Structural Insights into *Escherichia coli* Phosphopantothenoylcysteine Synthetase by Native Ion Mobility-Mass Spectrometry

Daniel Shiu-Hin Chan,[a] Jeannine Hess,[a] Elen Shaw,[a] Christina Spry,[a] Robert Starley,[a] Claudio Dagostin,[a] Marcio V. B. Dias,[b] Ramesh Kale,[a] Vitor Mendes,[b] Tom L. Blundell,[b] Anthony G. Coyne*,[a] and Chris Abell*[a]

[a] Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom.

[b] Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, United Kingdom.

Correspondence: Anthony G. Coyne (agc40@cam.ac.uk) or Chris Abell (ca26@cam.ac.uk)

Abstract

E. coli CoaBC, part of the vital coenzyme A biosynthetic pathway in bacteria, has recently been validated as a promising antimicrobial target. In this work, we employed native ion mobility-mass spectrometry to gain structural insights into the phosphopantothenoylcysteine synthetase domain of CoaBC. Moreover, native mass spectrometry was validated as a screening tool to identify novel inhibitors of this enzyme, highlighting the utility and versatility of this technique both for structural biology and for drug discovery.

Introduction

Coenzyme A (CoA) is a ubiquitous and essential cofactor that is synthesized from pantothenate (Pan, also known as vitamin B$_5$) through a conserved pathway consisting of five

enzymatic steps.[1] In the second step, phosphopantothenoylcysteine synthetase (PPCS) catalyzes the Mg^{2+}-dependent conversion of 4′-phosphopantothenate (PPan) and ʟ-cysteine into 4′-phosphopantothenoylcysteine (PPC) using either CTP (in bacteria) or ATP (in eukaryotes) for activation.[2] In bacteria, PPan is first coupled with the α-phosphate of CTP to give an acyl cytidylate intermediate, 4′-phosphopantothenoyl cytidylate, with the concomitant release of pyrophosphate (PP_i) (Figure 1a). The CMP-activated acyl group is then transferred to the amino group of ʟ-cysteine, and PPC and CMP are released.[2]

Bacterial PPCS, also referred to as CoaB, is typically fused to the next enzyme of the CoA biosynthetic pathway, phosphopantothenoylcysteine decarboxylase (PPCDC, also referred to as CoaC in bacteria) to form the bifunctional CoaBC protein (*coaBC* gene product, previously known as *dfp*).[3] The bacterial CoaBC enzyme has been proposed to be a dodecamer based on gel filtration experiments[4] and by analogy with the related flavoprotein EpiD, for which the X-ray crystal structure has been solved.[5]

The CoaB domain crystallizes as a homodimer and the residues responsible for dimerization are conserved in all domains of life.[6] Dimerization is largely mediated by a dimerization domain comprising two antiparallel β-sheets (Thr^{198} to Asn^{210} and Ala^{278} to Lys^{306} in *E. coli*).[7] The latter region, which contains the highly conserved Lys^{298}-X-Lys-Lys (X = Ile in *E. coli*) motif, reaches out to the other protomer as a partially disordered loop that is situated near the entry of the neighbouring active site, with Lys^{298} contacting the γ-phosphate of CTP.[7]

[Insert Figure 1]

2

Figure 1 (a) Conversion of PPan into PPC proceeding through an acyl cytidylate intermediate, as catalyzed by PPCS. (b) Structures of CoaB inhibitors previously reported.

Kinetic and structural studies have indicated that the nucleobase binding sites of bacterial PPCS enzymes[7] are substantially different from the corresponding human enzyme,[8] rendering PPCS as a promising antimicrobial target.[1] The bacterial CoaB enzyme has a strict requirement for CTP, in contrast to human CoaB which uses ATP about four times more efficiently than it does CTP.[9] In *E. coli* CoaB (*Ec*CoaB), the nucleotide binding pocket is quite rigid and shows strong shape complementarity to the cytosine moiety.[7] In contrast, human CoaB has an enlarged and also more flexible nucleotide binding pocket that confers a preference for ATP, but that still can accommodate the smaller CTP molecule.[8] Recently, Mizrahi and co-workers have established the essentiality of CoaBC in a murine model of *Mycobacterium tuberculosis* infection.[10] Silencing of *coaBC* was bactericidal regardless of whether knockdown was initiated at the time of infection or during either the acute or chronic stages of infection, demonstrating that CoaBC is needed for *M. tuberculosis* to grow and persist in mice.

The first CoaB inhibitors were reported by Dotson and co-workers in 2009.[11] These compounds are structural mimics of the acyl cytidylate intermediate that forms during catalysis, with the reactive acyl-phosphate moiety of the natural intermediate replaced by either a stable phosphodiester or sulfamate isostere (Figure 1b). These compounds are non-competitive, slow-onset and tight-binding inhibitors that exhibit nanomolar IC_{50} and K_i values against the CoaB enzymes of *E. coli*, *Enterococcus faecalis*, and *Staphylococcus pneumoniae*, as well as up to 1000-fold selectivity over human CoaB. More recently, Burkart, Strauss and co-workers have characterized the natural product CJ-15,801,[12] an antibiotic effective against drug-resistant

strains of *Staphylococcus aureus* (MRSA), as the precursor to a tight-binding inhibitor of CoaB. CJ-15,801 structurally resembles Pan (differing only because it has a *trans*-substituted double bond in the β-alanine moiety), and it can be phosphorylated by the *S. aureus* Pan kinase (PanK) to furnish a PPan-like compound (Figure 1b) that binds to the CoaB domain. Subsequent cytidylylation by CoaB generates a tight-binding structural mimic of the native acyl cytidylate intermediate with nanomolar potency against the *E. coli* CoaB domain and *S. aureus* CoaBC.[12] However, despite the CoaB enzyme having been characterized over ten years ago, no non-mimetic inhibitors of CoaB have been reported in the literature. A better understanding of the structure and interactions of CoaB could aid the development of therapeutic compounds against this promising antimicrobial target.

Native electrospray ionization-mass spectroscopy (ESI-MS), and particularly native nanoelectrospray ionization-mass spectrometry (nESI-MS), has emerged as a powerful technique for the study of biomolecular complexes.[13-15] Moreover, ESI-MS can be interfaced with ion mobility (IM) spectrometry, which can provide an orthogonal mode of separation of gaseous ions. In ion mobility-mass spectrometry (IM-MS), ions traverse a chamber of neutral gas particles under the influence of an electric field, and become separated on the basis of differences in their rotationally-averaged collision cross-section (CCS).[16-18] In this work, native nESI-IM-MS experiments were conducted to provide structural insights into the *Ec*CoaB domain. Our results indicate that CoaB exists in an equilibrium of monomers and dimers in solution, with dimer formation being favoured in the presence of CTP. Equilibrium titration experiments suggest that the binding of CTP to CoaB is cooperative, and also confirm the preference of *Ec*CoaB for CTP over ATP. IM-MS analysis was applied to characterize the size and unfolding profiles of the *Ec*CoaB monomer and dimer in the gas phase. Finally, native MS was used to

screen a small library of sulfonamide compounds against *Ec*CoaB, revealing a new class of small-molecule, non-mimetic inhibitors of CoaBC.

Materials and Methods

Mass spectrometry

nESI-IM-MS spectra were recorded on a Synapt HDMS mass spectrometer (Waters, Manchester, UK) modified for studying high masses and equipped with a traveling wave (TW) IM cell. *Ec*CoaB was exchanged into 200 mM ammonium acetate (pH 7.0, unless otherwise stated) using Micro Bio-Spin 6 chromatography columns (Bio-Rad, UK). Protein samples were equilibrated at room temperature in the absence or presence of the indicated concentrations of ligands for at least 15 min before analysis. An aliquot of protein (2.5 μL) was then injected into a gold-coated borosilicate emitter (Thermo Scientific, UK) for sampling. Typical conditions were capillary voltage 1.5 kV, cone voltage 50 V, trap collision voltage 20 V, transfer collision voltage 20 V, source temperature 20 °C, backing pressure 3–4 mbar, trap pressure $3–4 \times 10^{-2}$ mbar, IM (N_2) pressure $5–6 \times 10^{-1}$ mbar and time-of-flight (TOF) pressure $7–8 \times 10^{-7}$ mbar. For IM experiments, the mass spectrometer was operated in mobility TOF mode using a wave velocity of 250 ms^{-1} and a wave height of 10 V. To effect collision-induced unfolding (CIU), the trap collision voltage was raised in 2 V increments from 10 to 80 V. Avidin, concanavalin A, alcohol dehydrogenase and pyruvate kinase were used as calibrants for CCS determination. All mass spectra were calibrated externally with cesium iodide (10 mg/mL). Data acquisition and processing were performed using MassLynx 4.1 (Waters) and DriftScope 2.5 (Waters).

K_d determination

The dissociation constant (K_d) for the *Ec*CoaB monomer-dimer equilibrium was determined using the procedure of Zenobi and co-workers,[19] which was based on earlier work by the group of Gabelica.[20] Briefly, the equilibrium association constant (K_a) relating the *Ec*CoaB monomer (*M*) and dimer (*D*) concentration is described using equation 1,

$$K_a = \frac{[D]}{[M]^2} = \frac{1}{K_d} \tag{1}$$

The response factors (R_M and R_D) governing the relationship between the concentrations of *M* and *D* and their ion intensities (I_M and I_D respectively) are described using equations 2 and 3,

$$I_M = R_M \times [M] \tag{2}$$

$$I_D = R_D \times [D] \tag{3}$$

Given a total *Ec*CoaB monomer concentration of $[M_0]$ (where $[M_0] = [M] + 2[D]$), and setting the ratio of response factors of the dimer and monomer as *R*, the equations 4 and 5 can be derived.

$$R = \frac{R_D}{R_M} \tag{4}$$

$$\frac{I_M}{I_D} = \frac{1 + \sqrt{1 + 8K_a M_0}}{2R K_a M_0} \tag{5}$$

After determination of I_M and I_D from the mass spectra, equation 5 can be solved for *R* and K_a by double-parametric non-linear curve fitting. Non-linear parameter uncertainties were estimated using a Monte Carlo method as described by Si and co-workers.[21]

Ion mobility calibration

Calibration with ions of known CCS is required on a TW-IM instrument because the analytical equations governing ion mobility are complex owing to the non-linear time- and space-varying dependence of the electric field.[22, 23] Briefly, the drift times (t_D) (in ms) are corrected for mass-dependent flight time using equation 6,

$$t'_D = t_D - \frac{C\sqrt{(m/z)}}{1000} \tag{6}$$

where t'_D is the corrected drift time of the calibrant ions, C is the "EDC (Enhanced Duty Cycle) delay coefficient" and is equal to 1.57, and m/z is the mass-to-charge ratio of the observed ion. The calibrant CCS (Ω) values are then corrected for both ion charge state (z) and reduced mass (μ) according to equation 7,

$$\Omega' = \frac{\Omega}{(z \times \sqrt{(1/\mu)})} \tag{7}$$

where Ω' is the corrected CCS value and μ is the reduced mass of the ion-N_2 collision complex. Next, $\ln(t_D')$ is plotted against $\ln(\Omega')$ and the slope (X) of the graph is determined. A new corrected drift time (t''_D) is then calculated using equation 8,

$$t''_D = t'_D{}^X \times z \times \sqrt{(1/\mu)} \tag{8}$$

Finally, Ω is re-plotted against t''_D to generate a linear calibration plot that can then be used to determine CCS values for analyte ions. Unless otherwise stated, all CCS values are quoted as N_2 values. PULSAR[24] was used to perform CCS calibration,[23] generate CIU plots and calculate

CIU$_{50}$ values. As algorithms for CCS calculation from X-ray crystal structures are parameterized for helium collision gas,[22] theoretical CCS(He) values were first determined using the projection approximation (PA) algorithm implemented in DriftScope, and then converted to theoretical CCS(N$_2$) values as described by Bush *et al.*[25] However, the PA algorithm, which functions by finding the average projection (or "shadow") of an ion onto a defined plane, generally underestimates CCS values for biomolecules because it does not take into account the greater gas slow-down effect for cavities.[22] To account for this, an empirically-determined scaling factor of 1.14 is introduced.[14] Thus, the "corrected" theoretical CCS value (CCS$_{calc}$) was calculated using equation 9,

$$CCS_{calc} = 1.14 \times CCS_{PA} \times (\frac{M_{exp}}{M_{PDB}})^{\frac{2}{3}} \tag{9}$$

where CCS$_{PA}$ is the CCS calculated from the X-ray crystal structure using the PA algorithm, M_{exp} is the experimental mass of the protein and M_{PDB} is the mass calculated from the crystal structure. In the case of the *Ec*CoaB dimer, CCS$_{PA}$ was determined to be 3482 Å^2 from the reported X-ray crystal structure (PDB: 1U7U),[7] and M_{exp} and M_{PDB} were 27942 and 24834 Da, respectively. Thus, CCS$_{calc}$ was calculated to be 4258 Å^2.

Results and Discussion

Native MS of the EcCoaB domain

Although the bifunctional *Ec*CoaBC enzyme has been shown by size exclusion chromatography to exist as a dodecamer,[4] the isolated CoaB domain crystallizes as a homodimer.[6] Strauss and Begley have shown that the product of the PPCS (CoaB) reaction is released from *Ec*CoaBC

before it returns to a different active site for transformation by PPCDC (CoaC).[26] Therefore the functional significance of the CoaB-CoaC fusion is not currently apparent. As the isolated CoaB domain of *E. coli* is fully functional,[6] in this study we have used the isolated CoaB domain as a surrogate for CoaBC in binding and mechanistic studies. The dimeric *Ec*CoaB enzyme (56 kDa), being far smaller than the dodecameric CoaBC protein (*ca.* 500 kDa), is much more amenable to native MS analysis, and easier to express and purify. Doubly-His$_6$-tagged *Ec*CoaB, containing residues 182 to 406 of *Ec*CoaBC, was overexpressed and purified in a high yield from *E. coli*, and protein identity and purity were assessed by denaturing LC-MS (Figure S1) and SDS-PAGE analysis.

Typical buffers used for protein expression and purification are incompatible with ESI-MS owing to ion suppression as well as the formation of extensive protein-ion adducts. Therefore, *Ec*CoaB was first exchanged into 200 mM ammonium acetate (NH$_4$OAc) before native MS analysis. The native mass spectrum of *Ec*CoaB alone showed that *Ec*CoaB exists as an equilibrium of monomeric and dimeric species (Figure 2a). The *Ec*CoaB monomer, centered around the 10+ charge state, had an observed mass of 28002 ± 24 Da which was close to the theoretical mass of the protein (without its N-terminal methionine) and the mass observed by denaturing LC-MS (27942 Da; Figure S1). Meanwhile, the *Ec*CoaB dimer, centered around the 15+ charge state, had an observed mass of 55999 ± 42 Da, which was also close to the theoretical mass of 55883 Da for two molecules of *Ec*CoaB. The masses of the *Ec*CoaB monomer and dimer observed by native MS are slightly higher than the theoretical masses presumably due to the adduction of weakly-bound buffer or water molecules to the protein under the soft ionization conditions employed.

[Insert Figure 2]

Figure 2 The dimerization of *Ec*CoaB is concentration-dependent. Native mass spectra of *Ec*CoaB in 200 mM NH$_4$OAc at (a) 60 μM, (b) 50 μM, (c) 40 μM, (d) 30 μM and (e) 20 μM, with monomer (M) and dimer (D) charge states indicated.

Non-covalent interactions in native MS can be affected by the concentration of NH$_4$OAc buffer used.[27] In this study, the effect of NH$_4$OAc concentration on the appearance of the *Ec*CoaB native mass spectrum was investigated. At 50 mM NH$_4$OAc, higher-order species corresponding to trimers, tetramers, pentamers and hexamers of *Ec*CoaB were observed (Figure S2a). However, the abundance of the higher-order species (trimers and above) was greatly diminished at 200 mM NH$_4$OAc (Figure S2b) or 500 mM NH$_4$OAc (Figure S2c), suggesting that the oligomers arise as a consequence of non-specific aggregation of *Ec*CoaB. This finding is similar to a recent native MS study of CK2 kinase, where the oligomerization state of the protein was also found to be dependent on ionic strength.[28] The proportion of monomer and dimer species was similar at 200 and 500 mM NH$_4$OAc. However, the electrospray became more difficult to maintain at 500 mM NH$_4$OAc, therefore a concentration of 200 mM NH$_4$OAc was chosen for subsequent experiments.

Determination of affinity of *Ec*CoaB dimerization

The peak intensities of the different species in the mass spectrum are often considered to be proportional to the concentration of those species in solution. However, this assumes that the probability of ions being ionized, transmitted, and detected is the same for all species, which may not be true for the monomer and dimer of *Ec*CoaB given their different sizes. A more accurate equilibrium dissociation constant can be determined by taking into account the fact that the

response factors, which are the products of efficiencies of all the processes affecting the signal intensities of the ions, may be different for different species. In this study, the method of Zenobi and co-workers[19] (based on an earlier work by Gabelica *et al.*[20]) was used to fit the equilibrium association constant (K_a) of the EcCoaB monomer-dimer equilibrium (equation 1) as well as the ratio (R) between the individual dimer (R_D) and monomer (R_M) response factors (equation 4) from the measured peak ion intensities of the dimer (I_D) and monomer (I_M) according to equation 5.

To derive R and K_a, the native mass spectrum of EcCoaB was recorded at different concentrations of the enzyme (Figure 2a-e). The results show that as the concentration of protein was lowered, the abundance of the EcCoaB dimer decreased relative to the monomer. Using nonlinear regression on equation 5, R and K_a were determined to be 1.8 and 3.5×10^4 M^{-1} (corresponding to K_d = 29 µM), respectively. Figure S3 shows a graph of experimental I_M/I_D values against total EcCoaB monomer concentration (M_0) as well as the fitted curve derived from nonlinear regression of equation 5. The value of R (1.8) is within the range of 0.5 to 6.7 for biomolecules reported in the literature.[19]

Gabelica and co-workers emphasized that a very high quality data set from the titration experiment is required to obtain precise fitting values.[20] However, in a similar way to that observed by Zenobi and co-workers,[19] the I_M/I_D values showed significant scatter around the fitted curve. Therefore, the confidence intervals that were obtained by fitting the data using unconstrained values of R and K_a were relatively broad (Table 1, Entry 1). More precise K_a values may be obtained by first constraining the choice of R. With reference to Zenobi and co-workers,[19] the fitted K_a values when R was restricted to 1, 2 or 4, were determined to be 8.6×10^4, 2.9×10^4 and 1.1×10^4 M^{-1} respectively (Table 1, Entries 2–4). The K_a values decrease with

R as expected because increasing the value of R (the ratio of response factors between the dimer and the monomer) decreases the estimated dimer to monomer concentration ratio. The K_d values determined suggest that the interaction between two EcCoaB monomers is of low affinity.

Table 1. R and K_a values obtained by nonlinear regression of equation 5, using data obtained from the EcCoaB titration experiment (Figure 2). Corresponding K_d values (where $K_d = K_a^{-1}$) are shown to aid interpretation. 95% confidence intervals determined by the Monte Carlo method are shown in parentheses.

Entry	R	K_a / x 10^4 M^{-1}	K_d / μM
1[a]	1.8 (0.9-3.0)	3.5 (1.6-5.4)	29 (19-63)
2	1	8.6 (7.7-9.3)	12 (11-13)
3	2	2.9 (2.7-3.1)	34 (32-37)
4	4	1.1 (1.0-1.2)	91 (83-100)

[a]R was unconstrained. This gave broader confidence intervals for both R and K_a. In entries 2–4, R was constrained to 1, 2 or 4 respectively to generate more precise estimates for K_a.

Influence of solution pH on Ec*CoaB dimerization:*

Since pH governs the strength of many protein-protein interactions, it is of interest to determine the influence of solution pH on the dimerization of the EcCoaB domain. The native mass spectra of EcCoaB at pH values of 7.0, 5.0 and 3.0 are shown in Figure 3. The spectra show that the native mass spectra of EcCoaB at pH 7 (Figure 3a) and pH 5 (Figure 3b) are broadly similar, except for a slight decrease in EcCoaB dimer abundance and the appearance of higher

charge states of the monomer (13+ and 14+), suggesting the partial unfolding of the protein. However, the appearance of the spectrum was dramatically changed at pH 3 (Figure 3c). The dimer completely disappeared, and instead, a series of highly-charged monomeric *Ec*CoaB ions were observed and accompanied by the appearance of aggregates in the higher *m/z* region. All of these features are indicative of the chemical denaturation of *Ec*CoaB. As the protein unfolds, its surface area increases, which allows for a greater accumulation of positive charge on the molecule during the electrospray process. The exposure of interior lipophilic residues could allow the protein to interact via hydrophobic interactions to form higher-order aggregates. Additionally, the lowering of the pH is likely to disrupt polar interactions that are important for dimerization, leading to the disruption of the *Ec*CoaB protein-protein interface even while some of the monomer remains natively folded (as indicated by the intact charge envelope encompassing the 9+ to 12+ charge states of the monomer in Figure 3c).[7] This suggests that the tertiary structure of the *Ec*CoaB monomer is relatively more resistant to acid-induced denaturation than the quaternary interactions between the protomers in the dimer.

[Insert Figure 3]

Figure 3 Lowering of pH disrupts *Ec*CoaB dimerization and partially unfolds the protein. Native mass spectra of *Ec*CoaB (30 μM) in 200 mM NH₄OAc at (a) pH 7, (b) pH 5, and (c) pH 3, with monomer (M), dimer (D) and unfolded monomer (U) charge states indicated. For clarity, charge states for the unfolded monomer are not assigned in (c). The molecular masses of the aggregates in (c) were unable to be assigned. Note the change in *x*-axis scale in the higher *m/z* region.

Ion mobility-mass spectrometry of Ec*CoaB*

IM-MS is a technique that separates gaseous ions of different size and shape, based on differences in their mobility through a drift cell of neutral gas particles in the presence of an electric field.[16] In general, compact ions (low CCS) travel faster through the drift cell than extended ions (high CCS), due to fewer collisions with the neutral gas. IM-MS therefore provides an orthogonal dimension of resolution compared to MS alone, and can provide topological information that is complementary to other structural biological techniques.[18] In this study, IM-MS experiments were performed to investigate the structural characteristics of EcCoaB in the gas phase.

Table 2 presents CCS values for individual charge states of the EcCoaB monomer and dimer at either 0 or 5% (*v/v*) DMSO. CCS values generally increase with charge state, which could be due to the increased internal electrostatic repulsion of protein ions at higher charge. At 0% DMSO, the population-weighted average CCS of the EcCoaB monomer and dimer ions are 2556 ± 77 and 4147 ± 124 Å^2, respectively. For comparison, the theoretical CCS$_{calc}$ of the EcCoaB dimer calculated from the X-ray crystal structure (PDB: 1U7U)[7] was 4258 Å^2. As the experimental and theoretical CCS values are in agreement within the limits of experimental error, this suggests that the EcCoaB dimer remains natively folded in the gas-phase. In contrast, the experimental CCS of the EcCoaB monomer (2576 ± 79 Å^2) was 5% higher than the theoretical CCS (2460 Å^2) of a model of the monomer generated by artificial truncation of the dimer structure. This may be because the EcCoaB monomer adopts a slightly more extended conformation when it is not bound and stabilized by the second subunit.

Table 2. CCS values of *Ec*CoaB monomer and dimer ions as determined by IM-MS at 0 and 5% (*v/v*) DMSO. The population-weighted average CCS values of the monomer and dimer across all charge states are also presented. Errors in CCS are estimated to be 3%.[14, 25]

Monomer			Dimer		
Charge state	CCS / Å^2 (0% DMSO)	CCS / Å^2 (5% DMSO)	Charge state	CCS / Å^2 (0% DMSO)	CCS / Å^2 (5% DMSO)
8+	-	2307	11+	-	3674
9+	2358	2407	12+	-	3747
10+	2502	2542	13+	-	3841
11+	2755	-	14+	3938	3945
12+	3028	-	15+	4031	-
			16+	4199	-
			17+	4442	-
Average	2556	2382	Average	4147	3803

In the presence of 5% (*v/v*) DMSO, the charge state distributions of both the *Ec*CoaB monomer and dimer shifted to lower charge states (Figure S4). This is accompanied with a compaction of the protein structure (average CCS values are lowered by 7–8%, Table 2), and is consistent with previous studies on the effects of DMSO on protein size both in solution and in the gas phase.[29-31] However, the relative abundance of monomers and dimers was not substantially changed by the presence of DMSO.

Collision-induced unfolding of EcCoaB

The use of collisional activation to generate unfolding plots (or "fingerprints") is a recently developed technique that has been used to study the gas-phase stability of proteins.[24, 32]

Collision-induced unfolding (CIU) fingerprints are obtained by recording the native mass spectra of a protein at incrementally higher acceleration voltages within the collision cell of a mass spectrometer. As protein ions unfold through collisions with neutral gas molecules, changes in their CCS can be tracked as a function of collision voltage, allowing the quantitation of a protein's resistance to unfolding in the gas phase. CIU IM-MS experiments have been used, for example, to evaluate the effects of ligands on protein stability,[24] or to differentiate between Type I and Type II kinase inhibitors of Abl kinase via comparison of fingerprints.[33]

Figure 4. CIU IM-MS reveals gas-phase unfolding transitions of *Ec*CoaB. CIU fingerprints of indicated charge states of the *Ec*CoaB (a) monomer (M) and (b) dimer (D). For the 9+ state of the monomer, native-like (**N**), intermediate (**I**) and extended (**E**) forms of the protein are labeled.

The CIU fingerprints for *Ec*CoaB monomeric (Figure 4a) and dimeric (Figure 4b) ions are complex and reveal multiple unfolding transitions over the range of collision voltages tested. For instance, for the 9+ state of the monomer, the native-like protein (N, CCS $\approx$ 2400 Å^2) initially transitions into a partially unfolded intermediate form (I, CCS $\approx$ 2800 Å^{2+}) at moderate collision voltages, then shifts into a more extended form (E, CCS $\approx$ 3100 Å^{2+}) as the collision energy is increased further. It is assumed that E represents the most unfolded form of the monomeric 9+ ion as populations with larger CCS values were not observed even as the collision voltage was increased to 150 V, at which point protein degradation also began to occur (data not shown). The CIU50 values of the native-like (N) and intermediate (I) forms of the monomeric 9+ ion (i.e. the collision voltage needed to deplete 50% of the population of the conformation) were determined

to be 34 and 62 V, respectively. For the other monomeric and dimeric *Ec*CoaB charge states, multiple intermediate species were observed in the CIU plots, indicative of complicated unfolding pathways for those ions.

CIU plots for *Ec*CoaB monomeric 9+ and 10+ states are broadly similar in the presence of 5% (*v/v*) DMSO (Figure S5a) compared to the respective plots in the absence of DMSO. The 8+ state of the monomer, accessible only via charge reduction conditions with DMSO, did not unfold even at the highest collision voltages tested. The greater CIU stability of lower charge states is known [31, 34] and can be attributed to the lesser internal electrostatic repulsion experienced by proteins of lower charge. For the *Ec*CoaB dimer in the presence of DMSO, the 11+ state did not unfold, while the 12+ and 13+ states underwent one and three unfolding transitions respectively over the collision voltage range tested (Fig. S5b).

Interaction of EcCoaB with CTP

Since *Ec*CoaB utilizes CTP to activate the carboxylate group of PPan, the interaction between CTP and *Ec*CoaB was studied by native MS. Interestingly, the addition of 50 µM CTP to *Ec*CoaB (25 µM) (Figure 5b) increased the proportion of dimeric *Ec*CoaB relative to the proportion present in the absence of CTP (Figure 5a). With CTP present, the K_d value for the monomer-dimer equilibrium was estimated to be 4 µM, which is substantially lower than the corresponding K_d value of 29 µM determined in the absence of CTP (Table 1). This result indicates that CTP stabilizes the dimeric form of *Ec*CoaB, and is consistent with the observation that CTP contacts both subunits of the dimer in the X-ray co-crystal structure of the *Ec*CoaB

domain.[7] The proportion of dimer was further increased at higher concentrations of CTP (Figure S6a-e), and at 200 μM CTP, only a small amount of the monomer could be detected.

At 50 μM CTP, 66% of the *Ec*CoaB monomer was unbound while 34% was singly-bound by CTP (Figure 5b). In contrast, the majority of the dimer was observed in singly-bound (32%) or doubly-bound (38%) states. This can also be rationalized on the basis that the CTP-binding site of *Ec*CoaB is formed of residues from both subunits of the dimer.[7] Therefore, the binding affinity of CTP for the monomer would be expected to be weaker than for the dimer.

[Insert Figure 5]

Figure 5 *Ec*CoaB selectively binds CTP over ATP, and CTP stabilizes the dimer. Native MS of *Ec*CoaB (25 μM) in the (a) absence, or presence of (b) CTP (50 μM) or (c) ATP (50 μM). In (a), the monomer (M) and dimer (D) charge states are indicated. In (b) and (c), the number of ligands bound to each monomer or dimer species is indicated.

A small amount (17%) of the *Ec*CoaB dimer was bound to three molecules of CTP (Figure 5b). As each dimer only has two CTP-binding sites, this suggests that the third molecule of CTP is bound non-specifically to the protein. The observation that non-specific binding is observed at only a mild excess of CTP over protein (50 μM of CTP *vs* 25 μM of *Ec*CoaB) could possibly be accounted for by the fact that electrostatic interactions, which are presumably significant between the doubly-anionic CTP molecule and the positively-charged *Ec*CoaB protein, are

19

generally thought to be strengthened in the gas phase. In previous native MS studies, a high degree of non-specific binding was also observed between creatine kinase and ATP,[35] and between GroEL and ATP.[36] With *Ec*CoaB, non-specific binding was even more pronounced at higher CTP concentrations, with up to four CTP molecules bound per dimer at 200 µM of CTP, and two molecules of CTP bound per monomer (Figure S6e).

Several methods have been developed to tackle the issue of non-specific binding in native ESI-MS experiments. van der Rest and co-workers[35] developed a deconvolution method that modeled the specific binding as a binomial distribution and the non-specific binding as a Poisson distribution (which was originally proposed by Klassen and co-workers).[37] Alternatively, Guan and co-workers have modeled the non-specific component as a power-law distribution,[38] while in yet another approach, the group of Sharon and Horovitz have modeled all non-specific sites with a single binding constant.[39]

In this study, to isolate the effect of non-specific binding in this system, the interaction between *Ec*CoaB and ATP was monitored using native MS. The native mass spectrum of *Ec*CoaB in the presence of ATP (50 µM) is shown in Figure 5c. As expected, the proportion of *Ec*CoaB dimer was not increased by the addition of ATP, unlike with CTP. This indicates that ATP does not stabilize the *Ec*CoaB dimer, presumably because it is too large to enter into the nucleobase binding pocket of the enzyme. Moreover, ATP showed weaker binding to the dimer compared to CTP, as reflected by the lower proportion of ligand-bound protein for ATP compared to CTP at the same concentration. Based on the findings of Strauss *et al.*,[3] which showed that *Ec*CoaB enzyme activity was not promoted even at up to 10 mM of ATP, we assumed that *Ec*CoaB-ATP complexes observed by native MS were non-specific. With this assumption, the specific binding component of CTP can be estimated by subtracting the

occupancy of ATP from the occupancy of CTP at each site of the protein. Applying this method to the data shown in Fig. 5 revealed that the specific binding constant of CTP to the monomer is 380 μM, while the average specific binding constant of CTP to the two sites of the dimer is 70 μM. The latter affinity is composed of a specific binding constant of 580 μM for the first CTP molecule binding to the dimer, and a specific binding constant of 9 μM for the second CTP molecule binding to the dimer, which indicates some degree of positive cooperativity in the binding. The average specific binding constant (70 μM) is also consistent with the K_M value of *Ec*CoaB for CTP, which has been reported to be 48 μM.[12]

Screening for EcCoaB inhibitors by native MS

Due to its ability to detect protein-ligand interactions, native MS is finding increasing use as a compound screening technique, including in fragment-based drug discovery (FBDD).[40-42] In line with our continuing interest in the application of FBDD methods to develop new antimicrobial compounds, native MS was applied to screen a small library of fragment-sized sulfonamides against *Ec*CoaB. These analogues were designed and synthesized based on the structure of sulfonamide fragment hit **1** that was identified in a fragment screen using a combination of differential scanning fluorimetry and ligand-based NMR studies (data not reported). The library of analogues **2-11** retain the sulfonamide core of fragment **1**, but have simpler aromatic motifs in place of the synthetically challenging aniline and pyrimidine moieties of 1 (Figure 6).

[Insert Figure 6]

Figure 6 Structures of sulfonamide compounds **1-11** examined in this work.

21

Additionally, the structure-activity relationships (SAR) around the phenyl substituent attached to the sulfonamide nitrogen atom were investigated by the installation of a carboxyl group and/or a bromine atom at various positions of the ring (except for compound **9**, which carries a 2-methoxypyridin-3-yl substituent in place of the phenyl ring). All of the library compounds (with the exception of naphthalene compound **11**) have either 19 or 20 heavy atoms, putting them close to the upper limit in size of typical fragments. The synthesis and characterization of compounds **2**–**11** are described in the Electronic Supplementary Information.

To minimize non-specific binding, the compounds were tested at a concentration of 30 µM with *Ec*CoaB at 15 µM, giving only a 2:1 excess of ligand to protein. The native mass spectra of *Ec*CoaB incubated with compounds **1-11** are shown in Figure 7. The DMSO concentration was 1% (*v/v*). In general, the compounds did not increase the ratio of the dimer to the monomer, indicating that unlike CTP they do not stabilize the *Ec*CoaB dimer. This result is not surprising as the compounds are significantly smaller than CTP and thus less likely to be able to simultaneously contact both subunits of the dimer, even if they were binding to the CTP-binding site.

The native MS screen revealed five hits (compounds **4**, **5**, **6**, **9** and **11**) that bound to *Ec*CoaB, as indicated by the appearance of additional peaks corresponding to the mass of the protein-ligand complex in the spectra. Due to the relatively small size of the ligands, only ligands bound to the monomer could be discretely resolved. For compounds **6** and **11**, a small proportion of doubly-bound *Ec*CoaB monomer species was observed, which could either represent non-specific binding, or possibly two fragments binding at the same time to the active site.

To validate the screening hits, they were tested for *Ec*CoaB inhibitory activity using an enzymatic assay. In this assay, the production of pyrophosphate that occurs concomitantly with

the first half-reaction catalyzed by *Ec*CoaB (the formation of 4'-phosphopantothenoyl-CMP) was monitored. As *Ec*CoaB was used as a surrogate for the bifunctional *Ec*CoaBC in the MS studies, *Ec*CoaBC (16 nM) was used in the enzymatic assay to ensure that inhibitors identified were biologically relevant. At a concentration of 1 mM (and CTP, phosphopantothenate and L-cysteine at concentrations of 125, 125 and 500 µM, respectively (see Supporting Information)), the initial sulfonamide fragment **1** showed the weakest *Ec*CoaBC inhibitory activity of 6 ± 3%, while the most potent compound was the naphthalene analogue **11**, which inhibited *Ec*CoaBC activity by 92 ± 5% (Table S1). Pleasingly, there was good agreement between the results of the native MS assay and the biochemical assay. The five native MS hits (compounds **4**, **5**, **6**, **9** and **11**) were also the five most active compounds in the biochemical assay, with an average *Ec*CoaBC inhibition of 57% at 1 mM (range of 29 to 92%) (Figure S7). Conversely, the compounds that were not detected to bind by native MS (compounds **1**, **2**, **3**, **7**, **8** and **10**) showed the lowest activity in the biochemical assay, with an average inhibitory value of 18% (range of 6 to 25%). Taken together, these results validate the use of native MS with monofunctional *Ec*CoaB as a simple and rapid screening technique for identifying inhibitors of *Ec*CoaBC. In future work, native MS could be used to derive binding affinities for the inhibitors identified here (as was done for CTP) and the correlation between MS-determined binding affinities for *Ec*CoaB and inhibitory activity against *Ec*CoaBC explored. Moreover, CIU IM-MS could be used to investigate the effect of the inhibitors and other known ligands on *Ec*CoaB stability.

[Insert Figure 6]

Figure 7 Screening of sulfonamide compounds **1-11** against *Ec*CoaB by native MS. In each experiment, the protein concentration was 15 µM and the ligand concentration was 30 µM. For

the screening hits (compounds **4**, **5**, **6**, **9** and **11**), the number of ligands bound to each monomer species is indicated.

Conclusions

In this work, native nESI-IM-MS has been conducted to provide structural insights into the *Ec*CoaB domain of *Ec*CoaBC. While *Ec*CoaB crystallizes as a dimer, our experiments demonstrate that in solution an equilibrium exists between the monomer and dimer. Dimer formation can be promoted by CTP, but not ATP, which is consistent with previous structural and biochemical data. The gas-phase unfolding behavior of monomeric and dimeric *Ec*CoaB ions was studied by IM-MS, revealing complex CIU fingerprints for most charge states. Finally, native MS was applied to screening of a small library of fragment-sized sulfonamide compounds against *Ec*CoaB, and the hit compounds were validated using a biochemical assay. Good agreement was observed between the native MS binding studies with *Ec*CoaB and the enzyme inhibition assays performed with *Ec*CoaBC. The active sulfonamide compounds represent a new class of small-molecule, non-mimetic inhibitors of CoaBC, a recently validated antimicrobial target. This study thus highlights the utility and versatility of native MS both for structural biology and for drug discovery.

Abbreviations

4′-phosphopantothenate (PPan), pyrophosphate (PP_i), cytidine triphosphate (CTP), Adenosine triphosphate (ATP), phosphopantothenoylcysteine synthetase (PPCS), phosphopantothenoylcysteine decarboxylase (PPCDC), Multidrug resistant staphylococcus aureus (MRSA), Pantothenate kinase (PanK), Native electrospray ionization-mass spectroscopy (ESI-MS), native nanoelectrospray ionization-mass spectrometry (nESI-MS), ion mobility-mass spectrometry (IM-MS), collision cross-section (CCS), Dimethyl sulfoxide (DMSO), Collision-

induced unfolding (CIU), Dissociation constant (K_d), fragment-based drug discovery (FBDD), structure-activity relationships (SAR)

Present Addresses

CS: Current address: Research School of Biology, The Australian National University, ACT 2601, Australia.

MVBD: Current address: Department of Microbiology, Institute of Biomedical Science, University of São Paulo, Av. Prof. Lineu Prestes, 1374, 05508-000, São Paulo-SP, Brazil.

Author Contributions

AGC, TLB and CA designed and supervised the project. DSHC and ES performed and analysed the mass spectrometric experiments. The compounds used in this study were synthesized and characterised by JH, RS, CD and RK. The biochemical assay was devised and carries out by CS and JH. The cloning and expression of CoaB was carried out by MVBD and VM. DSHC wrote the first version of the manuscript and revised it with help from AGC and JH. All authors commented on the final version of the manuscript.

Funding Sources

Competing Interests

The Authors declare that there are no competing interests associated with the manuscript.

REFERENCES

1. Moolman, W. J. A., de Villiers, M. and Strauss, E., (2014), Recent advances in targeting coenzyme A biosynthesis and utilization for antimicrobial drug development, *Biochem. Soc. Trans.*, **42**, 1080-1086.

2. Yao, J., Patrone, J. D., and Dotson, G. D., (2009), Characterization and kinetics of phosphopantothenoylcysteine synthetase from *Enterococcus faecalis*, *Biochemistry*, **48**, 2799-2806.

3. Strauss, E., Kinsland, C., Ge, Y., McLafferty, F. W., and Begley, T. P., (2001), Phosphopantothenoylcysteine synthetase from *Escherichia coli.* Identification and characterization of the last unidentified coenzyme A biosynthetic enzyme in bacteria, *J. Biol. Chem.*, **276**, 13513-13516.

4. Kupke, T., Uebele, M., Schmid, D., Jung, G., Blaesse, M., and Steinbacher, S., (2000), Molecular characterization of Lantibiotic-synthesizing enzyme EpiD reveals function for bacterial Dfp proteins in Coenzyme A biosyntheis, *J. Biol. Chem.*, **275**, 31838-31846.

5. Blaesse, M., Kupke, T., Huber, R. and Steinbacher, S., (2000), Crystal structure of the peptidyl-cysteine decarboxylase EpiD complexed with a pentapeptide substrate, *The EMBO Journal*, **19**, 6299-6310.

6. Kupke, T., (2002), Molecular characterization of the 4'-phosphopantothenoylcysteine synthetase domain of bacterial Dfp flavoproteins, *J. Biol. Chem.*, **277**, 36137-36145.

7. Stanitzek, S., Augustin, M. A., Huber, R., Kupke, T., and Steinbacher, S., (2004) , Structural basis of CTP-dependant peptide bond formationin Coenzyme A biosyntheis cataylsed by *Escherichia coli* PPC Synthetase, *Structure*, **12**, 1977-1988.

8. Manoj, N., Strauss, E., Begley, T. P. and Ealick, S. E., (2003), Structure of human phosphopantothenoylcysteine synthetase at 2.3 Å resolution, *Structure*, **11**, 927-936.

9. Daugherty, M., Polanuyer, B., Farrell, M., Scholle, M., Lykidis, A., de Crécy-Lagard, V., and Osterman, A., (2002), Complete reconstitution of the human coenzyme A biosynthetic pathway via compariative genomics, *J. Biol. Chem.*, **277**, 21431-21439.

10. Evans, J. C., Trujillo, C., Wang, Z., Eoh, H., Ehrt, S., Schnappinger, D., Boshoff, H. I. M., Rhee, K. Y., Barry C. E. and Mizrahi, V., (2016), Validation of CoaBC as a bactericidal target in the Coenzyme A pathway of *Mycobacterium tuberculosis*, *ACS Infect. Dis.*, 958-968.

11. Patrone, J. D., Yao, J., Scott, N. E., and Dotson, G. D., (2009), Selective inhibitors of bacterial phosphopantothenoylcysteine synthetase, *J. Am. Chem. Soc.*, **131**, 16340-16341.

12. van der Westhuyzen, R., Hammons, J. C., Meier, J. L., Dahesh, S., Moolman, W. J., Pelly, S. C., Nizet, V., Burkart, M. D. and Strauss, E., (2012), The antibiotic CJ-15,801 is an antimetabolite that hijacks then inhibits CoA biosynthesis, *Chem. Biol.*, **19**, 559-571.

13. Hernande, H., and Robinson, C. V., (2007), Determing the stiochiometry and interactions of macromolecular assemblies from mass spectrometry, *Nat. Protocols*, **2**, 715-726.

14. Benesch, J. L. P. and Ruotolo, B. T., (2011), Mass spectrometry: An approach come-of-age for structural and dynamical biology, *Curr. Opin. Struct. Biol.*, **21**, 641-649.

15. Leney, A. C. and Heck, A. J. R. (2017), Native mass spectrometry: Whats in a name? *J. Am. Soc. Mass. Spectrom.*, **28**, 5-13.

16. Lanucara, F., Holman, S. W., Gray, C. J. and Eyers, C. E. (2014), The power of ion mobility-mass spectrometry for structural charachterization and the study of conformational dynamics, *Nat. Chem.*, **6**, 281-294.

17. Rajabi, K., Ashcroft A. E., and Radford, S. E., (2015), Mass spectrometric methods to analyze the structural organization on macromolecular complexes, *Methods*, **89**, 13-21.

18. Konijnenberg, A., Butterer, A. and Sobott, F. (2013), Native ion mobility-mass specctrometry and related methods in structural biology, *Biochim. Biophys. Acta*, **1834**, 1239-1256.

19. Boeri Erba, E., Barylyuk, K., Yang Y. and Zenobi, R. (2011), Quantifying protein-protein interactions within noncovalent comlexes using electrospray inization mass spectrometry, *Anal. Chem.*, **83**, 9251-9259.

20. Gabelica, V., Galic, N., Rosu, F., Houssier, C. and De Pauw, E., (2003), Influence of response factors on determining equilibrum association constants of non-covalent complexes by electrospray ionization mass spectrometry, *J. Mass Spectrom.*, **38**, 491-501.

21. Hu, W., Xie, J., Chau, H. W.and Si, B. C., (2015), Evaluation of parameter uncertanties in nonlinear regression using Microsoft Excel spreadsheet, *Environ. Syst. Res.*, **4**, 4.

22. D'Atri, V., Porrini, M., Rosu, F. and Gabelica, V., (2015), Linking molecular models with ion mobility experiments. Illustration with a rigid nucleic acid scructure. *J. Mass Spectrom.*, **50**, 711-726.

23. Ruotolo, B. T., Benesch, J. L. P., Sandercock, A. M., Hyung, S.-J. and Robinson, C. V. (2008), Ion mobility-mass spectrometry analysis of large protein complexes, *Nat. Protocols*, **3**, 1139-1152.

24. Allison, T. M., Reading, E., Liko, I., Baldwin, A. J., Laganowsky, A. and Robinson, C. V. (2015), Quantifying the stabalizing effects of protein-ligand interactions in the gas phase, *Nat. Commun.*, **6**, 8551.

25. Bush, M. F., Hall, Z., Giles, K., Hoyes, J., Robinson, C. V. and Ruotolo, B. T., (2010) Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology, *Anal. Chem.*, **82**, 9557-9565.

26. Strauss, E. and Begley, T. P., (2001), Mechanistic studies on phosphopantothenoylcysteine decarboxylase, *J. Am. Chem. Soc.*, **123**, 6449-6450.

27. Gavriilidou, A. F. M., Gülbakan, B. and Zenobi, R. (2015), Influence of ammonium acetate concentration on receptor-ligand binding affinities measured by native nano ESI-MS: A systematic study, *Anal. Chem.*, **87**, 10378-10384.

28. Seetoh, W.-G., Chan, D. S.-H., Matak-Vinković, D. and Abell, C. (2016), Mass spectrometry reveals protein kinase CK2 higher order oligomerisation via the circular and linear assembly, *ACS Chem. Biol.*, **11**, 1511-1517.

29. Voets, I. K., Cruz, W. A., Moitzi, C., Lindner, P., Arêas, E. P. G. and Schurtenberger, P. (2010), DMSO-induced denaturation of hen egg white lysozyme, *J. Phys. Chem. B*, **114**, 11875-11883.

30. Sterling, H. J., Prell, J. S., Cassou, C. A. and Williams, E. R. (2011), Protein conformation and supercharging with DMSO from aqueous solution, *J. Am. Soc. Mass. Spectrom.*, **22**, 1178-1186.

31. Chan, D. S.-H., Matak-Vinković, D., Coyne, A. G. and Abell, C. (2016), Effect of DMSO an protein structure and interactions accessed by collision-induced dissociation and unfolding, *ChemistrySelect*, **1**, 5686-5690.

32. Eschweiler, J. D., J. N. Rabuck-Gibbons, J. N., Y. Tian and B. T. Ruotolo, (2015), CIUSuite:A quantitatice analysis package for collision induced unfolding measurements of gas-phase protein ions, *Anal. Chem.*, **87**, 11516-11522.

33. Rabuck, J. N., Hyung, S.-J., Ko, K. S., Fox, C. C., Soellner, M. B. and Ruotolo, B. T. (2013), Activation state-selective kinase inhibitor assay based on ion mobility-mass spectrometry, *Anal. Chem.*, **85**, 6995-7002.

34. Bornschein, R. E., Niu, S., Eschweiler, J. and Ruotolo, B. T., (2016), Ion mobility-mass spectrometry reveals highly-compact intermediates in the collision induced dissociation of charge-reduced protein complexes, *J. Am. Soc. Mass. Spectrom.*, **27**, 41-49.

35. Daubenfeld, T., Bouin, A.-P. and van der Rest, G., (2006), A deconvolution method for the separation of specific versus nonspecific interactions in noncovalent protein-ligand complexes analyzed by ESI-FT-ICR mass spectrometry, *J. Am. Soc. Mass. Spectrom.*, **17**, 1239-1248.

36. Dyachenko, A., Gruber, R., Shimon, L., Horovitz, A. and Sharon, M. (2013), Allosrteric mechanisms can be distinguished using structural mass spectrometry, *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 7235-7239.

37. Wang, W., Kitova E. N. and Klassen, J. S. (2005), Nonspecific protein-carbohydrate complexes produced by nanoelectrospray ionization. Factors influencing their formation and stability, *Anal. Chem.*, **77**, 3060-3071.

38. Guan, S., Trnka, M. J., Bushnell, D. A., Robinson, P. J. J., Gestwicki, J. E. and Burlingame, A. L. (2015), Deconvolution method for the specific and nonspecific binding of ligand to multiprotin complex by native mass spectrometry, *Anal. Chem.*, **87**, 8541-8546.

39. Shimon, L., Sharon, M. and Horovitz, A. (2010), A method for removing effects of non-specific binding on the distribution of binding stiochiometries: application to mass spectroscopy data, *Biophys. J.*, **99**, 1645-1649.

40. Poulsen, S.-A. (2013), Fragment screening by native mass spectrometry, *Aust. J. Chem.*, **66**, 1495-1501.

41. Pedro, L. and Quinn, J. R. (2016), Native mass spectrometry in fragment-based drug discovery, *Molecules*, **21,** 984.

42. Chan, D. S.-H., Whitehouse, A. J., Coyne, A. G. and Abell, C. (2017), Mass spectrometry for fragment screening, *Essays Biochem.*, **61**, 465-473.

(a)

Ppan
CTP → PPi
PPCS

4'-phosphopantothenoyl cytidylate
cysteine → CMP
PPCS

PPC

(b)

phosphodiester isostere

sulfamate isostere

phospho-CJ-15,801

(a) apo

M11+

M10+

M9+

D17+

D16+

D15+

D14+

(b) 50 µM CTP

(c) 50 µM ATP

m/z

1600   2000   2400   2800   3200   3600   4000   4400   4800
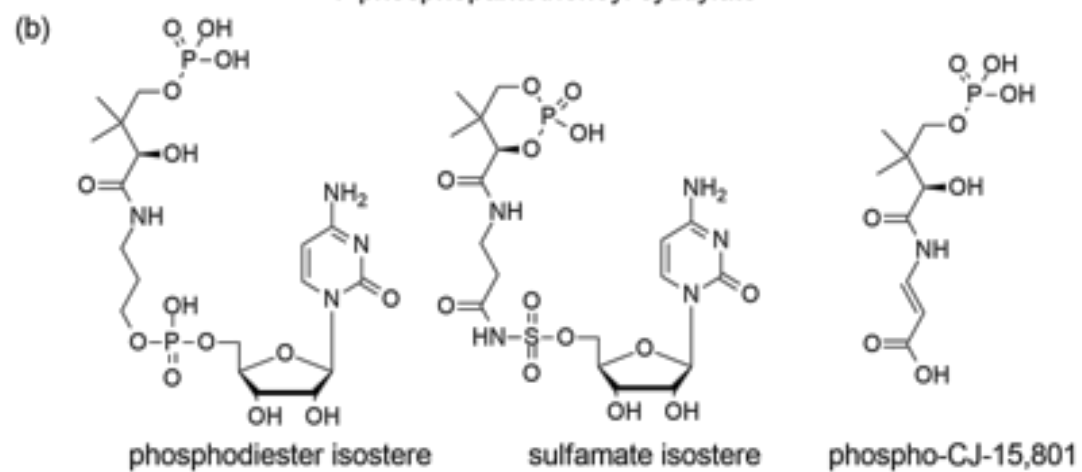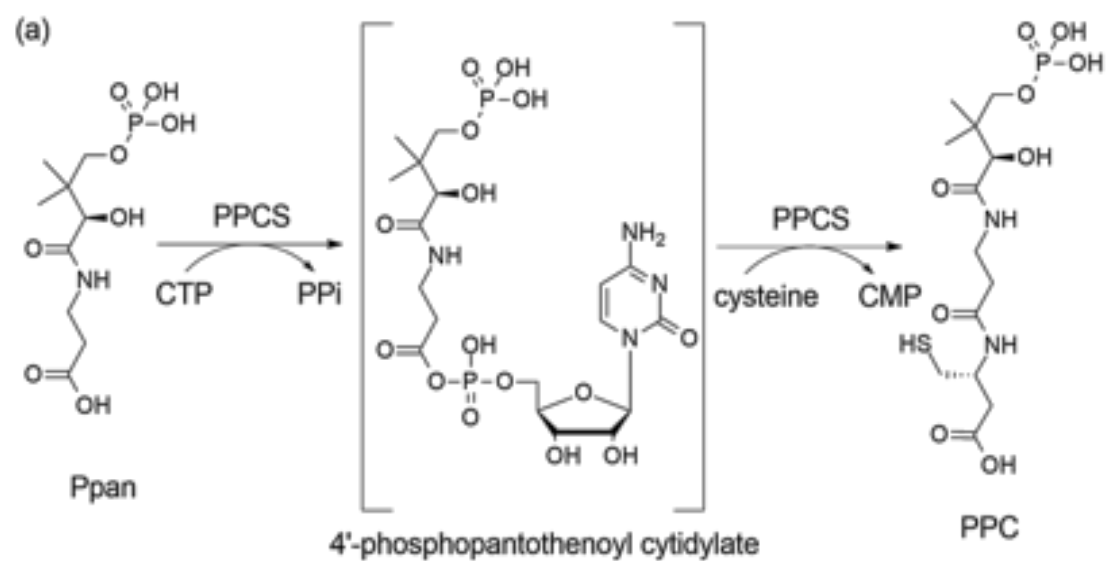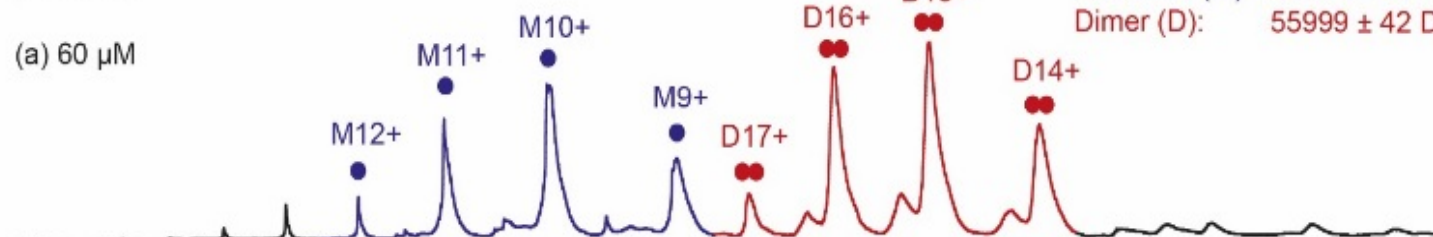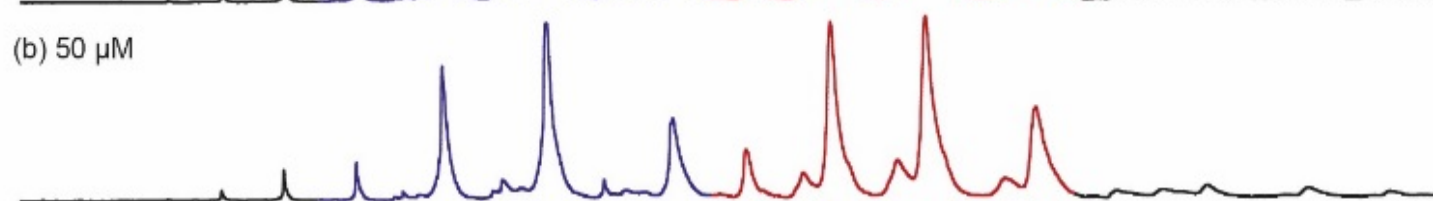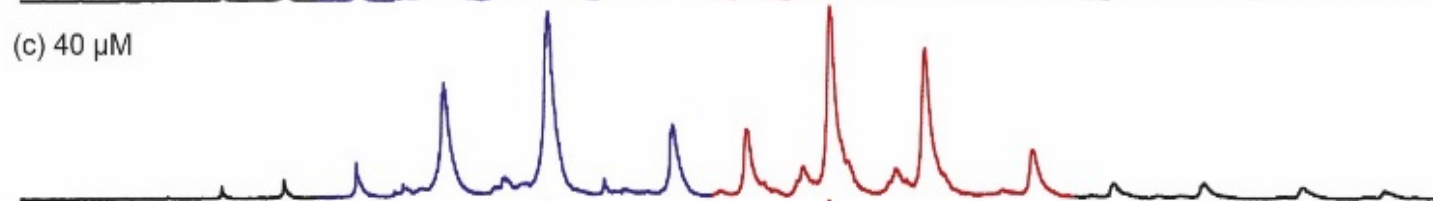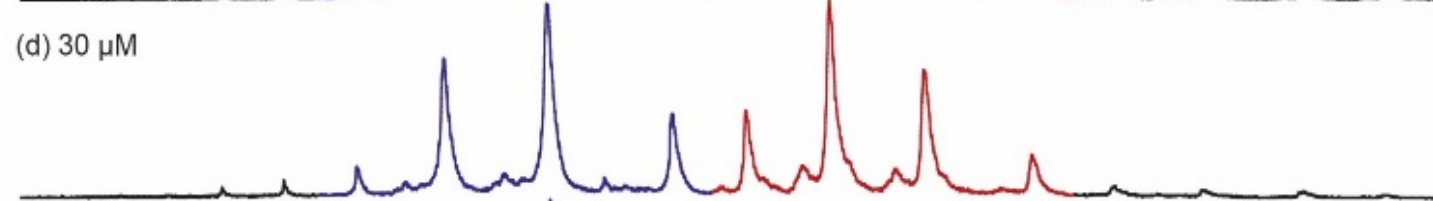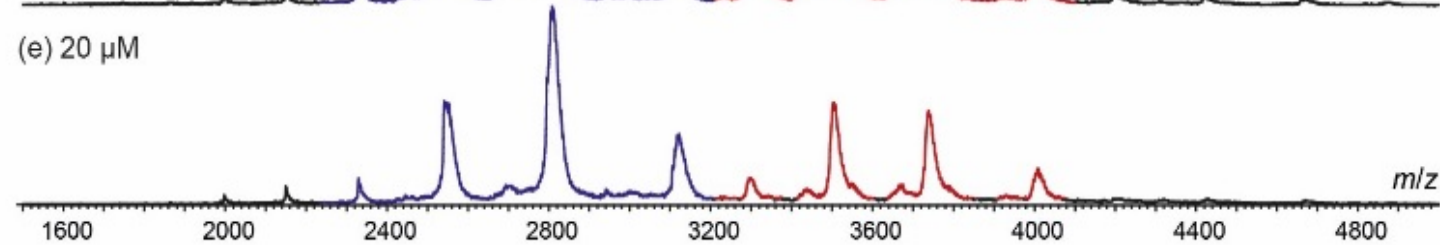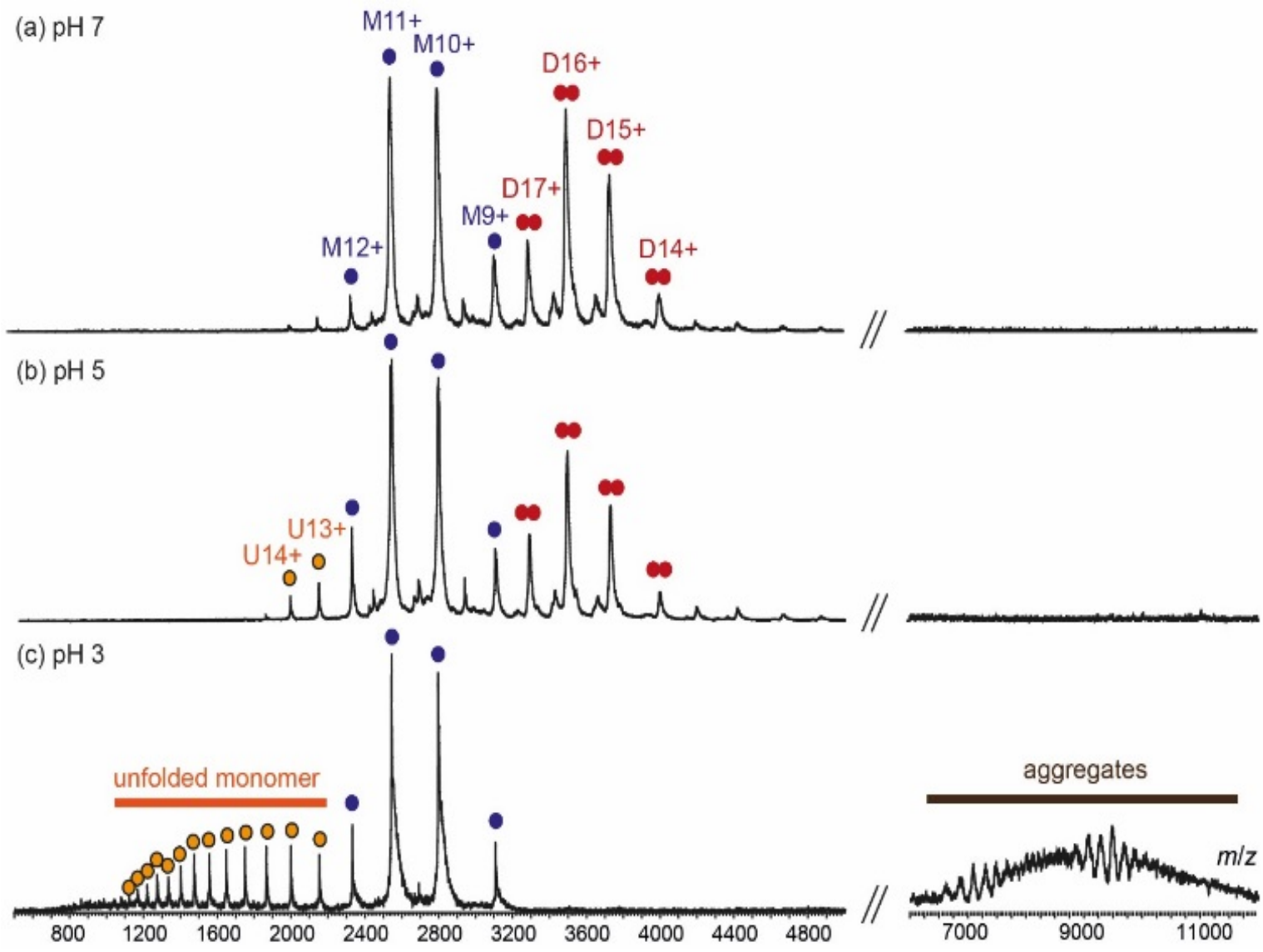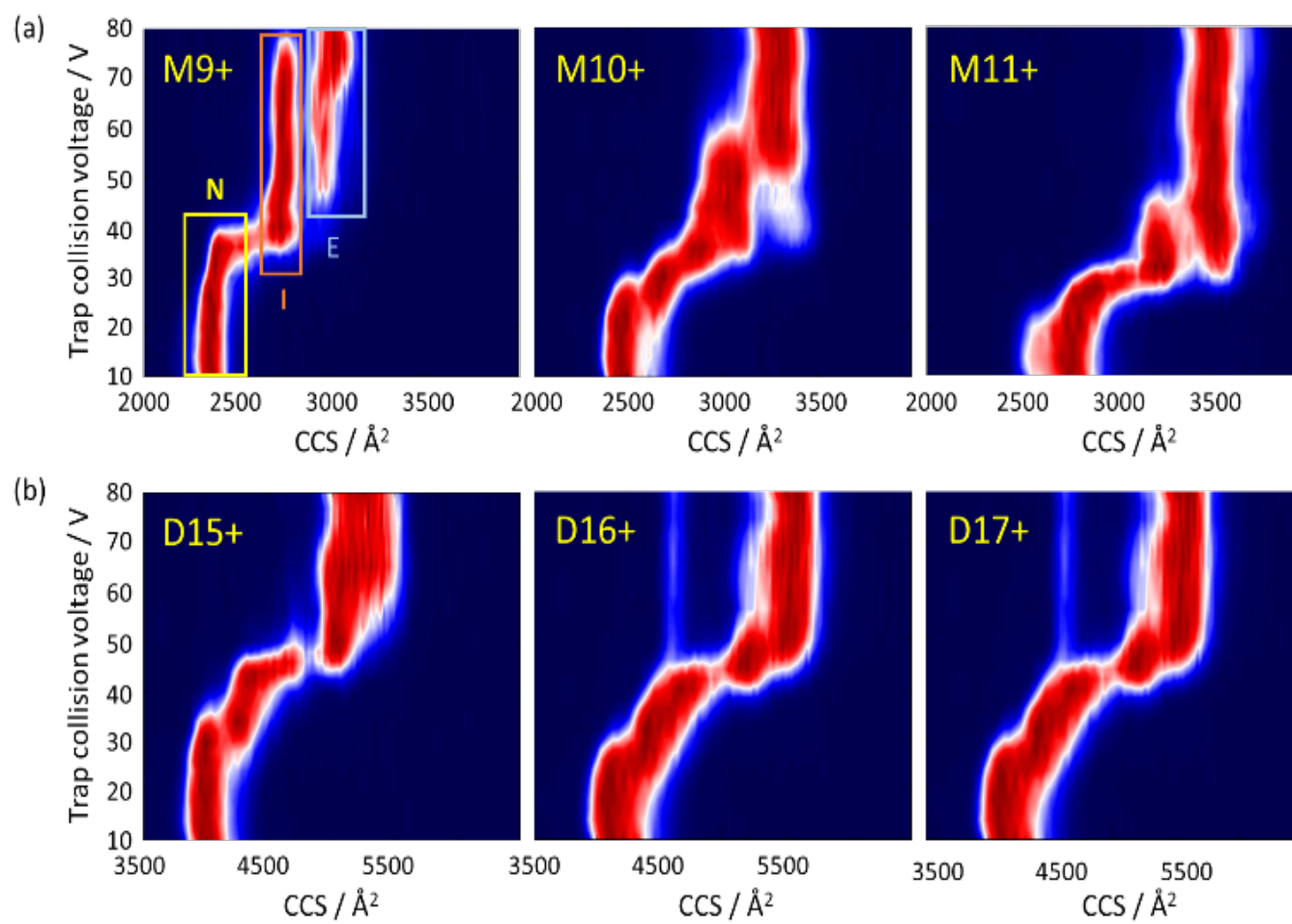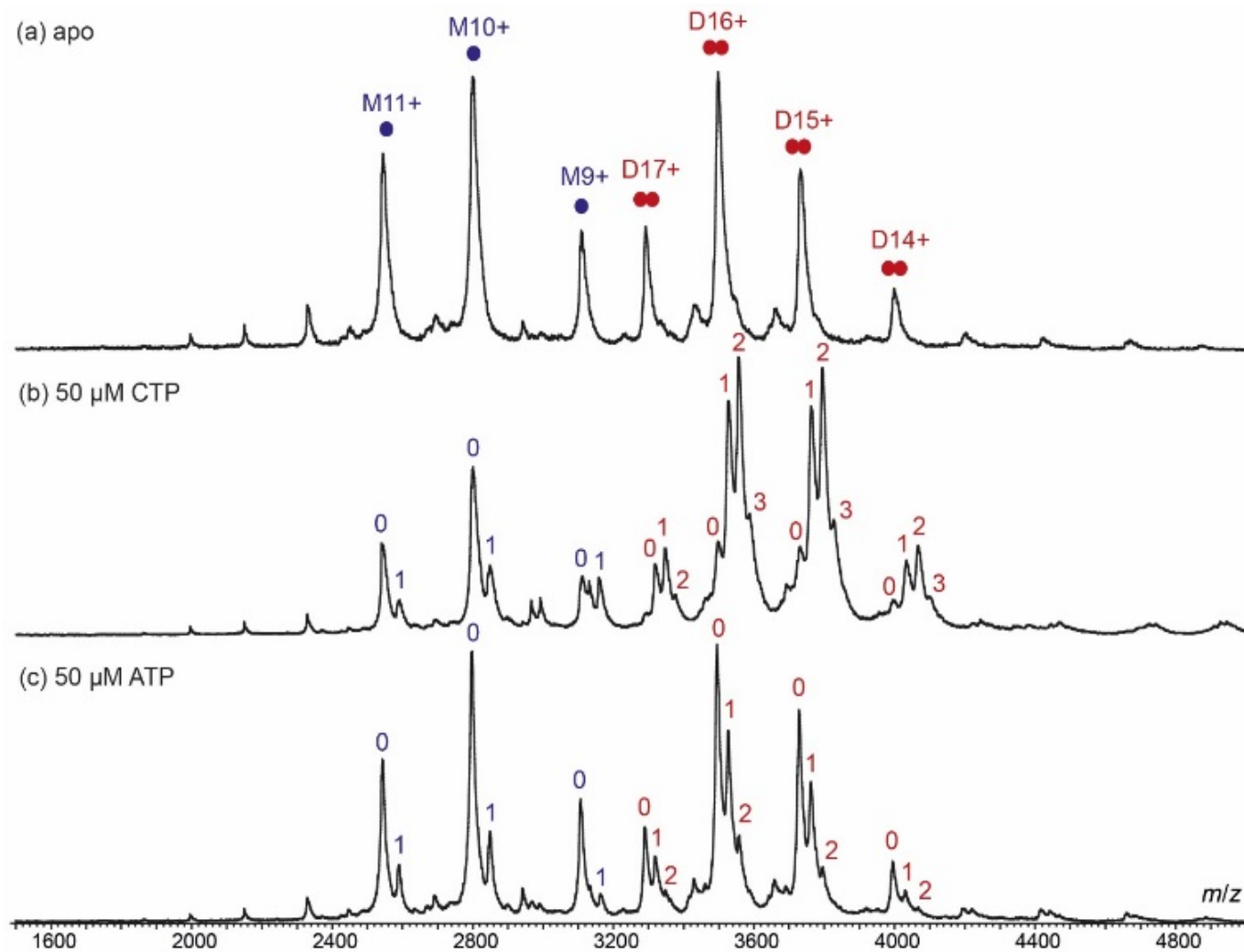
Initial fragment hit

1
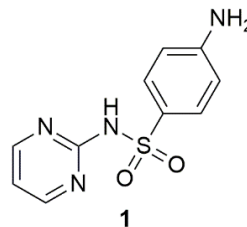
Phenylsulfonamide series

2 3 4 5

6 7 8 9
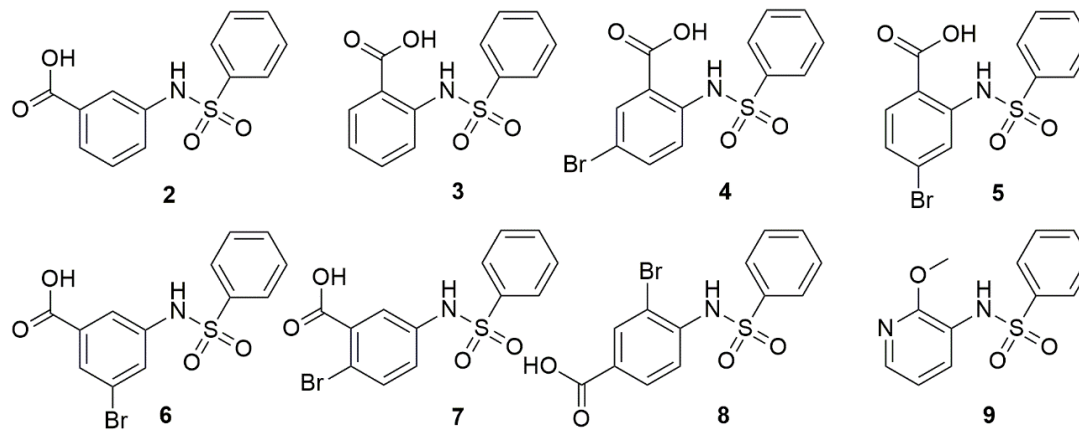
Pyridinesulfonamide

10
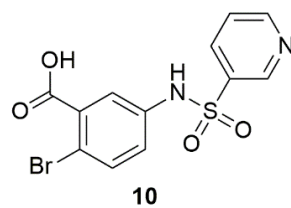
Naphthalenesulfonamide

11
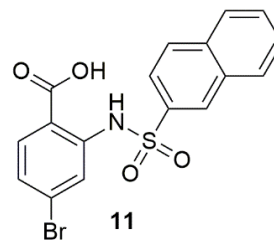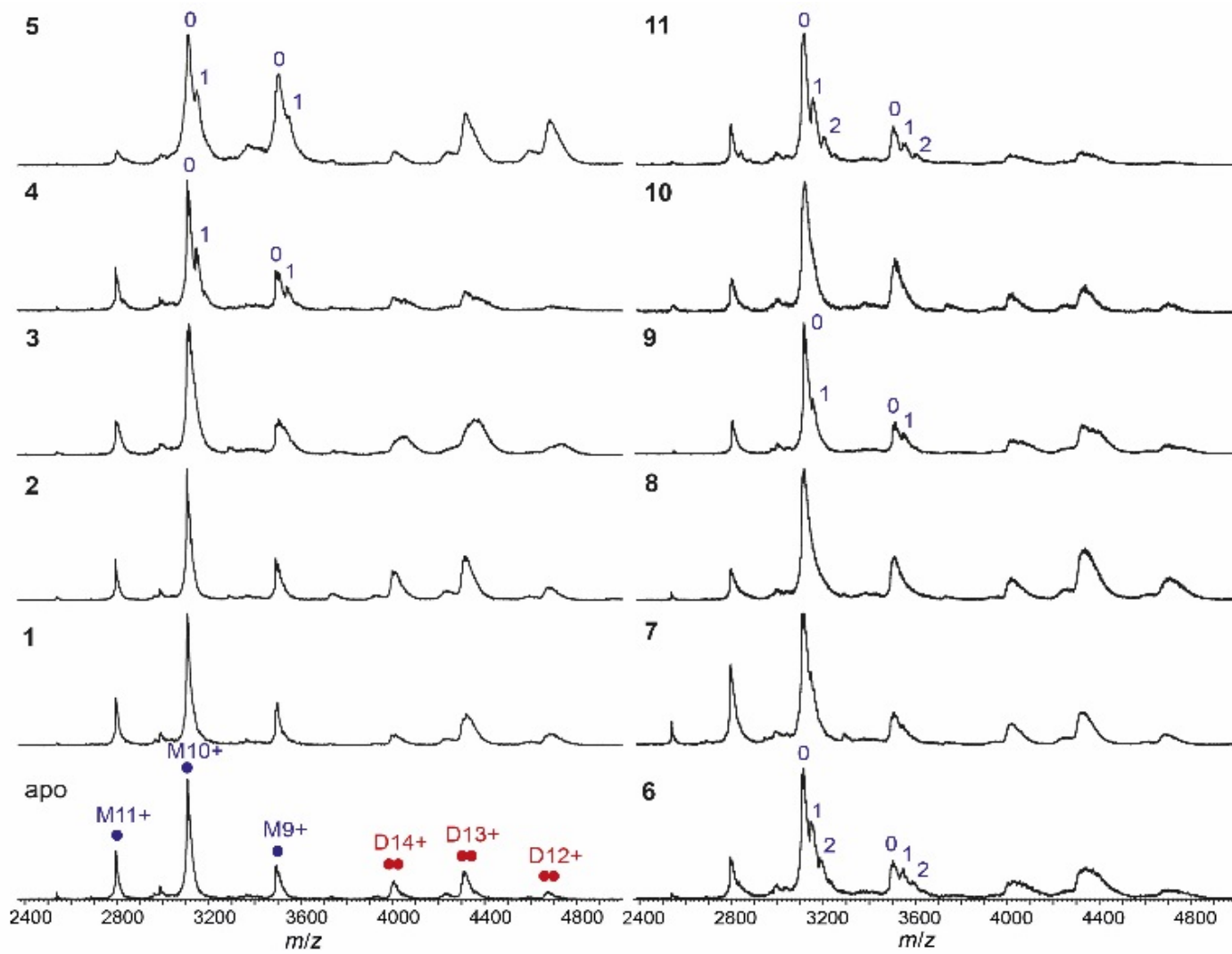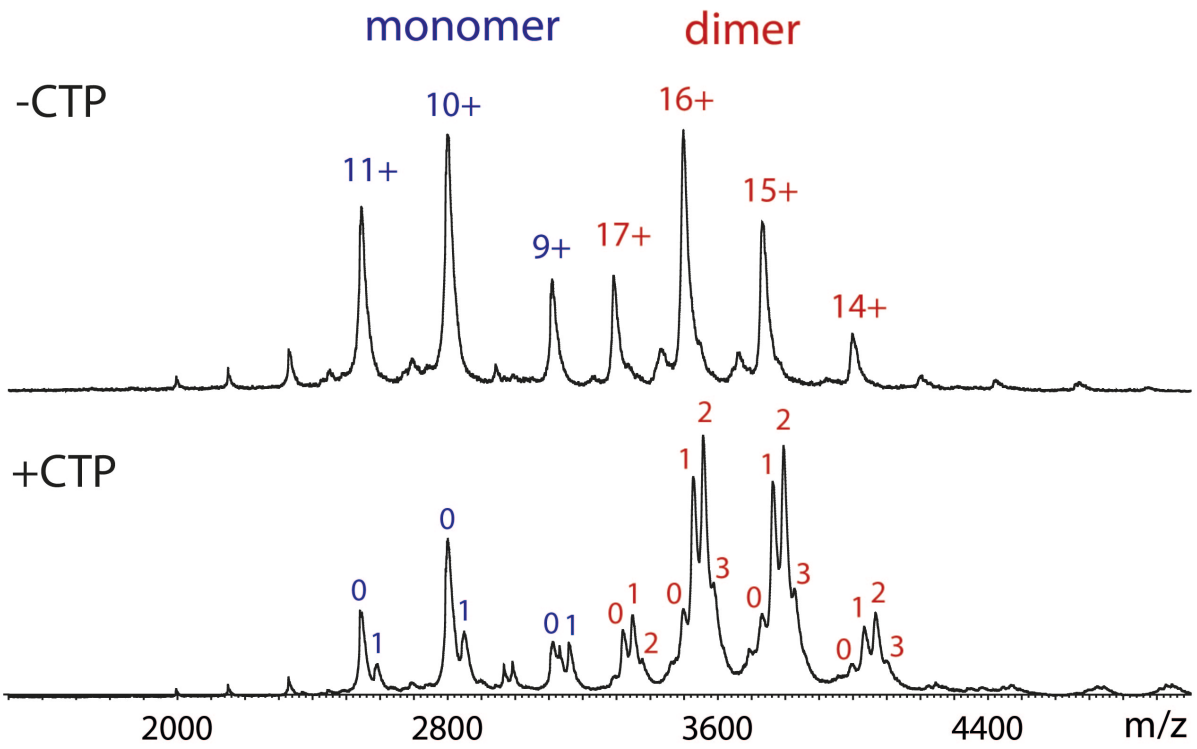
Escherichia coli phosphopantothenoylcysteine synthetase (CoaB), an emerging antibacterial drug target, has been studied by native ion mobility mass spectrometry.

Structural Insights into *Escherichia coli* Phosphopantothenoylcysteine Synthetase by Native Ion Mobility-Mass Spectrometry

Daniel Shiu-Hin Chan,[a] Jeannine Hess,[a] Elen Shaw,[a] Christina Spry,[a,] Robert Starley,[a] Claudio Dagostin,[a] Marcio V. B. Dias,[b,] Ramesh Kale,[a,] Vitor Mendes,[b] Tom L. Blundell,[b] Anthony G. Coyne, [*a] and Chris Abell[*a]

[a.]   Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom.

[b.]   Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, United Kingdom.

*   Corresponding authors: Email: ca26@cam.ac.uk and agc40@cam.ac.uk

Table of content

General Methods

Synthetic chemistry methods: All chemicals were of reagent grade quality or better, obtained from commercial suppliers and used without further purification. 4'-Phosphopantothenate was synthesised following a literature procedure (*Eur. J. Org. Chem.*, **2010**, 1728). Unless otherwise stated, reactions were conducted under positive pressure of a nitrogen atmosphere. Anhydrous solvents were purchased directly from commercial sources and used without further purification. Analytical thin layer chromatography (TLC) was performed using Merck glass-backed silica (Kieselgel 60 F254 0.25 mm) plates. Compounds were visualized using shortwave (254 nm) ultraviolet light. Flash column chromatography was performed using an Isolera Spektra One/Four purification system and the appropriately sized Biotage SNAP column containing KP-silica gel (50 μm). Solvents are reported as volume/volume eluent mixtures. Reactions were monitored by TLC and liquid chromatography-mass spectrometry (LC-MS) to determine consumption of starting materials. Nuclear magnetic resonance (NMR) spectra were recorded at 298 K unless otherwise stated, using either a Bruker 400 MHz AVANCE III HD Smart Probe, 400 MHz QNP cryoprobe, or 500 MHz DCH cryoprobe spectrometer. All spectra were recorded in the deuterated solvent indicated. Data are reported as chemical shift in parts per million (δ ppm) relative to the residual protonated solvent resonance peak. The relative integral, multiplicity and coupling constants (*J* Hz) of peaks has been provided where possible. LC-MS was carried out using an AQUITY UPLC H-class system (Waters, Manchester UK). Samples were run using a gradient of water (1−5%) (+0.1% formic acid) in acetonitrile over a period of 4 min. High resolution mass spectra (HRMS) were recorded using a Micromass Quadrapole-Time-of-Flight (Q-Tof) spectrophotometer or an Orbitrap LC-MS spectrometer attached to a Dionex Ultimate 3000 HPLC. The mass to charge ratio (*m/z*) of the molecular ion and difference from calculated mass (δ ppm) are quoted.

Overexpression and purification of *E. coli* CoaB (*Ec*CoaB). A pET-28b plasmid carrying the coding sequence for the CoaB domain of *E. coli* CoaBC (*Ec*CoaBC) was used. The plasmid encodes N and C-terminally hexa-histidine-tagged *Ec*CoaB (MGSSHHHHHHSSGLVPRGSHM-Pro[182]-Arg[406]-LEHHHHHH, CoaBC residues indicated). *E. coli* BL21(DE3) cells transformed with the construct were grown in LB supplemented with 30 μg/mL kanamycin until an OD_{600} of approximately 0.6 was reached. Immediately thereafter overexpression was induced by the addition of 1 mM isopropyl-β-D-

1-thiogalactopyranoside (IPTG). After overnight incubation at 37 °C, the bacteria were harvested by centrifugation (4000 × g, 15 min) and the cell pellet was stored at –20 °C. The harvested cells were lysed in 20 mM Tris, 500 mM NaCl, pH 8.0 containing one EDTA-free cOmplete protease inhibitor cocktail tablet (Roche) per 50 mL lysate, using an EmulsiFlex-C5 High Pressure Homogenizer (Avestin). The cells were passed through the pressure cell three times at a pressure of 40 000 psi, after which the lysate was centrifuged at 35 000 × g for 30 min. For nickel-affinity chromatography, the clarified lysate was applied to a 5 mL HisTrap FF column (GE Healthcare) and, following washing with lysis buffer, the His-tagged protein was eluted using a linear gradient of imidazole (0–0.5 M). Elution was monitored at A$_{280}$ and by SDS-PAGE analysis. Fractions containing EcCoaB were pooled and concentrated using a Vivaspin 20 centrifugal concentrator with a 10 kDa molecular weight cut-off before being subjected to gel filtration using a Superdex75 26/60 column (GE Healthcare) previously equilibrated with 20 mM Tris, 150 mM NaCl, pH 8.0. Gel filtration was performed at a flow rate of 2.5 mL/min and elution was again monitored at A$_{280}$ and by SDS-PAGE analysis. Fractions containing purified His-tagged EcCoaB were pooled and concentrated as before, and aliquots of the concentrated protein were flash-frozen in liquid N$_2$ and stored at –80 °C. The identify and purity of EcCoaB was assessed by denaturing LC-MS (Figure S1), which showed a mass for the protein that was identical to the theoretical mass of the protein without the N-terminal methionine (27942 Da).

Overexpression and purification of EcCoaBC. EcCoaBC was overexpressed in *E. coli* BL21(DE3) cells from a codon-optimized pET-28a construct. Transformed cells were grown at 37 °C in LB culture medium containing 50 μg/mL kanamycin to an OD$_{600}$ of approximately 0.5, then expression was induced with 500 μM IPTG. The culture was incubated for a further 20 h at 20 °C, and the cells were pelleted by centrifugation (4000 × g, 20 min). Cell pellets are frozen and stored at -80 °C. Cells were resuspended in lysis buffer (50 mM Hepes, pH 7.4, 1 M NaCl, 25 mM imidazole, 1 Pierce Protease Inhibitor tablet), and lyzed by sonication for 3.5 min (10 s on/20 s off, amplitude 50%). Insoluble material was removed from the lysate by centrifugation at 30 000 × g for 30 min at 4 °C. The supernatant was then loaded onto a 5 mL HisTrap FF column (GE Healthcare), which had been pre-equilibrated with lysis buffer. The column was further washed with lysis buffer, and then EcCoaBC was eluted from the column by increasing the imidazole concentration to 500 mM. The eluate was concentrated to ≤10 mL and injected onto a Superdex-200 Hiload 16/60 FPLC column (GE Healthcare) previously equilibrated with 50 mM Tris (pH 8.0), 250 mM

NaCl, and 5 mM 2-mercaptoethanol. The purified protein was concentrated and aliquots were flash-frozen in liquid N_2 and stored at −80 °C.

Phosphopantothenoylcysteine synthetase assay. Phosphopantothenoylcysteine synthetase activity was monitored by the method of Webb (*Proc. Natl. Acad. Sci. USA*, 1992, **89**, 4884-4887) using the commercially available EnzChek pyrophosphate assay kit (E-6645) (Life Technologies). The assay couples production of pyrophosphate to the phosphorolysis of 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG), which can be monitored by measuring absorbance at 360 nm. Reactions were performed at 25 ˚C in the wells of a 96-well plate (Greiner), in a final volume of 150 μL. The final composition of reactions was 16 nM *Ec*CoaBC, 125 μM CTP, 125 μM 4'-phosphopantothenate, 500 μM L-cysteine, 0.03 U/mL inorganic pyrophosphatase, 1 U/mL purine nucleoside phosphorylase, 200 μM MESG, 100 mM Tris, pH 7.6, 1 mM $MgCl_2$, 1 mM TCEP, and 1 mM of test compound (initially dissolved to 50 mM in DMSO) or an equivalent concentration of DMSO (2%, *v/v*). Initially, 100 mM Tris, pH 7.6, 1 mM $MgCl_2$, 1 mM TCEP with CTP, 4'-phosphopantothenate, L-cysteine, MESG and the test compound/DMSO, each at twice the final concentration, in a final volume of 75 μL, was added in triplicate to the wells of a 96-well plate, and the plate was pre-incubated at 25 ˚C for 10 min. A 100 mM Tris, pH 7.6, 1 mM $MgCl_2$, 1 mM TCEP solution containing *Ec*CoaBC, inorganic pyrophosphatase and purine nucleoside phosphorylase, each at twice the final concentration, was also prepared and similarly pre-incubated at 25 ˚C for 10 min. Reactions were subsequently initiated by the addition of 75 μL of the enzyme solution to each well, and enzyme activity was immediately monitored by measuring absorbance at 360 nm at 20 s intervals over 30 min in a ClarioStar microplate reader (BMG Labtech). Assays were performed in triplicate, and included wells without 4'-phosphopantothenate, which served as "no reaction" controls, as well as wells containing 2% (*v/v*) DMSO instead of test compound, which served as "uninhibited" controls. Initial reaction rates were determined from the linear portion of the time courses. To convert the reaction rates measured in the presence of each test compound to a percentage inhibition, the average rate determined for the "no reaction" wells was first subtracted from all initial rates. Thereafter the background-corrected rates obtained for the wells containing test compound were subtracted from the average of the background-corrected uninhibited controls, and then the resulting difference in rate divided by the average of the uninhibited control.
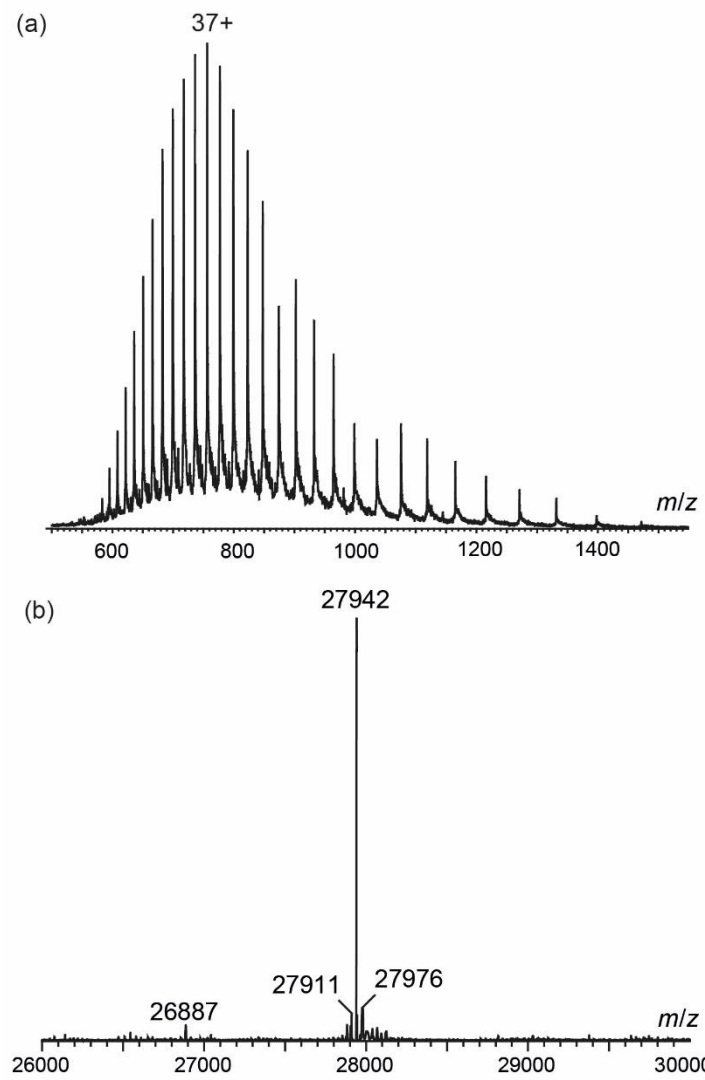
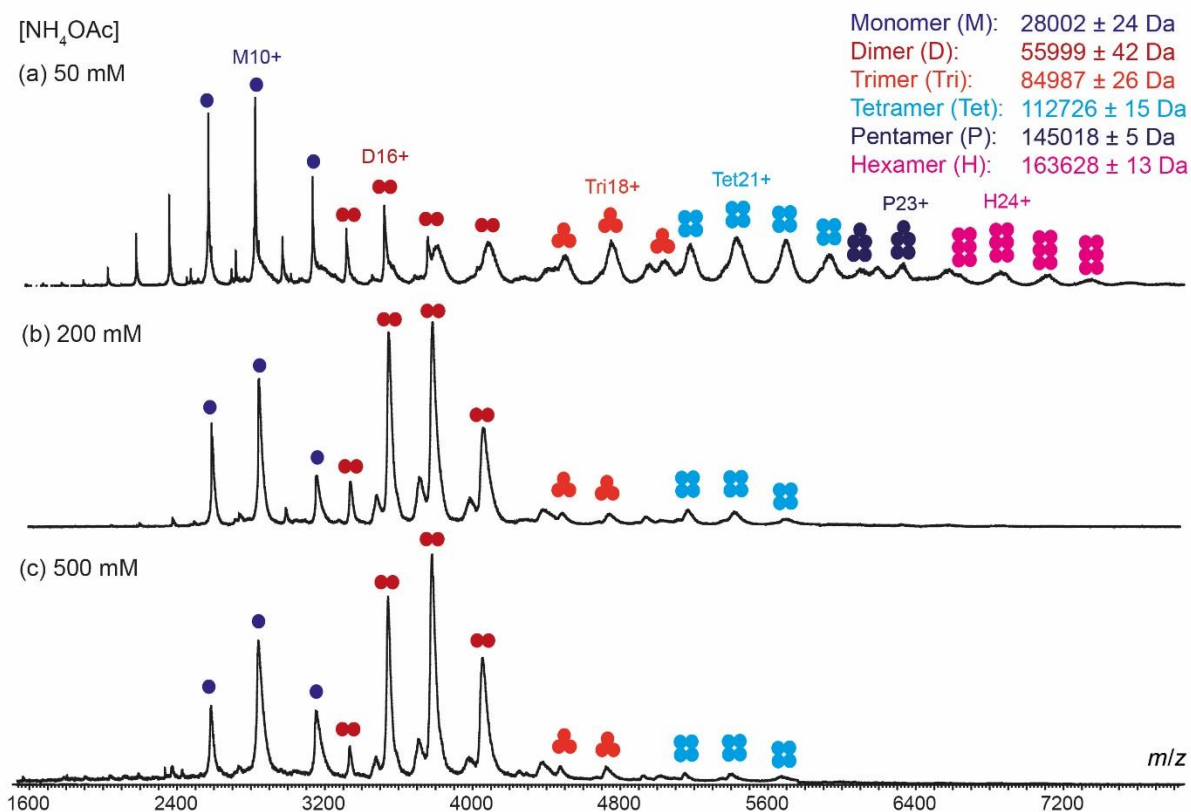**Figure S1.** Denaturing LC-MS spectrum (a) and deconvoluted LC-MS spectrum (b) of *Ec*CoaB.

Figure S2. The oligomerization of *Ec*CoaB is NH₄OAc concentration dependent. Native mass spectra of *Ec*CoaB in (a) 50 mM, (b) 200 mM and (c) 500 mM NH₄OAc buffer, with monomer (M), dimer (D), trimer (Tri), tetramer (Tet), pentamer (P) and hexamer (H) charge states indicated.
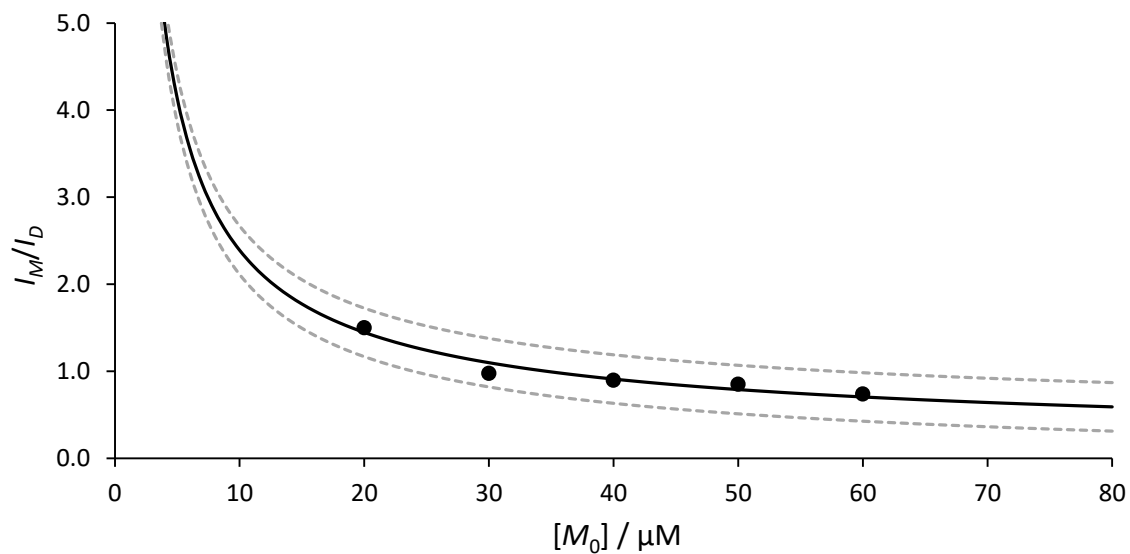
Figure S3. Plot of the ratio (I_M/I_D) of the monomer (I_M) and dimer (I_D) peak intensities (filled circles) against the total concentration of EcCoaB monomer (M_0) (see Fig. 2 for titration experiment). The solid line represents the fitted curve obtained by non-linear regression of the data using equation 5, from which R and K_a were determined to be 1.8 and $3.5 \times 10^4 \ M^{-1}$, respectively. The dashed lines represent the 95% confidence interval of the regression curve.
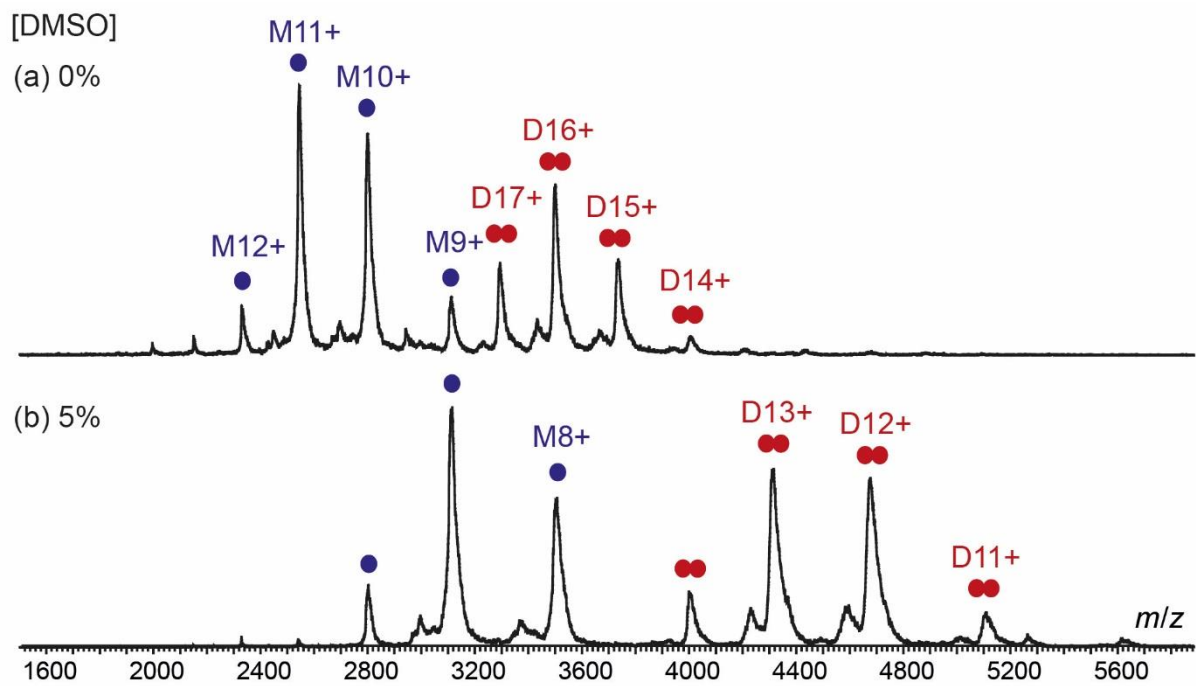
Figure S4. DMSO shifts charge state distributions of *Ec*CoaB to lower charge states. Native mass spectra of *Ec*CoaB (30 μM) in 200 mM NH₄OAc in the presence of (a) 0% or (b) 5% (*v/v*) DMSO.
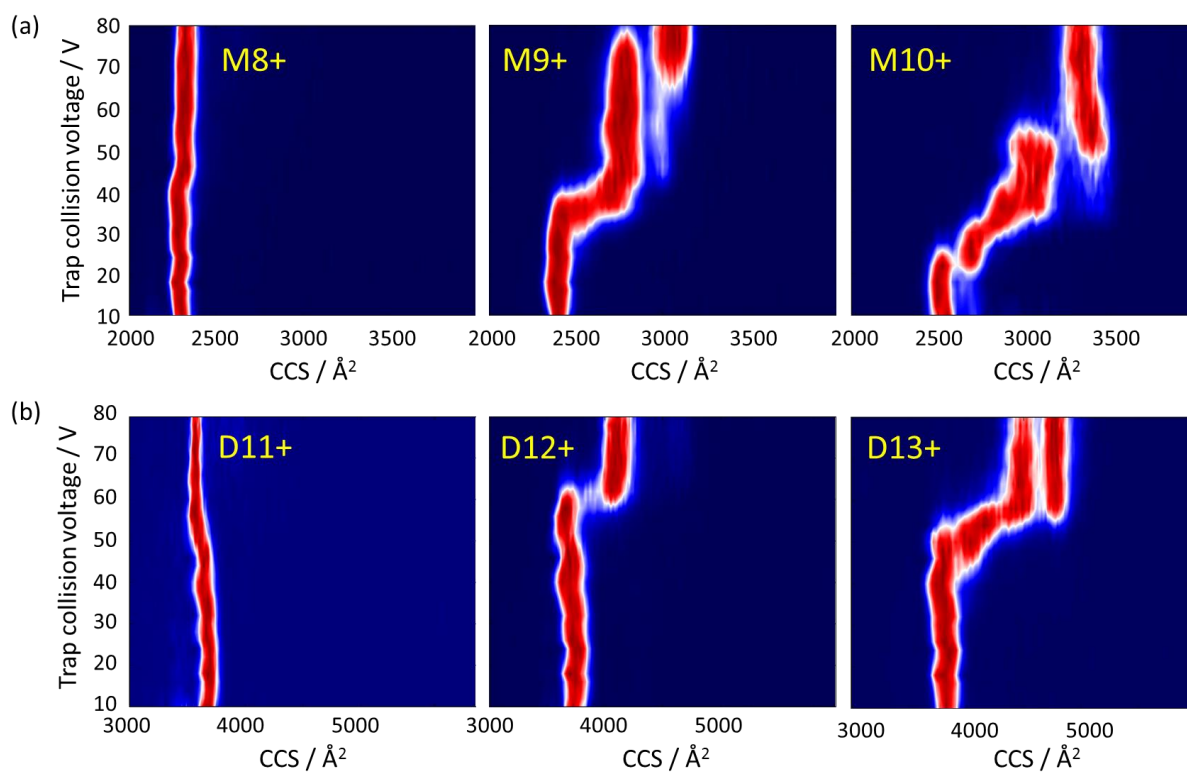
Figure S5. Collision-induced unfolding (CIU) ion mobility-mass spectrometry (IM-MS) of *Ec*CoaB in the presence of DMSO. CIU plots of indicated charge states of the *Ec*CoaB (a) monomer (M) (8+ to 10+) and (b) dimer (D) (11+ to 13+) with 5% (*v/v*) DMSO.
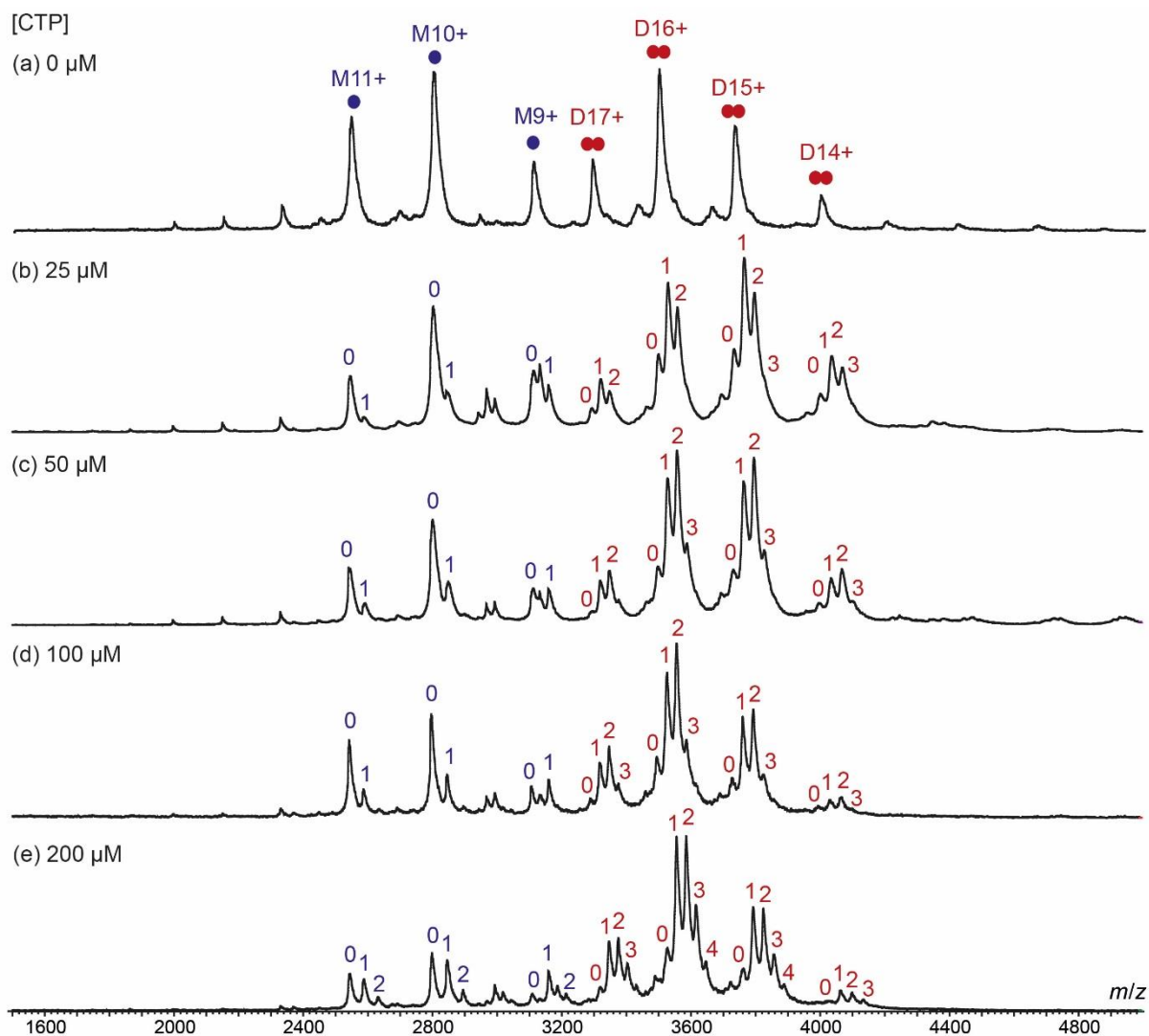
Figure S6. Titration of CTP into *Ec*CoaB reveals non-specific binding at higher CTP concentrations. Native mass spectra of *Ec*CoaB (30 μM) in the presence of (a) 0 μM, (b) 30 μM, (c) 50 μM, (d) 100 μM and (e) 200 μM of CTP. In (a), the monomer (M) and dimer (D) charge states are indicated. In (b)–(e), the number of CTP molecules bound to each monomer or dimer species is indicated.
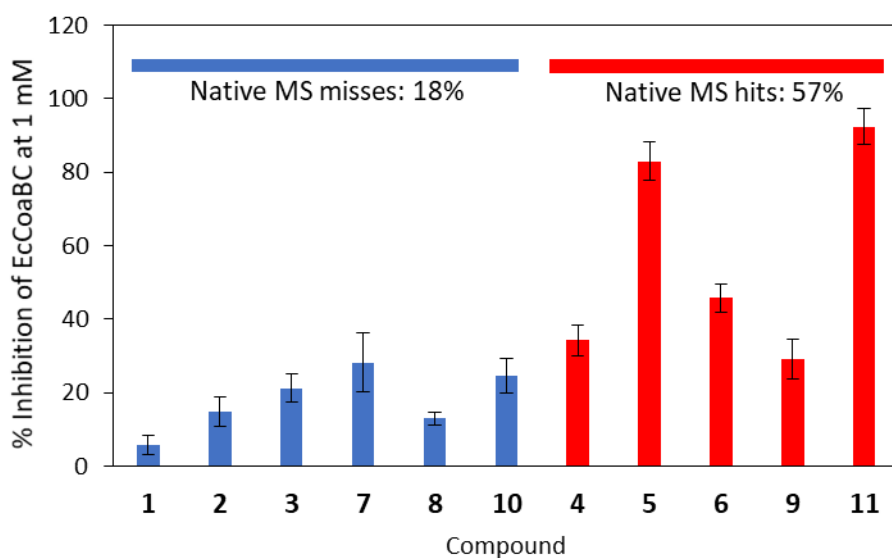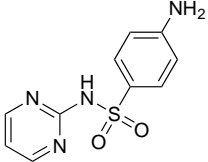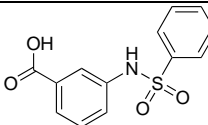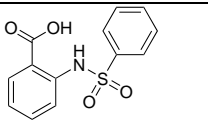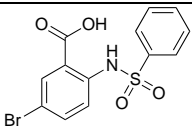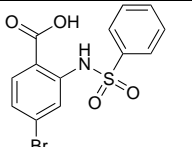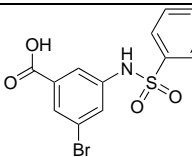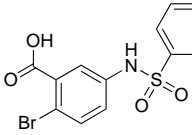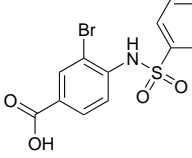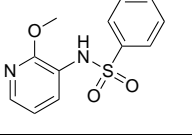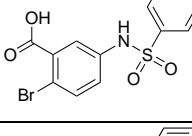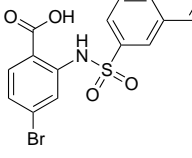
Figure S7. Inhibition of the phosphopantothenoylcysteine synthetase activity of *Ec*CoaBC by compounds **1–11** at 1 mM as measured by an enzymatic assay. *Ec*CoaBC, CTP, phosphopantothenate and L-cysteine concentrations of 16 nM, 125 μM, 125 μM and 500 μM, respectively. Native MS hits (**4**, **5**, **6**, **9** and **11**) were observed to inhibit enzymatic activity by an average of 57%. Native MS misses (**1**, **2**, **3**, **7**, **8** and **10**) inhibited enzymatic activity by an average of 18%. The errors bars represent the standard deviation of three experiments.

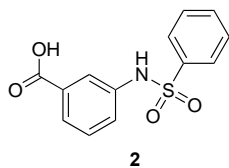**Table S1.** *Ec*CoaBC inhibition activity of compounds (**1-11**) at 1 mM as measured by an enzymatic assay.

Compound	Structure	% inhibition of *Ec*CoaBC at 1 mM	Native MS hit
1		6 ± 3	-
2		15 ± 4	-
3		21 ± 4	-

| 4		34 ± 4	Yes
5		83 ± 5	Yes
6		46 ± 4	Yes
7		28 ± 8	-
8		13 ± 2	-
9		29 ± 5	Yes
10		25 ± 5	-
11		92 ± 5	Yes

Synthesis of compounds

3-(Phenylsulfonamido)benzoic acid (2)

2

3-Aminobenzoic acid (0.10 g, 0.73 mmol) was suspended in H_2O (7.5 mL) and benzenesulfonyl chloride (0.15 g, 0.87 mmol) was added to the suspension with vigorous stirring. 1,4-Dioxane (2 mL) was added and the reaction mixture was heated a 37 °C for 2 h, and then allowed to cool to room temperature. The aqueous layer was extracted with EtOAc and the combined organic layers were dried with $MgSO_4$, filtered, and the solvent was evaporated to afford the crude product, which was purified on a flash silica column using an Isolera (Biotage) (SNAP KP-Sil 12 g column, 15 mL/min, A = CH_2Cl_2; B = MeOH, 0%B 5CV, 0%B-3%B 4CV, 3%-8%B 7.5CV, 8%B 4.5CV) to give the desired product 3-(phenylsulfonamido)benzoic acid as a colorless powder. Yield: 53% (0.11 g, 0.39 mmol).

^{1}H NMR (d$_6$-acetone, 400 MHz) δ 9.22 (s, 1H), 7.91 (t, J = 3.6 Hz, 1H), 7.83-7.81 (m, 2H), 7.73 (dt, J = 7.6 Hz, J = 2.6 Hz, 1H), 7.63-7.59 (m, 1H), 7.55-7.51 (m, 2H), 7.50-7.47 (m, 1H), 7.42-7.38 (m, 1H). ^{13}C NMR (d$_6$-acetone, 100 MHz) δ 167.0, 140.6, 139.1, 133.8, 132.5, 130.3, 130.0, 127.9, 126.4, 125.7, 122.4. LC-MS (ESI$^-$) m/z 276.1 [M-H]$^-$, retention time 1.61 min (100%). HRMS (ESI$^-$) m/z (calcd $C_{13}H_{10}NO_4S$ [M-H]$^-$ = 276.0336), obsd. 276.0335. IR (neat, cm^{-1}): 3254, 1682, 1588, 1301, 1267, 1158.

2-(Phenylsulfonamido)benzoic acid (3)

3
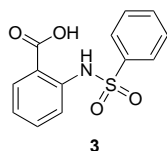
2-Aminobenzoic acid (0.10 g, 0.73 mmol) was suspended in H_2O (7.5 mL) and benzenesulfonyl chloride (0.15 g, 0.87 mmol) was added to the suspension with vigorous stirring. 1,4-Dioxane (2 mL) was added and the reaction mixture was heated at 37 °C for 2 h, and then allowed to cool to room temperature. The aqueous layer was extracted with EtOAc and the combined organic layers were dried with $MgSO_4$, filtered, and the solvent was evaporated to afford the crude product, which was purified on a reverse phase flash silica

column using an Isolera (Biotage) (SNAP C18 12 g column, 15 mL/min, A = H_2O with 0.1% formic acid; B = CH_3CN, 5%B 1CV, 5%B-95%B 10CV, 95%B 4CV) to give the desired product 2-(phenylsulfonamido)benzoic acid as a colorless powder. Yield: 34% (0.07 g, 0.25 mmol).

^{1}H NMR (d$_6$-acetone, 400 MHz) δ 8.01 (dd, J = 8.1 Hz, J = 1.6 Hz, 1H), 7.90-7.87 (m, 2H), 7.72-7.70 (m, 1H), 7.64-7.52 (m, 4H), 7.12 (td, J = 8.1 Hz, J = 1.1 Hz, 1H). ^{13}C NMR (d$_6$-acetone, 100 MHz) δ 170.6, 141.7, 140., 135.4, 134.1, 132.6, 130.1, 128.0, 123.9, 119.4, 117.1. LC-MS (ESI⁻) m/z 276.1 [M-H]⁻, retention time 1.88 min (100%). HRMS (ESI⁻) m/z (calcd $C_{13}H_{10}NO_4S$ [M-H]⁻ = 276.0336), obsd. 276.0329. IR (neat, cm⁻¹): 3120 (br), 1679, 1580, 1493, 1390, 1257, 1158.
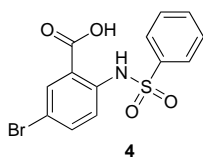
Ethyl 5-bromo-2-(phenylsulfonamido)benzoate (4a)
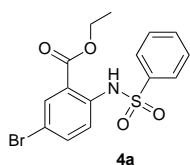
Ethyl 2-amino-5-bromobenzoate (0.202 g, 0.829 mmol) was dissolved in pyridine (2 mL) and cooled to 0 °C using an ice bath. Benzenesulfonyl chloride (0.12 mL, 0.940 mmol) was added dropwise to the reaction and the mixture was stirred at 0 °C for 1.5 h. Ice was added. The pink precipitate that formed was then isolated by vacuum filtration. The crude product was purified on a flash silica column using an Isolera (Biotage) (SNAP KP-Sil 24 g column, 12 mL/min, A = petroleum ether; B = EtOAc, 20%B 1CV, 20%B-58%B 6.2CV) to give the product ethyl 5-bromo-2-(phenylsulfonamido)benzoate as a colorless solid. Yield: 75% (0.239 g, 0.62 mmol).

^{1}H NMR (d$_6$-DMSO, 400 MHz) δ 10.35 (s, 1H), 7.92 (d, J = 2.6 Hz, 1H), 7.79-7.75 (m, 3H), 7.68 (m, 1H), 7.59-7.55 (m, 2H), 7.39 (d, J = 8.9 Hz, 1H), 4.27 (q, J = 7.1 Hz, 2H), 1.28 (t, 3H). ^{13}C NMR (d$_6$-DMSO, 100 MHz) δ 165.7, 138.5, 137.5, 136.81, 133.7, 133.1, 129.5, 126.9, 122.7, 121.0, 116.0, 62.0, 13.8. LC-MS (ESI⁻) m/z 384.0 [M-H]⁻, retention time 2.48 min (100%). HRMS (ESI⁺) m/z (calcd $C_{15}H_{15}BrNO_4S$ [M+H]⁺ = 383.9905), obsd. 383.9892. IR (neat, cm⁻¹): 3151 (br), 2995, 1685, 1576, 1481, 1446.

5-Bromo-2-(phenylsulfonamido)benzoic acid (4)

Ethyl 5-bromo-2-(phenylsulfonamido)benzoate (0.100 g, 0.259 mmol) was dissolved in ethanol (2 mL) and sodium hydroxide solution (2 M in H_2O, 2 mL, 4 mmol) was added. The mixture was heated at 50 °C for 1 h. The solvent was removed *in vacuo* and H_2O (10 mL) and DCM (10 mL) were added. The layers were separated and then the aqueous layer was acidified with hydrochloric acid solution (3 M) and a white precipitate formed. The product 5-bromo-2-(phenylsulfonamido)benzoic acid was collected by vacuum filtration as a colorless powder. Yield: 53% (0.049 g, 0.14 mmol).

^{1}H NMR (d_6-DMSO, 400 MHz) δ 11.11 (s, 1H), 7.95 (d, *J* = 2.5 Hz, 1H), 7.82-7.80 (m, 2H), 7.75-7.72 (m, 1H), 7.68-7.65 (m, 2H), 7.46 (d, *J* = 8.8 Hz, 1H). ^{13}C NMR (d_6-DMSO, 100 MHz) δ 168.4, 138.9, 138.4, 137.0, 133.8, 133.6, 129.6, 126.9, 120.8, 119.2, 115.1. LC-MS (ESI⁻) *m/z* 356.0 [M-H]⁻, retention time 2.16 min (100%). HRMS (ESI⁺) *m/z* (calcd $C_{13}H_{10}BrNO_4S$ [M+Na]⁺ = 377.9406), obsd. 377.9400. IR (neat, cm⁻¹): 3216 (br), 1703, 1667, 1573, 1478, 1448.

4-Bromo-2-(phenylsulfonamido)benzoic acid (5)
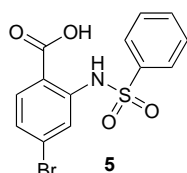
2-Amino-4-bromobenzoic acid (0.15 g, 0.69 mmol) was suspended in H_2O (5 mL) and benzenesulfonyl chloride (0.15 g, 0.83 mmol) was added to the suspension with vigorous stirring. Dioxane (5 mL) was added and the reaction mixture was heated at 37 °C overnight, and then allowed to cool to room temperature. The aqueous layer was extracted with EtOAc and the combined organic layers were dried with $MgSO_4$, filtered, and the solvent was evaporated to afford the crude product, which was purified on a reverse phase flash silica column using an Isolera (Biotage) (SNAP C18 30 g column, 20 mL/min, A = H_2O with 0.1% formic acid; B = CH_3CN, 5%B 1CV, 5%B-95%B 12CV, 95%B 2CV) to give the desired product 4-bromo-2-(phenylsulfonamido)benzoic acid as a colorless powder. Yield: 8% (0.020 g, 0.056 mmol).

^{1}H NMR (d$_6$-acetone, 400 MHz) δ 7.94-7.86 (m, 4H), 7.66-764 (m, 1H), 7.61-7.57 (m, 2H), 7.30 (dd, J = 8.5 Hz, J = 1.9 Hz, 1H). ^{13}C NMR (d$_6$-acetone, 100 MHz) δ 169.9, 142.9, 140.2, 134.4, 134.2, 130.3, 129.0, 128.0, 126.9, 121.9. LC-MS (ESI$^-$) m/z 356.1 [M-H]$^-$, retention time 2.16 min (92%). HRMS (ESI$^-$) m/z (calcd C$_{13}$H$_9$BrNO$_4$S [M-H]$^-$ = 353.9435), obsd. 353.9445. IR (neat, cm^{-1}): 3171 (br), 1669, 1386, 1247, 1168.
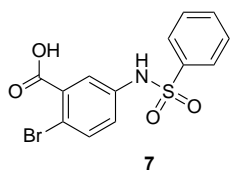
2-Bromo-5-(pyridine-3-sulfonamido)benzoic acid (6)
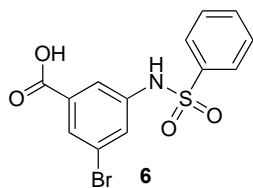
5-Amino-2-bromobenzoic acid (0.20 g, 0.93 mmol) was suspended in H$_2$O (4 mL) and pyridine-3-sulfonyl chloride (0.20 g, 1.13 mmol) was added to the suspension with vigorous stirring. 1,4-Dioxane (8 mL) was added and the reaction mixture was heated at 30 °C overnight, and then allowed to cool to room temperature. The aqueous layer was extracted with EtOAc and the combined organic layers were dried with MgSO$_4$, filtered, and the solvent was evaporated to afford the crude product, which was purified on a reverse phase silica column using an Isolera (Biotage) (SNAP C18 30 g column, 20 mL/min, A = H$_2$O with 0.1% formic acid; B = CH$_3$CN, 5%B 1CV, 5%B-40%B 4CV, 40%B 4CV, 40%B-65%B 2.4CV, 65%B-95%B 1CV, 95%B 2CV) to give the desired product 2-bromo-5-(pyridine-3-sulfonamido)benzoic acid as a colorless powder. Yield: 89% (0.29 g, 0.82 mmol).

^{1}H NMR (d$_6$-acetone, 400 MHz) δ 7.89-7.88 (m, 1H), 7.87-7.86 (m, 1H), 7.86-7.85 (m, 1H), 7.81 (t, J = 1.6 Hz, 1H), 7.66-7.63 (m, 2H), 7.59-7.56 (m, 2H). ^{13}C NMR (d$_6$-acetone, 100 MHz) δ 165.7, 140.7, 140.3, 134.5, 134.2, 130.2, 128.7, 127.9, 127.4, 123.0, 120.8. LC-MS (ESI$^+$) m/z 356.0 [M-H]$^-$, retention time 1.83 min (100%). HRMS (ESI$^-$) m/z (calcd C$_{13}$H$_9$BrNO$_4$S [M-H]$^-$ = 353.9435), obsd. 353.9444. IR (neat, cm^{-1}): 3251, 1697, 1336, 1298, 1155.

2-Bromo-5-(phenylsulfonamido)benzoic acid (7)

5-Amino-2-bromobenzoic acid (0.20 g, 0.93 mmol) was suspended in H_2O (4 mL) and benzenesulfonyl chloride (0.20 g, 1.11 mmol) was added to the suspension with vigorous stirring. Dioxane (8 mL) was added and the reaction mixture was heated at 30 °C overnight, and then allowed to cool to room temperature. The aqueous layer was extracted with EtOAc and the combined organic layers were dried with $MgSO_4$, filtered, and the solvent was evaporated to afford the crude product, which was purified on a reverse phase flash silica column using an Isolera (Biotage) (SNAP C18 30 g column, 20 mL/min, A = H_2O with 0.1% formic acid; B = CH_3CN, 5%B 1CV, 5%B-95%B 10CV, 95%B 3CV) to give the desired product 2-bromo-5-(phenylsulfonamido)benzoic acid as a colorless powder. Yield: 66% (0.22 g, 0.62 mmol).

^{1}H NMR (d6-acetone, 400 MHz) δ 9.32 (s, 1H), 7.85-7.83 (m, 2H), 7.73 (d, *J* = 2.7 Hz, 1H), 7.65-7.60 (m, 2H), 7.58-7.54 (m, 1H), 7.32 (dd, *J* = 8.7 Hz, *J* = 2.7 Hz, 1H). ^{13}C NMR (d6-acetone, 100 MHz) δ 166.6, 140.4, 138.4, 135.9, 134.4, 134.1, 130.1, 127.9, 125.1, 123.7, 116.5. LC-MS (ESI-) *m/z* 356.0 [M-H]⁻, retention time 1.72 min (100%). HRMS (ESI⁻) *m/z* (calcd $C_{13}H_9BrNO_4S$ [M-H]⁻ = 353.9435), obsd. 353.9442. IR (neat, cm⁻¹): 3151, 1695, 1473, 1291, 1157.

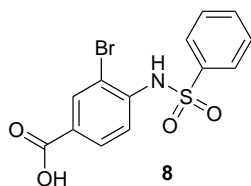## Methyl 3-bromo-4-(phenylsulfonamido)benzoate (8a)
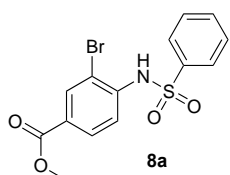
Methyl 4-amino-3-bromobenzoate (0.153 g, 0.665 mmol) was dissolved in pyridine (2 mL) and cooled to 0 °C using an ice bath. Benzenesulfonyl chloride (90 μL, 0.705 mmol) was added dropwise to the reaction and the mixture was stirred at 0 °C for 2 h. Ice was added and the product was extracted with EtOAc (2 × 10 mL). The combined organic layers were dried with $MgSO_4$, filtered, and the solvent removed *in vacuo*. The crude product was purified on a

flash silica column using an Isolera (Biotage) (SNAP KP-Sil 24 g column, 15 mL/min, A = petroleum ether B = EtOAc, 20%B 1CV, 20%B-51%B 7.6CV) to give the product methyl 3-bromo-4-(phenylsulfonamido)benzoate as a colorless powder. Yield: 27% (0.067 g, 0.18 mmol).

^{1}H NMR (d$_6$-DMSO, 400 MHz) δ 10.25 (s, 1H), 8.04 (d, J = 1.9 Hz, 1H), 7.85 (dd, J = 8.3 Hz, J = 1.80 Hz, 1H), 7.78 (d, J = 7.4 Hz, 2H), 7.67 (t, J = 7.4 Hz, 1H), 7.58 (t, J = 7.7 Hz, 2H), 7.39 (d, J = 8.3 Hz, 1H), 3.82 (s, 3H). ^{13}C NMR (d$_6$-DMSO, 100 MHz) δ 164.5, 140.1, 139.6, 133.8, 133.3, 129.4, 129.2, 128.1, 126.7, 125.7, 118.0, 52.5. LC-MS (ESI$^+$) m/z 270.0 [M-H]$^-$, retention time 2.11 min (100%). HRMS (ESI$^+$) m/z (calcd C$_{14}$H$_{12}$BrN$_4$O$_3$S [M+H]$^+$ = 369.9749), obsd. 369.9749. IR (neat, cm^{-1}): 3447, 3364, 3113 (br), 1726, 1651, 1616, 1559, 1450, 1412.

3-Bromo-4-(phenylsulfonamido)benzoic acid (8)

Methyl 3-bromo-4-(phenylsulfonamido)benzoate (0.036 g, 0.096 mmol) was dissolved in ethanol (1 mL) and sodium hydroxide solution (2 M in H$_2$O, 1 mL, 2 mmol) was added. The mixture was heated to 50 °C for 1 h. The solvent was removed *in vacuo* and H$_2$O (10 mL) and DCM (10 mL) were added. The layers were separated and then the aqueous layer was acidified with hydrochloric acid solution (3 M) upon which a white precipitate formed. The product 3-bromo-4-(phenylsulfonamido)benzoic acid was collected by vacuum filtration as a colorless solid. Yield: 45% (0.016 g, 0.045 mmol).

^{1}H NMR (d$_6$-DMSO, 400 MHz) 13.23 (s, 1H), 10.20 (s, 1H), 8.04 (d, J = 1.9 Hz, 1H), 7.85 (dd, J = 8.2 Hz, J = 1.9 Hz, 1H), 7.78 (d, J = 7.4 Hz, 2H), 7.67 (t, J = 7.4 Hz, 1H), 7.58 (t, J = 7.7 Hz, 2H), 7.36 (d, J = 8.2 Hz, 1H). ^{13}C NMR (d$_6$-DMSO, 100 MHz) δ 165.8, 140.5, 139.5, 134.21, 133.6, 129.7, 129.6, 127.0, 126.1, 118.4. LC-MS (ESI$^+$) m/z 356.0 [M-H]$^-$, retention time 1.76 min (100%). HRMS (ESI$^+$) m/z (calcd C$_{13}$H$_{10}$BrNO$_4$SNa [M+Na]$^+$ = 377.9406), obsd. 377.9398. IR (neat, cm^{-1}): 2158 (br), 1977, 1706.
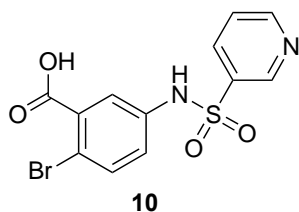
N-(2-Methoxypyridin-3-yl)benzenesulfonamide (9)

9
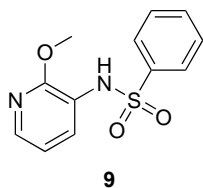
2-Methoxy-4-pyridin-3-amine (0.142 g, 1.14 mmol) was dissolved in pyridine (2 mL) and cooled to 0 °C using an ice bath. Benzenesulfonyl chloride (0.16 mL, 1.25 mmol) was added dropwise to the reaction and the mixture was stirred at 0 °C for 1 h. Ice was added and a brown precipitate formed. The product was isolated by vacuum filtration as a brown powder. Yield: 22% (0.066 g, 0.25 mmol).

^{1}H NMR (d$_6$-DMSO, 400 MHz) δ 9.83 (s, 1H), 7.92–7.90 (m, 1H), 7.73-7.71 (m, 2H), 7.64-7.60 (m, 1H), 7.57-7.52 (m, 3H), 6.95-6.92 (m, 1H), 3.59 (s, 3H). ^{13}C NMR (d$_6$-DMSO, 100 MHz) δ 156.7, 143.2, 140.2, 132.8, 132.7, 129.0, 126.6, 120.5, 117.1, 53.1. LC-MS (ESI$^+$) *m/z* 263.1 [M-H]$^-$, retention time 1.72 min (100%). HRMS (ESI$^+$) *m/z* (calcd C$_{12}$H$_{12}$N$_2$O$_3$S [M+Na]$^+$ = 286.0383), obsd. 286.0376. IR (neat, cm^{-1}): 3256, 1596, 1474, 1430.

2-Bromo-5-(pyridine-3-sulfonamido)benzoic acid (10)

10

5-Amino-2-bromobenzoic acid (0.20 g, 0.93 mmol) was suspended in H$_2$O (4 mL) and pyridine-3-sulfonyl chloride (0.20 g, 1.13 mmol) was added to the suspension with vigorous stirring. Dioxane (8 mL) was added and the reaction mixture was heated at 30 °C overnight, and then allowed to cool to room temperature. The aqueous layer was extracted with EtOAc and the combined organic layers were dried with MgSO$_4$, filtered, and the solvent was evaporated to afford the crude product, which was purified on a reverse phase flash silica column using an Isolera (Biotage) (SNAP C18 30 g column, 20 mL/min, A = H$_2$O with 0.1% formic acid; B = CH$_3$CN, 5%B 1CV, 5%B-40%B 4CV, 40%B 4CV, 40%B-65%B 2.4CV, 65%B-95%B 1CV, 95%B 2CV) to give the desired product 2-bromo-5-(pyridine-3-sulfonamido)benzoic acid as a colorless powder. Yield: 89% (0.29 g, 0.82 mmol).

^{1}H NMR (d$_6$-acetone, 400 MHz) δ 8.96 (dd, J = 2.7 Hz, J = 0.8 Hz, 1H), 8.79 (dd, J = 4.8 Hz, J = 1.6 Hz, 1H), 8.17 (dq, J = 8.1 Hz, J = 1.6 Hz, J = 0.8 Hz, 1H), 7.74 (d, J = 2.7 Hz, 1H), 7.65 (d, J = 8.6 Hz, 1H), 7.59 (qd, J = 8.8 Hz, 4.02 Hz, J = 2.6 Hz, 1H), 7.35 (dd, J = 8.6 Hz, J = 2.7 Hz, 1H). ^{13}C NMR (d$_6$-acetone, 100 MHz) δ 166.5, 154.7, 148.5, 137.7, 136.7, 136.1, 135.7, 134.5, 125.6, 125.0, 124.1, 117.2. LC-MS (ESI$^+$) m/z 359.0 [M+H]$^+$, retention time 1.53 min (100%). HRMS (ESI$^-$) m/z (calcd C$_{12}$H$_{10}$BrN$_2$O$_3$S [M+H]$^+$ = 356.9539), obsd. 356,9537. IR (neat, cm^{-1}): 3266, 1683, 1328, 1237, 1168.

4-Bromo-2-(naphthalene-2-sulfonamido)benzoic acid (11)
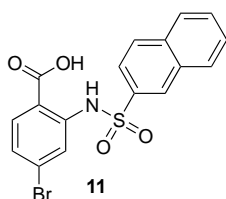
2-Amino-4-bromobenzoic acid (0.20 g, 0.93 mmol) was suspended in H$_2$O (3 mL) and naphthalene-2-sulfonyl chloride (0.25 g, 1.10 mmol) was added to the suspension with vigorous stirring. Dioxane (7 mL) was and the reaction mixture was heated at 37 °C for 5 h, then left stirring at room temperature overnight. The aqueous layer was extracted with EtOAc and the combined organic layers were dried with MgSO$_4$, filtered and the solvent was evaporated to afford the crude product, which was purified by reverse phase flash silica column using an Isolera (Biotage) (SNAP C18 30 g column, 20 mL/min, A = H$_2$O (0.1% formic acid) B = CH$_3$CN, 5%B 2CV, 5%B-95%B 12CV, 95%B 2CV) to give the desired product 4-bromo-2-(naphthalene-2-sulfonamido)benzoic acid as a colorless powder. Yield: 9% (0.035 g, 0.090 mmol).

^{1}H NMR (d$_6$-acetone, 400 MHz) δ 8.61-8.60 (d, J = 1.8 Hz, 1H), 8.19-8.17 (m, 1H), 8.09 (d, J = 8.7 Hz, 1H), 8.03-8.00 (m, 1H), 7.94 (d, J = 1.9 Hz, 1H), 7.89 (d, J = 8.5 Hz, 1H), 7.84 (dd, J = 8.7 Hz, J = 1.9 Hz, 1H), 7.78 (dd, J = 8.5 Hz, J = 1.9 Hz, 1H). ^{13}C NMR (d$_6$-acetone, 100 MHz) δ 169.8, 142.7, 136.8, 136.0, 134.1, 133.0, 130.8, 130.3, 129.8, 129.4, 128.9, 128.8, 127.0, 122.8, 121.9, 115.7. LC-MS (ESI$^+$) m/z 406.0 [M-H]$^-$, retention time 2.41 min (100%). HRMS (ESI$^-$) m/z (calcd C$_{17}$H$_{12}$BrNO$_4$SNa [M+Na]$^+$ = 427.9568), obsd. 427.9562. IR (neat, cm^{-1}): 1675, 1387, 1246, 1164.