OBC
www.rsc.org/obc

Heterologous expression in *Saccharopolyspora erythraea* of a pentaketide synthase derived from the spinosyn polyketide synthase †

Christine J. Martin,*[a] Máire C. Timoney,[a,d] Rose M. Sheridan,[a] Steven G. Kendrew,[a] Barrie Wilkinson,[a] James Staunton[a,c] and Peter F. Leadlay[a,b]

[a] *Biotica Technology Ltd, Building 5, Chesterford Research Park, Little Chesterford, Nr Saffron Walden, Essex, UK CB10 1XL. E-mail: Christine.Martin@biotica.co.uk*

[b] *Cambridge Centre for Molecular Recognition and Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge, UK CB2 1GA*

[c] *Cambridge Centre for Molecular Recognition and Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, UK CB2 1EW*

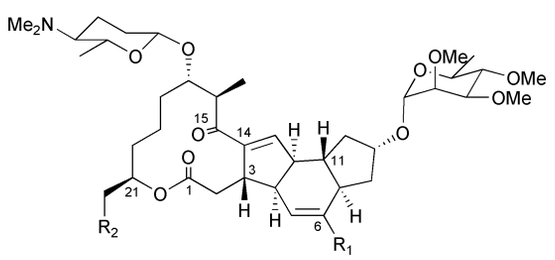
[d] *Pfizer Global Pharmaceuticals, Walton Oaks, Dorking Road, Tadworth, UK KT20 7NS*

Received 5th September 2003, Accepted 14th October 2003
First published as an Advance Article on the web 29th October 2003

A truncated version of the spinosyn polyketide synthase comprising the loading module and the first four extension modules fused to the erythromycin thioesterase domain was expressed in *Saccharopolyspora erythraea*. A novel pentaketide lactone product was isolated, identifying cryptic steps of spinosyn biosynthesis and indicating the potential of this approach for the biosynthetic engineering of spinosyn analogues. A pathway for the formation of the tetracyclic spinosyn aglycone is proposed.

The biosynthetic engineering of natural products in order to access novel chemical space around existing, commercially valuable molecules has emerged as a powerful new technology[1] and provided a paradigm shift in how natural products are utilised within discovery programs.[2] This is particularly evident in the area of polyketide biosynthesis.[3] We have recently applied these technologies to alter the glycosylation pattern around the spinosyns (spn),[4] a family of agriculturally important molecules which form the active constituents of the potent yet environmentally benign insecticide Tracer™ (Dow AgroSciences).[5] To explore the potential for biosynthetic engineering of the spn polyketide synthase (PKS), and thus access novel analogues of these molecules, we have expressed a portion of the spn PKS in the well defined heterologous host *Saccharopolyspora erythraea* JC2.[6]

The *spn* biosynthetic gene cluster has been cloned from *Saccharopolyspora spinosa* and sequenced.[7] Biosynthesis of the polyketide core is proposed to involve the formation of a decaketide precursor, the putative first enzyme-free intermediate, followed by an enzyme-mediated process resembling an intramolecular cycloaddition. The resulting tetracyclic aglycone is then glycosylated with L-rhamnose which is subsequently tri-*O*-methylated, and finally a second sugar, D-forosamine, is added to yield the biologically active compounds **1** and **2** along with other related metabolites in minor amounts (Fig. 1).[7] In the spn structure C15 bears a keto group, although the PKS appears to contain an active ketoreductase domain (KR) in module 4, which governs the fourth cycle of chain elongation, and this KR4 ought in theory to reduce the β-keto group corresponding to C15 to a hydroxy group. Construction of a truncated version of the spn PKS and characterisation of the product(s) formed was expected both to determine the viability of heterologous expression of the PKS, and to indicate the extent of β-keto processing that actually occurs at C15 during polyketide chain assembly, therefore identifying whether the KR domain of module 4 is active or not. If it is active, it would imply hitherto-unrecognised additional steps in spn biosynthesis to bring about the observed oxidation state in the final product.



1 R_1= H, R_2= CH_3, Spinosyn A
2 R_1= CH_3, R_2= CH_3, Spinosyn D
6 R_1 = H, R_2 = H, Spinosyn E

Fig. 1 Structures of spinosyns A, D and E.

A plasmid was engineered in which the erythromycin thioesterase domain (TE) was fused in-frame with, and downstream of, the spn PKS modules 1–4[8] (Fig. 2): an approach used previously for the expression of other truncated PKS systems.[9,10] The truncated, hybrid PKS gene was placed under the control of the actinorhodin promoter (P_{actI}) in a vector containing the *actII*-ORF4 activator, to give the plasmid designated pCJR308. This system is designed to direct expression of the PKS in *S. erythraea* strains.[6] Transformation of *S. erythraea* JC2 with this plasmid,[11] followed by homologous integration *via* the erythromycin TE portion of the PKS yields the strain *S. erythraea* JC2/pCJR308, which was designated BIOT-0966. Cultures of BIOT-0966 were grown and extracted as described previously,[6] and shown by LC-MS and LC-ELS methods to produce six novel compounds that are not present in the control strain *S. erythraea* JC2.

The major novel compound **3** (Fig. 3A) displayed chromatographic and mass spectrum (MS) fragmentation characteristics[12] appropriate for a pentaketide lactone with a hydroxy group, rather than a keto group at C3. This hydroxy group derives from the action of the KR domain in module 4 of the PKS and corresponds to the C15 position of the full spinosyn structure. To verify the structure, this compound was isolated

† Electronic supplementary information (ESI) available: Further details of the construction of pCJR308, the fermentation of BIOT-0966 and the isolation of pentaketide lactone, **3**, and figures showing the ^{13}C NMR and ^{1}H COSY spectra of **3**. See http://www.rsc.org/suppdata/ob/b3/b310740j/

DOI: 10.1039/b310740j

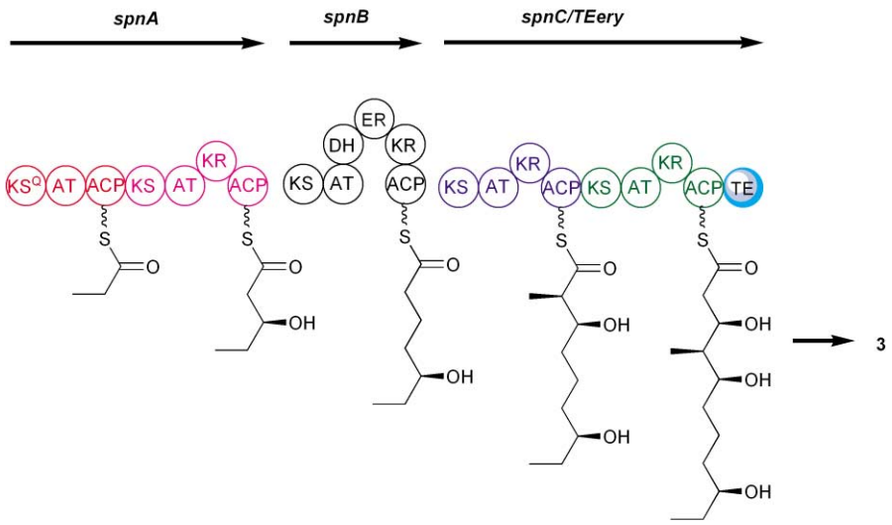


Fig. 2 Modular organisation of the truncated-hybrid pentaketide PKS—the *spn* genes and their direction of transcription are indicated by the bold arrows. The domain structure with their functions in polyketide chain assembly are illustrated below the genes; KS, β-ketoacyl synthase; AT, acyl transferase; ACP, acyl carrier protein; DH, dehydratase; ER, enoyl reductase; KR, β-ketoreductase; TE, thioesterase. Each complete module is indicated by a different colour, with the TE domain of the erythromycin PKS indicated as a solid sphere. The various enzyme bound intermediates produced during chain assembly are shown attached to the relevant ACP domain.

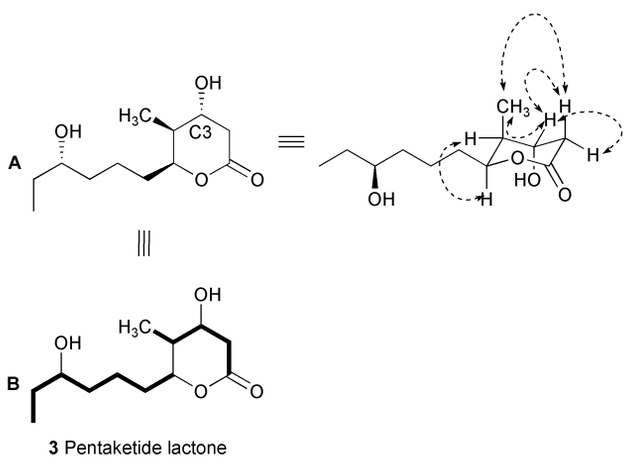


Fig. 3 Absolute configuration of the pentaketide-lactone **3**; **A**, dashed arrows show NOESY correlations; **B**, bold lines show COSY correlations.

from the fermentation broth of BIOT-0966 to yield 20 mg of pure product (see electronic supplementary information †). Electrospray ionization-high resolution-Fourier transform ion cyclotron resonance-MS analysis (ESI-HR-FTICR-MS) identified the molecular formula as $C_{12}H_{22}O_4$ and detailed NMR experiments were performed to provide structural characterisation. The ^{1}H, ^{13}C and DEPT spectra were consistent with the expected structure for a pentaketide lactone.[13] The COSY spectrum of **3** provided all of the scalar couplings expected, apart from one between H4 and H5 (Fig. 3B). Based on the absolute configuration of **1** & **2** the stereochemistry at all positions of **3** can be deduced, except for that at C3, the equivalent of which (C15) bears a keto group in **1** & **2**. Therefore, only two isomers of **3** are possible. The coupling constants observed for **3** are most consistent with the absolute configuration as shown (Fig. 3) and this structure was confirmed using a NOESY experiment. Of most importance was the strong NOE enhancement for H-3 to the C4-methyl group, and the enhancement for H-4 to H-5 (Fig. 3A). Compound **3** is therefore identified as 3*R*-hydroxy-4*S*-methyl-5*S*,9*S*-hydroxy-undecanoic acid-δ-lactone, indicating that the KR domain of module 4 is indeed functional.

Recent publications have addressed identification of amino acid residues in KR domains that are indicators of the stereochemical outcome of ketoreduction.[14,15] An alignment of the spinosyn KR domains (Fig. 4) shows that these motifs predict that in all positions where there is further processing of the hydroxy group, and the outcome of reduction is obscure, ketoreduction results in a "B-type" stereochemical outcome as expected. In the case of spinosyn module 4, the stereochemical outcome of the ketoreduction is predicted by these motifs to be "A-type"[14] and that is what is observed in the isolated product **3**. Although we cannot completely exclude the possibility of reduction by an exogenous enzyme, as has been observed when a polyketide synthase was expressed in *S. coelicolor*,[16] since all of the active-site residues expected for an active KR are present in spinosyn KR4,[15] we consider this to be unlikely.

```
A-type
.....................................................W.........
DEBS  KR2  HAAGLPQQVAINDMDE   .....   FSSGAGVWGSARQGAY
SPN   KR1  HAAGVIETGDAAAMSL   .....   FSSVSGVWGAGGHGAY
SPN   KR3  HAAGVGRLLPLAETDQ   .....   FSSGAGVWGGGGQGAY
SPN   KR4  HAAGVGEAGDVVEMGL   .....   FSSVSGVWGAGGQGAY
SPN   KR6  HAAGVVEVGDVASMCL   .....   FSSVSGVWGAGGQGAY
SPN   KR7  HTAGVGTPANLAETTL   .....   FSSISGVWGAGGQGAY

B-type
          .....LDD.............................P...N.
DEBS  KR1  HAAATLDDGTVDTLTG   .....   FSSFASAFGAPGLGGY
SPN   KR2  HAAGVLDDGVSESLTV   .....   FSSVSGVLGSGGQGNY
SPN   KR5  HTAGVLADGVIGSLSA   .....   FSSFSGIAGAAGQANY
SPN   KR8  HAAGALADGVVESLTP   .....   FSSFSGVAGAAGQGNY
SPN   KR9  HAAGALADGVIESLSP   .....   FSSLSGVAGAPGQGNY
SPN   KR10 HAAGVLDDGVLMSMSP   .....   FSSVAGLFGGAGQSNY
```

Fig. 4 Alignment of the 10 spinosyn KR domains with DEBS KRs 1 and 2. The regions shown are amino acid residues 88–103 and 134–149. Residues highlighted in bold type are those proposed by Caffrey[14] to be indicators of KR specificity.

The identification of a C15-hydroxy group in the immediate product of the spn PKS in turn means that subsequent oxidation must occur to establish the keto group at C15 during spinosyn biosynthesis. This has interesting consequences for the mechanism by which the bridged tetracyclic aglycone **4** is produced from the putative first enzyme free intermediate **5** (Fig. 5). In a recent report regarding the biosynthesis of similarly bridged tetracyclic fungal polyketides,[17] Oikawa suggests a mechanism whereby oxygenation of olefinic moieties within a putative polyolefinic precursor is utilised for the generation of carbocations, thereby giving rise to tetracyclic systems *via* oxidative cyclisation (and isomerisation) in a process akin to those involved in the biosynthesis of terpenoids.[17] Oikawa has drawn an analogy to the similarly-bridged tetracyclic spn system, and

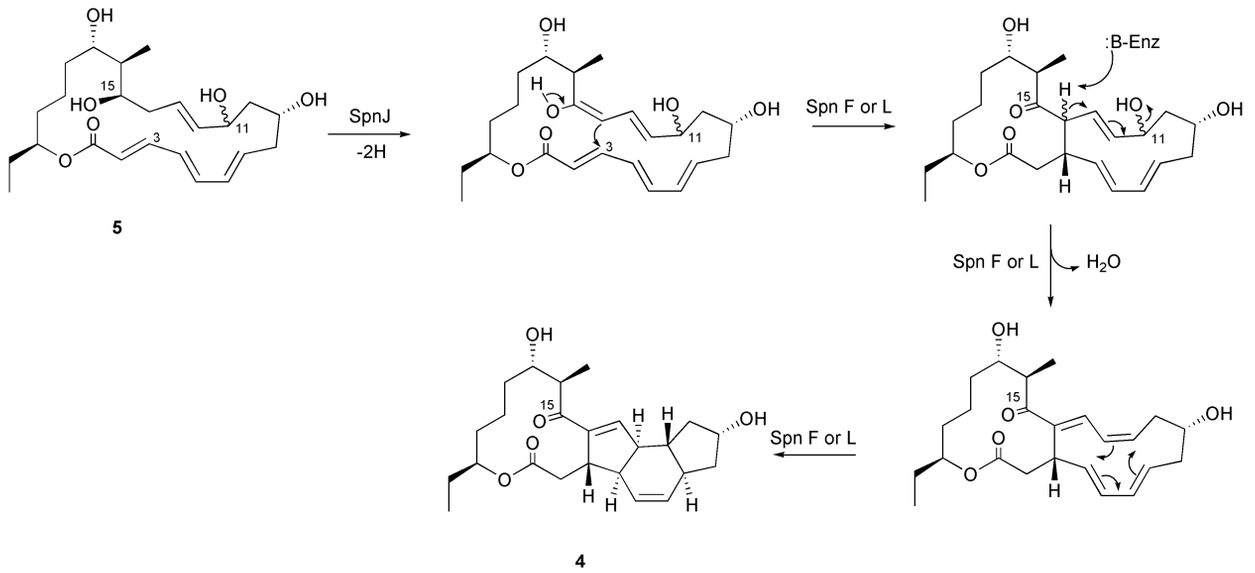


Fig. 5 Proposed mechanism for formation of the spn tetracyclic aglycone **4** from the putative PKS product **5**.

suggested a similar process here. However, we believe the process for spn to be more complicated: first, generation of the bridged spn system would appear to require two independent processes, with a C14–C3 carbon bridge most probably being formed by Michael attack from C14 upon C3 (Fig. 5). The remaining carbon bridges could then be formed through a Diels–Alder cyclisation if the remaining C14 proton is removed with concomitant isomerisation of the C13–C12 double bond and elimination of the hydroxyl group at C11. The driving force for such a reaction sequence would be provided by deprotonation at C14. This would be facilitated by the presence of a keto group at C15 that increases the acidity of the C14 protons and allows enolisation to drive the proposed Michael addition in the formation of the first bridge. The formation of the C15 keto group after full polyketide chain assembly may arise in order to ensure that the timing and order of the proposed bridging reaction, isomerisation/dehydration and subsequent cyclisation are correct. However, in the absence of experimental confirmation, the order of these steps remains speculative and equally plausible alternative schemes can be drawn.

Nevertheless, the series of biosynthetic steps proposed here would be consistent with the putative gene products identified within the spn biosynthetic cluster and shown to be involved in the bridging process.[7] After cyclisation and release from the PKS the first enzyme-free intermediate would be oxidised to give a keto group at C15 by SpnJ. SpnJ shares closest homology to hexose oxidases and berberine bridge forming oxidases.[7] In the next step one of the putative carbon–carbon bond forming enzymes SpnF or SpnL (which share closest sequence similarity with methyl transferases (MeT) from secondary metabolism, but which lack some of the key MeT residues) could catalyse an intramolecular Michael addition reaction, linking C15 and C3.[7] The final step of the process would involve the elimination/isomerisation process followed by a Diels–Alder cyclisation as described above, and could be catalysed by one of SpnL or SpnF.

In addition to **3**, we also observed the presence of other novel products produced by BIOT-0966. These compounds were produced at low levels (<5% total novel products) and therefore not isolated. Analysis by LC-MS indicated that they are most probably analogues of **3** in which some of the AT domains have incorporated alternative extender units (for example the misincorporation of malonyl-CoA rather than methylmalonyl-CoA by the loading module of SpnA—this is analogous to the production of spinosyn E **6** by *S. spinosa* (Fig. 1)). Also present were open-chain acid forms of these lactones and **3**.

In summary, we have demonstrated that spn PKS genes from *S. spinosa* can be expressed heterologously in *S. erythraea* JC2 and can be used as part of a functional hybrid PKS. These findings indicate that the heterologous expression of engineered versions of the spn PKS may provide a viable approach to access novel analogues of these important and valuable molecules. In addition the isolation and structural elucidation of the major new product **3** has led to the determination that the KR domain of module 4 is very likely active as predicted from its structure, and therefore identifies the otherwise cryptic step of C15 oxidation that must occur during spn biosynthesis.

We thank Dow AgroSciences for the kind provision of cosmids and the stimulating exchange of ideas, Dr Rachel Lill (Biotica Technology Ltd) for assistance with structural determination, Dr Paul Gates (University of Cambridge) for help with FT-ICR-MS/MS experiments and Mr John Lester and the DNA-sequencing facility (University of Cambridge).

Notes and references

1 M. D. Burkart, *Org. Biomol. Chem.*, 2003, **1**, 1–4.
2 C. Khosla, *J. Org. Chem.*, 2000, **65**, 8127–8133.
3 J. Staunton and B. Wilkinson, *Curr. Opin. Chem. Biol.*, 2001, **5**, 159–164.
4 S. Gaisser, C. J. Martin, B. Wilkinson, R. M. Sheridan, R. E. Lill, A. J. Weston, S. J. Ready, C. Waldron, G. D. Grouse, P. F. Leadlay and J. Staunton, *Chem. Commun.*, 2002, 618–619.
5 G. D. Crouse and T. C. Sparks, *Rev. Toxicol. (Amsterdam)*, 1998, **2**, 133–146.
6 C. J. Rowe, J. Cortés, S. Gaisser, J. Staunton and P. F. Leadlay, *Gene*, 1998, **216**, 215–223.
7 C. Waldron, P. Matsushima, P. R. Rosteck Jr., M. C. Broughton, J. Turner, K. Madduri, K. P. Crawford, D. J. Merlo and R. H. Baltz, *Chem. Biol.*, 2001, **8**, 487–499.
8 An *Mfe*I site was introduced between the spinosyn PKS module 4 acyl carrier protein domain and the erythromycin module 6 thioesterase by PCR using the following pair of oligonucleotides: SP01 TGCCGGGGCTGAGCACGTGGTGTTGA SP02 TTTCAATTGGGCGCAGAGGTGACGCGACAGCGACT. The truncated PKS was constructed as a DNA fragment with an *Nde*I site overlapping with the start codon and an *Xba*I site following the stop codon to facilitate cloning into pCJR24,[6] an appropriate expression vector for this system. In all cases the cosmids pRHB9A6 and pRHB3E11[7] were used as templates for PCR and as sources for isolating DNA fragments. Further details are provided in the electronic supplementary information †.
9 J. Cortés, K. E. H. Wiesmann, G. A. Roberts, M. J. B. Brown, J. Staunton and P. F. Leadlay, *Science*, 1995, **268**, 1487–1489.
10 Z. A. Hughes-Thomas, C. B. W. Stark, I. U. Böhm, J. Staunton and P. F. Leadlay, *Angew. Chem., Int. Ed.*, 2003, **42**, 4475–4478.

11 Adapted from: J. M. Weber and R. Losick, *Gene*, 1988, **68**, 173–180.
12 K. J. Weissman, G. C. Kearney, P. F. Leadlay and J. Staunton, *Rapid Commun. Mass Spectrom.*, 1999, **13**, 2103–2108.
13 The preparation and fermentation of strain BIOT-0966 along with the isolation procedure for **3** is described in the electronic supplementary information. † The homogeneity of **3** was verified using HPLC coupled to UV and MS detection methods. NMR spectra were recorded in $CDCl_3$ with chemical shift values (δ) referenced to the residual proton of $CHCl_3$ (7.26 ppm) or the carbon atom of $CDCl_3$ (77.0 ppm). *J* values are given in Hz. δ_H (500 MHz) 0.94 (3 H, t, ^3J 7.3, 11-CH_3), 0.95 (3 H, d, ^3J 7.3, 4-CH_3), 1.42 (2 H, m, 10-H), 1.50 (2 H, m, 7-H), 1.50 (1 H, m, 6-H_a), 1.65 (2 H, m, 8-H), 1.74 (1 H, m, 6-H_b), 1.95 (1 H, dddq, ^3J 3.4, 7.3 and 8.1, ^4J 0.9, 4-H), 2.54 (1 H, ddd, ^2J 18.3, ^3J 3.4, ^4J 0.9, 2-H_{eq}), 2.82 (1 H, dd, ^2J 18.3, ^3J 5.6, 2-H_{ax}), 3.55 (1 H, m, 9-H), 4.06 (1 H, ddd, ^3J 3.4, 3.4 and 5.6, 3-H) and 4.72 (1 H, ddd, ^3J 2.6, 4.7 and 8.1, 5-H); δ_C (125 MHz; ^{1}H decoupled) 9.8, 10.2, 21.6, 30.2, 31.7, 35.7, 36.4, 37.2, 68.5, 73.0, 78.0 and 170.6; *m*/*z* (ESI-HR-FT-ICR-MS) 253.1422 (MNa^+. $C_{12}H_{22}O_4Na$ requires 253.1410); *m*/*z* (ESI-LCMS) 135.1 (7%), 153.0 (5), 177.1 (9), 195.2 ($[MH - 2H_2O]^+$ (18)), 213.3 ($[MH - H_2O]^+$ (27)), 231.2 (MH^+ (57)), 253.3 (MNa^+ (100)), 269.3 (MK^+ (10)) and 482.9 ($[M_2Na]^+$ (58)).
14 P. Caffrey, *ChemBioChem*, 2003, **4**, 654–657.
15 R. Reid, M. Piagentini, E. Rodriguez, G. Ashley, N. Viswanathan, J. Carney, D. V. Santi, C. R. Hutchinson and R. McDaniel, *Biochemistry*, 2003, **42**, 72–79.
16 C. M. Kao, M. McPherson, R. N. McDaniel, H. Fu, D. E. Cane and C. Khosla, *J. Am. Chem. Soc.*, 1997, **119**, 11339–11340.
17 H. Oikawa, *J. Org. Chem.*, 2003, **68**, 3552–3557.