Natural Metabolic Diversity Encoded by the Enterocin Biosynthesis Gene Cluster

Jörn Piel,[†,‡] Khanh Hoang,[†] and Bradley S. Moore*[,†,§]

Department of Chemistry, Box 351700
University of Washington
Seattle, Washington 98195-1700
Division of Medicinal Chemistry
College of Pharmacy, University of Arizona
Tucson, Arizona 85721-0207

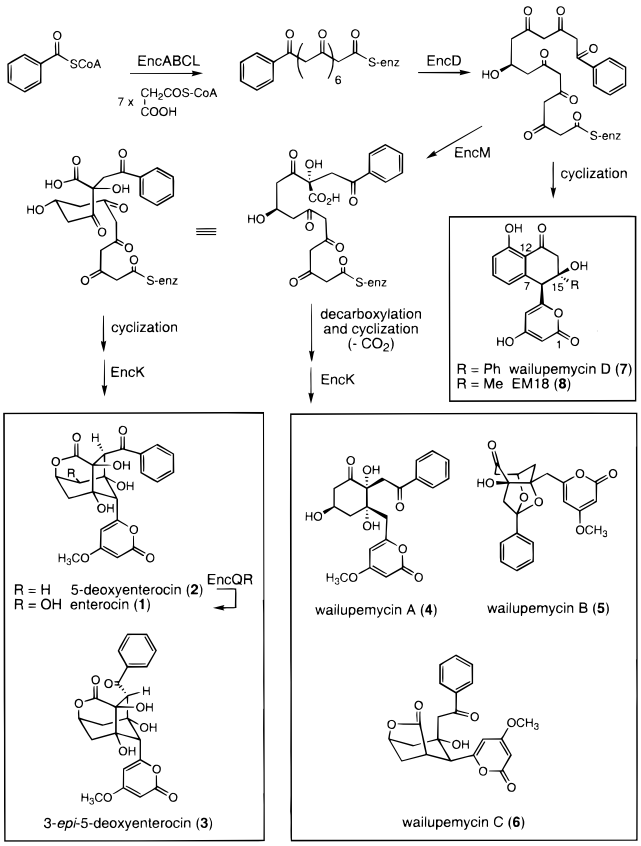
Received February 7, 2000

The marine bacterium "*Streptomyces maritimus*" produces a structurally diverse series of bacteriostatic polyketides known as the enterocins and wailupemycins (Scheme 1).[1,2] The predominant metabolite is the branched octaketide enterocin (**1**), which has additionally been isolated from edaphic streptomycetes[3] and a marine invertebrate.[4] 5-Deoxyenterocin (**2**) and 3-*epi*-5-deoxyenterocin (**3**) differ from enterocin only in hydroxyl group functionalization at C5 and the stereochemistry at C3. Three additional α-pyrone containing metabolites, wailupemycins A−C (**4**-**6**), are hypothetically derived through decarboxylation of a common branched intermediate. This family of metabolites is possibly produced from a single biosynthetic pathway that has numerous metabolic options in creating molecular diversity. Does "*S. maritimus*" generate this unusual set of polyketides from a distinct pathway, and if so, what is the genetic organization of the encoding biosynthetic gene cluster? Examination of how metabolic diversity is generated in natural systems should provide insight into engineering a wide range of novel molecules through combinatorial biosynthesis.

Enterocin (**1**) is derived from an uncommon benzoate starter unit and seven malonate molecules and undergoes a rare Favorskii-like carbon rearrangement.[3c] The α-pyrone moiety is reminiscent of a large number of mutant and recombinant aromatic polyketides generated from iterative type II polyketide synthase (PKS) systems and served to indicate that this family of metabolites is probably synthesized via an aberrant type II rather than a modular type I polyketide pathway.[5] Southern blot hybridization analysis with a type II PKS gene probe (*act*I-ORF1) revealed three distinct gene sets in the genome. End-sequencing of *Bam*HI restriction fragments from corresponding cosmids identified a minimal PKS gene set (*enc*) clustered with relevant genes putatively involved in the formation of the benzoic acid starter unit.

The cosmid clone pJP15F11 containing the *enc* cluster was heterologously expressed in the genetically engineered host strain *S. lividans* K4-114.[6] LC-MS analysis of organic extracts from K4-114/pJP15F11-1 unequivocally demonstrated that the transconjugant yielded **1** as the major polyketide metabolite (Figure 1).[7]

Scheme 1. Structure and Proposed Biosynthesis of Enterocin (**1**), 5-Deoxyenterocin (**2**), 3-*epi*-5-Deoxyenterocin (**3**), and Wailupemycins A−D (**4**−**7**) in "*S. maritimus*"[a]



[a] The carbons in **7** are labeled according to their number in the polyketide backbone. Proposed functions of the *enc* gene products are shown for each catalytic reaction.[13] The structure of EM18 (**8**) is provided for comparison with **7**. The stereochemistry of **7** is relative and is unknown in **8**.

Although **4**−**6** were not detected in either the transconjugant or the wild type "*S. maritimus*" under these growth conditions, **2** and several previously uncharacterized analogues were clearly evident by LC-MS analysis. To readily ascertain the breadth of expressed polyketides, [ring-d_5]benzoic acid was administered to K4-114/pJP15F11-1, and components from the resulting extract were analyzed for deuterium enrichment by LC-MS. Ten metabolites were indeed enriched with labeled benzoate and arranged into three isomeric groups of mass 414, 346, and 364. Consequently, all of the genetic information required for *S. lividans* K4-114 to synthesize enterocin and a large series of structural analogues was encoded on the cosmid pJP15F11.

* Address correspondence to this author at the University of Arizona. E-mail: moore@pharmacy.arizona.edu.
† University of Washington.
‡ Present Address: Max Planck Institute of Chemical Ecology, Carl-Zeiss-Promenade 10, 07745 Jena, Germany.
§ University of Arizona.

(1) Sitachitta, N.; Gadepalli, M.; Davidson, B. S. *Tetrahedron* **1996**, *52*, 8073.
(2) The bacterium has recently been described as a new species within the genus *Streptomyces* on the basis of cultural characteristics, physiological characteristics, and 16S rDNA analysis (AF233338).
(3) Miyairi, N.; Sakai, H. I.; Konomi, T.; Imanaka, H. *J. Antibiot.* **1976**, *29*, 227. (b) Tokuma, Y.; Miyairi, N.; Morimoto, Y. *J. Antibiot.* **1976**, *29*, 1114. (c) Seto, H.; Sato, T.; Urano, S.; Uzawa, J.; Yonehara, H. *Tetrahedron Lett.* **1976**, 4367.
(4) Kang, H.; Jensen, P. R.; Fenical, W. *J. Org. Chem.* **1996**, *61*, 1543.
(5) For recent reviews on type II iterative PKSs, see: Hopwood, D. A. *Chem. Rev.* **1997**, *97*, 2465. Rawlings, B. J. *Nat. Prod. Rep.* **1999**, *16*, 425.
(6) *S. lividans* K4-114 was generously provided by Dr. M. C. Betlach, Kosan Biosciences, Inc., Hayward, CA. This host lacks the entire actinorhodin biosynthesis gene cluster and displays a negligible level of background polyketide synthesis. (Ziermann, R.; Betlach, M. C. *BioTechniques* **1999**, *26*, 106).
(7) *S. lividans* K4-114/pJP15F11-1 was grown on R2YE agar plates containing 100 μg/mL apramycin for 1 week. Half of the plates were overlaid with d_5-benzoic acid (0.82 mM). As a control, the untransformed host was grown on an R2YE plate, and the wild-type enterocin producer "*S. maritimus*" was grown on a d_5-benzoate-containing A1 plate. The crude EtOAc extracts were subjected to silica gel flash chromatography and eluted with EtOAc and MeOH. The fractions containing enterocin-like compounds were combined, dried, dissolved in MeOH (0.5 mL), and analyzed (3 μL). A Hewlett-Packard HP1100 LC pump with a diode-array detector was linked to a Bruker HP Esquire LC mass spectrometer operating in the negative ion mode. A Beckman Ultrasphere C18 column (4.6 mm × 15 cm) was used at a flow rate of 1 mL/min with a linear solvent gradient of 5% MeOH in 0.15% trifluoroacetic acid−water to 100% MeOH over a period of 45 min.

10.1021/ja000445n CCC: $19.00
Published on Web 05/20/2000

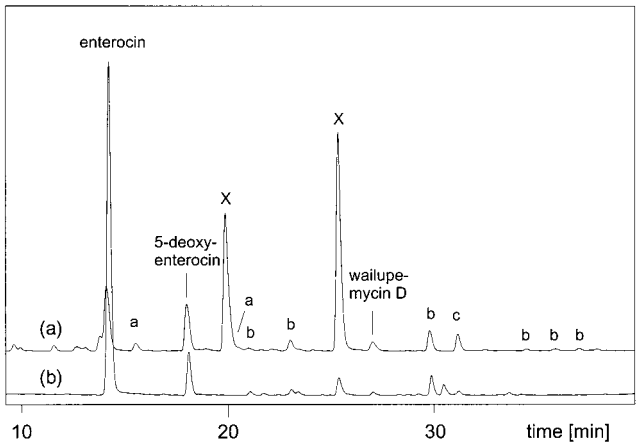


Figure 1. LC-MS analysis of the polyketide fraction of (a) *S. lividans* K4-114/pJP15F11-1 and (b) "*S. maritimus*". HPLC was monitored at 254 nm. Enc-derived polyketides were labeled with [ring-d_5]benzoic acid to aid in MS identification. Peaks labeled with letters "a" [$(M - 1)^+ =$ 413/418], "b" [$(M - 1)^+ = 363/368$], and "c" [$(M - 1)^+ = 345/350$] are uncharacterized. The two peaks in chromatogram (a) marked "X" are not shifted by 5 mass units and thus are not *enc*-derived products.

Fermentation of "*S. maritimus*" in the saltwater-based medium A1 followed by extraction and chromatography provided a new polyketide in 2.5 mg/L that was also present in the transconjugant.[8] Wailupemycin D (**7**) is optically active ($[\alpha]^{25}_D$ −5.9°, *c* 0.05, MeOH) and has the molecular composition $C_{21}H_{16}O_6$ as established by high-resolution FAB MS (*m*/*z* $[M + H]^+$ 365.1023 obs, 0.2 mmu error). Analysis of the proton and carbon NMR spectra, with the aid of gradient-enhanced HMBC data, clearly indicated that **7** contained α-pyrone and monosubstituted benzene residues common to enterocin.[9] Comparison with the structurally related *act*/*whiE* recombinant octaketide EM18[10] (**8**) indicated that they only differed in the substitution at C15, where **7** carries a phenyl unit and **8** contains a methyl group. Key HMBCs unambiguously verified the structure of the α-tetralone unit and the attachment of the two six-membered rings. Furthermore, the NMR analysis suggests that C8, C9, C10, and C12 in the structurally analogous **8** are misassigned. The relative stereochemistry of **7** was established by 2D NOESY providing an anti relationship between the pyrone and phenyl rings. MS analysis of K4-114/pJP15F11-1 indicated that six additional isomers of **7** and one dehydrated analogue were also produced.

Wailupemycin D (**7**) is probably derived from a linear polyketide intermediate, as in the case of EM18 (**8**),[10] and provides further evidence that the enterocin family of polyketides is derived from a carbon rearrangement of a type II PKS-generated linear poly-β-ketide. We propose that the unprecedented α-oxidation and Favorskii-like rearrangement take place on the linear octaketide (Scheme 1). This rare metabolic process has been proposed on the basis of labeling experiments in only a few other natural product pathways[11] and conceivably involves the action of an oxygenase ("favorskiiase"). The favorskiiase probably complexes with the minimal *enc* PKS since poly-β-ketides are notoriously reactive[12] and would likely spontaneously cyclize if passed onto a nonassociated protein. The rearrangement, however, probably takes place after the C9 ketoreduction to account for the formation of **7**. In addition to catalyzing the Favorskii-like rearrangement, the favorskiiase may induce the cyclization of the branched intermediate by limiting its conformational mobility and directing the intramolecular aldol condensations. Spontaneous decarboxylation and cyclization of this highly labile intermediate would account for the branch in the pathway providing the minor byproducts **3**−**6**.

The nucleotide sequence of the *enc* cluster has recently been completed and consists of 20 open reading frames spanning approximately 21.3 kb.[13] Noticeably absent from the octaketide-homologous type II PKS gene cluster are cyclase or aromatase genes. This significant architectural difference between the *enc* cluster and all known bacterial type II PKS gene clusters provides insight to the observed spontaneity of the *enc* pathway. Cyclases are integral components of type II PKS systems and catalyze regiospecific aldol condensations from a myriad of chemically feasible reactions on highly reactive poly-β-keto intermediates. When cyclases are unnaturally removed from their coupled minimal PKS, spontaneous chemistry can result as the minimal PKS relies on the stabilizing effects of the cyclase and other subunits to ensure proper assembly.[12] Furthermore, all of the genetic information for the synthesis and attachment of the rare bacterial benzoyl-CoA starter unit is arranged on the *enc* cluster.

In summary, the *enc* gene set encodes one of the most versatile polyketide biosynthetic pathways investigated to date in naturally generating metabolic diversity. Its small size and novel features provide the foundation for engineering hybrid expression systems with more typical type II PKS gene sets in the production of diverse ranges of novel compounds for use in drug discovery.[14] This work also represents the first expression of a marine bacterial polyketide in a terrestrial heterologous host[15] and validates that marine bacterial metabolites can be produced in nonmarine expression systems.

Acknowledgment. We thank B. S. Davidson (Utah State University) for the strain "*S. maritimus*" and standard samples of enterocin and 5-deoxyenterocin, M. Sadilek (University of Washington) for LC-MS assistance, and G. Trimurtulu (University of Hawaii) for measuring the optical rotation of wailupemycin D. This work was supported by the National and Washington Sea Grant programs (NA76RG01149, project no. R/B-28). J.P. is a recipient of a Feodor Lynen Postdoctoral fellowship from the Humboldt Foundation.

JA000445N

(8) "*S. maritimus*" was grown in 500 mL Erlenmeyer flasks containing 100 mL of A1 media on a rotary shaker at 200 rpm for 3 d at 26 °C. The liquid culture (2 L) was extracted with EtOAc (3 × 500 mL), and the resulting crude extract was subjected to silica flash chromatography using hexane/EtOAc/MeOH mixtures. Further purification of mixed fractions was achieved by reversed-phase HPLC (Alltech Ultrasphere C18, 5 μ, 4.6 mm × 25 cm) with a 63−100% MeOH in 0.1% acetic acid solvent gradient run at 4.5 mL/min to yield wailupemycin D (5 mg).

(9) ^{1}H NMR (DMSO-d_6) δ (multiplicity, assignment, coupling constants in Hertz, *HMBCs*, **NOESYs**): 3.04 (d, H14$_S$, J = 18.1 Hz, *C6, C12, C13, C15, C16*, **H14$_R$**), 3.65 (d, H14$_R$, J = 18.1 Hz, *C6, C13, C15*, **H6**, **H14$_S$**, **H17/17′**), 4.72 (s, H6, *C4, C5, C7, C8, C12, C14, C15, C16*, **H4**, **H8**, **H9**, **H17/17′**), 5.01 (br s, H2, *C1, C3, C4*), 6.00 (s, C15 hydroxyl), 6.09 (br s, H4, *C2, C5, C6*, **H6**, **H8**, **H9**), 6.60 (d, H8, J = 8.2 Hz, *C6, C10, C12*, **H4**, **H6**, **H9**), 6.89 (d, H10, J = 8.2 Hz, *C8, C11, C12*, **H9**), 7.22 (m, H19, *C17/17′*, **H18/18′**), 7.31 (m, H18/18′, *C16, C18*, **H17/17′**, **H19**), 7.48 (m, H17/17′, *C15, C17, C19*, **H6**, **H14$_R$**, **H18/18′**), 7.50 (m, H9, *C7, C11*, **H8**, **H10**), 12.4 (s, C11 hydroxyl, *C10, C11, C12*); ^{13}C NMR (DMSO-d_6) δ 51.5 (C14), 55.4 (C6), 76.2 (C15), 89.9 (C2), 105.3 (C4), 117.2 (C12), 117.3 (C10), 120.4 (C8), 126.3 (C17/17′), 128.2 (C19), 128.9 (C18/18′), 137.4 (C9), 143.3 (C7), 145.9 (C16), 162.3 (C11), 164.1 (C5), 164.7 (C1), 171.6 (C3), 204.5 (C13).

(10) Alvarez, M. A.; Fu, F.; Khosla, C.; Hopwood, D. A.; Bailey, J. E. *Nat. Biotechnol.* **1996**, *14*, 335.

(11) Kakinuma, K.; Uzawa, J.; Uramoto, M. *Tetrahedron Lett.* **1982**, *23*, 5303. Wright, J. L. C.; Hu, T.; McLachlan, J. L.; Needham, J.; Walter, J. A. *J. Am. Chem. Soc.* **1996**, *118*, 8757. Fuchser, J.; Zeeck, A. *Liebigs Ann./Recl.* **1997**, 87.

(12) Shen, Y.; Yoon, P.; Yu, T.-W.; Floss, H. G.; Hopwood, D. A.; Moore, B. S. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3622.

(13) The GenBank accession number for the *enc* cluster is AF254925.

(14) Cane, D. E.; Walsh, C. T.; Khosla, C. *Science* **1998**, *282*, 63. Hutchinson, C. R. *Curr. Opin. Microbiol.* **1998**, *1*, 319.

(15) Moore, B. S. *Nat. Prod. Rep.* **1999**, *16*, 653.