Stereospecific Acyl Transfers on the Erythromycin-Producing Polyketide Synthase

Andrew F. A. Marsden, Patrick Caffrey, Jesus F. Aparicio, Mark S. Loughran, James Staunton, Peter F. Leadlay*

During assembly of complex polyketide antibiotics like erythromycin A, molecular recognition by the multienzyme polyketide synthase controls the stereochemical outcome as each successive methylmalonyl–coenzyme A (CoA) extender unit is added. Acylation of the purified erythromycin-producing polyketide synthase has shown that all six acyltransferase domains have identical stereospecificity for their normal substrate, (2*S*)-methylmalonyl-CoA. In contrast, the configuration of the methyl-branched centers in the product, that are derived from (2*S*)-methylmalonyl-CoA, is different. Stereoselection during the chain building process must, therefore, involve additional epimerization steps.

The clinically important antibiotic erythromycin A is derived from one propionyl-CoA and six methylmalonyl-CoA units that are sequentially incorporated into a growing polyketide chain (*1*). As in all complex polyketides, the stereochemistry of chain growth differs between extension cycles (*2–5*), but the molecular basis for this is unknown (*3, 6*). Understanding this unusual stereochemical divergence would facilitate attempts to design hybrid polyketide antibiotics (*7, 8*).

The 6-deoxyerythronolide B synthase (DEBS) consists of three giant multienzymes DEBS 1, DEBS 2, and DEBS 3 (*7, 9, 10*), which catalyze formation of the macrolide core of erythromycin A from simple fatty acyl precursors in successive rounds of condensation and (to a variable extent) reduction. Each DEBS multienzyme is thought to contain all the activities necessary for two of these six chain extension cycles (Fig. 1). DEBS was purified from an erythromycin-producing strain of *Saccharopolyspora erythraea* and separated into its three constituent multienzymes (*10*). To confirm previous work with recombinant DEBS 3 purified from *Escherichia coli* (*11*), we incubated DEBS 3 from *S. erythraea* for 5 minutes with a fivefold molar excess of (2*R*,*S*)-[1-^{14}C]methylmalonyl-CoA and then subjected it to SDS–polyacrylamide gel electrophoresis. Radioactivity was found associated with the DEBS 3 polypeptide (Fig. 2A, lane 1). The extent of labeling decreased with longer incubation time, falling to zero after about 2 hours (Fig. 2A, lane 2) (*11*). Identical results were obtained with DEBS 1 and DEBS 2. The simplest interpretation of these data is that the six methylmalonyltransferase components of the synthase, two of which are found in each of DEBS 1, DEBS 2, and DEBS 3, are acylated and, in the absence of chain growth, the acyl group is subsequently hydrolyzed. Analogous results have been found for acylation of fatty acid synthase by malonyl-CoA and acetyl-CoA (*12*). Addition of further (2*R*,*S*)-[1-^{14}C]methylmalonyl-CoA after 2 hours of incubation restored the original amount of radiolabeling of DEBS 3 (Fig. 2A, lane 4), after which the intensity of labeling again decreased. The products of this reaction, as determined by thin-layer chromatography analysis, were methylmalonate and CoA (*13*).

If only one isomer of methylmalonyl-CoA was to acylate the enzyme, the unreactive isomer would remain after extended incubation with DEBS 3. We tested this by adding, after such an extended incubation, purified methylmalonyl-CoA epimerase from *Propionibacterium shermanii* (*14, 15*) that specifically and rapidly interconverts (2*R*)- and (2*S*)-methylmalonyl-CoA. This treatment restored about half the original amount of radiolabeling of purified recombinant DEBS 3 (Fig. 2A, lane 5), demonstrating (i) that only one isomer reacts with the enzyme and (ii) that DEBS 3 does not possess an intrinsic methylmalonyl-CoA epimerase activity. The same results were obtained for DEBS 1 and DEBS 2.

The DEBS multienzymes all competed for the same isomer of methylmalonyl-CoA as substrate. This was shown by preincubation of radiolabeled (2*R*,*S*)-methylmalonyl-CoA with one of the three DEBS multienzymes from *S. erythraea* until labeling was no longer observed (Fig. 2B, lanes 6 and 7). If one of the other two components was then added, it was not labeled (Fig. 2B, lanes 8 and 9). However, addition of methylmalonyl-CoA epimerase allowed subsequent labeling of all DEBS multienzymes present in the incubation mixture (Fig. 2B, lanes 11 to 13).

To discover which of the two stereoisomers of methylmalonyl-CoA is the preferred substrate for all of the DEBS multienzymes, we prepared samples of (2*R*)- and (2*S*)-methylmalonyl-CoA under conditions where the nonenzymatic epimerization is negligible (*15*). This was done in situ (Fig. 2C); the unwanted isomer was removed with enzymes that act specifically on the (2*R*) and the (2*S*) isomer, respectively. The remaining isomer was used immediately in the experiments with DEBS.

All three enzymes, DEBS 1, DEBS 2, and DEBS 3, either individually or in combination, were labeled by (2*S*)-methylmalonyl-CoA (Fig. 2C). For DEBS 1, both acyltransferases were labeled, as shown by limited proteolysis and subsequent labeling of the digests (Fig. 3). Similar results were obtained for DEBS 2 and DEBS 3. Only those proteolytic fragments containing at least one acyltransferase were labeled (*16*). The (2*R*) isomer did not label either DEBS 2 or DEBS 3 and only labeled DEBS 1 weakly. Even this weak labeling of DEBS 1

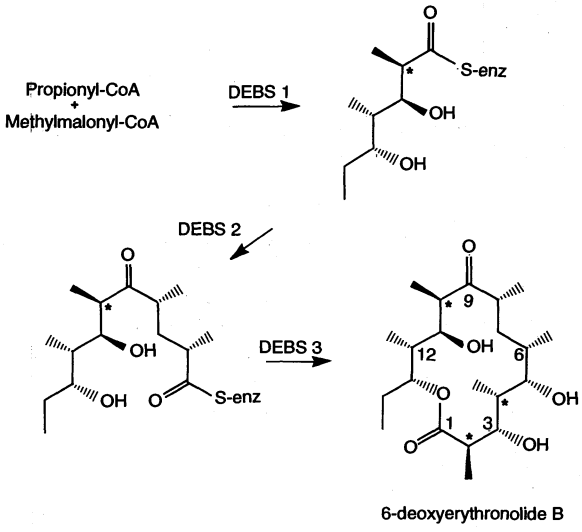

Fig. 1. The reaction catalyzed by DEBS from *S. erythraea*. The first two cycles of chain assembly are thought to be catalyzed by DEBS 1, the third and fourth by DEBS 2, and the final two by DEBS 3 (*7, 9*). Asterisks denote the three asymmetric centers shown to be derived from (2*S*)-methylmalonyl-CoA (*3*). The different configuration in the product is thought to result from an inversion of configuration as for fatty acid biosynthesis (*20*). S-enz denotes covalent attachment to an unspecified enzyme thiol.

A. F. A. Marsden, P. Caffrey, J. F. Aparicio, P. F. Leadlay, Cambridge Centre for Molecular Recognition, and Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, United Kingdom. M. S. Loughran and J. Staunton, Cambridge Centre for Molecular Recognition, and Department of Organic Chemistry, University Chemical Laboratory, University of Cambridge, Cambridge CB2 1EW, United Kingdom.

*To whom correspondence should be addressed.

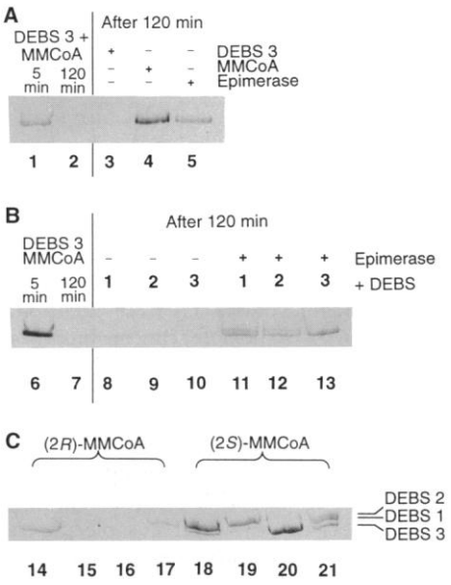

Fig. 2. Autoradiograms produced by the radiolabeling of DEBS multienzymes with substrate methylmalonyl-CoA (MMCoA). (**A**) Incubation of purified DEBS 3 (10 pmol) from *S. erythraea* with an approximately fivefold molar excess of (2*R*,*S*)-[1-^{14}C]methylmalonyl-CoA. The preparations were incubated for 5 min (lane 1) or 120 min (lane 2). After an initial incubation of 120 min, further additions were made: DEBS 3 7 pmol (lane 3), or [1-^{14}C]methylmalonyl-CoA 50 pmol (lane 4), or methylmalonyl-CoA epimerase 1 mU (lane 5) (*14*). (**B**) As in (A), except that a 20-fold molar excess of radiolabel was used. Incubations were done for 5 min (lane 6) or 120 min (lane 7). After 120 min, additions of DEBS 1, 2, or 3 were made (lanes 8, 9, and 10, respectively) and the incubations were continued for 5 min. In lanes 11, 12, and 13, after an initial incubation of 120 min, methylmalonyl-CoA epimerase (1 mU) was added and then DEBS 1, 2, and 3, respectively. (**C**) (2*R*)-methylmalonyl-CoA was incubated with DEBS 1, 2, or 3 (lanes 14 to 16, respectively) or all three multienzymes (lane 17). Analogous experiments were done with (2*S*)-methylmalonyl-CoA (lanes 18 to 21).

was not due to acylation by (2*R*)-methylmalonyl-CoA. The enzymatic preparation of this isomer involves propionyl-CoA as a by-product, the bulk of which is removed by specific hydrolysis. The traces that remained in this particular experiment account for all of the observed labeling. The NH_2-terminal loading domain of DEBS 1 was strongly and specifically labeled by [1-^{14}C]propionyl-CoA (Fig. 3) (*16*). Also, preincubation of DEBS 1 with unlabeled propionyl-CoA abolished the weak labeling seen in Fig. 2C, but did not interfere with the binding of (2*S*)-methylmalonyl-CoA (*16*). Taken together, these results confirm that all of the six individual methylmalonyltransferases in the synthase react with the same (2*S*) isomer of methylmalonyl-CoA.

There are several plausible mechanisms by which the stereochemistry of methyl branching during polyketide chain growth might be controlled. One proposal in particular has gained attention: that individual acyltransferases in the multienzymes might have different but absolute substrate specificity for one or the other stereoisomer of the extender unit methylmalonyl-CoA. This has been proposed (*7*, *17*) on the basis of two findings: DNA sequencing of the DEBS genes (*7*, *9*) showed that there is a separate methylmalonyltransferase for each cycle of chain extension, and replacement of the acyltransferase domain in the final extension cycle catalyzed by DEBS 3 leads to the incorporation of an acetate unit at this position to yield 2-norerythromycin (*18*). Alternatively, methylmalonyl transfer to the multienzyme might be promiscuous, from either (2*R*)- or (2*S*)-methylmalonyl-CoA. This would require a "proofreading" activity to hydrolyze those acyl groups added to the wrong sites, analogous to the discrimination exercized by certain aminoacyl-tRNA synthetases (*19*). Finally, transfer to the enzyme might be from the (2*R*) isomer or the (2*S*) isomer only. In this event, additional enzyme-catalyzed steps would be required during chain growth to epimerize at least some of the methyl-branched centers to the configuration found in the antibiotic product.

The finding reported here, that all of the units used in the construction of erythromycin apparently come from (2*S*)-methylmalonyl-CoA, allows a clear choice between these alternatives. The observed stereoselection during polyketide chain growth is not due to discrimination by the acyltransferases (*7*) nor to the operation of a proofreading mechanism, but must be the effect of additional epimerization steps catalyzed by the synthase. Because chain assembly is processive (*20*), epimerization must, therefore, occur either in the methylmalonyl group immediately before condensation or in the ketide product immediately after condensation. Previous attempts to address this question, by feeding isotopically labeled precursors to whole cells and then examining the distribution of label in the polyketide products, have led to ambiguous results (*3*, *6*). For example, a specifically deuterated propionate precursor of (2*S*)-methylmalonyl-CoA was incorporated into erythromycin A (*3*) in which residual deuterium was found attached only at C-2, C-4, and C-10 of the macrolide ring (Fig. 1). This result implicated the second, fifth, and sixth extension units as having arisen from incorporation of (2*S*)-methylmalonyl-CoA, with inversion of configuration as found for fatty acid synthase (*21*), but left the origin of the remaining units obscure.

The stereochemical variation along the polyketide chain of erythromycin is shared by more than 150 macrolide antibiotics (*2*, *22*), and similar patterns of variation have been discerned in polyethers (*4*, *22*). If, as seems probable, chain assembly in all these closely related polyketide synthases requires only one stereoisomer of methylmalonyl-CoA, this must be taken into account in future attempts (*7*, *8*) to reprogram such polyketide synthases and produce hybrid antibiotics.

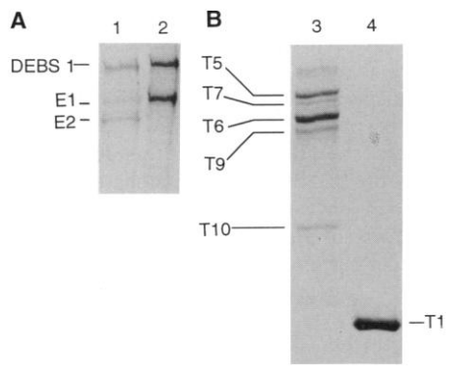

Fig. 3. Autoradiograms produced by the incubation of proteolyzed DEBS 1 with radiolabeled substrate. Similar data were obtained for DEBS 2 and DEBS 3, showing that the radiolabel was confined exclusively to polypeptides containing acyltransferase domains. (**A**) DEBS 1 was partially digested with elastase (*16*) and briefly incubated with [1-^{12}C]propionyl-CoA, then (2*R*,*S*)-[1-^{14}C]methylmalonyl-CoA (lane 1), or [1-^{14}C]propionyl-CoA alone (lane 2). Polypeptide fragments contained the following activities (*16*): DEBS 1—$AT_1ACP_1KS_1AT_2KR_1ACP_2KS_2AT_3KR_2ACP_3$; E1—$AT_1ACP_1KS_1AT_2KR_1ACP_2$; and E2—$KS_2AT_3KR_2ACP_3$. (**B**) DEBS 1 was partially digested with trypsin (*16*) and briefly incubated with [1-^{12}C]propionyl-CoA, then (2*R*,*S*)-[1-^{14}C]methylmalonyl-CoA (lane 3), or [1-^{14}C]propionyl-CoA alone (lane 4). Polypeptide fragments contained the following activities: T1—AT_1ACP_1; T5—$KR_1ACP_2KS_2AT_3KR_2ACP_3$; T6—$ACP_2KS_2AT_3KR_2ACP_3$; T7—$KR_1ACP_2KS_2AT_3$; T9—$KR_1ACP_2KS_2AT_3KR_2$; and T10—$KS_1AT_2$. ACP, acyl carrier protein; AT, acyltransferase; KR, β-ketoreductase; KS, β-ketoacyl ACP synthase. Numbering of activities corresponds to proximity to NH_2-terminus of DEBS 1.

REFERENCES AND NOTES

1. D. A. Hopwood and D. H. Sherman, *Annu. Rev. Genet.* **24**, 37 (1990); J. Staunton, *Angew. Chem. Int. Ed. Engl.* **30**, 1302 (1991).
2. W. D. Celmer, *J. Am. Chem. Soc.* **87**, 1801 (1965).
3. D. E. Cane, T. Liang, P. B. Taylor, C. Chang, C.-C. Yang, *ibid.* **108**, 4957 (1986).
4. D. E. Cane, W. D. Celmer, J. W. Westley, *ibid.* **105**, 3594 (1983).
5. D. E. Cane, R. H. Lambalot, P. C. Prabhakaran, W. R. Ott, *ibid.* **115**, 522 (1993).
6. J. M. Bulsing *et al.*, *J. Chem. Soc. Chem. Commun.* **19**, 1301 (1984); M. M. Sherman and C. R. Hutchinson, *Biochemistry* **26**, 438 (1987); G. R. Sood, J. A. Robinson, A. A. Ajaz, *J. Chem. Soc. Chem. Commun.* **21**, 1421 (1984).
7. S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson, L. Katz, *Science* **252**, 675 (1991).
8. D. A. Hopwood, *Curr. Top. Biotechnol.*, in press.

9. J. Cortes, S. F. Haydock, G. A. Roberts, D. J. Bevitt, P. F. Leadlay, *Nature* **348**, 176 (1990); D. J. Bevitt, J. Cortes, S. F. Haydock, P. F. Leadlay, *Eur. J. Biochem.* **204**, 39 (1992).
10. P. Caffrey, D. J. Bevitt, J. Staunton, P. F. Leadlay, *FEBS Lett.* **304**, 225 (1992).
11. G. A. Roberts, J. Staunton, P. F. Leadlay, *Eur. J. Biochem.* **214**, 305 (1993).
12. S. J. Wakil, *Biochemistry* **28**, 4523 (1989).
13. A. F. A. Marsden, unpublished data.
14. P. F. Leadlay, *Biochem. J.* **197**, 413 (1981).
15. J. Q. Fuller and P. F. Leadlay, *ibid.* **213**, 643 (1983).
16. J. F. Aparicio, P. Caffrey, A. F. A. Marsden, J. Staunton, P. F. Leadlay, *J. Biol. Chem.*, in press.
17. S. Donadio, J. B. McAlpine, P. J. Sheldon, M. Jackson, L. Katz, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7119 (1993).
18. J. B. McAlpine *et al.*, *J. Antibiot.* **40**, 1115 (1987).
19. A. R. Fersht, *Enzyme Structure and Mechanism*, (Freeman, New York, ed. 2, 1985), pp. 347–358.
20. D. E. Cane and C.-C. Yang, *J. Am. Chem. Soc.* **109**, 1255 (1987); S. Yue, J. C. Duncan, Y. Yamamoto, C. R. Hutchinson, *ibid.*, p. 1253.
21. B. Sedgwick *et al.*, *Eur. J. Biochem.* **75**, 481 (1977).
22. D. O'Hagan, *Nat. Prod. Rep.* **6**, 205 (1989).
23. Supported by grants from the Science and Engineering Research Council (United Kingdom) (Molecular Recognition Initiative and Biotechnology Directorate–LINK). J.F.A. was the recipient of a European Community bursary. We thank J. R. Knowles and D. A. Hopwood for their comments during preparation of the manuscript.

30 August 1993; accepted 30 November 1993

Rational Design of Potent, Bioavailable, Nonpeptide Cyclic Ureas as HIV Protease Inhibitors

Patrick Y. S. Lam,* Prabhakar K. Jadhav, Charles J. Eyermann, C. Nicholas Hodge, Yu Ru, Lee T. Bacheler, James L. Meek, Michael J. Otto, Marlene M. Rayner, Y. Nancy Wong, Chong-Hwan Chang, Patricia C. Weber, David A. Jackson, Thomas R. Sharpe, Susan Erickson-Viitanen*

Mechanistic information and structure-based design methods have been used to design a series of nonpeptide cyclic ureas that are potent inhibitors of human immunodeficiency virus (HIV) protease and HIV replication. A fundamental feature of these inhibitors is the cyclic urea carbonyl oxygen that mimics the hydrogen-bonding features of a key structural water molecule. The success of the design in both displacing and mimicking the structural water molecule was confirmed by x-ray crystallographic studies. Highly selective, preorganized inhibitors with relatively low molecular weight and high oral bioavailability were synthesized.

Knowledge of the HIV protease (HIV PR) mechanism of action and substrate specificity has been extensively used to design a variety of transition state–based inhibitors with inhibition constants in the nanomolar or subnanomolar range (*1*, *2*). The symmetry of the HIV PR dimer guided the design of twofold (C2) symmetric and pseudosymmetric inhibitors (*3*). However, these inhibitors retain substantial peptide character, and despite many elegant structure-activity studies, it has been difficult to combine adequate potency with oral bioavailability (*3*, *4*). The difficulty in developing such leads into useful therapeutics is challenging, for in addition to the traditional barriers encountered in the drug development process, peptide-based molecules are in general biologically unstable, poorly absorbed, and rapidly metabolized (*5*). This challenge is not unique to HIV PR; transition from peptide-based leads to therapeutics has proven formidable for other enzymes such as renin inhibitors (*6*, *7*).

We have previously explored a series of potent, linear C2-symmetric inhibitors in which the transition state mimetic was a diol (*8*). We were unable to overcome the poor oral bioavailability of these peptide molecules and consequently sought other approaches. The technique of searching databases containing three-dimensional (3D) molecular structures has been used to identify synthetic frameworks that can serve as the starting point for the design of nonpeptide inhibitors, and this approach has been explored with HIV PR. Unfortunately, the HIV PR inhibitors designed to date on the basis of 3D database searches (*9*, *10*) have yielded inhibitors with only micromolar potency.

Our current design of nonpeptide inhibitors (*11*) began with structural information available from published x-ray crystal structures of HIV-PR inhibitor complexes (*12*–*15*). A common feature observed is the presence of a tetracoordinated structural water molecule linking the bound inhibitor to the flexible glycine-rich β strands or "flaps" of the HIV PR dimer (Fig. 1). This water molecule accepts two hydrogen bonds from backbone amide hydrogens of HIV PR residues Ile 50 and Ile 50′ and donates two hydrogen bonds to carbonyl oxygens of the inhibitor, thus inducing the fit of the flaps over the inhibitor (*16*). Its relevance to the generation of HIV PR inhibitors has been noted (*12*, *13*).

We hypothesized that incorporation of the binding features of this structural water molecule into an inhibitor would be beneficial because its displacement should be energetically favorable (*17*). In addition, conversion of a flexible, linear inhibitor into a rigid, cyclic structure with restricted conformations should provide a positive entropic effect. Finally, incorporation of a mimic for the structural water within the inhibitor should ensure specificity for the HIV PR as against other aspartic acid proteases, because this water molecule is unique to retroviral proteases. We reasoned that these effects might provide highly potent and specific binding and reduce the need for multiple interactions at the specificity pockets. This should permit design of smaller (<600 daltons) inhibitors with improved oral bioavailability.

Extensive structure-activity relations (SARs) established for C2-symmetric diols indicated that the diol imparts significant potency as compared with corresponding mono-ol transition state analogs (*3*, *8*). Thus, we wanted to incorporate this feature of the diol–HIV PR interaction into a 3D pharmacophore model. However, no x-ray structure of a C2-symmetric diol-protease complex was available when this work was initiated; two independent reports of structures have since appeared (*18*). Therefore, computer models for C2-symmetric diols bound to the active site of HIV-1 PR were developed from the crystal structure coordinates of a hydroxyethylene inhibitor bound to HIV PR (*15*) by means of distance geometry (*19*) (Fig. 2, A and B) and several pharmacophores were generated.

The simplest pharmacophore model (Fig. 2C) was based on two key intramolecular distances: that between symmetric hydrophobic groups, designated P1 and P1′, that occupy corresponding enzyme pockets S1 and S1′ and that from P1 and P1′ to a hydrogen bond donor/acceptor group (or groups) that binds to catalytic aspartates. A 3D database search (*20*, *21*) with this pharmacophore model yielded the "hit" (*22*) shown in Fig. 2D, which not only met the initial search criteria, but also included an oxygen that matched the structural water found in HIV PR–inhibitor complexes. This 3D search indicated that a phenyl ring could properly position groups to interact with aspartates 25 and 25′ as well as to mimic the structural water (Fig. 2E). However, because a phenyl ring might not properly position all substituents in the

Departments of Virology Research and Chemical and Physical Sciences, The DuPont Merck Pharmaceutical Company, Wilmington, DE 19880.

*To whom correspondence should be addressed.