Electronic Supplementary Material (ESI) for Chemical Communications.

Supporting Information

Synchronous intramolecular cycloadditions of the polyene macrolactam polyketide heronamide C

Thomas J. Booth[1], Silke Alt[1], Robert J. Capon[2] and Barrie Wilkinson[1*]

[1]Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK

[2]Institute for Molecular Bioscience, The University of Queensland, St. Lucia, Queensland, Australia 4072

Table of Contents **Page**

Figure S1. Formation of heronamide A (**1**) from the hypothetical pyrolinidol (**7**) via [6π + 4π] cycloaddition and Cope rearrangement of the hypothetical [4π + 2π] adduct. [1]

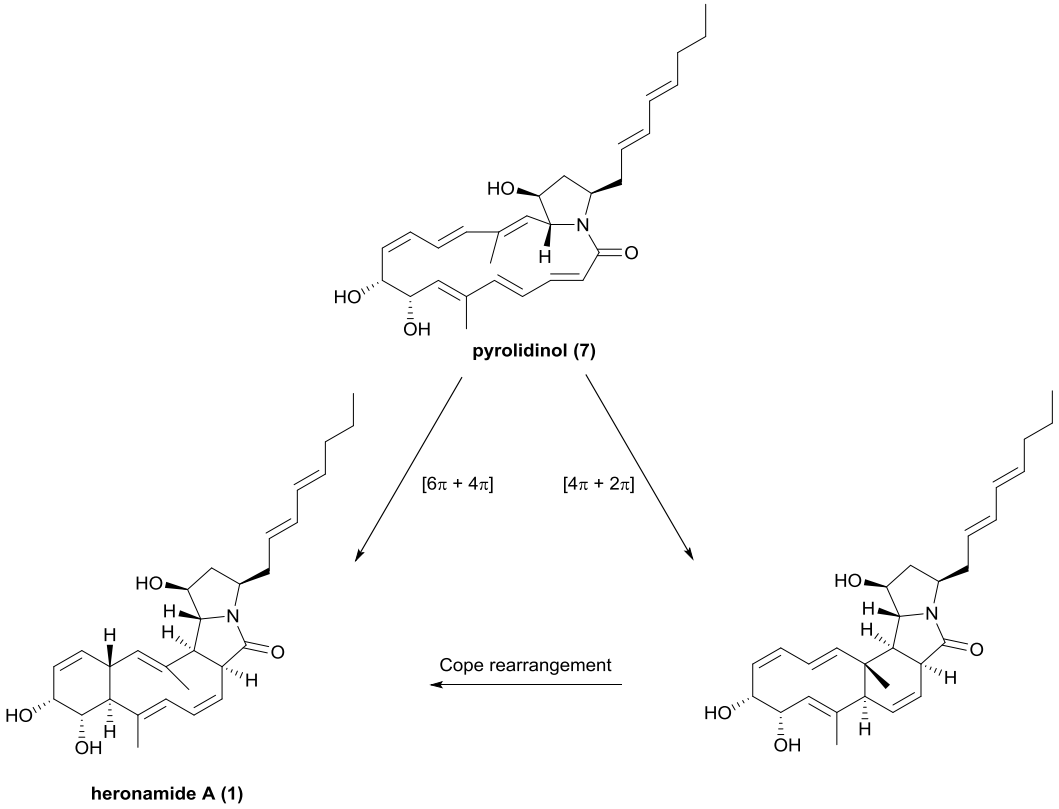


Figure S2. Woodward-Hoffman analysis of heronamide A (**1**) and B (**2**) formation. The Woodward Hoffman rules were applied to heronamide C (**3**), the hypothetical pyrrolidinol intermediate (**7**). As the π-systems are not conjugated these reactions were treated as cycloadditions**.**

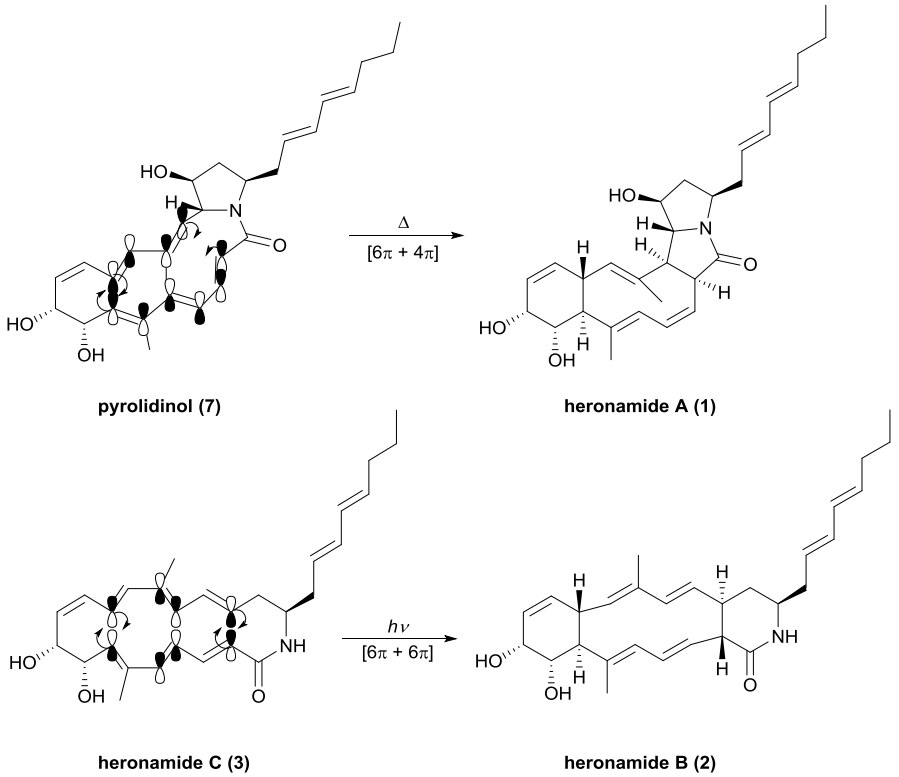


Figure S3. Photochemical conversion of ciromycin A to ciromycin B[2]

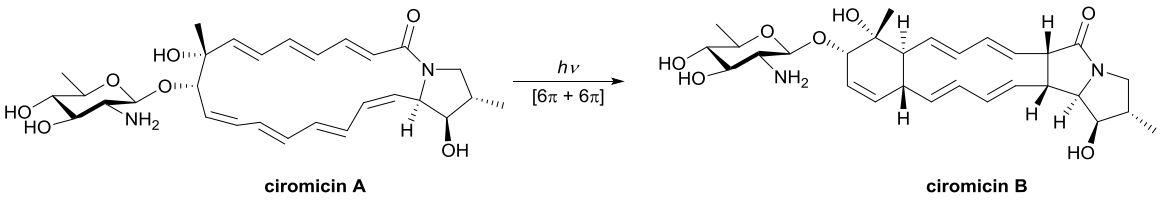
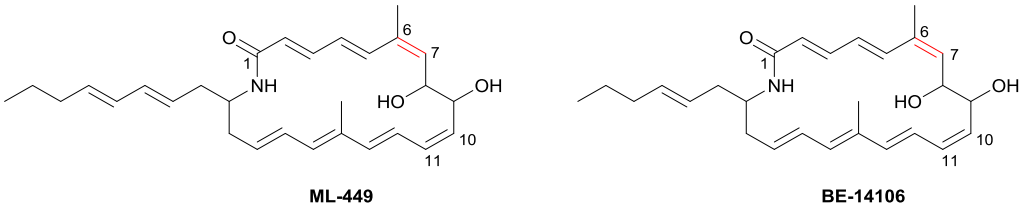


Figure S4. Reported structures of ML-449[3] and BE-14106;[4] the C6-C7 olefin (highlighted in red) is a *trans*-double bond.



General Remarks

Unless otherwise stated, all reagents were supplied by Sigma-Aldrich or Fischer Scientific. All solvents were of HPLC grade or equivalent.

Chemical analysis

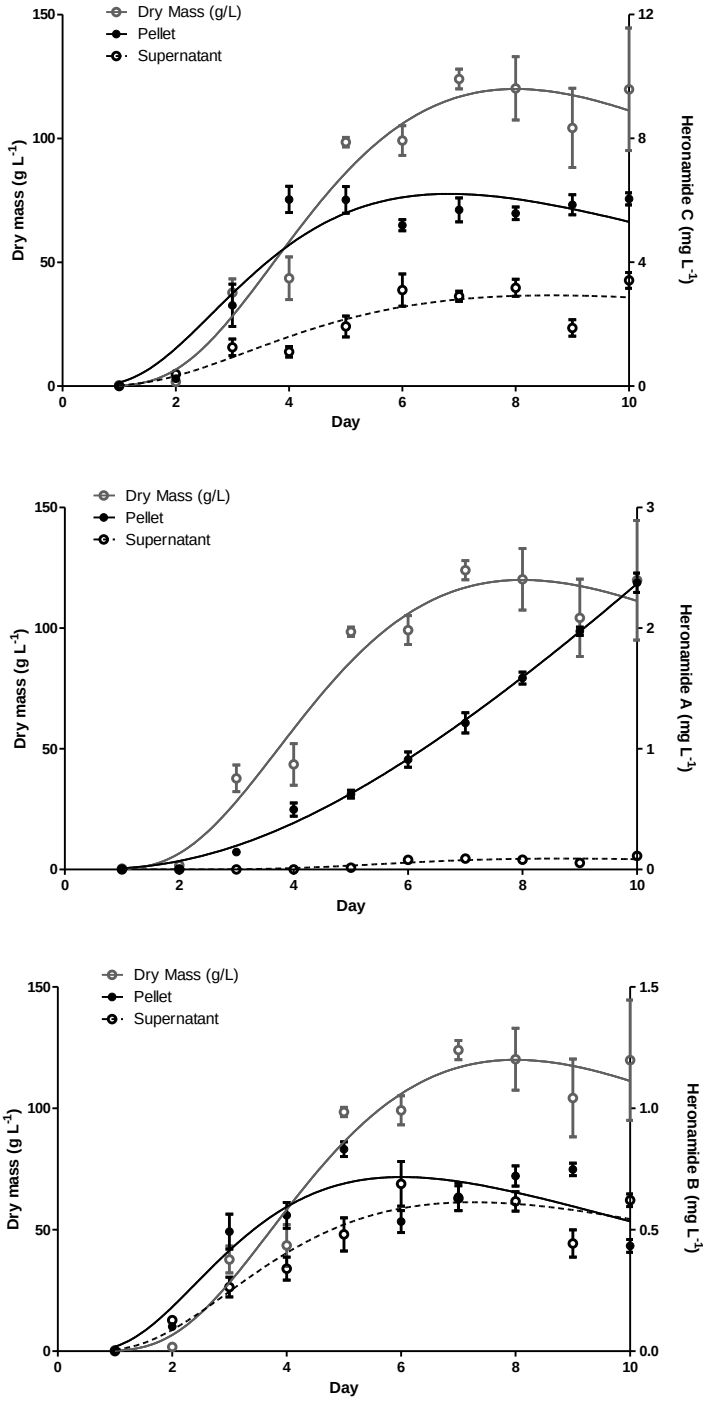
Analytical HPLC was performed on an Agilent 1100 series HPLC system with a Gemini-NX C_{18} 110A column (150 × 4.6 mm, 3 μm; Phenomenex) using a two-step water-MeOH gradient (10-50% MeOH over 3 min followed by 50-100% over 15 min with a 2 min hold at 100%).

Analysis of *Streptomyces sp.* CMB-0406 growth and heronamide production

Cultures of *Streptomyces sp.* CMB-0406 were grown in SV2 medium (1.5 % glucose, 1.5 % glycerol, 1.5 % soya peptone, 0.3 % NaCl, 0.1 % $CaCO_3$; pH 7.0). Mycelial stocks were streaked on SFM agar (soya flour 2.0 %, mannitol 2.0 %, Lab M No 1 agar 2.0%) and incubated until sporulation at 30 °C. Spore suspensions were produced according to the method described by Kieser *et al*.[5]

A single colony of *Streptomyces sp.* CMB-0406 was used to inoculate 250 ml SV2 seed culture. After 5 days, five Erlenmeyer flasks containing SV2 (400 mL in 2 L flasks) were inoculated with 400 μL seed culture. The cultures were incubated at 30°C and shaken at 250 rpm. Aliquots were taken from each culture every 24 h over a 10 day period. A portion of each sample (5 mL) was dried at 60°C before measuring the mass. The remaining material (1 mL) was centrifuged and the mycelial pellet and supernatant were each extracted separately with EtOAc (1 mL) and the solvent removed under reduced pressure. The resulting material was dissolved in methanol (500 μL) and analysed by HPLC as described above. The data from these experiments are plotted below (Figure S5).

Figure S5: Comparison of dry cell weight to heronamide titres during the fermentation of *Streptomyces* sp. CMB-0406.



Production and isolation of heronamides

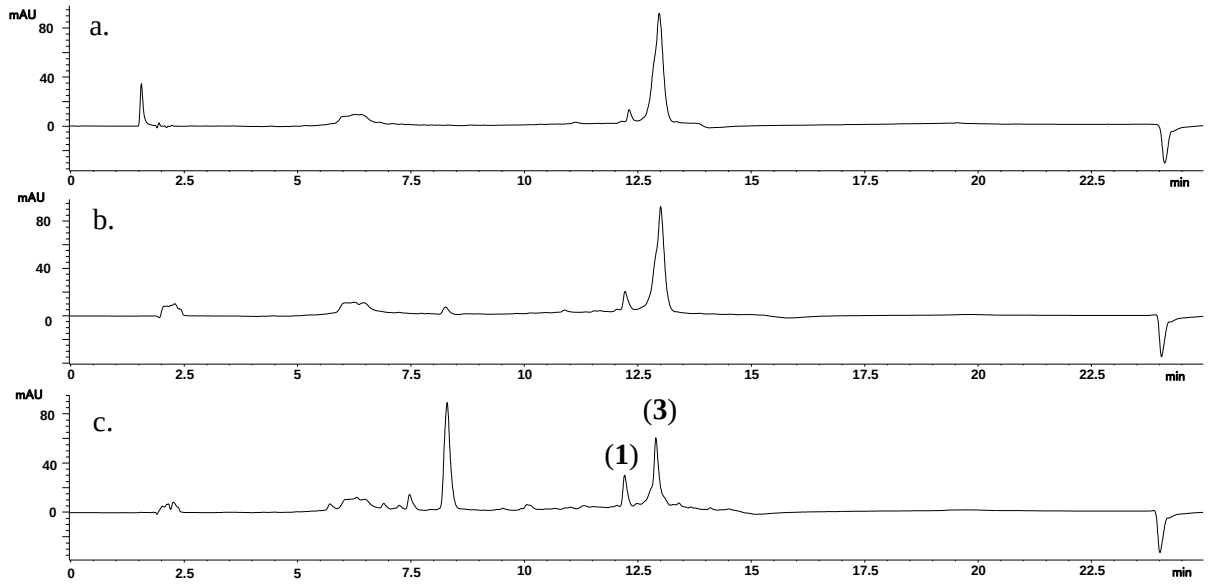
For large scale purification six SV2 cultures (400 mL in 2 L Erlenmeyer flasks) were grown as described above. The resulting broth (2.4 L) was extracted with an equal volume of EtOAc. The phases were separated by centrifugation at 5000 rpm for 15 min and the solvent removed under reduced pressure to yield a yellow extract (1.8 g). This was dissolved in excess methanol (50 mL) and triturated with an equal volume of water. The insoluble fraction (67 mg) was separated and dissolved in methanol and then fractionated by preparative HPLC (see method below) to yield heronamide A (**1**) (1.2 mg), heronamide B (**2**) (0.3 mg) and heronamide C (**3**) (4.8 mg).

Preparative HPLC was performed on an UltiMate 3000 LC system with a Gemini-NX C_{18} 110A column (150 × 21.2 mm, 5 μm; Phenomenex) using a two-step MeOH gradient (10-50% MeOH over 5 min followed by 50-100% over 15 min and a 2 min hold at 100%). Typical injection volumes were between 1-1.75 mL and fractions were collected at 10.1-10.4 min (heronamide A, **1**), 11.5 – 12 min (heronamide C, **3**) and 13-13.4 min (heronamide B, **2**). LCMS was performed on a Shimadzu LC–MS platform (equipped with a NexeraX2 liquid chromatograph (LC30AD), a Prominence photo diode array detector (SPD-M20A) and an LCMS-IT-TOF mass spectrometer) with chromatography over a Kinetex C_{18} 100A column (100 × 2.1 mm, 2.6 μm; Phenomenex) using a water-MeOH gradient (20-100% MeOH over 12 min with 1 min hold at 100%).

Spontaneous transformation of heronamide C (3) to heronamide A (1)

Heronamide C (**3**) was dissolved in methanol containing 10% DMSO at a final concentration of 225 μM while minimizing exposure to light. Aliquots (1 mL) were then incubated at 4°C, 30°C and 60°C for 7 days with all light excluded. Samples were then analyzed by HPLC using the method described above and the molar yields for each component calculated by comparison to UV calibration curve generated with isolated material. The data are shown in Figure S6 below.

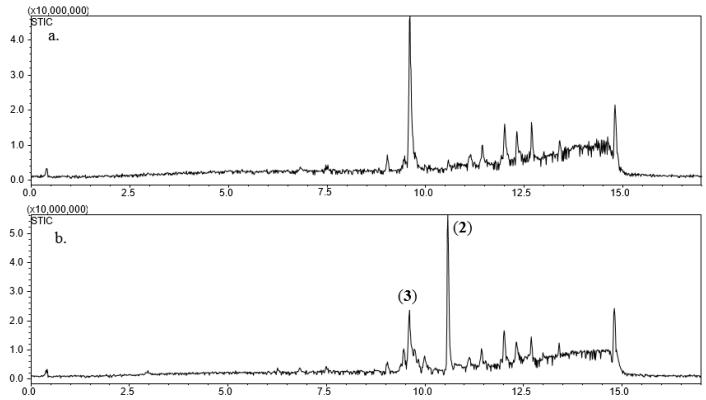
Figure S6: HPLC chromatogram (254 nm) of samples of heronamide C (**3**) following 7 days incubation at (a) 4°C, (b) 30°C and (c) 60°C.



Photochemical conversion of heronamide C (3) to heronamide B (2)

Heronamide C (**3**) was dissolved in methanol at a final concentration of 225 μM while minimizing the exposure to light. Aliquots (100 μL) were then placed into borosilicate vials and incubated under ambient conditions either exposed to bright sunlight by standing on a bench, or placed in a box so that light was excluded and located adjacent to the first sample. After 60 min the samples were analyzed by LCMS and compared to calibration curves of authentic material in order to quantify the species present. Representative chromatograms are shown in Figure S7 below.

Figure S7: TIC of heronamide C (**3**) following 1 h incubation with light excluded (a.) or exposed to ambient light (b.).



In a second set of experiments aliquots of **3** in methanol (1 mL, 225 μM) were placed into quartz cuvettes and exposed to varying wavelengths of UV radiation (330, 375, 405 nm) for a range of exposure times (10, 20, 40, 80, 160 s) using an Atlas Photonics LUMOS 43 photoreactor. The photon flux at each wavelength was calculated using ferioxilate chemical actinometry to allow for normalization across wavelengths. From these data, the half-life, rate constant and quantum yields were determined, the latter using the methods described below:

$$\Phi_{apparent} = \frac{N_{mole}}{N_{photons}}$$

The apparent quantum yield ($\Phi_{apparent}$) was calculated at different time points with the above equation, where N_{mole} represents the number of moles of heronamide B (**2**) at a given time point. $N_{photons}$ represents the number of photons absorbed by heronamide C (**3**) and was calculated by the following equation.

$$N_{photons} = q\,\chi\,t$$

Where *q* is the number of photons delivered to the cuvette per second, χ is the proportion of light absorbed by the sample and t is the exposure time.

Isolation of genomic DNA from *Streptomyces sp.* CMB-0406

High molecular weight genomic DNA was extracted according to the salting out procedure of the *Streptomyces* manual[5] with the modifications as described here: wet mycelium (0.5 mL) from a 30 h old SV2 culture was washed with 10% sucrose (10 mL) before resuspension in SET buffer (5 mL; 75 mM NaCl, 25 mM EDTA, 20 mM TrisHCl pH 8.0) to which lysozyme (200 μL; 50 mg/mL) and ribonuclease A (15 μL; 10 mg/mL) were added. The cells were incubated overnight at 37°C; fresh lysozyme (300 μL) was added after ca. 17 h followed by an additional 2 h incubation. The subsequent steps were performed according to Kieser *et al.*[5]

Genome sequencing, cluster annotation and bioinformatics analysis

Genomic DNA was sequenced with Pacific Biosciences (PacBio) RSII SMRT technology (commissioned to The Genome Analysis Centre (TGAC) Norwich, UK).[6] The assembly obtained five contigs of 8585775 bps, 13002 bps, 11991 bps, 11345 bps, and 10240 bps whereof the biggest contig contained all four smaller contigs. The 8.59 MB contig was therefore submitted to BASys (Bacterial Annotation System), Prodigal to identify protein coding sequences (PCSs), and RAST (Rapid Annotation using Subsystem Technology) to assign putative gene functions.[7,8,9] Putative gene clusters for the biosynthesis of specialised metabolites were identified with antiSMASH 3.0.[10]

Table S1: General characteristics of the *Streptomyces sp.* CMB-0406 genome

Assembled chromosome size	8585775 bp
Estimated chromosome size	8691436 bp
Estimated terminal inverted repeats (TIR)	115994 bp
Coding DNA	7566372 bp
Chromosome topology	Linear
Chromosome G+C content	70.27%
Protein coding sequences (PCS)	7578
rRNA operons	6
tRNA genes	64
Putative biosynthetic gene clusters for specialised metabolites	25

Table S2: COG (Cluster of Orthologous Genes) functional categories of chromosomal protein coding sequences identified in *Streptomyces sp.* CMB-0406 (as calculated by Basys[7]).

	COG functional categories	**Percentage**
C	Energy production and conversion	8.4
D	Cell division and chromosome partitioning	0.7
E	Amino acid transport and metabolism	10.4
F	Nucleotide transport and metabolism	1.8
G	Carbohydrate transport and metabolism	10.2
H	Coenzyme metabolism	3.8
I	Lipid metabolism	5.9
J	Translation, ribosomal structure and biogenesis	4
K	Transcription	8.4
L	DNA replication, recombination and repair	3.3
M	Cell envelope biogenesis, outer membrane	4
N	Cell motility	0.2
O	Posttranslational modification, protein turnover, chaperones	2.7
P	Inorganic ion transport and metabolism	4.8
Q	Secondary structure	3.1
R	General function prediction only	9.5
S	COG of unknown function	3.2
T	Signal transduction	5.9
	Unknown	6.3

Table S3. Putative biosynthetic gene clusters for specialised metabolites in *Streptomyces sp.* CMB-0406[10]

antiSMASH Cluster No.	Gene cluster type	Position		Gene clusters of compounds with highest similarity (%)
		From	To	
1	Nrps	61570	119251	Streptolydigin (10%)
2	Melanin	353228	363716	Melanin (100%)
3	Arylpolyene-Terpene	362554	411888	2-methylisoborneol (100%)
4	Terpene-Otherks-T1pks	429305	487988	Isorenieratene (85%)
5	T3pks	499477	540604	Herboxidiene (2%)
6	Nucleoside	574155	594589	Reveromycin (9%)
7	Nrps	597760	648748	Coelichelin (100%)
8	T1pks	737055	801594	Reveromycin (21%)
9	Terpene	1035427	1056401	-
10	Terpene	1381714	1408492	Hopene (76%)
11	Terpene-Bacteriocin	1481394	1517147	-
12	Other	1743111	1783809	Meilingmycin (5%)
13	Bacteriocin	1893447	1904775	-
14	Siderophore	2370109	2384856	-
15	Other	2744569	2785783	Stambomycin (16%)
16	Nrps-Lantipeptide-T1pks	3967300	4046374	Guadinomine (11%)
17	Terpene	4634133	4656319	Lipstatin (14%)
18	Nrps	6824991	6871418	-
19	Ectoine	6930581	6940979	Ectoine (75%)
20	Lassopeptide	7375142	7397656	SSV-2083 (62%)
21	Otherks-T1pks	7491509	7544137	-
22	**T1pks**	**7641469**	**7758787**	**Heronamide cluster BE-14106 (96%)**
23	T3pks	7942737	7983846	Alkylresorcinol (66%)
24	Lantipeptide	8037336	8068867	Meilingmycin (2%)
25	T3pks	8139618	8180721	Herboxidiene (3%)

Identification, Organisation and Bioinformatic Analysis of the Heronamide Biosynthetic Gene Cluster (BSGC) (*hrn*)

Two of the twenty-five clusters were identified as type-I polyketide synthase clusters, of which the heronamide BSGC was readily identifiable due to high similarity (96%) with the BEC-14106 BSGC.[11] The heronamide BSGC was analyzed in detail using Artemis for visualization and annotation;[12] gene function was assigned based on a combination of searches with BLAST at NCBI server,[13] Pfam at Sanger server,[14] and Conserved Domain Database (CDD) at NCBI server;[15] modules and domains of the polyketide synthase were assigned on the basis of Pfam, CDD and antiSMASH search output. The frame Plot function of Artemis was used to confirm the extent of PCSs and for identification of putative frame-shifts due to errors in the sequence.[16] One obvious frame-shift was identified causing the split of the PKS gene *hrnB* in two; in comparison with the amino acid sequences of the ML-449 and BE-14106 biosynthetic gene

cluster, a missing G causing the frame-shift was identified. The sequence of the heronamide biosynthetic gene cluster has been deposited in the European Nucleotide Archive under the accession number LT548273. The *hrn* gene cluster spans 80166 base pairs and bioinformatic analysis identified 21 genes, *hrnA* to *hrnQ*. Predicted protein sequences for the ketoreductase and dehydratase domains of the polyketides, heronamide C, ML-449 (FJ872525.1), vicenistatin (AB086653.1), hitachimycin (LC008143.1), BEC-14106 (FJ872523.1), incednine (AB767280.1), cremimycin (AB818354.1), bacillaene (AJ634060.2), rifamycin (AJ223012.1), rhizoxin (AM411073.1), and rapamycin (X86780.1) were extracted from their antismash 3.0 outputs and aligned, respectively, using clustalW2.[17] The alignment revealed that all heronamide KR domains were canonical B-type KR domains[18] containing the conserved LDD motif and P144 and N148 residues with the exception of the KR domain in module 4 of the macrolactam PKS responsible for the C10-C11 *cis*-double bond where the LDD sequence was replaced by VDN and the substitution of alanine at position 144; this was assigned as an A-type KR domain. The KR domain of module 6 in the ML-449 BSGC was 98.9% identical to its counterpart in the heronamide BSGC strongly indicating that the assignment of a C6-C7 *cis*-double bond is erroneous. In addition, our analysis supports the findings of Zhu *et al*.[19] who reported the absence of a 'shift module,' in the heronamide F BSGC. Furthermore, there is no consistent phylogenetic signal and as such, the basis of the αβ – βγ double bond shifts remains unclear.

Table S4: Description of proteins encoded in genes identified in the heronamide biosynthetic gene cluster, proposed function, and homologs

Protein	size (aa)	Proposed function	% Similarities/ % Identities Homolog from		
			ML-449 cluster[x]	**BE-14106 cluster[x]**	***her* cluster[x]**
HrnH	950	LuxR-type transcriptional regulator	MlaH 98/95	BecH 94/88	HerH 74/64
HrnA1	6334	Polyketide synthase type I	MlaA1 96/94	BecA 90/86	HerA1 74/64
HrnA2	1043	Polyketide synthase type I	MlaA2 98/97	BecA 84/77	HerA2 83/76
HrnI	363	Glycine oxidase/FAD-dependent oxidoreductase	MlaI 95/93	BecI 88/79	HerI 83/73
HrnC	695	Polyketide synthase type I	MlaC 97/95	BecC 90/84	HerC 84/75
HrnU	187	putative NRPS accessory protein	MlaU 100/99	BecU 95/89	HerU 88/81
HrnB	3526	Polyketide synthase type I	MlaB 96/95	BecB 92/87	HerB 85/78
HrnJ	532	putative AMP-dependent acyl-CoA synthetase/ ligase	MlaJ 97/97	Bec 92/87	HerJ 87/79
HrnK	313	Acyltransferase	MlaK 97/96	Bec 90/84	HerK 85/77
HrnS	78	Peptidyl carrier protein	MlaS 98/95	BecS 94/92	HerS 91/82
HrnL	504	NRPS adenylation domain	MlaL 96/94	BecL 89/83	HerL 89/81
HrnM	199	TetR-type transcriptional regulator	MlaM 98/96	BecM 89/86	HerM 87/81
HrnN	523	MFS-type efflux pump	MlaN 97/95	BecN 93/87	HerN 85/77
HrnO	411	P450 monooxygenase	MlaO 98/96	BecO 95/91	HerO 87/76
HrnD	3368	Polyketide synthase type I	MlaD 96/95	BecD 92/88	HerD 83/77
HrnP	311	Putative L-amino acid amidase/ proline iminopeptidase	MlaG 98/98	BecP 91/83	HerP 90/83
HrnG	2001	Polyketide synthase type I	MlaG 95/94	BecG 89/85	HerG 80/72
HrnF	3382	Polyketide synthase type I	MlaF 97/94	BecF 93/88	HerF 84/76
HrnE	1637	Polyketide synthase type I	MlaE 95/93	BecE 91/87	HerE 81/72
HrnT	88	SimX-like protein	MlaT 92/90	BecT 80/75	---
HrnQ	256	Thioesterase type II	MlaQ 98/96	BecQ 93/86	---

References

1. P. Yu, A. Patel and K. N. Houk, *J. Am. Chem. Soc.*, 2015, **137**,13518–13523.

2. D. K. Derewacz, B. C. Covington, J. A. McLean and B. O. Bachmann, *ACS Chem Biol.*, 2015, **18**, 1998-2006.

3. H. Jørgensen, K. F. Degnes, A. Dikiy, E. Fjaervik, G. Klinkenberg and S. B. Zotchev, *Appl. Environ. Microbiol.* 2010, **76**, 283–293.

4. K. Kojiri, S. Nakajima, H. Suzuki, H. Kondo and H. Suda, *J. Antibiot.*, 1992, **45**, 868–74; H. Jørgenson, K. F. Degnes, H. Sletta, E. Fjærvik, A. Dikiy, L. Ferfindal, P. Bruheim, G. Klinkenberg, H. Bredholt, G. Nygård, S. O. Døskeland, T. E. Ellingsen and S. B. Zotchev, *Chem. Biol.*, 2009, **16**, 1109-1121.

5. T. Kieser, M. J. Bibb, M. J. Buttner, K. F. Chater and D. A. Hopwood, Practical *Streptomyces* Genetics. 2000, Norwich, UK. John Innes Foundation.

6. J. Eid *et al*., *Science*, 2009, **323**, 133-138.

7. G. H. Van Domselaar, P. Stothard S. Shrivastava, J. A. Cruz, A. Guo, X. Dong, P. Lu, D. Szafron, R. Greiner and D. S. Wishart, *Nucleic Acids Res.* 2005, **33**, 455-459.

8. D. Hyatt, G. –L. Chen, P. F. Locascio, M. L. Land, F. W. Larimer and L. J. Hauser, *BMC Bioinformatics* 2010, **11**, 119.

9. R. K. Aziz *et al.*, *BMC genomics,* 2009, **9**, 75.

10. T. Weber, K. Blin, S. Duddela, D. Krug, H. U. Kim, R. Bruccoleri, S. Y. Lee, M. A. Fischbach, R. Müller, W. Wohlleben, R. Breitling, E. Takano and M. H. Medema, *Nucleic Acids Res.* 2015, DOI: 10.1093/nar/gkv437.

11. H. Jørgensen, K. F. Degnes, H. Sletta, E. Fjaervik, A. Dikiy, L. Herfindal, P. Bruheim, G. Klinkenberg, H. Bredholt, G. Nygård, S. O. Døskeland, T. E. Ellingsen and S. B. Zotchev, *Chem Biol*., 2009, **16**, 1109-1121.

12. K. Rutherford, J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream and B. Barrell, *Bioinformatics* 2000, **16**, 944-945.

13. S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman, *Nucleic Acids Res.* 1997, **25**, 3389-3402.

14. R. D. Finn, A. Bateman, J. Clements, P. Coggill, R. Y. Eberhardt, S. R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E. L. L. Sonnhammer, J. Tate, and M. Punta, *Nucleic Acids Res.* 2014, **42**, 222-230.

15. A. Marchler-Bauer, M. K. Derbyshire, N. R. Gonzales, S. Lu, F. Chitsaz, L. Y. Geer, R. C. Geer, J. He, M. Gwadz, D. I. Hurwitz, C. J. Lanczycki, F. Lu, G. H. Marchler, J. S. Song, N. Thanki, Z. Wang, R. A. Yamashita, D. Zhang, C. Zheng and S. H. Bryant, *Nucleic Acids Res.* 2015, **43**, 222-226.

16. M. J. Bibb, M. J. Bibb, J. M. Ward and S. N. Cohen, *Mol. Gen. Genet.* 1985, **199**, 26-36.

17. M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson and D. G. Higgins, *Bioinformatics* 2007, **23**, 2947-2948.

18. P. Caffrey, *ChemBioChem.*, 2003, **4,** 654–657.

19. Y. Zhu, W. Zhang, Y. Chen, C. Yuan, H. Zhang, G. Zhang, L. Ma, Q. Zhang, X. Tian, S. Zhang and C. Zhang, *ChemBioChem.*, 2015, **14**, 2086-2093.