

ORIGINAL ARTICLE

Identification of micromonolactam, a new polyene macrocyclic lactam from two marine *Micromonospora* strains using chemical and molecular methods: clarification of the biosynthetic pathway from a glutamate starter unit

Elizabeth Julia Skellam¹, Allison Kathleen Stewart¹, Wendy Karen Strangman¹ and Jeffrey Lawson Cameron Wright¹

Through a combination of chemical and molecular analysis, a new polyene macrolactam named micromonolactam was obtained from two marine-derived *Micromonospora* species. This new polyene metabolite is a constitutional isomer of salinilactam A but contains a different polyene pattern and one *cis* double bond, in contrast to the all *trans* structure reported for salinilactam A. The molecular analysis data also established that micromonolactam is a hybrid polyketide derived from 11 polyketide units and a modified glutamate starter unit.

The Journal of Antibiotics (2013) 66, 431–441; doi:10.1038/ja.2013.34; published online 15 May 2013

Keywords: biosynthesis; genome mining; marine actinomycetes; *Micromonospora*; polyene macrolactam

INTRODUCTION

Polyene macrolactams are an underexplored group of natural products that have only been found in actinomycetes. Examples isolated from *Streptomyces* sp., include viridenomycin,¹ hitachimycin (stubomycin),² BE-14106,³ aureoverticillalactam⁴ and ML-449,⁵ all of which display antibacterial and cytotoxic activities. Vicenistatin and incednine^{6,7} display potent antitumor activity, and cyclomenol A, another streptomycete product, displays potent anti-inflammatory activity.^{8,9} More recently, sceliphrolactam, an antifungal from a wasp-associated *Streptomyces* sp., has been reported.¹⁰ Micromonosporin A is the first polyene macrolactam to be isolated from a *Micromonospora* sp.;¹¹ however, because of its instability and spontaneous degradation, the natural product itself was unable to be evaluated for biological activity. Interestingly, the fully reduced derivative exhibited antimalarial and antimycobacterial activities.¹¹ Salinilactam A was identified in the obligate marine actinomycete genera *Salinispora*, but due to overlapping olefinic resonances and spontaneous degradation, the structure was completed by analysis of the genome sequence in combination with analytical data.¹² The wide range of potent activities exhibited by polyene macrolactams described to date make this class of compounds of interest for drug discovery efforts as well as synthetic endeavors.

Macrolactams are polyketide synthase/nonribosomal peptide synthase (PKS/NRPS) hybrids possessing varying degrees of unsatu-

ration. Hitachimycin (stubomycin) was proposed to arise from a modified phenylalanine starter unit following incorporation of ¹³C-labeled phenylalanine,² and a similar process can be proposed for viridenomycin (Figure 1a). An alternative starter unit is employed in the assembly of BE-14106,³ which is constructed from an unsaturated C₉ β-aminoacyl starter unit of PKS/NRPS origin; established following gene inactivation studies, heterologous expression and incorporation studies.¹³ Again it is reasonable to propose that a similar pathway operates in the biosynthesis of aureoverticillalactam and ML-449 (Figure 1b). Vicenistatin and incednine are both biosynthesized from a modified glutamic acid starter as determined by labeling studies.^{14,15} Cyclamenol A and micromonosporin A (Figure 1c) also appear to be constructed from a modified glutamic acid starter unit, while a lysine starter unit has been suggested for salinilactam A.¹²

During routine chemical screening (LC-DAD-ESIMS) of our chemical fraction library of marine actinobacteria, a UV signature characteristic of a polyene moiety (λ_{max} 304, 342, 362 nm) was observed in chemical fractions of two microbes identified as *Micromonospora* strains CMS I1-30 and CMS I2-32. The peak of interest from both strains was attributed to a parent ion of identical MW and fragmentation, suggesting that both strains produced the same metabolite. Strain CMS I2-32 also exhibited potent biological activity in other fractions, making it a priority for genome sequencing. Here

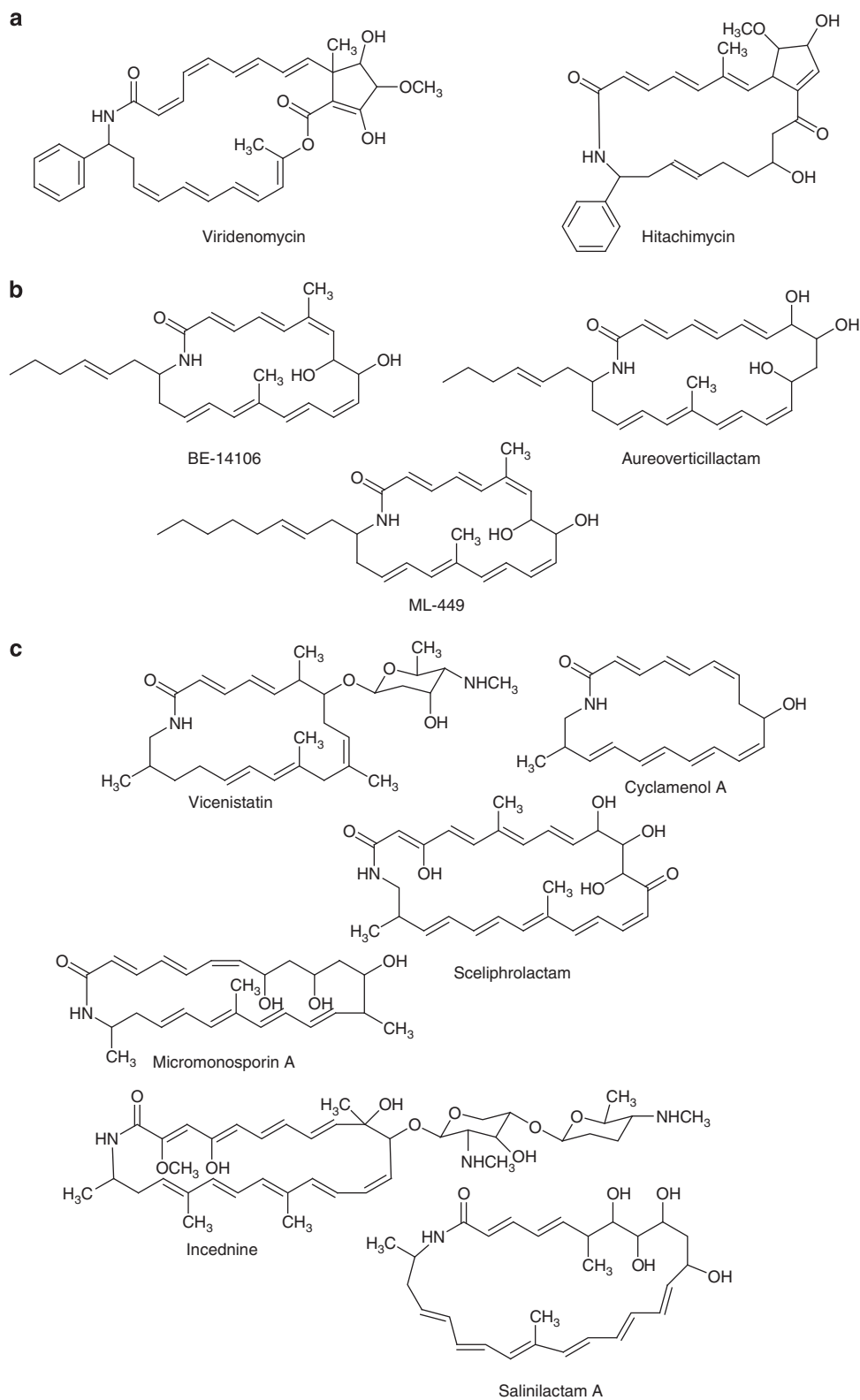


Figure 1 Structures of polyene macrolactams with (a) phenylalanine derived starter units, (b) aminoacyl starter units and (c) glutamate-derived starter units.

we report the identification of a polyene macrolactam, micromonolactam, in two *Micromonospora* strains, CMS I1-30 and CMS I2-32, as well as the identification and annotation of the biosynthetic genes for

micromonolactam in strain CMS I2-32, which we propose is biosynthesized from a modified glutamate starter unit with 11 polyketide extensions.

MATERIALS AND METHODS

General experimental procedures

UV spectra were recorded on an HP1100 Diode Array Detector (Hewlett Packard (Agilent), Santa Clara, CA, USA) and ESI mass spectra were measured on a Waters Micromass ZQ with ESI (capillary voltage 3.5 kV; cone voltages 30 V/50 V; source temperature 100 °C; desolvation temperature 400 °C; cone gas flow 60 l h⁻¹; desolvation gas flow 450 l h⁻¹). All samples were run in positive ion mode with a mass range of 100–1400 Da and a scan time of 0.5 s. Accurate MW data were obtained using a quadrupole time-of-flight mass spectrometer (Applied Bioscience-MDS Sciex QStar XL, Applied Bioscience, Indianapolis, IN, USA). ¹H, ¹³C and 2D homonuclear and heteronuclear NMR spectra were measured on an Avance 500 MHz NMR spectrometer (Bruker, Karlsruhe, Germany) using a 1.7-mm TXI probe head. Genomic DNA was sequenced using a Roche 454 Junior Sequencer (Roche).

Organisms

Strains CMS I1-30 and CMS I2-32 were isolated from ocean floor sediment samples collected off the coast of North Carolina. Sediment samples were collected in sterile Falcon tubes and treated with heat (55 °C for 1 h) and agitation (vortexed, 1 min) in order to select for actinomycetes. Treated samples were then plated (100 µl) on duplicate agar plates, using various specialized saline media conditions to further select for marine actinomycetes. Plates were incubated for 4–6 weeks, after which potential actinomycetes were continually transferred to fresh media until pure colonies were obtained. Strains CMS I1-30 and CMS I2-32 were isolated using M6 medium (mannitol 500 mg, peptone 100 mg, agar 18 g, 29.25 p.p.t. seawater 1 l).

Fermentation

For the production of micromonolactam, CMS I1-30 and CMS I2-32 were cultured in P1 medium (starch 10 g, yeast 1 g, peptone 1 g, 29.25 p.p.t. seawater 1 l).

DNA isolation

Isolation of genomic DNA from *Micromonospora* sp. CMS I2-32 was adapted from standard methods using the CTAB (cetrimonium bromide) procedure.¹⁶ A starter culture (5 ml) using M3 media (starch 10 g, yeast extract 4 g, peptone 2 g, 29.25 p.p.t. seawater 1 l) was incubated for 5 days at 30 °C. An aliquot (2.5 ml) of this culture was transferred to 50 ml of M3 media and incubated for a further 5 days at 30 °C with shaking at 200 r.p.m. The culture was centrifuged, and the collected cells were used for DNA isolation.

DNA sequencing and assembly

The genomic DNA sequence (10 × coverage, single-end reads) of *Micromonospora* strain CMS I2-32 was obtained from shotgun pyrosequencing by the Roche 454 GS Junior using titanium chemistry. The sequence reads were assembled using the Newbler software package¹⁷ into 868 contigs, with an average length of 8000 bp and total length of 6 944 130 bp.

Genome annotation

Annotation was performed using the RAST server.¹⁸ Prodigal and Artemis were also used to investigate open reading frames (ORFs) manually and in more detail.^{19,20} Amino-acid sequences of potential gene products were compared with the SWISS-PROT and PIR databases using BLAST.²¹

Nucleotide sequence accession number

The DNA sequences determined in this study are deposited in the DDBJ, EMBL and GenBank databases under the accession numbers provided in Supplementary Information.

RESULTS

Taxonomy of strains

Strain CMS I1-30 was identified as a sub species of *Micromonospora carbonacea* with a 98.2% sequence homology by phylogenetic analysis of a 500-bp segment of the 16S rDNA. Strain CMS I2-32 was

identified as a sub species of *Micromonospora peucetia* with a 97.9% sequence homology by phylogenetic analysis of a 500-bp segment of the 16S rDNA. Phylogenetic analysis of the strains was performed by Accugenix (Newark, DE, USA) using partial 16S rDNA sequence comparison. Construction of a phylogenetic tree for comparison with sequence data for recently described marine actinomycete strains (e.g., *Salinispora*) in the literature suggests that CMS I1-30 and CMS I2-32 are closely related and both occur in a unique clade that falls between the obligate marine actinomycete genera *Salinispora* and *Verrucosipora*.

Large scale fermentation, extraction and isolation of micromonolactam

Strain CMS I1-30 was grown in seed medium for 7 days at 26 °C and an aliquot (2.0 ml) was used to inoculate 200 ml of fermentation medium in 1-L shake flasks, which were cultured at 200 r.p.m. and 30 °C for 5 days. XAD-16 resin (20 g l⁻¹) was added to each flask at day 5, and the fermentation was continued for 24 h. The resin and cells were filtered using cheesecloth, washed with water to remove salts and extracted with a methanol gradient and a final elution of acetone to afford a crude extract (1.41 g), which was fractionated by reversed-phase C18 liquid chromatography eluting with increasing amounts of methanol in water. The fraction eluted with 80% methanol in water (105.9 mg) was purified by reversed-phase HPLC (Waters SunFire C18, Waters, Milford, MA, USA; 5 µm; 10 × 250 mm; flow rate of 2.5 ml min⁻¹; detection at 214 and 304 nm; 20–35% CH₃CN linear gradient over 5 min then 35–55% CH₃CN linear gradient over 20 min). Final purification of the semi-pure polyene (t_R = 17.8 min) was performed using reversed-phase HPLC (Waters SunFire C18; 5 µm; 4.6 × 150 mm; flow rate of 1.0 ml min⁻¹; detection at 214 and 304 nm; 20–40% CH₃CN linear gradient over 5 min then 40–45% CH₃CN linear gradient over 10 min to afford pure **1** (t_R = 10.8 min, 0.6 mg): light yellow solid; UV (MeOH) λ_{max} 304, 342, 362 nm; ESIMS *m/z* 452 (M-H₂O + H)⁺, 470 (M + H)⁺, 492 (M + Na)⁺, 939 (2M + H)⁺, 961 (2M + Na)⁺; HRESITOFMS *m/z* 492.2719 (calcd for C₂₈H₃₉NO₅Na⁺, 492.2726).

The culture, production and initial extraction methods for strain CMS I2-32 were identical to the protocols used for CMS I1-30. A slight modification of the final steps led to the isolation of pure **1** as follows: The fraction eluted with 80% methanol in water (129.4 mg) was purified by reversed-phase HPLC (Waters SunFire C18; 5 µm; 10 × 250 mm; flow rate of 2.0 ml min⁻¹; detection at 245 and 304 nm; 44% CH₃CN isocratic conditions for 15 min then 44–100% CH₃CN linear gradient over 5 min). Final purification of the semi-pure polyene (t_R = 9.4 min) was performed using reversed-phase HPLC (Phenomenex Gemini C18, Phenomenex, Torrance, CA, USA; 5 µm; 250 × 4.6 mm; flow rate of 1.0 ml min⁻¹; detection at 245 and 304 nm; 35% CH₃CN isocratic conditions for 20 min then 35–100% CH₃CN linear gradient over 5 min to afford pure **1** (t_R = 17.0 min, 0.4 mg): light yellow solid; UV (MeOH) λ_{max} 304, 342, 362 nm; ESIMS *m/z* 452 (M-H₂O + H)⁺, 470 (M + H)⁺, 492 (M + Na)⁺, 939 (2M + H)⁺, 961 (2M + Na)⁺; HRESITOFMS *m/z* 492.2720 (calcd for C₂₈H₃₉NO₅Na⁺, 492.2726). As compounds of this type can be quite labile, micromonolactam purification was performed in subdued light (laboratory lights were off and collection vials were covered in foil). Samples were also stored in a freezer in vials filled with N₂ gas until needed for processing or analysis.

Structural elucidation of micromonolactam

The molecular formula of **1** was found to be C₂₈H₃₉NO₅ by HRMS (10 double bond equivalents) and confirmed by the NMR data. The

Table 1 ^1H and ^{13}C NMR shift assignments of micromonolactam (**1**) in CD_3OD

Position	δ_{H}	δ_{C}	COSY	HMBC	TOCSY	NOESY
1		168.1				
2	5.71 (1H, m)	126.3	H3	C1, C4	H3, H4, H5, H6	H4
3	6.93 (1H, dd, 14.9, 10.8)	140.2	H2, H4		H2, H4, H5, H6	H5
4	5.98 (1H, m)	130.8	H3, H5		H2, H3, H5, H7, H8, H27	H2, H6
5	6.41 (1H, m)	139.6	H4, H6		H2, H3, H4, H6, H7	H3, H7
6	6.04 (1H, m)	131.7	H5, H7		H4, H5, H7, H8, H9, H27	H4, H27
7	5.80 (1H, m)	140.9	H6, H8		H4, H5, H6, H8, H9, H27	H5, H9/H10, H27
8	2.38 (1H, m)	40.9	H7, H9, H27		H6, H7, H9, H27	H7, H9/H10, H27
9	3.57 (1H, m)	76.7				
10	3.57 (1H, m)	76.7				
11	3.69 (1H, m)	69.9	H10, H12		H9, H10, H12, H13, H14, H15	H9/H10, H13
12	2.03 (1H, m)	40.6	H11, H13		H9, H10, H11, H13, H14, H15	H14
	1.83 (1H, m)	40.6	H11, H13		H9, H10, H11, H13, H14, H15	H14
13	4.35 (1H, dd, 9.7, 4.1)	72.1	H12, H14		H10, H11, H12, H14, H15	H11, H15
14	5.68 (1H, m)	136.6	H13, H15		H11, H12, H13, H15	H12, H16
15	6.36 (1H, m)	133.5			H11, H12, H13, H14	
16	6.36 (1H, m)	128.8				
17	6.36 (1H, m)	139.0		C18, C19	H19, H28	H19, H28
18		136.0				
19	6.22 (1H, m)	133.5	H20, H28	C17, C28	H20, H28	H17, H21
20	6.49 (1H, m)	129.8	H19, H21	C19	H19, H21, H22, H23, H24	
21	6.59 (1H, m)	131.3	H20, H22		H20, H22, H23, H24	H24
22	6.22 (1H, m)	133.5	H21, H23	C21, C24	H20, H21, H23, H24, H25, H26	H23
23	5.54 (1H, m)	128.8	H22, H24	C21	H21, H22, H24, H25, H26	H22
24	2.59 (1H, m)	35.0	H23, H25		H21, H22, H23, H25, H26	H21, H23, H25, H26
	2.28 (1H, m)	35.0	H23, H25		H21, H22, H23, H25, H26	H21, H23, H25, H26
25	3.80 (1H, m)	47.9	H24, H26		H22, H23, H24, H26	H23
26	1.29 (3H, d, 7.1)	20.9	H25	C24, C25	H22, H23, H24, H25	H23, H24, H25
27	1.09 (3H, d, 6.8)	17.6	H8	C5, C6, C7	H5, H6, H7, H8, H9	H6, H7, H8, H9/H10
28	1.85 (3H, s)	12.6	H19	C17, C18, C19	H17, H19	H17, H19, H20

characteristic UV absorption ($\lambda_{\text{max}} = 304, 342, 362 \text{ nm}$) was indicative of a polyene moiety and together these data suggested that this compound was a constitutional isomer of salinilactam A. The NMR data (Table 1) confirmed the presence of 28 carbons, including one carbonyl, 1 quaternary sp^2 , 21 methines (15 olefinic), 2 methylenes, and 3 methyl carbons. Three spin systems (bold lines in Figure 2a) present in **1** were determined by COSY, TOCSY and HMBC correlations (Figure 2a). A triene system, indicated by COSY and TOCSY correlations between six olefinic protons, was conjugated with a carbonyl, evidenced by the HMBC correlation from H-2 to a carbon at δ_{C} 168.1 (Table 1). In turn, H-2 was coupled to H-3, which appeared as a doublet of doublets at δ_{H} 6.93 ($J_{2,3} = 14.9, 10.8 \text{ Hz}$), and established the E geometry of this double bond. Due to the complexity of the ^1H NMR spectra in the olefinic region, coupling constants for protons H-4 through H-7 could not be ascertained, but NOESY correlations clearly illustrate the E geometry for all three double bonds in this triene system (Figure 2b). Additional 2D homonuclear correlations further extended this system from H-7 (δ_{H} 5.80) through to the H-9 oxymethine (δ_{H} 3.57), which was confirmed by HMBC correlations from the H-27 methyl doublet (δ_{H} 1.09 $J_{8,27} = 6.8 \text{ Hz}$) to carbons at δ_{C} 40.9 (C-8), 76.7 (C-9) and 140.9 (C-7). COSY and TOCSY data indicated a second spin system containing three oxymethines, a methylene and two olefinic protons. The ^1H resonance at δ_{H} 3.57 (assigned to H-9) integrated for two protons and the HSQC data indicated that these overlapping signals correlated with two identical carbon shifts at δ_{C} 76.7, making

assignment more difficult. Within this spin system, an oxymethine resonance at δ_{H} 4.35 was assigned to H-13 based on coupling with an olefinic proton assigned as H-14 and an NOE correlation with a second olefinic proton assigned as H-15. An additional NOE correlation between the δ_{H} 4.35 resonance with another oxymethine at δ_{H} 3.69 helped identify this second signal as H-11, and it was deduced from the COSY data that H-11 and H-13 were separated by the single methylene group H₂-12. An NOE correlation from H-11 was likely to H-9, rather than H-10, and this was supported by subsequent examination of the PKS domain organization in the biosynthetic cluster that indicated two chain extensions lacking dehydrogenase (DH) domains resulting in two sequential hydroxyl groups in the spin system. Thus another module also lacking a DH domain accounted for the third hydroxyl group at H-13 in this spin system. Notably, a P450 enzyme was identified within the micromonolactam biosynthetic cluster, and based on homology of the biosynthetic genes with those for salinilactam A in *Salinispora tropica*, most likely resulted in hydroxylation at C-10.¹²

COSY, TOCSY and HMBC data identified a third spin system that extended from C-17 through the C-26 methyl group. Five of the remaining olefinic protons were assigned to a triene moiety based on COSY and TOCSY data, which could be extended to include the H₂-24 methylene at δ_{H} 2.28/2.59, a methine at δ_{H} 3.80 (H-25) and a doublet methyl at δ_{H} 1.29 (H-26). The chemical shift of the H-25 methine (δ_{H} 3.80, δ_{C} 47.9) indicated proximity to a heteroatom, most likely nitrogen. NOESY correlations indicated 18E and 20E, whereas a

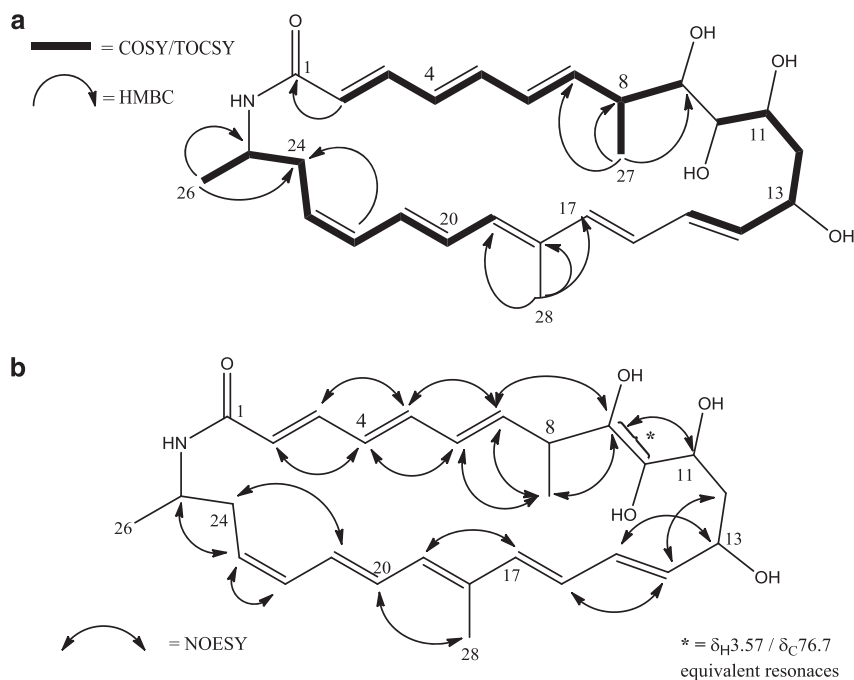


Figure 2 (a) Key COSY, TOCSY and HMBC correlations establishing the planar structure of **1**. (b) Key NOESY correlations for micromonolactam (**1**).

strong NOESY correlations between H₂-24 and H-21, as well as H-22 and H-23, established the geometry as 23Z. This spin system could be extended beyond C-19 that showed an HMBC correlation with the H₃-28 methyl protons, which appeared as a singlet (δ_H 1.85). Additional HMBC correlations from the singlet methyl resonance to a quaternary sp² carbon (δ_C 136.0, C-18), the C-19 olefinic carbon (δ_C 133.5) and an as yet unassigned olefinic carbon at δ_C 139.0 indicated that this spin system continued with further conjugation. HSQC data showed that the latter olefinic carbon (δ_C 139.0, H-17) was bound to a proton at δ_H 6.36, which overlapped with two additional olefinic methines at 6.36 p.p.m. with carbon shifts of δ_C 128.8 and 133.5, the latter of which had been already assigned to H-15 (δ_C 133.5) in the second spin system. Further analysis of the biosynthetic genes for **1** supported the presence of a pentaene moiety consistent with the number of olefinic shifts observed. Thus the remaining unassigned olefinic resonance (δ_H 6.36, δ_C 128.8) was assigned to position 16 in the molecule, thereby completing the planar structure.

UV-LCMS analysis of surrounding chromatographic fractions indicated four additional minor compounds with nominal masses of 469 and similar retention times and UV profiles to that of micromonolactam. The two most minor components were still present in complex mixtures and could not be identified further. The remaining two were present in enough quantity to be further purified and examined by ¹H NMR analysis, which indicated that they are likely isomers of micromonolactam. None of the ¹H data collected corresponded to the reported values for salinilactam. Nonetheless, the possibility remains that salinilactam is present in CMS I2-32 but is expressed in minute amounts by the organism.

Bioactivity assessment of micromonolactam

Micromonolactam (**1**) was tested for antimicrobial activity using a disc diffusion assay. Test organisms included Gram-positive (*Bacillus subtilis* and *Mycobacterium smegmatis*) and Gram-negative

(*Escherichia coli*) bacteria, yeast (*Candida kefyr*) and a filamentous fungus (*Aspergillus niger*). No activity was observed at the concentrations tested (0.5 $\mu\text{g}\mu\text{l}^{-1}$ and 1.0 $\mu\text{g}\mu\text{l}^{-1}$).

Identification of the micromonolactam gene cluster

Shotgun sequencing of *Micromonospora* sp. CMS I2-32 total DNA, resulted in 868 contigs after genome assembly with the Newbler program.¹⁷ The high %GC content and large genome size made sequencing and assembly difficult. The mean size of the contigs was 8000 bp, which separated natural product gene clusters over multiple contigs, and there was no indication of how the contigs related to the original genome sequence. In addition, sequencing errors introduced frame shifts within some of the larger genes.

Structural similarities between salinilactam A and micromonolactam suggested that the PKS genes, the amino-acid starter unit and tailoring enzymes involved in their biosynthesis should be similar. The genome of *S. tropica* has been sequenced and contains only a single modular PKS responsible for the biosynthesis of salinilactam A.¹²

Prodigal, a prokaryotic gene recognition and translation initiation site identification program, was used to identify ORFs within the contigs. The rationale for this was that modular biosynthetic enzymes are often encoded by genes that range from 4 kb up to 20 kb, the largest genes known in bacteria. By manually scanning the Prodigal output, it was simple to identify proteins over 1500 aa. Contigs 00164, 00306, 00396 and 00402 were identified as encoding such large proteins using Prodigal.¹⁹ The translated sequences were searched against BLAST and annotated using Artemis.^{21,20} The sequences ranged from 79% to 88% identity with *S. tropica* CNB-440 proteins Strop_2768 and Strop_2778 –Strop_2781, the PKS proteins involved in salinilactam A biosynthesis (Table 2).

To identify the entire micromonolactam gene cluster, the draft genome sequence was uploaded to the RAST server, an automatic annotation tool for prokaryotic genomes.¹⁸ Contigs 00287, 00410,

Table 2 Annotated contigs identified as containing genes homologous to the salinilactam gene cluster are listed in an order that highlights how the gene cluster organization observed in *Micromonospora* sp. CMS I2-32 aligns very closely with the order of the genes in *S. tropica*

Contig, (size in bp)	Gene designation	ORF start	ORF stop	Number of amino acids	Proposed function in micromonolactam biosynthesis	Closest similar protein, origin, (% identity/similarity)
00114 (15348)	mmlX	420	49	123	2-Oxoglutarate oxidoreductase	Pyruvate flavodoxin/ferredoxin oxidoreductase VAB18032_08040, <i>V. maris</i> AB-18-032, (94/94)
	mmlV	810	2318	502	Serine/threonine protein kinase	Serine/threonine protein kinase, FrCN3DRAFT_3673, <i>Frankia</i> sp. CN3, (204/507)
	mmlU	2337	3641	434	Histidine kinase	Histidine kinase Strop_2757, <i>S. tropica</i> CNB-440, (359/421)
	mmlW	3605	4267	220	Regulator	Response regulator receiver Strop_2758, <i>S. tropica</i> CNB-440, (212/220)
	mmlY	6955	4391	854	Unknown	Hypothetical protein Strop_2759, <i>S. tropica</i> CNB-440, (734/854)
	mmlS	7251	6955	98	ABC transporter	ABC transporter-like protein Strop_2760, <i>S. tropica</i> CNB-440, (90/98)
	mmlS	7733	7254	159	ABC transporter	ABC transporter-like protein Strop_2760, <i>S. tropica</i> CNB-440, (113/129)
	mmlR	10577	7911	888	LuxR regulator	Regulatory protein LuxR Strop_2761, <i>S. tropica</i> CNB-440, (746/888)
	N/A ^a	10599	10976	126	N/A	Hypothetical protein EGK_01709, <i>Macaca mulata</i> , (29/87)
	N/A ^a	11575	11832	86	N/A	Hypothetical protein ML5_5878, <i>Micromonospora</i> sp. L5, (41/85)
mmlT	12667	11891	258	Thioesterase	Thioesterase Strop_2763, <i>S. tropica</i> CNB-440, (216/258)	
mmlQ	14793	13069	125	Major Facilitator Transporter	Major facilitator transporter Strop_2765, <i>S. tropica</i> CNB-440, (472/572)	
00402 (5303)	mmlA	> 5302	3161	714	Modular PKS—partial sequence (AT, DH domains)	Beta-ketoacyl synthase Strop_2768, <i>S. tropica</i> CNB-440, (535/652)
	mmlA	3513	< 1	1171	Modular PKS—partial sequence (KR, ACP, KS domains)	Beta-ketoacyl synthase Strop_2768, <i>S. tropica</i> CNB-440, (918/1075)
00632 (2341)	mmlA	> 2339	< 1	780	Modular PKS—partial sequence (KR, ACP domains)	Beta-ketoacyl synthase Strop_2768, <i>S. tropica</i> CNB-440, (651/780)
00858 (368)	mmlA	> 367	< 1	122	Modular PKS—partial sequence (KS domain)	Beta-ketoacyl synthase Strop_2768, <i>S. tropica</i> CNB-440, (94/102)
00396 (5400)	mmlA	> 5400	< 1	1800	Modular PKS—partial sequence (AT, DH, KR, ACP domains). The AT domain contains the GHSQGE and YASH motifs, indicating a methylmalonate extender unit. ²²	Beta-ketoacyl synthase Strop_2768, <i>S. tropica</i> CNB-440, (1165/1462)
00601 (2679)	mmlA	> 2677	< 1	892	Modular PKS—partial sequence (KS, AT domains)	Beta-ketoacyl synthase Strop_2768, <i>S. tropica</i> CNB-440, (550/907)
00493 (3980)	mmlA	< 1	2262	753	Modular PKS—partial sequence (DH domain)	Beta-ketoacyl synthase Strop_2768, <i>S. tropica</i> CNB-440, (466/562)
	mmlA	2340	> 3974	756	Modular PKS—partial sequence (KR, ACP domains)	Beta-ketoacyl synthase Strop_2768, <i>S. tropica</i> CNB-440, (426/547)
00328 (6928)	N/A ^a	51	2024	657	N/A	Hypothetical protein SSAG_04916, <i>Streptomyces</i> sp. Mg1, (256/554)
mmlF	2397	2161	78	78	Ferredoxin	Hypothetical protein Strop_2769, <i>S. tropica</i> CNB-440, (67/75)
mmlG	3581	2394	395	395	Cytochrome P450 hydroxylase	Cytochrome P450 Strop_2770, <i>S. tropica</i> CNB-440, (376/395)
mmlH	3947	5293	448	448	Lysine 2,3 aminomutase	Hypothetical protein Strop_2771, <i>S. tropica</i> CNB-440, (418,447)
mmlI	5323	6579	418	418	Aminotransferase	Class I and II aminotransferase Strop_2772, <i>S. tropica</i> CNB-440, (376/418)
00641 (2 254)	mmlJ	93	641	182	Malonyl CoA-acyl carrier protein transacylase	Hypothetical protein Strop_2773, <i>S. tropica</i> CNB-440, (165/181)

Table 2 (Continued)

Contig, (size in bp)	Gene designation	ORF start	ORF stop	Number of amino acids	Proposed function in micromonolactam biosynthesis	Closest similar protein, origin, (% identity/similarity)
	mmLK	666	2210	514	Long chain fatty acid CoA ligase	AMP-dependent synthetase and ligase Strop_2774, <i>S. tropica</i> CNB-440, (467/512)
00287 (8 015)	mmLL	1664	33	543	Long chain fatty acid CoA ligase	AMP-dependent synthetase and ligase Strop_2775, <i>S. tropica</i> CNB-440, (487/531)
	mmLM	1789	2037	82	PP-binding protein	Hypothetical protein Strop_2776, <i>S. tropica</i> CNB-440, (68/82)
	mmLN	3065	2199	288	Aldo-keto reductase	Oxidoreductase MCAG_04791, <i>Micromonospora</i> sp. ATCC 39149, (232/288)
	mmLO	3201	3644	148	Transcriptional regulator	Transcriptional regulator MCAG_04790, <i>Micromonospora</i> sp. ATCC 39149, (110/143)
	mmLP	4580	3675	301	Proline iminopeptidase	Proline-specific peptidase Strop_2777, <i>S. tropica</i> CNB-440, (263/301)
	mmLE	5212	4592	206	Modular PKS—partial sequence (TE domain)	Beta-ketoacyl synthase Strop_2778, <i>S. tropica</i> CNB-440, (185/206)
	mmLE	>7985	5196	929	Modular PKS—partial sequence (DH, KR, ACP domains)	Beta-ketoacyl synthase Strop_2778, <i>S. tropica</i> CNB-440, (573/675)
00306 (7488)	mmLE	1167	>7475	2102	Modular PKS—partial sequence (KS, AT, DH, KR, ACP, KS domains)	Beta-ketoacyl synthase Strop_2778, <i>S. tropica</i> CNB-440, (1673/2101)
	mmLD	83	1108	341	Modular PKS—partial sequence (KR, ACP domains)	Acyl transferase domain-containing protein Strop_2779, <i>S. tropica</i> CNB-440, (301/341)
00701 (1694)	mmLD	>1692	967	241	Modular PKS—partial sequence (KS domain)	Acyl transferase domain-containing protein Strop_2779, <i>S. tropica</i> CNB-440, (288/311)
	mmLC	1343	<1	447	Modular PKS—partial sequence (KR, ACP domains)	Beta-ketoacyl synthase Strop_2780, <i>S. tropica</i> CNB-440, (212/241)
00164 (12271)	mmLB	>12269	8364	1301	Modular PKS—partial sequence (AT, DH, KR, ACP domains)	Beta-ketoacyl synthase Strop_2781, <i>S. tropica</i> CNB-440, (1032/1292)
	mmLC	8319	<1	2273	Modular PKS—partial sequence (KS, AT, KR, ACP, KS, AT, KR domains)	Beta-ketoacyl synthase Strop_2780, <i>S. tropica</i> CNB-440, (2221/2801)
00569 (3 058)	N/A ^a	979	>1	326	N/A	Superoxide dismutase YP_004084678.1, <i>Micromonospora</i> sp. L5, (176/244)
	mmLB	1101	>3058	652	Modular PKS—partial sequence (KS domain)	Beta-ketoacyl synthase Strop_2781, <i>S. tropica</i> CNB-440, (546/639)
00410 (5 188)	mmLE	<1	>5188	1729	Modular PKS—partial sequence (AT, DH, KR, ACP, KS domains)	Beta-ketoacyl synthase Strop_2781, <i>S. tropica</i> CNB-440, (1032/1292)

Abbreviations: ORF, open reading frame; PKS, polyketide synthase.

^aWhere a gene has been designated N/A, it is not believed to be involved in micromonolactam biosynthesis.

00493, 00569, 00601, 00632, 00652, 00701 and 00858 were identified as encoding additional salinilactam-like PKS domains (Table 2). To discover the remaining genes, the salinilactam cluster in *S. tropica* was studied, spanning from Strop_2757 to Strop_2281 (Accession numbers YP_001159577.1 to YP_001159601.1). Analysis of the salinilactam gene cluster in *S. tropica* identified a discrete Type II thioesterase (Strop_2763). The sequence of this protein was used to BLAST against CMS I2-32 in RAST, identifying contig00114 which encodes 12 genes, including a discrete type II thioesterase, as well as 7 genes which are homologous to the salinilactam gene cluster range of Strop_2757 to Strop_2765 (Table 2). The remaining 7 kb of the cluster was found in a similar manner; the aminotransferase (Strop_2772) was used to BLAST *Micromonospora* sp. CMS I2-32, identifying contig00328, and the AMP-ligase (Strop_2774) was used in a BLAST search, identifying contig00641 (Table 2).

DISCUSSION

The presence of micromonolactam, a polyene macrocyclic lactam in *Micromonospora* sp. CMS I2-32 was established using a combination of chemical and bioinformatics data. Over 80 kb of DNA homologous to the salinilactam gene cluster was observed, located over 15 contigs of the draft genome sequence of *Micromonospora* sp. CMS I2-32.

Analysis of the micromonolactam PKS domains in *Micromonospora* sp. CMS I2-32 identified five PKS genes encoding 48 domains, contained within 11 modules that were located over 12 contigs. Further analysis of the PKS fragments obtained from *Micromonospora* sp. CMS I2-32 indicated that at least 6 more domains are required for a complete PKS sequence, putting the total number of domains as 54 (Figure 3). The salinilactam PKS in *S. tropica* is reported to have 10 modules and 49 domains,¹² suggesting a different amino-acid starter unit than micromonolactam, despite the noticeable similarities.

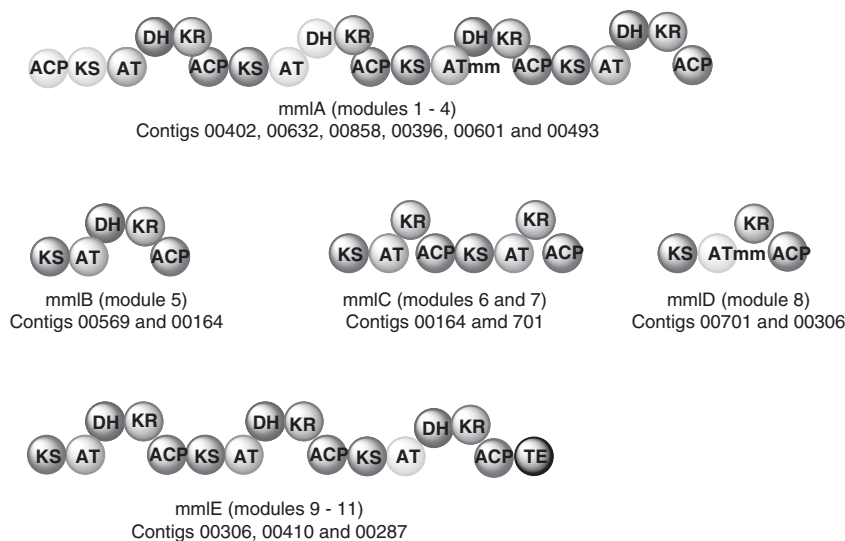


Figure 3 Proposed PKS module and domain organization for micromonolactam biosynthesis. AT domains that activate malonate are labeled AT, while those that activate methylmalonate are labeled AT_{mm}. The domains shaded in grey (The ACP and KS domains in module 1; the AT and DH domains in module 2; the AT domain in module 8; and the AT domain in module 11) are necessary for biosynthesis, based on the structure of micromonolactam, but were not observed in the genome sequence data. mmlA is proposed to consist of four modules that account for the introduction of four double bonds and a methylmalonate unit, similar to Strop_2768 in salinilactam A biosynthesis. mmlB is speculated to consist of a single module accounting for the fifth double bond in micromonolactam. mmlC is proposed to consist of at least two modules that introduce two hydroxyl groups in the sixth and seventh round of chain extension. mmlD is proposed to introduce the third hydroxyl group and second methylmalonate extender unit. mmlE is proposed to consist of at least two modules that introduce the sixth and seventh double bond. Module 11 introduces the final double bond and initiates polyketide cyclization and release. A full color version of this figure is available at *The Journal of Antibiotics* journal online.

Aware that the differences in the number of PKS domains between the two macrolactams may be due to errors introduced by sequence assembly, the partial PKS sequences of micromonolactam were compared with the original sequence of Strop_2768 (Accession number YP_001159588.1). Examination of Strop_2768 also revealed a fourth PKS module, containing KS, AT, DH, KR, ACP domains, which had not been reported previously. This indicates that the biosynthesis of salinilactam A in *S. tropica* requires an additional round of polyketide chain elongation. Previous studies had proposed that salinilactam A was derived from a 5-amino-hex-2-enoate starter unit and condensed with eight malonate and two methylmalonate units.¹² However, our analysis of the highly similar PKS sequences from the two different organisms indicated that there are nine malonate and two methylmalonate units in both molecules, suggesting an alternate C₅ starter unit, such as glutamate. Examination of other genes in the micromonolactam gene cluster identified a 2-oxoglutarate oxidoreductase (mmlX), which pointed towards a glutamate starter unit. Although a 2-oxoglutarate oxidoreductase gene is not present in the salinilactam gene cluster of *S. tropica*, a homolog is present elsewhere (Strop_0699).

Other macrolactams thought to utilize glutamate as the starter unit include vincenistatin, cyclamenol, micromonosporin and incednine (Figure 1c). Vicenistatin has been shown to be synthesized from glutamate via (2S,3S)-3-methylaspartate by labeling studies,²³ while labeling studies have shown that incednine is biosynthesized from glutamate via β-glutamic acid,¹⁵ indicating two different mechanisms for the generation of the glutamate-derived starter unit (Figure 4).

Considering that micromonolactam and salinilactam A are more structurally similar to incednine than vincenistatin, based on the position of the C-26 methyl group relative to the amide group, it is likely that micromonolactam and salinilactam A process glutamate in the same manner as observed in incednine biosynthesis. Both gene clusters contain ORFs identified as encoding a class I and II

aminotransferase (mmlI and Strop_2772) and an L-lysine 2,3-aminomutase (mmlH and Strop_2771) based on sequence homology.

If this group of related macrocyclic lactams are biosynthesized from a glutamate starter unit, then a glutamate 2,3-aminomutase enzyme is required. The first report of a glutamate 2,3-aminomutase was characterized in *Clostridium difficile*.²⁴ The protein was homologous to lysine 2,3-aminomutase in many other species, but closer inspection revealed the lack of the lysyl-binding residues Asp293 and Asp330.²⁴ Using this information, mmlH, the protein assigned as lysine 2,3-aminomutase in *Micromonospora* sp. CMS I2-32, was aligned with glutamate 2,3-aminomutase proteins identified from the NCBI protein database. The alignment includes lysine 2,3-aminomutase from *Clostridium subterminale* Sb4,²⁵ in which the lysine-binding residues have been confirmed from the X-ray crystal structure (Figure 5).

The alignment shows the conservation of the key active site residues characteristic of radical SAM enzymes (CxxxCxxC).²⁷ The lysine-binding residues, characterized in the lysine 2,3-aminomutase, isolated from *C. subterminale* Sb4 (CSUBT in Figure 4), include Arg134, which binds the α-carboxylate group, Asp293 and Asp330, which bind the ε-ammonium group.²⁵ The eight proteins identified as glutamate 2,3-aminomutase have Lys and Asn in place of Asp293 and Asp330, respectively, similar to the glutamate 2,3-aminomutase characterized in *C. difficile*, which binds glutamate, not lysine.²⁴ Analysis of mmlH has identified the radical SAM motif CxxxCxxC and Arg134, indicative of an amino-acid substrate.²⁴ However, mmlH lacks Asp293 and Asp330, indicating that it is unlikely to bind a lysine substrate.²⁴

Therefore, we speculate that mmlH, identified as lysine 2,3-aminomutase in BLAST, is actually a glutamate 2,3-aminomutase and is involved in converting glutamic acid to β-glutamic acid. Then, decarboxylation by mmlI, a PLP (pyridoxal-phosphate)-dependent aminotransferase,²⁸ would generate 3-aminobutyric acid

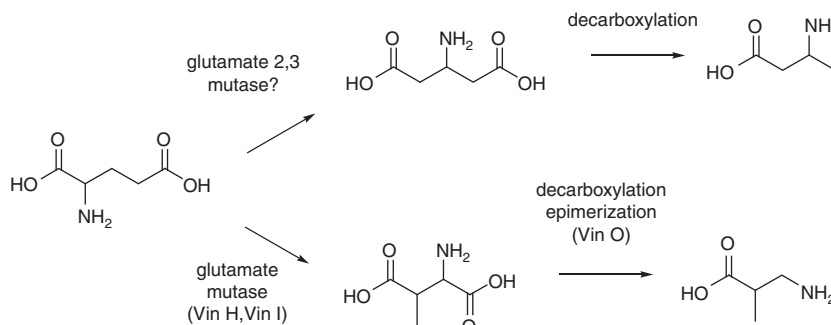


Figure 4 Two different pathways for the rearrangement of glutamate resulting in 3-aminobutyric acid ((a) salinilactam, incednine pathway) and 3-methylaspartate ((b) vincenistatin, cyclamenol pathway).

```

DMERI  ----MAVELQERSGTFRQYALKRADELKQMLPFINNPKVPTGFL--SDAYESRRGNILSYFNATLEDMENWQMAHRI TEAHV NALLPGLTMDQCEEIERVGKTYRWA VSPYYLSL
DACID  ----MAVQLYGLDNTNRQFSLRRASELKKQIVPYLEKCLTIPTGFLN--AALIQKRREHILNYFNATLQDQEDW WQLKHRVQSAND KALFPQLSABEQCAVEHIGRSFRWA I SPYYLSL
SGLYC  ----MALEDKQIYEELSNRYSILRTEELKKKIEPYLKSLETIPAGLKS--EELFQINKNNILAYFKAEPEKWN DW WQLKHRITD VET DGI VG-LSAVQKKEISKVGRVYRWA I SPYYLSL
DDEHA  ----MAVEFLPNNPQASQARALELQKQASYNKCKETIPCGLAL--SEEYNNRDFILDQLDAGLEHWQDWKWLKNNR QDAEN STLLP-LTPQKREHIDEVVKIYRWA VSPYYLSL
DREDU  MSVHLKQEEFRLRQEKRAALKRAKELKARITDYLENKDQIKNGFEV--QEYNNRAKQTLNLFNADNEQENW WQMANR KDKVVISQL D-LSPA EKEAIEKVGQRVYRWA VSPYYMAL
DDICH  ----MAVEFLPNNPREASQARALELQMAQPYFKCKETIPSGLAL--SEEYNNRDRILQGF DANLDQDWQRWQKLNRI QNPET SNLLP-LTIKQIHEISEV GKAYRWA VSPYYLSL
MTHER  ----MGREAKREIALDRAAE LKARIVDYLEEREKIASGLEA--AAEIEASKQRI LAYFGAGEAEQDWQRWQLT HRTISVATIELIP-LTEAEKEA I LKVERTYRWA VSPYYLSL
IPOLY  ----MNRREI SLERAGELKSRIEDYLSVRDSIPTGPFSEKVEQYKKSKEILMKHFGASEEQNDYKWLINRISDIET SLIVN-LTEKEEKEEKNV GATYRWA I SPYYAAL
CSUBT  -----XINRRYELFKDVSDA WNDWRWQVNRN IETVEE KKYIP-LTKEEEEGVAQCCKSLRXA ITPYYLSL
CMSI232  -----MSLLVSTTADPGRFHSFGPHHIDELGARYGLSADNLQAIRTISQVLPFRVNEYVLS LVDWDRIPDDPIFRLVFP-----QRGMLAADDER I LGDLLRA
STROP  -----MSLLVSTTSDPGRFRSFGPHHIDELGARYGLSADNVAIRTISRVLPFRVNEYVLS LLDWDRIPDDPIFRLVFP-----QRGMLAADDER I LGDLLGA
    
```

Cys 125, 129, 132 and Arg 134

```

DMERI  MSDTDPLDPIRLQGLPTPLAEDFG-EDDPMG GITSAPACI RRYDRIL IINVTNMCAMYCRICORR RNIGENDTHSPRTQ KAALDYIFANPE I RDVLVTGGDALLNDLTDWLL
DACID  MDITEPSDPIRLQGLPSLAALSDDDFG-EEDPMG ALTNPAPCI RRYDRIL IINVTNMCAMYCRICORR RNIGETDLHAKMTDKTALDYIFANPE I RDVLVTGGDALLNDHTD WLL
SGLYC  ADFSNDPLDPIRLQGLPTPLAEDFG-EDDPMG ALTSPAPCI RRYDRIL IINVTNMCAMYCRICORR RNIGEDSHKNRQDI SAALAYV ENPE I RDVLVTGGDALLSDETD WLL
DDEHA  IDKDDPRDIRLQSLPSMEILDDSG-EVDPMG EYTSAPASI RRYDRIL IINVTNMCAMYCRICORR RNIGIDLHETRAN EALDYV SNPE I RDVLVTGGDALLSDQM DWLL
DREDU  AMVSGSGGFWLQAIPCIEVKDRYG-VEDPMG EYTSAPAGI RRYDRIL IINVTNMCAMYCRICORR RNIGEDVHKS RKV EGALQYIFENKE I RDVLVTGGDALLSDRQIEWLL
DDICH  IDKDDPDIRLQSLPSIQIRDSGG-EEDPMG EFTSPAPCI RRYDRIL IINVTNMCAMYCRICORR RNIGETDLHETRAN EALDYV SNPE I RDVLVTGGDALLSDQM DWLL
MTHER  MGPEPD-CPIRRQALPSAALEDNNG-VLDPMD ELTSPAPAI RRYDRIL IINVTNMCAMYCRICORR RNIGETDLHETRAN EALDYV SNPE I RDVLVTGGDALLSDQM DWLL
IPOLY  IDPENKYSIRLLSVPTGSAAHPEG-EVDPMG EFTNPAGSI RRYDRIL IINVTNMCAMYCRICORR RNIGEDTHKSDAVIMESIDYIFNPE I RDVLVTGGDALLSDKRIEWLL
CSUBT  IDPNDFNDPVRKQAIPTALELNKAAADLEDPLHEDTDSVPVGLHRYDRVLLITIXCSXYCRICTRF RFAGQSDSDSXPERIDKAI D Y I HNTFQVRDVLSSGGDALLSDDET EYII
CMSI232  GDRKGLRGEVARI RAGLNPHSPGQQQ-HNVPSLDGQELPGMHKRYETVLYFPQGGTCHAYOTYCFRFAQFVGDADLRFAAPGPEQLVTVLHRHPAVTDVLTGGDPMMIMTER RSHV
STROP  GDRTGLRVEVARI RAGLNPHSPGQQQ-HNVPHLDGHELPGMHRYETVLYFPQGGTCHAYOTYCFRFAQFVGDADLRFAAPGPEQLVTVLHRHPAVTDVLTGGDPMMIMTER RSHV
    
```

```

DMERI  GELHAI PHVEIKRIGTRVPTLPMRIT-----DELCSILAYYPPIYINTQFNHPKEVTKDAK-KAADKLIASGVVIGNQAVLLFGINHPPEVMKKI NHELKIRVRPYYIFHAKNV
DACID  SELHSI PHVEIKRIGTRVPTLPMRIN-----DHLCSILAYYPPIYINTQFNHPQEI EDAK-KAADKLIASGVVIGNQAVLLFGINHPQEI MKKI NHELKIRIRPYYIFHAKNV
SGLYC  NELHQI PHVEIKRIGTRVPTLPMRIT-----DHLVNI LAYYPPIYINTQFNHPPIEVL EAK-QAVDR IASGVVIGNQAVLLFGINHPNIMKKI NHELKIRVRPYYIFHAKNI
DDEHA  GELHEI KHVEIKRIGTRVPTLPMRIT-----DELCAVLEIYYPPIYINTQFNHPKEV EETK-KAADRIKAGVIGNQAVLLFGINHPRIPIMKKI NHELKIRVRPYYIFHAKNV
DREDU  TELDNI PHVEIKRIGTRVPTLPMRIT-----PELCKLEIYYPPIYINTQFNHPLEVI PEAK-KACDMIKAGVIGNQAVLLFGINHPQVPMKKI NQSLTIRVRPYYIFHAKAV
DDICH  GELGHI PHVEIKRIGTRVPTLPMRIT-----ENLCKLEIYYPPIYINTQFNHPKEV EETK-KAADRIKAGVIGNQAVLLFGINHPQVPMKKI NHELKIRVRPYYIFHAKNV
MTHER  TELDNI PHVEIKRIGTRVPTLPMRIT-----PELCRVLAHPYIYINTQFNHPREVI AAK-EACDR I VQAGVIGNQAVLLFGINHPFVPMKKI NHELKIRVRPYYIFHAKPV
IPOLY  KELKSI PHVDYIRIGTRVPTLPMRIT-----DELVDMKAYQPIYINTQFNHPKEI IPEVK-EACDKLANGGISIGNQAVLLFGINHPNDKYVMRL NHEM KIRVRPYYIFHAKV
CSUBT  AKLREI PHVEIVRIGSTPVLPMRIT-----PELVNKKIYHPVWLNHFENHPNEI EEST-RAQQLADAGVIGNQAVLLFGINHPNDKVVXKELVNKLVKIRVRPYYIFHAKDLS
CMSI232  EPLL RVDVTRVFRGKAVAYWYR FVSDSDADLLRLFEQVVASGRSVAMAHFSHPRELSTE VAGRAIARIRSTGAVVYQAPLIRYVNDARAWSEMWRAEI AVGAVPYMVERDIT
STROP  EPLL RVDVTRVFRGKAVAYWYR FVSDSDADLLRLFAQVVAACRNVAVMAHFSHPRELATEIATRAIARIRSTGAVVYQAPLIRYVNDPHAWSDMWRAEI ALGAVPYMVERDIT
    
```

Asp293

```

DMERI  KGTRHFVPRIQDLKVDQLRGYTSGLAVFTY--IINAPKGGGKTPILPQYLVSLDQEHAVLRTWEGKIVHYETP-----
DACID  KGTRHFVPSI KEGLEV KHLRGYTSGLAVFTY--IINAPNGGGKTPILPQYLVSLKENQALKTWEDKIVQYETP-----
SGLYC  KGTRHFIPSI QEGSLAVMEHLRGYTSGLAVFTY--IINAPKGGGKIPILLPQYLLSLENKAVFRSEGWNVVHYETP-----
DDEHA  KGTSHFIPRI QDGLKINENLRGYTSGLAVFTY--IINAPGGGGKTPILPQYLISLNDDEAVIRTWEGKVVHYPNH-----
DREDU  KGTRHFITGVDGDAINEQLRGYTSGLAVFTY--IINAPNGYGTTPVLPQYI IENKNDQVTLRTWEEKRIIPYNSIGKH-----
DDICH  KGTSHFIPRI QDGLKINDHLRGYTSGLAVFTY--IINAPGGGGKTPILPQYLVSLNEHEALRTWEGKIVHYPNH-----
MTHER  KGTRHFITSIEEGVEINDKLRGYTSGLAVFTY--IINAPHGLGKTPILPQYVIARNHQVILRTWEEKRIIFYSNLGRQKEQA-----
IPOLY  KGTLHFNTSVDDSEIEMEYLRGYTSGLAVFTY--IINAPKGGKTPIMPQYLLSRKSNKSVKIRTWEGVVIDYENHPTIDIKEYL-----
CSUBT  LGLEHFRTPVSKIEIIEGLRGHTSGYCVFTF--VVDHGGGGKTPVXNYVISQSHDKVILRNFEQVITTYSEPIINTPGCNCVDTGKKKVHKVGVAGLLNNGEGXALEPVGLERNKRH
CMSI232  GPRDYFQVPLTRAADIFRTAYQDLPGIARTVRGVPVMSATPGKVLVDGVERTPDGDFQLRLIQARDRI VGRFRFRARWSADA AWLSLELDPTAPADVLA AVS PAARPGA AVTT-----
STROP  GPRDYFQVPLTRAADIFRTAYQDLPGIARTVRGVPVMSATPGKVLVDGVERTPHGFEFQLRLIQARDRI VGRFRFRARWSADA AWLSLDQDPAADILA AVS PAARPGA AVTA-----
    
```

Asp330

Figure 5 Amino-acid sequences for glutamate 2,3 aminomutases aligned using ClustalW.²⁶ DMERI—*Desulfosporosinus meridiei* DSM 13257 (AFQ42647.1), DACID—*Desulfosporosinus acidiphilus* SJ4 (AFM39794.1), SGLYC—*Syntrophobotulus glycolicus* DSM 8271 (ADY56980.1), DDEHA—*Desulfitobacterium dehalogenans* ATCC 51507 (YP_006429495.1), DREDU—*Desulfatocamulaculum reducens* MI-1 (ABO50665.1), DDICH—*Desulfitobacterium dichloroeliminans* LMG P-21439 (AGA68574.1), MTHER—*Moorella thermoacetica* ATCC 39073 (ABC18788.1), IPOLY—*Ilyobacter polytrophicus* DSM 2926 (YP_003966579.1), CSUBT—lysine 2,3-aminomutase from *C. subterminale* Sb4 (2A5H_A), CMSI232 (mmIH)—*Micromonospora* sp. CMSI232, STROP (Strop_2771)—*Salinispora tropica* (ABP55213.1). A full color version of this figure is available at *The Journal of Antibiotics* journal online.

(Figure 4a). This is consistent with the labeling studies of incednine by Eguchi and colleagues who speculated that salinilactam could be synthesized from glutamate and noted that an additional module for polyketide chain extension would be required.¹⁵ This module has now

been identified in both the salinilactam A and micromonolactam gene clusters.

The proposed biosynthesis of micromonolactam is shown below (Figure 6). Glutamic acid is mutated to β-glutamate by mmlH.

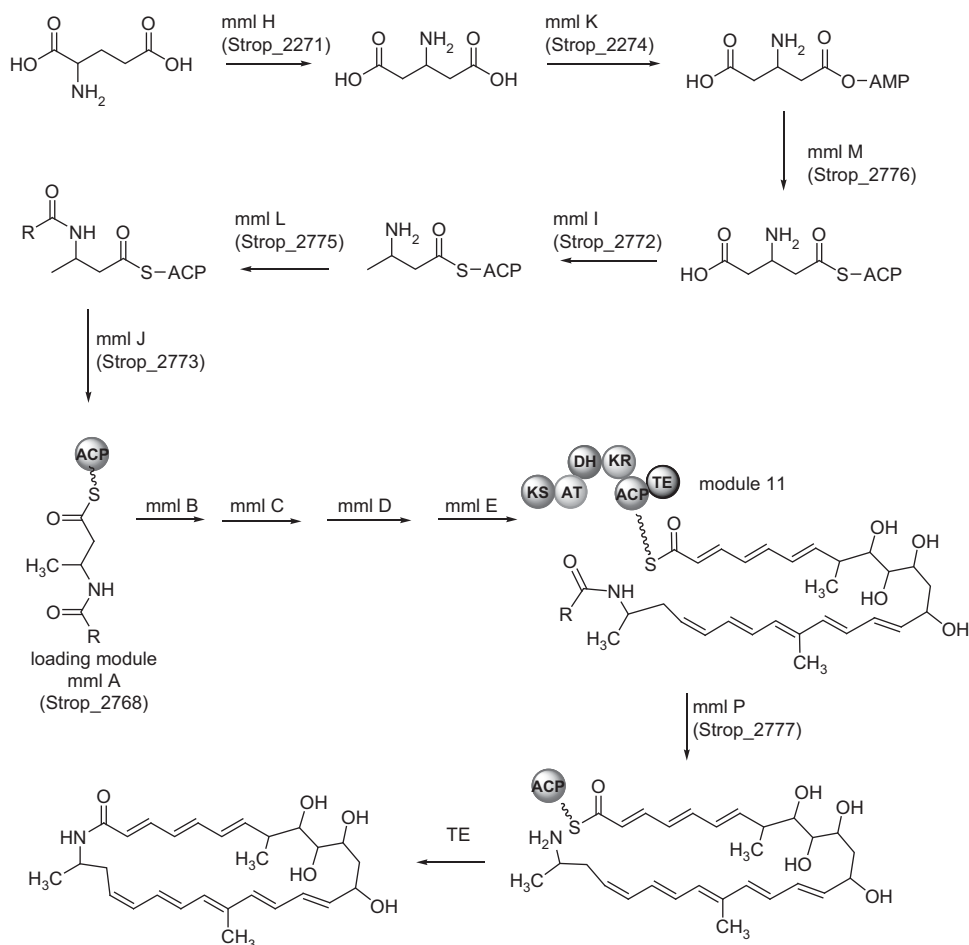


Figure 6 Proposed biosynthesis of micromonolactam in *Micromonospora* sp. CMS I2-32. A full color version of this figure is available at *The Journal of Antibiotics* journal online.

β -Glutamate is activated by mmlK, an AMP-dependent synthetase and ligase, and transferred to mmlM, an acyl-carrier protein. MmlI decarboxylates this intermediate to generate the 3-aminobutyric acid thioester. Eguchi and colleagues have discovered that the aminoacyl-ACP intermediate is protected by the addition of an acyl chain transferred by a second ATP-dependent ligase.¹⁴ The acyltransferase, mmlJ, transfers the primed starter unit to the PKS system, mmlA. Polyketide biosynthesis then proceeds with four chain elongation cycles, the third of which uses methylmalonate instead of malonate, and the intermediate is transferred to a second PKS system, beginning with mmlB (Figure 6).

In conclusion, we have isolated and identified a new macrocyclic lactam, named micromonolactam, obtained from two different *Micromonospora* species isolated from marine sediments. The genome sequencing data for one strain revealed the presence of a number of biosynthetic gene clusters, including one cluster for a polyunsaturated hybrid polyketide identified as micromonolactam. Together, the chemical, analytical and genetic data were used to establish the structure of the compound. The genetic data also revealed that the biosynthetic cluster contained 11 polyketide modules, indicating that the compound is biosynthesized from 9 malonate and 2 methylmalonate units. In addition, a series of genes associated with processing a glutamate starter unit were also observed, which is in line with a proposal for the related hybrid lactam incednine.¹⁵ Retro analysis of

the genetic data for salinilactam A from *S. tropica* has uncovered an additional PKS module, establishing a common biosynthetic pathway for these metabolites.

ACKNOWLEDGEMENTS

EJS is a Business of Biotech Post Doctoral Fellow supported by the State of North Carolina MARBIONC program. AKS, WS and JLCW thank MARBIONC for additional funding. Funding for the Roche 454 Sequencer came from the North Carolina Biotech Center (Grant 2011-IDG-1006) and we wish to acknowledge funding (Grant NSF-CHE-1039784) for the purchase of the HR TOF mass spectrometer.

- 1 Nakagawa, M, Toda, Y, Furihata, K, Hayakawa, Y & Seto, H. Studies on viridenomycin, a novel 24-membered macrocyclic polyene lactam antibiotic. *J. Antibiot.* **45**, 1133–1137 (1992).
- 2 Omura, S, Nakagawa, A, Shibata, K & Sano, H. The structure of hitachimycin, a novel macrocyclic lactam involving β -phenylalanine. *Tetrahedron Lett.* **23**, 4713–4716 (1982).
- 3 Kojiri, K, Nakajima, S, Suzuki, H, Kondo, H & Suda, H. A new macrocyclic lactam antibiotic, BE-14106 I. taxonomy, isolation, biological activity, and structure elucidation. *J. Antibiot.* **45**, 868–874 (1992).
- 4 Mitchell, S, Nicholson, B, Teisan, S, Lam, K & Potts, B. Aureoverticillactam, a novel 22-atom macrocyclic lactam from the marine actinomycete *Streptomyces aureoverticillatus*. *J. Nat. Prod.* **67**, 1400–1402 (2004).
- 5 Jorgensen, H *et al.* Insights into the evolution of macrolactam biosynthesis through cloning and comparative analysis of the biosynthetic gene cluster for a novel macrocyclic lactam, ML-449. *Appl. Environ. Microbiol.* **76**, 283–293 (2010).

- 6 Shindo, K, Kamishohara, M, Odagawa, A, Matsuoka, M & Kawai, H. Vicenistatin, a novel 20-membered macrocyclic lactam antitumor antibiotic. *J. Antibiot.* **46**, 1076–1081 (1993).
- 7 Futamura, Y *et al.* Discovery of incednine as a potent modulator of the anti-apoptotic function of Bcl-xL from microbial origin. *J. Am. Chem. Soc.* **130**, 1822–1823 (2008).
- 8 Mueller, H *et al.* Fermentative manufacture of cyclamenol as an inflammation inhibitor. German Patent 4231289A1 (1994). US Patent 5,565,561 (1996).
- 9 Nazare, M & Waldmann, H. Enantiospecific synthesis of the (9S, 18R)-diastereomer of the leukocyte adhesion inhibitor cyclamenol A. *Chem. Europ. J.* **7**, 3363–3376 (2001).
- 10 Oh, D, Poulsen, M, Currie, C & Clardy, J. Sceliphrolactam, a polyene macrocyclic lactam from a wasp-associated *Streptomyces* sp. *Org. Lett.* **13**, 752–755 (2011).
- 11 Thawai, C *et al.* Micromonosporin A, a novel 24-membered polyene lactam macrolide from *Micromonospora* sp. Isolated from peat swamp forest. *Chem. Biodiversity* **1**, 640–645 (2001).
- 12 Udway, D *et al.* Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proc. Natl Acad. Sci.* **104**, 10376–10381 (2007).
- 13 Jorgenson, H *et al.* Biosynthesis of macrolactam BE-14106 involves two distinct PKS systems and amino acid processing for generation of the aminoacyl starter unit. *Chem. Biol.* **16**, 1109–1121 (2009).
- 14 Shinohara, Y, Kudo, F & Eguchi, T. A natural protecting group strategy to carry an amino acid starter unit in the biosynthesis of macrolactam polyketide antibiotics. *J. Am. Chem. Soc.* **133**, 18134–18137 (2011).
- 15 Takaishi, M, Kudo, F & Eguchi, T. A unique pathway for the 3-aminobutyrate starter unit from L-glutamate through β -glutamate during biosynthesis of the 24-membered macrolactam antibiotic, incednine. *Org. Lett.* **14**, 4591–4593 (2012).
- 16 Keiser, T, Bibb, MJ, Buttner, MJ, Chater, KF & Hopwood, DA. Practical Streptomyces Genetics. *The John Innes Foundation* (Norwich, UK, 2000).
- 17 Miller, JR, Koren, S & Sutton, G. Assembly algorithms for next generation sequencing data. *Genomics* **95**, 315–327 (2010).
- 18 Aziz, RK *et al.* The RAST server: rapid annotations using subsystems technology. *BMC Genomics* **9**, 75 (2008).
- 19 Hyatt, D *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).
- 20 Berriman, M & Rutherford, K. Viewing and annotating sequence data with Artemis. *Brief Bioinform* **4**, 124–132 (2003).
- 21 Altschul, SF *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nuc. Acids Res.* **25**, 3389–3402 (1997).
- 22 Reeves, CD *et al.* Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. *Biochemistry* **40**, 15464–15470 (2001).
- 23 Nishida, H, Eguchi, T & Kakinuma, K. Amino acid starter unit in the biosynthesis of macrolactam polyketide antibiotic vicenistatin. *Tetrahedron* **57**, 8237–8242 (2001).
- 24 Ruzika, FJ & Frey, PA. Glutamate 2,3-aminomutase: a new member of the radical SAM superfamily of enzymes. *Biochim. Biophys. Act* **1774**, 286–296 (2007).
- 25 Lepore, BW, Ruzika, FJ, Frey, PA & Ringe, D. The x-ray crystal structure of lysine 2,3-aminomutase from *Clostridium subterminale*. *Proc. Natl Acad. Sci.* **102**, 13819–13824 (2005).
- 26 Thompson, J, Higgins, D & Gibson, T. CLUSTAL W: improving sensitivity of progressive multiple sequencing alignment through sequence weighting, position-specific gap penalties, and weight matrix choice. *Nuc. Acids Res.* **22**, 4673–4680 (1994).
- 27 Sofia, HJ, Chen, G, Hetzler, BG, Reyes-Spindola, JF & Miller, NE. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nuc. Acids Res.* **29**, 1097–1106 (2001).
- 28 Lima, S *et al.* The crystal structure of the *Pseudomonas dacunhae* aspartate-beta-decarboxylase dodecamer reveals an unknown oligomeric assembly for a pyridoxal-5'-phosphate-dependent enzyme. *J. Mol. Biol.* **388**, 98–108 (2009).

Supplementary Information accompanies the paper on The Journal of Antibiotics website (<http://www.nature.com/ja>)