Electronic Supplementary Material (ESI) for ChemComm.

Supporting information for :

X-ray Crystallographic Structure of a Teixobactin Analogue Reveals Key Interactions of the Teixobactin Pharmacophore

Hyunjun Yang, Derek R. Du Bois, Joseph W. Ziller, and James S. Nowick*
Email: jsnowick@uci.edu

Department of Chemistry, University of California, Irvine,
Irvine, California 92697-2025, USA

Table of Contents

Materials and Methods

General information

Methylene chloride (CH_2Cl_2) was passed through alumina under argon prior to use. Amine-free *N,N*-dimethylformamide (DMF) was purchased from Alfa Aesar. Fmoc-D-*allo*-Ile-OH was purchased from Santa Cruz Biotechnology. Other protected amino acids were purchased from CHEM-IMPEX. Preparative reverse-phase HPLC was performed on a Beckman Gold Series P instrument equipped with an Agilent Zorbax SB-C18 column. Analytical reverse-phase HPLC was performed on either an Agilent 1200 or an Agilent 1260 Infinity II instrument, both equipped with a Phenomonex Aeris PEPTIDE 2.6μ XB-C18 column. HPLC grade acetonitrile (MeCN) and deionized water (18 MΩ) containing 0.1% trifluoroacetic acid (TFA) were used as solvents for both preparative and analytical reverse-phase HPLC. Deionized water (18 MΩ) was obtained from a Barnstead NANOpure Diamond water purification system. All teixobactin homologues were prepared and studied as the trifluoroacetate salts.

Synthesis of teixobactin homologues

Ac-$\Delta_{1\text{-}5}$-Arg_{10}-teixobactin and other teixobactin homologues were synthesized as the trifluoroacetate salts following procedures we have previously reported.[1] Dry DMF was used instead of a mixture of MeCN/THF/CH_2Cl_2 for the cyclization step. For the acetylation reaction, glacial acetic acid (3.0 μL, 0.90 mmol, 10 equiv) was coupled with coupling reagent HCTU (142 mg, 0.46 mmol, 4 equiv) in 20% (v/v) collidine in dry DMF (5 mL). For the synthesis of D-Dap_8,Arg_{10}-teixobactin, Fmoc-D-Dap(Alloc)-OH was used instead of Fmoc-D-Thr-OH, and the Alloc protecting group was deprotected using $Pd(PPh_3)_4$ (0.10 equiv) and $PhSiH_3$ (20 equiv) in CH_2Cl_2 prior to coupling Fmoc-Ile_{11}-OH.[2]

MIC assays of teixobactin homologues

MIC assays of Ac-$\Delta_{1\text{-}5}$-Arg_{10}-teixobactin and other teixobactin homologues were performed following procedures we have previously reported.[1]

Crystallization of Ac-$\Delta_{1\text{-}5}$-Arg_{10}-teixobactin[3]

Ac-$\Delta_{1\text{-}5}$-Arg_{10}-teixobactin was dissolved in 0.1 M $Na_4P_2O_7$ (sodium pyrophosphate) at pH 7.00 (adjusted with HCl and NaOH) to make a 10 mg/mL stock solution. Crystallization conditions were screened using the hanging-drop vapor-diffusion method with three crystallization kits (Hampton Index, PEG/Ion, and Crystal Screen) in 96-well plates. Using a TTP LabTech Mosquito® liquid handling instrument, three 150-nL hanging drops with differing ratios of peptide to well solution (1:1, 1:2, and 2:1 peptide/well solution) were made per condition in each 96-well plate, for a total of 864 experiments. Crystals of Ac-$\Delta_{1\text{-}5}$-Arg_{10}-teixobactin grew rapidly (~24 h) with a well solution of 0.2 M ammonium tartrate dibasic and 20% polyethylene glycol 3,350. Crystallization conditions were further optimized using a 4x6 matrix Hampton VDX 24-well plate, varying the concentration of ammonium tartrate dibasic (0.12, 0.16, 0.20, 0.24, 0.28, and 0.32 M) in the columns and the concentration of polyethylene glycol 3,350 (10, 15, 20, and 25%) in the rows. The 0.24 M ammonium tartrate dibasic and 20% polyethylene glycol 3,350 condition afforded colorless parallelogram-shaped crystals suitable for X-ray diffraction.

X-ray crystallographic data collection, data processing, and structure determination

A colorless crystal of approximate dimensions 0.030 x 0.130 x 0.200 mm was mounted in a cryoloop and transferred to a Bruker SMART APEX II diffractometer. The APEX2[4] program package was used to determine the unit-cell parameters and for data collection (180 sec/frame scan time for a sphere of diffraction data). The raw frame data was processed using SAINT[5] and

SADABS[6] to yield the reflection data file. Subsequent calculations were carried out using the SHELXTL[7] program. The diffraction symmetry was 2/*m* and the systematic absences were consistent with the monoclinic space groups *C*2, *Cm* and *C*2/*m*. It was later determined that space group *C*2 was correct.

The structure was solved by direct methods and refined on F^2 by full-matrix least-squares techniques. The analytical scattering factors[8] for neutral atoms were used throughout the analysis. Hydrogen atoms were either located from a difference-Fourier map and refined (x,y,z and U_{iso}) or were included using a riding model. There were 1.5 molecules of water solvent present per formula-unit. One water molecule was located on a twofold rotation axis. Water hydrogen atoms were refined with d(O-H) = 0.85Å.

At convergence, wR2 = 0.0878 and Goof = 1.016 for 520 variables refined against 7914 data (0.80Å), R1 = 0.0424 for those 6389 data with $I > 2.0\sigma(I)$. The absolute structure was assigned by refinement of the Flack parameter.[9]

There was a single residual (1.23e^-) present in the final difference-Fourier map. It was not possible to determine the nature of the residual. The SQUEEZE[10] routine in the PLATON[11] program package was used to account for the electrons associated with the solvent accessible voids.

Definitions:
$wR2 = [\Sigma[w(F_o^2-F_c^2)^2] / \Sigma[w(F_o^2)^2]]^{1/2}$
$R1 = \Sigma||F_o|-|F_c|| / \Sigma|F_o|$
$Goof = S = [\Sigma[w(F_o^2-F_c^2)^2] / (n-p)]^{1/2}$ where n is the number of reflections and p is the total number of parameters refined.

The thermal ellipsoid plot is shown at the 50% probability level.

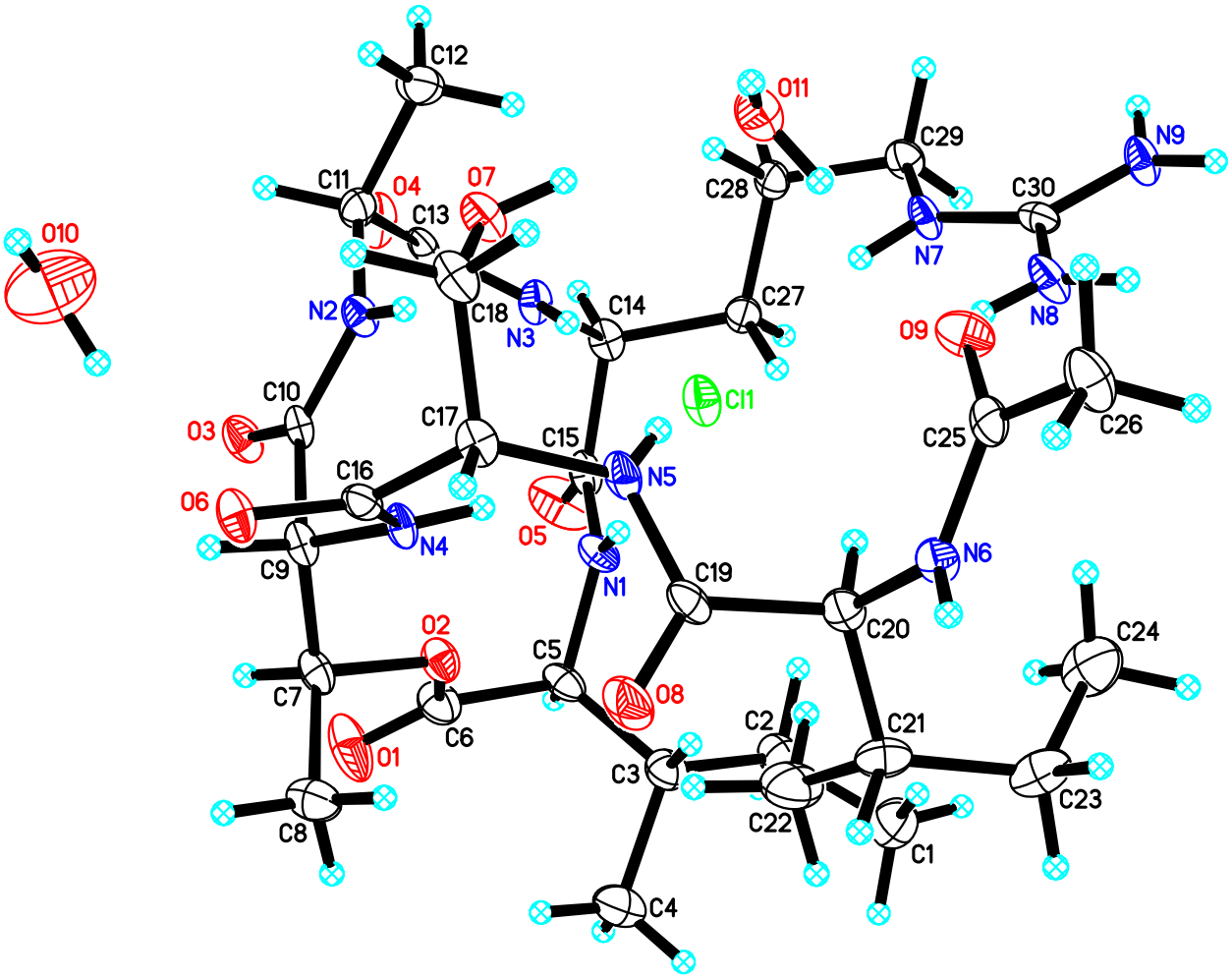
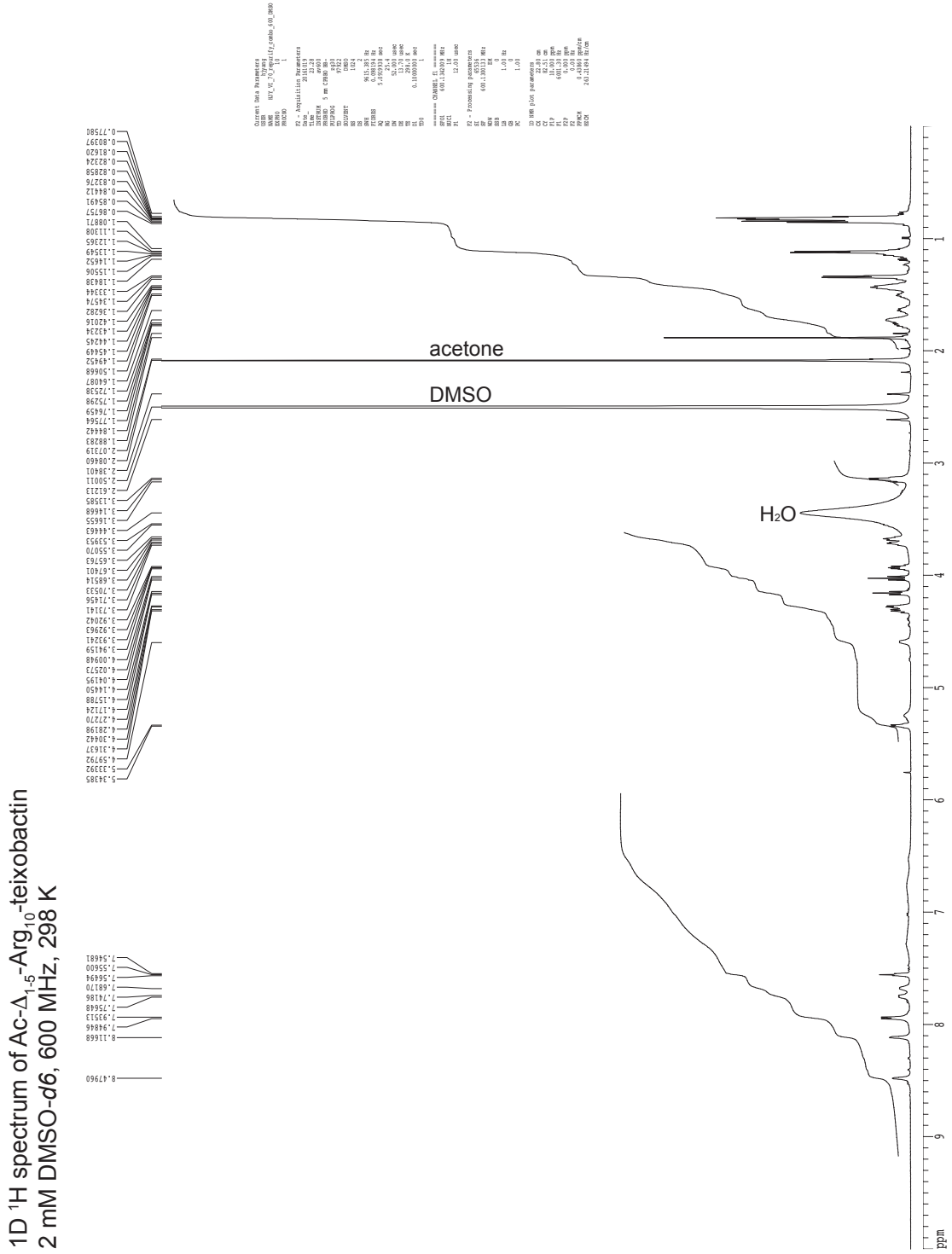


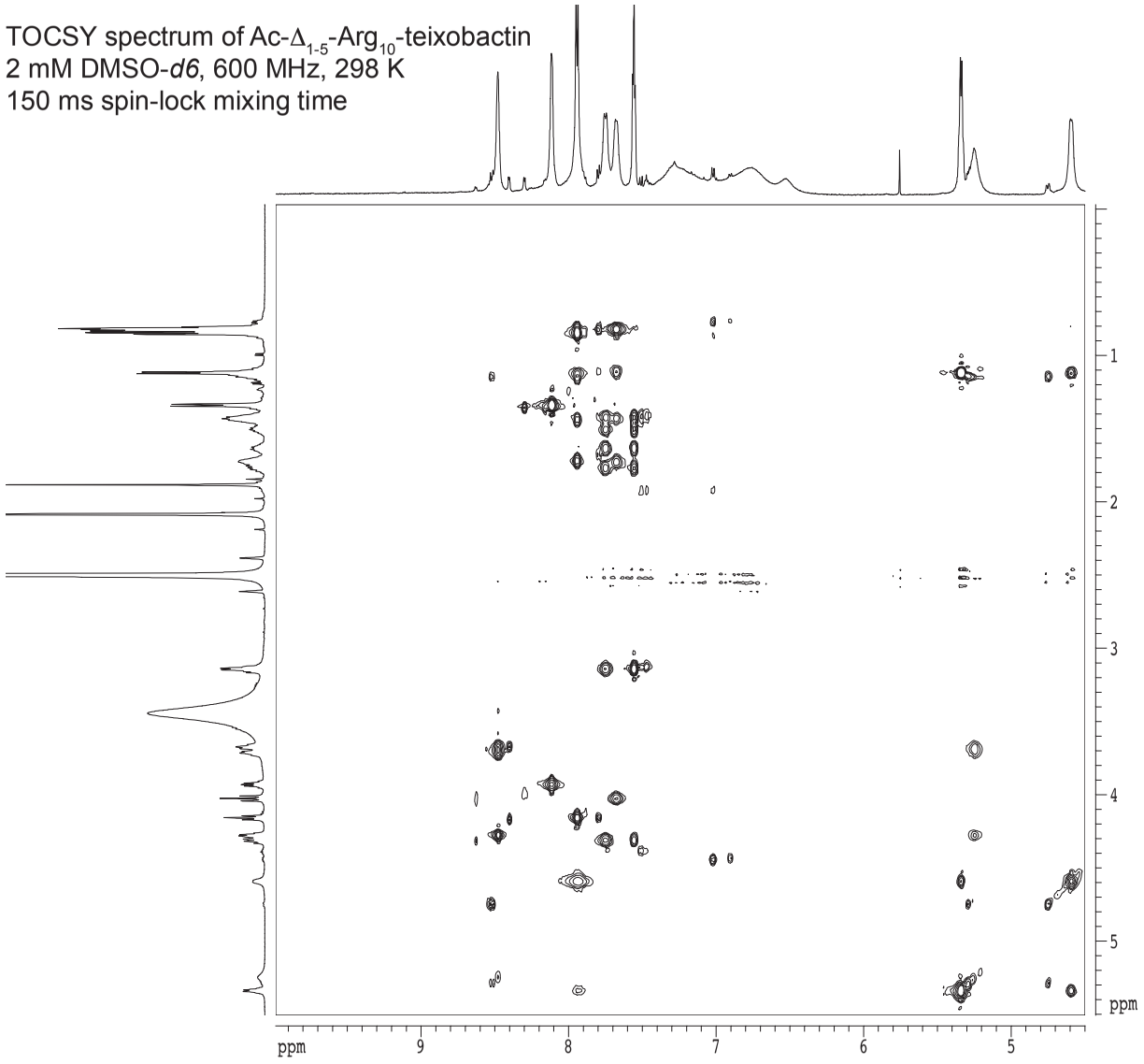
Table S1 Crystal data and structure refinement for Ac-$\Delta_{1\text{-}5}$-Arg_{10}-teixobactin.

Empirical formula	$C_{30}\ H_{54}\ Cl\ N_9\ O_9 \bullet 1.5(H_2O)$	
Formula weight	747.29	
Temperature	88(2) K	
Wavelength	0.71073 Å	
Crystal system	Monoclinic	
Space group	*C*2	
Unit cell dimensions	a = 19.376(3) Å	α= 90°.
	b = 12.405(2) Å	β= 94.809(3)°.
	c = 16.135(3) Å	γ = 90°.
Volume	3864.5(12) $Å^3$	
Z	4	
Density (calculated)	1.284 Mg/m^3	
Absorption coefficient	0.163 mm^{-1}	
F(000)	1604	
Crystal color	colorless	
Crystal size	0.200 x 0.130 x 0.030 mm^3	
Theta range for data collection	1.951 to 26.393°	
Index ranges	$-24 \le h \le 24$, $-15 \le k \le 15$, $-20 \le l \le 20$	
Reflections collected	21510	
Independent reflections	7914 [R(int) = 0.0477]	
Completeness to theta = 25.500°	100.0 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	0.8620 and 0.8121	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	7914 / 4 / 520	
Goodness-of-fit on F^2	1.016	
Final R indices [I>2sigma(I) = 6389 data]	R1 = 0.0424, wR2 = 0.0811	
R indices (all data, 0.80 Å)	R1 = 0.0625, wR2 = 0.0878	
Absolute structure parameter	0.04(4)	
Largest diff. peak and hole	0.193 and -0.348 $e.Å^{-3}$	

^{1}H and TOCSY NMR spectra of Ac-$\Delta_{1\text{-}5}$-Arg$_{10}$-teixobactin



TOCSY spectrum of Ac-$\Delta_{1\text{-}5}$-Arg_{10}-teixobactin
2 mM DMSO-*d6*, 600 MHz, 298 K
150 ms spin-lock mixing time




```
                    Current Data Parameters
                USER                  hjyang
NAME            HJY_VI_70_repurify_combo_600_DMSO
                EXPNO                     11
                PROCNO                     1

                    F2 - Acquisition Parameters
                Date_              20161019
                Time                  23.30
                INSTRUM                av600
                PROBHD    5 mm CPBBO BB-
                PULPROG               mlevph
                TD                      1912
                SOLVENT                 DMSO
                NS                        32
                DS                        16
                SWH                 6009.615 Hz
                FIDRES              3.143104 Hz
                AQ                  0.1591284 sec
                RG                       25.4
                DW                     83.200 usec
                DE                     10.00 usec
                TE                     298.0 K
                D0                  0.00007156 sec
                D1                  2.00000000 sec
                D9                  0.15000001 sec
                D12                 0.00002000 sec
                IN0                 0.00016640 sec
                L1                        90

                ======= CHANNEL f1 =======
                SFO1               600.1330007 MHz
                NUC1                        1H
                P1                     12.00 usec
                P5                     16.67 usec
                P6                     25.00 usec
                P7                     50.00 usec
                P17                  2500.00 usec

                    F1 - Acquisition parameters
                ND0                         1
                TD                        256
                SFO1                600.133 MHz
                FIDRES             23.475060 Hz
                SW                    10.014 ppm
                FnMODE                   TPPI

                    F2 - Processing parameters
                SI                       1024
                SF                600.1300133 MHz
                WDW                      QSINE
                SSB                         2
                LB                     0.00 Hz
                GB                          0
                PC                     1.00

                    F1 - Processing parameters
                SI                       1024
                MC2                      TPPI
                SF                600.1300133 MHz
                WDW                      QSINE
                SSB                         2
                LB                     0.00 Hz
                GB                          0

                    2D NMR plot parameters
                CX2                    15.00 cm
                CX1                    15.00 cm
                F2PLO                  9.985 ppm
                F2LO                5992.24 Hz
                F2PHI                  4.499 ppm
                F2HI                2699.86 Hz
                F1PLO                  5.506 ppm
                F1LO                3304.34 Hz
                F1PHI                 -0.029 ppm
                F1HI                  -17.38 Hz
                F2PPMCM              0.36574 ppm/cm
                F2HZCM             219.49181 Hz/cm
                F1PPMCM              0.36900 ppm/cm
                F1HZCM             221.44806 Hz/cm
```

Ac-$\Delta_{1\text{-}5}$-Arg$_{10}$-teixobactin

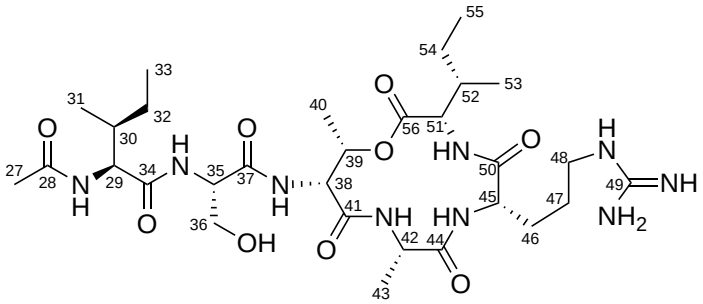
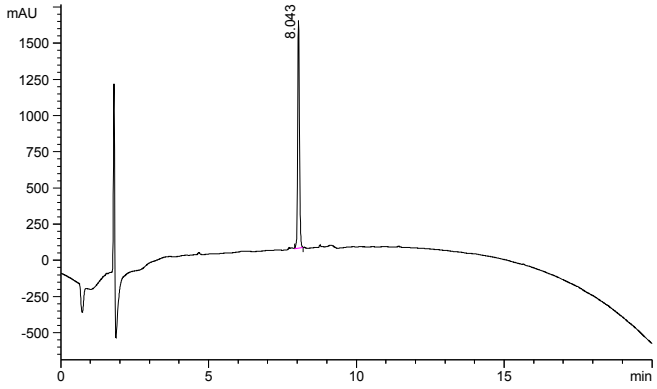


Table S2 NMR data of Ac-$\Delta_{1\text{-}5}$-Arg$_{10}$-teixobactin

	Ac	27	1.88 (3H, s)	Residue 9	Ala	42	3.90 (1H, qd, 7.5, 5.6)
		28	N/A			42-NH	8.12 (1H, m)
Residue 6	Ile	29	4.16 (1H, t, 8.2)			43	1.34 (3H, d, 7.5)
		29-NH	7.94 (1H, d, 8.0)			44	N/A
		30	1.75 (1H, m)	Residue 10	Arg	45	4.30 (1H, m)
		31	0.85 (3H, m)			45-NH	7.75 (1H, m)
		32	1.14 (1H, m)			46	1.78 (1H, m)
			1.44 (1H, m)				1.65 (1H, m)
		33	0.81 (3H, m)			47	1.52 (1H, m)
		34	N/A				1.43 (1H, m)
Residue 7	Ser	35	4.28 (1H, q, 5.7)			47-NH	not observed
		35-NH	8.48 (1H, m)			48	3.14 (2H, m)
		36	3.72 (1H, m)			48-NH	7.56 (1H, t, 5.4)
			3.67 (1H, m)			49	N/A
		36-OH	5.25 (1H, br)			49-NH	not observed
		37	N/A			50	N/A
Residue 8	D-Thr	38	4.59 (1H, m)	Residue 11	Ile	51	4.03 (1H, t, 9.5)
		38-NH	7.94 (1H, m)			51-NH	7.68 (1H, m)
		39	5.34 (1H, m)			52	1.73 (1H, m)
		40	1.12 (3H, d, 6.2)			53	0.82 (3H, m)
		41	N/A			54	1.43 (1H, m)
							1.11 (1H, m)
						55	0.83 (3H, m)
						56	N/A

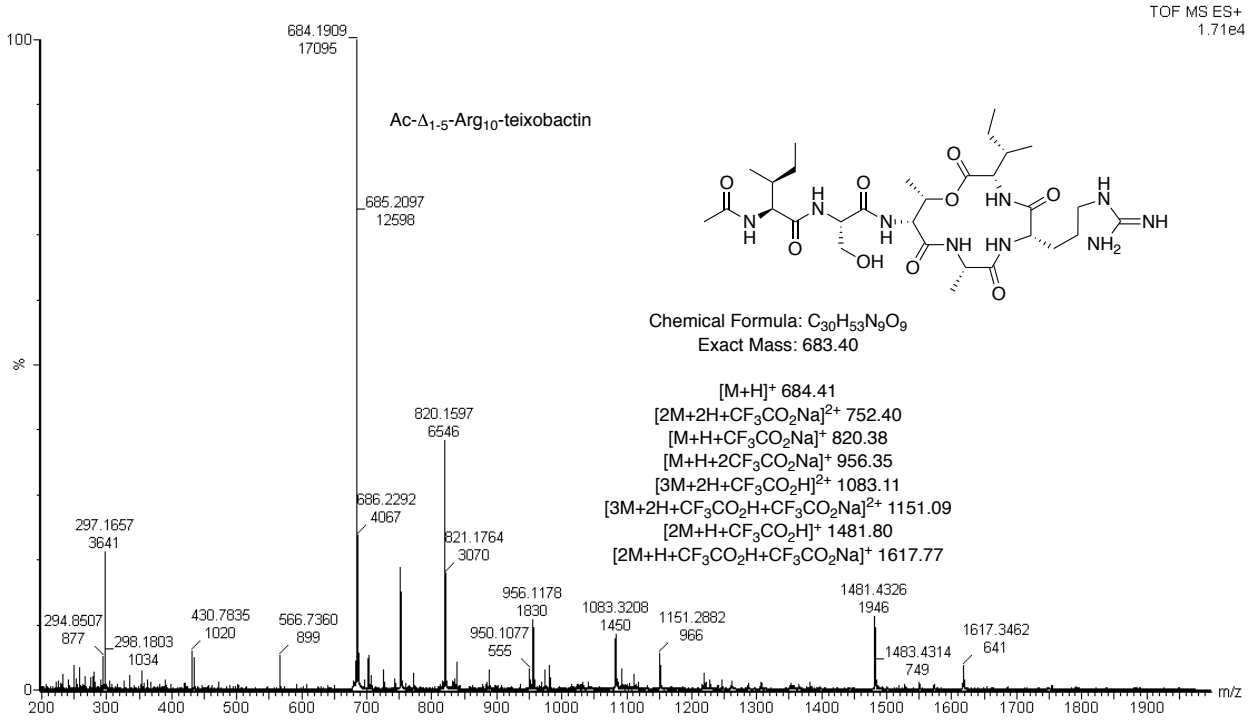
HPLC Traces and Mass Spectra of Teixobactin Homologues

Ac-$\Delta_{1\text{-}5}$-Arg$_{10}$-teixobactin : Analytical RP-HPLC and mass spectrum



Signal 1:VWD1 A, Wavelength=214 nm

Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	8.043	MM	0.071	6722.876	100.000	100.000

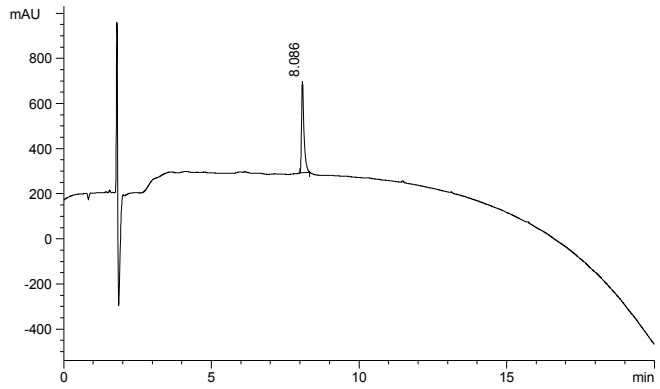


Ac-$\Delta_{1\text{-}5}$-Arg$_{10}$-teixobactin

Chemical Formula: $C_{30}H_{53}N_9O_9$
Exact Mass: 683.40

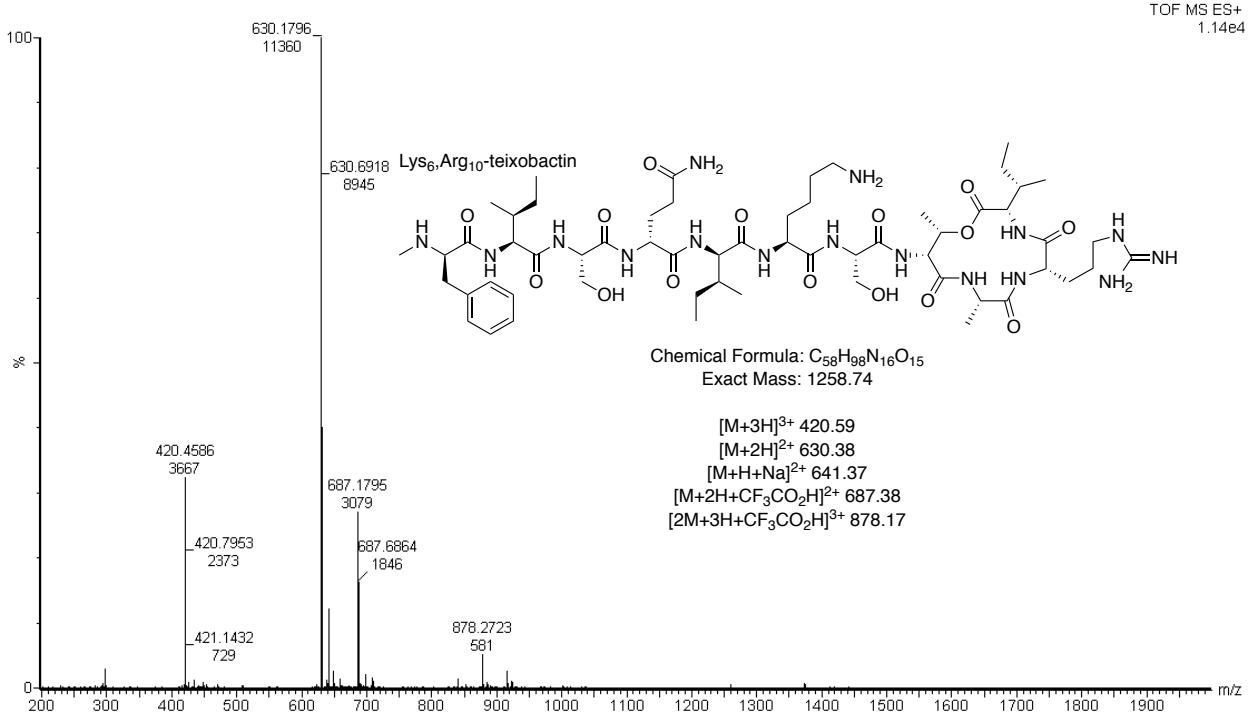
$[M+H]^+$ 684.41
$[2M+2H+CF_3CO_2Na]^{2+}$ 752.40
$[M+H+CF_3CO_2Na]^+$ 820.38
$[M+H+2CF_3CO_2Na]^+$ 956.35
$[3M+2H+CF_3CO_2H]^{2+}$ 1083.11
$[3M+2H+CF_3CO_2H+CF_3CO_2Na]^{2+}$ 1151.09
$[2M+H+CF_3CO_2H]^+$ 1481.80
$[2M+H+CF_3CO_2H+CF_3CO_2Na]^+$ 1617.77

Lys_6,Arg_{10}-teixobactin : Analytical RP-HPLC and mass spectrum



Signal 1:VWD1 A, Wavelength=214 nm

Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	8.086	MM	0.094	2297.094	100.000	100.000

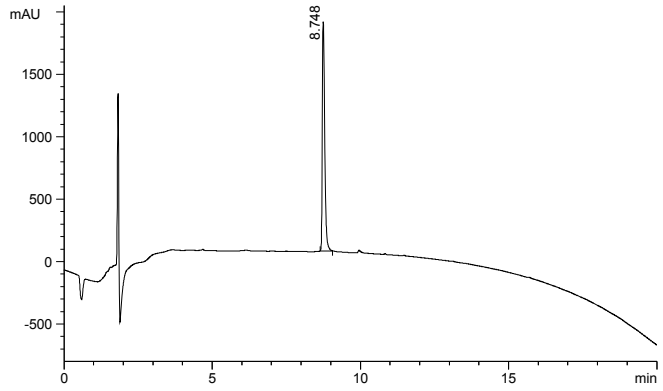


Lys_6,Arg_{10}-teixobactin

Chemical Formula: $C_{58}H_{98}N_{16}O_{15}$
Exact Mass: 1258.74

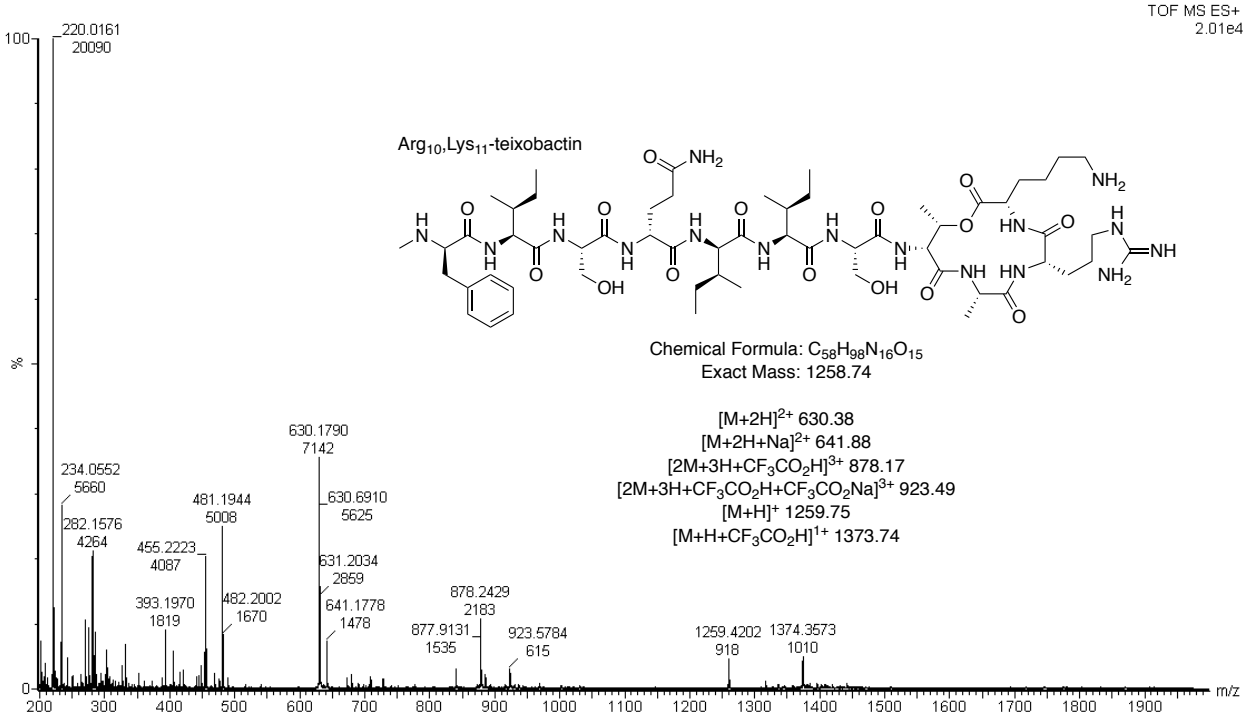
$[M+3H]^{3+}$ 420.59
$[M+2H]^{2+}$ 630.38
$[M+H+Na]^{2+}$ 641.37
$[M+2H+CF_3CO_2H]^{2+}$ 687.38
$[2M+3H+CF_3CO_2H]^{3+}$ 878.17

Arg_{10},Lys_{11}-teixobactin : Analytical RP-HPLC and mass spectrum



Signal 1:VWD1 A, Wavelength=214 nm

Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	8.748	MM	0.095	10503.242	100.000	100.000

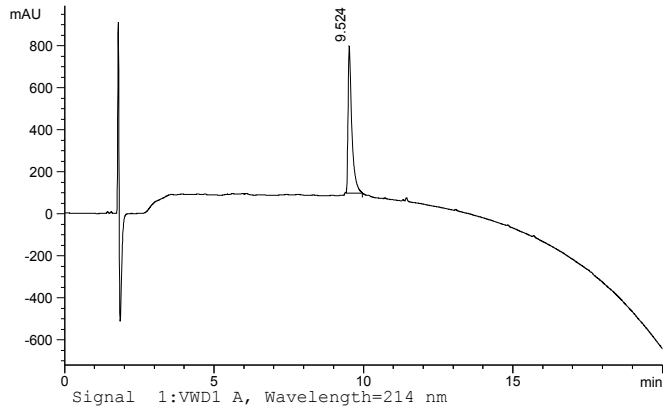


Arg_{10},Lys_{11}-teixobactin

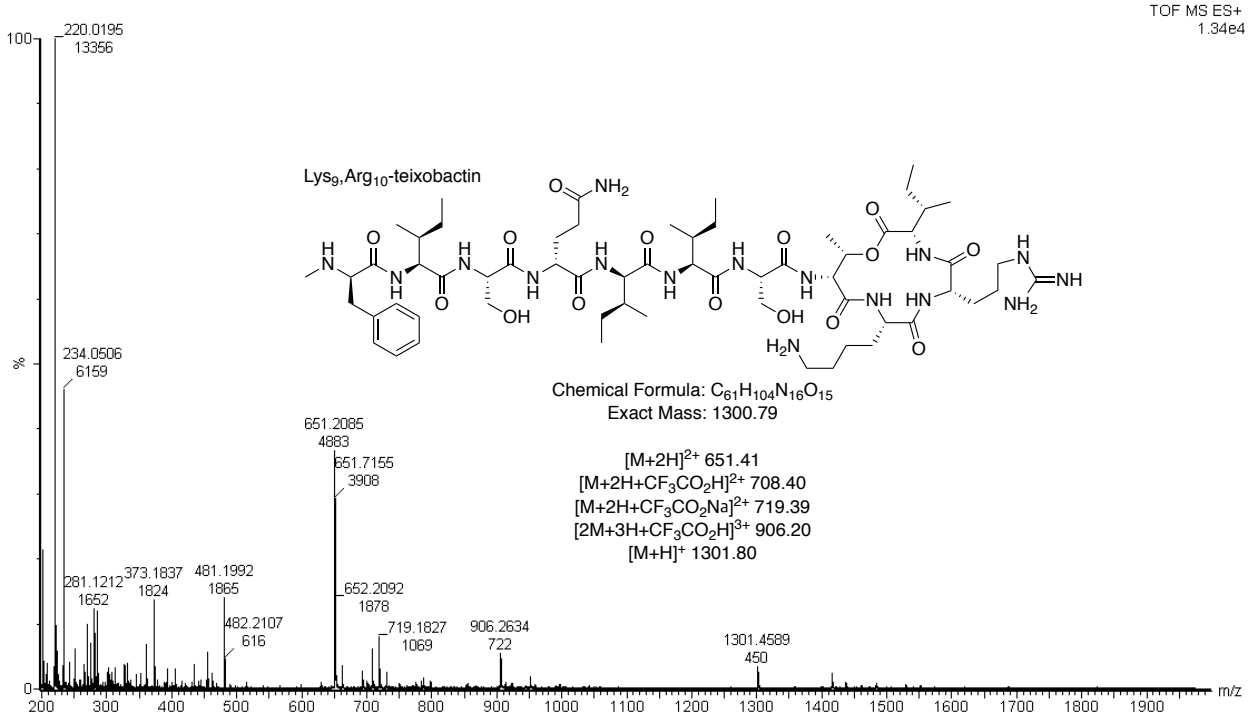
Chemical Formula: $C_{58}H_{98}N_{16}O_{15}$
Exact Mass: 1258.74

$[M+2H]^{2+}$ 630.38
$[M+2H+Na]^{2+}$ 641.88
$[2M+3H+CF_3CO_2H]^{3+}$ 878.17
$[2M+3H+CF_3CO_2H+CF_3CO_2Na]^{3+}$ 923.49
$[M+H]^{+}$ 1259.75
$[M+H+CF_3CO_2H]^{1+}$ 1373.74

Lys_9,Arg_{10}-teixobactin : Analytical RP-HPLC and mass spectrum



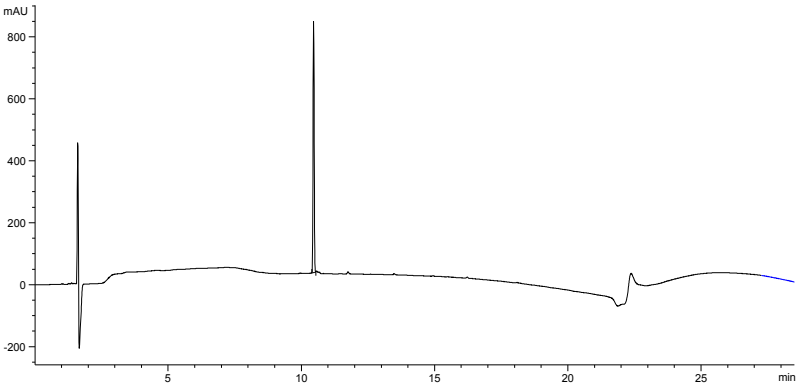
Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	9.524	MM	0.144	6068.178	100.000	100.000



Chemical Formula: $C_{61}H_{104}N_{16}O_{15}$
Exact Mass: 1300.79

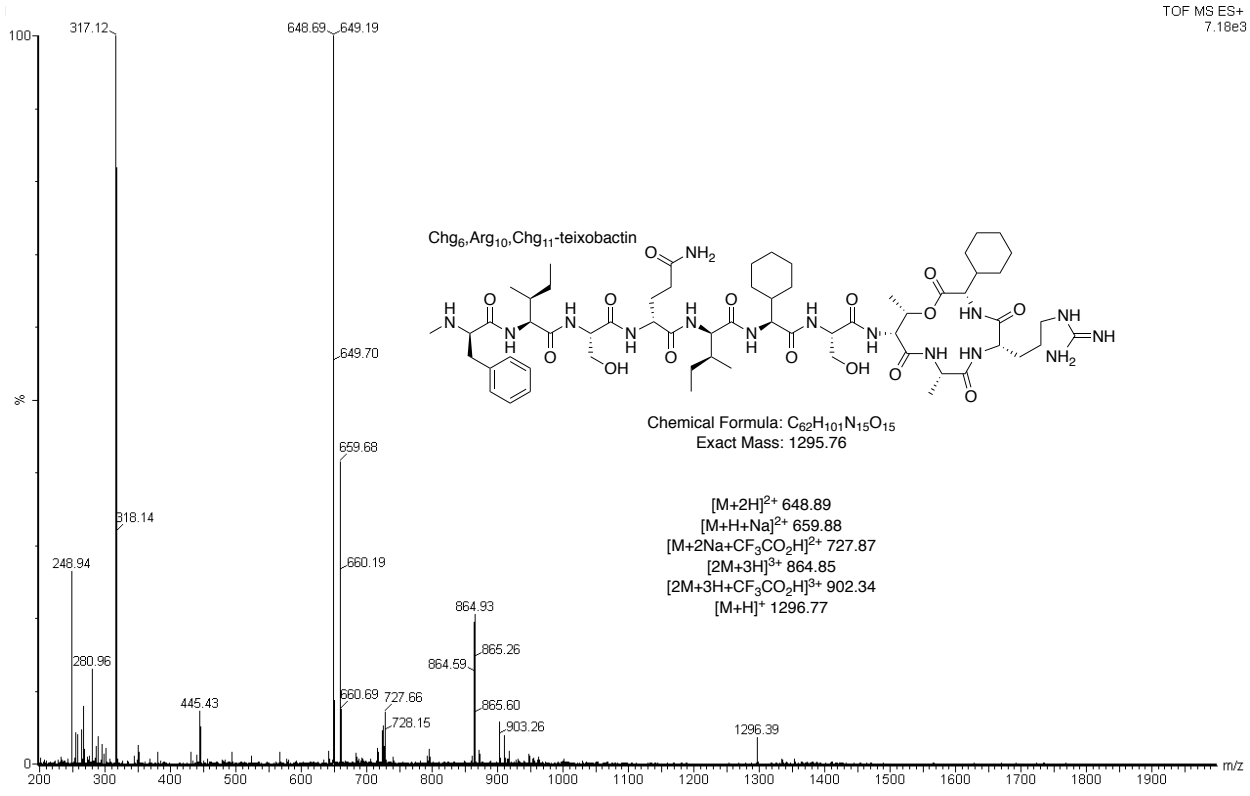
$[M+2H]^{2+}$ 651.41
$[M+2H+CF_3CO_2H]^{2+}$ 708.40
$[M+2H+CF_3CO_2Na]^{2+}$ 719.39
$[2M+3H+CF_3CO_2H]^{3+}$ 906.20
$[M+H]^{+}$ 1301.80

Chg_6,Arg_{10},Chg_{11}-teixobactin : Analytical RP-HPLC and mass spectrum



Signal 1: MWD1 A, Sig=214,4 Ref=off

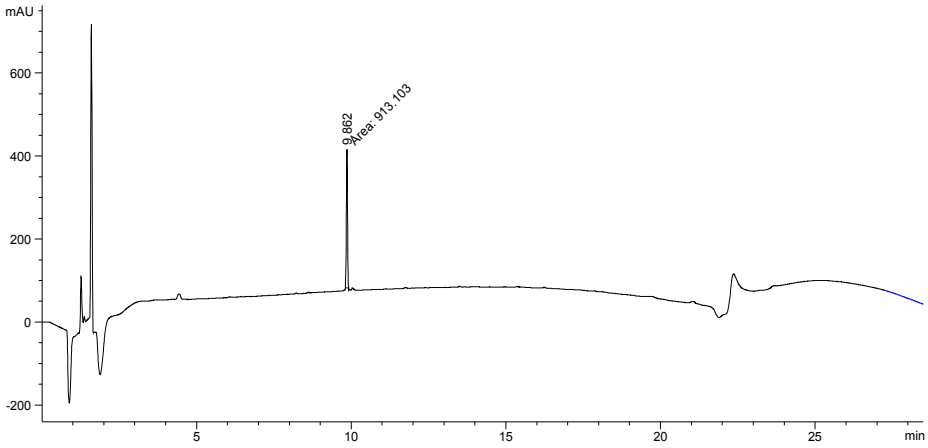
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Area %	Name
1	10.461	MM	0.0468	2296.97607	100.0000	?



Chemical Formula: $C_{62}H_{101}N_{15}O_{15}$
Exact Mass: 1295.76

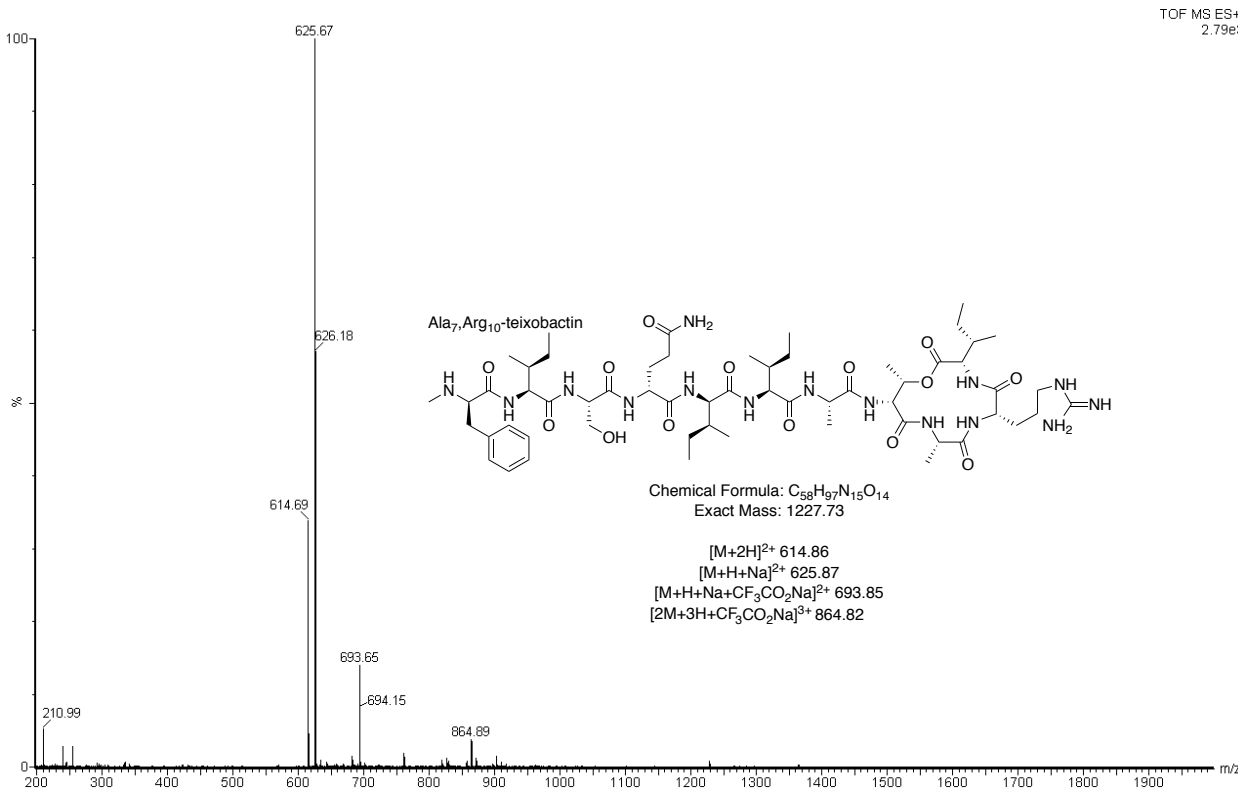
$[M+2H]^{2+}$ 648.89
$[M+H+Na]^{2+}$ 659.88
$[M+2Na+CF_3CO_2H]^{2+}$ 727.87
$[2M+3H]^{3+}$ 864.85
$[2M+3H+CF_3CO_2H]^{3+}$ 902.34
$[M+H]^{+}$ 1296.77

Ala_7,Arg_{10}-teixobactin : Analytical RP-HPLC and mass spectrum

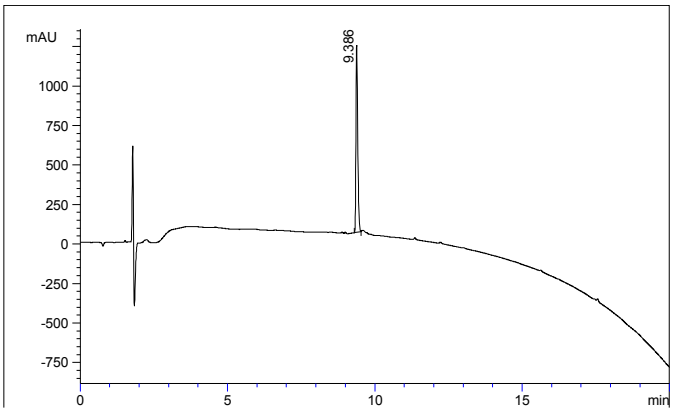


Signal 1: MWD1 A, Sig=214,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	9.862	MM	0.0452	913.10345	336.54343	100.0000

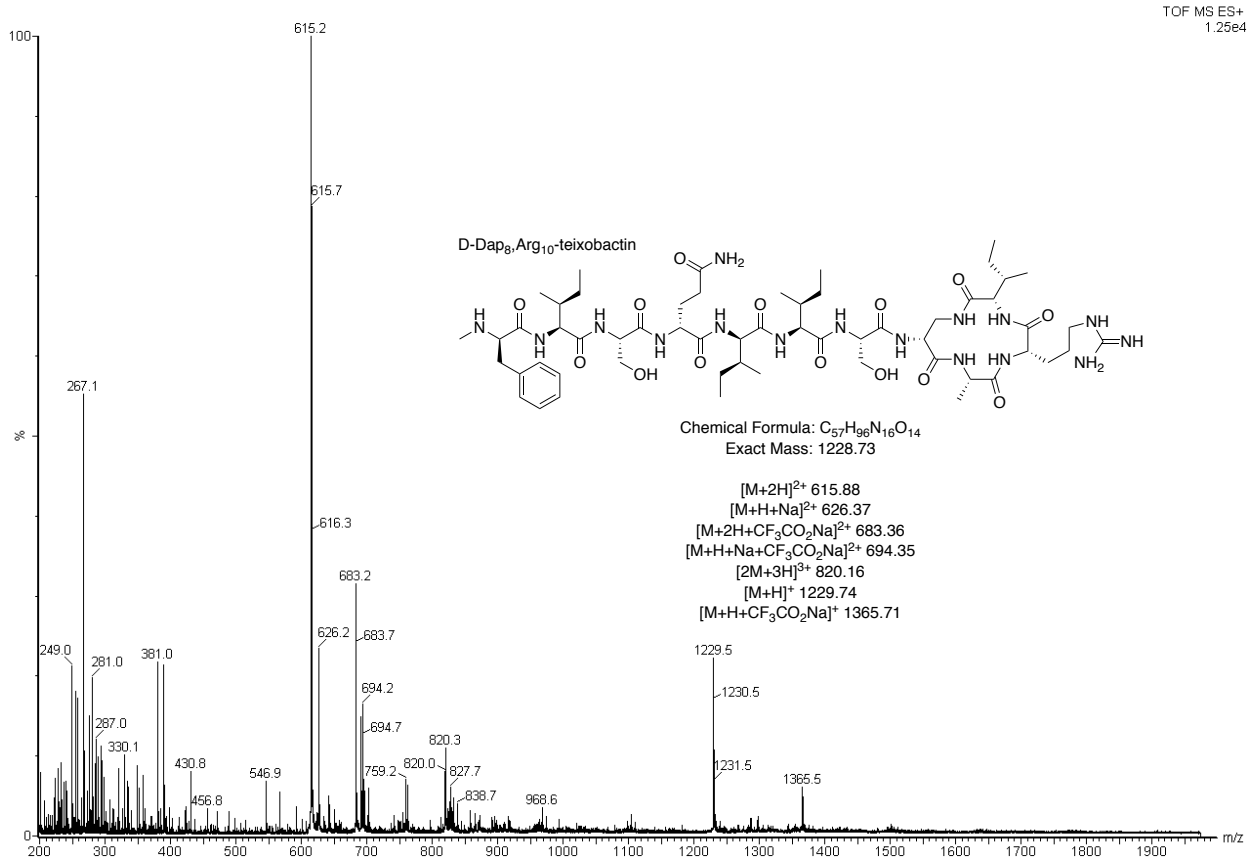


D-Dap$_8$,Arg$_{10}$-teixobactin : Analytical RP-HPLC and mass spectrum



Signal 1:VWD1 A, Wavelength=214 nm

Peak #	RT [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	9.386	MM	0.070	5032.440	100.000	100.000



Notes and References

[1] H. Yang, K, H. Chen, and J. S. Nowick, *ACS Chem. Biol.*, 2016, **11**, 1823-1826.

[2] N. Thieriet, J. Alsina, E. Giralt, F. Guibé and F. Albericio, *Tetrahedron Lett.,* **38**, 7275-7278.

[3] The procedure in this section is adapted from and in some cases taken verbatim from R. K. Spencer, A. G. Kreutzer, P. J. Salveson, H. Li and J. S. Nowick, *J. Am. Chem. Soc.*, 2015, **137**, 6304-6311 and A. G. Kreutzer, S. Yoo, R. K. Spencer and J. S. Nowick, *J. Am. Chem. Soc.*, 2017, **139**, 966-975.

[4] APEX2 Version 2014.11-0, Bruker AXS, Inc.; Madison, WI 2014.

[5] SAINT Version 8.34a, Bruker AXS, Inc.; Madison, WI 2013.

[6] G. M. Sheldrick, SADABS, Version 2014/5, Bruker AXS, Inc.; Madison, WI 2014.

[7] G. M. Sheldrick, SHELXTL, Version 2014/7, Bruker AXS, Inc.; Madison, WI 2014.

[8] International Tables for Crystallography 1992, Vol. C., Dordrecht: Kluwer Academic Publishers.

[9] S. Parsons, H. D. Flack, and Wagner, T. *Acta Cryst.*, 2013, **B69**, 249-259.

[10] A.L. Spek, *Acta Cryst.*, 2015, **C71**, 9-19.

[11] A. L. Spek, *Acta. Cryst.*, 2009, **D65**, 148-155