Wang et al. "Arginyltransferase is an ATP-independent self-regulating enzyme that forms distinct functional complexes in vivo." Supplemental Information.

Supplemental Figure Legends.

Supplemental Figure S1. a. Coomassie stained gel of the typical preparations of ATE1 isoforms 1-4 (lanes 1-4) and E. coli RRS (lane 5) used in the in vitro assays. b. Time course of the incorporation of [^{3}H]-Arg into BSA mediated by ATE1-2 (blue diamonds); control curves show arginylation reaction with no added ATE1 (pink squares), no RRS (red triangles), and no BSA substrate (green crosses). c. RRS-mediated incorporation of [^{3}H]-Arg into tRNA (red squares) during arginylation reaction. In the absence of heating, the ester bond between [^{3}H]-Arg and tRNA is not destroyed, producing a significant signal during scintillation counting, estimated to be ~10 times higher than the signal from the arginylated substrate (dark blue diamonds).

Supplemental Figures S2. Mass spectra of the modified peptides identified in BSA. The mass spectra in panels A-D illustrate the sequence (top) and the fragmentation pattern (bottom) of the arginylated BSA-derived peptides listed in the main text. See Supplemental Table 2 for the sequences and parameters of the identified peptides.

Supplemental Figure S3. a. Comparison of 2D-autoradiographs of proteins arginylated by each ATE1 isoform (marked 1-1 through 1-4, respectively) in the *Ate1* KO cell extract. Each comparison was made pairwise between gels shown in the top row and directly underneath it. See Supplemental Table 1 for spot markings and fold change. b. Coomassie-stained gel (left) and

Western blot (right) of the four ATE1 isoforms purified by Ni-NTA affinity chromatography after expression in *E.coli* and probed with rat monoclonal anti-ATE1. Faint bands of lower molecular weight are present in all four preparations and represent ATE1 proteolytic fragments, as evidenced from the Western blot analysis with the ATE1 antibodies.

Supplemental Figure S4. Left, autoradiographs of the gels on right, showing incorporation of [^{14}C]-Arg into α-lactalbumin by each of the ATE1 isoforms (labeled 1-4 on top). C-control lane showing arginylation reaction without the addition of ATE1. Addition of small doses of cell extract enhances the ability of ATE1-3/4 to arginylate α-lactalbumin, however larger amounts of cell extracts have inhibitory effects, presumably due to the competition of α-lactalbumin with other substrates found in the cell extracts. See also Fig. 4 in the main text.

Supplemental Figure S5. Mass spectra of the modified peptides identified in ATE1 isoforms 1-4 (a-d, respectively). The mass spectra in each panel illustrate the sequence (top) and the fragmentation pattern (bottom) of the arginylated ATE1-derived peptides listed in the main text. See Supplemental Table 2 for the sequences and parameters of the identified peptides.

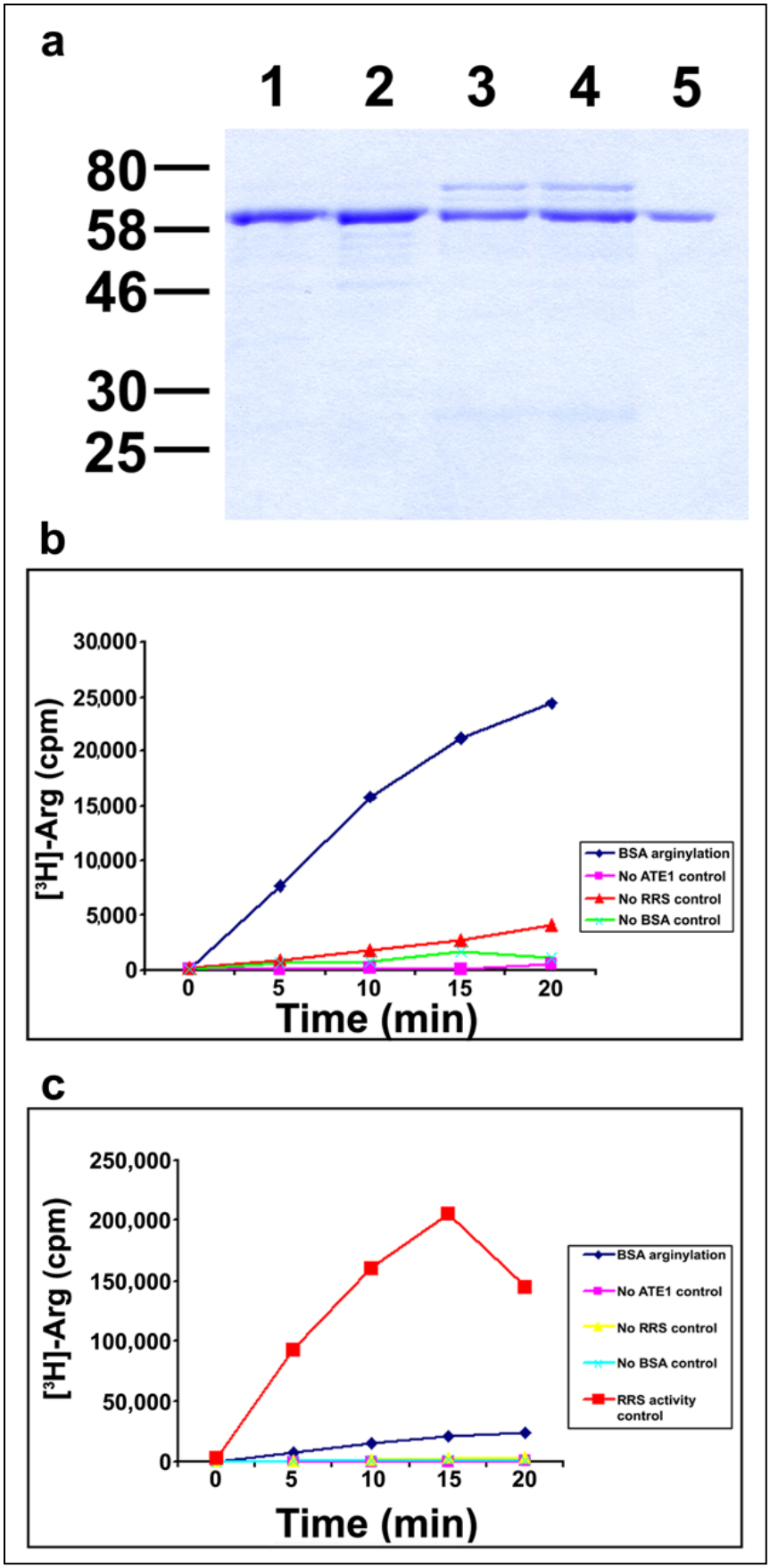
Supplemental Figure S6. a. Quantifiaction of the distribution of different intracellular markers in the fractions of the sucrose gradient shown in Fig. 6a of the main text. Numbers represent overall intensity distribution for each marker as a fraction of the total intensity of this marker in all fractions (taken as 1). Three independent runs quantified in triplicates were performed to obtain the average numbers and the error bars (SEM) shown in the figure. b.

Western blots of the fractions of liver extracts fractionated by differential centrifugation followed by 5-20% sucrose gradient, probed with antibodies to ATE1, translation initiation factor eIF1B2, dynein intermediate chain (DIC), and tubulin. Lanes labeled S1-3 and P1-3 represent supernatants and pellets from 1,500g, 16,000g, and 67,917g centrifugation steps, respectively. S3 was loaded onto the linear 5-20% sucrose gradient. Lanes labeled 1-20 represent sucrose gradient fractions.

Supplemental Table 1. Summary of Results for 2D Autoradiographic Film comparison showing *Ate1* knockout mouse cell extracts labeled with ^{14}C-Arg in the presence of four ATE1 isoforms. Reference spot numbering, pI, and MW, and fold change are shown for polypeptide spots that were differentially labeled in the four gels shown in Fig. S3. The differences are calculated from spot percentages (individual spot density divided by total density of all measured spots) as described in the Materials and Methods. Polypeptide spots increased by > 2.0 are highlighted in blue in the difference column, while spots with a fold decrease of ≤ -2.0 are highlighted in red.

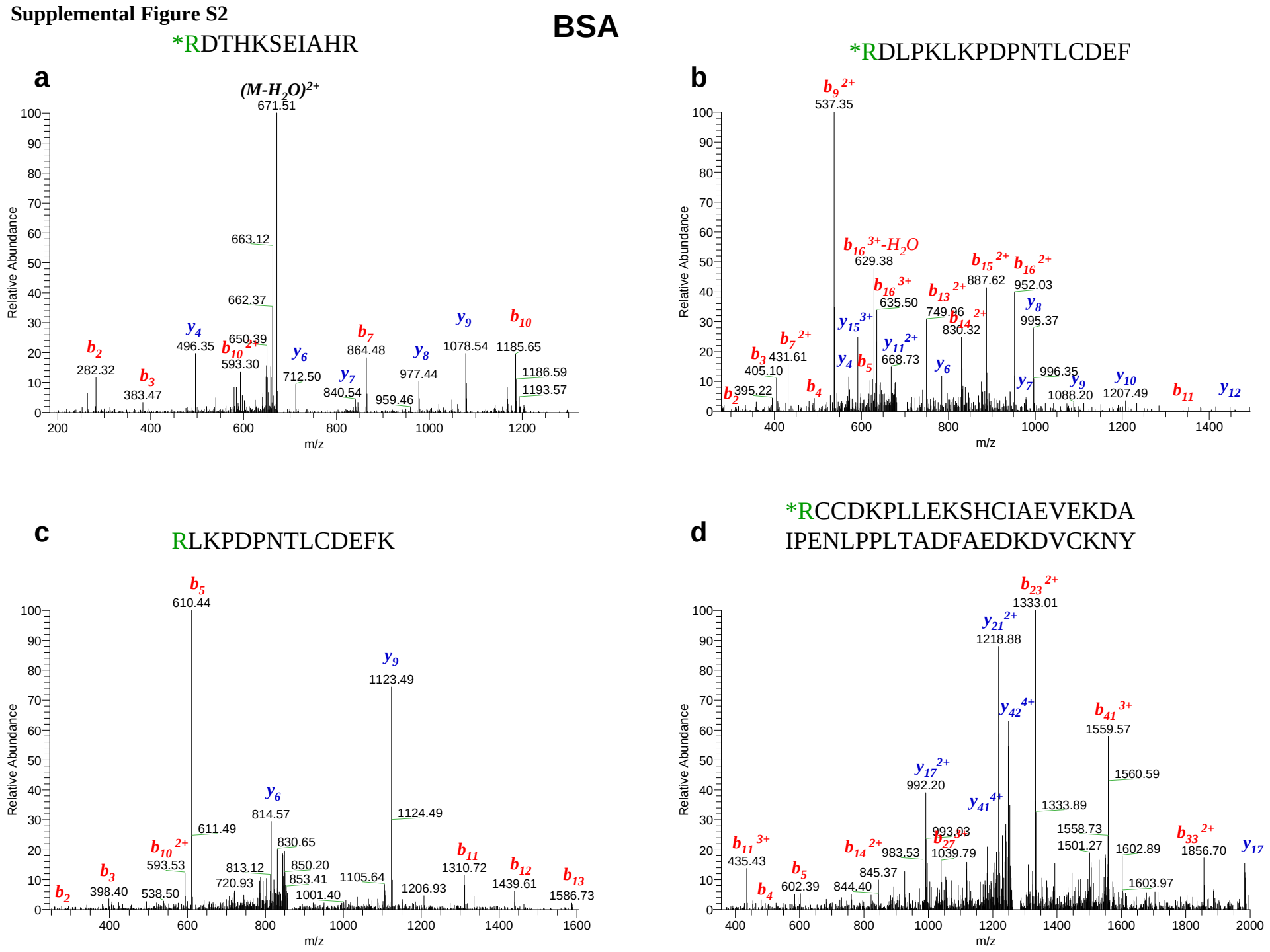
Supplemental Table 2. Peptide sequences and mass spectrometry database search parameters used for the identification of the arginylated peptides in the ATE1 and BSA preparations.

Supplemental Figure S1.

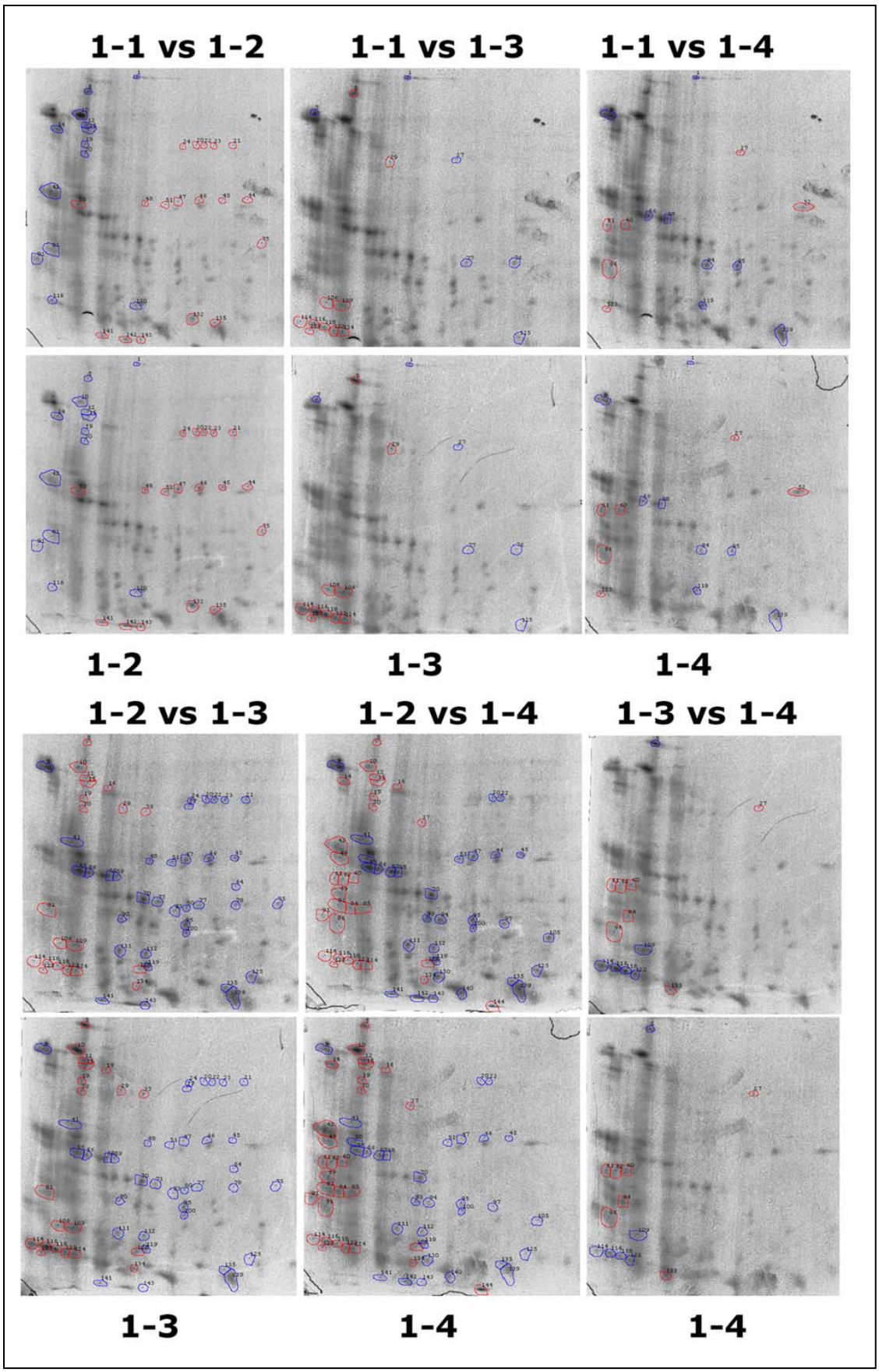


Supplemental Figure S2

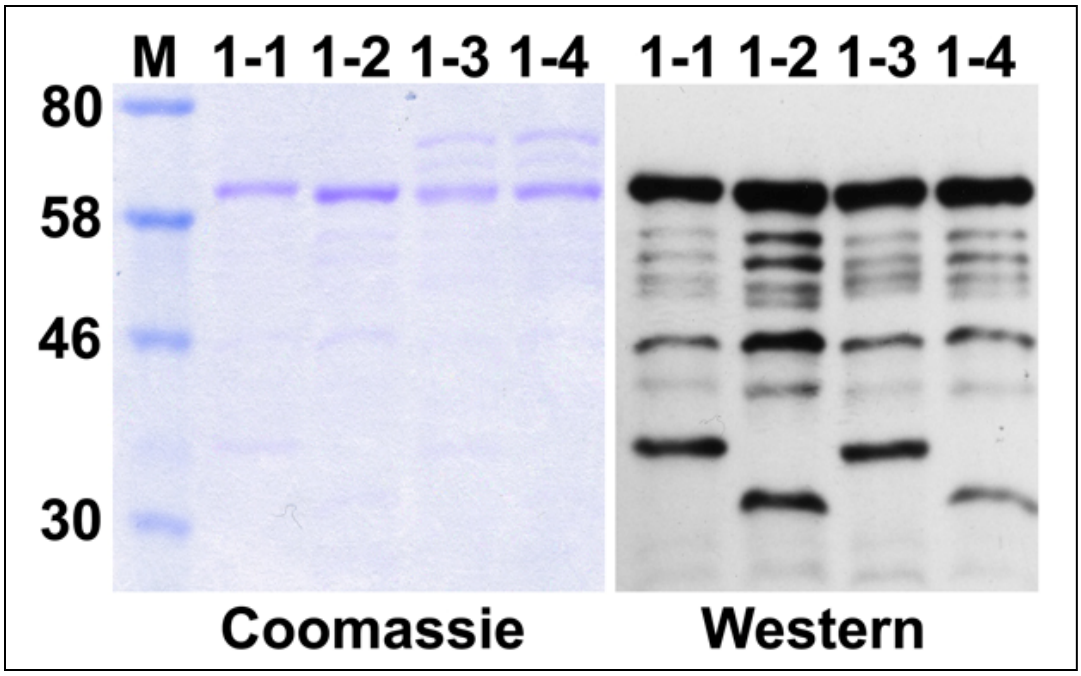
BSA



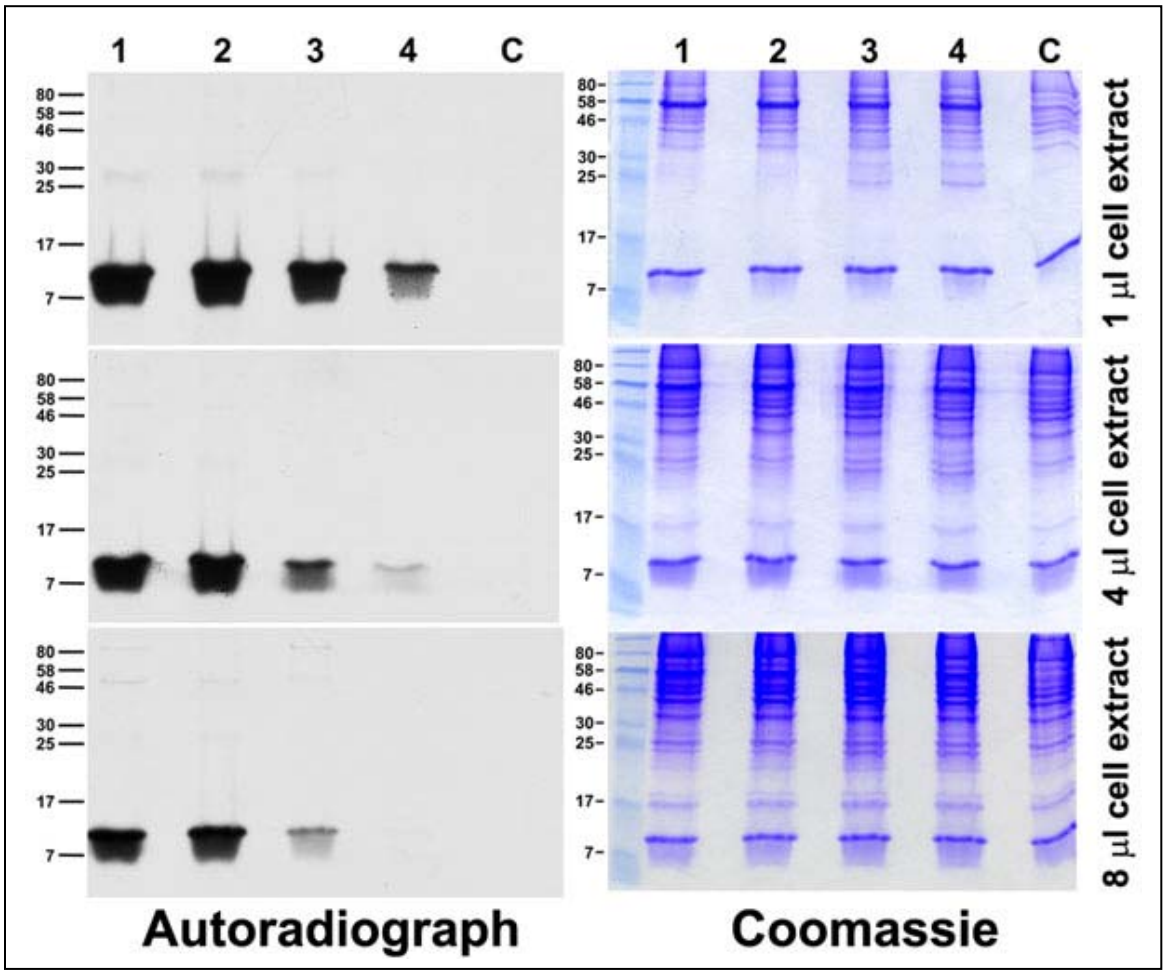
Supplemental Figure S3a



Supplemental Figure S3b



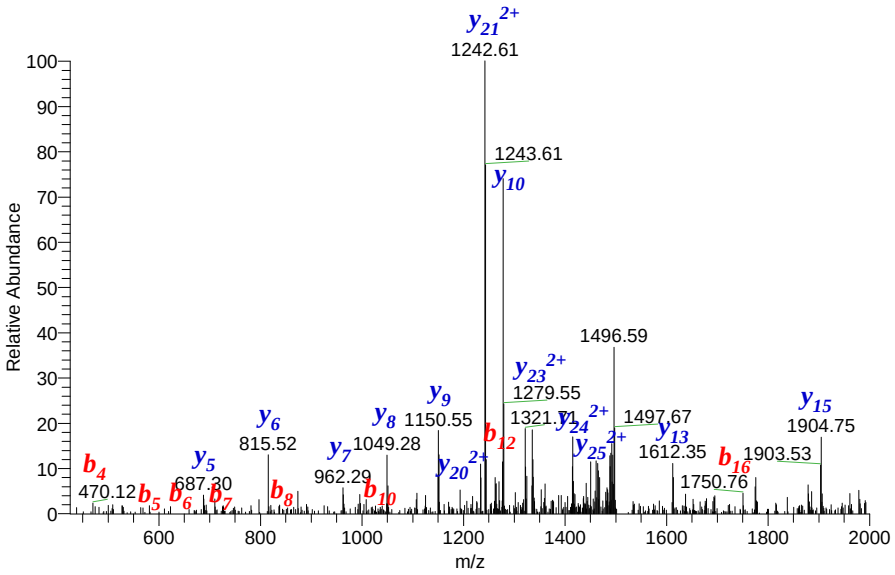
Supplemental Figure S4.



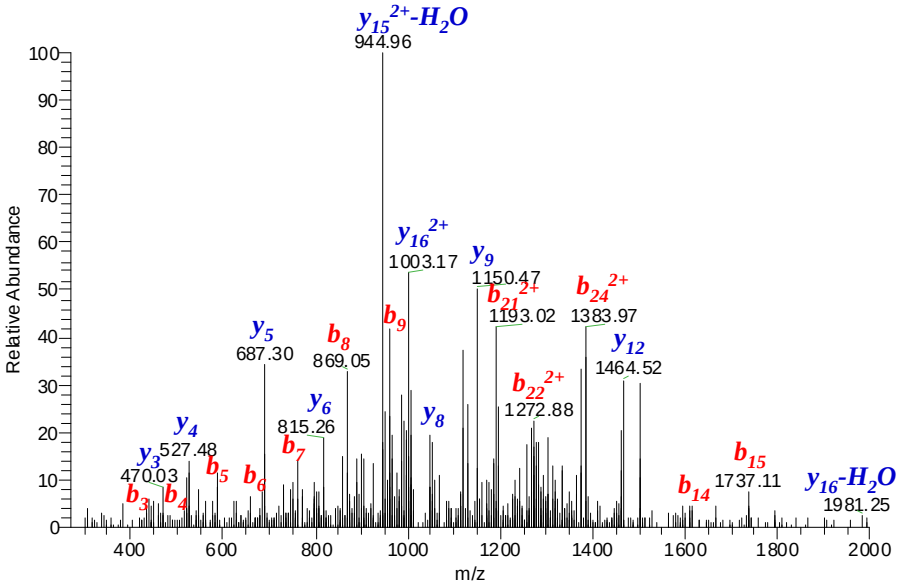
Supplemental Figure S5a

Ate 1-1

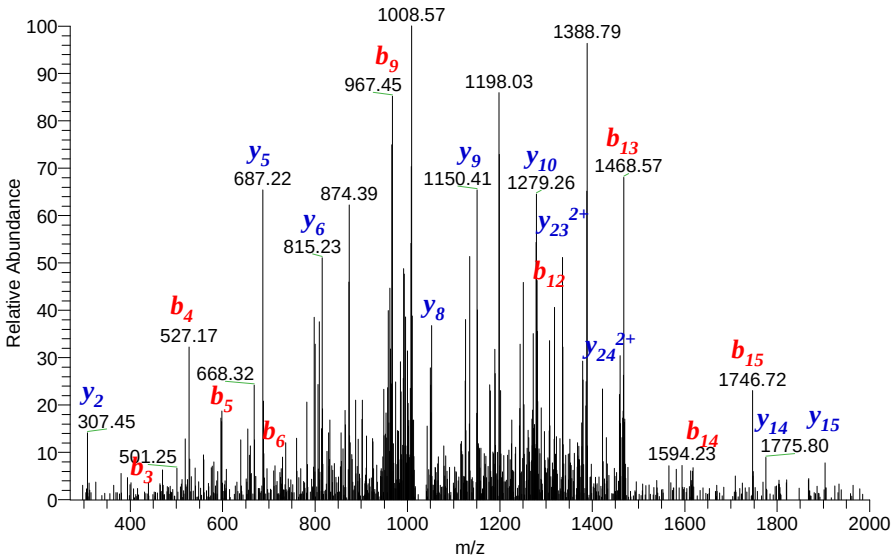
(42.0106)ASWSAPSPSLVEYFEGQTSFQCGYCK



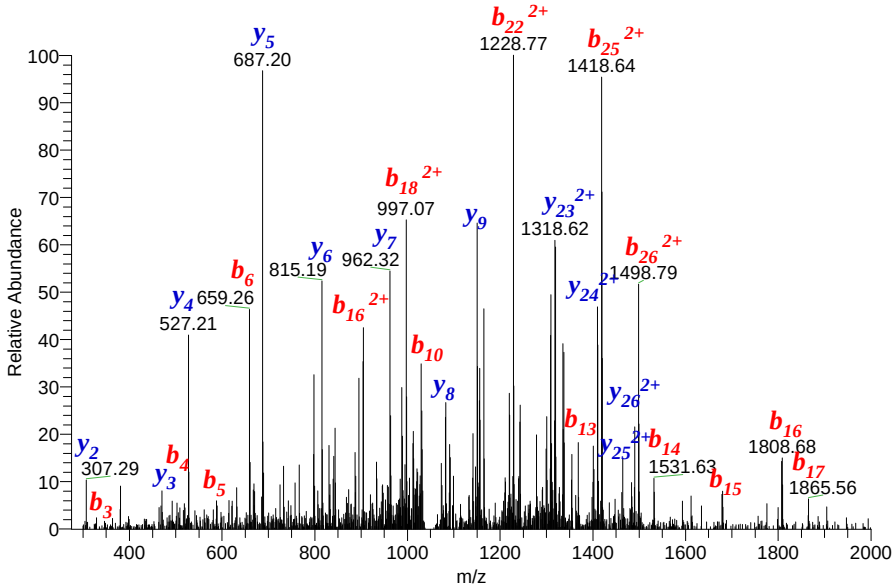
RSWSAPSPSLVEYFEGQTSFQCGYCK



*RSWSAPSPSLVEYFEGQTSFQCGYCK



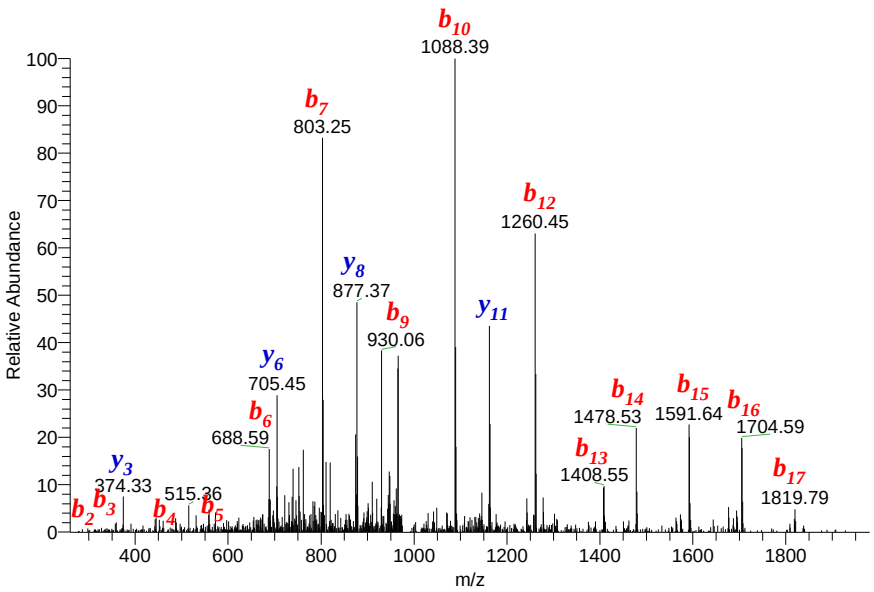
RASWSAPSPSLVEYFEGQTSFQCGYCK



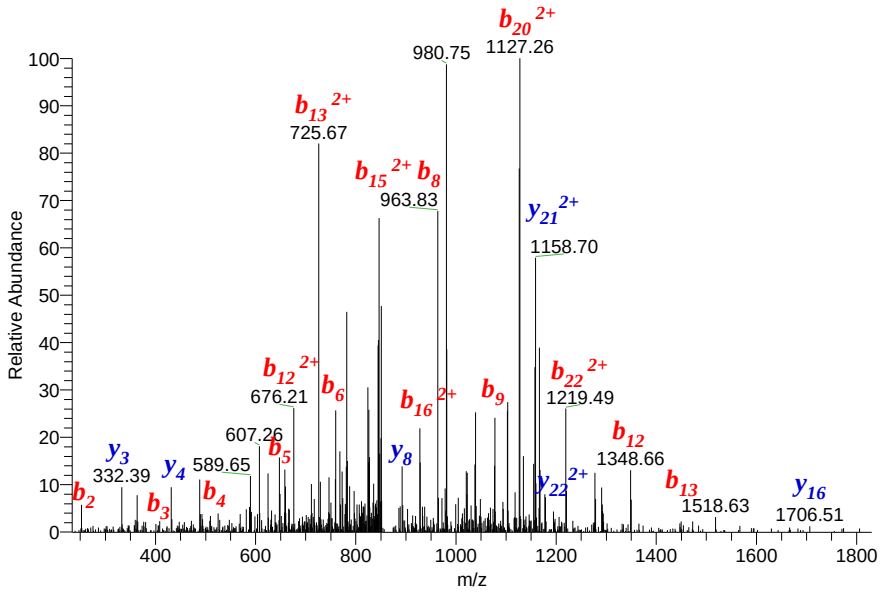
Supplemental Figure S5a

Ate 1-1

RDSTVEDAVDGDFALINK



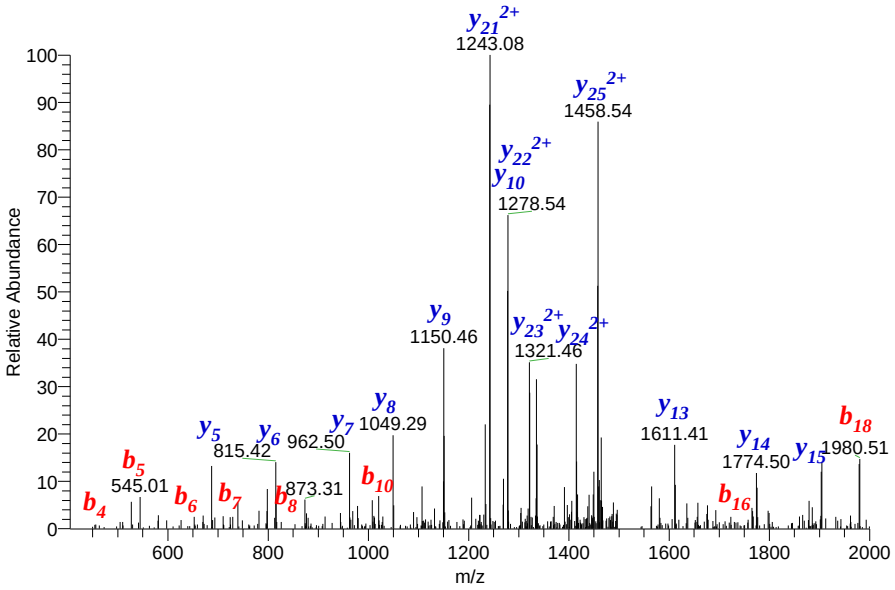
RNHQEDPSEEAGVLEYANLVGQK



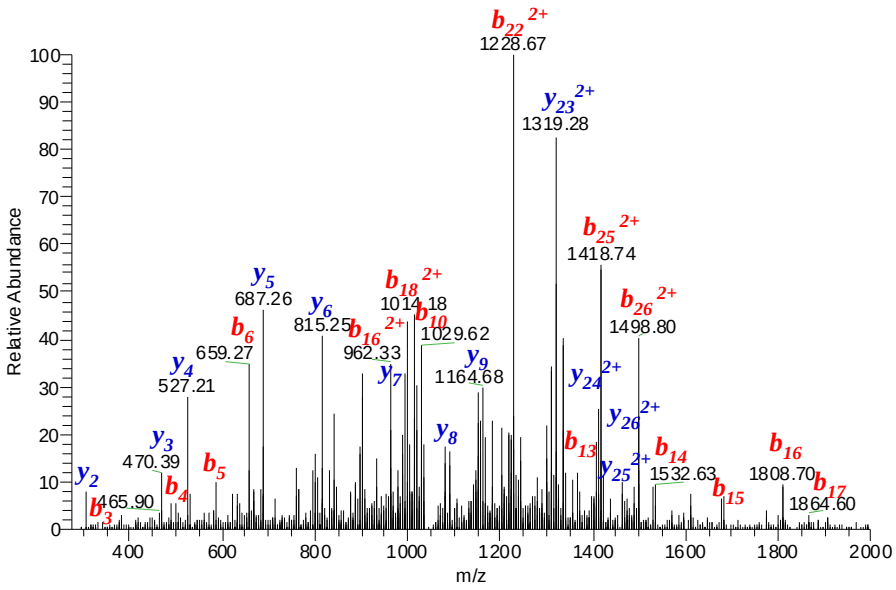
Supplemental Figure S5b

Ate 1-2

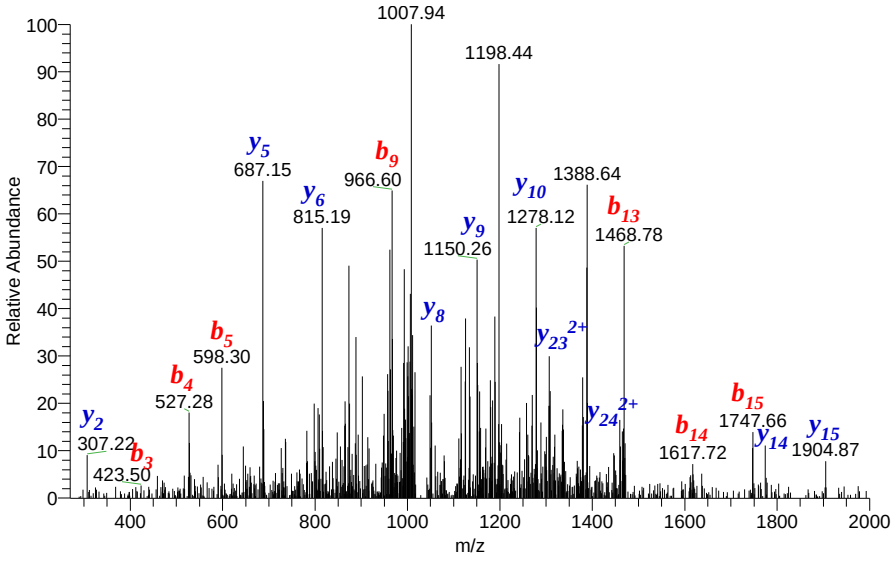
(42.0106)ASWSAPSPSLVEYFEGQTSFQCGYCK



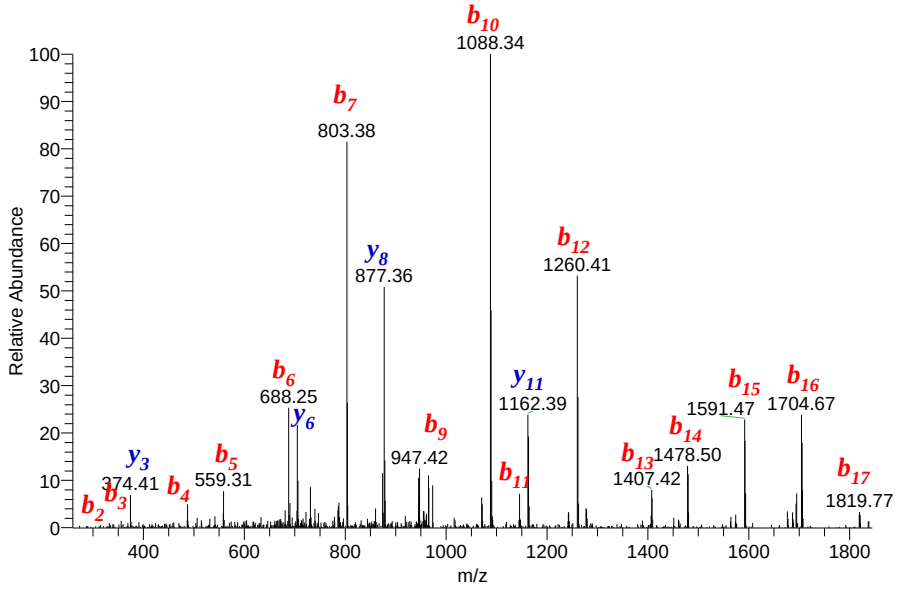
RASWSAPSPSLVEYFEGQTSFQCGYCK



*RSWSAPSPSLVEYFEGQTSFQCGYCK



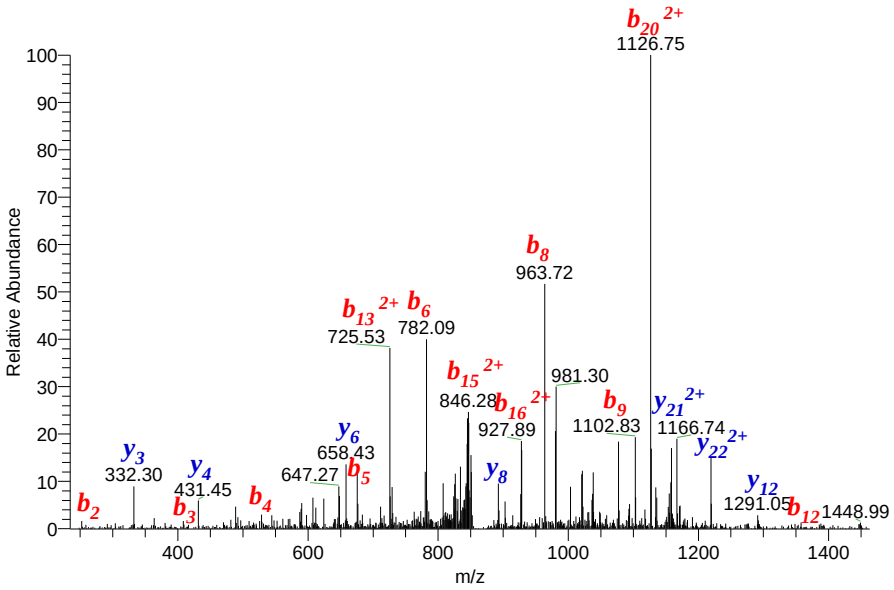
RDSTVEDAVDGDFALINK



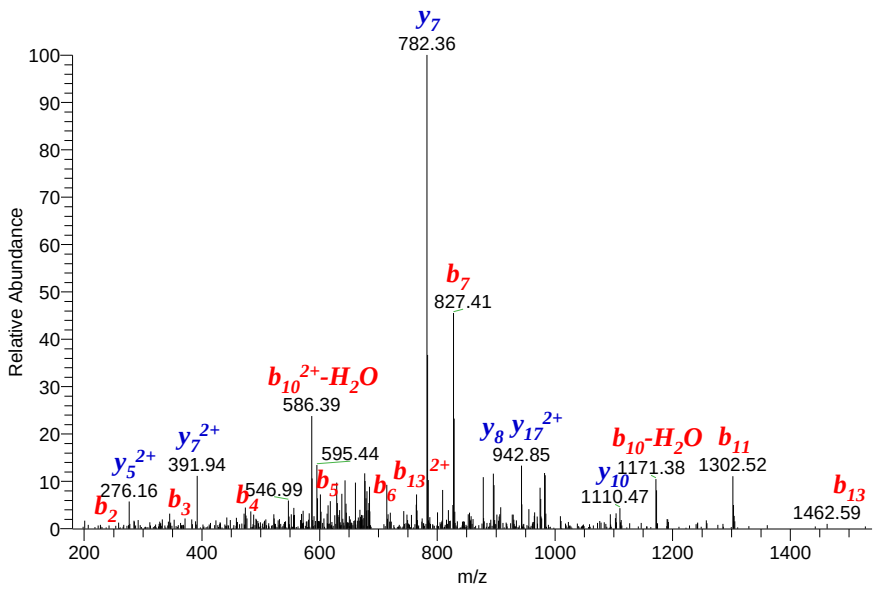
Supplemental Figure S5b

Ate 1-2

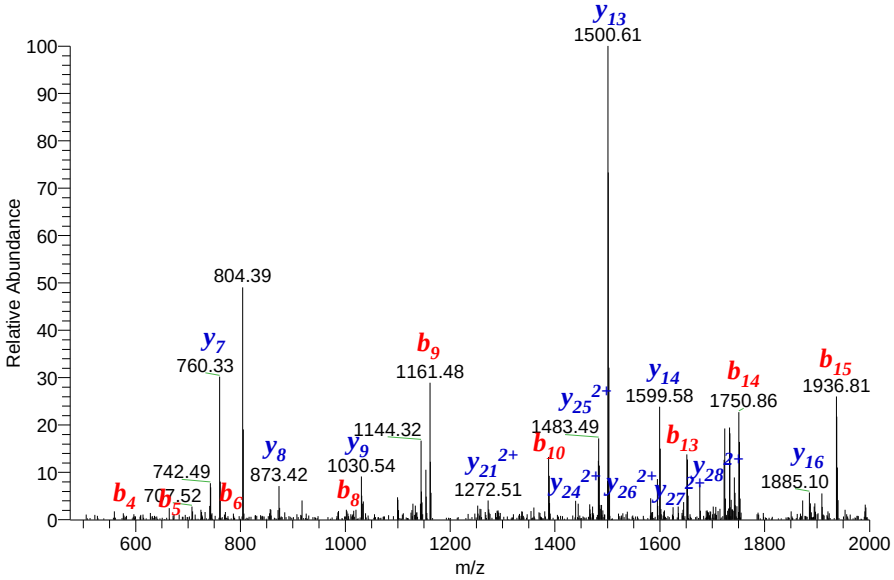
RNHQEDPSEEAGVLEYANLVGQK



RSLEDLIFQSLPENASHK



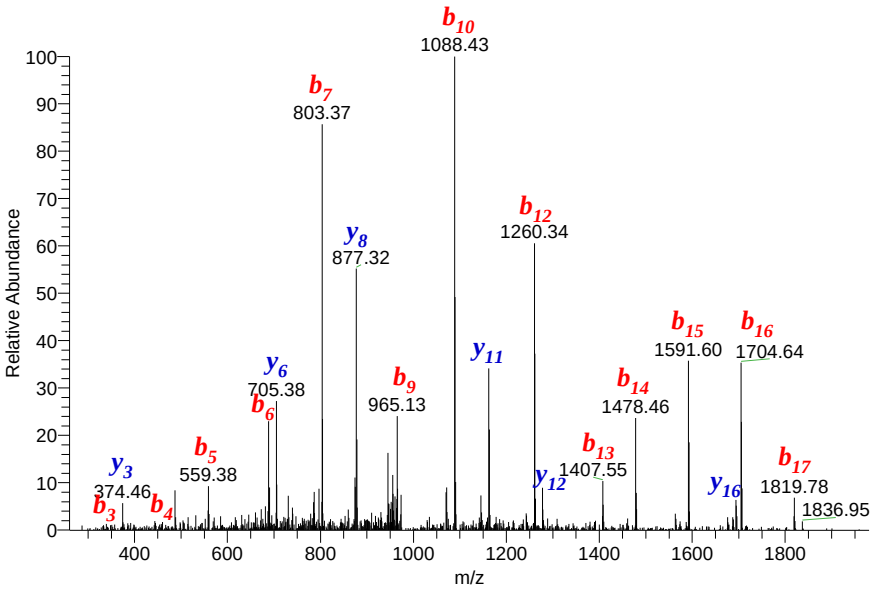
RYRPSDLLCPETYVWVPIEQCLPSLDNSK



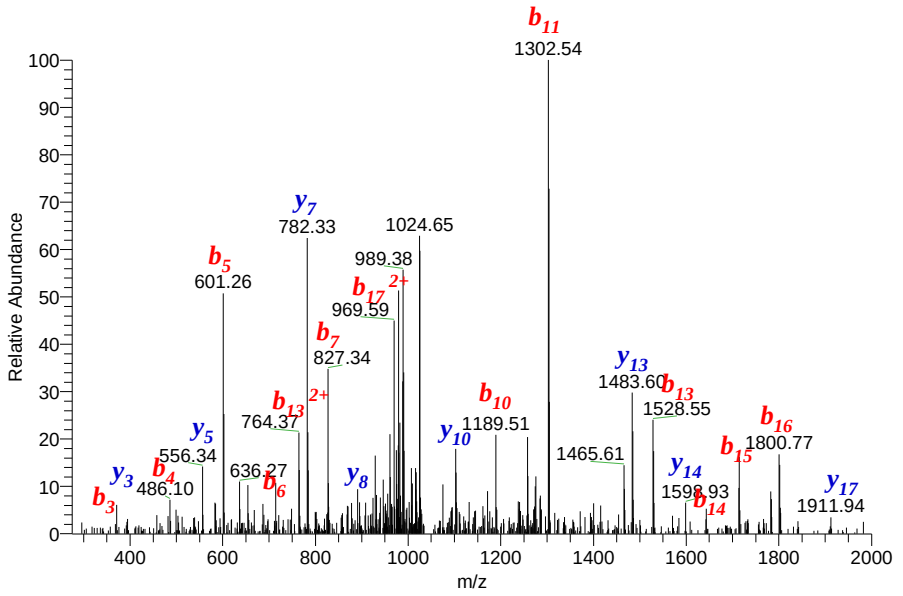
Supplemental Figure S5c

Ate 1-3

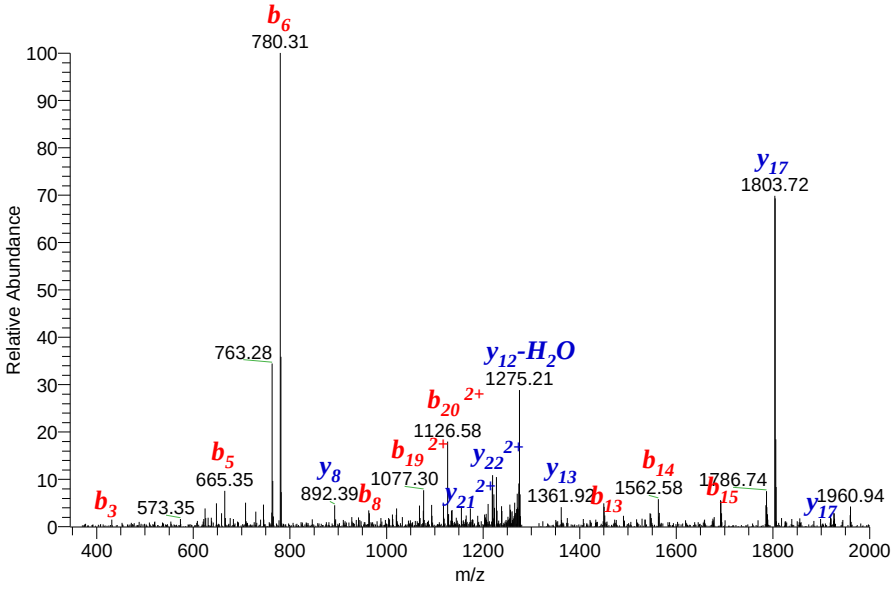
RDSTVEDAVDGDFALINK



RSLEDLIFQSLPENASHK



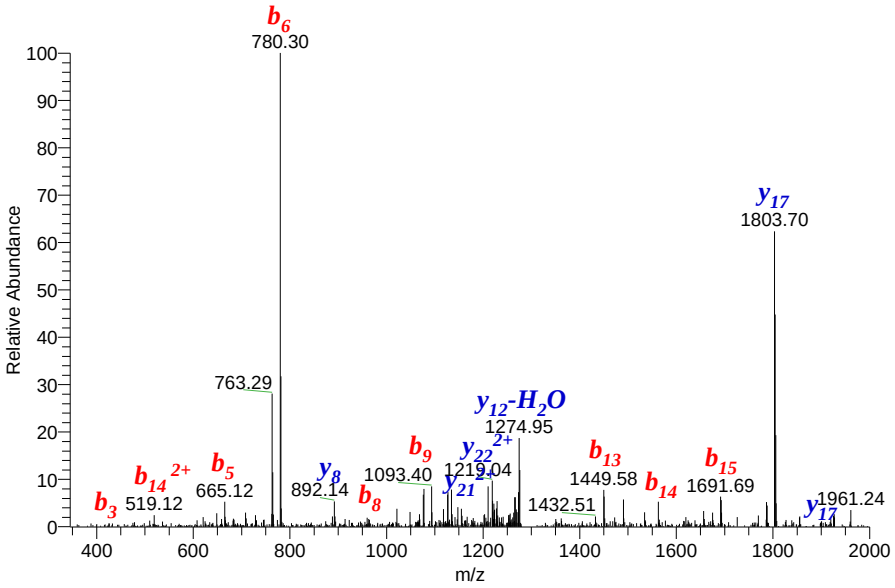
RNHQEDPSEEAGVLEYANLVGQK



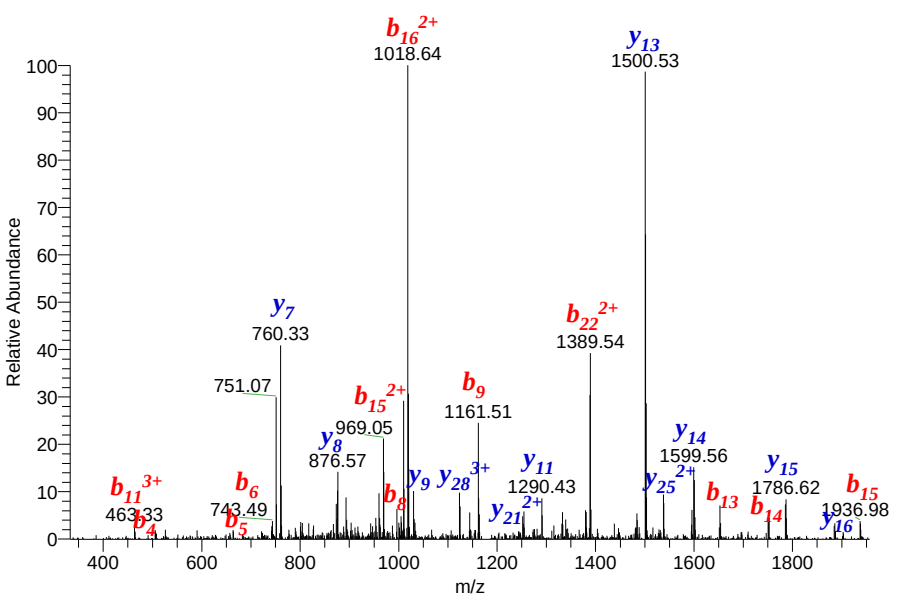
Supplemental Figure S5d

Ate 1-4

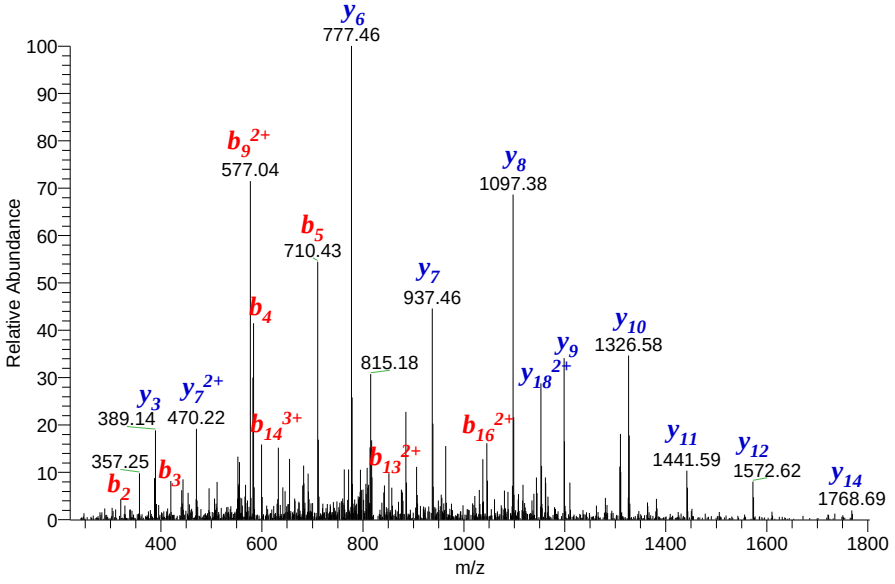
RNHQEDPSEEAGVLEYANLVGQK



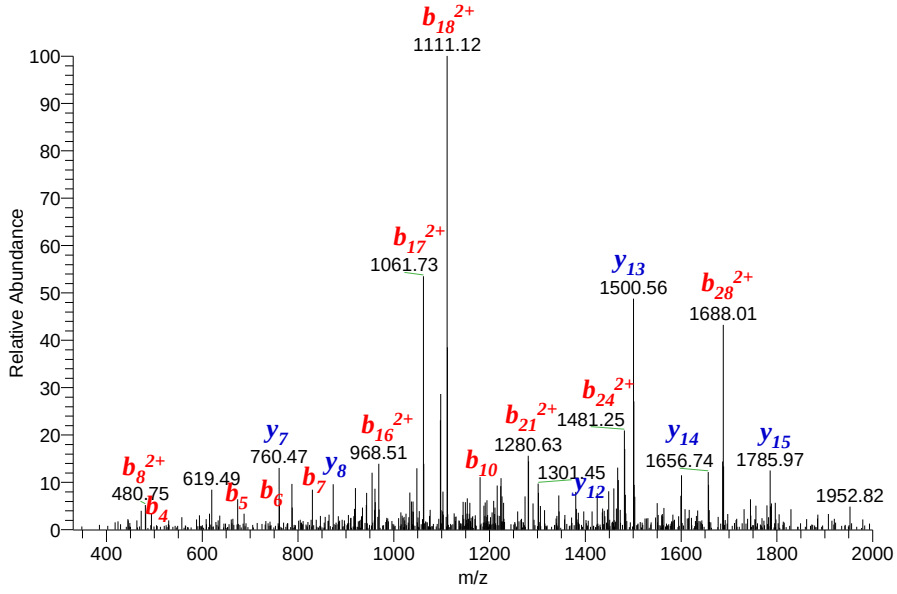
RYRPSDLLCPETYVWVPIEQCLPSLDNSK



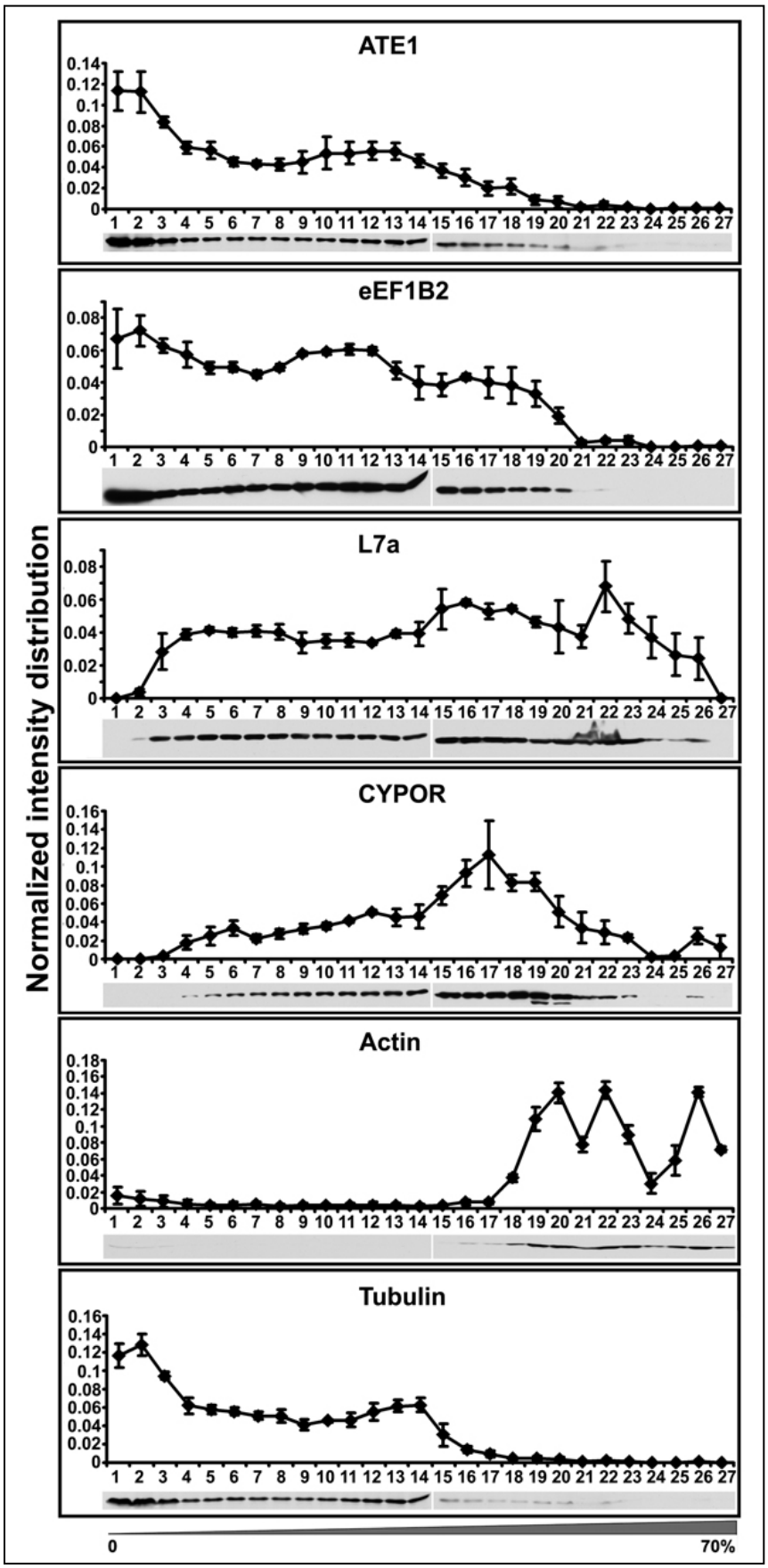
RYVYKPVMDQTCCPQYTIR



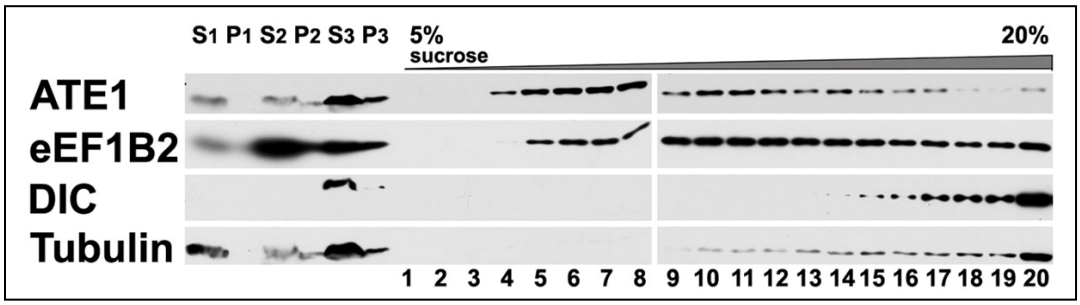
RGQYRPSDLLCPETYVWVPIEQCLPSLDNSK



Supplemental Figure S6a.



Supplemental Figure S6b.



Experimental Procedures:

2D gel electrophoresis and comparison. Two-dimensional electrophoresis was performed according to the carrier ampholine method (O'Farrell, 1975) by Kendrick Labs, Inc. (Madison, WI) as follows: isoelectric focusing was carried out in a glass tube of inner diameter 3.0 mm using 2.0% pH 3.5-10 ampholines (GE Healthcare, Piscataway, NJ) for 20,000 volt-hrs. One μg of an IEF internal standard, tropomyosin, was added to the sample. This protein migrates as a doublet with lower polypeptide spot of MW 33,000 and pI 5.2; an arrow on the stained gel marks its position. The enclosed tube gel pH gradient plot for this set of ampholines was determined with a surface pH electrode. After equilibration for 10 min in buffer "0" (10% glycerol, 50mM dithiothreitol, 2.3% SDS and 0.0625 M tris, pH 6.8), the tube gel was sealed to the top of a stacking gel that overlaid a 10% acrylamide slab gel (1.0 mm thick). SDS slab gel electrophoresis was carried out for about 5 hrs at 25 mA/gel. The following proteins (Sigma Chemical Co., St. Louis, MO) were added as molecular weight standards to a well in the agarose that sealed the tube gel to the slab gel: myosin (220,000), phosphorylase A (94,000), catalase (60,000), actin (43,000) carbonic anhydrase (29,000) and lysozyme (14,000). These standards appear as bands at the basic edge of the Coomassie Brilliant Blue R-250-stained 10% acrylamide slab gel. The gels were stained with Coomassie Brilliant Blue R-250. The cold Coomassie blue-stained gels were dried between sheets of cellophane with the acid edge to the left. The Coomassie blue-stained gels containing ^{14}C label proteins were treated with En^{3}Hance® (New England Nuclear) for one hour, rehydrated with ultrapure water for 30 minutes, and dried onto filter paper with the acid end to the left. Fluorography was carried out using Kodak Biomax film

with an exposure of 16 days at -80^{o}C. The films were developed using Kodak developer and fixer.

Ribosome cosedimentation assays.

To determine whether ATE1 is ribosome associated, co-sedimentation assay was carried using mouse liver lysate as the native ATE1 source and rabbit reticulocyte translation lysate (RRL) (Promega) as the ribosome source, following the centrifugation procedure described in (Hundley et al., 2005). Filtered liver homogenates were prepared as described in the Experimental Procedures section of the main text and centrifuged sequentially at 16,000 x g for 15 min at 4°C using a table top microcentrifuge and at 163,170 x g for 90 min using TL-100 tabletop ultracentrifuge with TLA-100.3 rotor (Beckman) at 4°C to remove the ribosomes. Mixtures of 125 μl liver lysate and 25 μl lysis buffer (ATE1 control), 25 μl RRL and 125 μl lysis buffer (ribosome control), and 125μl liver lysate and 25 μl RRL (ATE1 + ribosomes) were loaded onto 900 μl 0.5 M sucrose cushion and centrifuged at 163,170 x g for 90 min at 4°C. Equivalent amounts of samples from supernatants and pellets were analyzed by Western blot with antibodies against ATE1 and the ribosomal protein L7a. Western blot results and quantifications performed using ImageQuant TL (GE Healthcare) are shown in Fig. 6b of the main text.

Mass spectrometry and database searches.

Protein Digestion: The TCA-precipitated protein sample was redissolved in Invitrosol protein solubilizer, reduced and alkylated (Xu et al., 2009); and then split into two aliquots for separate digestions with Trypsin and Lys-C. Digestions were carried out at 37^{o}C for 16 hrs at

1:100 enzyme to substrate ratio. Formic acid was added to the combined reactions to a final concentration of 5%.

Multidimensional Protein Identification Technology (MudPIT): The protein digest was pressure-loaded onto a Kasil-fritted fused silica capillary column (250-µm i.d.) packed with 3 cm 5-µm Partisphere strong cation exchange resins (SCX, Whatman, Clifton, NJ) and 3 cm 5-µm Aqua C18 resins (RP, Phenomenex, Ventura, CA). The column was then washed with buffer containing 95% water, 5% acetonitrile, and 0.1% formic acid. After desalting, this sample-loaded back-end column was then connected to a 100-µm i.d. capillary column with a 5-µm pulled tip packed with 10 cm 3-µm Aqua C18 material through a zero-dead-volume union (UpChurch Scientific, Oak Harbor, WA), and the entire three phase column was placed inline with an Agilent 1200 quaternary HPLC (Agilent, Palo Alto, CA) and analyzed using a modified 12-step separation described previously. The buffer solutions used were 5% acetonitrile/0.1% formic acid (buffer A), 80% acetonitrile/0.1% formic acid (buffer B), and 500 mM ammonium acetate/5% acetonitrile/0.1% formic acid (buffer C). Step 1 consisted of a 70 min gradient from 0-100% buffer B. Steps 2-11 had the following profile: 5 min of 100% buffer A, 5 min of X% buffer C, a 10 min gradient from 0-15% buffer B, and a 80 min gradient from 15-45% buffer B, a 10 min gradient from 45% to 100% buffer B, and a 10 min equilibration of 100% Buffer A. The 5 min buffer C percentages (X) were 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100% respectively for the 2-11 step analysis. The final step, the gradient contained: 5 min of 100% buffer A, 20 min of 90% buffer C and 10% buffer B, a 10 min gradient from 0-15% buffer B, and a 85 min gradient from 15-100% buffer B. As peptides were eluted from the microcapillary column, they were electrosprayed directly into a hybrid LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific, San Jose, CA) with the application of a distal 2.5 kV spray voltage. A cycle of one FT full-scan

mass spectrum (400-1400 *m/z*, 60000 resolution) followed by 8 data-dependent LTQ MS/MS spectra at a 35% normalized collision energy was repeated continuously throughout each step of the multidimensional separation. Application of mass spectrometer scan functions and HPLC solvent gradients was controlled by the Xcalibur data system (Thermo Fisher Scientific, San Jose, CA).

Database searches and validation: MS2 spectra were analyzed using the software analysis protocol as previously described (Xu et al., 2009). Briefly, MS2 spectra were searched with the ProLucid algorithm (Xu, 2006) against the EBI mouse IPI database (ftp://ftp.ebi.ac.uk/pub/databases/IPI/, version 3.06), that was concatenated to a decoy database in which the sequence for each entry in the original database was reversed. The search parameters include a static cysteine modification of 57.02146 amu and six N-terminal differential modifications for R (156.1011 amu), heavy R (*R, 166.1093 amu), Acetylation (42.0106 amu), R plus Acetylation (198.1117 amu), heavy R plus Acetylation (208.1199 amu) and N-terminal Alkylation (57.02146 amu). No enzyme specificity was considered for any search. ProLuCID results were assembled and filtered using the DTASelect (version 2.0) program (Cociorva et al., 2007; Tabb et al., 2002). DTASelect 2.0 uses a linear discriminant analysis to dynamically set XCorr and DeltaCN thresholds for the entire dataset to achieve a user-specified false discovery rate (5% in this analysis). The false discovery rates are estimated by the program from the number and quality of spectral matches to the decoy database. The MS/MS spectra for the modified peptides were further filtered based on the delta mass (the difference between the measured mass and the theoretical mass). The P value for each peptide precursor mass is calculated as the delta mass for each modified peptide against the distribution of the delta masses of all the non-modified peptides. For each modified peptide, the delta mass P

value defines the chance that the observed mass corresponds to this particular peptide and not a highly similar one created by a similar mass shift as a result of another modification. A true hit should have sufficiently high P value (0.05 in this analysis), based on the assumption that the delta mass distribution of modified peptides should be similar to the distribution of the unmodified peptides. Further manual evaluation was also conducted using criteria reported previously (Wong et al., 2007; Xu et al., 2009).

References

Cociorva, D., D, L. T., and Yates, J. R. (2007). Validation of tandem mass spectrometry database search results using DTASelect. Curr Protoc Bioinformatics *Chapter 13*, Unit 13 14.

Hundley, H. A., Walter, W., Bairstow, S., and Craig, E. A. (2005). Human Mpp11 J protein: ribosome-tethered molecular chaperones are ubiquitous. Science *308*, 1032-1034.

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. J Biol Chem *250*, 4007-4021.

Tabb, D. L., McDonald, W. H., and Yates, J. R., 3rd (2002). DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J Proteome Res *1*, 21-26.

Wong, C. C. L., Xu, T., Rai, R., Bailey, A. O., Yates, J. R., Wolf, Y. I., Zebroski, H., and Kashina, A. (2007). Global Analysis of Posttranslational Protein Arginylation. PLoS Biology *5*, e258.

Xu, T., Venable, J.D., Park, S.K., Cociorva, D., Lu, B., Liao, L., Wohlschlegel, J., Hewel, J., Yates, J. R., IIIrd (2006). ProLuCID, a fast and sensitive tandem mass spectra-based protein identification program. Molecular & Cellular Proteomics *5*, S174.

Xu, T., Wong, C. C. L., Kashina, A., and Yates, J. R., III (2009). Identification of posstranslationally arginylated proteins and peptides by mass spectrometry. Nature Protocols *43*, 325-332.