Genetic Interception and Structural Characterization of Thiopeptide Cyclization Precursors from *Bacillus cereus*

Albert A. Bowers, Christopher T. Walsh*, and Michael G. Acker

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

Supporting Information

Contents:

- **1.** Materials and General Methods
- 2. Procedures for Generation of a *tcIM* Knockout and Subsequent Reintroduction
- **3.** Extraction of Intermediates
- 4. LC-MS and MS/MS Analysis
- **5.** Alignment of TcIM homologs.
- 6. LC-MS, Extracted Ion Chromatograms, and MS/MS of Mutant Compounds
 - a. MKO
 - **b.** *E*_{*in*}*MKO*
 - c. MKI
 - d. T3A_{in}MKO
- **7.** Table of ¹H and ¹³C NMR assignment for compound **10**.
- **8.** 1D & 2D NMR spectra of **10**.
- 9. 1D & 2D NMR spectra of 11.

1. Materials and General Methods

All molecular biology, recombinant DNA manipulation and microbiological assays were performed following the protocols of Sambrook et al.¹ Unless otherwise specified, all chemicals were purchased from Sigma-Aldrich. Restriction enzymes and Quick Ligase were purchased from New England Biolabs (Boston, MA). Pfu Turbo DNA Polymerase was purchased from Invitrogen (Carlsbad, CA) and Pag 5000 DNA polymerase from Stratagene (La Jolla, CA).DNA oligonucleotide primers were synthesized by Integrated DNA technologies (Coralville, IA). PCR was performed on a Biorad MyCycler thermal cycler. DNA sequencing was performed by the Molecular Biology Core Facilities at the Dana Farber Cancer Institute (Boston, MA). Top10 chemically competent E. coli cells were purchased from Invitrogen. Restriction endonuclease cleanup and gel extraction of DNA fragments were performed with QiaQuick PCR cleanup kit from Qiagen. Recombinant plasmids were isolated using the QiaPrep Spin Miniprep Kit from Qiagen. B. cereus ATCC 14579 genomic DNA was isolated from cultures using the DNeasy Kit from Qiagen. Analytical RP-HPLC was performed on a Beckman System Gold (Beckman Coulter) instrument using a Phenomenex Luna 5 µm C18(2) 100 Å 250 x 4.6 mm column, monitoring eluent absorption at 220 and 350 nm. Preparative RP-HPLC was performed on a Beckman System Gold (Beckman Coulter) instrument using a Phenomenex Luna 10 µm C18(2) 100 Å 250 x 21.20 mm column. ¹H NMR spectra were recorded on a Varian 600 MHz spectrometer. Chemical shifts are reported in ppm from tetramethylsilane with the solvent resonance resulting from incomplete deuteration as the internal standard (CDCI3 δ 7.26, D2O δ 4.79, CD3OD δ 3.31). Data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, g = quartet, br = broad, m = multiplet), coupling constants (Hz), and integration. Software and methods used for generation of solution phase structures from NMR data are described in the text.

2. Procedures for Generation of a *tcIM* Knockout and Subsequent Reintroduction

In a manner similar to that previously described for plasmid pMGA-tcl Δ E-H,² pMKO was generated from plasmid pKM082³, containing ampicillin and erythromycin resistance cassettes and employed in excision of *tclM via* double crossover homologous recombination. Two regions of homology were cloned into this plasmid: 1) the 1kb sequence immediately upstream of the *tclM* start codon and 2) the 1 kb sequence beginning 15bp upstream of the *tclN* start codon. A rescue plasmid, pMKI, was further generated from pMKO itself, by excision of the 3'-homology region and subsequent ligation with sequence containing *tclM* and a new 3'-homology region (ABprimer112 and 130).

Removal of *tcIM* from the genome of *B. cereus* ATCC 14579 was accomplished through two individual rounds of homologous recombination. In the first round, the entire plasmid was integrated into the chromosome. Transformation was effected into both wild type B. cereus ATCC 14579 and our previously reported tclE-H knockout strain; the transformation protocol, including conditions for growth, inoculation, and electroporation of competent cells from B. cereus ATCC14579 has been previously published.⁴ Positive transformants were selected for on MLS LB-agar plates (containing 1µg/ml erythromycin and 25µg/mL lincomycin) for 36 hours at 30 °C. Individual colonies were transferred to LB liquid cultures without antibiotic and incubated for 24 hours at 30 °C. The liquid cultures were diluted 1:1000 in antibiotic free LB and again incubated for 24 hours at 30 °C. The second recombination removes the plasmid, the gene of interest, and the erythromycin resistance cassette together with it. Thus, after 7-9 rounds of dilution and growth, the cultures were diluted 10⁻⁵, plated on LB agar and grown up over 24 hours at 30 °C. Colonies were then replica-plated via sterile velvet onto MLS LB-agar plates. Both the original, antibiotic-free plates and the fresh, MLS-supplemented plates were grown up over 24 hours at 30 °C. Colonies exhibiting growth on the antibiotic-free plates, but not on the MLS plates were re-struck for verification. Subsequent colony PCR in presence of ABprimer109 and 131 was used to confirm the tclM knockout; the knockout strain exhibited a loss of 924 bp in the PCR product, which was further verified by DNA sequencing.

Upon confirmation of the *tclM* knockout, the above procedure was again employed for reconstitution with pMKI. Introduction of a single copy of *tclE* and the mutant *tclE-T3A* into the *E-H,MKO* strain was achieved by our previously reported protocol.⁵

Oligo	Sequence	Role
ABprimer109	5'-GAT CGG ATC CGT TTT TTA ATA AAG GAA TGA TTA TAT G-3'	MKO-5'-homology
ABprimer110	5'-GAT CCT CGA GTT AAT CAT CCC TTT CTA CTC TTA TAC-3'	MKO-5'-homology
ABprimer112	5'-GAT CCT CGA GGG ATT TTG GTA AAG GGA GGG ATA ATA-3'	MKO-3'-homology
ABprimer131	5'-GAT CGC ATG CAA TAT TCA AAA AAT CAG ACA AAA AG-3'	MKO-3'-homology
ABprimer130	5'-GAT CCT CGA GAT GGA GCA GTA TCA TAA AAT TG-3'	MKI-3'-homology
MGAprimer 209	5'-GAA ATT ATG GGA GCG TCA TGT GCG ACA TGC GTA TGT ACA TGC AG-3'	T3A mutagenesis
MGAprimer 210	5'-CTG CAT GTA CAT ACG CAT GTC GCA CAT GAC GCT CCC ATA ATT TC-3'	T3A mutagenesis

SI Table 2.1. Oligonucleotides used for cloning and *tclM* knockout

3. Extraction of Thiocillin Compounds

Starter cultures (5 mL) were grown in LB for 20 hours at 30 °C. Larger cultures (0.5 L LB in 2 L culture baffles culture flasks) were inoculated with 300 μ L of starter culture and grown for 68 hours at 30 °C with shaking at 200 rpm. (*tclE* mutant strains *E_{in}MKO* and *T3A_{in}MKO* were grown in media supplemented with 1 μ g/mL erythromycin and 25 μ g/mL lincomycin.) Cultures were harvested and both the cell pellet and spent media were saved. To the pellet, 50 mL methanol was added along with 15 g sodium sulfate. The mixture was vortexed vigorously and allowed to sit for at least 10 minutes. The mixture was then filtered through Whatman filter paper (no. 1) and the methanol was removed by vacuum. Solid was solubilized in 10 mL 33%

acetonitrile in water for HPLC analysis. *tclE* mutants that produced compound at low levels were grown in a 5L fermenter in ECPM1 media lacking glycerol (20 g N-Z amine; 3 g Yeast Extract; 1 g KH₂PO₄; 4 g K₂HPO₄; 1 g NH₄Cl; 2.4g K₂SO₄ in 1 L supplemented with 10 mL 100X Trace Elements (5 g EDTA; 0.5 g FeCl₃•6H₂O; 0.05 g ZnO; 0.01 g CuCl₂•2H₂O; 0.01 g Co(NO₃)₂•6H2O; 0.01 g (NH₄)₆ Mo₇O₂₄ in 1 L) and 2 mL of 500X Mg/Ca solution (203 g MgCl₂; 66.2 g CaCl₂ in 1 L). Cells and media were harvested after 24 hours and extraction was performed as detailed above, scaled accordingly.

Further purification was accomplished by ethyl acetate extraction. Solvents were removed from the crude compound extracts on a rotary evaporator. The crude residue was then dissolved in 40 mL of 1:1 EtOAc: water. The biphasic solution was transferred to a 60mL separatory funnel, shaken and the organic layer removed. The aqueous layer was washed with a further 20 mL of EtOAc and the combined organics were dried over Na₂SO₄, filtered through a 60 mL coarse fritted glass funnel, and evaporated to dryness. For purposes of assessing the thiocillin content of the individual layers, the residue from the organic layer was redissolved in 10 mL of acetonitrile. 180 μ L of the acetonitrile solution was combined with 180 μ L of water and 300 μ L of this solution was injected onto the analytical HPLC. 300 μ L of the aqueous layer was also injected, being careful to avoid the surface organics retained from the extraction.

4. LC-MS and MS/MS Analysis.

High-resolution LC-MS data was collected in positive ion mode, on an Agilent 6520 Accurate-Mass Q-TOF Mass Spectrometer fitted with an electrospray ionization (ESI) source. The capillary voltage was set to 3500 kV, and the fragmentor voltage at 250 V. The drying gas temperature was maintained at 350°C with a flow rate of 12 L/min and a nebulizer pressure of 45 psi. Separation was effected on a Gemini-NX C18 reverse phase column (5µm, 110A, 2.0 x 50 mm, Phenomonex) for crude mixtures and a Kinetex C18 reverse phase column (2.6µm, 100A, 2.10 x 50 mm, Phenomonex) for chromatographically pure samples. Compounds were eluted in a gradient of solvents A (0.1% formic acid in water) and B (0.1% formic acid in acetonitrile): 2 min. isocratic 2%B, then increasing to 100%B over 10 min., and finally isocratic at 100%B for 2 min. before returning to 2%B and reequilibrating over 4 min. The order of elution relative to tailored states of the final products was conserved across variants, except where the short gradient created elution overlap. At least two analytical runs were performed for extracts from each mutant: crude extract was used in the first run in order to better search for the presence of trace quantities of all tailored states and purified compounds were examined in a second run to obtain high resolution masses with lower ppm error than those observed in the crude runs. Additional structural analysis was accomplished by targeted CID-MS/MS. For all samples examined, the collision energy was varied between 35 and 50 eV, with optimum fragmentation generally being observed at 45 eV. Representative spectra are illustrated below. Essential diagnostic peaks have been labeled.

(1) Sambrook, J.; Fitsch, E. F.; Maniatis, T. *Molecular Cloning. A Laboratory Manual. 3rd ed*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2001.

(2) Acker, M. G.; Bowers, A. A.; Walsh, C. T. J. Am. Chem. Soc. 2009, 131, 17563.

(3) Brown, L. C.; Acker, M. G.; Clardy, J.; Walsh, C. T.; Fischbach, M. A. *Proc Natl Acad Sci U S A* **2009**, *106*, 2549.

(4) Turgeon, N.; Laflamme, C.; Ho, J.; Duchaine, C. *J. Microbiol. Methods.* **2006**, 67, 543.

(5) Bowers, A. A.; Acker, M. G.; Koglin, A.; Walsh, C. T. *J. Am. Chem. Soc.* articles ASAP.

Alignments:

All protein sequences were acquired from GenBank according to published accession numbers or compound-producing species. Protein alignments were performed with ClustalW2 (<u>http://www.ebi.ac.uk/Tools/clustalw2/index.html</u>) using the default parameters. Alignment files were analyzed using GeneDoc software version 2.6.002 for PC (<u>http://www.psc.edu/biomed/genedoc</u>) with similarity groups enabled.

Table 1: Results of protein alignment of TcIM with proposed Diels-Alderase enzymes from other thiazolyl peptide natural product gene clusters.

Compound	Producing Organism (Diels-Alderase)	%Identity/Similarity*
Thiocillin	Bacillus cereus ATCC 14579 (TclM)	N/A
Nosiheptide	Streptomyces actuosus (NosO)	13/27
Thiostrepton	Streptomyces laurentii (TsrE)	14/31
Siomycin	Streptomyces sioyaensis (SioL)	14/31
GE2270A	Nonomuraea sp. WU8817 (TpdD)	12/32
Thiomuracin	Nonomuraea sp. Bp3714-39 (TpdD)	11/27
Cyclothiazamycin	Streptomyces hygroscopicus (CltG)	12/29

*%Identity/Similarity compared to the Thiocillin Diels-Alderase, TcIM

Alignment 1: TclM (Thiocillin) and NosO (Nosiheptide). Identical residues are highlighted in red. Similar residues are highlighted in gray.



Alignment 2: TsrE (Thiostrepton) and SioL (Siomycin). Identical residues are highlighted in red. Similar residues are highlighted in gray.



Alignment 3: TpdD (GE2270A), TpdD (Thiomuracin) and CltG (Cyclothiazamicin). Identical residues are highlighted in red. Similar residues (similarity limit = 50%) are highlighted in gray.



Alignment 4a: Predicted Diels-Alderase enzymes in thiazolyl peptide antibiotic biosynthesis gene clusters, depicting amino acid identity. 100% identity = red; 75% identity = blue; 50% identity = gray.

	*	20	*	40	*	60	*		
Thiocillin	•		MEOYHKTVI	TGSNAETM	NTEP	v	AVKFINNY		33
Nosibertide	•МТС		CAWLETCLD			DAFDCADV	POFFLECV	:	63
Thisstreptice		GE GONE ADAAIIA	ACTICEORDI			FOURT	CWEVILEN	:	10
Signation		MC31	ACCEPTIN				OWEVILEN	:	40 E1
GEORGE					WEDA	ACDOM DT	DDUIEDWI	:	67
GE2270A	MINIMPL DVDVUDO	DLDGLIL-ATRPI	LAGTPGRGWF		WFDA	-AQPSWERI	RDVLEPWL	:	67
	. MIWIKLKVDINDG	PMDDLILDALRP-	MADWARADWE		NEDGOOD	-DAIPI	VRAVE	•	47
Cyclothiazamycin	:		-M-PWAEAAWF(SKHWLKGPHLK	NFRCRGTL	WEERVRPTV	TGIVIDIL	:	4/
	00	+ 1/		100		± 1	40		
mhi a ai 11i a	80								100
Thiocillin	: KGFFIVEKISKDF	PIIDVIINNKIV	PENQLNKI QN:	SGARIRIIRNS	TENETOGNE	GMLGDKILA.	EFERNE	: 1	106
Nosineptide	: GAAQPALVVQLEV	TPGTDLAEPYA	KARALAAG GLI	PVQVAAGRATL	PLAGSVFA		AALAAVCP	: 1	136
Thiostrepton	: HGGRPFLRLRFAS	RSPSVERRL	KSRILAHVGPT.	LDAGDVFTYQP	NHEHDW G	GTAGLGIAE	NEWTETTP	: 1	118
Siomycin	: YGGRPFLRLRFAT	RS SVERRLI	KSRIMEHIG-SA	AASDDPFEHQPS	NHEHDW	GEAGLGIAE	APWTETTP	: 1	120
GE2270A	: RVNPSRARIDRDR	LLAQHRHLAZ	AERIDEPILPI	FYADNTLHRAAI	RSRAHV	CPAAEEIIFH	DEHTTASA	: 1	137
Thiomuracin	: AACPSAGTTDPQA	LLPLHERLAP	ELEGERGP	WAPDNTVTAEPI	?GI	DTELDRFLA	DEYADTTE	: 1	127
Cyclothiazamycin	: RARPSAARLDEGA	LAPVHARLAI	ELEMETGPRHP	WVPDNTVLER PY	DHRLPV	SLRASELLA	G <mark>F</mark> LSD <mark>T</mark> NG	: 1	117
		_							
	* 16	0 *	180	*	200	*	22		
Thiocillin	: IS NICNQNFESY	NKKIEFALEIMLI	ISAHYNYDS		-IKK <mark>G</mark> YLSY	ASHVNG FT	RWKDPNKI	: 1	165
Nosiheptide	: ALLTATEAAEQGR	PALIAS <mark>A</mark> AELMSA	HLRAVSVSAA	PGPRQWEELREC	SVPL <mark>G</mark> FLS	R <mark>SHAE</mark> AFLA	SSRDPKAA	: 2	209
Thiostrepton	: L <mark>ALDTLRATR</mark> GNR	ALRLAV <mark>A</mark> F <mark>D</mark> FLV(CTGVMLAPHLP1	PSIAKFO	GYKA <mark>G</mark> Y <mark>LS</mark> Y	LATF <mark>E</mark> GYML	LIR <mark>DP</mark> EGT	: 1	186
Siomycin	: L <mark>ALRTL</mark> RAT <mark>R</mark> GDR	ALRLAA <mark>A</mark> F <mark>D</mark> FLV0	CSGVLLAPHLPI	PPVAKFO	FKA <mark>G</mark> YLS	LATF <mark>E</mark> GYML	LIR <mark>DP</mark> EGT	: 1	188
GE2270A	: VAYDELDAVRAGE	SR-LVM <mark>AL</mark> DLMV	AAHAHAE	GG	-VRG <mark>G</mark> FV <mark>S</mark> E	R <mark>SHAE</mark> AFL <mark>A</mark>	SA <mark>P</mark> G-L	: 1	193
Thiomuracin	: A <mark>A</mark> FDA <mark>L</mark> GRV <mark>R</mark> AGT	PL-PGI <mark>A</mark> F <mark>D</mark> LVV	ATAHDLSE	GG	-LPTART <mark>S</mark> I	R <mark>SHAE</mark> AYLS	RL <mark>P</mark> GGV	: 1	184
Cyclothiazamycin	: Q <mark>A</mark> FRAYERV <mark>R</mark> AGG	AL-SLLALDLMW	TTSVAAVPF	TTGGEI	PIER <mark>G</mark> F <mark>LS</mark> I	R <mark>SH</mark> ADAFLS	RTRDPVAV	: 1	182
			_						
	0 *	240	*	260	*	280	*		
Thiocillin	: RDIFHKN <mark>Y</mark> LNNKE	Y <mark>l</mark> eskvseiidnn	INRSSLS			ELSDIITEM	KKEMTTDI	: 2	214
Nosiheptide	: QAMMDAKYTRAAA	T <mark>L</mark> ERLVDGVLTQ0	CEERG		-PVVSLPAF	RQ <mark>W</mark> YEAM <mark>R</mark> AA	KPAVTE F	: 2	264
Thiostrepton	: RAKHAQRYEKNRE	L <mark>LRP</mark> RLRTLVEQ	ISEPDGE	I	TDVPELAF	EWLVRLRDY	VPALQKGF	: 2	244
Siomycin	: RAKHAQRYEQNRN	LIRPRLRALVEO	QDPEGD	I	LADVPELAF	EWAVRLRGY	LPAIRKGF	: 2	246
GE2270A	: RERWDAEYAARAG	ALRARITAVVAG	PRGR			AWAGLLDRF	ADRGDE	: 2	240
Thiomuracin	: RAKWOAHYERNOE	PITARIKALTGA	SEPG			AWLRTIRAT	RDRGRTII	: 2	230
Cvclothiazamvcin	: RAAFDDRFRROET	VICERLRSVEAAI	SDGATEGDGAI	DRSDRSDRSEAU	GDVVPFVI		ORIAHPLL	: 2	255
0/0100000000000000000000000000000000000					0201111		x		
	300	*	320	* 340)	*	360		
Thiocillin	: EKC-NHVFNIEL	LOKPGERD		FLEKSOF	KTILNNPE	FSNFMNKDI		: 2	267
Nosiheptide	: RAGTO A DTEEO	PPD GPDGKG		LSESAF	RIVEGSDO		SFLATRLL	: 3	319
Thiostrepton	DEC-REY YATTPR	KAELAKLTPSPD	TYRRPDVEWL	SDLPEPPVAGT	RATADNTY	YOGMTREDR	RELASRIA		316
Siomycin	· DEC-REY YATER	KAETAKLTPSPD	TYRRPKVEWL	ADLPEAPVACT	RATADNTY	YOGMTREDR	RELASRLA		318
GE2270A	· ASC-ALVEDACD	DAVARD			PALPOND	WHEFULDCA		: 3	288
Thiomuracin	DEC-DECCVATD				PNI PTOFE	WILLS VEROR			200
Cuclothiagamucin	· ASC-EVENCCAAD				AVIDED		MEASEDIM		202
Cyclothiazamycin	. 110 EV SMOGAAR			KI SE	DRODRO		- ACC NUM	• •	
	* 3	80 *	400	*	420	*			
Thiocillin	: TVF YLLIRNIET	ONKORYLLCYYTY		TLELIRDFGKGF	RDNNVEDT	RY	: 3	25	
Nosiheptide	TSLLYLSLSSVET	ALAEBYFLCYAV	SBACESTEDTD	T.TVI.SGLARTS	STAS		: 3	70	
Thiostrepton		LLADBYT FYLT	BAFFERVOID	AAT.TROUPPET	EVAC		3	68	
Siomycin			BAFFERVOID	AET.TROVICEEP	EVSC			70	
CE2270A	· INT VTOTOTOT		SAVEOEXCITC	TETAMOCA				22	
				C				10	
Thiomuracin Coolethicerers	: LNCAYLELTRLEL	TPDQRF ICH A						19	
Cyciotniazamycin	: MINYLY IINRLCL	RPV KA ICH A		VGSFQRIVAS	DPSSERPE	WKKIGEAWA	AGGAG : J	15	

Alignment 4b: Predicted Diels-Alderase enzymes in thiazolyl peptide antibiotic biosynthesis gene clusters, depicting amino acid similarity. 100% similarity = red; 75% similarity = blue; 50% similarity = gray.

		*	20	*	40		*	60	*		
Thiocillin	:			ME Y	HKIVLTGS	NAETMLIK	N <mark>I</mark> EP		- <mark>V</mark> AVKFINNY	:	33
Nosiheptide	:	MTSGPG	QAPAEAAHA	AGAAWLE:	IGLDAPAD	AVPALVAG	V <mark>V</mark> RPLLRE	PAEPGA	P <mark>V</mark> PGFFLRGV	:	63
Thiostrepton	:			MSTSECKI	DLTVSVPW	SVQEDLL	DVAAPLLE	ESVELG	TDSWFYLREN	:	48
Siomycin	:		MSA	MSSSE	DLTVSVP <mark>W</mark>	SVQEDLL	D <mark>V</mark> AAPTLE	ESVALG	TESWF YLR EN	:	51
GE2270A	: MSWRRVDV	AYHDPDLD	GLIL-ATRE	LLAGTPGI	RGWFQRH <mark>W</mark>	VRGPH <mark>L</mark> E <mark>L</mark> I	WFDA	-AQPSW	RIRDVLEPWL	:	67
Thiomuracin	: MTWTRLRV	DYHDGPMD	DLILDALRE	AWHEIN	RGYFLRH <mark>W</mark>	VCGPH <mark>L</mark> R <mark>I</mark>	F <mark>V</mark> DG	-DAT	P <mark>I</mark> VRAVERHL	:	63
Cyclothiazamycin	:			-MAPWAE	AAWFGRH <mark>W</mark>	LRGPH <mark>L</mark> R <mark>L</mark>	FRCRGTE	WEERVRP	T <mark>V</mark> TGIVTDYL	:	47
				_			_		_		
	80		* _1	.00	_ *	_ 120		*	140		
Thiocillin	: KGFFYVF <mark>K</mark>	YSKDFPII	dvyinn <mark>k</mark> iv	TENQ <mark>L</mark> NK	ILQNSGAK	YKIYKNSI	FNETQGNE	GMLGDKY	LAE <mark>FEKKT</mark> NE	:	106
Nosiheptide	: GAAQPALV	/ <mark>V</mark> QLEVTPG	TDLAEPYA?	RARALAA	GIGLPVQV	AAGRATLV	PLAGSVFA	GAALGP <mark>V</mark>	TRAALAAVCP	:	136
Thiostrepton	: HGGRPFL	LRFASRS-	PSVE <mark>R</mark> RI	KSRI <mark>L</mark> AHV	VGPTIDAG	DVFTYQPY1	NHEHDWIG	GTAGLGL	AEN <mark>F</mark> WTE <mark>T</mark> TP	:	118
Siomycin	: YGGRPFL	RFATRS-	PSVE <mark>R</mark> RI	KSRI <mark>M</mark> EH	IG-SAASD	DPFEHQPY1	NHEHDWIG	GEAGLGL	AEA <mark>F</mark> WTE <mark>T</mark> TP	:	120
GE2270A	: RVNPSRA	IDRDRLL-	AQHRHLA	AAERIDEI	PLLPFYAD	NTLHRAAP	RSRAHVIG	GPAAEEL	FHD <mark>F</mark> HTTASA	:	137
Thiomuracin	: AACPSAGI	TDPQALL-	PLHERLA	ELEGERGI	PLIPWAPD	NTVTAEPP(GI	DTELDRF	LADFYADTTE	:	127
Cyclothiazamycin	: RARPSAA	LDEGALA-	PVHA <mark>R</mark> LA	ELEMETG	PRHPWVP	NTVLERPY	DHRLPV	SLRASE	LAGELSDING	:	117
	+	1.60	-			+	200		+ 00		
Thiogillin	. TOTAT	LOU	* • • • • • • • • • • • • • • • • • • •					ACUIDIO			165
Nagihortida		AFO DDAT	TACADET			OMEET DEC			ACCODDENA	:	102
Mosineptide	· ALLTATER	ALQCRPAL	ASAAL MS		VSAAPGPR	QWEELREG	VPLGELSI		ASSEDENAA	:	209
Ciemucia				CIGVMLA		SIARFG			TITERDECT	:	100
SIOMYCIN CE22702		TRGDRA R			PHLPP	PVAREG		DONAFAR		:	100
GE2270A	· ANEDALCE				E	GG			ASAPG-	:	104
Cuclothiagamucin	· ONFRANCE		STIATOTME		2 VDF		TEDCET ST			:	102
Cyclothiazamycin	: QAFRAILF	(VKACGA -	ST ST N		v P E	IIGGEP	TEKGT 191	RSHADAL		•	102
	0	*	240	*	26	0	*	280	*		
Thiocillin	: RDI HKN	LNNKEY	SKVSETTON	INNRSSLS				ELSDITT	EMKKEMTTD	:	214
Nosiheptide		TRAAATLE	RLVDGVLTC	CEERG		;	PVVSLPAF	OWYEAMR	AAKPAVTELF	:	264
Thiostrepton	: RAKHAORY	EKNRELLR	PRERTIVEC	SEPDGE		L	TDVPELAF	EWLVRLR	DYVPA OKGF		244
Siomycin	: RAKHAORY	EONRNLLR	PRLRALVEC	ODPEGD		L	ADVPELAF	EWAVRLR	GYLPATRKGF	:	246
GE2270A		AARAGALR						AWAGLID	RFADRGDELI	:	240
Thiomuracin	: RAKWOAHY	ERNOEPLT	ARTKALTGA	GEPG				AWLRTIR	ATRDRGRTLI	:	230
Cvclothiazamvcin	: RAAFDDRE	RROETVIC	ERLESVEAA	SDGATE	GDGADRSD	RSDRSEAV	GDVVPFVI	EWAAAVR	HHORIAHPLL	:	255
-1 1 -		~							~		
	300)	*	320	*	340		* _	360		
Thiocillin	: EKG-NLH	FNIELLQK	PGERD			FLEK <mark>SQFH</mark>	KTILNN PE	SNF <mark>M</mark> NK	DINFLGSRLI	:	267
Nosiheptide	: RAGTDLAI	DTEEQPPD	TGPDGKG			-lse <mark>s</mark> afh	RIVEGSDO	LRDFLDR	DPSFLATRLL	:	319
Thiostrepton	: DEG-RFYI	YATPRKAE	TAKLTPSPD	GLYRRPD	VEWLSDLP	EPPVAGI <mark>H</mark>	RAIADNTY	YQGMIRE	DRR <mark>F</mark> LAS <mark>RL</mark> A	:	316
Siomycin	: DEG-RFYI	YATPRKAE	TAKLTPSPD	GLYRRPK	VEWLADLP	EAPVAGIH	RAIADNTY	YQGMIRE	DRR <mark>F</mark> LAS <mark>RL</mark> A	:	318
GE2270A	: AS <mark>G-AL</mark> L	EPAGPDAV	ARP			D <mark>T</mark> AFH	RALRG <mark>N</mark> RI	WHEE <mark>V</mark> LR	SAP <mark>F</mark> RRY <mark>RL</mark> L	:	288
Thiomuracin	: DE <mark>G-RL</mark> SI	GYATDGPS	RPPL			-AAVSPFH	RNLETDEF	W-LALKD	TPA <mark>F</mark> AAY <mark>RL</mark> A	:	281
Cyclothiazamycin	: AS <mark>G</mark> -E <mark>V</mark> SM	GGAARAPR	MPTR			RT <mark>S</mark> E <mark>FH</mark>	AV <mark>I</mark> RSDHO	HEDF <mark>V</mark> RT	DDW <mark>F</mark> ASF <mark>RL</mark> M	:	305
	+	200			400		400		+		
Thiogillin	· mtz			VETTON			420 100000000000000000000000000000000000	DV		375	
Iniocillin Naaibaatida	. TVETTLL	RNLGIQNK	DRYLLCYY1	INII DK		LKDEGKGR	DININ VEDUÇ	KI	:	225 270	
Nosineptide Thiostroptor	· TSLL			ADAP		DOUDDEN				2/U 260	
intostrepton			DRUTLETLI	ARAPEEE		DOVODEN				308 970	
CE2270A		SPT CUNAY	OBALLCHER	ASATECE		MCCA-	EvSG		:	270 222	
Thiomuracia	· INCAVI									210	
Cuelethieremusi-	. LINCALLE	NDT OT VDV			RAVAS	FORVIN)19)75	
Cycrotniazamycin	: MIN L	MATCH VDA			ng <mark>vDa</mark> vGS	FORIVASV	DESSERAF	WKKIGEA	WAAGGAG :	515	

LC/MS Figure 1: Extracts from MKO cultures.



Compound	RT	Expected (M+H) ⁺	Observed $(M+H)^{+}$	Error (ppm)	Expected (M+Na) ⁺	Observed (M+Na) ⁺	Error (ppm)
9 (N-succ-Gly-Ala)	9.03	1421.3447	1421.3444	0.21	1443.3266	1443.3365	-6.86
10 (N-succ-Ala)	9.29	1364.3232	1364.3254	1.61	1386.3054	1386.3154	-7.21
11 (N-term ketone)	9.64	1194.2541	1194.2534	-0.59	1216.236	1216.2433	-6.00
12 (N-term hydroxyl)	9.38	1196.2697	1196.27	0.25	1218.2516	1218.2445	5.83



S8

LC/MS Figure 2: Extracts from *EinMKO* cultures.



980 1000 1020 1040 1060 1080 1100

 1480 1500 1520 1540 1560 1580 1600

LC/MS Figure 3: Extracts from *MKI* cultures.





LC/MS Figure 4: Extracts from T3AinMKO cultures.



Compound	RT	Expected (M+H) ⁺	Observed $(M+H)^{+}$	Error (ppm)	Expected (M+Na) ⁺	Observed (M+Na) ⁺	Error (ppm)
13 (N-succ-Gly-Ala)	9.26	1391.3341	1391.3322	1.37	1413.316	1413.32	-2.83
14 (N-succ-Ala)	9.44	1334.3126	1334.3091	-2.62	1356.2946	1356.2913	2.43
15 (N-term ketone)	9.95	1164.2435	1164.244	0.43	1186.2254	1186.2274	-1.69
16 (N-term hydroxyl)	9.48	1166.2591	1166.2568	-1.97	1188.2411	1188.2388	1.94





Fragment	Observed (M+H) ⁺	Expected (M+H) ⁺	Error (ppm)
Parent- H ₂O	1346.3285	1346.31268	-11.75
1	1193.2908	1193.27008	-17.36
1-H₂O	1175.2807	1175.25952	-18.02
16-2	1118.2197	1118.20167	-16.12
16-2-CH₃OH	1086.194	1086.17546	-17.07
14-2	1052.209	1052.19111	-17.00
5	940.231	940.21797	-13.86
5-H₂O	922.2233	922.20741	-17.23
14-4	883.1745	883.16012	-16.28
14-4-H₂O	865.1616	865.14956	-13.92
14-4-H ₂ O-CH ₃ OH	833.1368	833.12334	-16.16
14-6+NH ₂	799.1502	799.13899	-14.03
14-6+NH₂-CH₃OH	767.1225	767.11278	-12.67
12-6	547.1353	547.12504	-18.75
12-6-CH₃OH	515.1054	515.09883	-12.75
11-H₂O	376.0943	376.08964	-12.39
16-12	319.0338	319.03179	-6.30
6-2	254.0625	254.05939	-12.24
6-2-CO	226.0664	226.06447	-8.54



— 515.1065 547.1331

500

600

1346.3319

1346.3350

•

865.1653

833.1386

1770

833

900

1000

767.1261 - 799.1551

800

700

Counts vs. Mass-to-Charge (m/z)

1118.2210

1086.1957

1100

175.2799

1200

1300

1400

. •



x10³

3.5-3.

2.5

2-

1.5-

1-

0.5 0.

100

Compound 10, CID: 50 V

— 226.0684 254.0625

200

319.0338 376.0943

300

400



MSMS Figure 2: Compound **11** (1194.2541).

Fragment	Observed (M+H)*	Expected (M+H)*	Error (ppm)
Parent-H₂O	1176.2572	1176.24354	-11.61
Parent-H ₂ O-CH ₃ OH	1144.2334	1144.21732	-14.05
14	1119.1996	1119.18569	-12.43
14-CH₃OH	1087.1707	1087.15947	-10.33
12+NH ₂	1053.1847	1053.17512	-9.10
3	940.2309	940.21797	-13.75
3-H ₂ O	922.2191	922.20741	-12.68
12-2-H₂O	865.1602	865.14956	-12.30
12-2-H ₂ O-CH ₃ OH	833.1347	833.12334	-13.64
12-4+NH ₂	799.147	799.13899	-10.02
12-4+NH ₂ -CH ₃ OH	767.1235	767.11278	-13.97
10-4	547.1332	547.12504	-14.91
10-4-CH₃OH	515.1031	515.09883	-8.29
9-H ₂ O	376.0942	376.08964	-12.12
14-10	319.0349	319.03179	-9.75
4-CO	227.0501	227.04849	-7.09



Table of 1H & 13C NMR shifts (ppm) for Compound 10 (1364.3232) (600 MHz, DMSO- $d6$).	
(⁴¹	

58	42 41 48 0 39 ^S 38 35	65 \ 25 \	22 40 15)H <mark>64</mark>
$HO_{2}C \xrightarrow{62}{0} M \xrightarrow{55}{55} H \xrightarrow{55}{55} H \xrightarrow{50}{50} H$	H 45 N H 43 N 37 N	31 S 30 24 23 S	21 H 16 N 14	$ \begin{pmatrix} N & 11 \\ 10 & N & 6 \\ H & H & 0 \\ \end{pmatrix} $	1
63 61 61 H	0 /46 34 ^a 33		Ш О _{Н^{18b} Н^{18a}}	5 1 <u>2</u> 9	

Shift (ppm)	Integration	Proton	HSQC	COSY1	COSY2	TOCSY1	TOCSY2	HMBC1	HMBC2	HMBC3
10.41	1H	19	N-H	х	х	6.43	x	166.3	х	х
10.05	1H	43	N-H	х	х	6.70	x	131.0	х	х
9.64	1H	55	N-H	x	х	6.03	x	173.2	х	х
9.53	1H	9	N-H	х	х	1.69	6.51	159.9	х	х
8.86	1H	48	N-H	4.46-4.48	х	4.29	x	х	х	х
8.75	1H	35	N-H	5.08	x	0.86	0.95	161.2	х	х
8.60	1H	59	N-H	4.42-4.44	х	1.29	x	173.3	х	х
8.53	1H	12 (Thz 12)	126.0	x	х	х	x	162.0	151.2	х
8.45	1H	15 (Thz 11)	120.4	x	х	х	x	166.4	147.6	х
8.43	1H	22 (Thz 9)	126.7	x	х	х	x	172.9	149.3	х
8.38	1H	30 (Thz 7)	126.2	х	х	х	x	173.8	148.3	х
8.36	1H	27	N-H	5.47	x	4.34-4.36	x	161.0	56.2	х
8.32	1H	51 (Thz 2)	125.9	x	х	х	x	165.5	150.0	х
8.20	1H	38 (Thz 5)	124.9	x	х	x	x	168.6	149.8	х
7.94	1H	4	N-H	3.00-3.03	3.07-3.11	0.99	3.68-3.71	165.0	х	х
6.70	1H	41	126.9	1.67	х	10.05	x	168.0	131.0	14.2
6.51	1H	7	128.7	1.58-1.62	х	9.53	x	165.0	131.5	13.8
6.43	1H	18a	104.0	x	х	10.41	x	166.4	134.2	х
6.03	1H	54a	105.9	x	х	9.64	x	165.5	134.9	х
5.73	1H	18b	104.0	x	х	х	x	166.4	х	х
5.60	1H	54b	105.9	x	х	х	x	165.5	х	х
5.47	1H	24	56.2	4.34-4.36	8.36	1.28	x	16.4	х	х
5.08	1H	32	57.6	2.42-2.46	8.75	0.86	0.95	173.8	х	х
4.46-4.48	1H	45	60.8	4.29	8.86	1.22	х	х	х	х
4.42-4.44	1H	57	50.1	1.29	8.6	х	x	173.7	18.0	х
4.34-4.36	1H	25	77.6	1.28	5.47	8.36	x	х	х	х
4.29	1H	46	67.2	1.22	4.46-4.48	8.86	x	х	х	х
3.68-3.71	1H	2	65.8	3.00-3.03	3.07-3.11	7.94	x	21.6	х	х
3.07-3.11	1H	3a,b	47.6	3.68-3.71	7.94	0.99	x	165.0	х	х
3.00-3.03	1H	3a,b	47.6	3.68-3.71	7.94	0.99	x	165.0	х	х
3.27	3H	65	57.5	х	х	х	x	х	х	х
2.68	1H	62a,b	27.0	2.25-2.29	x	x	x	172.0	34.0	х
2.54	1H	62a,b	27.0	2.25-2.29	х	х	x	172.0	34.0	х
2.42-2.46	1H	33	32.7	0.86	0.95	x	x	173.6	58.0	40.3
2.25-2.29	2H	61a,b	33.9	2.54	2.68	х	x	173.3	27.0	х
1.69	3H	8	13.8	6.51	x	x	x	130.5	128.3	х
1.67	3H	42	14.2	6.70	x	x	x	68.0	29.9	11.5
1.29	3H	58	18.1	4.42-4.44	x	8.60	x	173.2	50.1	х
1.28	3H	26	16.4	4.34-4.36	х	5.47	x	77.6	56.2	х
1.22	3H	47	21.4	4.29	x	4.46-4.48	x	60.8	67.2	х
0.99	3H	1	21.6	3.68-3.71	х	3.00-3.03	3.07-3.11	65.8	47.6	х
0.95	3H	34a,b	19.2	2.42-2.46	x	5.08	8.75	57.7	32.7	20.1
0.86	3H	34a,b	20.1	2.42-2.46	x	5.08	8.75	57.6	32.9	19.2

Carbon Shift (ppm) 1 21.6 2 65.8 3 47.6 165.0 5 6 131.5 7 128.7 8 13.8 10 159.9 151.2 11 12 126.0 13 162.0 147.6 14 15 120.4 16 166.4 17 134.2 104.0 18 166.3 20 21 149.3 126.7 22 23 172.9 24 56.2 25 77.6 26 16.4 28 161.0 29 148.3 30 126.2 173.8 31 32 57.6 33 32.7 34a 19.2 34b 20.1 36 161.2 37 149.8 124.9 38 39 168.6 131.0 40 126.9 41 42 14.2 45 60.8 46 67.2 47 21.4 50 150.0 51 125.9 52 165.5 53 134.9 54 105.9 56 173.2 57 50.1 58 18.1 60 173.3 61 34.0 27.0 62 63 172.0

65

57.5

a Carbon shifts are based on 2D HSQC/HMBC data. We tentatively assign the stereochemistry at carbons 24, 25, 32, 45, 46, and 57 on analogy to that present in Micrococcin P1. Additionally, we note that this is the stereochemistry of the natural amino acids encoded by the structural genes, tclE-H. The strong TOCSY between acetamide N-Hs 9 and 43 with their respective vinylic protons, 7 and 41, indicates a *cis*-double bond geometry at these positions, as illustrated. Due to bond angle constraint, correlations between Dhb/Dha N-Hs and *cis*-vinylic protons do not typically appear in a 2D TOCSY.

NMR Figure 1: Compound 10 (1364.3232) 1H NMR (600 MHz, DMSO-d6)















NMR Figure 6: Compound **11** (1194.2541) 1H NMR (600 MHz, DMSO-*d*6)