Supporting Information for:

Multiplexed Quantification for Data-Independent Acquisition

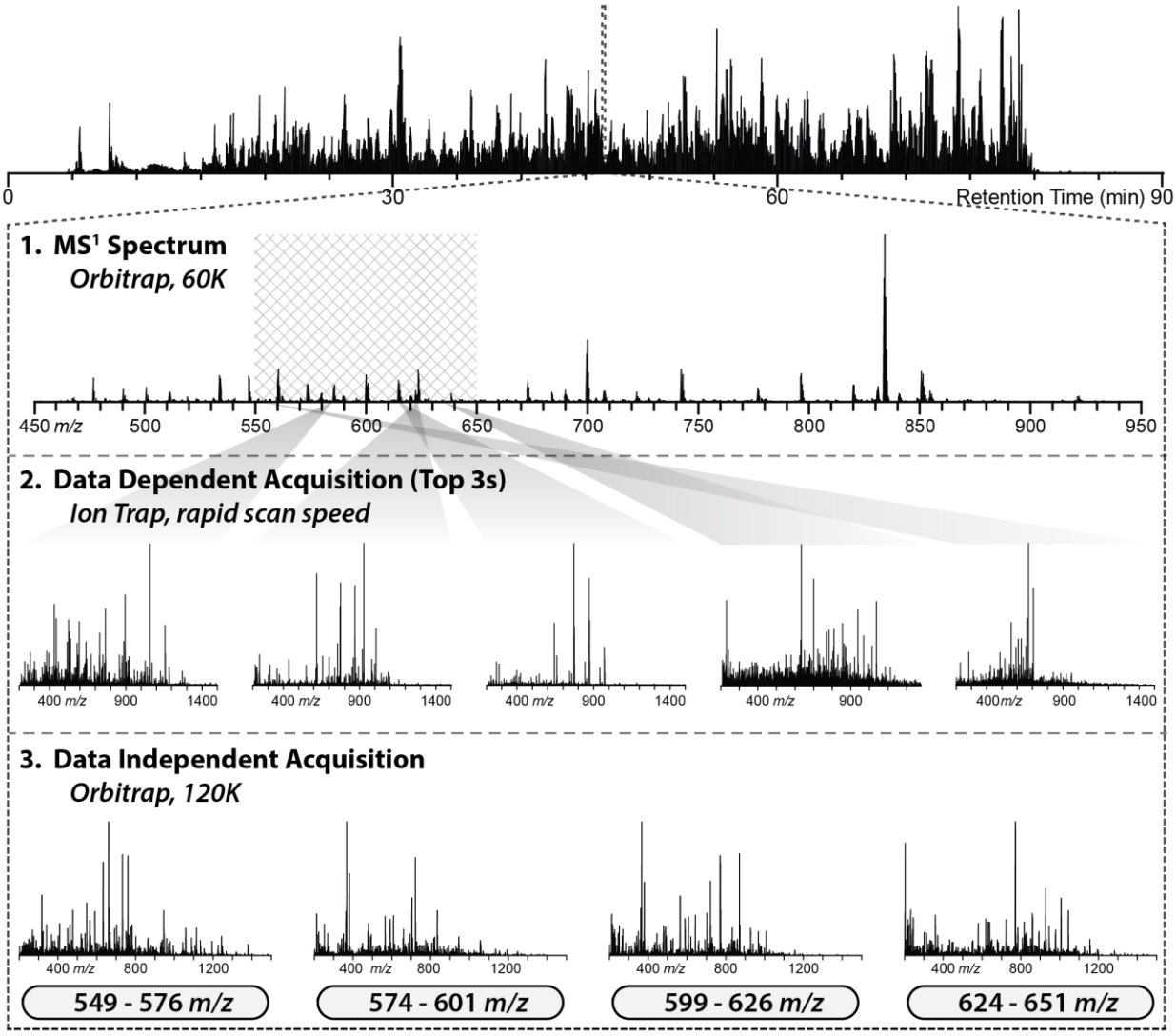
Catherine E. Minogue[1,4], Alexander S. Hebert[2,4], Jarred W. Rensvold[4], Michael S. Westphall[2], David J. Pagliarini[4], and Joshua J. Coon[1,2,3*]

Department of Chemistry,[1] Genome Center of Wisconsin,[2] Department of Biomolecular Chemistry,[3] and Department of Biochemistry,[4] University of Wisconsin, Madison, WI 53706

**Email: jcoon@chem.wisc.edu*

Supporting Information Table of Contents

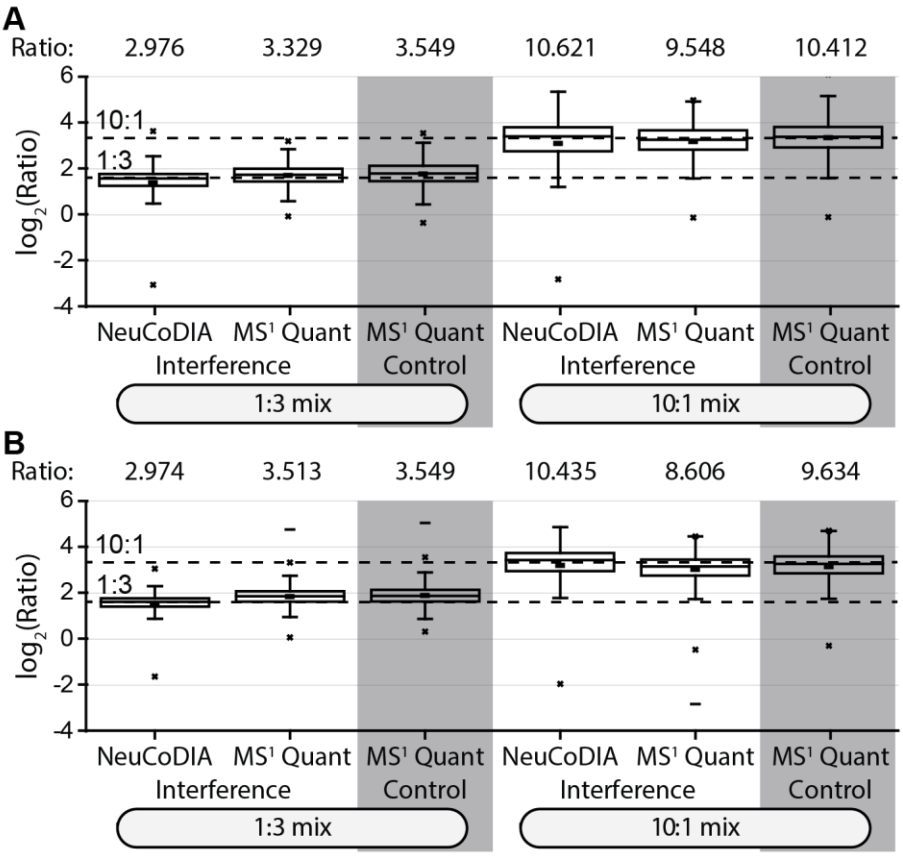
Supplementary Figures:



Supplementary Figure 1. NeuCoDIA scan sequence.

Each NeuCoDIA experiment is responsible for the analysis of precursor ions from a user-designated *m/z* range (in this example, the 550-650Th range is being analyzed, illustrated above with cross-hatching). Each NeuCoDIA scan cycle is comprised of three parts: 1. An MS^1 scan, analyzed in the Orbitrap at a resolution of 60k (at *m/z* 200); 2. A set of data-dependent MS^2 scans (for the generation of a target peptide list) that sample only from the precursor region of interest, analyzed in the ion trap at rapid scan speed using "top 3 second" acquisition; 3. Four data-independent MS^2 scans that result from the sequential isolation, fragmentation, and mass analysis

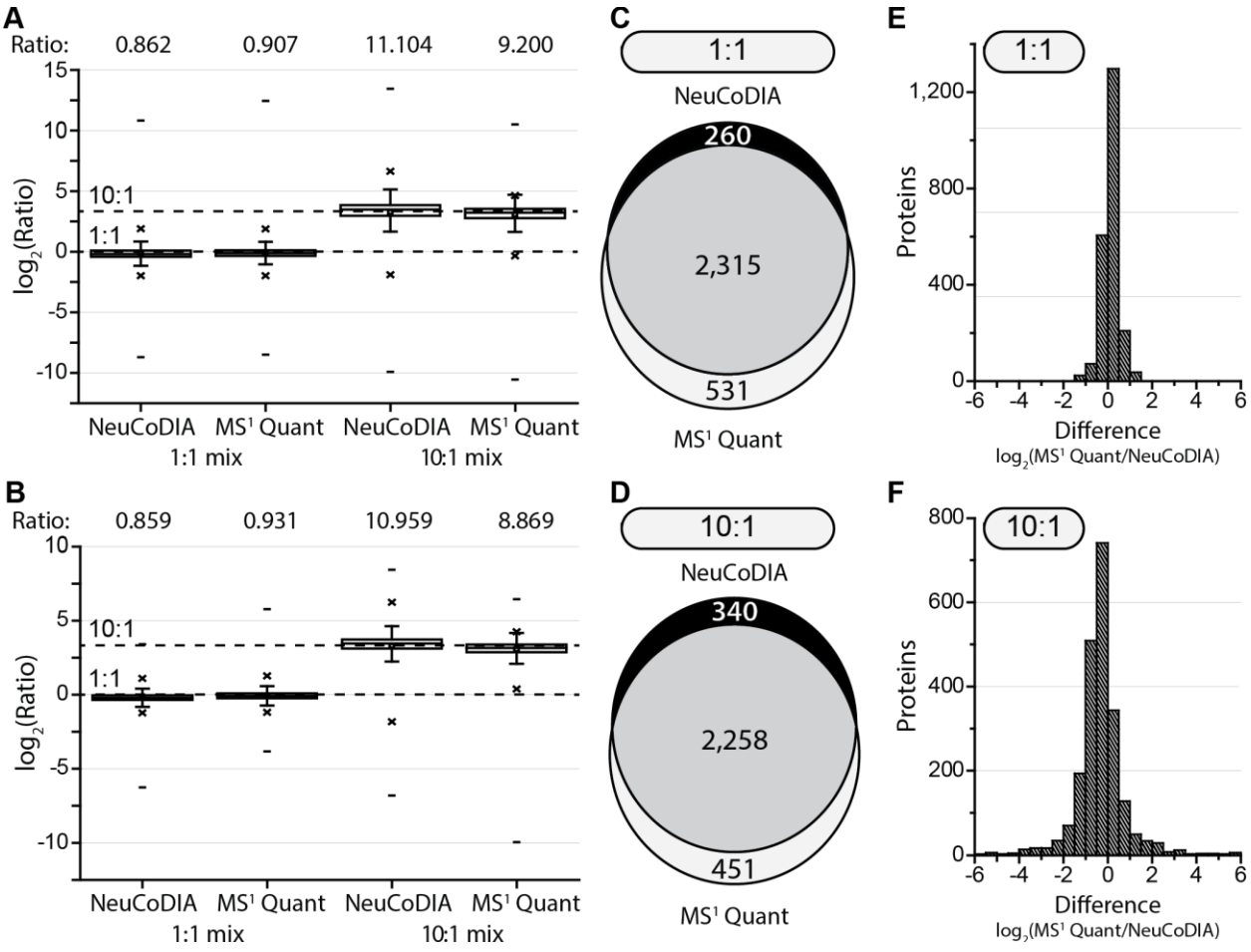
of distinct, equally-spaced *m/z* sections that span the precursor region of interest (in this example, the DIA regions sampled are 549-576Th, 574-601Th, 599-626Th, and 624-651Th), analyzed in the Orbitrap at a resolution of 120k (at *m/z* 200)).



Supplementary Figure 2. Precursor and product ion interference is negligible in both NeuCoDIA and MS[1] Quant (DDA) quantification.

Boxplots above show the measured (box and whiskers) and true (dashed lines) yeast peptide (A) and yeast protein (B) ratios for NeuCoDIA and MS[1] Quant (DDA) analyses of the 1:3 and 10:1 mixed-ratio samples with (white background) and without (grey background) human peptide interference (n = 7,375, 4,831, 7,702, 7,097, 5,007, 7,488, and n = 1,531, 1,232, 1,735, 1,555, 1,345, 1,779, left to right, respectively). Boxplots indicate the median (stripe), 25[th]-75[th] percentile (interquartile range, box), 1.5 times the interquartile range (whiskers), the 1[st] and 99[th] percentiles (x symbols), and the minimum/maximum values (lines, when applicable). Median ratios are provided above the boxplots. Standard deviations for the peptide distributions in (A) are 1.038, 0.634, 0.757, 1.705, 0.911, 1.046 and protein distributions in (B) are 0.745, 0.591, 0.657, 1.433, 0.848, 0.909 (left to right).

As is illustrated in the boxplots above, neither MS^1 Quant nor NeuCoDIA ratios are compressed as a result of spiking the 1:1 NeuCode human peptides into the 1:3 and 10:1 NeuCode yeast peptide mixtures, demonstrating that precursor and product ion interference, respectively, do not affect quantification within either method (this result contrasts with the significant effect precursor interference has on quantification with tandem mass tags). We do observe slight differences in the median ratio values obtained for the two methods, but we attribute this mostly to the different nature of the quantification strategies and not to the presence of interference (a higher ratio is measured by NeuCoDIA for the 10:1 sample while a higher median ratio is measured by MS^1 Quant for the 1:3 sample, indicating a bias towards one of the NeuCode channels for one, or both, of the quantification methods). Thus, requiring product ions to have at least three residues to be included in NeuCoDIA quantification prevents the incorporation of non-specific product ions into our analysis while still allowing us to use the majority of product ions for peptide/protein quantification.

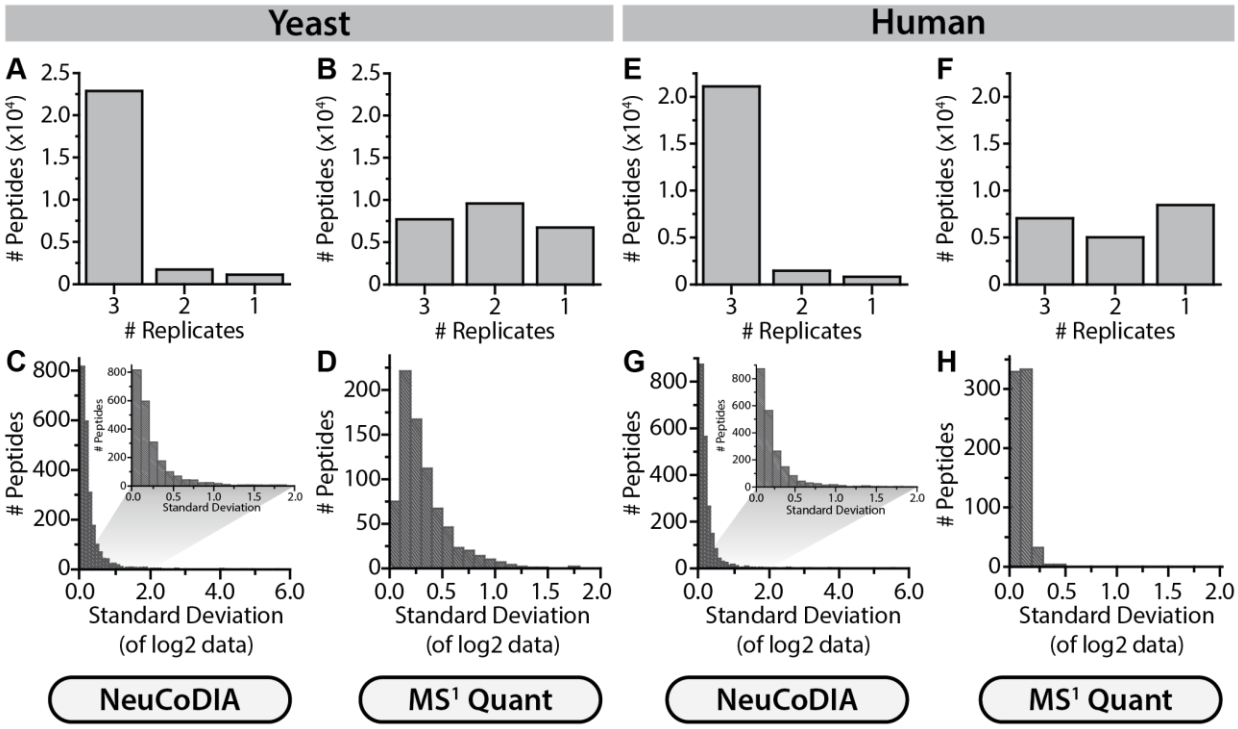


Supplementary Figure 3. Quantitative results for NeuCoDIA and MS[1] Quant (DDA) analyses of the 1:1 and 10:1 mixed-ratio yeast samples.

Boxplots show the measured (box and whiskers) and true (dashed lines) peptide (A) and protein (B) ratios for NeuCoDIA and MS[1] Quant (DDA) analyses at mixing ratios of 1:1 and 10:1 (n = 12,986, 13,393, 11,803, 12,112 and n = 2,575, 2,846, 2,598, 2,709, left to right, respectively). Boxplots indicate the median (stripe), 25[th]-75[th] percentile (interquartile range, box), 1.5 times the interquartile range (whiskers), the 1[st] and 99[th] percentiles (x symbols), and the minimum/maximum values (lines, when applicable). Median ratios are provided above the boxplots. Standard deviations (log_2) for the peptide distributions in (A) are 0.679, 0.629, 1.39 and 0.914 and protein distributions in (B) are 0.431, 0.435, 1.159, and 0.699 (left to right). Venn diagrams show the overlap in quantifiable protein identifications for NeuCoDIA and MS[1] Quant (DDA) analyses of the 1:1 (C) and 10:1 (D) mixed-ratio yeast samples. Histograms (E) and (F)

show the distribution of ratio differences between the NeuCoDIA and MS^1 Quant (DDA) methods for each quantifiable identification they had in common in (C) and (D), respectively.

It should be noted that the standard deviations of the NeuCoDIA quantification measurements are consistently slightly higher than those for the MS^1-based quantification measurements. This trend is also observed in the interference sample analyses presented in **Supplementary Figure 2**. We speculate that this discrepancy is largely due to the sophistication of the quantification software used to analyze the NeuCode data obtained by the different strategies.



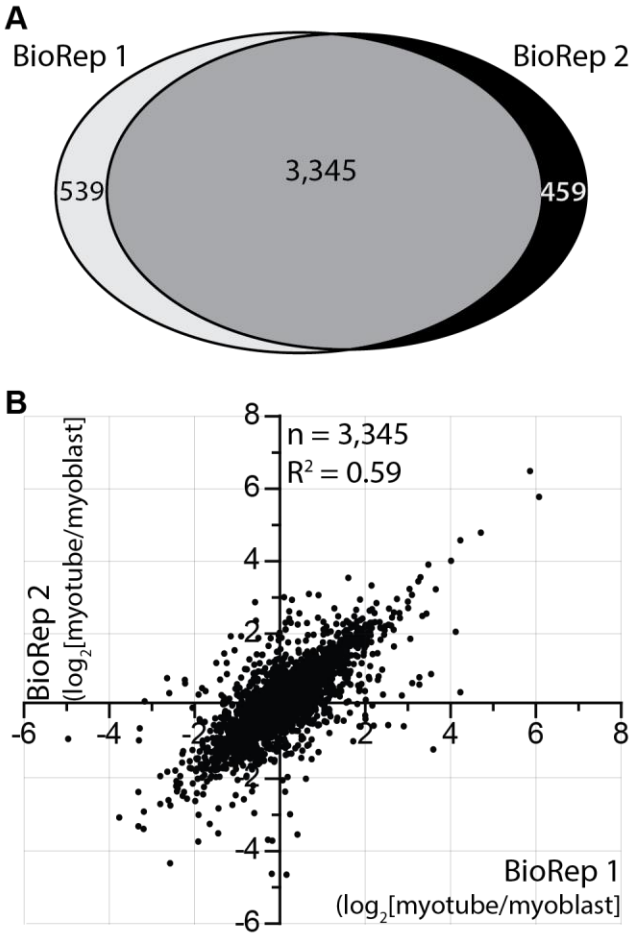
Supplementary Figure 4. NeuCoDIA acquires quantifiable identifications with greater reproducibility than the MS[1] Quant (DDA) method.

To evaluate the reproducibility of NeuCoDIA and MS[1]-based quantification methods, the 1:3 mixed-ratio yeast/human interference sample was analyzed by both strategies in technical triplicate. Only precursors from the 550-650Th precursor mass range were included in this assessment. The peptide target list (n = 2,587 for yeast and n = 2,339 for human) used for NeuCoDIA quantification was composed of all identifications made across the three replicate NeuCoDIA analyses. MS[1] Quant quantification was performed only on peptides identified within each of the MS[1] Quant analyses (n = 1,909 for yeast and n = 2,036 for human), as is standard for a typical DDA experiment.

Bar graphs A, B, E, and F illustrate the number of yeast (A, B) or human (E, F) peptides that could be quantified across one, two, or all three of the replicate NeuCoDIA (A, E) and MS[1] Quant (B, F) experiments. NeuCoDIA was able to quantify 88.4% (2,287) of the yeast peptides and 90.2% (2,110) of the human peptides across all three replicates, a significantly higher

number than MS^1 Quant, which was only able to quantify 39.4% (753) of the yeast peptides and 34.4% (700) of the human peptides across all three replicates.

NeuCoDIA was not only able to quantify a high percentage of peptides across replicate experiments, but the quantification values obtained for the observed peptides were also consistent between replicates. The standard deviation across the $\log_2$ intensity values acquired by the NeuCoDIA experiments was obtained for each peptide quantified across all three replicates, and the distribution of these standard deviations within the yeast and human sample sets are presented in histograms C and G. Similar histograms (D and H) were generated for the peptides quantified across all three MS^1 Quant experiments. Although the NeuCoDIA distributions contained a few more outliers than the MS^1 Quant datasets (which is to be expected given the fact that there are three times as many peptides included in the NeuCoDIA histogram, the majority of which escaped identification and quantification by all three MS^1 Quant analyses), the distribution of standard deviations for the 1:3 yeast measurements are incredibly similar for the NeuCoDIA and MS^1 Quant datasets. The standard deviation distributions obtained for the 1:1 human measurements slightly differ between the methods (with those values measured by MS^1 Quant being more consistent than those measured by NeuCoDIA); however, we believe this is largely the result of sample size discrepancy and less of a reflection on the quantitative reproducibility of the two methods. Overall, these data show that run-to-run reproducibility is a distinct advantage of NeuCoDIA over MS^1 Quant.



Supplementary Figure 5. NeuCoDIA analysis obtains comparable protein identification and protein quantification across two biological replicates of myogenic differentiation.

A) Venn diagram illustrates the overlap in quantifiable protein identifications obtained from NeuCoDIA analysis of each of the biological replicates. B) Scatter plot shows the ratios (log_2(myotube/myoblast)) associated with all proteins commonly quantified between the two biological replicates. Note that the R^2 value of 0.59 shown above is comparable to the R^2 values (0.60 and 0.69) observed for SILAC and MS[1] NeuCode quantification, respectively, of myogenic biological replicates in previous studies.[1]

Experimental Section:

Cell culture, differentiation, and lysis. Saccharomyces cerevisiae strain BY4741 Lys1Δ was grown in defined synthetic complete (SC, Sunrise Science) drop out media with either heavy $^{13}C_6{}^{15}N_2$ lysine (+8.0142 Da, Cambridge Isotopes, Cambridge, MA) or heavy D_8 (+8.0502 Da, Cambridge Isotopes, Cambridge, MA). Cells were propagated to a minimum of ten doublings. At mid-log phase, cells were harvested by centrifugation at 3,000 ×g for 3 minutes then washed three times with chilled ddH_2O. Cell pellets were re-suspended in 5mL lysis buffer (50mM Tris (pH8), 8M urea, 75mM NaCl, 100mM sodium butyrate, 1mM sodium orthovanadate, mini EDTA-free protease inhibitor (Roche Diagnostics, Indianapolis, IN), and phosSTOP phosphatase inhibitor (Roche Diagnostics, Indianapolis, IN)), and protein was extracted by glass bead milling (Retsch, Haan, Germany).

C2C12 cells were grown in DMEM lysine and arginine dropout culture medium (Cambridge Isotopes) supplemented with 10% dialyzed FBS, antibiotics, 100 mg/L unlabeled L-arginine and 100 mg/L of either heavy [$^{13}C_6$,$^{15}N_2$]L-lysine or heavy [$^{2}H_8$]L-lysine for six passages. 1.3×10^6 cells from these plates were seeded onto fresh plates with the same medium type and allowed to grow for 2 d. Cells grown with heavy [$^{2}H_8$]L-lysine were harvested; cells grown with heavy [$^{13}C_6$,$^{15}N_2$]L-lysine were refed with DMEM supplemented with 2% dialyzed FBS and then allowed to differentiate for 5 additional days before harvesting. Cells were pelleted and washed with ice-cold PBS. The cell pellets were resuspended in 8 M urea, 50 mM Tris, pH 8.0, 5 mM $CaCl_2$ and protease inhibitors (Roche Diagnostics, Indianapolis, IN). Cells were lysed by sonication.

HeLa S3 cells (ATCC CCL-22; ATCC, Manassas, VA) were grown in DMEM lysine and arginine dropout culture medium (Cambridge Isotopes) supplemented with 10% dialyzed FBS, antibiotics, 100 mg/L unlabeled L-arginine and 100 mg/L of either heavy [$^{13}C_6$,$^{15}N_2$]L-lysine or heavy [$^{2}H_8$]L-lysine for three passages. Cells were harvested at >80% confluency through centrifugation at 300xg for 5 minutes at 4°C. The supernatant was removed, and cells were washed with phosphate-buffered saline (PBS) and centrifuged at 300xg for 5 minutes at 4°C. The resulting pellet was stored at -80°C. HeLa S3 pellets were re-suspended in lysis buffer containing 8 M urea, 50 mM tris (pH 8), 5 mM CaCl2, 30 mM NaCl, 1 µL benzonase, and protease (Roche Diagnostics, Indianapolis, IN) and phosphatase (Roche Diagnostics, Indianapolis, IN) inhibitor tablets. The pellet was lysed by sonication.

Protein content within each of the yeast, mouse, and human samples was evaluated using a BCA assay (Thermo Fisher Scientific, San Jose, CA).

Protein digestion. Yeast, mouse, and human proteins were subjected to cysteine residue reduction by adding 5mM dithiothreitol and incubating for 30 minutes at 58°C. Free thiols were subjected to alkylation by adding 15mM iodoacetamide and incubating in the dark, at ambient temperature, for 30 minutes; this reaction was quenched by adding an additional 5mM dithiothreitol and incubating at ambient temperature for 15 minutes.

Yeast, mouse, and human proteins were enzymatically digested by adding Lys-C (Wako Chemicals, Richmond, VA) at a ratio of 1:50 (enzyme:protein) and incubating at ambient temperature for approximately 16 hours overnight. An additional 1:100 aliquot of Lys-C was added the following morning and incubated at ambient temperature for 1 hour. Both digests were quenched by bringing the pH ~2 with trifluoroacetic acid and immediately de-salted using C18 solid-phase extraction columns (SepPak, Waters, Milford, MA).

Sample preparation. Mixed-ratio yeast samples were prepared by mixing $^{13}C_6{}^{15}N_2$ (+8.0412 Da)-labeled and D_8 (+8.0502 Da)-labeled yeast peptides in 1:1 or 10:1 ratios, by mass. The mouse myoblast/myotube differentiation samples were prepared by mixing biological replicates of $^{13}C_6{}^{15}N_2$ (+8.0412 Da)-labeled myotube peptides and D_8 (+8.0502 Da)-labeled myoblast peptides in a 1:1 ratio, by mass.

Mixed-ratio yeast/human interference samples were generated to evaluate the effect of precursor and product ion interference on NeuCode quantification with MS^1 Quant and NeuCoDIA methods, respectively. $^{13}C_6{}^{15}N_2$ (+8.0142 Da)-labeled and D_8 (+8.0502 Da)-labeled yeast peptides were mixed in a 10:1 ratio or a 1:3 ratio by mass to generate two different mixed-ratio samples. An aliquot of each was put aside and analyzed using DDA (MS^1 Quant) to evaluate mixture ratios prior to the addition of interference. $^{13}C_6{}^{15}N_2$ (+8.0142 Da)-labeled and D_8 (+8.0502 Da)-labeled human peptides were mixed in a 1:1 ratio by mass. The HeLa mixture was then spiked into each of the yeast mixed-ratio samples in a 1:1 ratio by mass.

Liquid chromatography-mass spectrometry. All experiments were performed using an EASY-nLC system (Thermo Fisher Scientific, San Jose, CA) coupled to an Orbitrap Fusion (Q-OT-qIT)

mass spectrometer (Thermo Fisher Scientific, San Jose, CA). Reverse-phase columns were made in-house by packing a fused silica capillary (75µm i.d., 360 µm o.d, with a laser-pulled electrospray tip) with 1.7µm diameter, 130 Å pore size Bridged Ethylene Hybrid C18 particles (Waters) to a final length of 30cm. The column was heated to 60°C for all experiments. Samples were loaded onto the column for 12 minutes in 95:5 buffer A [water, 0.2% formic acid, and 5% DMSO]:buffer B [acetonitrile, 0.2% formic acid, and 5% DMSO] at a flow-rate of 0.30µL/min. Peptides were eluted using the following gradient: an increase to 7% B over 1 min, followed by a 42 min linear gradient from 7% to 18% B, followed by a 28 min linear gradient from 18% to 27% B, followed by a final 1 min ramp to 75% B which was held for 3 minutes. The column was equilibrated with 5% buffer B for an additional 25 min. Precursor peptide cations were generated from the eluent through the utilization of a nanoESI source.

MS^1 NeuCode Quantification: MS^1 Quant instrument methods consisted of a survey scan (450-950Th) analyzed in the Orbitrap at a resolution of 240k (at 200m/z) and an ion accumulation target of $5x10^5$. Tandem MS scans were collected at top speed mode with 5 second cycles; precursor peptides were isolated at 1.4Th with the quadrupole, subjected to HCD fragmentation (NCE 30), and analyzed in the ion trap at rapid scan speed. An ion accumulation target of $1x10^4$ was used for all tandem MS scans.

NeuCoDIA Quantification: For this proof-of-principle study, DDA peptide identification and NeuCoDIA quantification were combined into a single instrument method. Each NeuCoDIA experiment sampled only from a designated mass range. For the yeast mixed-ratio and interference experiments, we analyzed four discrete ranges in total: 450-550Th, 550-650Th, 650-750Th, 750-850Th. For the myogenesis experiments there were eight discrete ranges in total: 300-400Th, 400-500Th, 500-600Th, 600-700Th, 700-800Th, 800-900Th, 900-1,000Th, 1,000-1,400Th. Note that the mass ranges used for the three sets of experiments were selected based on the objective of each experiment. The yeast mixed-ratio and interference experiments were designed to show the accuracy and precision of the NeuCoDIA method, particularly in comparison with the previously described NeuCode MS^1 quantification approach. A subset of peptides from the mixed-ratio samples were sufficient to perform this comparison; therefore, for this analysis we selected the four 100Th precursor windows where the majority of LysC peptides reside. The myogenesis experiments, however, were conducted to evaluate the performance of NeuCoDIA quantification on a real, complex biological sample. For this analysis we therefore

wanted to analyze as many peptides/proteins from the sample as possible, as this is the goal of any typical large-scale biological study. For this reason, we performed NeuCoDIA analysis on the entire 300-1400Th range for the myogenesis study. Each NeuCoDIA experiment consisted of a survey scan (which covered all mass ranges) analyzed in the Orbitrap at a resolution of 60k (at 200*m/z*) and an ion accumulation target of $5x10^5$. This survey scan was followed by tandem MS scans collected at top speed mode with 1 second cycles that only selected precursors from the 100Th mass range of interest; precursor peptides were isolated at 1.4Th with the quadrupole, subjected to HCD fragmentation (NCE 30), and analyzed in the ion trap at rapid scan speed. An ion accumulation target of $1x10^4$ was used for these tandem MS scans. These scans were then followed by tandem MS scans which sequentially isolated (quadrupole), fragmented (HCD, NCE 30), and analyzed (Orbitrap, 120k resolution at 200m/z) four identically-sized *m/z* windows within their designated mass range (e.g. the 450-550 analysis will conduct MS/MS analyses on the 449-476Th, 474-501Th, 499-526Th, and 524-551Th ranges). Note that the isolation windows each overlap by 2Th to ensure complete isolation of the mass range of interest.

Data analysis. Data was processed using the in-house software suite COMPASS.[2] OMSSA[3] (version 2.1.8) searches were performed against a target-decoy database (Uniprot (yeast), www.uniprot.org, July 29, 2011; UniProt (mouse), www.uniprot.org, October 1, 2011; Uniprot (concatenated yeast and human), www.uniprot.org, April 4, 2014). Searches were conducted using a 150 ppm precursor mass tolerance and a 0.35 Da product mass tolerance. The fixed modifications specified were carbamidomethylation of cysteine residues and modification of lysine residues (+8.0322, average of the $^{13}C_2{}^{15}N_6$ and $^{2}H_8$ isotopologues); the variable modification specified was the oxidation of methionine residues. A maximum of 3 missed LysC cleavages were allowed. For each set of experiments, data was collectively filtered to 1% FDR at the peptide level.

MS^1 Quantification was performed as described in Hebert et. al.[1] NeuCoDIA quantification was achieved using in-house software. Briefly, retention time windows for peptide elution were acquired based on peptide identification. A set of MS^2 NeuCoDIA scans were then extracted for each precursor based on retention time and precursor *m/z* information. Using this set of scans, extracted ion chromatograms (XICs) could then be obtained for each theoretically quantifiable fragment ion pair (each fragment containing a lysine residue that could

theoretically be resolved at 120K resolution). Retention time elution windows were narrowed based on the chromatographic alignment of XIC base peaks. Fragment XICs were only utilized for peptide quantification if they had a base peak that 1) had a S/N above 3 (to maintain S/N threshold consistency with NeuQuant MS[1]-based quantification software[1,4]) and 2) aligned with the majority of the peptide fragments. When three or more fragment ions pairs were extracted for quantification, Dixon's Q-test was utilized for outlier removal. If both XICs could be detected/extracted for the majority of fragment ion pairs of a given peptide, those fragment ions for which both XICs could not be found were removed from quantification.

Protein grouping was performed within the COMPASS software suite. Only one peptide was required for each protein identification; however, a peptide was only used for protein inference if it uniquely mapped to a given protein group (for the sake of quantitative accuracy). Following protein grouping, data was collectively filtered to 1% FDR at the protein level. For the evaluation of our myogenesis study, enzymes belonging to major catabolic pathways were manually curated from Gene Ontology, KEGG and BioCyc annotations.

Theoretical Calculations. A list of fragment ions was generated using all peptide targets identified by NeuCoDIA analysis of the 10:1 yeast mixture. The theoretical full width at 10% maximum peak height (FWTM) for each fragment ion at resolving powers (R) of 30K, 60K, 120K, and 240K was calculated using the equation:

$$FWTM = 1.822616 \times \frac{m/z}{R \times \sqrt{\frac{200}{m/z}}}$$

where resolving power is defined as the minimum *m/z* difference that can be resolved at 200 *m/z* and the coefficient is derived from Gaussian peak shape modeling. The *m/z* difference ($\Delta m/z$) for each theoretical isotope doublet, assuming an isotopologue mass difference of 36 mDa, is given by:

$$\Delta m/z = \frac{36 \times n}{z}$$

Where n is the number of lysine residues in the fragment ion. Note that we only used z = 1 for each fragment ion in our analysis. An isotopologue pair is considered resolvable at the tested isotopologue mass difference and resolving power only if $\Delta m/z >$ FWTM.

References

(1) Hebert, A. S.; Merrill, A. E.; Bailey, D. J.; Still, A. J.; Westphall, M. S.; Strieter, E. R.; Pagliarini, D. J.; Coon, J. J. *Nat Meth* **2013**, *10*, 332-334.
(2) Wenger, C. D.; Phanstiel, D. H.; Lee, M.; Bailey, D. J.; Coon, J. J. *Proteomics* **2011**, *11*, 1064-1074.
(3) Geer, L. Y.; Markey, S. P.; Kowalak, J. A.; Wagner, L.; Xu, M.; Maynard, D. M.; Yang, X.; Shi, W.; Bryant, S. H. *Journal of proteome research* **2004**, *3*, 958-964.
(4) Merrill, A. E.; Hebert, A. S.; MacGilvray, M. E.; Rose, C. M.; Bailey, D. J.; Bradley, J. C.; Wood, W. W.; El Masri, M.; Westphall, M. S.; Gasch, A. P.; Coon, J. J. *Molecular & Cellular Proteomics* **2014**, *13*, 2503-2512.