

# Mechanism and specificity of the terminal thioesterase domain from the erythromycin polyketide synthase

Rajesh S Gokhale<sup>1</sup>, Daniel Hunziker<sup>1</sup>, David E Cane<sup>2</sup> and Chaitan Khosla<sup>3</sup>

**Background:** Polyketides are important compounds with antibiotic and anticancer activities. Several modular polyketide synthases (PKSs) contain a terminal thioesterase (TE) domain probably responsible for the release and concomitant cyclization of the fully processed polyketide chain. Because the TE domain influences qualitative aspects of product formation by engineered PKSs, its mechanism and specificity are of considerable interest.

**Results:** The TE domain of the 6-deoxyerythronolide B synthase was overexpressed in *Escherichia coli*. When tested against a set of *N*-acetyl cysteamine thioesters the TE domain did not act as a cyclase, but showed significant hydrolytic specificity towards substrates that mimic important features of its natural substrate. Also the overall rate of polyketide chain release was strongly enhanced by a covalent connection between the TE domain and the terminal PKS module (by as much as 100-fold compared with separate TE and PKS 'domains').

**Conclusions:** The inability of the TE domain alone to catalyze cyclization suggests that macrocycle formation results from the combined action of the TE domain and a PKS module. The chain-length and stereochemical preferences of the TE domain might be relevant in the design and engineered biosynthesis of certain novel polyketides. Our results also suggest that the TE domain might loop back to catalyze the release of polyketide chains from both terminal and pre-terminal modules, which may explain the ability of certain naturally occurring PKSs, such as the picromycin synthase, to generate both 12-membered and 14-membered macrolide antibiotics.

## Introduction

Polyketides are a structurally diverse class of natural products that are important as antibiotics, anticancer drugs and other pharmaceutically active agents [1]. Biosynthesis of these compounds is achieved by repetitive condensations of simple carboxylic-acid monomers, a process that closely parallels fatty-acid synthesis. In contrast to the fatty-acid synthases (FAS), polyketide synthases (PKSs) generate an enormous variety of different products by using a broad palette of both primer and extender units and by varying the degree of processing after each condensation step, as well as controlling the stereochemistry of the reduction of the intermediate  $\beta$ -ketoacyl thioesters. Although the random combination of these basic biochemical reactions could potentially generate complex mixtures of numerous polyketide metabolites, naturally occurring PKSs typically produce a single product or a small group of closely related compounds. The possibility of altering PKSs in a controlled manner to obtain rationally novel analogues of known natural products has generated considerable interest. The understanding of the fundamental molecular recognition properties within these complex catalytic systems is a crucial prerequisite towards achievement of this goal.

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

Correspondence: Chaitan Khosla  
E-mail: ck@chemeng.stanford.edu

**Key words:** 6-deoxyerythronolide synthase, hydrolytic specificity, *N*-acetyl cysteamine thioester, polyketide synthase, thioesterase

Received: 9 October 1998

Revisions requested: 4 November 1998

Revisions received: 13 November 1998

Accepted: 7 December 1998

Published: 22 January 1999

**Chemistry & Biology** February 1999, 6:117–125  
<http://biomednet.com/elecref/1074552100600117>

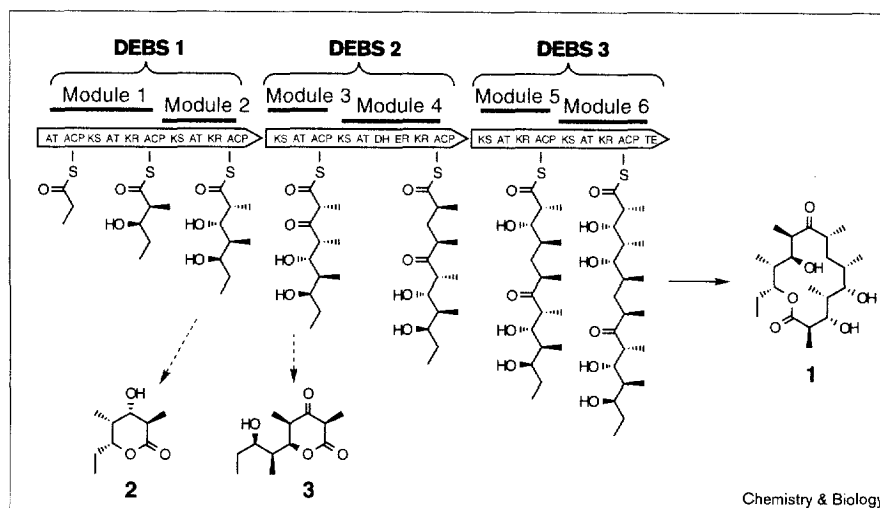
© Elsevier Science Ltd ISSN 1074-5521

6-Deoxyerythronolide B synthase (DEBS) (Figure 1), which catalyzes the formation of 6-deoxyerythronolide B (6-dEB, **1**), the macrolide aglycone of erythromycin, is a hexameric protein complex ( $\alpha_2\beta_2\gamma_2$ ) [2–5]. Each constituent polypeptide chain (DEBS 1, 2 and 3) consists of two modules and functions as a homodimer. A combination of genetic, protein chemical and chemical approaches has provided fundamental new insights into many properties of this remarkable biocatalyst, and has allowed the rational biosynthesis of many novel polyketides (reviewed in [6–8]). Notwithstanding these advances, a good deal about the molecular mechanism of polyketide biosynthesis is still poorly understood.

Here we focus on the properties of the terminal thioesterase (TE) domain of DEBS, believed to be responsible for release and concomitant cyclization of the mature heptaketide chain, resulting in the formation of the macrolactone 6-dEB (**1**). Previous experiments have shown that, when fused to the end of an appropriate module, the TE can release enzymatic intermediates bound to full-length or truncated PKSs [9–13]. By analogy with homologous TE domains from fatty-acid synthases, the mechanism of polyketide chain release by TE is believed to involve acyl transfer of the PKS-bound polyketide chain from the

**Figure 1**

Genetic model of DEBS. DEBS consists of three large polypeptides, each of which contains two 'modules'. Each module catalyzes the elongation of the enzyme-bound polyketide chain by a single methylmalonyl-CoA extender unit, along with the reductive modification steps. The amino terminus of module 1 contains a loading domain for a propionyl CoA starter unit. Each module contains a set of domains including a ketosynthase (KS), acyltransferase (AT) and acyl carrier protein (ACP). All modules except module 3 have active keto reductase (KR) domains and module 4 contains additional dehydratase (DH) and enoyl reductase (ER) domains. Module 6 bears the thioesterase (TE) domain, which is responsible for the cyclization of the heptaketide intermediate to form 6-dEB (1). Truncated DEBS systems generate smaller cyclic products. The triketide lactone 2 arises from an engineered DEBS1 + TE mutant, whereas the tetraketide lactone 3 is formed by DEBS1 and module 3 + TE.

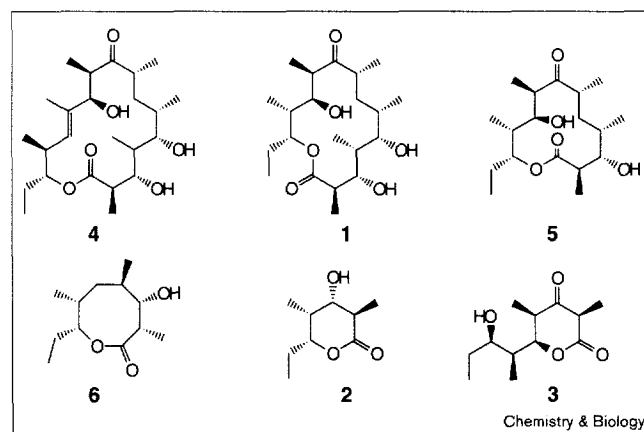


Chemistry &amp; Biology

proximal acyl carrier protein (ACP) domain to an active-site serine residue in the TE. This is followed by an intramolecular nucleophilic attack on this acyl-enzyme intermediate by an appropriate secondary hydroxyl group of the polyketide acyl chain, resulting in lactonization, or alternatively, attack by an external nucleophile, most often water, resulting in net hydrolysis and release of the polyketide product. The three-dimensional structure of myristoyl-ACP-specific TE from *Vibrio harvey*, a member of this enzyme family, has been solved [14]. The tertiary fold of this enzyme closely resembles the consensus fold of the superfamily of  $\alpha/\beta$ -hydrolases, although there is no detectable sequence similarity with any of its members at the amino-acid level. Figure 2 shows some of the lactone structures (1–6), ranging in size from 6–16-membered rings, that have been prepared using such TE-catalyzed reactions. These compounds vividly illustrate the catalytic tolerance of the TE when present as part of a larger PKS while raising several important questions. For example, to what extent is the catalytic capability of the TE influenced by its covalent attachment to the final module of the PKS? Is the TE capable of catalyzing cyclization by itself, or do PKSs form macrolactones as a result of the combined action of the TE and the terminal PKS module? Is acylation or deacylation the rate-limiting step in the TE catalytic cycle? And what features are important in the recognition of potential substrates by the TE domain?

Some insight into the answers to these questions has come from two recent studies [15,16] in which the TE enzyme was expressed as part of a bifunctional ACP6-TE protein and assayed against a number of synthetic substrates. Although the enzyme was able to hydrolyze the *p*-nitrophenyl (pNP)

esters of a variety of carboxylic acids, it was unable to cyclize long-chain substrates possessing a terminal ( $\omega$ ) hydroxyl group. The ability of the enzyme to discriminate between structurally different acyl chains was also probed. Only apparently minor (1.7–4.5-fold, depending on assay conditions) kinetic differences were detected among pNP esters of the four diastereomers of 2-methyl-3-hydroxy-pentanoic acid, of which the (2*S*,3*S*) stereoisomer was judged to

**Figure 2**

Chemistry &amp; Biology

Various cyclic polyketide products synthesized by different chimeric DEBS mutants with thioesterase domain at the carboxyl terminus of the polypeptide chain. Six-membered rings are being formed in the cases of triketide lactone 2 and tetraketide lactone 3. Apart from the natural 14-membered ring product 6-dEB (1), the palette further includes 16-membered (tylosin analog 4), 12-membered (methymycin analog 5) and 8-membered (compound 6) rings. A 10-membered ring compound has not been observed so far.

be the best substrate. Longer-chain pNP esters designed to mimic the length of the natural heptaketide intermediate did not undergo appreciable TE-catalyzed hydrolysis. Here, in addition to presenting new insights into the mechanism and specificity of the TE domain, we compare and contrast our own findings with these earlier reports.

## Results and discussion

### Overexpression of the thioesterase domain as an isolated protein

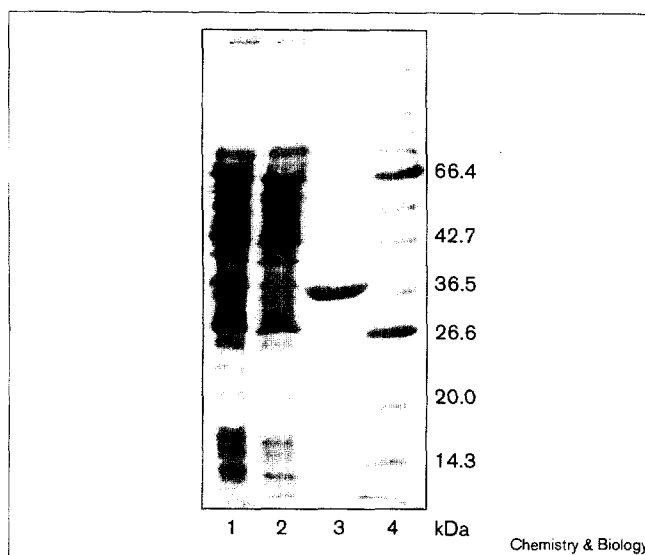
The DEBS TE domain has previously been expressed in *Escherichia coli* as part of an ACP6-TE didomain protein [17]. To facilitate study of the TE as an isolated enzyme, we overexpressed protein corresponding to just the TE domain in *E. coli* using the pET expression system. The natural stop codon at the end of TE domain of DEBS3 was deleted so that the TE could be overproduced as a carboxy-terminal (His)<sub>6</sub>-tagged fusion protein. The protein was purified using a Ni-agarose column (Figure 3). A typical 11 culture produced ~50 mg of purified protein. The recombinant TE domain was stable in solution even at very high concentrations and was active as an independent protein. Its activity was followed a hydrolytic assay using the *N*-acetyl cysteamine (NAC) analog of the natural diketide (**10**; Figure 4) as substrate and Ellman's reagent (5,5'-dithio-2-nitrobenzoic acid, DTNB) to detect the released free thiol by monitoring the progress of the reaction spectrophotometrically at 412 nm. The NAC thioesters are good mimics of the phosphopantetheine arm of the ACP and serve as surrogates for the natural ACP-bound substrate. To decrease the background absorbance, the assays were carried out in the absence of any thiol reducing agents. The TE domain has two free thiols, whose modification by DTNB did not have any detectable effect on its catalytic activity, as judged by independent control experiments (data not shown).

As for vertebrate fatty-acid synthases [18,19], individual modules of DEBS exist as dimers with two independent catalytic centers [20]. Indeed, the ACP6-TE didomain protein is also dimeric [21]. We therefore investigated the quaternary structure of the recombinant TE polypeptide generated as described above using both gel-filtration and cross-linking experiments. Sephadex 200 gel filtration chromatography showed that the TE domain elutes with an apparent molecular weight of 60 kDa (calculated mass 33 kDa), suggesting that the TE domain exists as a dimer. This was also confirmed by glutaraldehyde cross-linking, which exclusively cross-linked the TE domain to form a dimer (data not shown).

### Substrate specificity of the thioesterase

Rates for the hydrolysis of various NAC thioester substrates by TE were proportional to protein concentration, indicating that the observed reaction is indeed enzyme catalyzed. As the TE domain had been used in conjunction

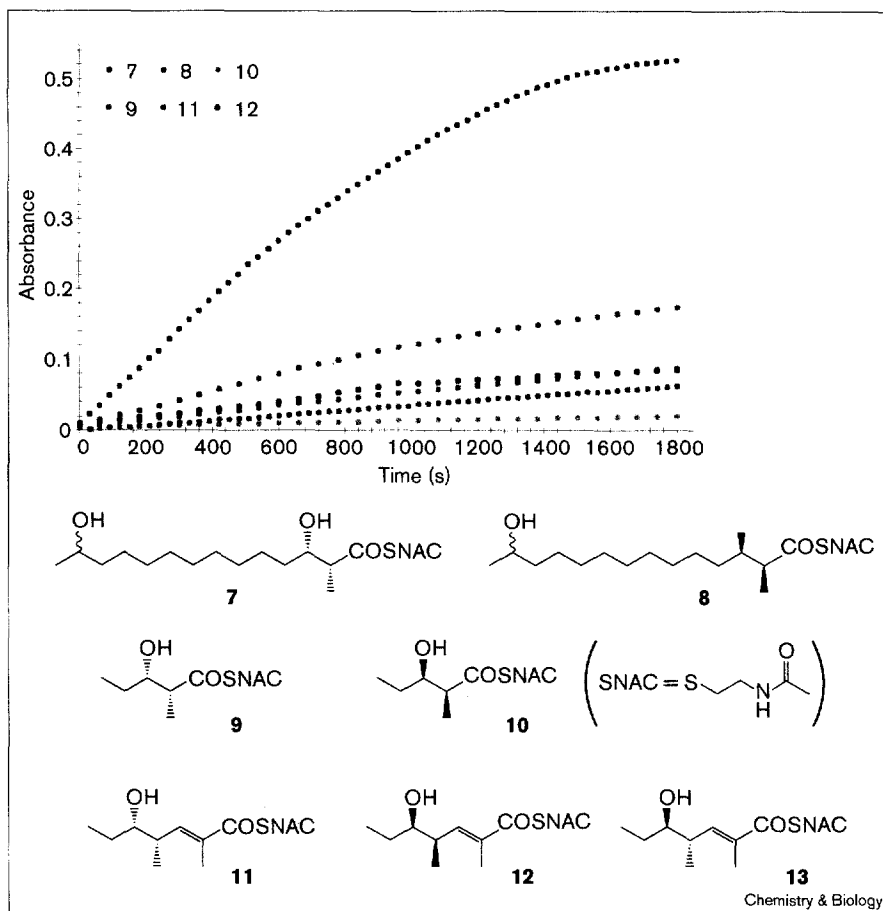
**Figure 3**



Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) of the recombinant TE domain showing purity of protein fractions at different stages of purification. Recombinant His-tagged TE protein was purified from crude extracts using a Ni-agarose column. Protein was eluted with an imidazole gradient. Lane 1, crude extract; lane 2, flow through; lane 3, protein eluted from the Ni-agarose column at 50 mM imidazole; lane 4, molecular weight markers.

with several engineered DEBS modules to generate a variety of different cyclic products (Figure 2), we wished to probe its ability to discriminate between substrates of different chain length and absolute configuration. Compounds **7–13** were chosen for this study. Diketide **10** is the natural diketide synthesized by DEBS module 1 (Figure 1), whereas diketide **9**, its enantiomer, resembles the natural heptaketide substrate of the TE at the  $\alpha$  and  $\beta$  positions with respect to both functionality and stereochemistry. Both of the latter substrates had previously been evaluated as the corresponding *p*-nitrophenolate (pNP) esters in the independent studies of Staunton, Leadlay and coworkers [16]. Thioesters **7** and **8** are the closest mimics of the natural substrate of the TE domain (normally produced by DEBS module 6) that have been evaluated thus far. In addition to possessing the 2-methyl-3-hydroxy functionality, each substrate also carries a secondary hydroxyl group at the C-13 position. (In the natural substrate of the TE, the corresponding hydroxyl group attacks the acyl–enzyme intermediate to form the 14-membered macrolactone 6-dEB). Moreover, because the **9–10** pair closely mimics the **7–8** pair, with respect to substitution pattern and stereochemistry, this subset of substrates was expected to cast light on the ability of the TE to recognize chains of differing lengths. Finally, the rates of hydrolysis of triketides **11**, **12** and **13** were also measured to assess the relative importance of chain length, functionality and stereochemistry.

Figure 4



Comparison of rates of hydrolysis of various substrates catalyzed by recombinant TE domain. For details, see text.

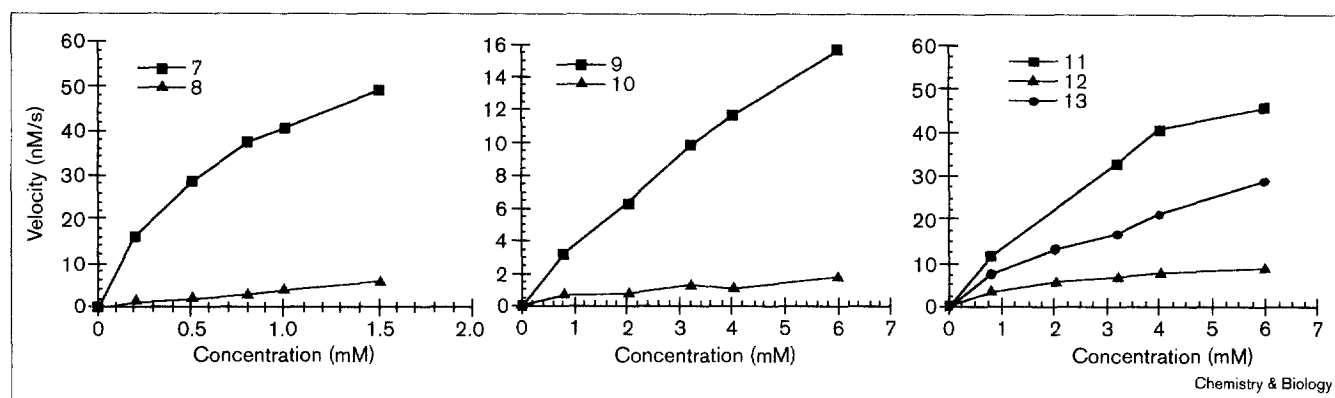
Figure 4 compares the time course of hydrolysis of compounds 7 through 12 assayed at 1 mM substrate concentration and an enzyme concentration of 13.6  $\mu\text{M}$ . (The buffer composition, enzyme concentration and substrate concentration are comparable with those employed previously to study of the corresponding pNP esters [16]). Analysis of the results illustrated in Figure 4 revealed several important trends: NAC thioesters such as 8 and 10 with the (2*S*,3*R*)-configuration are hydrolyzed 13-fold and 4-fold slower than the enantiomeric esters 7 and 9, respectively. This observation suggests that the specificity of the TE domain for substrates with the (2*R*,3*S*)-2-methyl-3-hydroxy substitution pattern and stereochemistry of the natural heptaketide-ACP substrate is significantly greater than originally suggested. Interestingly, when the same stereogenic centers are separated from the critical thioester carbonyl group by an additional *trans* double bond, as in the triketides 11 and 12, the difference in hydrolytic rates was less pronounced (2.2-fold preference for the (4*S*,5*R*)- over the (4*R*,5*S*)-configuration). Finally, comparison of the hydrolytic rates of 9 and 10 with those of their longer-chain analogs 7 and 8, respectively, suggests that the TE domain has a significant preference for longer-chain substrates. Indeed, the

rates of hydrolysis of triketides relative to those of diketides also support this finding. Earlier studies with the myristoyl-ACP-specific TE have also indicated a chain-length dependence for the acyl-ACP cleavage [22].

More quantitative information concerning TE specificity was obtained by direct measurement of the specificity parameter  $k_{\text{cat}}/K_{\text{M}}$  for compounds 7–13. These results are shown in Figure 5, and the calculated rate constants are summarized in Table 1. Again, the importance of chain length is underscored by a comparison of the specificity constants for 7 and 9, which differ by a factor of 24, and for 8 and 10, which differ by a factor of 14. Likewise, preference for the  $\alpha,\beta$  stereochemistry of the natural substrate of the TE domain is clearly illustrated by a 15-fold difference in  $k_{\text{cat}}/K_{\text{M}}$  between 9 and 10, and a 25-fold difference in  $k_{\text{cat}}/K_{\text{M}}$  between 7 and 8.

Due to limits in solubility of the NAC-thioesters 7–13 in aqueous buffer, the absolute values of  $k_{\text{cat}}$  and  $K_{\text{M}}$  could not be measured for every substrate. These values could be estimated for 7 (which is the best substrate among those tested) and 9 (whose solubility was adequately high so as to

Figure 5



Rate versus concentration plots for the hydrolysis of (a) heptaketide analogs **7** and **8**, (b) diketides **9** and **10** and (c) triketides **11**, **12** and **13** by recombinant TE. The scale of rate axis is the same for all compounds. The concentration range for **7** and **8** was 0.2 to 1.5 mM

and 0.8 to 6.0 mM for all other compounds. The  $k_{\text{cat}}/K_M$  for each substrate was calculated from the linear slope of each plot at low substrate concentration.

permit kinetic measurements under substrate-saturating conditions), however. Compound **7** had a  $K_M \sim 0.8$  mM and a  $k_{\text{cat}} \sim 0.7$  min<sup>-1</sup>, whereas the short-chain diketide **9** had a  $K_M \sim 12.4$  mM and a  $k_{\text{cat}} \sim 0.44$  min<sup>-1</sup>. These results suggest that chain length primarily affects  $K_M$ , whereas the catalytic rate for hydrolysis remains largely unaffected. Using a reasonable kinetic model of TE-mediated hydrolysis in which formation of the Michaelis complex is followed by acylation of the active-site serine and subsequent hydrolysis of the covalent acyl-thioester intermediate, it is evident that the acylation step is the most sensitive to the structure and stereochemistry of the substrate, consistent with the relative variations in  $k_{\text{cat}}/K_M$ . When the active site is saturated, the effect on the observed  $k_{\text{cat}}$  is modulated by the rate of deacylation (hydrolysis), which would appear to be less sensitive to substrate structure, consistent with the smaller observed differences in  $k_{\text{cat}}$  among the various NAC-thioester substrates. It is also possible that in the earlier reported studies of TE specificity [16], some of the

differences in substrate selectivity were masked by the use of the intrinsically more reactive pNP esters. Thus for the well-documented and closely mechanistically related case of serine proteases, the use of such reactive leaving groups can mask kinetic differences in the acylation step of enzyme-catalyzed hydrolysis [23,24]. Previous mechanistic studies with the TE domain of chicken fatty-acid synthases have also suggested that the formation of the tetrahedral adduct leading to the acyl-enzyme intermediate is the rate-limiting step in the hydrolysis of acyl CoAs [25].

#### Can the thioesterase domain alone act as a cyclase?

In order to investigate the ability of the TE domain to act as a cyclase, preparative incubations were carried out with substrates **7** and **8**. Because previous studies had failed to observe TE-catalyzed lactonization of substrates containing primary ( $\omega$ ) hydroxyl groups, we chose to focus instead on substrates possessing secondary hydroxyl groups at C13. Notably, the known macrolactone product, **8**, 8a-deoxyoleandolide, which is synthesized from an acetate primer unit by DEBS, arises from an analogous substrate carrying such an ( $\omega-1$ )-hydroxyl function [26].

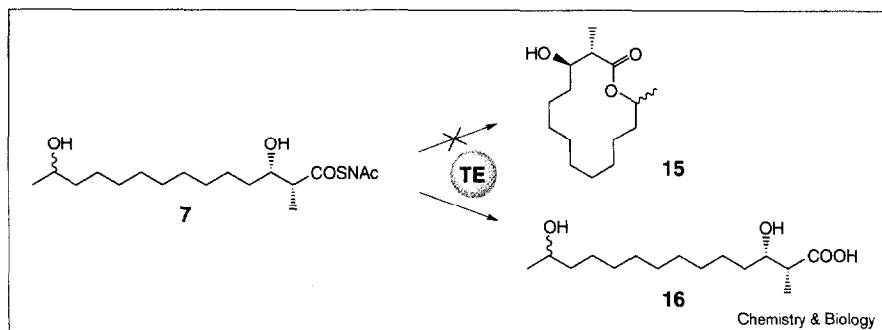
Following incubation of TE with **7** and **8** for 15 h, reaction mixtures were extracted and product formation was analyzed using liquid-chromatography-mass spectrometry (LC-MS). No cyclic product could be detected, however, as judged by mass spectrometric analysis of the extracts, as well as comparison of the retention times with a synthetic cyclic standard **15** (Figure 6), prepared as described in the Materials and methods section. Thin layer chromatography (TLC), as well as <sup>1</sup>H-nuclear magnetic resonance (NMR) analysis of the extract, indicated complete hydrolysis of both **7** and **8** within 15 h as judged by comparison with synthetic standard **16** (Figure 6). The isolated TE domain

Table 1

#### Kinetic parameters of TE-domain-catalyzed thioester hydrolysis.

Compound	$k_{\text{cat}}/K_M$	$K_M$ (mM)	$k_{\text{cat}}$ (min <sup>-1</sup> )
<b>7</b>	14.34	0.8	0.70
<b>8</b>	0.57	–	–
<b>9</b>	0.60	12.2	0.44
<b>10</b>	0.04	–	–
<b>11</b>	2.76	–	–
<b>12</b>	0.80	–	–
<b>13</b>	0.70	–	–

Figure 6



Investigation of cyclase activity of the TE domain. Incubation of heptaketide analog **7** with TE led exclusively to formation of the free acid **16** *in vitro*. No cyclic products, such as teradecanotide **15**, were observed as judged by comparison to authentic standards.

therefore appears to be incapable of acting as a cyclase in order to mediate macrolactone formation. It is conceivable that the presentation of a substrate to the TE as an ACP-bound chain might be necessary for cyclization. Alternatively, effective macrolactonization of the long-chain ( $\omega$ -1)-hydroxyacyl thioester might require additional internal conformational constraints intrinsic to the natural substitution pattern and stereochemistry of the native seco-ester substrate.

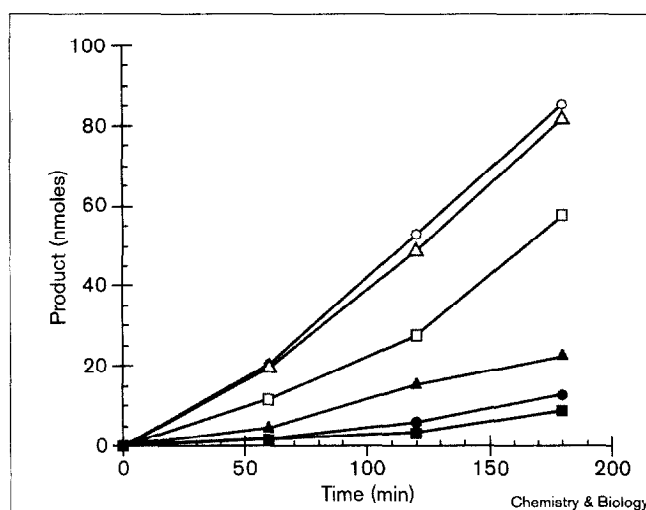
#### Intramolecular and intermolecular interactions of the thioesterase with PKS modules

Earlier studies had shown that although expression of DEBS1 alone in *Streptomyces coelicolor* led to biosynthesis of small quantities of the expected triketide lactone **2**, fusion of the TE domain to the carboxy-terminal end of DEBS1 significantly enhanced the level of production of **2** [9,11,26]. These results implicated the TE domain as an active participant in polyketide chain release. To assess the importance of the covalent connection between DEBS1 and the TE domain, DEBS1 and DEBS1 + TE were purified to homogeneity from *S. coelicolor* CH999/pCK9 [26] and CH999/pCK12 [10], respectively, using protocols described previously [27]. Incubation of DEBS1 with the substrates, [ $^{14}$ C]-propionyl CoA, (2-*RS*) methylmalonyl CoA and NADPH, did not produce detectable quantities of **2**. Under the same incubation conditions, DEBS1 + TE catalyzed triketide lactone synthesis with an apparent  $k_{\text{cat}}$  of  $5.4 \text{ min}^{-1}$  [27]. The ability of recombinant TE polypeptide to interact with DEBS1 was also tested. At a concentration ratio of TE:DEBS1 of 20:1 the apparent  $k_{\text{cat}}$  for the formation of **2** was  $0.058 \text{ min}^{-1}$ . This result points to the virtual absence of any specific protein-protein interactions between module 2 and TE. This structural independence and integrity of the TE domain bodes well for the future genetic engineering of novel polyketides, as this domain could be used as a general chain terminator in various polyketide synthases. Interestingly, recent studies on the nonribosomal peptide synthetase responsible for surfactin biosynthesis have shown that disconnection of the corresponding TE domain from

the rest of the peptide synthetase also results in loss of product formation [28].

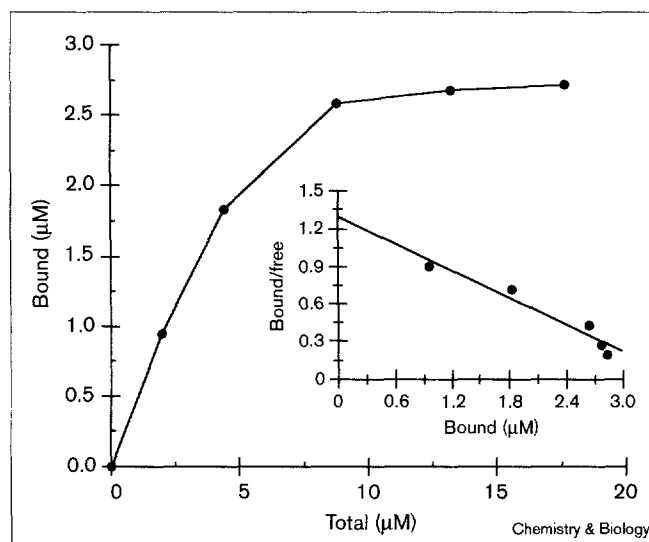
In certain naturally occurring PKSs, such as the PKS responsible for biosynthesis of picromycin (a 14-membered macrolide) and methymycin (a 12-membered macrolide) in *Streptomyces venezuelae*, a single TE domain present the end of module 6 may catalyze the release of both the heptaketide chain from module 6 as well as the hexaketide chain from module 5 (R. McDaniel, Kosan Biosciences, personal communication). Recently we made a similar observation during the course of *in vitro* studies using truncated DEBS systems [27]. In particular, it was observed that when purified DEBS1 and module 3 + TE were reconstituted, not only did they synthesize the expected tetraketide lactone **3** (Figure 1), but the triketide

Figure 7



Time course for the synthesis of polyketide products **2** and **3** at varying ratios of DEBS1 and module 3 + TE concentrations. Open symbols represent tetraketide lactone **3**, whereas filled symbols correspond to triketide lactone **2**. Molar ratios of DEBS 1 and module 3 + TE are as follows: ■, 1:1, **2**; □, 1:1, **3**; ●, 1:2, **2**; ○, 1:2, **3**; ▲, 1:4, **2**; △, 1:4, **3**.

Figure 8



Analysis of interaction between DEBS1 and module 3 + TE. A typical saturation curve is observed for the interaction of these two proteins. Linear transformation of the binding data were carried out by using Scatchard plot analysis, as shown in the inset.

lactone **2** (Figure 1) was produced as well. As DEBS1 alone is unable to synthesize detectable quantities of **2** (see above), we investigated the mechanism of triketide lactone formation by the mixture of DEBS1 and module 3 + TE. Specifically, we wished to discriminate between two alternative hypotheses: whether the TE domain of module 3 + TE might associate nonspecifically with DEBS1 to catalyze triketide release, or upon specific noncovalent association of module 3 + TE with DEBS1, the TE domain could loop back to interact with DEBS1.

DEBS1 was titrated with increasing amounts of module 3 + TE, and the interaction between the two proteins was studied by monitoring the rates of synthesis of both polyketides **2** and **3** in the presence of saturating concentrations of [ $^{14}\text{C}$ ]-propionyl CoA, (*2-R/S*) methylmalonyl CoA, and NADPH. At a fixed concentration of DEBS1, increasing the concentration of module 3 + TE resulted in the saturation of the rate of tetraketide lactone synthesis (Figure 7). The  $K_D$  for association between DEBS1 and module 3 + TE was estimated to be 2.6  $\mu\text{M}$  by Scatchard plot analysis (Figure 8). Although the rate of triketide lactone formation was relatively small compared with that for tetraketide lactone formation (Figure 7), the maximum observed rate of triketide lactone synthesis by DEBS1 and module 3 + TE was 10-fold higher than the maximum rate of 0.058  $\text{min}^{-1}$  for release of the triketide from DEBS1 catalyzed by recombinant TE alone. We therefore suggest that the TE domain has sufficient conformational flexibility so as to be able to catalyze chain release not only from terminal modules but also from certain preterminal modules

of multimodular PKSs (the second hypothesis above). The fact that the TE domain prefers (*2R,3S*) substrates might further enhance the ability of this trimodular PKS to synthesize **2** in addition to **3**. Perhaps this might also suggest how the TE domain at the terminus of module 6 of the picromycin PKS could be capable of releasing methymycin from module 5 as well.

## Significance

Modular polyketide synthases (PKSs), such as 6-deoxyerythronolide synthase (DEBS), catalyze the biosynthesis of structurally complex natural products by repetitive condensation of simple carboxylic acid monomers. Terminal thioesterase (TE) domains of modular polyketide synthases play a key role in the release and concomitant cyclization of enzyme-bound intermediates during the biosynthesis of a number of important macrolide antibiotics. Overexpression of the TE domain of the 6-deoxyerythronolide B synthase has allowed us to systematically investigate the properties of this enzyme. Although the isolated protein is incapable of acting as a cyclase, it shows hydrolytic specificity towards substrates that mimic important features of its natural substrate, including the preference for a long-chain acyl group and a (*2R,3S*) configuration for its 2-methyl-3-hydroxy substituents. Although the observed selectivity is relatively modest, in some cases it might be sufficient to alter the qualitative and/or quantitative aspects of polyketide production by recombinant modular PKSs.

## Materials and methods

### Reagents, chemicals and synthetic compounds

DL-2-[Methyl- $^{14}\text{C}$ ]-methylmalonyl-Coenzyme A (CoA) (56.4 mCi/mmol) was purchased from ARC Corporation (St. Louis, MO). [ $^{1-14}\text{C}$ ]-Propionyl-CoA (54 mCi/mmol) was obtained from Moravak Biochemicals (Brea, CA.) Reagents for assays and for chemical synthesis were obtained from Aldrich in the highest purity available and used without further purification. Synthesis of (*2R,3S,13R/S*)-3,13-dihydroxy-2-methylmyristic acid *N*-acetylcysteamine thioester (**7**), as well as of (*2S,3R,13R/S*)-3,13-dihydroxy-2-methylmyristic acid *N*-acetylcysteamine thioester (**8**) will be reported elsewhere. Synthesis of diketides **9** and **10** as well as of triketides **11**, **12** and **13** has been reported previously [13]. Synthesis of a diastereomeric mixture of 2,13-dimethyl-3-hydroxytetradecanolides **15** as a cyclic reference material for cyclase activity studies of TE was carried out by reduction of (*2R/S,13S*)-2-methyl-3-oxo-tetradecanolide as described previously [29]. (*2S,3R,13R/S*)-3,13-Dihydroxy-2-methylmyristic acid (**16**) was obtained for reference purposes from the corresponding *bis*-TBDMS protected derivative, which is an intermediate in the synthesis of **7**:  $R_f$ : 0.43 (ethyl acetate containing 1% acetic acid);  $^1\text{H-NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  3.96–3.91 (m, 1H); 3.83–3.79 (m, 1H), 2.59 (qd, 1H,  $J_1 = 7.2$ ,  $J_2 = 3.6$ ), 1.54–1.23 (m, 18H), 1.21 (d, 3H,  $J = 7.2$ ), 1.19 (d, 3H,  $J = 5.8$ ).

### Preparative assays for TE cyclase activity

Preparative *in vitro* reactions were set up in a total volume of 500  $\mu\text{l}$  containing 50 mM Tris buffer at pH = 7.4, a total of 15  $\mu\text{l}$  dimethylsulfoxide (DMSO) and 1 mM substrate. 50  $\mu\text{l}$  of a solution of TE was added. Reactions were incubated at 37°C for 20 h. Note that DTNB (Ellman's reagent) was omitted in these reactions, but the enzyme was stabilized with dithiothreitol (DTT). Following incubation, the mixture was extracted 3 $\times$  with an equal volume of ethyl acetate containing 1% acetic acid and the combined organic extracts were concentrated and

dried *in vacuo*. The residue was dissolved in methanol and subjected to TLC and LC-MS analysis in order to detect trace amounts of products of interest. Mass spectra were compared with data for a synthetic reference sample, obtained as described above.

#### Determination of kinetic parameters

Hydrolysis of NAC-thioester substrates was followed spectrophotometrically by observation of formation of 3-thio-5-nitrobenzoate at 412 nm ( $\epsilon = 13,600 \text{ M}^{-1} \text{ cm}^{-1}$ ). Each assay contained (in a total volume of 500  $\mu\text{l}$ ) 50 mM phosphate buffer, pH = 8.0 or 50 mM Tris buffer, pH = 8.4, a variable amount of a stock solution of the appropriate substrate, 6.45  $\mu\text{M}$  TE, 5  $\mu\text{l}$  of a saturated solution of DNTB in 50 mM Tris buffer at pH = 8.4. Stock solutions of compounds **7** and **8** were 50 mM in DMSO, whereas stocks of the other compounds were 200 mM in DMSO. In all cases, the total content of DMSO in the reaction was adjusted to 15  $\mu\text{l}$ . The reactions were initiated by brief mixing and the rate of free thiol formation was observed continuously with an interval time of 15 s. Rates of hydrolysis were calculated from the initial linear portion of the curves. In all cases, rates were corrected for the background rate of chemical hydrolysis in the absence of enzyme. The concentration range analyzed was 0.2 to 1.5 mM for **3** and **4** (concentrations above 1.5 mM are not accessible due to insufficient solubility) and 0.8 to 6 mM for diketides and triketides. In order to remove any residual free thiols in the triketide preparations, compounds **13**, **14** and **15** were filtered through a short silica gel column impregnated with copper sulfate prior to analysis with TE.

#### Overexpression of the thioesterase domain

The TE domain was cloned using a previously engineered *SpeI* restriction site [26] and the natural *EcoRI* restriction site into the expression vector pET21c by replacing the *NheI/EcoRI* fragment. The resultant plasmid pRSG33 was used to transform *E. coli* BL21 (DE3) for expression of the TE gene. 1 l cultures of *E. coli* BL21 (DE3)/pRSG33 were grown in 2 l flasks using LB medium. Protein was induced at  $A_{595} = 0.6$  by adding 1 mM of IPTG. Cells were harvested after 14 h, centrifuged at 2500 g and were resuspended in 20 ml of 100 mM Tris pH 8.4. Cells were crushed using a French press at 13,000 psi. The supernatant was collected by centrifugation at 10,000 g and was diluted to 50 ml and loaded onto the previously equilibrated Ni-agarose column (3 cm, Qiagen). The column was washed with 10 mM imidazole in 100 mM Tris pH 8.4; the TE eluted at 20 mM to 50 mM imidazole. Imidazole was removed by exchanging buffer using Centripeg 10 (Amicon). A typical batch of 1 l produced 50 mg of protein.

#### Purification of the DEBS enzymes

Homogenous preparations of DEBS were obtained by using the reported purification protocol [30]. DEBS1 and module 3-TE could be partially separated from each other on a gel filtration column. These were then further purified on a Resource Q column. Module 3-TE eluted at 215 mM NaCl, whereas DEBS 1 eluted at 260 mM NaCl.

#### DEBS assay conditions

A typical assay mixture contained 150–250  $\mu\text{g/ml}$  of protein and saturating concentrations of each substrate (propionyl-CoA, methylmalonyl-CoA, NADPH) as described earlier [27]. Polyketide products were extracted with ethyl acetate and separated using TLC. Quantitative measurements were performed on a Packard InstantImager<sup>®</sup>.

#### Measurement of kinetic constants

Time courses for the formations of triketide and tetraketide lactones were performed at varying module 3-TE concentrations for a fixed concentration of DEBS1. The reactions were quenched by adding ethyl acetate to the reaction mixture. The products were extracted in ethyl acetate and were separated on TLC plates. The amount of <sup>14</sup>C-labeled were calculated from the radioactivity incorporated into the relevant polyketides using Packard InstantImager. Protein concentrations were measured by the Bradford method [31]. The association constant between DEBS 1 and module 3-TE was estimated by using Scatchard analysis.

## Acknowledgements

This work was supported by grants from the National Institutes of Health (CA 66736 to C.K. and GM22172 to D.E.C.) and by a David and Lucile Packard Fellowship in Science and Engineering to C.K. D.H. is a recipient of a Swiss National Science Foundation Postdoctoral Fellowship. We thank Larry Weiler, (University of British Columbia, Vancouver, Canada) for a generous gift of (2R/S, 13S)-2-methyl-3-oxo-tetradecanolid.

## References

- Hopwood, D.A. (1997). Genetic contributions to understanding polyketide synthases. *Chem. Rev.* **97**, 2465-2497.
- 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 erythraea*. *Nature* **348**, 176-178.
- 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.
- Caffrey, P., Bevitt, D.J., Staunton, J. & Leadlay, P.F. (1992). Identification of DEBS1, DEBS2 and DEBS3, the multienzyme polypeptides of the erythromycin-producing polyketide synthase from *Saccharopolyspora erythraea*. *FEBS Lett.* **304**, 225-228.
- Pieper, R., Luo, G., Cane, D.E. & Khosla, C. (1995). Cell-free synthesis of polyketides by recombinant erythromycin polyketide synthase. *Nature* **378**, 263-266.
- Katz, L. (1997). Manipulation of modular polyketide synthases. *Chem. Rev.* **97**, 2557-2575.
- Khosla, C. (1997). Harnessing the biosynthetic potential of modular polyketide synthases. *Chem. Rev.* **97**, 2577-2590.
- Staunton, J. & Wilkinson, B. (1997). Biosynthesis of erythromycin and rapamycin. *Chem. Rev.* **97**, 2611-2629.
- Cortes, J., Wiesmann, K.E.H., Roberts, G.A., Brown, M.J., Staunton, J. & Leadlay, P.F. (1995). Repositioning of a domain in a modular polyketide synthase to promote specific chain cleavage. *Science* **268**, 1487-1489.
- 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.
- 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.
- 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. 2. engineered biosynthesis of an eight-membered ring tetraketide lactone. *J. Am. Chem. Soc.* **119**, 11339-11340.
- 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.
- Lawson, D.M., et al., & Derewenda, Z.S. (1994). Structure of a myristoyl-ACP-specific thioesterase from *Vibrio harveyi*. *Biochemistry* **33**, 9382-9388.
- Aggarwal, R., Caffrey, P., Leadlay, P.F., Smith, C.J. & Staunton, J. (1995). The thioesterase of the erythromycin-producing polyketide synthase: mechanistic studies *in vitro* to investigate its mode of action and substrate specificity. *J. Chem. Soc. Chem. Commun.* 1519-1520.
- Weissman, K.J., et al., & Leadlay, P.F. (1998). The thioesterase of the erythromycin-producing polyketide synthase: influence of acyl chain structure on the mode of release of substrate analogues from the acyl enzyme intermediates. *Angew. Chem. Int. Ed.* **37**, 1437-1440.
- Caffrey, P., Green, B., Packman, L.C., Rawlings, B., Staunton, J. & Leadlay, P.F. (1991). An acyl carrier protein-thioesterase domain from the 6-deoxyerythronolide B synthase of *Saccharopolyspora erythraea* - high level production purification and characterization in *E. coli*. *Eur. J. Biochem.* **195**, 823-830.
- Wakil, S.J. (1989). Fatty acid synthase - a proficient multifunctional enzyme. *Biochemistry* **28**, 4523-4530.
- Joshi, A.K. & Smith, S. (1995). Construction, expression and characterization of a mutated animal fatty acid synthase deficient in the dehydratase function. *J. Biol. Chem.* **268**, 22508-22513.
- Kao, C.M., Pieper, R., Cane, D.E. & Khosla, C. (1996). Evidence for two catalytically active clusters of active sites in a functional modular polyketide synthase. *J. Am. Chem. Soc.* **118**, 12363-12368.
- Staunton, J., Caffrey, P., Aparicio, J.F., Roberts, G.A., Bethell, S.S. & Leadlay, P.F. (1996). Evidence for a double-helical structure for modular polyketide synthases. *Nat. Struct. Biol.* **3**, 188-192.
- Ferri, S.R. & Meighen, E.A. (1991). A lux-specific myristoyl transferase in luminescent bacteria related to eukaryotic serine esterases. *J. Biol. Chem.* **266**, 12852-12857.



23. Bender, M.L. & Kezdy, F.J. (1965). Mechanism and action of proteolytic enzymes. *Ann. Rev. Biochem.* **34**, 49-76.
24. Walsh, C. (1979). *Enzymatic Reaction Mechanisms*. W.H. Freeman & Co., New York.
25. Pazirandeh, M., Chirala, S.S., Wakil, S.J. (1991). Site-directed mutagenesis studies on the recombinant thioesterase domain of chicken fatty acid synthase expressed in *Escherichia coli*. *J. Biol. Chem.* **266**, 20946-20952.
26. 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.
27. 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.
28. Guenzi, E., Galli, G., Grgurina, I., Pace, E., Ferranti, P. & Grandi, G. (1998). Coordinate transcription and physical linkage of domains in surfactant synthetase are not essential for proper assembly and activity of the multienzyme complex. *J. Biol. Chem.* **273**, 14403-14410.
29. Neeland, G., Ounsworth, J.P., Sims, R.J. & Weiler, L. (1994). Synthesis, conformational analysis, and stereoselective reduction of 14-membered ring 3-keto lactones. *J. Org. Chem.* **59**, 7383-7394.
30. Gokhale, R.S., Lau, J., Cane, D.E. & Khosla, C. (1998). Functional orientation of the acyltransferase domain in a module of the erythromycin polyketide synthase. *Biochemistry* **37**, 2524-2528.
31. Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254.

---

**Because Chemistry & Biology operates a 'Continuous Publication System' for Research Papers, this paper has been published via the internet before being printed. The paper can be accessed from <http://biomednet.com/cbiology/cmb> – for further information, see the explanation on the contents pages.**