Electronic Supplementary Material (ESI) for Organic Chemistry Frontiers. This journal is © the Partner Organisations 2023

Supplementary Information

Revision of the full stereochemistry of telomycin

Sandra Resa, Marta González, Fernando Reyes, Ignacio Pérez-Victoria*

Fundación MEDINA, Centro de Excelencia en Investigación de Medicamentos Innovadores en Andalucía, Parque Tecnológico de Ciencias de la Salud, Avda. del Conocimiento 34, 18016 Armilla, Granada, Spain.

*Corresponding author: ignacio.perez-victoria@medinaandalucia.es

Supplementary Information Contents

Experimental procedures:

a) General methods

b) Phylogeny-based bioinformatic analysis of the condensation-like domains of telomycin NRPS

c) Telomycin hydrolysis

d) NMR analysis of telomycin hydrolyzate and comparison with standards

e) Advanced Marfey's method applied to telomycin hydrolyzate

f) NMR analyses of intact telomycin

Table S1. NMR data of telomycin.

Figure S1. Screenshots of the NaPDos web server employed for the phylogeny-based bioinformatic analysis of the condensation like domains in telomycin NRPSs.

Figure S2. Screenshots of the NaPDos web server showing the classification of the different condensation- like domains identified in telomycin NRPSs.

Figures S3-S7. Overlay of the multiplicity-edited HSQC spectra (selected expanded regions) of telomycin hydrolyzate and various amino acid standards.

Figure S8. Extracted-ion chromatograms corresponding to the advanced Marfey's analysis applied to telomycin acidic hydrolyzate.

Figure S9. Zoomed region of the ROESY spectrum of telomycin highlighting the key correlations determining the stereochemistry in the $\Delta_{2,3}$ -Trp residue of telomycin.

Figure S10. Diagnostic ¹³C chemical shifts determining the sequence position of L-*cis*-3-Hyp and L-*trans*-3-Hyp epimers in telomycin.

Figure S11. *J*-based configurational analysis (JBCA) approach for determining the sequence position of L-Thr and D-*allo*-Thr residues in telomycin.

Figure S12. *J*-based configurational analysis (JBCA) approach for determining the relative configuration of the D-*erythro*- β -MeTrp and L-*erythro*- β -Hyl residues in telomycin.

Figures S13-S21. NMR spectra of telomycin

References

Experimental procedures

a) General methods.

Telomycin was purchased from Bioaustralis fine chemicals. Marfey's reagent L-FDVA (*N*-(2,4-dinitro-5-fluorophenyl)-L-valinamide) and its enantiomer D-FDVA were purchased from SIGMA. Amino acid standards L-Asp, D-Asp, L-Ser, D- Ser, L-Thr, D-Thr, racemic DLallo-Thr, L-Ala, D-Ala, L-*cis*-3-Hyp and L-*trans*-3-Hyp were from SIGMA. Amino acid standards L-*erythro*- β -MeTrp and L-*threo*- β -Hyl, prepared by enzymatic synthesis using an engineered variant of the tryptophan synthase PfTrpB¹ and the L-threonine transaldolase ObiH ², respectively, were kindly donated by Dr. Andrew Büller (University of Wisconsin-Madison). HCl and thioglycolic acid were from SIGMA. HPLC grade solvents were from Merck. LC-UV-MS analyses were performed on an Agilent 1100 MSD single quadrupole LC-MS system. NMR spectra were recorded on a Bruker Avance III spectrometer (500 and 125 MHz for ¹H and ¹³C NMR, respectively) equipped with a 1.7mm MicroCryoprobe. Telomycin spectra were acquired at 40 °C in DMSO-d₆ using the residual solvent signal as internal reference (δ_H 2.51 and δ_C 39.0 ppm). Amino acid standards and telomycin hydrolyzate spectra were acquired at 24 °C in D₂O using thioglycolic acid as external reference.

b) Phylogeny-based bioinformatic analysis of the condensation-like domains of telomycin NRPS.

Identification of the enantioselectivity (^LC_L vs. ^DC_L) of the canonical condensation domains in the different modules of telomycin NRPSs was carried out with the bioinformatic tool NaPDos (Natural Product Domain Seeker)^{3, 4}. The gene sequence reported for the telomycin gene cluster, with accession number KT881498,⁵ was downloaded in FASTA format and employed directly as query in the web interface of NaPDos. The results displayed by the web interface (Figs. S1-S2) were copied and edited for preparing Table 1 in the body of the article.

c) Telomycin hydrolysis

0.5 mg of commercial telomycin were hydrolyzed with 0.6 mL of HCl 6N, in the presence of 5% thioglycolic acid, for 16 h at 110 °C in a sealed reaction vial. Thioglycolic acid was added to mitigate the degradation of the β -MeTrp residue during the harsh acidic hydrolysis.⁶ The hydrolysis reaction was evaporated to dryness under a stream of nitrogen for 24 h, not only to eliminate the solvent but also to considerably reduce the thioglycolic acid excess. It is important to remark that the conditions employed for telomycin acid hydrolysis of peptides,⁷ including those bearing both Thr and *allo*-Thr residues.⁸ The results observed in the NMR and advanced Marfey's analyses of

telomycin hydrolyzate confirmed negligible amino acid racemization during the acid hydrolysis of the antibiotic.

d) NMR analysis of telomycin hydrolyzate and comparison with standards

The dried residue from telomycin hydrolysis was redissolved in 50 μ L of D₂O and transferred to a 1.7 mm NMR tube for spectral acquisition. A set of 1D ¹H NMR and 2D NMR spectra including COSY, TOCSY, HSQC and HMBC spectra were acquired to ensure the correct assignment of the different amino acid signals in the HSQC spectrum. Since excess of thioglycolic acid (in reduced and disulfide form) was very prominent in the NMR spectra of telomycin hydrolyzate, the ¹H and HSQC NMR of the different amino acid standards were acquired (also in D₂O) using samples prepared from their corresponding hydrochloride salts which contained a similar excess of thioglycolic acid for a faithful comparison.

e) Advanced Marfey's method applied to telomycin hydrolyzate

The different L- amino acid standards were derivatized with L-FDVA with the following protocol: A 1% (w/v) solution (50 μ L) of L-FDVA in acetone was added to an aliquot (25 μ L) of a 50 mM solution of each amino acid standard. After addition of 10 μ L of 1 M NaHCO₃ solution, each mixture was incubated for 60 min at 40 °C. The reactions were quenched by addition of 5 μ L of 1 N HCl, and the crude mixtures were diluted with 700 µL of acetonitrile and 50 µL of water before analysis by LC-MS. These L- amino acid standards were derivatized in the same manner with D-FDVA (as an equivalent of the derivatization of the corresponding D- amino acid enantiomers with L-FDVA). The telomycin hydrolyzate recovered from the previous NMR analysis was evaporated to dryness under a nitrogen stream and then redissolved in 100 μ L of water. Such telomycin hydolyzate solution was split in two parts, one for derivatization with L-FDVA and the other for derivatization with D-FDVA, as required for the advanced Marfey's method, using the following protocol: A 1% (w/v) solution (50 μ L) of L-FDVA (or D-FDVA) in acetone was added to 50 μ L of telomycin hydrolyzate solution. After addition of 10 μ L of 1 M NaHCO₃ solution, each mixture was incubated for 60 min at 40 °C. The reactions were quenched by addition of 5 µL of 1 N HCl, and the crude mixtures were diluted with 350 μ L of acetonitrile and 50 μ L of water before analysis by LC-MS. Separations were carried out on a Waters XBridge C18 column (4.6 \times 150 mm, 5 μ m), maintained at 40 °C. A mixture of two solvents, A (10% CH₃CN, 90% H₂O) and B (90% CH₃CN, 10% H₂O) , both containing 1.3 mM TFA and 1.3 mM ammonium formate, was used as the mobile phase under a linear gradient elution mode (isocratic 20% B for 1 min, 20-60% B in 19 min, isocratic 60% B for 1 min, 60-100% B in 0.5 min, isocratic 100% B for 4 min, 100-20% in 0.5 min, and then isocratic 20% B for 4 min) at a flow rate of 0.8 mL/min. Retention times for the derivatized amino acid standards and those observed for the derivatized amino acids from the telomycin hydrolyzate were identified by selective ion extraction in the LC-MS chromatographic traces and are summarized in Table 2 in the body of the article. The corresponding chromatograms are included in the Supporting Information (Fig. S9).

f) NMR analyses of intact telomycin

Telomycin (ca. 0.5 mg) was dissolved in DMSO-d₆ (50 μ L) to create a solution ca. 7 mM which was transferred to a 1.7 mm NMR tube. A full set of spectra including 1D ¹H and ¹³C NMR and 2D NMR spectra (*J*-Resolved, COSY, TOCY, ROESY, HSQC and HMBC) were acquired. The key COSY, TOCSY and HMBC correlations determining the connectivity of telomycin could be corroborated.⁸ The full NMR data and assignments (Table S1) showed an excellent agreement with those reported previously.⁹ TOCSY spectrum was acquired with a spin-lock time of 90 ms while the ROESY spectrum was acquired with a mixing time of 400 ms. Two HMBC spectra were acquired, optimized for a long-range coupling of 8 Hz and 2 Hz, respectively. From such spectra it was possible to qualitatively classify the key heteronuclear coupling constants as large, medium or small based on the intensity of the corresponding cross-peaks,^{10, 11} as we have recently reported for the caniferolides macrolides,¹² to apply the JBCA approach.¹³

Position	δc , type	δ _н , mult. (<i>J</i> in Hz)	Position	δc , type	δ _H , mult. (<i>J</i> in Hz)
L-Asp			Z-∆₂,₃-Trp		
1	170.3, C		1	163.3, C	
2	50.9 <i>,</i> CH	3.94, m (ov.)	2	122.3, C	
3	37.6, CH ₂	2.73, dd (12.0, 12.0)	3	122.4 <i>,</i> CH	7.47, br s
		2.33, dd (12.3, 3.0)			
4	171.2, C		1'-NH		11.7, br s
NH ₂		n. d.	2'	127.6, CH	7.93, d (2.3)
			3'	117.0, C	
L-Ser	170.0.0		3a	125.7, C	7(1 d (9 0))
	170.9, C	4.10 m	4 F'	117.0, CH	7.01, 0 (8.0)
2	57.5, CH	4.10, 11 3.75 m (ov.)	5 6'	120.0, CH	7.00, III (0V.)
5	00.5, CH2	3.66 m (ov.)	0	121.5, CH	7.15, 00 (8.0, 7.5)
NH		8.60 d (4.3)	7'	111 7 СН	7 38 d (8 0)
		0.00, 0 (4.3)	, 7a'	135.3. C	7.50, 0 (0.0)
L-Thr			NH	, _	10.02. s
1	168.7, C				, -
2	57.4, CH	4.47, m (ov.)	D- <i>erythro</i> -β-MeTrp		
3	70.5, CH	4.98, dq (6.2, 3.0)	1	171.5, C	
4	15.4, CH₃	1.15, d (6.2)	2	59.5, CH	4.49, m (ov.)
NH		8.68, d (9.2)	3	32.6 <i>,</i> CH	3.65 <i>,</i> m (ov.)
			3'	117.0, C	
D- <i>allo</i> -Thr			1'-NH		10.7, br s
1	168.9, C		2'	122.6 <i>,</i> CH	7.07, m (ov.)
2	58.1 <i>,</i> CH	4.08, dd (8.5, 8.5)	3a'	125.7, C	
3	66.6 <i>,</i> CH	3.84, m (ov.)	4'	118.9, CH	7.56, d (8.0)
4	20.8, CH ₃	1.07, d (6.2)	5'	118.1 <i>,</i> CH	6.93, dd (7.8, 7.7)
NH		7.38, d (7.9)	6'	120.6, CH	7.01, dd (7.8, 7.7)
			7	111.4, CH	7.27, d (8.0)
L-Ala	474 7 6		/a′	136.5, C	
1	1/1./, C		3- IVIE	19.2, CH ₃	1.24, d (7.0)
2	47.9, CH	4.45, m (ov.)	INH		7.74, d (7.9)
5 NH	17.2, CH3	7.64 d (7.5)	L oruthro R Hul		
INFI		7.04, u (7.3)	1 L- <i>егуцпго-</i> р-пу	170 G C	
Gly			2	51 8 CH	15 dd (86 67)
1 Cly	168.8 C		3	74.6 CH	3.32 m (ov)
2	40.9. CH ₂	4.43. m (ov.)	4	28.6. CH	1.95. m (ov.)
_		3.80. m (ov.)		_0.0, 0.1	,
NH		8.77, dd (4.5, 4.5)	5	15.9, CH₃	0.79, d (6.6)
			6	20.0, CH₃	0.85, d (6.6)
L-3-trans-Hyp			NH		6.62, d (2.9)
1	171.3, C				
2	66.8 <i>,</i> CH	4.19, d (5.8)	∟-3- <i>cis</i> -Hyp		
3	73.0, CH	4.22, m (ov.)	1	169.0, C	
4	33.3, CH ₂	2.18, app. sextet (5.6)	2	61.6 <i>,</i> CH	4.69, d (7.3)
		1.98, m (ov.)			
5	44.3, CH ₂	3.80, m (ov.)	3	69.0, CH	4.75, app. dd (14.5, 7.5)
		3.54, app. dd (15.8, 7.5)			
			4	31.5, CH ₂	2.03, m (ov.)
					1.76, dddd (14.0, 11.0,
					8.5, 8.5)
			5	44.4, CH ₂	5.94, III (OV.)
1	1				1.3.70.111007.1

Table S1. NMR data for telomycin (DMSO-d₆, 40 °C, 500 MHz).





Description Description tomo Overview Tutorial Run Analysis Pathways Contact US CMSCNRPS Domain Search Image: Search Search Search Image: Search Sear	NaPDOS Natural Product Domain Seeker Home Overview Tutorial Run Analysis Pathways Contact Us
Natural Product Domain Seeker Yom Overview Tutorial Run Analysis Pathways Contact Us YKS/NRPS Domain Search Image: Standard Settings can be used to customize program parameters. Standard Advanced Settings can be used to customize program parameters. Ionain type KS domains C domains C domains C domains C domains Predicted context program parameters. C domains Res domains C domains C domains C domains Res domains C domains Predicted context program parameters. Contact (DNA) Res domains C domains C domains C domains Aury sequence Statistical domains (DNA) Res domains on the spectration of the size of the spectration of the size analysis experiments of the size of the size analysis of the size a	Natural Product Domain Seeker Home Overview Tutorial Run Analysis Pathways Contact Us
tome Overview Totorial Run Analysis Pathways Contact Use PKCS/NRPS Domain Search <th>Home Overview Tutorial Run Analysis Pathways Contact U</th>	Home Overview Tutorial Run Analysis Pathways Contact U
PCKS/NRPS Domain Search mose a domain and query type, then enter you data to identify candidate KS andor C domains. advanced Setting's can be used to customize program parameters. Image: Comman (Comman	
Joinal Advanced Settings can be used to customize program parameters.	PKS/NRPS Domain Search
with speet KS domains Image: KS domains KS domains KS domains C domains Image: KS domains Second	ptional Advanced Settings can be used to customize program parameters.
■ Kis domains ● C domains uery type ● Predicted protein sequences (amino acid) ● Predicted coding sequences or PCR products (DNA) ● Genome or metagenome contigs (DVA) ■ Carbon or metagenome contigs (DVA) ■ pload a file: Examinar	Domain type
uery type Predicted ording sequences or PCR products (DNA) @ Genome or metagenome contigs (DNA) @ Lery sequence Enter or paste sequence(s) in FASTA format, or upload a FASTA file. Please avoid using sequence id numbers ordinating non-standard characters (e.g. commas, slashes, colons, parentheses, ampersands, etc.), as well as imbiguity codes. "R" and "Y" for nucleic acid sequence files. vote: Current size limit for uploaded files is 30 MB. For larger files, please consider our suggestions for file size nanagement before submission. pload a file: Examinar, KT881498.fasta SEEEK RESET >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>	O KS domains © C domains
Predicted coding sequences (amino acid) Predicted coding sequences or PCR products (DNA) Genome or metagenome contigs (DNA) Cuery sequence Ther or paste sequence(s) in FASTA format, or upload a FASTA file. Please avoid using sequence id numbers containing non-standard characters (e.g. commas, slashes, colons, parentheses, ampersands, etc), as well as mibguity codes "R" and "Y" for nucleic acid sequence files. Use: Current size limit for uploaded files is 30 MB. For larger files, please consider our suggestions for file size nanagement before submission. Ipload a file: Examinar KT881498.fasta SEEK RESET b ID 316393149 timated processing time: "4 seconds. IPMIT 10B arch Parameters • query type = genome • number of input sequences = 1 • omain type = C • timm: cutoff = 1e-5 • nana saved his per query = 1 • efference sequences = 1 • of main type if do main type if the side acid sequence if acid sequence if the sequences in the or by the resentence in the fore input sequences = 1 • of main type if do main type if the sequences = 1 • of main type if the domain type if the sequences = 1 • of main type if the domain type if the sequences = 1 • of main type if the domain type if the sequences = 1 • of main type if the domain type if the sequences = 1 • of main type if the domain type if the sequences = 1 • of main type if the domain type if the domain type if the sequences = 1 • of main type if the domain type if the sequences = 1 • of main type if the domain type if the d	uery type
Predicted coding sequences or PCR products (DNA) Cuery sequence There or metagenome contigs (DNA) Cuery sequence There or paste sequence(s) in FASTA format, or upload a FASTA file. Please avoid using sequence id numbers containing non-standard characters (e.g. commas, slashes, colons, parentheses, ampersands, etc), as well as minipuly codes? K* and Y* for nucleic acid sequence files. Vote: Current size limit for uploaded files is 30 MB. For larger files, please consider our suggestions for file size nanagement before submission Jpload a file Examinar, KT881496.fasta SEEK RESET Do ID 316393149 Imated processing time: -4 seconds. UBMIT JOB arch Parameters • umber of input sequences = 1 • ummat file fight = 200 aa • inmer wake for ding the 200 aa • inmer wake sequences = 1 • of man type = C • inmound the fight = 200 aa • inmer wake for ding the 200 aa • inmer wake for ding the 200 aa • inmer wake sequences = 1 • of man type = C • inmound the fight = 200 aa • inmer wake for ding the 200 aa • inmer wake sequences = 1 • of man type = C • inmound the fight = 200 aa • inmer wake sequences = 1 • of man type = C • inmound the fight = 200 aa • inmer wake for ding the 200 aa • inmer wake for ding the 200 aa • inmer wake sequences = 1 • of man type = C • inmound the fight = 200 aa • inmer wake for ding the 200 aa • inmer wake for ding the 200 aa • inmer wake for ding the 200 aa • inmer wake sequences = 1 • of man type = C • inmound the fight = 200 aa • inmer wake for ding the 200 aa • inmer wake sequences = 1 • of man type = C • inmound the fight = 200 aa • inmer wake sequences = 1 • of man type = C • inmound the fight = 200 aa • inmer wake sequences = 1 • of man type = C • inmound the fight = 200 aa • inmer wake sequences = 1 • of man type = C • inmound the fight = 200 aa • inmer wake sequences = 1 • of man type = C •	O Predicted protein sequences (amino acid)
Duery sequence Enter or pasts sequence(s) in FASTA format, or upload a FASTA file. Please avoid using sequence id numbers ontaining non-standard characters (e.g. commas, stashes, colons, parentheses, ampersands, etc.), as well as influguity codes. "Frand "Y" for nucleic acid sequence files." Note: Current size limit for uploaded files is 30 MB. For larger files, please consider our suggestions for file size nanagement before submission. Imagement before submis	 Predicted coding sequences or PCR products (DNA) Genome or metagenome contigs (DNA)
Duery sequence Enter or paste sequence(s) in FASTA format, or upload a FASTA file. Please avoid using sequence id numbers containing on-standard characters (e.g. commas, slashes, colons, parentheses, ampersands, etc.), as well as inhibiguity codes "R" and "Y" for nucleic acid sequence files. Vote: Current size limit for uploaded files is 30 MB. For larger files, please consider our suggestions for file size nanagement before submission. Imagement before submission. Ipload a file: Examinar KT881498.fasta SEEK RESET Imated processing time: -4 seconds. Immuted processing time: -4 seconds. Immuted processing time: -4 seconds. • query type = genome • number of input sequences = 1 • domain type = C • minimum match length = 200 aa	
<pre>inter or paste sequence(s) in FASTA format, or upload a FASTA file. Please avoid using sequence id numbers containing non-standard characters (e.g. commas, slashes, colons, parentheses, ampersands, etc.), as well as imbjuity codes "R" and "Y" for nucleic acid sequence files. Wote: Current size limit for uploaded files is 30 MB. For larger files, please consider our suggestions for file size nanagement before submission. Jpload a file: Examinar, KT881498.fasta SEEK RESET >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>></pre>	Query sequence
<pre>ofinalming itorisatical characters (eg colimitas, stantes, colors, parentifieses, ampersands, etc.), as wen as minipuity codes "R" and "\" to nucleic acid sequence files. Note: Current size limit for uploaded files is 30 MB. For larger files, please consider our suggestions for file size nanagement before submission. pload a file: Examinar, KT881498.fasta set (RESET b) ID 316393149 timated processing time: ~4 seconds. UBMIT JOB arch Parameters</pre>	Enter or paste sequence(s) in FASTA format, or upload a FASTA file. Please avoid using sequence id numbers
<pre>vote: Current size limit for uploaded files is 30 MB. For larger files, please consider our suggestions for file size nanagement before submission. Jpload a file: Examinar KT881498.fasta SEEK RESET bl ID 316393149 timated processing time: -4 seconds. UBMIT JOB arch Parameters</pre>	ambiguity codes "R" and "Y" for nucleic acid sequence files.
Jpload a file: Examinar KT881498.fasta SEEK RESET Sb ID 316393149 timated processing time: ~4 seconds. UBMIT JOB arch Parameters • query type = genome • number of input sequences = 1 • domain type = C • Imm cutoff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max save for db ma	Note: Current size limit for uploaded files is 30 MB. For larger files, please consider our suggestions for file size management before submission.
Jpload a file: Examinar, KT881498.fasta SEEK RESET bl ID 316393149 timated processing time: ~4 seconds. UBMIT JOB arch Parameters • query type = genome • number of input sequences = 1 • domain type = C • himm cutoff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max saved his per query = 1 • reference sequences = all_C_public_12062011 faa • druthene markin = all_C_public_12062011 faa	
Jpload a file: Examinar KT881498.fasta SEEK RESET bb ID 316393149 timated processing time: ~4 seconds. UBMIT JOB arch Parameters • query type = genome • number of input sequences = 1 • domain type = C • hinm cutoff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max saved hits per query = 1 • reference sequences = all C_public_12062011.faa • druthene merging = all @C_public_12062011.faa	
Jpload a file: Examinar KT881498.fasta SEEK RESET Db ID 316393149 timated processing time: ~4 seconds. UBMIT JOB tarch Parameters • query type = genome • number of input sequences = 1 • domain type = C • himmour doff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max save dhits per query = 1 • reference sequences = all_C_public_12062011.faa	
Jpload a file: Examinar, KT881498.fasta SEEK RESET Db ID 316393149 timated processing time: ~4 seconds. UBMIT JOB tearch Parameters • query type = genome • number of input sequences = 1 • domain type = C • hmmc cutoff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max saved hits per query = 1 • reference sequences = all_C_public_12062011.faa • detabanes urgen and 0.0 Ache an	
Jpload a file: Examinar KT881498.fasta SEEK RESET Db ID 316393149 timated processing time: ~4 seconds. UBMIT JOB arcch Parameters • query type = genome • number of input sequences = 1 • domain type = C • hmm cutoff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max save dh its per query = 1 • reference sequences = all_C_public_12062011.faa • dotabene uremain = all_0 @ adice	
Jpload a file: Examinar KT881498.fasta SEEK RESET bb ID 316393149 timated processing time: ~4 seconds. UBMIT JOB Arcch Parameters • query type = genome • number of input sequences = 1 • domain type = C • hminicuuff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • minimum match length = 1200 aa • min e-value for db match = 1e-5 • minimum match length = 1200 aa • min e-value for db match = 1e-5 • max saved hits per query = 1 • reference sequences = all C_public_12062011.faa • dottehene surgring = pin 02 order	
Jpload a file: Examinar KT881498.fasta SEEK RESET Db ID 316393149 timated processing time: ~4 seconds. UBMIT JOB search Parameters • query type = genome • number of input sequences = 1 • domain type = C • hmm cutoff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max saved hits per query = 1 • reference sequences = al. C_public_12062011.faa • dotsheav surgina = take 00 actool 12062011.faa	
Jpload a file: Examinar, KT881498.fasta SEEK RESET Db ID 316393149 timated processing time: ~4 seconds. UBMIT JOB search Parameters • query type = genome • number of input sequences = 1 • domain type = C • hmm cutoff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max saved hits per query = 1 • reference sequences = alC_public_12062011.faa • dothene vargine rate 00 action 12062011.faa	
Jpload a file: Examinar KT881498.fasta SEEK RESET Db ID 316393149 timated processing time: ~4 seconds. UBMIT JOB search Parameters • query type = genome • number of input sequences = 1 • domain type = C • hmm cutoff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max saved hits per query = 1 • reference sequences = all_C_public_12062011.faa • dtothene vargine = the C0 = c02011.faa	
SEEK RESET SEEK RESET Sob ID 316393149 timated processing time: ~4 seconds. UBMIT JOB March Parameters • query type = genome • number of input sequences = 1 • domain type = C • hmm cutoff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max saved hits per query = 1 • reference sequences = all_C_public_12062011.faa	Upload a file: Examinar KT881498.fasta
SEEK RESET bb ID 316393149 timated processing time: ~4 seconds. UBMIT JOB sarch Parameters • query type = genome • number of input sequences = 1 • domain type = C • hmm cutoff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max saved hits per query = 1 • reference sequences = all_C_public_12062011.faa • dotsheav surgina = the 02 acde	
bb ID 316393149 timated processing time: ~4 seconds. UBMIT JOB arcch Parameters • query type = genome • number of input sequences = 1 • domain type = C • hmm cutoff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max saved hits per query = 1 • reference sequences = all_C_public_12062011.faa • detehene surgina = the 02 action	SEEK RESET
timated processing time: ~4 seconds. UBMIT JOB Parch Parameters • query type = genome • number of input sequences = 1 • domain type = C • hmm cutoff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max saved hits per query = 1 • reference sequences = all_C_public_12062011.faa • dotabase varging = the 00 actions = acti	ob ID 316393149
uBMIT JOB arch Parameters • query type = genome • number of input sequences = 1 • domain type = C • hmm cutoff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max saved hits per query = 1 • reference sequences = all_C_public_12062011.faa • dotebreav surgion = pin 02 order	timated processing time: ~4 seconds.
<pre>arch Parameters • query type = genome • number of input sequences = 1 • domain type = C • hmm cutoff = 1e-5 • minimum match length = 200 aa • min e-value for db match = 1e-5 • max saved hits per query = 1 • reference sequences = all_C_public_12062011.faa • detebrace vergina = ben 02 action = ben 02 • detebrace vergina = ben 02 • of the detebrace vergina</pre>	UBMIT JOB
 earch Parameters query type = genome number of input sequences = 1 domain type = C hmm cutoff = 1e-5 minimum match length = 200 aa min e-value for db match = 1e-5 max saved hits per query = 1 reference sequences = all_C_public_12062011.faa detabace version = pto 02 actoc 	
 query type = genome number of input sequences = 1 domain type = C hmm cutoff = 1e-5 minimum match length = 200 aa min e-value for db match = 1e-5 max saved hits per query = 1 reference sequences = all_C_public_12062011.faa datebase version = pic_02_adea 	earch Parameters
 number of input sequences = 1 domain type = C hmm cutoff = 1e-5 minimum match length = 200 aa min e-value for db match = 1e-5 max saved hits per query = 1 reference sequences = all_C_public_12062011.faa detabase version = pto 0.2 order 	• query type = genome
 hmm cutoff = 1e-5 minimum match length = 200 aa min e-value for db match = 1e-5 max saved hits per query = 1 reference sequences = all_C_public_12062011.faa datebace version = pto_02_adeo 	 number of input sequences = 1 domain type = C
 min e-value for db match = 1e-5 max saved hits per query = 1 reference sequences = all_C_public_12062011.faa detabase varging = ptg_02_cdec 	hmm cutoff = 1e-5 minimum match length = 200 aa
• reference sequences = all C_public_12062011.faa • database variase pice 02_dece	min e-value for db match = 1e-5 max saved hits per guery = 1
	reference sequences = all_C_public_12062011.faa database variase = ks_02_sdase

Figure S1. Screenshots of the NaPDos web server employed for the phylogeny-based bioinformatic analysis of the condensation like domains in telomycin NRPSs. The Genebank accession number of telomycin gene cluster is KT881498 and its sequence in FASTA format was uploaded for the query.

Genome Search Results

Summary

- Domain type: C
 Minimum sequence length: 200
 HMM cutoff: -T -91.5
 Domains found: 13

Domain Locations

domain id	parent seq	start	end	read frame	cand_id
C1	KT881498.1	24852	25742	3	KT881498.1_3_6_94_390
C2	KT881498.1	27984	28898	3	KT881498.1_3_6_1138_1442
C3	KT881498.1	31161	32066	3	KT881498.1_3_6_2197_2498
C4	KT881498.1	34368	35279	3	KT881498.1_3_6_3266_3569
C5	KT881498.1	37587	38582	3	KT881498.1_3_6_4339_4670
C6	KT881498.1	39158	40066	2	KT881498.1_2_19_87_389
C7	KT881498.1	42293	43198	2	KT881498.1_2_19_1132_1433
C8	KT881498.1	45491	46426	2	KT881498.1_2_19_2198_2509
C9	KT881498.1	48710	49636	2	KT881498.1_2_19_3271_3579
C10	KT881498.1	51920	52831	2	KT881498.1_2_19_4341_4644
C11	KT881498.1	55103	56065	2	KT881498.1_2_19_5401_5722
C12	KT881498.1	56626	57519	1	KT881498.1_1_24_788_1085
C13	KT881498.1	59782	60669	1	KT881498.1_1_24_1834_2135

Further Analysis Options

Click on the link below to analyze pathways associated with these candidates.

GET MORE INFO

Database Search Results

C domain matches were found for 13/13 query sequences. Use check boxes to select candidates for further analysis

Select All

	Query id	Database match id	percent identity	align length	e-value	pathway product	domain class	^
6	KT881498.1_1_24_788_1085 [C12]	cdaps3_C1_DCL	55	292	8e-80	calcium- dependent antibiotic	DCL	
	KT881498.1_1_24_1834_2135 [C13]	cyclom1C3_LCL	59	294	6e-91	cyclomarin	LCL	
D	KT881498.1_2_19_87_389 [C6]	act3_C1_DCL	52	297	2e-72	actinomycin	DCL	
	KT881498.1_2_19_1132_1433 [C7]	act3_C2_LCL	57	296	4e-72	actinomycin	LCL	
6	KT881498.1_2_19_2198_2509 [C8]	cdaps1_C2_LCL	59	307	1e-81	calcium- dependent antibiotic	LCL	
	KT881498.1_2_19_3271_3579 [C9]	prist2_C3_LCL	61	305	5e-88	pristinamycin	LCL	
	KT881498.1_2_19_4341_4644 [C10]	act3_C2_LCL	59	298	2e-94	actinomycin	LCL	
	KT881498.1_2_19_5401_5722 [C11]	tioR_C2E	44	313	3e-48	thiocoraline	epim	
þ	KT881498.1_3_6_94_390 [C1]	bacil2_C1_start	39	290	1e-47	bacillibactin	start	
	KT881498.1_3_6_1138_1442 [C2]	cdaps2_C2_LCL	55	298	1e-67	calcium- dependent antibiotic	LCL	~
	KT881498.1_3_6_1138_1442 [C2]	cdaps2_C2_LCL	55	298	1e-67	calcium- dependent antibiotic	LCL	
	KT881498.1_3_6_2197_2498 [C3]	act3_C2_LCL	56	296	1e-75	actinomycin	LCL	
	KT881498.1_3_6_3266_3569 [C4]	act3_C3_LCL	61	298	2e-80	actinomycin	LCL	
	KT881498.1_3_6_4339_4670 [C5]	cdaps1_C4E	45	322	2e-59	calcium- dependent antibiotic	epim	~
<							>	

Figure S2. Screenshots of the NaPDos web server showing the classification of the different condensationlike domains identified in telomycin NRPSs. The information retrieved in the Database Search Results table was employed for preparing Table 1 in the body of the article.



Figure S3. Overlay of the multiplicity-edited HSQC spectra (expanded methyl region) of telomycin hydrolyzate (blue cross-peaks) and Thr standard (green cross-peak). The perfect match of the amino acid C_{β} -methyl cross-peak in both spectra confirms that Thr is a constituent amino acid in telomycin.



Figure S4. Overlay of the multiplicity-edited HSQC spectra (expanded methyl region) of telomycin hydrolyzate (blue cross-peaks) and *allo*-Thr standard (green cross-peak). The close match of the amino acid C_{β} -methyl cross-peak in both spectra confirms that *allo*-Thr is a constituent amino acid in telomycin.



Figure S5. Overlay of the multiplicity-edited HSQC spectra (expanded C4 methylene region) of telomycin hydrolyzate (red cross-peaks), *cis*-3-Hyp standard (cyan cross-peaks) and *trans*-3-Hyp standard (magenta cross-peak). The close match of the C4–methylene cross-peaks when comparing the hydrolyzate and the standards spectra confirms that both *cis*-3-Hyp and *trans*-3-Hyp are constituent amino acids in telomycin.



Figure S6. Overlay of the multiplicity-edited HSQC spectra (expanded C3 methine region) of telomycin hydrolyzate (blue cross-peaks), *cis*-3-Hyp standard (green cross-peak) and *trans*-3-Hyp standard (brown cross-peak). The close match of the C3–methine cross-peaks when comparing the hydrolyzate and the standards spectra confirms that both *cis*-3-Hyp and *trans*-3-Hyp are constituent amino acids in telomycin.



Figure S7. Overlay of the multiplicity-edited HSQC spectra (expanded methyl region) of telomycin hydrolyzate (blue cross-peaks) and *threo*- β -Hyl standard (green cross-peak). The absence of resonance frequency match between the β -Hyl methyls cross-peaks when comparing both spectra indirectly confirms that *erythro*- β -Hyl is a constituent amino acid in telomycin.

Aspartic acid

1200000 -	<i>m/z</i> = 412, [M-H]	7.90	L-Asp-L-FDVA standard
800000 -			
e00000 -			
400000 -			
200000 -			
0 -			· · · · · · · · · · · · · · · · · · ·
260000 -	<i>m/z</i> = 412, [M-H]	7.91	L-FDVA derivatized hydrolyzate
150000 -			

٨٨



Serine

250000 -	<i>m/z</i> = 386, [M+H] ⁺	7.50	L-Ser-L-FDVA standard
200000 -			
150000 -			
100000 -			
50000 -			
0			· · · · · · · · · · · · · · · · · · ·
120000 -			I-FDVA
100000 -	<i>m/z</i> = 386, [M+H] ⁺	7.49	derivatized hydrolyzate
80000 -			
e0000 -			



Threonine





allo-Threonine



Alanine





trans-3-Hydroxyproline





cis-3-Hydroxyproline





erythro-β-Methyl-Tryptophan





threo-β-Hydroxyleucine

40000 - 380000 - 380000 -	<i>m/z</i> = 853, [2M-H]	10.76	L- <i>threo</i> -β-Hyl-L-FDVA standard
250000 -			
200000 -			
150000 -			
100000 -			
50000 -			
•			

600000 - - - 400000 -	<i>m/z</i> = 853, [2M-H]	10.95	L-FDVA derivatized hydrolyzate
300000 -			
200000 -			
100000 -			
0-			·····





Figure S9. Zoomed region of the ROESY spectrum of telomycin highlighting the key correlations determining the *Z* geometry of the olefinic double bond in the $\Delta_{2,3}$ -Trp residue.



Figure S10. Diagnostic ¹³C chemical shifts determining the sequence position of L-*cis*-3-Hyp and L-*trans*-3-Hyp epimers in telomycin.



Figure S11. *J*-based configurational analysis (JBCA) approach for determining the sequence position of L-Thr and D-*allo*-Thr residues in telomycin.



Figure S12. *J*-based configurational analysis (JBCA) approach for determining the relative configuration of the D-*erythro*- β -MeTrp and L-*erythro*- β -Hyl residues in telomycin.



Figure S13. ¹H NMR spectrum of telomycin (DMSO-d₆, 40 °C, 500 MHz).



Figure S14. ¹³C NMR spectrum of telomycin (DMSO-d₆, 40 °C, 125 MHz).



Figure S15. J-Resolved spectrum of telomycin.



Figure S16. COSY spectrum of telomycin.



Figure S17 TOCSY spectrum of telomycin.







Figure S19. HSQC spectrum of telomycin.



Figure S20. HMBC ("8 Hz") spectrum of telomycin.



Figure S21. HMBC ("2 Hz") spectrum of telomycin.

References

- 1. C. E. Boville, R. A. Scheele, P. Koch, S. Brinkmann-Chen, A. R. Buller and F. H. Arnold, Engineered biosynthesis of β-alkyl tryptophan analogues, *Angew. Chem. Int. Ed.*, 2018, 57, 14764-14768.
- 2. P. Kumar, A. Meza, J. M. Ellis, G. A. Carlson, C. A. Bingman and A. R. Buller, L-Threonine transaldolase activity is enabled by a persistent catalytic intermediate, *ACS Chem. Biol.*, 2021, 16, 86-95.
- 3. N. Ziemert, S. Podell, K. Penn, J. H. Badger, E. Allen and P. R. Jensen, The natural product domain seeker NaPDos: a phylogeny based bioinformatic tool to classify secondary metabolite gene diversity, *PLoS One*, 2012, 7, e34064.
- 4. w. NaPDoS.
- 5. D. Y. Liu, Y. Li and N. A. Magarvey, Draft genome sequence of *Streptomyces canus* ATCC 12647, a producer of telomycin, *Genome Announc.*, 2016, 4, e00173-00116.
- 6. Y. Yokote, K. M. Arai and K. Akahane, Recovery of tryptophan from 25-minute acid hydrolysates of protein, *Anal. Biochem.*, 1986, 152, 245-249.
- 7. E. Bayer, E. Gil-Av, W. A. Koenig, S. Nakaparksin, J. Oro, W. Parr, Retention of configuration in the solid phase synthesis of peptides, *J. Am. Chem. Soc.*, 1970, 92, 1738-1740.
- S. Nakamukai, K. Takada, K. Furihata, Y. Ise, S. Okada, Y. Morii, N. Yamawaki, T. Takatani, O. Arakawa, K. R. Gustafson, S. Matsunaga, Stellatolide H, a cytotoxic peptide lactone from a deep-sea sponge *Discodermia* sp., *Tetrahedron Lett.*, 2018, 59, 2532-2536.
- 8. C. W. Johnston, M. A. Skinnider, C. A. Dejong, P. N. Rees, G. M. Chen, C. G. Walker, S. French, E. D. Brown, J. Bérdy, D. Y. Liu and N. A. Magarvey, Assembly and clustering of natural antibiotics guides target identification, *Nat. Chem. Biol.*, 2016, 12, 233-239.
- 9. C. Fu, L. Keller, A. Bauer, M. Brönstrup, A. Froidbise, P. Hammann, J. Herrmann, G. Mondesert, M. Kurz, M. Schiell, D. Schummer, L. Toti, J. Wink and R. Müller, Biosynthetic studies of telomycin reveal new lipopeptides with enhanced activity, *J. Am. Chem. Soc.*, 2015, 137, 7692-7705.
- 10. P. E. Hansen, Assignment of the natural abundance carbon-13 spectrum of proteins using carbon-13proton detected heteronuclear multiple-bond correlation NMR spectroscopy: structural information and stereospecific assignments from two- and three-bond carbon-hydrogen coupling constants, *Biochemistry*, 1991, 30, 10457-10466.
- 11. W. Bermel, K. Wagner and C. Griesinger, Proton-detected C,H correlation via long-range couplings with soft pulses; determination of coupling constants, *J. Magn. Reson.*, 1989, 83, 223-232.
- I. Pérez-Victoria, D. Oves-Costales, R. Lacret, J. Martín, M. Sánchez-Hidalgo, C. Díaz, B. Cautain, F. Vicente, O. Genilloud and F. Reyes, Structure elucidation and biosynthetic gene cluster analysis of caniferolides A–D, new bioactive 36-membered macrolides from the marine-derived *Streptomyces caniferus* CA-271066, *Org. Biomol. Chem.*, 2019, 17, 2954-2971.
- 13. N. Matsumori, D. Kaneno, M. Murata, H. Nakamura and K. Tachibana, Stereochemical determination of acyclic structures based on carbon–proton spin-coupling constants. a method of configuration analysis for natural products, *J. Org. Chem.*, 1999, 64, 866-876.