Supporting Information

Thomas E. J. Chavas, Matthew J. Fuchter, Peter A. DiMaggio Jr, "Unbiased Mass Spectrometry Elucidation of the Targets and Mechanisms of Activity-Based Probes: A Case Study involving Sulfonyl Fluorides"

Supplementary Material 1: Calculations for determining the extent of ^{18}O incorporation into the tryptic peptide containing S200

Supplementary Figure 1: Workflow for unconstrained sequencing based on variable probe mass (variable 1) and variable protein target residues (variable 2).

Supplementary Material 2: see Excel file (Supplementary_Material_2_DiMaggio.xlsx)

Supplementary Material 1. Computing the yield of ^{18}O incorporation into the S200 peptide

The isotopic distributions of the intact and ^{18}O-labelled peptides are predicted from their respective elemental composition; those are shown in Figure S1.1 and Figure S1.2, respectively. The isotopic distribution observed in the treated experiment (shown in Figure 3A of the main article) can be viewed as a superposition of the two theoretical isotopic distributions (labelled and unlabelled). Thus, computing the relative contribution of each that is required to match the observed distribution is one way to establish the ratio of ^{18}O-labelled versus naturally-occurring species in the experiment.

Elemental Composition: **C61 H101 N20 O27 S2**
Monoisotopic M/Z: **805.33269**
Total Abundance: **100.00%**

Isotope Number	m/z	Percent Total	Percent Maximum
0	805.33269	39.26	100.00
1	805.83407	31.31	79.74
2	806.33443	17.97	45.76
3	806.83494	7.66	19.52
4	807.33538	2.70	6.89
5	807.83587	0.81	2.07
6	808.33640	0.21	0.55
7	808.83698	0.05	0.13
8	809.33762	0.01	0.03
9	809.83831	0.00	0.00
10	810.33899	0.00	0.00

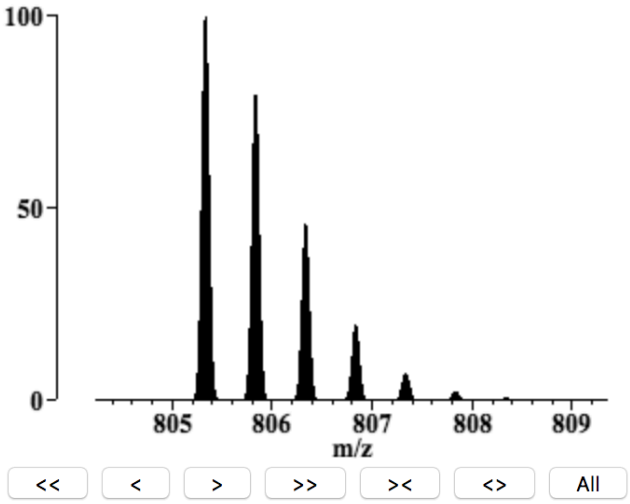

Figure S1.1 – Predicted isotopic distribution of the naturally-occurring S200-containing peptide (unlabelled).

Elemental Composition: **C61 H101 N20 O26 S2 18O1**
Monoisotopic M/Z: **806.33481**
Total Abundance: **100.00%**

Isotope Number	m/z	Percent Total	Percent Maximum
0	806.33481	39.36	100.00
1	806.83619	31.37	79.70
2	807.33655	17.92	45.53
3	807.83705	7.61	19.34
4	808.33749	2.67	6.79
5	808.83797	0.80	2.03
6	809.33850	0.21	0.53
7	809.83907	0.05	0.12
8	810.33969	0.01	0.03
9	810.84034	0.00	0.00
10	811.34110	0.00	0.00

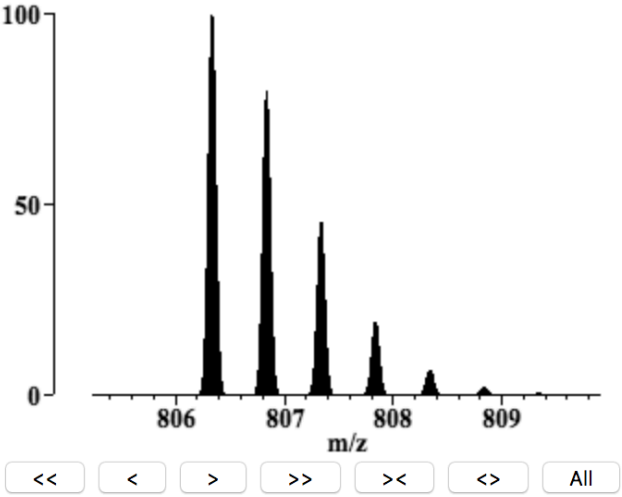

Figure S1.2 – Predicted isotopic distribution of the S200-containing peptide labelled with one ^{18}O atom.

The linear regression problem at hand is essentially to work out the weighting coefficients w_1 and w_2 that satisfy Equation S1.1, below; where d corresponds to the predicted or observed isotopic distributions of the ^{18}O-labelled (treated) or naturally-occurring (^{16}O) peptides.

$$w_1 \times d_{16O,theoretical} + w_2 \times d_{18O,theoretical} = d_{18O,observed}$$

Equation S1.1 – Computing the contributions of the naturally-occurring and ^{18}O-containing peptides to the observed isotopic distribution.

As an approximation for the distributions, the intensity of each of the +0, +1, +2, +3, +4 and +5 Da isotopic peaks are recorded and normalised to the highest intensity peak in that distribution. This allows us to set up the equation and compute coefficients w_1 and w_2. Those are essentially a measure of the relative amounts of the intact versus ^{18}O-labelled species giving rise to the isotopic distribution observed.

Thus, the intensity of the 6 most intense peaks in the five MS^1 spectra closest to the chromatographic maximum were sampled, for both the treated and control experiments. The intensities were normalised to the highest intensity peak in each spectrum (805.36 *m/z* for the control, 806.36 *m/z* for the treated) and the ratios thus obtained were used to determine the weighting coefficients w_1 and w_2 for each spectrum.

The results are presented in Table S1.1 and Table S1.2, below.

Table S1.1 – Computing the w_1 and w_2 coefficients for the MS1 trace obtained in the control experiment (shown in Figure 3B).

Trypsin_control _04Oct16	Ratio of peak height over highest intensity peak at retention time						
M/Z	20.09 min	20.14 min	20.18 min (chromatographic apex)	20.23 min	20.27 min		
805.345	1.000	1.000	1.000	1.000	1.000		
805.845	0.921	0.818	0.758	0.838	0.926		
806.345	0.516	0.510	0.448	0.484	0.559		
806.845	0.219	0.217	0.199	0.228	0.271		
807.345	0.091	0.080	0.072	0.083	0.096		
807.845	0.029	0.029	0.027	0.026	0.031		
Explained Variance	0.990	0.999	0.999	0.999	0.988		
r^2	0.990	0.999	0.999	0.999	0.988		
Mean Squared Error	0.002	0.000	0.000	0.000	0.002	**AVERAGE**	**STANDARD DEVIATION**
Predicted naturally-occurring peptide contribution	0.990	0.961	0.998	0.983	0.941	**0.975**	**0.0230**
Predicted 18O-peptide contribution	0.010	0.039	0.002	0.017	0.059	**0.025**	**0.0230**

Table S1.2 – Computing the w_1 and w_2 coefficients for the MS1 trace obtained in the treated experiment

Trypsin_treated_23Sep16	Ratio of peak height over highest intensity peak at retention time						
M/Z	20.70 min	20.75 min (chromatographic apex)	20.79 min	20.84 min	20.89 min		
805.345	0.277	0.263	0.283	0.278	0.304		
805.845	0.241	0.227	0.230	0.271	0.267		
806.345	1.000	1.000	1.000	1.000	1.000		
806.845	0.773	0.752	0.693	0.752	0.725		
807.345	0.429	0.427	0.427	0.378	0.384		
807.845	0.193	0.205	0.175	0.211	0.171		
Explained Variance	0.999	1.000	0.995	0.992	0.998		
r^2	0.999	1.000	0.995	0.992	0.998		
Mean Squared Error	0.000	0.000	0.000	0.001	0.000	**AVERAGE**	**STANDARD DEVIATION**
Predicted naturally-occurring peptide contribution (w_1)	0.232	0.220	0.251	0.253	0.280	**0.247**	**0.0228**
Predicted 18O-peptide contribution (w_2)	0.768	0.780	0.749	0.747	0.720	**0.753**	**0.0228**

In the control sample, the model predicts 2.5% incorporation of exogenous ^{18}O atom for this peptide, with a standard deviation of 2.3%. This is in line with the fact that, being a control experiment, this peptide sample should not contain any exogenous ^{18}O atom.

Furthermore, the model predicts that a 75% incorporation of exogenous ^{18}O atom for this peptide in the treated sample, with a standard deviation of 2.3%.

Supplementary Figure 1: Workflow for unconstrained sequencing based on variable probe mass (variable 1) and variable protein target residues (variable 2).