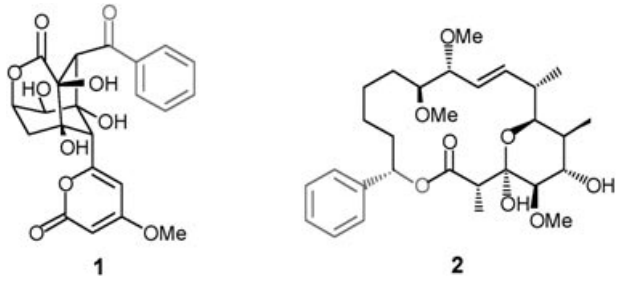
Engineered Biosynthesis of Phenyl-Substituted Polyketides

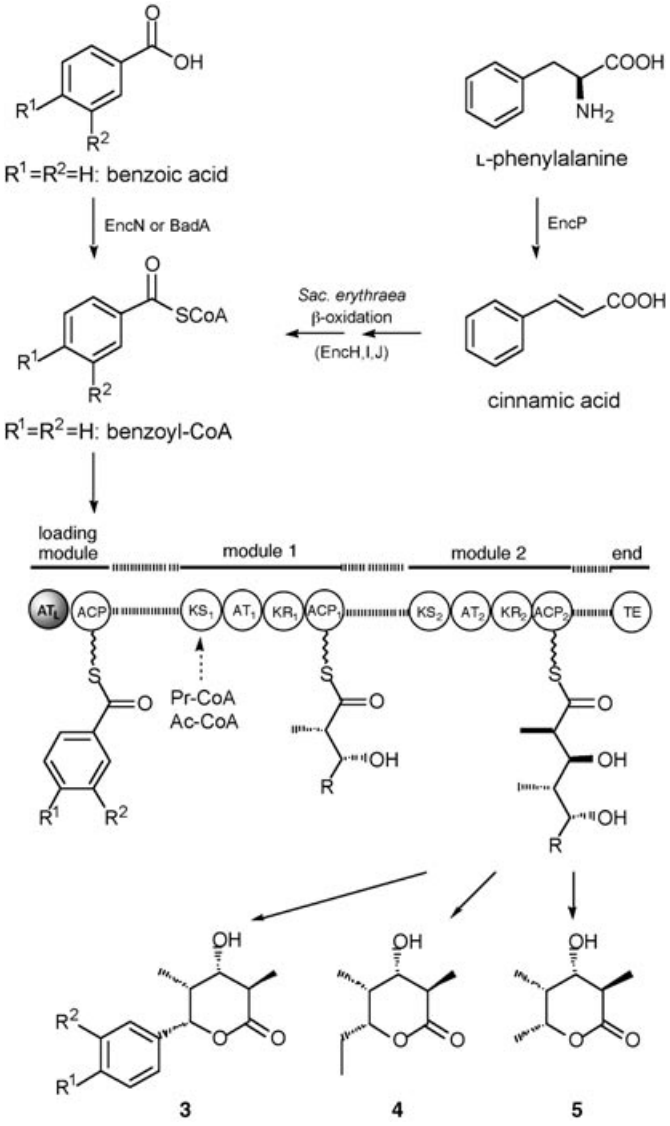
José Garcia-Bernardo,[a] Longkuan Xiang,[b] Hui Hong,[c] Bradley S. Moore,*[b] and Peter F. Leadlay*[a]

Benzoic acid, activated as its coenzyme A thioester (benzoyl-CoA) serves as a building block in the biosynthetic pathways to a number of important natural products, including zaragozic acid A (squalestatin S1)[1] from fungi and paclitaxel (taxol)[2] and cocaine[3] from plants. It also provides the starter unit for the polyketides enterocin (**1**)[4] and soraphen (**2**)[5] from prokaryotic micro-organisms (Scheme 1). We have previously shown that engineering the loading module of modular polyketide synthases (PKSs) to accept alternative starter units provides an effective route to novel analogues of clinically useful polyketide drugs, such as erythromycin A.[6] We now describe results that promise to broaden the scope for such biosynthetic engineering of new polyketide derivatives containing benzoate starter units.



Scheme 1. *Structure of benzoate-primed (shaded) polyketides enterocin (**1**) and soraphen (**2**).*

We recently reported that using the acyltransferase (AT) domain of the loading module of the soraphen-producing PKS of *Sorangium cellulosum*[7] to replace the propionate-selective AT in the loading module of a truncated bimodular derivative (DEBS1-TE) of the erythromycin-producing PKS of *Saccharopolyspora erythraea*[8] led to the production of 5-phenyl-substituted triketide lactone **3**, along with the normal products of DEBS1-TE **4** and **5**.[9] Lactone **3** is formed by polyketide chain extension of benzoyl units attached to the hybrid loading module, while **4** and **5** are thought to be formed by competitive direct loading of, respectively, propionyl-CoA or acetyl-CoA directly onto the ketosynthase domain of extension module 1 of DEBS1-TE (Scheme 2). It was found to be essential for the production of **3**, that the recombinant *S. erythraea* also expressed the *badA* gene of the anaerobic pseudomonad *Rhodopseudomonas palustris*, which encodes one of the benzoate:CoA ligases of that organism,[10] and also that the medium was supplemented with high concentrations of benzoic acid. Unfortunately, the observed deleterious effects of 3 mM benzoic acid on the growth of *S. erythraea*[9] make this a suboptimal approach for mutasynthesis. Surprisingly, cinnamic acid, which was expected to give rise to intracellular benzoyl-CoA by β-oxidation,[11] was found to be toxic for *S. erythraea* even at low



Scheme 2. *Alternative routes for supply of benzoyl-CoA to the* sor-ery *hybrid triketide synthase in* Saccharopolyspora erythraea. *Pr-CoA = propionyl-CoA, Ac-CoA = acetyl-CoA, MeMal-CoA = methylmalonyl-CoA, KS = ketosynthase, AT = acyltransferase, KR = ketoreductase, ACP = acyl carrier protein, TE = thioesterase. Shaded part of PKS is derived from soraphen PKS. R^1 and R^2 = H, F, and OH; see Table 1 for explanation.*

[a] *Dr. J. Garcia-Bernardo, Prof. Dr. P. F. Leadlay*
Department of Biochemistry, University of Cambridge
80 Tennis Court Road, Cambridge, CB2 1GA (UK)
Fax: (+44) 1223-766091
E-mail: pfl10@mole.bio.cam.ac.uk

[b] *Dr. L. Xiang, Prof. Dr. B. S. Moore*
College of Pharmacy and the Department of Chemistry,
University of Arizona
Tucson, Arizona 85721 (USA)
Fax: (+1) 520-626-2466
E-mail: moore@pharmacy.arizona.edu

[c] *Dr. H. Hong*
Department of Chemistry, University of Cambridge
Lensfield Road, Cambridge, CB2 1EW (UK)

 DOI: 10.1002/cbic.200400007

levels (0.1 mM).[12] In contrast, in *S. cellulosum*, cinnamate and cinnamate analogues have recently been shown to be effective precursors of soraphens through a β-oxidation pathway;[13] while in *"Streptomyces maritimus"*, both cinnamate and *p*-fluorocinnamate have been shown to be incorporated as precursors of analogues of enterocin (**1**) and the related wailupemycins.[14]

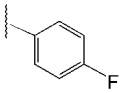
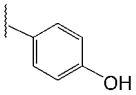
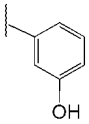
An alternative strategy is to replace genes in the strain with suitable biosynthetic ones that promote the intracellular production of the key biosynthetic building blocks. Thus, inserting genes for cyclohexanecarboxylic acid-CoA production into *Streptomyces avermitilis* supported efficient synthesis of doramectin.[15]

We now report that replacement of *badA* by the alternative benzoate:CoA ligase *encN* from the biosynthetic pathway to **1** in the marine isolate "*S. maritimus*"[16] leads to a threefold higher incorporation of added benzoic acid into **3**, as judged by GC-MS analysis of the products. The increased efficiency may, for example, be due to better expression of the *enc*N gene. Furthermore, this replacement did not lead to any inhibition of strain growth at comparable concentrations of added benzoate.

In parallel experiments, EncN was also found to convert added *p*-fluorobenzoate and both *m*- and *p*-hydroxybenzoates into the corresponding coenzyme A esters, since these compounds served as substrates for the hybrid PKS and gave rise to analogues of **3**. In each case the amounts were small, but the $[M+NH_4]^+$ ion (ammonia CI) was observed, and the MS/MS fragmentation pattern of this ion was fully consistent with the proposed structure of the respective triketide lactone products (Table 1). In vitro, *p*-fluorobenzoate is activated nearly ten times more efficiently than *p*- or *m*-hydroxybenzoate,[14] but the levels of production of the lactone derivatives are broadly similar. Other monosubstituted benzoates, including cyclohex-1-enecarboxylic acid, and 2- and 3-thiophenecarboxylic acids, which are each activated by EncN as efficiently as *p*-fluorobenzoate gave little or no detectable aromatic triketide product. These results apparently reflect the specificity of the polyketide synthase, in particular the loading AT domain derived from the soraphen PKS,[13] rather than that of EncN itself.

We particularly wished to determine whether co-expression of the hybrid PKS gene together with the gene *encP* of "*S. maritimus*", which encodes an L-phenylalanine ammonia lyase,[17] might allow production of the target 5-phenyl triketide lactone **3** even in the absence of added benzoate (Scheme 2). The results of this experiment (Table 1) showed that inclusion of *encP* in an expression cassette with the hybrid PKS gene in *S. erythraea* did indeed lead, as hoped, to significant accumulation of **3**, albeit in amounts some eightfold lower than when *encN* was present and coexpressed (together with an optimal initial concentration of added benzoate). This remarkable result means that *S. erythraea* must contain the enzymes necessary to activate cinnamate as its CoA thioester and to degrade this by β-oxidation to benzoyl-CoA (Scheme 2); and it therefore allows the production of intracellular benzoyl-CoA without the need to add precursors to the medium. Given that we have observed an apparent toxicity of added cinnamate,[12] it is possible that coexpression in the engineered *S. erythraea* of the genes encoding further enzymes from the enterocin feeder pathway from phenylalanine to benzoyl-CoA[11] might in future further increase flux through this plant-like precursor pathway.

Table 1. *GC retention times (t_r) and MS measurements of aromatic tri-ketides generated from strains of* S. erythraea *engineered to produce intracellular benzoyl-CoA, or monosubstituted derivatives.*

Starter unit source	Ar	t_r [min] $[M+NH_4]^+$
benzoic acid phenyl alanine		20.1 238 (220, 202, 157, 142)
p-fluorobenzene		19.1 256 (238, 220, 175, 160)
p-hydroxybenzoate		17.5 268[a] (253, 236, 219, 176)
m-hydroxybenzoate		18.5 268[a] (253, 236, 219, 176)

Some evidence was obtained from MS analysis for inefficient production of analogues of **3** when either 2- or 3-thiophene carboxylic acid was fed to *S. erythraea*. Nicotinic acid, 2-furoic acid, cyclohex-1-enecarboxylic acid and *p*-toluic acid were not incorporated. [a] Analysed as methoxybenzoate triketide lactone. Mass values in brackets represent primary peaks in the MS-MS spectrum.

In summary, expression in *S. erythraea* of a single additional gene from the enterocin pathway allows either production of phenyl-substituted polyketides without medium supplementation (*encP*); or incorporation of a range of substituted phenyl starter units into such polyketides (*encN*). These findings further expand the prospects for engineered biosynthesis of complex polyketides as leads in drug discovery.

Experimental Section

Coexpression of *encN* and *sor-ery* hybrid PKS gene *tkscjw3*: The gene *encN* and its ribosome binding site (RBS) were PCR amplified from the cosmid clone pJP15F11[4] with the primers 5′-G*TTAATTAA*TGCCGCGACCGGGGGCAG-3′ (forward) and 5′-G*CATATG*TGTCCTCCTGGTCACACGGCGTGGGC-3′ (reverse), cloned into pCR^R-Blunt (Invitrogen), sequence verified, digested with *Pac*I and *Nde*I (the introduced restriction sites are indicated by italics), and cloned into the appropriate sites upstream of the *sor-ery* hybrid PKS gene *tkscjw3* in the expression plasmid pCJW45[9] to yield pBM17. The underlined region in the reverse primer corresponds to the RBS associated with *tkscjw3*. Expression of pBM17 (and of pBM18, see below) in *S. erythraea* JC2, feeding of sodium benzoate and analogues (3 mM), and GC-MS analysis of the resulting triketide lactones followed published procedures.[9] The extracts were analysed by GC-MS on a Finnigan GCQ instrument (ThermoFinnigan). A Phenomenex Zebron ZB5 5% phenyl-polysiloxane column was used with helium as carrier gas. The GC-MS was performed in chemical ionisation (CI) mode with NH_3 as reagent gas. For the feeding experiments with *p*- and *m*-hydroxybenzoic acid, the crude extracts were derivatized with trimethylsilyldiazomethane before GC-MS analysis.

Coexpression of *encP* and *tkscjw3*: In a similar fashion to the construction of pBM17, *encP* and its RBS were PCR amplified from pJP15F11 with the primers 5′-G*TTAATTAA*CCCGTCGAGTCCACCG-3′ (forward) and 5′-G*CATATG*TGTCCTCCTTCCAGGTGCTGCTTCAG-3′ (reverse) in the construction of the pCJW45-derivative pBM18 harbouring *encP*.

Acknowledgements

We gratefully acknowledge the support of the European Union, through a Marie Curie Postdoctoral Fellowship (J.G.-B.); of the UK Biotechnology and Biological Sciences Research Council via a project grant (P.F.L.); and of the National Institutes of Health (grant AI47818) (B.S.M.).

Keywords: enterocin · enzymes · loading module · mutasynthesis · polyketides

[1] a) J. D. Bergstrom, C. Dufresne, G. F. Bills, M. Nallinomstead, K. Byrne, *Annu. Rev. Microbiol.* **1995**, *49*, 607 – 639; b) C. A. Jones, P. J. Sidebottom, R. J. Cannell, D. Noble, B. A. Rudd, *J. Antibiot.* **1992**, *45*, 1492 – 1498.
[2] a) E. Baloglu, D. G. I. Kingston, *J. Nat. Prod.* **1999**, *62*, 1448 – 1472; b) K. Walker, R. Croteau, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 13591 – 13596.
[3] J. A. Bjorklund, E. Leete, *Phytochemistry* **1992**, *31*, 3883 – 3887.
[4] J. Piel, K. Hoang, B. S. Moore, *J. Am. Chem. Soc.* **2000**, *122*, 5415 – 5416, and references therein.
[5] K. Gerth, N. Bedorf, H. Irschik, G. Höfle, H. Reichenbach, *J. Antibiot.* **1994**, *47*, 23 – 31.
[6] a) A. F. A. Marsden, B. Wilkinson, J. Cortés, N. J. Dunster, J. Staunton, P. F. Leadlay, *Science* **1998**, *279*, 199; b) M. S. Pacey, J. P. Dirlam, R. W. Geldart, P. F. Leadlay, H. A. I. McArthur, E. L. McCormick, R. A. Monday, T. N. O'Connell, J. Staunton, T. J. Winchester, *J. Antibiot.* **1998**, *51*, 1029 – 1034; P. F. Long, C. J. Wilkinson, C. P. Bisang, J. Cortés, N. Dunster, M. Oliynyk, E. McCormick, H. A. I. McArthur, C. Mendez, J. A. Salas, J. Staunton, P. F. Leadlay, *Mol. Microbiol.* **2002**, *43*, 1215 – 1225; C. Bisang, P. F. Long, J. Cortés, J. Westcott, J. Crosby, A. L. Matharu, R. J. Cox, T. J. Simpson, J. Staunton, P. F. Leadlay, *Nature* **1999**, *401*, 502 – 505; B. S. Moore, C. Hertweck, *Nat. Prod. Rep.* **2003**, 19, 70 – 99, and references therein.
[7] a) T. Schupp, C. Toupet, B. Cluzel, S. Neff, D. S. Hill, J. J. Beck, J. M. Ligon, *J. Bacteriol.* **1995**, *177*, 3673 – 3679; b) J. M. Ligon, D. S. Hill, J. J. Beck, R. Zirkle, I. Molnar, J. Zawodny, S. Money, T. Schupp, *Gene* **2002**, 285, 257 – 267.
[8] J. Cortés, K. E. H. Wiesmann, G. A. Roberts, M. J. B. Brown, J. Staunton, P. F. Leadlay, *Science* **1995**, *268*, 1487 – 1489.
[9] C. J. Wilkinson, E. J. Frost, J. Staunton, P. F. Leadlay, *Chem. Biol.* **2001**, *8*, 1197 – 1208.
[10] P. G. Egland, J. Gibson, C. S. Harwood, *J. Bacteriol.* **1995**, *177*, 6545 – 6551.
[11] a) C. Hertweck, B. S. Moore, *Tetrahedron* **2000**, *56*, 9115 – 9120; b) C. Hertweck, A. P. Jarvis, L. Xiang, B. S. Moore, N. J. Oldham, *ChemBioChem* **2001**, *2*, 784 – 786; c) L. Xiang, B. S. Moore, *J. Bacteriol.* **2003**, *185*, 399 – 404.
[12] J. Garcia-Bernardo, H. Hong, P. F. Leadlay, unpublished results.
[13] a) A. M. Hill, B. L. Thompson, J. P. Harris, R. Segret, *Chem. Commun.* **2003**, 1358 – 1359; b) A. M. Hill, B. L. Thompson, *Chem. Commun.* **2003**, 1360 – 1361.
[14] J. A. Kalaitzis, M. Izumikawa, L. Xiang, C. Hertweck, B. S. Moore, *J. Am. Chem. Soc.* **2003**, *125*, 9290 – 9291.
[15] T. A. Cropp, D. J. Wilson, K. A. Reynolds. *Nat. Biotech.* **2000**, *18*, 980 – 983.
[16] J. Piel, C. Hertweck, P. R. Shipley, D. M. Hunt, M. S. Newman, B. S. Moore, *Chem. Biol.* **2000**, *7*, 943 – 955.
[17] L. Xiang, B. S. Moore, *J. Biol. Chem.* **2002**, *277*, 32505 – 32509.

Received: January 14, 2004

