Supplementary Figures, Tables, Data and Note

Multiplexed MS/MS for Improved Data Independent Acquisition

Jarrett D. Egertson[1], Andreas Kuehn[2], Gennifer E. Merrihew[1], Nicholas W. Bateman[3], Brendan X. MacLean[1], Ying S. Ting[1], Jesse D. Canterbury[4], Donald M. Marsh[1], Markus Kellmann[2], Vlad Zabrouskov[4], Christine C. Wu[3], Michael J. MacCoss[1]

[1]Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington, USA

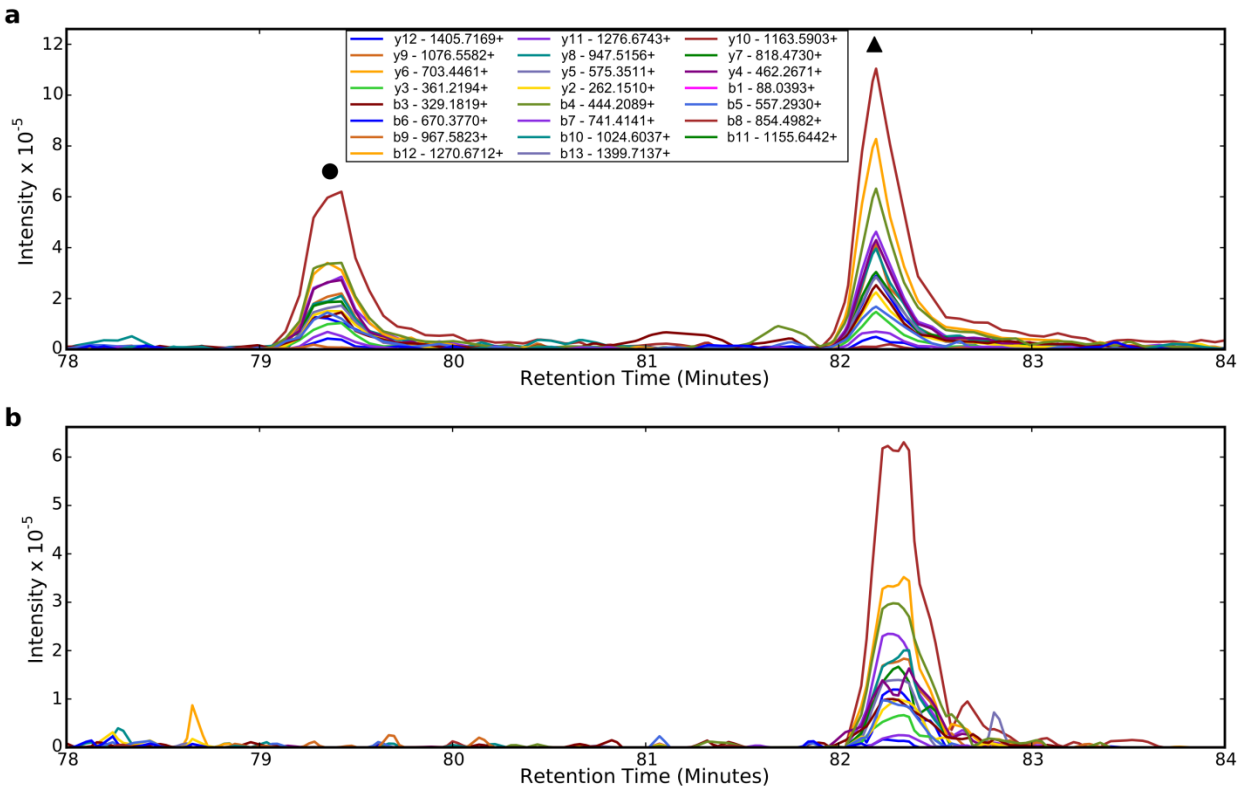
[2]Thermo Fisher Scientific (Bremen) GmbH, Bremen, Germany

[3]Department of Cell Biology and Physiology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA

[4]Thermo Fisher Scientific, San Jose, California, USA

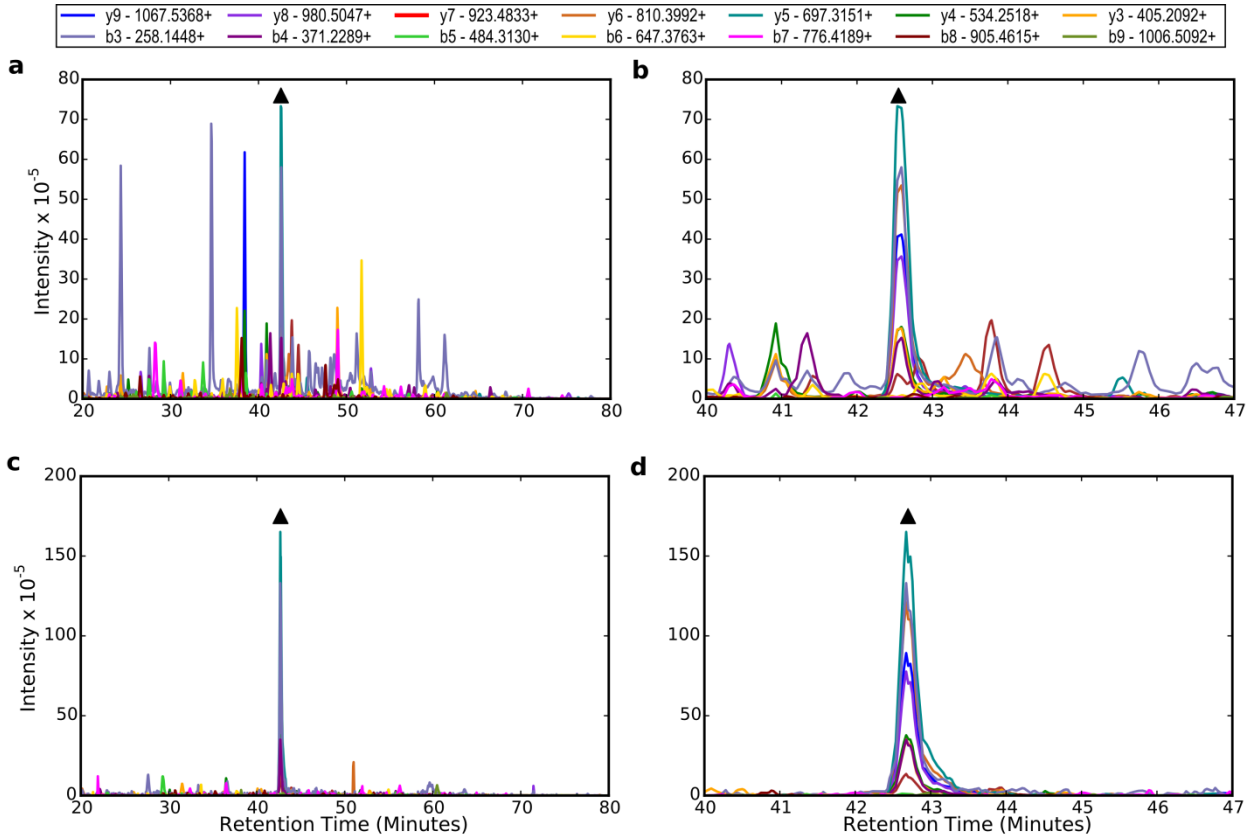
Supplementary Figures

Supplementary Figure 1: Improved Selectivity of MSX over 10 *m/z* DIA



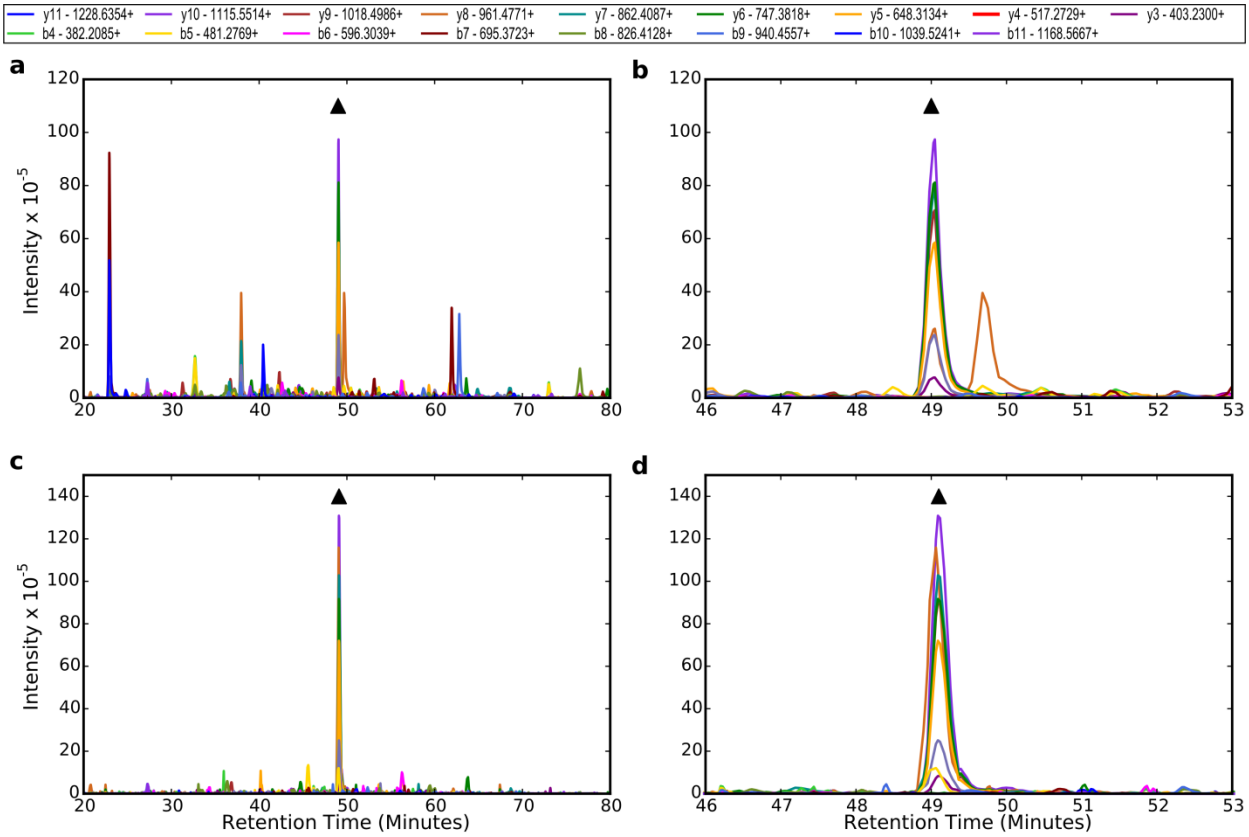
Both a) and b) show extracted fragment ion chromatograms for the peptide SLQDIIAILGMDELSEEDKLTVSR+++ from an analysis of a *C. elegans* soluble lysate sample on a Q-Exactive. Panel **a** shows the extracted fragment ion traces from a 10 *m/z* wide window DIA approach in which the mass range from 500 *m/z* to 900 *m/z* is analyzed with consecutive 10 *m/z* wide isolation window targeted MS/MS scans. The peak labeled with a triangle is the unmodified peptide, the peak labeled with the circle is the peptide with an oxidized methionine (validated using data from a DDA analysis of the same sample on the same column). Panel **b** shows the same fragment ion traces from an MSX experiment. The oxidized methionine peak has been removed due to the improved precursor selectivity of the MSX method.

Supplementary Figure 2: 10 *m/z* DIA and MSX Analysis of ISGLIYEETR++ Peptide



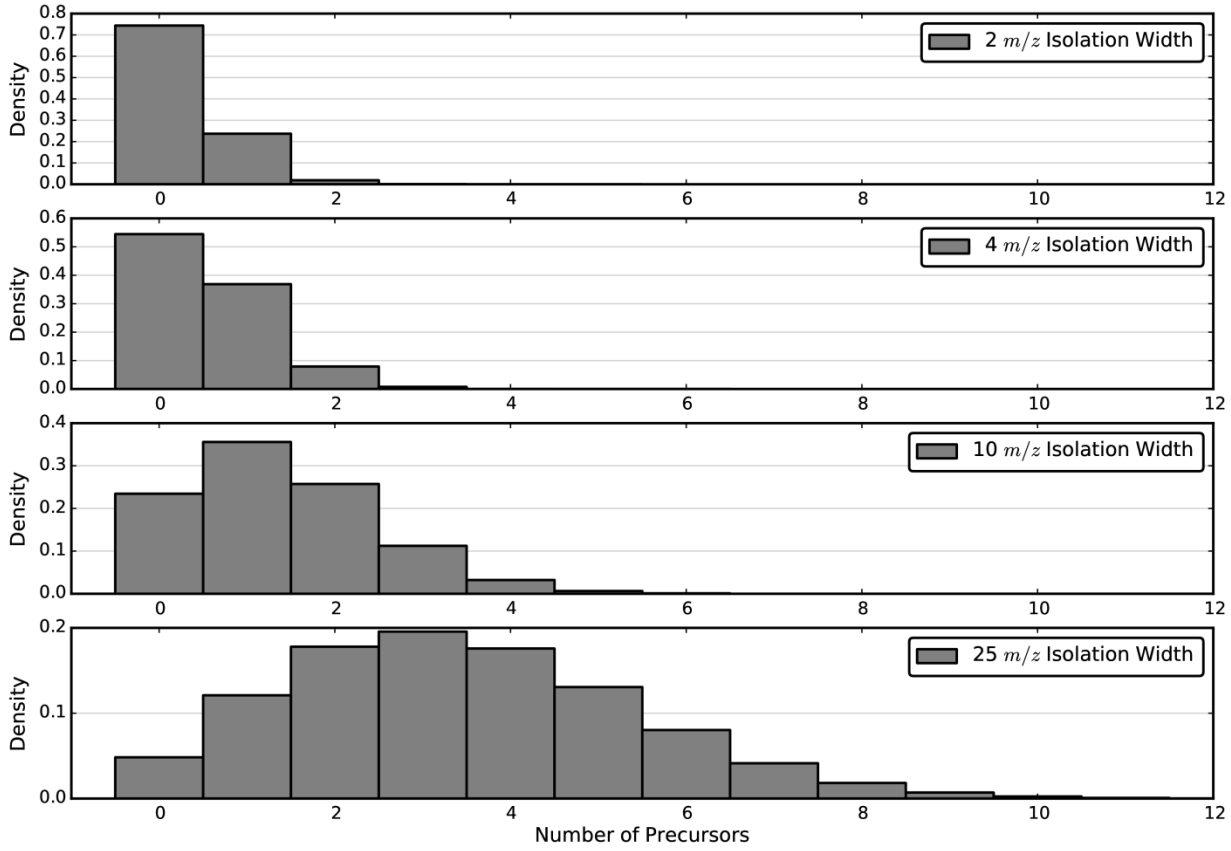
A *C. elegans* soluble lysate was analyzed on the Q-Exactive using a 10 *m/z* wide isolation window DIA approach in which the mass range from 500 *m/z* to 900 *m/z* is analyzed with consecutive 10 *m/z* wide isolation window targeted MS/MS scans and a 5 x 4 *m/z* MSX approach covering the same mass range. Fragment ion chromatograms for the peptide ISGLIYEETR++ were extracted using Skyline from both the wide-window DIA (**a** and **b**) and MSX data (**c** and **d**). The peak with the triangle hovering over it is the peak corresponding to this peptide as validated in a DDA analysis. In both the zoomed-out views (panels **a** and **c**) and zoomed-in views (panels **b** and **d**), the improved precursor selectivity of MSX results in much less non-target peaks.

Supplementary Figure 3: 10 *m/z* DIA and MSX Analysis of NIPGVDVMNVER++ Peptide



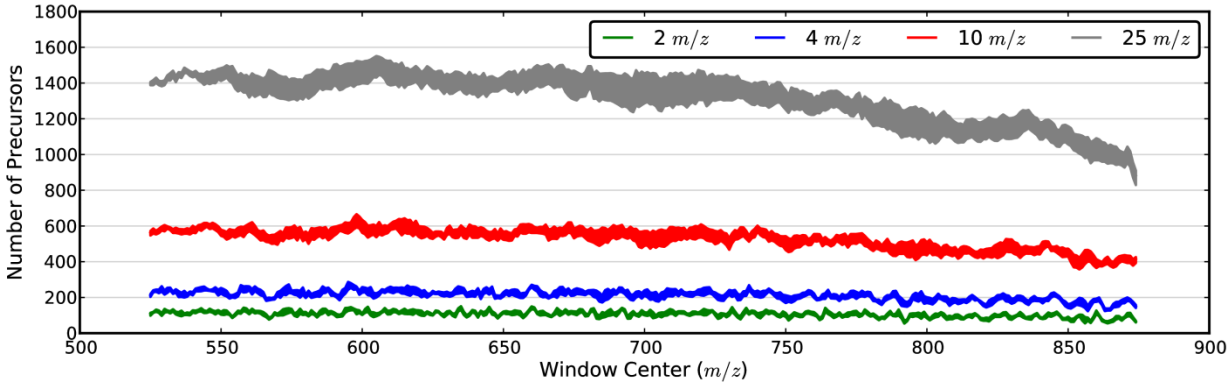
A *C. elegans* soluble lysate was analyzed on the Q-Exactive using a 10 *m/z* wide isolation window DIA approach in which the mass range from 500 *m/z* to 900 *m/z* is analyzed with consecutive 10 *m/z* wide isolation window targeted MS/MS scans and a 5 x 4 *m/z* MSX approach covering the same mass range. Fragment ion chromatograms for the peptide NIPGVDVMNVER++ were extracted using Skyline from both the wide-window DIA (**a** and **b**) and MSX data (**c** and **d**). The peak with the triangle hovering over it is the peak corresponding to this peptide as validated in a DDA analysis. In both the zoomed-out views (panels **a** and **c**) and zoomed-in views (panels **b** and **d**), the improved precursor selectivity of MSX results in much less non-target peaks.

Supplementary Figure 4: Simulation of the Impact of Isolation Width on Co-fragmentation



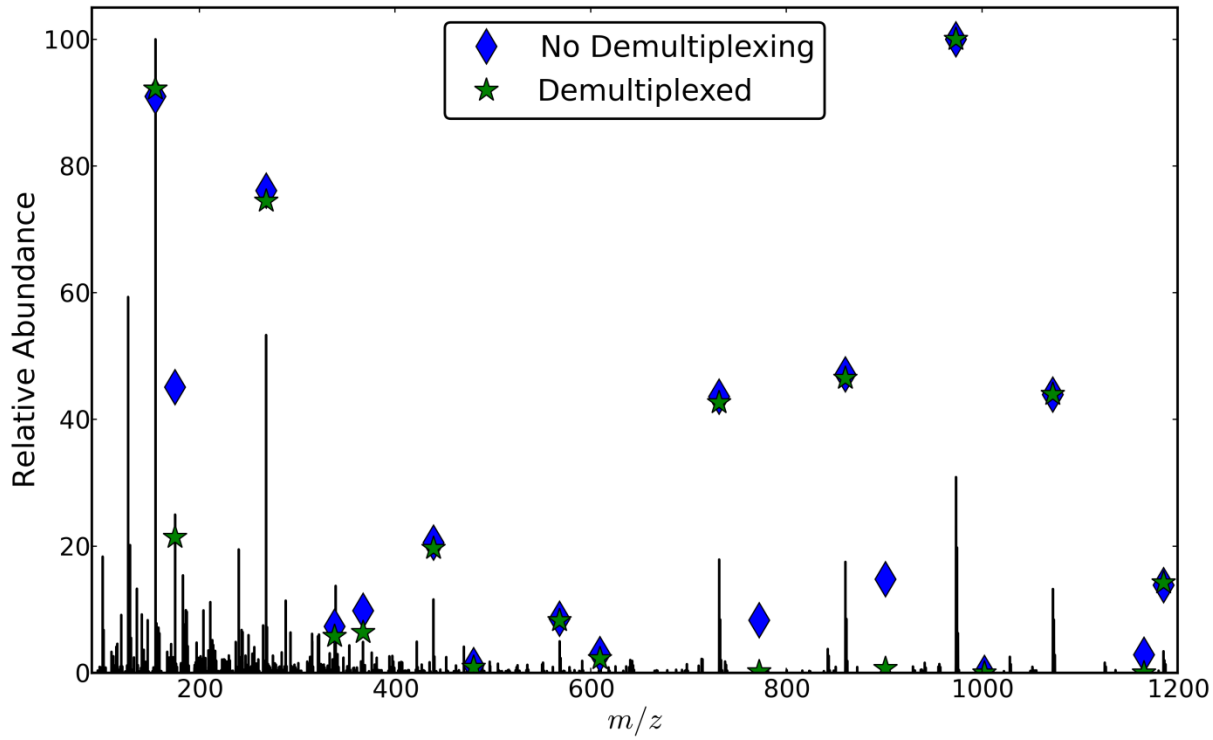
MS1 scans from 6 replicate MSX analyses of the *S. cerevisiae* lysate used as the background matrix in the spike-in experiments (see Methods section) were analyzed using Hardklör (Hoopmann M.R., *et. al.* Analytical Chemistry, 2007) and Krönik to detect persistent peptide isotope distributions (PPID). Next, isolation was simulated at every point in time of the instrument run where peptides elute (between 25 to 90 minutes) for isolation windows centered at integer values between 525 and 875 *m/z*. In each panel, the simulation is done with a different isolation width, and a histogram of the number of peptides co-fragmented is plotted. When 2 *m/z* wide isolation windows are used, more than 70% of the isolation windows do not contain a peptide precursor detected in the MS scan. When 25 *m/z* wide windows are used, ~85% of the windows isolate and fragment two or more peptide precursors and on average, 3.4 peptide precursors are co-fragmented in each window.

Supplementary Figure 5: Number of Persistent Peptide Isotope Distributions Detected in 2, 4, 10, and 25 *m/z*-wide Isolation Windows



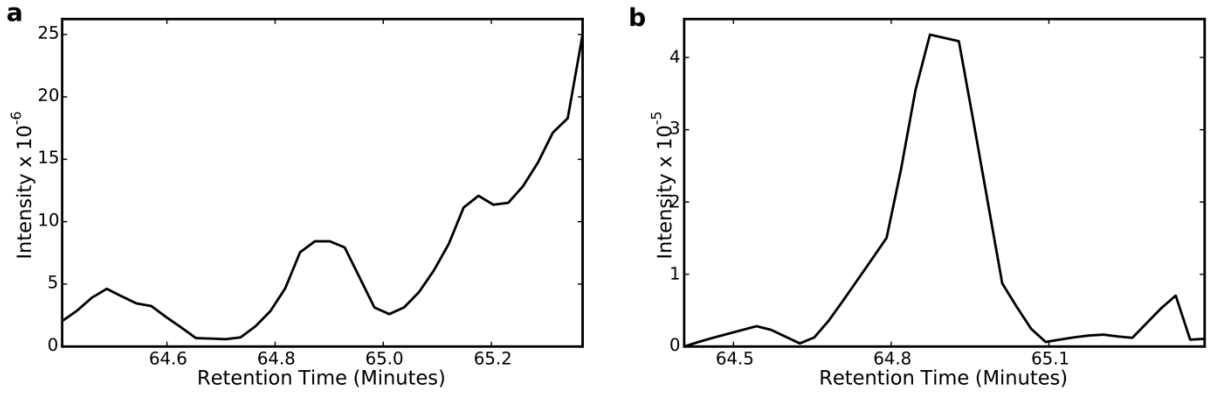
MS1 scans from six MSX runs of *S. cerevisiae* lysate MS were analyzed to detect persistent peptide isotope distributions (PPID) in the same manner as in Supplementary Figure 4. The number of PPID's detected in simulated isolation events with retention time between 25 and 90 minutes using windows of widths 2, 4, 10, and 25 *m/z* centered at every integer *m/z* between 525 and 875 is plotted. The upper and lower bounds of the lines for each isolation width is the maximum and minimum number of PPID's observed out of the six MSX runs analyzed.

Supplementary Figure 6: Demultiplexing Removes Fragment Ion Intereference and Increases Similarity to a DDA Spectrum



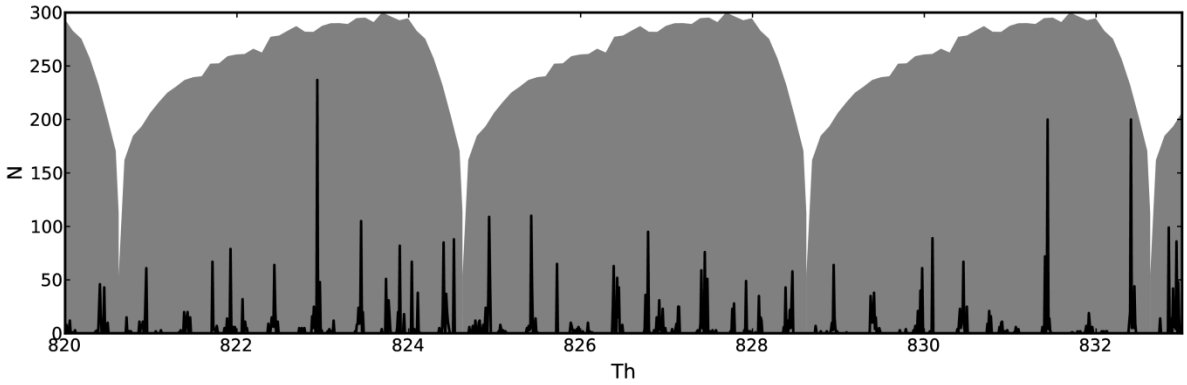
MSX and DDA data were acquired on a *S. cerevisiae* cell lysate (see Methods). The MS/MS spectrum from the DDA data acquired with a 2 *m/z* wide isolation window for the peptide GPLVLEYETYR++ is plotted. This MS/MS spectrum was confidently matched to GPLVLEYETYR (q<0.01) using a target-decoy database searching strategy with the SEQUEST search engine (Eng J. *et. al.*, JASMS, 1994) and Percolator (Käll L. *et. al.*, Nature Methods, 2007) for post-processing. Data for fragment ions of GPLVLEYETYR were extracted from the MSX data and integrated with and without demultiplexing (Methods, Figure 2). The relative abundance of the fragment ions from the MSX data with and without demultiplexing is overlaid on the DDA spectrum as green stars and blue diamonds, respectively. The relative abundances of the demultiplexed fragment ions match the DDA data better than the non-demultiplexed fragments (0.96 vs. 0.94 dot product). This is due to the removal of fragment ion interference by the demultiplexing and the relative lack of interference in the DDA spectrum due to its narrow isolation window (2 *m/z*).

Supplementary Figure 7: Summed Ion Current of MS1 and MSX Signals for LVNELTEFAK++ Peptide



Serum Albumin (part of a commercial six protein digest) was spiked into a complex matrix (soluble *S. cerevisiae* lysate) at amounts ranging from 50 attomoles – 100 femtomoles on column. MSX data were acquired with an MS1 scan interleaved every 10 scans. The summed ion current from the M, M+1, and M+2 masses used for MS1 quantitation is plotted in (**a**) at 1.02 femtomoles (the lower limit of detection by MS1). The summed ion current from the transitions used for quantifying the same peptide at the same spike in amount using MSX is plotted in (**b**).

Supplementary Figure 8: Optimized Placement of Isolation Window Edges



As described in the Methods section, the edges of the isolation windows for DIA experiments were optimized to place the edges in regions where peptides are unlikely to occur to minimize splitting of precursor signal and inefficient isolation. The black line is a histogram of the number of precursor peptides (+1, +2, and +3 charged) in the Bibliospec spectral library. The grey overlay shows the 4 *m/z* windows targeting a subset of the 500 – 900 *m/z* range covered in our MSX experiment. The quadrupole mass filter in the Q-Exactive does not isolate with equal efficiency across an isolated *m/z* window (i.e. the efficiency near the edges is less than at the center). The quadrupole isolation efficiency on our Q-Exactive is superimposed in grey.

Supplementary Tables

Supplementary Table 1: Transitions Used for Bovine Spike-In Protein Quantification

β-Lactoglobulin

Peptide	Transitions
GLDIQK+	b+(4) y+(2-5)
IPAVFK+	b+(1,3-5) y+(2-5)
TPEVDDEALEK++	b+(6-7) y+(4,7-9) y++(10)
VLVLDTDYK++	b+(3,6) y+(4-8)
VYVEELKPTPEGDLEILLQK+++	b+(2) y+(3-7,9-10) y++(11-14)

Carbonic Anhydrase

Peptide	Transitions
DFPIANGER++	y+(3-7) y++(7)
EPISVSSQQMLK++	b+(5,8) y+(3) b++(4,10) y++(9)
HNGPEHWHK++	b+(1-2,7) y+(2-4,7-8) b++(6)
QSPVDIDTK++	y+(3,5-6) y++(7)
VGDANPALQK++	b+(5) y+(2-9) y++(5,9)

Glutamate Dehydrogenase

Peptide	Transitions
DDGSWEVIEGYR++	b+(3) y+(3-10) b++(9,11) y++(8)
DSNYHLLMSVQESLER+++	b+(4-8) y+(2,5-6,8)
ELEDFK+	b+(3-5) y+(2-5)
GASIVEDK+	b+(3,5-6) y+(2-5)
HGGTIPIVPTAEFQDR++	b+(1-5,8) y+(2-4,6,9-13) y++(8,11)
HGGTIPIVPTAEFQDR+++	b+(1-2,5) y+(2-3,6-9) b++(6) y++(6,8)
LQHGTILGFPK++	b+(2-6,8-9) y+(3-9) y++(9)
NYTDEDLEK++	b+(2-3) y+(3,6-8)
TAAYVNAIEK++	y+(2-6,8-9)
VYNEAGVTFT++	b+(2-3) y+(2) b++(9)
YSTDVSVDEVK++	b+(1,4) y+(2-5,7,9-10) b++(4)

Lactoperoxidase

Peptide	Transitions
DYLPIVLGSEMQK++	y+(3-10) y++(10)
FGHMEVPSTVSR++	y+(6-7) b++(8) y++(5,7,11)
FWWENPGVFTEK++	b+(2-4) y+(3-10) y++(3,7)
GLQTVLK+	b+(4) y+(3-6)
IHGFDLAAINLQR++	b+(3,6,8,10) y+(4-12) y++(12)
IVGYLDEEGVLDQNR++	b+(3) y+(3,8) b++(6,10,12-13)

Serum Albumin

Peptide	Transitions
ATEEQLK+	b+(4-6) y+(2,5-6)
DAFLGSFLYEYSR+++	b+(6-7,10) b++(6,10)
DDSPDLPK+	b+(1,5-6) y+(2-3,5-6)
HLVDEPQNLIK++	b+(1-7,10) y+(3-10) b++(6-7)
HPEYAVSVLLR++	y+(6,8-10)
LGEYGFQNALIVR++	y+(2-6,9-12)
LVNELTEFAK++	b+(4-5) y+(3,5-9) y++(5,8)
LVVSTQTALA++	b+(8) y+(9) b++(4)
NYQEAK+	b+(2-5) y+(3-5)
TPVSEK+	b+(3-5) y+(2-5)

Normalization Peptides (Background matrix)

Peptide	Transitions
DNSQVFGVAR++	b+(2-3,5) y+(2-6)
ESTLHLVLR++	b+(2-3,5-7) y+(2-8) y++(5-8)

Supplementary Table 2: Window Centers for MSX Isolation

All windows have width 4.0018 *m/z*

Start	Stop	Center
500.4774	504.4792	502.4783
504.4792	508.481	506.4801
508.481	512.4828	510.4819
512.4828	516.4847	514.4837
516.4847	520.4865	518.4856
520.4865	524.4883	522.4874
524.4883	528.4901	526.4892
528.4901	532.4919	530.491
532.4919	536.4937	534.4928
536.4937	540.4956	538.4947
540.4956	544.4974	542.4965
544.4974	548.4992	546.4983
548.4992	552.501	550.5001
552.501	556.5028	554.5019
556.5028	560.5047	558.5038
560.5047	564.5065	562.5056
564.5065	568.5083	566.5074
568.5083	572.5101	570.5092
572.5101	576.5119	574.511
576.5119	580.5138	578.5128
580.5138	584.5156	582.5147
584.5156	588.5174	586.5165
588.5174	592.5192	590.5183
592.5192	596.521	594.5201
596.521	600.5229	598.5219
600.5229	604.5247	602.5238
604.5247	608.5265	606.5256
608.5265	612.5283	610.5274
612.5283	616.5301	614.5292
616.5301	620.5319	618.531
620.5319	624.5338	622.5329
624.5338	628.5356	626.5347
628.5356	632.5374	630.5365
632.5374	636.5392	634.5383
636.5392	640.541	638.5401
640.541	644.5429	642.5419
644.5429	648.5447	646.5438

648.5447	652.5465	650.5456
652.5465	656.5483	654.5474
656.5483	660.5501	658.5492
660.5501	664.552	662.551
664.552	668.5538	666.5529
668.5538	672.5556	670.5547
672.5556	676.5574	674.5565
676.5574	680.5592	678.5583
680.5592	684.561	682.5601
684.561	688.5629	686.562
688.5629	692.5647	690.5638
692.5647	696.5665	694.5656
696.5665	700.5683	698.5674
700.5683	704.5701	702.5692
704.5701	708.572	706.5711
708.572	712.5738	710.5729
712.5738	716.5756	714.5747
716.5756	720.5774	718.5765
720.5774	724.5792	722.5783
724.5792	728.5811	726.5801
728.5811	732.5829	730.582
732.5829	736.5847	734.5838
736.5847	740.5865	738.5856
740.5865	744.5883	742.5874
744.5883	748.5902	746.5892
748.5902	752.592	750.5911
752.592	756.5938	754.5929
756.5938	760.5956	758.5947
760.5956	764.5974	762.5965
764.5974	768.5992	766.5983
768.5992	772.6011	770.6002
772.6011	776.6029	774.602
776.6029	780.6047	778.6038
780.6047	784.6065	782.6056
784.6065	788.6083	786.6074
788.6083	792.6102	790.6093
792.6102	796.612	794.6111
796.612	800.6138	798.6129
800.6138	804.6156	802.6147
804.6156	808.6174	806.6165
808.6174	812.6193	810.6183

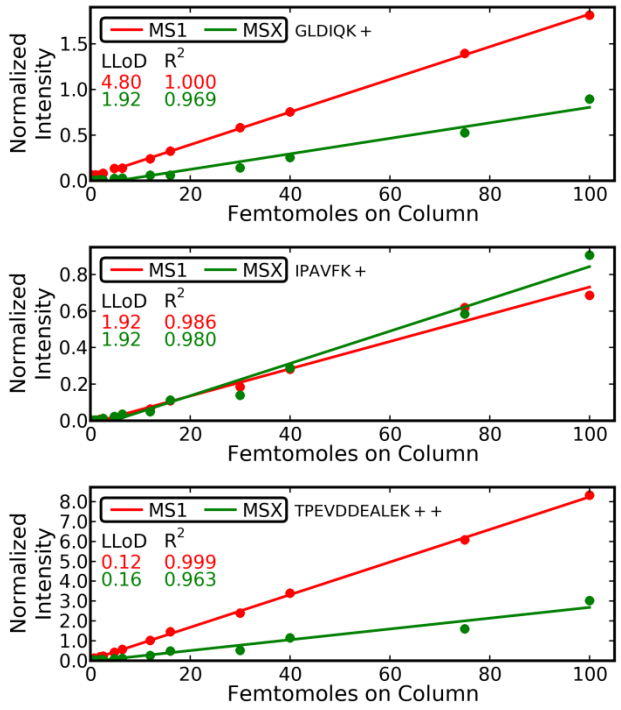
812.6193	816.6211	814.6202
816.6211	820.6229	818.622
820.6229	824.6247	822.6238
824.6247	828.6265	826.6256
828.6265	832.6284	830.6274
832.6284	836.6302	834.6293
836.6302	840.632	838.6311
840.632	844.6338	842.6329
844.6338	848.6356	846.6347
848.6356	852.6374	850.6365
852.6374	856.6393	854.6384
856.6393	860.6411	858.6402
860.6411	864.6429	862.642
864.6429	868.6447	866.6438
868.6447	872.6465	870.6456
872.6465	876.6484	874.6475
876.6484	880.6502	878.6493
880.6502	884.652	882.6511
884.652	888.6538	886.6529
888.6538	892.6556	890.6547
892.6556	896.6575	894.6565
896.6575	900.6593	898.6584

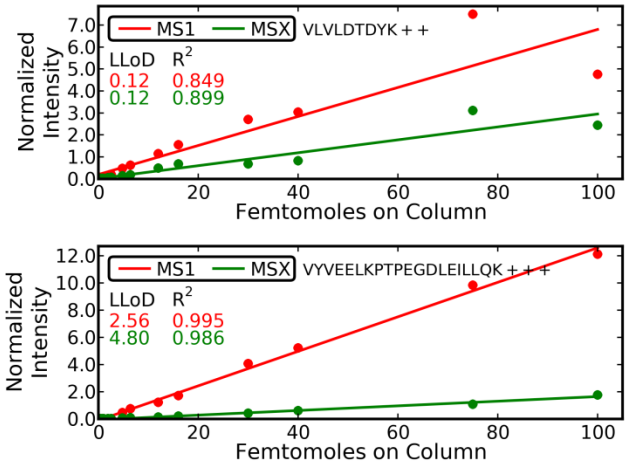
Supplementary Data Legends

Supplementary Data: Spike in data for 36 peptides (5 proteins)

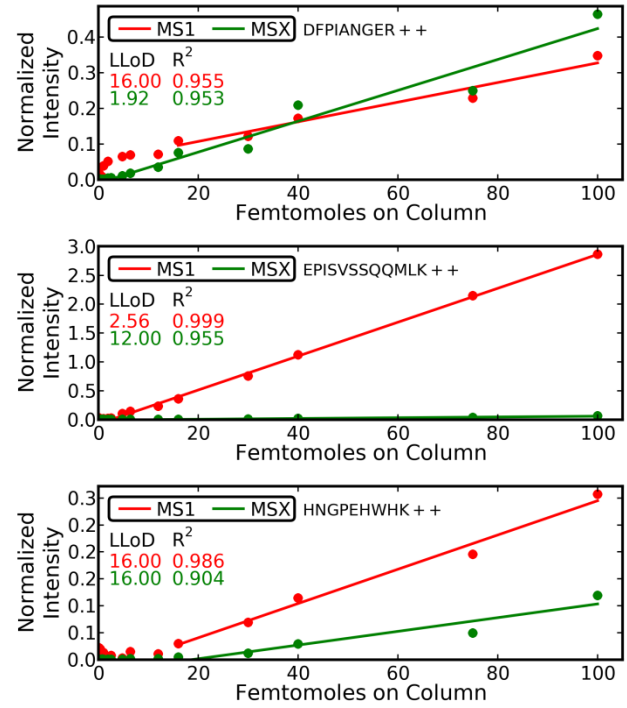
A commercial six protein digest was spiked into a complex matrix (soluble *S. cerevisiae* lysate) at amounts ranging from 50 attomoles – 100 femtomoles on column. MSX data were acquired with an MS1 scan interleaved every 10 scans. The normalized signal intensity for each spike in point is plotted for each peptide as well as the regression line fit to all points greater than or equal to the lower limit of detection.

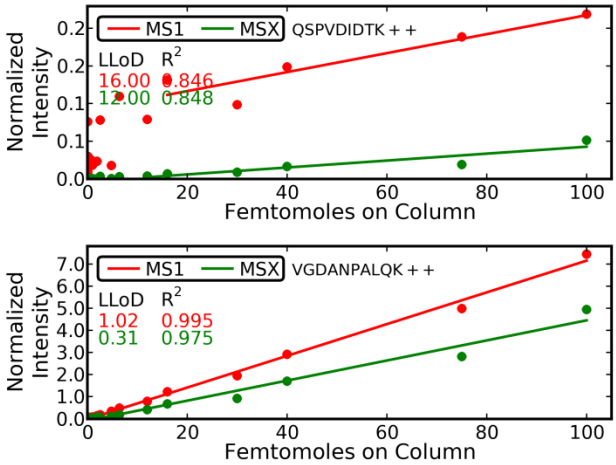
β-Lactoglobulin



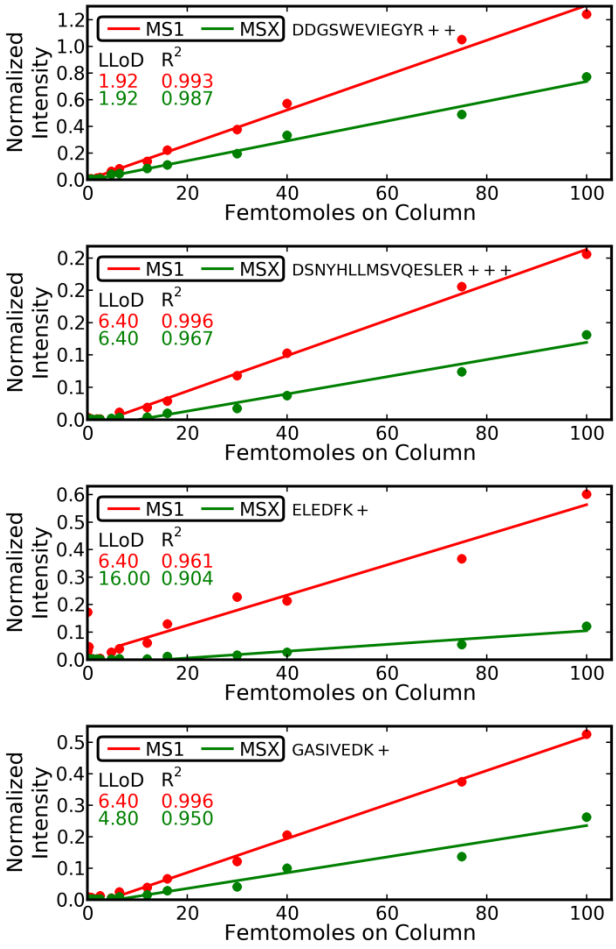


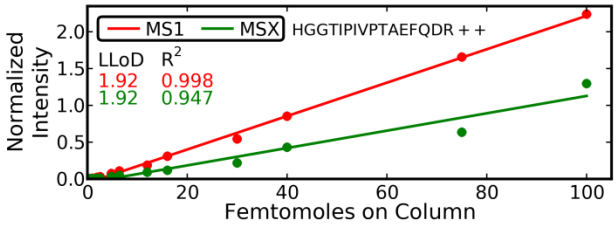
Carbonic Anhydrase

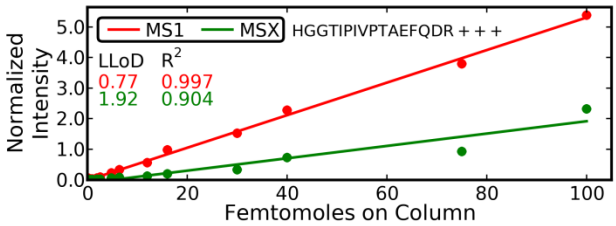


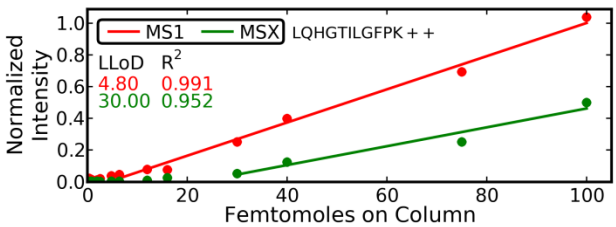


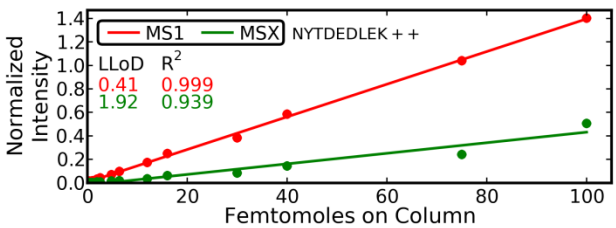
Glutamate Dehydrogenase

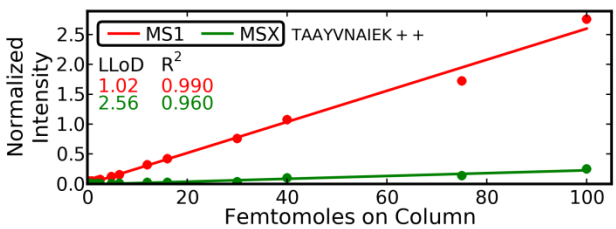


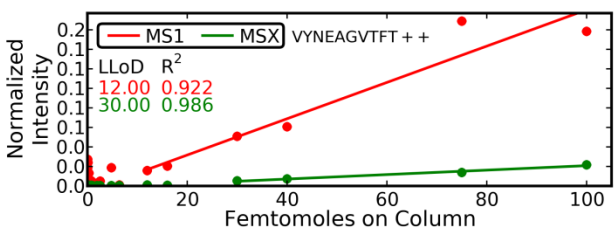

MS1
MSX
HGGTIPIVPTAEFQDR + +
LLoD R2
1.92 0.998
1.92 0.947
Normalized Intensity
Femtomoles on Column

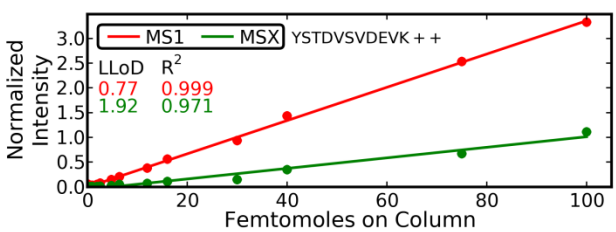

MS1
MSX
HGGTIPIVPTAEFQDR + + +
LLoD R2
0.77 0.997
1.92 0.904
Normalized Intensity
Femtomoles on Column


MS1
MSX
LQHGTILGFPK + +
LLoD R2
4.80 0.991
30.00 0.952
Normalized Intensity
Femtomoles on Column

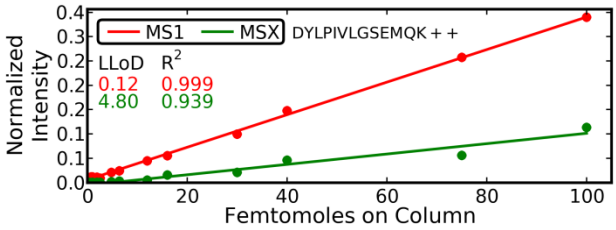

MS1
MSX
NYTDEDLEK + +
LLoD R2
0.41 0.999
1.92 0.939
Normalized Intensity
Femtomoles on Column

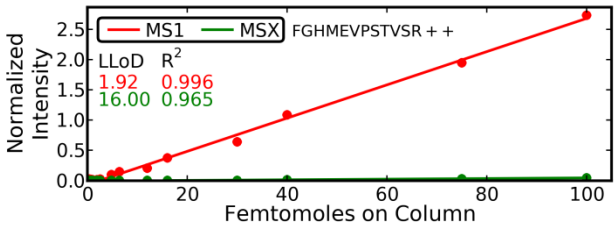

MS1
MSX
TAAYVNAIEK + +
LLoD R2
1.02 0.990
2.56 0.960
Normalized Intensity
Femtomoles on Column

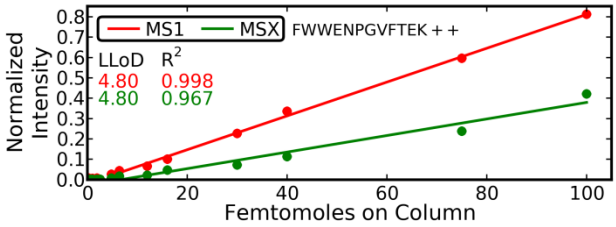

MS1
MSX
VYNEAGVTFT + +
LLoD R2
12.00 0.922
30.00 0.986
Normalized Intensity
Femtomoles on Column

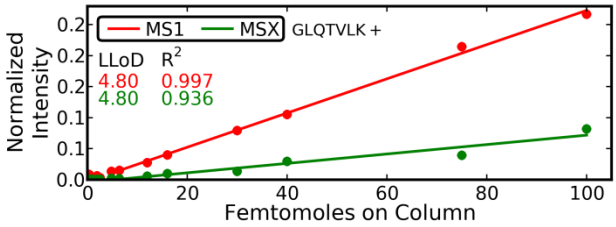

MS1
MSX
YSTDVSVDEVK + +
LLoD R2
0.77 0.999
1.92 0.971
Normalized Intensity
Femtomoles on Column

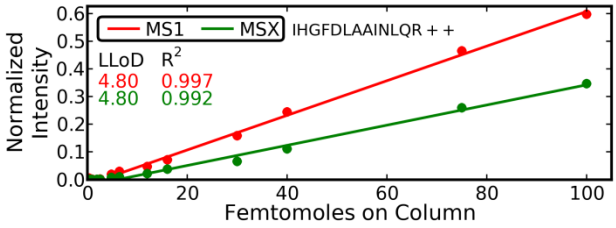
Lactoperoxidase

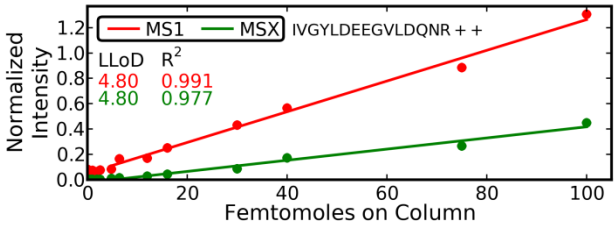




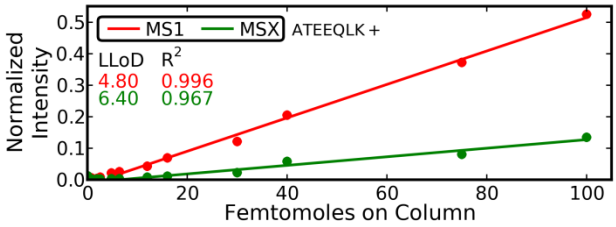


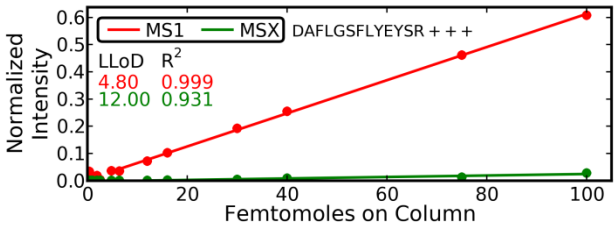


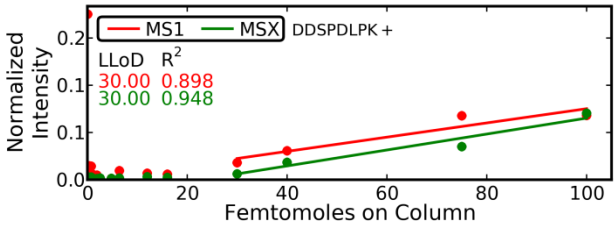


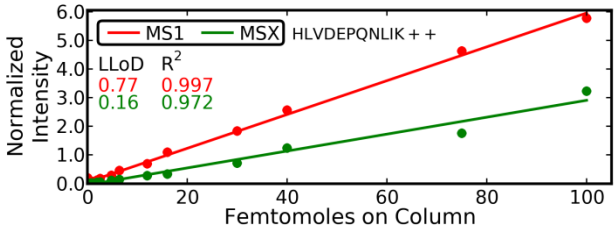


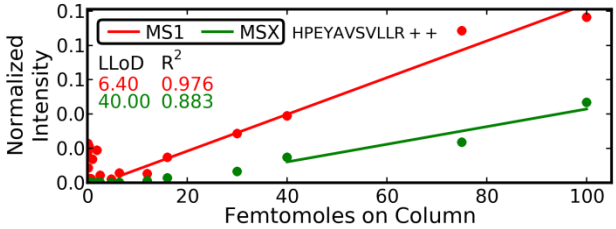
Serum Albumin

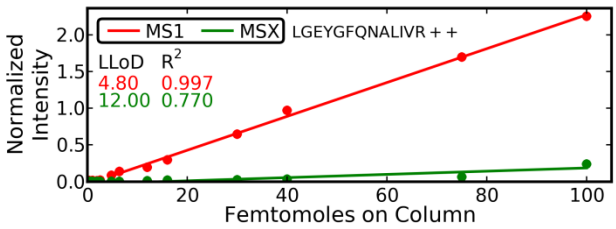


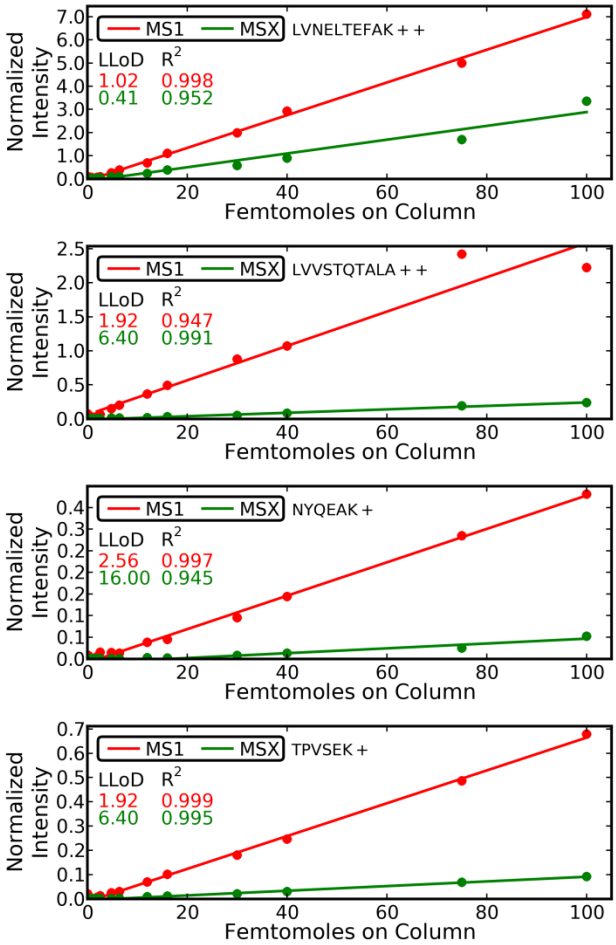











MS1
MSX
LVNELTEFAK + +
LLoD R2
1.02 0.998
0.41 0.952
LVVSTQTALA + +
LLoD R2
1.92 0.947
6.40 0.991
NYQEAK +
LLoD R2
2.56 0.997
16.00 0.945
TPVSEK +
LLoD R2
1.92 0.999
6.40 0.995
Normalized Intensity
Femtomoles on Column

Supplementary Note: Frequently Asked Questions

What do we mean by precursor selectivity in the context of a DIA experiment?

In tandem mass spectrometry, there are three basic steps:

1) Isolation of a population of precursor ions

2) Fragmentation of that ion population

3) Mass analysis of the fragment ion population

Precursor selectivity refers to the resolution of the isolation of precursor ions in step 1. In this manuscript, this resolution is the isolation width used on the mass selective quadrupole in the Q-Exactive instrument. Methods that use smaller isolation widths isolate a targeted *m/z* with higher resolution, thus providing more specific knowledge of what any precursors isolated may be.

Why is improved precursor selectivity necessary?

Improved precursor selectivity helps DIA analysis in two ways:

1) Improved qualitative confirmation of a target peptide because of reduced precursor co-isolation

Once the fragment ion chromatograms for a peptide of interest have been extracted, if the peptide is present, there will be a set of co-eluting fragment ion peaks from that peptide. However, in almost every case there will be more than

one peak with co-eluting fragment ions, making it difficult to determine which is the best match to the peptide of interest. These multiple peaks occur because other peptide precursors fall in the same isolation window, and have multiple fragment ions with close, or exactly equal *m/z* (e.g. post translational modifications) as those from the peptide of interest. With higher precursor selectivity, this happens less frequently, but is still difficult to avoid.

Supplementary Fig. 1 shows an example of this. In panel **a**, fragment ion chromatograms extracted from 10-*m/z* wide isolation window DIA data are shown for the peptide SLQDIIAILGMDELSEEDKLTVSR+++. The peak annotated with a triangle is the peak for the unmodified peptide, the peak with the circle is the peptide with a methionine oxidation. It is difficult to determine which peak is the modified vs. unmodified peptide because they share similar fragmentation patterns (this particular case was identified with the aide of DDA data). The modified and unmodified precursor are 5.3 *m/z* apart, and were co-isolated in the 10 *m/z* wide isolation window. With a smaller isolation window (higher precursor selectivity), say 4 *m/z*, these two forms would not be co-isolated and it would be easier to select the correct peak for the unmodified peptide. This is the case in the MSX analysis (panel **b**).

The number of MS1 features in each isolation window can give a rough estimate of the number of precursors isolated in each MS/MS spectrum (**Supplementary Fig. 4**). It is not a perfect estimate, because some features may not be observable in the MS spectrum but generate observable peaks in

the MS/MS spectrum. Regardless, it is clear that even with 4 *m/z* wide isolation windows, there is an appreciable number of scans fragmenting more than one precursor. Additionally, **Supplementary Fig. 5** shows the total number of precursors observed over the course of an experiment with different isolation widths.

2) Improved quantitation due to reduced chemical noise

Once the correct peak is known, the MS/MS signal can be used for quantitation by integrating the area under the curve for a set of extracted fragment ion chromatograms (transitions). With low precursor selectivity, it is more likely that one or more precursors eluting at the same time with overlapping fragment ion masses will be co-fragmented, causing interference. These transitions with interference are usually not used for quantitation. By improving precursor selectivity, more transitions can be used for quantitation. Having more transitions for quantitation is preferable due to increased signal-to-noise.

There have been significant efforts in our field to improve selectivity. We know that with the 0.7 *m/z* isolation used in standard SRM experiments we can get co-eluting and co-fragmenting species. Thermo built quadrupoles that could isolate at 0.2 *m/z* because unit resolution wasn't sufficient in some cases. AB Sciex and others have added differential ion mobility to the front end of their triple quadrupoles to provide an extra dimension of separation prior to the Q1

isolation. Furthermore, some labs have had to resort to MS3 to get sufficient selectivity for quantitative analysis.

What is the difference between precursor selectivity improvements in the context of a database search vs. precursor selectivity here?

In the context of a database search, algorithms such as XDIA (Carvalho P.C. *et. al.*, Bioinformatics; 2010) and Bullseye (Hsieh E.J. *et. al.*, JPR, 2010) can improve precursor selectivity. XDIA assigns precursor masses and charges likely to have been fragmented in a given DIA spectrum using an analysis of charged-reduced precursors in the MS/MS spectrum and nearby MS1 spectra. Bullseye determines a more accurate precursor *m/z* to MS/MS spectra by analyzing high-resolution MS spectra. This improves precursor selectivity for the database search because only candidate spectra corresponding to the precursors detected, rather than all in the isolation window, are searched. However, the fragmentation signals contributed by each precursor remain mixed. The fragment ion spectrum is unchanged (save removing unfragmented precursor by XDIA) because the precursor selectivity of the spectrum was not improved. If spectrum precursor selectivity were improved, the complexity of the resulting fragment ion spectrum should be reduced because it represents contributions from fewer precursors.

The MSX algorithm unmixes the original fragmentation spectrum into component spectra corresponding to each of the 5 4 *m/z* windows that were isolated together. In this process, precursor selectivity is improved from 20

m/z to 4 *m/z* and the fragment ion signal is unmixed. This unmixing of the fragment ion signal is key because it removes fragment ion interference which improves the accuracy of quantitation.

How does MS1 quantitation compare to MSX quantitation?

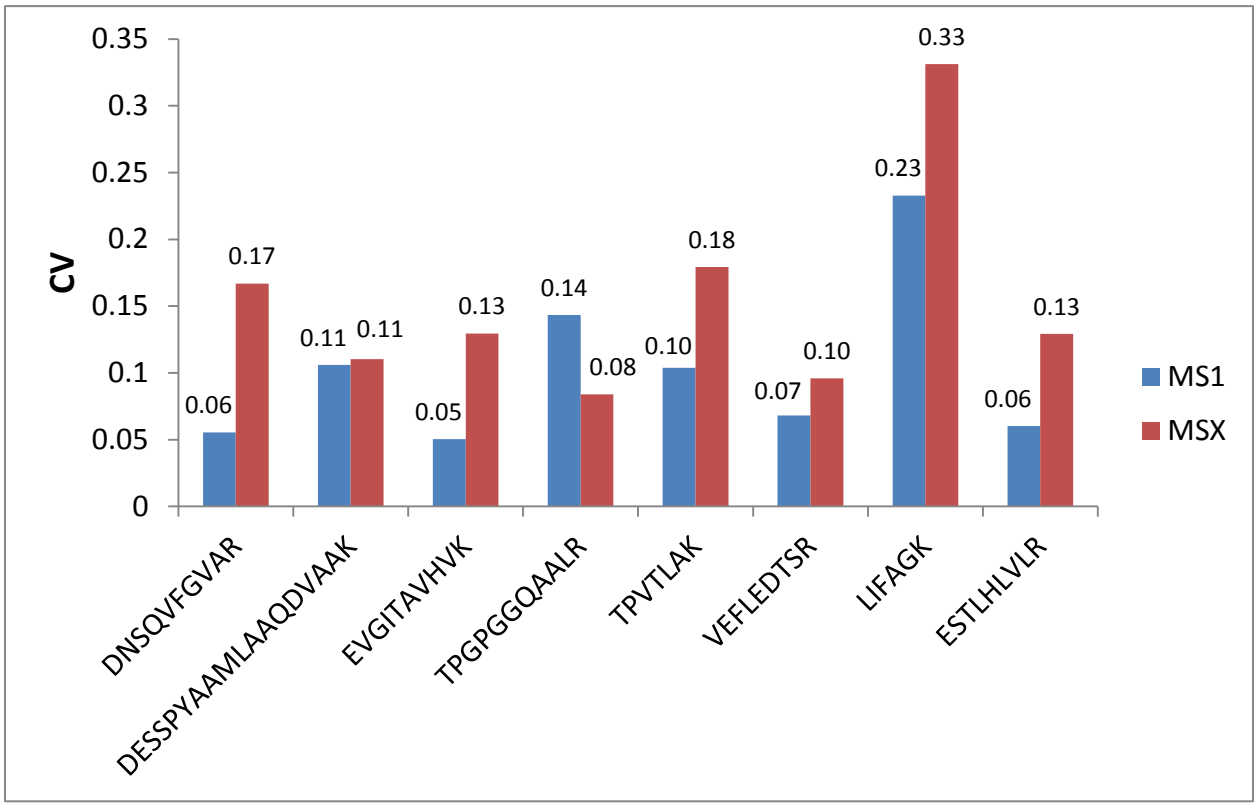
In our spike-in experiment with 36 peptides, we found that the lower limit of detection was 8.66 and 4.98 femtomoles for MSX and MS1 respectively. All peptides had a linear response above the limit of detection with R^2 values of the regression lines averaging 0.95 and 0.98 for MSX and MS1 respectively. The CV of 18 replicate measurements of 6 peptides averaged 0.1 and 0.15 by MS1 and MSX respectively. The standard deviation in these measurements was 1.56×10^8 and 2.29×10^7 for MS1 and MSX respectively. Figures depicting reproducibility are shown at the end of this response.

We compared the quantitation by MSX to that using MS1 because integration of the area under MS1 peaks is a common technique for peptide quantitation. In our experiments, MSX and MS1 data were acquired in the same sample runs (see **Methods**). This means that MSX and MS1 data can be combined and the advantages of both techniques realized. MS1 quantitation should perform better in low-complexity regions where there is less chemical noise (examples in **Supplementary Data**). MSX data provides the structural selectivity necessary for peptide identification, and quantitation in high-complexity regions (**Fig. 3**, **Supplementary Fig. 7**). For 7 of the 36 peptides quantified, MSX quantitation was more sensitive than MS1 quantitation with a

3.4 fold improvement in the limit of detection on average (**Supplementary Data**).

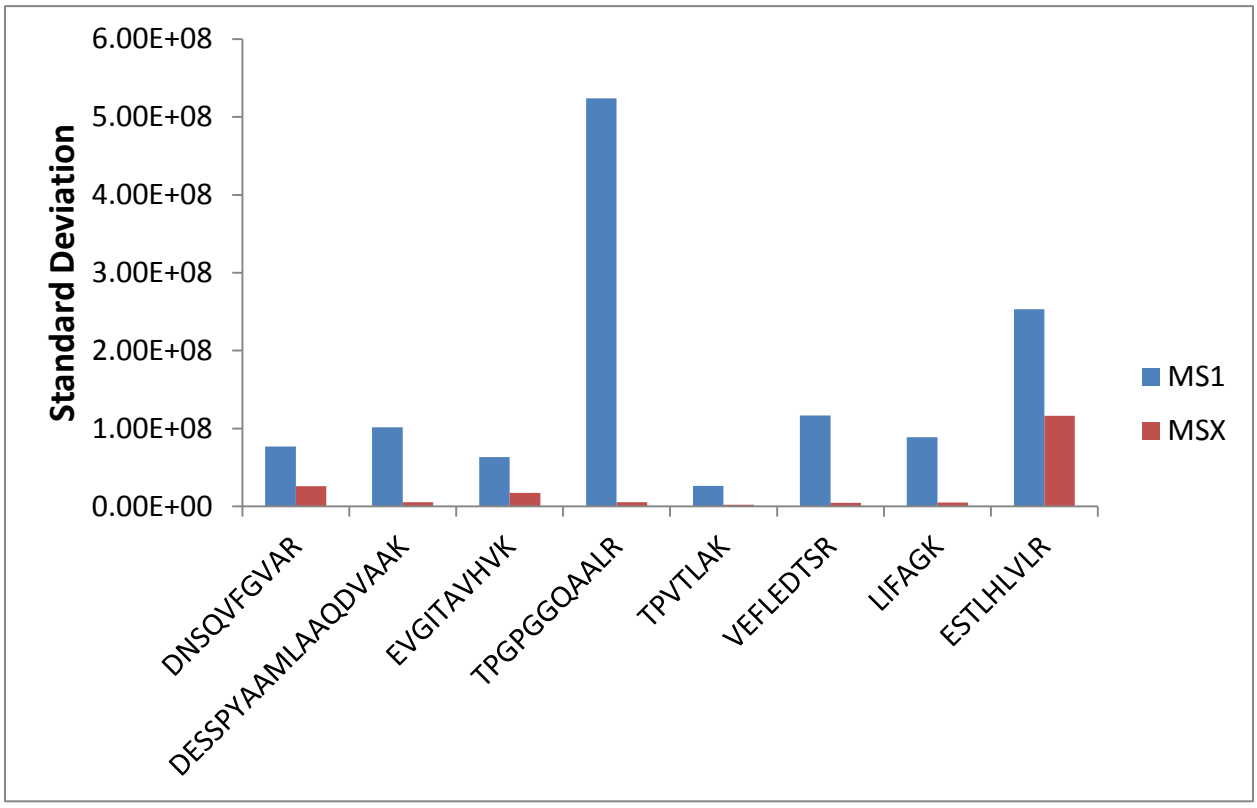
If there is no chemical noise present, MS1 will be more sensitive. Consider an ion beam for a peptide precursor ion of 10^6 ions/sec. If MS/MS is performed on this ion beam with 100% fragmentation of the precursor, zero losses, and the signal split between 20 equally abundant fragments, then in this absolute best case scenario each fragment would have an ion beam of $5x10^4$ ions/sec. If there is no chemical noise in the MS measurement, it will perform better because all of the ions are in a single beam. However, in more complex regions, with more chemical noise, chemical interferences should decrease in intensity faster than the signal from the target analyte and MS/MS quantification should improve sensitivity. The 1983 text "Tandem Mass Spectrometry" by Fred McLafferty is an excellent reference for more information on this topic.

Reproducibility of Peptide Quantitation (CV)



The reproducibility of quantitation by MS1 is compared to MSX by calculating the CV in peak area for peptides in the *S. cerevisiae* background matrix from the spike-in experiments over 18 injections. MS1 and MSX data were acquired in the same instrument runs. The average CV by MS1 and MSX is 0.1, and 0.15 respectively. The signal for MS/MS-based quantitation is reduced, causing an increase in the relative error (CV).

Reproducibility of Peptide Quantitation (StDev)



The reproducibility of quantitation by MS1 is compared to MSX by calculating the standard deviation in the peak area for peptides in the *S. cerevisiae* background matrix from the spike-in experiments over 18 injections. MS1 and MSX data were acquired in the same instrument runs. The standard deviation in peak areas by MSX is lower, as expected due to reduced variance in peaks of lower intensity. The average standard deviation is 1.56×10^8 and 2.29×10^7 for MS1 and MSX respectively.

Can fragment ion based quantitation from MSX be combined with MS1 quantitation?

Yes, and we think it should be. See the response to the question "How does MS1 quantitation compare to MSX quantitation?"

How does MSX compare to quantitation by DDA MS/MS?

Oftentimes, quantitation by DDA MS/MS uses the acquired MS/MS scans for peptide identification, and maps those to MS1 peaks for quantitation. For a comparison of MS1 and MSX quantitation, please see the "How does MS1 quantitation compare to MSX quantitation?" answer.

In MSX, we collect the product ions of all the peptides in the selected mass range all of the time. In DDA, we semi-randomly select peptides based on intensity. Therefore in MSX we can extract pseudo SRM traces for any peptide of interest, modified or unmodified, and if it was within the limit of detection integrate a peak. In contrast, only peptides that have been selected for MS/MS and identified using a search engine can be quantified. We like to think of MSX (and all other DIA methods) as essentially being a targeted SRM assay on all peptides.

How does MSX compare to SWATH and MS^E? What samples is it most applicable to?

The three techniques are all DIA techniques which acquire high mass accuracy fragment ion measurements which can be used for identifying or quantifying peptides. One difference between all of these methods is precursor selectivity. MSX has the greatest precursor selectivity and MS^E (Plumb R.S., *Rapid Communications in Mass Spectrometry,* 2006), has the worst. In particular, SWATH (Gillet L.C., *Molecular and Cellular Proteomics*, 2012) and MS^E will both struggle when analyzing complex samples where peptide modifications and peptides with similar sequence will often be isolated and fragmented in the

same MS/MS scan. With 25-*m/z* isolation windows, peptides and their modified forms (eg. n-terminal acetylation, oxidized methionine) may be isolated in the same window and will be difficult to differentiate due to overlapping fragmentation patterns. If they co-fragment, substantial fragment ion interference will hinder quantitation.

MSX covers a wide precursor mass range (500 - 900 *m/z*) with better precursor selectivity than competing DIA methods such as SWATH and MSe. The increased precursor selectivity will improve quantitative performance in the analysis of complex biological samples (cell culture, cell/tissue lysates) where the improved selectivity will reduce fragment ion interference. However, in less complex samples such as immunoaffinity enrichments, MSX may not offer an advantage over existing techniques. With that being said, it should be noted that the reported MSX technique is easily modified to use wider isolation windows which would make it more applicable to use in less complex samples.

How do the limits of identification and quantitation compare to other existing approaches?

It is difficult to make comparisons between these methods because all methods are only general frameworks for data acquisition, each of which are rarely implemented in the same manner between labs.

It is our opinion that the best data acquisition technique depends on the sample being analyzed, instrument platform, and chosen figures of merit. Combined with the lack of a single optimal implementation of any one of these

techniques, it is difficult to make broad comparisons without a very rigorous study outside the scope of this manuscript.

A summary of some of the specific parameters impacting the performance of each data acquisition technique follows:

DDA -- number of dependent scans per parent scan, MS1 resolution, MS2 resolution, maximum fill time/AGC target for trapping instruments, number of microscans, charge-state based precursor selection, precursor intensity threshold, dynamic exclusion list length, dynamic exclusion time, exclusion list settings, collision energy

SWATH -- isolation window width, scan time (# of microscans for signal averaging), mass range covered, collision energy

MS^E -- Collision energy, scan time (# of microscans for signal averaging)

MSX -- isolation window width, number of isolation windows per scan, mass range to cover, MS1 resolving power, MS2 resolving power, AGC target, maximum fill time, collision energy

What types of instrumentation is the MSX approach applicable to?

Performing an MSX experiment only makes sense on hardware where the isolation and collisional activation of the peptides is fast relative to the mass analysis. In this case, we can isolate and activate 5 different precursor windows in the same time it takes to collect a spectrum at 17,500 resolving power in the Orbitrap. In contrast, this would not make any sense to perform

on a Q-TOF type hybrid instrument where the precursor isolation and activation is slow relative to the time-of-flight mass analyzer.

As far as the required hardware configuration to perform a MSX DIA experiment, the Q-Exactive is currently the only commercially available instrument that can perform this experiment. We require an instrument that can isolate a precursor mass range, activate it, and store the fragments in an ion trap. The ion trap must be able to receive multiple fills with minimal ion losses of the ions already stored in the trap. This ion trap is then used to send the mixed ion population to a high resolution mass analyzer with a large dynamic range.

How does AGC work for a multiplexed approach? How is space charging avoided?

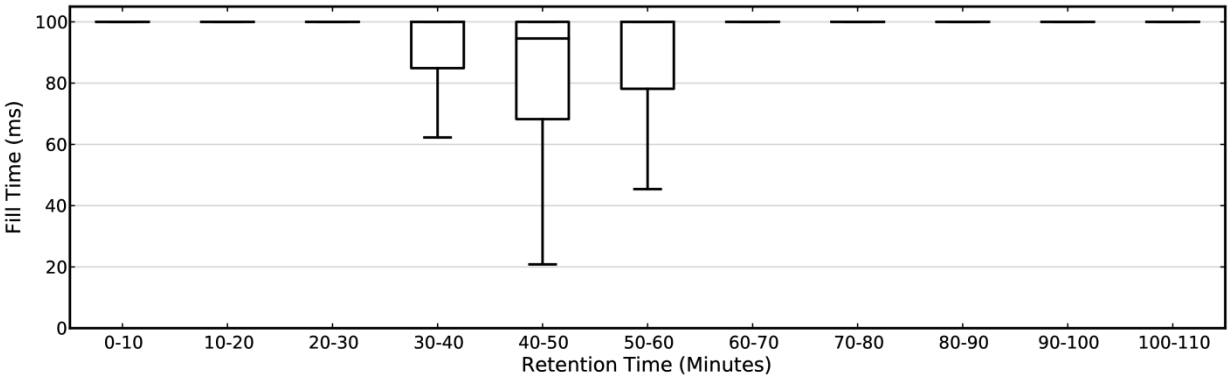
To avoid space charging, automated gain control (AGC) is used to avoid overfilling the ion trap. With AGC, a target number of ions is chosen (500,000 in this manuscript) as a maximum number of ions that are allowed to enter the trap in a given scan. For MSX, AGC is a two-step process. For each MS/MS scan:

1) The most recent MS scan is analyzed to determine the ion current (ions / second) for the 5 windows to be isolated. This analysis is done by summing up the total MS1 signal in the 5 discontiguous 4 *m/z* windows that will be isolated in the MS/MS scan.

2) The fill time is calculated by dividing the AGC target (# ions) by the ion current calculated in step 1 (ions / second). This is the amount of time that each of the 5 windows should be filled for. Each window is filled for this amount of time, with a maximum of 20 ms.

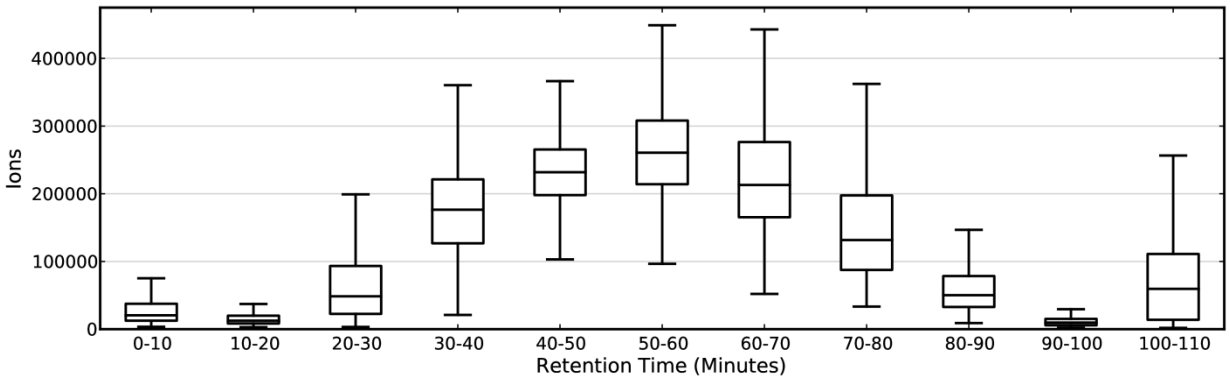
There are no ion trap or space charge issues in these data. In fact, with the data acquired here we are significantly underfilling the trap. We have a maximum inject time of 100 milliseconds for the 5 separate precursor windows (20 milliseconds for each individual window). This relatively low inject time means that a majority of our precursor windows (85% of the scans) are maxed at the maximum inject time. The fill times and number of ions entering the trap are plotted below.

MSX Fill Times



Fill times from an MSX analysis of the soluble *S. cerevisiae* lysate used as a background for the spike-in experiments are plotted. The fill times are plotted as the total time spent collecting ions for all five windows in a scan combined. The majority of the scans (85%) reach the maximum fill time (100 ms). The whiskers are the most extreme data point within 1.5 times the inner quartile range.

MSX Ion Counts



The number of ions in the trap for MS/MS scans from an MSX analysis of the soluble *S. cerevisiae* lysate used as background in the spike-in experiment is plotted. Less ions enter the trap at lower and higher retention times due to less total ion current at the extremes of the acetonitrile gradient used to separate the peptides. The AGC target for these data is 500,000 ions. The whiskers are the most extreme data point within 1.5 times the inner quartile range.

MSX variable time between measurements (randomness) could result in up to 7 seconds between scans. How does this affect the results?

This could happen, but it would only be that way for one scan cycle. Our mean +/- SD is 3.64 +/- 1.49 seconds, so while a 7 second gap between measurements is possible, it is extremely rare.

How does bias for transmission away from the isolation window center impact quantitation of ions near the edges of the isolation window?

Bias for transmission away from the window happens, but will not impact quantitation. If the transmission on the edge of the window is 95% of the transmission at the center of the window in every single scan, then there will

be no effect on the quantitation because the bias will remain consistent every time the window is isolated.

If the peptide at the edge of the isolation window it may generate a response 90% of what it would be in the center, but it will do that consistently regardless of the abundance of that peptide. The relationship between peptide abundance and signal intensity does not change, just the absolute signal intensity. For example, it is well known that with any mass spectrometer, the instrument ion optics may transmit lower *m/z* values differently than higher *m/z* values or an electron multiplier responds differently to low *m/z* ions than high *m/z* ions -- just like transmission might be different on the edge of a window relative to the center. These are all well-known effects but do not affect quantitation and certainly wouldn't cause any specific quantitation issues. In SRM, with a 0.7 *m/z* Q1 isolation width, every doubly charged peptide loses isotope peaks -- we don't think this is a problem. It is no different than with a 4 *m/z*-wide window. Likewise, if an isotope distribution is split between two windows, there will be only a signal loss but no effect on quantitation -- as long as the split is consistent in every scan.

We put the edges into "forbidden" zones (Frahm J.L. *et. al.*, JASMS; 2006) where it is extremely unlikely for a peptide monoisotopic *m/z* to be located. An example of the use of these forbidden zones in an unrelated application is presented in Egertson J.D. *et al*, JASMS; 2012. **Supplementary Table 2** lists the locations of the 100 x 4 *m/z* windows (the windows are actually 4.002 *m/z*

wide). **Supplementary Fig. 8** shows the "forbidden zones" in a narrow m/z range and demonstrates that the bias against the monoisotopic mass is minimal as long as the window locations are put in the correct spot.

Would a higher MS/MS resolving power improve the results of an MSX analysis or a non-multiplexing DIA analysis?

Product ion resolution is useful to some point but it is limited and in many ways less useful than precursor ion resolution. Unfortunately, a lot of peptides have similar sequence but different precursor mass. In these cases, no amount of product ion resolution can help because the product ion fragments have the exact same elemental composition. Consider the case of an oxidized methionine (**Supplementary Fig. 1**). In our case, the MSX DIA data would not show signal from the oxidized methionine when targeting the unmodified peptide precursor *m/z*. However, with 10 *m/z* wide isolation windows, signals show up in the data for both peptides and 50% of the predicted fragments are identical. Using a 25 *m/z*-wide window, peptides with modifications of ~50 and ~75 amu difference in mass for doubly and triply charged precursors respectively would greatly interfere with the qualitative detection of unmodified peptides. An n-terminal acetylation is easy to distinguish between using MSX but would be difficult with 25 *m/z* wide windows because the y-ion series between the modified and unmodified peptides would be identical.

A 4 *m/z* wide window means that we only have interference from peptides ~8 amu and ~12 amu wide away with the same fragment ions for doubly and

triply charged peptides. This relatively high selectivity on the precursor means that the only prevalent modification that we struggle with is deamidation -- which is also a problem using unit resolution precursor isolation.

Additionally, on the Q-Exactive, increased resolving power comes at the cost of instrument scan speed. The maximum resolving power that could be used with 20 *m/z* windows is 35,000 on this instrument due to the necessity of having a duty cycle of 3.5 seconds or less. At 35,000 resolving power, the orbitrap takes 128 ms to acquire a transient. With the simplifying assumption that each scan takes 128 ms (no scan overhead time, AGC scan time unaccounted for), the amount of time it would take to perform the 20 MS/MS scans required to cover 500-900 *m/z* is 2.56 seconds. The next highest resolving power (70,000) would bring the duty cycle time up to 5.12 seconds at the minimum, which is much too slow on an HPLC time scale. It may be possible to acquire data at 35,000 R.P. rather than 17,500 as in this manuscript, but certainly not any higher.

It is important to note that an increase of resolving power to 35,000 is applicable to both the MSX and continuous window approaches. The MSX method and the continuous 20 *m/z* wide isolation window approach both isolate 20 *m/z* of the mass range per scan and therefore require the same number of scans to cover the 500-900 *m/z* range. Additionally, for both approaches, the scan time is limited by the amount of time dedicated to mass analysis in the orbitrap since ion fills and mass analysis occur simultaneously.

Therefore, the cycle time for both approaches should be the same, as well as the selectivity on the MS/MS level. However, the MSX technique would have improved precursor selectivity over the continuously acquired wide windows.

Why not cover a wider *m/z* range than 500-900 *m/z*? Why 5 4-*m/z* windows per scan instead of a different combination like 3 8-*m/z* windows?

We agree that if there was no trade off in covering a wider range it would be advantageous. It turns out that most of the best responding peptides from a tryptic digest fall in this 500-900 *m/z* range (Canterbury JD *et. al.*; ASMS 2010). Once the mass range is expanded outside of this range, there are diminishing returns in the number of peptides analyzed due to the non-uniform distribution of strong responding peptides across the *m/z* range. Is it worth a significant loss in selectivity to cover a mass range where few peptides are going to be measurable? This question will be answered differently by different investigators depending on the biological question they are trying to address.

We reduced the mass range in comparison to Gillet *et. al.* as a compromise between mass range coverage, and duty cycle time. It would be possible to cover a wider mass range with these techniques by increasing precursor isolation window size (i.e. 5 x 6 *m/z* wide windows), increasing the number of precursors isolated per scan (i.e. 6 x 4 *m/z* wide windows), or simply accepting a slower duty cycle. However, selecting the optimal mass range to cover,

isolation width, and number of isolation windows per scan is a non-trivial exercise which will depend on the complexity of the sample being analyzed.

Where can the source code for MSX deconvolution be found?

The code is available as part of the open source Skyline software codebase developed by our lab which itself is part of the larger Proteowizard project. The proteowizard code can be perused in a web browser at: http://sourceforge.net/p/proteowizard/code or checked out from its SVN repository: https://svn.code.sf.net/p/proteowizard/code/trunk/pwiz. The code is available under a liberal Apache 2.0 Open Source license.