

SUPPLEMENTARY MATERIAL

Unearthing a Cryptic Biosynthetic Gene Cluster for the Piperazic Acid-Bearing Depsipeptide Diperamycin in the Ant-dweller *Streptomyces* sp. CS113

Coral García-Gutiérrez^{1,2}, Ignacio Pérez-Victoria³, Ignacio Montero^{1,2}, Jorge Fernández-De la Hoz,¹ Mónica G. Malmierca¹, Jesús Martín³, José A. Salas,^{1,2} Carlos Olano^{1,2}, Fernando Reyes³, and Carmen Méndez^{1,2,*}

¹ Departamento de Biología Funcional e Instituto Universitario de Oncología del Principado de Asturias (I.U.O.P.A), Universidad de Oviedo, 33006 Oviedo, Spain

² Instituto de Investigación Sanitaria de Asturias (ISPA), 33011 Oviedo, Spain

³ Fundación MEDINA, Centro de Excelencia en Investigación de Medicamentos Innovadores en Andalucía, 18016 Granada, Spain

* Correspondence: cmendezf@uniovi.es

Table S1. Predicted functions of gene products from the *Streptomyces* sp. CS113 diperamycin biosynthesis gene cluster (*dpn*)

Gene product	aa	Predicted function	<i>Streptomyces tendae</i>	<i>Streptomyces</i> sp. NRRL F-5650	<i>Streptomyces</i> sp. PSAA01
DpnL	286	Fatty acyl-AMP ligase	WP_282090654 (97.37)	WP_051852084 (92.48)	WP_237510125 (86.47)
DpnZ	214	Piperazate synthase	WP_282090653 (95.77)	WP_031046733 (88.32)	MCG0283772 (84.11)
DpnA	266	Amidinotransferase	WP_282090652 (98.5)	WP_031046736 (93.98)	WP_237510127 (88.68)
DpnB1	261	Thioesterase	WP_282090651 (94.25)	WP_031046739 (91.94)	WP_237510128 (84.12)
DpnS1	1081	NRPS	WP_282090685 (95.93)	WP_199807506 (90.79)	WP_272933258 (82.69)
DpnO1	517	FAD-dependent monooxygenase	WP_282090650 (98.44)	WP_234332323 (97.29)	WP_272933259 (92.0)
DpnP4	1045	Type I PKS	WP_282090684 (ND)	WP_037836816 (ND)	WP_237510131 (84.3)
DpnP3	2189	Type I PKS	WP_282090925 (ND)	WP_037837542 (ND)	NI
DpnP2	1871	Type I PKS	WP_282090926 (ND)	WP_234332486 (93.57)	NI
DpnP1	1036	Type I PKS	WP_282090898 (93.39)	WP_037837494 (92.95)	WP_237510737 (83.51)
DpnR1	1047	BTAD domain-containing transcriptional regulator	WP_282090897 (97.13)	WP_031050885 (95.8)	WP_237510738 (87.28)
DpnT1	551	Putative exporter	WP_282090896 (97.51)	WP_199807543 (94.83)	WP_237510739 (86.21)
DpnT2	323	ABC transporter ATP-binding protein	WP_282090895 (99.04)	WP_234332468 (96.59)	WP_237510740 (85.54)
DpnB2	238	Thioesterase	WP_282090894 (98.74)	WP_234332470 (98.32)	WP_237510741 (80.67)
DpnO2	435	Ornithine N-monooxygenase	WP_282090893 (98.39)	WP_031050897 (97.24)	WP_237510742 (87.13)
DpnS4	1266	NRPS	WP_282090892 (ND)	WP_234332471 (92.05)	WP_237510743 (88.39)
DpnS3	2569	NRPS	WP_282085203 (ND)	NI	WP_237510744 (83.91)
DpnS2	2888	NRPS	WP_282085205 (97.33)	WP_031042662 (94.91)	WP_272933399 (87.08)
DpnO3	395	NAD(P)/FAD-dependent oxidoreductase	WP_282085206 (99.49)	WP_031042659 (96.46)	WP_237510745 (92.41)
DpnR2	227	LmbU family transcriptional regulator	WP_282085207 (96.04)	WP_031042656 (94.59)	WP_272933409 (83.48)
DpnM	72	MbtH family protein	WP_282085209 (98.57)	WP_031042654 (98.61)	WP_237510747 (87.50)
DpnC	460	Crotonyl-CoA carboxylase/reductase	WP_282085210 (98.91)	WP_051851967 (97.17)	WP_237510748 (92.39)
DpnT3	273	ABC transporter ATP-binding protein	WP_282085212 (98.90)	WP_051851966 (96.34)	WP_237510749 (87.50)
DpnT4	253	ABC transporter permease	WP_282085214 (99.21)	WP_031042645 (98.81)	WP_272933410 (93.36)
DpnK1	274	SDR family oxidoreductase	WP_282085940 (99.64)	WP_031042642 (97.81)	MCG0284487 (90.51)
DpnQ	902	NRPS	WP_282085216 (96.78)	WP_037836225 (91.01)	WP_237510751 (81.58)
DpnR3	205	LmbU family transcriptional regulator	WP_282085218 (98.86)	WP_051851965 (96.59)	WP_272933400 (87.50)

Numbers between brackets indicate percentage of identical amino acids to the corresponded Dpn protein. ND, not determined. NI, not identified. aa, number of amino acids of Dpn proteins.

Table S2. Oligonucleotides used for PCR amplification

PRIMER	SEQUENCE 5'-3'
C9C2	<u>GATATCATGCCGTCTCGCACTCCC</u>
C9C1	<u>GAATTCCTAGGCGAAGAGGTCCAAGG</u>
113C1orf18-indC-F	AAAAT CTAGAG AGCCGATACCAGATGGAG
113C1orf18-indC-R	AAA AGATAT CGCTGTCTCCAAATCGGCA
113C1orf18-indC-comp-R	CGTTGTAGGTCCGTGACG
M13-F	GTAAACGACGGCCAG
113C1orf18A-F	<u>CTTGACATTGGGGAATTGACGTCGACATGAAGGTCACG</u>
113C1orf18A-R	<u>CGGGGACCTGCAGGTCGACTCAGGAAGTTCGCCAGTGC</u>
113C1orf18B-F	<u>CCGCCGAAAGTTCCTCGAAGGCGAGCTGACCAATTCGTC</u>
113C1orf18B-R	<u>GGCATCAGTTACCGTGAGCGTACGGAGTACCTCGTGC</u>
113C1orf18-comp-F	TGATCGACCGGTACTTCTCG
113C1orf18-comp-R	CAACGGGATCATCGACGCTC
NRPSint113_I_RV	CGAGGACGTCGATGG
NRPSint113_D_FW	GCACGACAGGTGTTC
113C1orf2A-F	<u>CTTGACATTGGGGAATTGACGGCAACAGTCCTTCGCCGACAG</u>
113C1orf2A-R	<u>CGGGGACCTGCAGGTCGACTCACGGGCACACAGTTCGTAG</u>
113C1orf2B-F	<u>CCGCCGAAAGTTCCTCGAAGGAGAACGACGACTTGATGGC</u>
113C1orf2B-R	<u>GGCATCAGTTACCGTGAGCGGACTGCGAGTGATCAGTTC</u>
113C1orf2-comp-F	GGATCCGCACTGACGCTGGACAACAG
113C1orf2-comp-R	GGATCCTGGGCAGATCACGAGAGTGC
113C1orf15A-F	<u>CTTGACATTGGGGAATTGACGAGTGGGGAGATCGTGGAATC</u>
113C1orf15A-R	<u>CGGGGACCTGCAGGTCGACTTGACGCATGAACAGGTCTC</u>
113C1orf15B-F	<u>CCGCCGAAAGTTCCTCGAAGCGGCTTTCTGGCAGATGTTG</u>
113C1orf15B-R	<u>GGCATCAGTTACCGTGAGCGGGACCGAAACCGATACCG</u>
113C1orf15-comp-F	GGATCCACGCCACACATGACGAAATG
113C1orf15-comp-R	GGATCCTCAACTCCGATCACTGGTCC
113C1orf2-OE-Tc-F	<u>ACAATCGTGCCGGTTGGTAGGCACTGACGCTGGACAACAG</u>
113C1orf2-OE-Tc-R	<u>ACAGCTATGACATGATTACGTGGGCAGATCACGAGAGTGC</u>
ApraGib-F	GCTCACGGTAACTGATGCC
ApraGib-R	CGTCAATTCCCAATGTCAAG

* Restriction sites are in bold; nucleotide sequences for Gibson assembly of DNA are underlined.

Table S3. NMR data of diperamycin in DMSO-d₆ (500 MHz, 24 °C)

Position	δ_c , type	δ_H , mult. (J in Hz)	Position	δ_c , type	δ_H , mult. (J in Hz)
L-Thr			L-Pip		
CO	167.7, C		CO	172.0, C	
C $_{\alpha}$	48.2, CH	5.56, br s	C $_{\alpha}$	48.2, CH	5.57, br s
C $_{\beta}$	71.4, CH	4.76, br s	C $_{\beta}$	24.4, CH ₂	1.84, m (ov.) 2.14, m (ov.)
C $_{\gamma}$	15.9, CH ₃	1.24, d (6.2)	C $_{\gamma}$	20.9, CH ₂	1.51, m (ov.)
NH		8.12, d (9.0)	C $_{\delta}$	46.5, CH ₂	2.66, m (ov.) 3.02, br d (11.5)
			NH		4.94, br d (12.5)
D-Pip			N-OH-D-Ala		
CO	174.5, C		CO	169.3, C	
C $_{\alpha}$	48.4, CH	5.33, dd (6.7, 2.8)	C $_{\alpha}$	53.4, CH	5.07, br m
C $_{\beta}$	24.3, CH ₂	1.84, m (ov.) 2.07, m (ov.)	C $_{\beta}$	13.5, CH ₃	1.24, d (6.4)
C $_{\gamma}$	20.6, CH ₂	1.51, m (ov.)	NOH		10.29, s
C $_{\delta}$	46.3, CH ₂	2.67, m (ov.) 3.02, br d (11.5)			
NH		5.00, br d (11.9)			
N-OH-β-MeO-L-Ala			Polyketide		
CO	167.2, C		1	177.2, C	
C $_{\alpha}$	57.3, CH	5.14, dd (8.1, 5.5)	2	75.6, C	
C $_{\beta}$	67.1, CH ₂	3.70, dd (10.8, 5.5) 3.75, dd (10.8, 8.1)	3	98.8, C	
CH ₃ O	57.9, CH ₃	3.26, s	4	27.0, CH ₂	1.53, m (ov.) 1.61, m (ov.)
NOH		10.29, s	5	23.7, CH ₂	1.32, m (ov.) 1.62, m (ov.)
Gly			6	41.1, CH	0.99, m (ov.)
CO	168.6, C		7	70.0, CH	3.54, dq (8.0, 6.3)
C $_{\alpha}$	41.5, CH ₂	3.81, dd (17.8, 4.3) 4.20, dd (17.8, 3.5)	8	20.2, CH ₃	1.24, s
NH		6.97, m (ov.)	9	31.2, CH ₂	0.95, m (ov.) 1.32, m (ov.)
			10	25.7, CH ₂	1.15, m (ov.) 1.32, m (ov.)
			11	29.1, CH ₂	1.23, m (ov.)
			12	31.2, CH ₂	1.24, m (ov.)
			13	22.1, CH ₂	1.26, m (ov.)
			14	14.0, CH ₃	0.86, t (6.3)

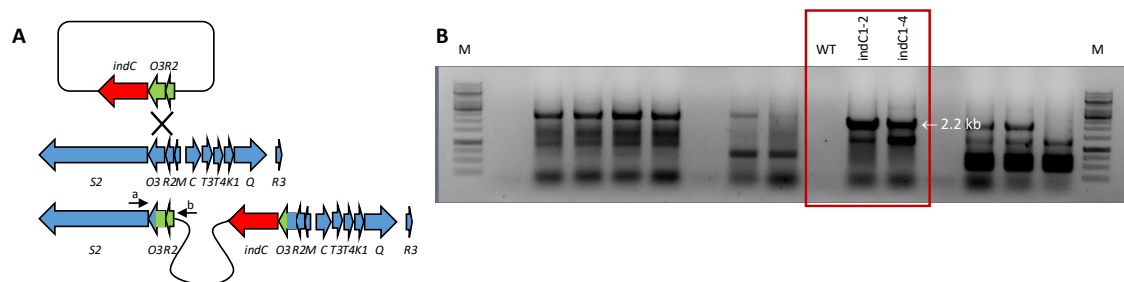


Figure S1. Generation of CS113R-indC1 strain. (A) Graphical representation of the generation of the recombinant strain; (B) PCR analysis of CS113R-indC1 mutant strain (lanes framed in red). PCR products from the wild type (WT) strain and from the CS113R-indC1 mutant strain (indC1) using oligonucleotides 113C1orf18-indC-comp-R (a) and M13-F (b). indC1-2 and indC1-4 correspond to two independent mutant strains. M, 1 kb ladder. *indC*, indigoidine reporter gene.

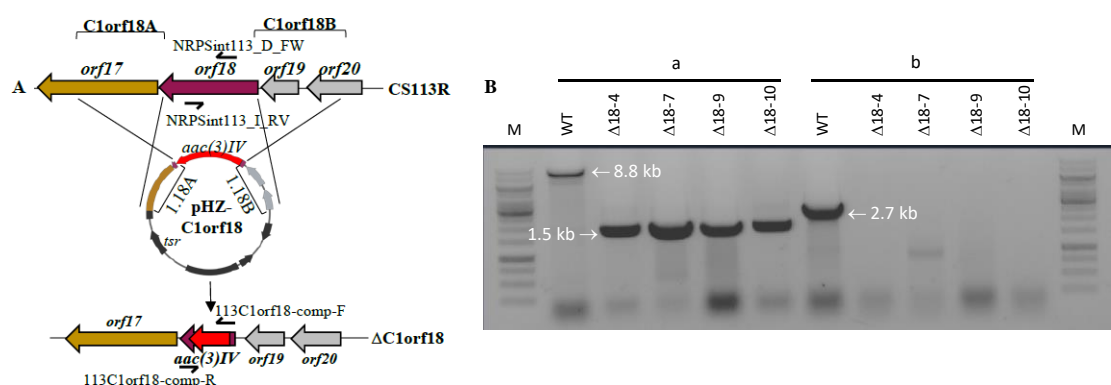


Figure S2. Generation of CS113R- Δ C1orf18 mutant. (A) Graphical representation of the generation of the mutant strain. *Orf17*, *orf18*, *orf19* and *orf20* correspond to *dpnS3*, *dpnS2*, *dpnO3* and *dpnR2*, respectively; (B) PCR analysis of CS113R- Δ C1orf18 mutant strain. PCR products from the wild type (WT) strain and from CS113R- Δ C1orf18 mutant strain (Δ 18) using oligonucleotides 113C1orf18-comp-F and 113C1orf18-comp-R (a); and NRPSint113_I_RV and NRPSint113_D_FW (b). Δ 18-4, Δ 18-7, Δ 18-9 and Δ 18-10 correspond to four independent mutants. M, 1 kb ladder. *aac(3)/IV*, apramycin resistance gene.

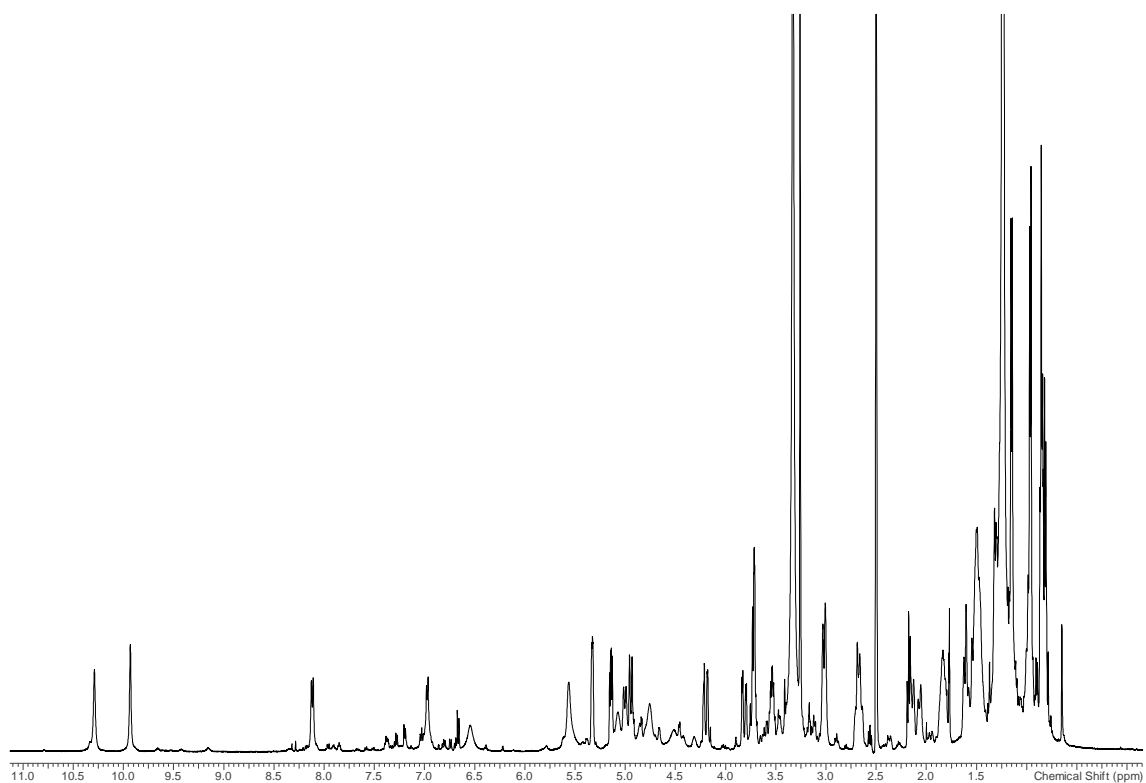


Figure S3. ^1H NMR spectrum of diperamycin ($\text{DMSO-}d_6$, 500 MHz, 24 °C).

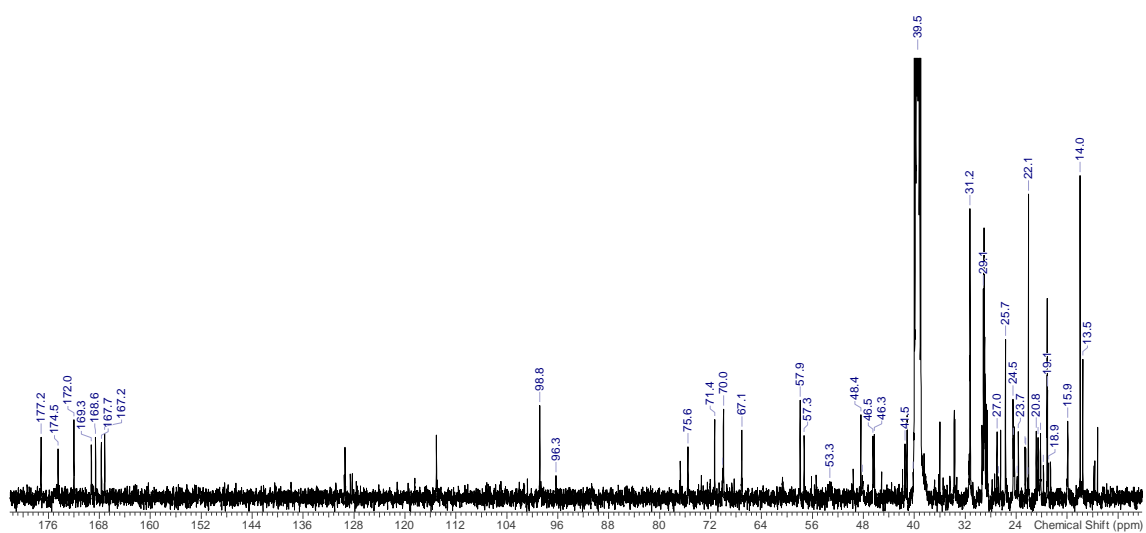


Figure S4. ^{13}C NMR spectrum of diperamycin ($\text{DMSO-}d_6$, 125 MHz, 24 °C).

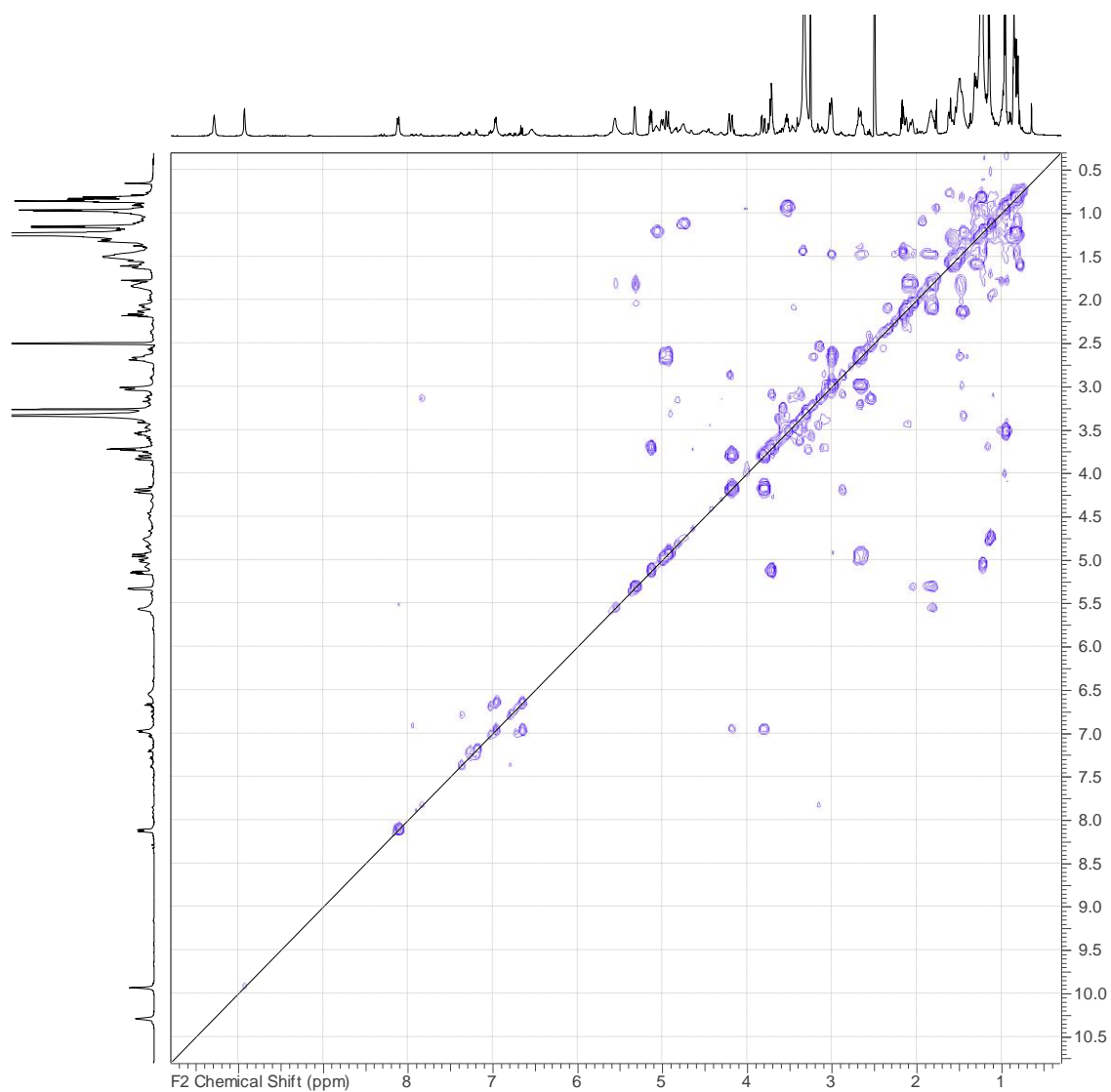


Figure S5. COSY spectrum of dipramycin.

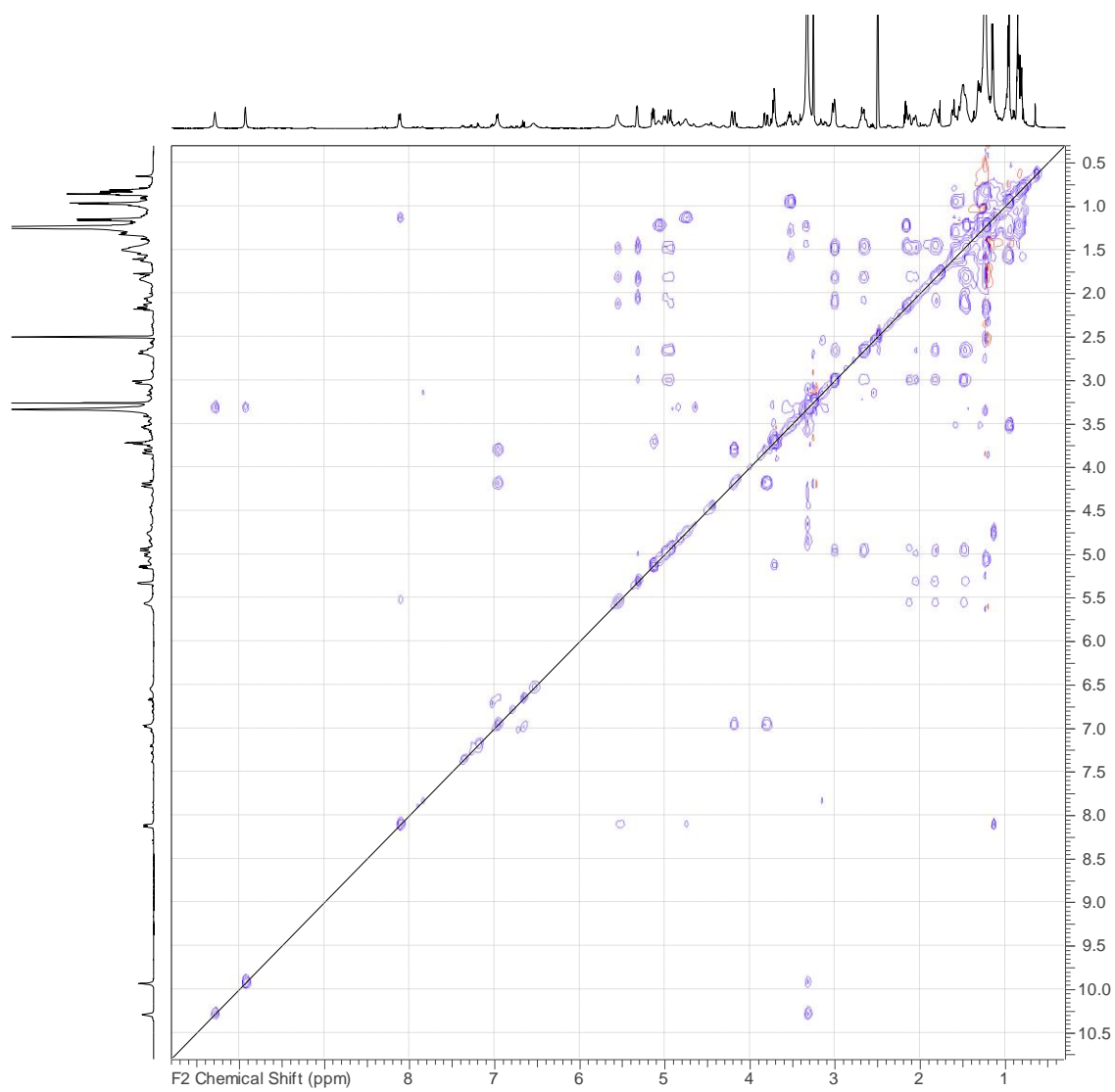


Figure S6. TOCSY spectrum of diperamycin.

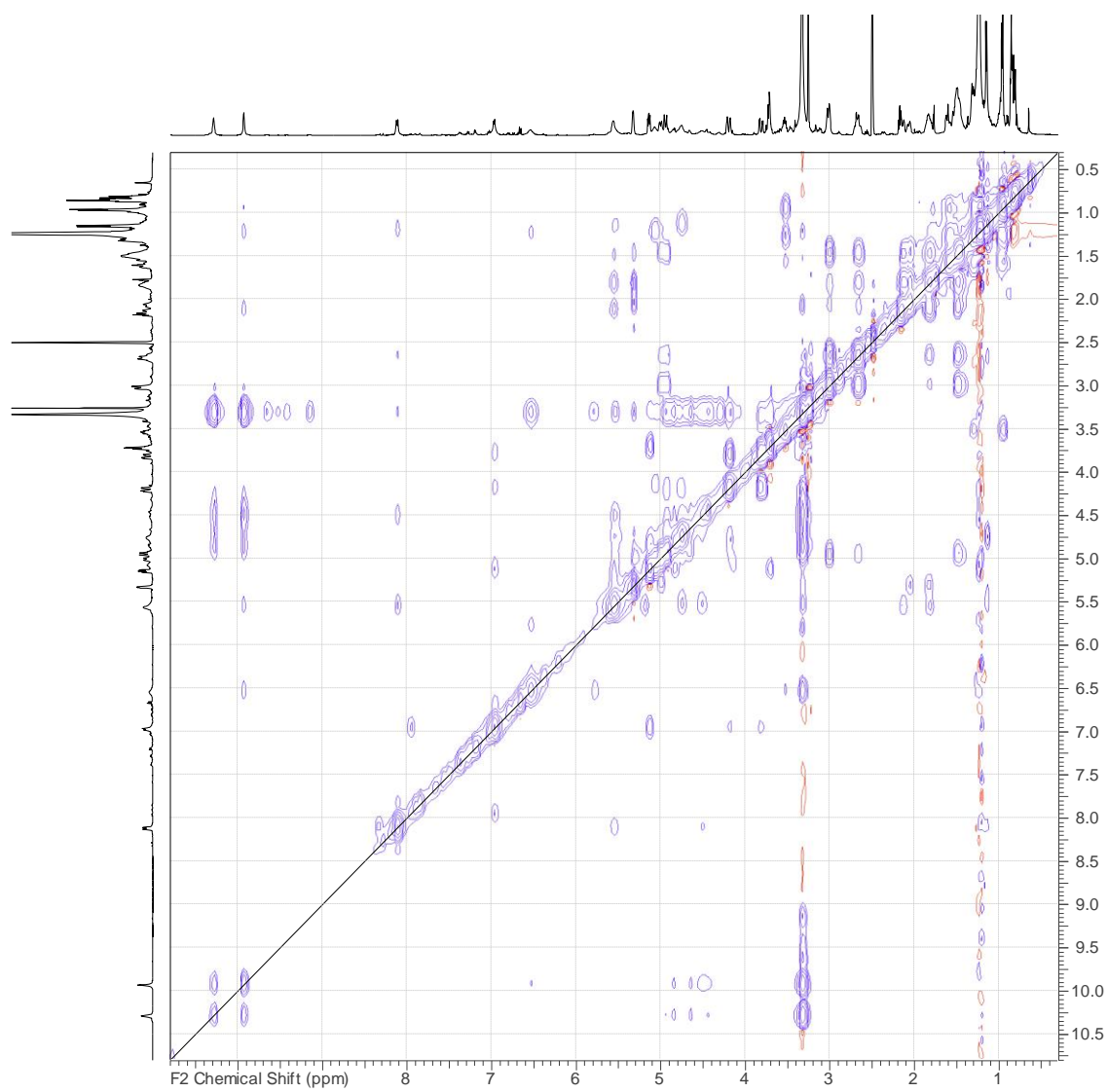


Figure S7. NOESY spectrum of diperamycin.

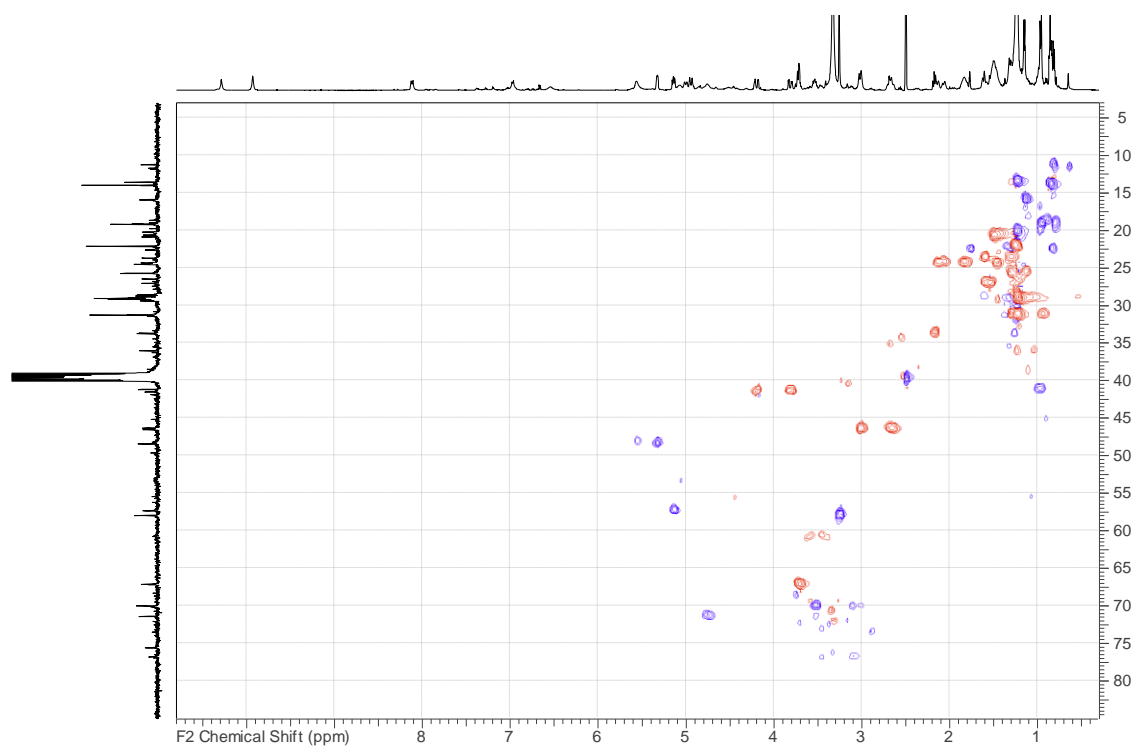


Figure S8. Multiplicity-edited HSQC spectrum of diperamycin.

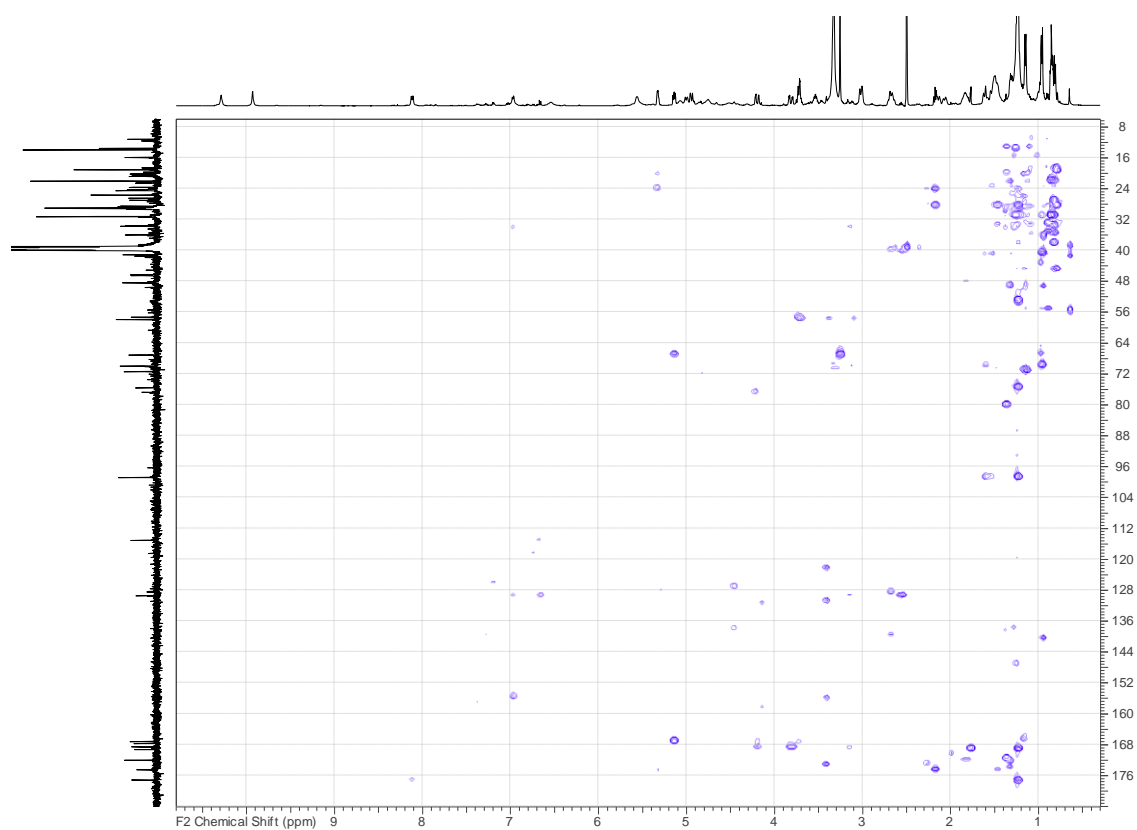


Figure S9. HMBC spectrum of diperamycin

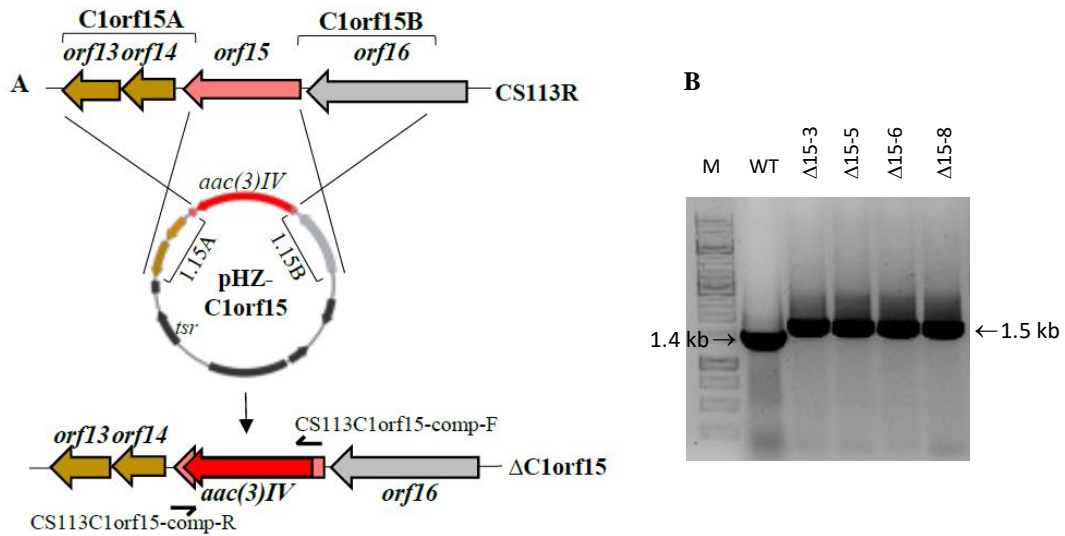


Figure S10. Generation of CS113R- Δ C1orf15 mutant. (A) Graphical representation of the generation of the mutant strain. *Orf13*, *orf14*, *orf15* and *orf16* correspond to *dpnT2*, *dpnB2*, *dpnO2* and *dpnS4*, respectively; (B) PCR analysis of CS113R- Δ C1orf15 mutant strain. PCR products from the wild type (WT) strain and from the CS113R- Δ C1orf15 mutant strain (Δ 15) using oligonucleotides 113C1orf15-comp-F and 113C1orf15-comp-R. Δ 15-3, Δ 15-5, Δ 15-6 and Δ 15-8 correspond to four independent mutants. M, 1 kb ladder. *aac(3)IV*, apramycin resistance gene.

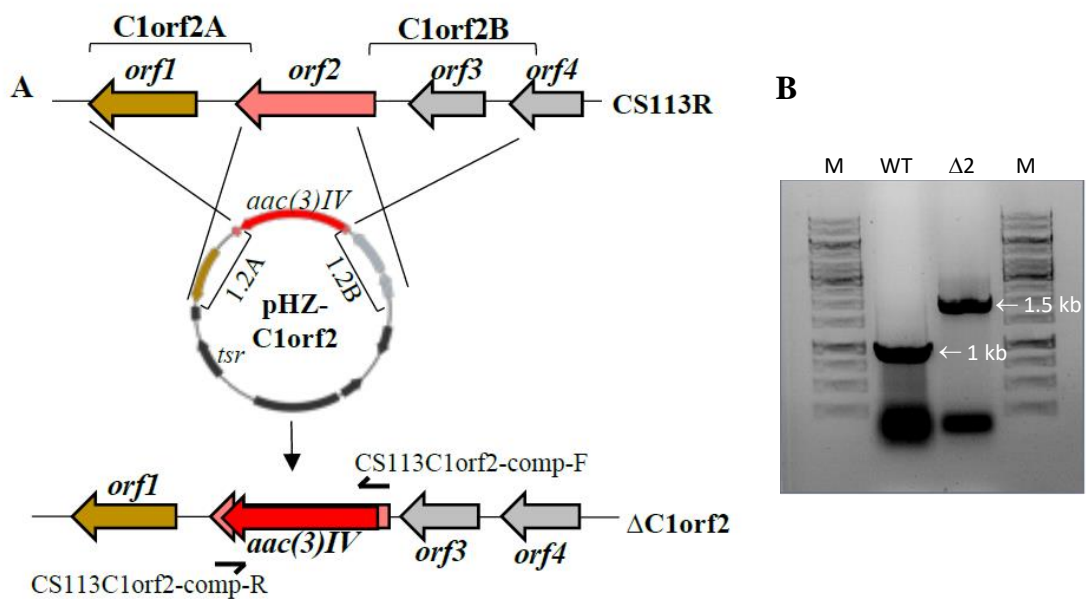
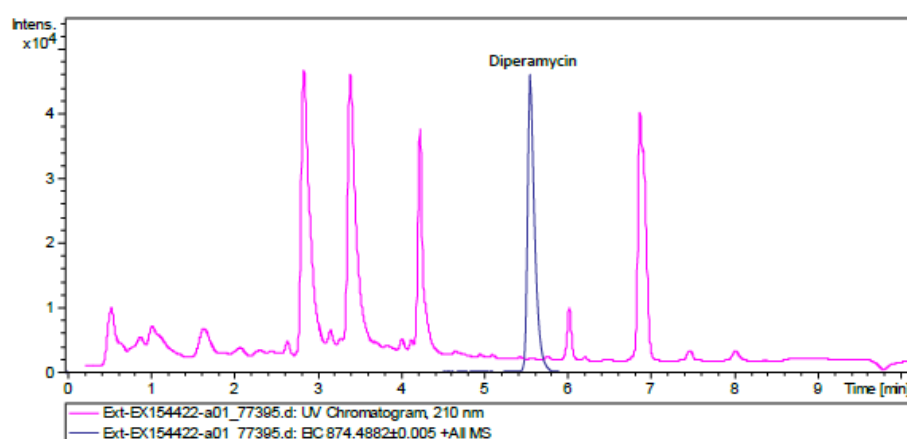


Figure S11. Generation of CS113R- Δ C1orf2 mutant. (A) Graphical representation of the generation of the mutant strain. *Orf1*, *orf2*, *orf3* and *orf4* correspond to *dpnL*, *dpnZ*, *dpnA* and *dpnB1*, respectively; (B) PCR analysis of CS113R- Δ C1orf2 mutant strain. PCR products from the wild type (WT) strain and from the CS113R- Δ C1orf2 mutant strain (Δ 2) using oligonucleotides 113C1orf2-comp-F and 113C1orf2-comp-R. M, 1 kb ladder. *aac(3)/IV*, apramycin resistance gene.



P1

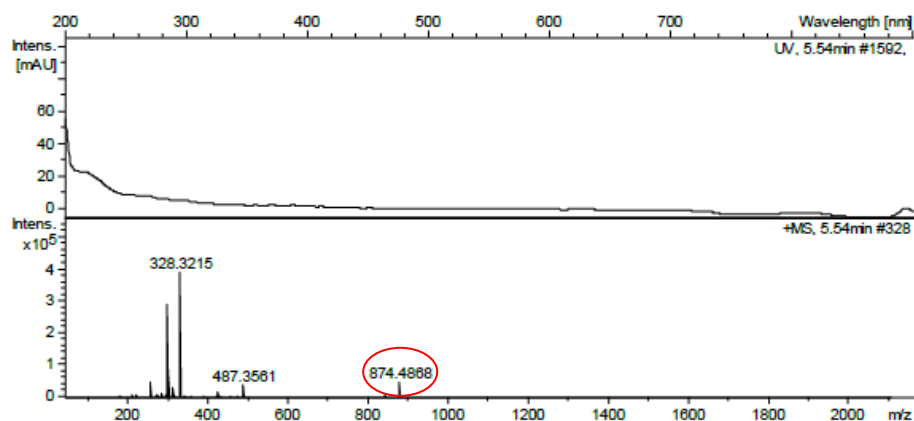


Figure S12. Complementation of CS113R- Δ C1orf2 mutant with pSETETc-C1orf2. Chromatogram at 210 nm and extracted ion ($[M+NH_4]^+$) chromatogram and HRMS spectrum of diperamycin.