1997). Precursor-directed biosynthesis of erythromycin analogs by an engineered polyketide synthase. *Science* **277**, 367-369.
32. Cane, D.E., Liang, T.-C., Taylor, P.B., Chang, C. & Yang, C.-C. (1986). Macrolide biosynthesis. 3. Stereochemistry of the chain-elongation steps of erythromycin biosynthesis. *J. Am. Chem. Soc.* **108**, 4957-4964.
33. Ruan, X., *et al.*, & Katz, L. (1997). Acyltransferase domain substitutions in the erythromycin polyketide synthase yield novel erythromycin derivatives. *J. Bacteriol.* **179**, in press.
34. Pieper, R., Gokhale, R.S., Luo, G, Cane, D.E. & Khosla, C. (1997). Purification and characterization of bimodular and trimodular derivatives of the erythromycin polyketide synthase. *Biochemistry* **36**, 1846-1851.