CHEMBIOCHEM

Supporting Information

A Defined and Flexible Pocket Explains Aryl Substrate Promiscuity of the Cahuitamycin Starter Unit–Activating Enzyme CahJ

Ashootosh Tripathi,[a, e] Sung Ryeol Park,[a, h] Andrew P. Sikkema,[a, b, i] Hyo Je Cho,[d] Jianfeng Wu,[c] Brian Lee,[a] Chuanwu Xi,[c] Janet L. Smith,[a, b] and David H. Sherman*[a, e, f, g]

cbic_201800233_sm_miscellaneous_information.pdf

Author Contributions

A.T. Conceptualization: Equal; Data curation: Equal; Investigation: Equal; Writing – original draft: Lead; Writing – review & editing: Lead
S.P. Conceptualization: Equal; Data curation: Equal; Investigation: Equal
A.S. Conceptualization: Equal; Data curation: Equal; Investigation: Supporting; Writing – original draft: Supporting.

Experimental Procedures

General experimental procedures. Optical rotation measurements were obtained on a Perkin-Elmer 241 Polarimeter calibrated using a Rudolph Quartz Control Plate Calibration Standard at sodium D line (at +11.502°). Ultraviolet spectra were obtained on a UV-visible Molecular Devices SpectraMax M5 spectrophotometer using 1-ml cuvettes with 1.0 cm path lengths at room temperature in solvent methanol (MeOH). Spectrophotometric assays were performed on Molecular Devices SpectraMax M5 384 variable wavelength spectrometer. All NMR spectra were acquired on a Varian INOVA 600MHz spectrometer at the NMR Facility, Life Sciences Institute, University of Michigan. HRESIMS spectra were measured at the University of Michigan core facility in the Department of Chemistry using an Agilent 6520 Q-TOF mass spectrometer equipped with an Agilent 1290 HPLC system. RP-HPLC was performed using Econosil C18 10 mm 22×250-mm column and Agilent ZORBAX RX-C8 5 mm 9.4×250-mm column and a solvent system of MeCN and H_2O. The LC-MS analysis of HPLC fractions was performed on a Shimadzu 2010 EV APCI spectrometer.

Protein Expression, Purification. The *cahJ* gene was amplified by PCR from genomic DNA of *S. gandocaensis* using the primer SR235 (5'-GCCGCGCGGCAGCCATATGCTCGATGGGTGGGTG-3') and SR236 (5'-TCGAGTGCGGCCGCAAGCTTCAGTGGACACCGCCGTC -3'). Cloning of *cahJ* gene into pET28b (Novagen) for expression of N-terminal His_6-tagged CahJ was performed using Gibson assembly [1] master mix (New England Biolabs). The amplified PCR product of *cahJ* and pET28b, which had been cut by *Eco*RI and *Hind*III, were incubated at 50°C for 2 h. Following incubation, the samples were transformed into *E. coli* DH5α and isolated. The accuracy of *cahJ* in pET28b was confirmed by restriction digestion and sequencing. Synthetic codon-optimized DNA encoding the CahA ArCP was purchased from IDT, inserted into the pMCSG7 vector using LIC cloning [2], and verified by sequencing.

Protein Expression and Purification. Recombinant plasmid pET28b-*cahJ* was utilized to transformed into *Escherichia coli* BL21(DE3)/pRARE cells. The resulting transformant was grown at 37 °C in 1 liter of LB broth containing kanamycin (50 μg/ml) until OD_{600} reached 0.6~1.0. Then, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.2 mM to induce gene expression and cells were cultured at 18°C overnight. Cells were harvested by centrifugation at 5,000 × g for 10 min at 4°C to harvest cells. The cell pellet was resuspended in 70 ml of lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, 10% glycerol, 0.1mM EDTA (pH8.0), 0.2mM TCEP, pH7.6) and lysed by sonication. Cell debris was removed by centrifugation for 20 min at 45,000 × g. The supernatant was mixed with 1 ml of Ni-NTA agarose (Qiagen) for 1 h at 4 °C and loaded onto an empty column. The column was washed with 50-100 ml of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, 10% glycerol, 0.2mM TCEP, pH7.6). The bound His_6-tagged proteins were eluted with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, 10% glycerol, 0.2mM TCEP, pH 8.0). The CahJ (59 kDa) proteins were further purified and concentrated with 30K size exclusion filters (Amicon). Buffer exchange with storage buffer (50 mM NaH_2PO_4, 10% glycerol, pH7.3) using a PD-10 column (GE Healthcare) was used to lower the sample's salt concentration. Protein concentration was determined using NanoDrop spectrophotometer. The molar extinction coefficient for CahJ at A_{280} is 52,370 $M^{-1}cm^{-1}$.

Expression and Purification of CahA ArCP and CahJ for crystallization. Cultures were grown in either Terrific Broth (TB) medium containing 4% glycerol or in M9 minimal medium supplemented with nutrient mix (AthenaES) and L-methionine (Sigma) as indicated. Ampicillin (100 μg/mL, pMCSG7), spectinomycin (50 μg/mL, pRARE), and kanamycin (50 μg/mL, pET28b) were used with the corresponding plasmids in *E. coli* cultures. *E. coli* strain BAP1 [3] was transformed with pMCSG7-*cahA_ArCP*, and strain BL21-AI (Thermofisher) was transformed with pET28b-*cahJ*. Both

strains contained the pRARE plasmid [4]. Cultures were grown at 37°C in 500 mL TB with 4% glycerol to an OD_{600} of 1.5, cooled to 20°C for at least 1 hr, induced with IPTG (final concentration 0.2 mM), and grown for an additional 14-18 hr before harvesting by centrifugation. Cell pellets were stored at -20°C until used for protein purification. Minimal media cultures were moved to 20°C at an OD600 of 0.6, but were otherwise treated identically to rich media cultures.

Cell pellets were resuspended in approximately 5 mL Buffer A (50 mM HEPES pH 7.6, 250 mM NaCl, 5% glycerol) containing 15 mM imidazole per 1 g cell paste. All purification steps were performed at 4°C. Resuspended cells were incubated with DNase (4 mg), lysozyme (10 mg), and $MgCl_2$ (4 mM) for 30 min, then lysed by three passes through an Avestin EmulsiFlex-C3 homogenizer. The lysate was clarified by centrifugation at 30,000xg in a Sorvall F21-8x50y rotor. The soluble fraction was filtered through 0.45 µm and 0.22 µm filters and then loaded onto a 5-mL HisTrap Ni NTA column (GE Healthcare) and eluted with an imidazole gradient up to 300 mM in Buffer A. The eluate was concentrated to approximately 5 mL using a 50 kDa Amicon Ultra-15 membrane, centrifuged at 20,000xg for 30 min, then loaded onto a HiLoad 16/600 Superdex 200 column (GE Healthcare) equilibrated with Buffer B (20 mM Tris pH 7.5, 150 mM NaCl, 10% glycerol) and isocratically eluted. Fractions were pooled, concentrated as above to 20-30 mg/mL, centrifuged at 20,000xg for 30 min, flash frozen in 20-50 µL aliquots with liquid N_2, and stored at -80°C. CahJ co-purified with bound SA adenylate from cells grown in rich media and with bound benzoyl adenylate from cells grown in minimal media. These co-purifying ligands were removed from CahJ by inserting the following steps while the protein was bound to the HisTrap Ni NTA column, prior to elution with imidazole: 10 CV linear gradient of 0 to 500 mM urea in Buffer A; 5 CV wash; 10 CV linear gradient of 500 mM to 0 urea in Buffer A containing 1 mM AMP; and 5 CV wash with Buffer A containing 30 mM imidazole.

CahA ArCP was purified identically to CahJ except that the buffer was 100 mM Tris pH 7.5, 500 mM NaCl, 10% glycerol, 5 mM TCEP. The eluate was concentrated using a 10 kDa Amicon Ultra-15 to approximately 10 mg/mL. CahA was then incubated for 16 hr at 30°C with the phosphopantetheinyl transferase Sfp [5] at a 1:40 ratio in a reaction mix containing 16 mM ATP, 4 mM pantethine, 4 mM DDT, and 0.2 mg/mL of CoaA, CoaD, and CoaE [6]. After incubation, protein was centrifuged at 20,000 xg for 30 min and then loaded onto a HiLoad 16/600 Superdex 75 column (GE Healthcare) equilibrated with 100 mM Tris pH 7.5, 250 mM NaCl, 5% glycerol, 5 mM TCEP and isocratically eluted. Fractions were pooled, concentrated to 10 mg/mL, centrifuged at 20,000 xg for 30 min, flash frozen in 20-30 µL aliquots with liquid N_2, and stored at -80°C.

Crystallization. Frozen CahJ stock solution was thawed on ice and glycerol was removed by buffer exchange into 20 mM Tris pH 7.5, 150 mM NaCl using a 50K Amicon Ultra 0.5 membrane. After addition of the appropriate substrates (final concentration for all 1.5 mM, Table 1), the mixture was diluted to a final concentration of 20 mg/mL CahJ, incubated on ice for 30 min, and clarified by centrifugation at 20,000xg for 30 min at 4°C. Crystals grew in approximately one week at 20°C by vapor diffusion in 1.5-µL sitting drops containing equal volumes of protein stock and reservoir solution (0.08-0.14 M sodium cacodylate pH 6.5, 1.5 - 2.0 M sodium acetate). Crystals were harvested in nylon loops, cryoprotected in reservoir solution containing 20% glycerol, and flash cooled in liquid N_2.

Diffraction Data Collection and Structure Determination. Diffraction data were collected at 100 K on GM/CA beamline 23ID-D at the Advanced Photon Source (APS) at Argonne National Laboratory (Argonne, IL). For each crystal, 360º of data were collected at an X-ray energy of 12 keV with 0.2º rotation and 0.2 sec exposure per image. Data were processed using XDS [7]. The structure of CahJ with bound AMP and salicylic acid was solved by molecular replacement (MR) with Phaser [8] in the PHENIX [9] software suite from the DhbE structure [10] (58% identity, PDB code: 1MD9) with the C-terminal domain removed. The initial model from MR was modified with AutoBuild [11] to generate a 75% complete model of CahJ. The remaining portion of model was completed manually using Coot [12]. Refinement was performed using phenix.refine [13]. Structures of

CahJ with other ligands were solved by difference Fourier methods as the crystals were isomorphous. Restraints for all substrates, except AMP, were generated using eLBOW [14]. Images and figures were prepared using PyMOL [15].

Determination of Substrate Specificity and Kinetic Parameters. The substrate selectivity of CahJ was determined using a non-radioactive colorimetric assay [16]. All assays were performed in 384-well reaction plates. To perform the assays, benzoic acid substrates were prepared in DMSO and 10 µl was transferred in triplicate to a 384-well reaction plate. Reactions were initiated by adding the mixture (40 µl) containing recombinant CahJ (6 µM), inorganic pyrophosphatase (0.2 U/ml), ATP (0.5 mM), NaCl (100 mM), $MgCl_2$ (10 mM), substrates (20 µM), and 50 mM Tris–HCl (pH 7.5) and incubated at RT. The reactions were terminated after 10 min by the addition of molybdate/malachite green reagent (10 µl). After 10 min of color development, the released P_i concentration was monitored at 600 nm on a microplate reader (SpectraMax®). For the determination of K_m and K_{cat} for formation of SA adenylate (SA-AMP) and 6-MSA adenylate (6-MSA-AMP), reactions were done with varying concentrations of SA and 6-MSA (1.25 µM, 2.5 µM, 5.0 µM, 10.0 µM, and 20.0 µM) in the same condition as described above. The reactions were incubated at RT and stopped after 0, 5 and 10 min by the addition of molybdate/malachite green reagent (10 µl). A reaction mixture without the recombinant CahJ domain was used as a negative control. The experiments were carried out in triplicate for each substrate concentration.

CahJ Mass Spectrometry Assays with ArCP. Reactions in 25 µL mixtures (50 mM Tris pH 7.8, 100 mM NaCl, 10 mM $MgCl_2$, 500 µM ATP, 2 µM CahJ, and 20 µM ArCP) were initiated by addition of 500 µM substrate in DMSO (final concentration 0.5% DMSO), incubated at 25°C for 1 hr, quenched by the addition of 10% formic acid, and stored at 4°C prior to analysis. 1 µL of reaction mixture was used for LC-MS analysis (Agilent 6545 Q-TOF equipped with Agilent 1290 HPLC system). Samples were separated for intact protein analysis by RP-HPLC (Aeris widepore C4 column 3.6 µM, 50 X 2.10 mm) with a 10 min gradient of 10-95% acetonitrile with 0.1% formic acid at a flow rate of 0.5 mL/min. Data were analyzed using the Agilent Mass Hunter Qualitative Analysis software with the maximum entropy deconvolution algorithm. A reaction mixture without CahJ was used as a negative control.

Generation of Cahuitamycin Analog through Mutasynthesis. Mutant strain of Δ*cahI* (DHS334) was first pre-cultured in 3ml R2YE liquid medium for 15 days at 28 °C and then 3ml of the seed culture was used to inoculate 100 ml of the same medium, followed by cultivation for 15 days at 28 °C. Salicyclic acid (as positive control) and 4-methylsalicylic acid were added every alternate day to separate 100 mL cultures of Δ*cahI* at a final concentration of 500 µM for 10 days. The products were first extracted using Amberlite XAD-16. The resin was separated and subjected to organic extraction using MeOH:EtOAc (1:1) for LC-MS analysis as described above.

The substrate-fed crude extract was then further purified by RP-HPLC on a gradient of 10–75% ACN and was followed by ultraviolet–visible photodiode array detection at 215 nm to yield semi-pure compounds **1** (4.3 mg). Compounds were again subjected to re-purification over RP-HPLC on isocratic condition of 35% MeOH (0.1% FA) using C-8 column to get compounds **1** (2.1 mg).

Cahuitamycin F (**1**). Amorphous powder; ultraviolet (ACN: H_2O) λ_{max} 204, 240, 252 and 304 nm; ^1H and ^{13}C NMR, see Table S2; HRESIMS m/z 650.2706 $[M+H]^+$ (calculated for $C_{28}H_{40}N_7O_{11}$, 650.2786).

Minimum Inhibitory Concentration Test. Cultures of *A. baumannii* ATCC 17978 grown overnight were diluted into fresh 10% Mueller Hinton II broth to get 105 colony-forming units per ml by determining optical density at OD_{600}. A volume of 100 ml of diluted cultures were incubated in individual wells of 96-well plate containing 1% of DMSO and various concentrations of cahuitamycins A (**3**) and F (**1**). The plate was incubated at 30°C with shaking at 150 r.p.m. The optical density of each well was monitored every 1 hr up to 24 hrs using a microtiter plate reader (Synergy HT, Bio-Tek, Winooski, VT). The number of viable bacterial cells was determined by counting colony-forming units.

Biofilm Dispersal Assay. The biofilm dispersal assay was performed as previously reported [17]. Briefly, *A. baumannii* ATCC 17978 biofilm was developed as described in minimum inhibitory concentration test. After 16 h incubation at 30 °C, the suspension in each well was removed by pipetting gently. A volume of 100 ml of 1% PBS buffer containing cahuitamycins A (**2**) and F (**1**) at various concentrations was placed into each well. The plate was incubated at 30 °C for 12 h to allow cahuitamycins to disassemble pre-formed biofilms. Then, the suspension was removed by pipetting. The remaining attached biofilm was quantified by crystal violet assay.

.

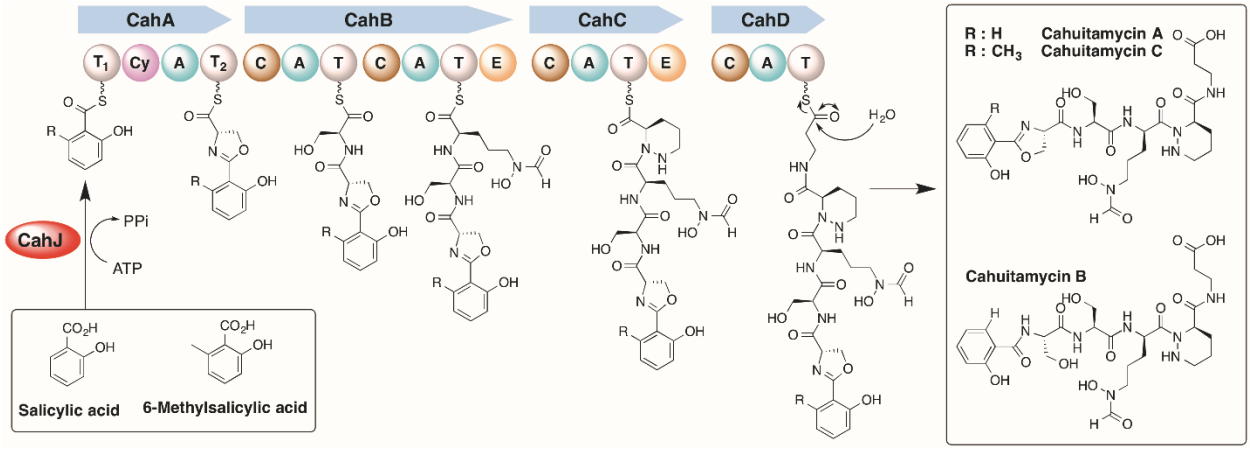

Figure S1: Proposed biosynthesis of cahuitamycins A-C in *S. gandocaensis*. The CahJ-dependent adenylation of salicylic acid and 6-methylsalicylic acid in cahuitamycin assembly is an ATP-dependent process that leads to the release of pyrophosphate (PP_i). C, condensation domain; Cy, cyclization domain; E, epimerization domain T, thiolation domain.

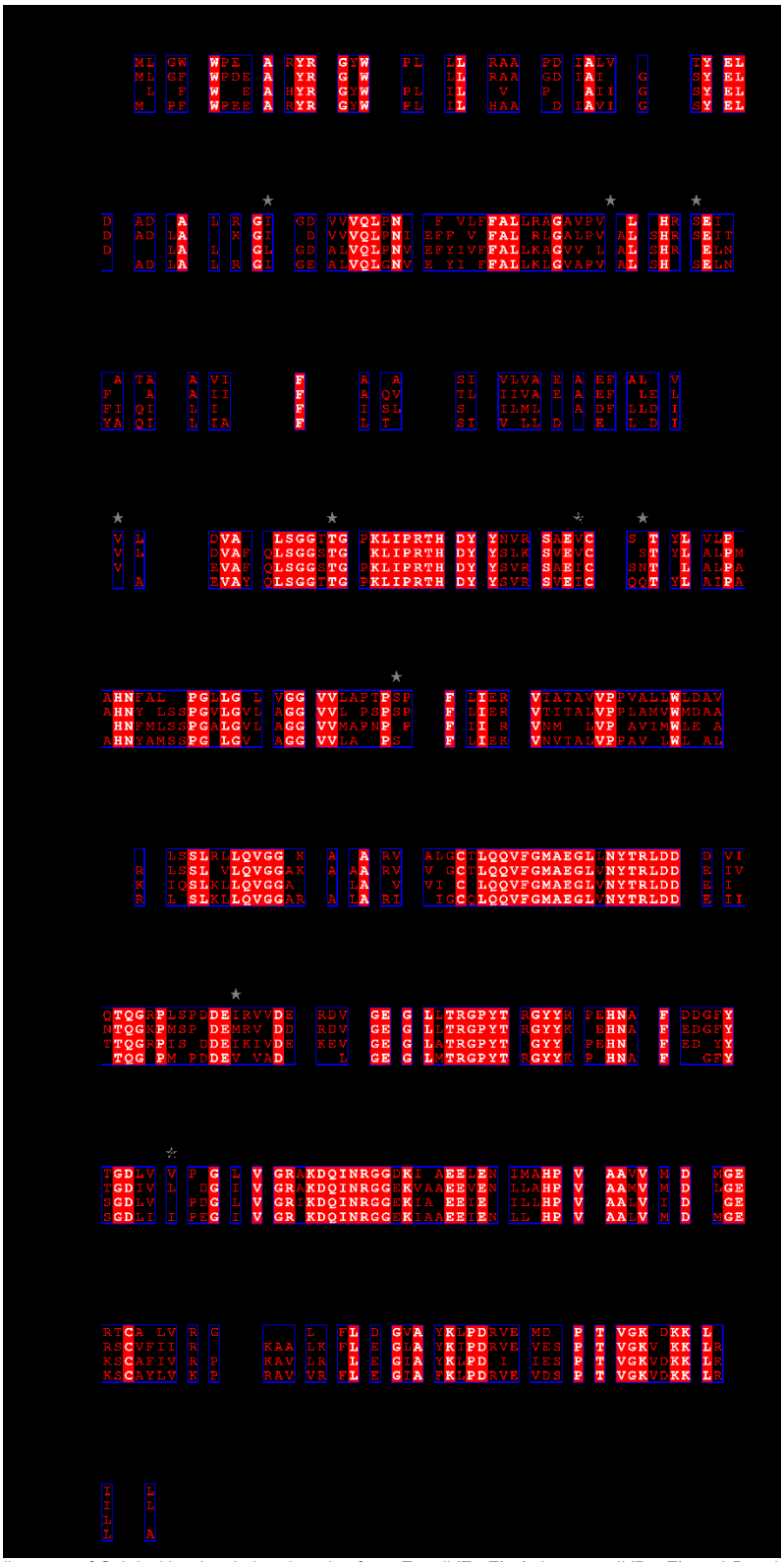

Figure S2. Structure based sequence alignment of CahJ with adenylation domains from *E. coli* (EntE), *A. baumannii* (BasE), and *B. subtilis* (DhbE). Conserved residues are shown in white on a red background and similar residues are shown in red font. Secondary structure for CahJ is shown above the alignment (α for α-helix and β for β-strand, TT for β-turn and η for 3_{10} helix).

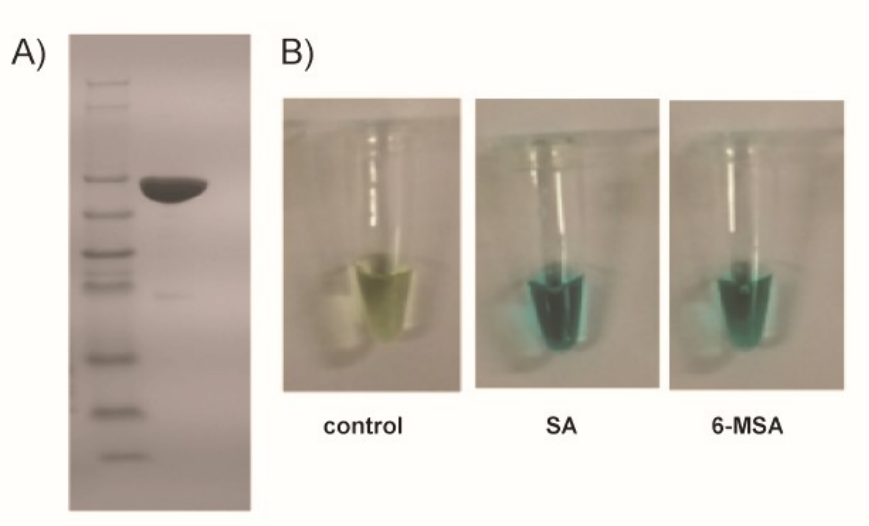

Figure S3: A) Coomassie blue-stained SDS-PAGE gel ran in Bis-Tris buffer for CahJ. B) Representative data of CahJ activity for SA and 6-MSA by malachite green assay. SA and 6-MSA samples show an obvious color change from yellow to green compared to blank control.

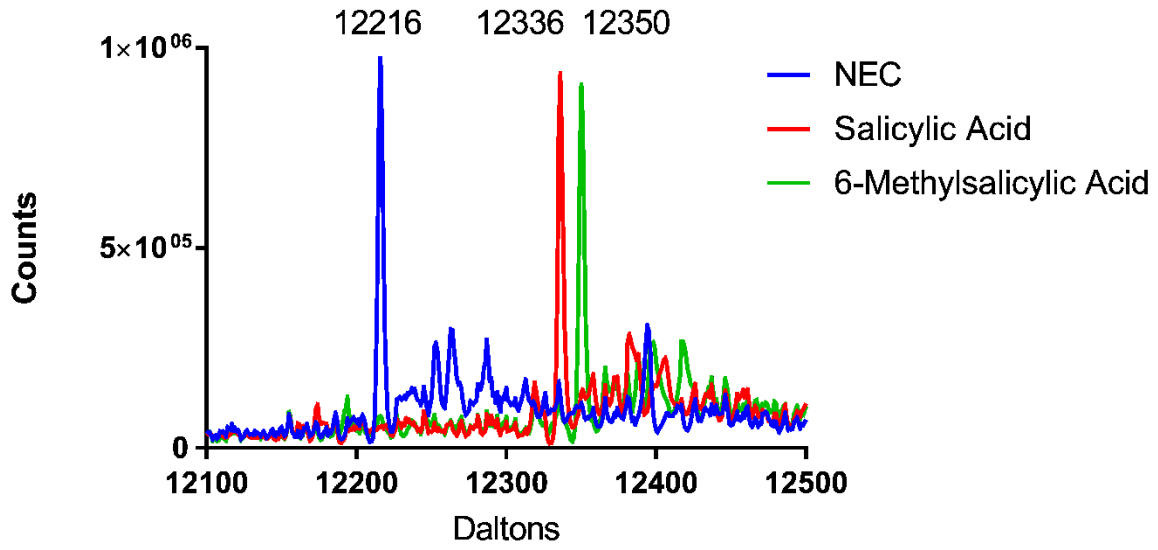

Figure S4: Intact protein mass spectra of CahJ loading of SA and 6-MSA onto CahA ArCP. CahJ was incubated 1 hr with either SA or 6-MSA and ArCP, in both cases leading to complete conversion of holo-ArCP to aryl-ArCP. NEC is a no enzyme control. Calculated and observed masses for ionized species are listed in Supplemental Table S3.

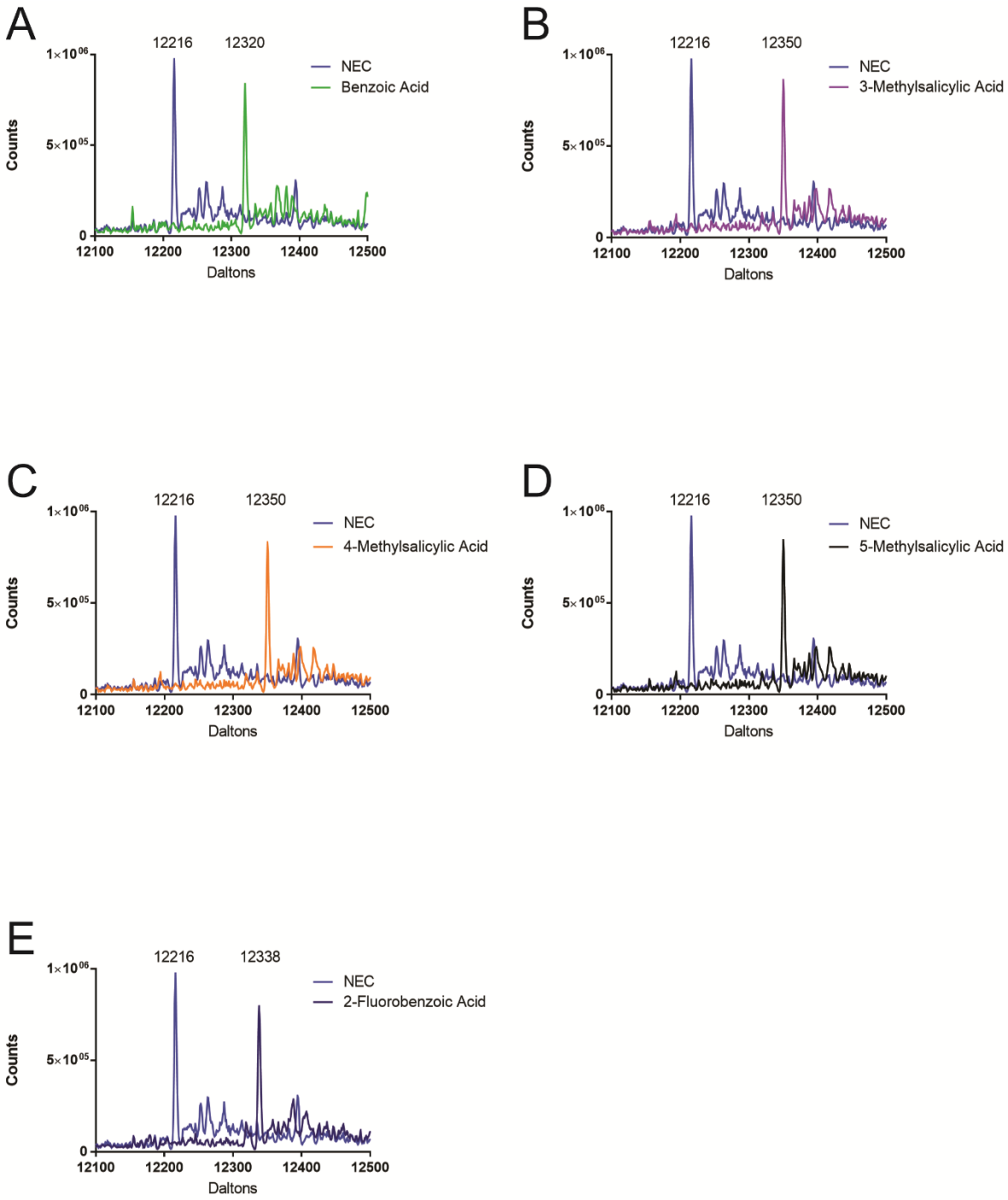

Figure S5: Intact protein MS of CahJ loading of unnatural substrates onto CahA ArCP. CahJ was incubated 1 hr with benzoic acid (A), 3-MSA (B), 4-MSA (C), 5-MSA (D), or 2-fluorobenzoic acid and ArCP, in all cases there was complete conversion of holo-ArCP to substrate loaded ArCP. NEC is a no enzyme control. Calculated and observed masses for ion species can be found in Table S3.

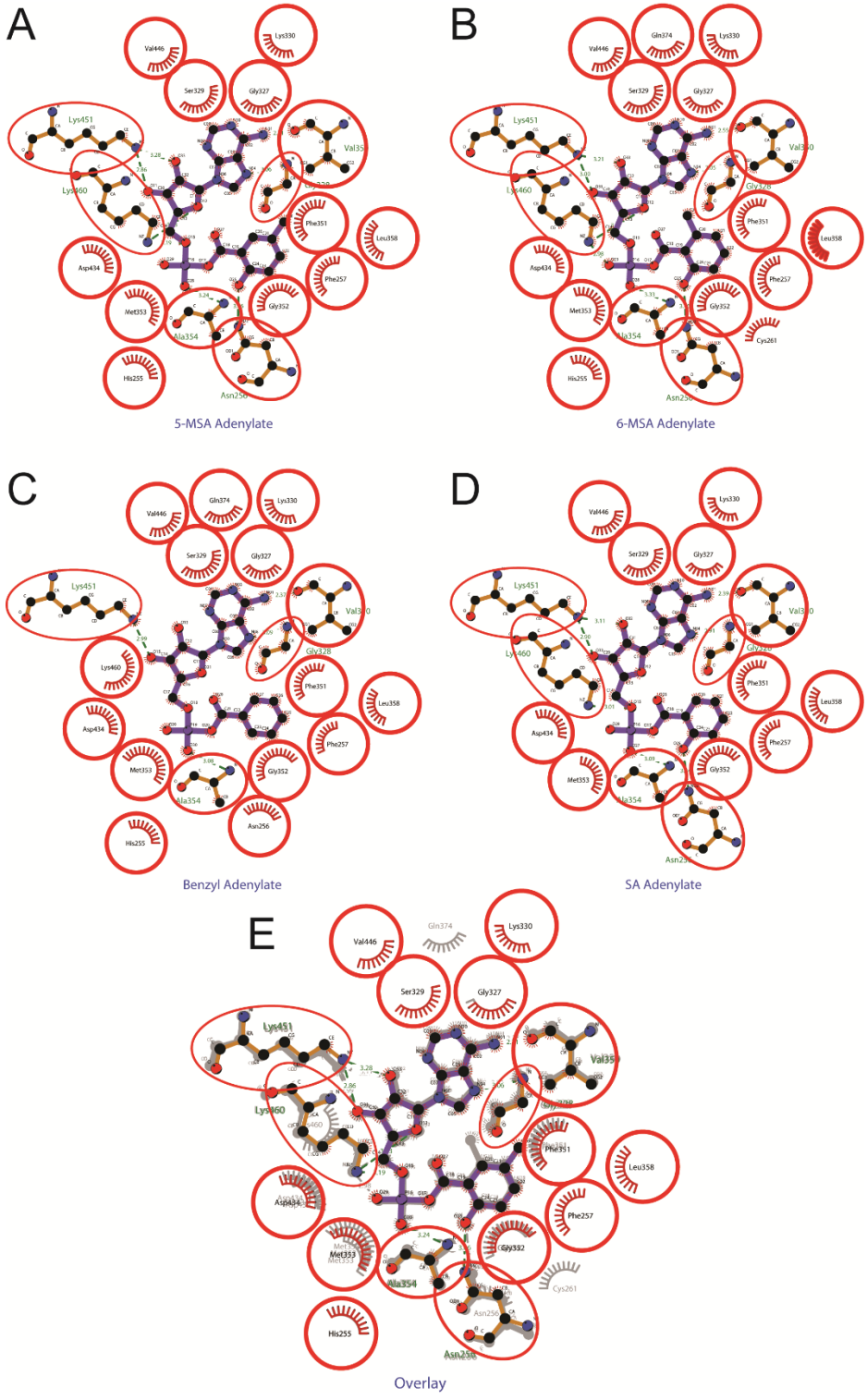

Figure S6: CahJ interaction with substrate adenylates, generated by LigPlot+ (Laskowski and Swindells, 2011). A) 5-MSA adenylate. B) 6-MSA adenylate. C) Benzyl adenylate. D) SA adenylate. E) Overlay of the four previous LigPlots. Red circles indicate contacts observed in multiple substrate complexes.

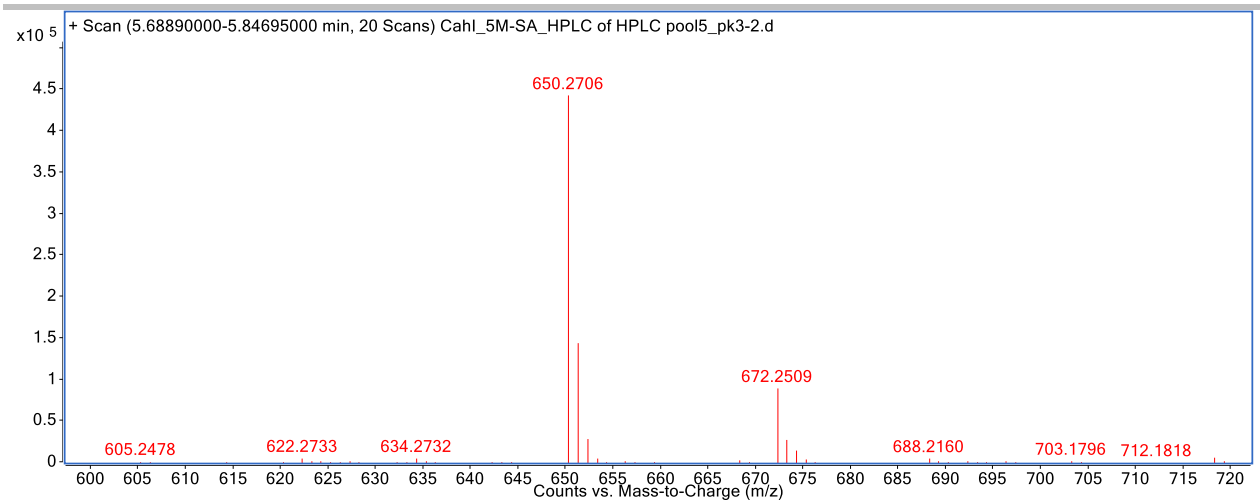

Figure S7: HRRESIMS $[M+H]^+$ chromatogram of cahuitamycin F (1).

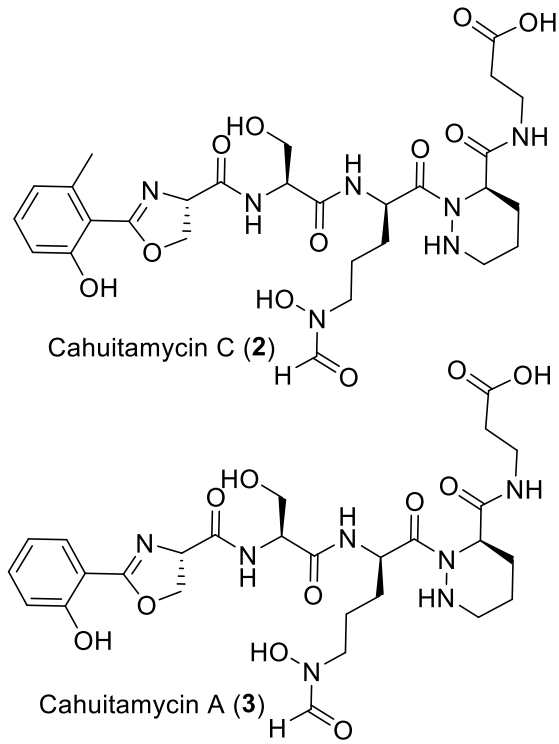

Figure S8: Structures of cahuitamycins C and A (**2**–**3**) with absolute stereochemistry.

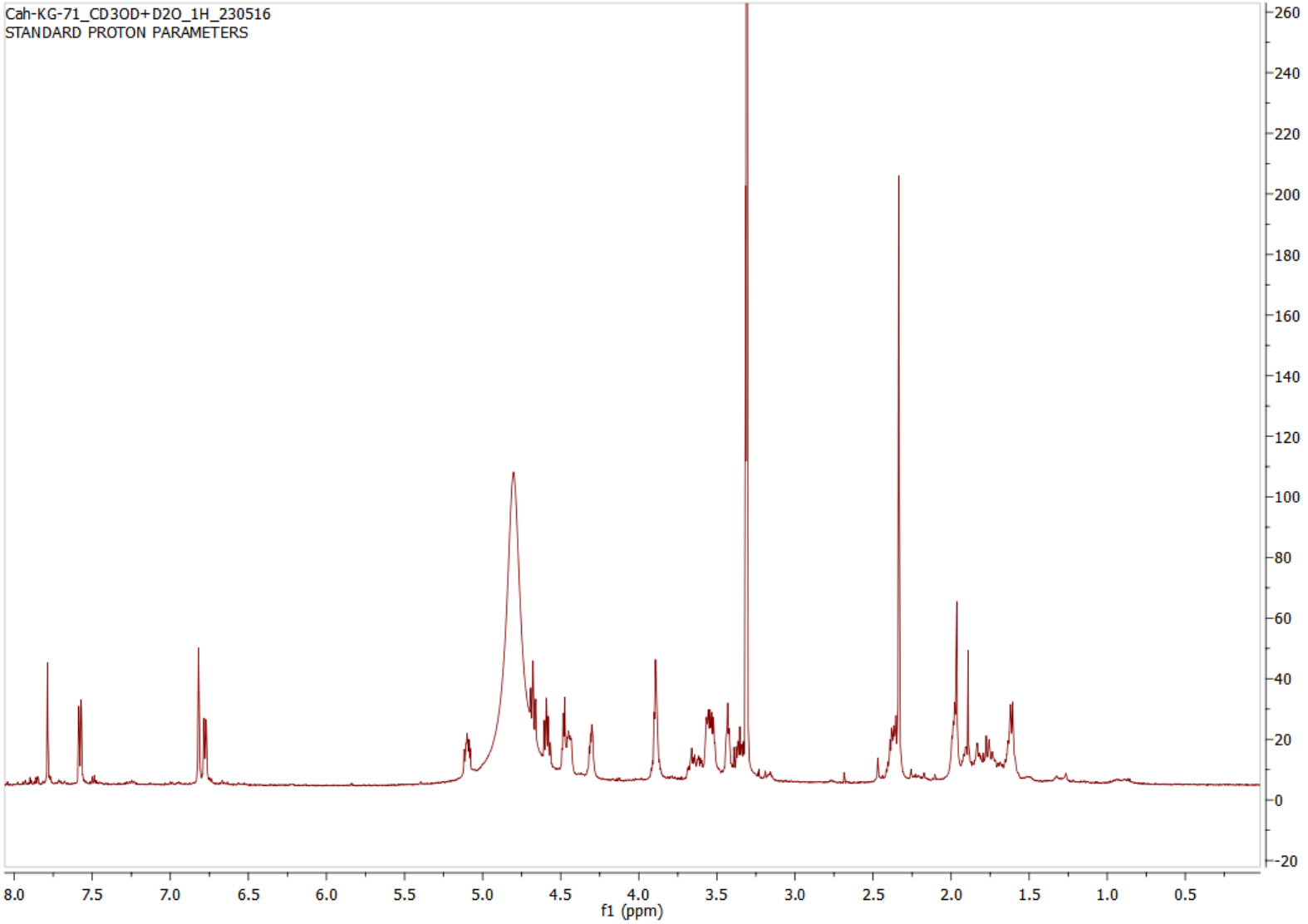

Figure S9: ^{1}H NMR (600 MHz, CD_3OD) spectrum of cahuitamycin F (**1**).

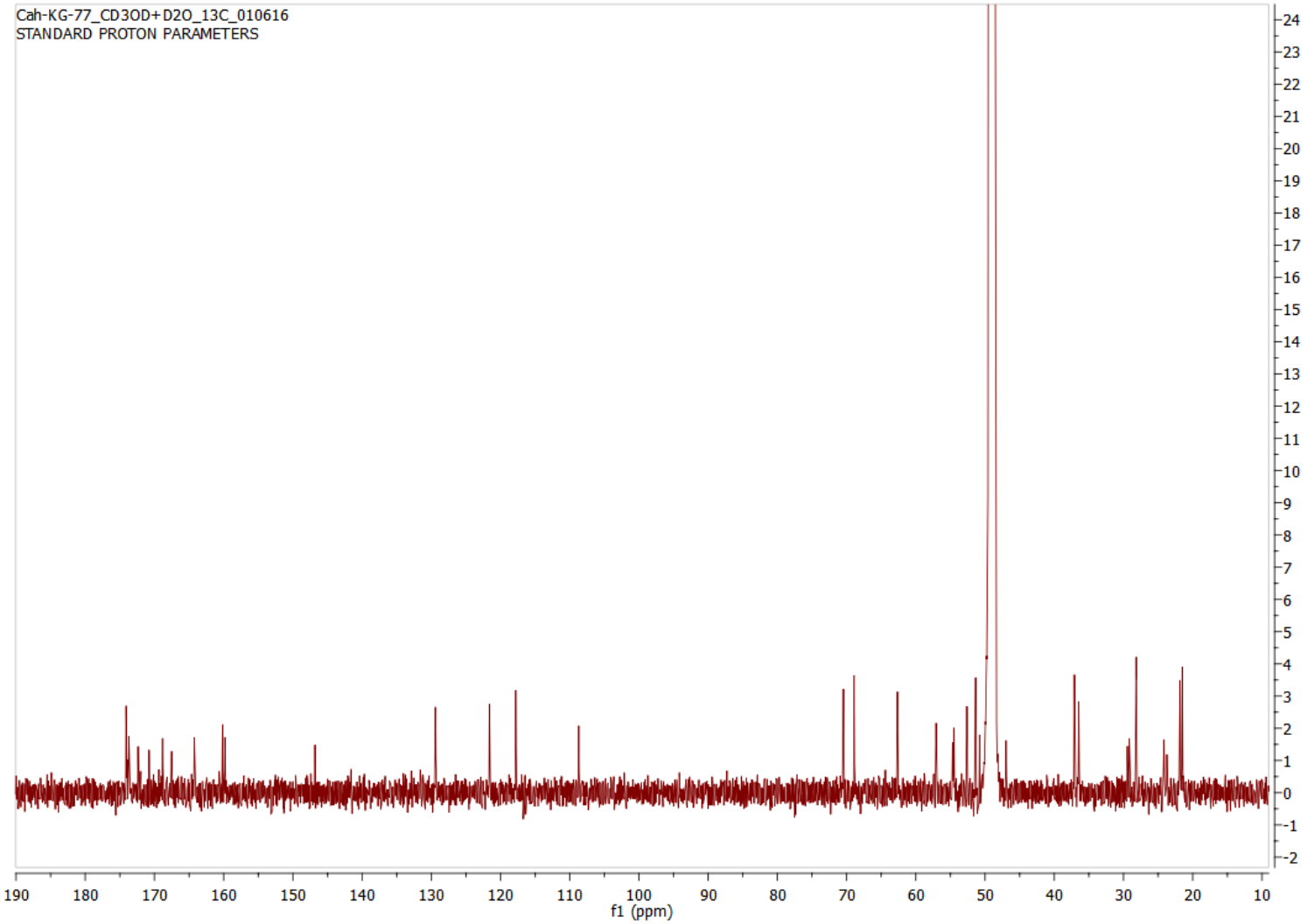

Figure S10: ^{13}C NMR (145 MHz, CD_3OD) spectrum of cahuitamycin F (**1**).

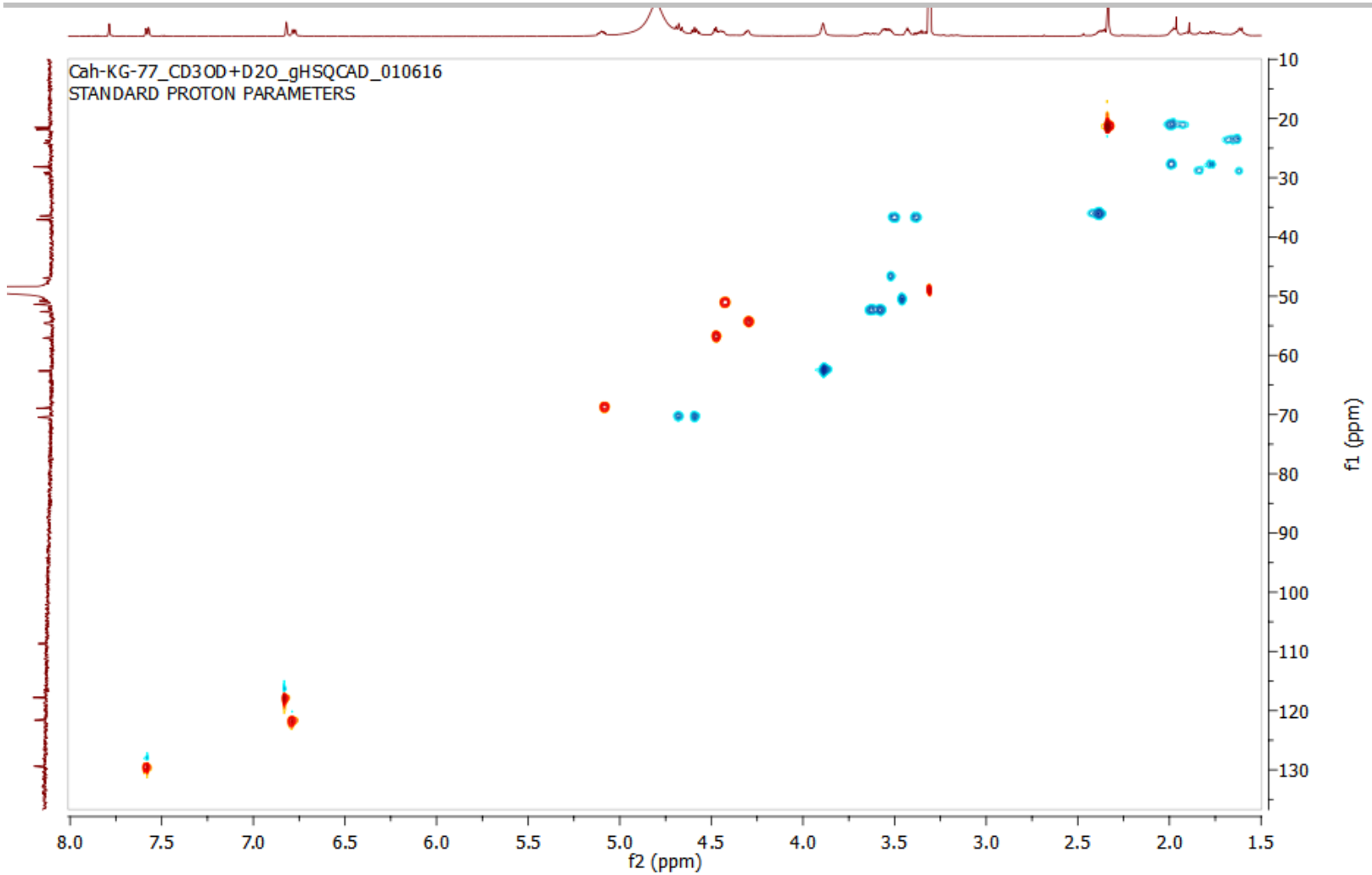

Figure S11: 2D HSQC NMR spectrum of cahuitamycin F (1).

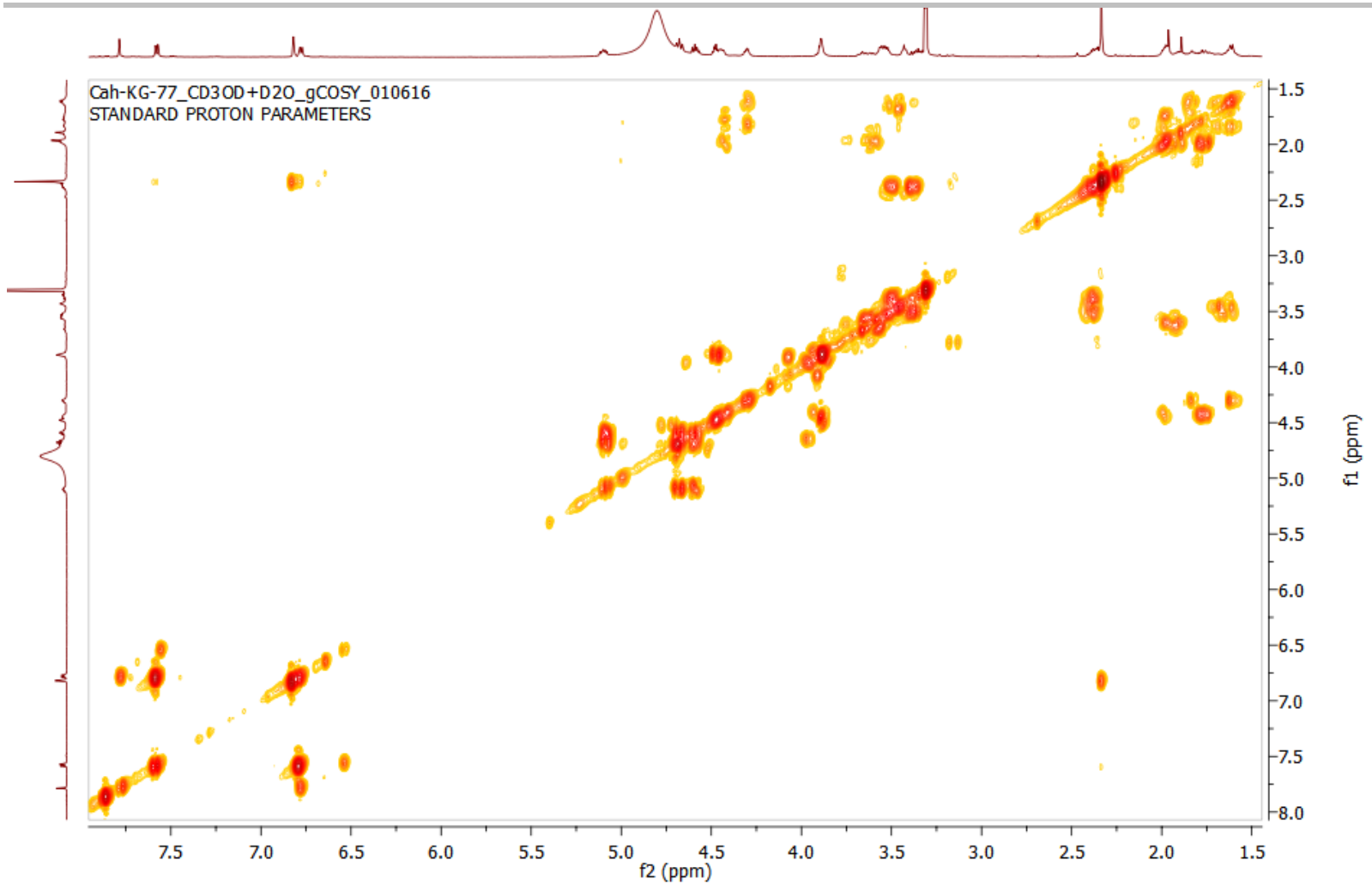

Figure S12: ^{1}H-^{1}H COSY NMR spectrum of cahuitamycin F (**1**).

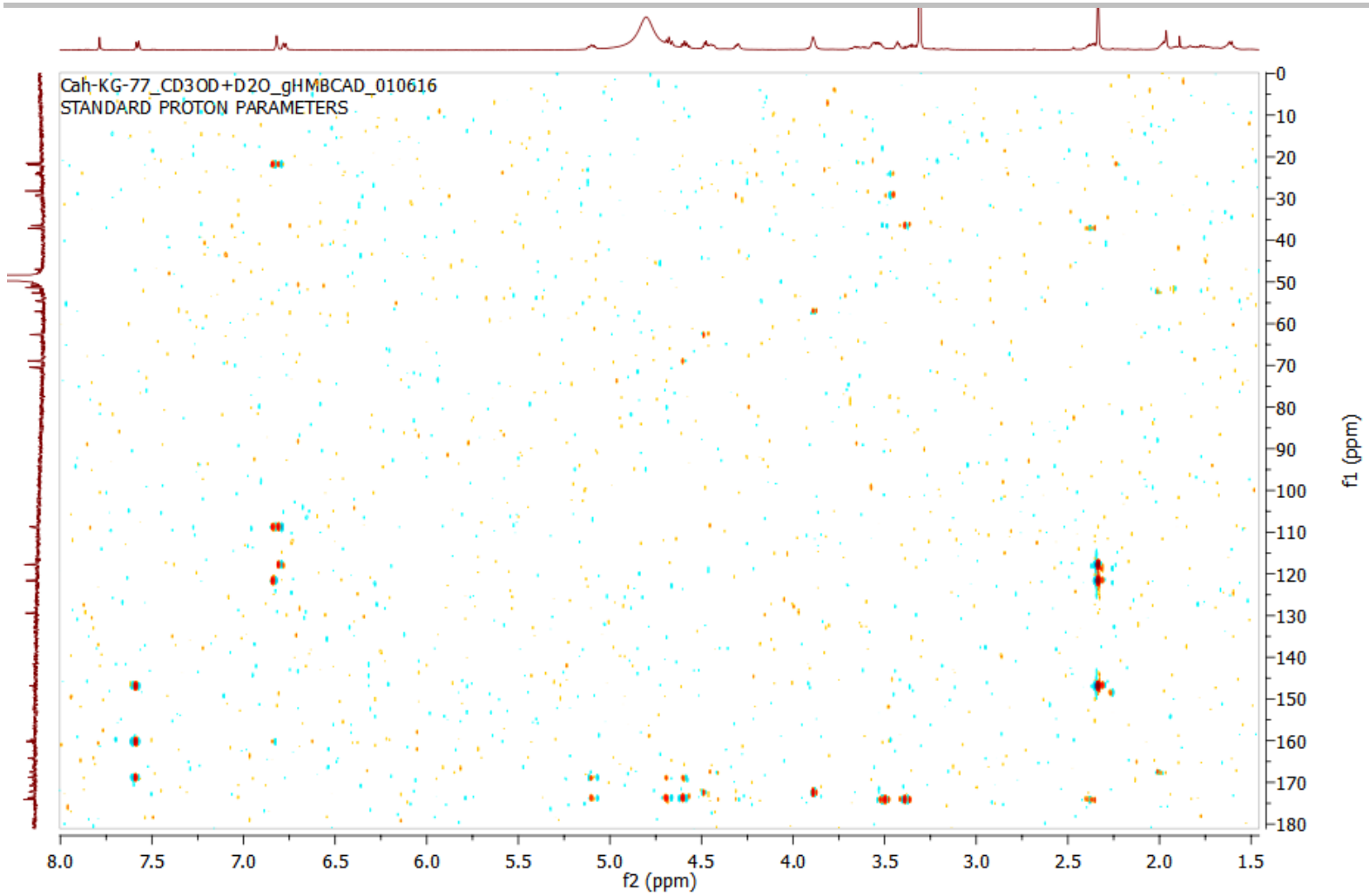

Figure S13: 2D HMBC NMR spectrum of cahuitamycin F (**1**).

Table S1: Protein production and crystal growth conditions plus data collection and refinement statistics for CahJ structures.

[a]Values in parentheses pertain to outermost shell of data

CahJ complex	AMP + SA	Salicyl adenylate	5-MSA adenylate	6-MSA adenylate	Benzoyl adenylate	AMP
CahJ production	Rich media, urea treatment	Rich media	Rich media, urea treatment	Rich media, urea treatment	Minimal media	Rich media, urea treatment
Added ligands	AMP, SA	none	ATP, 5-MSA	ATP, 6-MSA	none	AMP
Data Collection						
Space group	$P3_221$	$P3_221$	$P3_221$	$P3_221$	$P3_221$	$P3_221$
Cell dimensions a/b,c (Å)	122.1, 87.7	122.4, 88.1	122.0, 87.7	122.2, 88.2	121.7, 87.9	122.0, 88.0
X-ray source	APS 23ID-D	APS 23ID-D	APS 23ID-D	APS 23ID-D	APS 23ID-D	APS 23ID-D
Wavelength (Å)	1.033	1.033	1.033	1.033	1.033	1.033
d_{min} (Å)	1.55 (1.64-1.55)[a]	1.68 (1.78-1.68)	1.80 (1.91-1.80)	1.78 (1.89-1.78)	2.0 (2.12-2.0)	1.78 (1.88-1.78)
R_{merge}	0.079 (0.955)	0.120 (0.584)	0.097 (1.20)	0.090 (0.822)	0.140 (0.75)	0.098 (0.777)
Wilson B factor	24.8	29.6	33.6	31.4	28.6	27.2
Avg I/σ(I)	20.47 (2.43)	5.86 (1.54)	8.96 (1.19)	10.41 (1.50)	8.35 (1.71)	11.37 (1.76)
Completeness (%)	99.9 (99.6)	98.9 (98.4)	98.9 (99.2)	99.1 (99.3)	99.4 (99.4)	96.1 (95.1)
Multiplicity	10.02 (9.71)	3.39 (3.32)	3.40 (3.40)	3.93 (3.84)	3.39 (3.35)	4.06 (4.02)
Total observations	1,093,062 (168,861)	291,083 (45,428)	233,697 (37,664)	283,329 (44,413)	171,821 (27,161)	284,544 (44,625)
$CC_{1/2}$	0.999 (0.637)	0.985 (0.663)	0.996 (0.414)	0.997 (0.517)	0.986 (0.422)	0.996 (0.458)
CC*	0.999 (0.882)	0.996 (0.893)	0.999 (0.765)	0.999 (0.823)	0.996 (0.770)	0.999 (0.793)
Refinement						
Data range (Å)	35.60-1.55	35.76-1.68	18.51-1.80	35.76-1.78	35.64-2.0	35.68-1.78
Reflections used in refinement	109,137	85,969	69,218	72,133	50,752	70,015
R_{work}/R_{free} (%)	13.0/16.0	17.4/20.4	15.6/19.3	15.5/19.4	13.8/17.5	14.3/17.5
Twin fraction	0.29	0.04	0.12	0.17	0.49	0.47
Non-hydrogen atoms	4,701	5,001	4651	5,025	4,506	4481
protein	4,125	4,079	4056	4,083	4,070	4047
ligands	43	58	47	68	36	23
water	533	864	548	874	400	411
Amino acid residues	536	537	536	537	536	536
Deviation from ideality						
bond lengths (Å)	0.006	0.007	0.008	0.008	0.007	0.007
bond angles (°)	0.84	0.923	0.97	0.84	0.87	0.86
Average B-factor ($Å^2$)	20.3	27.5	32.1	23.8	23.1	20.3
macromolecules	18.4	24.7	30.5	21.2	22.4	19.6
ligands	27.0	23.5	29.1	21.3	14.7	17.2
solvent	35.3	40.9	44.2	36.1	30.5	27.3
Ramachandran plot						
favored (%)	96.9	97.8	98	97.6	97.0	97.2
allowed (%)	2.9	2.0	1.8	2.0	2.8	2.6
outliers (%)	0.2	0.2	0.2	0.4	0.2	0.2

Table S2: NMR spectroscopic data for cahuitamycins F (**1**).

	Cahuitamycin F (1)			
	δ_C	δ_H, multi (*J* in Hz)	COSY	HMBC
1	174.1			
2	36.5	2.37, m	3	1
		2.41, m		
3	36.9	3.37, m	2	1
		3.51, m	2	4
4	167.3			
5	50.8	4.41, dd (4.5, 9.5)	6	4, 9
6	28.1	1,78, m	5, 7	
		1.98, m	5, 7	
7	20.9	1.92, m	6, 8	5
		1.99, m	6, 8	
8	52.5	3.57, m	7	9
		3.62, m	7	6
9	173.2			
10	54.4	4.30, dt (6.0, 14.0)	11	9, 11, 15
11	29.1	1.62, m	12, 10	
		1.84, m	10	
12	23.5	1.63, m	11, 13	13, 10
		1.68, m	11, 13	10
13	51.1	3.43, t (6.4, 13.0)	12	11, 14
14	163.1	7.78, s		
15	172.5			
16	56.9	4.47, m	17	15, 18
17	62.6	3.89, d (5.0)	16	15
18	173.8			
19	68.9	5.10, dd (7.5, 10.0)	20	18
20	70.4	4.59, dd (7.5, 9.0)	19	18, 21
		4.68, dd (8.0, 9.0)	19	21
21	168.7			
22	146.5			
23	129.3	7.58, d (7.0)	24	21, 22, 28
24	121.6	6.77, d (7.0	23	26, 27
25	108.6			
26	21.3	2.31, s		24, 25, 27
27	117.7	6.82, s	28	24, 25, 26
28	160.1			

Table S3: Calculated and observed ion masses from intact protein MS assay.

Species	Calculated mass (Da)	Observed Mass (Da)
holo-ArCP	12216	12216
Salicylyl ArCP	12336	12336
6-MSA ArCP	12350	12350
Benzyl ArCP	12320	12320
3-MSA ArCP	12350	12350
4-MSA ArCP	12350	12350
5-MSA ArCP	12350	12350
2-Fluorobenzyl ArCP	12338	12338

References

[1] D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, H. O. Smith, *Nat Meth* **2009**, *6*, 343-345.
[2] L. Stols, M. Gu, L. Dieckman, R. Raffen, F. R. Collart, M. I. Donnelly, *Protein expression and purification* **2002**, *25*, 8-15.
[3] B. A. Pfeifer, S. J. Admiraal, H. Gramajo, D. E. Cane, C. Khosla, *Science* **2001**, *291*, 1790-1792.
[4] J. R. Whicher, S. S. Smaga, D. A. Hansen, W. C. Brown, W. H. Gerwick, D. H. Sherman, J. L. Smith, *Chemistry & biology* **2013**, *20*, 1340-1351.
[5] L. E. Quadri, P. H. Weinreb, M. Lei, M. M. Nakano, P. Zuber, C. T. Walsh, *Biochemistry* **1998**, *37*, 1585-1595.
[6] A. S. Worthington, M. D. Burkart, *Organic & biomolecular chemistry* **2006**, *4*, 44-46.
[7] W. Kabsch, *Acta crystallographica. Section D, Biological crystallography* **2010**, *66*, 125-132.
[8] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *Journal of applied crystallography* **2007**, *40*, 658-674.
[9] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta crystallographica. Section D, Biological crystallography* **2010**, *66*, 213-221.
[10] J. J. May, N. Kessler, M. A. Marahiel, M. T. Stubbs, *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 12120-12125.
[11] T. C. Terwilliger, R. W. Grosse-Kunstleve, P. V. Afonine, N. W. Moriarty, P. H. Zwart, L. W. Hung, R. J. Read, P. D. Adams, *Acta crystallographica. Section D, Biological crystallography* **2008**, *64*, 61-69.
[12] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 486-501.
[13] P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, *Acta Crystallogr. D Biol Crystallogr.* **2012**, *68*, 352-367.
[14] N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams, *Acta crystallographica. Section D, Biological crystallography* **2009**, *65*, 1074-1080.
[15] Schrodinger, LLC, **2015**.
[16] T. J. McQuade, A. D. Shallop, A. Sheoran, J. E. Delproposto, O. V. Tsodikov, S. Garneau-Tsodikova, *Anal. Biochem.* **2009**, *386*, 244-250.
[17] T. Böttcher, I. Kolodkin-Gal, R. Kolter, R. Losick, J. Clardy, *J. Am. Chem. Soc.* **2013**, *135*, 2927-2930.

Contributions

* Contributed equally to this manuscript.