Quantitative shotgun proteome analysis by direct infusion

Jesse G. Meyer, Natalie M. Niemi, David J. Pagliarini and Joshua J. Coon

Liquid chromatography–mass spectrometry (LC–MS) delivers sensitive peptide analysis for proteomics but requires extensive analysis time, reducing throughput. Here, we demonstrate that gas-phase peptide separation instead of LC enables fast proteome analysis. Using direct infusion–shotgun proteome analysis (DI-SPA) by data-independent acquisition mass spectrometry (DIA-MS), we demonstrate the targeted quantification of over 500 proteins within minutes of MS data collection (-3.5 proteins per second). We show the utility of this technology in performing a complex multifactorial proteomic study of interactions between nutrients, genotype and mitochondrial toxins in a collection of cultured human cells. More than 45,000 quantitative protein measurements from 132 samples were achieved in only ~4.4 h of MS data collection. Enabling fast, unbiased proteome quantification without LC, DI-SPA offers an approach to boost throughput, critical to drug and biomarker discovery studies that require analysis of thousands of proteomes.

Shotgun proteomics methods using LC–MS achieve the greatest depth and breadth of proteome coverage\(^1\)\(^-\)\(^^\)\(^2\). The time required for such comprehensive proteome analysis, once a major burden, has been driven down by technological adaptations. Just over a decade ago, weeks of MS data collection were required to quantify nearly all expressed yeast proteins\(^1\); in 2016, our group accomplished the same task in just over one hour\(^1\). More recent advancements in DIA and fast LC have further reduced analysis times and enabled routine protein quantification at rates of up to 15,000 non-unique proteins per hour\(^3\). Still, as the fields of proteomics and metabolomics push for higher throughput, the requirement for liquid-phase separations inevitably requires time that in turn limits throughput. This is amplified by time needed to load and re-equilibrate the LC column.

In theory, omitting LC prefractionation could decrease analysis time\(^4\). Several papers have described qualitative analysis of peptides from simple mixtures by direct infusion, an approach that is already common in metabolomics\(^5\)\(^-\)\(^^\)\(^7\). Twenty-five years ago, direct infusion of peptides from trypsin-digested gel bands or standard proteins was common but offered limited depth, typically of fewer than 60 peptides\(^8\)\(^-\)\(^^\)\(^9\). As LC and MS coevolved, LC–MS became the premier technology for the analysis of the tremendously complex mixture of peptides that results from whole-proteome digestion. Although direct infusion was recently used to profile histone modifications in one minute\(^1\), it has not been able to interrogate peptide mixtures from the human proteome, which contain well over 100,000 distinct peptide sequences\(^1\). Several factors may hinder detection of peptides from such complex mixtures by direct infusion electrospray ionization, including peptide polarity\(^1\), mobile-phase composition\(^2\)\(^-\)\(^^\)\(^3\), ion suppression\(^2\)\(^) and ion competition\(^2\). However, recent advancements in MS around accurate mass measurement, sensitivity and speed have inspired us to revisit the concept of peptide identification without LC.

Among recent MS advances, ion mobility has enabled an additional dimension of gas-phase peptide cation separation\(^4\)\(^-\)\(^^\)\(^6\) that complements fractionation by quadrupole selection\(^4\)\(^-\)\(^^\)\(^6\). Unlike liquid separations that work on the principle of hydrophobicity, the ion-mobility separations sort gas-phase peptide ions on the basis of their charge and shape. High-field asymmetric waveform ion mobility spectrometry (FAIMS) can permit very rapid gas-phase separation through a device placed between the electrospray emitter and atmospheric pressure inlet of a mass spectrometer. FAIMS filters ions through inner and outer electrodes on the basis of their differential mobility in high or low asymmetric fields. Analyte separation by FAIMS and other ion-mobility methods may improve the analysis of complex peptide mixtures without LC.

Here, we show that gas-phase separation can substitute for LC to deliver expeditious analysis of complex peptide mixtures from the human proteome. We name this strategy DI-SPA. Peptide samples are directly infused and ionized by electrospray, and the resulting peptide cations are separated in the gas phase before detection by DIA with high resolution MS/MS. We explored DI-SPA data-collection parameters and found that as the extent of gas-phase separations is positively correlated with the depth of observable proteome coverage. Strategies for peptide identification and quantification with DI-SPA were validated with standard mixtures of known heavy and light protein ratios and compared with traditional LC–MS peptide quantification. The utility of DI-SPA for high-throughput biological screening was demonstrated by quantifying proteomic responses of human cells to a complex multifactorial experiment grid of mitochondrial toxins, genotypes and nutrients. Application of DI-SPA to quantify proteins from mitochondrial subcellular fractions is also demonstrated. Altogether, the results show that DI-SPA enables fast proteome analysis without LC separation, permitting rapid quantification of biologically relevant proteome changes in cells and purified mitochondria.
Results

We first sought to determine how effectively gas-phase fractionation by FAIMS and a quadrupole mass filter purify peptide candidates using computational calculations. Precursor m/z values and maximum FAIMS compensation voltage (CV) transmission for human peptide identifications from LC–FAIMS–MS/MS were compiled and used for this theoretical gas-phase fractionation. The data were composed of 112,742 unique peptide precursors with maximum CV values from −20 to −120, and precursor m/z values ranging from 300 to 1,350. The number of peptides in each theoretical quadrupole isolation range (m/z 4 or 2), and FAIMS CV (10–V steps from −20 V to −120 V), was plotted as a stacked histogram (Extended Data Fig. 1). x-axis bin widths are analogous to the isolation width used for the quadrupole mass filter during MS analysis. Using a theoretical isolation width of 4 m/z without FAIMS, over 1,000 precursors were observed in a single 4–m/z window. By reducing the isolation width to 2 m/z, the maximum number of peptide precursors in any bin dropped to 564; coupling this reduced isolation width to FAIMS, complexity dropped even further to a maximum of 164 peptide precursors in any bin. This theoretical analysis indicated that, indeed, increased gas-phase fractionation substantially decreases the complexity of peptide precursors in any given channel. Smaller quadrupole selection ranges linearly reduce the number of peptides selected, and FAIMS selection enables a complementary but non-linear reduction in the number of peptide ions. Even with small quadrupole selection windows and FAIMS selection, multiple peptide ions are predicted to be present in every window. Fragmentation, a means to identify those coselected peptide ions, produces chimeric fragment ion spectra with signals that distinguish the original peptides. Altogether, this computational analysis reveals that gas-phase fractionation can theoretically reduce the complexity of peptide ions for analysis without LC.

On the basis of these theoretical results, we experimentally tested whether proteins and peptides could be identified and quantified with gas-phase fractionation instead of LC. Peptides were delivered to the nanospray emitter by direct infusion (DI) (Extended Data Fig. 2), and electrospayed ions were fractionated by FAIMS, subjected to quadrupole selection and dissociated using beam-type collisional activation (HCD), followed by product ion detection in the orbitrap (Fig. 1a). To perform peptide identification from these multiplexed spectra, we used MSPLIT-DIA33, a technique that applies the concept of spectral projection34 to extract library-spectra fragment ions within the instrument mass-accuracy range from candidate spectra. Examples of raw spectra, along with their projected spectra and library spectra for identified peptides (<1% false-discovery rate (FDR)), reveal that extraction of projected spectra with 10 ppm fragment mass tolerance (Supplementary Fig. 1) effectively produces matches with good spectral correlation (Fig. 1bc).

To optimize identifications, we tested a grid of mass-spectrometer settings in parameter-scouting experiments using peptide samples from whole human proteome proteolysis (MCF7 cells) (Extended Data Fig. 3a). Parameters included use of FAIMS, width of the Q1 quadrupole isolation window (Fig. 2a), maximum ion accumulation time (Fig. 2b) and orbitrap resolution (Fig. 2c). In each experiment, peptide solution was delivered to the source by DI, and each FAIMS CV of interest was held constant while stepping through Q1 isolation windows that spanned the precursor mass range of interest (Extended Data Fig. 3b). The number of peptide identifications increases as width of the Q1 isolation window decreases, ion injection time increases and orbitrap resolution increases. For almost all tested parameters, the addition of FAIMS increased the identified peptides. At the optimal setting, use of FAIMS increased peptide identifications by up to threefold (Fig. 2b). An increased number of peptide identifications due to use of FAIMS and smaller mass-filter selection windows indicates that the extent of gas-phase fractionation is a key determinant in the success of DI-SPA.

These data culminated in the discovery of the optimal settings for peptide analysis without chromatography.

We next examined the relationship between detectable peptide precursor ion features (MS1) and peptide identifications by DI-SPA. The same solution of peptides from the MCF7 proteome was infused, and precursor ion scans were collected with the same FAIMS gas-phase fractions. Thrash feature identification35, as implemented in decontools36, was used to identify a total of 1,477 MS1 features (excluding +1 ions). Peptide precursor masses identified by DI-SPA were compared with the observed precursor feature masses (Supplementary Fig. 2). DI-SPA enabled identification of 1,435 unique peptides with 55% overlap in identified precursor masses with the precursor masses observed in the MS1-only experiment. Comparison of the matched features as a function of deconvoluted mass and relative abundance from the MS1 revealed that many of the features found by DI-SPA were low abundance (Supplementary Fig. 2). Thus, compared with DDA, which requires observation of a mass in the MS1 spectra, the DI-SPA strategy picked up more low-abundance ions, probably due to the scan sequence that waits for a signal at each possible mass window.

To better understand the potential utility of DI-SPA, we examined the character of peptide identifications. The same MCF7 peptide sample was analyzed by traditional nano-scale LC–MS/MS (nLC–MS/MS) to perform label-free quantification (LFQ), and peptide identifications from both methods were compared (Fig. 2d). Peptide identifications from DI-SPA were a sample of the most abundant peptides present in the proteome sample. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of the 395 protein identifications (protein-level FDR < 1.25%) from the same DI-SPA analysis revealed substantial coverage of several pathways central to biology, including central carbon metabolism (glycolysis, pentose phosphate pathway, tricarboxylic acid cycle) and protein synthesis and degradation (splicosome, ribosome, proteasome and proteome processing in the endoplasmic reticulum) (Fig. 2e, Extended Data Fig. 4 and Supplementary Table 1).

The robustness and reproducibility of DI-SPA was assessed by consecutively analyzing the same MCF7 peptide sample 100 times (Extended Data Fig. 5). Because DI-SPA is so fast, these data were collected overnight. An average of 1,542 peptides were identified per injection, with a s.d. of 63.1 peptides (4.1% or the mean). Of all peptides identified, exactly 869 were identified from all 100 injections (56% of average peptide identifications), and 1,264 were identified in at least 80/100 injections (82% of average peptide identifications). The reproducibility of DI-SPA analysis over 100 replicates therefore exceeds that of pairs of technical replicates from some DDA strategies37.

DI-SPA was then challenged with one of the most difficult sample matrices, human plasma. Two different sources of human plasma either were not depleted or had the top 12 most abundant proteins depleted, and the samples were analyzed by DI-SPA (Extended Data Fig. 6). DI-SPA performed similarly on frozen and on lyophilized human plasma samples, and found more peptides and proteins from depleted samples. Compared with the standard few hundred protein identifications by traditional nLC–MS/MS studies38, DI-SPA identified 745 peptides that map to 109 human proteins (Supplementary Table 2). Although this proof-of-principle plasma analysis by DI-SPA was done with the scouting method in >15 min (as shown in Extended Data Fig. 3), subsequent methods could target only one peptide per plasma protein and achieve the same analysis in only about 31 s (~0.28 s per scan).

Next, a quantitative DI-SPA strategy was evaluated using defined mixtures of A549 cells labeled with heavy or light arginine and lysine (Extended Data Fig. 7). Heavy-labeled and light (natural abundance) proteome samples were combined at various ratios, proteolytically digested with trypsin and analyzed by standard nLC–MS/MS and MaxQuant to verify the labeling ratios
and provide a baseline of quantitative values for comparison (Fig. 3a). One sample of A549 peptides was then analyzed with DI-SPA using the optimal peptide identification settings determined by the scouting experiments, resulting in the identification of 2,248 unique peptides (Supplementary Table 3). To enable quantification from heavy and light y-type ion pairs using DI-SPA, a new data-collection method was built that coisolated light and heavy SILAC precursor pairs from doubly charged peptides using the ion multiplexing feature (MSX) of the Orbitrap Fusion Lumos (Supplementary Fig. 3). Peptides were quantified from MSX DI-SPA data by taking the median heavy/light ratio of the y-type ions in the fragment ion spectra. Only the three most abundant y-type ions that matched to the spectral library were used for quantification. From these data, we conclude that DI-SPA can effectively quantify peptides (Fig. 3b). Finally, a DI-SPA method was developed to quantify selected proteins more quickly. This method targeted heavy and light masses of peptide precursors for 1 peptide from each of the 552 identified proteins (Extended Data Fig. 8 and Supplementary Table 4). Up to 525 proteins were quantified in only ~2 min of MS data collection, resulting in a rate of 3.5 proteins quantified per second. The quantitative quality of this fast, targeted quantification experiments, in which 1,384 peptides were identified (peptide-level FDR = 0.93%) from 451 proteins (protein-level FDR = 2.8%). Next, a quantitative targeted DI-SPA assay was generated to expeditiously monitor the most intense peptide from each of the 451 proteins. This method was used to analyze all 132 samples, generating 132 proteotypes and requiring only ~2 min of MS data collection per sample. A total of 341 out of 451 protein targets were quantified across all 132 samples in only ~4.4 h of total MS data acquisition, which corresponds to over 45,000 proteins quantified at a rate of nearly 3 proteins per second (Fig. 4b).

This dataset of 44 unique cellular states revealed many interesting changes due to mitochondrial toxin treatments (Supplementary Dataset 1 and Supplementary Table 5). For example, nearly all glycolytic proteins were upregulated upon treatment with the mito-}


deferoxamine (DFO) compared with the appropriate controls (Fig. 4c). Supporting our observation, DFO has previously been reported to increase glycolysis. Next, to understand relationships between various treatments, we summarized these data with uniform manifold approximation and projection (UMAP) (Fig. 4d). UMAP shows clear segregation of the treatments into 24-h and 6-h groups, and within the 6-h group, the proteotypes easily segregated from the WT and PPTC7-KO cells. Within the 6-h WT group, the
different media had a minimal influence on the proteotype. In relation to the 6-h controls, complex I inhibitors rotenone and antimycin A were most similar. Toxins that influence mitochondrial membrane potential, carbonyl cyanide m-chlorophenyl hydrazone (CCCP) and oligomycin produced more (and comparable) proteome rearrangement. Valinomycin, which diffuses potassium-ion...
gradients across membranes, induced the most profound proteome perturbation relative to controls. Finally, UMAP analysis revealed that bardoxolone (CDDO) treatment is medium-dependent.

Data from this DI-SPA experiment also revealed proteome differences due to the PPTC7-KO genotype, including a lower citrate synthase activity across 24-h controls and treatments than in WT 293T cells (Fig. 4e). On the basis of this observation and previous results, we hypothesized that these cells may have a mitochondrial defect or lower mitochondrial quantity. Mitochondria were enriched from WT and KO cells, and crude mitochondria fractions were analyzed by DI-SPA. Quantitative DI-SPA ratios for these samples were compared with those computed with MaxQuant analysis of LC–MS data, revealing excellent agreement and data quality (Extended Data Fig. 9). In this organelle-specific analysis, 351 proteins were quantified across all samples, and there was a clear trend toward lower quantities of most mitochondrial proteins (Extended Data Fig. 10, 149 annotated mitochondrial according to Gene Ontology Cellular Component). In DI-SPA data from the mitochondrial...
fractions, 4 proteins were significantly downregulated in PPTC7-KO cells compared to WT controls (Benjamini–Hochberg-corrected P < 0.05): Acetyl-CoA acetyltransferase THIL, 10-kDa heat-shock protein CH10, prohibitin PHB and, again, citrate synthase CISY (Supplementary Table 6). To validate our hypothesis from DI-SPA data, we measured mitochondrial function with Seahorse respirometry and found that PPTC7-KO cells indeed have lower oxygen consumption than WT 293T cells (Fig. 4f).

Discussion
Here we describe and validate DI-SPA, a qualitative and quantitative MS-based proteomics method that does not use LC. DI-SPA instead separates peptides in the gas phase with three primary technologies: (1) ion mobility (FAIMS), (2) m/z-based quadrupole mass-filter isolation and (3) ion dissociation. The complex and chimeric MS/MS spectra from DI-SPA are analyzed using the projected spectrum concept and chimeric MS/MS spectra from DI-SPA are analyzed using the projected spectrum concept (Fig. 1). Ion mobility is a key determinant of success; addition of FAIMS tripled peptide identifications. DI-SPA identified peptides that reflect a subset of the most abundant peptides found by traditional nLC–MS/MS. While DI-SPA identifies fewer proteins than traditional nLC–MS/MS, it provides a very simple and robust interface for a superficial view of the proteome — including important cellular pathways — within minutes (Fig. 2).

Our data demonstrate quantitative analysis by DI-SPA with samples containing stable-isotope-labeled protein standards (such as SILAC), which is achieved by comparing ratios of peptide-fragment ions. This enables protein quantification at speeds of up to 3.5 proteins per second. The quantitative values obtained by DI-SPA are similar to those from standard LC–MS (Fig. 3 and Extended Data Fig. 9). The proteome changes reflect known biology, such as an upregulation of glycolysis due to DFO treatment, and depletion of mitochondrial function due to PPTC7 KO (Fig. 4). Our results further suggest that DI-SPA is an ideal choice for analysis of simplified mixtures, such as subcellular fractions of mitochondria and proteins from coimmunoprecipitation, for example. Currently, DI-SPA makes rapid proteome quantification a viable option for a range of high-throughput studies, including drug and biomarker discovery.

Many recent reports aim to improve the speed and throughput of proteome analysis by pushing for shorter LC separations. DI-SPA takes the concept of shorter LC separation to the logical extreme by completely omitting LC. A combination of several solutions were required to enable DI-SPA: (1) an additional dimension of separation by ion mobility, (2) data collection by DIA, (3) peptide identification with the projected spectrum concept and (4) the coisolation of heavy-labeled standard peptides to enable quantification from fragments. Compared with those recent studies that focus on faster analysis through shorter LC separation, DI-SPA quantifies proteins at a similar pace (up to 3.5 proteins per second). Some shortcomings of this first iteration of DI-SPA are that it is not yet adapted to perform label-free quantification, and it has not yet been applied to high-throughput quantification of proteins from biofluids.

The method proposed here may seem at odds with prior calls for better chromatography to drive the field of proteomics to more thorough analysis and better depth. In our view, there are many applications where the proteomic depth is not required, but rather speed and reproducibility are the driving figures of merit. Here, we demonstrate how this LC-free paradigm can fill this technological need in certain example cases: (1) obtaining quick quantitative prototype profiles revealing mechanisms of toxins, and (2) quantifying isolated mitochondria proteotypes. We expect continued advancements in the speed and sensitivity of MS to be beneficial for subsequent iterations of the DI-SPA strategy, improving the depth and breadth of LC-free proteome coverage.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41592-020-00999-z.

Received: 20 May 2020; Accepted: 21 October 2020; Published online: 23 November 2020

References
1. Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. Nature 537, 347–355 (2016).
2. Meyer, J. G. & Schilling, B. Clinical applications of quantitative proteomics using targeted and untargeted data-independent acquisition techniques. Expert Rev. Proteom. 14, 419–429 (2017).
3. de Godoy, L. M. F. et al. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. Nature 455, 1251–1254 (2008).
4. Hebert, A. S. et al. The one hour yeast proteome. Mol. Cell. Proteom. 13, 339–347 (2014).
5. Bache, N. et al. A novel LC system embeds analytes in pre-formed gradients for rapid, ultra-robust proteomics. Mol. Cell. Proteom. 17, 2284–2296 (2018).
6. Kelstrup, C. D. et al. Performance evaluation of the Q Exactive HF-X for shotgun proteomics. J. Proteome Res. 17, 727–738 (2018).
7. Gochumi, G., Purves, R. W., Hopf, C. & El-Aneel, A. Fast quantification without conventional chromatography, the growing power of mass spectrometry. Anal. Chem. 92, 8628–8637 (2020).
8. Chekmeneva, E. et al. Optimization and application of direct infusion nanoelectrospray HRMS method for large-scale urinary metabolic phenotyping in molecular epidemiology. J. Proteome Res. 16, 1646–1658 (2017).
9. Koulman, A. et al. High-throughput direct-infusion ion trap mass spectrometry: a new method for metabolomics. Rapid Commun. Mass Spectrom. 21, 421–428 (2007).
10. Wilm, M. et al. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature 379, 466–469 (1996).
11. Chen, S. Rapid protein identification using direct infusion nanoelectrospray ionization mass spectrometry. Proteomics 6, 16–25 (2006).
12. Pereira-Medrano, A. G., Sterling, A., Snijders, A. P. L., Reardon, K. F. & Wright, P. C. A systematic evaluation of chip-based nanoelectrospray ionization for rapid identification of proteins from a complex mixture. J. Am. Soc. Mass Spectrom. 18, 1714–1725 (2007).
13. Chen, J., Canales, L. & Neal, R. F. Multi-segment direct inject nano-ESI-LTQ-FT-ICR-MS/MS for protein identification. Proteome Sci. 9, 38 (2011).
14. Jiang, Y. & Koomen, J. M. Evaluation of direct infusion-multiple reaction monitoring mass spectrometry for quantification of heat shock proteins. Anal. Chem. 84, 1981–1986 (2012).
15. Kretschy, D. et al. High-throughput flow injection analysis of labeled peptides in cellular samples—ICP-MS analysis versus fluorescence based detection. Int. J. Mass Spectrom. 307, 105–111 (2011).
16. Sudjli, S. et al. One minute analysis of 200 histone posttranslational modifications by direct injection mass spectrometry. Genome Res. 29, 978–987 (2019).
17. Michalski, A., Cox, J. & Mann, M. More than 100,000 detectable peptide species eluted in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC–MS/MS. J. Proteome Res. 10, 1785–1793 (2011).
18. Cech, N. B. & Enke, C. G. Relating electrospray ionization response to nonpolar character of small peptides. Anal. Chem. 72, 2717–2723 (2000).
19. Meyer, J. G. & A. Komives, E Charge state coalescence during electrospray ionization improves peptide identification by tandem mass spectrometry. Mol. Cell. Proteom. 16, 2284–2296 (2018).
20. Ogóralek Loo, R. R., Lakshmanan, R. & Loo, J. A. What protein charging (and supercharging) reveal about the mechanism of electrospray ionization. J. Am. Soc. Mass Spectrom. 25, 1675–1693 (2014).
21. Annesley, T. M. Ion suppression in mass spectrometry. Clin. Chem. 49, 1041 (2003).
22. Sarvin, B. et al. Fast and sensitive flow-injection mass spectrometry metabolomics by analyzing sample-specific ion distributions. Nat. Commun. 11, 3186 (2020).
23. Webb, I. K. et al. Experimental evaluation and optimization of structures for lossless ion manipulations for ion mobility spectrometry with time-of-flight mass spectrometry. Anal. Chem. 86, 9169–9176 (2014).
24. Giles, K. et al. A cyclic ion mobility–mass spectrometry system. Anal. Chem. 91, 8564–8573 (2019).
25. Meier, F. et al. Parallel accumulation–serial fragmentation (PASEF): multiplying sequencing speed and sensitivity by synchronized scans in a trapped ion mobility device. J. Proteome Res. 14, 5378–5387 (2015).

26. Swearingen, K. E. et al. Nanospray FAIMS fractionation provides significant increases in proteome coverage of unfractionated complex protein digests. Mol. Cell. Proteom. 11, M111.014985 (2012).

27. Hebert, A. S. et al. Comprehensive single-shot proteomics with FAIMS on a hybrid orbitrap mass spectrometer. Anal. Chem. 90, 9529–9537 (2018).

28. Nagy, G. et al. Separation of β-amyloid tryptic peptide species with isomerized and racemized l-aspartic residues with ion mobility in structures for lossless ion manipulations. Anal. Chem. 91, 4374–4380 (2019).

29. Melani, R. D. et al. Direct measurement of light and heavy antibody chains using differential ion mobility spectrometry and middle-down mass spectrometry. MAbs 11, 1351–1357 (2019).

30. Horn, D. M., Zubarev, R. A. & McLafferty, F. W. Automated reduction and interpretation of high resolution electrospray mass spectra of large molecules. J. Am. Soc. Mass Spectrom. 14, 1746–1748 (2010).

31. Jaitly, N. et al. Decon2LS: an open-source software package for automated processing and visualization of high resolution mass spectrometry data. BMC Bioinform. 10, 87 (2009).

32. Tabb, D. L. et al. Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. J. Proteome Res. 9, 761–776 (2010).

33. Ng, S. S. et al. Quantification of lysine acetylation and succinylation stoichiometry in proteins using mass spectrometric data-independent acquisitions (SWATH). J. Am. Soc. Mass Spectrom. 27, 1758–1771 (2016).

34. Ivanov, M. V. et al. DirectMS1: MS/MS-free identification of 1000 proteins of cellular proteomes in 5 minutes. Anal. Chem. 92, 4326–4333 (2020).

35. Bekker-Jensen, D. B. et al. A compact quadrupole-orbitrap mass spectrometer with FAIMS interface improves proteome coverage in short LC gradients. Mol. Cell Proteom. 19, 716–729 (2020).

36. Oexle, H., Gnaiger, E. & Weiss, G. Iron-dependent changes in cellular energy metabolism: influence on citric acid cycle and oxidative phosphorylation. Biochim. Biophys. Acta 1413, 99–107 (1999).

37. McInnes, L., Healy, J. & Melville, J. UMAP: uniform manifold approximation and projection for dimension reduction. Preprint at https://arxiv.org/abs/1802.03426 (2018).

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020
Methods
Theoretical analysis of gas-phase fractionation. Data from FAIMS compensation voltage stepping experiments using peptides from tryptic–catalyzed proteolysis of the yeast proteome described by Hebert et al. was re-analyzed with MS-Fragger to identify peptides using the default settings, except that a fragment ion tolerance of 0.35 Daltons was used. The distributions of m/z values for identified peptides were plotted as histograms across m/z space with differing bin widths to visualize the complexity reduction that was possible with quadrupole isolation widths. Subsets of identifications from single compensation voltage analysis were plotted to visualize the complexity reduction afforded by FAIMS fractionation.

Samples for parameter scouting and SILAC validation experiments. MCF7 cells were grown to 80% confluence adherently on a T-175 flask, rinsed once with 1x D-PBS and then detached from the cell-culture plate using 1x trypsin solution. The trypsin was quenched by the addition of medium, and then the cells were pelleted by centrifugation at 150 × g. The cells were washed twice with ice-cold 1x D-PBS, and the supernatant was aspirated to remove any medium components. The cell pellet was then frozen at −80 °C until lysis. A549 cells for SILAC quantification experiments were grown in medium supplemented with 10% dialyzed bovine serum and heavy or light lysine and arginine for more than 10 population doublings to completely label cells (Thermo Scientific SILAC Protein Quantitation Kit, catalog no. A33972). Completely labeled cells were then collected by addition of trypsin, washed with cold PBS and counted to determine accurate cell numbers. Various ratios of heavy- and light-labeled cells were combined to reach a final number of 100,000 total cells. Combined cells were pelleted by centrifugation, PBS was aspirated and pellets were frozen and stored at −80 °C until lysis.

Lysis, digestion and desalting. Frozen cells were lysed by addition of 8 M urea with 50 mM TEAB buffer at pH 8.5 containing 10 mM TCEP and 10 mM chloroacetamide. The pellets were vortexed until homogenous with lysis buffer, and then were kept on ice. The larger lysis of MCF7 cells for infusion scouting experiments was sonicated on ice using a probe tip for 3 cycles of 10 s. The small-volume lysis of SILAC-labeled A549 cell samples was sonicated in a Qsonica water bath maintained at 4 °C. After sonication, lysis buffer was diluted to 2 M urea using 50 mM TEAB, and catalytic hydrolysis of proteins was initiated by addition of trypsin (Promega) and LysC (Wako) at a ratio of 1:100 protease:substrate by weight. Proteome proteolysis was incubated overnight (approximately 18h) at room temperature. Peptides were desalted using Strata reversed-phase cartridges, and then dried completely in a vacuum centrifuge. Between 0.5–1 mg ml−1 dried peptides were resuspended in 50%/49.8%/0.2% ACN/Water/FA (vol/vol/vol) for direct-infusion experiments, or at the same concentration in water with 2% ACN and 0.2% FA for nLC–MS/MS experiments.

Data collection. An orbitrap Fusion Lumos mass spectrometer was operated in targeted MS2 (MS2) mode with quadrupole isolation windows spanning the range from 400–1,000 Thompsons. Peptides were infused into a 75-µm inner diameter capillary tip from new objective (part no. P-560-70-10-N-5) that was packed with approximately 1 cm of C8 particles (Ivytyp, 5-µm particle size) to prevent clogging of the fritted tip by small particles. To ensure that this did not result in peptide retention or chromatography, we examined extracted ion chromatograms from several random multiply charged masses and found identical patterns of signal with time (Figure S2). Using a 50-µm ID LC-pump were used to automate the infusions. Injections of 1–4 µl (depending on the length of the entire experiment cycle) were infused at a rate of 100–300 min−1 using 50% ACN, 49.8% H2O and 0.2% FA as a carrier fluid. Several data-collection parameters were varied, including the maximum ion-accumulation time, the orbitrap resolution and the width of the quadrupole isolation window. Final optimal data-collection parameters were 120,000 orbitrap resolution with 246 ms maximum ion-accumulation time and 2–m/z precursor isolation windows. Mass spectrometry data were collected using Thermo Fisher Foundation software version 3.1 SP4, Freestyle version 1.3 SP2 and Xcalibur 4.0.

Peptide and protein identification. The spectral library is available with the MS data on massive18, and was created from a “biff–library” spectral library made with Skyline from a database search with MS-Fragger of data from FAIMS-fractionated human peptide samples. BLibToMs2 from Proteowizard was used to convert bif to bll2 format, which was then converted to mgf with msconvert. Custom Python code (fixMGFlib.ipynb; available on GitHub https://github.com/igneyerucid/D12A/blob/master/Python/fixMGFlib.ipynb) was used to then fix the mgf library by adding back the peptide sequence lines. Decoy was added to the spectral library by the spectral library processor included with MSPLIT-DIA19.

RAW files were converted to mzXML using msconvert with the default settings except that fragmentation was used. Converted files were searched against the human spectral library that included decoy spectra using MSPLIT-DIA with precursor tolerance equal to the isolation window width and fragment tolerance of 10 ppm. Peptides were scored by cosine similarity of experimental projected spectra with spectral library spectra using MSPLIT-DIA. Peptide identifications were sorted by their cosine match score, filtered to keep only the best score per peptide, and the peptide-level false-discovery rate was computed using the target-decoy strategy.

Although peptide identification and quantification were the focus of this study, for some experiments, protein-level FDR was computed using the target-decoy strategy with the best peptide cosine score as the protein score as described in the original MSPLIT-DIA paper20.

Untargeted protein quantification MS method. To first determine whether quantification from SILAC experiments would be possible, a general method to coisolate all heavy and light peptide precursor mass pairs for doubly charged peptide precursor ions was developed. The optimal peptide identification settings determined from the optimization grid were used in a scouting experiment to identify peptides from the 1:16 (heavy:light) A549-derived peptide sample. These identifications were used to determine peptide quantification targets in subsequent experiments.

Theoretical heavy masses were predicted from all the peptides identified from analysis of the 1:16 (heavy:light) SILAC sample (Supplementary Fig. 3). This analysis revealed that, as expected from tryptic peptides, most mass shifts result from incorporation of a single heavy lysine, followed by a single heavy arginine. Therefore, a simple multiplexed tandem MS (MSX) method was designed that coisolated every precursor mass M simultaneously with mass M + 4.5, which is between the mass of a single doubly charged heavy lysine (8.014199/2 = 4) or single heavy arginine (10.00827/2 = 5). This method fills the ion routing multipole for half the time with both specified ions before fragment mass analysis, ensuring that spectra contain fragments from any presence of both the light and heavy peptide. All light masses between 400–1,000 m/z were coisolated and fragmented with their heavy partner (M and M + 4.5, each with Q1 isolation of 2 m/z) using steps of 1.5 m/z.

Targeted protein quantification MS method. Data-collection methods were designed that targeted a single peptide from each protein identification using custom scripts written in R and Python, which are available from https://github.com/igneyerucid/D12A. First, peptide identifications were matched to proteins in a FASTA database. To be conservative, only peptides that matched a single protein entry were kept for FDR calculation using the target-decoy method. Specifically, the peptide from each protein with the best cosine score was kept, and that cosine score was used as the protein score (for example, some algorithms combine multiple peptide scores into one protein score to strengthen it). A protein target list was then generated that consisted of the peptide from each protein that was identified with the highest MS/MS spectra intensity from the scouting experiment. Predicted precursor light and heavy m/z for each peptide was then determined on the basis of the charge state and the counts of arginine and lysine residues, and the FAIMS CV that produced the identification was gathered from the mzXML scan header. Lists of target peptides at each FAIMS CV were then generated using the predicted light and heavy m/z, and custom data-collection methods were built that coisolated the light and heavy m/z signal from each peptide using the ion multiplexing (MSX) option of the Orbitrap Fusion Lumos. Fragment ions were measured in the orbitrap with 120,000 resolution and 246 ms maximum ion-accumulation time, unless otherwise noted.

Plasma experiment. Frozen liquid plasma treated with sodium heparin was purchased from BioIVT. Lyophilized plasma treated with citrate buffer was purchased from Sigma Aldrich (P9523–1ML) and resuspended in 1 mL of sterile deionized water immediately before use. Both plasma types were depleted in parallel with Top12 spin columns (Pierce no. 835165), according to the manufacturer’s instructions. Eluted plasma protein samples from the spin columns were concentrated, and buffer was exchanged into denaturation buffer (8 M Urea with 50 mM TEAB, pH 8.5) to approximately 40 µl with a 10-KDa (0.5 ml) Amicon ultrafiltration device. Undepleted plasma was diluted 17.5-fold into the same denaturation buffer. Protein concentrations from depleted and not depleted plasma samples in denaturation buffer were determined using the BCA assay. The protein concentration of all samples was adjusted to 1 mg ml−1 of total of 40 µl, and TCEP and chloroacetamide were added to a final concentration of 10 mM. After protein reduction and alkylation for 30 min, the urea was diluted to 2 µl with 50 mM TEAB buffer, and enzymatic hydrolysis of proteins was initiated by the addition of 0.8 µg of LysC and trypsin and was allowed to proceed overnight at room temperature. The reaction was stopped in the morning by the addition of 16% TFA, and m/z spectra were desalted with Phenomenex Strata-X 33 µm polymeric reversed-phase cartridges (10 mg sorbent, 1-ml tube, part no. 8B-S100-AAK). DI-SPA analysis was performed using the best parameter scouting method.

Mitox experiment, cell culture. 293T cells were purchased from ATCC (no. CRL-3216) and maintained in DMEM (FBS and 1x penicillin–streptomycin (100 U ml−1 final concentration). Human Plasma-Like Medium (HPLM) was kindly provided by the Cantor laboratory, and was supplemented with 10% dialyzed FBS (Thermo no. 26000036) and 1x penicillin–streptomycin (100 U ml−1 final concentration). For heavy labeling, 293T cells were labeled using the
Generation of PPTC7-knockout 293T cells. PPTC7 knockout in 293T cells was performed using the Akr system (Integrated DNA Technologies/IDT) for delivery of CRISPR-Cas9 reagents. A single guide RNA was selected toward exon 1 of PPTC7 (5'-CTCGGTCTCCTGTACGAGG-3') using the crispr.mit.edu tool, and was ordered as an Alt-R CRISPR-Cas9 crRNA (IDT). This crRNA, along with ATTO550-TracrRNA (IDT no. 107592) were used to generate a TracrRNA–crRNA complex, which was incubated in equimolar amounts (1 μm each) with AltR Cas9 V3 Nuclease (IDT no. 1081058). The specificity of this antibody was validated using wild type and PPTC7−/− mouse embryonic fibroblasts derived from a previously generated Ppctc7−/− mouse model.

MitoTox screen conditions. 293T or PPTC7-knockout 293T cells were split and plated in 24-well plates at 7.5 × 10^4 cells per well. Cells were allowed to adhere overnight, and medium was replaced with fresh DMEM or HPLM for a total of 24 h before harvest and collection of cells. Compound treatments were grouped into 6-h or 24-h incubations, with 6-h compound treatments occurring in the last 6 h of the 24-h medium change, and 24-h compound treatments occurring for the entire 24 h of medium treatment. Compounds used for 6 h include antimycin A (5 μM final concentration, Sigma no. A8674), rotenone (5 μM final concentration, Sigma no. R8875), oligomycin (2.5 μM final concentration, Sigma no. O4876), CCCP (10 μM final concentration, Sigma no. C2759), valinomycin (1 μM final concentration, Sigma no. V0627) and CDDO (2.5 μM final concentration, Cayman Chemical no. 11883). Compounds used for 24 h include doxycycline (10 μg ml−1 final concentration, VWR no. 75844-668) and DFO (100 μM final concentration, Sigma no. D9533). One compound, 4-NB (1 mM final concentration, Sigma no. 461091), requires 6–9 d for efficacy, and thus cells were treated with this compound for 5 d before being split to 7.5 × 10^4 cells per well and grouped with the 24-h incubations. Control, untreated 293T cells were split and collected with both the 6- and 24-h compound treatment sets. All conditions were plated and collected in three replicate wells.

To generate an internal control for each sample, SILAC-heavy-labeled 293T cells (see ‘Cell Culture’ for details) were spiked into lysis buffer in at ~1:1 ratios of signal to light samples, as determined by mass spectrometry (corresponding to 8 × 10^4 heavy-labeled cells per well of light cells). Heavy cells were counted and resuspended at a final concentration of 8 × 10^4 heavy cells in 100 μl of this buffer containing 50 mM HEPES, 150 mM NaCl, 10 mM glucose, 1 mM pyruvate and 2 mM glutamine. Oxygen-consumption rates (OCR) and extracellular acidification rate was monitored on a Seahorse xFp96 basally, and in the presence of a Seahorse XF Cell Mito Stress Test (Agilent cell–strectomy in the Stress Test, cells were treated with oligomycin (1 μM final concentration), FCCP (1 μM final concentration) and rotenone and antimycin A (0.5 μM final concentration). After the assay, the cells were fixed with 1% glutaraldehyde and stained with 1.5% crystal violet, and, after release of the stain with 10% acetic acid, each well was read at an absorbance of 590 nm (ref. 14). These absorbance values were used to normalize each assayed well within the Wave software (version 2.6.0). Data were exported from the Wave software and analyzed using Prism (version 8).

Peptide quantification from DI-SPA. Peptides were quantified using custom code written in Python and R available from: https://github.com/igmeyerucsd/DI-SPA-Technologies and access to specific files for quantification in Python. To perform quantification, at least one of the three most abundant y-ions (either heavy or light) was required to be observed within 10 ppm, unless otherwise noted. The median ratios of heavy/light were determined from those y-ions (up to three of the most abundant). If the heavy or light partner was not detected, then the average value of the ten least abundant peaks in the MS/MS spectra was used as noise for the missing partner ion to compute a ratio. For the whole-cell MitoTox samples, at least one heavy or light y-ion was required to be observed within 12 ppm of the expected mass to compute quantification. For the enriched mitochondria samples, data were collected with a maximum ion-injection time of 502 ms and a resolution of 240,000 in the orbitrap. This higher resolution data was analyzed with more stringent requirements; all 3 pairs of the 3 most abundant heavy and light ions were required to be detected within 10 ppm to report quantification. For the MitoTox whole cell and enriched mitochondria experiments, distributions of heavy-light ratios from each sample were scaled to have the same center before statistical analysis using the sklearn feature StandardScaler.

Statistics. Unless otherwise noted, statistical tests used for data presented in main and extended data figures were independent two-sample, two-tailed t-tests assuming equal variance. Exactly three replicate biological samples from independent cell cultures were compared in all statistical tests (for example, separate wells in a multiwell plate). Replicates were from one independent experiment. Exact P values are available in the legend or source data table, and experiments were not replicated. The supplementary data file contains tables of ANOVA with F statistics, P values and degrees of freedom for all quantified proteins, compared across factors and interactions in the multifactorial experiment.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
All raw data (along with the Excel sheet giving details of each file), filtered and unfiltered search results and quantification files are available on MassIVE under the dataset identifier MSV000085156 (https://doi.org/10.25355/CSM686), The raw data files are relevant for the human FANTOM v7 database 2019-03-14-td-UP000005640.fasta. Detailed descriptions of the RAW data files are on MassIVE under the folder ‘other’ in the Excel file ‘Raw data files descriptions v3.xlsx’. The massive repository includes the human spectral libraries for use with MSPLIT-DIA and the files used to create libraries. Source data is provided with this paper.

Code availability
All code availability is written in Python and R and is available on GitHub from https://github.com/igmeyerucsd/DI2A or from Zenodo (https://doi.org/10.5281/ZENODO.4119303).
53. Gupta, N. & Pevzner, P. A. False discovery rates of protein identifications: a strike against the two-peptide rule. *J. Proteome Res.* 8, 4173–4181 (2009).

54. Shanmugam, A. K., Yocum, A. K. & Nesvizhskii, A. I. Utility of RNA-seq and GPMDB protein observation frequency for improving the sensitivity of protein identification by tandem MS. *J. Proteome Res.* 13, 4113–4119 (2014).

55. Forsman, U., Sjöberg, M., Turunen, M. & Sindelar, P. J. 4-Nitrobenzoate inhibits coenzyme Q biosynthesis in mammalian cell cultures. *Nat. Chem. Biol.* 6, 515–517 (2010).

56. Kueng, W., Silber, E. & Eppenberger, U. Quantification of cells cultured on 96-well plates. *Anal. Biochem.* 182, 16–19 (1989).

57. Levitsky, L. I., Klein, J. A., Ivanov, M. V. & Gorshkov, M. V. Pyteomics 4.0: five years of development of a Python proteomics framework. *J. Proteome Res.* 18, 709–714 (2019).

58. Meyer, J. Code for analysis of data from direct infusion shotgun proteome analysis (DI-SPA) (Zenodo, 2020); [https://doi.org/10.5281/ZENODO.4115930](https://doi.org/10.5281/ZENODO.4115930)

Acknowledgements
We thank D. Hwang for help with preparing figures, A. Williams for help with writing, J. Mabry for assistance in the generation of the PPTC7 knockout cell line, A. Hebert for helpful discussions and the Cantor Lab for their generous gift of HPLM. This work was supported by the following NIH grants: T15 LM007359 (J.G.M.), P41 GM108538 (J.J.C.) and R01 DK098672 (D.J.P.).

Author contributions
Conceptualization, J.G.M., N.M.N. and D.J.P. Data curation, J.G.M. Formal analysis, J.G.M. Funding acquisition, J.G.M., D.J.P. and J.J.C. Investigation, J.G.M. and N.M.N. Methodology, J.G.M. and N.M.N. Project administration, J.G.M., D.J.P. and J.J.C. Resources, J.G.M., D.J.P. and J.J.C. Software, J.G.M. Supervision, J.G.M., D.J.P. and J.J.C. Validation, J.G.M. and N.M.N. Visualization, J.G.M. Writing—original draft, J.G.M. Writing—review and editing, J.G.M., N.M.N., D.J.P. and J.J.C.

Competing interests
J.J.C. is a consultant for Thermo Fisher Scientific. J.G.M., N.M.N. and D.J.P. have no competing interests.

Additional information
Extended data is available for this paper at [https://doi.org/10.1038/s41592-020-00999-z](https://doi.org/10.1038/s41592-020-00999-z).

Supplementary information is available for this paper at [https://doi.org/10.1038/s41592-020-00999-z](https://doi.org/10.1038/s41592-020-00999-z).

Correspondence and requests for materials should be addressed to J.G.M. or J.J.C.

Peer review information Arunima Singh was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints).
Extended Data Fig. 1 | Theoretical Analysis of Peptide Complexity Reduction by Gas-phase Fractionation quadrupole isolation width and FAIMS compensation voltage (CV). This analysis uses the maximum CV signal for all peptide precursor masses identified from stepped CV analysis of the human proteome. Stacked barplots show the number of peptide precursor masses per bin split by the contribution from each FAIMS CV fraction. The top panel shows precursor masses per 4 m/z isolation bin. The middle panel shows a roughly linear decrease in the maximum number of peptide precursor masses when the isolation width is decreased to 2 m/z. The bottom panel shows a nonlinear decrease in the number of peptide precursor masses due to selection by FAIMS gas phase fractionation at constant quadrupole isolation width.
Extended Data Fig. 2 | See next page for caption.
Extended Data Fig. 2 | Examples of infusion data traces. Tryptic peptides from the MCF7 proteome (1mg/mL) were infused as described for DI-SPA analysis, but precursor ions (MS1) were measured. a, Description of general flowgram parameters over time. b, MS1 trace of the no FAIMS experiment (top) and extracted ion chromatograms of various randomly chosen multiply charged m/z values (±10ppm) show a consistent pattern of elution for all masses. This suggests that peptides are not retained or separated in our setup. c, Comparison of the signal from without FAIMS versus FAIMS using each CV setting from −30 V to −80 V. d, Example flow-gram from DI-SPA-PRM-MS of 100 fmol/μL angiotensin II showing the typical smooth trace of mass- and FAIMS-selected peptide precursor.
Extended Data Fig. 3 | DI-SPA scouting experiments for untargeted peptide identification. a, Fixed and varied parameters for each of the parameter scouting experiments in Fig. 2a–c. Values highlighted in yellow were varied with the other values in that row fixed. b, Schematic of scouting experiment with actual data. Peptides were directly infused into the mass spectrometer over the duration of a scouting experiment. The first selection is performed by FAIMS according to compensation voltage (CV). FAIMS CV is fixed at a value between -30 volts and -80 volts while cycling through the second selection by m/z with the first quadrupole isolation window. is stepped across the m/z range of interest (400–1,000 here) to isolate specific subfractions of the peptide population. The FAIMS and quadrupole-selected peptides are fragmented by HCD, and finally the fragment ions are detected in the orbitrap to produce a tandem mass spectra. No precursor ion scans (MS1) are collected. MS/MS spectra from DI-SPA are identified by spectral library search.
Extended Data Fig. 4 | Enriched KEGG pathways including all protein members of those pathways identified by DI-SPA (matching Fig. 2e). Pathway enrichment analysis was done in Cytoscape with the plugin clueGO. Larger circles correspond to lower corrected p-value of term enrichment, and the colored portion of the circle gives the proportion of proteins in that pathway that were identified.
Extended Data Fig. 5 | Robustness and reproducibility of DI-SPA. Tryptic peptides from the MCF7 proteome (1 mg/mL) were analyzed 100 times with a shortened version of the parameter scouting method (Extended Data Fig. 3). a, TIC traces of the infusion data from injection #1, #25, #50, #75, #100, and those five overlaid. b, The number of peptide identifications from MSPLIT-DIA per analysis (FDR < 0.01) and (c) the distribution of peptide identifications summarized as a boxplot. The boxplot shows the median (percentile 50%) with an orange line, and the box represents the inner quartile range (IQR) Q1 and Q3 (percentiles 25 and 75). Whiskers show Q1 - 1.5*IQR and Q3 + 1.5*IQR.
Extended Data Fig. 6 | Application of DI-SPA to human plasma. Two different purchased human plasma samples were analyzed by DI-SPA-MS using the parameter scouting method strategy shown in Extended Data Fig. 3. The number of identifications for the two sources of human plasma were compared with and without depletion.
Extended Data Fig. 7 | Workflow for preparation of standard samples to assess quantitative DI-SPA. A549 cells were grown in DMEM media containing either light lysine and arginine (LIGHT) or $^{13}$C$_6$-$^{15}$N$_2$-lysine and $^{13}$C$_6$-$^{15}$N$_4$ L-arginine (HEAVY) and then combined at various ratios including: 1:8, 1:4, 1:2, 1:1, 2:1, 4:1, and 8:1 (HEAVY:LIGHT). Samples were then lysed proteins were reduced and alkylated, and proteolysis was initiated with trypsin. Peptides from trypsin digestion were desalted and then data was collected in parallel with either traditional nanoLC-MS/MS to verify SILAC ratios and provide a benchmark, or with DI-SPA to determine quantitative quality. Data from nanoLC-MS/MS was analyzed using MaxQuant to identify and quantify peptides, and data from DI-SPA was analyzed with MSPLIT-DIA and custom code in python and R.
Extended Data Fig. 8 | Examples of relationships between DI-SPA data collection settings for different peptides and their corresponding proteins. Peptides that uniquely identify proteins are found with a combination of gas-phase fractionation by FAIMS and precursor mass isolation with the first quadrupole (Q1). In this example, two unique peptides from different proteins are co-isolated with FAIMS compensation voltage (CV) of $-50 \text{ V}$ and Q1 set to 437 m/z. A single peptide is isolated with CV of $-70 \text{ V}$ and the same Q1 setting of 437 m/z, and two more peptides are isolated with a CV of $-80 \text{ V}$ at Q1 set to 438.5 m/z. Library spectra are shown for each peptide. The three most abundant singly charged y-ions in the library spectra are used for peptide quantification unless otherwise noted.
Extended Data Fig. 9 | See next page for caption.
Extended Data Fig. 9 | Comparison of Quantification from peptides shared between LC-MS (MaxQuant) and DI-SPA analysis. Data are from peptides quantified with both methods from enriched mitochondria. Bands around the regression line show the 95% confidence interval.
Extended Data Fig. 10 | See next page for caption.
Extended Data Fig. 10 | DI-SPA quantification of proteins from mitochondria subcellular fractions. a, Overlap of 37 mitochondrial proteins quantified: 149 from the purified mitochondria experiment and 56 from the whole cell experiment. b, Heatmap showing trend of general decrease in the 149 proteins annotated mitochondrial from DI-SPA analysis of the purified mitochondria. Light is the signal from the experimental condition and heavy is the signal from the SILAC standard protein. *Benjamini-Hochberg adjusted p-value <0.05, exact corrected p-values: PHB=0.023, CISY=0.023, THIL=0.036, CH10=0.036. n = 3 independent biological replicates of mitochondria preparations from 293 T cells from one independent experiment. p-values are from a two-tailed T-test assuming equal variance, and corrected p-value is from Benjamini-Hochberg multiple hypothesis testing correction. Source Data is available as Supplementary Table 6.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- An indication of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
  *Give P values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

*Our web collection on statistics for biologists* contains articles on many of the points above.

Software and code

Policy information about availability of computer code

**Data collection** Mass spectrometry data was collected using ThermoFisher Foundationt software version 3.1 SP4, FreeStyle version 1.3 SP2, and Xcalibur 4.0.

**Data analysis**

Seahorse data was analyzed using Wave software version 2.6.0 and Prism version 8. Peptides were identified from raw mass spec data using the MSPLIT-DIA software version 1.0. Identifications were filtered to FDR thresholds and assigned to proteins with custom code written in R (version 3.5) and Python (version 3.7) that is available from github www.github.com/jgmeyeruusd/DIA

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. Github). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All raw data, filtered and unfiltered search results, and quantification files are available on massive under the dataset identifier MSV000085156 [https://doi.org/10.25345/CSM686]. The massive repository also contains the relevant human FASTA database under the directory "updates/2020-09-25_jgmeyer_53bcc7df/sequence/2019-03-14-td-UP0000005640.fasta". All MS data and result files are described in the file "Raw data files descriptions v3.xlsx" at the massive repository under the folder "other". The massive repository includes the human spectral libraries for use with MSPLIT-DIA, and the files used to create that library and enable creation of other libraries.
Field-specific reporting

Please select the one that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences  [ ] Behavioural & social sciences  [ ] Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/mr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Sample size |
|-------------|
| No sample size calculation was performed before performing experiments. For the qualitative parameter scouting experiment, one technical replicate sample was collected to survey gross differences in qualitative peptide analysis. For the quantitative method validation experiment, one technical replicate of the LC-MS and targeted DI-SPA method were collected, and triplicate technical replicates of the untargeted DI-SPA experiment were collected to compare gross differences in quantitative method quality. For the multifactorial experiment demonstrating the application of our method, we chose 3 samples per condition to enable a reasonable estimate of variance and enable statistical comparison between treatments. This choice was sufficient because we detect subtle effects due to good measurement reproducibility. For the mitochondria enrichment experiment, 3 samples of separate mitochondria enrichments were collected each for the WT and PPTC7 knockout 293T cells. This was chosen to look for gross quantitative differences between mitochondria proteins. For the seahorse respirometry experiment, 11 replicate wells were collected for each condition to ensure that subtle differences would be detected. |

| Data exclusions |
|----------------|
| No data was excluded from our studies. |

| Replication |
|-------------|
| The parameter exploration experiments were replicated at least twice and replication was successful. Replication was not attempted for the results from the multi-factorial experiment, or the mitochondria enrichment experiment because the focus was to demonstrate potential application rather than prove new biology. Replication of the quantitative validation experiment was not carried out because the disparate methods showed excellent agreement. |

| Randomization |
|---------------|
| All data collection was done in random order for the parameter scouting and mitochondria toxins experiments, and samples were treated randomly during sample preparation. |

| Blinding |
|---------|
| We did not use blinding in our study because we treated the samples randomly and had no expectation about the outcomes that might unconsciously bias results. |

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
| ☑   | Antibodies            |
| ☑   | Eukaryotic cell lines |
| ☑   | Palaeontology and archaeology |
| ☑   | Animals and other organisms |
| ☑   | Human research participants |
| ☑   | Clinical data |
| ☑   | Dual use research of concern |

| n/a | Involved in the study |
|-----|-----------------------|
| ☑   | ChIP-seq |
| ☑   | Flow cytometry |
| ☑   | MRI-based neuroimaging |

Eukaryotic cell lines

Policy information about cell lines

| Cell line source(s) |
|---------------------|
| 293T cells were obtained from ATCC (catalog #CRL-3216), A549 cells were from ATCC [catalog # CCL-185], and MCF7 cells were from ATCC [catalog # HTB-22]. |

| Authentication |
|----------------|
| Cells were not authenticated. |

| Mycoplasma contamination |
|--------------------------|
| All cell lines used in the study tested negative for mycoplasma contamination. |
Commonly misidentified lines (See ICLAC register)

No commonly misidentified cell lines were used in this summary.