Unbiased Mass Spectrometry Elucidation of the Targets and Mechanisms of Activity-Based Probes: A Case Study involving Sulfonyl Fluorides

Thomas E. J. Chavas,[†] Matthew J. Fuchter,[†] and Peter A. DiMaggio Jr.[*,‡]

[†]*Department of Chemistry, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom*

[‡]*Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom*

E-mail: p.dimaggio@imperial.ac.uk

Phone: +44 (0)20 75945589

Abstract

The elucidation of protein/drug interactions remains a major challenge in drug discovery. Liquid chromatography tandem mass spectrometry has emerged as a tremendously powerful technology for this endeavor, but its full potential has yet to be realised owing in part to unresolved challenges in data analysis. Herein, we demonstrate how tandem mass spectrometry can comprehensively map small molecule/peptide adducts when combined with unconstrained sequencing. Using a published sulfonyl fluoride activity-based probe as a model system, this method enabled the discovery of several unreported sites of interaction with its target proteins. Crucially, this probe was found to undergo quantitative displacement and hydrolysis from the target protein's active

site. Isotopic labelling experiments provided a mechanistic rationale for the observed hydrolysis that involves neighbouring-group participation. A chemical biology tagging strategy that leverages the probe's observed lability was developed and shown to be compatible with the original small molecule inhibitor in discovery profiling experiments.

Keywords

Activity-Based Probes, Sulfonyl Fluorides, Chemical Proteomics, Tandem Mass Spectrometry

Introduction

Within drug development programs, there is a critical need to comprehensively identify the biological targets of a candidate drug as early as possible.[1] The ability to rigorously identify and validate the full spectrum of drug/protein interactions would significantly help in circumventing the commonly observed stalling in phase II clinical trials, where drug candidates are expected to demonstrate that modulation of a particular target leads to the desired phenotypic effect.[2] In addition, identifying a drug's off-targets may be useful to predict safety liabilities,[1] and to understand a drug's side-effects.[3] However, the task of fully elucidating small molecule interactions with the specific residues of its target protein remains a grand challenge and presents a major technology gap.[1]

In this context, activity-based protein profiling (ABPP) has garnered considerable interest from the research community over the past two decades.[4] This technology consists in designing a covalent mimic of the parent drug, to which is appended an affinity handle. Thus, the activity-based probe (ABP) offers the prospect of targeting, isolating and identifying its protein targets in complex environments. Sulfonyl fluorides (SF) in particular have become very popular warheads for ABPs, as demonstrated in a number of studies.[5–8] Their wide-use has prompted a recent foray exploring the factors influencing their reactivity as well as the

stability of the resulting adducts, using protected amino acids as model nucleophiles.[9]

The primary endpoint assay coupled to ABPP workflows is liquid chromatography tandem mass spectrometry (LC-MS/MS) since it provides unbiased identification of the protein targets and can, at least in principle, identify the specific residue associated with the ABP interaction based upon the mass shift of the probe. Despite recent advancements in the development of cleavable linkers for selective enrichment,[10] the relative number of unambiguous MS/MS identifications of ABP/peptide adducts reported in this field is disproportionately low.[11–15] A possible explanation for the general lack of ABP/peptide MS/MS annotations is that the adduct may undergo further chemical modification in the gas phase and/or during sample processing, resulting in an unanticipated change in its total mass.[13,16] This represents a major hurdle in the analysis of mass spectrometry data, as peptide identification algorithms are reliant upon some form of *a priori* knowledge regarding what protein modifications are expected to be present, the amino acids they target, and their exact elemental composition. Recent studies utilising mass-tolerant search strategies have highlighted that more than half of the MS/MS spectra collected in a standard shotgun proteomics experiment remain unannotated,[16] partly owing to these search constraints.

Using a published sulfonyl fluoride probe (SFABP) as a case study,[5] we demonstrate how unbiased mass spectrometry combined with unconstrained sequencing can be used to characterise unexpected probe-protein interactions to an unprecedented level. The fundamental tenet behind this approach is that mass spectrometry data contains all the necessary information, but that the target residue and elemental composition of the probe must be treated as variables for it to be fully exploited. Using this approach, we found SFABP to react with several lysine, tyrosine and serine residues on trypsin, chymotrypsin and GST variants. Unconstrained sequencing also identified that the SFABP bound to the active site serine of trypsin (S200) is entirely displaced by hydrolysis, and heavy water labelling experiments elucidated two competing mechanisms.

We have created a public resource called the SFABP database, which contains the anno-

tated MS/MS resulting from our unconstrained sequencing approach and provides structural information regarding the proteins studied and the peptide/ABP adducts observed. The findings described in this article hyperlink directly to their corresponding entries in the SFABP database, which can be accessed at `https://dimaggiolab.github.io/SFABP_DB/`.

Results and Discussion

Discovery of Non-Active Site Residues Labelled with SFABP

Our first experiment involved incubating the SFABP (Figure 1A) with purified trypsin using the conditions reported in Shannon et al..[5] After labelling, excess probe was removed via MeOH/$CHCl_3$ precipitation and digestion of the labelled protein was performed. The resulting peptides were analysed by LC-MS/MS and unconstrained site sequencing of the MS/MS was performed using variable target residues to identify sites labelled by SFABP that were not present in vehicle-treated samples, as described in the Experimental Section.

SFABP is expected to bind covalently to nucleophilic residues, through displacement of the fluoride leaving group (Figure 1B). Although the expected probe mass of 277 Da (SFABP minus hydrogen fluoride) was not observed on the peptide containing the active site serine (S200), the search unambiguously identified MS/MS annotations corresponding to such SFABP labelling of trypsin lysines K114, K159, K162 and K240 (Figure 1 C, D). Inspection of the crystal structure for trypsin revealed that these lysine residues are surface-exposed and not proximal to the active site (Figure 1D). Given that no lysine labelling by SFABP had been reported by Shannon et al., it was considered this reactivity might be an artefact resulting from the purified conditions of the assay. To reduce the extent of non-specific labelling, the experiment was repeated in varying backgrounds of nuclear proteins extracted from HeLa cells. The sample was processed for MS proteomics as described above, and we repeatedly observed labelling of residues K114, K162 and K240 in a five-fold excess (w/w) of nuclear protein extract to recombinant trypsin. SFABP labelling of K114 and K162

was confirmed even in the presence of a 20-fold excess (w/w) of nuclear protein extract to purified bovine trypsin.

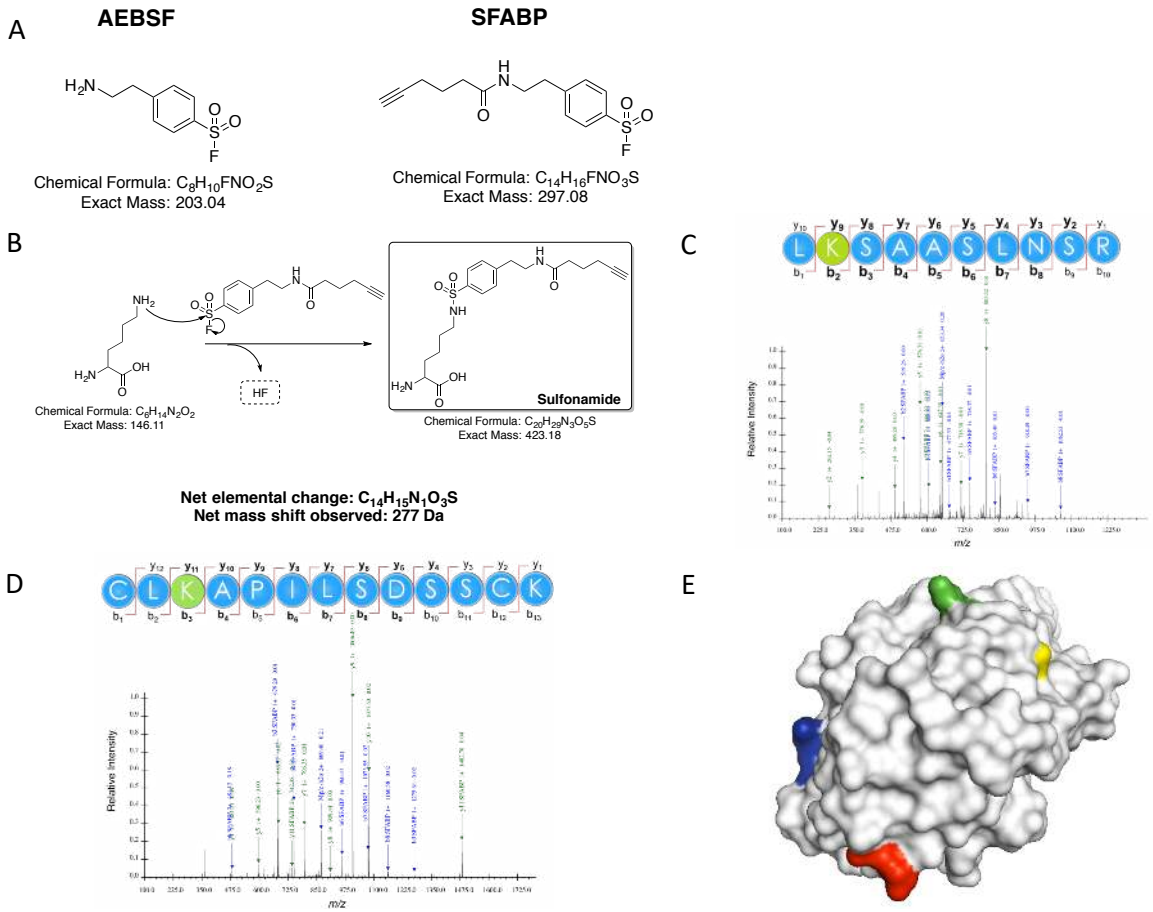


Figure 1: Labelling of bovine trypsin's lysine residues by the SFABP. (a) Structures of the serine protease inhibitor AEBSF and of the corresponding activity-based probe SFABP. (b) Mechanistic rationale for labelling of the SFABP by lysine residues, resulting in sulfonamide formation. (c) MS/MS annotation supporting labelling of bovine trypsin's K114 by the SFABP (peptide L113-R222). (d) MS/MS annotation supporting labelling of bovine trypsin's K162 by the SFABP (peptide C160-K172). (e) Lysine residues K114, K159, K162 and K240 were mapped onto the crystal structure of bovine trypsin (PDB 1S0Q) and are visualized in red, yellow, green and blue, respectively.

We then examined labelling of recombinant bovine chymotrypsin by the SFABP based on Shannon et al.'s observation that this molecule could also serve as an ABP for this protease.[5] Unambiguous MS/MS annotations confirming the expected SFABP labelling of Y146, K169 and Y171 were identified, and all three residues were confirmed to be surface-

exposed based on the crystal structure for chymotrypsin. SFABP labelling of chymotrypsin's active site serine (S195) was not identified. Although the chymotrypsin peptide containing S195 is rather large (25 amino acids), unambiguous MS/MS annotations of the corresponding unmodified peptide were identified. This either suggests that the expected SFABP adduct is not present on the active site serine or that it is inhibiting MS/MS fragmentation to a considerable extent.

Since other studies have reported that glutathione transferases (GSTs) can also be labelled by sulfonyl fluoride probes,[8] the experimental workflow described above was repeated using a mixture of purified equine GSTs. Performing a general database search against the horse proteome using Proteome Discoverer (version 1.3) of the tandem MS identified the following GST equine variants to be present: F6XQQ9, F6RZ56, F7BHA2 and F6VSN2. In an attempt to identify the expected SFABP adduct on GST peptides, unconstrained site sequencing allowing for variable target residues was performed. Three tyrosine residues (Y9, Y82 and Y109) on variants F6XQQ9, F6RZ56 and F6VSN2, respectively, were unambiguously identified as targets of the SFABP. As none of these GST variants studied have been crystallised to date, the crystal structures of homologous GST proteins were used to map the approximate location of the identified target residues. Interestingly, the homologous residue to Y109 of GST variant F6VSN2 (YVTLIY*TNYEAGK) was recently identified to react with the SFABP in a mouse liver and pancreas proteome (GST variant P19157, YVTLIY*TNYENGK).[8] It should be mentioned that these two GST variants share more than 90% sequence identity (as determined by Clustal 2.1), and that the mouse variant P19157 has been crystallised in complex with a nitrobenzyl glutathione inhibitor (PDB 1GLQ). Interestingly, the nitrobenzyl moiety of the inhibitor lies in close proximity to Y109's phenolic oxygen. This is in line with the observed reactivity of the SFABP's sulfonyl fluoride moiety with the homologous phenolic oxygen, which suggests that both inhibitors could adopt a similar binding pose. F6XQQ9 and F6RZ56, on the other hand, respectively share 78% and 79% sequence identity with a mutant human GSTA1 whose crystal structure

has been resolved (PDB 3I69).[17] Consistent with our observation, the homologous tyrosine residues in this mutant GSTA1 are both surface-exposed, whilst the homologous residue to Y9 of GST F6XQQ9 lines the protein's GSH-binding pocket.

In contrast to the previous experiments characterising the targets of trypsin and chymotrypsin, SFABP labelling of serine residue S142 was observed on either GST variant F6RZ56 or F7BHA2, which share full sequence homology for this peptide. A third GST variant (F6XQQ9) with the same peptide sequence, save for one residue at the C-terminus, was also found to be labelled by the SFABP at serine residue S142. It was found that F6XQQ9, F6RZ56 and F7BHA2 share 78%, 79% and 78% sequence identity with the aforementioned mutant human GSTA1.[17] This mutant protein and F6XQQ9 have an identical S142-containing tryptic peptide and, consistent with our observations, this homologous residue is surface-exposed.

Interestingly, in replicate experiments an adjacent tyrosine residue, Y147, was found to be labelled with the probe instead of serine S142 on the same peptide. We rationalized this by invoking the potential instability of sulfonylated serine residues,[18–20] which is discussed in the next section. It should be noted that, although they are both surface-exposed, S142 and Y147 are not in immediate spatial proximity based upon the crystal structure of GSTA1.

Active Site Labelling by the SFABP

As mentioned above, our analysis did not identify any of the expected SFABP adduct on the active site serine residues of trypsin (S200) or chymotrypsin (S195). However, analysis of the LC-MS/MS data corresponding to SFABP-labelled trypsin using unconstrained mass sequencing (i.e. considering a modification of variable mass on the peptide containing S200) resulted in the unambiguous identification of an unexplained adduct of 193 Da on S200 (Figure 2A), as opposed to the 277 Da expected for the SFABP adduct. The formation of this adduct is probe-induced since it is not observed in control (vehicle-treated) samples.

In line with the observation that sulfonylated serine residues are susceptible to nucle-

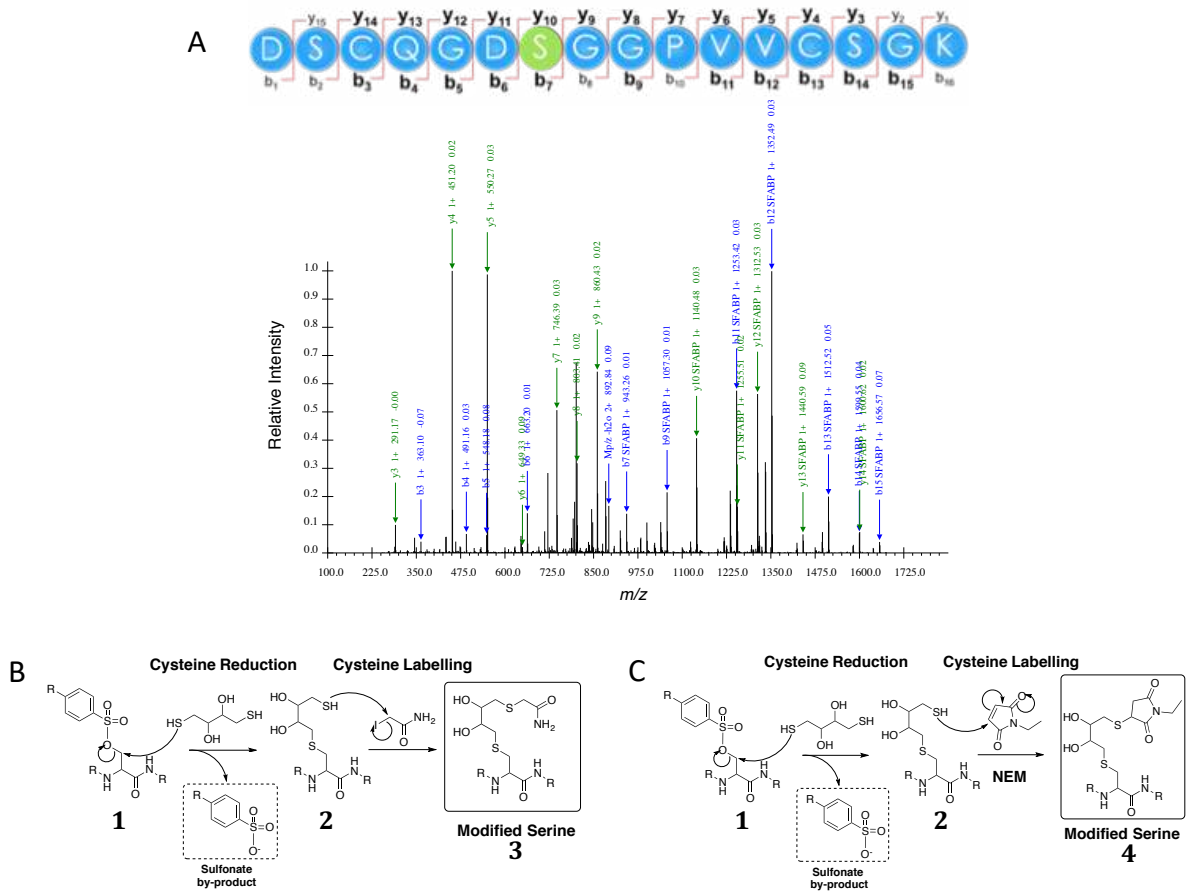


Figure 2: Labelling of bovine trypsin's active-site serine with the SFABP, followed by nucleophilic displacement of the sulfonate by DTT. (a) MS/MS annotation supporting the formation of an adduct of +193 Da on S200 (peptide D194-K209). (b) Mechanistic rationale for the formation of a +193 Da adduct on S200 during standard sample processing (DTT reduction followed by IA treatment). (c) Mechanistic rationale for the observed formation of a +261 Da adduct on S200 when IA was substituted for NEM during sample processing.

ophilic attack,[19] we surmised that the reducing agent dithiothreitol (DTT), which is used in sample processing prior to enzymatic digestion, could act as a nucleophile (Figure 2B). We postulated that the resulting thiol **2** could then react with iodoacetamide (IA) in the subsequent protein alkylation, where this reagent is used to cap reduced cysteine residues. The resulting thioether **3** on active site serine S200 accounts for the observed modification mass of 193 Da at this position. Consistent with this hypothesis, substituting iodoacetamide for N-ethylmaleimide (NEM) as the reagent for protein alkylation results in a change in the

mass of the adduct observed on S200 that is line with the formation of DTT-NEM adduct **4** (Figure 2C). In this experiment, a DTT-NEM adduct of mass 261 Da was identified at S200, whilst the 193 Da adduct (corresponding to adduct **3**) found in the previous experiment was not observed. Furthermore, when tris-carboxyethylphosphine (TCEP) was used instead of DTT as the reducing agent prior to protein alkylation of SFABP-labelled trypsin, neither adducts **3** or **4** were observed to have formed at S200. To confirm that S200 labelling by SFABP (followed by formation of adducts **3** or **4**) was dependent upon the integrity of the active site, thermal-denaturation experiments were carried out and adducts were identified via MS/MS as described above. Non-specific SFABP labelling of surface-exposed lysines K114, K162 and K240 was consistently observed in these experiments, whereas S200 labelling (in the form of the aforementioned adduct) was lost, which implies that interaction of the SFABP with S200 is dependent on the integrity of the active site.

Given that the SFABP interaction with serine residues is specific to S200 and that sulfonyl fluoride moieties are reported to be highly reactive,[6] it was surprising to find that adducts **3** and **4** observed on S200 were low in abundance (signal intensity of the modified peptide was found to be approximately 0.6% and 1.5% of unmodified peptide, respectively). A plausible explanation for this is that degradation (e.g. hydrolysis) of the SFABP-labelled S200 **1** occurred to a significant extent during sample processing and/or LC-MS/MS analysis. To test this hypothesis, 'heavy' PBS was prepared using $^{18}OH_2$ and SFABP-labelled trypsin was incubated in this buffer. To prevent the trypsin-mediated incorporation of ^{18}O at the C-terminus of the peptides, which occurs during proteolysis in heavy water,[21] the protein was precipitated and re-suspended in light PBS for in-solution digestion following the standard protocol. The isotopic envelope of the 'unmodified', S200-containing peptide revealed an unambiguous exogenous incorporation of one ^{18}O atom (Figure 3A). In the corresponding control experiments, where vehicle-treated trypsin was incubated in 'heavy' ^{18}O PBS (Figure 3B) and SFABP-labelled trypsin was incubated in 'light' ^{16}O PBS (Figure 3C), the isotopic envelope for the same peptide is observed to only contain naturally-occurring isotopes. This

data indicates that significant hydrolysis of the SFABP/S200 adduct takes place during sample processing.

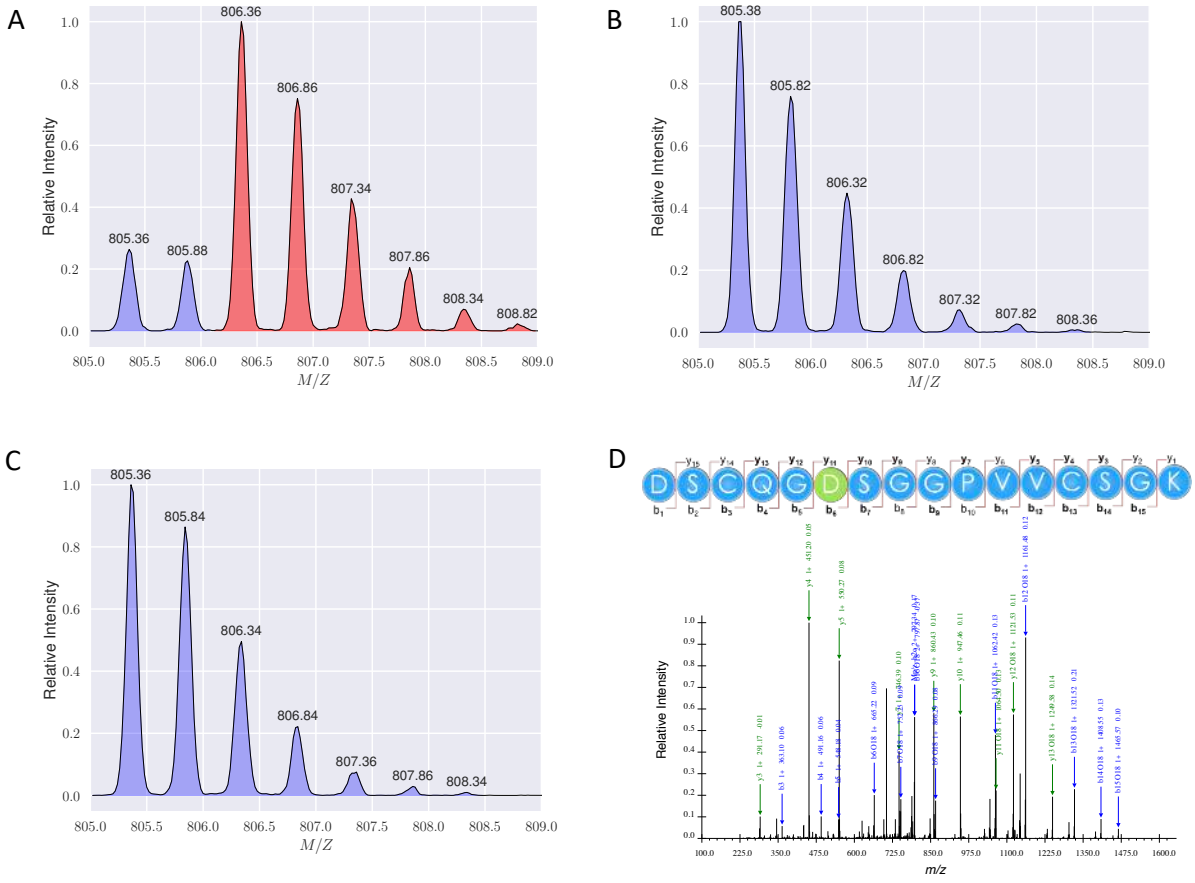


Figure 3: $^{18}OH_2$-mediated hydrolysis of SFABP-labelled bovine trypsin. (a) MS spectrum of the active-site peptide after SFABP-treatment of the parent protein and incubation in heavy PBS (made with $^{18}OH_2$). (b) MS spectrum of the active-site peptide after vehicle-treatment of the parent protein and incubation in heavy PBS (made with $^{18}OH_2$). (c) MS spectrum of the active-site peptide after SFABP-treatment of the parent protein and incubation in light PBS (made with $^{16}OH_2$). (d) MS/MS annotation supporting the incorporation of a ^{18}O atom on D199, the residue adjacent to the active-site serine (peptide D194-K209).

The extent of ^{18}O incorporation was computed using linear regression, where the theoretical isotopic envelopes for the S200-containing peptide with no and full ^{18}O incorporation served as basis functions to fit the observed isotopic envelope (Figure 3A). Based on the weights estimated by linear regression, it appears that $(75 \pm 2)\%$ of ^{18}O incorporation has

taken place at the time of analysis (Supplementary Material 1). Assuming the extent of hydrolysis of the SFABP-S200 adduct to be 100%, it follows that the yield of ^{18}O incorporation reflects the yield of SFABP labelling. Given that we observed none of the expected SFABP adduct bound to S200 and that DTT/IA adduct **3** was always found to be of very low abundance (2-3 orders of magnitude lower in intensity compared to that of the corresponding unmodified peptide), we believe this assumption to be accurate.

Mechanisms for Hydrolysis of the SFABP: 5-*exo*-tet versus lactonisation

To confirm the site of ^{18}O incorporation on the S200-containing peptide, the corresponding MS/MS data was subjected to unconstrained sequencing starting with a +2 Da modification at S200. This analysis revealed that ^{18}O incorporation had occurred at D199 rather than at S200, as shown in the unambiguous MS/MS annotation (Figure 3D). Seminal work by Fahrney and Gold,[22] who studied the inhibition of chymotrypsin by phenylmethanesulfonyl fluoride (PMSF) and its subsequent, spontaneous reactivation, proposed a mechanism for hydrolysis of the sulfonylated serine that involves a 5-*exo*-tet cyclisation to form oxazoline intermediate **6** (Figure 4A). In our experiment with heavy water, this mechanism would lead to the formation of peptide **7** and the incorporation of one ^{18}O atom at the amide moiety of the aspartic acid residue, which is consistent with our observation.

However, the MS/MS of this same peptide shows that the *b*- and *y*- fragment ions containing residue D199 uniquely exhibit characteristic neutral losses that are both 18 and 20 Da lighter than the parent ion. The loss of 18 Da is to be expected since it is the result of water loss from the sidechain of aspartic acid, which is commonly seen in tandem mass spectra.[23] By analogy, the loss of 20 Da would indicate that a heavy oxygen atom (^{18}O) has been incorporated into the sidechain of the residue, yet this is not consistent with Fahrney and Gold's proposed mechanism and the resulting peptide **7**.[22] Thus, we propose lactonisation as a competing mechanism, where the carboxylic acid sidechain of D199 attacks the sulfony-

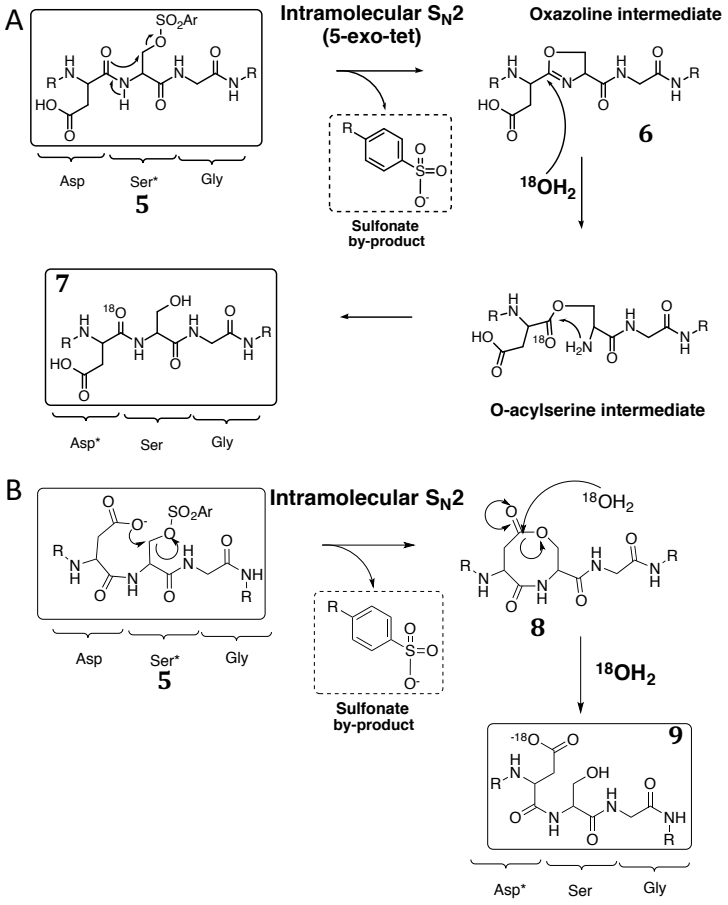


Figure 4: Competing mechanisms for the incorporation of ^{18}O at D199 following SFABP-labelling of S200. (a) 5-*exo*-tet cyclisation followed by oxazoline hydrolysis, as proposed by Fahrney and Gold. (b) Neighbouring-group participation: lactonisation followed by hydrolysis.

lated serine to yield 8-membered lactone **8** (Figure 4B). Subsequent hydrolysis of lactone **8** by heavy water leads to the formation of peptide **9**, which has an identical sequence to the original peptide but has incorporated ^{18}O on the sidechain of D199. However, it should be noted that due to tautomerism, this resulting peptide **8** is prone to both heavy ($^{18}OH_2$) and light ($^{16}OH_2$) water losses during CID fragmentation (Figure 5A).

To determine if the contribution of the lactonisation mechanism of hydrolysis is significant, the relative intensities of the MS/MS fragment ions were further examined. Figure 5A represents 1) how the individual contributions of the peptides **7** and **9** combine to form a

single ion peak (Signal D), 2) the respective 18 and 20 Da neutral loss peaks from this peak (Signals B and A, respectively), and 3) the contribution of the +2 Da naturally occurring isotope from the unlabelled peptide (Signal C, which serves as a control ion peak). If a loss of heavy water ($^{18}OH_2$; 20 Da) is indeed significant, which in turn implies that lactonisation is a competing mechanism, then we would expect the ratio of Signal A over Signal C to increase for D199-containing fragment ions relative to control experiments.

It should be noted that certain fragment ions were excluded from the analysis if they had observed interferences with other fragment ions. For example, y_{11}^+ and b_{11}^+ were excluded as their corresponding signals are close in m/z (i.e. 1062.47 *versus* 1060.38, respectively). Similarly, a few ions were also excluded (e.g. y_7) if they were found to interfere with uncharacterised MS/MS ions. In total, seven D199-containing ions were observed to be free of interference from other species. Figure 5B shows that for the seven D199-containing fragment ions (b_6^+, b_7^+, b_9^+, b_{12}^+, b_{13}^+, y_{12}^+, y_{13}^+), the ratio of Signal A over Signal C was found to be significantly larger ($p < 0.01$) in SFABP labelling experiments conducted in heavy ^{18}O PBS, relative to vehicle-treated experiments carried out in the same buffer (Figure 5B and Supplementary Material 2.A, 2.B). Three fragment ions that did not contain D199 (b_4^+, y_4^+, y_{10}^+) were used as a reference control, and the ratio of Signal A to Signal C did not change significantly for these ions (Figure 5B and Supplementary Material 2.B). Thus, this data supports the hypothesis that lactonisation is a competing mechanism for hydrolysis of the SFABP, as losses of heavy water are occurring from the side chain of D199.

The model in Figure 5A was also used to estimate the relative contributions of the 5-*exo*-tet and lactonisation mechanisms to hydrolysis of the SFABP/S200 adduct. First, the average intensity of the loss of water peak relative to its parent ion were estimated from control experiments for the fragment ions listed in Figure 5B, where only the loss of water (18 Da) from Signal C is observed in Signal A (Supplementary Material 2.C). These relative intensities were then used to compute the approximate contribution of Signal C to Signal A in treated samples (i.e. the green contribution to signal A). The remaining intensity of Signal

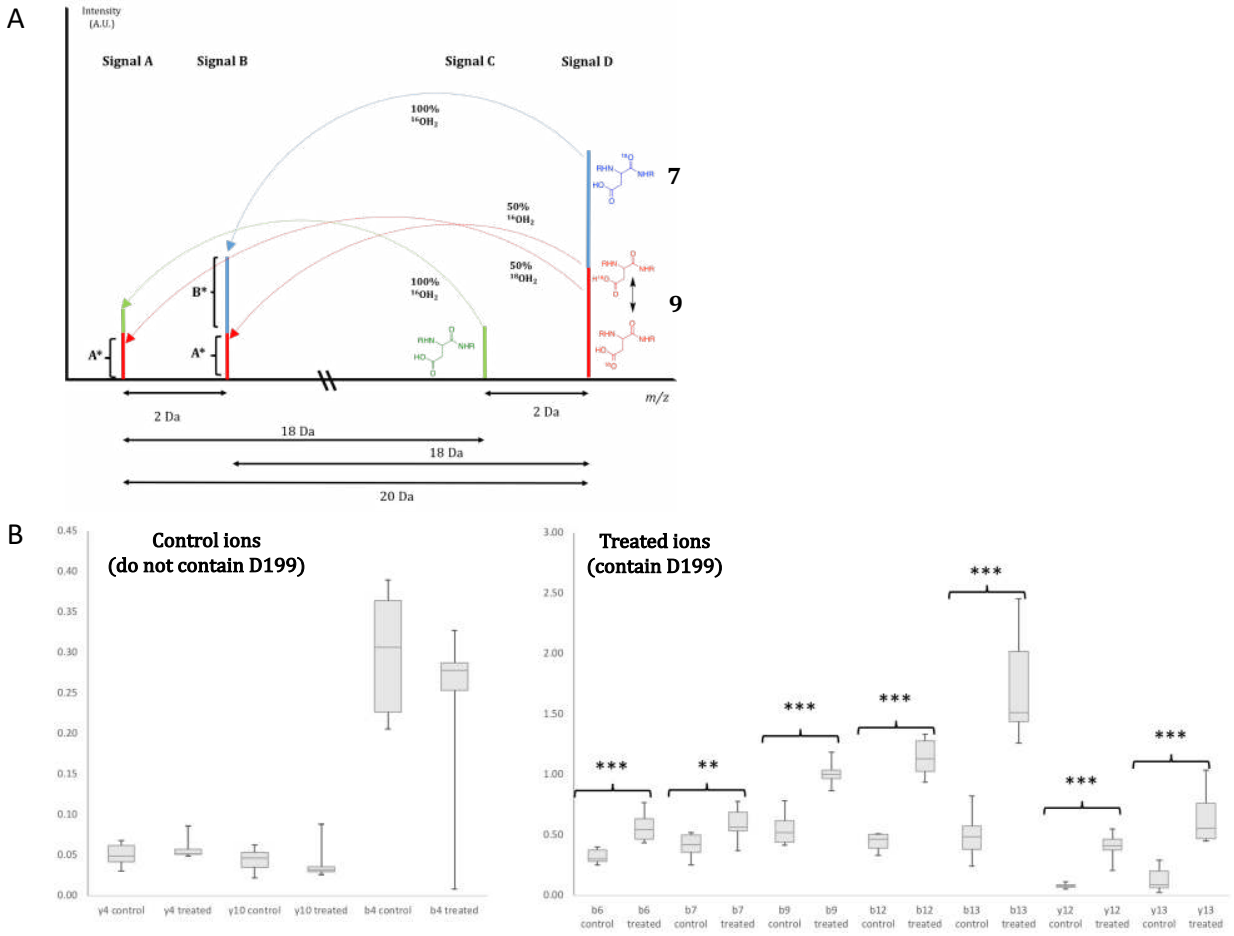


Figure 5: Assessing the contribution of the two proposed mechanisms through examination of water loss from the aspartic acid side chain. (a) Schematic representation of the MS/MS of the +2 Da parent ion of bovine trypsin's D199-containing peptide after SFABP-treatment of the parent protein in heavy PBS (i.e. containing $^{18}OH_2$). Contributions to signal D are due to the presence of ^{18}O on the backbone (blue contribution) or on the side chain (red contribution) of the residue. Contributions to signal C are due to isobaric parent peptides which bear a mixture of naturally-occurring isotopes elsewhere in the peptide, but whose fragment ion does not contain those heavy isotopes. (b) For every non-interfering fragment ion present in the MS/MS of the +2 Da parent, S200-containing peptide, the ratio of signals A/C was computed, and compared with the corresponding values from the control (vehicle-treated) experiment. This ratio is higher in treated experiments, for all fragment ions containing D199 (i.e. not b_4^+, y_4^+, and y_{10}^+ ions), indicating that lactonisation is a significant competing mechanism in the hydrolysis of the sulfonylated active site serine.

A, A^*, could be attributed to the loss of heavy water from the lactonisation contribution (i.e. peptide **9**) to Signal D. Since tautomerism results in a random loss of both heavy and light water from the lactonisation product ion, we approximated the contribution of A^* to be

equal in Signals A and B, and thus the relative contribution of the lactonisation mechanism to hydrolysis of the SFABP is given in Equation 1:

$$\text{Relative contribution of lactonisation mechanism } = \frac{2A^*}{2A^* + B^*} \quad (1)$$

Based on this model, we estimate that lactonisation contributes to (31 ± 16) % of hydrolysis of the SFABP from the active site serine (Supplementary Material 2.D).

A Chemical Proteomics Approach to Harness SFABP Lability

The data presented thus far demonstrates that a considerable proportion of the SFABP-labelled active site serine undergoes hydrolysis during sample processing. This considerably limits its use as an ABP to profile the activity of trypsin and, more generally, serine proteases. We postulated that the susceptibility of a sulfonylated serine to undergo nucleophilic attack could be leveraged to selectively replace the SFABP with a functional group that increases the likelihood of mass spectrometry identification of probe-binding to the serine.

Based on the pioneering work by Koshland and Neet[19] and recent applications of ligand-directed tosyl chemistry for labelling proteins,[24] 3-bromothiophenol was examined to see if it would be sufficiently nucleophilic to attack sulfonylated serine **1**, thus yielding the adduct **10** (Figure 6A). Notably, the presence of a bromine atom on the labelled residue would enable rapid detection of the exogenous adduct through inspection of the MS isotopic distribution of the peptide (Figure 6B). To confirm our hypothesis, SFABP-labelled trypsin was dissolved in a solution of 3-bromothiophenol (100 mM in MeOH). The sample was then re-suspended in water, alkylated, digested and analyzed using LC-MS/MS following the standard protocol. The resulting data was searched for labelling of S200 by 3-bromothiophenol and unambiguous MS/MS annotations corresponding to the two expected adducts (bearing ^{79}Br and ^{81}Br) were identified. This bromine labelling results in a unique isotopic distribution for the parent peptide, which is characteristic of any bromine-containing molecule (Figure 6C). The

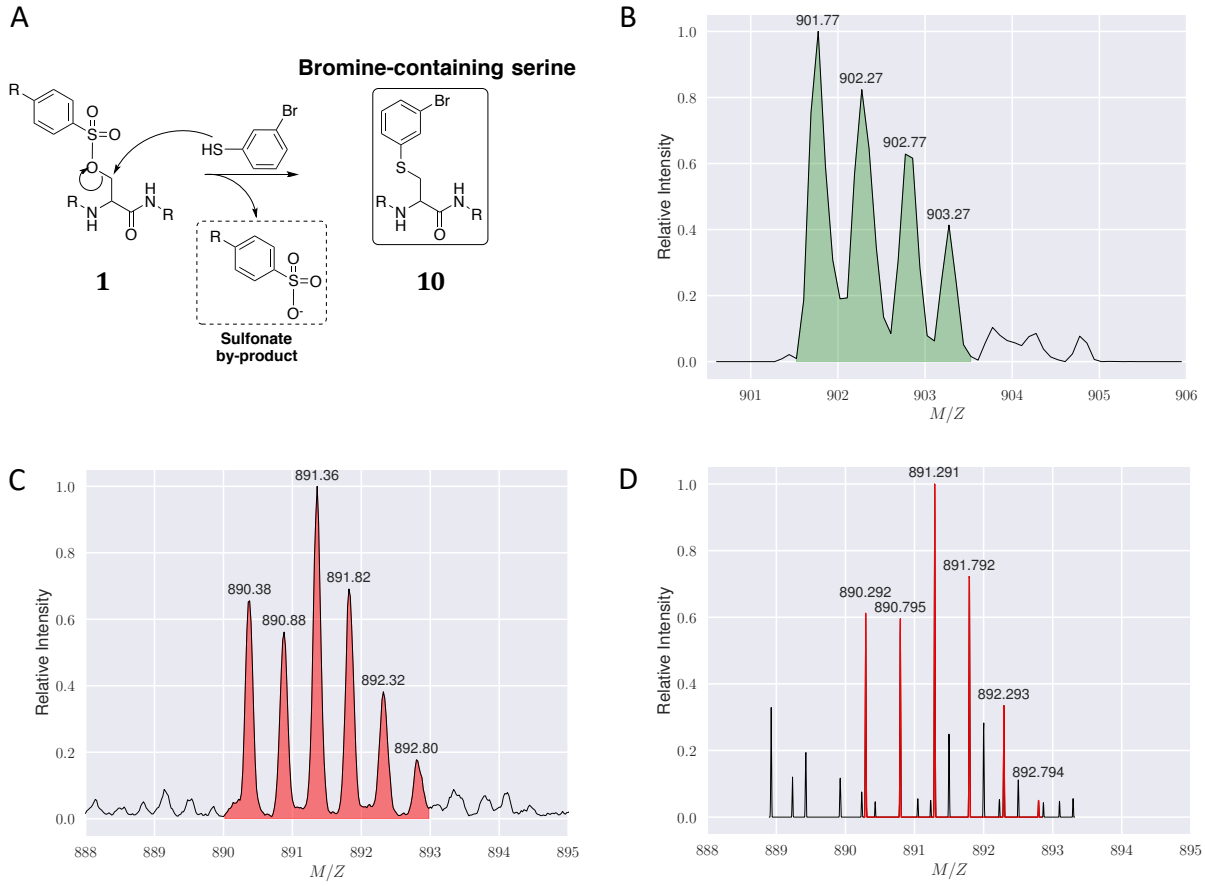


Figure 6: Labelling of bovine trypsin's active-site serine with the SFABP, followed by nucleophilic displacement of the sulfonate by 3-bromothiophenol. (a) Mechanistic rationale for the formation of a brominated thioether from a SFABP-labelled serine. (b) MS spectrum of bovine trypsin's active-site peptide after SFABP-treatment of the parent protein and standard processing of the protein (DTT and IA treatment). (c) MS spectrum of bovine trypsin's active-site peptide after SFABP-treatment of the parent protein followed by 3-bromothiophenol treatment and standard processing of the protein (DTT and IA treatment). (d) MS spectrum of bovine trypsin's active-site peptide after SFABP-treatment of the parent protein, spiked in a HeLa nuclear extract, followed by 3-bromothiophenol treatment and standard processing of the protein sample (DTT and IA treatment).

exogenous nature of bromine further enables detection of labelled peptides, regardless of whether or not they have been selected for MS/MS fragmentation by the mass spectrometer (e.g., due to low abundance).

To test this approach in a more complex protein background, the above experiment was repeated in HeLa nuclear protein extract spiked with bovine trypsin. The MS/MS

data collected was searched for the expected adducts at S200 of bovine trypsin and 3-bromothiophenol labelling was again confirmed at this residue. The high-mass accuracy of the MS/MS fragment ions combined with the isotopic splitting pattern of the parent ion (Figure 6D) provide sufficient confirmation of probe-labelling at this residue.

We hypothesised that since SFABP displacement by 3-bromothiophenol was observed to be effective, this approach could in principle be extended to analyse the targets of AEBSF directly without the need for an ABP. Thus, we repeated the labelling experiments with bovine trypsin using AEBSF and confirmed 3-bromothiophenol incorporation at S200. Interestingly, we also observed on the same peptide sequence intact AEBSF labelling of S195, which was not displaced by 3-bromothiophenol. When repeating AEBSF treatment in HeLa protein extract, S200 labelling of trypsin was again confirmed and we also identified 3-bromothiophenol displacement at S283 of a novel protein target (T2R14). This demonstrates that this approach can be applied directly using the original small molecule inhibitor and that residue-specific identification can be achieved in complex backgrounds without the need for pull-down or enrichment of the probe-bound peptides. Furthermore, this workflow can be combined with existing software such as IsoStamp[14] that identify bromine-containing peptides based on their isotopic signature.

The Power and Challenges of Unconstrained Sequencing

The identification of the novel sites of SFABP labelling and the unexpected adducts resulting from probe displacement at the active site serine of trypsin were made possible by the use of unconstrained sequencing. The overall premise for using unconstrained sequencing is that mass spectrometry is an unbiased analytical tool. Therefore, if an adduct is present it should be detectable, even if the resulting composition or mass of the corresponding adduct is not correctly predicted a priori. The iterative convergence of our unconstrained sequencing approach is currently dependent upon user intervention at the outer loop of the algorithm (see Supplemental Figure 1), which is required to account for competing degrees of freedom

in the search (variable position and mass) and to correctly distinguish true probe adducts from false positive matches. For example, the following adducts were also identified in our data using unconstrained sequencing.

Side-products of protein alkylation. When using iodoacetamide to alkylate cysteine residues, carbamidomethylation was found to occur at serine and lysine residues (+57 Da), which is in agreement with the literature.[25,26] Interestingly, when substituting iodoacetamide for N-ethylmaleimide (NEM) for protein alkylation, NEM-labelling of serine residues (+125 Da) was also identified.

Cysteine DTT-IA labelling. A significant number of cysteines were observed to contain an adduct of observed mass +209 Da in both the control and probe-treated experiments, which consists of a sulfur-sulfur bond between a cysteine residue and a DTT molecule that has been capped by IA at the other thiol moiety. It should be noted that such adducts have previously been reported on cysteine residues using error-tolerant database searches.[27] Importantly, this modification was also observed on a tryptic peptide with an N-terminal cysteine, C160, in combination with apparent carbamidomethylation of an amine group proximal to the N-terminus, resulting in a collective adduct mass of +266 Da. Since this same peptide sequence was observed to also contain either the expected SFABP adduct (+277 Da) on a neighboring lysine (K162) or just a DTT-IA adduct on C160 (+209 Da), it highlights the requirement for manual interpretation of the search results to accurately converge the unconstrained sequencing method and distinguish true probe adducts from false positive matches.

Conclusion

Sulfonyl fluorides such as AEBSF or PMSF are routinely used in cell-based assays as inhibitors for serine proteases, which can cause significant degradation of the proteome during lysis.[28] The SFABP used in this work was designed to reflect the reactivity of AEBSF (Figure 1A). Our findings demonstrate that this probe does not function as a true ABP since it

does not form a stable adduct with trypsin's active-site serine. Thus, to rigorously characterise the proteome targets of this class of inhibitors, alternative warheads that result in the formation of a stable bond between the ABP and its target serine residues will need to be developed; recent work involving fluorosulfates seems to be promising in that regard, even if stability issues are still latent.[29] Alternatively, chemical proteomic strategies that exploit the stability of the ABP with its target residue, as described in this work, provide a basis for distinguishing between transient and stable interactions and are directly compatible with the inhibitor of interest. Interestingly, the SF probes developed and used by Mukherjee et al. were found to be unsuitable for the durable covalent inhibition of cysteine residues due to the observed instability of the resulting adduct.[9] In contrast, lysine and tyrosine residues reacted with the same SF probes to form adducts that were stable with respect to hydrolysis, highlighting the need to improve the hydrolytic stability of the ABP/peptide adducts.

We envisage that further advancements in mass spectrometry-based identification tools such as unconstrained sequencing will continue to push the frontiers in our understanding of protein/drug interactions. Current algorithms are not able to fully accommodate both variable mass adducts and variable sites of modification as this significantly reduces the power of scoring functions, thus requiring manual inspection at some stage of the search. As the knowledge base of chemical modifications continues to grow thorough rigorous interpretation of MS/MS data, the availability of resources such as the SFABP database will provide a foundation for the development of automated software for unconstrained sequencing based on techniques such as Bayesian inference.

Methods

Standard protein labelling with SFABP

To a mixture of 42.5 μL PBS and 2.5 μL SFABP stock (10 mM in DMSO, final concentration 500 μM) was added 5.0 μL of protein stock (bovine trypsin, bovine chymotrypsin:

2.5 mg mL^{-1} in 1 mM HCl; equine GST mixture: 2.5 mg mL^{-1} in PBS). The labelling was allowed to proceed at room temperature, in the dark, for 45 min. After that time, a MeOH/$CHCl_3$ precipitation was carried out, and the residue was re-suspended in 50 μL PBS. A standard in-solution digestion was carried out, followed by desalting of the sample and LC-MS/MS analysis. Labelling experiments using AEBSF were also performed following this protocol.

MeOH/$CHCl_3$ protein precipitation

The organic solvents used in this protocol were stored at -20 °C when not in use. To a 50 μL sample of protein in aqueous buffer was added 200 μL of MeOH, followed by 50 μL $CHCl_3$. The sample was vortexed and 150 μL water was added to it, causing separation of the organic and aqueous phases. The sample was vortexed again and then centrifuged at 18 kg for 10 min. The top layer was removed without perturbing the protein sitting at the interface, and an additional 200 μL MeOH was added to the sample. Centrifugation at 18 kg for 10 min, followed by removal of the supernatant by vacuum centrifugation afforded the purified protein sample.

Protein digestion

Reduction using DTT and alkylation using iodoacetamide

To a protein sample in 50 μL or 100 μL PBS (or suitable aqueous buffer, with a $7.0 < pH < 8.5$) was added 10 μL of DTT stock (30 or 55 mM in PBS). After incubating the sample at 51 °C for 1 hour, it was allowed to cool down. 10 μL of iodoacetamide stock (100 or 168 mM in water) was then added to the sample. The reaction was allowed to proceed at room temperature, in the dark, for 45 min. The pH was then adjusted to 8 if necessary, and to the protein mixture was added sequencing-grade trypsin to a final protease/protein ratio of 1/10 (w/w). The sample was incubated at 37 °C for 6 hours. A standard desalting

procedure was then followed.

Reduction using TCEP and alkylation using NEM

In some experiments, TCEP and NEM were used as substitutes for DTT and IA, respectively, in the in-solution digestion procedure. In those cases, the standard in-solution digestion protocol reported above was followed, but replacing DTT stock (30 mM in PBS) with TCEP stock (30 mM in water), and the IA stock (100 mM in water) with NEM stock (100 mM in water).

Peptide desalting

Standard procedure:

To the peptide sample was added 200 μL of the acid wash (99.9% OH_2, 0.1% AcOH). The sample was then loaded on tips packed with C18 solid phase extraction (3M Empore, UK) that had previously been washed and equilibrated with 200 μL of MeOH followed by 200 μL of the acid wash, respectively. The sample was then washed with 200 μL of the acid wash, and eluted with 100 μL of the elution buffer (75% MeCN, 20% water, 5% AcOH). The desalted peptides were then dried down to completion and stored at -80 °C prior to LC-MS/MS analysis.

^{18}O desalting procedure:

The standard procedure reported above was followed, but the acid wash (99.9% OH_2, 0.1% AcOH) and elution buffer (75% MeCN, 20% water, 5% AcOH) were prepared with $^{18}OH_2$.

SFABP labelling of denatured proteins

In two separate Eppendorf tubes, 5 μL of bovine trypsin stock (2.5 mg mL^{-1} in 1 mM HCl) were added to 42.5 μL PBS. The control aliquot was left standing at room temperature

for 20 min, whilst the treated aliquot was heated to 90 °C for 20 minutes to allow protein denaturation, as per the protocol reported by Park and Russell.[30] Both samples were then equilibrated back to room temperature and 2.5 μL of SFABP stock (10 mM stock in DMSO) was added to both treated and control samples, and the labelling was allowed to proceed for 45 minutes. After that time, a MeOH/$CHCl_3$ precipitation was carried out, and the residue was re-suspended in 50 μL PBS. Both samples were then digested, desalted and analysed by LC-MS/MS separately, following the protocols above.

SFABP labelling of HeLa nuclear extracts spiked with bovine trypsin

A nuclear lysate was extracted from HeLa cells following the protocol reported by Dignam et al..[31] The buffer composition of the lysate was of 20 mM HEPES pH 7.9, 25% glycerol, 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, with a reported protein concentration of 8.5 mg mL^{-1} . To 120.5 μL PBS were added 20 μL of the aforementioned nuclear extract, 2.0 μL of the bovine trypsin stock (2.5 mg mL^{-1} in 1 mM HCl) and 7.5 μL of SFABP stock (10 mM in DMSO). The final SFABP concentration (500 μM) is identical to that used in our standard labelling experiments. The labelling was allowed to proceed for 45 minutes. After that time, a MeOH/$CHCl_3$ precipitation was carried out, and the residue was re-suspended in 50 μL PBS. A standard in-solution digestion was carried out, followed by desalting of the sample.

SFABP labelling in heavy water

Preparation of heavy PBS

Heavy PBS was prepared following the salt concentrations of the Cold Spring Harbor Protocol.[32] In an Eppendorf tube were dissolved 4 mg NaCl, 0.1 mg KCl, 0.7 mg Na_2HPO_4 and 0.1 mg KH_2PO_4 in 500 μL $^{18}OH_2$ (97% isotopic purity). The solution was used as is in the relevant experiments.

^{18}O labelling of trypsin

To a mixture of 42.5 μL PBS and 2.5 μL SFABP stock (10 mM in DMSO) was added 5.0 μL of bovine trypsin stock (2.5 mg mL^{-1} in 1 mM HCl). The labelling was allowed to proceed at room temperature, in the dark, for 45 min. After that time, a MeOH/$CHCl_3$ precipitation was carried out, and the residue was re-suspended in 50 μL heavy PBS. The protein sample was incubated at 37 °C for 6 hours, time after which 10 μL of DTT stock (5 mM in heavy PBS) was added. The sample was incubated at 51 °C for 1 hour and then allowed to cool down. 10 μL of iodoacetamide stock (15 mM in $^{18}OH_2$) was added and the reaction was allowed to proceed at room temperature, in the dark, for 45 min. After that time, the sample was dried down to completion and re-suspended in 50 μL H_2O (i.e. 'light' water). pH was adjusted to 8, and to the protein sample was added 1.25 μg of sequencing grade Promega trypsin. The sample was incubated at 37 °C for 6 hours. A standard desalting procedure was then followed.

Thiophenol labelling

Of purified trypsin:

To a mixture of 42.5 μL PBS and 2.5 μL SFABP stock (10 mM in DMSO) was added 5.0 μL of bovine trypsin stock (2.5 mg mL^{-1} in 1 mM HCl). The labelling was allowed to proceed at room temperature, in the dark, for 45 min. After that time, a MeOH/$CHCl_3$ precipitation was carried out, and the residue was re-suspended in 50 μL of a 100 mM solution of 3-bromothiophenol in MeOH. The reaction was allowed to proceed with vortexing at room temperature, for 30 minutes. After that time, the sample was dried down to completion, re-suspended in 50 μL PBS and sonicated for 10 min. The sample was centrifuged at 5000 g for 30 seconds, and the supernatant was collected. Another 50 μL PBS was added to the pellet, which was sonicated for 10 min and centrifuged at 5000 g for 30 seconds. The supernatant was pooled with the previous one, and the combined supernatant fractions were

brought forward for protein digestion. The pellet containing water-insoluble thiophenol was discarded. The protein sample was made up to 100 μL with PBS, digested, desalted and analysed by LC-MS/MS following the standard protocols.

Of HeLa extracts spiked with purified trypsin:

To 141 μL HeLa cell extract (containing 0.55 g/L protein) were added 1.5 μL of the bovine trypsin stock (2.5 mg mL^{-1} in 1 mM HCl) and 7.5 μL of SFABP or AEBSF stock (10 mM in DMSO). The labelling was allowed to proceed at room temperature for 45 minutes. After that time, a MeOH/$CHCl_3$ precipitation was carried out, and the residue was re-suspended in 50 μL of a solution of 3-bromothiophenol (1 M in 1/1 MeOH/MeCN). The sample was allowed to react at 37 °C for 1 hour, and was further vortexed at room temperature for 1 hour. After that time, the sample was dried down to completion, re-suspended in 100 μL H_2O and excess probe was removed by MeOH/$CHCl_3$ precipitation of the protein. The protein pellet was further washed with ice-cold MeOH (2 x 1.5 mL), dried and re-suspended in 50 μL PBS. The sample then processed for protein digestion.

Liquid chromatography tandem mass spectrometry - LTQ Velos Pro Ion Trap

Approximately 5 μg of sample reconstituted in 0.5% MeCN and 0.1% TFA was loaded via autosampler onto an Acclaim PepMap100 C18 column (3 μm particle size, 100 Å pore size, 75 μm internal diameter, 150 mm length; Thermo Fisher). The peptides were chromatographically resolved using an Ultimate 3000 RSLCnano System (Dionex) with a flow rate of 300 nL/min and a linear gradient starting at 1% B (95% MeCN, 0.1% FA) and 99% A (0.1% FA in H_2O) increased to 50% B over 88 min. Real time mass spectra were acquired on an LTQ Velos Pro linear ion trap (Thermo Scientific) over a 300-2000 m/z scan range. CID tandem mass spectra (35% collision energy) were collected using a top 6-10 data-dependent acquisition with a dynamic exclusion list (repeat count of 2, repeat duration of 15 s, exclusion

list size 100 and exclusion duration of 40 s).

Liquid chromatography tandem mass spectrometry - LTQ Velos Pro OrbiTrap

Approximately 5 μg of sample reconstituted in 2% MeCN and 0.05% TFA was loaded via autosampler onto an EASY-Spray C18 column (2 μm particle size, 100 Å pore size, 75 μm internal diameter, 50 cm length; Thermo Fisher). The peptides were chromatographically resolved using an Ultimate 3000 RSLCnano System (Dionex) with a flow rate of 275 nL/min and a linear gradient starting at 8% B (75% MeCN, 5% DMSO, 0.1% FA) and 92% A (5% DMSO, 0.1% FA in H_2O) increased to 45% B over 105 min. Real time mass spectra were acquired on an LTQ Velos Pro Orbitrap (Thermo Scientific) over a 350-1700 m/z scan range at a resolution of 60k. Ion trap CID tandem mass spectra (35% collision energy) were collected using a top 10 data-dependent acquisition with a dynamic exclusion list (repeat count of 1, repeat duration of 10s, exclusion list size 200 and exclusion duration of 30 s).

Iterative approach to unconstrained sequencing

Unconstrained sequencing was iteratively performed according to the flowchart presented in Supplementary Figure 1. Starting from an expected elemental composition for a peptide modification of interest (variable 1) and a candidate target residue (variable 2), the program searches all tandem MS within a large tolerance of error around the expected precursor m/z, which allows for the possibility of significant and unexpected changes to adduct composition. These tandem MS are then sorted based on a scoring function that considers the intensity and consecutive matches between observed and expected fragment ions. For each partially annotated spectrum, the user will manually assess the quality of the fragment ion matches in comparison to the annotated tandem MS for the unmodified peptide sequence (including elimination of false positives based on enzymatic missed cleavages). If the difference between

the observed and expected precursor mass is less than some tolerance (0.5 Da for ion trap and 10 ppm for Orbitrap), then the expected elemental composition of the modification (variable 1) is determined to be correct and the user will iteratively modify the position of the target residue (variable 2) to maximise the number of matches between observed and expected fragment ions in the tandem MS. Else if the difference between the observed and expected precursor mass is larger than the tolerance, then the user will modify the elemental composition of the modification (variable 1) based on the observed mass difference and repeat the aforementioned search procedure.

Acknowledgement

This work was supported by an EPSRC-funded Institute of Chemical Biology CDT studentship (Award Reference 1372421). The authors thank E. Tate and S. Lovell for providing the SFABP, which was synthesised according to the original protocol.[5] The authors are very grateful to B. Snijders and A. Borg for their assistance in performing measurements with a LTQ Velos Pro Orbitrap mass spectrometer. The authors are also very grateful to S. Lukauskas for his help in setting up the SFABP database.

Supporting Information Available

The following files are available free of charge.

- Supplementary Material 1: PDF document containing the calculations for determining the extent of ^{18}O incorporation into the tryptic peptide containing S200.
- Supplementary Material 2: Microsoft Excel spreadsheet (.xlsx) containing: (A) the raw MS/MS data for Figure 5B; (B) the processed MS/MS data for Figure 5B; (C) the MS/MS data used to compute the average fragment ion height corresponding to

a loss of H_2O; (D) the calculations for the relative contributions of the 5-*exo*-tet and lactonisation mechanisms to hydrolysis of the SFABP/S200 adduct.

- Supplemental Figures: PDF document containing algorithmic flow chat for unconstrained sequencing.

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

1. Bunnage, M. E., Gilbert, A. M., Jones, L. H., and Hett, E. C. (2015) Know your target, know your molecule. *Nat. Chem. Biol. 11*, 368–372.

2. Bunnage, M. E. (2011) Getting pharmaceutical R&D back on target. *Nat. Chem. Biol. 7*, 335–339.

3. Klaeger, S., Gohlke, B., Perrin, J., Gupta, V., Heinzlmeir, S., Helm, D., Qiao, H., Bergamini, G., Handa, H., Savitski, M., Bantscheff, M., Médard, G., Preissner, R., and Kuster, B. (2016) Chemical Proteomics Reveals Ferrochelatase as a Common Off-target of Kinase Inhibitors. *ACS Chem. Biol. 11*, 1245–1254.

4. Sanman, L., and Bogyo, M. (2014) Activity-Based Profiling of Proteases. *Annu. Rev. Biochem. 83*, 249–273.

5. Shannon, D., Gu, C., McLaughlin, C., Kaiser, M., van der Hoorn, R., and Weerapana, E. (2012) Sulfonyl Fluoride Analogues as Activity-Based Probes for Serine Proteases. *ChemBioChem 13*, 2327–2330.

6. Narayanan, A., and Jones, L. (2015) Sulfonyl fluorides as privileged warheads in chemical biology. *Chem. Sci. 6*, 2650–2659.

7. Zhao, Q., Ouyang, X., Wan, X., Gajiwala, K., Kath, J., Jones, L., Burlingame, A., and Taunton, J. (2017) Broad-Spectrum Kinase Profiling in Live Cells with Lysine-Targeted Sulfonyl Fluoride Probes. *J. Am. Chem. Soc. 139*, 680–685.

8. Gu, C., Shannon, D., Colby, T., Wang, Z., Shabab, M., Kumari, S., Villamor, J., McLaughlin, C., Weerapana, E., Kaiser, M., Cravatt, B., and van der Hoorn, R. (2013) Chemical Proteomics with Sulfonyl Fluoride Probes Reveals Selective Labeling of Functional Tyrosines in Glutathione Transferases. *Chem. Biol. 20*, 541–548.

9. Mukherjee, H., Debreczeni, J., Breed, J., Tentarelli, S., Aquila, B., Dowling, J., Whitty, A., and Grimster, N. (2017) A study of the reactivity of S(VI)-F containing warheads with nucleophilic amino-acid side chains under physiological conditions. *Org. Biomol. Chem. 15*, 9685–9695.

10. Lohse, J., Schindl, A., Danda, N., Williams, C., Kramer, K., Kuster, B., Witte, M., and Medard, G. (2017) Target and identify: triazene linker helps identify azidation sites of labelled proteins via click and cleave strategy. *Chem. Commun. 53*, 11929–11932.

11. Yang, J., Gupta, V., Carroll, K., and Liebler, D. (2014) Site-specific mapping and quantification of protein S-sulphenylation in cells. *Nat. Commun. 5*, 4776.

12. Kim, H.-Y., Tallman, K., Liebler, D., and Porter, N. (2009) An Azido-Biotin Reagent for Use in the Isolation of Protein Adducts of Lipid-derived Electrophiles by Streptavidin Catch and Photorelease. *Mol. Cell. Proteomics 8*, 2080–2089.

13. Yip, G., Chen, Z.-W., Edge, C., Smith, E., Dickinson, R., Hohenester, E., Townsend, R., Fuchs, K., Sieghart, W., Evers, A., and Franks, N. (2013) A propofol binding site on mammalian GABAA receptors identified by photolabeling. *Nat. Chem. Biol. 9*, 715–720.

14. Palaniappan, K., Pitcher, A., Smart, B., Spiciarich, D., Iavarone, A., and Bertozzi, C. (2011) Isotopic Signature Transfer and Mass Pattern Prediction (IsoStamp): An Enabling Technique for Chemically-Directed Proteomics. *ACS Chem. Biol. 6*, 829–836.

15. Woo, C., Iavarone, A., Spiciarich, D., Palaniappan, K., and Bertozzi, C. (2015) Isotope-targeted glycoproteomics (IsoTaG): a mass-independent platform for intact N- and O-glycopeptide discovery and analysis. *Nat. Methods 12*, 561–567.

16. Chick, J., Kolippakkam, D., Nusinow, D., Zhai, B., Rad, R., Huttlin, E., and Gygi, S. (2015) A mass-tolerant database search identifies a large proportion of unassigned spectra in shotgun proteomics as modified peptides. *Nat. Biotechnol. 33*, 743–749.

17. Balogh, L., Trong, I. L., Kripps, K., Tars, K., Stenkamp, R., Mannervik, B., and Atkins, W. (2009) Structural Analysis of a Glutathione Transferase A1-1 Mutant Tailored for High Catalytic Efficiency with Toxic Alkenals. *Biochemistry 48*, 7698–7704.

18. Weiner, H., White, W., Hoare, D., and Koshland, D. (1966) The Formation of Anhydrochymotrypsin by Removing the Elements of Water from the Serine at the Active Site. *J. Am. Chem. Soc. 88*, 3851–3859.

19. Neet, K., and Koshland, D. (1966) The conversion of serine at the active site of subtilisin to cysteine: a "chemical mutation". *Proc. Natl. Acad. Sci. U.S.A. 56*, 1606–1611.

20. Strumeyer, D., White, W., and Koshland, D. (1963) Role of serine in chymotrypsin action. Conversion of the active serine to dehydroalanine. *Proc. Natl. Acad. Sci. U.S.A. 50*, 931–935.

21. Miyagi, M., and Rao, K. (2007) Proteolytic 18O-labeling strategies for quantitative proteomics. *Mass Spectrom. Rev. 26*, 121–136.

22. Gold, A., and Fahrney, D. (1964) Sulfonyl Fluorides as Inhibitors of Esterases. II. Formation and Reactions of Phenylmethanesulfonyl α-Chymotrypsin. *Biochemistry 3*, 783–791.

23. Harrison, A., and Tu, Y.-P. (1998) Ion chemistry of protonated aspartic acid derivatives. *J. Mass Spectrom. 33*, 532–542.

24. Tsukiji, S., Miyagawa, M., Takaoka, Y., Tamura, T., and Hamachi, I. (2009) Ligand-directed tosyl chemistry for protein labeling in vivo. *Nat. Chem. Biol. 5*, 341–343.

25. Muller, T., and Winter, D. (2017) Systematic Evaluation of Protein Reduction and Alkylation Reveals Massive Unspecific Side Effects by Iodine-containing Reagents. *Mol. Cell. Proteomics 16*, 1173–1187.

26. Wilhelm, M. et al. (2014) Mass-spectrometry-based draft of the human proteome. *Nature 509*, 582–587.

27. Chalkley, R., Baker, P., Medzihradszky, K., Lynn, A., and Burlingame, A. (2008) In-depth Analysis of Tandem Mass Spectrometry Data from Disparate Instrument Types. *Mol. Cell. Proteomics 7*, 2386–2398.

28. Powers, J., Asgian, J., Ekici, Ö., and James, K. (2002) Irreversible Inhibitors of Serine, Cysteine, and Threonine Proteases. *Chem. Rev. 102*, 4639–4750.

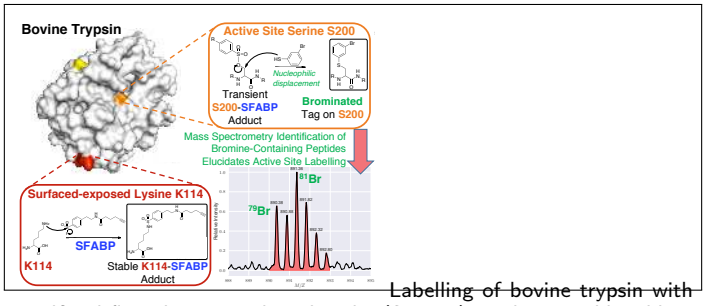
29. Fadeyi, O., Hoth, L., Choi, C., Feng, X., Gopalsamy, A., Hett, E., Kyne, R., Robinson, R., and Jones, L. (2017) Covalent Enzyme Inhibition through Fluorosulfate Modification of a Noncatalytic Serine Residue. *ACS Chem. Biol. 12*, 2015–2020.

30. Park, Z.-Y., and Russell, D. (2000) Thermal Denaturation: A Useful Technique in Peptide Mass Mapping. *Anal. Chem. 72*, 2667–2670.

31. Dignam, J., Lebovitz, R., and Roeder, R. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res. 11*, 1475–1489.

32. (2006) Phosphate-buffered saline (PBS). *Cold Spring Harb. Protoc.*

Graphical TOC Entry



Labelling of bovine trypsin with a sulfonyl fluoride activity-based probe (SFABP) results in stable adduct formation with lysine residues (K114 shown here) but is displaced from the active site serine (S200) by hydrolysis. A chemical proteomics approach levering this nucleophilic displacement was developed to label the transient interactions between SFABP and the active site with a bromine-tag, which facilitates mass spectrometry identification of sites of probe binding.