SUPPORTING INFORMATION

An optimized workflow for multiplexed phosphorylation analysis of TMT-labeled peptides using high-Field Asymmetric waveform Ion Mobility Spectrometry (FAIMS)

Devin K Schweppe [1,*,#], Scott F. Rusin [1,#], Steven P. Gygi [1], Joao A. Paulo [1,*]

[1] Department of Cell Biology, Harvard Medical School, Boston, MA 02115, United States

# Contributed equally

*Corresponding authors

Extended sample preparation. The HCT116 human cell line was propagated in DMEM supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin until 80% confluency was reached. The cells were washed twice with ice cold PBS, and harvested on plate with 8M urea, 200mM EPPS pH 8.5 plus 1X Roche c*O*mplete protease and PhosStop phosphatase inhibitors, syringe lysed (10 times with 21-gauge needle) and stored at -80°C until use.

Mouse brain tissue was a gift from Alexander S. Banks and all previous animal experiments were performed according to procedures approved by Beth Israel Deaconess Medical Center's Institutional Animal Care and Use Committee [1]. Brain tissue from mice (B6.Cg-Lepob/J-Jackson Labs) had been immediately stored in liquid nitrogen after dissection. Tissues were homogenized in lysis buffer (8 M urea, 200 mM EPPS pH 8.5, 1X Roche c*O*mplete protease, 1X Roche PhosStop phosphatase inhibitors) at a concentration of approximately 10-15 mg/mL using a polytron tissue grinder. The homogenized sample was passed 10 times through a 21-gauge (~3.2 cm long) needle. The homogenate was sedimented by centrifugation at 21,000 x g for 5 min. The supernatant was transferred to a new tube and stored at -80°C until use.

Following a BCA assay to estimate protein concentration, all lysates were reduced (20 min with 5mM TCEP at room temperature), alkylated (20 min with 10mM iodoacetamide, at room temperature in the dark), and quenched (20 min with 10mM DTT, at room temperature in the dark). Proteins were precipitated by chloroform-methanol precipitation, as described previously [2]. Precipitated proteins were resuspended in 200mM EPPS pH 8.5 (~1 mg/ml) and digested first with LysC (Wako) for 3 hr 37°C shaking on a vortex (speed =50%) followed by a 6 hr trypsin digestion at 37°C (sequencing grade, Pierce), both at a 100:1 protein-to-protease ratio. The digests were acidified and desalted (C18 SepPak, Waters) prior to enrichment.

Extended mass spectrometry data analysis. Spectra were converted to mzXML via MSconvert [3]. Database searching included all entries from the human or mouse Uniprot database (downloaded April 2014). Each database was concatenated with one composed of all protein sequences for that database in the reversed order. Searches were performed using a 50-ppm precursor ion tolerance for total protein level profiling. The product ion tolerance was set to 0.9 Da. These wide mass tolerance windows were chosen to maximize sensitivity in conjunction with SEQUEST v.28 (rev. 12) searches and linear discriminant analysis [4, 5]. Data analysis and visualization was performed in Microsoft Excel or R (packages: ggplot2, UpSetR, stringr, dplyr, reshape2, gridExtra).

References

1. Paulo, J. A.; McAllister, F. E.; Everley, R. A.; Beausoleil, S. A.; Banks, A. S.; Gygi, S. P., Effects of MEK inhibitors GSK1120212 and PD0325901 in vivo using 10-plex quantitative proteomics and phosphoproteomics. *Proteomics* **2015,** *15* (2-3), 462-73.

2. Paulo, J. A.; Gygi, S. P., A comprehensive proteomic and phosphoproteomic analysis of yeast deletion mutants of 14-3-3 orthologs and associated effects of rapamycin. *Proteomics* **2015,** *15* (2-3), 474-86.

3. Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M. Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P., A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotechnol* **2012,** *30* (10), 918-20.

4. Beausoleil, S. A.; Villen, J.; Gerber, S. A.; Rush, J.; Gygi, S. P., A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nature biotechnology* **2006,** *24* (10), 1285-92.

5. Huttlin, E. L.; Jedrychowski, M. P.; Elias, J. E.; Goswami, T.; Rad, R.; Beausoleil, S. A.; Villen, J.; Haas, W.; Sowa, M. E.; Gygi, S. P., A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **2010,** *143* (7), 1174-89.

Figure S1

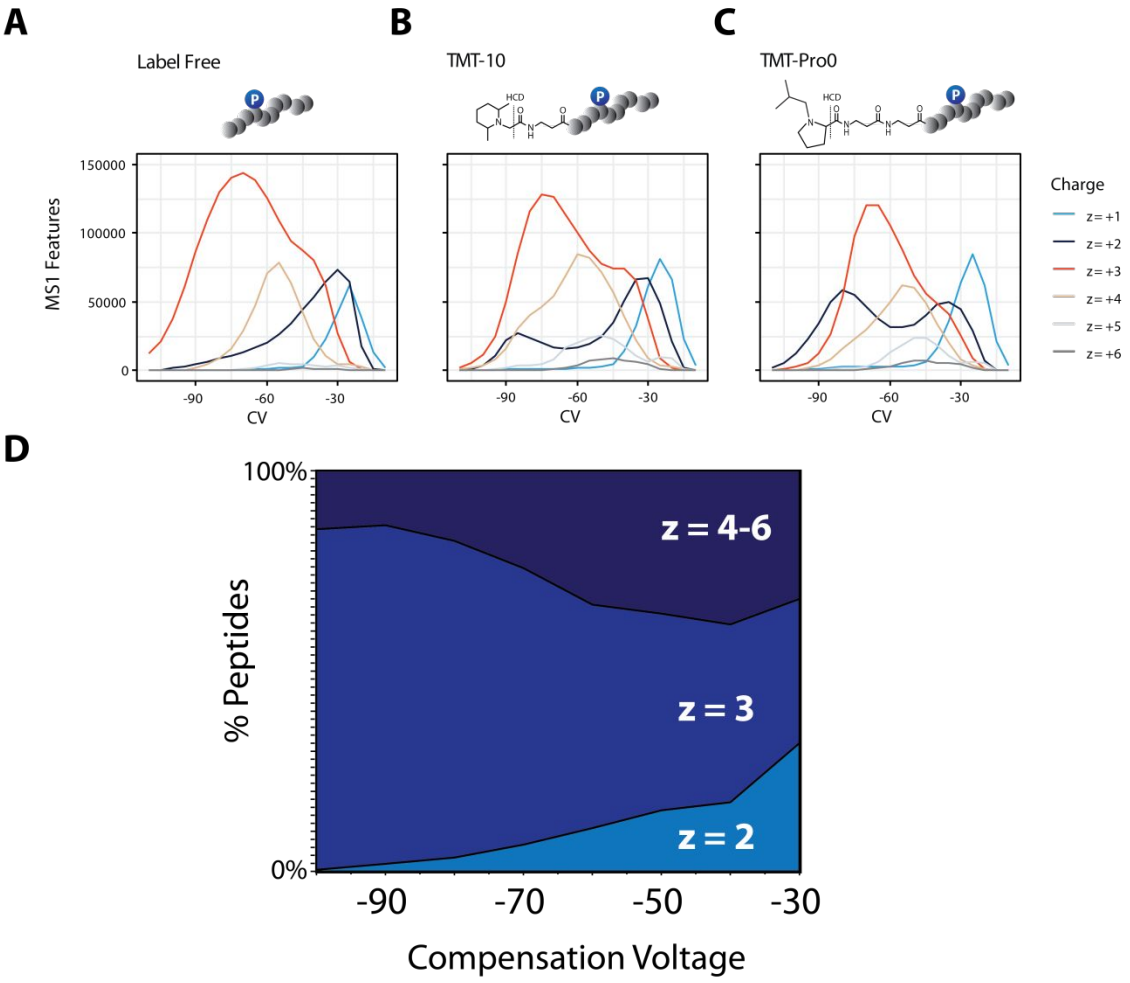


Figure S1: MS1 features across CVs for differentially derivatized peptide populations. Distribution of the MS1 features for different CVs for separate charge states for: **A.** label-free, **B.** TMT10-labeled, and **C.** TMTpro0-labeled peptides. We note differences with respect to charge state distribution across CVs, particularly between label-free and both TMT-derivatized samples. **D.** Proportion of identified peptide charge states across a FAIMS CV range of -100V to -30V.

Figure S2

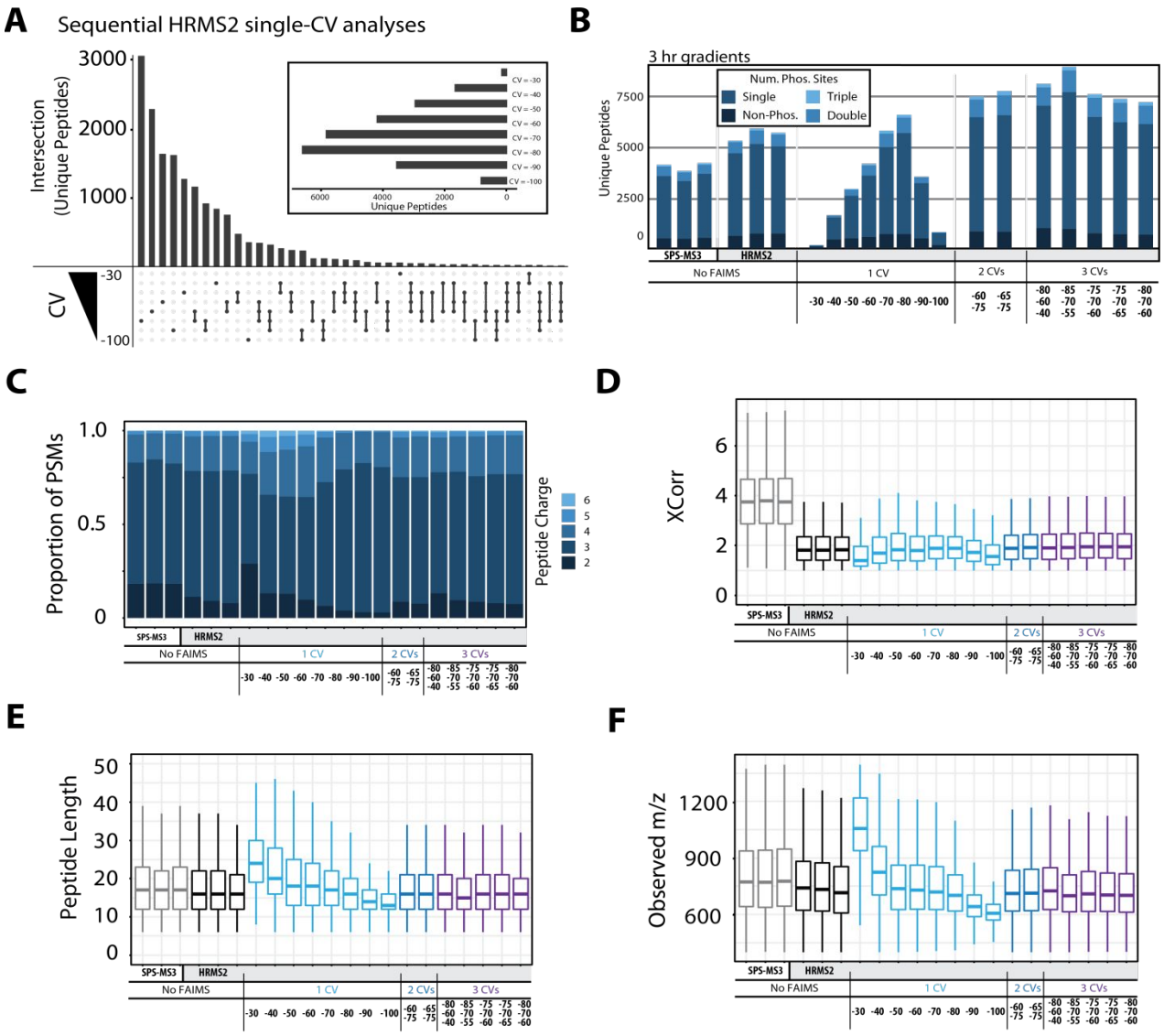
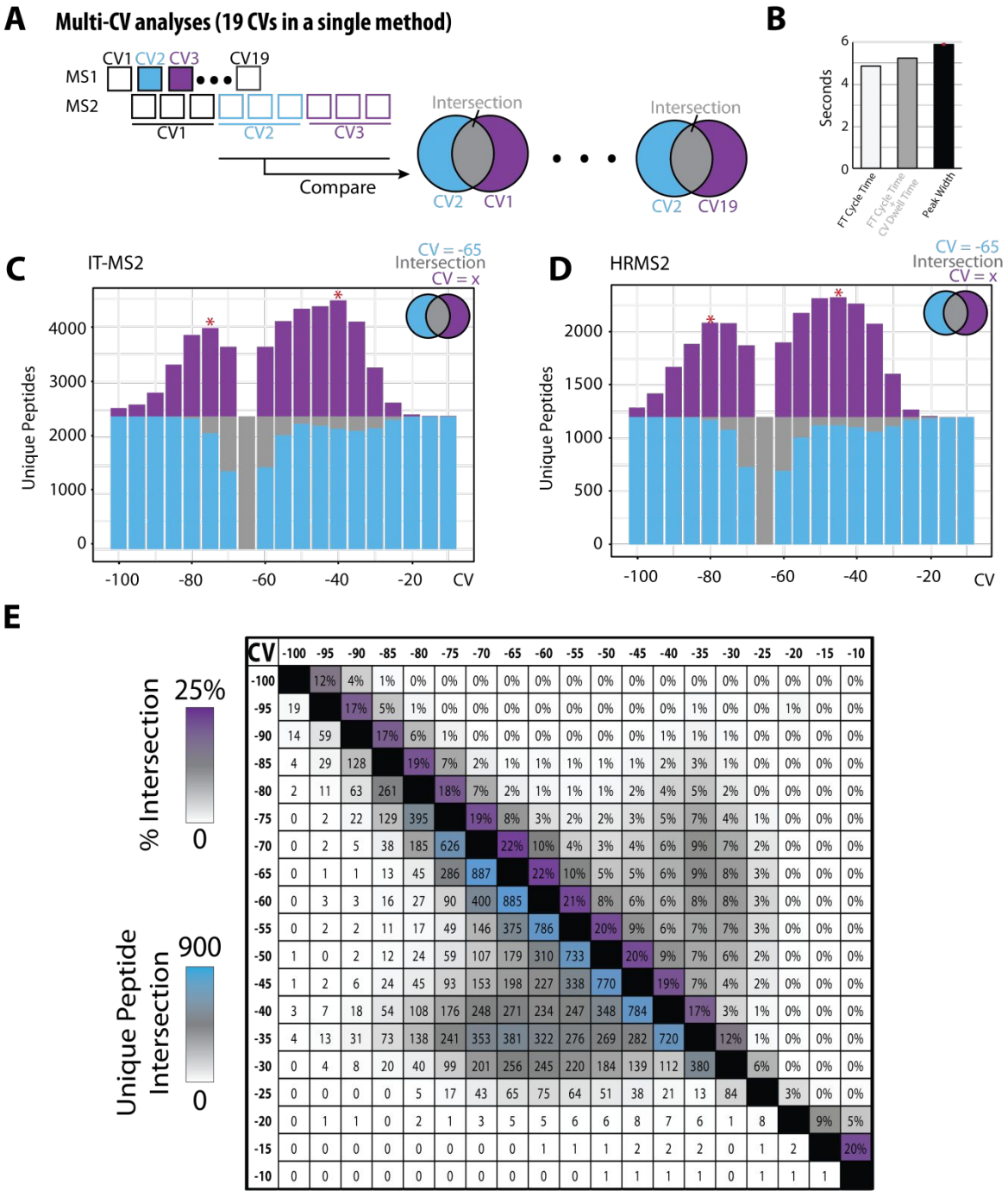


Figure S2: Extended optimization of FAIMS for phosphorylation analysis. A. Upset plot illustrates the intersection of phosphopeptides from 8 consecutive FAIMS analyses with -10V CV increments ranging from -10 to -100V. **B.** Unique peptides identified by HRMS2, SPS-MS3, and FAIMS-MS3-based methods using single, double, or triple CV combinations. **C.** Graph displaying the proportion of the charge state for the phosphopeptide population identified by each of the methods used in **Figure 1C**. Box plots depicting figures of merit for the data depicted in **Figure 1C**: **D.** XCorr SEQUEST scoring metric, **E.** peptide length, and **F.**. observed m/z (Th).

Figure S3



CV	-100	-95	-90	-85	-80	-75	-70	-65	-60	-55	-50	-45	-40	-35	-30	-25	-20	-15	-10
-100		12%	4%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
-95	19		17%	5%	1%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%	0%
-90	14	59		17%	6%	1%	0%	0%	0%	0%	0%	0%	1%	1%	1%	0%	0%	0%	0%
-85	4	29	128		19%	7%	2%	1%	1%	1%	1%	1%	2%	3%	1%	0%	0%	0%	0%
-80	2	11	63	261		18%	7%	2%	1%	1%	1%	2%	4%	5%	2%	0%	0%	0%	0%
-75	0	2	22	129	395		19%	8%	3%	2%	2%	3%	5%	7%	4%	1%	0%	0%	0%
-70	0	2	5	38	185	626		22%	10%	4%	3%	4%	6%	9%	7%	2%	0%	0%	0%
-65	0	1	1	13	45	286	887		22%	10%	5%	5%	6%	9%	8%	3%	0%	0%	0%
-60	0	3	3	16	27	90	400	885		21%	8%	6%	6%	8%	8%	3%	0%	0%	0%
-55	0	2	2	11	17	49	146	375	786		20%	9%	6%	7%	7%	3%	0%	0%	0%
-50	1	0	2	12	24	59	107	179	310	733		20%	9%	7%	6%	2%	0%	0%	0%
-45	1	2	6	24	45	93	153	198	227	338	770		19%	7%	4%	2%	0%	0%	0%
-40	3	7	18	54	108	176	248	271	234	247	348	784		17%	3%	1%	0%	0%	0%
-35	4	13	31	73	138	241	353	381	322	276	269	282	720		12%	1%	0%	0%	0%
-30	0	4	8	20	40	99	201	256	245	220	184	139	112	380		6%	0%	0%	0%
-25	0	0	0	0	5	17	43	65	75	64	51	38	21	13	84		3%	0%	0%
-20	0	1	1	0	2	1	3	5	5	6	6	8	7	6	1	8		9%	5%
-15	0	0	0	0	0	0	0	0	1	1	1	2	2	2	0	1	2		20%
-10	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	1	1	1	

Figure S3: Determining the combination of compensation voltages (CV) that generates the highest number of CV-exclusive peptides. A. Schematic of the analysis. A series of 20 CVs were selected that ranged from -10 to -100V in 5 V increments. We compared all CVs relative to CV=-65V. Blue represents peptides unique to CV=-65V. Grey represents the peptides of a given

CV that overlap with CV=-65V. **B**. Estimated FT (Fourier transform) cycle time plus CV dwell time for all 19 CVs is less than the average measured peak width for identified peptides. **C.** Low resolution MS2 (IT-MS2) and **D.** high-resolution MS2 (HRMS2) CV scanning data for are represented with stacked bar graph with color scheme described above. Note that -65V overlaps with itself completely, thus is represented by a grey bar. **E.** Matrix of phosphopeptide overlap among the CVs tested. The number of overlapping phosphopeptides and the percent overlap are listed for each pair of CVs below and above the diagonal, respectively.

Figure S4

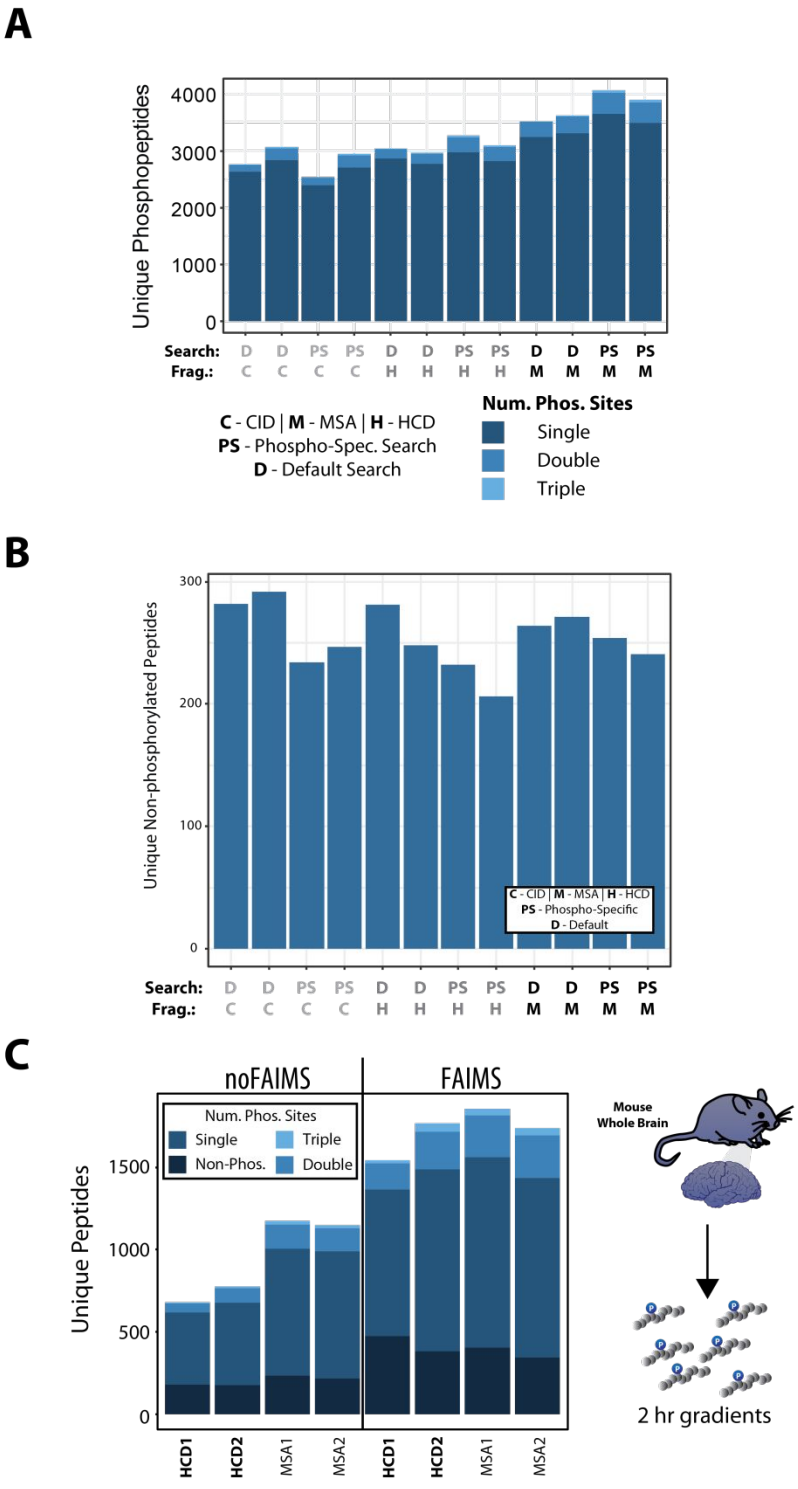


Figure S4: Extended comparison of methods using HCD and MSA fragmentation for MS2 analysis. A. Unique peptides identified via HCD and MSA. Phosphopeptide-enriched sample is of mouse origin, as used in **Figure 2B**. The samples were analyzed in duplicate, with or without

FAIMS voltage (i.e., CV=0V and CV=-40,-60,-80V) and with MSA or HCD. **B.** Analysis of non-phosphorylated peptides with methods as used in **Figure 2A**. TMT-labeled peptides from human cell lysate are analyzed in duplicate. These peptides are fragmented (frag) with CID (C), HCD (H), or MSA-CID (M) and searched with phosphopeptide-specific (PS) option in SEQUEST or not (D, default). The PS option allows neutral loss fragments to be considered by SEQUEST when scoring peptide matches. C. Comparison of non-FAIMS and FAIMS analysis of murine-derived phosphopeptides using either HCD or MSA fragmentation. FAIMS consistently improved the number of identified peptides.