EncM, a versatile enterocin biosynthetic enzyme involved in Favorskii oxidative rearrangement, aldol condensation, and heterocycle-forming reactions

Longkuan Xiang*, John A. Kalaitzis*, and Bradley S. Moore*†‡

*Department of Pharmacology and Toxicology, College of Pharmacy, and †Department of Chemistry, University of Arizona, Tucson, AZ 85721

Edited by Stephen J. Benkovic, Pennsylvania State University, University Park, PA, and approved September 27, 2004 (received for review July 29, 2004)

The bacteriostatic natural product enterocin from the marine microbe "Streptomyces maritimus" has an unprecedented carbon skeleton that is derived from an aromatic polyketide biosynthetic pathway. Its caged tricyclic, nonaromatic core is derived from a linear poly-β-ketide precursor that formally undergoes a Favorskii-like oxidative rearrangement. In vivo characterization of the gene encM through mutagenesis and heterologous biosynthesis demonstrated that its protein product not only is solely responsible for the oxidative C–C rearrangement, but also facilitates two aldol condensations plus two heterocycle forming reactions. In total, at least five chiral centers and four rings are generated by this multifaceted flavoprotein. Heterologous expression of the enterocin biosynthesis genes encABCDELMN in Streptomyces lividans resulted in the formation of the rearranged metabolite desmethyl-5-deoxyenterocin and the shunt products wailupemycins D–G. Addition of the methyltransferase gene encK, which was previously proposed through mutagenesis to additionally assist EncM in the Favorskii rearrangement, shifted the production to the O-methyl derivative 5-deoxyenterocin. The O-methyltransferase EncK seems to be specific for the pyrone ring of enterocin, because bicyclic polyketides bearing pyrone rings are not methylated in vivo. Expression of encM with different combinations of homologous actinorhodin biosynthesis genes did not result in the production of oxidatively rearranged enterocin–actinorhodin hybrid compounds as anticipated, suggesting that wild-type EncM may be specific for its endogenous type II polyketide synthase or for benzoyl-primed polyketide precursors.

Materials and Methods

General Experimental Details. NMR spectra were recorded on Bruker (Billerica, MA) DRX-300 and DRX-600 spectrometers. 1H and 13C chemical shifts were referenced to the solvent peak (DMSO-d6) δ 2.49 and 39.5 ppm, respectively. Standard parameters were used for 1D and 2D NMR spectra obtained, which included 1H, 13C, heteronuclear multiple quantum correlation (HMQC), and heteronuclear multiple-bond correlation (HMBC). Optical rotation was measured on a Jasco (Easton, MD) P-1010 polarimeter. A Waters 600 pump equipped with a Waters 2487 dual λ absorbance detector was used for semi-preparative HPLC separations. A Waters 996 photodiode array detector was used for analytical HPLC. High-resolution fast atom bombardment MS (HRFABMS) were recorded on a JEOL RX110A high resolution mass spectrometer at the Mass Spectrometry Facility, Department of Chemistry, University of Arizona.

Bacterial Strains and General Techniques for DNA Manipulation. Streptomyces lividans K4-114 was used as a host for transformation of plasmids (11). Escherichia coli XL1-Blue was used for the...

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: ACP, acyl carrier protein; KR, ketoreductase; PKS, polyketide synthase; HMBC, heteronuclear multiple-bond correlation; HRFABMS, high-resolution fast atom bombardment MS.

*To whom correspondence should be addressed. E-mail: moore@pharmacy.arizona.edu.

© 2004 by The National Academy of Sciences of the USA
manipulation of plasmid DNA. *Streptomyces coelicolor* A3(2) was used to extract genomic DNA for cloning act genes, and the cosmids clone pJP15F11 (5), which harbors the entire enc gene cluster, was used as the source of DNA for cloning enc genes. The *E. coli–Streptomyces* shuttle pSEK4 and pRM5-based vectors were used for all expression experiments in *Streptomyces* (3). Recombinant DNA procedures were performed by standard techniques (12, 13). Restriction enzyme-digested DNA fragments were recovered from agarose gels with the QIAquick DNA Purification Kit (Qiagen, Valencia, CA). Oligonucleotides were obtained from Sigma Genosys. PCR was carried out on a PTC-2000 Peltier thermal cycler (MJ Research, Cambridge, MA) with PfuTurbo (Stratagene) DNA polymerase. DNA sequencing by BigDye terminator cycle sequencing reagents was performed on a PTC-2000 Peltier thermal cycler (MJ Research, Cambridge, MA) with PfuTurbo (Stratagene) DNA polymerase. DNA sequencing was carried out on a PTC-2000 Peltier thermal cycler (MJ Research, Cambridge, MA) with PfuTurbo (Stratagene) DNA polymerase.

**Fig. 1.** Proposed biosynthetic pathway to enc-based polyketides (1-10) and structures of analogous act-derived polyketides (11-13). Path A involves monooxidation of 14 to the trione intermediate 16, whereas path A’ involves a dioxygenase role of EncM in which the C7–C12 olefin of the aldol product 15 is oxidatively cleaved to 16. The Favorskii-like rearrangement is speculated to occur on a putative ACP (EncC)-bound intermediate. The relative size of the reaction arrows reflects the flux in the pathway to the major product 1 as can be observed chromatographically (see Fig. 2). Dashed lines in structures 16 and 17 represent newly created bonds. In compounds 14–17, the carbon numbering is based on linear 14 in which priority is given to the carboxyl carbon; the numbering in compounds 1, 9, and 16 is in contrast based on that defined in ref. 18 for 9.

**Construction of Expression Plasmids.** pBM43 is a derivative of pMP6 (14) in which the actI promoter was replaced by a 1.6-kb fragment harboring encDK. The encDK cassette was PCR amplified from the cosmids clone pJP15F11 (5) with the primers 5'-GGTTAATTAACCGGCGG-3' (forward, *PacI* site underlined) and 5'-GGTTAATTAACCGGCGG-3' (reverse). The resulting 1.6-kb fragment was ligated into the *XbaI* and blunt-ended *NdeI* sites of pGFP7505 (15) to create pBM42. A 3.2-kb *PacI–HindIII* fragment containing encDK and the divergent act promoters was cloned into analogous sites of pMP6 to create pBM43.

pBM47 is a pSEK4 derivative in which a 6.1-kb *PacI/EcoRI* fragment harboring actI-ORFsI,II,III and encMN was cloned into the analogous sites. A 3.0-kb fragment containing actI-ORFsI,II,III was PCR amplified from *S. coelicolor* A3 (2) genomic DNA by using primers 5'-GGTTAATTAACGCGGCGG-3' (forward, *PacI* site underlined) and 5'-GGTTAATTAACGCGGCGG-3' (reverse, *XbaI* site underlined) and cloned into pCR-Blunt to create pBM44. The 3.0-kb *XbaI–HindIII* fragment harboring actI-ORFsI,II,III from pBM44 was ligated into the analogous sites of pBM45 to create pBM46. pBM45 is a pCR-Blunt-derived plasmid in which an *XbaI* and *NheI* 3.1-kb fragment containing encMN was cloned into the analogous sites. encMN was PCR amplified from pJP15F11 by using primers 5'-GACTCTAGAGCCCGTGCGCCGGC-3' (forward, *XbaI* site underlined) and 5'-GCCGGCTGCTGGTGCCG-3' (reverse, *NheI* site underlined).

PacI–HindIII fragments from pMP6 (harboring encD), pBM43 (harboring encDK), and pRM5 (harboring actIII) were separately cloned into the analogous sites of pBM47 to create pBM48, pBM49, and pBM50, respectively.

pBM56 and pBM57 are derivatives of pBM55 in which *PacI–HindIII* fragments of pMP6 and pBM43 harboring encD and encDK, respectively, were cloned into the analogous sites. pBM55 is a pSEK4-derived plasmid in which the act genes contained within the *PacI/EcoRI* site were replaced by a *PacI/EcoRI* cassette harboring actI-ORFsI,II,III-encMNK from pBM54. The actI-ORFsI,II,III-encMNK cassette was constructed from a *NheI–EcoRI* fragment harboring encK from pBMS1,
which was ligated into the pCR-Blunt-derived pBM45 (harboring *encMN*) to create pBM53. An *HindIII-XbaI* fragment from pBM52 harboring actI-ORFs I–III was then cloned into the analogous sites of pBM53 to create pBM54. pBM51 is a pCR-Blunt-derived plasmid harboring *encK* that was PCR amplified from pJP15F11 by using primers 5'–GTCGCTAGCTGATTCTGACGCTGCTAGTG-3' (forward, *NheI* underlined) and 5'–GGTGCCCCGGAGCTGATGC-3' (reverse).

Culture Conditions and Purification of Polyketides. Plasms were introduced into *S. lividans* K4-114, and transformants were selected and grown on solid R2YE containing 20 μg/ml thistreon as described (14). Spores from single colonies were used as inoculum for liquid fermentations in R2YE containing 20 μg/ml thistreon. Cultures were grown at 30°C with shaking (250 rpm) for 4–5 days. Where appropriate, benzoic acid (0.8 mM) was added to the culture after 24 h. Liquid cultures were acidified to pH 5 with 1 M HCl and then extracted with EtOAc. Extracts were dried over anhydrous MgSO4 and concentrated in vacuo.

Polyketide production was monitored by analytical HPLC as described (14).

The *S. lividans* K4-114/pBM43 crude extract (294 mg) from a 1-liter fermentation was subjected to stepwise normal-phase silica gel (Merck, 230–400 mesh) flash column chromatography with 1:1 hexane/EtOAc, EtOAc, 9:1 EtOAc/MeOH, 1:1 EtOAc/MeOH, and MeOH. Each fraction was analyzed by reversed-phase C18 analytical HPLC as described (14). The fraction eluting with 1:1 EtOAc/MeOH was subjected to HPLC purification. Compounds were purified by using a YMC-Pack ODS-A HPLC column (YMC, Kyoto) (250 × 20 mm, 10 μm), employing a gradient from 20% to 70% MeOH in 0.15% trifluoroacetic acid (TFA) over 75 min at a flow rate of 9.5 ml/min with UV detection at 254 nm. Desmethyl-5-deoxyenterocin (10, 9.3 mg) eluted between 12.5 and 14 min.

**Table 1. Plasmid constructions and resulting polyketide products in *S. lividans* K4-114**

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Genes</th>
<th>Major products</th>
</tr>
</thead>
<tbody>
<tr>
<td>pMP6</td>
<td>encA EncB EncL</td>
<td>5–8, 10*</td>
</tr>
<tr>
<td>pBM43</td>
<td>encB EncD EncM</td>
<td>5–10*</td>
</tr>
<tr>
<td>pBM47</td>
<td>actI-ORFs I–III-encMN</td>
<td>18, 19*</td>
</tr>
<tr>
<td>pBM48</td>
<td>actI-ORFs I–III-encMN</td>
<td>11–13*</td>
</tr>
<tr>
<td>pBM49</td>
<td>actI-ORFs I–III-encMN</td>
<td>11–13*</td>
</tr>
<tr>
<td>pBM50</td>
<td>actI-ORFs I–III-encMN</td>
<td>11–13*</td>
</tr>
<tr>
<td>pBM55</td>
<td>actI-ORFs I–III-encMN</td>
<td>18, 19*</td>
</tr>
<tr>
<td>pBM56</td>
<td>actI-ORFs I–III-encMN</td>
<td>11–13*</td>
</tr>
<tr>
<td>pBM57</td>
<td>actI-ORFs I–III-encMN</td>
<td>18–21</td>
</tr>
</tbody>
</table>

*In the presence of 0.8 mM benzoic acid.
†20 and 21 produced in trace amounts.
‡18 and 19 produced in trace amounts.

Bioconversion of Desmethyl-5-Deoxyenterocin to 5-Deoxyenterocin. Desmethyl-5-deoxyenterocin (1 mg) dissolved in DMSO (0.1 ml) was added to a 1-day-old 30-ml culture of *S. lividans* K4-114/pBM43. After another 24 h, the culture was extracted with EtOAc and analyzed by analytical HPLC as described above, showing an ≈30–40% conversion to 5-deoxyenterocin.

**Results**

Identification of *encM* as the Enterocin Favorksiiae-Encoding Gene. On the basis of our earlier report on the identification of the genes *encM* (FAD-dependent oxygenase) and *encK* (methyltransferase) as playing a joint role in the enterocin favorksiiae rearrangement in *S. maritimus* (7), we set out to heterologously express these genes with those encoding the enterocin minimal PKS to further elucidate this biosynthetic event. We earlier prepared a series of pRMS-based expression plasmids for expression in the engineered host strain *S. lividans* K4-114 and demonstrated that the minimal enterocin PKS requires its endogenous KR EncD for activity (14). With this in mind, we constructed the plasmid pBM43 by introducing the *encDK* gene cassette into pMP6 (14) downstream of the actI promoter (Table 1). This plasmid, which additionally carries the *encBC-CLMN* gene cassette downstream of the divergent *actI* promoter, contains the machinery for the synthesis of the starter unit benzyl-CoA from exogenous benzoic acid by the benzoate:CoA ligase *EncN* (16, 17), its loading onto the enterocin ACP *EncC*, the heterodimeric KS_{AB} *EncA*-B, and the acyltransferase *EncL*. This expression plasmid was introduced by means of transformation into the *S. lividans* K4-114 (11) and grown in a benzoate-enriched medium. As projected, the *S. lividans* K4-114/pBM43 transformant produced 5-deoxyenterocin (9) as the major product together with the nonrearranged polyketides waiupemycins D to G (5–8) (Fig. 24), thereby establishing that no other gene

Xiang et al.
product is required for the rearrangement and subsequent cyclization reactions. In addition to these compounds, a new compound (10) was produced when the transformant was fermented in liquid R2YE media rather than on solid agar plates in ~100-fold excess.

Reversed-phase HPLC purification of compound 10 from a 1-liter fermentation provided 9.5 mg for structure characterization. Comparison of the $^1$H and $^{13}$C NMR data with those for 5-deoxyenterocin (18) suggested that 10 was the desmethyl derivative. The absence of the methyl singlet at $\delta$ 3.80 in the $^1$H NMR spectrum supported the loss of the pyrone methyl group. Furthermore, a new, exchangeable proton signal at $\delta$ 11.7 showing HMBC correlations into the pyrone confirmed that the pyrone was not methylated, and thus the structure could be assigned as desmethyl-5-deoxyenterocin (10). This assignment was further established by HRFAABMS with a molecular formula of $C_{27}H_{19}O_{9}$ ($m/z = 415.1029$ [M+H]$^+$, $\Delta = -0.4$ millimass units).

HPLC analysis of the wild-type S. maritimus strain showed that 10 is a very minor natural constituent that was never characterized (Fig. 2B) (8).

**Assignment of EncK as an O-Methyltransferase.** In contrast to our earlier in vivo mutagenesis experiments (7), the heterologous expression experiment implies that the oxygenase EncM is solely responsible for catalyzing the Favorsskii rearrangement and that EncK may act only as a methyltransferase with rearranged polykete substrates (Fig. 1). To test this new hypothesis, we recultured the plasmid pSEK4 (3) by replacing the encMN and encKMN to give pBM43 and pBM55, respectively. In both cases, S. lividans K4-114 transformants produced only the known act compounds SEK4 and SEK4b (Fig. 3), suggesting that the enterocin favorsactin PKS, from producing aromatic polyketides to rather synthesize acetate-primed enterocin-like compounds. A series of hybrid act-enc gene combinations were thus constructed in pRM5 and pSEK4-based act PKS expression vectors (3) and expressed heterologously in S. lividans K4-114 (Table 1).

The nonreduced shunt octaketides SEK4 (18) and SEK4b (19) are products of the minimal act PKS ActI-ORFSl-III when expressed in the absence of the KR ActIII (1, 2). We reengineered the plasmid pSEK4 (3) by replacing the actVII (cyclase) and actIV (aromatase) genes with encMN and encKMN to give pBM47 and pBM55, respectively. In both cases, S. lividans K4-114 transformants produced only the known act compounds SEK4 and SEK4b (Fig. 3), suggesting that the enterocin favorssactin PKS and the O-methyltransferase EncK were not able to use the act minimal PKS products as substrates. To evaluate whether a C9 ketoreduced act substrate is preferred, much like in enc biosynthesis, we next expressed the act minimal PKS with either its endogenous KR ActIII or with the homologous enterocin KR EncD, which has been successfully used before in conjunction with the act PKS (5). Several plasmin containing the act minimal PKS and a KR with different combinations of the genes encKMN (Table 1) all yielded roughly the same polyketide profile when expressed in S. lividans K4-114 (data not shown). In each case, in addition to the KR-reduced act polyketides mutacin, dehydrolutacin, and EM18, the nonreduced SEK4 and SEK4b compounds were measured at various levels along with two new compounds. These molecules were isolated by reversed-phase HPLC and characterized by MS and NMR. Proton and carbon NMR data, along with gradient-enhanced heteronuclear multiple quantum correlation (HMQC) and HMBC spectroscopy, quickly revealed that 20 and 21 were dehydrated analogs of SEK4 and SEK4b, respectively, which was supported in each case by HRFAABMS. Dehydro SEK4 (aka B26, 20) was previously characterized from the heterologous expression of the naphthocyclinone PKS, although its structure was never assigned by NMR (20). Key HMBC correlations supporting the structure 20 were observed from the H14 olefin at $\delta$ 6.01 to the aromatic carbon at $\delta$ 113.7 (C12) and the methyl carbon at $\delta$ 19.4 (C16). Similarly, the placement of the olefin in dehydro SEK4b (21) was established through HMBC correlations from the H-8 olefin at $\delta$ 6.09 into the benzene ring and the C-6 methylene bridge to the

---

**Fig. 2.** HPLC analysis at 254 nm of crude extracts from S. lividans K4-114/pBM43 plus benzoic acid (trace A), wild-type S. maritimus (trace B), S. lividans K4-114/pMP6 plus benzoic acid (trace C), and the biotransformation of 10 to 9 in S. lividans K4-114/pBM43 (trace D). Cinnamic acid is an intermediate in the biosynthesis of the benzoyl-CoA starter unit from $\gamma$-phenylalanine (17) as seen in trace A.

The role of EncK as an $O$-methyltransferase was further confirmed through a biotransformation experiment in which 10 was added to a culture of S. lividans K4-114/pBM43 (encABCLMN/encDK). In the absence of supplemental benzoate, this transformant does not produce enc-based polyketides. Thus, when pure 10 was added to the fermentation, we observed by HPLC the production of 5-deoxyenterocin (Fig. 2D), thereby confirming that methylation is the penultimate step in enterocin biosynthesis. EncK is most homologous to AurI, another Streptomyces polyketyl pyrone $O$-methyltransferase (19), with a similarity/identity score of 54/36 and has been shown to complement an aural genetic mutation in aureothin biosynthesis (J. He, C. Hertweck, L.X., and B.S.M., unpublished observation).
pyrone. Interestingly, the levels of the nonreduced polyketides SEK4 and SEK4b and their dehydrated counterparts increased relative to the KR-reduced compounds as the number of enc genes was enlarged. The enc-tailoring proteins EncK and EncM may compete with the KR ActIII or EncD in complexing with the minimal acr PKS, thereby perturbing the ketoreduction process. In no case, unfortunately, did we detect oxidatively rearranged or O-methylated acr products.

Discussion

The encM gene encodes a putative 464-aa flavoprotein with remarkable catalytic properties. Sequence analysis suggests that the “favorskiiase” EncM has a covalent attachment of FAD at a conserved histidine residue at position 78, which is reminiscent of many oxidases (21, 22), including those involved in the biosynthesis of the natural products mitomycin (23) and griseorhodin A (24). Flavoenzymes are very versatile catalysts and are involved not only in the activation of oxygen for oxidation and hydroxylation reactions, but also in dehydrogenation reactions, in light emission, and in one- and two-electron transfer reactions (25). Flavoprotein-catalyzed oxidative reactions associated with post-PKS modifications include hydroxylations, epoxidations, Baeyer–Villiger rearrangements (26), and now Favorskii rearrangements.

The enterocin flavoprotein EncM is implicated in this study not only to catalyze the Favorskii-like oxidative rearrangement, but also must facilitate the two aldol condensation and two heterocycle-forming reactions in the formation of desmethyl-deoxyenterocin (10). In the course of this series of chemical transformations, at least five new chiral centers and four new rings are generated. The presumed EncM precursor is the linear enc PKS C9-reduced octaketide 14 that is oxidized at C12 to form the 11,12,13-trione intermediate 16 (Fig. 1, path A). If EncM functions rather as a dioxygenase, then oxidation of the cyclohexenone intermediate 15 similarly yields 16 (Fig. 1, path A′). Hydration of the C13 carbonyl and scission of the C13–C14 bond with concomitant formation of a new C12–C14 linkage generates the hypothetical Favorskii product 17 bearing a chiral α-hydroxy acid center. This branched intermediate formally undergoes two aldol condensations between C6–C11 and C7–C14 and in the process generates chiral centers at each position. Although cyclases catalyze such aldol condensation reactions during aromatic polyketide assembly (1, 2), the enc gene cluster is a notable exception among all other known type II PKS gene sets for its absence of typical cyclase and aromatase encoding genes (5). We hypothesize by virtue of the formation of 10 that EncM likely facilitates these C–C-forming reactions by orienting in a conducive conformation, because linear poly-β-ketides are extremely reactive and spontaneously cyclize (27, 28). The lactone C–O bond between the carboxylate oxygen and C9 is also generated during the course of 10 biosynthesis. On the basis of the presumed 9R-configuration that is inferred from the structures of the wailupemycin A and B shunt products (8, 29), the C–O linkage should be generated with inversion of stereochemistry with displacement of the acetate-derived C9 hydroxyl group. Pyrone formation with hydroxylation of the thioester linkage to the EncC ACP completes the biosynthetic transformation to 10. Because the activity of EncM has not yet been reconstituted in vitro, there remains the possibility that other host proteins may contribute to the transformation in vivo.

We previously proposed on the basis of a gene knockout experiment that the encK gene product also participates in the rearrangement reaction (7). Further experimentation in this study, however, conclusively showed this not to be the case with the characterization of 10. Primary sequence analysis of EncK demonstrates that it belongs to a distinct family of S-adenosyl-L-methionine-dependent O-methyltransferases involved in the methylation of macrolide antibiotics and plant phenylpropanoids.

Although the encM gene product has now been shown to solely catalyze a remarkable series of biosynthetic events during enterocin biosynthesis, its use as a combinatorial biosynthetic reagent to generate “unnatural” natural products may unfortunately be limited. This statement is, however, based upon a single observation with the actinorhodin PKS in which the addition of encM did not redirect the post-PKS tailoring reactions of the act pathway, suggesting that wild-type EncM may be specific for its endogenous type II PKS or for benzoyl-primed polyketide precursors. An in-depth biochemical characterization of EncM will be necessary to probe its mechanism and to evolve it into a suitable reagent for the engineered biosynthesis of new chemical entities.

S. lividans K4-114 was generously provided by Kosan Biosciences, Inc. (Hayward, CA), and pRM5 and pSEK4 were kindly provided by C. Khosla (Stanford University, Stanford, CA). This work was supported by National Institutes of Health (NIH) Grant AI47818. Partial support for the 600-MHz NMR spectrometer in the College of Pharmacy was provided by NIH National Center for Research Resources Shared Instrumentation Grant 1S10RR16659.