ELSEVIER

Chemistry & Biology 8 (2001) 329–340

Chemistry & Biology

www.elsevier.com/locate/chembiol

Research Paper

Molecular basis of Celmer's rules: role of the ketosynthase domain in epimerisation and demonstration that ketoreductase domains can have altered product specificity with unnatural substrates

Ines E. Holzbaur [a], Anand Ranganathan [b,1], Iain P. Thomas [a,b], Dominic J.A. Kearney [a], James A. Reather [a], Brian A.M. Rudd [c], James Staunton [a,*], Peter F. Leadlay [b]

[a] *Cambridge Centre for Molecular Recognition and University Chemical Laboratory, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK*
[b] *Cambridge Centre for Molecular Recognition and Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1GA, UK*
[c] *Bioprocessing Unit, Glaxo Wellcome Research and Development, Medicines Research Centre, Gunnels Wood Road, Stevenage, Herts SG1 2NY, UK*

Received 22 August 2000; revisions requested 17 October 2000; revisions received 8 January 2001; accepted 26 January 2001
First published online 6 March 2001

Abstract

Background: Polyketides are structurally diverse natural products with a range of medically useful activities. Non-aromatic bacterial polyketides are synthesised on modular polyketide synthase multienzymes (PKSs) in which each cycle of chain extension requires a different 'module' of enzymatic activities. Attempts to design and construct modular PKSs that synthesise specified novel polyketides provide a particularly stringent test of our understanding of PKS structure and function.

Results: We show that the ketoreductase (KR) domains of modules 5 and 6 of the erythromycin PKS, housed in the multienzyme subunit DEBS3, exert an unexpectedly low level of stereochemical control in reducing the keto group of a synthetic analogue of the diketide intermediate. This led us to construct a hybrid triketide synthase based on DEBS3 with ketosynthase domain ketosynthase (KS)5 replaced by the loading module and KS1. The construct in vivo produced two major triketide stereoisomers, one expected and one surprising. The latter was of opposite configuration at three out of the four chiral centres: the branching alkyl centre was that produced by KS1 and, surprisingly, both hydroxyl centres produced by the reduction steps carried out by KR5 and KR6 respectively.

Conclusions: These results demonstrate that the epimerising activity associated with module 1 of the erythromycin PKS can be conferred on module 5 merely by transfer of the KS1 domain. Moreover, the normally precise stereochemical control observed in modular PKSs is lost when KR5 and KR6 are challenged by an unfamiliar substrate, which is much smaller than their natural substrates. This observation demonstrates that the stereochemistry of ketoreduction is not necessarily invariant for a given KR domain and underlines the need for mechanistic understanding in designing genetically engineered PKSs to produce novel products.

Keywords: Polyketide synthase; Celmer's rules; Hybrid polyketide synthase; Epimerisation; KR domain

1. Introduction

Rapid progress has been made in recent years in producing altered polyketide metabolites by genetic engineering of the enzymes responsible for their biosynthesis. In the case of erythromycin, for example, the polyketide synthase (PKS) has been modified to produce erythromycins with altered starter acids, related polyketides with missing branching methyl groups and also macrolide structures with different levels of oxidation at some of the oxygenated centres [1–7].

This success in altering structure has been assisted by the modular nature of the proteins responsible for assembly of the macrolide core of erythromycin and similar natural products. In the case of erythromycin, it was shown that the PKS is encoded by three giant genes called

[1] Present address: International Centre for Genetic Engineering and Biotechnology, P.O. Box 10504, Aruna Asaf Ali Marg, New Delhi 110067, India.

* Correspondence: James Staunton;
E-mail: js24@cam.ac.uk

1074-5521/01/$ – see front matter
PII: S1074-5521(01)00014-X

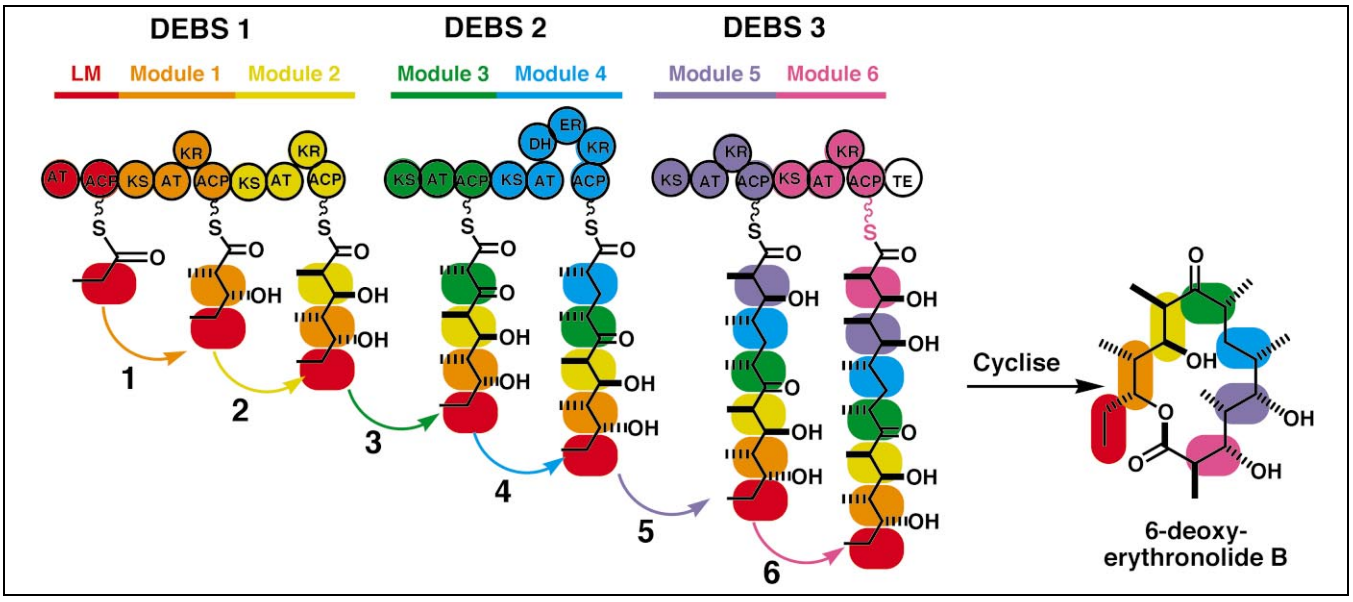

Fig. 1. The erythromycin-producing PKS; primary organisation of the genes and their corresponding protein domains. There are six chain extension modules, each beginning with a KS domain, that are responsible for adding the six successive C_3 units. The loading module (LM) specifies propionate as the starter unit. The completed chain is released as a macrolactome by the TE domain. KR, β-ketoacyl reductase.

ery AI, *ery AII* and *ery AIII* respectively. Each of these is approximately 11 kbp in length and each encodes a giant protein consisting of approximately 3000 amino acids [8,9], which is isolated as a homodimer [9–11]. The erythromycin PKS, therefore, comprises three giant multifunctional proteins, which are called DEBS1, DEBS2 and DEBS3 (Fig. 1). The active sites (domains) housed in these giant proteins are organised into chain extension modules, six in total, one for each chain extension step involved in assembly of the macrolide core. The first module is preceded by a bidomain that selects and loads a propionate from the pool of intracellular propionyl CoA. The last (sixth) module in DEBS3 is followed by a thioesterase domain (TE), which releases the complete polyketide chain, presumably by macrolide ring formation.

There is a direct correlation in the erythromycin PKS between the domains present in each chain extension module and the chemical processes needed to develop the structure of the corresponding extension unit in the polyketide chain. This correspondence can be understood on the basis of the known chemistry of fatty acid biosynthesis which proceeds by a series of repeated cycles of condensa-

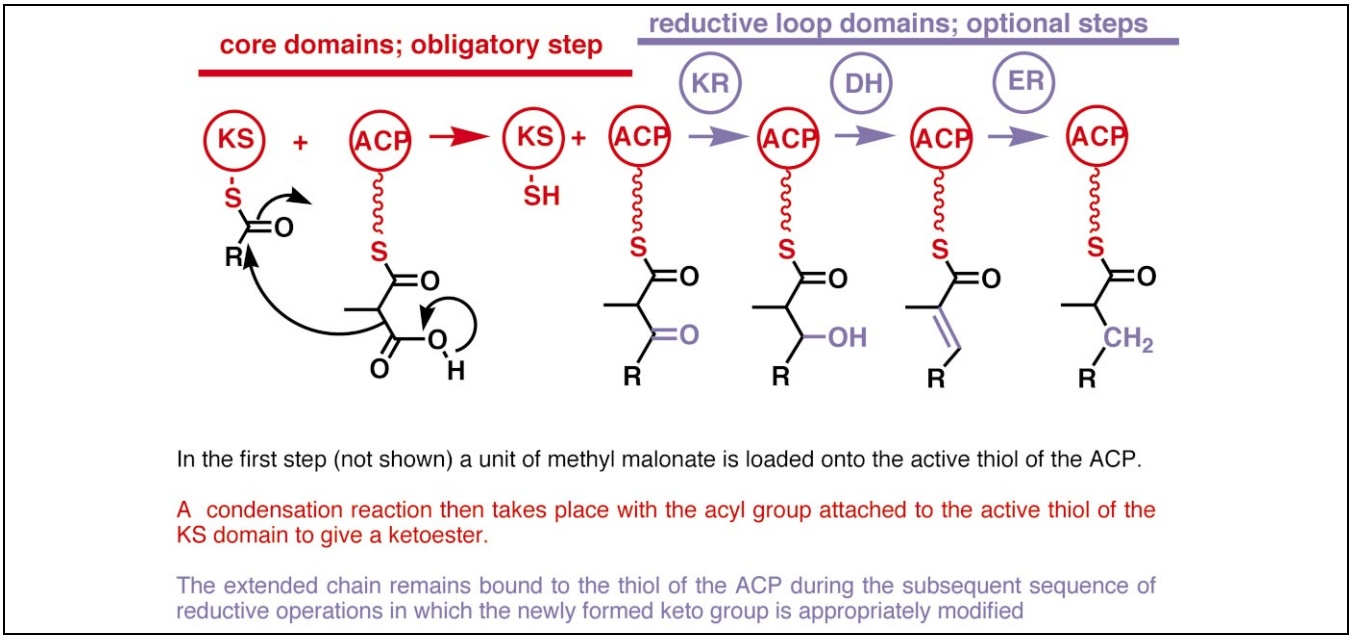

Fig. 2.

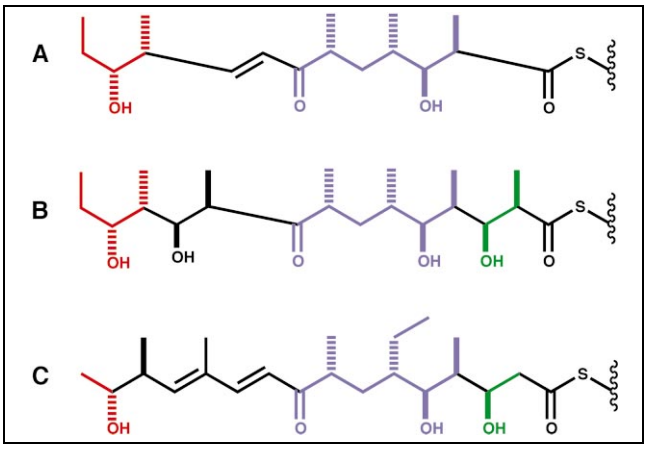

Fig. 3. Comparison of the open chain of the polyketide cores of the related macrolides: (A) methymycin (12 membered macrolide); (B) erythromycin (14-membered macrolide) and; (C) tylosin (16-membered macrolide).

tion and reduction (Fig. 2) in which the initially formed keto group is converted to a methylene in every chain extension cycle. In a modular PKS, however, the extent of reductive modification can be truncated in modules, which lack one or more of the relevant domains. For example, if the enoyl reductase (ER) domain is missing, the chain extension product is a double bond; if both ER and dehydratase (DH) are missing, as in modules 2, 5 and 6 of the erythromycin PKS then a hydroxyl group is formed in that chain extension unit; and if all three reductive domains (DH, ER and KR) are missing, for example in module 3 of the erythromycin PKS, a carbonyl group remains in the developed polyketide chain. This correlation makes it possible to predict, with some confidence, the product a given PKS may make. It also gives a rational basis for genetic engineering to alter the oxidation level at appropriate sites of the macrolide ring [12–14]. In contrast, the stereochemistry at carbon centres bearing hydroxyl and alkyl groups produced by a PKS cannot be predicted as yet from the gene product responsible for chain assembly.

In a fully reduced unit, the stereochemistry at the branching alkyl centre is determined by the reduction process carried out by the ER domain. In the case of a hydroxy intermediate the chirality at the hydroxyl group must be dictated by the KR domain. However, in this case and in the case of a unit with a surviving keto group, the agent of stereocontrol at the alkyl branch is not clear. The work described in this paper was designed to contribute to a better mechanistic understanding of stereochemical control in the early steps in each reduction cycle. This in turn might allow rational alterations in PKSs, to produce predicted variations in the stereochemistry at alkyl branching centres.

As early as 1965, Celmer noted that there is a strong, position-specific, structural and stereochemical homology running through families of macrolides of various sizes. Examples are shown in Fig. 3 in which representative 12-, 14- and 16-membered macrolides are compared [15]. This first suggested that there is a genetic origin for stereochemical control at alkyl branching centres and, therefore, that an understanding of how the genetic information is correlated with chemical structure should provide a rational basis for designing modifications in the PKS systems aimed at producing altered stereochemistry in polyketide products.

In a given chain extension cycle when either a propionate or butyrate chain extender is used, the first chiral centre is established in the condensation step at the branching methyl group in the β-keto acyl residue. Subsequently, the keto group is reduced by a KR and a second chiral centre is established. It is notable, in the case of erythromycin, that both configurations of methyl centres are present in the polyketide chain. Referring to Fig. 1, it can be seen that the methyl branching centre created by module 1 has the methyl group (*S*)-configuration (pointing down), whereas the corresponding hydroxy extension units in modules 2, 5 and 6 all have the methyl group in the (*R*)-configuration (pointing up). Therefore, in erythromycin, there must be a way of developing both configurations at methyl branching sites prior to keto group reduction.

Preliminary insights into the mechanism of this stereochemical control have been gained from in vitro studies with a truncated version of the erythromycin PKS called DEBS1-TE [16,17]. In this protein, the erythromycin PKS was engineered so that the TE normally resident at the end of module 6 was attached at the terminus of module 2 (Fig. 4). This gives a self-sufficient single multifunctional protein which is capable of carrying out the first two chain extensions of the erythromycin biosynthetic pathway. The triketide products are released as δ-lactones (**1** and **2**). This model system produces both stereochemical possibilities at the methyl branching centres, and is therefore an ideal model system for probing the molecular mechanisms of stereochemical control. The protein was isolated in pure form and it was shown to produce the δ-lactone product in vitro in good yield when supplied with propionyl CoA, methylmalonyl CoA and NADPH [18,19].

In the first experiments on the mechanism of stereocontrol, it was proved that (*S*)-methylmalonyl CoA was used in both chain extension units [18,20]. In contrast, (*R*)-methylmalonyl CoA was not incorporated. This showed conclusively that the control of stereochemistry at the methyl branching centres was not exercised by selection of the appropriate chiral form of the chain extension building block, as had been proposed on the basis of classical in vivo experiments [21], and re-proposed on the basis of the existence of separate acyltransferase (AT) domains for each cycle of chain extension in DEBS [2,22]. Instead, the control must be exercised in the mechanisms of the chain extension processes themselves.

Subsequently, experiments were carried out with (*S*)-methylmalonyl CoA labelled at the chiral centre with deu-

terium. The deuterium label was retained at the methyl branching chiral centre in the chain extension residue produced by module 2 but was completely lost in the chain extension process carried out by module 1 [23]. These experiments also showed that the chain extension reaction was carried out with inversion of configuration in module 2. It was suggested, to explain the loss of the deuterium from the chiral centre in module 1, that the equivalent condensation step might also take place in the same stereochemical sense with inversion of configuration to give the (2*R*) isomer of the keto ester and that this isomer would then be epimerised to the (2*S*) isomer by enolisation prior to reduction by the KR domain.

These results demonstrate that module 1 holds the key as far as alteration of stereochemistry is concerned in two respects. First, it has to have a mechanism for carrying out the epimerisation of the initially formed (2*R*) isomer of the keto ester to the (2*S*) isomer. Second, some agent must be present which selects only the (2*S*) isomer from the epimerised mixture; this is predicted to be the KR domain. Two further experiments confirmed these postulates. First, a diketide synthase was created in which there was only one chain extension module, which combined KS1 with the reductive loop corresponding to module 2 [24]. This produced a diketide in which the chirality of the methyl was opposite to that produced in module 1. This supports the proposal that the KS1 is capable of delivering both enantiomers of the keto ester for reduction and that KR1 selects the (*S*)-isomer in module 1. Further evidence to support the proposed substrate selectivity of KR1 came from direct studies of the stereochemistry of reduction using a synthetic diketide keto ester analogue as substrate. It was shown that KR1 does, indeed, select from a racemic mixture of the *N*-acetyl cysteamine (NAC) derivative of the keto ester diketide only one enantiomer (2*S*) and it produces the single product corresponding to the stereochemistry of reduction in module 1 [25].

What remains to be established after these experiments is the factor in module 1 that is responsible for the epimerisation process. We suspected that it might be the KS domain but it could have been any of the other domains or even a part of the protein in the reductive loop that has

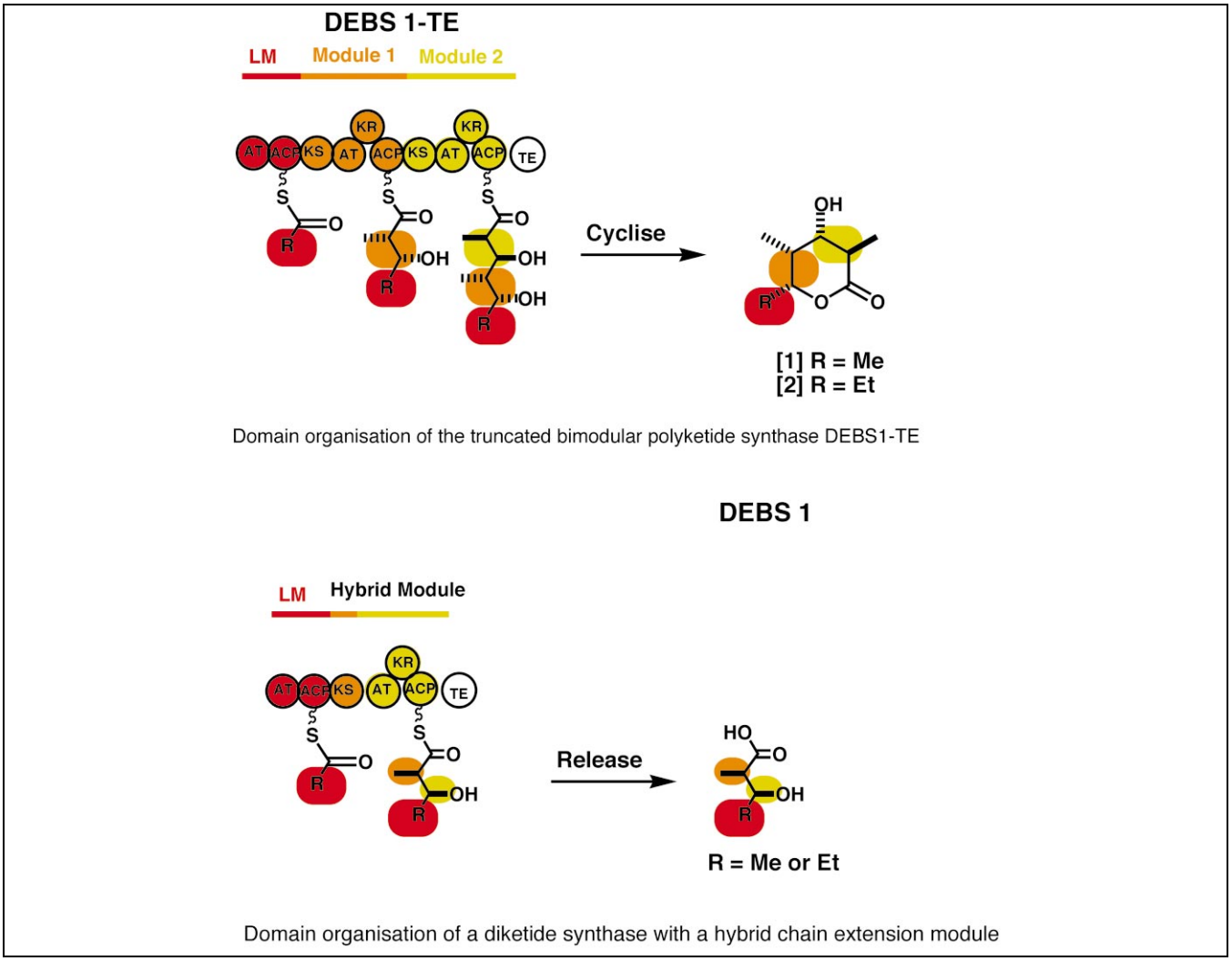

Fig. 4.

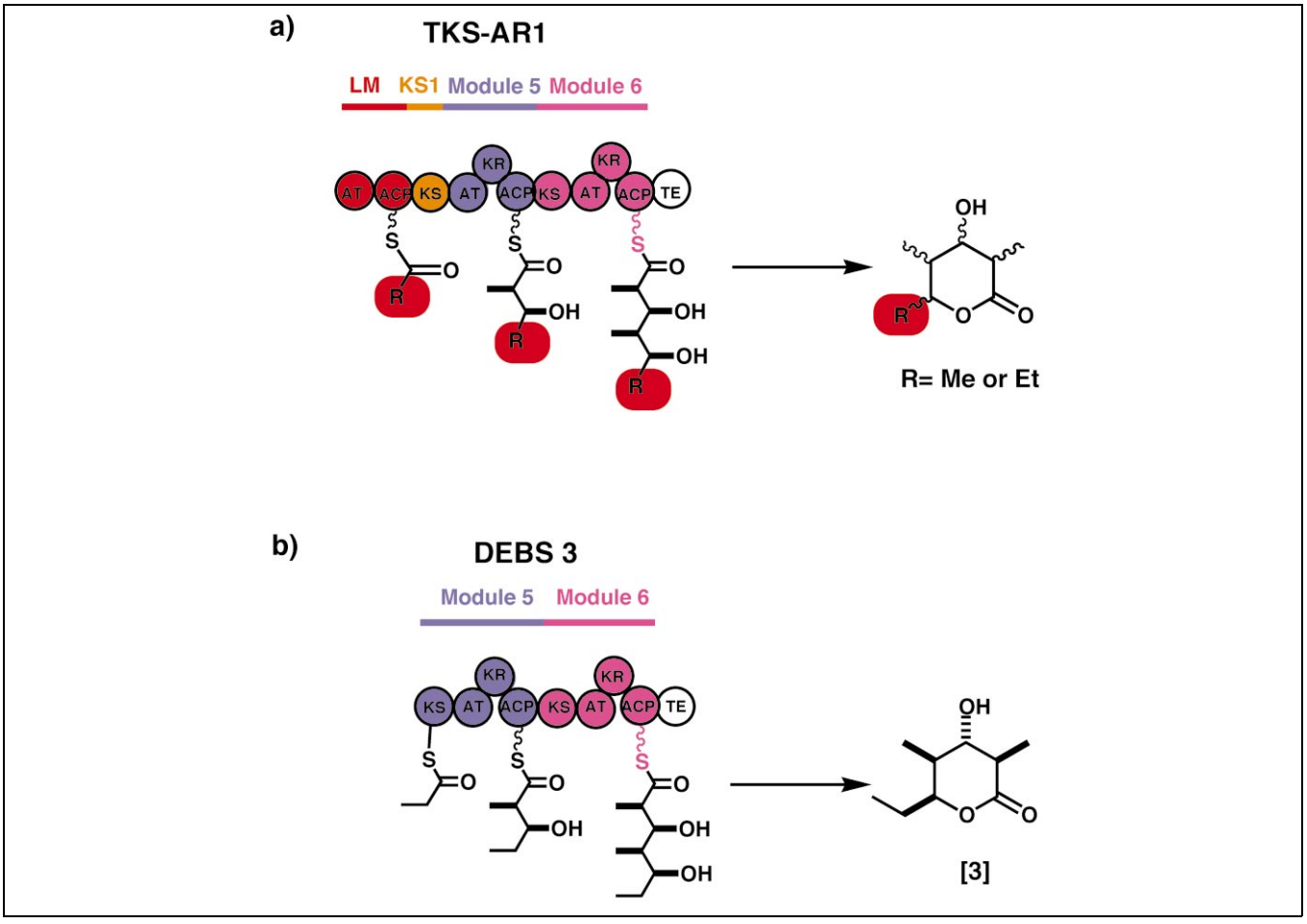

Fig. 5. Arrangement of the domains in: (a) the hybrid PKS TKS–AR1 and; (b) DEBS3.

not been assigned any specific function. In order to test the possibility that KS1 might be the active determinant of epimerisation, we designed a hybrid PKS based on DEBS3 in which the first KS, KS5, has been replaced by KS1, and, to make the construct completely efficient, we also incorporated the DEBS loading module at the front of the truncated PKS. This construct, TKS–AR1, is shown in Fig. 5a. Since chain initiation is carried out by the normal erythromycin starter module, it was anticipated that both acetate and propionate would be used. The stereochemical consequences of replacing KS1 by KS5 could not be predicted for any of the chiral centres. For comparison, the reported pattern of behaviour of modules 5 and 6 in an unmodified version of DEBS3 is shown in Fig. 5b [26]. Since this lacks a competent starter module, it was forced to rely on an aberrant mechanism of chain initiation that specified propionate by decarboxylation of methyl malonate. Only the single enantiomer (**3**) was reported, corresponding to the normal stereochemical operations of modules 5 and 6.

We were guided to this choice of model system TKS–AR1 by the results of experiments to probe the stereoselectivity and stereospecificity of the KR domains of DEBS3. As will be explained later, the results suggested that one or both of these domains shows a low degree of stereospecificity in carrying out reduction of the keto group of the model keto ester. This observation raised the interesting possibility that one or both KRs might accept both stereoisomers of the keto ester form of the diketide intermediate and so provide a very effective approach to localise the epimerisation activity.

2. Results and discussion

2.1. Stereospecificity of keto group reduction by DEBS3

The experiments here followed closely on the earlier experiments with DEBS1-TE [25]. DEBS3 was overexpressed in *Escherichia coli* [27]. The resulting enzyme was not efficiently phosphopantetheinylated on the two acyl carrier protein (ACP) domains and so was not capable of chain extension chemistry. However, the KR domains were expected to be fully competent. The enzyme was incubated in the presence of the NAC analogue of the keto form of the erythromycin PKS diketide under the conditions described earlier to produce reduced diketides (Fig. 6a). The reduction, monitored by consumption of

NADPH, was much slower than the equivalent process with DEBS1-TE. The reduction products were then isolated and analysed by chiral high performance liquid chromatography (HPLC); surprisingly, all four diastereoisomers of the hydroxy diketide analogue were formed in approximately equal amounts (Fig. 6b). Clearly, one or both of the KR domains showed relaxed reaction stereochemistry, either in terms of substrate selection – either (2*R*) or (2*S*) keto ester is an acceptable substrate – or in terms of stereochemistry of reduction – hydride addition to the *re* or *si* face of the carbonyl group.

The low stereoselectivity exhibited by KR5 and KR6 towards the unnatural diketide substrate can be explained in terms of the stereoelectronic imperatives of the reduction process, which are illustrated in Fig. 7. Hydride delivery (from the nicotinamide coenzyme) takes place to the neighbouring face of the carbonyl group. The choice of face is dictated by selective binding of R_1 and R_2 into appropriate pockets. The developing charge on the oxygen requires stabilisation from a strategically placed Lewis acid (e.g. a zinc species). The extent of spatial selectivity and hence the enantiomeric purity of the resulting carbinol centre is determined by the ability of the enzyme active site to discriminate between R_1 and R_2.

This analysis is extended in Fig. 8 to a consideration of the binding of the diketide analogue at the active site of a KR domain. The normal mode of binding of the natural keto ester substrate of KR5 is mode A with hydride delivery from the lower face. In this mode the enantiomer of the diketide is similarly bound. Normally KR5 is not confronted by the 2*S*-configuration at C-2 so there is no reason to expect this site to be able to discriminate between

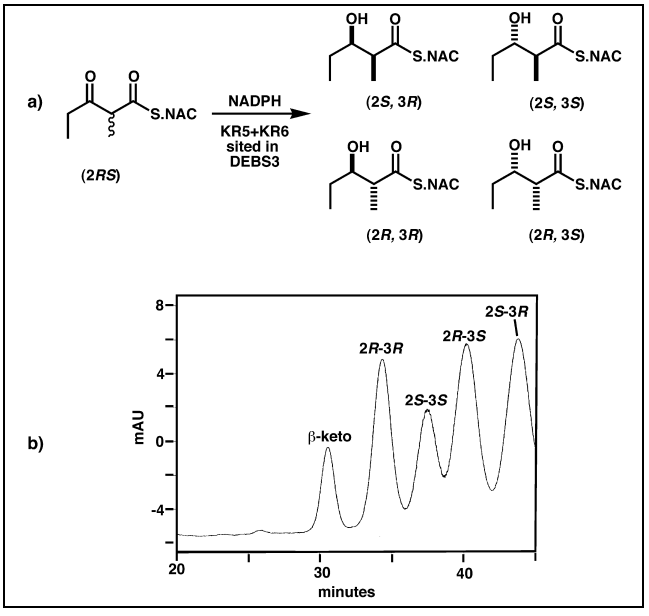

Fig. 6. (a) Incubation of the NAC analogue of the keto form of the diketide intermediate of the erythromycin PKS with DEBS3 yields four reduced products. (b) Chiral HPLC trace of the product mixture produced.

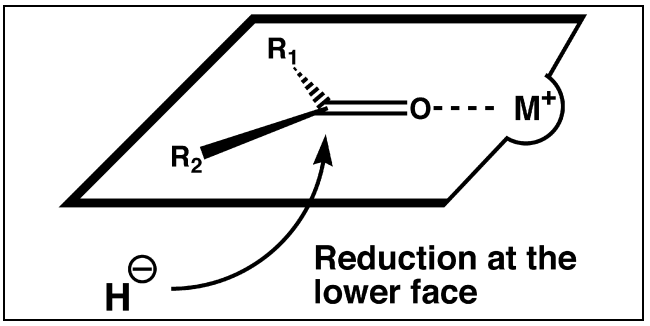

Fig. 7. Stereoelectronics of reduction of a keto group by a KR. The keto group position is determined by coordination to a metal, commonly zinc. Hydride is delivered to either the upper or the lower face depending on the site of binding of the nicotinamide coenzyme (the lower face is arbitrarily selected in this illustration). The stereoselectivity of the reduction process depends on the ability of the active site to discriminate between R_1 and R_2.

modes A and B. The formation of the opposite enantiomer at C-3 can be explained by binding modes C and D in which the direction of binding is reversed. This is equivalent to a loss of ability to discriminate between the groups R_1 and R_2 in Fig. 7. Note that these two groups are similar in the diketide.

This analysis assumes that the thioester residue remains outside the keto ester binding pocket and, therefore, does not play a role in controlling direction of binding. It is possible, however, that distance constraints on the phosphopantetheine group of the ACP would prevent the reverse binding mode. The lack of selectivity in the model experiments may not apply, therefore, when the diketide intermediate is generated on the carrier thiol group of ACP5. So, the planned experiments test not only the ability of KS1 to effect epimerisation in the hybrid module, but also the distance constraints imposed by the flexible arm of the ACP, which carries the growing polyketide chain.

2.2. Design of the hybrid PKS TKS–AR1 based on DEBS3

The triketide synthase used as the test system was essentially DEBS3 fronted by the loading module (AT0+ACP0) from the erythromycin PKS but with KS5 replaced by KS1 (Fig. 5). In order to join this *ery* loading 'tridomain' (consisting of loading module and KS of *ery* module 1) with DEBS3 (minus the KS domain of module 5), a unique *Hin*dIII site was created in the linker region (Fig. 9). A genetically engineered diketide synthase possessing a similar junction (between the loading 'tridomain' and KR domain of *ery* module 2) has been found to be catalytically active [24]. The unique *Hin*dIII site was created in the gene fragment encoding DEBS3 by PCR. The forward and reverse primers for PCR had *Blp*I sites. The fragment was then cloned in pUC-18 vector, following which the 513 bp *Blp*I fragment from *ery* module 5 was replaced by the *Blp*I fragment that contained the engineered *Hin*dIII site, to yield a plasmid that contains the gene encoding DEBS3.

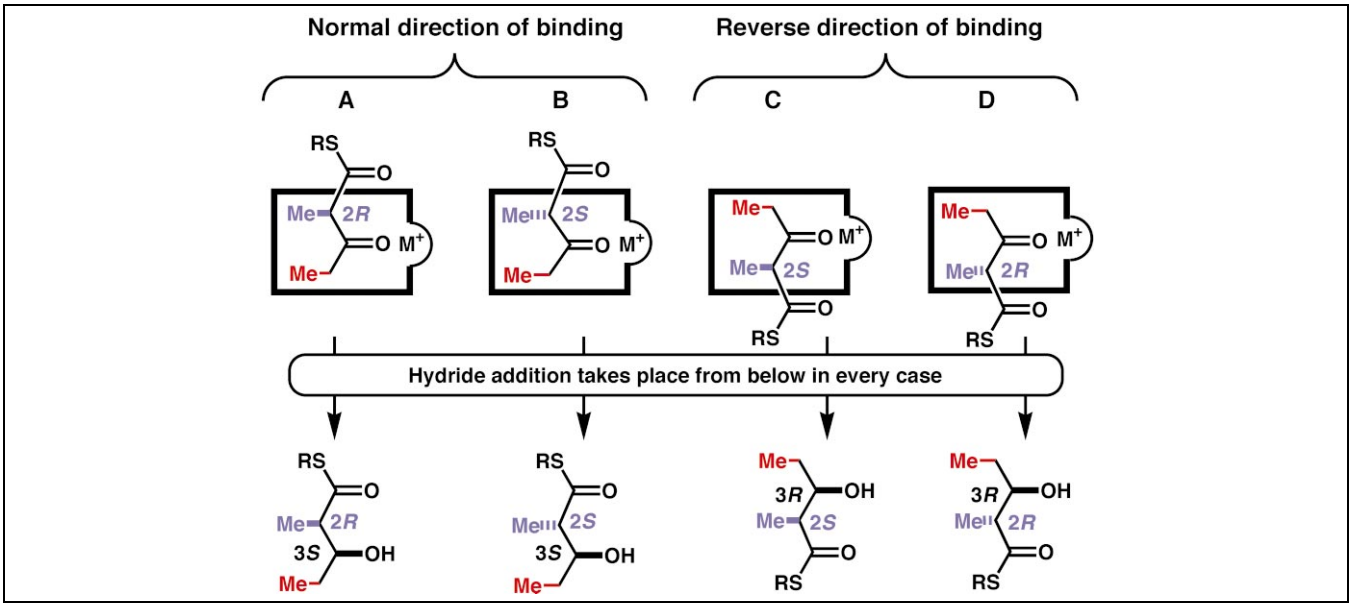

Fig. 8. Alternative binding modes for the ketone carbonyl of the (*R*)- and (*S*)-diketides at the active site of KR5. Hydride reduction takes place at the lower face of the keto group in every case. A combination of all four modes produces all four stereoisomers of the reduction product.

Finally, the *Hin*dIII fragment possessing this hybrid gene was inserted into vector pIB015. This plasmid has the gene that encodes a diketide synthase DKS1-2 (comprised of *ery* loading 'tridomain' joined with *ery* module 2 and *ery* TE) [24]. The resulting plasmid p015-5624 contains the gene that would encode the hybrid PKS TKS–AR1 having the *ery* loading 'tridomain', a truncated *ery* module 5 followed by module 6 and *ery* TE (Fig. 5a; see Section 4 for details of the construction of the expression plasmid p015-5624).

2.3. *Production of the triketide lactones by the hybrid synthase TKS–AR1*

Saccharopolyspora erythraea/JC2 protoplasts were transformed with plasmid pO15-5624 and colonies were selected by their resistance to thiostrepton upon integration of the plasmid into the *S. erythraea* chromosome. *S. erythraea*/JC2 is a mutant strain of the wild-type *S. erythraea*/*NRRL2338* that lacks the entire *ery*-PKS genes except for the TE DNA fragment [28]. Single transformants were picked and grown on 'tap-water medium' plates supplemented with thiostrepton, following which starter cultures of TSB medium were inoculated from *Streptomyces* spore suspensions or from patches on tap water medium plates and growth maintained at 30°C for 24–30 h. Solid medium was used for growth of *S. erythraea*. Typically 500 cm^2 sterile plates containing solid SM3 medium [29] were spread with the starter culture (10 ml TSB) and the plates incubated at 30°C for 10–15 days. The solid medium was then vortexed in water and extracted with ethyl acetate at pH 4.0. Combined organic extracts were washed with sodium bicarbonate solution and evaporated to dryness in vacuo. The resulting solid was resuspended in a small amount of ethyl acetate. The solvent was evaporated and a sample of each extract was analysed by gas chromatography-mass spectrometry (GC-MS; gradient 40–145°C in 23 min). A typical spectrum is shown in Fig. 10. The total ion current spectrum (trace a) showed the typical complex mixture isolated from cultures grown under these conditions. The two lower traces, b and c respectively, show the spectrum edited to select *m*/*z* 190 corresponding to triketide lactone stereoisomers with a propionate starter unit (**3** and **5**) and *m*/*z* 176 corresponding to equivalent compounds with acetate starter units (**4** and **6**). In both edited spectra, all the peaks are in the range of retention times characteristic of triketide lactone products. The two strongest peaks in each spectrum gave ms/ms spectra that were consistent with triketide lactone products. The relative amounts of the products varied from one culture to an-

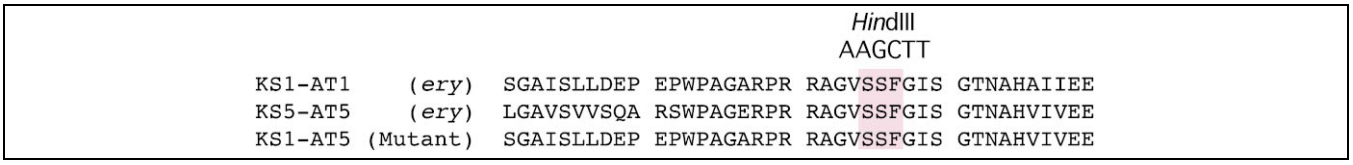

Fig. 9. Amino acid sequence around the unique *Hin*dIII fusion site introduced in the linker regions between the KS and AT domains in *ery* modules 1 and 5.

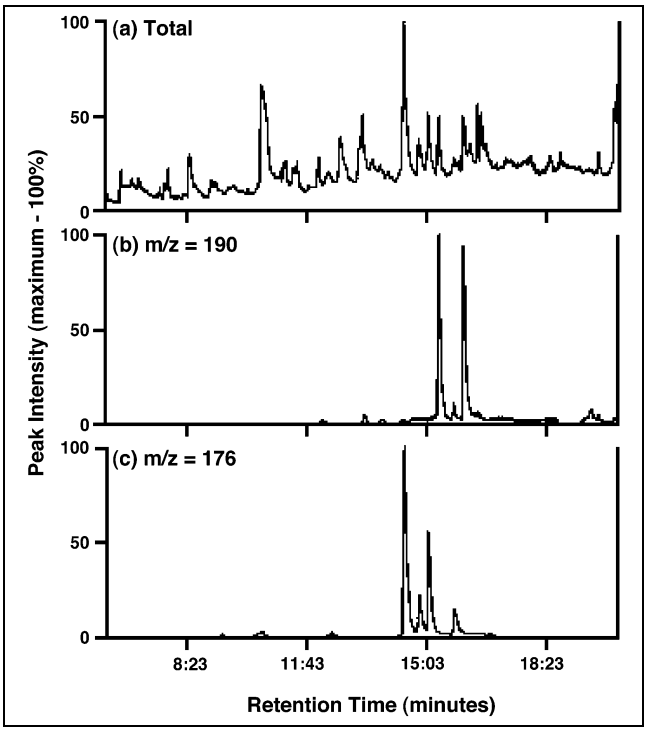

Fig. 10. (a) The total ion current of a typical GC-MS trace of an ethyl acetate extraction of *S. erythraea*/JC2-p015-5624. (b) GC-MS selected for $m/z = 190$ corresponding to compounds **3** and **5**. (c) GC-MS selected for $m/z = 176$ corresponding to compounds **4** and **6**.

other. Usually, the compounds with propionate starter units rather than acetate were the dominant species. The relative amounts of the two main propionate-derived products (**3** and **5**) remained constant at about 1:1. The minor components were not identified but are probably stereoisomers of the major components.

The two main propionate compounds (**3** and **5**) were isolated by preparative HPLC/MS using a VG platform mass spectrometer. The combined extracts from eight plates were loaded onto a column of silica (14 cm × 14 cm) and successive fractions eluted with diethyl ether (400 ml), 25% ethyl acetate in diethyl ether (500 ml), 50% ethyl acetate in diethyl ether (500 ml), 75% ethyl acetate in diethyl ether (400 ml) and ethyl acetate (400 ml). The eluent was collected in 200 ml fractions that were analysed by GC-MS for the presence of triketide lactone. The appropriate fractions were then concentrated under vacuum to dryness and dissolved in 1.5 ml of HPLC grade acetonitrile. To this was added 0.5 ml of Milli-Q water. Batches (500 μl) of the filtered semi-pure extract were then loaded onto a preparative LC system using a reverse phase ODS column (25 cm × 2.2 cm) and eluted using a linear gradient of 5–30% acetonitrile in 45 min. Fractions corresponding to the two major propionate peaks were collected and concentrated to dryness in preparation for NMR analysis.

One of the products (**3**) proved to be the stereoisomer

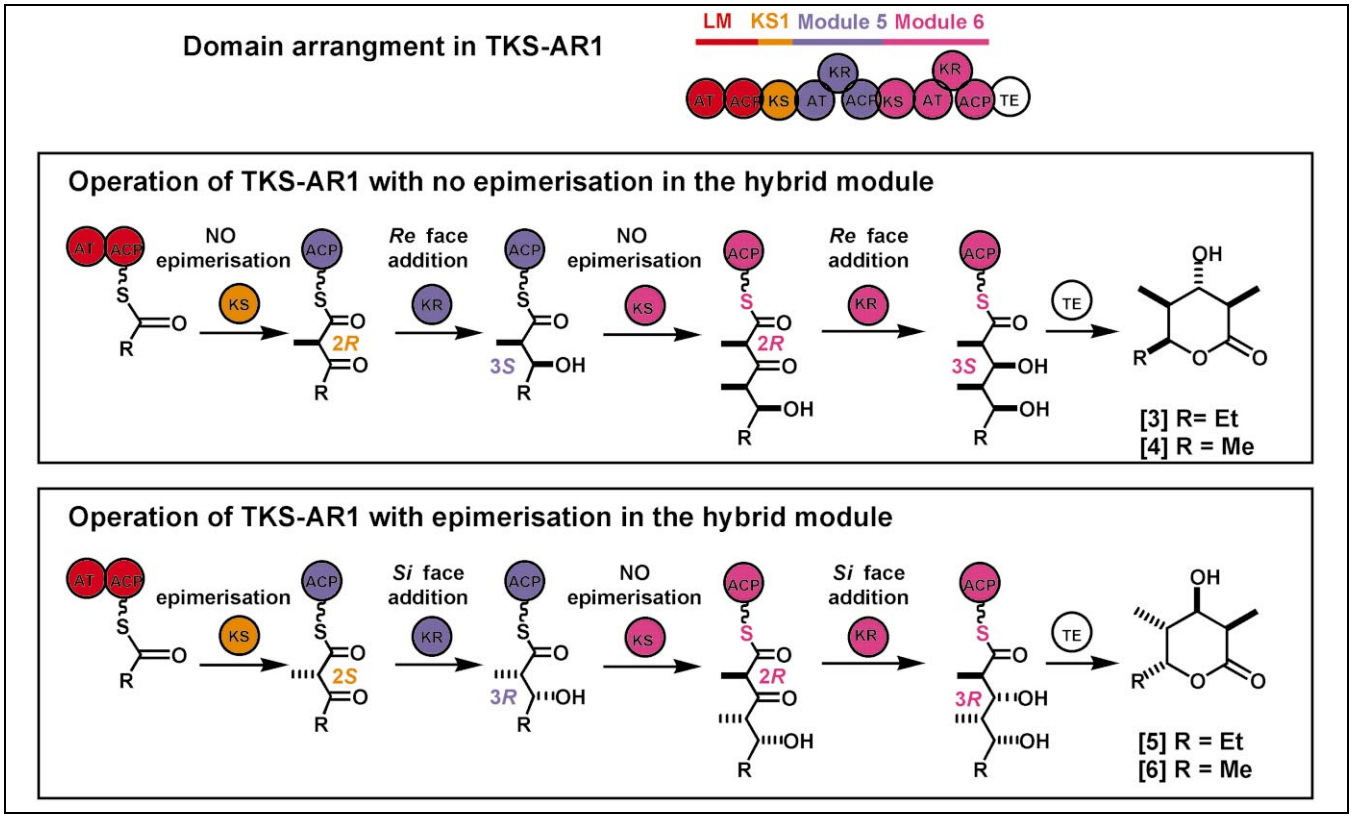

Fig. 11. Condensation and reduction cycles performed by TKS–AR1 in the biosynthesis of compounds **3**–**6**.

[(2*R*), (3*S*), (4*S*), (5*S*)][2] formed by the normal operations of modules 5 and 6 (Fig. 11) which had been isolated earlier in studies on triketide formation by an unmodified version of DEBS3 [26]. This result confirms that KS1 can deliver the unepimerised (2*R*)-keto ester to its partner KR domain in the same module. The second major isomer (**5**) was identical by proton NMR to another stereoisomer [(2*R*), (3*R*), (4*R*), (5*R*)] also isolated in earlier work [30]. The relative stereochemistries at C-2, C-3, C-4 and C-5 were confirmed by molecular modelling.

There are two possible absolute configurations for this compound, either stereoisomer **5** or its enantiomer. We feel confident that the enantiomeric structure can be excluded for two reasons. First, formation of the enantiomer of **5** would require that module 5 behaves normally, whereas module 6 behaves abnormally. Module 5 has two domains with compromised stereospecificity and so is much more likely to be the source of stereochemical divergence. Second, only one diastereoisomer was observed in equivalent studies with DEBS3 [31]. If the enantiomer of **5** had been formed in the present investigation, it would surely have also been formed and isolated as a major metabolite in this earlier work. Work is in progress to confirm the absolute configuration of **5**. For simplicity in the following discussion, we will assume that **5** is the correct configuration.

Compound **5** differs from the 'normal' product at three chiral centres. The most significant point of difference in the context of the current investigation was the stereochemistry at C4 which corresponds to the (2*S*)-isomer produced by epimerisation of the keto ester at the methyl branching site of the diketide formed in module 5. In accepting this opposite epimeric form of the keto ester, KR5 reduces the keto group not in its usual sense by hydride addition to the *re* face to give the (3*S*)-3-hydroxy product but in the opposite sense to give the (3*R*)-product (**5**). Surprisingly, the reverse reduction behaviour then continues in module 6. The condensation proceeds as expected to give the non-epimerised (2*R*)-keto ester but KR6 then reduces this by hydride addition not in its normal mode to the *re* face but to the opposite *si* face. Consequently, the stereochemical relationship between C-2 and C-3 in the second chain extension unit is *anti* rather than *syn*. Remarkably, this altered stereospecificity in the reduction step is triggered not by alteration of the configuration at the branching methyl centre generated in the current chain extension reaction, but by anomalous chiral centres generated in the previous chain extension module.

That both KR5 and KR6 can reduce diketide and triketide intermediates generated in situ with opposite stereochemistries proves that substrate recognition and reduction at the active site are less effectively controlled with the abnormal substrates, even when the intermediates are constrained by the phosphopantetheine chain rather than the model NAC thioester model system.

The most significant outcome of these experiments, however, is the demonstration that module 5 can be turned from a non-epimerising environment to one capable of epimerisation merely by replacement of KS5 with KS1. We believe that the two loading domains can be excluded from this role because of the earlier work by Katz who showed that removal of these domains from the front of module 1 in the otherwise intact erythromycin PKS did not destroy the epimerising activity of module 1 [32]. Our results clearly implicate KS1 as the source of epimerising activity, although it remains to be proved that it produces this change by acting as an active catalyst of the epimerisation process.

In earlier NMR experiments with the keto ester diketide as its NAC thioester, it was shown that the chiral centre rapidly exchanged deuterium with D_2O under conditions which are considered physiological [33]. There is, therefore, a possibility that epimerisation is a spontaneous chemical event which is allowed in modules which epimerise and is prevented in modules which do not support epimerisation. The mechanism of prevention could, for example, rely on rapid transfer of the keto ester intermediate from the KS active site to the KR active site through a channel or tunnel which excludes water and so suppresses the spontaneous epimerisation process. Definition of the precise site and timing of epimerisation will require studies of the structure and function of the KS1 at a detailed molecular level.

While the precise nature of the epimerisation process remains to be defined, the factor that differentiates one module from another in this respect is associated with the KS domain. The results presented here demonstrate that epimerisation activity can be transferred from one module to another by genetic engineering but, depending on the character of the receptor module, the stereochemical outcome of subsequent reduction steps may be both surprising and unexpected to the extent that an *anti* product is formed by a KR domain which normally forms a *syn* product (cf. [29]).

3. Significance

Modular PKSs are potentially valuable for the biosynthesis of novel complex polyketides and many successful modifications to the structure have been achieved to date. In contrast, there have been few modifications to stereochemistry at hydroxyl centres and none at all at methyl branching centres. The results here show that the KS domain plays a key role in control of stereochemistry at

[2] The assignment of C-4 as either *R* or *S* in the open chain form of the triketides discussed in this work is opposite to the assignment of the corresponding carbon centre in the lactone form. This does not indicate an inversion at C-4, but is merely a consequence of the rules of priority used to define *R* or *S* to a particular stereocentre. There has been confusion on this point in some of the earlier literature.

methyl branching sites and that methyl stereochemistry can be altered by transfer of an appropriate KS domain. Of equal significance is the demonstration that KR domains differ from one another in their fidelity when faced with unnatural substrates. Some KR domains, such as KR5 and KR6 in this work, cannot be relied upon to carry out keto ester reduction, either in their normal configurational sense or with their normal fidelity. In one of our constructs a module is allowed to produce both *syn* and *anti* products. This appreciation of the underlying factors determining stereochemical outcome may help to explain why, despite confident predictions, some engineered PKSs have failed to produce the expected product.

4. Materials and methods

4.1. Strains and plasmids

S. erythraea/JC2 was maintained as described previously [28]. Routine cloning and transformation procedures were as previously described for *E. coli* [34] and for *Streptomyces* [35,36]. Chemicals were of analytical grade or the best commercially available. Electrocompetent cells of *E. coli* DH10B strain were made as described previously [37]. X-gal was obtained from Novabiochem. Agarose (electrophoresis grade) was obtained from Life Technologies Inc., casamino acids and tryptone were from Difco laboratories. Glucidex (MD30E), soya flour and beet molasses for SM3 culture media were the kind gifts of Glaxo Wellcome, UK. All antibiotics were bought from Sigma Chemical Company, St. Louis, MO, USA.

4.2. DNA manipulations

PCR reactions were performed on a programmable Robo Cycler Gradient 40 (Stratagene, USA) according to a modification of a protocol described earlier [38]. Automated DNA sequencing was carried out on double-stranded DNA templates using an automated ABI 373A sequencer (Applied Biosystems). RecA protein was purchased from Promega.

4.3. Chemical analysis

^{1}H NMR spectra were recorded at 600 MHz on a Bruker DRX-600, at 500 MHz on a Bruker DRX-500, at 400 MHz on a Bruker AM-400 or Bruker DRX-400, and at 250 MHz on a Bruker AC-250. ^{13}C NMR spectra were recorded at 100 MHz on a Bruker AM-400 or at 400 MHz on a Bruker DRX-400. ^{1}H and ^{13}C NMR spectra were referenced internally to $CHCl_3$ (7.27 and 77.5 ppm respectively). Chromatography was carried out using Merck Kieselgel 60 (40–63 μm). GC-MS was performed on a Finnigan MAT GCQ instrument. Analytical and preparative reverse phase HPLC-MS analysis was carried out using Phenomenex Prodigy 5μ ODS3 100 Å columns with the following dimensions: 250 mm × 4.6 mm (analytical), 250 mm × 10.0 mm (semi-preparative) and 250 mm × 21.2 mm (preparative) on a Finnigan MAT LCQ instrument. Mass spectra were recorded on a Kratos MS 890 double focussing magnetic sector MS (EI) and a Bruker Bio Apex II Fourier transform ion cyclotron resonance (FT-ICR, 4.7 T), (ESI). Solvents were dried under standard conditions.

4.4. Growth medium

Special medium 3 (SM3) (for 1 l); 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.

4.5. Construction of plasmid pKS5

The 513 bp fragment of the gene encoding *ery* module 5 extending from nucleotide 12 069 to nucleotide 12 566 of *eryAIII* [1] was amplified by PCR, and the PCR product inserted in *Sma*I-cut/dephosphorylated pUC18. The oligonucleotides 5′-AAGCTCAGCATC**AAGCTT**CGGCATCAGCGGCACCAAC-3′ and 5′-GCCGCTGAGCAGGTCGGTCAGCGACCAGTC-3′ were used as primers. A unique *Hin*dIII site was engineered in the forward primer. The PCR product was precipitated and ligated to the *Sma*I-cut pUC-18 vector and *E. coli* DH10B cells were transformed with the ligation mixture. Selected colonies were screened for their plasmid content using ampicillin as the resistance marker. Plasmid pKS5 was identified by its restriction map and subsequently sequenced using pUC forward and reverse primers to check for misincorporation of nucleotides during PCR. The 513 bp sequence was found to be accurate.

4.6. Construction of plasmid pKS5A

Plasmid pKS5 was cut with *Blp*I and the 513 bp fragment containing the unique *Hin*dIII site was ligated to the vector pARE24 that had been cut with *Blp*I. Plasmid pARE24 contains the DEBS3 gene fused at its 5′-end with the *ery* loading module, and its construction is described elsewhere [29]. *E. coli* DH10B cells were transformed with the ligation mixture and selected colonies were screened for their plasmid content. The desired plasmid pKS5A was identified by restriction mapping.

4.7. Construction of the expression plasmid pO15-5624

Plasmid pKS5A was cut with *Hin*dIII and the 9.2 kbp fragment containing the gene encoding DEBS3 (minus the KS5 domain) was ligated to the vector pIBO15 [24] that had been cut with *Hin*dIII. *E. coli* DH10B cells were transformed with the ligation mixture and selected colonies were screened for their plasmid content. The desired plasmid pO15-5624 was identified by restriction mapping [30,39].

Acknowledgements

We thank Dr. I. Böhm for the kind gift of the plasmid pIB015, Dr. U. Hanefeld for providing a synthetic stan-

dard of compound (**3**) and Mr. J.B. Lester for DNA sequencing. A.R. wishes to thank Pembroke College, Cambridge and the Cambridge Commonwealth Trust, Cambridge. I.E.H. thanks the Cambridge Commonwealth Trust, Cambridge, Fonds pour la formation de chercheurs et l'aide à la recherche (FCAR) and Glaxo Wellcome plc. D.J.A.K. wishes to thank Professor Ray Jones of The Open University for helpful advice. The project also received financial support from BBSRC.

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.G. Swan, A.M. Rodríguez, C. Vilches, C. Méndez, J.A. Salas, Characterization of a *Streptomyces antibioticus* gene encoding a type I polyketide synthase which has an unusual coding sequence, Mol. Gen. Genet. 242 (1994) 358–362.

[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. Stautnon, 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] D.J. MacNeil, J.L. Occi, K.M. Gewain, T. MacNeil, P.H. Gibbons, C.L. Ruby, S.J. Danis, Complex organization of the *Streptomyces avermitilis* genes encoding the avermectin polyketide synthase, Gene 115 (1992) 119–125.

[7] 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.

[8] 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.

[9] P. Caffrey, D.J. Bevitt, J. Staunton, P.F. Leadlay, Identification of DEBS 1, DEBS 2 and DEBS 3, the multienzyme polypeptides of the erythromycin-producing polyketide synthase from *Saccharopolyspora erythraea*, FEBS Lett. 304 (1992) 225–228.

[10] 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.

[11] 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, Nat. Struct. Biol. 3 (1996) 188–192.

[12] C.M. Kao, M. McPherson, R.N. McDaniel, H. Fu, D.E. Cane, C. Khosla, Gain of function mutagenesis of the erythromycin polyketide synthase. 2. Engineered biosynthesis of eight-membered ring tetraketide lactone, J. Am. Chem. Soc. 119 (1997) 11339–11340.

[13] R. McDaniel, C.M. Kao, H. Fu, P. Hevezi, C. Gustafsson, M. Betlach, G. Ashley, D.E. Cane, C. Khosla, Gain-of-function mutagenesis of a modular polyketide synthase, J. Am. Chem. Soc. 119 (1997) 4309–4310.

[14] J.R. McCormick, E.P. Su, A. Driks, R. Losick, Growth and viability of *Streptomyces coelicolor* mutant for the cell division gene *ftsZ*, Mol. Microbiol. 14 (1994) 243–254.

[15] W.D. Celmer, Macrolide configurational model, J. Am. Chem. Soc. 87 (1965) 1801.

[16] 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.

[17] C.M. Kao, G.L. Luo, L. Katz, D.E. Cane, C. Khosla, Manipulation of macrolide ring size by directed mutagenesis of a modular polyketide synthase, J. Am. Chem. Soc. 117 (1995) 9105–9106.

[18] 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.

[19] R. Pieper, G.L. Luo, D.E. Cane, C. Khosla, Cell free synthesis of polyketides by recombinant erythromycin polyketide synthases, Nature 378 (1995) 263–266.

[20] A.F.A. Marsden, P. Caffrey, J.F. Aparicio, M.S. Loughran, J. Staunton, P.F. Leadlay, Stereospecific acyl transfers on the erythromycin-producing polyketide synthase, Science 263 (1994) 378–380.

[21] D.E. Cane, T.C. Liang, P.B. Taylor, C. Chang, C.C. Yang, Macrolide biosynthesis. 3. Stereochemistry of the chain-elongation steps of erythromycin biosynthesis, J. Am. Chem. Soc. 108 (1986) 4957–4964.

[22] L. Katz, S. Donadio, in: L.C. Vining, C. Stuttard (Eds.), Macrolides, Chapter 14, Butterworth-Heinemann, London, 1995, pp. 406–408.

[23] K.J. Weissman, M. Timoney, M. Bycroft, P. Grice, U. Hanefeld, J. Staunton, P.F. Leadlay, The molecular basis of Celmer's rules: the stereochemistry of the condensation step in chain extension on the erythromycin polyketide synthase, Biochemistry 36 (1997) 13849–13855.

[24] I. Böhm, I.E. Holzbaur, U. Hanefeld, J. Cortés, J. Staunton, P.F. Leadlay, Engineering of a minimal modular polyketide synthase, and targeted alteration of the stereospecificity of polyketide chain extension, Chem. Biol. 5 (1998) 407–412.

[25] I.E. Holzbaur, R.C. Harris, M. Bycroft, J. Cortés, C. Bisang, J. Staunton, B.A.M. Rudd, P.F. Leadlay, Molecular basis of Celmer's rules: the role of two ketoreductase domains in the control of chirality by the erythromycin modular polyketide synthase, Chem. Biol. 6 (1999) 189–195.

[26] J.R. Jacobsen, D.E. Cane, C. Khosla, Spontaneous priming of a downstream module in 6-deoxyerythronolide B synthase leads to polyketide biosynthesis, Biochemistry 37 (1998) 4928–4934.

[27] G.A. Roberts, J. Staunton, P.F. Leadlay, Heterologous expression in *Escherichia coli* of an intact multienzyme component of the erythromycin-producing polyketide synthase, Eur. J. Biochem. 214 (1993) 305–311.

[28] 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.

[29] A. Ranganathan, M. Timoney, M. Bycroft, J. Cortés, I.P. Thomas, B. Wilkinson, L. Kellenberger, U. Hanefeld, I.S. Galloway, J. Staunton, P.F. Leadlay, Knowledge-based design of bimodular and trimodular polyketide synthases based on domain and module swaps: a route to simple statin analogues, Chem. Biol. 6 (1999) 731–741.

[30] C.M. Kao, M. McPherson, R.N. McDaniel, H. Fu, D.E. Cane, C. Khosla, Alcohol stereochemistry in polyketide backbones is controlled by the β-ketoreductase domains of modular polyketide synthases, J. Am. Chem. Soc. 120 (1998) 2478–2479.

[31] N. Wu, F. Kudo, D.E. Cane, C. Khosla, Analysis of the molecular recognition features of individual modules derived from the erythromycin polyketide synthase, J. Am. Chem. Soc. 122 (2000) 4847–4852.

[32] 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.

[33] I.E. Holzbaur, R.C. Harris, M. Bycroft, J. Cortés, C. Bisang, J. Staunton, B.A.M. Rudd, P.F. Leadlay, Molecular basis of Celmer's rules: the role of two ketoreductase domains in the control of chirality by the erythromycin modular polyketide synthase, Chem. Biol. 6 (1999) 189–195.

[34] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

[35] H. Yamamoto, K.H. Maurer, C.R. Hutchinson, Transformation of *Streptomyces erythraeus*, J. Antibiot. 39 (1986) 1304–1313.

[36] D.A. Hopwood, M.J. Bibb, K.F. Chater, T. Kieser, C.J. Bruton, H.M. Kieser, D.J. Lydiate, C.P. Smith, J.M. Ward, H. Schrempf, Genetic manipulation of Streptomyces, a laboratory manual, John Innes Institution, Norwich, 1985.

[37] 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.

[38] R.K. Saiki, D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis, H.A. Erlich, Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase, Science 239 (1988) 487–491.

[39] 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.