Organic & Biomolecular Chemistry

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/c5ob00537j

Synergism between genome sequencing, tandem mass spectrometry and bio-inspired synthesis reveals insights into nocardioazine B biogenesis†

Norah Alqahtani,‡[a] Suheel K. Porwal,§‡[a] Elle D. James,[b] Dana M. Bis,[b] Jonathan A. Karty,[c] Amy L. Lane*[b] and Rajesh Viswanathan*[a]

Marine actinomycete-derived natural products continue to inspire chemical and biological investigations. Nocardioazines A and B (**3** and **4**), from *Nocardiopsis* sp. CMB-M0232, are structurally unique alkaloids featuring a 2,5-diketopiperazine (DKP) core functionalized with indole C3-prenyl as well as indole C3- and *N*-methyl groups. The logic of their assembly remains cryptic. Bioinformatics analyses of the *Nocardiopsis* sp. CMB-M0232 draft genome afforded the *noz* cluster, split across two regions of the genome, and encoding putative open reading frames with roles in nocardioazine biosynthesis, including cyclodipeptide synthase (CDPS), prenyltransferase, methyltransferase, and cytochrome P450 homologs. Heterologous expression of a twelve gene contig from the *noz* cluster in *Streptomyces coelicolor* resulted in accumulation of *cyclo*-L-Trp-L-Trp DKP (**5**). This experimentally connected the *noz* cluster to indole alkaloid natural product biosynthesis. Results from bioinformatics analyses of the *noz* pathway along with challenges in actinomycete genetics prompted us to use asymmetric synthesis and mass spectrometry to determine biosynthetic intermediates in the *noz* pathway. The structures of hypothesized biosynthetic intermediates **5** and **12–17** were firmly established through chemical synthesis. LC-MS and MS-MS comparison of these synthetic compounds with metabolites present in chemical extracts from *Nocardiopsis* sp. CMB-M0232 revealed which of these hypothesized intermediates were relevant in the nocardioazine biosynthetic pathway. This established the early and mid-stages of the biosynthetic pathway, demonstrating that *Nocardiopsis* performs indole C3-methylation prior to indole C3-normal prenylation and indole N1′-methylation in nocardioazine B assembly. These results highlight the utility of merging bioinformatics analyses, asymmetric synthetic approaches, and mass spectrometric metabolite profiling in probing natural product biosynthesis.

Received 19th March 2015,
Accepted 29th April 2015

DOI: 10.1039/c5ob00537j

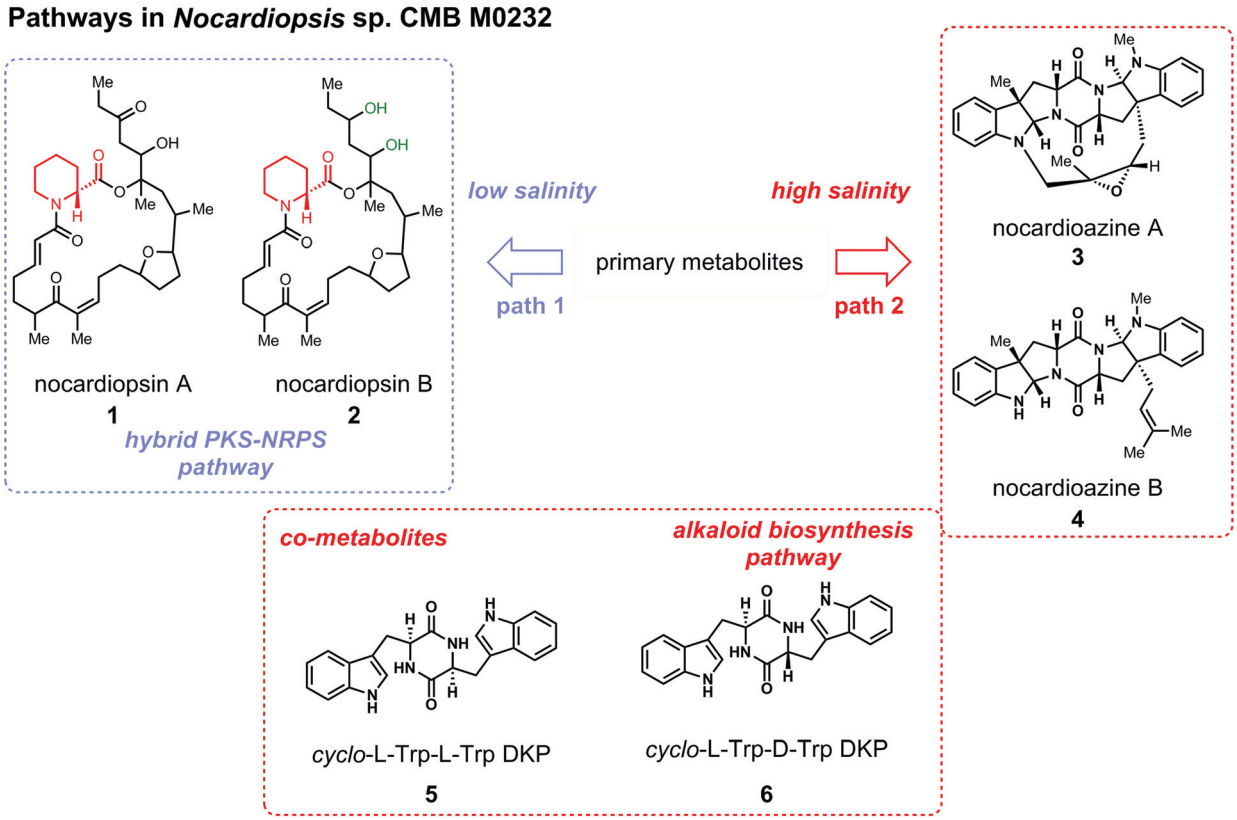
www.rsc.org/obc

Introduction

Marine actinomycetes continue to be rich sources of structurally diverse natural products endowed with promising pharmacological properties.[1] Recently, Capon and co-workers reported the isolation and structural characterization of nocardiopsins[2] (*e.g.* **1** and **2**) and diketopiperazine (DKP) containing nocardioazine alkaloids[3] (**3–6**) from the marine-derived actinomycete *Nocardiopsis* sp. CMB-M0232 (Scheme 1). Intriguingly, under low salinity fermentation conditions, gene regulatory mechanisms predominantly favour the biosynthesis of the hybrid polyketide and nonribosomal peptide-derived nocardiopsins, whose biosynthetic pathway we recently established (path 1, Scheme 1).[4] Under relatively high salinity, DKPs including nocardioazines A and B (**3** and **4**) are dominant (path 2, Scheme 1).

Nocardioazines A and B possess a dimerized tryptophan DKP core. The skeleton comprises seven fused rings (A–B–C–D–C′–B′–A′) in a 6–5–5–6–5–5–6 diannulated manner forming a pyrroloindoline–DKP–pyrroloindoline assembly. Among DKP natural products, nocardioazines A and B (**3** and **4**) stand out as the only C3-prenylated DKPs reported from a bacterial

[a]*Department of Chemistry, Case Western Reserve University, Millis Science Center: Rm 216, 2074, Adelbert Road, Cleveland OH 44106-7078, USA. E-mail: rajesh.viswanathan@case.edu, amy.lane@unf.edu*
[b]*Department of Chemistry, University of North Florida, 1 UNF Drive, Jacksonville, FL 32224, USA*
[c]*Mass Spectrometry Facility, Indiana University Department of Chemistry, 800 E. Kirkwood Ave., Bloomington, IN 47405, USA*
†Electronic supplementary information (ESI) available. See DOI: 10.1039/c5ob00537j
‡These authors contributed equally to this work.
§Present address: Faculty member, Department of Chemistry, DIT University, Dehradun 248009, India.



Scheme 1 Structures of nocardiopsins A (**1**) and B (**2**), nocardioazines A (**3**) and B (**4**), *cyclo*-L-Trp-L-Trp (**5**) and *cyclo*-L-Trp-D-Trp (**6**).

source and the first indole-C3-normal prenylated DKP from any source, implicating a unique biosynthetic pathway. The co-isolation of *cyclo*-L-Trp-L-Trp DKP (**5**) and *cyclo*-L-Trp-D-Trp DKP (**6**) alongside **3** and **4** alludes to **5** or **6** (or one of their epimers, *cyclo*-D-Trp-D-Trp DKP, *ent*-**5**) as likely precursors for the more complex congeners **3** and **4**.[3] The first reported synthesis of nocardioazine B by Wang *et al.* corrected the originally assigned stereochemistry of the natural product, alluding to the possibility of *ent*-**5** as a likely intermediate.[5]

The first enantioselective synthesis of **3** (in addition to other related isoprenylated indole alkaloids), recently reported by the Reisman group, constituted an ingenious strategy towards synthetically assembling the macrocyclic ring E.[6] Given the lack of any prior studies on the characterization of their gene cluster, biosynthetic intermediates and enzymes, the molecular logic of the nocardioazine assembly remains poorly understood. Herein we report the identification of the *noz* gene cluster encoding nocardioazine B biosynthesis from the draft genome sequence of *Nocardiopsis* sp. CMB-M0232 and characterize pathway intermediates. Our approach of employing bio-inspired synthetic molecules to elucidate the molecular logic of natural product assembly represents a relatively overlooked alternative to conventional gene-knockout-guided approaches. As we demonstrate herein, this strategy is particularly valuable in many cases where organisms are not amenable to genetic manipulation.

Results

Nocardiopsis sp. CMB-M0232 draft genome sequence and bioinformatics-based prediction of the *noz* gene cluster

Sequencing and assembly of the *Nocardiopsis* sp. CMB-M0232 genome yielded a ~6.4 Mbp draft with >5500 open reading frames (ORFs) (see the ESI†). The putative *noz* biosynthetic genes are clustered across two separate regions of the *Nocardiopsis* sp. CMB-M0232 chromosome (Fig. 1). Bioinformatics analyses of the ORFs revealed candidate enzymes for nocardioazine biosynthesis (Table 1). BLASTP analyses of individual predicted ORFs in the entire draft genome revealed both putative nonribosomal peptide synthetases (NRPSs) and a cyclodipeptide synthase (CDPS) as candidates for the assembly of the DKP core during the early stage of nocardioazine biosynthesis. However, bioinformatics analyses of adenylation domains from putative NRPSs revealed that none predicted to accept two tryptophan substrates.[7] Further, additional genes clustered with these putative NRPS-encoding genes were strongly suggestive of the biosynthesis of hybrid polyketide synthase-nonribosomal peptide synthase (PKS-NRPS) products[4] and other classes of secondary metabolites, rather than prenylated diketopiperazine alkaloids. Distinctly, a single putative CDPS (NozA) identified in the draft genome represents the most plausible candidate for assembly of *cyclo*-L-Trp-L-Trp DKP (**5**) (Fig. 1, Table 1).

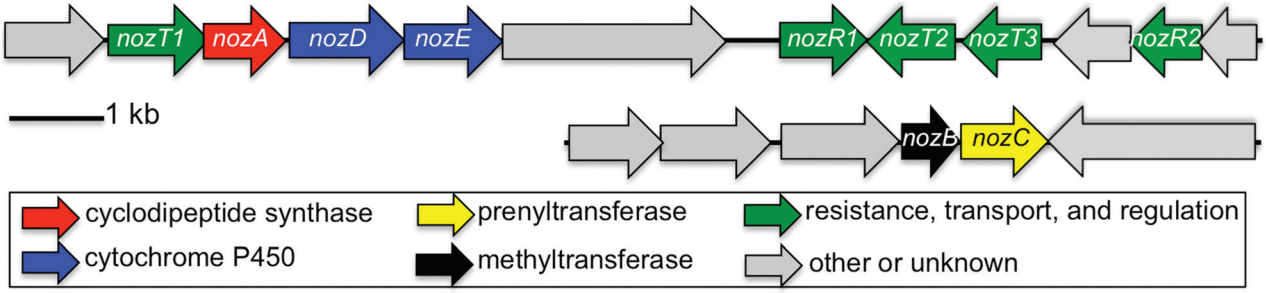


Fig. 1 Organization of the two clusters of *Nocardiopsis* sp. CMB-M0232 biosynthetic genes (*noz*) predicted to play roles in nocardioazine biosynthesis.

Analyses of the *Nocardiopsis* sp. CMB-M0232 genome revealed a single putative prenyltransferase, NozC (Table 1), as the sole candidate for a C3′-normal prenylation of the DKP core. NozC shares homology with enzymes previously annotated as prenyltransferases but for which the biosynthetic function is yet to be experimentally confirmed. However, little homology was noted between NozC and biochemically characterized prenyltransferases including the dimethylallyltryptophan synthases FgaPT2[8] and AnaPT.[9,10] This observation is potentially explained by the unique regioselectivity of NozC as the sole prenyltransferase yielding C3′-normal prenylation. The *nozC* gene is located within a cluster of biosynthetic genes chromosomally distinct from *nozA* (Fig. 1). The *nozC* prenyltransferase gene is located within the same operon as *nozB*, which encodes a putative methyltransferase that is a candidate for C- and N-methylation of the DKP scaffold. Although the regioselectivity of NozB remains unknown, BLASTP analyses revealed that NozB possesses residues conserved among SAM-dependent methyltransferases.[11] In Fig. 1 and Table 1, all genes predicted by bioinformatic analyses to play enzymatic or regulatory roles in nocardioazine biosynthesis pathway are assigned as "*noz*" genes. Following typical conventions (for annotation of gene clusters), those genes annotated with no apparent role in nocardioazine biogenesis are listed as "orfs" and many of these correspond to hypothetical proteins whose function remain unclear.

NozA is a cyclodipeptide synthase homologous to Amir4627 as revealed through bioinformatics

The putative CDPS, NozA, identified by bioinformatics analyses as the most plausible candidate for assembly of *cyclo*-L-Trp-L-Trp DKP (**5**), was compared with sequences of known characterized CDPSs. Amino acid sequence alignment revealed 35% identity between NozA and Amir_4627, a CDPS from *Actinosynnema mirum* and the only known example of a CDPS incorporating two Trp residues (NCBI accession #YP_003102306; Fig. 2).[12] NozA includes residues conserved among related biochemically characterized, catalytically functional CDPSs[13] including Amir_4627.[12] Beyond the conserved active site residues (highlighted in yellow), correlations are also apparent between NozA and Amir_4627 for residues implicated in recognition and binding of NozA to aminoacyl-charged tRNA substrates (highlighted brown). Similar predicted secondary and tertiary structural features are noticeable between the two enzymes (Fig. 2). Given this prediction, we next sought to establish the connection of the gene cluster harboring *nozA* towards the production of *cyclo*-L-Trp-L-Trp (**5**) through heterologous expression in *S. coelicolor*.

Heterologous expression connects contig #1 to *cyclo*(L-Trp-L-Trp) biosynthesis

From the SuperCos 1 cosmid library generated from genomic DNA, the cosmid clone pAL557 was found to carry ~40 kb of the *Nocardiopsis* sp. CMB-M0232 genome, including the entirety of contig #1 (Fig. 1, Table 1). After ensuring that the host organism lacks nocardioazine-like pathway genes, cosmid pAL557 was adapted with genetic elements required for integration into the *Streptomyces* genome and heterologous expression (see the ESI†).[14] This yielded a plasmid (pAL5571), which was introduced by intergeneric conjugation from *E. coli* into *S. coelicolor* M1146, a host engineered for optimized heterologous expression of actinomycete gene clusters.[15] M1146 treatment cultures were fermented in parallel with M1146 controls lacking these biosynthetic genes. Metabolite profiles of chemical extracts from these cultures were compared by HPLC with diode array detection, revealing a signal at 11.2 min as the sole discernable metabolite present in treatment cultures and absent from controls (Fig. 3). The retention time of this metabolite matched that of synthetic *cyclo*(L-Trp-L-Trp), generation of which is described below. Further, high-resolution LC/MS supported the assignment of the molecular formula of this metabolite as $C_{22}H_{20}N_4O_2$ (*m/z* 373.1691 $[M + H]^+$), corresponding with the formula of *cyclo*(L-Trp-L-Trp). Based on bioinformatics-predicted functions of proteins encoded by contig #1 (Fig. 1, Table 1), NozA represents a plausible candidate for catalyzing *cyclo*(L-Trp-L-Trp) biosynthesis. Ongoing investigations are directed at experimentally establishing the function of NozA.[16]

Assembly of predicted downstream *noz* pathway intermediates

Two specific reasons prompted us to turn to synthesis and tandem MS for furthering our knowledge of the *noz* pathway. First, given the recent advancements (in the post-genomics

Table 1 Predicted functions of putative nocardioazine biosynthetic enzymes based on bioinformatics analyses. Two chromosomally distinct gene clusters (contig 1–2) encode these enzymes. #aa = number of amino acid residues; ID = % identity; Sim = % Similarity

Contig #1	#aa	BLASTP annotation	NCBI accession number of homolog	Organism	ID/Sim (%)
orf1	421	Hypothetical PLP-dependent protein	WP_012786924	*Catenulispora acidiphila* DSM 44928	51/62
nozT1	402	transporter	AEF16056	*Streptomyces vinaceusdrappus* NRRL 2363	58/72
nozA	234	Cyclodipeptide synthase	YP_003102306	*Actinosynnema mirum* DSM 43827	35/54
nozD	385	Cytochrome P450	KCP45129	*Mycobacterium tuberculosis* BTB09-382	30/42
nozE	398	Cytochrome P450	WP_026248114	*Streptomyces* sp. MspMP-M5	36/52
orf6	825	Adenylosuccinate synthase	WP_016473091	*Streptomyces* sp. HPH0547	57/70
nozR1	307	XRE family transcriptional regulator	WP_020869817	*Streptomyces rapamycinicus*	80/89
nozT2	362	ABC transporter	WP_013153583	*Nocardiopsis dassonvillei*	67/77
nozT3	289	ABC transporter	WP_017566834	*Nocardiopsis synnemataformans*	76/88
orf10	277	Hypothetical protein	WP_018657142	*Actinomadura flavalba*	59/69
nozR2	136	MerR family transcriptional regulator	WP_017544945	*Nocardiopsis prasina*	76/84
orf12	216	Hypothetical protein	ERT00338	*Sporothrix schenckii*	36/52
Contig #2	#aa	BLASTP annotation	NCBI accession number of homolog	Organism	ID/Sim (%)
orf13	371	Hypothetical protein	WP_017620974	*Nocardiopsis giba* YIM 90087	73/82
orf14	318	3-Ketoacyl-ACP reductase	WP_018724690	*Salinispora pacifica* CNS055	63/74
orf15	456	Family 1 glycosyltransferase	WP_017620972	*Nocardiopsis giba* YIM 90087	85/92
nozB	212	Indole C3′ and N1′ methyltransferase	WP_026123683	*Nocardiopsis chromatogenes* YIM 90109	84/92
nozC	358	Indole C3′ prenyltransferase	WP_017620970	*Nocardiopsis giba* YIM 90087	80/84
orf18	781	Hypothetical protein	WP_017625718	*Nocardiopsis chromatogenes* YIM 90109	68/80

era) in the employment of tandem-MS to connect molecules to individual gene clusters,[17] we anticipated that assembly of synthetic intermediates may lead to conclusive evidence to support the *noz* pathway. Further, the use of tandem-MS-guided strategies can illuminate the biosynthetic relationships between multiple pathways encoded by respective gene cluster families.[18] Second, our initial efforts to probe nocardioazine biosynthesis focused on the conventional approach of generating *Nocardiopsis* sp. CMB-M0232 gene replacement mutants with the intention to employ them as tools for determining biosynthetic intermediates. Thus far, *Nocardiopsis* sp. has proven resistant to select gene knockout experiments. Therefore, we turned to the alternative bio-guided synthesis and tandem-MS-centric strategy presented herein to experimentally establish nocardioazine biosynthetic intermediates predicted through bioinformatics analyses.

To provide synthetic standards for the *in vitro* characterization of NozA-catalyzed *cyclo*(L-Trp-L-Trp) and for the assembly of downstream pathway intermediates, we constructed **5** and *ent*-**5**. *Cyclo*-(L-Trp-L-Trp) (**5**) was constructed through a four-step sequence, starting with protection of the amino functionality of L-Trp with the benzyloxycarbonyl (Cbz) group (Scheme 2). Treatment with Cbz-Cl along with sodium bicarbonate–sodium carbonate in acetonitrile–water (2 : 3, v/v) as solvents over 3 h resulted in **8** providing the western half of the DKP. Similarly, treatment of L-Trp under thionyl chloride in methanol at reflux over 18 h resulted in the formation of the L-Trp methyl ester (**9**) in near quantitative yield, providing the eastern half of the DKP. BOPCl-mediated coupling of **8** and **9** in the presence of triethylamine as a base in THF resulted in amide **10** in 93% yield. BOPCl-mediated activation of the carboxylic acid functionality of **8** proved the most efficient for the isolation of a high yield of amide product **10**. Deprotection of the Cbz group in **10** under hydrogenating conditions in the presence of Pd–C in MeOH (with a trace amount of water) yielded deprotected amine precursor **15** which also contained an ester functionality as an intramolecular reactive partner. The DKP ring system was then formed through treatment of **11** under 14 M ammonia in methanol at 60 °C for 8 h, resulting in *cyclo*-(L-Trp-L-Trp) (**5**) in 95% yield. Likewise, an identical sequence was applied starting from D-Trp (through protection resulting in *ent*-**8** and ester *ent*-**9**, followed by coupling to give *ent*-**10**, finally with deprotection–cyclization steps), resulting in the formation of *cyclo*-(D-Trp-D-Trp) (*ent*-**5**) in excellent overall yield. The four-step sequence was reproduced consistently with identical % yields for either antipode, as shown in Scheme 2A. As shown in Scheme 2B, we were able to mono-methylate the N1′ position of **5** to synthesize **14**.

Given the bioinformatics-based prediction and homology comparisons of enzymes, we collectively identified **5**, *ent*-**5**, and **12–17** as candidates for *in vivo* intermediates in the *noz* pathway. NozB and NozC are expected to catalyze prenylation and methylation steps to yield six unique potential intermediates (**12–17**) depending on the order of reactions (as described later in Scheme 5). Additionally, we expected the synthetic endeavour to afford relevant intermediates for future *in vitro*

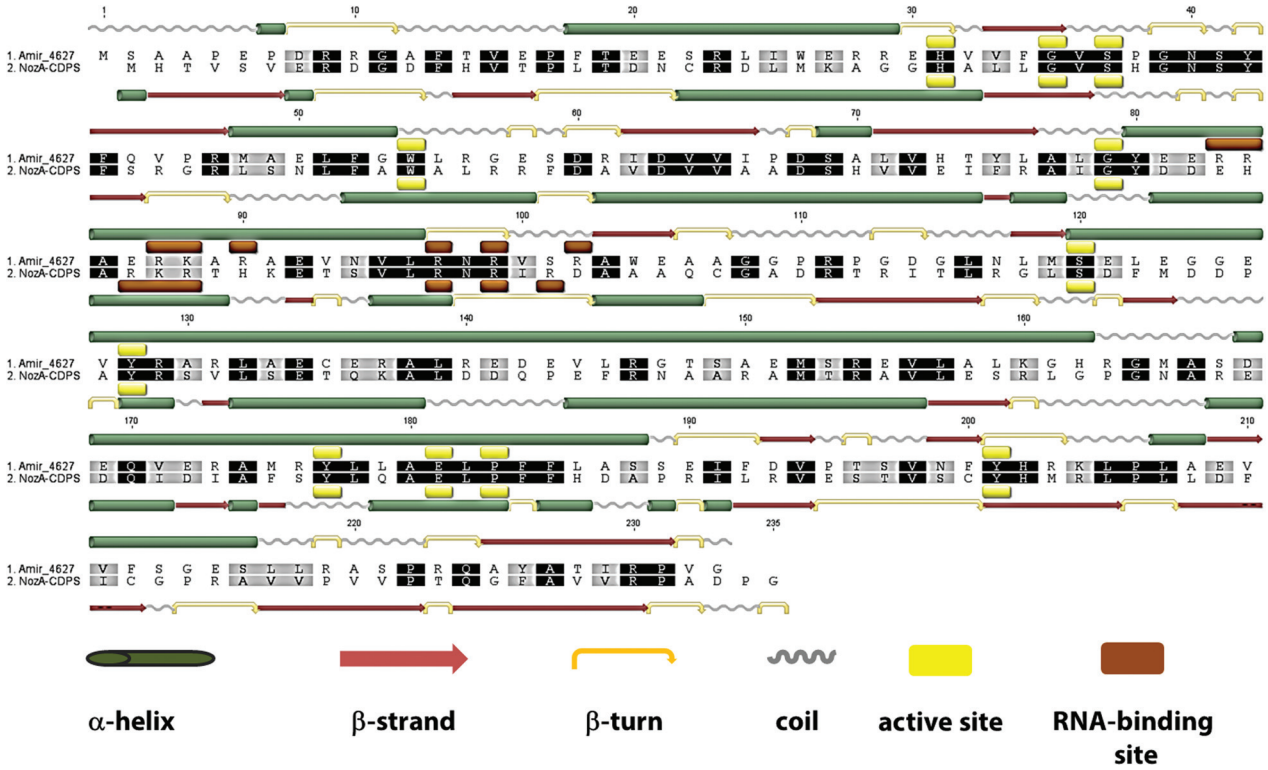


Fig. 2 Amino acid sequence alignment between NozA and Amir_4627 and bioinformatics model of NozA generated using Geneious™. Clustal was used for basic sequence alignment.

and *in vivo* reconstitution assays of individual steps catalyzed by NozA, NozB and NozC in the nocardioazine pathway.

Due to the relative complexity of the proposed intermediates, regio- and stereoselective C3′-prenylation, C3-methylation and N1′-methylation presented significant challenges. As illustrated in Scheme 3A, 3-methyl indole (**18**) and enamide (**19**) served as reasonable starting points to employ an enantio- and diastereoselective indole-enamide [3 + 2] cycloaddition reaction in the presence of (*S*)-BINOL and tin(IV) chloride as a key step to install the C3-methyl functionality.[19] En route to employing this step as a strategy towards assembling **13** and **16**, enamide **19** was prepared from L-serine through the conversions involving the corresponding *O*-Boc derivative (see the ESI†). The [3 + 2] cycloaddition between **18** and **19** proceeded with a 12 : 1 diastereomeric ratio favouring the *exo* isomer **20a** over the minor *endo* isomer **20b**. Each diastereomer exhibited a 2 : 3 ratio of conformational isomers (caused by the Cbz group at N11 position) as revealed by the presence of equivalent sets of ^{1}H NMR signals. The overall yield of the [3 + 2] cycloaddition product **20** is 61%. Considering the relative stereochemical disposition of substituents at C2, C3 and C9 in the major *exo* isomer **20a**, we initially attempted an LDA-mediated deprotonation–reprotonation sequence to invert the C9 center. By virtue of the lack of an allyl protecting group at N1 (that was present in a previous report[6]), a retro-Michael addition occurred, resulting in degradation of **20a** under strongly basic conditions. Therefore, a revision of appropriate conditions to achieve the correct relative stereochemistry was imminent. After careful screening and optimization, treatment with excess lithium hydroxide in a mixture of methanol, water and THF (1 : 1 : 1; v/v/v) affected this transformation efficiently to yield **20c** (Scheme 3A). Concomitant with the epimerization, we observed base-mediated hydrolysis of the carboxymethyl ester functionality. It proved to be a beneficial outcome, as the next step *en route* to **13** involved an amide bond forming coupling to an L-Trp-containing partner **9**. Similarly to the formation of **10**, we observed smooth peptide bond formation under BOP-Cl-mediated activation of **20c** followed by nucleophilic participation of the amino functionality of **9** resulting in the coupled product **22** in 90% yield in THF as the solvent (Scheme 3B). Hydrogenative deprotection of the Cbz group of **22** was affected smoothly to result in **23**. During this deprotection of the Cbz group under Pd–C, we observed direct intramolecular cyclization resulting in the formation of **13** in ~10% efficiency. However, this low conversion rate motivated the employment of a relatively stronger base[20] involving methanolic NH_3 to result in the formation of *cyclo*-C3-Me-L-Trp-L-Trp DKP (**13**) in high efficiency through the participation of the secondary amino functional group through an internal nucleophilic substitution reaction. The overall synthetic sequence is 5 linear steps starting from 3-methyl indole (**18**). The overall yield for formation of **13** was 24.9%. Similarly, the assembly of **16** began with **20a** undergoing a tandem epimerization–hydrolysis event under aqueous lithium hydroxide,

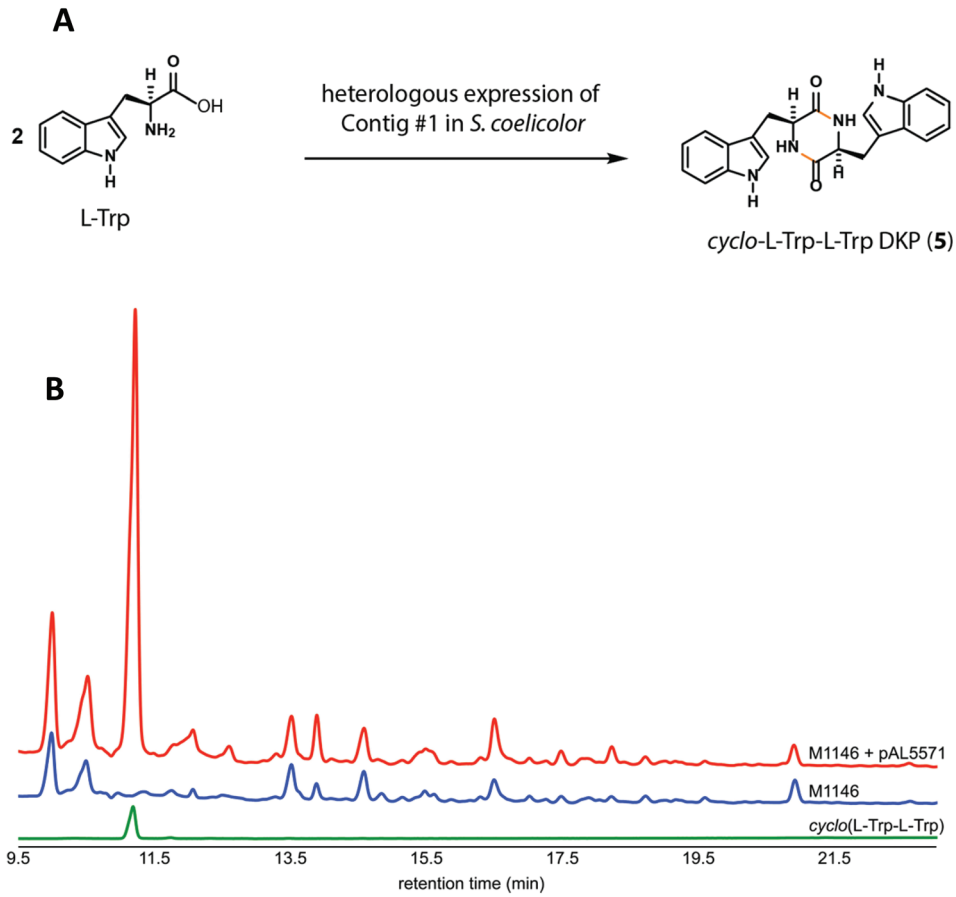
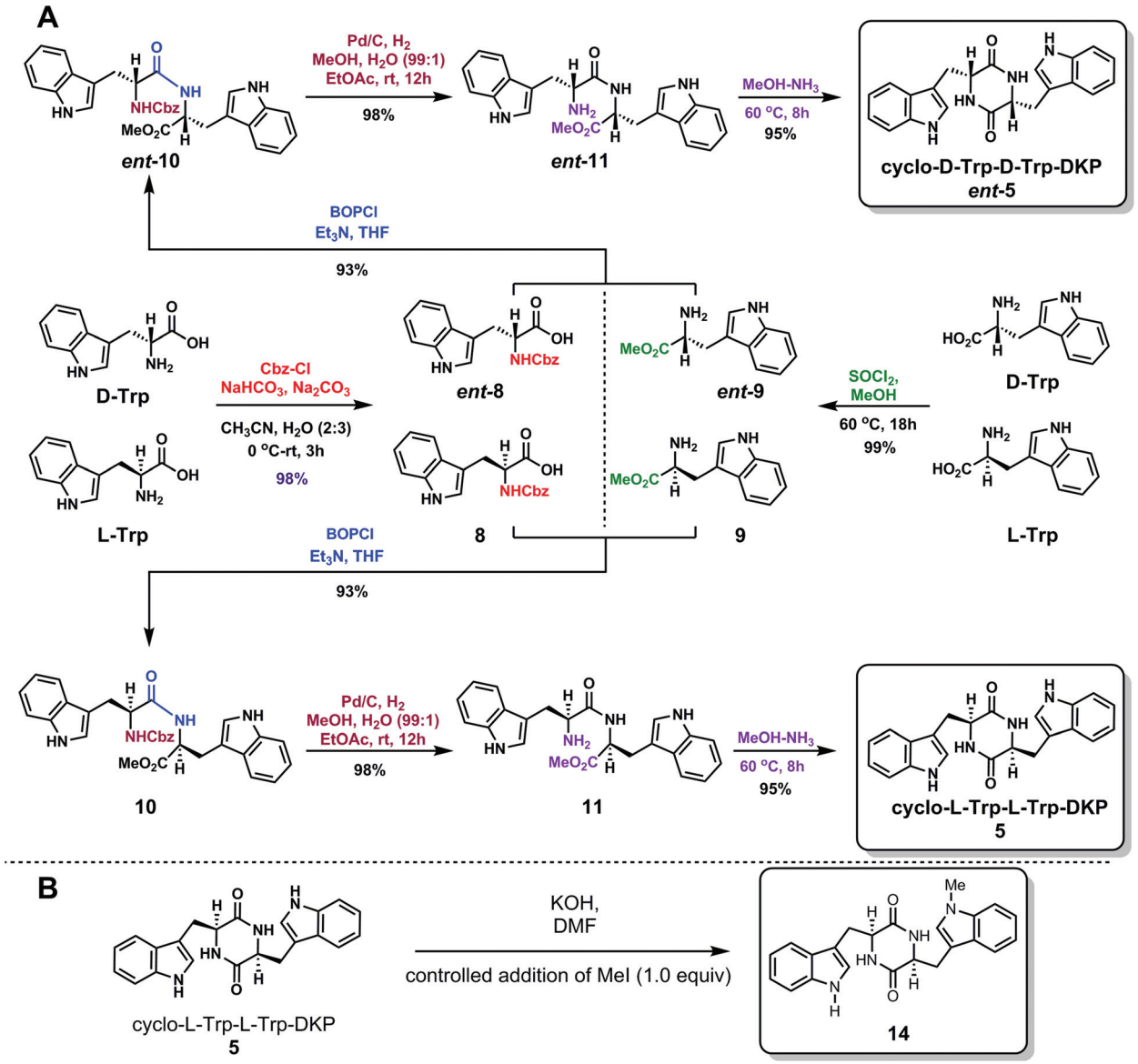


Fig. 3 (A) Formation of *cyclo*-L-Trp-L-Trp (**5**). (B) HPLC chemical profiles of *S. coelicolor* M1146 treatment with pAL5571 (shown in red), control M1146 (shown in blue), and *cyclo*(L-Trp-L-Trp) (**5**) standard (shown in green). UV detection at 280 nm revealed *cyclo*(L-Trp-L-Trp) (**5**) produced by treatment cultures carrying contig #1 genes but absent from controls lacking these biosynthetic genes. This suggested that NozA catalyzes the biosynthesis of this nocardioazine precursor.

yielding **20c**. N1′-methylated L-Trp (**21**) was synthesized (from L-Trp, see the ESI†) for its engagement in a coupling step with **20c**. Likewise, **20c** under BOP-Cl activation and triethylamine gave **24** as the product in 81% yield. Similarly to the non-methylated counterpart **22**, we could effect a hydrogenative deprotection followed by a base-mediated intramolecular cyclization event on **24** to result in **16** (*via* **25**) in fairly high efficiency (91% yield) in 5 linear steps from commercially available **18**. The overall yield for the formation of **16** was 31.0%.

As illustrated in Scheme 4A, we aimed at *cyclo*-L-Trp-C3′-^{n}prenyl-L-Trp DKP (**12**) and its N1′-methylated variant *cyclo*-L-Trp-N1′-Me-C3′-^{n}prenyl-L-Trp DKP (**15**) as synthetic targets. Through a biomimetic prenylation method we published recently,[21] employment of the methyl ester of L-tryptophan (**9**) served as a precursor to engage in a domino process initiated by a C3′-prenylation event (with prenyl bromide as the electrophile), subsequently resulting in a C–N bond-forming pyrroloindoline cyclization, under sodium acetate–acetic acid conditions (pH = 2.7) at room temperature, to result in the formation of **27a** and **27b** as a 4 : 1 mixture of *exo* and *endo* diastereomers. The overall yield for this transformation was 67% considering full recovery of unreacted **9**. The fact that **27a** and **b** were accessed through a single biomimetic step afforded direct access to the C3′-normal prenylated scaffold of nocardioazines. The stereochemical relationship between C3′-^{n}prenyl substitution, C2′–H and C9′-carboxymethyl substituent for the major diastereomer **27a** was established through NOESY correlations (see the ESI†). Upon treatment of **27a** with *N*-phthalyl-protected L-Trp-acid **29** (prepared previously using a one-step protection reaction with phthalic anhydride, see the ESI†) under BOP-Cl activation and basic conditions, we obtained the coupled product **30** (comprising the carbon skeleton of target **12**) in 90% yield. Gratifyingly, the coupled product **30** underwent a tandem sequence initiated by a hydrazine hydrate-mediated deprotection of the phthalyl group followed by an intramolecular cyclization in methanol–dichloromethane and resulted in a 70% yield of *cyclo*-C3′-^{n}prenyl-L-Trp-L-Trp DKP (**12**). The NOESY experiment showed a 2.98% enhancement between C8′–H and olefinic C2″–H; a 1.78% enhancement between protons at 2′ and 8′-α CH; and finally a 3.05% enhancement between protons at 8′β CH and 9′ positions. These confirmed the stereochemistry to

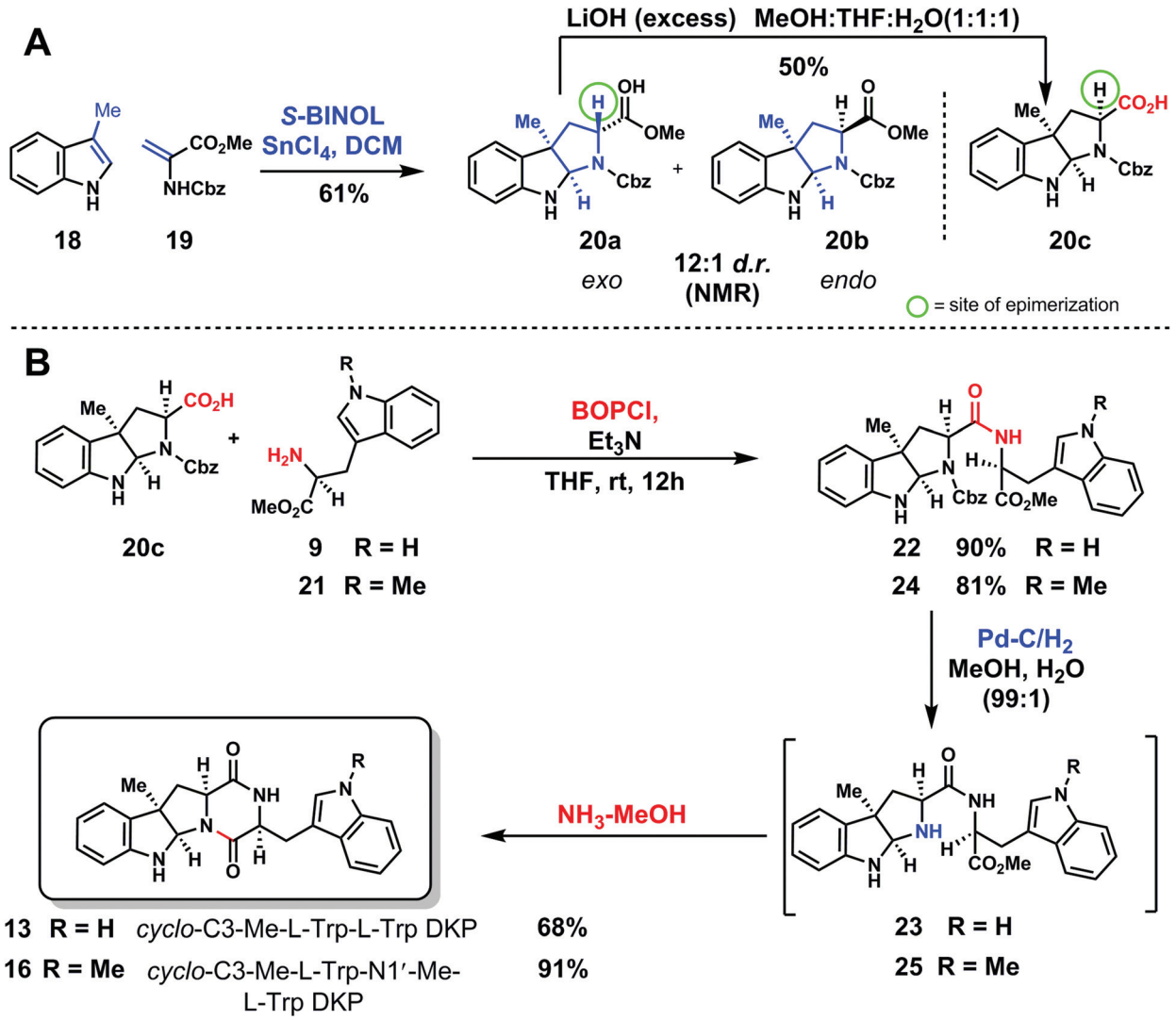


Scheme 2 (A) Synthesis of *cyclo*(L-Trp-L-Trp) (**5**) and *cyclo*(D-Trp-D-Trp) (*ent*-**5**). (B) Synthesis of N1′-Me-*cyclo*(L-Trp-L-Trp) (**14**) from **5**. Predicted candidate intermediates of the *noz* pathway are presented in box.

be *cis* across the DKP ring system and an overall *exo* arrangement for the B′–C′ pyrroloindoline ring fusion. Likewise, engagement of N1′-methylated-L-Trp carboxymethyl ester (**26**) in a one-step prenylation (in aqueous solution) using prenyl bromide resulted in 72% overall yield of C3′-prenylated **28a** (major) and **28b** (minor) based on recovery of unreacted **26**. Similarly to the formation of **30**, upon subjecting **28a** to a coupling reaction with **29** using BOP-Cl and triethylamine in THF, we obtained **31** which upon subjecting to a hydrazine hydrate-mediated deprotection–cyclization sequence resulted in the B′–C′ ring-forming process leading to *cyclo*-L-Trp-N1′-Me-C3′-^{n}prenyl-L-Trp DKP (**15**) in 74% yield. The overall yields for the formation of **12** and **15** were 42.41% and 44.6% over 3 linear steps respectively. *Cyclo*-C3-Me-L-Trp-L-Trp DKP (**13**) underwent C3′-prenylation (similar to prenylations on **9** and **26**) to result in *des*-N1′-Me-Nocardioazine B (**17**). In addition to NMR indicating the presence of a mixture of diastereomers, the identity of **17** for biosynthetic characterization is supported by HPLC (Fig. 5), HRMS (Table 2) and LC-MS-MS (Fig. 4).

Evaluation of the biosynthetic relevance of synthesized intermediates through NMR, LC-MS and HR-tandem MS reveals precursor–product relationships for nocardioazine B biosynthesis

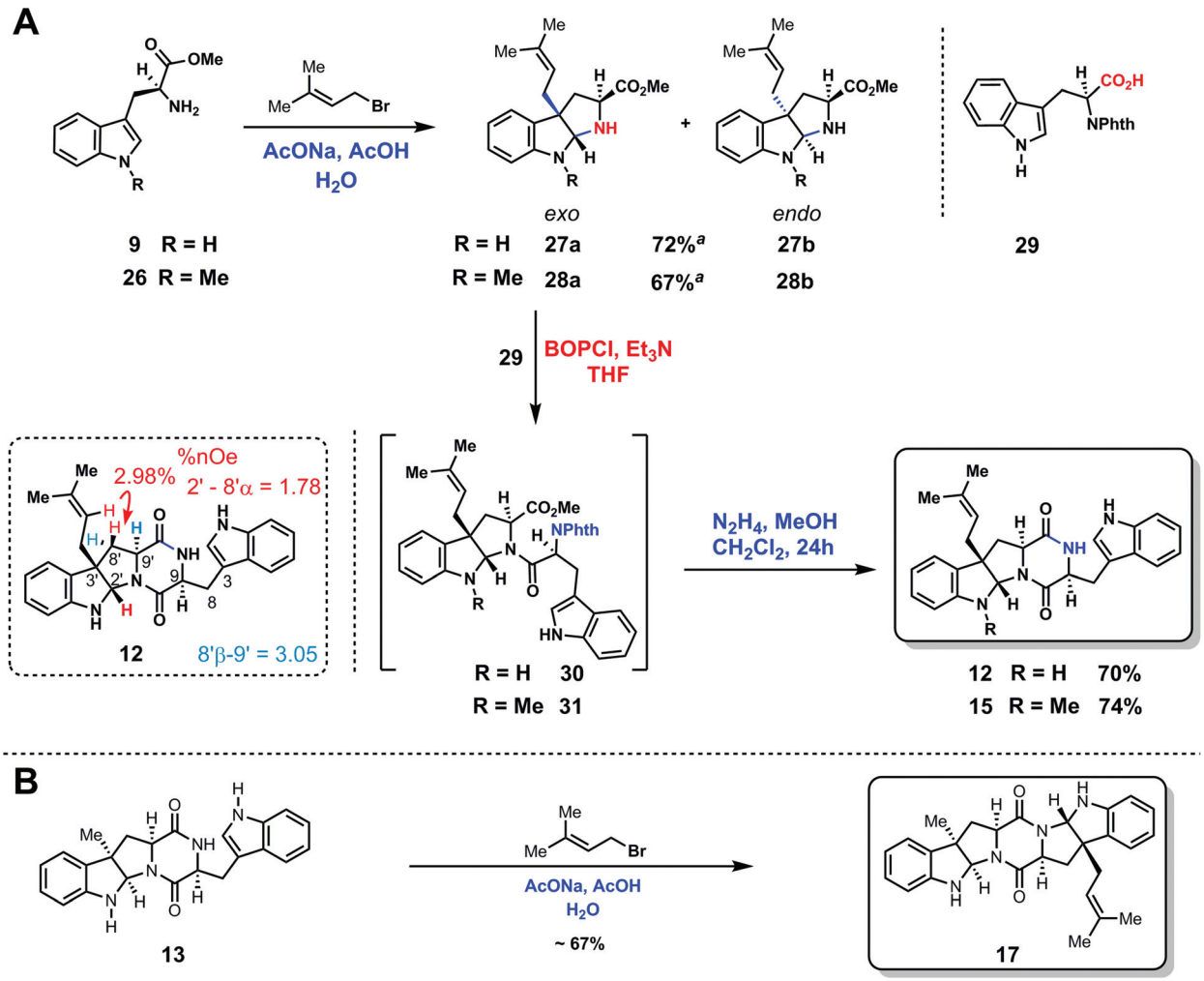
Having synthesized candidate intermediates of the *noz* pathway, we applied LC-coupled-tandem-MS as a tool to establish nocardioazine alkaloidal biosynthetic intermediates. While relatively simpler L-Trp-L-Trp DKP (as products of cyclo-



Scheme 3 (A) Asymmetric C3-methylation to yield **20**. (B) Synthesis of *cyclo*-C3-Me-L-Trp-L-Trp DKP (**13**) and *cyclo*-C3-Me-L-Trp-N1′-Me-L-Trp DKP (**16**). Predicted candidate intermediates of the *noz* pathway are presented in box.

dipeptide synthase biosynthesis) and other dimeric amino acid DKPs have been analysed through tandem mass spectrometry,[22] complex DKPs like **12–17** were thus far not investigated through mass spectrometry, adding further importance to this study. EIC traces (Fig. 4A) indicated that HPLC-MS profiles uniquely separated and distinguished most synthesized intermediates. Unique signatures are observable in MS^2 spectra for each biosynthetic metabolite (Fig. 4B–H). Specifically, Fig. 4B shows the presence of *cyclo*-L-Trp-L-Trp DKP (**5**). Its $[M + H]^+$ ion (at 373.1662 Da in positive ion mode ESI-MS profile) and its $[M - H]^-$ ion (at 371.1530 Da in negative ion mode ESI-MS) are noticeable (see the ESI†). Characteristic Trp fragments were observed as fingerprints of **5** through MS^2 fragmentation (Fig. 4B and Table S2†). The product molecular ions arise out of neutral losses of a Trp moiety (129 Da), along with sequential loss of CO (28 Da) and/or $HCONH_2$ (45 Da), in various combinations. Neutral loss of HCN (27 Da) from ions at m/z = 130.0654 accounted for the presence of m/z = 103.0547 ion. Likewise, we mapped tandem MS signatures of *cyclo*-L-Trp-C3′-^{n}prenyl-L-Trp DKP (**12**); *cyclo*-C3-Me-L-Trp-L-Trp DKP (**13**); *cyclo*-N1′-Me-L-Trp-L-Trp DKP (**14**); *cyclo*-L-Trp-C3′-^{n}prenyl-N1′-Me-L-Trp DKP (**15**); *cyclo*-C3-Me-L-Trp-N1′-Me-L-Trp DKP (**16**) and *des*-N1′-Me Nocardioazine B (**17**), as illustrated in Fig. 4 (additionally in the ESI†).

Next, we looked for signatures of **12–17** directly from cultures of *Nocardiopsis* sp. CMB-M0232 to detect their presence as biosynthetic intermediates *in vivo*. Reverse-phase HPLC uniquely identified synthetic **12–16** (Fig. 5A). Comparison of retention times of these synthetic compounds with the alkaloidal fractions of *Nocardiopsis* sp. CMB-M0232 revealed that **12**, **14** and **15** were *not* relevant biosynthetic products or intermediates. Interestingly, we detected the presence of three relevant metabolites in *Nocardiopsis* sp. extracts, namely, **13**, **16** and **17**. Further supporting its biosynthetic relevance, TLC pat-



Scheme 4 (A) Synthesis of *cyclo*-L-Trp-C3′-^{n}prenyl-L-Trp DKP (**12**) and *cyclo*-L-Trp-N1′-Me-C3′-^{n}prenyl-L-Trp DKP (**15**). *a* represents % isolated yield based on the recovered starting material. (B) Synthesis of *des*-N-Me-Nocardioazine B (**17**). Predicted candidate intermediates of the *noz* pathway are presented in box.

terns of extracts showed the presence of **13** (see Fig. S11†). We compared the MS^2 fragmentation pattern of **13** extracted from the bacterial culture to that of the synthetic standard. The ESI-TOF-MS-MS data for the synthesized **13** overlapped precisely with that of the extracted metabolite (Fig. 5C and D). The ion at *m*/*z* = 256.110 (C3-methyl group containing Trp-DKP) from the *m*/*z* 385.2 precursor ion after loss of a neutral Trp unit (129 Da) was observed both in the synthesized standard as well as from the extract. Further, ^{1}H and ^{13}C NMR analyses (of LC-derived extracts) confirmed the structure of this metabolite as **13** (Fig. S12 and S13, ESI†). Despite attempting several solvent conditions and flow rates, we were unable to distinctly separate **17** out of overlap in retention time from **13** and **16**. Overall, this approach of combining synthesis, LC and tandem MS gave a global picture of the biosynthetic map for the *noz* pathway. The map in Scheme 5 was derived from LC-MS^2 investigations for all of the intermediates. An intermediate was considered "observed" if its retention time and MS^2 pattern seen in bacterial extracts matched those of the synthetic standard.

Upon consideration of the three mid-stage enzymatic steps in a simple permutation fashion (Scheme 5), the multi-pronged approach reveals the relevance of C3-methylation as a step preceding C3′-prenylation and N1′-methylation. Specifically, if *Nocardiopsis* sp. CMB-M0232 were to employ an indole C3′-prenyltransferase (hypothetical NozC) to install a dimethyl allyl group on DKP **5**, then the product of this biosynthetic reaction is expected to be *cyclo*-L-Trp-C3′-^{n}prenyl-L-Trp DKP (**12**). The formation of pyrroloindoline cycle (of **12**) during this prenyltransfer step is based on fungal precedents such as FgaPT2.[9] Alternatively, if the indole C3-methyltransferase (NozB) were operative on the basic early-stage intermediate DKP **5**, the product expected out of this transformation is represented by *cyclo*-C3-Me-L-Trp-L-Trp DKP (**13**). The corresponding N1′-methylated product from the action of a methyltransferase (NozB, but regioselectively on the N1′ posi-

Table 2 LC-MS and MS-MS data for synthetic and extracted biosynthetic intermediates and products. **NF** – not found; for full detailed listing and corresponding formulas of molecular ions, see the ESI. Note: metabolites with M_w = 369, 383, 482, 466, 468 and 452 were identified as DKPs in Capon's study.[2,4] Ions shown in **bold** are identified from extracts of *Nocardiopsis* sp. CMB-M0232. * No synthetic standard, exact mass match only. Shaded entries represent those metabolites experimentally observed in *Nocardiopsis* extracts

Name of metabolite	Molecular formula	ESI HR-MS $[M]^+$ $[M + H]^+$ $[M - H]^-$ (expected) Found	LC retention time found (synth.) (min)	MS^2 fragmentation pattern	Biosynthetic role	Observations
Cyclo-L-Trp-L-Trp DKP (**5**) and *cyclo*-D-Trp-D-Trp DKP (*ent*-**5**)	$C_{22}H_{20}N_4O_2$	(372.1586) (373.1659) (371.1513) **373.1665 (+) and 371.1530 (−)**	**7.06** (7.06)	242.0925; 144.0805 and 130.0654	Early stage intermediate	$[M + H]^+$ and $[M - H]^-$ observed in extracts Matches with synthetic standard
Cyclo-L-Trp-C3′-^{n}prenyl-L-Trp DKP (**12**)	$C_{27}H_{28}N_4O_2$	(440.2212) (441.2285) (439.2139) **NF**	**NF** (23.2)	373.1671; 242.0931; 113.0337; 198.1288; 183.1044; 130.0658	Mid-stage product of C3′-prenyltransfer on **5**	Not detected
Cyclo-C3-Me-L-Trp-L-Trp DKP (**13**)	$C_{23}H_{22}N_4O_2$	(386.1743) (387.1816) (385.1670) **387.1825 and 385.1711**	**11.8 (11.8)**	385.1690; 256.110; 130.065	Mid-stage product of C3-methyltransfer on **5**	$[M + H]^+$ and $[M - H]^-$ observed in extracts Matches with synthetic standard
Cyclo-N1′-Me-L-Trp-L-Trp DKP (**14**)	$C_{23}H_{22}N_4O_2$	(386.1743) (387.1816) (385.1670) **387.1825 385.1711**	**NF** (9.5)	242.0932, 184.0761, 144.0814	Mid-stage product of N1′M methyltransfer on **5**	Not detected
Cyclo-L-Trp-N1′-Me-C3′-^{n}prenyl-L-Trp DKP (**15**)	$C_{28}H_{30}N_4O_2$	(454.2369) (455.2442) (453.2296) **NF**	**NF** (18.0)	399.1806; 212.1441; 144.0813; 130.0657	Mid-stage product of indole N1′-methyltransfer on **12**	Not detected
Cyclo-C3-Me-L-Trp-N1′-Me-L-Trp DKP (**16**)	$C_{24}H_{24}N_4O_2$	(400.1899) (401.1972) (399.1826) **401.1972**	**12.2** (12.2)	401.1981, 256.1089, 184.0761, 144.0813	Mid-stage indole N1′-methyltransferase product from **13**	$[M + H]^+$ observed in extracts Matches with synthetic standard
Des-N1′-Me-nocardioazine B (**17**)	$C_{28}H_{30}N_4O_2$	(454.2369) (455.2442) **455.2442**	**12.6** (12.6)	256.1089; 184.0761; 144.0813	Putative precursor to secondary metabolite product **4**	$[M + H]^+$ observed in extracts
Nocardioazine B (**4**)	$C_{29}H_{32}N_4O_2$	(468.2525) (469.2598) **469.2695**	**18.44***	186.0914; 156.0812; 144.0805 and 130.0654	Putative precursor to secondary metabolite product **3**	$[M + H]^+$ observed in extracts
Nocardioazine A (**3**)	$C_{29}H_{30}N_4O_3$	(482.2318) (483.2396) (481.224) **483.2396**	**8.80***	483.2396	Secondary metabolite product	$[M + H]^+$ observed in extracts

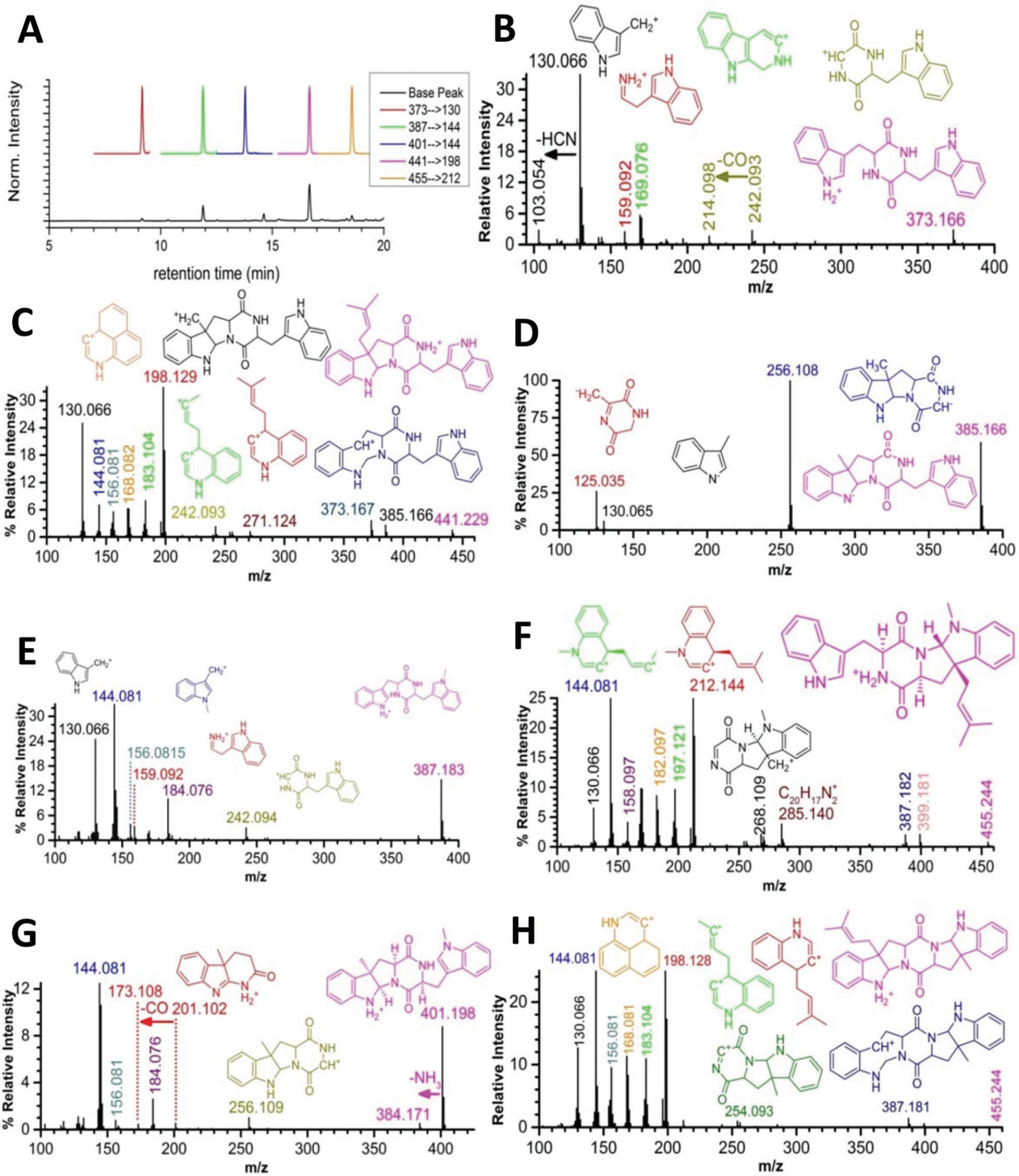


Fig. 4 (A) Extracted ion chromatogram (EIC) traces of synthetic standards. ESI TOF-MS^2 fragmentation data: (B) *cyclo*-L-Trp-L-Trp DKP (**5**) $(M + H)^+$; (C) *cyclo*-L-Trp-C3′-^{n}prenyl-L-Trp DKP (**12**) $(M + H)^+$; (D) *cyclo*-C3-Me-L-Trp-L-Trp DKP (**13**) $(M - H)^-$; (E) *cyclo*-N1′-Me-L-Trp-L-Trp DKP (**14**) $(M + H)^+$; (F) *cyclo*-L-Trp-N1′-Me-C3′-^{n}prenyl-L-Trp DKP (**15**) $(M + H)^+$; (G) *cyclo*-C3-Me-L-Trp-N1′-Me-L-Trp DKP (**16**) and (H) *des*-N1′-Me nocardioazine B (**17**).

tion) would be *cyclo*-L-Trp-N1′-Me-L-Trp DKP (**14**). Products **15**, **16** and **17** represent further increase of complexity through a subsequent enzymatic event. Their relevance in a combinatorial way is discussed in Scheme 5. N1′-methylation of **13** will lead to production of **16** and therefore we anticipated the presence of **16** from cultured *Nocardiopsis* sp. CMB-M0232. Indeed, LC-ESI-(+)TOF MS-MS analysis indicated the presence of **16** at R_t = 12.2 min (Fig. S6 in the ESI†). HRMS verification of its

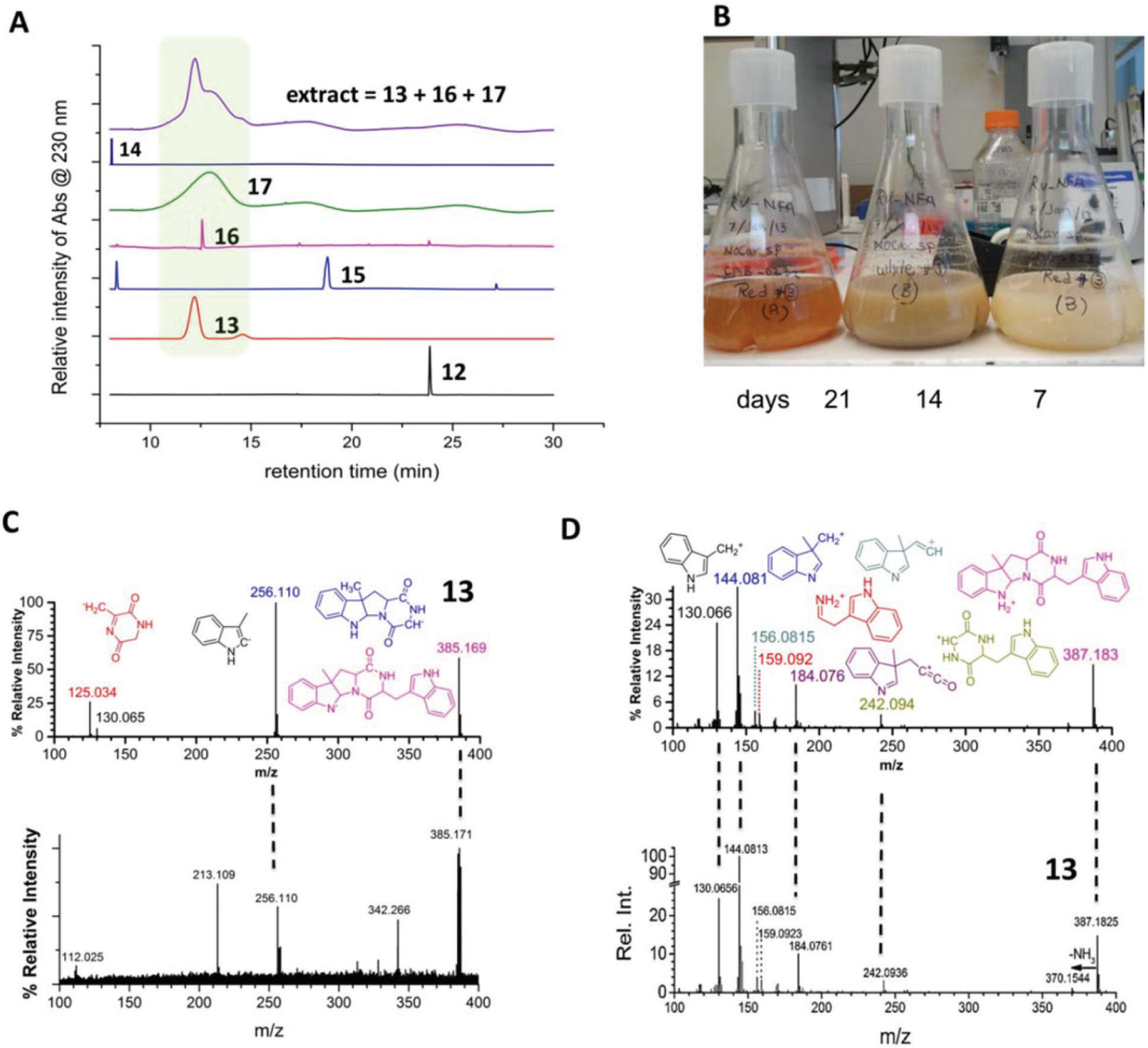


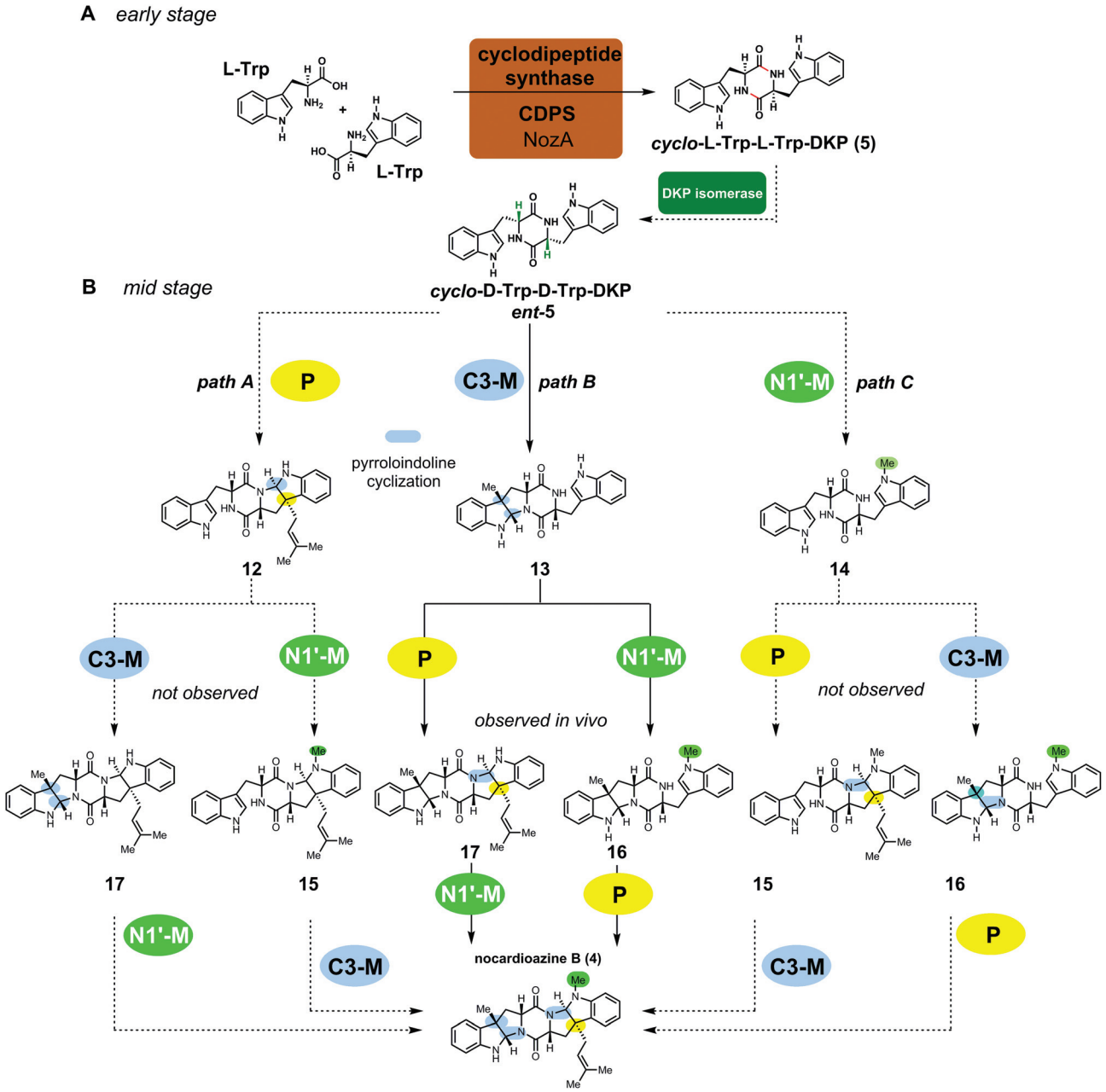
Fig. 5 (A) Reverse phase HPLC traces for synthesized and extracted metabolites from *Nocardiopsis* sp. CMB-M0232. (B) Culture of *Nocardiopsis* sp. CMB-M0232 at 7, 14 and 21 days. (C) ESI-(–)-TOF-MS-MS spectrum of **13** from *Nocardiopsis* sp. CMB-M0232 (top) matched with the spectrum of synthesized **13** (bottom). (D) ESI-(+)-TOF-MS-MS spectrum of **13** from *Nocardiopsis* sp. CMB-M0232 (top) matched with the spectrum of synthesized **13** (bottom).

presence was confirmed through the observation of an ion at m/z = 401.1972 and, furthermore, MS^2 fragmentation revealed the presence of characteristic ions at m/z = 256.109, 184.076, and 144.081 that matched well with the synthesized **16** (Table 2). The presence of *des*-N1′-Me nocardioazine B (**17**) was detected through the identification of a broad LC peak at R_t ~ 12.6 min that corresponded to an HR-MS signal at m/z = 455.2442 (Δm = 0 ppm). Its corresponding MS^2 spectra revealed signature peaks at m/z 256.1089 (seen in fragmentation of **13** + 14 Da), 184.0761, and 144.0813 typically observed for all synthetic standards possessing the C3-methyl substitution and two Trp units of the DKP ring system (Table 2). Significantly, **12**, **14** and **15** were *not* identified from cultures of *Nocardiopsis* sp. CMB-M0232 as verified through the conspicuous absence of signature prenylated ions at m/z = 198.129 ($C_{14}H_{16}N^+$) observed only in **12** and a corresponding ion seen only for **15** at m/z = 212.144 ($C_{15}H_{18}N^+$). Pathway specific metabolites identified through LC and their MS-MS fragmentation (in this study) are highlighted in green in Table 2.

Discussion

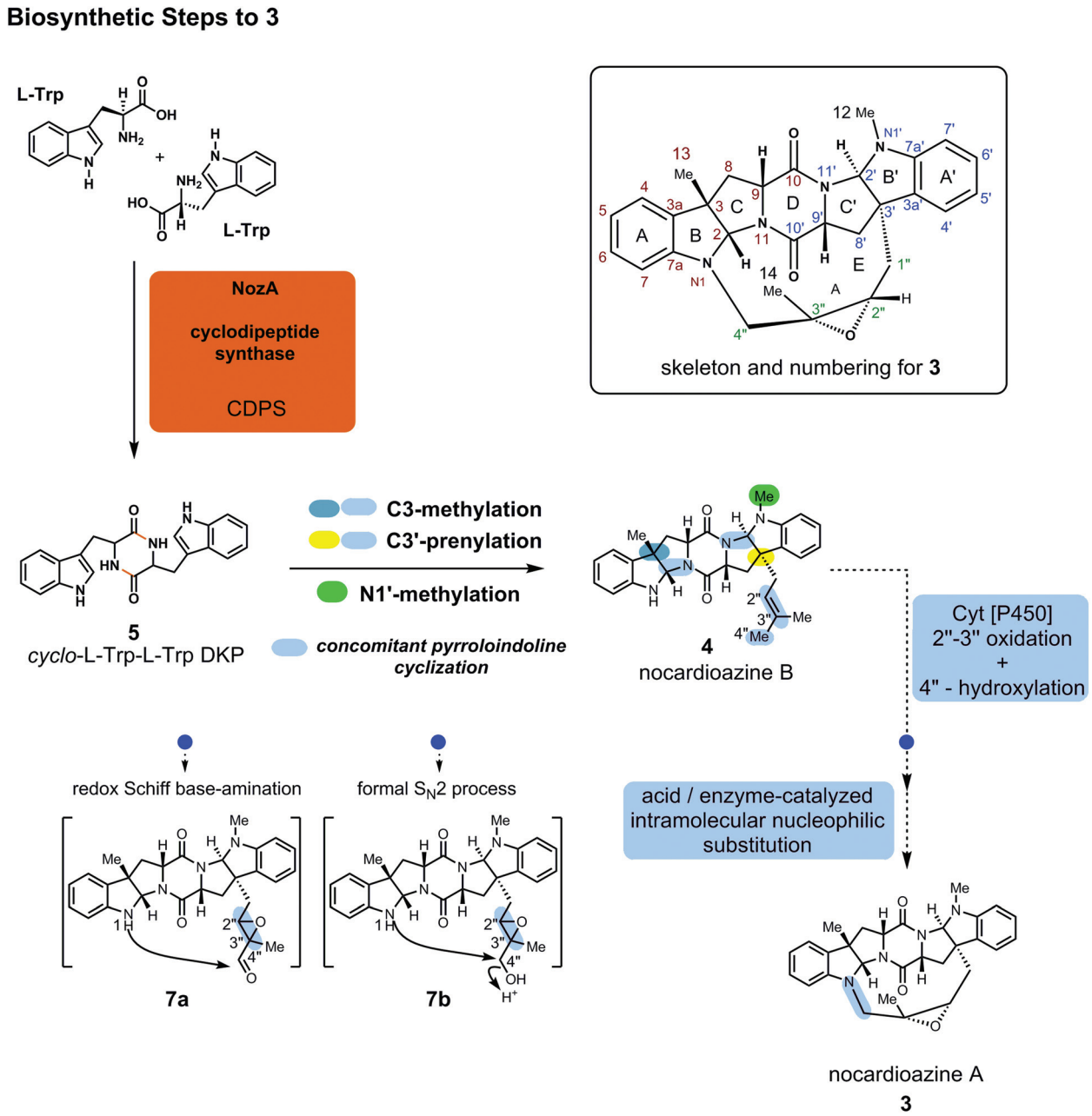
Synergistic approach establishes the pathway to nocardioazine B

Microbial systems continue to inspire the discovery of novel biocatalysts for the synthesis of organic molecules with unique



Scheme 5 Biosynthetic steps for early- and mid-stages of the *noz* pathway. Dotted lines show hypothesized possibilities and bold lines show the path that is evident from HPLC, LC-MS, and MS^n analyses for all relevant intermediates.

structural and biological properties.[23] We present evidence that points to early stage assembly of *cyclo*-Trp-Trp DKP as an intermediate that undergoes a regio- and stereoselective C3-methyltransfer step resulting in the formation of a subsequent intermediate that is C3′-normal prenylated and N1′-methylated by respective enzymes encoded in the *noz* pathway with reasonable promiscuity in the order of their occurrence. These results illuminate the specific precursor–product relationships in the nocardioazine alkaloid biosynthetic pathway and are expected to guide future genetic and enzymatic studies to further probe the *noz* pathway. The latent symmetry present in the DKP ring system of **3** and **4** enables numbering (N1–N11 and N1′–N11′) of the skeletal constituents comprising the 6–5–5–6–5–5–6 skeleton that includes rings A–B–C–D–C′–B′–A′ respectively (Scheme 6). Nature has further decorated the western half of the DKP core through a methyltransferase-catalyzed regio- and stereoselective indole C3-methylation event resulting in the pyrroloindoline B–C ring fusion. The fungal prenyltransferase enzymology offers a precedent for pyrroloindoline formation through enzymatic functionalization of the indole-C3 position accompanied by a concomitant cyclization event between N11 and C2 positions.[9] The B′–C′ rings, on the



Scheme 6 Steps in the *noz*-encoded pathway illustrated as a function of known and unknown stages. Nocardioazine numbering is illustrated in box.

eastern side of the DKP core, are functionalized through a prenyltransferase-catalyzed regio- and stereoselective indole C3′-normal prenylation event (with a 3′–1″ head-to-head connectivity) and a concomitant N11′–C2′ bond-forming cyclization generating *des-N*1′-Me-nocardioazine B (**17**). Further, indole N1′-methylation (at C12) is observed as a likely event catalyzed by an *N*-methyltransferase leading to **3** and **4**. Overall, these three pivotal biosynthetic events create the asymmetry in the two annulated pyrroloindoline moieties of **3** and **4**. The mid-stages of the pathway offered a reasonably sized cohort of synthetically tractable intermediates that could be used as standards for HPLC and LC-MS^2 analyses. These experiments facilitated the identification of the order of biosynthesis *in vivo*.

Thus far, bioinformatics analyses by previous researchers have predicted more than 50 gene clusters to encode the CDPS machinery for assembly of DKP natural products in various species spanning both prokaryotes and eukaryotes.[24] However, far fewer of these CDPSs have been biochemically characterized.[12,13,25,26] Only a single experimentally characterized CDPS, Amir_4627 from *Actinosynnema mirum*, has been established to yield *cyclo*-L-Trp-L-Trp DKP as the dominant product.[12] Homologs of NozA are evident in a range of *Nocardiopsis* strains

whose CDPS-containing gene clusters are available in publicly deposited genomes. For example, *Nocardiopsis alba* encodes an enzyme (AlbC) possessing 40% sequence identity to NozA.[25] To date, biochemically characterized CDPSs have been reported to catalyze the formation of DKPs exclusively from L-amino acids. This is due to the mechanism of CDPSs, which employ aminoacyl-charged tRNAs from primary metabolism as substrates in catalyzing the formation of the DKP scaffold. Hence, if nocardioazines A–B (**3–4**), featuring D-amino acid stereochemistry, are indeed CDPS-derived, then an unidentified isomerase is also expected as a required component of their biosynthetic pathway to isomerize *cyclo*-L-Trp-L-Trp DKP (**5**) into its antipode *ent*-**5**.

As illustrated in Scheme 5, it is evident that the methyltransferase step, likely encoded by NozB, in *Nocardiopsis* sp. CMB-M0232 is successively processing methylations of **5** and **13**. The possibility of recruitment of a promiscuous indole C3′-normal prenyltransferase that could prenylate either **13** or **16** leading to **17** or **4** is plausible. Nocardioazine A (**3**) has an additional isoprenoid-tethered DKP scaffold comprised of an 11-membered macrocycle (ring E) bridged between N1 and C3′ (tether numbered as 4″-3″-2″-1″). P-glycoprotein-mediated efflux pump (P-gp) inhibition is shown specifically by **3**, by virtue of the macrocycle E.[3] The biosynthetic pathway to nocardioazine A probably incorporates **4** as a reasonable intermediate, and employs a few additional oxidative transformations in tethering two annulated pyrroloindoline rings with a 5-carbon isoprenoid moiety. Cytochrome P450 homologs[27] NozD and NozE (Fig. 1, Table 1) represent candidates for oxidative transformation of **3** to afford **4**. According to Raju *et al.*[3] the C2″–C3″ olefinic bond is mono-oxidized into an oxirane and the C4″ position participates in an intramolecular cyclization event with the indolic nitrogen of the B-ring on **7** to close the macrocycle. Two possible intermediates (**7A** or **B**) for this late stage of the pathway are presented in Scheme 6. 2,5-Diketopiperazines of α-amino acids are valuable structural cores that have inspired natural product research.[28] Bio-inspired synthesis complementing genetic studies of such a privileged core, as detailed herein, therefore has potential to allow new synthetic pathways towards the creation of structural analogs through chemo enzymatic pathways and mutasynthesis.

Conclusion

Nocardioazines A and B (**3** and **4**), as the first indole-C3-normal prenylated DKPs from any biological source, present a poorly understood pathway. In this study, we laid the chemical foundations of nocardioazine biosynthesis by synthesizing an exhaustive set of putative, bioinformatics-predicted intermediates. Structural verification through 1D and 2D NMR and analyses through HPLC-MS-MS and HRMS methods established the framework for evaluation of the biological relevance of specific intermediates in the proposed *noz* (*n*ocardio*a*zine) pathway *in vivo*. Upon comparing HPLC and tandem mass spectrometry data between synthesized standards and alkaloidal fractions extracted from *Nocardiopsis* sp. CMB-M0232, it is conclusively evident that indole C3-methylation leading to **13** is a biosynthetic event that precedes indole C3′-normal prenylation and a second methyl transfer to the N1′ position. In addition, through bioinformatics analyses of the draft genome of *Nocardiopsis* sp. CMB-M0232, heterologous expression of Contig #1 was shown to result in the assembly of *cyclo*-L-Trp-L-Trp (**5**) as a precursor to the nocardioazine alkaloids. Future efforts are necessary to unveil the complete genetic and enzymatic-underpinning of nocardioazine A and B biosynthesis. Collectively, these results highlight the utility of synergizing bioinformatics analyses, asymmetric synthesis, and mass spectrometric metabolite profiling in guiding natural product biosynthesis studies.

Competing financial interests

The authors declare no competing financial interests.

Additional information

None.

Acknowledgements

We gratefully thank Prof. Robert J. Capon (University of Queensland, Australia) for providing the *Nocardiopsis* sp. CMB-M0232 isolate. We are grateful to Drs. Mervyn J. Bibb and Juan Pablo Gomez-Escribano for providing *S. coelicolor* M1146. RV, NA and SKP thank Case Western Reserve University's Department of Chemistry for funding portions of this project. NA thanks the Saudi Arabian Ministry of Higher Education for graduate financial aid. Saudi Arabian Cultural Mission (SACM) and Drs. Rashed M. Alsedran (Director) and Elbreir Mohammed (academic Advisor) are thanked by NA for academic support. Authors thank Dr. Dale Ray for expert NMR data collection. Authors thank Dr. Ormond Brathwaite for technical support for culturing *Nocardiopsis* sp. CMB-M0232 strain for metabolic extraction. We thank Prof. Robert Salomon and Dr. Mikhail Linetsky for expert technical support. Authors thank Angela Hansen and Taylor Harmon for support with the collection of mass spectrometry data. Authors thank William Frank of SIRC (CWRU) for instrumentation support. ALL is grateful for funding from a Research Corporation Cottrell College Science Award, a UNF Transformational Learning Opportunity Award, and a UNF Academic Affairs Scholarship grant.

References

1 J. W. Blunt, B. R. Copp, R. A. Keyzers, M. H. G. Munro and M. R. Prinsep, *Nat. Prod. Rep.*, 2013, **30**, 237–323.

2 (*a*) R. Raju, A. M. Piggott, M. Conte, Z. Tnimov, K. Alexandrov and R. J. Capon, *Chem. – Eur. J.*, 2010, **16**, 3194–3200; (*b*) R. Raju, A. M. Piggott, M. Quezada and R. J. Capon, *Tetrahedron*, 2013, **69**, 692–698.
3 R. Raju, A. M. Piggott, X.-C. Huang and R. J. Capon, *Org. Lett.*, 2011, **13**, 2770–2773.
4 D. M. Bis, Y. H. Ban, E. D. James, N. Alqahtani, R. Viswanathan and A. L. Lane, *ChemBioChem*, 2015, **16**, 990–997.
5 M. Wang, X. Feng, L. Cai, Z. Xu and T. Ye, *Chem. Commun.*, 2012, **48**, 4344–4346.
6 H. Wang and S. E. Reisman, *Angew. Chem., Int. Ed.*, 2014, **53**, 6206–6210.
7 M. Roettig, M. H. Medema, K. Blin, T. Weber, C. Rausch and O. Kohlbacher, *Nucleic Acids Res.*, 2011, **39**, W362–W367.
8 (*a*) L. Y. P. Luk, Q. Qian and M. E. Tanner, *J. Am. Chem. Soc.*, 2011, **133**, 12342–12345; (*b*) N. Mahmoodi, Q. Qian, L. Y. P. Luk and M. E. Tanner, *Pure Appl. Chem.*, 2013, **85**, 1935–1948; (*c*) N. Mahmoodi and M. E. Tanner, *ChemBioChem*, 2013, **14**, 2029–2037.
9 (*a*) S.-M. Li, *Phytochemistry*, 2009, **70**, 1746–1757; (*b*) W.-B. Yin, A. Grundmann, J. Cheng and S.-M. Li, *J. Biol. Chem.*, 2009, **284**, 100–109.
10 D. Pockrandt and S.-M. Li, *ChemBioChem*, 2013, **14**, 2023–2028.
11 H. L. Schubert, R. M. Blumenthal and X. D. Cheng, *Trends Biochem. Sci.*, 2003, **28**, 329–335.
12 T. W. Giessen, A. M. von Tesmar and M. A. Marahiel, *Biochemistry*, 2013, **52**, 4274–4283.
13 P. Belin, M. H. Le Du, A. Fielding, O. Lequin, M. Jacquet, J.-B. Charbonnier, A. Lecoq, R. Thai, M. Courcon, C. Masson, C. Dugave, R. Genet, J.-L. Pernodet and M. Gondry, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 7426–7431.
14 M. J. Smanski, J. Casper, R. M. Peterson, Z. Yu, S. R. Rajski and B. Shen, *J. Nat. Prod.*, 2012, **75**, 2158–2167.
15 J. P. Gomez-Escribano and M. J. Bibb, *Microbiol. Biotechnol.*, 2011, **4**, 207–215.
16 E. James, N. Alqahtani, R. Viswanathan and A. L. Lane, *Planta Med.*, 2014, **80**, PJ5.
17 J. Y. Yang, L. M. Sanchez, C. M. Rath, X. Liu, P. D. Boudreau, N. Bruns, E. Glukhov, A. Wodtke, R. de Felicio, A. Fenner, W. R. Wong, R. G. Linington, L. Zhang, H. M. Debonsi, W. H. Gerwick and P. C. Dorrestein, *J. Nat. Prod.*, 2013, **76**, 1686–1699.
18 D. D. Nguyen, C.-H. Wu, W. J. Moree, A. Lamsa, M. H. Medema, X. Zhao, R. G. Gavilan, M. Aparicio, L. Atencio, C. Jackson, J. Ballesteros, J. Sanchez, J. D. Watrous, V. V. Phelan, C. van de Wiel, R. D. Kersten, S. Mehnaz, R. De Mot, E. A. Shank, P. Charusanti, H. Nagarajan, B. M. Duggan, B. S. Moore, N. Bandeira, B. O. Palsson, K. Pogliano, M. Gutierrez and P. C. Dorrestein, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, E2611–E2620.
19 (*a*) L. M. Repka, J. Ni and S. E. Reisman, *J. Am. Chem. Soc.*, 2010, **132**, 14418–14420; (*b*) L. M. Repka and S. E. Reisman, *J. Org. Chem.*, 2013, **78**, 12314–12320.
20 S. Takase, Y. Itoh, I. Uchida, H. Tanaka and H. Aoki, *Tetrahedron*, 1986, **42**, 5887–5894.
21 K. Thandavamurthy, D. Sharma, S. K. Porwal, D. Ray and R. Viswanathan, *J. Org. Chem.*, 2014, **79**, 10049–10067.
22 (*a*) Y.-C. Guo, S.-X. Cao, X.-K. Zong, X.-C. Liao and Y.-F. Zhao, *Spectroscopy*, 2009, **23**, 131–139; (*b*) N. A. J. C. Furtado, R. Vessecchi, J. C. Tomaz, S. E. Galembeck, J. K. Bastos, N. P. Lopes and A. E. M. Crotti, *J. Mass Spectrom.*, 2007, **42**, 1279–1286.
23 M. L. Micallef, D. Sharma, B. M. Bunn, L. Gerwick, R. Viswanathan and M. C. Moffitt, *BMC Microbiol.*, 2014, **14**, 213.
24 (*a*) L. Aravind, R. F. de Souza and L. M. Iyer, *Biology Direct*, 2010, **5**, 48; (*b*) L. Bonnefond, T. Arai, Y. Sakaguchi, T. Suzuki, R. Ishitani and O. Nureki, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 3912–3917; (*c*) T. W. Giessen, A. M. von Tesmar and M. A. Marahiel, *Chem. Biol.*, 2013, **20**, 828–838.
25 M. W. Vetting, S. S. Hegde and J. S. Blanchard, *Nat. Chem. Biol.*, 2010, **6**, 797–799.
26 M. Gondry, L. Sauguet, P. Belin, R. Thai, R. Amouroux, C. Tellier, K. Tuphile, M. Jacquet, S. Braud, M. Courcon, C. Masson, S. Dubois, S. Lautru, A. Lecoq, S.-I. Hashimoto, R. Genet and J.-L. Pernodet, *Nat. Chem. Biol.*, 2009, **5**, 414–420.
27 P. Belin, M. Moutiez, S. Lautru, J. Seguin, J.-L. Pernodet and M. Gondry, *Nat. Prod. Rep.*, 2012, **29**, 961–979.
28 A. D. Borthwick, *Chem. Rev.*, 2012, **112**, 3641–3716.

