SUPPLEMENTAL INFORMATION

A Gcn5-related N-acetyltransferase (GNAT) capable of acetylating polymyxin B and colistin antibiotics *in vitro*

Mateusz P. Czub[#1,2], Brian Zhang[#3], M. Paul Chiarelli[4], Karolina A. Majorek[1,2], Layton Joe[3], Przemyslaw J. Porebski[1,2], Alina Revilla[3], Weiming Wu[3], Daniel P. Becker[4], Wladek Minor[1,2], and Misty L. Kuhn[*3]

[1]Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, VA 22908, USA

[2]Center for Structural Genomics of Infectious Diseases (CSGID), University of Virginia, 1340 Jefferson Park Avenue, Charlottesville, VA 22908, USA

[3]Department of Chemistry and Biochemistry, San Francisco State University, San Francisco, CA 94132, USA

[4]Department of Chemistry and Biochemistry, Loyola University Chicago, Chicago, IL 60660, USA

SUPPLEMENTAL METHODS

Materials

Acetyl coenzyme A (AcCoA; #A2181) and coenzyme A (CoA; #C3019) trilithium salts, and polymyxin B (#P4932) and colistin (#C4461) sulfate salts were purchased from Sigma-Aldrich. All other reagents for biochemical assays were purchased at the highest quality available.

Protein expression and purification for crystallization

The ampicillin resistant plasmid containing the *PA3944* gene was transformed into chloramphenicol resistant *E. coli* BL21(DE3) CodonPlus-RIL cells (Agilent). A starter culture of LB with 100 µg/ml ampicillin and 34 µg/ml chloramphenicol was inoculated and grown overnight at 37°C at 200 rpm. The following morning 10 ml of starter culture was used to inoculate 1 L of LB. Cells were grown until the OD_{600nm} reached 0.8 and then cooled to 16°C. Protein expression was induced with 1 mM IPTG and expressed overnight at 16°C with shaking. The next day, the cells were harvested by centrifugation (25 min at 4000 x g at 4°C in a Beckman Coulter Avanti J-26 XP centrifuge) and the supernatant was discarded. The pellet was resuspended in lysis buffer (50 mM Tris-HCl pH 7.8, 2 mM imidazole, 150 mM sodium chloride) and homogenized in a glass homogenizer. One tablet of protease inhibitor (cOmplete Mini, EDTA-free Protease Inhibitor Cocktail Tablets, Roche) and a pinch of lysozyme (Sigma product # L-6876) and 2 µl of Benzonase Nuclease (Sigma product # E1014) were then added to the cell suspension, sonicated on ice, and centrifuged (45 min at 142,400 x g, and 4°C in a Beckman Coulter Optima L-80 XP ultracentrifuge). The supernatant was loaded onto a gravity flow Ni-NTA column (Qiagen) previously equilibrated with lysis buffer and the flow-through was collected and reapplied on the column three times to extend the length of interaction time between the protein and resin. The resin was then washed with wash buffer (50 mM Tris-HCl pH 7.8, 10 mM imidazole, 600 mM sodium chloride) and protein was eluted with elution buffer (50 mM Tris-HCl pH 7.8, 250 mM imidazole, 200 mM sodium chloride). To remove the polyhistidine tag, the eluted protein was combined with tobacco etch virus (TEV) protease in a 15:1 ratio and dialyzed in a 10 kD MWCO Slide-A-Lyzer dialysis cassette (Thermo Fisher Scientific) against 1 L of dialysis buffer (50 mM Tris-HCl pH 7.8, 150 mM sodium chloride, 14 mM BME) overnight at 4°C. The dialyzed protein was loaded onto the nickel-affinity column that had been equilibrated with dialysis buffer. The flow-through was collected, concentrated with an Amicon Ultra Centrifugal Filter with a 10K MWCO, and loaded onto a HiLoad 16/60 Superdex 200 gel filtration column using an AKTA FPLC (GE Healthcare) in 100 mM Tris-HCl pH 7.5 and 150 mM NaCl buffer. Polyhistidine tag cleavage was confirmed by SDS-PAGE.

Protein crystallization

The PA3944 protein was crystallized using hanging drop vapor diffusion at 16°C in 3-Well Midi Crystallization plates (Swissci) that were set using a Mosquito crystallization robot (TTP Labtech). Crystals used for data collection and structure determination were from two different preparations of protein but in the same crystallization condition (MCSG1 screen, well C11: 100 mM Tris-HCl pH 7.0, 200 mM calcium acetate monohydrate, 20%w/v PEG 3000): 1) cleaved PA3944 protein at 14 mg/ml co-crystallized in the presence of 5 mM puromycin and CoA with a ratio of protein

to mother liquor of 1:1; ethylene glycol was used as a cryoprotectant. 2) cleaved PA3944 protein at 10 mg/ml co-crystallized in the presence of 5 mM colistin and CoA with a ratio of protein to mother liquor of 3:2; no cryoprotectant was added when harvesting crystals for data collection. Crystals were harvested using Hampton CryoLoops and dried for 10-15 minutes over 1 M sodium chloride (a slow dehydration technique) and then frozen in liquid nitrogen.

Size-exclusion chromatography

The oligomeric state of the PA3944 protein was determined using size exclusion chromatography with a HiLoad 16/60 Superdex 75 gel filtration column (GE Healthcare). A calibration curve was prepared from retention times and molecular weights of albumin (66.5 kDa), ovalbumin (44.3 kDa), chymotrypsin (25 kDa) or ribonuclease A (13.7 kDa) in 100 mM Tris-HCl pH 7.0 buffer. To determine whether pH affected the oligomeric state of the PA3944 protein, separate measurements were determined in the following buffers that contained 150 mM NaCl: 100 mM potassium phosphate pH 6.0; 100 mM Tris-HCl at pH 7.0 or 8.0; and 100 mM CHES at pH 9.0. We also assessed whether the presence of AcCoA affected the oligomeric state of PA3944 by analyzing the retention times of the PA3944 protein in the presence of a 2-fold excess of AcCoA: protein. 15 nmol of protein was injected onto the column at a flow rate of 1.5 ml/min for each measurement.

Sequence and structural comparison of homologs

To identify PA3944 homologs, we queried the SwissProt database using the PA3944 sequence and single iteration BLAST (E-value cutoff of 10). From these results, we selected sequences of proteins with 3D structures and experimentally verified functions or preliminary substrates. These included two polyamine acetyltransferases (Uniprot IDs ATDA_VIBCH and ATDA_ECOLI) and a fusarinine C siderophore acetyltransferase (Uniprot ID: SIDG_ASPFU)[1-3]. Two sequences of protein acetyltransferase RimI from *E. coli* and *S.typhimurium* (Uniprot IDs Q8ZJW4)SALTY, RIMI_ECO57, respectively) and three sequences of GNAT acetyltransferases from *P. aeruginosa* (Uniprot IDs: Q9HUU7, and Q9HV14) that have been found to acetylate L-methionine sulfoximine and L-methionine sulfone[4], and a C-terminal lysine residue of a peptide[5], respectively, as well as UniProt ID: Q9I0Q8, which was included in our broad substrate screening[6] but does not have a known function. The sequence alignment (**Figure 3**) was performed using the program PROMALS3D[7,8]. A structural alignment of all GNATs from *P. aeruginosa* with crystal structures was performed using secondary-structure-matching (SSM) with the program SUPERPOSE from the CCP4 suite[9]. The acceptor substrate binding site in all structures was visually identified.

SUPPLEMENTAL RESULTS

Mass spectrometry of acetylated colistin and polymyxin B

Molecular weight and product ion spectra were acquired from each reaction mixture. Product ion analyses of colistin and polymyxin indicate that the primary site of enzymatic acetylation occurs (just once) on the primary amine on the side chain closest to the cyclic peptide at the end of the chain. Molecular weight analysis of reaction mixtures containing AcCoA and

polymyxin in presence and absence of enzyme are shown in Figure S2. These spectra strongly suggest polymyxin acetylation only occurs in the presence of the PA3944 enzyme. In Figure S2A, $(M+H)^+$ and $(M+Na)^+$ ions formed from polymyxin B1 are observed at m/z 1204 and m/z 1226 and those for polymyxin B2 are observed at m/z 1190 and m/z 1211, respectively. No acetylated polymyxin ions are observed in the spectrum of the control reaction in absence of enzyme. Figure S2B shows monoacetylated polymyxin B1 at m/z 1246 $(M+H)^+$, m/z 1268 $(M+Na)^+$, and m/z 1290 $(M-H+2Na)^+$, respectively. No multiply acetylated polymyxin molecules are observed.

Product ion spectra of the $(M+H)^+$ ion at m/z 1246 formed from the monoacetylated polymyxin is shown in Figure S3. An ion observed at m/z 904.8 is formed by cleavage on the side chain resulting in the loss of a neutral fragment containing the (unmodified) primary amine that is farthest away from the cyclic peptide ring. The m/z 904.8 fragment ion contains the acetyl group. The product ion observed at m/z 762.4 is the cyclic peptide fragment formed by loss of the side chain analogous to the m/z 728.3 in the product ion spectrum of the singly-acetylated polymyxin B in Figure S3 (see also Scheme S1). If polymyxin B was acetylated on the primary amine closest to the end of the fatty acyl chain (Dab1) we would expect to see an ion at m/z 863. This ion is not observed, so we conclude that the primary amine closest to the peptide ring in polymyxin is the main site of acetylation.

Molecular weight analysis of reaction mixtures containing AcCoA and colistin in the presence and absence of enzyme are shown in Figure S4. Spectra of the control reaction show $(M+H)^+$ and $(M+Na)^+$ of unacetylated colistin E2 (m/z 1156 and m/z 1178) and colistin E1 molecules (m/z 1170 and m/z 1192), respectively (Figure S4A). Ions formed from monoacetylated colistin E2 are observed in Figure S4B at m/z 1198 $(M+H)^+$ and m/z 1220 $(M+Na)^+$. A smaller abundance of the $(M+Na)^+$ formed from monoacetylated colistin E1 is observed at m/z 1234. These spectra strongly suggest the acetylation of colistin only occurs in the presence of the PA3944 enzyme. No m/z values consistent with the $(M+H)^+$ and $(M+Na)^+$ ions formed from multiply acetylated colistin are observed. We do not know the identities of the additional ions observed at m/z 1313, 1327, and 1449 but each appear to become acetylated in the presence of the enzyme due to the observation of ions at m/z 1356, 1370, and 1491. Since colistin sulfate is supplied as a mixture of molecules, we cannot conclusively state the identity of these ions; however, it is clear that the apparently acetylated molecules are only observed in the presence of the PA3944 enzyme.

The protonated molecule ion at m/z 1198 was subjected to product ion analysis (Figure S5) to locate the position of acetyl modification on the colistin molecule (Scheme S2). The product ion spectra suggest that the primary position of acetylation is on the primary amine on the side chain closest to the peptide ring. The observation of an acylium ion formed by the cleavage of colistin produces an ion at m/z 997.4 (containing the acetyl group) that is formed by the loss of a fragment containing the unmodified primary amine farthest away from the cyclic peptide ring. The formation of a product ion at m/z 728.3, corresponding to the unacetylated cyclic peptide fragment, strongly suggests that acetylation occurs on the primary amine closest to the cyclic peptide (Scheme S2). Modification of the amine closest to the end of the peptide chain would have yielded an acylium ion at m/z 955 formed by the loss of a fragment containing the acetyl group and this ion is not observed in any significant abundance. Therefore, we conclude the primary amine closest to the peptide ring is the principal site of modification.

SUPPLEMENTAL TABLES, FIGURES, and SCHEMES

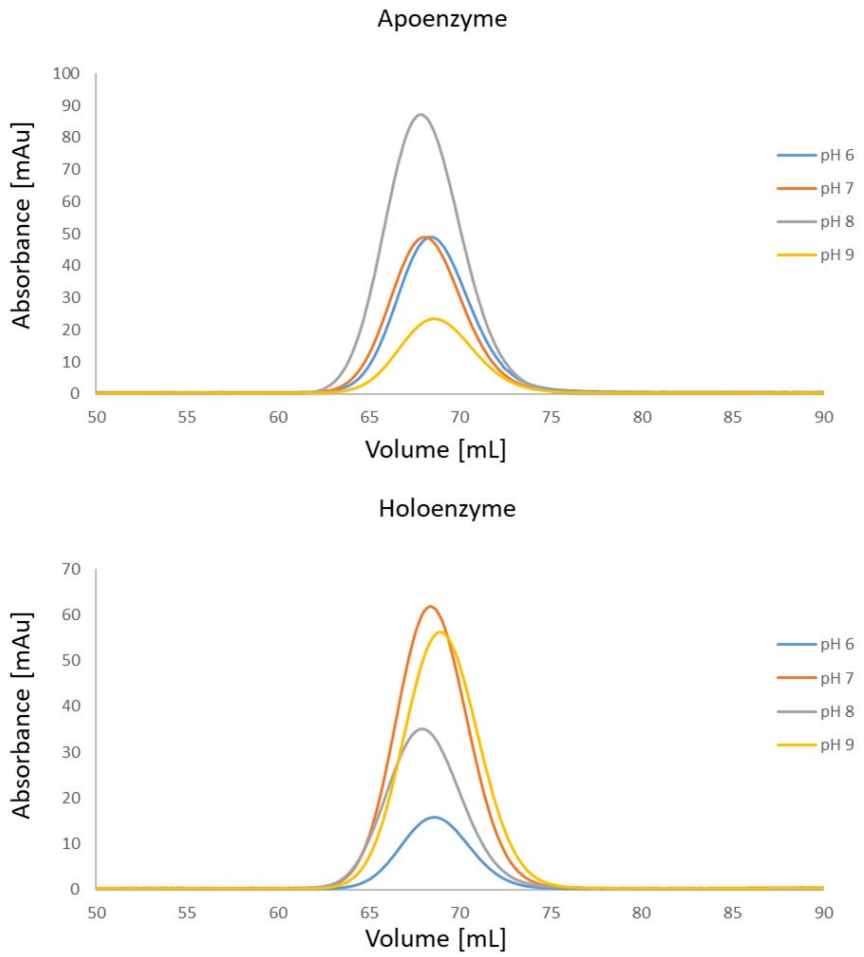
Table S1. Summary of *Pseudomonas aeruginosa* GNAT genes, proteins, and structures.

Pseudomonas gene	UniProt ID	PDB IDs of crystal structures
PA0115	Q9I717	1xeb
PA0249	Q9I6P0	
PA0478	Q9I640	2fe7
PA0483	Q9I635	
PA0711	Q9I5L7	
PA1062	Q9I4R2	
PA1377	Q9I3W7	2vi7
PA1428	Q9I3R6	
PA1472	Q9I3P0	
PA1749	Q9I2Y8	
PA1885	Q9I2L2	
PA1928	Q9I2H6	
PA1943	Q9I2G1	
PA2271	Q9I1K2	
PA2578	Q9I0Q8	3owc
PA2631	Q9I0K7	
PA3127	Q9HZ97	
PA3248	Q9HYZ2	
PA3270	Q9HYX1	1yre
PA3368	Q9HYN3	
PA3460	Q9HYE6	
PA3944	Q9HX72	
PA3945	Q9HX71	
PA4026	Q9HX01	2eui
PA4114	Q9HWR6	
PA4166	Q9HWL5	
PA4166	Q9HWL5	
PA4534	Q9HVP3	4ubr, 5ib0
PA4678	Q9HVB7	
PA4794	Q9HV14	3kkw, 4kot, 4oae, 4l8a, 4oad, 4kub, 4klw, 4kow, 4klv, 4kua, 4kov, 4l89, 4kos, 4kor, 4kou, 4koy, 3pgp, 4m3s, 4kox
PA4863	Q9HUV0	
PA4866	Q9HUU7	1yvo, 2j8m, 2j8n, 2j8r, 2bl1
PA5204	P22567	
PA5432	Q9HTD4	
PA5433	Q9HTD3	
PA5475	Q9HT95	

* Based on predictions from InterPro, each protein listed is either annotated as a GNAT or contains a GNAT domain.

Figure S1. Size exclusion chromatograms of PA3944 protein with different buffers in the absence (apoenzyme) and presence of AcCoA (haloenzyme) (A), calculated molecular weights (B), and calibration curve (C). See Methods for more specific details.

A



B

Apoenzyme				
pH	6	7	8	9
Volume [ml]	68.5	68.1	67.9	68.5
Calculated mass [kDa]	24.2	24.7	25.0	24.2
Holoenzyme				
pH	6	7	8	9
Volume [ml]	68.6	68.4	67.9	69.0
Calculated mass [kDa]	24.0	24.3	25.0	23.5

C

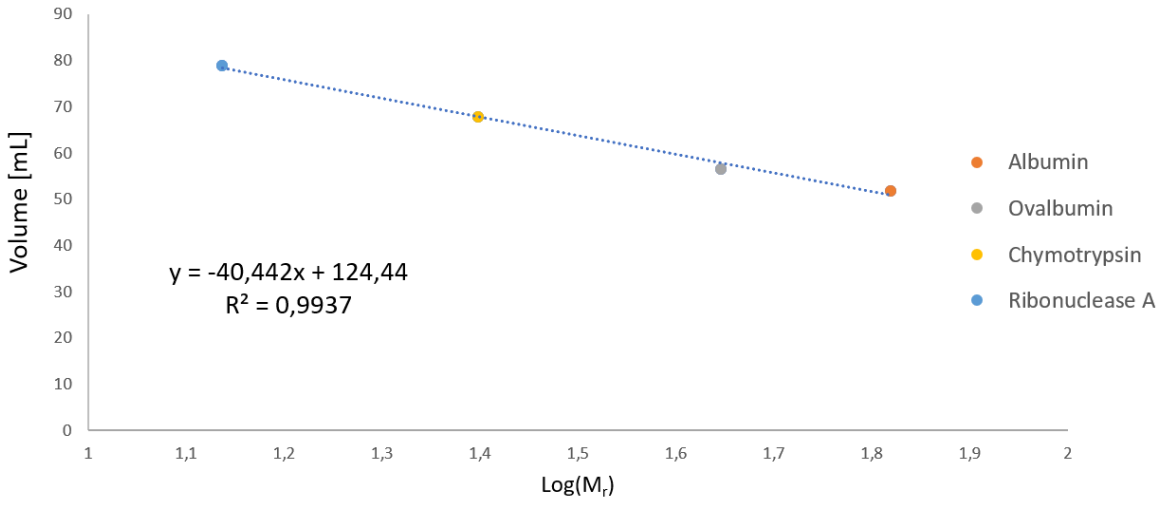

90
80
70
60
50
40
30
20
10
0
Volume [mL]
y = -40,442x + 124,44
R² = 0,9937
1
1,1
1,2
1,3
1,4
1,5
1,6
1,7
1,8
1,9
2
Log(Mr)
Albumin
Ovalbumin
Chymotrypsin
Ribonuclease A

Figure S2. Full scan ESI mass spectrum of a reaction mixture containing polymyxin B and AcCoA in A) absence of PA3944 enzyme and B) presence of PA3944 enzyme. The electrospray mass spectra of the polymyxin peptide mixture (Figure S2A) show protonated $(M+H)^+$ and cationized $(M+Na)^+$ molecule ions generated from form B1 at m/z 1203.9 and at m/z 1225.8, respectively. Smaller abundances of the protonated molecule ion formed from the B2 form are observed at m/z 1189.7. No acetylated products are observed without the enzyme present in the reaction mixture. When the enzyme is present in the reaction mixture (Figure S2B) only ions generated from the acetylated polymyxin are observed. Protonated molecule ions from the B2 and B1 forms modified by single acetyl groups are observed at m/z 1231.3 and m/z 1245.3. Ions containing one $(M+Na)^+$ and two sodium ions $(M+2Na)^+$ are formed from the B1 containing one acetyl group are observed at m/z 1267.8 and m/z 1289.7. No unmodified polymyxin or polymyxin molecules modified by multiple acetyl groups are observed in the Figure 2B spectra. See Methods for more specific details about sample preparation (positive ion mode).

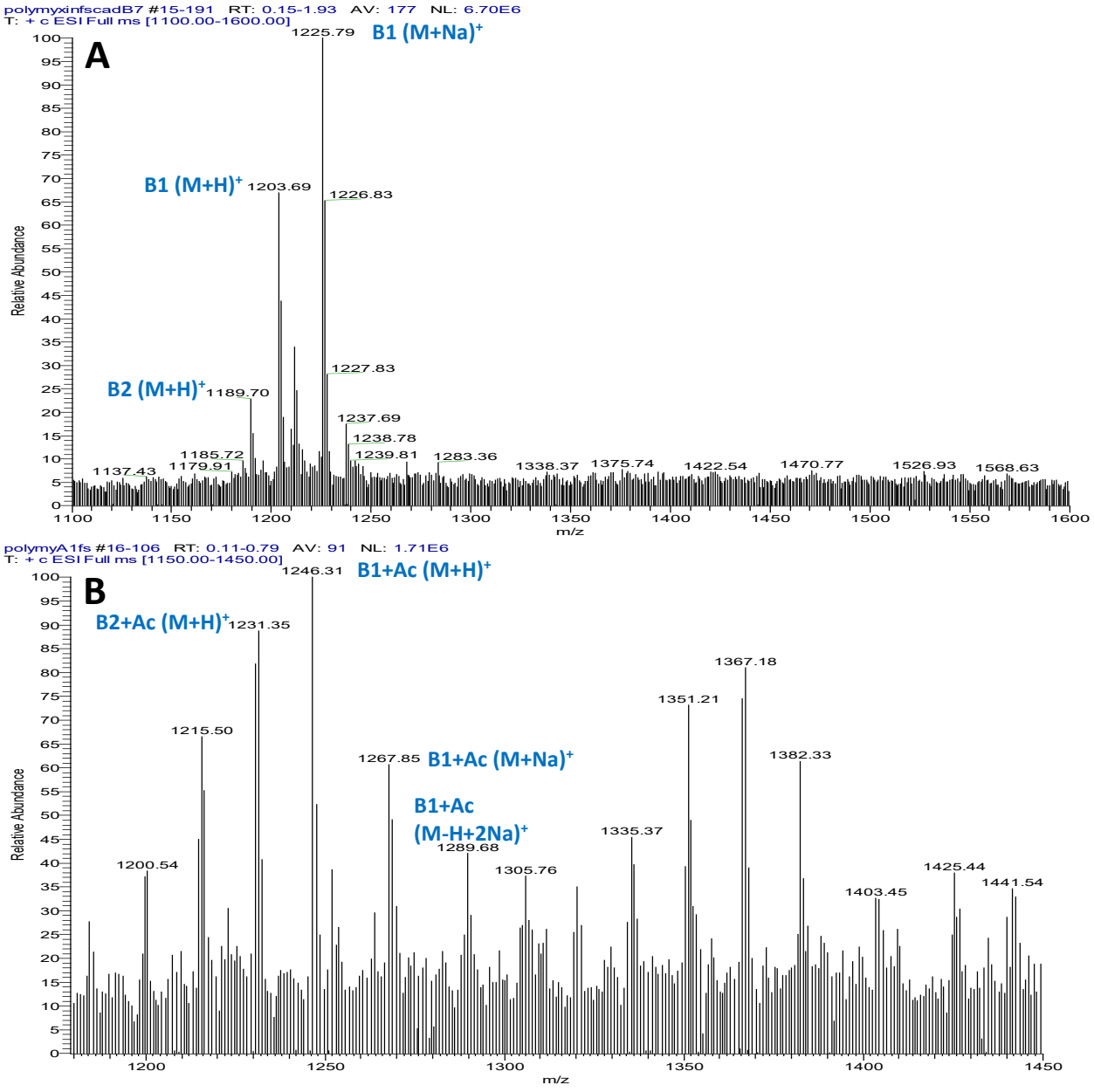


Figure S3. Product ion mass spectra of the protonated molecule ion of **A)** unmodified polymyxin B that was observed at m/z 1204 and **B)** polymyxin containing one acetyl group that was observed at m/z 1246 (positive ion mode), showing loss of water as the primary fragmentation in both cases.

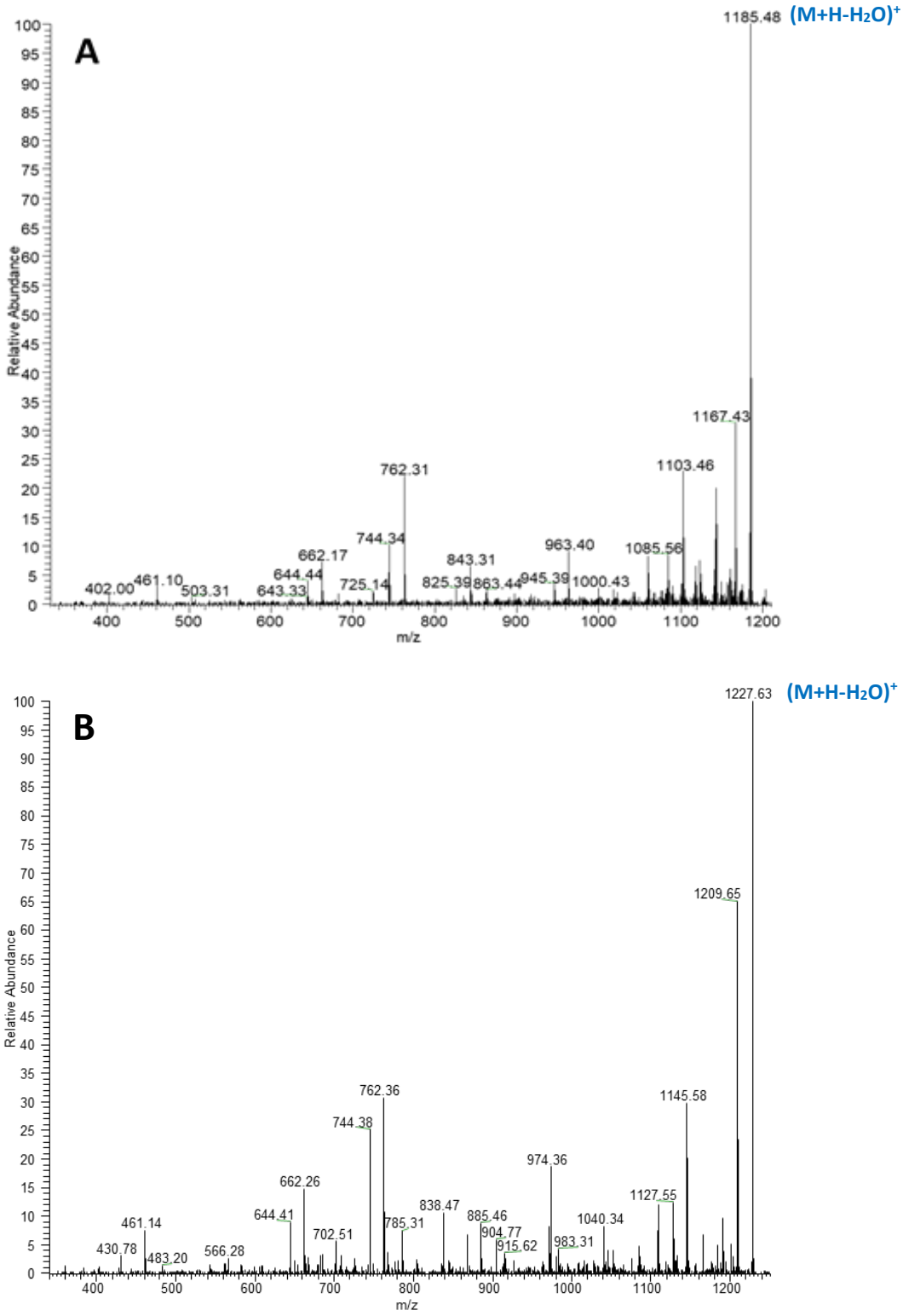


Figure S4. Full scan ESI mass spectra of a reaction mixture containing colistin A and B and AcCoA A) in the absence of PA3944 enzyme and B) in the presence of PA3944 enzyme. See Methods for more specific details about sample preparation (positive ion mode).

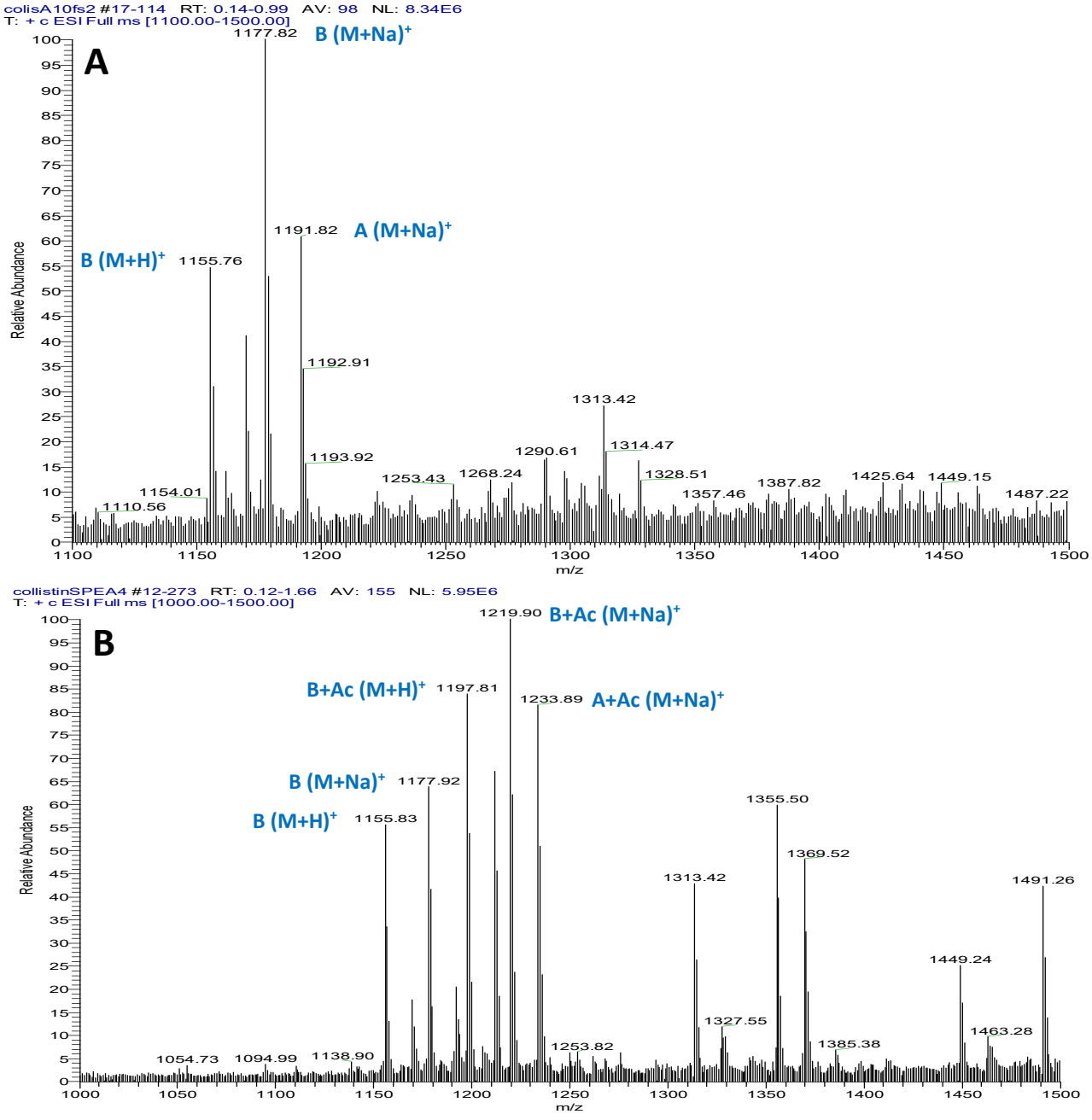


Figure S5. Product ion mass spectra of the protonated molecule ion of **A)** unmodified colistin that had been observed at m/z 1156 and **B)** of the colistin ion containing one acetyl group (positive ion mode) that had been observed at m/z 1198. Loss of water was observed as the primary fragmentation in both cases.

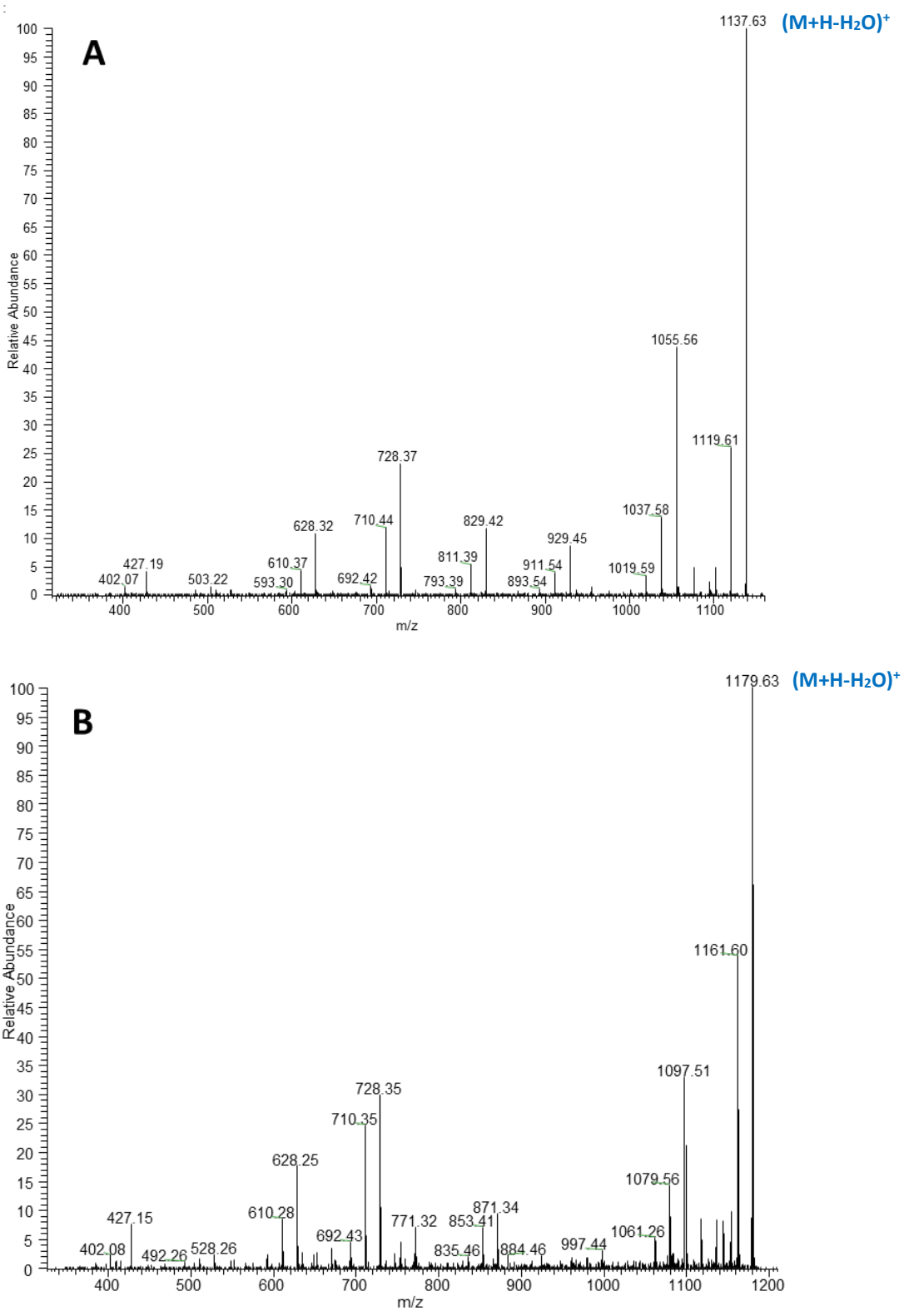
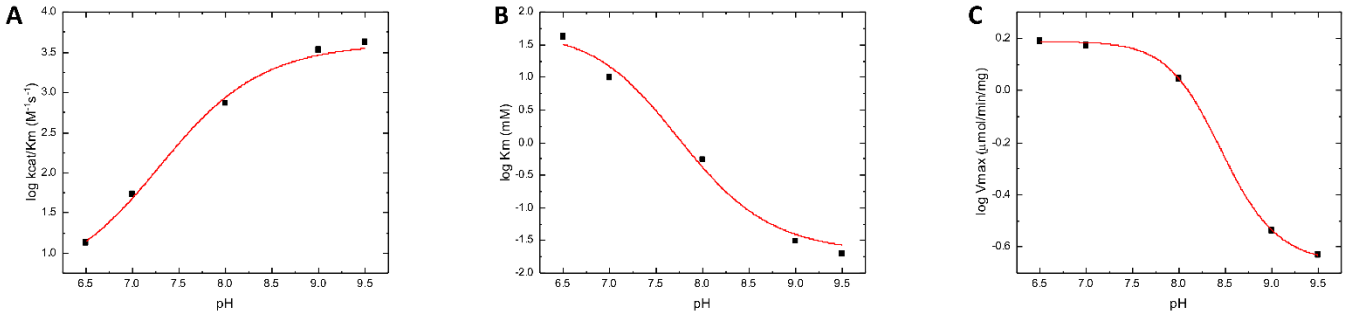
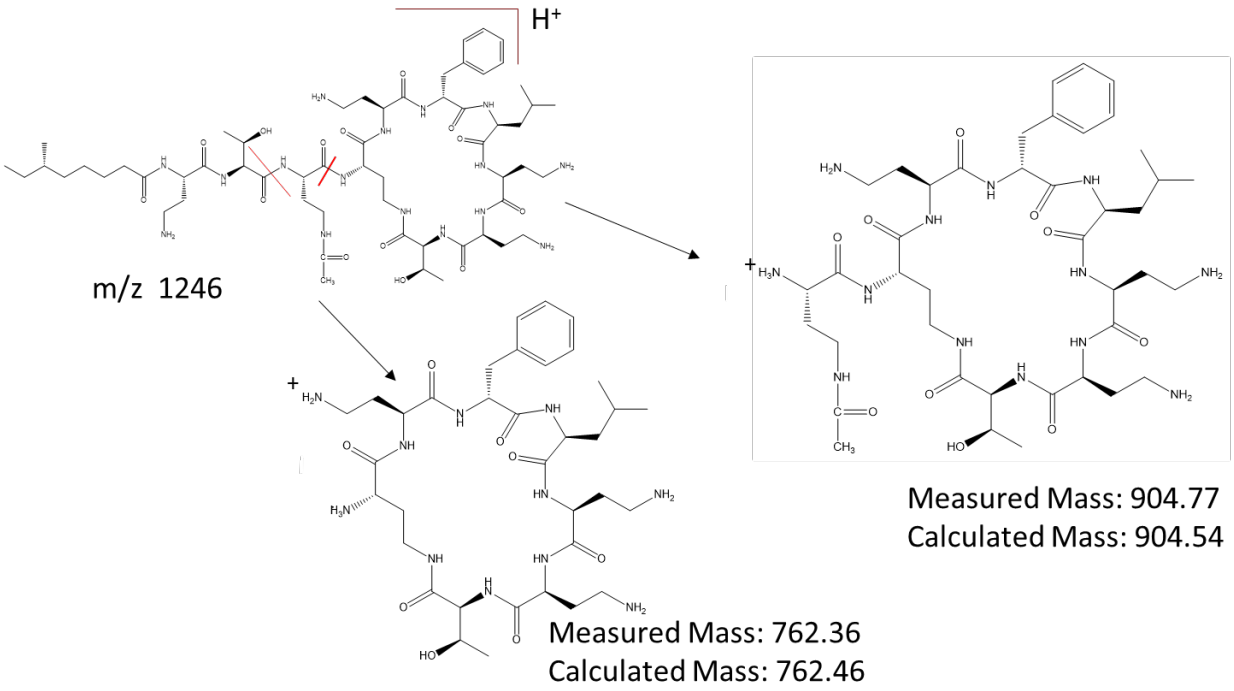


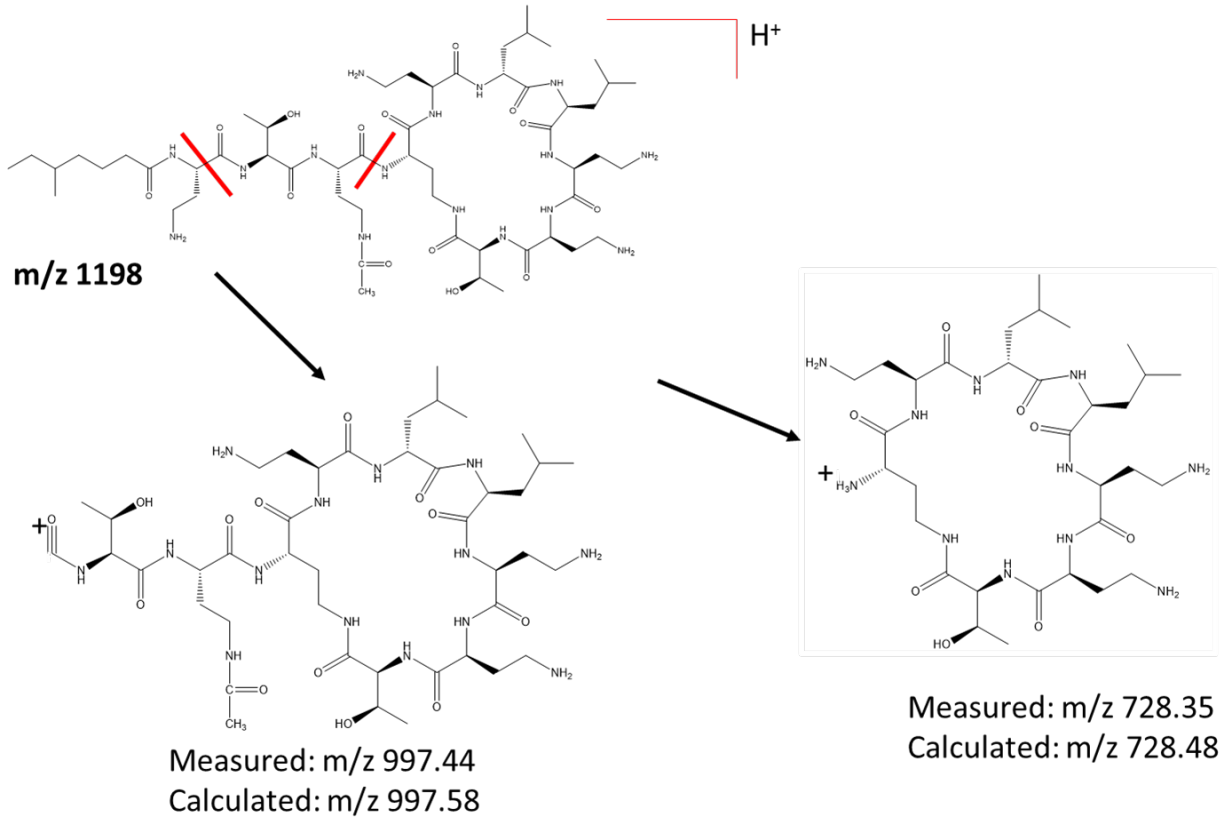
Figure S6. PA3944 pH profiles with polymyxin B as acceptor substrate. A) log k_{cat}/K_m ($M^{-1}s^{-1}$) as a function of pH, B) log K_m (mM) as a function of pH, and C) log V_{max} (µmol/min/mg) as a function of pH. Kinetic parameters used for the plots were calculated from non-linear regression fittings of the substrate saturation curves in Figure 6B to the Hill equation as described in materials and methods.



Scheme S1. Product ion formation indicative of the position of acetyl group attachment in the enzyme catalyzed acetylation of polymyxin B.



Scheme S2. Product ion formation process indicative of the position of acetyl group attachment in the enzyme catalyzed colistin acetylation.



References

1. Filippova, E. V.; Kuhn, M. L.; Osipiuk, J.; Kiryukhina, O.; Joachimiak, A.; Ballicora, M. A.; Anderson, W. F. (2015) A Novel Polyamine Allosteric Site of SpeG from Vibrio cholerae Is Revealed by Its Dodecameric Structure. *J. Mol. Biol. 427*, 1316-1334.

2. Blatzer, M.; Schrettl, M. F.; Sarg B FAU - Lindner, Herbert,H.; FAU, L. H.; Pfaller, K. F.; Haas, H. (2011) SidL, an Aspergillus fumigatus transacetylase involved in biosynthesis of the siderophores ferricrocin and hydroxyferricrocin. *Appl. Environ. Microbiol. 77,* 4959-66.

3. Fukuchi, J.; Kashiwagi, K. F.; Takio, K. F.; Igarashi, K. (1994) Properties and structure of spermidine acetyltransferase in Escherichia coli. *J. Biol. Chem. 269,* 22581-5.

4. Davies, A. M.; Tata, R.; Beavil, R. L.; Sutton, B. J.; Brown, P. R. (2007) L-Methionine sulfoximine, but not phosphinothricin, is a substrate for an acetyltransferase (gene PA4866) from Pseudomonas aeruginosa: structural and functional studies. *Biochemistry 46*, 1829-1839.

5. Majorek, K. A.; Kuhn, M. L.; Chruszcz, M.; Anderson, W. F.; Minor, W. (2013) Structural, functional, and inhibition studies of a Gcn5-related N-acetyltransferase (GNAT) superfamily protein PA4794: a new C-terminal lysine protein acetyltransferase from pseudomonas aeruginosa. *J. Biol. Chem. 288*, 30223-30235.

6. Kuhn, M. L.; Majorek, K. A.; Minor, W.; Anderson, W. F. (2013) Broad-substrate screen as a tool to identify substrates for bacterial Gcn5-related N-acetyltransferases with unknown substrate specificity. *Protein Sci. 22*, 222-230.

7. Pei, J.; Tang M FAU - Grishin, Nick,V.; Grishin, N. V. (2008) PROMALS3D web server for accurate multiple protein sequence and structure alignments. *Nucleic Acids Res. 36,* W30-4.

8. Pei, J.; FAU, K. B.; Grishin, N. V. (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res. 36,* 2295-300.

9. Krissinel, E.; Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst. D60,* 2256-68.