Supporting Information

Size Matters and How You Measure It: A Gram-Negative Antibacterial Example Exceeding Typical Molecular Weight Limits

Authors: Fiorella Ruggiu, Shengtian Yang, Robert L. Simmons, Anthony Casarez, Adriana K. Jones, Cindy Li, Johanna M. Jansen, Heinz E. Moser, Charles R. Dean, Folkert Reck, Mika Lindvall*

Novartis Institutes for BioMedical Research, 5300 Chiron Way, Emeryville, CA, 94608

Contents

I. Synthetic procedure and characterization of compound 1

General Conditions

Mass spectra were acquired on LC-MS, SFC-MS, or GC-MS systems using electrospray, chemical and electron impact ionization methods from a range of instruments of the following configurations: Waters ACQUITY UPLC system and equipped with a ZQ 2000 or SQD MS system where (M+1) refers to the protonated molecular ion of the chemical species, (M+) refers to the unprotonated quaternary ammonium cation, (M+Na) refers to the sodium-incorporated ion and (M-1) refers to the deprotonated molecular ion of the chemical species.

NMR spectra were run either on Bruker AVANCE 500MHz using ICON-NMR, under TopSpin program control or Varian 400MHz NMR spectrometers using VNMRJ program. Spectra were measured at 298K, unless indicated otherwise, and were referenced relative to the solvent resonance. Commercial reagents and solvents (anhydrous) were used as received without additional purification. Reactions were performed under an inert atmosphere unless noted otherwise.

Instrumentation

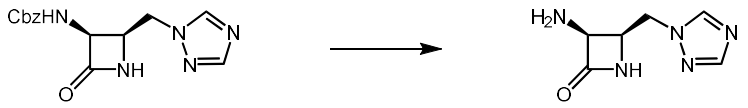
MS Methods: Using Agilent 1100 HPLC systems with an Agilent 6110 Mass Spectrometer

Method 2m_acidic:

Column	Kinetex C18 50 x 2.1 mm, 2.6 µm
Column Temperature	50 °C
Eluents	A: H_2O, B: acetonitrile, both containing 0.1% TFA
Flow Rate	1.2 mL/min
Gradient	2% to 88% B in 1.30 min, 0.15 min 95% B

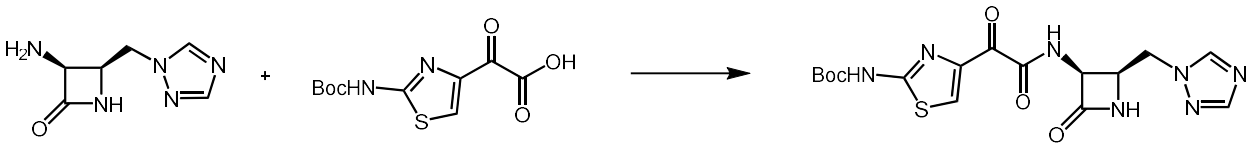
Synthetic procedure

(3*S*,4*R*)-4-((1*H*-1,2,4-triazol-1-yl)methyl)-3-aminoazetidin-2-one



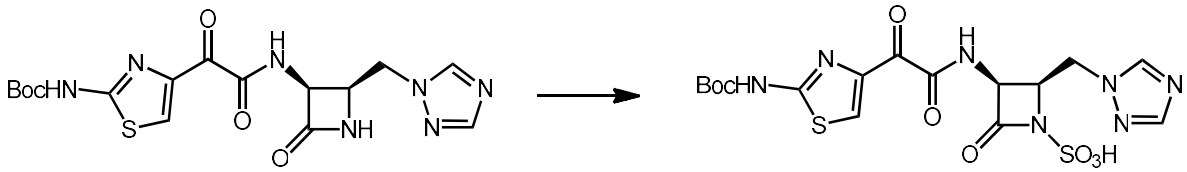
A flask was charged with benzyl ((2*R*,3*S*)-2-((1*H*-1,2,4-triazol-1-yl)methyl)-4-oxoazetidin-3-yl)carbamate (2.511 g, 8.33 mmol, synthesis described in US2015266867) and Pd-C (896 mg, 0.842 mmol) was purged with N_2 and the reagents were slurried in EtOH (80 mL)/MeOH (40 mL). The system was evacuated and backfilled with H_2 (3x), stirred vigorously for 2.5 h then purged with N_2 and filtered through celite. The eluent was concentrated in vacuo, dissolved in toluene and re-concentrated (2x). The crude residue was used directly without purification.

tert-Butyl (4-(2-(((2*R*,3*S*)-2-((1*H*-1,2,4-triazol-1-yl)methyl)-4-oxoazetidin-3-yl)amino)-2-oxoacetyl)thiazol-2-yl)carbamate.



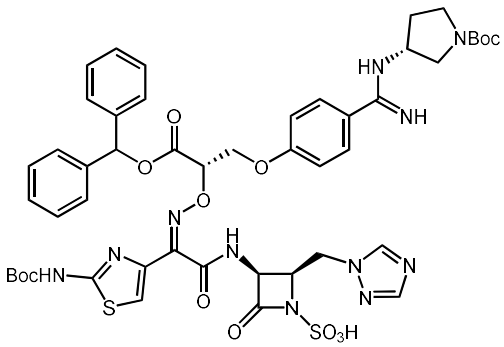
To a slurry of 2-(2-((*tert*-butoxycarbonyl)amino)thiazol-4-yl)-2-oxoacetic acid (2.72 g, 9.99 mmol) and HATU (3.80, 10.0 mmol) in DCM:DMF (3:1, 33.3 mL) at 0 °C was added DIPEA (2.91 mL, 16.7 mmol). A soln of (3*S*,4*R*)-4-((1*H*-1,2,4-triazol-1-yl)methyl)-3-aminoazetidin-2-one (1.39 g, 8.33 mmol) in DCM:DMF (1:1, 32 mL) was added followed by a DMF (3 mL) wash. After stirring for 48 h the dark solution was diluted with EtOAc (150 mL)/brine (140 mL) and the layers were separated. The aqueous was extracted with EtOAc (3x) and the combined organic layers were washed with brine (70 mL). The brine layer wash was re-extracted with EtOAc. The combined organic layers were dried over Na_2SO_4 and concd in vacuo. The crude residue was purified via silica gel chromatography (MeOH-DCM, 0-10%), affording the title compound (2.38 g, 68%) as a red solid. LCMS: R_t = 0.59 min, m/z = 422.0 (M+1) Method 2m_acidic; ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.85 (s, 1H), 9.70 (d, *J* = 9.3 Hz, 1H), 8.52 (d, *J* = 1.6 Hz, 1H), 8.48 (s, 1H), 8.44 (s, 1H), 7.96 (s, 1H), 5.28 (ddd, *J* = 9.3, 5.2, 1.5 Hz, 1H), 4.45 (dd, *J* = 14.2, 5.3 Hz, 1H), 4.36 (dd, *J* = 14.1, 7.6 Hz, 1H), 4.18 (dt, *J* = 7.6, 5.3 Hz, 1H), 1.47 (s, 9H).

(2*R*,3*S*)-2-((1*H*-1,2,4-triazol-1-yl)methyl)-3-(2-(2-((*tert*-butoxycarbonyl)amino)thiazol-4-yl)-2-oxoacetamido)-4-oxoazetidine-1-sulfonic acid.



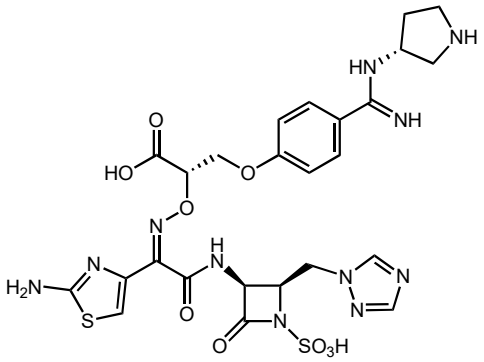
To a solution of *tert*-Butyl (4-(2-(((2*R*,3*S*)-2-((1*H*-1,2,4-triazol-1-yl)methyl)-4-oxoazetidin-3-yl)amino)-2-oxoacetyl)thiazol-2-yl)carbamate (200 mg, 0.475 mmol) in DMF (4.75 mL) at 0 °C was added SO_3•DMF (367 mg, 2.40 mmol). After 16 h of stirring it was concentrated in vacuo and purified with HP21 resin (ACN-water, 0-50%), affording the title compound (110 mg, 46%) as a pale yellow solid. LCMS: R_t = 0.54 min, m/z = 501.9 (M+1) Method 2m_acidic; ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.87 (s, 1H), 9.75 (d, *J* = 9.2 Hz, 1H), 8.94 (s, 1H), 8.44 (s, 1H), 8.34 (s, 1H), 5.25 (dd, *J* = 9.1, 5.4 Hz, 1H), 4.75 (dd, *J* = 14.3, 4.9 Hz, 1H), 4.61 (dd, *J* = 14.4, 7.6 Hz, 1H), 4.43 (dt, *J* = 7.6, 5.2 Hz, 1H), 1.49 (s, 9H).

(2*R*,3*S*)-2-((1*H*-1,2,4-triazol-1-yl)methyl)-3-((*Z*)-2-((((*S*)-1-(benzhydryloxy)-3-(4-(*N*-((*R*)-1-(*tert*-butoxycarbonyl)pyrrolidin-3-yl)carbamimidoyl)phenoxy)-1-oxopropan-2-yl)oxy)imino)-2-(2-((*tert*-butoxycarbonyl)amino)thiazol-4-yl)acetamido)-4-oxoazetidine-1-sulfonic acid.



To a solution of (2*R*,3*S*)-2-((1*H*-1,2,4-triazol-1-yl)methyl)-3-(2-(2-((*tert*-butoxycarbonyl)amino)thiazol-4-yl)-2-oxoacetamido)-4-oxoazetidine-1-sulfonic acid (60%, 267 mg, 0.320 mmol) in DCM-MeOH (1:1, 3.2 mL) was added *tert*-butyl (*R*)-3-(4-((*S*)-2-(aminooxy)-3-(benzhydryloxy)-3-oxopropoxy)benzimidamido)pyrrolidine-1-carboxylate (184 mg, 0.32 mmol, synthesis described in WO2013110643). After stirring at rt for 18 h it was concentrated in vacuo then dissolved in DCM and purified via silica gel chromatography (MeOH-DCM, 0-20%), affording the title compound (205 mg, 60%) as an off white solid. LCMS: R_t = 0.94 min, m/z = 1058.8 (M+1) Method 2m_acidic.

(*S*)-2-(((*Z*)-(2-(((2*R*,3*S*)-2-((1*H*-1,2,4-triazol-1-yl)methyl)-4-oxo-1-sulfoazetidin-3-yl)amino)-1-(2-aminothiazol-4-yl)-2-oxoethylidene)amino)oxy)-3-(4-(*N*-((*R*)-pyrrolidin-3-yl)carbamimidoyl)phenoxy)propanoic acid.



SIMMORO1-011-EXP044-001 (212 mg, 0.200 mmol) was dissolved in DCM (2.0 mL) and cooled to 0°C, whereupon TFA (772 µl, 10.0 mmol) was added drop wise. The solution was stirred at 0°C for 10 min then at rt for 2 h. It was diluted with DCM and concentrated under in vacuo. The crude residue was dissolved in water (6mL) and purified by reverse phase prep HPLC (Xselect CSH, 30 x 100 mm, 5 □M, C18 column; ACN-water with 0.1% formic acid modifier, 60 mL/min). LCMS: R_t = 0.39 min, m/z = 692.3 (M+1) Method 2m_acidic_polar; ^{1}H NMR (500 MHz, DMSO-d_6): δ 9.52 (d, *J* = 7.6 Hz, 3 H), 9.11 (br s, 1 H), 8.92-8.78 (m, 2 H), 8.40 (s, 1 H), 7.92 (s, 1 H) 7.67 (d, *J* = 8.8 Hz, 2 H), 7.30 (br s,

2 H), 6.90 (d, *J* = 9.1 Hz, 2 H), 6.80 (s, 1 H), 5.11-5.04 (m, 2 H), 4.51-4.58 (m, 1 H), 4.48-4.35 (m, 4 H), 4.25-4.19 (m, 1 H), 3.53-3.44 (m, 1 H), 2.35-2.26 (m, 1 H), 2.29-2.11 (m, 1 H).

Hydrolytic stability $t_{1/2}$ 226h.

High Resolution Mass Spectrometry by LC-MS of compound 1

ESI-MS data were recorded using a LTQ-XL Orbitrap mass spectrometer (ThermoFisher Scientific) with electrospray ionization source. The resolution of the MS system was approximately 30000. The drug candidate was infused into the mass spectrometer by UPLC (Acquity, Waters) from sample probe. The separation was performed on Acquity UPLC BEH C18 1x50 mm column at 0.15 mL/min flow rate with the gradient from 5% to 95% in 3 min. Solvent A was Water with 0.1% Trifluoroacetic acid and solvent B was 75% Methnol and 25% Isopropyl alcohol with 0.1% Trifluoroacetic acid. The mass accuracy of the system has been found to be <5 ppm.

Structure Name	Formula	Isotopic Mass	calculated mass for MH+	measured mass for MH^+	error(ppm)	RT(min)
Compound 1	C25H29N11O9S2	691.1591	692.1664	692.1662	-0.3	0.47

II. 1D ^{1}H NMR and 2D ^{1}H-^{1}H NOESY NMR

All NMR spectra were recorded using a Bruker AVANCE III-500 NMR spectrometer operating at a frequency of 500.08 MHz for ^{1}H, 125.76 MHz for 13C, 50.68 MHz for 15N. The instrument was equipped with a 5 mm broadband (BBFO) Cryo-probe with a Z-gradient. Deuterium Oxide was used as the NMR solvent for all the experiments. Chemical shifts for ^{1}H spectra were referenced to the D_2O solvent peak at 4.79 ppm. For the 2D-NOESY experiments, each spectrum was acquired at 1024x256 size, 64 scan, 2 sec delay time, and at 300ms, 500ms, and 800ms mixing time each. All spectra were recorded at a temperature of 296K.

Figure S1. Compound 1 structure with annotated protons corresponding to NMR spectra

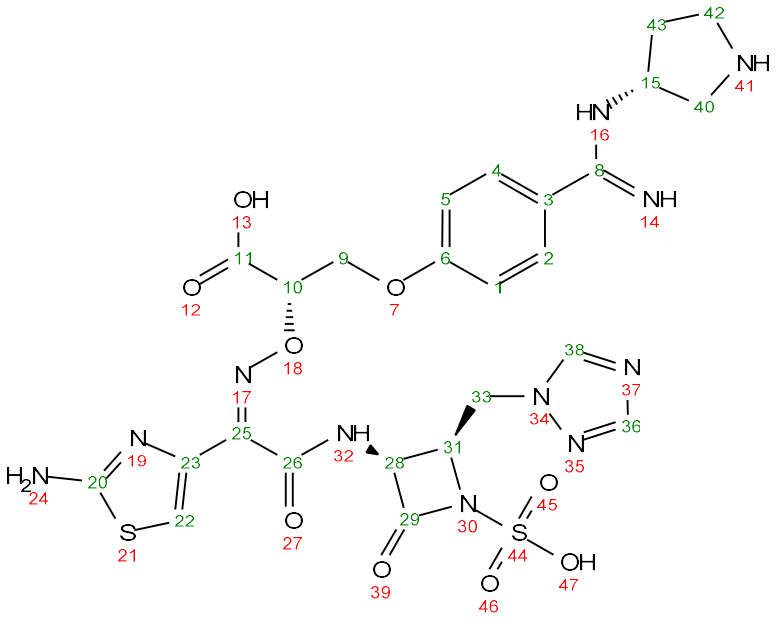
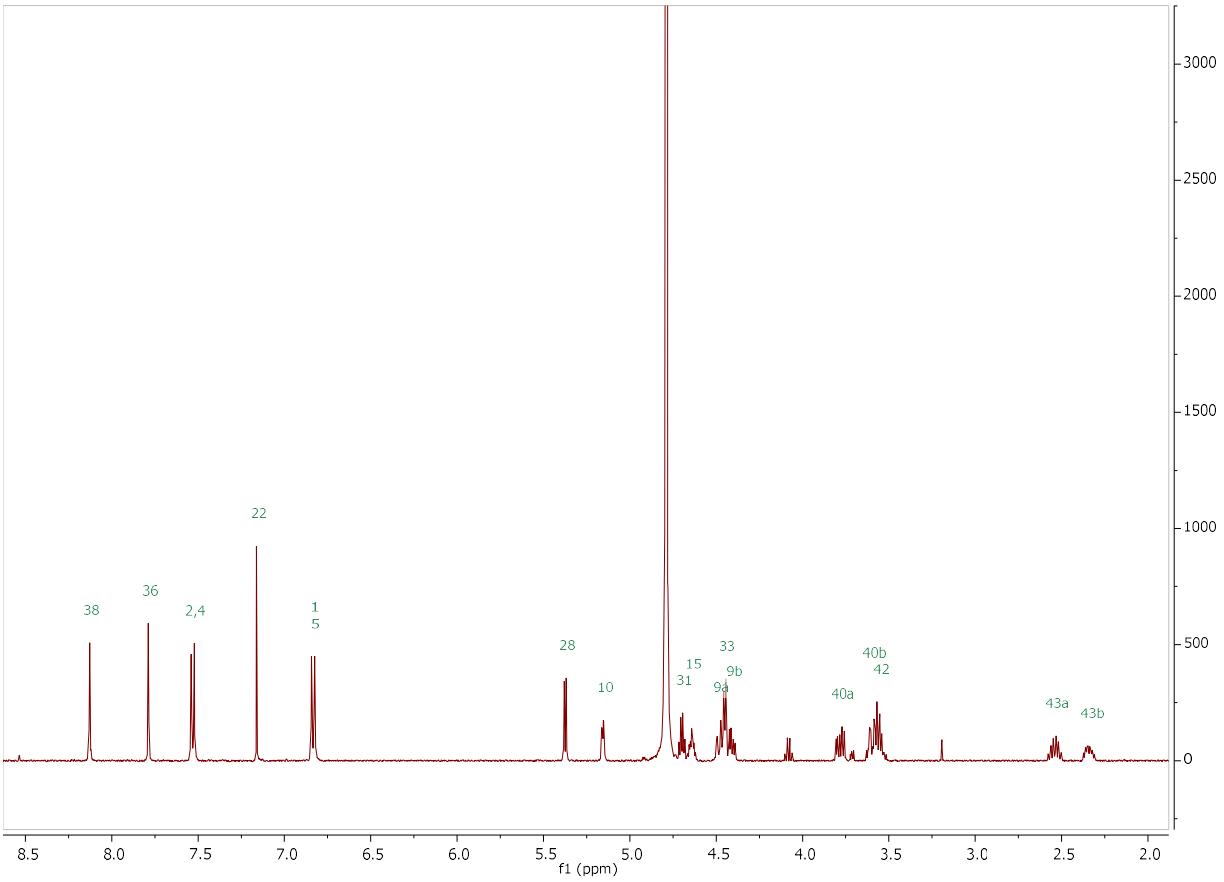


Figure S2. [1]H-NMR spectra of compound 1



The H1 NMR shows sharp peaks of expected splitting pattern and no signs of overlaps of similar peaks, suggesting it is a monomer and that the dimer is unlikely to exist.

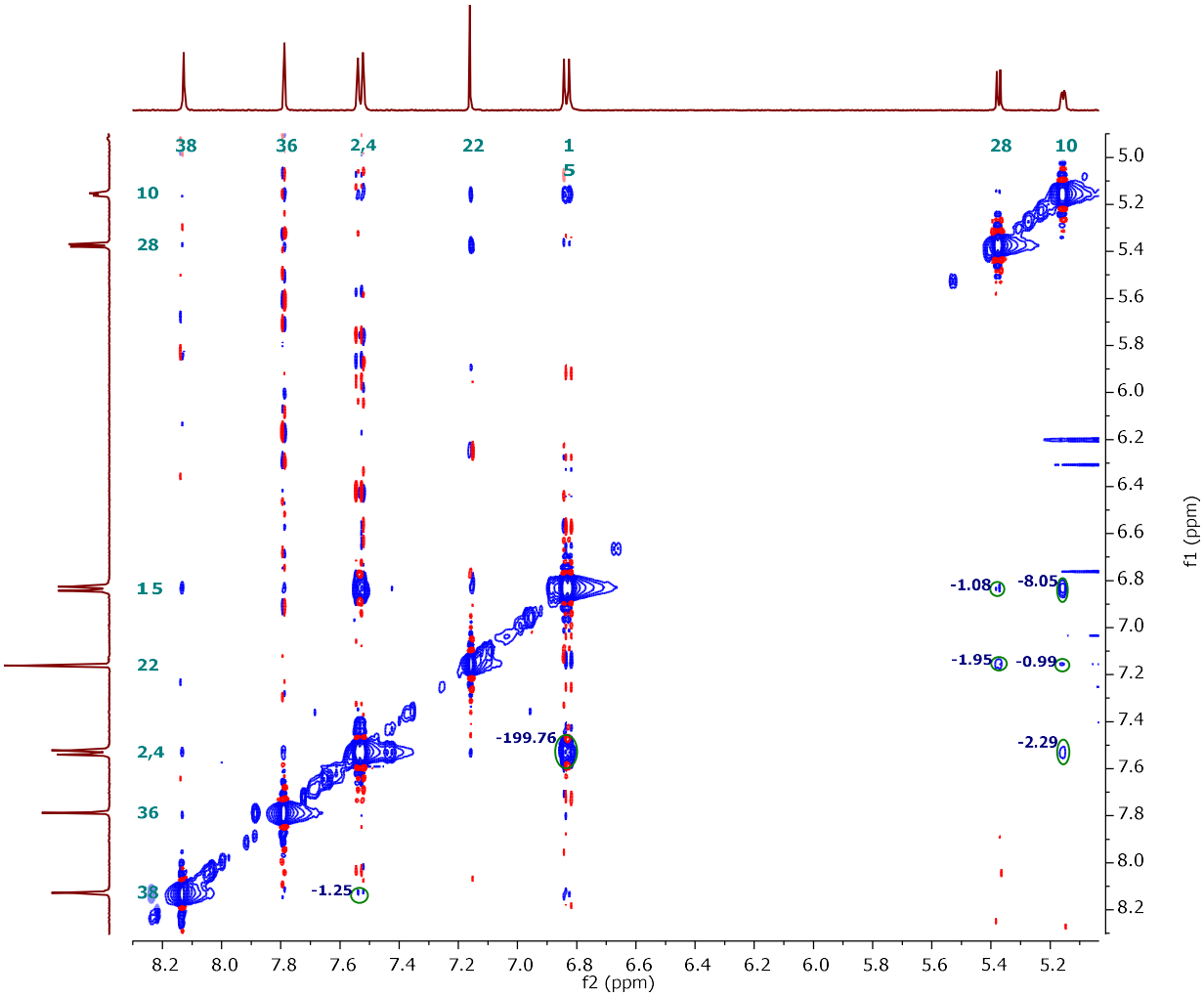


Figure S3. 2D-NOESY NMR spectra in water (D_2O) of compound 1

2D-NOESY NMR experiment was done at different mixing times. The long range distances signals, restraint 1 (H2,H4 to H38) and restraint 2 (H28 to H1,H5), are observed at 2 different ones, indicating that these signals are not noise.

The intensity of signal between proton H33 and H28 is set at 100. The distance between those very fixed protons is estimated at 2.5Å. Assuming a r^{-6} decay of the signal and the intensities of the long range distances to be 1, then the estimated distance between those long range distance is ~5.4Å. Those distances usually have a high error bar ~1Å. We decided to look for conformers where the corresponding distance for restraint 1 and 2 is ≤ 7Å.

Table S1. Observed 2D-NOESY signals at different mixing times

		Mixing times				
Restraint	Protons	300ms	500ms	800ms	1D	
	1,5 to 2,4	224.17	199.76	180.77	1	
1	2,4 to 38	1.78	1.25		s	
2	28 to 1,5		1.08	1.27		

	28 to 22	2.02	1.95	2.23	s	
	10 to 1,5	8.22	8.05	8.94	s	
	10 to 2,4	1.21	0.99	2.95	n	
	9a to 1,5	83.14	60.35	48.69		
	9b to 1,5	75.35	53.06	42.09		
	9a to 2,4	7.22	3.79	6.97	s	
	9b to 2,4	4.56	3.57	4.12	s	
	33 to 36	1.54		0.97		
	33 to 38	22.97	15.89	10.5	L	
	9a to 10	59.73	43.15	32.28		
	9b to 10	69.33	54.23	38.01		
	31 to 28	100	100	100		Reference intensity set to 100
	33 to 28	9.69	8.28	10.41		The diagonal is set negative and the NOE peaks also negative.
	33 to 31	99.97	49.71	38.24		
	2,4 to 36				s	
	2,4 to 22				m	

III. MIC testing

MIC testing was performed using CLSI broth microdilution according to: CLSI M7-A10. 2015. Methods for dilution susceptibility tests for bacteria that grow aerobically; approved standard, tenth edition CLSI document M7-A10. Clinical and Laboratory Standards Institute, Wayne, PA

Table S2. CLSI M7-A10, 2015, approved MIC QC ranges

CLSI Approved MIC QC Ranges (µg/mL):		
Antibiotic	E. coli ATCC 25922	P. aeruginosa ATCC 27853
Aztreonam	0.06-0.25	2-8
Cefepime	0.015-0.12	0.5-4
Ceftazidime	0.06-0.5	1-4
Chloramphenicol	2-8	-
Ciprofloxacin	0.004-0.015	0.25-1
Tigecycline	0.03 - 0.25	-
Cefiderocol	0.06-0.5	0.06-0.5

IV. Moldiscovery Moka 2.6.5 pKa predictions of compound 1

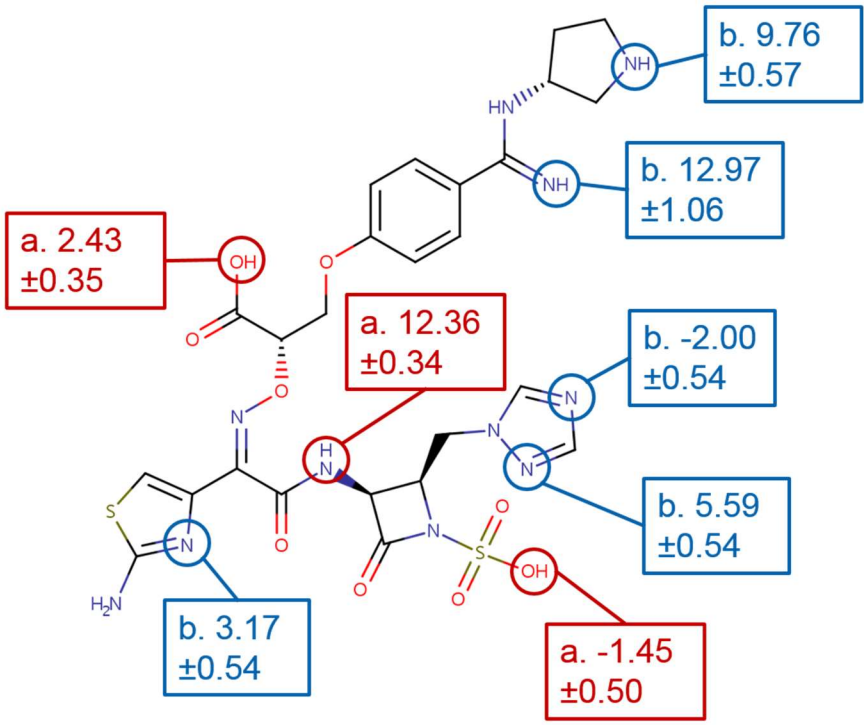


Figure S4. pKa prediction from Moldiscovery moka 2.6.5 internal pKa model. The type of center is indicated with an "a" for acidic (in red) and "b" for basic (in blue), followed by the pKa prediction and the standard deviation for the prediction.

V. CCG MOE conformer generation and crystal structure overlay

Conformer generation and correspondence with restraint results

Conformers were generated with a svl script: file_confsearch.svl (Scientific Vector Language (SVL) source code provided by Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2018.) and the following commands:

- With electrostatic term:
  moebatch -exec
  "file_confsearch['compound1_pKa.sdf','sdf','compound1_e80_MOEEle_100.mdb','mdb',
  [maxconf:100, solDielectric:80,eleEnable:1]]"
- Without electrostatic term:
  moebatch –exec
  "file_confsearch['compound1_pKa.sdf','sdf','compound1_e80_MOEnoEle_100.mdb','mdb',
  [maxconf:100, solDielectric:80,eleEnable:0]]"

The resulting SDFs were then used as input to check_restraint_dist.py to calculate restraint distance and flag conformers with both restraints ≤ 7Å. The script uses the RDKit 2018.03.4.0 library and was run in Anaconda 3.5-4.0.0. It superposes a structure encoded in SMILES where the hydrogen atoms are manually labeled to the conformers. It then calculates the restraints by averaging the distances between H2/H4 and H38 for restraint 1 and the distances between H1/H5 and H28 for restraint 2.

6 out of the 32 conformers generated when using the electrostatic term had both restraints ≤ 7Å but only one where their Boltzmann population >1% and 48 out of the 100 conformers generated without using the electrostatic term had both restraints ≤ 7Å but only 3 where their Boltzmann population >1% (See Table S3). Note: Boltzmann populations were calculated separately for each conformer generation strategy, i.e. with and without electrostatic term.

Table S3. Conformer generation and restraints details

Conf. nb.	Electrostatic term used in conformer generation	Boltz-mann population (%)	Restraint 1 (Å)	Restraint 2 (Å)
1	True	16	6.7	6.5
2	False	32	5.9	6.3
3	False	21	5.8	6.1
4*	False	2	6.4	6.9

* conformer fitting in *E.coli* OmpF crystal structure 4gcp when superimposed with ampicillin

Conformer overlay with crystal structure 4gcp

E.coli OmpF crystal structure PDB ID 4gcp was loaded into MOE2018.01 and prepared using Compute → Protonate 3D. The β-lactam + amide motif was then selected and Compute → Molecule → Superpose was used to superpose the conformer molecular databases (loaded from the SDF conformer files).

VI. Cheminformatics script for size evaluation and dipole moment calculations

The conformers were saved in the mol2 format from MOE to export coordinates and partial charges from the EHT:Amber10 force field. The cheminformatics script min_dim_molecule.py (available on github https://github.com/fio-ruggiu/G-monobactam_size) was used to find the minimal area rectangle of the projections and output the minimum dimensions along with the angle. The python script uses RDKit 2018.03.4.0, numpy 1.15.1, pandas 0.23.4 (and scipy 1.1.0) and was run in Anaconda 3.5-4.0.0. The cartesian coordinate system (x,y,z) is centered at the molecules' center. The molecule is then rotated in two dimensions (x, y) by increments of 5º and the x, y projections in the third dimension (z) are measured taking into account the van der Waals radii of the atoms. The minimal area of the rectangle formed by the projected dimensions is identified and corresponding area, length of molecules and angle between dipole moment and z-axis are outputted. The dipole moment is calculated from the partial charges in the mol2 file. Results for the four stable conformers in agreement with the restraints are reported in Table S4.

Table S4. Stable conformers' geometric parameters

Conf. nb.	Minimum rectangle area ($Å^2$)	Minimum dimension 1 (Å)	Minimum dimension 2 (Å)	Angle perpendicular to minimum rectangle (º)	Dipole moment magnitude from EHT:Amber10 (D)	Dipole moment magnitude from AM1 (D)
1	100	8.7	11.5	52	4.2	12.7
2	108	9.5	11.4	9	15.2	36.1
3	108	8.9	12.1	10	15.9	36.1
4*	109	9.7	11.3	27	12.3	33.8

*conformer fitting in *E.coli* OmpF crystal structure 4gcp when superimposed with ampicillin