ELSEVIER

Chemistry & Biology 8 (2001) 1197–1208

Chemistry & Biology

www.elsevier.com/locate/chembiol

Research Paper

Chain initiation on the soraphen-producing modular polyketide synthase from *Sorangium cellulosum*

Christopher J. Wilkinson [a], Elizabeth J. Frost [b], James Staunton [b], Peter F. Leadlay [a,*]

[a] *Cambridge Centre for Molecular Recognition and Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1GA, UK*
[b] *Cambridge Centre for Molecular Recognition and Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 9EY, UK*

Received 28 June 2001; revisions requested 17 September 2001; revisions received 21 September 2001; accepted 21 September 2001
First published online 7 November 2001

Abstract

Background: Polyketides are structurally diverse natural products with a wide range of useful activities. Bacterial modular polyketide synthases (PKSs) catalyse the production of non-aromatic polyketides using a different set of enzymes for each successive cycle of chain extension. The choice of starter unit is governed by the substrate specificity of a distinct loading module. The unusual loading module of the soraphen modular PKS, from the myxobacterium *Sorangium cellulosum*, specifies a benzoic acid starter unit. Attempts to design functional hybrid PKSs using this loading module provide a stringent test of our understanding of PKS structure and function, since the order of the domains in the loading and first extension module is non-canonical in the soraphen PKS, and the producing strain is not an actinomycete.

Results: We have constructed bimodular PKSs based on DEBS1-TE, a derivative of the erythromycin PKS that contains only extension modules 1 and 2 and a thioesterase (TE) domain, by substituting one or more domains from the soraphen PKS. A hybrid PKS containing the soraphen acyltransferase domain AT1b instead of extension acyltransferase domain AT1 produced triketide lactones lacking a methyl group at C-4, as expected if AT1b catalyses the addition of malonyl-CoA during the first extension cycle on the soraphen PKS. Substitution of the DEBS1-TE loading module AT domain by the soraphen AT1a domain led to the production of 5-phenyl-substituted triketide lactone, as well as the normal products of DEBS1-TE. This 5-phenyl triketide lactone was also the product of a hybrid PKS containing the entire soraphen PKS loading module as well as part of its first extension module. Phenyl-substituted lactone was only produced when measures were simultaneously taken to increase the intracellular supply of benzoyl-CoA in the host strain of *Saccharopolyspora erythraea*.

Conclusions: These results demonstrate that the ability to recruit a benzoate starter unit can be conferred on a modular PKS by the transfer either of a single AT domain, or of multiple domains to produce a chimaeric first extension module, from the soraphen PKS. However, benzoyl-CoA needs to be provided within the cell as a specific precursor. The data also support the respective roles previously assigned to the adjacent AT domains of the soraphen loading/first extension module. Construction of such hybrid actinomycete–myxobacterial enzymes should significantly extend the synthetic repertoire of modular PKSs.

Keywords: Acyltransferase; Chain initiation; Erythromycin; Polyketide; Soraphen

1. Introduction

Macrocyclic polyketides form a large and structurally diverse class of natural products and include numerous clinically important antibacterial, antifungal, anthelminthic, antitumour and immunosuppressant compounds. They are synthesised on type I modular polyketide synthases (PKSs) that consist of giant multienzymes [1–8] housing sets or modules of enzymatic activities. Each module contains the enzymatic domains required to catalyse a particular cycle of polyketide chain extension and, where appropriate, β-keto processing. For chain extension, an acyltransferase (AT) domain selects the appropriate dicarboxylic acid, activated as its coenzyme A (CoA) ester. For chain initiation, starter units derived from monocar-

Abbreviations: PKS, polyketide synthase; DEBS, 6-deoxyerythronolide B synthase; TKS, triketide synthase; AT, acyltransferase; ACP, acyl carrier protein; TE, thioesterase; KSQ, β-ketoacyl synthase-like domain with glutamine (Q) substituted for the active site cysteine residue; ::, integrated into the genome

* Corresponding author.
E-mail address: pfl10@mole.bio.cam.ac.uk (P.F. Leadlay).

1074-5521/01/$ – see front matter
PII: S1074-5521(01)00087-4

boxylic acids are selected by one of several alternative mechanisms, as explained below. The enzymatic activities that govern chain initiation are housed in a dedicated loading module (LM or Mod0) located immediately adjacent to the first extension module, normally as part of the same multienzyme protein but in some cases as a discrete entity [9–11]. Likewise, there is normally a dedicated domain C-terminal of the final extension module that governs chain termination, consisting often of a thioesterase (TE) domain which may act as a cyclase. Alternatively, the fully extended acyl chain may serve as the substrate for further enzymes, as in the biosynthesis of rapamycin where an amino acid is added by a module of non-ribosomal peptide synthetase activities [4,12].

Most of the biosynthetic gene clusters for macrocyclic polyketides that have been sequenced so far have been cloned from the actinomycete order of Gram-positive bacteria [13], and each PKS has proved to be closely patterned on the structure originally described for the 6-deoxyerythronolide B synthase (DEBS) of *Saccharopolyspora erythraea* which produces the precursor of the 14-membered macrolide antibiotic erythromycin A (Fig. 1). More recently, several modular PKS clusters have been identified from non-actinomycete bacteria, and they often show puzzling differences in the organisation of their modules, from the canonical organisation established for streptomycete PKSs, which have yet to be satisfactorily explained. For example, the PKSs found in the *Bacillus* genome (accession number AL009126) apparently contain no integral extender AT domains, and this is also a feature of the putative pseudomonic acid synthase from *Pseudomonas fluorescens* [14].

Other modular PKSs have been recently characterised from the Gram-negative myxobacteria [15,16]. These include those for the antifungal soraphen [17] (Fig. 1) and the antitumour compound epothilone [18,19], both obtained from different strains of *Sorangium cellulosum*. The sequence of the soraphen biosynthetic gene cluster is available in the patent literature (EMBL accession number I88042). It contains two type I modular PKSs, SorA and SorB. An unusual feature of this PKS is the lack of a clearly demarcated loading module for the unusual benzoate starter unit. Instead, the domains of the loading and first extension module appear to be combined together, so that the AT domain of the loading module is located within the first extension module. Recent genetic engineering of PKSs to create novel antibiotics has heavily emphasised the key role of chain initiation [20] and a number of novel antibiotics have been made by this route [21,22]. We therefore hoped that a closer investigation of the loading module of the soraphen PKS would not only yield insight into the role of individual domains in the scrambled loading module, but might also provide an alternative route to the biosynthesis of hybrid macrolides from a benzoate starter unit.

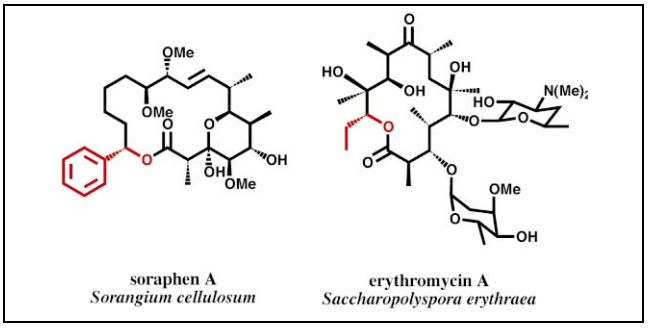

Fig. 1. The structures of soraphen A and erythromycin A. The starter ketide units introduced by the respective loading modules are highlighted.

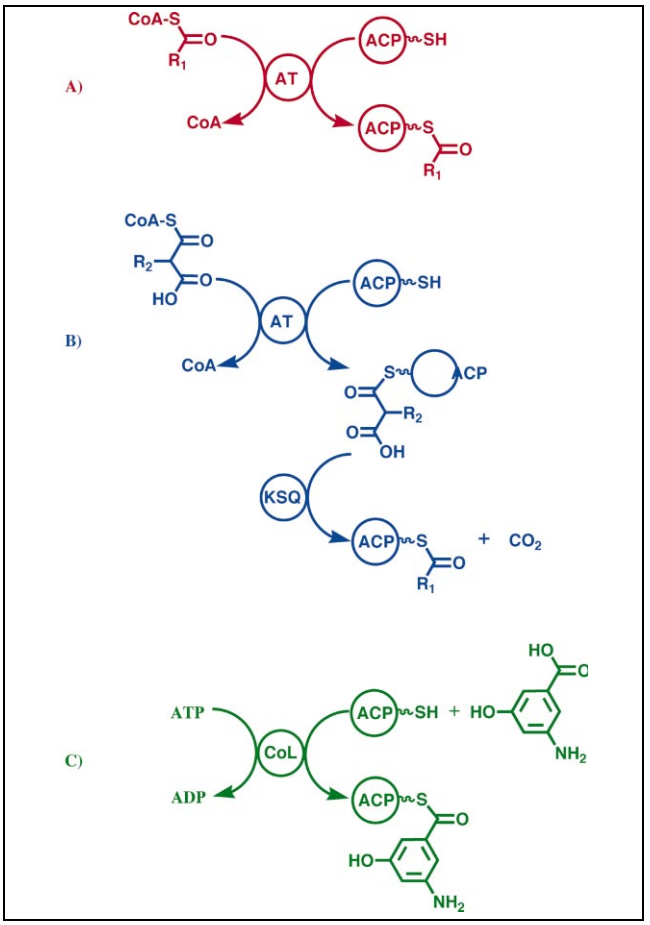

Fig. 2. Mechanisms of chain initiation on modular PKSs: (A) acyltransfer from an acyl-CoA ester to an acyl carrier protein (ACP) domain catalysed by an AT domain; (B) acyltransfer from the CoA ester of a dicarboxylic acid to an ACP domain catalysed by an AT domain, followed by in situ decarboxylation mediated by the KSQ domain; (C) ATP-dependent activation of a carboxylic acid to an acyladenylate, followed by transfer to an ACP domain. KS, β-ketoacyl synthase; KR, β-ketoacyl reductase; DH, dehydratase; ER, enoylreductase; TE, thioesterase; CoL, CoA ligase-like enzyme; KSQ, KS-like domain with glutamine (Q) substituted for the active site cysteine residue; R_1, Me or Et; R_2, Me or H.

The starter units for most complex polyketides are derived from simple aliphatic acids such as acetate and propionate. These are selected, as shown in Fig. 2, by an

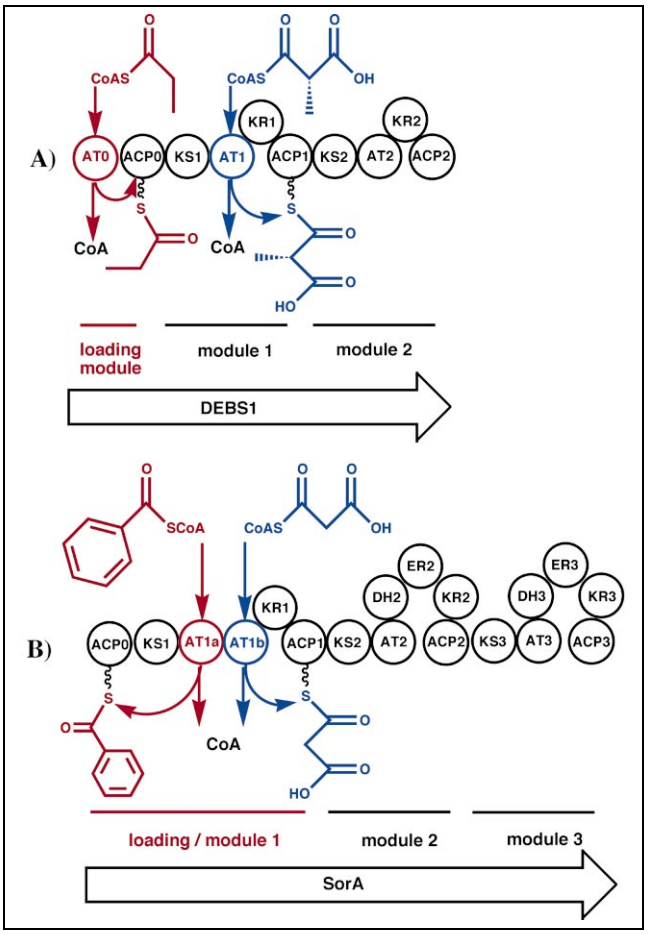

Fig. 3. Domain organisation of (A) multienzyme DEBS1-TE derived from the erythromycin-producing PKS and (B) multienzyme SorA, which catalyses chain initiation and two rounds of chain extension on the soraphen-producing PKS. The proposed function of individual AT domains in initiation is shown.

N-terminal loading module which may consist either of an AT followed by an acyl carrier protein (ACP) domain [23], as in erythromycin or avermectin; or more commonly of KSQ, AT and ACP domains, where the AT recruits a dicarboxylic acid onto the ACP and the KSQ domain accomplishes its in situ decarboxylation [24]. More complex moieties, for example the alicyclic and aromatic starter units of rapamycin and rifamycin respectively, are selected by CoA ligase-like domains which activate the carboxylic acid to the corresponding acyl adenylate, using ATP, before their covalent attachment to the loading module via an ACP domain [4,25,26].

Here we show that when the N-terminus of the soraphen PKS, bearing the loading module and most of the first extension module, was fused to an appropriately truncated version of DEBS1-TE, the chimaeric enzyme was capable of synthesising low levels of the predicted 5-phenyl triketide lactone. When the AT1a domain alone was moved to the orthodox position for a loading AT in the DEBS1-TE system (Fig. 3), the phenyl-substituted product was also detected. Likewise, the AT1b domain was found to substitute for the function of the first extension AT domain within the DEBS1-TE. These results accord with predictions made on the basis of alignments of their respective protein sequences with authentic loading and extender AT domains from other PKSs. The low yields of novel hybrids obtained are due at least in part to the difficulty of providing adequate intracellular levels of the benzoate precursor.

2. Results and discussion

2.1. *Comparisons of the protein sequences of type I PKS AT domains*

The crystal structure of the malonyl-CoA:ACP AT of the *Escherichia coli* fatty acid synthase [27] reveals that arginine residue 117 lies in the active site, where it might provide an ionic interaction with the carboxy group of the malonyl-CoA substrate, and favour the binding of this over the alternative substrate for this AT, acetyl-CoA. Smith and his colleagues [28] have shown, using site-directed mutagenesis, that the counterpart of this specific arginine residue in animal fatty acid synthase AT indeed contributes to the discrimination in favour of carboxylated over non-carboxylated CoA substrates. The alignment in Fig. 4 shows, strikingly, that in the AT domains of the loading modules of avermectin PKS (AVES) and erythromycin PKS (DEBS), which are the only loading module ATs known to be specific for non-carboxylated aliphatic CoA derivatives [23,29,30], the arginine at this position is replaced in both cases by a tryptophan residue.

Similar considerations can be used for preliminary assignment of roles to the two N-terminal SorA AT domains. When the protein sequences of these two ATs (designated AT1a and AT1b) are included in the sequence alignment (Fig. 4), there are significant differences between the two. First, AT1a is most similar to the loading AT domains of AVES and DEBS, while AT1b is grouped with malonyl-CoA-specific extender ATs. Specifically, at position 117 (*E. coli* numbering) AT1b has an arginine residue, while, in contrast, AT1a has a tyrosine residue. On this basis, AT1a is proposed to be specific for benzoyl-CoA and to select the starter unit, while AT1b selects malonyl-CoA and this results in the observed incorporation of acetate into soraphen as the first act of chain extension [31] (Fig. 3).

2.2. *Production of 5-phenyl triketide lactones by a hybrid bimodular PKS incorporating both sor AT domains into DEBS1-TE*

Initially, we attempted to introduce the soraphen loading module and part of the first extension module of the soraphen PKS into DEBS1-TE, replacing the first four domains of DEBS1-TE so as to preserve the unorthodox protein architecture in the resulting hybrid triketide lac-

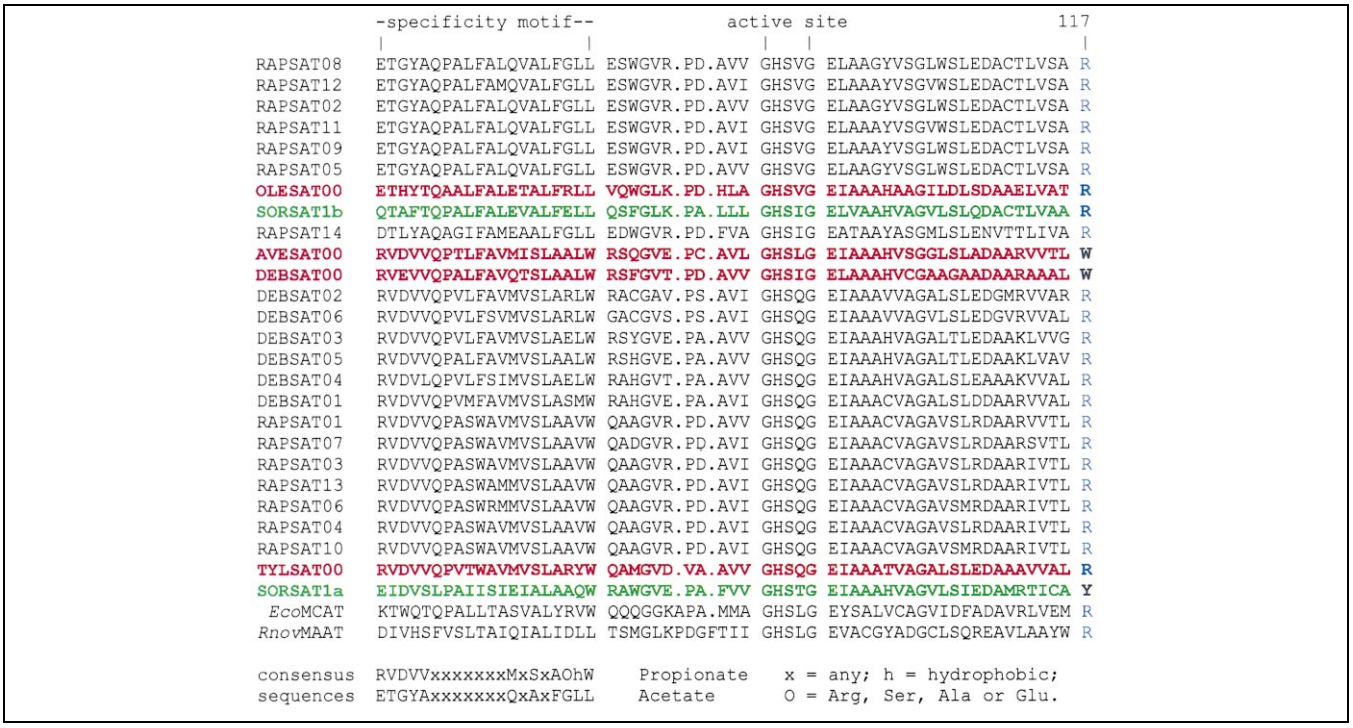

Fig. 4. Comparison of the protein sequences of AT domains. Loading module ATs are in red, bold font; the soraphen-producing PKS ATs of unknown specificity and role are in green, bold font. The specificity motif is defined by Haydock et al. [59]; the arginine residue conserved in all ATs specific for (methyl)malonate (blue) was observed by Rangan et al. [28]. DEBS, erythromycin PKS; RAPS, rapamycin PKS; SORS, soraphen PKS; OLES, oleandomycin PKS; TYLS, tylosin PKS; *Rnov*MAAT, rat fatty acid synthase AT; *Eco*MCAT, *E. coli* fatty acid synthase AT.

tone synthase. This synthase, TKSCJW1, contains both SorA AT1a and AT1b. The fusion site between SorA and DEBS1-TE lies between SorA AT1b and the DEBS KR1 [32], so the first extension module is chimaeric. If the SorA domain arrangement is required for incorporation of benzoate as starter unit and acetate as extender unit respectively, then the lactone **5** (Fig. 5) is the expected product.

An *S. erythraea* strain JC2 [33], from which the DEBS genes have been deleted, was transformed with pCJW38, an expression vector for TKSCJW1 based on the integrative plasmid pCJR24 [33]. The resulting recombinant strain appeared to make no lactone products, as judged by GC-MS analysis of extracts of culture broths (data not shown). One plausible reason for this is the lack of provision within the cells of the specific precursor of benzoate units, assumed to be benzoyl-CoA. There is no evidence that *S. erythraea* generates benzoyl-CoA as part of its normal metabolism. In parallel, therefore, *S. erythraea* JC2 was transformed with a plasmid (pCJW40) that contained both the hybrid TKS gene TKSCJW1 and the *badA* gene from *Rhodopseudomonas palustris*. This anaerobic bacterium can use benzoic acid as sole carbon source, and has a dedicated gene cluster to enable it to utilise benzoic acid [34]. The first enzyme in this catabolic pathway is BadA, an ATP-dependent benzoic acid:CoA ligase [35]. A clone of this gene, pPE202, was the kind gift of Professor Caroline S. Harwood, Department of Microbiology, University of Iowa, Iowa City, IA, USA. The *badA* gene was expressed under the strong, constitutive promoter, P_{ermE*} [36] (Fig. 6).

S. erythraea JC2 cells transformed with plasmids containing this gene cassette were grown in SM3 medium with a 3 mM final concentration of sodium benzoate. Concentrations higher than this were found to retard or even completely inhibit the growth of *S. erythraea* on solid medium (data not shown). This recombinant strain initially appeared to make none of the expected lactones. However, analysis by the more sensitive method of GC-MS/MS revealed a small peak (by relative ion current) that had the same retention time and mass spectrum as the synthetic standard (Fig. 7). This peak was not present in a control strain which lacked the *badA* gene or was not fed with benzoate (data not shown). The expected benzoate starter lactone product had therefore been produced, albeit at less than 0.2 mg/l, representing less than 0.5% of normal rates of triketide production by a control strain containing DEBS1-TE under the same conditions.

Analysis using SDS–PAGE of cell extracts of the recombinant *S. erythraea* did not reveal the presence of significant levels of a high molecular weight protein, indicating that poor expression levels of the hybrid PKS were also an important contributing factor in the poor production of the expected lactone. We therefore turned to ex-

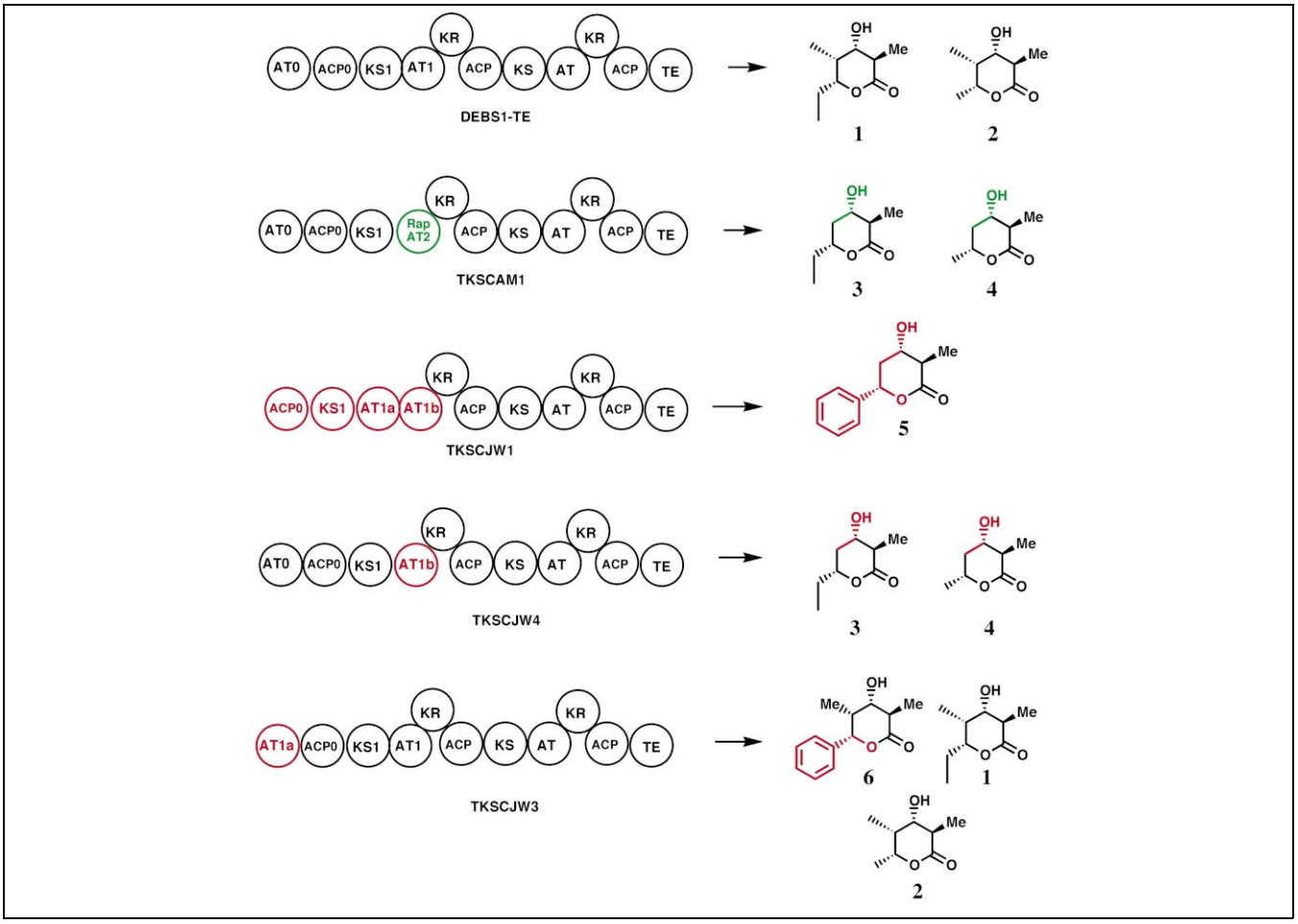

Fig. 5. Domain organisation of DEBS1-TE (Cortes et al., [60]), TKSCAM1 (Oliynyk et al. [32]), and the TKSs constructed in this study. Domains from the soraphen PKS are in red; rapamycin AT2 is in green.

amining DEBS1-TE hybrids in which individual *sor* PKS AT domains have been introduced, both to confirm their individual roles in the soraphen PKS, and possibly to obtain more stable and productive hybrids.

2.3. *AT1b of the soraphen PKS functions as an extender module AT in DEBS1-TE*

Using the system of Oliynyk et al. [32], AT1b of the soraphen PKS was substituted for AT1 of DEBS1-TE. *S. erythraea* JC2 cells expressing this synthase, TKSCJW4, produced the expected C4 desmethyl triketide lactones that would result from incorporation of acetate as the first extension unit, as judged by GC-MS analysis. Yields were low, around 0.5 mg/l, less than 1% of the levels of triketide lactones normally produced under these conditions by DEBS1-TE. The simplest explanation for these results is that the introduced AT either decreases the stability of the hybrid PKS, or is not correctly folded against the flanking domains. Despite this, these data serve to establish that AT1b of the soraphen PKS can indeed function as a bona fide extender AT with the expected specificity for malonyl-CoA.

2.4. *Production of 5-phenyl triketide lactones by a hybrid bimodular PKS incorporating the sor AT1a domain in place of the loading module AT of DEBS1-TE*

Restriction sites were engineered in the gene for DEBS1-TE, flanking the sequence for the loading module AT, to allow this domain to be specifically replaced by the *sor* AT1a domain. These sites were introduced in the same positions as those flanking the DNA sequence of AT1 of DEBS1-TE in the engineered synthases produced by Oliynyk et al. [32]. In that work, an *Msc*I site was introduced without alteration in the encoded amino acid sequence at the margin of the highly conserved N-terminus of the AT domain. An *Avr*II site was introduced in the sequence for the interdomain region C-terminal of the AT domain, again without changing the encoded amino acid sequence. In the DEBS loading module AT, the protein sequence at this C-terminal splice site is not conserved between the DEBS and *sor* AT domains. Introduction of the *Avr*II site was therefore mutagenic (altering amino acids from DD to LG). It was important to determine if this was detrimental to enzyme activity, especially since charged residues had been replaced with hydrophobic ones. In a

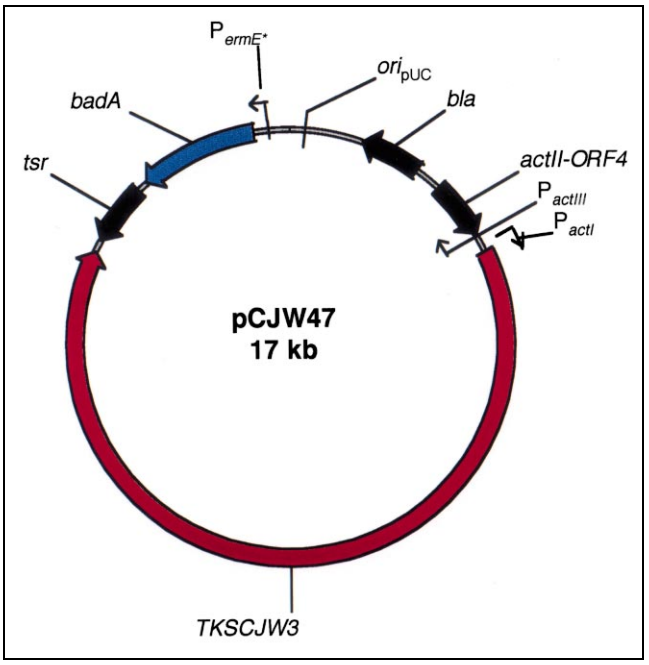

Fig. 6. Map of pCJW47, plasmid used for co-expression of the hybrid synthase TKSCJW3 (red) and of *badA* (blue). *tsr*, thiostrepton resistance gene; *actII-ORF4* and $P_{actI/III}$, *Streptomyces coelicolor* actinorhodin cluster activator–promoter system; ori_{pUC}, pUC18 origin of replication; *bla*, ampicillin resistance gene.

control experiment, *S. erythraea* JC2 cells expressing this mutated synthase were found to produce the expected triketide lactones in vivo in yields that were qualitatively similar (40–50 mg/l) to those from strains expressing DEBS1-TE. The apparent indifference of the PKS to the introduced mutations presumably reflects the fact that they occur in a mobile flexible linker region [37].

The loading module AT (AT0) of DEBS1-TE was then replaced with AT1a of SorA and the hybrid enzyme expressed in *S. erythraea* JC2. GC-MS analysis of the culture broths revealed three triketide lactones to have been produced (Fig. 7). One was the expected benzoate starter lactone, whose identity was confirmed by comparison with a synthetic standard. The other two triketide lactones had incorporated the starter units normally selected by the DEBS loading module, that is, propionate and acetate respectively. All of these three products were again present in low yield, less than 1% of the levels of triketide lactones normally produced under these conditions by DEBS1-TE.

This experiment was repeated in the absence of *badA*. The polyketide products of this strain, *S. erythraea* JC2::pCJW45, were lactones with either an acetate or propionate starter but not with a benzoate starter, as judged by GC-MS and GC-MS/MS analysis of broth extracts (Fig. 7). BadA was therefore directly shown to be important in the supply of substrate to the TKS. Unfortu-

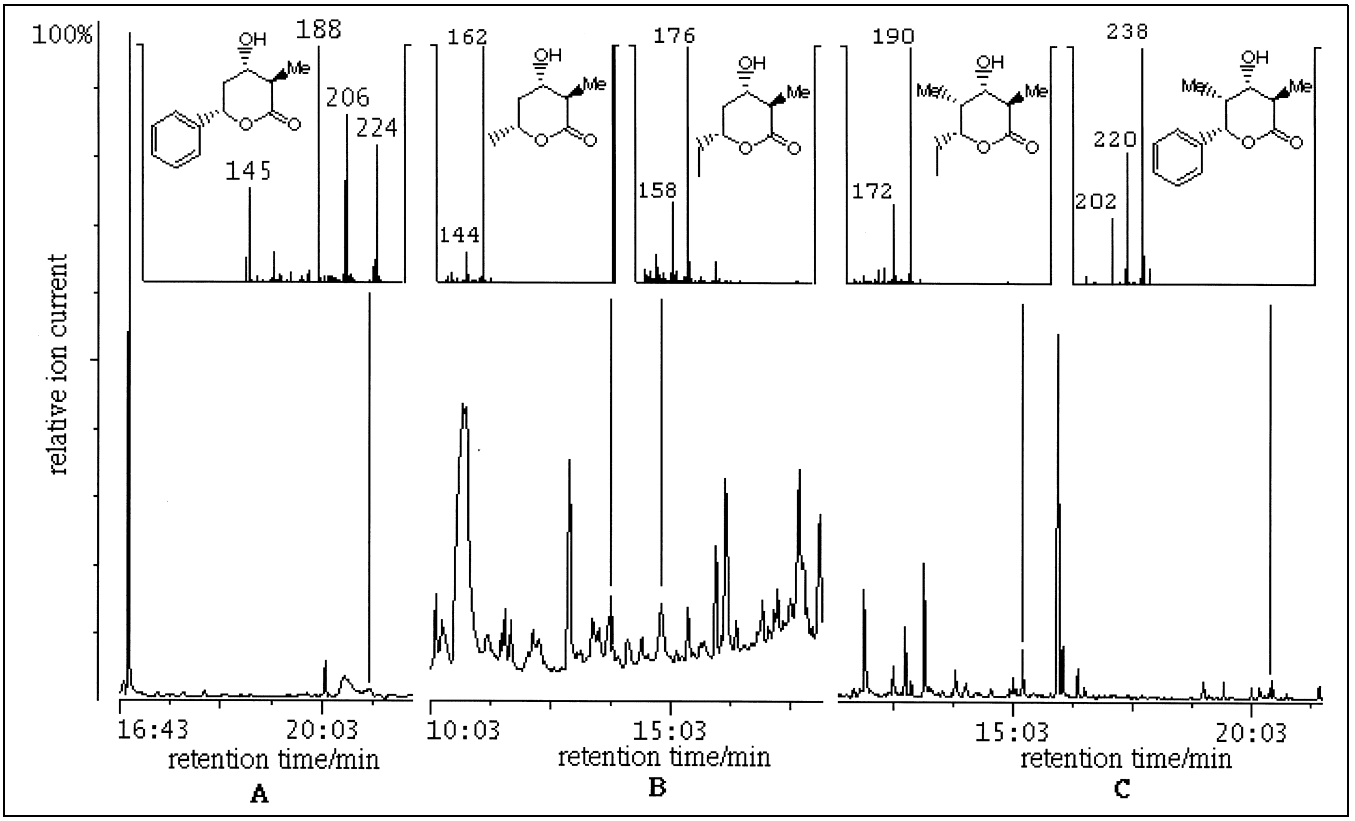

Fig. 7. The total ion current of typical GC-MS traces of ethyl acetate extracts from (A) *S. erythraea*::pCJW40; (B) *S. erythraea*::pCJW49; (C) *S. erythraea*::pCJW47. The structures of the compounds producing the peaks highlighted are also shown together with their mass spectra.

nately, the toxicity of sodium benzoate limited the use of higher concentrations of this precursor in the medium.

An alternative means of supplying the loading AT with substrate was sought, in the form of the *N*-acetylcysteamine (NAC) thioester of benzoic acid (**7**). NAC thioesters act as CoA mimics and can diffuse into the cell [38–41]. When *S. erythraea* JC2::pCJW47 cells were fermented in the presence of this compound, the relative levels of the 5-phenyl triketide lactone **5** appeared to increase as judged by GC-MS and GC-MS/MS (Fig. 7), again indicating that substrate supply to AT1a is contributing to limit the yields in this system. Recent work on the origin of the benzoate starter of the PKS for the aromatic polyketide enterocin [42,43] has revealed that the biosynthetic gene cluster contains genes not only for a BadA analogue, but also for phenylalanine ammonia lyase. In future, co-expression of a heterologous phenylalanine ammonia lyase gene in *S. erythraea* may provide a more efficacious way to enhance the intracellular pool of benzoate, since cinnamate produced from L-phenylalanine is readily catabolised to benzoic acid. Similarly, feeding the relevant bacterial host strain phenyl-substituted fatty acids with an odd-numbered chain length may offer a route to the generation in the cell of benzoyl-CoA and derivatives through the β-oxidation pathway, which itself was investigated using such compounds in the very first biochemical labelling experiments by Franz Knoop in 1904 [44].

2.5. The hybrid loading module containing the sor AT1a domain is required for incorporation of benzoate starter units by DEBS1-TE

Propionate starter units have previously been shown to be incorporated by the DEBS through direct loading onto the ketosynthase active site in extension module 1 [45] (N. Dunster and P.F. Leadlay, unpublished data). If the hybrid loading module containing the *sor* AT1a domain is not functioning efficiently, possibly due to problems with protein interactions, dimerisation, or substrate supply, this alternative route for chain initiation might be favoured, and could readily account for the appearance of lactones **1** and **2** in these experiments, without having to postulate that the *sor* AT1a is able to accommodate propionyl- and acetyl-CoA as substrates as well as benzoyl-CoA.

To test for such 'direct loading' as a mechanism for the generation of the benzoate starter lactone, a version of DEBS1-TE containing an inactive ACP domain in the loading module was studied in vivo under the same conditions. pND4 is a pCJR24-based expression plasmid for a DEBS1-TE derivative in which the loading module ACP (ACP0) has been inactivated by mutation of a specific serine residue to alanine, abolishing the phosphopantetheine attachment site. In vivo, this mutant enzyme makes small (less than 5% of normal levels) amounts of acetate and propionate starter triketide lactones (N. Dunster and P.F. Leadlay, unpublished data). Since the perturbation of the protein structure by this point mutation and hence the effects on expression and folding of the PKS are expected to be very small, these data give an indication of the levels of direct loading that occur when the loading module is totally inactivated.

S. erythraea JC2::pND4 cells were grown in SM3 medium supplemented with benzoyl-NAC. When crude extracts of these broths were analysed by GC-MS both lactones **1** and **2**, bearing respectively a propionate and an acetate starter unit, were observed at levels previously found without supplementation. No benzoate starter lactone was observed, within the limits of detection (data not shown), indicating that the 5-phenyl triketide lactone **6** is unlikely to have arisen by direct loading onto KS1.

3. Perspective

The results of these experiments taken together establish that in principle functional hybrid PKSs constructed from myxobacterial DNA and actinomycete DNA can be created and used to make novel polyketides, which might considerably increase the structural diversity of the products that can be obtained by PKS engineering. However, in order to realise this potential, it is clear from this work that improvements to the levels of expression of the hybrid proteins are required; as well as the engineering of appropriate metabolic pathways to provide an intracellular supply of specific building blocks such as benzoyl-CoA, before these procedures can be carried out on a preparative scale. Alternative routes to such products have already been described or may become available, either through feeding of the appropriate diketide NAC thioesters to PKSs in which KS1 has been inactivated [46] or through the use of other heterologous loading modules such as avermectin [29], rifamycin [26,47] or rapamycin [48].

Nevertheless, study of these hybrid synthases has provided support for the proposed functions of the SorA AT1a and AT1b domains. It has also been possible to separate the two ATs and still maintain their respective function. On the available evidence, the 5-phenyl triketide lactone is the product of a conventional mechanism for chain initiation in which the benzoyl moiety is transferred to the loading module ACP by the soraphen PKS AT1a.

Sequencing of other myxobacterial PKS gene clusters has revealed the same unusual tandem arrangement of AT domains in the loading and first extension modules, even where the starter unit is an aliphatic carboxylic acid such as 3-methylbutyrate [16]. If the folding of the PKS allows the active site of the loading AT1a domain to be reached directly by the swinging arm of the loading ACP, this must require a three-dimensional arrangement of domains within the myxobacterial enzymes that is rather different from current models for modular PKS structure [49,50].

These results have also strengthened the view that argi-

nine 117 (*E. coli* numbering) is an important residue for the ability of PKS AT domains to discriminate between the CoA esters of mono- and dicarboxylic acids. In those systems whose mechanism of chain initiation is unclear [9,10], the loading module AT also contains an arginine residue at this position. Structural and kinetic experiments with highly purified domains in vitro, together with site-directed mutagenesis, will be required to quantify the contribution of this residue and establish its mechanistic importance.

4. Significance

Modular PKSs are responsible for the biosynthesis of a wide range of clinically important natural products, including antibacterial, antitumour, antiparasitic and immunosuppressant compounds. They are increasingly being uncovered in bacteria other than the actinomycetes, with the myxobacteria proving to be a particularly rich source of such multienzymes. We have shown here that it is possible to analyse the specificity of individual myxobacterial AT domains, implicated in chain initiation and the first round of chain extension, by transplanting them into a well-characterised actinomycete PKS test system derived from the erythromycin-producing PKS. Our results support the view that benzoate units are recruited by the soraphen PKS by a conventional mechanism involving acyl transfer from benzoyl-CoA, or its equivalent, to an ACP domain. The synthesis of detectable amounts of 5-phenyl triketide lactone by the hybrid enzymes shows in principle how we can further increase the structural diversity of products from engineered modular PKSs by creating such hybrids. At the same time, the high-level expression of these hybrid PKSs and the intracellular provision of their precursors remain significant challenges if these opportunities are to be fully exploited.

5. Materials and methods

5.1. Strains

All DNA manipulations were performed in *E. coli* DH10B (Gibco) using standard culture conditions [51]. *E. coli* cells were transformed by electroporation [52]. *S. erythraea* JC2 [33] was used as host for expression of PKS genes.

5.2. Plasmids

pUC18 [53], pSET152 [54] and pCJR24 [33] were used as cloning and expression vectors. pCJR26 was from Rowe et al. [33]. pVKM15 was obtained from the NRRL culture collection. pIB139 and pPFL28, whose construction will be described elsewhere, were the kind gifts of Drs I. Böhm and P.F. Long. The details of construction of other plasmids are given below.

psormod1: pVKM15 was digested with *Bgl*II and *Hin*dIII. The fragment that encodes SorA module 1 was isolated and ligated to *Bam*HI-*Hin*dIII pUC18 backbone.

pCJW31: a fragment encoding SorA domains 1–3 was amplified by PCR using psormod1 as template and oligonucleotides oCJW4.1 and oCJW4.2 as primers. The fragment was 5′-phosphorylated with T4 polynucleotide kinase and ligated to *Sma*I- and SAP-treated pUC18.

pCJW33: a fragment encoding the SorA AT1b was amplified by PCR using psormod1 as template and oligonucleotides oCJW4.5 and oCJW4.6 as primers. The fragment was 5′-phosphorylated with T4 polynucleotide kinase and ligated to *Sma*I- and SAP-treated pUC18.

pCJW34: a fragment encoding *badA* was amplified by PCR using pPE202 as template and oligonucleotides oCJW4.3 and oCJW4.4 as primers. The fragment was 5′-phosphorylated with T4 polynucleotide kinase and ligated to *Sma*I- and SAP-treated pUC18.

pCJW35: pCJW31 and pCJW33 were both digested with *Msc*I and *Eco*RI. The fragment encoding SorA AT1b was ligated to the backbone of pCJW31.

pCJW36: pIB139 and pCJW34 were both digested with *Nde*I and *Eco*RI. The *badA* fragment was ligated to the backbone of pIB139. This places *badA* under the control of P_{ermE*}. This plasmid encodes the *badA* expression cassette.

pCJW38: an *Pac*I-*Avr*II fragment encoding SorA domains 1–4 from pCJW35 was ligated to *Avr*II-*Pac*I pCJR26 backbone.

pCJW40: pCJW38 and pCJW36 were both digested with *Hin*dIII. The *badA* expression cassette was ligated to pCJW38. Orientation: *badA* is transcribed in the same direction as *tsr*.

pCJW41: a fragment encoding the N-terminal ~100 amino acids of DEBS1 was amplified by PCR using pKW9 as template and oligonucleotides oCJW4.9 and oCJW4.10 as primers. The fragment was 5′-phosphorylated with T4 polynucleotide kinase and ligated to *Sma*I- and SAP-treated pUC18.

pCJW42: a fragment encoding the SorA AT1a was amplified by PCR using psormod1 as template and oligonucleotides oCJW4.7 and oCJW4.8 as primers. The fragment was 5′-phosphorylated with T4 polynucleotide kinase and ligated to *Sma*I- and SAP-treated pUC18.

pCJW43: pCJW41 and pCJW42 were both digested with *Msc*I and *Eco*RI. The fragment encoding Sor AT1a was ligated to the backbone of pCJW41.

pCJW44: pPFL28 and pCJW43 were both digested with *Nde*I and *Avr*II. The fragment encoding DEBS1 N-terminus-Sor AT1a was ligated to the backbone of pPFL28. This plasmid encodes TKSCJW3 as an *Nde*I-*Xba*I fragment.

pCJW45: pCJR24 and pCJW44 were both digested with *Nde*I and *Xba*I. The fragment encoding TKSCJW3 was ligated to the backbone of pCJR24. This plasmid is an expression plasmid for TKSCJW3.

pCJW47: pCJW40 and pCJW47 were both digested with *Spe*I and *Eco*RI. The fragment encoding *actII-ORF4*;P_{act};TKSCJW3 was ligated to the backbone of pCJW40. This plasmid is an expression plasmid for TKSCJW3 that also includes the *badA* expression cassette from pCJW36. Orientation: *badA* is transcribed in the same direction as *tsr*.

pCJW49: an *Msc*I-*Avr*II fragment encoding SorA AT1b from pCJW33 was ligated to *Avr*II-*Msc*I pCJR26 backbone.

pCJW80: pCJW45 and pCJW80 were digested with *Msc*I and *Avr*II and the fragment encoding DEBS AT0 ligated to pCJW45 plasmid backbone. This is an expression plasmid for TKSCJW8.

This is a DEBS1-TE derivative, the gene for which has unique *Msc*I and *Avr*II sites flanking the loading module AT.

pCJW81: a fragment encoding the DEBS AT0 was amplified by PCR using pCJR65 as template and oligonucleotides oCJW8.1 and oCJW8.2 as primers. The fragment was 5′-phosphorylated with T4 polynucleotide kinase and ligated to *Sma*I- and SAP-treated pUC18. Orientation: anti.

pIB139: this plasmid is derived from pSET152. P_{ermE*} was cloned by PCR introducing an *Nde*I site overlapping the start codon downstream of the promoter. This fragment was then inserted into the multiple cloning site of pSET152 as an *Xba*I fragment.

pPFL28: this plasmid encodes a DEBS1-TE derivative with AT2 from the rapamycin PKS fused to DEBS1-TE before ACP0. The splice site is an *Avr*II recognition sequence in the same position in pCJW43.

5.3. Oligonucleotides

The following synthetic oligonucleotides were used in this study (incorporated restriction sites are shown in bold): oCJW4.1, 5′-**TTAATTAA**GGAGGACACATATGACAAAGGAGTACACGCGTCCGCAGTCGGCGCCGTT-3′ (*Pac*I site); oCJW4.2, 5′-TGGCTTCCT**TGGCCA**GTAAAGAGAATGGCGA-3′ (*Msc*I site); oCJW4.3, 5′-AGGAGGAA**CATATG**AATGCAGCCGCGGT-3′ (*Nde*I site); oCJW4.4,5′-TTGC**GAATTCAAGCTT**CAGCCCAACACACCCT-3′ (*Eco*RI and *Hin*dIII sites); oCJW4.5, 5′-TCGCCATTCTCTTTAC**TGGCCA**AGGAAGCCA-3′ (*Msc*I site); oCJW 4.6: 5′-AGCGAAGGGC**CCTAGG**AAAGCGTTCCAGT-3′ (*Avr*II site); CJW4.7: 5′-TT**TGGCCA**GGGGGCGCAGTGGTTCGGCATGG-3′ (*Msc*I site); oCJW4.8: 5′-AA**CCTAGG**TAGACATTGTCCCACCGCACCGGCG-3′ (*Avr*II site); oCJW4.9: 5′-**CATATG**GCGGACCTGTCAAAGCTCTCCGAC-3′ (*Nde*I site); oCJW4.10: 5′-AA**TGGCCA**GGGAAGACGAACACCACCGGCGCGGA-3′ (*Msc*I site); oCJW8.1, 5′-TT**TGGCCA**GGGCGCGCAATGGGCCGGGATGGCGG-3′ (*Msc*I site); oCJW8.2: 5′-AA**CCTAGG**TAGGCGGCGGTCCAGTCGACGGCCACG-3′ (*Avr*II site).

5.4. Manipulation of DNA

Routine cloning and transformation procedures were as previously described for *E. coli* [51]. PCR was performed using *Pfu* polymerase (Promega) according to the manufacturer's instructions using a programmable Robocycler Gradient 40 (Stratagene, USA). Automated DNA sequencing was carried out on double-stranded DNA using an automated ABI 373A sequencer (Applied Biosystems).

5.5. Growth of S. erythraea

Transformation and propagation of *S. erythraea* was carried out using standard methods [55–57]. SM3 was used as production medium. For certain strains, it was supplemented with sodium benzoate at 3 mM or benzoyl-NAC (**7**) at 1.4 mM [46]. SM3 medium: glucose 5 g, MD30E/glucidex 50 g, soya bean flour 25 g, beet molasses 3 g, K_2HPO_4 0.25 g, $CaCO_3$ 2.5 g, Milli-Q™ water to 1 l and pH adjusted to 7.0 with KOH.

5.6. Analysis of hybrid triketide lactone synthases in vivo

S. erythraea JC2-derived strains were grown at 30°C in TSB (Gibco) and SM3 media containing thiostrepton (5 mg/l; Sigma). *S. erythraea* cultures were grown for 5 days in 50 ml SM3 and 10 ml of this fermentation broth was then used for analysis. The broth was first acidified to pH 3 with formic acid before being extracted three times with an equal volume of ethyl acetate. These crude extracts were then analysed by GC-MS on a GCQ instrument (Finnigan, MAT). A Restek Corporation Crossbond™ 5% diphenyl-, 95% dimethyl-polysiloxane column was used with helium as carrier gas. The constant gas velocity was 50 cm/s and the injector temperature was 250°C. The GC-MS was used in CI (collision ionisation) mode with NH_3 as reagent gas.

5.7. Chemical synthesis

5.7.1. Synthesis of (1S)-1-phenyl-but-3-en-1-ol (**8**)

To a solution of (−)-β-methoxy-diisopinylcampheylborane (4.68 g, 14.8 mmol, 1 eq) in diethyl ether (12 ml), at −78°C allylmagnesium bromide (14.8 ml of 1 M soln. in diethyl ether, 14.8 mmol, 1 eq) was added dropwise with stirring. The solution was stirred for 15 min at −78°C, then allowed to warm to room temperature over 1 h. The solution was re-cooled to −78°C and benzaldehyde (1.64 ml, 14.8 mmol, 1 eq) added. This solution was stirred at −78°C for 1 h, and then at room temperature for 1 h. 3 M sodium hydroxide (11 ml) and hydrogen peroxide (4.5 ml 30% soln.) were added and the solution refluxed for 1 h. The organic layer was separated and washed with water (15 ml), brine (15 ml) and dried ($MgSO_4$). The solvent was removed under reduced pressure and the crude product subjected to distillation and silica column chromatography (DCM 50% in 40–60 petroleum ether), total separation was not achieved, and the impure product product **8** was carried through to the next step.

δ_H(250 MHz, $CDCl_3$) 7.35–7.26 (5H, m, Ph), 5.79 (1H, ddt, *J* 7.0, 10.2, 17.1 Hz, 2-H), 5.14 (2H, m, 2×1-H), 4.69 (1H, t, *J* 6.5 Hz, 4-H), 2.49 (2H, dd, *J* 6.7, 6.7 Hz, 3-H), 1.22–0.90 (impurities); **δ_C(62.5 MHz, $CDCl_3$)** 144.0 (C-5), 134.6 (C-2), 128.4, 127.5, 125.9 (C-6, C-7, C-8, C-9, C-10), 118.3 (C-1), 73.4 (C-4), 43.8 (C-3).

5.7.2. Synthesis of 4-[(tert-butyldimethylsilyl)oxy]-4-phenyl-but-1-en (**9**)

To a solution of (1*S*)-1-phenyl-but-3-en-1-ol (**8**) (0.3 g, impure approx. 2.0 mmol, 1 eq) in DCM (20 ml) at −78°C were added diisopropylethylamine (0.7 ml, 4.0 mmol, 2 eq) and *tert*-butyldimethylsilyl triflate (0.5 ml, 6.8 mmol, 1.1 eq) successively. The resultant solution was stirred until the reaction was shown to be complete by TLC (2 h), and the volatiles were removed under reduced pressure. The crude product was purified by flash column chromatography (DCM) to yield **9** (0.51 g, 97%).

δ_H(500 MHz, $CDCl_3$) 7.31–7.24 (5H, m, Ph), 5.77 (1H, m, 2-H), 5.04 (1H, m, 1-H), 4.98 (1H, m, 1-H), 4.68 (1H, dd, *J* 5.3, 7.1 Hz, 4-H), 2.43 (2H, m, 3-H), 0.89 (9H, s, *t*-Bu), 0.03 (3H, s, 1×Si-CH_3), −0.13 (3H, s, 1 x Si-CH_3), 1.19–1.02 (impurities); **δ_C(100 MHz, $CDCl_3$)** 145.1 (C-5), 135.3 (C-2), 128.0, 126.9, 125.9 (C-6, C-7, C-8, C-9, C-10), 116.8 (C-1), 75.0 (C-4), 45.5 (C-3), 25.8 (Si-C(CH_3)$_3$), 18.2 (Si-C(CH$_3$)$_3$), −4.7 (Si-CH_3), −4.9 (Si-CH_3).

*5.7.3. Synthesis of (3S)-3-[(tert-butyldimethylsilyl)oxy]-3-phenyl-propionaldehyde (**10**)*

To a solution of **9** (0.6 g, 2.3 mmol, 1 eq) in dioxane/water (12 ml, 5:1 dioxane:water) were added sodium periodate (1.81 g, 8.5 mmol, 3.7 eq) and a catalytic quantity of osmium tetroxide (0.292 g of 1% weight on polymer support, 0.005 eq). After 5 h no reaction was observed, and a further portion of sodium periodate was added (1.81 g, 8.5 mmol, 3.7 eq). The reaction was stirred for 14 h at room temperature, then filtered, and the filtrate extracted with diethyl ether (200 ml). The organic solvents were removed under reduced pressure and the crude product purified by silica column chromatography (15% ethyl acetate in 40–60 petroleum ether). To give **10** (0.21 g, 0.8 mmol, 35%). The acid was detected as a major byproduct.

MS (ESI) *m/z* $C_{15}H_{24}O_2Na$ $[M+Na]^+$, calculated 287.1437, found 287.1438. **δ_H(500 MHz, $CDCl_3$)** 9.79 (1H, t, *J* 2.3 Hz, 1-H), 7.35–7.31 (5H, m, Ph), 5.21 (1H, dd, *J* 4.2, 8.1 Hz, 3-H), 2.85 (1H, ddd, *J* 2.7, 8.1, 15.7 Hz, 2-H), 2.62 (1H, ddd, *J* 2.0, 4.2, 15.7 Hz, 2-H), 0.87 (9H, s, *t*-Bu), 0.04 (3H, s, 1 x Si-CH_3), −0.14 (3H, s, 1×Si-CH_3), 1.22–1.14 (impurities); **δ_C(100 MHz, $CDCl_3$)** 201.4 (C-1), 143.7 (C-4), 128.4, 127.6, 125.7 (C-5, C-6, C-7, C-8, C-9), 70.7 (C-3), 54.0 (C-2), 25.7 (Si-C(CH_3)$_3$), 18.1 (Si-*C*(CH_3)$_3$), −4.7 (Si-CH_3), −4.9 (Si-CH_3).

*5.7.4. Synthesis of (4R)-3-[(2′R,3′S,4′R,5′S)-3′-hydroxy-2′,4′-methyl-5′-[(tert-butyldimethylsilyl)oxy]-5′-phenylpentanoyl]-4-phenylmethyl-2-oxazolidinone (**11**)*

Dibutylboron triflate (0.26 ml, 1.0 mmol, 1.25 eq) was added dropwise to a solution of (4*R*)-3-propionyl-4-phenylmethyl-2-oxazolidinone (0.22 g, 0.9 mmol, 1.1 eq) in DCM (5 ml, dry) at 0°C. To this solution triethylamine (0.16 ml, 1.2 mmol, 1.5 eq) was added dropwise. The solution was stirred for 20 min at −78°C, then (3*S*)-[(*tert*-butyldimethylsilyl)oxy]-3-phenylpropionaldehyde (**10**) (0.20 g, 0.8 mmol, 1 eq) was added. This solution was stirred at −78°C for 20 min and then at 0°C for a further 2 h. The reaction was quenched by addition of pH 7 phosphate buffer (1 ml) and methanol (3 ml). A mixture of methanol and aqueous hydrogen peroxide (2:1, 3 ml 27 wt% H_2O_2) was added slowly dropwise. The internal temperature of the reaction flask was maintained below 10°C by cooling with an ice bath. The solution (cloudy) was stirred for 1 h and the volatiles removed under reduced pressure (25–30°C). The resulting slurry was extracted with diethyl ether (3×20 ml). Combined organic extracts were washed with 5% sodium sulphite solution (40 ml, 2 M) until all peroxides were destroyed. The aqueous layer was extracted with ether (20 ml) and the resulting organic extracts combined, dried ($MgSO_4$), and concentrated under reduced pressure. The resulting oil was subjected to flash column chromatography (15% ethyl acetate in 40–60 petroleum ether) to yield an impure yellow oil (**11**) (0.27 g). The impure product was taken on to the next stage in the synthesis.

MS (ESI) *m/z* $C_{28}H_{39}O_5NaNSi$ $[M+Na]^+$, calculated 520.2505, found 522.2490.

*5.7.5. Synthesis of 4-hydroxy-3-methyl-6-phenyl-tetrahydropyran-2-one (**5**)*

Aqueous hydrogen peroxide (0.19 ml, 27 wt% in H_2O, 1.5 mmol, 4.5 eq), and lithium hydroxide, (58 mg, 2.4 mmol, 7.1 eq) were added to a solution of (4*R*)-3-[(2′*R*,3′*S*,4′*R*,5′*S*)-3′-hydroxy-2′,4′-methyl-5′-[(*tert*-butyldimethylsilyl)oxy]-5′-phenylpentanoyl]-4-phenylmethyl-2-oxazolidinone (**11**) (0.17 g, 0.34 mmol, 1 eq) in a mixture of tetrahydrofuran and water (3:1, 10 ml) at 0°C. The resulting solution was stirred at room temperature for 20 min, and then at 40°C until the reaction was complete by TLC (4 h). The mixture was then cooled to 0°C and quenched with sodium sulphite (3.3 ml, 1.5 M, 15 eq). The volatile solvents were removed under reduced pressure, and the aqueous slurry extracted with DCM (3×20 ml). The organic extracts were combined and the solvent removed under reduced pressure. The chiral auxiliary was separated from the acid by flash column chromatography (10% ethyl acetate in 40–60 petroleum ether).

The resultant acid (4 mg, 0.02 mmol) was stirred in a solution of HCl (aq)/THF (10 ml, pH 4 HCl (aq)) and THF (2 ml) at room temperature for 18 h, 40°C for 4 h, then 80°C for 4 h, when the reaction was shown to have gone to completion by TLC. The solution was extracted with ethyl acetate (3×10 ml). The organic extracts were combined, washed with brine, dried ($MgSO_4$), and the solvent removed under reduced pressure, to yield the lactone (**5**) as a colourless oil (2 mg, 0.1 mmol, 50% over final step).

MS (CI-NH_4) 224 (M+18); **δ_H(250 MHz, $CDCl_3$)** 7.43–7.25 (5H, m, Ph), 5.26 (1H, dd, *J* 3.1, 12.0 Hz, 5-H), 3.95 (1H, dddd, *J* 4.0, 4.7, 9.8, 11.0 Hz, 3-H), 2.50 (1H, qd, *J* 7.1, 9.8 Hz, 2-H), 2.46 (1H, ddd, J 3.1, 4.0 13.5 Hz,1 x 4-H), 2.03 (1H, ddd, *J* 11.0, 12.0, 13.5 Hz, 1×4-H), 1.48 (3H, d, *J* 7.1 Hz, 2-CH_3).

*5.7.6. Synthesis of (2R,3S,4R,5S)-3,5-dihydroxy-2,4-dimethyl-5-phenyl-n-pentanoic acid-δ-lactone (**6**)*

This was as described in Weissman et al. [58].

*5.7.7. Synthesis of 2′-(acetylamino)ethylbenzoylthioate (**7**)*

4-Dimethylaminopyridine (97 mg, 0.8 mmol, 0.2 eq) was added to a solution of benzoic acid (0.50 g, 4.0 mmol, 1 eq) in dry DCM (50 ml), followed by NAC (0.64 ml, 6.0 mmol, 1.5 eq) and 1,3-dicyclohexylcarbodiimide (4.4 ml of a 1 M solution in DCM, 4.4 mmol, 1.1 eq) at 0°C. The resulting solution was stirred at 0°C for 10 min and then 3 days at room temperature. The solution was filtered through Celite and eluted with ethyl acetate. After filtration through copper sulphate-impregnated silica, the crude product was purified by recrystallisation from ethyl acetate, precipitated with 40–60 petroleum ether to provide a white crystalline product (**7**) (0.79 g, 3 mmol, 75%).

m.p. 84°C; **ν_{max} (thin film)/cm^{-1}** 1660 (C=O), 1552 (NH); **δ_H(250 MHz, $CDCl_3$)** 7.94–7.36 (5H, m, Ph), 6.47 (1H, s, NH), 3.48 (2H, dt, *J* 6.2, 6.2 Hz, CH_2N), 3.18 (2H, t, *J* 6.2 Hz, CH_2S), 1.93 (3H, s, $COCH_3$); **δ_C(62.5 MHz, $CDCl_3$)** 191.3 (C-1), 171.0 (NCO), 136.8 (C-2), 133.6 (C-5), 128.7, 127.2 (C-3, C-4, C-6, C-7), 39.4 (CH_2N), 28.4 (CH_2S), 23.1 ($COCH_3$).

Acknowledgements

We thank Professor Caroline Harwood (Department of Microbiology, University of Iowa, Iowa City, IA, USA) for the kind gift of a clone of the *badA* gene, Jesus Cortés and Marc Roddis for helpful discussions, John Lester and Kate Pennock for DNA sequence analysis and Hui Hong for assistance with GC-MS analysis. This research was funded by grants from the Biotechnology and Biological Sciences Research Council to P.F.L. and J.S. C.J.W. and

E.J.F. were supported by studentships from the Biotechnology and Biological Sciences Research Council.

References

[1] J. Cortés, S.F. Haydock, G.A. Roberts, D.J. Bevitt, P.F. Leadlay, An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*, Nature 348 (1990) 176–178.

[2] S. Donadio, M.J. Staver, J.B. McAlpine, S.J. Swanson, L. Katz, Modular organization of genes required for complex polyketide biosynthesis, Science 252 (1991) 675–679.

[3] D.J. Bevitt, J. Cortés, S.F. Haydock, P.F. Leadlay, 6-Deoxyerythronolide B synthase 2 from *Saccharopolyspora erythraea*. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme, Eur. J. Biochem. 204 (1992) 39–49.

[4] T. Schwecke, J.F. Aparicio, I. Molnár, A. König, L.E. Khaw, S.F. Haydock, M. Oliynyk, P. Caffrey, J. Cortés, J.B. Lester, G.A. Böhm, J. Staunton, P.F. Leadlay, The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin, Proc. Natl. Acad. Sci. USA 92 (1995) 7839–7843.

[5] J.F. Aparicio, I. Molnár, T. Schwecke, A. König, S.F. Haydock, L.E. Khaw, J. Staunton, P.F. Leadlay, Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of the enzymatic domains in the modular polyketide synthase, Gene 169 (1996) 9–16.

[6] S. Kuhstoss, M. Huber, J.R. Turner, J.W. Paschal, R.N. Rao, Production of a novel polyketide through the construction of a hybrid polyketide synthase, Gene 183 (1996) 231–236.

[7] S.J. Kakavas, L. Katz, D. Stassi, Identification and characterization of the niddamycin polyketide synthase genes from *Streptomyces caelestis*, J. Bacteriol. 179 (1997) 7515–7522.

[8] H. Ikeda, T. Nonomiya, M. Usami, T. Ohta, S. Omura, Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*, Proc. Natl. Acad. Sci. USA 96 (1999) 9509–9514.

[9] J.F. Aparicio, A.J. Colina, E. Ceballos, J.F. Martin, The biosynthetic gene cluster for the 26-membered ring polyene macrolide pimaricin – a new polyketide synthase organization encoded by two subclusters separated by functionalization genes, J. Biol. Chem. 274 (1999) 10133–10139.

[10] S. Zotchev, K. Haugan, O. Sekurova, H. Sletta, T.E. Ellingsen, S. Valla, Identification of a gene cluster for antibacterial polyketide-derived antibiotic biosynthesis in the nystatin producer *Streptomyces noursei* ATCC 11455, Microbiology 146 (2000) 611–619.

[11] P. Caffrey, S. Lynch, E. Flood, S. Finnan, M. Oliynyk, Amphotericin biosynthesis in *Streptomyces nodosus*: deductions from analysis of polyketide synthase and late genes, Chem. Biol. 8 (2001) 713–723.

[12] A. König, T. Schwecke, I. Molnar, G.A. Bohm, P.A.S. Lowden, J. Staunton, P.F. Leadlay, The pipecolate-incorporating enzyme for the biosynthesis of the immunosuppressant rapamycin - nucleotide sequence analysis, disruption and heterologous expression of *rapP* from *Streptomyces hygroscopicus*, Eur. J. Biochem. 247 (1997) 526–534.

[13] D.A. Hopwood, Genetic contributions to understanding polyketide synthases, Chem. Rev. 97 (1997) 2465–2497.

[14] C.A. Whatling, J.E. Hodgson, M.K.R. Burnham, N.J. Clarke, F.C.H. Franklin, C.M. Thomas, Identification of a 60-kb region of the chromosome of *Pseudomonas fluorescens* NCIB-10586 required for the biosynthesis of pseudomonic acid (mupirocin), Microbiology 141 (1995) 973–982.

[15] S. Beyer, B. Kunze, B. Silakowski, R. Muller, Metabolic diversity in myxobacteria: identification of the myxalamid and the stigmatellin biosynthetic gene cluster of *Stigmatella aurantiaca* Sg a15 and a combined polyketide-(poly)peptide gene cluster from the epothilone producing strain *Sorangium cellulosum* So ce 90, Biochim. Biophys. Acta 1445 (1999) 185–195.

[16] B. Silakowski, H.U. Schairer, H. Ehret, B. Kunze, S. Weinig, G. Nordsiek, P. Brandt, H. Blocker, G. Hofle, S. Beyer, R. Muller, New lessons of combinatorial biosynthesis from myxobacteria – The myxothiazol biosynthetic gene cluster of *Stigmatella aurantiaca* DW4/3-1, J. Biol. Chem. 274 (1999) 37391–37399.

[17] J.M. Ligon, T. Schupp, J.J. Beck, D.S. Hill, S. Neff, J.A. Ryal, Genes for the biosynthesis of soraphen. United States Patent No. 5716849 (1998).

[18] I. Molnár, T. Schupp, M. Ono, R.E. Zirkle, M. Milnamow, B. Nowak-Thompson, N. Engel, C. Toupet, A. Stratmann, D.D. Cyr, J. Gorlach, J.M. Mayo, A. Hu, S. Goff, J. Schmid, J.M. Ligon, The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So ce 90, Chem. Biol. 7 (2000) 97–109.

[19] B. Julien, S. Shah, R. Ziermann, R. Goldman, L. Katz, C. Khosla, Isolation and characterization of the epothilone biosynthetic gene cluster from *Sorangium cellulosum*, Gene 249 (2000) 153–160.

[20] J. Staunton, B. Wilkinson, Combinatorial biosynthesis of polyketides and nonribosomal peptides, Curr. Opin. Chem. Biol. 5 (2001) 159–164.

[21] A.F.A. Marsden, B. Wilkinson, J. Cortés, N.J. Dunster, J. Staunton, P.F. Leadlay, Engineering broader specificity into an antibiotic-producing polyketide synthase, Science 279 (1998) 199–202.

[22] M.S. Pacey, J.P. Dirlam, R.W. Geldart, P.F. Leadlay, H.A.I. McArthur, E.L. McCormick, R.A.O. Monday, T.N. Connell, J. Staunton, T.J. Winchester, Novel erythromycins from a recombinant *Saccharopolyspora erythraea* strain NRRL 2338 pIG1 - I. Fermentation, isolation and biological activity, J. Antibiot. 51 (1998) 1029–1034.

[23] K.E.H. Wiesmann, J. Cortés, M.J.B. Brown, A.I. Cutter, J. Staunton, P.F. Leadlay, Polyketide synthesis *in vitro* on a modular polyketide synthase, Chem. Biol. 2 (1995) 583–589.

[24] C. Bisang, P.F. Long, J. Cortes, J. Westcott, J. Crosby, A.L. Matharu, R.J. Cox, T.J. Simpson, J. Staunton, P.F. Leadlay, A chain initiation factor common to both modular and aromatic polyketide synthases, Nature 401 (1999) 502–505.

[25] P.R. August, L. Tang, Y.J. Yoon, S. Ning, R. Muller, T.W. Yu, M. Taylor, D. Hoffmann, C.G. Kim, X.H. Zhang, C.R. Hutchinson, H.G. Floss, Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amycolatopsis mediterranei* S699, Chem. Biol. 5 (1998) 69–79.

[26] S.J. Admiraal, C.T. Walsh, C. Khosla, The loading module of rifamycin synthetase is an adenylation-thiolation didomain with substrate tolerance for substituted benzoates, Biochemistry 40 (2001) 6116–6123.

[27] L. Serre, E.C. Verbree, Z. Dauter, A.R. Stuitje, Z.S. Derewenda, The *Escherichia coli* malonyl-CoA:acyl carrier protein transacylase at 1.5-angstrom resolution – crystal structure of a fatty acid synthase component, J. Biol. Chem. 270 (1995) 12961–12964.

[28] V.S. Rangan, S. Smith, Alteration of the substrate specificity of the malonyl-CoA/acetyl-CoA:acyl carrier protein *S*-acyltransferase domain of the multifunctional fatty acid synthase by mutation of a single arginine residue, J. Biol. Chem. 272 (1997) 11975–11978.

[29] C.J. Dutton, S.P. Gibson, A.C. Goudie, K.S. Holdom, M.S. Pacey, J.C. Ruddock, J.D. Bu'Lock, M.K. Richards, Novel avermectins produced by mutational biosynthesis, J. Antibiot. 44 (1991) 357–365.

[30] J. Lau, D.E. Cane, C. Khosla, Substrate specificity of the loading didomain of the erythromycin polyketide synthase, Biochemistry 39 (2000) 10514–10520.

[31] T. Schupp, C. Toupet, B. Cluzel, S. Neff, S. Hill, J. Beck, J.M. Ligon, A *Sorangium cellulosum* (Myxobacterium) gene cluster for the biosynthesis of the macrolide antibiotic Soraphen A – cloning, characterization, and homology to polyketide synthase genes from Actinomycetes, J. Bacteriol. 177 (1995) 3673–3679.

[32] M. Oliynyk, M.J.B. Brown, J. Cortés, J. Staunton, P.F. Leadlay, A hybrid modular polyketide synthase obtained by domain swapping, Chem. Biol. 3 (1996) 833–839.
[33] C.J. Rowe, J. Cortés, S. Gaisser, J. Staunton, P.F. Leadlay, Construction of new vectors for high-level expression in Actinomycetes, Gene 216 (1998) 215–223.
[34] P.G. Egland, D.A. Pelletier, M. Dispensa, J. Gibson, C.S. Harwood, A cluster of bacterial genes for anaerobic benzene ring biodegradation, Proc. Natl. Acad. Sci. USA 94 (1997) 6484–6489.
[35] P.G. Egland, J. Gibson, C.S. Harwood, Benzoate-coenzyme A ligase, encoded by *badA*, is one of three ligases able to catalyze benzoyl-coenzyme A formation during anaerobic growth of *Rhodopseudomonas palustris* on benzoate, J. Bacteriol. 177 (1995) 6545–6551.
[36] M.J. Bibb, G.R. Jansse, J.M. Ward, Cloning and analysis of the promoter region of the erythromycin resistance gene (*ermE*) of *Streptomyces erythraeus*, Gene 38 (1985) 215–226.
[37] J.F. Aparicio, P. Caffrey, A.F.A. Marsden, J. Staunton, P.F. Leadlay, Limited proteolysis and active site studies of the first multienzyme component of the erythromycin-producing polyketide synthase, J. Biol. Chem. 269 (1994) 8524–8528.
[38] J. Staunton, A.C. Sutkowski, The polyketide synthase (PKS) of aspyrone biosynthesis – evidence for the enzyme bound intermediates from incorporation studies with N-acetylcysteamine thioesters in intact cells of *Aspergillus melleus*, J. Chem. Soc. Chem. Commun. (1991) 1110–1112.
[39] D.E. Cane, R.H. Lambalot, P.C. Prabhakaran, W.R. Ott, Macrolide biosynthesis. 7. Incorporation of polyketide chain elongation intermediates into methymycin, J. Am. Chem. Soc. 115 (1993) 522–526.
[40] S.W. Brobst, C.A. Townsend, The potential role of fatty acid initiation in the biosynthesis of the fungal aromatic polyketide aflatoxin B-1, Can. J. Chem. Rev. Can. Chim. 72 (1994) 200–207.
[41] D.E. Cane, G.G. Luo, C. Khosla, C.M. Kao, L. Katz, Erythromycin biosynthesis – highly efficient incorporation of polyketide chain elongation intermediates into 6-deoxyerythronolide B in an engineered *Streptomyces* host, J. Antibiot. 48 (1995) 647–651.
[42] C. Hertweck, B.S. Moore, A plant-like biosynthesis of benzoyl-CoA in the marine bacterium '*Streptomyces maritimus*', Tetrahedron 56 (2000) 9115–9120.
[43] J. Piel, C. Hertweck, P.R. Shipley, D.M. Hunt, M.S. Newman, B.S. Moore, Cloning, sequencing and analysis of the enterocin biosynthesis gene cluster from the marine isolate '*Streptomyces maritimus*': evidence for the derailment of an aromatic polyketide synthase, Chem. Biol. 7 (2000) 943–955.
[44] F. Knoop, Der Abbau aromatischer Fettsäuren im Thierkörper, Ernst Kuttruff, Freiburg, 1904.
[45] A. Pereda, R.G. Summers, D.L. Stassi, X.A. Ruan, L. Katz, The loading domain of the erythromycin polyketide synthase is not essential for erythromycin biosynthesis in *Saccharopolyspora erythraea*, Microbiology 144 (1998) 543–553.
[46] J.R. Jacobsen, C.R. Hutchinson, D.E. Cane, C. Khosla, Precursor-directed biosynthesis of erythromycin analogs by an engineered polyketide synthase, Science 277 (1997) 367–369.
[47] B.A. Pfeifer, S.J. Admiraal, H. Gramajo, D.E. Cane, C. Khosla, Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*, Science 291 (2001) 1790–1792.
[48] P.A.S. Lowden, B. Wilkinson, G.A. Bohm, S. Handa, H.G. Floss, P.F. Leadlay, J. Staunton, Origin and true nature of the starter unit for the rapamycin polyketide synthase, Angew. Chem. Int. Ed. 40 (2001) 777–779.
[49] J. Staunton, P. Caffrey, J.F. Aparicio, G.A. Roberts, S.S. Bethell, P.F. Leadlay, Evidence for a double-helical structure for modular polyketide synthases, Nature Struct. Biol. 3 (1996) 188–192.
[50] D.E. Cane, C.T. Walsh, The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases, Chem. Biol. 6 (1999) 319–325.
[51] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.
[52] W.J. Dower, J.F. Miller, C.W. Ragsdale, High efficiency transformation of *E. coli* by high voltage electroporation, Nucleic Acids Res. 16 (1988) 6127–6145.
[53] C. Yanisch-Perron, J. Vieira, J. Messing, Improved M13 phage cloning vectors and host strains – nucleotide sequences of the M13MP18 and pUC19 vectors, Gene 33 (1985) 103–109.
[54] M. Bierman, R. Logan, K. O'Brien, E.T. Seno, R. Nagaraja Rao, B.E. Schoner, Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp, Gene 116 (1992) 43–49.
[55] H. Yamamoto, K.H. Maurer, C.R. Hutchinson, Transformation of *Streptomyces erythraeus*, J. Antibiot. 39 (1986) 1304–1313.
[56] S. Gaisser, G.A. Böhm, J. Cortés, P.F. Leadlay, Analysis of seven genes from the *eryAI*/*eryK* region of the erythromycin biosynthetic gene cluster in *Saccharopolyspora erythraea*, Mol. Gen. Genet. 256 (1997) 239–251.
[57] T. Kieser, M.J. Bibb, M.J. Buttner, K.F. Chater, D.A. Hopwood, Practical *Streptomyces* Genetics, 2nd edn., John Innes Foundation, Norwich, 2000.
[58] K.J. Weissmann, M. Bycroft, A.L. Cutter, U. Hanefeld, E.J. Frost, M.C. Timoney, R. Harris, S. Handa, M. Roddis, J. Staunton, P.F. Leadlay, Evaluating precursor-directed biosynthesis towards novel erythromycins through in vitro studies on a bimodular polyketide synthase, Chem. Biol. 5 (1995) 743–754.
[59] S.F. Haydock, J.F. Aparicio, I. Molnár, T. Schwecke, L.E. Khaw, A. König, A.F.A. Marsden, I.S. Galloway, J. Staunton, P.F. Leadlay, Divergent sequence motifs correlated with the substrate specificity of (methyl)malonyl-CoA:acyl carrier protein transacylase domains in modular polyketide synthases, FEBS Lett. 374 (1995) 246–248.
[60] J. Cortés, K.E.H. Wiesmann, G.A. Roberts, M.J.B. Brown, J. Staunton, P.F. Leadlay, Repositioning of a domain in a modular polyketide synthase to promote specific chain cleavage, Science 268 (1995) 1487–1489.