Characterization of the Nocardiopsis Biosynthetic Gene Cluster Reveals Similarities and Differences to the Rapamycin and FK-506 Pathways

Dana M. Bis,[a] Yang H. Ban,[a,][b] Elle D. James,[a] Norah Alqahtani,[b] Rajesh Viswanathan,[b] and Amy L. Lane*[a]

Abstract: Macrolide-pipicolate natural products, such as rapamycin (1) and FK-506 (2), are renowned modulators of FKBP s. The nocardiopsins, from Nocardiopsis sp. CMB-M0232, are the newest members of this structural class. Herein, the biosynthetic pathway for nocardiopsins A-D (4-7) is revealed by cloning, sequencing, and bioinformatic analyses of the nsn gene cluster. In vitro evaluation of recombinant NsnF revealed this lysine cyclodeaminase catalyzes conversion of L-lysine to L-pipocolic acid incorporated into 4-5. Bioinformatic analyses supported that a linear nocardiopsin precursor is staged with the hydroxy group required for macroclide closure by a new strategy employing a P450 epoxidase (NsnF) and limonene epoxide hydrolase homolog (NsnG). The nsn cluster also encodes candidates for tetrahydrofuran group biosynthesis. The nocardiopsin pathway provides opportunities for engineering of FKBP-binding metabolites and probing novel enzymology in Nature’s polyketide tailoring arsenal.

Introduction

Rapamycin (1),[1] FK-506 (2),[2] and meridamycin (3)[3] are representatives of a small family of natural products featuring a macrolide core and pipicolate group (Scheme 1). Members of this family are noteworthy for their ability to bind immunophilin FKBP s, which interact with diverse proteins (e.g. mTOR) that play roles in many intracellular signaling pathways.[4] Consequently, the formation of binding complexes between these natural products and FKBP s affords myriad downstream biological effects. Members of this family are used clinically as immunosuppressants to prevent transplanted organ rejection and for treatment of renal cell carcinoma. This group of compounds also holds promise as future therapies for neurodegenerative diseases and other pathologies.[5]

Despite prominent biomedical enthusiasm about the rapamycin/FK-506 natural product family, genome mining and traditional isolation-based approaches have afforded only a handful of unique skeletons.[6] Nocardiopsins A-D (4-7), isolated from marine-derived Nocardiopsis sp. CMB-M0232, are the only new macrolide-pipicolate metabolites reported during the past decade (Scheme 1).[7] The nocardiopsins feature several structural differences that distinguish them from previously reported natural products. Notably, the nocardiopsins lack the C9-keto and C10-C14 hemiketal groups (rapamycin numbering) that play a critical role in the FKBP-binding capability of 1-3 and other family members.[8] Even without these groups, 4-5 were found to bind FKBP12 with Kd values in the low micromolar range.[9] Further structurally distinguishing them from other rapamycin and FK-506 family members, all nocardiopsins feature a cis-configured tetrahydrofuran (THF) group spanning C9-C12. The structures of 4-7 suggest that closure of the 21-membered macroclide occurs via a tertiary C6 hydroxy group of a macrolyketide synthase origin.[10] This is in contrast to the vast majority of macroclide ring closures, which occur through the intramolecular reaction of a PKS-derived hydroxy group with a PKS- or nonribosomal peptide synthetase (NRPS)-tethered thioester.[11] These structural features both distinguish the nocardiopsins from previously reported natural products and suggest that the nocardiopsin biosynthetic pathway includes key differences from characterized macrolide-pipicolate pathways.

Herein, we report the cloning, sequencing, and bioinformatic analyses of the nocardiopsin biosynthetic gene cluster (nsn). We provide evidence for the function of this gene cluster in nocardiopsin biosynthesis through the in vitro evaluation of recombinant lysine cyclodeaminase homolog NsnL. This enzyme is demonstrated to catalyze the formation of L-pipicolic acid, an uncommon amino acid featured in 4-5 and essential for FKBP12 binding.[12] Further, our results support that Nocardiopsis sp. stages a linear intermediate for macroclide ring closure via a unique route that includes the formation and hydrolysis of an epoxide. Our data also suggest an intriguing route to nocardiopsin THF functionalization, and offer testable hypotheses to fully unveil this biosynthetic strategy. Leveraging understanding of the molecular basis for nocardiopsin biosynthesis with knowledge of pathways for other macroclide-pipicolates is expected to offer opportunities for engineering of hybrid natural products of this medicinally promising group as well as increased understanding of Nature’s toolkit for polyketide tailoring.

Results and Discussion

Cloning and sequencing of the nocardiopsin gene cluster (nsn)

Open reading frames (ORFs) predicted from the ~6.4-Mbp draft genome sequence of Nocardiopsis sp. CMB-M0232 were evaluated for homology to previously characterized PKSs and NRPSs,[8] revealing a partial gene cluster hypothesized to be responsible for nocardiopsin biosynthesis. Cosmid clones carrying Nocardiopsis sp. genomic DNA were screened for these genes, and three overlapping cosmids were sequenced and assembled to yield a single 119-kb contig, from which 77-kb was annotated as the nocardiopsin gene cluster (nsn, Figure 1). The nsn cluster encodes enzymes homologous to characterized biosynthetic enzymes (Table 1), including PKSs (Nsna-NsnD), an NRPS (NsnE), cytochrome P450s (NsnF; NsnH), an oxidoreductase (NsnI), an epoxide hydrolase (NsnG), and a...
A cytochrome P450 and epoxide hydrolase are proposed to act upon a linear nocardioipsin precursor to stage it for macrolide closure. The structures of nocardioipsins A-D (4-7) imply that macrolide ring closure occurs through the reaction of a C6 hydroxy group with the NRPS-tethered pipercolyl or prolinyl group (Figure 2). For all previously characterized macrolide-pipeolate family pathways as well as the majority of macrolides, the requisite hydroxy group is formed by reduction of a ρ-ketoioester intermediate by a PKS KR domain. However, KR-catalyzed hydroxylation at C6 of the linear nocardioipsin precursor is not expected. Instead, module 2 of the PKS assembly line is expected to afford a C5-C6 alkenic group (Figure 2), suggesting a unique route to stage the molecule for macrolide closure.

The nsn cluster encodes two cytochrome P450 homologs (NsnF and NsnH) as well as an epoxide hydrolase homolog (NsnG) as the most plausible candidates for staging the linear nocardioipsin precursor with the hydroxy group needed for ring closure. Comparison of NsnF with characterized polyketide-pipeolate family pathways with the NRPS (Supporting Information Figure S4) revealed that NsnF is most similar to several epoxidases (e.g. 55% similarity to GfsF from the FD-891 pathway). The homology of NsnF to biochemically characterized epoxidases, along with observation of an epoxide hydrolase homolog (NsnG) within the cluster, implicate these two enzymes as the most plausible candidates to stage the linear precursor for ring closure (Figure 2). NsnF is predicted to act upon the linear nocardioipsin precursor, catalyzing epoxidation of the C5-C6 alkene. NsnG is then predicted to catalyze nucleophile ring opening to transform the C5-C6 epoxide into a vicinal diol, for which the tertiary alcohol group is recruited for macrolide closure via the NRPS NsnE (Figure 2).

NsnG is most closely related to members of the limonene epoxide hydrolase family (e.g. 57% similarity to Rv2740). This family of epoxide hydrolases is highly distinct from classic epoxide hydrolases involved in secondary metabolism. Classic epoxide hydrolases feature a highly conserved αβ hydrolase fold and a two-step mechanism that includes a covalent enzyme-substrate intermediate. Instead, X-ray crystallographic and computational studies of two limonene epoxide hydrolases,

lysine cyclodeaminase (NsnL). The cluster also encodes proteins with predicted roles in pathway regulation (NsnR1-NsnR4) and antibiotic transport (NsnT1-NsnT2).

Bioinformatic analyses support assembly of the nocardioipsin core by polyketide synthases NsnA-NsnD and nonribosomal peptide synthetase NsnE. The nsn cluster encodes four ORFs (NsnA-NsnD) sharing homology with characterized modular type I PKSs (Figure 1, Table 1).[9] Bioinformatic analyses revealed that the four PKSs encompass a predicted loading module as well as ten extender modules, corresponding to the number of extension cycles expected to afford the nocardioipsin polyketide core (i.e. C1-C23, Figure 2). Domains within each extender module and the substrate preference of each acyltransferase (AT) domain were proposed by bioinformatic analyses (Figure 2).[9] Each predicted ketoreductase (KR), dehydratase (DH), and enoyl reductase (EH) domain contains amino acid sequence signatures that correspond to common catalytically active PKS domains (Supporting Information Figures S1-S3). Together, these analyses support biosynthesis of the nocardioipsin polyketide core via a prototypical type I modular PKS, with assembly proceeding co-linearly to yield the parent polyketide chain of 4-7. Following PKS module 10, pipercollic acid (as in 4-5) or proline (as in 6-7) is incorporated into the polyketide chain through reactions catalyzed by the NRPS homolog NsnE. Sequence homology of the adenylation (A) domain of NsnE with those of NRPSs known to act upon pipercolic acid substrates supported the role of NsnE in pipercolate incorporation (Supporting Information Table S1). Like previously reported pipercolate-incorporating NRPSs,[10] NsnE exhibits relaxed substrate specificity, as evidenced by structural differences between the amino acid moieties of 4-5 vs. 6-7. NsnE includes two predicted condensation (C) domains (Figure 2). Analogous to the biosynthetic pathways for 1-3,[8, 10a, 11] one of these C domains is proposed to catalyze transfer of the polyketide chain to the pipercolate group while the other is expected to catalyze closure of the macrolide, with concomitant release of the macrolide from the NRPS.

A cytochrome P450 and epoxide hydrolase are proposed to

Scheme 1. Chemical structures of selected members of the rapamycin and FK-506 natural product family, including nocardioipsins A-D (4-7).
Unfortunately, Zotchev and co-workers employed mutagenesis approaches analogous to those employed by Gust and co-workers,\textsuperscript{[13]} as well as insertional mutagenesis approaches analogous to those employed by Zotchek and co-workers within the Nocardiosis genus.\textsuperscript{[14]} Unfortunately, Nocardiosis sp. CMB-M0232 proved entirely resistant to genetic manipulation via these methods.

While epoxide formation and hydrolysis are relatively common polyketide tailoring steps, to our knowledge, the current study offers the first example of a biosynthetic pathway that appears to apply these combined steps to poise a linear substrate for macrolide formation via stage an acyclic polyketide for macrolide formation. It is plausible that these enzymes act upon a megasynthase-tethered precursor, which would further add to their uniqueness (Figure 2). Relatively few tailoring enzymes acting upon megasynthase-tethered substrates have been reported.\textsuperscript{[14]} As one analogous example, the stambomycin pathway was recently reported as the first to stage an acyclic polyketide for macrolide formation via cytochrome P450-catalyzed hydroxylation of an unactivated carbon.\textsuperscript{[15a,16]}

\section*{Table 1. Proposed functions of proteins from the Nocardiosis sp. CMB-M0232 nocardiosin gene cluster (nsn) based on comparison of amino acid sequences with homologous enzymes. \textit{(nsn cluster deposited in GenBank as accession #KP339942.)}}

<table>
<thead>
<tr>
<th>gene</th>
<th>translated protein size (Raa)</th>
<th>annotation</th>
<th>protein homolog (NCBI accession number)</th>
<th>identity/similarity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>orf(3)</td>
<td>408</td>
<td>adenylylsulfate kinase</td>
<td>MB27_02760 (KHD79007)</td>
<td>47/58</td>
</tr>
<tr>
<td>orf(2)</td>
<td>644</td>
<td>phosphogluconate dehydrogenase</td>
<td>RHRU231_200054 (CD287071)</td>
<td>77/83</td>
</tr>
<tr>
<td>orf(1)</td>
<td>145</td>
<td>(2Fe-2S)-binding protein</td>
<td>SSIG_01081 (EWS90718)</td>
<td>57/67</td>
</tr>
<tr>
<td>nsnR1</td>
<td>238</td>
<td>AraC family transcriptional regulator</td>
<td>SGLAU_30775 (AISO2089)</td>
<td>47/59</td>
</tr>
<tr>
<td>nsnR2</td>
<td>198</td>
<td>TetR family transcriptional regulator</td>
<td>SIRAN742 (CDR0393)</td>
<td>91/94</td>
</tr>
<tr>
<td>nsnI</td>
<td>249</td>
<td>oxidoreductase</td>
<td>NBRGN_038_01510 (GAJ81480)</td>
<td>91/94</td>
</tr>
<tr>
<td>nsnH</td>
<td>414</td>
<td>cytochrome P450</td>
<td>AveE (BAA84477)</td>
<td>31/45</td>
</tr>
<tr>
<td>nsnT1</td>
<td>213</td>
<td>ABC transporter protein</td>
<td>IL38_13665 (KGI81075)</td>
<td>56/71</td>
</tr>
<tr>
<td>nsnT2</td>
<td>150</td>
<td>transporter protein</td>
<td>RradSPS_0169 (AHY45452)</td>
<td>80/88</td>
</tr>
<tr>
<td>orf7</td>
<td>873</td>
<td>ATPase</td>
<td>IL38_11110 (KGI81478)</td>
<td>61/77</td>
</tr>
<tr>
<td>nsnF</td>
<td>359</td>
<td>cytochrome P450 epoxidase</td>
<td>GlsF (BA16472)</td>
<td>44/55</td>
</tr>
<tr>
<td>nsnL</td>
<td>345</td>
<td>lysine cycloamine</td>
<td>PipA (CBW45757)</td>
<td>55/70</td>
</tr>
<tr>
<td>nsnE</td>
<td>1544</td>
<td>nonribosomal peptide synthetase</td>
<td>MerP (ABC7508)</td>
<td>49/62</td>
</tr>
<tr>
<td>nsnA</td>
<td>6409</td>
<td>polyketide synthase</td>
<td>PldAI (BAH2268)</td>
<td>50/61</td>
</tr>
<tr>
<td>nsnB</td>
<td>5166</td>
<td>polyketide synthase</td>
<td>MerC (AB97439)</td>
<td>51/61</td>
</tr>
<tr>
<td>nsnC</td>
<td>4937</td>
<td>polyketide synthase</td>
<td>ApoS (AEP0934)</td>
<td>49/60</td>
</tr>
<tr>
<td>nsnD</td>
<td>1801</td>
<td>polyketide synthase</td>
<td>LiiPKS3 (ABB05104)</td>
<td>51/61</td>
</tr>
<tr>
<td>nsnG</td>
<td>145</td>
<td>epoxide hydrolase</td>
<td>Rv2740 (2BNG_A)</td>
<td>49/57</td>
</tr>
<tr>
<td>nsnR3</td>
<td>230</td>
<td>LuxR transcriptional regulator</td>
<td>AURJ3M (ACD75765)</td>
<td>62/72</td>
</tr>
<tr>
<td>orf17</td>
<td>307</td>
<td>(2Fe-2S)-binding protein</td>
<td>DF19_12420 (KD77345)</td>
<td>52/66</td>
</tr>
<tr>
<td>orf18</td>
<td>431</td>
<td>glycosyltransferase</td>
<td>MPTA5024_33555 (ETK31740)</td>
<td>54/64</td>
</tr>
<tr>
<td>nsnR4</td>
<td>133</td>
<td>XRE family transcriptional regulator</td>
<td>M877_14005 (EST28853)</td>
<td>57/64</td>
</tr>
</tbody>
</table>

Rv2740 (PDB accession #2BNG) and LEH (PDB accession #1NWW), revealed this class of enzymes features a distinct structure and unique conserved active site triad that catalyzes epoxide ring opening through a one-step acyl-catalyzed mechanism.\textsuperscript{[15]} Amino acid sequence alignment of NsnG with Rv2740 and LEH revealed that NsnG possesses this active site triad (Figure 3).

To probe the \textit{in vivo} functions of \textit{nsnF} and \textit{nsnG}, we attempted to eliminate these genes from \textit{Nocardiosis} sp. CMB-M0232 using the PCR-targeted gene replacement strategy pioneered by Gust and co-workers,\textsuperscript{[13]} as well as insertional mutagenesis approaches analogous to those employed by Zotchek and co-workers within the \textit{Nocardiosis} genus.\textsuperscript{[14]}
The nocardiopsins (4-7) feature a THF group that is absent in all other macrolide-pippecolate family natural products described to date. Examination of ORFs encoded by the nsn cluster revealed NsnH, a cytochrome P450 homolog, as one potential candidate for catalyzing THF formation. NsnH shares 43% homology with AveE. AurH is a bifunctional cytochrome P450 from the aureothin pathway, and was demonstrated to catalyze the hydroxylation of an unactivated carbon atom as well as subsequent THF formation via a ring closure reaction involving this hydroxy group and an alkene. Likewise, AveE was previously postulated to catalyze the formation of the avermectin THF group by a similar mechanism. However, analysis of nocardiopsin PKS domains (i.e. within modules 4-5, Figure 2) suggest that the THF precursor features hydroxy groups at both C9 and C11. Hence, if NsnH is indeed operative in THF functionalization, it is expected to employ a mechanism fundamentally distinct from that of AveE and AurH.

Another possibility is that THF functionalization of nocardiopsin requires NsnH in addition to other enzyme(s) or, alternatively, that THF formation does not require NsnH. Of note, NsnI homologs have been biochemically characterized, leaving the role of NsnI in THF biosynthesis ambiguous.

Previous studies have demonstrated the biosynthesis of cyclic ethers through a variety of intriguing mechanisms in addition to those catalyzed by P450s. These include biosynthesis of the nonactin THF group by an enoyl CoA hydratase homolog and the biosynthesis of polyether ladder

Figure 2. Overview of the nocardiopsin biosynthetic pathway. (top) PKS domain organization and substrate specificity for each module, as proposed by bioinformatic analyses. Chain numbering is based upon numbering of final product. (bottom) To afford the C6 hydroxy group required for macrolide closure, the C5-C6 alkene formed via PKS module 2 is predicted to be converted to an epoxide (shown in red) by NsnF. This epoxide is presumably opened via epoxide hydratase NsnG to yield a C5-C6 vicinal diol suitable as a substrate for NRPS-catalyzed ring closure. Epoxide opening and formation is illustrated as occurring while the linear precursor is tethered to the NRPS; however, the actual timing of these steps remains cryptic.

Figure 3. Analysis of NsnG, a limonene epoxide hydratase homolog encoded by the nocardiopsin gene cluster. (a) ClustalW amino acid alignment of NsnG and limonene epoxide hydrolases. (b) Proposed mechanism of the NsnG catalytic triad in opening the nocardiopsin precursor epoxide, based on previous studies of P4740 and LEH. Amino acid numbering is based upon residues from LEH.
natural products such as lasalocid through a cascade of epoxide ring opening cyclization reactions. Further, tetrahydropyran (THP) group formation was recently demonstrated to occur via a pyran synthase (PS) domain within a trans-AT PKS assembly line (e.g. in pederin biosynthesis). Analogously, within cis-AT PKS assembly lines, a OH domain from the ambruticin pathway was recently demonstrated to catalyze THP formation in addition to dehydramination. A unique cyclase domain within the indanomycin PKS has also been suggested to play a role in THP biosynthesis. However, based on bioinformatic analyses of PKSs NsnA-NsnD and other ORFs encoded within the nsn cluster, none of these previously reported biosynthetic strategies appear plausible for nocardiopsin THF biogenesis.

**Pipelicolic acid biosynthesis is catalyzed by lysine cyclodeaminase NsnL.**

Previous studies demonstrated that 4-5 (featuring a pipelicolate group) exhibited significant binding with FKBP12 while 6-7 (featuring a proline group) were inactive. Hence, the pipelicolate moiety appears essential to the biological activity of nocardiopsins. NsnL, a lysine cyclodeaminase homolog (62% similarity to RapL from rapamycin pathway) was hypothesized to catalyze the formation of L-pipelicolic acid from L-lysine in the nocardiopsin pathway. Bioinformatic analyses revealed no other lysine cyclodeaminase homologs encoded by the Nocardiopsis sp. CMB-M0232 genome. To establish the function of NsnL (394 kDa), this enzyme was recombinantly expressed as a His-fusion protein in E. coli and purified (Supporting Information Figure S6). The typical yield of NsnL was ~15 mg per liter of culture.

NsnL was incubated with selected individual substrates (i.e. L-lysine, D-lysine, and L-ornithine) to probe its function and substrate specificity in vitro. Reaction products were derivatized and evaluated by gas chromatography with flame ionization detection (GC-FID). These analyses revealed that NsnL catalyzed the conversion of L-lysine to L-pipelicolic acid (Figure 4a). D-lysine was not accepted as a substrate (data not shown). This established NsnL as a source of pipelicolic acid for incorporation into nocardiopsins by NRPS NsnE (Figure 2), and support that the enzyme is stereospecific. NsnL also catalyzed the conversion of L-ornithine to L-proline, suggesting promiscuity with respect to the substrate carbon chain length (Figure 4b).

NsnL is expected to catalyze these cyclization reactions through a mechanism analogous to that established for homologous ornithine cycloaminases yielding proline in primary metabolism and that proposed for RapL, which catalyzes the formation of L-pipelicolic acid during rapamycin biosynthesis. In this mechanism, the α-amino group of lysine or ornithine is first oxidized to an imine, with nicotinamide adenine dinucleotide (NAD⁺) as a co-factor. The terminal amine then attacks the imine carbon to afford a cyclic tetrahedral intermediate. The loss of ammonia from this intermediate then affords a cyclic imino acid, which is reduced by NADH to yield pipelicolic acid and re-form NAD⁺. Supporting this mechanism, NsnL catalyzed the formation of pipelicolic acid or proline only when NAD⁺ was added to in vitro reactions (data not shown).

**Conclusions**

The discovery of nocardiopsins A-D (4-7) by Capon and co-workers marked the first novel members of the macroclide-pipelicolate class of FKBP modulators in over a decade and significantly enhanced the structural diversity of this class of medicinally valuable natural products. Establishment of the nocardiopsin biosynthetic gene cluster (nsn), presented in the current work, eludes to several unique pathway features. The further study of this is anticipated to afford increased understanding of Nature’s array of polyketide tailoring mechanisms as well as potential biocatalysts for the synthesis of biologically active polyketedes. Of particular interest are the two enzymes, a P450 epoxidase (NsnF) and epoxide hydrolase (NsnG) homolog, expected to act sequentially on linear polykete intermediates to equip these molecules with hydroxy groups necessary for macrolide closure. Given that macrokides are important determinants of biological activities for a variety of polyketedes, the characterization of NsnF and NsnG may offer strategies for the engineering of designer polyketyde macrokides. Further, additional enzymes encoded by the nsn pathway (e.g. NsnH and/or NsnI) represent candidates for biosynthesis of the THF moiety of 4-7 and suggest a unique route to formation of this group.

Accepted for publication on February 6, 2015 by ChemBioChem.
Experimental Section

*Nocardiopsis* sp. CMB-M0232 genomic DNA isolation and draft genome sequencing. *Nocardiopsis* sp. CMB-M0232 was maintained following previously described procedures.[5] Genomic DNA (gDNA) was isolated by standard procedures.[6] Sequencing was conducted on this gDNA by Cofactor Genomics (St. Louis, MO) using illumina with paired-end 80-bp reads and 454 with single-end 400-bp reads. ORFs were predicted using GeneMark.[7] To locate candidate enzymes for nocardipsin biosynthesis, ORFs were searched for homology to representative modular PKSs and NRPSs using BLASTp in Geneious (BioMatters).

Cloning and sequencing of the nocardipsin gene cluster (*nsn*). Individual contigs from the draft genome sequence included only partial NRPSs and PKS gene clusters. To complete the *nsn* pathway, a cosmide library of gDNA was constructed and PCR screened for these gene clusters as described in Supporting Information. The final *nsn* DNA sequence is deposited in GenBank as accession #KP339942.

Bioinformatic analyses of the nocardipsin gene cluster. ORFs within the *nsn* gene cluster were proposed using FrackeSequencer[8] and FGENESB (Softberry).[9] Each ORF was evaluated for homology to previously reported proteins using BLASTp. Individual modules and domains within PKSs and the NRPSs were predicted using the tool developed by Bachmann and Ravel.[10] AT and C domain substrate preference was assigned using methods developed by Khayatt et al.[11] and the SEARCHPKS program by Yadav et al.[12] Amino acid alignments were validated using ClustalW.[13]

Preparation of the construct for expression of lysine cyclodeaminase gene *nsnL* in *E. coli*. The gene encoding NsnL was PCR amplified using forward and reverse primers (*NsnL_pETF* and *NsnL_pETR*, Supporting Information Table S2) incorporating NspV and Ncol restriction sites for cloning into the corresponding sites of the pET30a(+) expression vector to yield a fusion protein with an N-terminal hexahistidine tag. The 1.1-kb PCR product was ligated into pGEM-T vector (Promega) and the resulting construct propagated in *E. coli* JM109. The construct was digested with NspV and Ncol to release the *nsnL* PCR product with sticky ends for ligation into pET30a(+) (Novagen). The *nsnL* gene was ligated into equivalently digested pET30a(+), the resulting construct propagated in JM109, and then introduced into *E. coli* BL21(DE3) for gene expression. DNA sequencing was conducted to confirm the *nsnL* gene insert sequence.

Overexpression and purification of lysine cyclodeaminase NsnL. A 6 h culture of BL21(DE3) carrying pET30a(+)/*nsnL* was diluted with 1:1000 in LB containing kanamycin (30 µg/mL) and incubated for ~4 h at 200 rpm and 37 °C to afford an optical density (OD600) of ~0.5-0.6. The culture was supplemented with IPTG (1 mM) to induce expression of NsnL. Cultures were incubated for an additional 13 h at 15 °C with shaking at 250 rpm. Cells were pelleted by centrifugation and stored at ~80 °C until protein purification.

To purify NsnL, cells (1 g wet mass) were lysed by mixing with B-PER (4 mL, Thermo Scientific) supplemented with lysozyme (1.5 mg/mL) and DNase (1.5 U/mL, Qiagen) and incubated at 30 °C for 30 min. Resulting lysates were centrifuged at 12,000 rpm for 30 min at 4 °C. The supernatants were passed through a 0.4 µm filter and supplemented to afford a final concentration of 400 mM NaCl, 10% glycerol, and 2 mM imidazole in Tris (25 mM, pH 8) for Ni affinity chromatography. To batch bind NsnL, the solution (4 mL) was incubated with Ni agarose slurry (0.5 mL, Qiagen) by shaking on ice at 200 rpm for 2 h. This mixture was poured into a column, and the above buffer (1.25 mL) was used to wash the column. The column was then eluted sequentially with aliquots of the above buffer (250 µL) supplemented with final concentrations of 5 mM, 20 mM, 40 mM, 60 mM, and 200 mM imidazole, respectively. Finally, the column was eluted with buffer (1.25 mL) containing imidazole (500 mM). In this final fraction, NsnL (39.4 kDa) was enriched to >90% purity, as determined by SDS-PAGE (Supporting Information).

The buffer was exchanged by two-step dialysis at 4 °C, first with Tris buffer (500 mL, 50 mM, pH 8) containing 100 mM NaCl (100 mM), EDTA (1 mM), and glycerol (10% v/v), followed by Tris buffer (500 mL, 50 mM, pH 8) containing NaCl (100 mM), glycerol (10% v/v), and TCEP (1 mM). NsnL was concentrated by ~4-fold using a 9 kD cutoff centrifugal concentrator, the protein concentration was determined by the BCA method, and NsnL was stored at ~80 °C.

Evaluation of the function of NsnL through *in vitro* assays and GC-FID. Treatment and control assays (100 µL) contained potassium phosphate buffer (100 mM, pH 8), NAD+ (400 µM), and DTT (3 mM). In addition to these components, treatment samples contained NsnL (17.5 µM) as well as of L-lysine, D-lysine, or L-ornithine (10 mM). Control samples contained individual amino acid substrates in the absence of NsnL, or NsnL in the absence of added substrate. Assays analogous to those described above were also setup without the addition of NAD+ to evaluate the effect of NAD+ addition on NsnL activity.

Assays were incubated overnight (~15 h) at 30 °C with shaking at ~200 rpm. Resulting reaction products were immediately extracted, derivatized, and subjected to GC-FID using the EZ:Faast GC-FID kit (Phenomenex), including column, and following manufacturer recommendations. Norvaline was included in all experiments as an internal standard. Synthetic L-lysopipeolic acid (Sigma) and L-proline (Sigma) were equivalently derivatized and used as standards.

Acknowledgements

We are grateful to Prof. Robert Capon for providing *Nocardiopsis* sp. CMB-M0232 and to Prof. Sergey B. Zotchev for providing pKE5. ALL acknowledges funding from a Research Corporation Cottrell College Science Award, University of North Florida (UNF) Transformational Learning Opportunity Awards, UNF Academic Affairs Grants, and a UNF Faculty Fellowship Award. YB was supported by a UNF SMART grant. RV and NA thank Case Western Reserve University for financial support.

Keywords: natural products • biosynthesis • polyketides • macrolide • pipeolic acid


Expanding Nature’s macrolide-pipeolate biosynthetic repertoire. Cloning, sequencing, and analysis of the nocardiopsin biosynthetic pathway revealed that Nocardiopsis sp. CMB-M0232 employs a novel strategy, featuring both a P450 epoxidase and epoxide hydrolase, to transform the alkene group of a linear polyketide precursor into the hydroxy group required for macrolactonization.

Dana M. Bis, Yang H. Ban, Elle D. James, Norah Alqahtani, Rajesh Viswanathan, Amy L. Lane*