The Thioesterase of the Erythromycin-producing Polyketide Synthase: Mechanistic Studies *in vitro* to Investigate its Mode of Action and Substrate Specificity

Ranjana Aggarwal,[a] Patrick Caffrey,[b] Peter F. Leadlay,[b] Cameron J. Smith[a] and James Staunton*[a]

[a] *Centre for Molecular Recognition, Department of Chemistry, Lensfield Road, Cambridge, UK CB2 1EW*
[b] *Centre for Molecular Recognition, Department of Biochemistry, Tennis Court Road, Cambridge, UK CB2 1QW*

In vitro studies of the mode of action of the thioesterase of the erythromycin-producing PKS have shown that it cleaves ester groups by forming acyl enzyme intermediates, and that it has a broad substrate specificity.

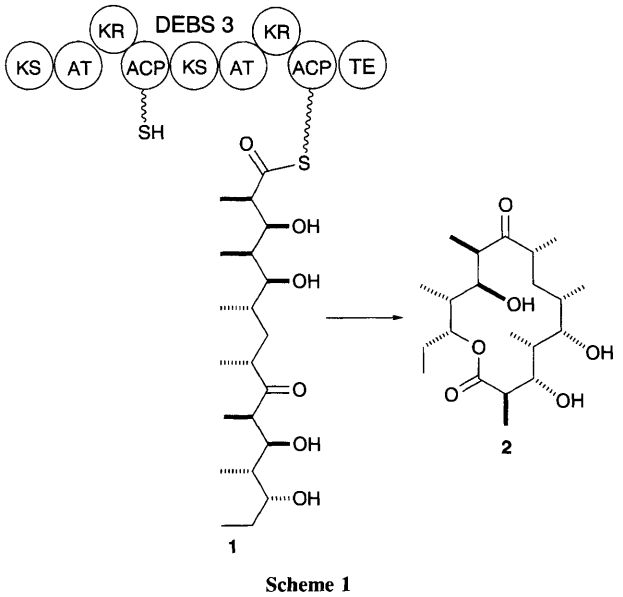
Rapid advances have been made recently in our understanding of erythromycin biosynthesis, largely as a result of genetic studies.[1] It is now known that the polyketide synthase (PKS) which assembles the polyketide chain consists of three giant multidomain proteins called DEBS 1, DEBS 2 and DEBS 3, which cooperate to form a highly organised synthetic production line. The predicted activities for the various domains are based on sequence homology with equivalent domains in fatty acid synthases (*e.g.* KS designates ketone synthase). The domains are sequentially organised into modules one for each chain extension cycle and the growing chain is thought to move from one module to the next until at the end of DEBS 3 the completed heptaketide chain **1** is ready for cyclisation (Scheme 1). The thioesterase domain (TE) conveniently located at this point provides a mechanism for release of the chain, probably by lactonisation to give the first enzyme-free intermediate, 6-deoxyerythronolide B **2**, as indicated, or possibly by transfer of the acyl chain to an external nucleophilic species.

Despite their structural complexity, the DEBS proteins have been genetically engineered in rational ways to produce new products. Most reports to date have described destructive alterations[2] which inevitably have very limited potential. More recently, a constructive alteration has been achieved by relocation of a functional domain, the thioesterase (TE) naturally located at the end of DEBS 3, to the C-terminal end of DEBS 1.[3] In mutant organisms containing this engineered construct (DEBS 1-TE) the relocated domain interrupted the chain extension process at the triketide intermediate **3** leading to its release as the corresponding lactone **4** (Scheme 2). Two products have been identified to date, one with an acetate starter (R = Me) the other with a propionate starter (R = Et), the relative amounts depending on the particular mutant.[4]

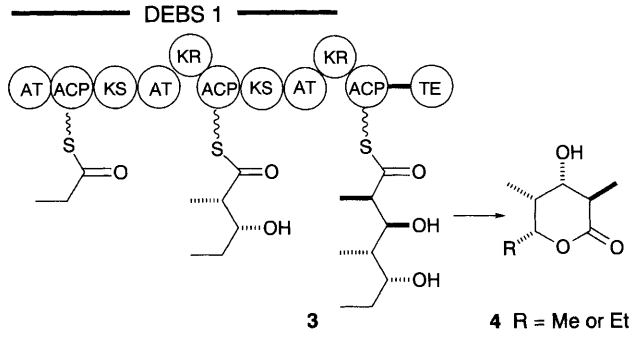
This demonstration that the TE domain can function in an unnatural position and with a radically different substrate points to the potential of using this reconstructive strategy to redirect polyketide chain assembly more widely even in other organisms with other PKS systems. We have therefore investigated the substrate specificity of the TE domain *in vitro* using a truncated version of the DEBS 3 protein, consisting of the C-terminal bidomain, acyl carrier protein-thioesterase (ACP-TE), which was over-expressed in *Escherichia coli*.[5] In this host, the acyl carrier protein domain was not activated by post-translational addition of a phosphopantetheine residue. The ability of the neighbouring TE domain to function independently as a thioesterase *in vitro* with added substrates was not be affected by this deficiency, however. In the earlier experiments the TE bound the typical protease inhibitor, phenylmethylsulfonyl fluoride (PMSF) in a manner consistent with its presumed thioesterase activity.[5]

Two types of acyl derivative were effective substrates for the TE. The first, the *N*-acetylcysteamine (NAC) thioester, mimics the phosphopantetheine thioester derivative of an ACP and therefore comes close to the normal substrate of the TE in its natural context. The second type, the *p*-nitrophenyl ester, was less appropriate in this respect, but the reaction could be more conveniently followed. In a typical experiment the test substrate was incubated with the ACP-TE in aqueous buffer at pH 7.4, containing 1% (*v*/*v*) of ethanol or acetonitrile to aid solubility of the substrate. The reactions of the NAC derivatives were followed by analysis of the products by ^{1}H NMR. The rate of cleavage of the phenyl esters was monitored by UV at 400 nm (λ_{max} for the phenolate anion) as well as by analysis of the products by ^{1}H NMR. In all cases control experiments were carried out in the absence of enzyme to determine the background rate of chemical hydrolysis.

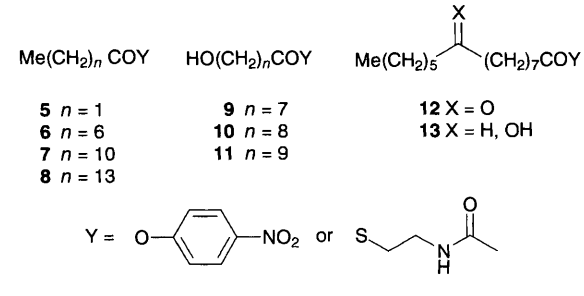
Initially small scale exploratory experiments at high dilution using UV detection were carried out with *p*-nitrophenyl derivatives of the acyl chains **5**–**13**. Compounds **5**–**8** are simple



Scheme 1



Scheme 2 Function of DEBS 1-TE in mutants



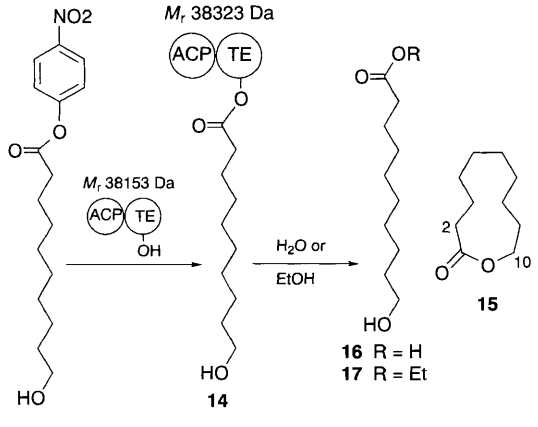
fatty acyl chains which completely lack the complex substitution of the normal substrate **1**; the ω-hydroxyacyl chains of **9–11** offer the possibility of lactone formation; in **12** there is a strategically placed keto group (at C-9 as in the natural substrate 1 of the esterase) which could assist binding to the TE, and in **13** this group is reduced to OH. Most of the test compounds were acceptable substrates for the TE, although there was a wide variation in the rate of reaction (Table 1). Generally, the rate fell as the chain length increased, but this effect could be compensated by extra polar groups (OH or carbonyl). The long list of successful substrates demonstrates that the TE of DEBS 3 can accomodate a wide range of structures. It is possible that the negative results with **7** and **8** were the result of reduced solubility (the reaction mixtures were hazy throughout) rather than an inherent inability of the longer acyl chain to enter the TE active site.

In view of the earlier evidence that the thioesterase reacts with the standard serine protease inhibitor, PMSF, we suspected that the mechanism would involve formation of an acyl enzyme intermediate.[5] The reaction of the hydroxyacyl derivative **11** was therefore monitored by withdrawing samples at intervals and analysing the protein content by electrospray mass spectrometry. Spectra were transformed to produce single peaks with *m/z* values corresponding to singly charged ions. At the start of the experiment a single peak was observed for the underivatised protein (M_r 38153 Da). As the hydrolysis proceeded the intensity of this peak diminished and a new peak appeared at higher mass (M_r 38323 Da). This is the expected value for the acyl enzyme intermediate **14** in which the active serine hydroxy of the TE had been acylated by the acyl group of the substrate (Scheme 3). Eventually the peak corresponding to acyl enzyme began to diminish in intensity and the original peak to increase until after about 10 min all the acyl enzyme had been hydrolysed. The TE therefore behaves like a serine protease in which the initial step of the ester cleavage involves transfer of the acyl chain to a serine hydroxy in its active site. Similar evidence for acyl-enzyme formation was found for **9** and **10**.

Table 1 Relative rates of hydrolysis of *p*-nitrophenyl esters catalysed by the thioesterase (TE)

Acyl chain	Rate[a]
5	3.0
6	4.6
7	0
8	0
9	21.6
10	20.2
11	14.9
12	0.36
13	0.43

[a] Measured at 37 °C for the appearence of *p*-nitrophenylate anion (λ_{max} 400 nm) after the reaction had reached the initial linear phase; substrate concentration, 0.125 mmol dm^{-3}; enzyme concentration, 0.16 nmol dm^{-3}; Tris buffer (10 mmol dm^{-3}) containing EDTA (0.2 mmol dm^{-3}) at pH 7.4.



Scheme 3 Formation and deacylation of the acyl-enzyme intermediate

Next, the experiment with the *p*-nitrophenyl ester **11** was performed on a larger scale to see if the macrolide **15** was formed on deacylation of the acyl-enzyme intermediate. The products were extracted with ethyl acetate and the crude mixture analysed by ^{1}H NMR. The spectrum showed peaks corresponding to a mixture of the corresponding carboxylic acid **16** and its ethyl ester **17**, formed by nucleophilic attack by water and ethanol respectively (Scheme 3). There were no other significant peaks in the spectrum. In particular there was no evidence for multiplets at δ 2.3 or 4.1, which are present for H-2 and H-10 respectively in the spectrum of a synthetic sample of the macrocyclic lactone **15**. A synthetic sample of the lactone was found to be stable under the conditions of the reaction. The ratio (2 : 1) of the amounts of **17** to **16** in the product mixture was much higher than the molar ratio (*ca.*1 : 300) of ethanol to water in the reaction medium, which shows that the esterase has a strong preference for use of an alcohol rather than water as the nucleophile for the deacylation step. Similar results were obtained the *p*-nitrophenyl esters **9** and **10**.

Finally, the *N*-acetylcysteamine thioesters corresponding to **9**, **10** and **11** were tested as substrates. In these experiments the nature and extent of reaction after a fixed period was again determined by isolation of the products by ethyl acetate extraction followed by ^{1}H NMR analysis of the crude mixture. The rate and product mixture for each NAC derivative was closely similar to that found for the corresponding *p*-nitrophenyl analogue. The NAC derivatives were stable in the absence of enzyme.

Our results establish that the TE has a broad substrate specificity as far as the acyl chain is concerned. It forms acyl enzyme intermediates and it has a similar versatility in the deacylation step in that it can use an external nucleophile rather than an internal OH group to release the chain. The TE is therefore a potentially versatile agent for release of diverse polyketide chains from engineered polyketide synthases.

We thank the Association of Commonwealth Universities for a Commonwealth Scholarship to C. J. S. and the SERC for financial support.

Received, 24th March 1995; Com. 5/01885D

References

1 J. Staunton, *Angew. Chem., Int. Ed. Engl.*, 1991, **30**, 1302.
2 S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson and L. Katz, *Science*, 1991, **252**, 675; S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson and L. Katz, *Gene*, 1992, **115**, 97; S. Donadio, J. B. McAlpine, P. J. Sheldon, M. Jackson and L. Katz, *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 7119.
3 J. Cortes, K. Wiesmann, G. A. Roberts, M. Brown, J. Staunton and P. F. Leadlay, *Science*, 1995, **268**, 1487.
4 M. J. B. Brown, J. Cortes, A. L. Cutter, P. F. Leadlay and J. Staunton, *J. Chem. Soc., Chem. Commun.*, 1995, 1517.
5 P. Caffrey, B. Green, L. C. Packman, B. J. Rawlings, J. Staunton and P. F. Leadlay, *Eur. J. Biochem.*, 1991, **195**, 823.