Development of Mass Spectrometry-based Methods for the Analysis of Complex Biological Mixtures

By

Noah M. Lancaster

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Chemistry)

at the

UNIVERSITY OF WISCONSIN-MADISON

2026

Date of final oral examination: 04/20/2026

The dissertation is approved by the following members of the Final Oral Committee:

Joshua J. Coon, Professor, Biomolecular Chemistry
Lloyd M. Smith, Professor, Chemistry
Lingjun Li, Professor, Pharmacy
Amy M. Weeks, Assistant Professor, Biochemistry

To Lindsay, with all my love.

ACKNOWLEDGEMENTS

The first chemistry question that truly captivated me was, ‘How do scientists figure out what chemicals are present?’. I remember mournfully reflecting in my high school chemistry on how sad it was that once you combined two chemicals, there would never be a way to get them apart again (Sort of the opposite of the Humpty Dumpty story). So later on in college, my mind was blown when my introductory analytical chemistry class started discussing ways for separating and analyzing mixtures of chemicals, and I was immediately hooked on diving deep into understanding how chemicals are separated, measured, and analyzed. The completion of my doctoral work at UW-Madison is the culmination of many years of work and pursuit of this curiosity, and I wouldn’t have made it here without the support and encouragement of countless individuals. I want to take some time here to acknowledge some of the many people who have supported me on my scientific journey.

First, I would like to thank my advisor, Dr. Joshua Coon, for all of his time and support. In addition to giving me many opportunities to work on exciting mass spectrometry research, you have been an enormous part of my development as a scientist and professional researcher. At many points throughout my Ph.D., you encouraged me to step back and look at the big picture, which has helped me break out of my tendency to get stuck over the technical details. I am thankful for your ability to envision grand scientific goals and for your commitment to supporting the pursuit of those goals. I also very much appreciate your valuable feedback and mentorship as I’ve grown in my ability to effectively communicate my work and how to successfully handle collaborations.

I would also like to thank the many excellent colleagues I’ve had in the Coon Lab: Dr. Evgenia Shishkova, Dr. Alex Hebert, Dr. Michael Westphall, Dr. Katie Overmyer, Dr. Scott Quarmby, Mitch Probasco, Laura Van Toll, Heather Coon, Dr. Yuchen He, Dr. Laura Muehlbauer,

Dr. Dain Brademan, Dr. Yunyun Zhu, Dr. Austin Salome, Dr. Kenny Lee, Dr. Trent Peters-Clarke, Dr. Pavel Sinitcyn, Dr. Lia Serrano, Dr. Keaton Mertz, William Beimers, Annie Jen, Salma Abou Elhassan, Corinne Moss, Meg Robinson, Dr. Marcel Morgenstern, Daniel Nesbitt, Jack Chylstek, Dr. Ben Chadwick, Dr. Benton Anderson, Li-Yu Chen, Bingnan Zhao, Drew Jordahl, Zhimin Ma, Dylan Forbes, Dr. Ross Soens, Emily Johnson, Van Ledermann, Riley Rowland, Joanne Du, and Dr. Srija Chakraborty.

Of these excellent colleagues, I've had the privilege of working more closely with several. To Dr. Evgenia Shishkova, thank you for your time training me and supporting my early projects as I worked toward becoming a more independent scientist. To Dr. Katie Overmyer and Dr. Scott Quarmby, thank you for being excellent collaborators and always being willing to chat about project directions and technical challenges. To Mitch Probasco, I appreciate how genuine of a person you are and would not have been able to finish my projects without your help fixing and making new instrument parts. To Dr. Laura Muehlbauer and Dr. Dain Brademan, thank you for answering all my questions and helping me learn things when I first started out in the lab, and thank you to Laura for the help in my post-graduate job search! To Dr. Austin Salome, thank you for being an excellent friend and a great colleague on my first project in the lab and for always being down to chat. To Dr. Pavel Sinitcyn, thank you for being an excellent co-author and helping me learn the paper writing process during our phospho project- your curiosity and passion for science is inspiring! To Dr. Trent Peters-Clarke and Dr. Lia Serrano, thank you for being great collaborators on all of the Astral projects – it was a pleasure. To Corinne Moss and Meg Robinson, it was great working with y'all to get the Orbitrap Astral up and running in the lab. To Li-Yu Chen, Bingnan Zhao, and Dr. Benton Anderson, I thoroughly appreciate your contributions and support

of the SynchroSep-MS project. To William Beimers, Annie Jen, Salma Abou Elhassan, Dr. Marcel Morgenstern, and Jack Chlystek – you have all been exceptional colleagues and I appreciated the opportunities to work together on various things in the lab. To all the newer members of the lab, I wish you all the best in your future scientific journeys!

Outside of the Coon Lab, there are many great folks I've gotten to work with over the years. The Coon Lab's ongoing research collaborations with Thermo Fisher Scientific were a significant part of my Ph.D., and there are many folks I'd like to thank for their support. For the Orbitrap Astral projects, I am very thankful to have gotten to work with great scientists, including Dr. Tabiwang Arrey, Dr. Anna Pashkova, Dr. Hamish Stewart, Dr. Christian Hock, Dr. Eugen Damoc, Dr. Amelia Peterson, and Dr. Vlad Zabrouskov. I would especially like to thank Tabiwang, Anna, Eugen, and Vlad for being exceptional hosts during my research visits to the Bremen factory. For other collaborations with Thermo Fisher Scientific, I would like to thank Dr. Graeme McAlister for being a great go-to person and for also giving helpful feedback on my projects. To Dr. John EP Syka, Dr. Christopher Mullen, and Dr. Josh Hinkle, thank you for providing support and always sharing your thoughtful feedback on research challenges during our calls and also for hosting me for my research internship at the San Jose R&D site – I learned so much and grew so much during my internship and felt that it was an incredibly valuable part of my Ph.D. journey. I would also like to thank Dr. Robin Sutka for being a great resource to the Coon Lab (both in terms of helping us get the supplies we needed and for bringing snacks when she visited!). There are many other external collaborators who were invaluable to my work. Thank you to Nicholas Arp, Dr. Jing Fan, Juli Hansen, Dr. Andrea Galmozzi, Dr. Julie Rojas, Dr. Audrey Gasch for providing the biological specimens for the phosphoproteomics study. Thank you to Dr. Patrick Forny for being an excellent

co-author and contributor to the phosphoproteomics study and to Dr. Caroline Fecher, Dr. Andrew Smith, and Dr. David Pagliarini for your support of that project. To Dr. Vadim Demichev, your contributions to the SynchroSep-MS project were essential and timely, and I would also like to thank Dr. Daniel Polasky and Dr. Alexey Nesvizhskii for software support.

At UW-Madison, I also benefitted from my experiences as a graduate ambassador for the Wisconsin Alumni Research Foundation (WARF) and want to thank Dr. Beth Fischer and Amanda Naas, as well as all my fellow WARF ambassadors and other WARF staff for making it a great experience.

To Dr. Shane Mangold, thank you for your generosity and willingness to participate in the UW-Madison Chemistry Industry Mentorship program – our conversations were helpful and meaningful as I was planning my next steps after graduate school, applying to jobs, and working to wrap up at UW-Madison. I appreciate your dedication to being a mentor for me and I hope to have a chance to pay it forward in the future.

During my 4th year in the Ph.D. program, I was fortunate to be the inaugural recipient of the Dahlberg-Klessig Distinguished Fellowship awarded for displaying creativity, curiosity, collegiality and perseverance in pursuit of their PhD thesis research and acknowledge the generous contributions of Dr. Jim Dahlberg and Dr. Dan Klessig to make the fellowship opportunity possible.

I would like to thank each of the members of my committee for being a valuable source of feedback as I've progressed through my Ph.D. work: Dr. Lloyd Smith, Dr. Lingjun Li, and Dr. Amy Weeks. Your thoughtful feedback and insight have been an essential part of my progress in graduate school.

During my time at UNC-Chapel Hill, I benefitted greatly from the opportunity to learn how to do research in the Lee Lab and am grateful to Dr. Andrew Lee and Dr. Paul Sapienza for their encouragement and support as I started out in science and learned all sorts of things about protein chemistry. I'd also like to thank Dr. Yoandris Del Toro Duany and Artjohn Villafania for their mentorship and support during my research internship at Merck and Co. I thank Dr. Gary Glish for his support – learning about mass spectrometry during your course was an outstanding experience and helped guide me to the excellent opportunities I've had at UW-Madison!

I would like to express gratitude for Dr. Peggy Morphew – your openness to my questions and curiosity helped inspire my passion for science during my time in your biology class at North Stanly High School.

I would like to express gratitude for my very supportive family. To my grandmothers, Ellen Lancaster and Peggy Byrd, thank you for your love and support. To my parents, Michael and Janet Lancaster, thank you for teaching me to be someone who values hard work and striving for excellence. I could not have made it here without all your love and support. To my twin brother, Jacob Lancaster, thank you for being the best brother and friend I could ever ask for. Your commitment to your calling as a minister and your love for others and the world inspires me!

Finally, I want to express my deepest love and gratitude to my beautiful fiancée, Lindsay Fisher. I'm so grateful to have someone so loving and supportive and who understands me in a way I'd never known or imagined before. I'm so thankful that we were brought together in Madison and I'm looking forward to the rest of my life with you, my love!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

%RSD	percent relative standard deviation
1D	one-dimensional
2D	two-dimensional
2D-E	Two-dimensional electrophoresis
ACN	Acetonitrile
AGC	automatic gain control
BAT	brown adipose tissue
BCA	bicinchoninic acid
BSA	bovine serum albumin
CAD	collisionally activated dissociation
CAMK	Ca2+/calmodulin-dependent protein kinase
CDKs	cyclin-dependent kinases
COVID-19	Coronavirus disease 2019
CPS1	carbamoyl phosphate synthetase I
CRP	C-reactive protein
DDA	data-dependent acquisition
DIA	data-independent acquisition
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
DNM1L	Dynamin-1-like protein

DOA	dominant optic atrophy
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal Growth Factor
ERKs	extracellular signal-regulated kinases
ESI	Electrospray ionization
ETD	electron transfer dissociation
FAIMS	high field asymmetric ion mobility spectrometry
FBS	Fetal Bovine Serum
FDP	false discovery proportion
FDR	false discovery rate
FWHM	full width half maximum
GDP	Guanosine diphosphate
GED	GTPase effector domain
GSEA	Gene Set Enrichment Analysis
GSK3A	glycogen synthase kinase-3 alpha
GTP	Guanosine triphosphate
HBB1	Hemoglobin subunit beta-1
HBB2	Hemoglobin subunit beta-2
HCD	Higher-energy Collisional Dissociation
HCTT	high-capacity transfer tube
hEGFR	human epidermal growth factor receptor

HEK293T	Human embryonic kidney 293 transformed
HER2	human epidermal growth factor receptor 2
HGMD	Human Gene Mutation Database
HPLC	High-performance liquid chromatogram
ID	Inner Diameter
IMAC	immobilized metal affinity chromatography
IMDM	Iscove's modified Dulbecco's medium
IMS	intermembrane space
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	knockout
LC	Liquid Chromatography or Liquid Chromatograph
LC-MS	Liquid Chromatography - Mass Spectrometry
LC-MS/MS	liquid chromatography-mass spectrometry/mass spectrometry
LDL-C	low density lipoprotein cholesterol
LDLR	LDL receptor
LOWESS	locally weighted scatterplot smoothing
m/z	mass-to-charge
MBR	match-between-runs
MFN1	mitofusin 1
MFN2	mitofusin 2
MIM	mitochondrial inner membrane

MIPS	Monoisotopic Precursor Selection
MOAC	metal oxide affinity chromatography
MOM	mitochondrial outer membrane
mRNA	Messenger Ribonucleic Acid
MS	mass spectrometry
MS/MS	mass spectrometry/mass spectrometry
MTS	mitochondrial transition signal
NAG	N-acetyl-L-glutamate
nanoESI	nano-electrospray ionization
nanoLC	nano-liquid chromatograph
nESI	nano-electrospray ionization
nLC	nano-liquid chromatograph
nLC-MS/MS	nanoflow liquid chromatography - mass spectrometry/mass spectrometry
OD	Outer Diameter
OPA1	Optic atrophy protein 1
PBS	Phosphate buffered saline
PCA	principle component analysis
PCKA/PCKB	Protein kinase C alpha/Protein kinase C beta
PCR	Polymerase Chain Reaction
PCSK9	proprotein convertase subtilisin/kexin type 9
PDB	Protein Data Bank

PHKG2	phosphorylase kinase catalytic subunit gamma 2
pLDDT	predicted local distance difference test
ppm	parts per million
PRTC	peptide retention time calibrant
PSM	peptide spectral match
PTM	post-translational modification
QC	Quality Control
RCSB	Research Collaboratory for Structural Bioinformatics
RF	radio frequency
RSD	relative standard deviation
RT	retention time
S/N	Signal-to-Noise Ratio
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SPE	Solid Phase Extraction
subWAT	Subcutaneous White Adipose Tissue
SWATH-MS	sequential window acquisition of all theoretical mass spectra
TCA	Tricarboxylic acid cycle
TCEP	tris(2-carboxyethyl)phosphine
TFA	trifluoroacetic acid
TIC	total ion current
Ti-IMAC	Titanium Immobilized Metal Affinity Chromatography

TIMS	trapped ion mobility spectrometry
TMT	tandem mass tags
TM	transmembrane domain
TOF	time-of-flight
t-SNE	t-distributed stochastic neighbor embedding
UHPLC	Ultra-High Performance Liquid Chromatography
UPLC	Ultra-Performance Liquid Chromatography
UV	Ultraviolet
visWAT	Visceral White Adipose Tissue
WAT	White Adipose Tissue
YPD	Yeast extract Peptone Dextrose

ABSTRACT

The analysis of mixtures is a fundamental challenge in analytical chemistry, and this challenge is especially manifest in the analysis of proteomes, which are among the most complex mixtures known in nature. Liquid chromatography-mass spectrometry/mass spectrometry (LC-MS/MS) has become a premier tool for the analysis of proteomes. This dissertation presents advances in LC-MS/MS methodologies to improve the comprehensive characterization of proteomics samples while improving throughput and maintaining high accuracy and reproducibility in order to improve biological discovery. **Chapter 1** provides a brief history of developments in analytical chemistry that enabled the analysis of complex proteomics samples and outlines some current challenges in proteomics analyses. **Chapter 2** describes the development of a new phosphoproteomics method using the novel Orbitrap Astral mass spectrometry platform and its application to generate an expanded atlas of protein phosphorylation in the mouse. **Chapter 3** introduces SynchroSep-MS, a new paradigm for sample multiplexing without the need for chemical derivatization. This method performs synchronized separations across multiple liquid chromatography columns with a single mass spectrometer to improve proteome analysis throughput. **Chapter 4** presents a detailed characterization of the effect of emitter positioning on signal intensity for nanoflow electrospray, provides detailed insight into emitter positioning effects, and suggests a path toward improving reproducibility for experimental setups with multiple emitters. Finally, **Chapter 5** presents conclusions, limitations, and future directions of the work presented in this dissertation as well as an outlook on how these developments fit into the broader proteomics field.

Chapter 1
Background and Introduction

Analyzing Chemical Mixtures

The field of analytical chemistry has played an integral role in the development of many key technologies. In the announcement of *Zeitschrift für analytische Chemie*, the first analytical chemistry-specific journal, Carl Fresenius writes that

> It requires little knowledge to realize that all the major developments in the field of chemistry are connected in some way with developments in the methods of analytical chemistry…. The development of analytical chemistry, therefore, always preceded the development of general chemistry. Just as new routes lead to new discoveries so improved analytical methods always yield new chemical results.[1,2]

The veracity of this claim is manifest in notable accomplishments including the development of atomic theory, which required accurate measurements of weight,[1] the discovery of oxygen, which relied on methods for gas analysis,[1] and the discovery of multiple elements in mineral samples using spectroscopic methods.[3] Each of these examples required the detection of specific chemical species amidst a mixture of other components. The challenge of mixture analysis was noted by Wilhelm Ostwald who writes in his seminal work *Scientific Foundations of Analytical Chemistry* that "…the task of recognizing any given substance … is always more or less easy of accomplishment … But the problem becomes far more complicated when we have to deal, not with a simple substance, but with a mixture…"[4] In the real world, even 'purified' chemicals exhibit trace level impurities that must be analyzed in the context of other components.[5] The analysis of chemical mixtures requires methods that can effectively distinguish between different chemical species. The degree to which a method can distinguish between different analytes (*i.e.* avoid interference) is known as its selectivity[6], a critical consideration when evaluating the suitability of a method for analyzing a mixture of different analytes. A further challenge in mixture analysis is obtaining quantitative values for the different components. An early example of the quantitative

analysis of a mixture was the characterization of the gaseous components of air by Henry Cavendish, which determined the percent composition of oxygen in air as well as detection of a trace amount of a gas later to be identified as argon.[1,7–9] Overall, the quantitative analysis of mixtures has long been a key measurement challenge due to the need for selective detection and accurate quantification.

Classical techniques for mixture analysis involved the use of many characteristic reactions for specific analytes to purify, isolate, and identify chemical analytes. The introduction of instrumental analysis techniques in the early 20th century transformed the practice of analytical chemistry by emphasizing direct measurement of the physical properties of analytes rather than characterization with chemical reactions.[10] For mixture analysis, instrumental analysis brought many advantages such as lower sample material requirements and higher selectivity for more direct analysis of substances without the need for extensive chemical purification and isolation.[10] One notable instrumental technique for the analysis of mixtures is liquid chromatography. The discovery of the chromatographic phenomena can be attributed to Friedrich Runge, who observed the separation of colored dyes on filter paper in 1855.[1,11,12] The development of this technique into its modern form is attributed to M.S. Tswett, who separated plant pigments using a filter funnel packed with a 'stationary phase'.[13,14] Later on, the introduction of high-performance liquid chromatography (HPLC) in the 1960s took advantage of advances in stationary phase technology and instrumentation development to offer high efficiency separations of a wide variety of analytes with short analysis times.[12] Notably, HPLC represents a separation approach that can be generally applied to a wide range of analytes and samples,[15] as compared to classical approaches for mixture analysis that relied on a suite of different chemical reactions that had to be selected based on the

specific analytes.[10] The later development of ultrahigh-performance liquid chromatography (UPLC or UHPLC) in late 1990s with its higher separation efficiency enabled much higher resolution analysis of complex mixtures.[16–18]

While liquid chromatography offers high resolution separation of mixtures, the detection and identification of analytes is another challenge in the analysis of complex mixtures. Mass spectrometry offers many advantages as a tool for detection, identification, and quantification of mixture components. This technology began with the work of J.J. Thomson, who was able to measure mass spectra of a few species, including hydrogen, carbon dioxide, and neon.[19] The further development of the instrument by F.W. Aston improved the spectral resolution, enabling the measurement of many isotopes, and paving the way for further instrumental development.[20] Later on, the coupling of liquid chromatography to mass spectrometry to analyze mixtures of non-volatile compounds was most effectively accomplished with electrospray ionization.[21] The development of tandem mass spectrometry, made more widely accessible by the invention of the triple-quadrupole mass spectrometer by Richard Yost and Christie Enke[22], offers a way to obtain detailed structural information about ionized analytes.[23] Combined with high-performance liquid chromatography, the high selectivity and molecular specificity of tandem mass spectrometry offers an approach for comprehensive mixture analysis. The introduction of high resolution/high mass accuracy mass analyzers further expands the capability of liquid chromatography-mass spectrometry/mass spectrometry (LC-MS/MS) workflows by improving the resolution of different analytes and increasing the confidence of compound identification.[24] The culmination of these developments in mass spectrometry combined with modern liquid chromatography have led to

LC-MS/MS becoming a leading approach for the characterization of complex mixtures, especially for the analysis of complex mixtures of proteins.

Measuring Proteins: An Introduction to Proteomics

Proteins are a diverse group of biomolecules known for carrying out diverse functions to support life. Within cells, genetic information stored on deoxyribonucleic acid (DNA) molecules is transcribed into messenger ribonucleic acid (mRNA) molecules. The information contained on mRNA is then translated into proteins by ribosomes. Proteins can perform diverse functions, ranging from structural support in cells, replication of DNA, to facilitating the synthesis of amino acids. Each of these distinct functions are essential for maintaining homeostasis, and by modulating protein expression rates, degradation rates, or the activity of a particular protein, cells and organisms can respond to stimuli in their environments. For this reason, proteins are often targeted by therapeutics. For example, overexpression of the human epidermal growth factor receptor 2 protein (HER2) is associated with increased severity for breast cancer. Trastuzumab, an anti-HER2 antibody inhibits HER2 activity to reduce HER2-mediated intracellular signal and prevent tumor growth.[25,26] As another example, familial hypercholesterolemia, a disease indicated by elevated levels of low density lipoprotein cholesterol (LDL-C), is often caused by a mutation in the gene encoding for the LDL receptor (LDLR) leading to reduced LDLR expression.[27] To address this issue, PCSK9 inhibitors, such as evolocumab, prevent PCSK9 from binding LDLR, reducing protein degradation and increasing LDLR levels.[28] In addition to serving as drug targets, measuring proteins can provide insight into the state of a biological system, as evident in the use of protein measurements for a wide variety of medical diagnostics. Some examples include the measurement of human chorionic gonadotropin levels by many common pregnancy tests[29,30],

detection of SARS-CoV-2 proteins by COVID-19 rapid antigen tests[31–34], and quantification of C-reactive protein (CRP) to assess coronary artery disease risk.[35–38]

As biopolymers comprised of 20 common amino acid monomers, proteins exhibit a wide range of chemical diversity that must be considered when developing analytical methods. The diverse chemical properties of amino acid sidechains lead to proteins exhibiting wide ranges of hydrophobicity and acidity, with total molecular weights ranging from a few kDa to multiple MDa.[39] Additionally, the presence of >700 known post-translational modifications (based on the UniProt database at the time of writing[40]), and alternative sequences/isoforms result in estimates of >1,000,000 different chemical forms of just human proteins.[41,42] Furthermore, small modifications of proteins can lead to large biological effects. For instance, one of the most common causes of sickle cell disease is caused by the substitution of a single glutamic acid residue to a valine on the β globin chain.[43–45] Human epidermal growth factor receptor (hEGFR) is known to be phosphorylated at multiple sites, with serine 229 phosphorylation affecting drug responsiveness due to nuclear translocation, tyrosine 1045 phosphorylation inducing EGFR degradation which plays an anti-cancer role, and phosphorylation of tyrosine 1068 and 1086 leading to Stat3 activation which is associated with cancer.[46–48] These examples indicate the importance of obtaining site-specific information about protein modifications. Consequently, to provide the most useful information, analytical methods to characterize proteins must distinguish between minor chemical differences across protein species.

Beyond the variation of chemical properties for proteins, biological systems can contain a few thousand (*i.e.* yeast[49]) to >16,000 proteins[50] expressed at any given time. The dynamic range of abundances for different proteins can exceed ten orders of magnitude for samples such as blood

plasma.[51,52] The broad diversity of molecular forms, large amount of different proteins, and wide ranges of concentrations make biological mixtures of proteins among the most complex mixtures known in nature.[53] This complexity presents a huge challenge for developing analytical mixtures that can distinguish between so many types of proteins while also providing reliable and reproducible quantification, making the analysis of protein mixtures one of the key frontiers in modern-day analytical chemistry research.

The deep biological insight that can be obtained when performing comprehensive analysis of these complex protein mixtures has driven the development of the field of proteomics. The term 'proteome' was first coined by Marc Wilkins to describe the characterization of all the proteins within a biological system[54], analogous to efforts by the genomics community to characterize all the genes within organisms. While the genomics community has greatly benefited from the signal amplification provided by the polymerase chain reaction (PCR)[55], no comparable technology exists for proteins. Accordingly, the field of proteomics involves a rich history of analytical innovations to enable effective analysis.[56]

For reliable detection and identification of proteins, characterization of the amino acid sequence is crucial as this sequence determines the structural, chemical, and functional characteristics of proteins. The first protein sequence was determined by Frederick Sanger in 1951[57,58], and the Edman degradation approach to sequencing proteins was introduced just a few years prior.[59,60] Edman degradation became a primary approach to protein sequencing with the introduction of automated machines.[61,62] The invention of the triple quadrupole mass spectrometer[22,63] paved the way for the development of tandem mass spectrometry-based protein sequencing. The Hunt group introduced such a method in 1986 that allowed peptide sequencing

by performing chemical or enzymatic digestion of proteins followed by HPLC fractionation to generate samples that could be ionized via liquid secondary-ion ionization and fragmented via collision-activated dissociation (CAD).[64] This approach was later combined with electrospray ionization (ESI) for high sensitivity peptide sequencing.[65] These mass spectrometry-based protein sequencing methods (known as bottom-up proteomics) were eventually adopted more widely due to the improved tolerance to mixtures of peptides and higher sensitivity than the existing Edman degradation approaches.[66] Altogether, sequencing is an essential part of protein analysis as subtle protein sequence differences can lead to divergent functional roles.

The earliest measurements of complex mixtures of proteins were accomplished using two-dimensional electrophoresis (2D-E).[67–69] This technique allowed resolution of complex protein mixtures with the primary challenge being identification of the detected species. Tandem mass spectrometry helped to address this challenge by analyzing peptides generated from tryptic digests of gel spots[70–72], thus providing large-scale detection and identification of complex mixtures of proteins.[73–75] Although 2D-E offers high resolving power, the dynamic range for protein detection can be limited and identification of gel spots requires coupling to mass spectrometry, increasing workflow complexity.[76,77] Early one-dimensional chromatographic methods lacked the chromatographic peak capacity to match the resolution provided by 2D-E, leading to the use of multi-dimensional separations to demonstrate improved peak capacities and enable high proteomic depth (*i.e.* detection of large portions of the proteome) without 2D-E.[77–79] The improved separations afforded by such techniques increased the method's ability to resolve proteins across the multiple orders of magnitude of abundances observed in proteomics samples.[80] To capture the theoretical improvements achievable with smaller particle size stationary phases, ultrahigh

pressure separations in capillary columns were first demonstrated in the Jorgenson group[16,17] and have been shown to yield improved peak capacities.[81,82] The introduction of commercial ultrahigh-performance liquid chromatography (UHPLC) systems made high peak capacity peptide separations more accessible to the field.[83,84] The use of UHPLC systems has been extensively combined with offline pre-fractionation of proteomics samples to enable detection of thousands of proteins.[85–97] The development of higher resolution, more sensitive, and faster mass spectrometers alongside continual improvements in chromatography has paved the way for single-shot analysis, that is, directly analyzing a whole proteome digest with a single LC-MS/MS run.[98–102] Because of mass spectrometry's ability to resolve multiple co-eluting chromatographic peaks with high resolution, faster and more sensitive instruments can enable the analysis of more complex protein mixtures without pre-fractionation. Note that the detection of lower abundance proteins is diminished without pre-fractionation due to mass spectrometer's ability to sample the greater number of peptide analytes present in the more complex samples being limited by MS/MS acquisition rates[103] and increased co-elution leading to analyte suppression either due to ion suppression during electrospray ionization[104–107] or limits on the intra-scan dynamic range of the mass spectrometer.[108–111] However, initial results provide detection of many thousands of proteins from complex proteome samples, generating datasets that still provide much useful biochemical information despite not being as comprehensive of an analysis.[99,112]

Work by the Mann group in 2018 introduced the BoxCar acquisition method to improve the limit of detection for intact peptide precursor analysis in the Orbitrap mass analyzer.[108] This study demonstrated MS^1-based evidence for more than 10,000 proteins in a mouse brain sample using a 100-minute gradient, representing a substantial proportion of the proteins expected to be

expressed in mouse brain at a given time.[113] This work demonstrated that there was sufficient chromatographic resolution of peptides to present evidence for most of the proteins expected in the sample; however, the detection of all of these peptides required sacrificing most of the acquisition time to acquire MS^1 data, and relied heavily on matching MS^1 features to the retention times and accurate mass values for peptides successfully identified by MS^2 analysis in separate fraction analysis. This observation suggests that focusing on improving mass spectrometry speed and sensitivity to allow for more effectively sampling and identification of eluting peptides would contribute greatly to the development of methods that provide comprehensive proteome characterization while maximizing throughput.

Alongside enhancements in mass spectrometry hardware, innovations in data acquisition schemes have bypassed the limitations of traditional data-dependent acquisition (DDA). In data-dependent acquisition, MS^1 measurements of intact peptide precursors are performed and the peptide ions with the greatest signal intensity are selected for fragmentation as these are most likely to yield high quality fragmentation spectra. Because the instrument is making real-time decisions as to which precursor ions to sample for fragmentation, ion sampling exhibits a random behavior biased toward high intensity peptides, leading to poor overlap across replicate runs and limiting the sampling depth of the mass spectrometry analysis.[103,114] Data-independent acquisition (DIA) schemes such as SWATH-MS function by repeatedly iterating over a defined *m/z* range with larger isolation windows than those used in DDA and fragmenting all the precursor ions in each window, leading to deterministic ion sampling.[115–117] Whereas ion sampling in DDA is strictly limited by the scan speed of the instrument, DIA methods parallelize ion fragmentation (at the expense of increased spectral complexity), allowing more extensive precursor sampling that is less

constrained by the mass spectrometer scan speed. This feature can lead to lower limits of detection by alleviating the sampling bias toward high abundance peaks.[118–120] For DIA methods, the depth of detection is limited then by the dynamic range of the LC-MS/MS method. On the mass spectrometer side, the dynamic range of the method (the range of analyte abundances that can be measured by the method) is limited by ion transmission efficiency[121–123], instrument duty cycle[124,125], intra-scan dynamic range[108], analysis selectivity, and mass analyzer detection sensitivity. Features such as ion transmission efficiency, instrument duty cycle, intra-scan dynamic range, and mass analyzer detection sensitivity all continue to improve as instrument manufacturers further develop their hardware. For analysis selectivity, there have been many notable innovations that led to improved method performance. Instruments with traps for ion accumulation enable higher duty cycles (the percentage of time that an instrument is actively sampling/measuring an analyte) leading to improved method sensitivity[124], but they can demonstrate limited dynamic range due to space charge capacity of the traps.[110] Consequently, the presence of high intensity ions can impede the ability to detect lower intensity ions. In these cases, reducing the complexity of the mass spectra with a higher selectivity method reduces the chance that the dynamic range of the ions sampled will exceed the intra-scan dynamic range. For data-independent acquisition, using narrower isolation windows or reducing the number of precursors isolated per scan decreases the spectral complexity and improves the dynamic range of the mass spectrometry method, enabling detection of a wider range of analyte abundances and eliminating spectral interference for more confident detection and quantification.[118,126–130] Note that the ability to use narrower isolation windows for reduced spectral complexity is constrained by the scan speed of the instrument, as the time required to cycle through all the isolation windows must be sufficiently fast to enable

adequate chromatographic peak sampling for reproducible detection and quantification.[116,131] The *m/z* range that is iterated over can also be reduced to allow narrower isolation windows for better detection depth.[132] Other approaches to reducing spectral complexity use gas-phase ion mobility as an additional separation prior to mass spectrometry analysis.[133,134] These innovations in the data acquisition approach lay the groundwork for much of the work presented in this dissertation, as will be shown in **Chapters 2** and **3**. In conclusion, improvements in mass spectrometry methods and hardware that yield improved scan speed, resolution, and detection sensitivity are expected to yield higher quality proteome characterization, and evaluating and developing such technology and methods is the focus of this dissertation.

Proteome Analysis with Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS)

A typical bottom-up proteomics workflow consists of a few steps. The first step is sample preparation, which typically involves sample lysis, protein extraction, digestion of proteins into peptides, and sample cleanup.[135] With optimization for the specific sample type, this workflow enables analysis of proteomes from many diverse sources.[136–144] For the analysis of PTMs, enrichment steps are typically performed to assist with detection of low stoichiometry modifications, as even low stoichiometry modifications can have functional implications.[145,146] For example, phosphorylated peptides are enriched using strategies such as immobilized metal ion affinity chromatography (IMAC), metal oxide affinity chromatography (MOAC), immunoprecipitation, and ion exchange chromatography.[147,148] Samples are then subjected to online LC-MS/MS to generate data and, owing to the large amount of data generated, automated data processing is required. The output of this workflow is data tables showing peptide and protein quantities across the samples analyzed. The comprehensive nature of these datasets allow the

investigation of many biological hypotheses that can enable discovery of new biological mechanisms by taking advantage of modern bioinformatic approaches.[149,150] The work in this dissertation primarily aims to address challenges in the data acquisition (*i.e.* LC-MS/MS) stage with a goal of improving detection depth and throughput while maintaining high accuracy and reproducibility.

In more detail, data acquisition begins with chromatographic separation of peptides followed by ionization. Nanoflow chromatography with electrospray ionization is a common choice for proteome analysis due to its high sensitivity.[151,152] Following ionization, gas-phase peptide ions are transmitted into the mass spectrometer via the atmospheric pressure interface. From here, the intact masses of the peptide ions can be measured via MS^1 analysis, or the peptide ions can be fragmented to provide sequence-informative fragment ions, enabling identification. Modern mass spectrometers incorporate many types of mass analyzers that offer unique advantages. For instance, Orbitrap analyzers provide high-resolution and mass accuracy[153], quadrupole mass filters enable isolation of ions for downstream analysis[154], quadrupole ion traps allow fast and sensitive detection[155], and time-of-flight (TOF) analyzers provide rapid scan speeds[156] with modern TOFs capable of high resolution and mass accuracy.[157]

After collecting LC-MS/MS data, data processing begins by performing an *in silico* digest of a protein sequence database to generate a list of peptides that could be observed. For each peptide, either the theoretically observable fragment ion masses are calculated or fragmentation spectra are predicted using machine learning methods.[158–162] Then, the mass spectral data is searched against these potential peptides with error rates being controlled via the target-decoy approach.[163] The intensity of peptide signals are integrated to generate quantitative values for each peptide, which

are then rolled up into ‘protein groups’ quantities, where protein groups are generated from the different proteins from which the peptide sequences could arise. This process, known as protein inference, involves a level of ambiguity as each digested protein can generate many peptides but peptide sequences might not be unique to a specific protein.[164] This ambiguity as to which proteins a peptide comes from is a key motivation for developments in ‘top-down proteomics’, where proteins are directly analyzed without a digestion step, the analytical challenges of which are beyond the scope of this dissertation.[165] Even with this ambiguity, bottom-up proteomics can provide comprehensive insight into the state of different biological systems. These workflows have been developed over many years and have already been successfully applied to many biological applications.[166–171] However, there remain outstanding challenges in the methodologies that could be improved to enhance data quality and the value of the proteomics methods.

Two key performance characteristics for proteomics analysis are depth and throughput. The depth of a proteomics method refers to the portion of the proteome a method can measure. Typically, proteomics methods are biased towards detection of higher abundant proteins, so a method with improved depth enables the detection of lower abundance proteins.[80] Achieving better proteomic depth increases the discovery potential of a study by enabling detection of a larger variety of proteins that might be involved with the biological mechanisms being investigated. As technology has developed, the depth achievable in a proteomics study has improved, although achieving comprehensive characterization frequently requires much longer analysis times, making deep proteome coverage out of reach for more routine analyses.[172] This point illustrates the constraints on analysis throughput when performing proteomics. Modern mass spectrometers are very expensive, and operation represents one of the most expensive parts of the workflow, so the

shorter it takes for data acquisition, the less the effective cost of the analysis due to decreased instrument usage. Additionally, larger numbers of samples are required to increase statistical power when working with human samples that exhibit substantial variance or trying to measure more subtle changes in the proteome, making analysis throughput more important to facilitate larger scale studies.[112] Generally, there exists a trade-off between the throughput and depth of analysis, but the demand for higher throughput methods also opens opportunities for innovation to improve the depth of detection achievable in shorter analysis methods. Overall, the improvement of the depth and throughput achievable would make comprehensive proteome analysis more accessible and enable larger-scale analyses for enhanced biological insight. Enabling deeper and more reproducible proteome coverage while enhancing throughput constitutes the main analytical challenges addressed in this dissertation as will be demonstrated by the following three studies.

As previously discussed, the presence of post-translational modifications (PTMs), such as phosphorylation, can have large effects on biological outcomes, making their measurement of great importance. However, deep phosphoproteome analysis with high throughput analysis remains challenging, with recent studies achieving limited coverage of the phosphoproteome[173] or requiring extensive library construction to achieve high depth.[134,174] To address some of the challenges in phosphoproteome analysis, **Chapter 2** introduces a new phosphoproteomics method developed on a novel mass spectrometer platform, the Orbitrap Astral, leading to substantial improvements in depth of coverage and analysis throughput through the use of narrow isolation window DIA. This method is shown to demonstrate high reproducibility, quantitative linearity, and site localization accuracy. Application of this method to characterize phosphorylation of different

mouse tissues results in an updated phosphorylation atlas of the mouse that serves as a valuable resource for the proteomics and broader biochemistry communities.

For standard proteome analysis, improved throughput is also desired. A common strategy for improved proteomics throughput is multiplexed analysis through chemical derivatization. A popular instantiation of this strategy is the use of isobaric tags, such as tandem mass tags (TMT), which involve derivatizing peptides with labeling reagents that impart a unique mass tag for each sample to be analyzed. This strategy is widely used and has been very successful, but suffers from multiple limitations, including high reagent cost, more complex sample preparation, and substantial quantitative bias.[175] Another approach to improved throughput has been focused on the chromatographic separation. As mass spectrometers have become faster and more sensitive, these instruments have become less of a constraint, leading to the chromatographic separation being more of a throughput bottleneck. The use of short chromatographic separations has been frequently employed[176–184], although shorter gradients generally decrease proteome depth and worsen quantitative precision and accuracy. For the nanoflow chromatography often used in proteomics for its high electrospray sensitivity, the low flow rates make sample loading and column washing/equilibration significant fractions of the total analysis time, especially as gradient separations become shorter. This factor has led some researchers to parallelize the sample loading and column washing/equilibration steps (often referred to as the LC overhead) on one column with the gradient separation on a second column.[185–198] While this approach does make better use of the mass spectrometer by minimizing dead time, the throughput can still be limited by gradient length, which must be long enough to yield acceptable method performance. To provide an alternative strategy for improved throughput, **Chapter 3** introduces the SynchroSep-MS method, which

introduces a new paradigm for sample multiplexing for proteome analysis to allow the analysis of multiple samples at the same time without the need for chemical derivatization. This method performs multiple liquid chromatography separations on multiple columns into a single mass spectrometer and uses a retention time offset to enable analysis of multiple samples at once. The initial implementation of this method enabled up to ~50% improvement in analysis throughput while maintaining most of the proteomic depth achievable with a conventional method and demonstrating high reproducibility and quantitative reliability. These results demonstrate how the increased scan capacities and high sensitivity of modern mass spectrometers can tolerate increased spectral complexity and enable this new approach to multiplexed sample analysis.

During the implementation of the SynchroSep-MS method, it was determined that the positioning of the emitters relative to the inlet was of particular importance to effective method implementation. Further review of the literature identified a need for a general characterization of how emitter positioning impacts signal intensity, especially for the high-capacity inlet capillary employed by multiple Thermo Fisher Scientific mass spectrometers. These platforms represent a significant portion of those used widely for proteome analysis.[199] **Chapter 4** presents a detailed characterization of the effect of emitter positioning on signal intensity in nanoflow electrospray ionization. This analysis utilizes a similar nanoflow setup to those widely employed in the field of proteomics and uses a tryptic digest of bovine serum albumin (BSA) to mimic typical bottom-up proteomic samples. The analysis demonstrates that emitter positioning results are consistent across the many peptides in the sample, suggesting that emitter positioning does not need to be separately optimized for different analytes, an important finding for the analysis of complex proteomics mixtures. Additionally, practical recommendations of positioning tolerances to improve

reproducibility in proteome analysis are provided. Finally, the results suggest approaches for minimizing variance across emitters to improve the performance of the SynchroSep-MS method.

The research described in each of these chapters addresses key challenges in proteome analysis via LC-MS/MS workflows. With proteomic samples representing some of the most complex mixtures found in nature, the work described in this dissertation presents a notable contribution to the field of analytical chemistry.

References

(1) Szabadváry, F.; Belcher, R.; Gordon, L. History of Analytical Chemistry: International Series of Monographs in Analytical Chemistry; Elsevier Science: Burlington, 1966.

(2) Fresenius, H. Zur Erinnerung an R. Fresenius. Z. Für Anal. Chem. 1897, 36 (1), III–XVIII. https://doi.org/10.1007/BF01348148.

(3) Weeks, M. E. The Discovery of the Elements. XIII. Some Spectroscopic Discoveries. J. Chem. Educ. 1932, 9 (8), 1413. https://doi.org/10.1021/ed009p1413.

(4) Ostwald, W.; McGowan, G. The Scientific Foundations of Analytical Chemistry, Treated in an Elementary Manner; London, New York, Macmillan and co., limited, 1900.

(5) Specker, H. Problems and Possibilities of Modern Trace Analysis. Angew. Chem. Int. Ed. Engl. 1968, 7 (4), 252–259. https://doi.org/10.1002/anie.196802521.

(6) Skoog, D. A.; Holler, F. J.; Crouch, S. R. Principles of Instrumental Analysis, Seventh edition.; Cengage Learning: Australia, 2018.

(7) Cavendish, H. XIII. Experiments on Air. Philos. Trans. R. Soc. Lond. 1784, No. 74, 119–153. https://doi.org/10.1098/rstl.1784.0014.

(8) Uprety, D. C.; Reddy, V. R.; Mura, J. D. Greenhouse Gases: A Historical Perspective. In Climate Change and Agriculture: A Historical Analysis; Uprety, D. C., Reddy, V. R., Mura, J. D., Eds.; Springer: Singapore, 2019; pp 31–41. https://doi.org/10.1007/978-981-13-2014-9_3.

(9) West, J. B. Henry Cavendish (1731–1810): Hydrogen, Carbon Dioxide, Water, and Weighing the World. Am. J. Physiol.-Lung Cell. Mol. Physiol. 2014, 307 (1), L1–L6. https://doi.org/10.1152/ajplung.00067.2014.

(10) Baird, D. Analytical Chemistry and the 'Big' Scientific Instrumentation Revolution. Ann. Sci. 1993, 50 (3), 267–290. https://doi.org/10.1080/00033799300200221.

(11) Touchstone, J. C. History of Chromatography. J. Liq. Chromatogr. 1993, 16 (8), 1647–1665. https://doi.org/10.1080/10826079308021679.

(12) Snyder, L. R.; Kirkland, J. J.; Dolan, J. W. Introduction. In Introduction to modern liquid chromatography; Wiley: Hoboken, N.J, 2010; pp 1–17. https://doi.org/10.1002/9780470508183.

(13) Ettre, L. S.; Sakodynskii, K. I. M. S. Tswett and the Discovery of Chromatography I: Early Work (1899–1903). Chromatographia 1993, 35 (3), 223–231. https://doi.org/10.1007/BF02269707.

(14) Ettre, L. S.; Sakodynskii, K. I. M. S. Tswett and the Discovery of Chromatography II: Completion of the Development of Chromatography (1903–1910). Chromatographia 1993, 35 (5), 329–338. https://doi.org/10.1007/BF02277520.

(15) Snyder, L. R. Modern Practice of Liquid Chromatography: Before and after 1971. J. Chem. Educ. 1997, 74 (1), 37. https://doi.org/10.1021/ed074p37.

(16) MacNair, J. E.; Lewis, K. C.; Jorgenson, J. W. Ultrahigh-Pressure Reversed-Phase Liquid Chromatography in Packed Capillary Columns. Anal. Chem. 1997, 69 (6), 983–989. https://doi.org/10.1021/ac961094r.

(17) MacNair, J. E.; Patel, K. D.; Jorgenson, J. W. Ultrahigh-Pressure Reversed-Phase Capillary Liquid Chromatography: Isocratic and Gradient Elution Using Columns Packed with 1.0-Mm Particles. Anal. Chem. 1999, 71 (3), 700–708. https://doi.org/10.1021/ac9807013.

(18) Tolley, L.; Jorgenson, J. W.; Moseley, M. A. Very High Pressure Gradient LC/MS/MS. Anal. Chem. 2001, 73 (13), 2985–2991. https://doi.org/10.1021/ac0010835.

(19) Thomson, J. J. Rays of Positive Electricity and Their Application to Chemical Analyses; Longmans, Green and Co, 1913.

(20) Beynon, J. H. The History of Mass Spectrometry and the Search for Zero. Biomed. Mass Spectrom. 1981, 8 (9), 380–383. https://doi.org/10.1002/bms.1200080906.

(21) Whitehouse, C. M.; Dreyer, R. N.; Yamashita, Masamichi.; Fenn, J. B. Electrospray Interface for Liquid Chromatographs and Mass Spectrometers. Anal. Chem. 1985, 57 (3), 675–679. https://doi.org/10.1021/ac00280a023.

(22) Yost, R. A.; Enke, C. G. Selected Ion Fragmentation with a Tandem Quadrupole Mass Spectrometer. J. Am. Chem. Soc. 1978, 100 (7), 2274–2275. https://doi.org/10.1021/ja00475a072.

(23) Watson, J. T.; Sparkman, O. D. Introduction to Mass Spectrometry: Instrumentation, Applications, and Strategies for Data Interpretation, 4th ed.; J. Wiley & sons: Hoboken (N.J.), 2007.

(24) Marshall, A. G.; Hendrickson, C. L. High-Resolution Mass Spectrometers. Annu. Rev. Anal. Chem. 2008, 1 (Volume 1, 2008), 579–599. https://doi.org/10.1146/annurev.anchem.1.031207.112945.

(25) Exman, P.; Tolaney, S. M. HER2-Positive Metastatic Breast Cancer: A Comprehensive Review. Clin Adv Hematol Oncol 2021, 19 (1), 40–50.

(26) Martín, M.; Pandiella, A.; Vargas-Castrillón, E.; Díaz-Rodríguez, E.; Iglesias-Hernangómez, T.; Martínez Cano, C.; Fernández-Cuesta, I.; Winkow, E.; Perelló, M. F. Trastuzumab Deruxtecan in Breast Cancer. Crit. Rev. Oncol. Hematol. 2024, 198, 104355. https://doi.org/10.1016/j.critrevonc.2024.104355.

(27) Raal, F. J.; Hovingh, G. K.; Catapano, A. L. Familial Hypercholesterolemia Treatments: Guidelines and New Therapies. Atherosclerosis 2018, 277, 483–492. https://doi.org/10.1016/j.atherosclerosis.2018.06.859.

(28) Shapiro, M. D.; Minnier, J.; Tavori, H.; Kassahun, H.; Flower, A.; Somaratne, R.; Fazio, S. Relationship Between Low-Density Lipoprotein Cholesterol and Lipoprotein(a) Lowering in Response to PCSK9 Inhibition With Evolocumab. J. Am. Heart Assoc. 2019, 8 (4), e010932. https://doi.org/10.1161/JAHA.118.010932.

(29) Cole, L. A. Immunoassay of Human Chorionic Gonadotropin, Its Free Subunits, and Metabolites. Clin. Chem. 1997, 43 (12), 2233–2243. https://doi.org/10.1093/clinchem/43.12.2233.

(30) Davies, S.; Byrn, F.; Cole, L. A. Human Chorionic Gonadotropin Testing for Early Pregnancy Viability and Complications. Clin. Lab. Med. 2003, 23 (2), 257–264. https://doi.org/10.1016/S0272-2712(03)00026-X.

(31) Aruleba, R. T.; Adekiya, T. A.; Ayawei, N.; Obaido, G.; Aruleba, K.; Mienye, I. D.; Aruleba, I.; Ogbuokiri, B.; Aruleba, R. T.; Adekiya, T. A.; Ayawei, N.; Obaido, G.; Aruleba, K.; Mienye, I. D.; Aruleba, I.; Ogbuokiri, B. COVID-19 Diagnosis: A Review of Rapid Antigen, RT-PCR and Artificial Intelligence Methods. Bioengineering 2022, 9 (4). https://doi.org/10.3390/bioengineering9040153.

(32) Rosella, L. C.; Agrawal, A.; Gans, J.; Goldfarb, A.; Sennik, S.; Stein, J. Large-Scale Implementation of Rapid Antigen Testing System for COVID-19 in Workplaces. Sci. Adv. 2022, 8 (8), eabm3608. https://doi.org/10.1126/sciadv.abm3608.

(33) Arshadi, M.; Fardsanei, F.; Deihim, B.; Farshadzadeh, Z.; Nikkhahi, F.; Khalili, F.; Sotgiu, G.; Shahidi Bonjar, A. H.; Centis, R.; Migliori, G. B.; Nasiri, M. J.; Mirsaeidi, M. Diagnostic Accuracy of Rapid Antigen Tests for COVID-19 Detection: A Systematic Review With Meta-Analysis. Front. Med. 2022, 9. https://doi.org/10.3389/fmed.2022.870738.

(34) Xie, J.-W.; He, Y.; Zheng, Y.-W.; Wang, M.; Lin, Y.; Lin, L.-R. Diagnostic Accuracy of Rapid Antigen Test for SARS-CoV-2: A Systematic Review and Meta-analysis of 166,943 Suspected COVID-19 Patients. Microbiol. Res. 2022, 265, 127185. https://doi.org/10.1016/j.micres.2022.127185.

(35) Rifai, N.; Ridker, P. M. High-Sensitivity C-Reactive Protein: A Novel and Promising Marker of Coronary Heart Disease. Clin. Chem. 2001, 47 (3), 403–411. https://doi.org/10.1093/clinchem/47.3.403.

(36) Zimmerman, M. A.; Selzman, C. H.; Cothren, C.; Sorensen, A. C.; Raeburn, C. D.; Harken, A. H. Diagnostic Implications of C-Reactive Protein. Arch. Surg. 2003, 138 (2), 220–224. https://doi.org/10.1001/archsurg.138.2.220.

(37) Casas, J. P.; Shah, T.; Hingorani, A. D.; Danesh, J.; Pepys, M. B. C-Reactive Protein and Coronary Heart Disease: A Critical Review. J. Intern. Med. 2008, 264 (4), 295–314. https://doi.org/10.1111/j.1365-2796.2008.02015.x.

(38) Bisoendial, R. J.; Boekholdt, S. M.; Vergeer, M.; Stroes, E. S. G.; Kastelein, J. J. P. C-Reactive Protein Is a Mediator of Cardiovascular Disease. Eur. Heart J. 2010, 31 (17), 2087–2091. https://doi.org/10.1093/eurheartj/ehq238.

(39) Walsh, G. Proteins: Biochemistry and Biotechnology; John Wiley & Sons, Incorporated: Newark, UNITED STATES, 2014.

(40) The UniProt Consortium. UniProt: The Universal Protein Knowledgebase in 2025. Nucleic Acids Res. 2025, 53 (D1), D609–D617. https://doi.org/10.1093/nar/gkae1010.

(41) Ponomarenko, E. A.; Poverennaya, E. V.; Ilgisonis, E. V.; Pyatnitskiy, M. A.; Kopylov, A. T.; Zgoda, V. G.; Lisitsa, A. V.; Archakov, A. I. The Size of the Human Proteome: The Width and Depth. Int. J. Anal. Chem. 2016, 2016 (1), 7436849. https://doi.org/10.1155/2016/7436849.

(42) Aebersold, R.; Agar, J. N.; Amster, I. J.; Baker, M. S.; Bertozzi, C. R.; Boja, E. S.; Costello, C. E.; Cravatt, B. F.; Fenselau, C.; Garcia, B. A.; Ge, Y.; Gunawardena, J.; Hendrickson, R. C.; Hergenrother, P. J.; Huber, C. G.; Ivanov, A. R.; Jensen, O. N.; Jewett, M. C.; Kelleher, N. L.; Kiessling, L. L.; Krogan, N. J.; Larsen, M. R.; Loo, J. A.; Ogorzalek Loo, R. R.; Lundberg, E.; MacCoss, M. J.; Mallick, P.; Mootha, V. K.; Mrksich, M.; Muir, T. W.; Patrie, S. M.; Pesavento, J. J.; Pitteri, S. J.; Rodriguez, H.; Saghatelian, A.; Sandoval, W.; Schlüter, H.; Sechi, S.; Slavoff, S. A.; Smith, L. M.; Snyder, M. P.; Thomas, P. M.; Uhlén, M.; Van Eyk, J. E.; Vidal, M.; Walt, D. R.; White, F. M.; Williams, E. R.; Wohlschlager, T.; Wysocki, V. H.; Yates, N. A.; Young, N. L.; Zhang, B. How Many Human Proteoforms Are There? Nat. Chem. Biol. 2018, 14 (3), 206–214. https://doi.org/10.1038/nchembio.2576.

(43) Ingram, V. M. Gene Mutations in Human Hæmoglobin: The Chemical Difference Between Normal and Sickle Cell Hæmoglobin. Nature 1957, 180 (4581), 326–328. https://doi.org/10.1038/180326a0.

(44) Ashley-Koch, A.; Yang, Q.; Olney, R. S. Sickle Hemoglobin (Hb S) Allele and Sickle Cell Disease: A HuGE Review. Am. J. Epidemiol. 2000, 151 (9), 839–845. https://doi.org/10.1093/oxfordjournals.aje.a010288.

(45) Kwiatkowski, D. P. How Malaria Has Affected the Human Genome and What Human Genetics Can Teach Us about Malaria. Am. J. Hum. Genet. 2005, 77 (2), 171–192. https://doi.org/10.1086/432519.

(46) Huang, W.-C.; Chen, Y.-J.; Li, L.-Y.; Wei, Y.-L.; Hsu, S.-C.; Tsai, S.-L.; Chiu, P.-C.; Huang, W.-P.; Wang, Y.-N.; Chen, C.-H.; Chang, W.-C.; Chang, W.-C.; Chen, A. J.-E.; Tsai, C.-H.; Hung, M.-C. Nuclear Translocation of Epidermal Growth Factor Receptor by Akt-Dependent Phosphorylation Enhances Breast Cancer-Resistant Protein Expression in Gefitinib-Resistant Cells*. J. Biol. Chem. 2011, 286 (23), 20558–20568. https://doi.org/10.1074/jbc.M111.240796.

(47) Levkowitz, G.; Waterman, H.; Ettenberg, S. A.; Katz, M.; Tsygankov, A. Y.; Alroy, I.; Lavi, S.; Iwai, K.; Reiss, Y.; Ciechanover, A.; Lipkowitz, S.; Yarden, Y. Ubiquitin Ligase Activity and Tyrosine Phosphorylation Underlie Suppression of Growth Factor Signaling by C-Cbl/Sli-1. Mol. Cell 1999, 4 (6), 1029–1040. https://doi.org/10.1016/S1097-2765(00)80231-2.

(48) Shao, H.; Cheng, H. Y.; Cook, R. G.; Tweardy, D. J. Identification and Characterization of Signal Transducer and Activator of Transcription 3 Recruitment Sites within the Epidermal Growth Factor Receptor1. Cancer Res. 2003, 63 (14), 3923–3930.

(49) Ghaemmaghami, S.; Huh, W.-K.; Bower, K.; Howson, R. W.; Belle, A.; Dephoure, N.; O'Shea, E. K.; Weissman, J. S. Global Analysis of Protein Expression in Yeast. Nature 2003, 425 (6959), 737–741. https://doi.org/10.1038/nature02046.

(50) Sinitcyn, P.; Richards, A. L.; Weatheritt, R. J.; Brademan, D. R.; Marx, H.; Shishkova, E.; Meyer, J. G.; Hebert, A. S.; Westphall, M. S.; Blencowe, B. J.; Cox, J.; Coon, J. J. Global Detection of Human Variants and Isoforms by Deep Proteome Sequencing. Nat. Biotechnol. 2023, 41 (12), 1776–1786. https://doi.org/10.1038/s41587-023-01714-x.

(51) Anderson, N. L.; Anderson, N. G. The Human Plasma Proteome: History, Character, and Diagnostic Prospects *. Mol. Cell. Proteomics 2002, 1 (11), 845–867. https://doi.org/10.1074/mcp.R200007-MCP200.

(52) Nanjappa, V.; Thomas, J. K.; Marimuthu, A.; Muthusamy, B.; Radhakrishnan, A.; Sharma, R.; Ahmad Khan, A.; Balakrishnan, L.; Sahasrabuddhe, N. A.; Kumar, S.; Jhaveri, B. N.; Sheth, K. V.; Kumar Khatana, R.; Shaw, P. G.; Srikanth, S. M.; Mathur, P. P.; Shankar, S.; Nagaraja, D.; Christopher, R.; Mathivanan, S.; Raju, R.; Sirdeshmukh, R.; Chatterjee, A.; Simpson, R. J.; Harsha, H. C.; Pandey, A.; Prasad, T. S. K. Plasma Proteome Database as a

Resource for Proteomics Research: 2014 Update. Nucleic Acids Res. 2014, 42 (D1), D959–D965. https://doi.org/10.1093/nar/gkt1251.

(53) Goodacre, R. The Blind Men and the Elephant: Challenges in the Analysis of Complex Natural Mixtures. Faraday Discuss. 2019, 218 (0), 524–539. https://doi.org/10.1039/C9FD00074G.

(54) Wasinger, V. C.; Cordwell, S. J.; Cerpa-Poljak, A.; Yan, J. X.; Gooley, A. A.; Wilkins, M. R.; Duncan, M. W.; Harris, R.; Williams, K. L.; Humphery-Smith, I. Progress with Gene-Product Mapping of the Mollicutes: Mycoplasma Genitalium. ELECTROPHORESIS 1995, 16 (1), 1090–1094. https://doi.org/10.1002/elps.11501601185.

(55) Bell, J. The Polymerase Chain Reaction. Immunol. Today 1989, 10 (10), 351–355. https://doi.org/10.1016/0167-5699(89)90193-X.

(56) Yates, J. R. I. The Revolution and Evolution of Shotgun Proteomics for Large-Scale Proteome Analysis. J. Am. Chem. Soc. 2013, 135 (5), 1629–1640. https://doi.org/10.1021/ja3094313.

(57) Sanger, F.; Tuppy, H. The Amino-Acid Sequence in the Phenylalanyl Chain of Insulin. 1. The Identification of Lower Peptides from Partial Hydrolysates. Biochem. J. 1951, 49 (4), 463–481. https://doi.org/10.1042/bj0490463.

(58) Sanger, F.; Tuppy, H. The Amino-Acid Sequence in the Phenylalanyl Chain of Insulin. 2. The Investigation of Peptides from Enzymic Hydrolysates. Biochem. J. 1951, 49 (4), 481–490. https://doi.org/10.1042/bj0490481.

(59) Edman, P. A Method for the Determination of the Amino Acid Sequence in Peptides. Arch Biochem 1949, 22 (3), 475–476.

(60) Edman, P. A Method for the Determination of the Amino Acid Sequence in Peptides. Acta Chem Scand 1950, 4 (2), 283–293.

(61) Edman, P.; Begg, G. A Protein Sequenator. Eur. J. Biochem. 1967, 1 (1), 80–91. https://doi.org/10.1111/j.1432-1033.1967.tb00047.x.

(62) [36] Automated Edman Degradation: The Protein Sequenator. In Methods in Enzymology; Academic Press, 1973; Vol. 27, pp 942–1010. https://doi.org/10.1016/S0076-6879(73)27039-8.

(63) Yost, R. A.; Enke, C. G. Triple Quadrupole Mass Spectrometry for Direct Mixture Analysis and Structure Elucidation. Anal. Chem. 1979, 51 (12), 1251–1264. https://doi.org/10.1021/ac50048a002.

(64) Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R. Protein Sequencing by Tandem Mass Spectrometry. Proc. Natl. Acad. Sci. 1986, 83 (17), 6233–6237. https://doi.org/10.1073/pnas.83.17.6233.

(65) Hunt, D. F.; Henderson, R. A.; Shabanowitz, J.; Sakaguchi, K.; Michel, H.; Sevilir, N.; Cox, A. L.; Appella, E.; Engelhard, V. H. Characterization of Peptides Bound to the Class I MHC Molecule HLA-A2.1 by Mass Spectrometry. Science 1992, 255 (5049), 1261–1263.

(66) Arnott, D.; Shabanowitz, J.; Hunt, D. F. Mass Spectrometry of Proteins and Peptides: Sensitive and Accurate Mass Measurement and Sequence Analysis. Clin. Chem. 1993, 39 (9), 2005–2010. https://doi.org/10.1093/clinchem/39.9.2005.

(67) Kaltschmidt, E.; Wittmann, H. G. Ribosomal Proteins. VII: Two-Dimensional Polyacrylamide Gel Electrophoresis for Fingerprinting of Ribosomal Proteins. Anal. Biochem. 1970, 36 (2), 401–412. https://doi.org/10.1016/0003-2697(70)90376-3.

(68) O'Farrell, P. High Resolution Two-Dimensional Electrophoresis of Proteins. J. Biol. Chem. 1975, 250 (10), 4007–4021. https://doi.org/10.1016/S0021-9258(19)41496-8.

(69) O'Farrell, P. H. The Pre-Omics Era: The Early Days of Two-Dimensional Gels. PROTEOMICS 2008, 8 (23–24), 4842–4852. https://doi.org/10.1002/pmic.200800719.

(70) Wilm, M.; Shevchenko, A.; Houthaeve, T.; Breit, S.; Schweigerer, L.; Fotsis, T.; Mann, M. Femtomole Sequencing of Proteins from Polyacrylamide Gels by Nano-Electrospray Mass Spectrometry. Nature 1996, 379 (6564), 466–469. https://doi.org/10.1038/379466a0.

(71) Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. Anal. Chem. 1996, 68 (5), 850–858. https://doi.org/10.1021/ac950914h.

(72) Shevchenko, A.; Tomas, H.; Havli, J.; Olsen, J. V.; Mann, M. In-Gel Digestion for Mass Spectrometric Characterization of Proteins and Proteomes. Nat. Protoc. 2006, 1 (6), 2856–2860. https://doi.org/10.1038/nprot.2006.468.

(73) Shevchenko, A.; Jensen, O. N.; Podtelejnikov, A. V.; Sagliocco, F.; Wilm, M.; Vorm, O.; Mortensen, P.; Shevchenko, A.; Boucherie, H.; Mann, M. Linking Genome and Proteome by Mass Spectrometry: Large-Scale Identification of Yeast Proteins from Two Dimensional Gels. Proc. Natl. Acad. Sci. 1996, 93 (25), 14440–14445. https://doi.org/10.1073/pnas.93.25.14440.

(74) Link, A. J.; Hays, L. G.; Carmack, E. B.; Yates III, J. R. Identifying the Major Proteome Components of Haemophilus Influenzae Type-Strain NCTC 8143. ELECTROPHORESIS 1997, 18 (8), 1314–1334. https://doi.org/10.1002/elps.1150180808.

(75) Aebersold, R.; Mann, M. Mass Spectrometry-Based Proteomics. Nature 2003, 422 (6928), 198–207. https://doi.org/10.1038/nature01511.

(76) Gygi, S. P.; Corthals, G. L.; Zhang, Y.; Rochon, Y.; Aebersold, R. Evaluation of Two-Dimensional Gel Electrophoresis-Based Proteome Analysis Technology. Proc. Natl. Acad. Sci. 2000, 97 (17), 9390–9395. https://doi.org/10.1073/pnas.160270797.

(77) Wolters, D. A.; Washburn, M. P.; Yates, J. R. An Automated Multidimensional Protein Identification Technology for Shotgun Proteomics. Anal. Chem. 2001, 73 (23), 5683–5690. https://doi.org/10.1021/ac010617e.

(78) Tong, W.; Link, A.; Eng, J. K.; Yates, J. R. Identification of Proteins in Complexes by Solid-Phase Microextraction/Multistep Elution/Capillary Electrophoresis/Tandem Mass Spectrometry. Anal. Chem. 1999, 71 (13), 2270–2278. https://doi.org/10.1021/ac9901182.

(79) Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R. Direct Analysis of Protein Complexes Using Mass Spectrometry. Nat. Biotechnol. 1999, 17 (7), 676–682. https://doi.org/10.1038/10890.

(80) Zubarev, R. A. The Challenge of the Proteome Dynamic Range and Its Implications for In-Depth Proteomics. PROTEOMICS 2013, 13 (5), 723–726. https://doi.org/10.1002/pmic.201200451.

(81) Lippert, J. A.; Xin, B.; Wu, N.; Lee, M. L. Fast Ultrahigh-Pressure Liquid Chromatography: On-Column UV and Time-of-Flight Mass Spectrometric Detection. J. Microcolumn Sep. 1999, 11 (9), 631–643. https://doi.org/10.1002/(SICI)1520-667X(199911)11:9<631::AID-MCS1>3.0.CO;2-I.

(82) Wu, N.; Lippert, J. A.; Lee, M. L. Practical Aspects of Ultrahigh Pressure Capillary Liquid Chromatography. J. Chromatogr. A 2001, 911 (1), 1–12. https://doi.org/10.1016/S0021-9673(00)01188-2.

(83) Liu, H.; Finch, J. W.; Lavallee, M. J.; Collamati, R. A.; Benevides, C. C.; Gebler, J. C. Effects of Column Length, Particle Size, Gradient Length and Flow Rate on Peak Capacity of Nano-Scale Liquid Chromatography for Peptide Separations. J. Chromatogr. A 2007, 1147 (1), 30–36. https://doi.org/10.1016/j.chroma.2007.02.016.

(84) Nováková, L.; Vlčková, H. A Review of Current Trends and Advances in Modern Bio-Analytical Methods: Chromatography and Sample Preparation. Anal. Chim. Acta 2009, 656 (1), 8–35. https://doi.org/10.1016/j.aca.2009.10.004.

(85) Shen, Y.; Jacobs, J. M.; Camp, D. G.; Fang, R.; Moore, R. J.; Smith, R. D.; Xiao, W.; Davis, R. W.; Tompkins, R. G. Ultra-High-Efficiency Strong Cation Exchange LC/RPLC/MS/MS

for High Dynamic Range Characterization of the Human Plasma Proteome. Anal. Chem. 2004, 76 (4), 1134–1144. https://doi.org/10.1021/ac034869m.

(86) Wang, Y.; Yang, F.; Gritsenko, M. A.; Wang, Y.; Clauss, T.; Liu, T.; Shen, Y.; Monroe, M. E.; Lopez-Ferrer, D.; Reno, T.; Moore, R. J.; Klemke, R. L.; Camp II, D. G.; Smith, R. D. Reversed-Phase Chromatography with Multiple Fraction Concatenation Strategy for Proteome Profiling of Human MCF10A Cells. PROTEOMICS 2011, 11 (10), 2019–2026. https://doi.org/10.1002/pmic.201000722.

(87) Manadas, B.; Mendes, V. M.; English, J.; Dunn, M. J. Peptide Fractionation in Proteomics Approaches. Expert Rev. Proteomics 2010, 7 (5), 655–663. https://doi.org/10.1586/epr.10.46.

(88) Geiger, T.; Wehner, A.; Schaab, C.; Cox, J.; Mann, M. Comparative Proteomic Analysis of Eleven Common Cell Lines Reveals Ubiquitous but Varying Expression of Most Proteins *. Mol. Cell. Proteomics 2012, 11 (3). https://doi.org/10.1074/mcp.M111.014050.

(89) Wiśniewski, J. R.; Zougman, A.; Mann, M. Combination of FASP and StageTip-Based Fractionation Allows In-Depth Analysis of the Hippocampal Membrane Proteome. J. Proteome Res. 2009, 8 (12), 5674–5678. https://doi.org/10.1021/pr900748n.

(90) Huttlin, E. L.; Jedrychowski, M. P.; Elias, J. E.; Goswami, T.; Rad, R.; Beausoleil, S. A.; Villén, J.; Haas, W.; Sowa, M. E.; Gygi, S. P. A Tissue-Specific Atlas of Mouse Protein Phosphorylation and Expression. Cell 2010, 143 (7), 1174–1189. https://doi.org/10.1016/j.cell.2010.12.001.

(91) Wiśniewski, J. R.; Ostasiewicz, P.; Duś, K.; Zielińska, D. F.; Gnad, F.; Mann, M. Extensive Quantitative Remodeling of the Proteome between Normal Colon Tissue and Adenocarcinoma. Mol. Syst. Biol. 2012, 8 (1), MSB201244. https://doi.org/10.1038/msb.2012.44.

(92) Beck, M.; Schmidt, A.; Malmstroem, J.; Claassen, M.; Ori, A.; Szymborska, A.; Herzog, F.; Rinner, O.; Ellenberg, J.; Aebersold, R. The Quantitative Proteome of a Human Cell Line. Mol. Syst. Biol. 2011, 7 (1), MSB201182. https://doi.org/10.1038/msb.2011.82.

(93) Nagaraj, N.; Wisniewski, J. R.; Geiger, T.; Cox, J.; Kircher, M.; Kelso, J.; Pääbo, S.; Mann, M. Deep Proteome and Transcriptome Mapping of a Human Cancer Cell Line. Mol. Syst. Biol. 2011, 7 (1), MSB201181. https://doi.org/10.1038/msb.2011.81.

(94) Fang, Y.; Robinson, D. P.; Foster, L. J. Quantitative Analysis of Proteome Coverage and Recovery Rates for Upstream Fractionation Methods in Proteomics. J. Proteome Res. 2010, 9 (4), 1902–1912. https://doi.org/10.1021/pr901063t.

(95) Manadas, B.; English, J. A.; Wynne, K. J.; Cotter, D. R.; Dunn, M. J. Comparative Analysis of OFFGel, Strong Cation Exchange with pH Gradient, and RP at High pH for First-Dimensional Separation of Peptides from a Membrane-Enriched Protein Fraction. PROTEOMICS 2009, 9 (22), 5194–5198. https://doi.org/10.1002/pmic.200900349.

(96) Xie, H.; Onsongo, G.; Popko, J.; de Jong, E. P.; Cao, J.; Carlis, J. V.; Griffin, R. J.; Rhodus, N. L.; Griffin, T. J. Proteomics Analysis of Cells in Whole Saliva from Oral Cancer Patients via Value-Added Three-Dimensional Peptide Fractionation and Tandem Mass Spectrometry*. Mol. Cell. Proteomics 2008, 7 (3), 486–498. https://doi.org/10.1074/mcp.M700146-MCP200.

(97) Nakamura, T.; Kuromitsu, J.; Oda, Y. Evaluation of Comprehensive Multidimensional Separations Using Reversed-Phase, Reversed-Phase Liquid Chromatography/Mass Spectrometry for Shotgun Proteomics. J. Proteome Res. 2008, 7 (3), 1007–1011. https://doi.org/10.1021/pr7005878.

(98) Han, J.; Ye, L.; Xu, L.; Zhou, Z.; Gao, F.; Xiao, Z.; Wang, Q.; Zhang, B. Towards High Peak Capacity Separations in Normal Pressure Nanoflow Liquid Chromatography Using Meter Long Packed Capillary Columns. Anal. Chim. Acta 2014, 852, 267–273. https://doi.org/10.1016/j.aca.2014.09.006.

(99) Thakur, S. S.; Geiger, T.; Chatterjee, B.; Bandilla, P.; Fröhlich, F.; Cox, J.; Mann, M. Deep and Highly Sensitive Proteome Coverage by LC-MS/MS Without Prefractionation*. Mol. Cell. Proteomics 2011, 10 (8), M110.003699. https://doi.org/10.1074/mcp.M110.003699.

(100) Nagaraj, N.; Alexander Kulak, N.; Cox, J.; Neuhauser, N.; Mayr, K.; Hoerning, O.; Vorm, O.; Mann, M. System-Wide Perturbation Analysis with Nearly Complete Coverage of the Yeast Proteome by Single-Shot Ultra HPLC Runs on a Bench Top Orbitrap*. Mol. Cell. Proteomics 2012, 11 (3), M111.013722. https://doi.org/10.1074/mcp.M111.013722.

(101) Köcher, T.; Swart, R.; Mechtler, K. Ultra-High-Pressure RPLC Hyphenated to an LTQ-Orbitrap Velos Reveals a Linear Relation between Peak Capacity and Number of Identified Peptides. Anal. Chem. 2011, 83 (7), 2699–2704. https://doi.org/10.1021/ac103243t.

(102) Yamana, R.; Iwasaki, M.; Wakabayashi, M.; Nakagawa, M.; Yamanaka, S.; Ishihama, Y. Rapid and Deep Profiling of Human Induced Pluripotent Stem Cell Proteome by One-Shot NanoLC–MS/MS Analysis with Meter-Scale Monolithic Silica Columns. J. Proteome Res. 2013, 12 (1), 214–221. https://doi.org/10.1021/pr300837u.

(103) Michalski, A.; Cox, J.; Mann, M. More than 100,000 Detectable Peptide Species Elute in Single Shotgun Proteomics Runs but the Majority Is Inaccessible to Data-Dependent LC−MS/MS. J. Proteome Res. 2011, 10 (4), 1785–1793. https://doi.org/10.1021/pr101060v.

(104) Luo, Q.; Tang, K.; Yang, F.; Elias, A.; Shen, Y.; Moore, R. J.; Zhao, R.; Hixson, K. K.; Rossie, S. S.; Smith, R. D. More Sensitive and Quantitative Proteomic Measurements Using Very Low Flow Rate Porous Silica Monolithic LC Columns with Electrospray Ionization-Mass Spectrometry. J. Proteome Res. 2006, 5 (5), 1091–1097. https://doi.org/10.1021/pr050424y.

(105) Cech, N. B.; Enke, C. G. Practical Implications of Some Recent Studies in Electrospray Ionization Fundamentals. Mass Spectrom. Rev. 2001, 20 (6), 362–387. https://doi.org/10.1002/mas.10008.

(106) Tang, K.; Page, J. S.; Smith, R. D. Charge Competition and the Linear Dynamic Range of Detection in Electrospray Ionization Mass Spectrometry. J. Am. Soc. Mass Spectrom. 2004, 15 (10), 1416–1423. https://doi.org/10.1016/j.jasms.2004.04.034.

(107) King, R.; Bonfiglio, R.; Fernandez-Metzler, C.; Miller-Stein, C.; Olah, T. Mechanistic Investigation of Ionization Suppression in Electrospray Ionization. J. Am. Soc. Mass Spectrom. 2000, 11 (11), 942–950. https://doi.org/10.1016/S1044-0305(00)00163-X.

(108) Meier, F.; Geyer, P. E.; Virreira Winter, S.; Cox, J.; Mann, M. BoxCar Acquisition Method Enables Single-Shot Proteomics at a Depth of 10,000 Proteins in 100 Minutes. Nat. Methods 2018, 15 (6), 440–448. https://doi.org/10.1038/s41592-018-0003-5.

(109) Phlairaharn, T.; Shannon, A. E.; Zeng, X.; Truong, D.-J. J.; Schoof, E. M.; Ye, Z.; Searle, B. C. Improving Proteomic Dynamic Range with Multiple Accumulation Precursor Mass Spectrometry. J. Proteome Res. 2025, 24 (10), 5116–5126. https://doi.org/10.1021/acs.jproteome.5c00469.

(110) Kaufmann, A.; Walker, S. Extension of the Q Orbitrap Intrascan Dynamic Range by Using a Dedicated Customized Scan. Rapid Commun. Mass Spectrom. 2016, 30 (8), 1087–1095. https://doi.org/10.1002/rcm.7530.

(111) Zhang, Z.; Chan, P. K.; Richardson, J.; Shah, B. An Evaluation of Instrument Types for Mass Spectrometry-Based Multi-Attribute Analysis of Biotherapeutics. mAbs 2020, 12 (1), 1783062. https://doi.org/10.1080/19420862.2020.1783062.

(112) Messner, C. B.; Demichev, V.; Wang, Z.; Hartl, J.; Kustatscher, G.; Mülleder, M.; Ralser, M. Mass Spectrometry-Based High-Throughput Proteomics and Its Role in Biomedical Studies and Systems Biology. PROTEOMICS 2023, 23 (7–8), 2200013. https://doi.org/10.1002/pmic.202200013.

(113) Giansanti, P.; Samaras, P.; Bian, Y.; Meng, C.; Coluccio, A.; Frejno, M.; Jakubowsky, H.; Dobiasch, S.; Hazarika, R. R.; Rechenberger, J.; Calzada-Wack, J.; Krumm, J.; Mueller, S.; Lee, C.-Y.; Wimberger, N.; Lautenbacher, L.; Hassan, Z.; Chang, Y.-C.; Falcomatà, C.; Bayer, F. P.; Bärthel, S.; Schmidt, T.; Rad, R.; Combs, S. E.; The, M.; Johannes, F.; Saur, D.;

de Angelis, M. H.; Wilhelm, M.; Schneider, G.; Kuster, B. Mass Spectrometry-Based Draft of the Mouse Proteome. Nat. Methods 2022, 19 (7), 803–811. https://doi.org/10.1038/s41592-022-01526-y.

(114) Liu, H.; Sadygov, R. G.; Yates, J. R. A Model for Random Sampling and Estimation of Relative Protein Abundance in Shotgun Proteomics. Anal. Chem. 2004, 76 (14), 4193–4201. https://doi.org/10.1021/ac0498563.

(115) Gillet, L. C.; Navarro, P.; Tate, S.; Röst, H.; Selevsek, N.; Reiter, L.; Bonner, R.; Aebersold, R. Targeted Data Extraction of the MS/MS Spectra Generated by Data-Independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis*. Mol. Cell. Proteomics 2012, 11 (6), O111.016717. https://doi.org/10.1074/mcp.O111.016717.

(116) Ludwig, C.; Gillet, L.; Rosenberger, G.; Amon, S.; Collins, B. C.; Aebersold, R. Data-independent Acquisition-based SWATH-MS for Quantitative Proteomics: A Tutorial. Mol. Syst. Biol. 2018, 14 (8), MSB178126. https://doi.org/10.15252/msb.20178126.

(117) Ting, Y. S.; Egertson, J. D.; Payne, S. H.; Kim, S.; MacLean, B.; Käll, L.; Aebersold, R.; Smith, R. D.; Noble, W. S.; MacCoss, M. J. Peptide-Centric Proteome Analysis: An Alternative Strategy for the Analysis of Tandem Mass Spectrometry Data*. Mol. Cell. Proteomics 2015, 14 (9), 2301–2307. https://doi.org/10.1074/mcp.O114.047035.

(118) Egertson, J. D.; Kuehn, A.; Merrihew, G. E.; Bateman, N. W.; MacLean, B. X.; Ting, Y. S.; Canterbury, J. D.; Marsh, D. M.; Kellmann, M.; Zabrouskov, V.; Wu, C. C.; MacCoss, M. J. Multiplexed MS/MS for Improved Data-Independent Acquisition. Nat. Methods 2013, 10 (8), 744–746. https://doi.org/10.1038/nmeth.2528.

(119) Bruderer, R.; Bernhardt, O. M.; Gandhi, T.; Xuan, Y.; Sondermann, J.; Schmidt, M.; Gomez-Varela, D.; Reiter, L. Optimization of Experimental Parameters in Data-Independent Mass Spectrometry Significantly Increases Depth and Reproducibility of Results*. Mol. Cell. Proteomics 2017, 16 (12), 2296–2309. https://doi.org/10.1074/mcp.RA117.000314.

(120) Muntel, J.; Gandhi, T.; Verbeke, L.; Bernhardt, O. M.; Treiber, T.; Bruderer, R.; Reiter, L. Surpassing 10 000 Identified and Quantified Proteins in a Single Run by Optimizing Current LC-MS Instrumentation and Data Analysis Strategy. Mol. Omics 2019, 15 (5), 348–360. https://doi.org/10.1039/C9MO00082H.

(121) Canterbury, J. D.; Merrihew, G. E.; MacCoss, M. J.; Goodlett, D. R.; Shaffer, S. A. Comparison of Data Acquisition Strategies on Quadrupole Ion Trap Instrumentation for Shotgun Proteomics. J. Am. Soc. Mass Spectrom. 2014, 25 (12), 2048–2059. https://doi.org/10.1007/s13361-014-0981-1.

(122) Hossain, M.; Kaleta, D. T.; Robinson, E. W.; Liu, T.; Zhao, R.; Page, J. S.; Kelly, R. T.; Moore, R. J.; Tang, K.; Camp, D. G.; Qian, W.-J.; Smith, R. D. Enhanced Sensitivity for Selected Reaction Monitoring Mass Spectrometry-Based Targeted Proteomics Using a Dual Stage Electrodynamic Ion Funnel Interface*. Mol. Cell. Proteomics 2011, 10 (2), S1–S9. https://doi.org/10.1074/mcp.M000062-MCP201.

(123) Kelly, R. T. Single-Cell Proteomics: Progress and Prospects. Mol. Cell. Proteomics 2020, 19 (11), 1739–1748. https://doi.org/10.1074/mcp.R120.002234.

(124) Loboda, A. V.; Chernushevich, I. V. A Novel Ion Trap That Enables High Duty Cycle and Wide m/z Range on an Orthogonal Injection TOF Mass Spectrometer. J. Am. Soc. Mass Spectrom. 2009, 20 (7), 1342–1348. https://doi.org/10.1016/j.jasms.2009.03.018.

(125) Wang, Z.; Mülleder, M.; Batruch, I.; Chelur, A.; Textoris-Taube, K.; Schwecke, T.; Hartl, J.; Causon, J.; Castro-Perez, J.; Demichev, V.; Tate, S.; Ralser, M. High-Throughput Proteomics of Nanogram-Scale Samples with Zeno SWATH MS. eLife 2022, 11, e83947. https://doi.org/10.7554/eLife.83947.

(126) Panchaud, A.; Scherl, A.; Shaffer, S. A.; von Haller, P. D.; Kulasekara, H. D.; Miller, S. I.; Goodlett, D. R. Precursor Acquisition Independent From Ion Count: How to Dive Deeper into the Proteomics Ocean. Anal. Chem. 2009, 81 (15), 6481–6488. https://doi.org/10.1021/ac900888s.

(127) Zhang, Y.; Bilbao, A.; Bruderer, T.; Luban, J.; Strambio-De-Castillia, C.; Lisacek, F.; Hopfgartner, G.; Varesio, E. The Use of Variable Q1 Isolation Windows Improves Selectivity in LC–SWATH–MS Acquisition. J. Proteome Res. 2015, 14 (10), 4359–4371. https://doi.org/10.1021/acs.jproteome.5b00543.

(128) Amodei, D.; Egertson, J.; MacLean, B. X.; Johnson, R.; Merrihew, G. E.; Keller, A.; Marsh, D.; Vitek, O.; Mallick, P.; MacCoss, M. J. Improving Precursor Selectivity in Data-Independent Acquisition Using Overlapping Windows. J. Am. Soc. Mass Spectrom. 2019, 30 (4), 669–684. https://doi.org/10.1007/s13361-018-2122-8.

(129) Heil, L. R.; Fondrie, W. E.; McGann, C. D.; Federation, A. J.; Noble, W. S.; MacCoss, M. J.; Keich, U. Building Spectral Libraries from Narrow-Window Data-Independent Acquisition Mass Spectrometry Data. J. Proteome Res. 2022, 21 (6), 1382–1391. https://doi.org/10.1021/acs.jproteome.1c00895.

(130) Heil, L. R.; Remes, P. M.; Canterbury, J. D.; Yip, P.; Barshop, W. D.; Wu, C. C.; MacCoss, M. J. Dynamic Data-Independent Acquisition Mass Spectrometry with Real-Time Retrospective Alignment. Anal. Chem. 2023, 95 (32), 11854–11858. https://doi.org/10.1021/acs.analchem.3c00903.

(131) Pino, L. K.; Just, S. C.; MacCoss, M. J.; Searle, B. C. Acquiring and Analyzing Data Independent Acquisition Proteomics Experiments without Spectrum Libraries. Mol. Cell. Proteomics 2020, 19 (7), 1088–1103. https://doi.org/10.1074/mcp.P119.001913.

(132) Fröhlich, K.; Furrer, R.; Schori, C.; Handschin, C.; Schmidt, A. Robust, Precise, and Deep Proteome Profiling Using a Small Mass Range and Narrow Window Data-Independent-Acquisition Scheme. J. Proteome Res. 2024, 23 (3), 1028–1038. https://doi.org/10.1021/acs.jproteome.3c00736.

(133) Meier, F.; Brunner, A.-D.; Frank, M.; Ha, A.; Bludau, I.; Voytik, E.; Kaspar-Schoenefeld, S.; Lubeck, M.; Raether, O.; Bache, N.; Aebersold, R.; Collins, B. C.; Röst, H. L.; Mann, M. diaPASEF: Parallel Accumulation–Serial Fragmentation Combined with Data-Independent Acquisition. Nat. Methods 2020, 17 (12), 1229–1236. https://doi.org/10.1038/s41592-020-00998-0.

(134) Skowronek, P.; Thielert, M.; Voytik, E.; Tanzer, M. C.; Hansen, F. M.; Willems, S.; Karayel, O.; Brunner, A.-D.; Meier, F.; Mann, M. Rapid and In-Depth Coverage of the (Phospho-)Proteome With Deep Libraries and Optimal Window Design for Dia-PASEF. Mol. Cell. Proteomics 2022, 21 (9). https://doi.org/10.1016/j.mcpro.2022.100279.

(135) Modern Proteomics – Sample Preparation, Analysis and Practical Applications; Mirzaei, H., Carrasco, M., Eds.; Advances in Experimental Medicine and Biology; Springer International Publishing: Cham, 2016; Vol. 919. https://doi.org/10.1007/978-3-319-41448-5.

(136) Han, M.-J.; Lee, S. Y. The Escherichia Coli Proteome: Past, Present, and Future Prospects. Microbiol. Mol. Biol. Rev. 2006, 70 (2), 362–439. https://doi.org/10.1128/mmbr.00036-05.

(137) Washburn, M. P.; Wolters, D.; Yates, J. R. Large-Scale Analysis of the Yeast Proteome by Multidimensional Protein Identification Technology. Nat. Biotechnol. 2001, 19 (3), 242–247. https://doi.org/10.1038/85686.

(138) Lundberg, E.; Fagerberg, L.; Klevebring, D.; Matic, I.; Geiger, T.; Cox, J.; Älgenäs, C.; Lundeberg, J.; Mann, M.; Uhlen, M. Defining the Transcriptome and Proteome in Three Functionally Different Human Cell Lines. Mol. Syst. Biol. 2010, 6 (1), MSB2010106. https://doi.org/10.1038/msb.2010.106.

(139) Shishkova, E.; Coon, J. J. Rapid Preparation of Human Blood Plasma for Bottom-up Proteomics Analysis. STAR Protoc. 2021, 2 (4), 100856. https://doi.org/10.1016/j.xpro.2021.100856.

(140) Schutzer, S. E.; Liu, T.; Natelson, B. H.; Angel, T. E.; Schepmoes, A. A.; Purvine, S. O.; Hixson, K. K.; Lipton, M. S.; Ii, D. G. C.; Coyle, P. K.; Smith, R. D.; Bergquist, J. Establishing the Proteome of Normal Human Cerebrospinal Fluid. PLOS ONE 2010, 5 (6), e10980. https://doi.org/10.1371/journal.pone.0010980.

(141) Dzieciatkowska, M.; D'Alessandro, A.; Moore, E. E.; Wohlauer, M.; Banerjee, A.; Silliman, C. C.; Hansen, K. C. Lymph Is Not a Plasma Ultrafiltrate: A Proteomic Analysis of Injured Patients. Shock 2014, 42 (6), 485. https://doi.org/10.1097/SHK.0000000000000249.

(142) Hu, S.; Loo, J. A.; Wong, D. T. Human Saliva Proteome Analysis. Ann. N. Y. Acad. Sci. 2007, 1098 (1), 323–329. https://doi.org/10.1196/annals.1384.015.

(143) Beasley-Green, A. Urine Proteomics in the Era of Mass Spectrometry. Int. Neurourol. J. 2016, 20 (Suppl 2), S70-75. https://doi.org/10.5213/inj.1612720.360.

(144) Cappellini, E.; Jensen, L. J.; Szklarczyk, D.; Ginolhac, A.; da Fonseca, R. A. R.; Stafford, T. W. Jr.; Holen, S. R.; Collins, M. J.; Orlando, L.; Willerslev, E.; Gilbert, M. T. P.; Olsen, J. V. Proteomic Analysis of a Pleistocene Mammoth Femur Reveals More than One Hundred Ancient Bone Proteins. J. Proteome Res. 2012, 11 (2), 917–926. https://doi.org/10.1021/pr200721u.

(145) Prus, G.; Hoegl, A.; Weinert, B. T.; Choudhary, C. Analysis and Interpretation of Protein Post-Translational Modification Site Stoichiometry. Trends Biochem. Sci. 2019, 44 (11), 943–960. https://doi.org/10.1016/j.tibs.2019.06.003.

(146) Niemi, N. M.; Pagliarini, D. J. The Extensive and Functionally Uncharacterized Mitochondrial Phosphoproteome. J. Biol. Chem. 2021, 297 (1), 100880. https://doi.org/10.1016/j.jbc.2021.100880.

(147) Low, T. Y.; Mohtar, M. A.; Lee, P. Y.; Omar, N.; Zhou, H.; Ye, M. Widening the Bottleneck of Phosphoproteomics: Evolving Strategies for Phosphopeptide Enrichment. Mass Spectrom. Rev. 2021, 40 (4), 309–333. https://doi.org/10.1002/mas.21636.

(148) Muneer, G.; Chen, C.-S.; Chen, Y.-J. Advancements in Global Phosphoproteomics Profiling: Overcoming Challenges in Sensitivity and Quantification. PROTEOMICS 2025, 25 (1–2), e202400087. https://doi.org/10.1002/pmic.202400087.

(149) Sinitcyn, P.; Rudolph, J. D.; Cox, J. Computational Methods for Understanding Mass Spectrometry–Based Shotgun Proteomics Data. Annu. Rev. Biomed. Data Sci. 2018, 1 (Volume 1, 2018), 207–234. https://doi.org/10.1146/annurev-biodatasci-080917-013516.

(150) Chen, C.; Hou, J.; Tanner, J. J.; Cheng, J. Bioinformatics Methods for Mass Spectrometry-Based Proteomics Data Analysis. Int. J. Mol. Sci. 2020, 21 (8), 2873. https://doi.org/10.3390/ijms21082873.

(151) Emmett, M. R.; Caprioli, R. M. Micro-Electrospray Mass Spectrometry: Ultra-High-Sensitivity Analysis of Peptides and Proteins. J. Am. Soc. Mass Spectrom. 1994, 5 (7), 605–613. https://doi.org/10.1016/1044-0305(94)85001-1.

(152) Marginean, I.; Kelly, R. T.; Prior, D. C.; LaMarche, B. L.; Tang, K.; Smith, R. D. Analytical Characterization of the Electrospray Ion Source in the Nanoflow Regime. Anal. Chem. 2008, 80 (17), 6573–6579. https://doi.org/10.1021/ac800683s.

(153) Perry, R. H.; Cooks, R. G.; Noll, R. J. Orbitrap Mass Spectrometry: Instrumentation, Ion Motion and Applications. Mass Spectrom. Rev. 2008, 27 (6), 661–699. https://doi.org/10.1002/mas.20186.

(154) Li, C.; Chu, S.; Tan, S.; Yin, X.; Jiang, Y.; Dai, X.; Gong, X.; Fang, X.; Tian, D. Towards Higher Sensitivity of Mass Spectrometry: A Perspective From the Mass Analyzers. Front. Chem. 2021, 9. https://doi.org/10.3389/fchem.2021.813359.

(155) March, R. E. Quadrupole Ion Traps. Mass Spectrom. Rev. 2009, 28 (6), 961–989. https://doi.org/10.1002/mas.20250.

(156) Guilhaus, M.; Selby, D.; Mlynski, V. Orthogonal Acceleration Time-of-Flight Mass Spectrometry. Mass Spectrom. Rev. 2000, 19 (2), 65–107. https://doi.org/10.1002/(SICI)1098-2787(2000)19:2<65::AID-MAS1>3.0.CO;2-E.

(157) Brais, C. J.; Ibañez, J. O.; Schwartz, A. J.; Ray, S. J. Recent Advances in Instrumental Approaches to Time-of-Flight Mass Spectrometry. Mass Spectrom. Rev. 2021, 40 (5), 647–669. https://doi.org/10.1002/mas.21650.

(158) Gessulat, S.; Schmidt, T.; Zolg, D. P.; Samaras, P.; Schnatbaum, K.; Zerweck, J.; Knaute, T.; Rechenberger, J.; Delanghe, B.; Huhmer, A.; Reimer, U.; Ehrlich, H.-C.; Aiche, S.; Kuster, B.; Wilhelm, M. Prosit: Proteome-Wide Prediction of Peptide Tandem Mass Spectra by Deep Learning. Nat. Methods 2019, 16 (6), 509–518. https://doi.org/10.1038/s41592-019-0426-7.

(159) Demichev, V.; Messner, C. B.; Vernardis, S. I.; Lilley, K. S.; Ralser, M. DIA-NN: Neural Networks and Interference Correction Enable Deep Proteome Coverage in High Throughput. Nat. Methods 2020, 17 (1), 41–44. https://doi.org/10.1038/s41592-019-0638-x.

(160) Tiwary, S.; Levy, R.; Gutenbrunner, P.; Salinas Soto, F.; Palaniappan, K. K.; Deming, L.; Berndl, M.; Brant, A.; Cimermancic, P.; Cox, J. High-Quality MS/MS Spectrum Prediction for Data-Dependent and Data-Independent Acquisition Data Analysis. Nat. Methods 2019, 16 (6), 519–525. https://doi.org/10.1038/s41592-019-0427-6.

(161) Sinitcyn, P.; Hamzeiy, H.; Salinas Soto, F.; Itzhak, D.; McCarthy, F.; Wichmann, C.; Steger, M.; Ohmayer, U.; Distler, U.; Kaspar-Schoenefeld, S.; Prianichnikov, N.; Yılmaz, Ş.; Rudolph, J. D.; Tenzer, S.; Perez-Riverol, Y.; Nagaraj, N.; Humphrey, S. J.; Cox, J. MaxDIA Enables Library-Based and Library-Free Data-Independent Acquisition Proteomics. Nat. Biotechnol. 2021, 39 (12), 1563–1573. https://doi.org/10.1038/s41587-021-00968-7.

(162) Zeng, W.-F.; Zhou, X.-X.; Willems, S.; Ammar, C.; Wahle, M.; Bludau, I.; Voytik, E.; Strauss, M. T.; Mann, M. AlphaPeptDeep: A Modular Deep Learning Framework to Predict Peptide Properties for Proteomics. Nat. Commun. 2022, 13 (1), 7238. https://doi.org/10.1038/s41467-022-34904-3.

(163) Elias, J. E.; Gygi, S. P. Target-Decoy Search Strategy for Increased Confidence in Large-Scale Protein Identifications by Mass Spectrometry. Nat. Methods 2007, 4 (3), 207–214. https://doi.org/10.1038/nmeth1019.

(164) M. Miller, R.; M. Smith, L. Overview and Considerations in Bottom-up Proteomics. Analyst 2023, 148 (3), 475–486. https://doi.org/10.1039/D2AN01246D.

(165) Chen, B.; Brown, K. A.; Lin, Z.; Ge, Y. Top-Down Proteomics: Ready for Prime Time? Anal. Chem. 2018, 90 (1), 110–127. https://doi.org/10.1021/acs.analchem.7b04747.

(166) Coscia, F.; Lengyel, E.; Duraiswamy, J.; Ashcroft, B.; Bassani-Sternberg, M.; Wierer, M.; Johnson, A.; Wroblewski, K.; Montag, A.; Yamada, S. D.; López-Méndez, B.; Nilsson, J.; Mund, A.; Mann, M.; Curtis, M. Multi-Level Proteomics Identifies CT45 as a Chemosensitivity Mediator and Immunotherapy Target in Ovarian Cancer. Cell 2018, 175 (1), 159-170.e16. https://doi.org/10.1016/j.cell.2018.08.065.

(167) Bar-Peled, L.; Kemper, E. K.; Suciu, R. M.; Vinogradova, E. V.; Backus, K. M.; Horning, B. D.; Paul, T. A.; Ichu, T.-A.; Svensson, R. U.; Olucha, J.; Chang, M. W.; Kok, B. P.; Zhu, Z.; Ihle, N. T.; Dix, M. M.; Jiang, P.; Hayward, M. M.; Saez, E.; Shaw, R. J.; Cravatt, B. F. Chemical Proteomics Identifies Druggable Vulnerabilities in a Genetically Defined Cancer. Cell 2017, 171 (3), 696-709.e23. https://doi.org/10.1016/j.cell.2017.08.051.

(168) Niu, L.; Geyer, P. E.; Gupta, R.; Santos, A.; Meier, F.; Doll, S.; Wewer Albrechtsen, N. J.; Klein, S.; Ortiz, C.; Uschner, F. E.; Schierwagen, R.; Trebicka, J.; Mann, M. Dynamic Human Liver Proteome Atlas Reveals Functional Insights into Disease Pathways. Mol. Syst. Biol. 2022, 18 (5), MSB202210947. https://doi.org/10.15252/msb.202210947.

(169) Rensvold, J. W.; Shishkova, E.; Sverchkov, Y.; Miller, I. J.; Cetinkaya, A.; Pyle, A.; Manicki, M.; Brademan, D. R.; Alanay, Y.; Raiman, J.; Jochem, A.; Hutchins, P. D.; Peters, S. R.; Linke, V.; Overmyer, K. A.; Salome, A. Z.; Hebert, A. S.; Vincent, C. E.; Kwiecien, N. W.; Rush, M. J. P.; Westphall, M. S.; Craven, M.; Akarsu, N. A.; Taylor, R. W.; Coon, J. J.; Pagliarini, D. J. Defining Mitochondrial Protein Functions through Deep Multiomic Profiling. Nature 2022, 606 (7913), 382–388. https://doi.org/10.1038/s41586-022-04765-3.

(170) Overmyer, K. A.; Shishkova, E.; Miller, I. J.; Balnis, J.; Bernstein, M. N.; Peters-Clarke, T. M.; Meyer, J. G.; Quan, Q.; Muehlbauer, L. K.; Trujillo, E. A.; He, Y.; Chopra, A.; Chieng, H. C.; Tiwari, A.; Judson, M. A.; Paulson, B.; Brademan, D. R.; Zhu, Y.; Serrano, L. R.; Linke, V.; Drake, L. A.; Adam, A. P.; Schwartz, B. S.; Singer, H. A.; Swanson, S.; Mosher,

D. F.; Stewart, R.; Coon, J. J.; Jaitovich, A. Large-Scale Multi-Omic Analysis of COVID-19 Severity. Cell Syst. 2021, 12 (1), 23-40.e7. https://doi.org/10.1016/j.cels.2020.10.003.

(171) Stukalov, A.; Girault, V.; Grass, V.; Karayel, O.; Bergant, V.; Urban, C.; Haas, D. A.; Huang, Y.; Oubraham, L.; Wang, A.; Hamad, M. S.; Piras, A.; Hansen, F. M.; Tanzer, M. C.; Paron, I.; Zinzula, L.; Engleitner, T.; Reinecke, M.; Lavacca, T. M.; Ehmann, R.; Wölfel, R.; Jores, J.; Kuster, B.; Protzer, U.; Rad, R.; Ziebuhr, J.; Thiel, V.; Scaturro, P.; Mann, M.; Pichlmair, A. Multilevel Proteomics Reveals Host Perturbations by SARS-CoV-2 and SARS-CoV. Nature 2021, 594 (7862), 246–252. https://doi.org/10.1038/s41586-021-03493-4.

(172) Martinez-Val, A.; Guzmán, U. H.; Olsen, J. V. Obtaining Complete Human Proteomes. Annu. Rev. Genomics Hum. Genet. 2022, 23 (1), 99–121. https://doi.org/10.1146/annurev-genom-112921-024948.

(173) Oliinyk, D.; Meier, F. Ion Mobility-Resolved Phosphoproteomics with Dia-PASEF and Short Gradients. PROTEOMICS 2023, 23 (7–8), 2200032. https://doi.org/10.1002/pmic.202200032.

(174) Wu, X.; Liu, Y.-K.; Iliuk, A. B.; Tao, W. A. Mass Spectrometry-Based Phosphoproteomics in Clinical Applications. TrAC Trends Anal. Chem. 2023, 163, 117066. https://doi.org/10.1016/j.trac.2023.117066.

(175) Tian, X.; Permentier, H. P.; Bischoff, R. Chemical Isotope Labeling for Quantitative Proteomics. Mass Spectrom. Rev. 2023, 42 (2), 546–576. https://doi.org/10.1002/mas.21709.

(176) Guzman, U. H.; Martinez-Val, A.; Ye, Z.; Damoc, E.; Arrey, T. N.; Pashkova, A.; Renuse, S.; Denisov, E.; Petzoldt, J.; Peterson, A. C.; Harking, F.; Østergaard, O.; Rydbirk, R.; Aznar, S.; Stewart, H.; Xuan, Y.; Hermanson, D.; Horning, S.; Hock, C.; Makarov, A.; Zabrouskov, V.; Olsen, J. V. Ultra-Fast Label-Free Quantification and Comprehensive Proteome Coverage with Narrow-Window Data-Independent Acquisition. Nat. Biotechnol. 2024. https://doi.org/10.1038/s41587-023-02099-7.

(177) Shen, Y.; Smith, R. D.; Unger, K. K.; Kumar, D.; Lubda, D. Ultrahigh-Throughput Proteomics Using Fast RPLC Separations with ESI-MS/MS. Anal. Chem. 2005, 77 (20), 6692–6701. https://doi.org/10.1021/ac050876u.

(178) Shen, Y.; Strittmatter, E. F.; Zhang, R.; Metz, T. O.; Moore, R. J.; Li, F.; Udseth, H. R.; Smith, R. D.; Unger, K. K.; Kumar, D.; Lubda, D. Making Broad Proteome Protein Measurements in 1−5 Min Using High-Speed RPLC Separations and High-Accuracy Mass Measurements. Anal. Chem. 2005, 77 (23), 7763–7773. https://doi.org/10.1021/ac051257o.

(179) Bache, N.; Geyer, P. E.; Bekker-Jensen, D. B.; Hoerning, O.; Falkenby, L.; Treit, P. V.; Doll, S.; Paron, I.; Müller, J. B.; Meier, F.; Olsen, J. V.; Vorm, O.; Mann, M. A Novel LC System Embeds Analytes in Pre-Formed Gradients for Rapid, Ultra-Robust Proteomics*.

Mol. Cell. Proteomics 2018, 17 (11), 2284–2296. https://doi.org/10.1074/mcp.TIR118.000853.

(180) Bian, Y.; Zheng, R.; Bayer, F. P.; Wong, C.; Chang, Y.-C.; Meng, C.; Zolg, D. P.; Reinecke, M.; Zecha, J.; Wiechmann, S.; Heinzlmeir, S.; Scherr, J.; Hemmer, B.; Baynham, M.; Gingras, A.-C.; Boychenko, O.; Kuster, B. Robust, Reproducible and Quantitative Analysis of Thousands of Proteomes by Micro-Flow LC–MS/MS. Nat. Commun. 2020, 11 (1), 157. https://doi.org/10.1038/s41467-019-13973-x.

(181) Bekker-Jensen, D. B.; Martínez-Val, A.; Steigerwald, S.; Rüther, P.; Fort, K. L.; Arrey, T. N.; Harder, A.; Makarov, A.; Olsen, J. V. A Compact Quadrupole-Orbitrap Mass Spectrometer with FAIMS Interface Improves Proteome Coverage in Short LC Gradients *. Mol. Cell. Proteomics 2020, 19 (4), 716–729. https://doi.org/10.1074/mcp.TIR119.001906.

(182) Messner, C. B.; Demichev, V.; Bloomfield, N.; Yu, J. S. L.; White, M.; Kreidl, M.; Egger, A.-S.; Freiwald, A.; Ivosev, G.; Wasim, F.; Zelezniak, A.; Jürgens, L.; Suttorp, N.; Sander, L. E.; Kurth, F.; Lilley, K. S.; Mülleder, M.; Tate, S.; Ralser, M. Ultra-Fast Proteomics with Scanning SWATH. Nat. Biotechnol. 2021, 39 (7), 846–854. https://doi.org/10.1038/s41587-021-00860-4.

(183) Ishikawa, M.; Konno, R.; Nakajima, D.; Gotoh, M.; Fukasawa, K.; Sato, H.; Nakamura, R.; Ohara, O.; Kawashima, Y. Optimization of Ultrafast Proteomics Using an LC-Quadrupole-Orbitrap Mass Spectrometer with Data-Independent Acquisition. J. Proteome Res. 2022, 21 (9), 2085–2093. https://doi.org/10.1021/acs.jproteome.2c00121.

(184) Szyrwiel, L.; Gille, C.; Mülleder, M.; Demichev, V.; Ralser, M. Fast Proteomics with Dia-PASEF and Analytical Flow-Rate Chromatography. PROTEOMICS 2024, 24 (1–2), 2300100. https://doi.org/10.1002/pmic.202300100.

(185) Bonneil, E.; Tessier, S.; Carrier, A.; Thibault, P. Multiplex Multidimensional nanoLC-MS System for Targeted Proteomic Analyses. ELECTROPHORESIS 2005, 26 (24), 4575–4589. https://doi.org/10.1002/elps.200500603.

(186) Livesay, E. A.; Tang, K.; Taylor, B. K.; Buschbach, M. A.; Hopkins, D. F.; LaMarche, B. L.; Zhao, R.; Shen, Y.; Orton, D. J.; Moore, R. J.; Kelly, R. T.; Udseth, H. R.; Smith, R. D. Fully Automated Four-Column Capillary LC−MS System for Maximizing Throughput in Proteomic Analyses. Anal. Chem. 2008, 80 (1), 294–302. https://doi.org/10.1021/ac701727r.

(187) Orton, D. J.; Wall, M. J.; Doucette, A. A. Dual LC–MS Platform for High-Throughput Proteome Analysis. J. Proteome Res. 2013, 12 (12), 5963–5970. https://doi.org/10.1021/pr400738a.

(188) Lee, H.; Lee, J. H.; Kim, H.; Kim, S.-J.; Bae, J.; Kim, H. K.; Lee, S.-W. A Fully Automated Dual-Online Multifunctional Ultrahigh Pressure Liquid Chromatography System for High-

Throughput Proteomics Analysis. J. Chromatogr. A 2014, 1329, 83–89. https://doi.org/10.1016/j.chroma.2013.12.084.

(189) Lee, H.; Mun, D.-G.; Bae, J.; Kim, H.; Yeon Oh, S.; Soo Park, Y.; Lee, J.-H.; Lee, S.-W. A Simple Dual Online Ultra-High Pressure Liquid Chromatography System (sDO-UHPLC) for High Throughput Proteome Analysis. Analyst 2015, 140 (16), 5700–5706. https://doi.org/10.1039/C5AN00639B.

(190) Webber, K. G. I.; Truong, T.; Johnston, S. M.; Zapata, S. E.; Liang, Y.; Davis, J. M.; Buttars, A. D.; Smith, F. B.; Jones, H. E.; Mahoney, A. C.; Carson, R. H.; Nwosu, A. J.; Heninger, J. L.; Liyu, A. V.; Nordin, G. P.; Zhu, Y.; Kelly, R. T. Label-Free Profiling of up to 200 Single-Cell Proteomes per Day Using a Dual-Column Nanoflow Liquid Chromatography Platform. Anal. Chem. 2022, 94 (15), 6017–6025. https://doi.org/10.1021/acs.analchem.2c00646.

(191) Kreimer, S.; Haghani, A.; Binek, A.; Hauspurg, A.; Seyedmohammad, S.; Rivas, A.; Momenzadeh, A.; Meyer, J. G.; Raedschelders, K.; Van Eyk, J. E. Parallelization with Dual-Trap Single-Column Configuration Maximizes Throughput of Proteomic Analysis. Anal. Chem. 2022, 94 (36), 12452–12460. https://doi.org/10.1021/acs.analchem.2c02609.

(192) Kreimer, S.; Binek, A.; Chazarin, B.; Cho, J. H.; Haghani, A.; Hutton, A.; Marbán, E.; Mastali, M.; Meyer, J. G.; Mesquita, T.; Song, Y.; Van Eyk, J.; Parker, S. High-Throughput Single-Cell Proteomic Analysis of Organ-Derived Heterogeneous Cell Populations by Nanoflow Dual-Trap Single-Column Liquid Chromatography. Anal. Chem. 2023, 95 (24), 9145–9150. https://doi.org/10.1021/acs.analchem.3c00213.

(193) Staes, A.; Boucher, K.; Dufour, S.; Maia, T. M.; Timmerman, E.; Haver, D. V.; Pauwels, J.; Demol, H.; Vandenbussche, J.; Gevaert, K.; Impens, F.; Devos, S. High-Throughput Nanoflow Proteomics Using a Dual-Column Electrospray Source. Anal. Chem. 2024, 96 (17), 6534–6539. https://doi.org/10.1021/acs.analchem.4c00845.

(194) Chen, L.; Zhang, Z.; Matsumoto, C.; Gao, Y. High-Throughput Proteomics Enabled by a Fully Automated Dual-Trap and Dual-Column LC-MS. Anal. Chem. 2024, 96 (24), 9761–9766. https://doi.org/10.1021/acs.analchem.3c03182.

(195) Xie, X.; Truong, T.; Huang, S.; Johnston, S. M.; Hovanski, S.; Robinson, A.; Webber, K. G. I.; Lin, H.-J. L.; Mun, D.-G.; Pandey, A.; Kelly, R. T. Multicolumn Nanoflow Liquid Chromatography with Accelerated Offline Gradient Generation for Robust and Sensitive Single-Cell Proteome Profiling. Anal. Chem. 2024, 96 (26), 10534–10542. https://doi.org/10.1021/acs.analchem.4c00878.

(196) Wang, H.; Hanash, S. M. Increased Throughput and Reduced Carryover of Mass Spectrometry-Based Proteomics Using a High-Efficiency Nonsplit Nanoflow Parallel Dual-

Column Capillary HPLC System. J. Proteome Res. 2008, 7 (7), 2743–2755. https://doi.org/10.1021/pr700876g.

(197) Hosp, F.; Scheltema, R. A.; Eberl, H. C.; Kulak, N. A.; Keilhauer, E. C.; Mayr, K.; Mann, M. A Double-Barrel Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) System to Quantify 96 Interactomes per Day*. Mol. Cell. Proteomics 2015, 14 (7), 2030–2041. https://doi.org/10.1074/mcp.O115.049460.

(198) Shen, Y.; Tolić, N.; Zhao, R.; Paša-Tolić, L.; Li, L.; Berger, S. J.; Harkewicz, R.; Anderson, G. A.; Belov, M. E.; Smith, R. D. High-Throughput Proteomics Using High-Efficiency Multiple-Capillary Liquid Chromatography with On-Line High-Performance ESI FTICR Mass Spectrometry. Anal. Chem. 2001, 73 (13), 3011–3021. https://doi.org/10.1021/ac001393n.

(199) Peters-Clarke, T. M.; Coon, J. J.; Riley, N. M. Instrumentation at the Leading Edge of Proteomics. Anal. Chem. 2024, 96 (20), 7976–8010. https://doi.org/10.1021/acs.analchem.3c04497.

Chapter 2

Fast and Deep Phosphoproteome Analysis with the Orbitrap Astral

Lancaster, N. M.; Sinitcyn, P.; Forny, P.; Peters-Clarke, T. M.; Fecher, C.; Smith, A. J.; Shishkova, E.; Arrey, T. N.; Pashkova, A.; Robinson, M. L.; Arp, N.; Fan, J.; Hansen, J.; Galmozzi, A.; Serrano, L. R.; Rojas, J.; Gasch, A. P.; Westphall, M. S.; Stewart, H.; Hock, C.; Damoc, E.; Pagliarini, D. J.; Zabrouskov, V.; Coon, J. J. Fast and Deep Phosphoproteome Analysis with the Orbitrap Astral Mass Spectrometer. *Nat Commun* **2024**, 15 (1), 7016. https://doi.org/10.1038/s41467-024-51274-0.

Abstract

Owing to its roles in cellular signal transduction, protein phosphorylation plays critical roles in myriad cell processes. That said, detecting and quantifying protein phosphorylation has remained a challenge. We describe the use of a novel mass spectrometer (Orbitrap Astral) coupled with data-independent acquisition (DIA) to achieve rapid and deep analysis of human and mouse phosphoproteomes. With this method, we map approximately 30,000 unique human phosphorylation sites within a half-hour of data collection. The technology is benchmarked to other state-of-the-art MS platforms using both synthetic peptide standards and with EGF-stimulated HeLa cells. We apply this approach to generate a phosphoproteome multi-tissue atlas of the mouse. Altogether, we detect 81,120 unique phosphorylation sites within 12 hours of measurement. With this unique dataset, we examine the sequence, structural, and kinase specificity context of protein phosphorylation. Finally, we highlight the discovery potential of this resource with multiple examples of novel phosphorylation events relevant to mitochondrial and brain biology.

Introduction

Protein phosphorylation is an essential post-translational regulatory mechanism for myriad cellular functions including apoptosis, inflammation, metabolism, proliferation, protein trafficking, and many others.[1] Global detection of which proteins, and perhaps most importantly which residues, are subject to this dynamic modification has been a key technological gap for decades.[2] In the early 2000s, key advancements in technologies to enrich phosphorylated peptides – prior to mass spectrometry (MS) analysis – enabled a new era of large-scale phosphorylation discovery experiments.[3–13] The evolution of that work, combined with gradual, but steady, improvements in both MS hardware and data analysis tools make it possible to map and quantify thousands of phosphorylation sites in a single experiment.[14–27] Still, current phosphoproteomic analyses are neither routine nor straightforward when compared to protein detection and quantification.

Phosphoproteome analysis remains challenging due to four main requirements: need for site localization, dynamic range, reproducibility, and throughput. The quality of the tandem mass spectra required to determine precisely upon which residue the phosphoryl group resides is higher than that required to identify an unmodified peptide. Alternative dissociation methods, such as electron transfer dissociation (ETD), and improved mass resolving power and accuracy are both strategies that can aid in improving spectral quality and increasing the likelihood that the detected site can be localized with high confidence.[15,21,28,29] Additionally, sensitivity and dynamic range of the mass analyzer can also be important for detection of low level, but critical, product ions.[11,23,27] Next, the often sub-stoichiometric amounts of phosphorylated protein elevates the difficulty.[2,13,17–19] Beyond that, previous studies demonstrate that this dynamic range problem is further

exacerbated in that ~10% of detected phosphopeptides account for ~80% of the observed signal.[30] This tremendous dynamic range necessitates both enrichment of phosphorylated peptides and chromatographic fractionation ahead of conventional capillary liquid chromatography tandem MS (nLC-MS/MS).[2,17,18]

Phosphorylation site quantification is often essential to elucidating biological insight.[17] The requirements outlined above also add challenges to achieving this goal. Quantification of a protein, for example, is done by summing the signals of multiple unique peptides all stemming from that single protein. Thus, if a single peptide is not reproducibly detected, the overall measurement can still be made. But unique phosphopeptides cannot be summed with other signals; they must be reproducibly and reliably detected from one sample to the next. Such demands make performing truly large-scale (*i.e.*, > 100 samples) comparisons of phosphoproteomes very difficult.[18] Finally, getting sufficient depth to detect targets of a particular kinase, for example, might require extensive fractionation, as discussed. However, scaling that experiment to multiple conditions or samples is often not possible from a throughput perspective, despite all the caveats noted. We conclude that global and quantitative phosphoproteomics technologies require improvement to enable routine and truly large-scale phosphoproteome measurements.

Recently, a new type of mass analyzer has been described – the Asymmetric Track Lossless analyzer (Astral™). The Astral analyzer can achieve high resolving powers (~80,000) and mass accuracy (5 ppm), single ion detection limit, and MS/MS scan speeds up to 200 Hz.[31,32] Here we describe the use of a quadrupole-Orbitrap™-Astral hybrid MS instrument for the analysis of phosphopeptides. Specifically, we examine the ability of this system to perform MS/MS scans of complex mixtures of phosphopeptides separated over durations ranging from 7 to 60 minutes. We

examine the performance of these methods for various peptide mass loads and data acquisition settings as a function of detected and localized phosphorylation sites and overall reproducibility. Validation of site localization and quantification was accomplished using synthetic phosphopeptide standards spiked into a yeast phosphopeptide background. Next, we benchmarked the performance of our Orbitrap Astral method against an Orbitrap Ascend and a previously described timsTOF Pro method by replicating an EGF stimulation of HeLa cells.[33] Finally, we leverage this technology to collect an atlas of phosphorylation in the mouse – generating tens of thousands of localized phosphorylation sites from each of twelve unique tissues. With these data we present, to our knowledge, the deepest mouse phosphoproteome collected in a single study. Using this atlas combined with AlphaFold predicted protein structure[34], we confirm existing hypotheses that most phosphorylation events are directed towards unstructured regions of proteins.[35–37] The incorporation of a previous kinome atlas allows investigation of tissue-specific kinase activity.[38] We additionally provide examples of how our novel resource can be mined for key phosphorylation events on biologically relevant proteins.

Results

Rapid phosphopeptide analysis with the Orbitrap Astral mass spectrometer.

Owing to its fast MS/MS scan rate and high sensitivity, we hypothesized that the Orbitrap Astral™ mass spectrometer could resolve many of the aforementioned challenges in analyzing phosphoproteomes. Specifically, the Orbitrap Astral MS comprises a conventional quadrupole-Orbitrap coupled with a new mass analyzer (Astral, **Figure 2.1A**). The combination of low ion losses and single ion detection drives the high sensitivity of the Astral analyzer, while very fast MS/MS scan rates allow it to cycle through large numbers of targets. In a typical Orbitrap Astral method, the Astral analyzer is set to generate 200 MS/MS spectra per second while the Orbitrap analyzer in parallel generates slower high resolution and high dynamic range MS data.[31]

To test the utility of the Orbitrap Astral MS for the analysis of phosphopeptides, we purified tryptic phosphopeptides from HEK239T cells, loaded them onto a nanoflow capillary column with a pulled electrospray emitter[39], and eluted them into the Orbitrap Astral mass spectrometer. For this initial experiment, we utilized a data-dependent acquisition (DDA) method wherein MS scans were acquired in the Orbitrap analyzer while MS/MS scans were collected using the Astral analyzer. **Figure 2.1B** represents an example single-scan tandem mass spectrum that was identified following a traditional MaxQuant database search.[25,40] Here, a triply-protonated precursor having *m/z* value of 623.6314 and a sequence of RPsQNAISFFNVGHSK was selected, dissociated with beam-type collisional activation (HCD), and analyzed in the Astral analyzer– all within 5 ms. In total, from this single 30-minute nLC-MS/MS DDA experiment, we collected 3,201 MS and 174,944 MS/MS scans, from which we identified 12,327 phosphopeptides corresponding to 9,537

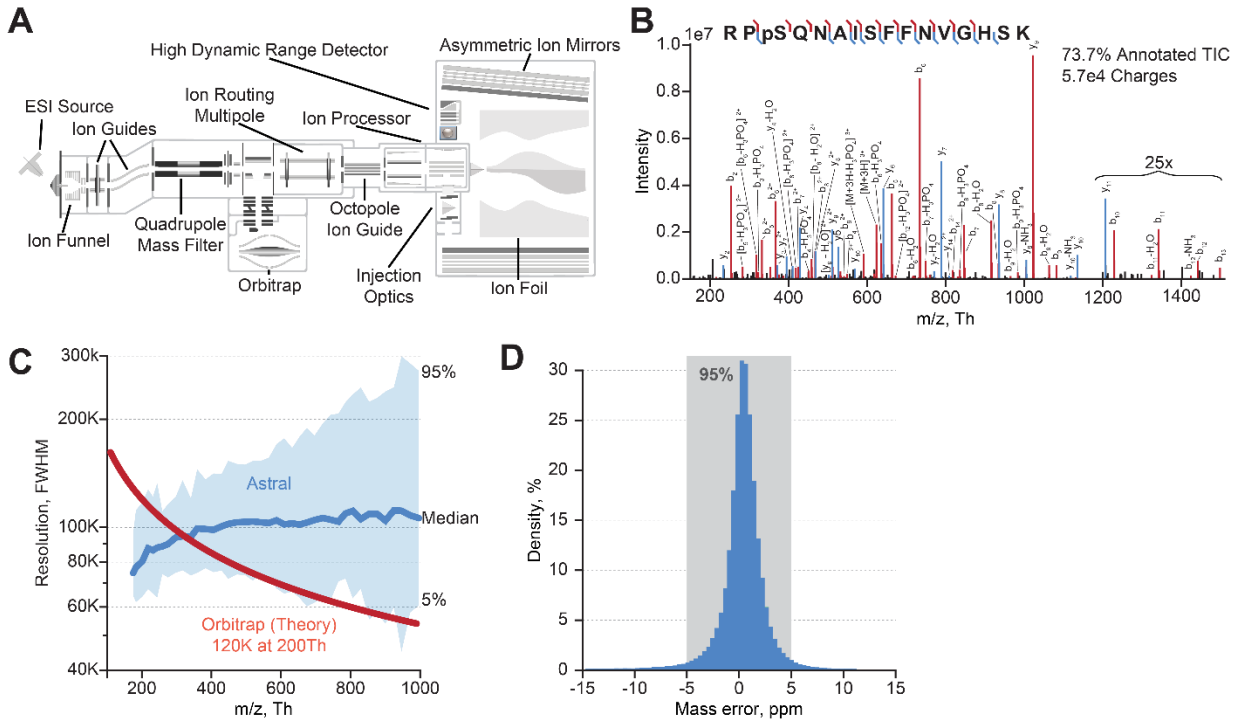

Figure 2.1. Overview of Orbitrap Astral MS and its key figures of merit. (A) Orbitrap Astral instrument schematic, highlighting the quadrupole, Orbitrap, and Astral analyzers. (B) Tandem mass spectrum of representative phosphopeptide collected using the Astral analyzer. (C) Distribution of Astral analyzer resolving power as a function of mass from phosphopeptide product ion spectra collected here. (D) Phosphopeptide product ion mass measurement error from Astral analysis. Image Description

unique sites of phosphorylation. Note a previous study using the Orbitrap Tribrid platform with a DDA method reported a depth of ~9,500 phosphosites following a 120-minute method.[20] These MS/MS scans were collected within ~6.3 ms on average, enabling fast sampling of selected precursors, and sometimes exhausting the precursors available to sequence. Next, we investigated the Astral analyzer's capabilities by plotting the measured resolving power (**Figure 2.1C**) and mass accuracy (**Figure 2.1D**) for these product ions. The upper bound of observed resolving powers likely correspond to peaks with low numbers of ions, whereas the lower bound presumably arises from high intensity peaks exhibiting space charging.[31] Note the Astral resolving power persists with increasing *m/z* value. Finally, mass accuracy, as measured for these product ions, was within 5 ppm for 95% of the measurements.

With the exceptional speed of the Astral mass analyzer, combined with its ability to deliver high mass resolution and accuracy, we next wondered how this instrument would perform for DIA analysis of phosphopeptides.[33,41–45] Accordingly, we analyzed the same tryptic phosphopeptide mixture using a DIA method. The speed of the Astral analyzer allowed for use of DIA windows as narrow as 2 *m/z* across the range of 380 to 980 *m/z* while still maintaining a cycle time of 1.5 seconds. This scan method is visualized in **Figure 2.2A** with the high-resolution Orbitrap MS scans (240K resolving power) denoted in blue (~0.6 s cycle time) and the Astral MS/MS scans depicted in red.

A 250-ng injection of tryptic phosphopeptides was analyzed with a 30-minute nLC separation and the DIA method described above (see **Figure 2.3A** for the base peak chromatogram). The resultant data was searched both with Spectronaut and a developmental build of CHIMERYS (see Methods). From the Spectronaut search, we localized 29,190 phosphorylation

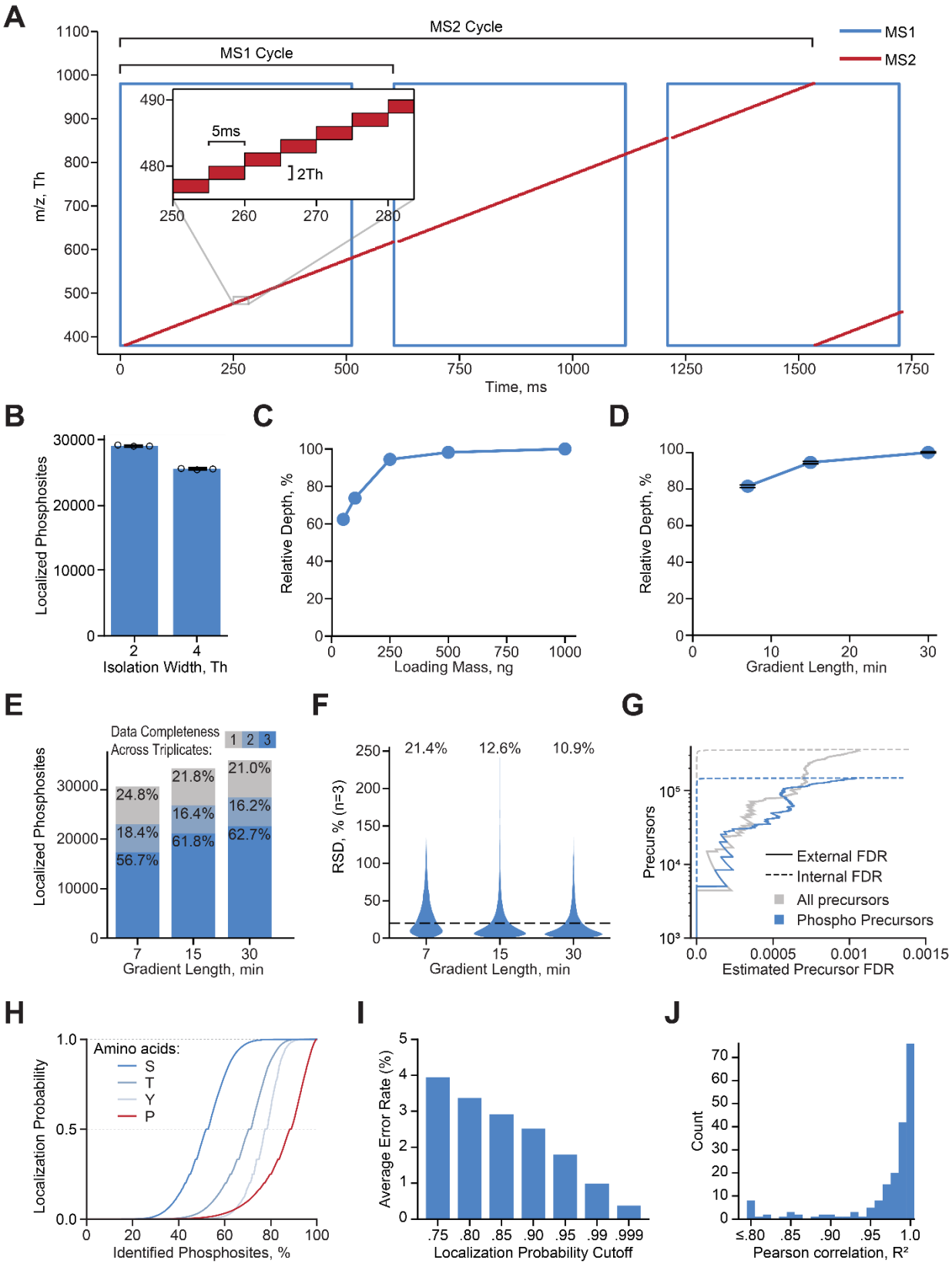
A
MS2 Cycle
MS1 Cycle
MS1
MS2
5ms
2Th
m/z, Th
Time, ms
B
Localized Phosphosites
Isolation Width, Th
C
Relative Depth, %
Loading Mass, ng
D
Relative Depth, %
Gradient Length, min
E
Data Completeness Across Triplicates:
1 2 3
24.8%
18.4%
56.7%
21.8%
16.4%
61.8%
21.0%
16.2%
62.7%
Localized Phosphosites
Gradient Length, min
F
21.4%
12.6%
10.9%
RSD, % (n=3)
Gradient Length, min
G
Precursors
External FDR
Internal FDR
All precursors
Phospho Precursors
Estimated Precursor FDR
H
Amino acids:
S
T
Y
P
Localization Probability
Identified Phosphosites, %
I
Average Error Rate (%)
Localization Probability Cutoff
J
Count
Pearson correlation, R²

Figure 2.2 Overview of DIA phosphoproteomics on the Orbitrap Astral MS. (A) Illustration of DIA acquisition scheme. (B) Average number of localized phosphorylation sites identified using DIA method with various isolation widths for n=3 injection replicates. Error bars indicate maximum and minimum observed values. (C) Evaluation of phosphopeptide loading mass on performance. Data was collected with a 30-minute active gradient and 2 *m/z* DIA isolation width. (D) Effect of gradient length on performance. Points represent the average across n=3 injection replicates. Error bars indicate maximum and minimum observed values. (E) Evaluation of phosphosite identification reproducibility (for results from (D)). (F) Evaluation of phosphosite quantitative precision. The relative standard deviation of phosphosite quantities is shown for phosphosites detected across triplicate injections with the median value displayed for each active gradient length. (G) Comparison of External (maize/human entrapment experiment) and Internal (Spectronaut) FDR on precursor level. (H) Phosphoproline Decoy Search to test reliability of localization algorithm. The cumulative distribution of localization probabilities is shown for phosphorylation events at different amino acids. (I) Average site localization error rate as a function of localization probability cutoff. Synthetic phosphopeptide standards spiked into a yeast phosphopeptide background were used as a ground truth for error rate determination. (J) The distribution of R^2 values for linear calibrations curves is shown for phosphopeptides standards detected in at least three concentrations across a five-point dilution series into a constant yeast phosphopeptide background. Image Description

sites (**Figure 2.2B,** see **Figure 2.4A** for comparison of Spectronaut and CHIMERYS, and **Figure 2.4B** and **Figure 2.4C** for phosphopeptide multiplicity and localization probability distributions, respectively). We note that the use of a wider DIA window (4 *m/z*) produced 14% fewer localized phosphorylation sites. Next, we examined the impact of loading amount on performance using a dilution series ranging from 1 µg to 50 ng loaded on column (**Figure 2.2C**). We note that similar results are obtained down to 250 ng loads. The sensitivity of the system is evident in that a 50-ng load produced ~60% of the detected sites observed with a 1 µg load. To determine whether the Astral scan speed would allow for increased throughput, we next evaluated a range of active gradient lengths (**Figure 2.2D, Figure 2.3A**). A four-fold reduction in gradient length (30 to 7 minutes) results in just a ~20% reduction in localized phosphorylation sites. Additional method parameters are justified in the **Methods**.

To determine the impact of gradient length on reproducibility, we performed two additional analyses. First, we examined identification reproducibility across triplicate injections (**Figure 2.2E**). The 7- ,15-, and 30-minute active gradient methods resulted in 56.7%, 61.8%, and 62.7% of phosphosites detected across all three injection replicates, respectively. Second, we assessed quantitative precision by determining percent relative standard deviation (RSD) of phosphorylation sites detected across all three replicates (**Figure 2.2F**). The 15- and 30-minute active gradients achieved median RSDs of 10.9 and 12.6%. Not unexpectedly, the 7-minute active gradient method has considerably reduced precision (21.4% median RSD), most likely due to insufficient sampling across the narrower elution peaks (**Figure 2.3B-E**).

Since these data were collected with a novel mass analyzer and new instrument, we next sought to ensure that the phosphosites were identified and localized accurately. To test this, we

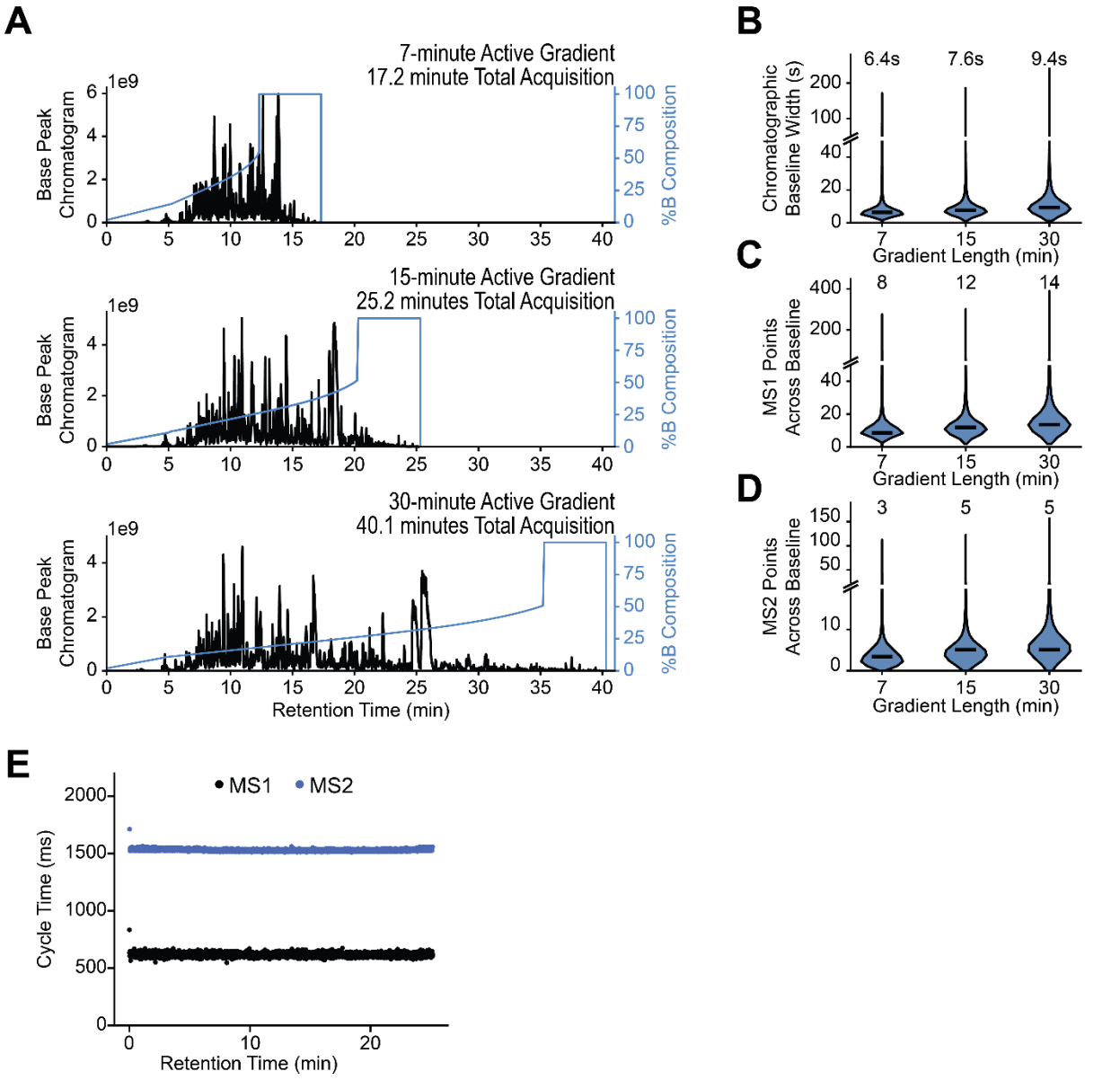

Figure 2.3 Chromatographic Properties of HEK293T Phosphoproteomics Datasets. (A) Representative chromatograms for different gradient lengths. The base peak chromatograms for the three different active gradient lengths are shown with the gradient composition curve plotted. The distribution of (B) chromatographic baseline widths, (C) MS^1 points across the baseline peak, and (D) MS^2 points across the baseline peaks are shown with the median value indicated in the violin plot and printed above for three different active gradient lengths. The values were estimated by multiplying chromatographic FWHM, MS^1 points across FWHM, and MS^2 points across FWHM reported by Spectronaut 17 by 1.7 (based on the 4σ baseline width of a Gaussian distribution where 4σ ≈1.7 * FWHM). (E) The cycle times for the MS^1 and MS^2 scan sequences are shown as a function of RT. Image Description

first performed an entrapment search with a combined human and maize protein database.[46] In this experiment, which is commonly used to evaluate internal false discovery rate (FDR), maize false identifications allow for the calculation of an external FDR. The results of this analysis are shown in **Figure 2.2G**, where internal and external FDR is plotted for all precursors and phosphopeptide precursors. For lower q-values the internal and external FDRs disagree; however, these curves converge at approximately the 1% FDR threshold, suggesting that the identifications reported here are reliable. We conclude that, as the DIA software tools continue to evolve and are adopted for the new data type, this discordance will be diminished.

In the case of phosphoproteomics, peptide identification is not the final result – the site of phosphorylation is ideally localized to a specific residue with confidence. To evaluate the quality of phosphosite localization, we searched these data with phosphorylated proline as a variable modification where any detected proline phosphorylation is false.[47] S, T, and Y all have a significant amount of highly confident localized phosphosites whereas the P does not (**Figure 2.2H, Figure 2.4H**). To further explore localization confidence, we analyzed a set of synthetic phosphopeptide standards that had been spiked into a complex mixture yeast tryptic phosphopeptides at various concentrations. Note this same 225 synthetic phosphopeptide mix has been used by others and thereby provides a means of both assessment and comparison[41,48]; specifically, the experiment allows for calculation of localization error rate based on ground truth knowledge. From these data we conclude that, depending on localization probability cutoff, the error rate ranges between 1 and 5 percent, a similar trend as the previous study (**Figure 2.2I, Figure 2.5A**).[41] **Figure 2.5B** demonstrates that we retain the majority of correct precursors even at more stringent localization probability cutoffs. Taken together, these results confirm that a library-free

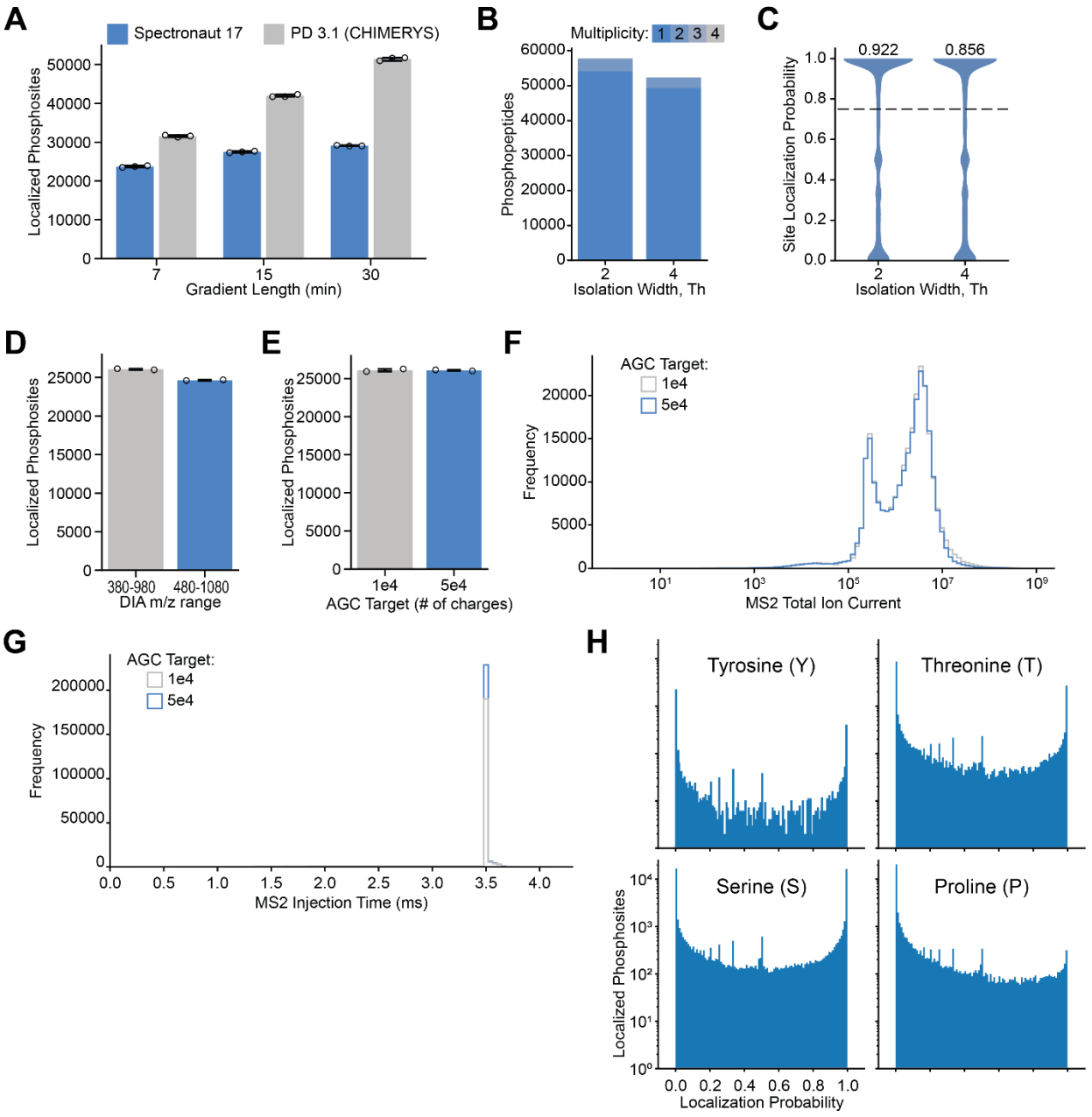

Figure 2.4. Evaluation of Phosphoproteomic Method and Processing Parameters. (A) The average number of localized phosphosites across triplicate injections for three different gradient lengths is shown for searches in Spectronaut (blue) and CHIMERYS in Proteome Discoverer 3.1 (gray). The error bars represent the maximum and minimum value across the triplicates. (B) The average number of phosphopeptides and multiplicity distributions are shown for the data in **Figure 2.2B**. (C) The site localization probability distributions are shown for the data in **Figure 2.2B** with the median value indicated above. (D) The average number of phosphosites detected across two replicate injections in EGF-stimulated HeLa cells are shown for two different DIA *m/z* ranges.

Error bars indicated the maximum and minimum number of phosphosites. (E) The average number of phosphosites detected across two replicate injections in EGF-stimulated HeLa cells are shown for two different AGC target values. Error bars indicated the maximum and minimum number of phosphosites. (F) The MS^2 total ion current distributions are indicated for the two AGC targets in **Figure 2.4E**. (G) The distributions of MS^2 injection times are indicated for the two AGC targets in **Figure 2.4E**. (H) Histograms of localization probabilities for S, T, Y, and P residues are shown as an alternative visualization for the phosphoproline decoy search results in **Figure 2.2H**. Image Description

DIA search, using MS/MS spectra from an Astral analyzer, can reliably detect and localize phosphorylation sites.

The aforementioned phosphopeptide standards also provided an opportunity to evaluate quantitative linearity.[42,48] To accomplish this, phosphopeptide standards were diluted from 10,000 to 39 attomole (per standard, on-column) into a constant complex mixture of yeast tryptic phosphopeptides. Using the 15-minute DIA Orbitrap Astral method, we analyzed the various spike-in samples and then calculated the linear fit between the observed MS intensities and the load amount. **Figure 2.5D** displays three randomly chosen curves while **Figure 2.2J** confirms that the majority of the phosphosite calibration curves have R^2 values greater than 0.95, indicating good quantitative linearity. **Figure 2.5C** confirms that even at loads as low as 39 attomole, half of the phosphorylation sites were detected across all three injections. Finally, we note that the phosphopeptide standard intensity distributions observed across the dilution series overlap with the yeast phosphopeptide intensity distribution (**Figure 2.5E**). Together, these results demonstrate that our method exhibits reliable detection and quantification over the concentration range typically observed in phosphoproteomic analysis.

Comparison of Orbitrap Astral to other phosphoproteomics platforms

To compare the performance of the Orbitrap Astral to other phosphoproteomics platforms we replicated the HeLa EGF stimulation experiment as previously reported by Skowronek et al. using the timsTOF Pro hybrid mass spectrometer with dia-PASEF acquisition.[33] Briefly, we cultured HeLa cells in triplicate followed by 15 minutes of EGF stimulation, cell lysis, protein extraction, trypsin digestion, and phosphopeptide enrichment. The resulting phosphopeptide samples were then analyzed with a 15-minute active gradient, closely replicating the total

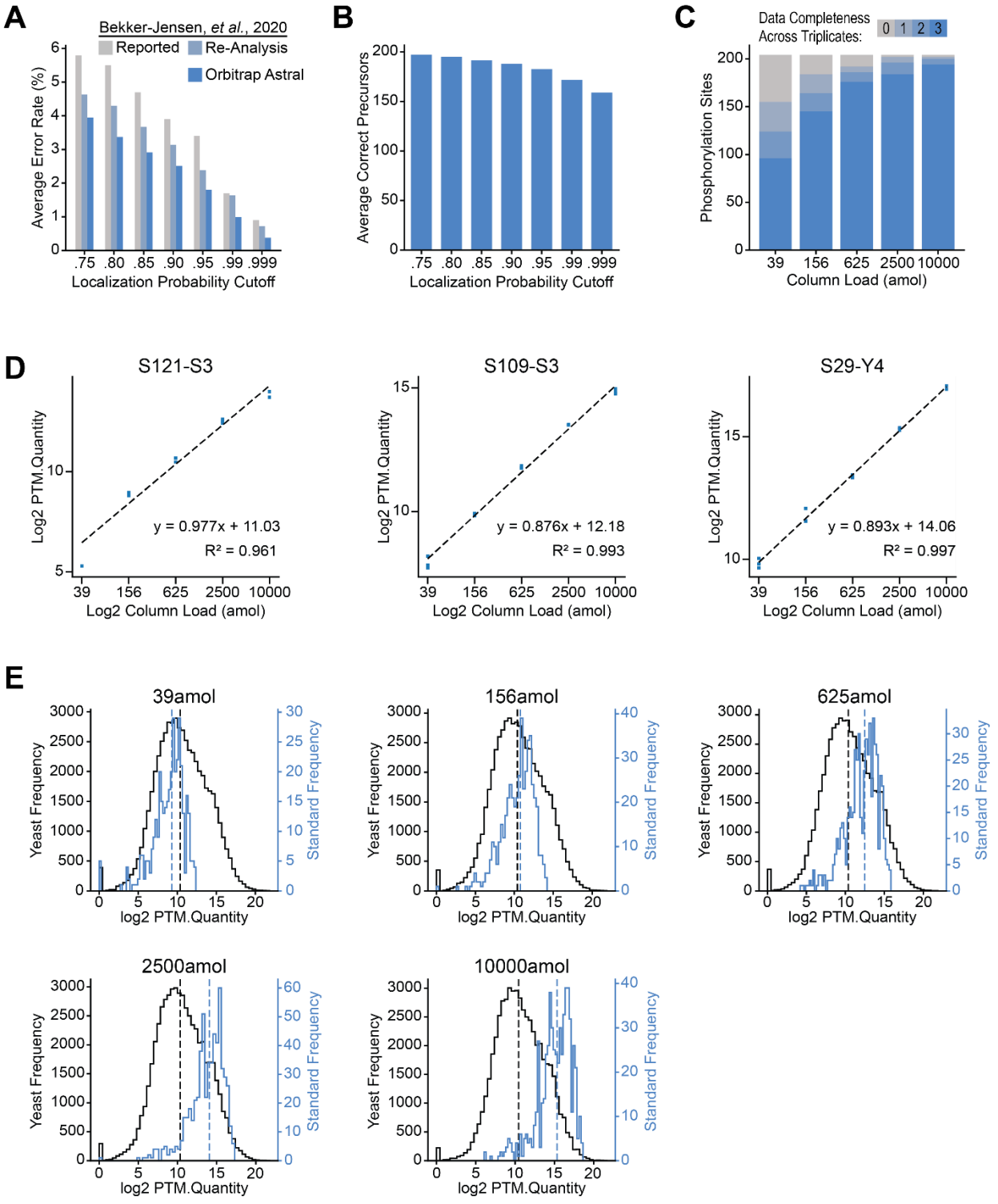

Figure 2.5. Validation of Site Localization and Quantification using Synthetic Phosphopeptide Standards. (A) The localization error rates shown in **Figure 2.2I** are compared to the values previously reported by Bekker-Jensen *et al.*[41] alongside the error rates calculated

following re-analysis of the data from Bekker-Jensen *et al.* with the same protein database and Spectronaut workflow used in this study. (B) The average number of correctly identified precursors as a function of localization probability cutoff are shown for the dataset in **Figure 2.2I**. (C) The detection completeness of synthetic phosphosites across triplicate injections is shown as a function of the column load for the data in **Figure 2.2J**. (D) Calibrations curves for three randomly selected synthetic phosphosites are shown from the synthetic phosphopeptides dilution series dataset. 'S121-S3' indicates the phosphorylation serine at position 3 for phosphopeptide standard 121. (E) The distribution of phosphosite intensities (Spectronaut's PTM.Quantity value) for the synthetic phosphopeptide standards and yeast phosphopeptides are shown in blue and black, respectively, for each of the dilution series points. The median intensity for the phosphopeptide standards and yeast phosphopeptides is indicated on each plot with a blue and black dotted line, respectively. Image Description

acquisition time for the Skowronek *et al*. publication, with detection using either the Orbitrap Astral or Orbitrap Ascend™ Tribrid mass spectrometers. To enable a direct comparison, we searched our raw files and those from the Skowronek *et al.* study with the same protein database and Spectronaut library-free method. **Figure 2.6A** shows the phosphoproteomic depth and completeness across triplicates for all three datasets. The Orbitrap Astral method provided the deepest phosphoproteomic analysis and yielded near 60% phosphosite completeness across replicates from approximately three-fold more localized sites as compared to either the timsTOF Pro or Orbitrap Ascend instruments. Note the Orbitrap Ascend is not optimally operated in DIA mode given its slower scan speed; however, we chose to use a DIA method for direct comparison. Interestingly, the Orbitrap Ascend generated the highest level of phosphosite overlap of all three instruments at 64%.

To examine the dynamic range of detected phosphorylation sites across platforms we plotted the intensity distributions of phosphorylation sites detected in all three replicates using only the EGF-stimulated data. **Figure 2.6B** and **Figure 2.6C** present the Orbitrap Astral intensity distributions with overlapping sites indicated for either the timsTOF Pro or Orbitrap Ascend datasets, respectively. Both plots reveal that the unique phosphosites in the Orbitrap Astral dataset are biased toward lower intensities. In contrast, unique phosphosites for the timsTOF Pro dataset are distributed more evenly across the intensity range (**Figure 2.6D**). These observations suggest that our Orbitrap Astral exhibits improved sensitivity, resulting in a larger dynamic range. Note the Orbitrap Astral and Ascend shows higher overlap as compared to the timsTOF Pro (**Figure 2.6E**); we suppose this difference is likely due to variation in EGF treatment and sample preparation.

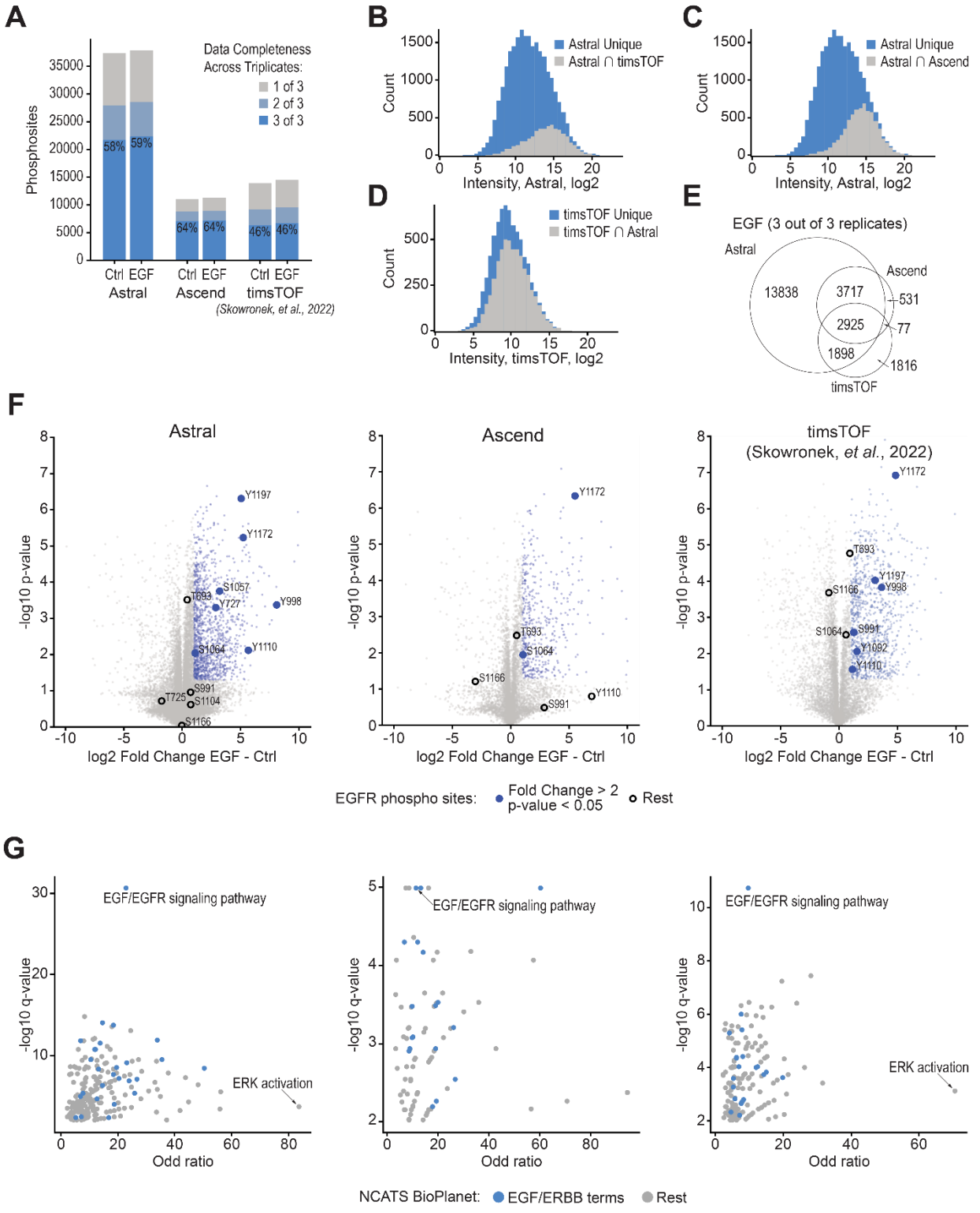

Figure 2.6. Biological validation of phosphoproteomics platforms using EGF stimulation. (A) Reproducibility of detected phosphosites across biological triplicates for three phosphoproteomics

platforms – Orbitrap Astral, Orbitrap Ascend, and timsTOF Pro.[33] Intensity distribution of phosphosites consistently detected (3 out of 3 EGF stimulated replicates) by Orbitrap Astral with overlap indicated for (B) timsTOF Pro and (C) Orbitrap Ascend. (D) Intensity distribution of phosphosites consistently detected (3 out of 3 EGF stimulated replicates) by timsTOF Pro with overlap indicated for Orbitrap Astral. (E) Venn diagram of phosphosites detected across all three mass spectrometry platforms. (F) Volcano plots of phosphosites between EGF stimulation and control across platforms with phosphosites meeting the differential expression criteria (fold-change >2, p-value < 0.05 via two-sided t-test without multiple testing correction) indicated in blue and EGFR phosphosites labeled. (G) Pathway enrichment analysis using the NCATS BioPlanet terms was performed for the three platforms with EGF/ERBB-associated terms indicated in blue. Image Description

To see how this performance variation translates to biological discovery we performed a differential expression analysis between control and EGF-stimulated samples for all three datasets (**Figure 2.6F**). Each datapoint in these plots presents a phosphorylation site with the differentially expressed ones highlighted in blue. Phosphorylation sites occurring directly on the EGFR are labeled, many of which are detected across all platforms. To further compare, we performed a pathway enrichment analysis from each dataset (**Figure 2.6G**) – all three platforms had the EGF/EGFR signaling pathway as the top enriched term. We conclude that despite differences in treatment, sample preparation, and instrumentation, all technologies could arrive at the same biological conclusions.

Deep phosphoproteome analysis of the mouse

Having demonstrated the speed and sensitivity of the Orbitrap Astral MS for reliable identification of phosphopeptides, we next sought to leverage this methodology to generate a comprehensive phosphoprotein atlas of the mouse. In 2010, Huttlin and co-workers described the first mouse phosphoproteome atlas, reporting 35,965 phosphorylation sites detected following 10 days of mass spectrometry analysis and serving as a valuable reference point for this study.[49] **Figure 2.7A** presents our overall experimental design wherein proteins were isolated, digested, and enriched for phosphorylation from 12 distinct tissues. For each tissue, the purified phosphopeptides were separated offline and concatenated into four fractions, each of which were analyzed using a 15-minute DIA nLC-MS/MS method on the Orbitrap Astral MS. In total, this experiment required 12 hours of active data acquisition time and resulted in the detection of 81,120 unique phosphorylation sites. Note to provide complementary protein abundance we separately

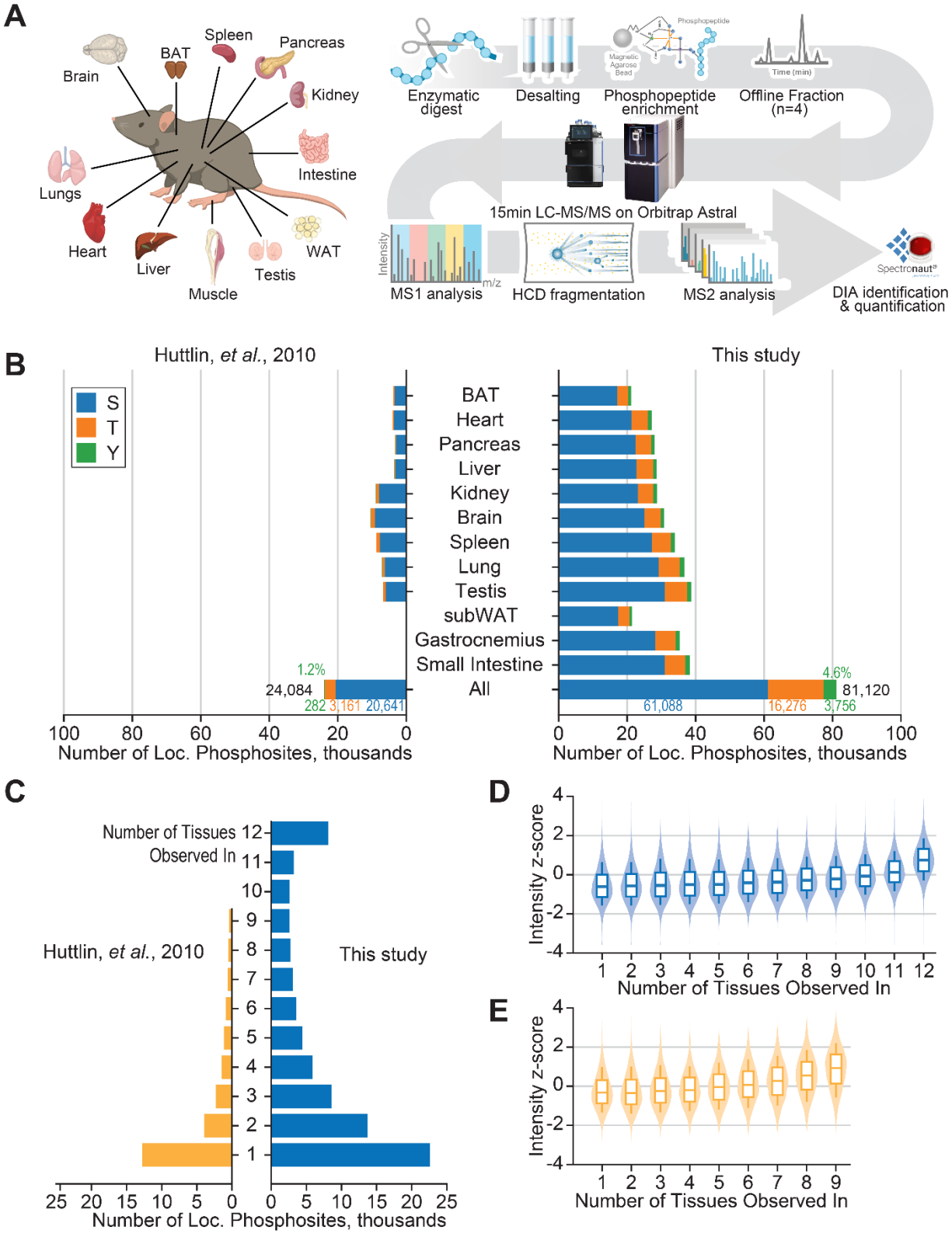
A
Brain
BAT
Spleen
Pancreas
Kidney
Intestine
Lungs
Heart
Liver
Muscle
Testis
WAT
Enzymatic digest
Desalting
Phosphopeptide enrichment
Offline Fraction (n=4)
15min LC-MS/MS on Orbitrap Astral
MS1 analysis
HCD fragmentation
MS2 analysis
DIA identification & quantification
B
Huttlin, et al., 2010
This study
S
T
Y
BAT
Heart
Pancreas
Liver
Kidney
Brain
Spleen
Lung
Testis
subWAT
Gastrocnemius
Small Intestine
All
1.2%
24,084
282 3,161 20,641
4.6%
81,120
61,088
16,276 3,756
Number of Loc. Phosphosites, thousands
C
Number of Tissues Observed In
Huttlin, et al., 2010
This study
Number of Loc. Phosphosites, thousands
D
Intensity z-score
Number of Tissues Observed In
E
Intensity z-score
Number of Tissues Observed In

Figure 2.7. Mouse Phosphorylation Atlas Workflow and Results. (A) Mouse Phosphorylation Atlas Workflow. (B) Mouse Tissue Phosphoproteomic Analysis. Numbers of unique phosphorylation sites are shown for each tissue and the total unique sites with the fraction of S,T, and Y localizations indicated. The Huttlin *et al.* results[49] were generated by researching the raw data in MaxQuant using the same protein database used in this study. (C) Tissue Specificity of Detected Phosphorylation sites. The y-axis indicates the number of tissues in which a phosphosite was detected. (D) and (E) Intensity distributions for phosphorylation sites detected in a given number of tissues for this study and by Huttlin *et al.*, respectively. For the distributions in (D) and (E), the 0.1, 0.25, 0.5, 0.75, and 0.9 percentiles are indicated with the lower error bar, lower box bound, center box line, upper box bound, and upper error bar, respectively. Each distribution was generated by down-sampling to 1000 datapoints from the phosphosite categories in (C). Image Description

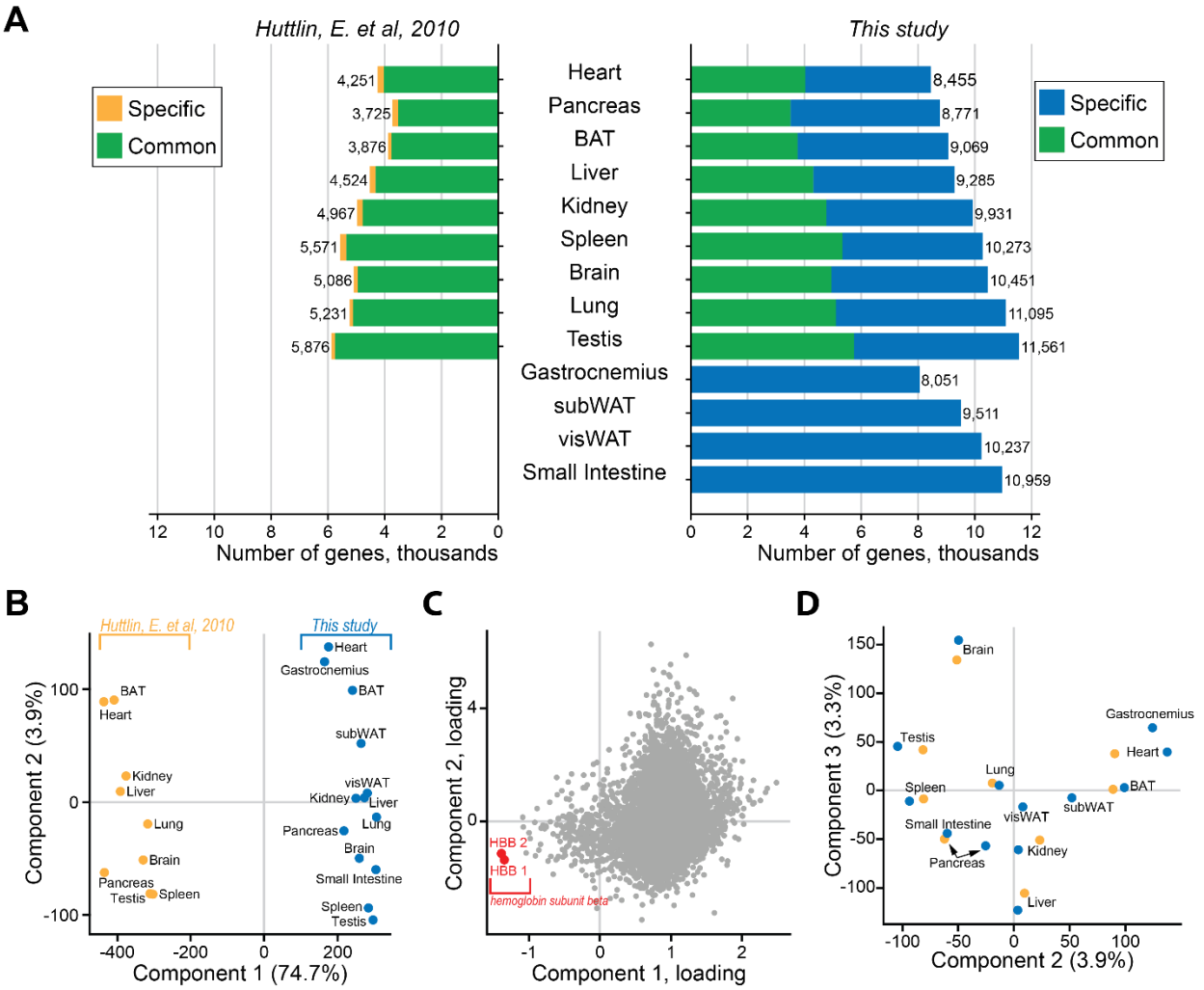

Figure 2.8. Proteomic Profiling of Mouse Tissues. (A) Results from Mouse Tissue Proteomic Analysis. The numbers of protein groups are indicated for each tissue, with common protein groups highlighted in green, those unique to Huttlin *et al.* in yellow, and those unique to this study in blue. The Huttlin *et al.* results were generated by analyzing raw data in MaxQuant, employing the same protein database utilized in this study. (B) Principal Component Analysis (PCA) based on protein expression of common proteins between two studies. (C) PCA loading plot of first two components from (B) with highlights of HBB1 and HBB2 proteins. (D) The same PCA as (B) but Component 2 and Component 3 visualized. Image Description

analyzed non-enriched peptides from each tissue (**Figure 2.8**).

Depending upon the source tissue, the number of identified phosphorylation sites ranged from approximately 20,000 to 40,000 sites (**Figure 2.7B**). Of the 81,120 unique sites identified, 61,088 were localized to S; 16,276 localized to T; and 3,756 to Y. It is noteworthy that ~4.6% of the total sites identified stem from pY. Previous studies have placed this number closer to ~2% and often rely on pY-specific antibodies for enrichment of this perhaps most functionally important phosphorylation site.[30,50] We suppose that the increased sensitivity and depth afforded by this new analyzer permits the detection of these low-expression phosphorylation events (**Figure 2.9C**). Also shown in **Figure 2.7B** are the results, by tissue, from the previous mouse atlas. Note these raw data were reprocessed using MaxQuant with the matching protein database and filtered using similar quality metrics. On a per tissue comparison, the method leveraged here provides a 5-fold boost in the number of localized phosphorylation sites in approximately 1/24th the time.

Next, we plotted the distribution, by tissue, of phosphorylation sites and whether they were detected in both studies (**Figure 2.9Figure *2.9*A**). Despite the age and strain differences between mice, and the dynamic nature of phosphorylation, we see good overlap in detection of sites (median 55.3%). Tissue-specificity was also plotted (**Figure 2.7C** and **Figure 2.9B**) and compared to Huttlin *et al.* In both studies, a large portion of phosphorylation sites appear to have single tissue specificity; however, our results indicate that a sizeable number of sites are detected across all twelve tissues. This result contrasts with the Huttlin work and likely stems from the increased sensitivity and reproducibility afforded by the new analyzer and DIA method. With this result, we expect that many sites may indeed be present across all tissues and that further improvements in sensitivity and reproducibility are likely needed to detect them. This hypothesis is further evinced

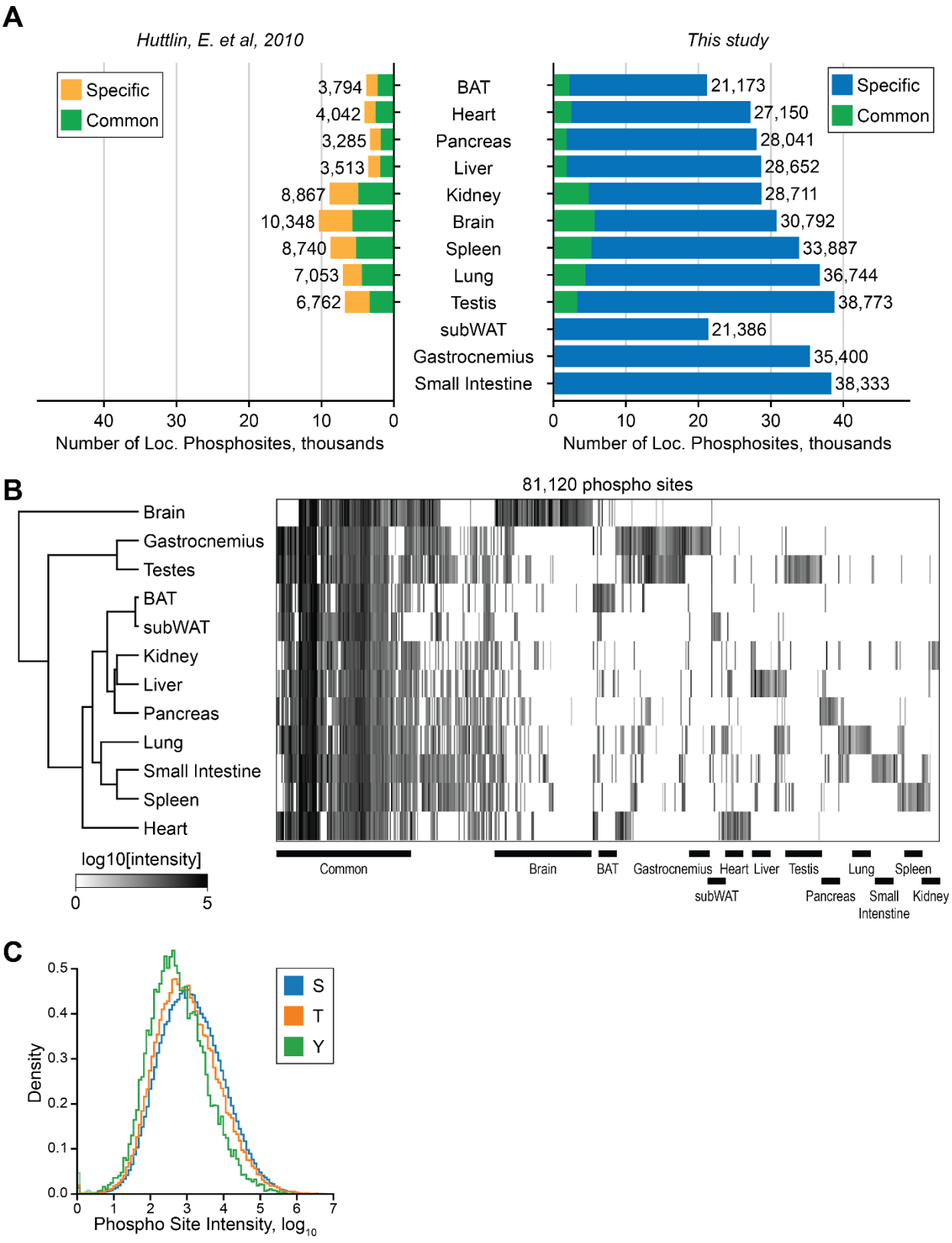
A
Huttlin, E. et al, 2010
This study
Specific
Common
3,794
4,042
3,285
3,513
8,867
10,348
8,740
7,053
6,762
BAT
Heart
Pancreas
Liver
Kidney
Brain
Spleen
Lung
Testis
subWAT
Gastrocnemius
Small Intestine
21,173
27,150
28,041
28,652
28,711
30,792
33,887
36,744
38,773
21,386
35,400
38,333
40 30 20 10 0
0 10 20 30 40
Number of Loc. Phosphosites, thousands
Number of Loc. Phosphosites, thousands
B
81,120 phospho sites
Brain
Gastrocnemius
Testes
BAT
subWAT
Kidney
Liver
Pancreas
Lung
Small Intestine
Spleen
Heart
log10[intensity]
0
5
Common
Brain
BAT
Gastrocnemius
Heart
Liver
Testis
Lung
Spleen
subWAT
Pancreas
Small Intenstine
Kidney
C
Density
0.0 0.1 0.2 0.3 0.4 0.5
0 1 2 3 4 5 6 7
Phospho Site Intensity, log10
S
T
Y

Figure 2.9. Mouse Phosphorylation Atlas Comparison to Huttlin *et al.* (A) Overlap of phosphorylation sites detected in this study and by Huttlin *et al.*[51] (B) Heatmap showing the relative intensities of phosphosites detected across tissues. Note that there are many phosphorylation sites with similar intensities across all tissues, but each tissue tends to have clusters of phosphosites that are uniquely high intensity (C) Normalized distribution of phosphorylation site intensities for S, T, and Y localizations. Image Description

by the results shown in **Figure 2.7D-E**, where we examined the detected peak intensities for each site as a function of how many tissues in which it was observed. From these data, both studies show concordance between number of tissues detected and overall signal strength (*i.e.*, abundance).

Furthermore, we compared our comprehensive phosphoprotein atlas of the mouse with the most recent multi-tissue study conducted by Giansanti and co-workers.[52] Similar to our sample preparation, the phosphopeptides from multiple tissues were offline fractionated and further concatenated into 4 samples. These samples were measured using a 90-minute gradient DDA on the Q Exactive Orbitrap HF MS system. Despite a 6-fold shorter measurement time, the Orbitrap Astral detected 2-3 times more phosphosites per tissue, with improved agreement with our dataset compared to Huttlin *et al.* (median 64.2% versus 55.3%; **Figure 2.10A**). Interestingly, Giansanti's study shows a similar trend towards a higher number of sites that are present across all tissues (**Figure 2.10B**). The portion of pY in this study is 1.1% (**Figure 2.10C**), which additionally highlights the capability of combining Astral sensitivity with DIA acquisition to enable the detection of 4.6% pY in our data.

Phosphorylation sites in the context of protein sequence and structure

Phosphorylation sites within proteins exhibit a specificity determined by both the sequence and structural motifs. In our investigation, we sought to unravel this specificity by clustering phosphorylation sites and their flanking regions. To achieve this, we computed all pairwise comparisons of phosphosite sequences (site and five amino acids in both C and N-terminal direction) and projected this comparison onto a 2D plane using the t-SNE algorithm (**Figure 2.11A**). Notably, our analysis highlighted the importance of the phosphorylated amino acid as a

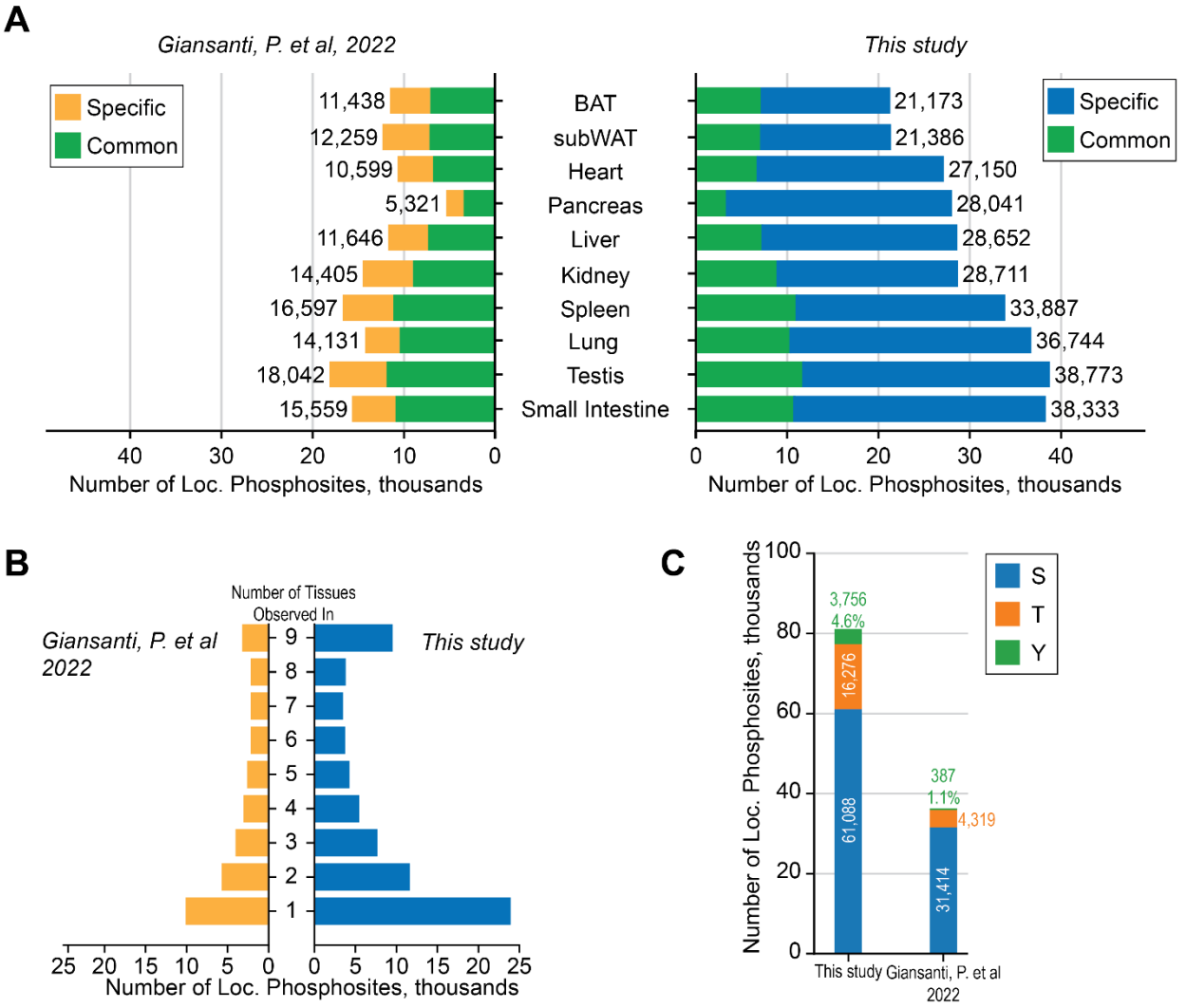

Figure 2.10. Mouse Phosphorylation Atlas Comparison to Giansanti *et al.* (A) Overlap of phosphorylation sites detected in this study and by Giansanti *et al*.[52] (B) Tissue Specificity of Detected Phosphorylation sites, excluding the pancreas, brain, and gastrocnemius tissues from comparison. The y-axis indicates the number of tissues in which a phosphosite was detected. (C) The total number of S, T, Y phosphorylation sites are shown for this study and Giansanti *et al*.[52] The percentage of the total sites occurring on phosphotyrosine residues is indicated for both studies. Image Description

major factor driving separation (**Figure 2.11B**). That is clusters 1-8 are Ser-based, while 9 is Tyr-based, and 10-12 are Thr-based. Intriguingly, upon repeating the analysis while excluding the phosphorylated amino acid from binary comparisons (**Figure 2.12A**), we identified predominant clusters organized by specific motifs (**Figure 2.12B**), such as S/T-P (cluster 1), and RXX-S/T (cluster 9). Proline-directed protein kinases demonstrate a preference for phosphorylating serine or threonine residues immediately preceding a proline residue in proteins. Prominent examples of such kinases include extracellular signal-regulated kinases (ERKs) and cyclin-dependent kinases (CDKs). Leveraging annotations from the PhosphoSitePlus database[53], our analysis revealed that the majority of ERK/CDK annotated phosphorylation sites belong to the S/T-P cluster (**Figure 2.11C**). Note there is limited representation of Y in clusters 1 and 9, even though the phosphorylated residue was not considered in this analysis (**Figure 2.11D**). These data confirm known biology that Y is not a substrate for P and R-directed protein kinases.[54]

To explore structural motifs, we mapped the 81,120 phosphorylation sites detected here onto protein structures predicted by AlphaFold.[34] The UniProt structural library (assembled using AlphaFold) contains structures with varying degrees of certainty as measured by a per residue confidence score (pLDDT) where 0 is very low and 100 is very high confidence. Note that lower confidence scores are indicative of either flexible or disordered protein regions.[55] Upon plotting the confidence score distribution for phosphorylated S, T, and Y sites, we observed a striking prevalence for these sites to be in low confidence regions (**Figure 2.12C**), especially when compared to all S, T, and Y residues or even all amino acids. Phosphorylation sites that were detected in all tissues had an even slightly higher preference for these low confidence regions. However, all amino acids observed in non-enriched proteomics experiments did not follow the

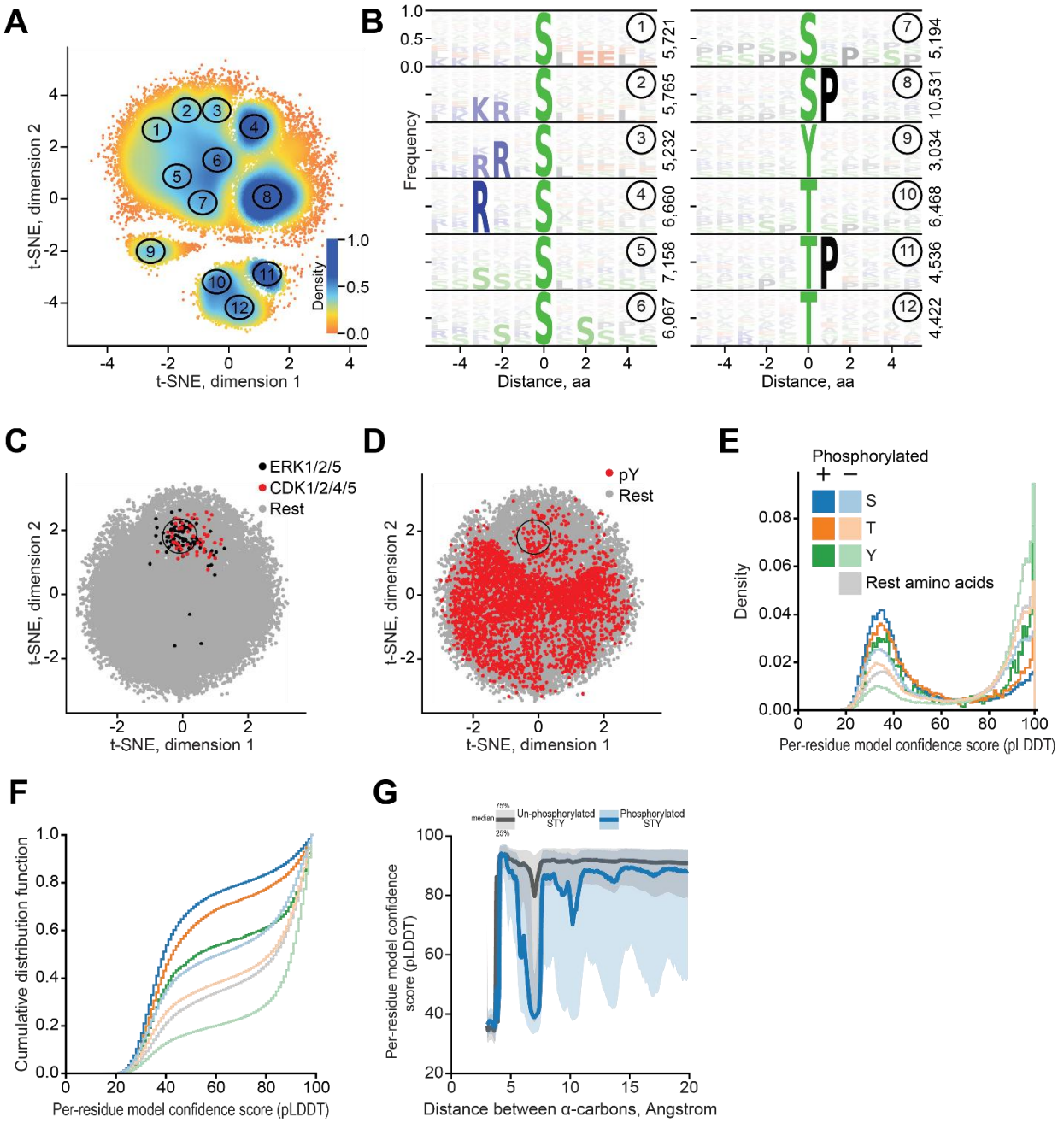

Figure 2.11. Sequence and Structural Context of Phosphosites. (A) Two-dimensional representation of all phosphosites and their 5-amino acid flanking sequences, including the central amino acid in the comparison. Each cluster has been manually selected to highlight the densest regions. (B) Sequence logo plot for all clusters depicted in (A). (C) Similar to **Figure 2.11A** with highlights of ERK and CDK kinases based on the PhosphoSitePlus database. (D) Similar to **Figure 2.11A** with highlights of all phosphorylated tyrosine sites. (E) Distribution of structural confidence scores for phosphorylated (+) and unphosphorylated STY sites. (F) Similar to (E) but presented as a cumulative distribution plot. (G) Confidence score distribution based on the distance between

alpha-carbons surrounding phosphorylation site and phosphorylation state – phosphorylated and un-phosphorylated S/T/Y. Image Description

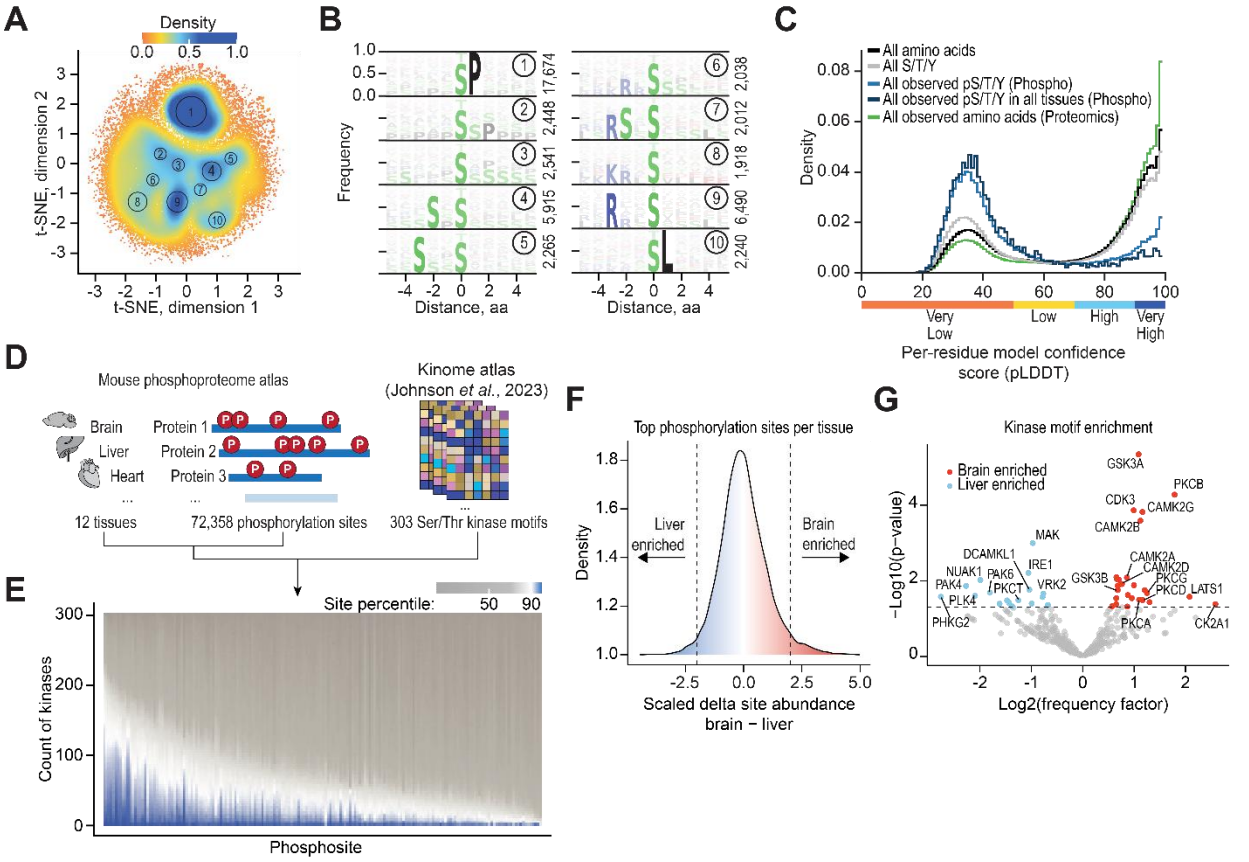

Figure 2.12. Sequence, Structural and Kinase Specificity Context of Phosphosites. (A) Two-dimensional representation of all phosphosites and their 5-amino acid flanking sequences, excluding the central amino acid from the comparison. Each cluster has been manually selected to emphasize the densest regions. (B) Sequence logo plot for all clusters depicted in (A). (C) Distribution of confidence scores for all amino acids, specifically S/T/Y, and for phosphorylated S/T/Y detected across all tissues. (D) Our mouse phosphoproteome data derived from nine tissues was applied to the kinome atlas search tool. (E), All phosphorylation sites detected in our study are plotted on the x-axis, sorted by the number of kinases that scored higher than 90 for a specific site. (F), Z-score transformed difference between abundances of shared phosphorylation sites in brain and liver tissue. Vertical dashed lines indicate thresholds for selection of phosphorylation sites that are used for kinase motif enrichment analysis. A chi-squared test was used to calculate p-values after applying Haldane's correction. (G), Based on the top sites per tissue, a motif enrichment analysis was performed, and the resulting frequency of how often a kinase was predicted to act on a site was plotted on the x-axis, along with the p-value on the y-axis. The scheme and types of analyses have been adapted from Reference 38. Image Description

same trend, emphasizing the specificity of this effect for phosphorylation sites. It is noteworthy that when plotted individually, S, T, and Y phosphorylation sites all exhibit the same trend (**Figure 2.11E** and **Figure 2.11F**). These global results confirm previous work that suggests phosphorylation sites are enriched in intrinsically disordered regions.[35,36] In fact, some efforts have used intrinsically disordered regions to refine phosphorylation site prediction models.[37] Finally, although phosphorylation is more likely to occur in less structured regions, examination of the surrounding environment shows the confidence scores are increasing with distance from the phosphorylated residue (**Figure 2.11G**). These data suggest further research on the role of structure and phosphorylation site is warranted.

Kinase predictions based on flanking regions of phosphorylation sites

Further leveraging our extensive dataset of phosphorylation sites across various tissues, we applied a recently published kinase prediction tool[38], which was developed by using synthetic peptide libraries to profile the substrate sequence specificity of 303 serine/threonine kinases (**Figure 2.12D**). As expected, a systematic analysis of all phosphorylation sites in our dataset revealed a trend where many phosphorylation sites were assigned to a limited number of putative kinases; however, half of all phosphorylation sites were predicted to be targeted by 24 or more kinases (**Figure 2.12E**). Furthermore, this approach enabled us to predict kinases for sites that showed tissue enrichment (**Figure 2.12F**). By comparing enriched phosphorylation motifs between tissue pairs, we identified kinases with established activities in certain tissues (**Figure 2.13**). For instance, comparing brain and liver enriched motifs revealed kinases such as GSK3A[56], CAMK subfamily 2 members[57–59], and PCKA/PCKB[60,61], which are known to modulate the brain

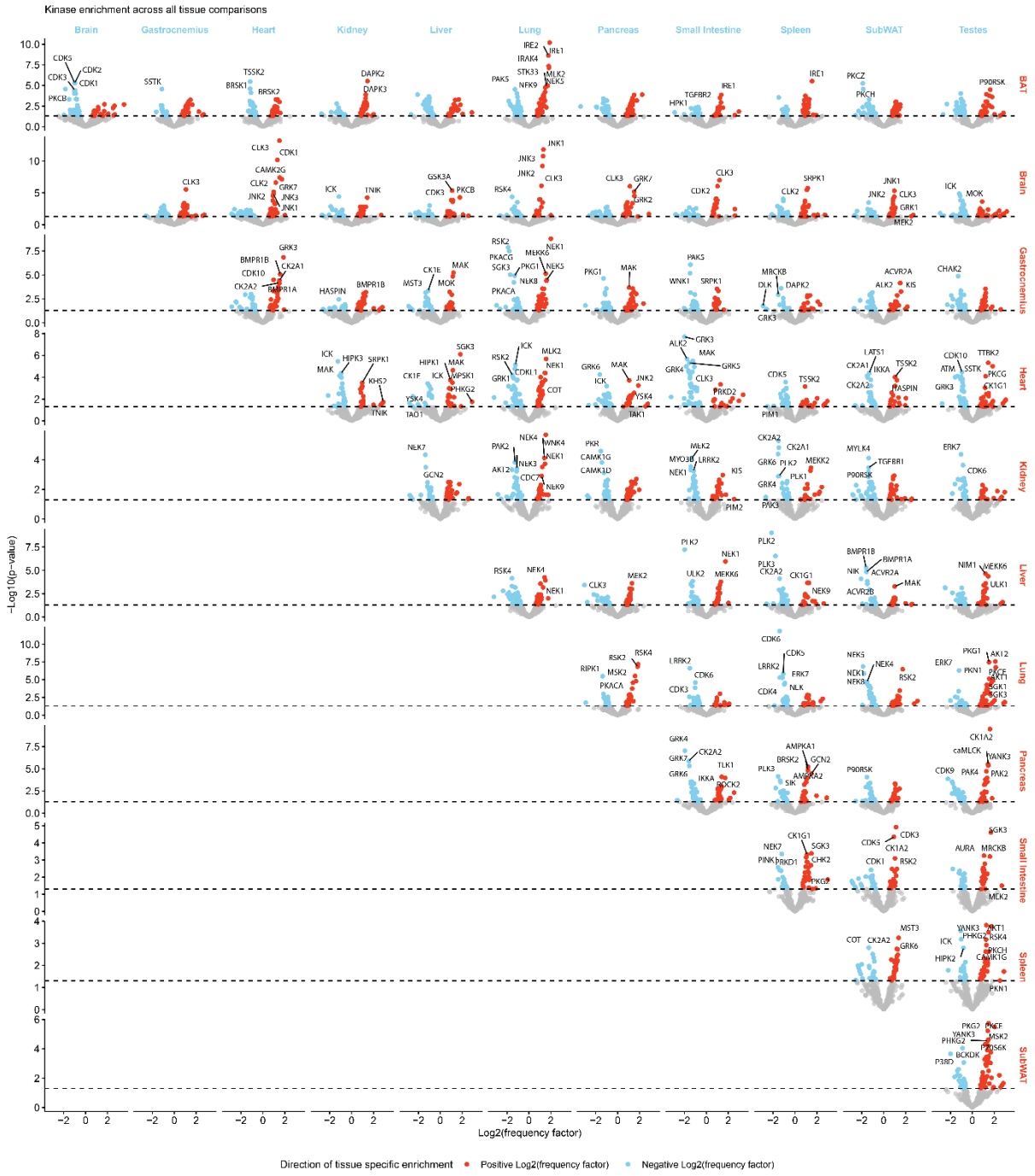

Figure 2.13. Kinase Enrichment Across Tissues. For each possible tissue-tissue pair, the top tissue-specific phosphorylation sites were selected to perform a kinase motif enrichment analysis. The frequency of how often a kinase was predicted to act on a site in a tissue was plotted on the x-axis. The p-value is shown on the y-axis. The red dots indicate significant kinases in the tissues that are labeled in red on the right side of the plot. The blue dots correspond to the tissues plotted across the top of the graph. The p-values were calculated using a chi-squared test on the data set

after applying Haldane's correction to compare the frequency factors for each kinase in each tissue. The frequency factor was calculated as the proportion of how often a kinase was predicted within the top phosphorylation sites versus the total of top phosphorylation sites, divided by the same ratio for the unchanged sites, *i.e.*, not ranked among the top/bottom 200 (with regards to z-score). Image Description

phosphoproteome and have been implicated in neurological disease states (**Figure 2.12G**). While these kinases have activities in other tissues, our dataset emphasized that the phosphorylation sites predicted to be clients of these kinases are strongly overrepresented in brain as compared to liver tissue. Additionally, we identified PHKG2 as the kinase with the strongest frequency factor for liver tissue, consistent with this enzyme's role as a liver isoform critical for glycogen breakdown.[62] Our analysis suggests additional kinases with probable tissue-specific activities, highlighting the potential of our resource to facilitate the discovery of such tissue-specific kinase functions.

Novel phosphorylation sites on mitochondrial proteins

We next sought to investigate how our comprehensive analysis of phosphorylation patterns in the mouse proteome can be further applied in the biochemical and biological context. We first compared our dataset to the mouse phosphorylation dataset from PhosphoSitePlus[53], revealing a capture of 38,454 previously unidentified phosphorylation sites, constituting 53% of the total identified sites within our dataset (**Figure 2.14A**). Upon categorizing phosphorylation sites, we observed a significant representation of mitochondrial proteins, with 55% of known mitochondrial proteins harboring at least one phosphorylation site (**Figure 2.14B**). This observation aligns with the growing appreciation for protein phosphorylation in regulating mitochondrial function.[63] An exemplary mitochondrial pathway whose protein phosphorylation sites showed the most dynamic pattern across all tissues was the tricarboxylic acid cycle (**Figure 2.15A**) with the majority of sites being detected for the first time in this study (**Figure 2.15B**), setting the stage for further exploration into the understudied regulation of this key metabolic pathway.[64]

The occurrence of phosphorylation sites was variable across tissues in the entire dataset

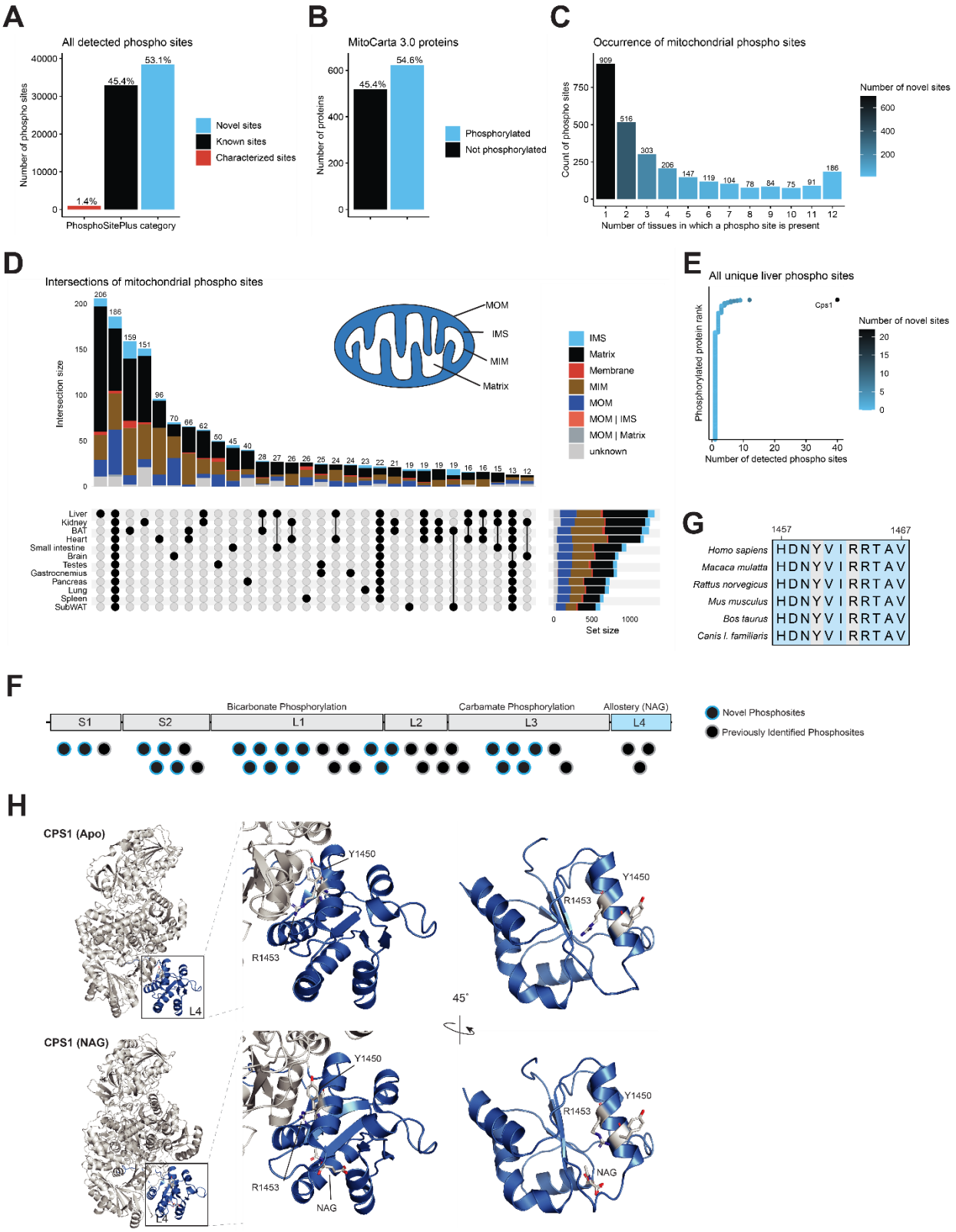
A
All detected phospho sites
Number of phospho sites
PhosphoSitePlus category
1.4%
45.4%
53.1%
Novel sites
Known sites
Characterized sites
B
MitoCarta 3.0 proteins
Number of proteins
45.4%
54.6%
Phosphorylated
Not phosphorylated
C
Occurrence of mitochondrial phospho sites
Count of phospho sites
Number of tissues in which a phospho site is present
Number of novel sites
909
516
303
206
147
119
104
78
84
75
91
186
D
Intersections of mitochondrial phospho sites
Intersection size
MOM
IMS
MIM
Matrix
IMS
Matrix
Membrane
MIM
MOM
MOM | IMS
MOM | Matrix
unknown
Liver
Kidney
BAT
Heart
Small intestine
Brain
Testes
Gastrocnemius
Pancreas
Lung
Spleen
SubWAT
Set size
E
All unique liver phospho sites
Phosphorylated protein rank
Cps1
Number of detected phospho sites
Number of novel sites
G
1457
1467
Homo sapiens HDNYVIRRTAV
Macaca mulatta HDNYVIRRTAV
Rattus norvegicus HDNYVIRRTAV
Mus musculus HDNYVIRRTAV
Bos taurus HDNYVIRRTAV
Canis l. familiaris HDNYVIRRTAV
F
Bicarbonate Phosphorylation
Carbamate Phosphorylation
Allostery (NAG)
S1
S2
L1
L2
L3
L4
Novel Phosphosites
Previously Identified Phosphosites
H
CPS1 (Apo)
Y1450
R1453
L4
45˚
CPS1 (NAG)
NAG

Figure 2.14. Mitochondrial phosphoproteomics reveals novel liver-specific phosphorylation site. (A) Bar plot of all detected phosphorylation sites in our study stratified into categories that are directly derived from the PhosphoSitePlus database (downloaded August 22, 2023). (B) Mouse MitoCarta 3.0[65] proteins with at least one detected phosphorylation site in our data versus the remainder. (C) Bar plot indicating the number of mitochondrial phosphorylation sites that occur in a specific number of tissues. (D) Intersections of phosphorylation sites on mitochondrial proteins per tissue subsets. Intersection sizes of 12 or more are shown. Sub-mitochondrial localization is derived from MitoCarta 3.0. (E) Dot plot displaying all proteins that harbor phosphorylation sites unique to liver tissue. The highest ranked protein harbors the most phosphorylation sites as indicated on the x-axis. Novel sites are according to the PhosphoSitePlus database. (F) Schematic representation of the CPS1 peptide chain with all detected phosphorylation sites indicated as novel or previously identified. (G) Sequence alignment of CPS1 orthologues using Clustal Omega[66]. Patient variant residue and phosphorylation site residue of interest are in grey. (H) Structural modeling based on structures (PDB: 5DOT (Apo), 5DOU (NAG)) from RCSB PDB[67]. Image Description

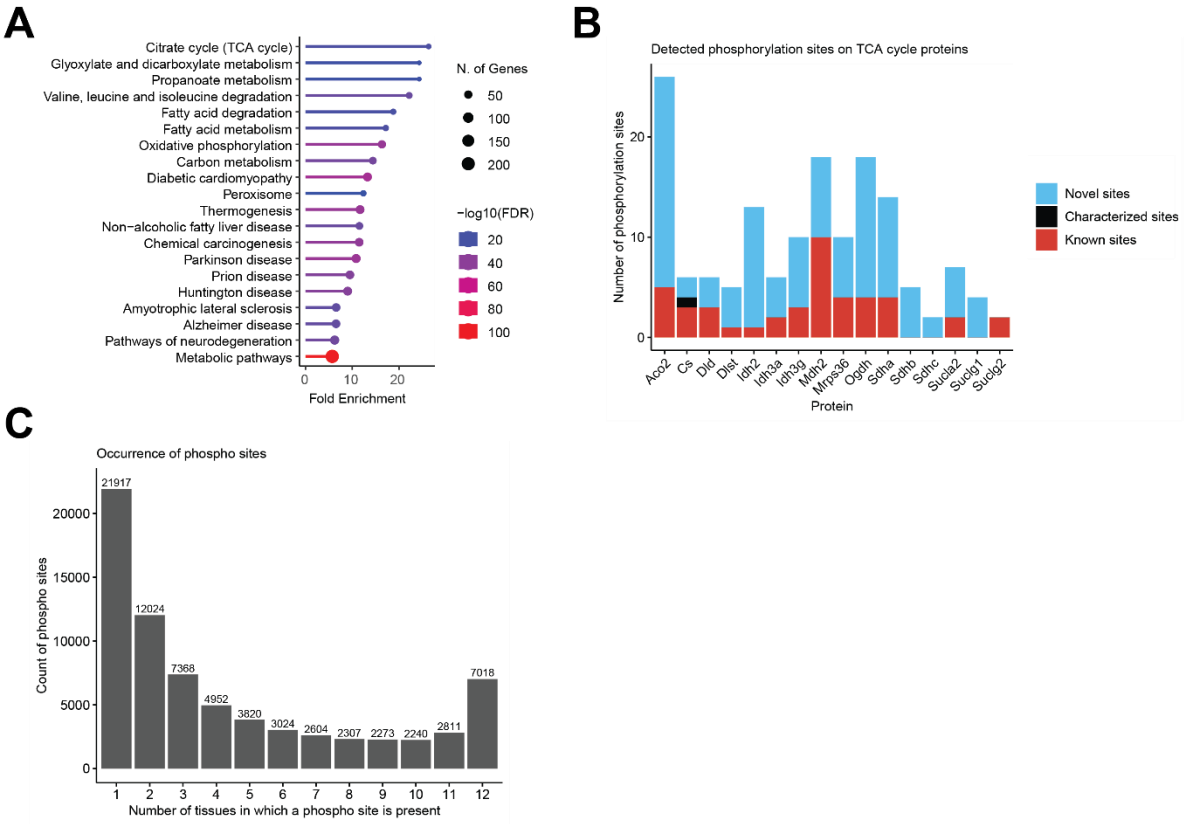

Figure 2.15. TCA cycle phosphorylation events and occurrence of phosphorylation sites across tissues. (A) GSEA using the KEGG pathway database, FDR 0.05, minimal pathway size 2, through the ShinyGO app.[68] Input: Proteins whose phosphorylation sites showed an absolute z-score above 2. (B) Number of phosphorylation sites detected on tricarboxylic cycle proteins (subset based on MitoCarta 3.0 pathways). Sites that occur in multiple tissues are summed. Characterized sites according to PhosphoSitePlus are also known sites and are only counted in the characterized sites category. (C) Bar plot indicating the number of phosphorylation sites that occur in a specific number of tissues. Image Description

(**Figure 2.15C**) as well as in the mitochondrial subset of our data (**Figure 2.14C**), indicating that tissue-specific phosphorylation patterns exist within the murine mitochondrial phosphoproteome. To investigate this in more detail, we calculated the intersections of all mitochondrial phosphorylation sites across tissues and found the liver to harbor the most unique phosphorylation sites (**Figure 2.14D**). Moreover, these unique liver phosphorylation sites showed the largest proportion of mitochondrial matrix proteins, suggesting that phosphoregulation of proteins in this sub-compartment is particularly important in this tissue. We identified carbamoyl phosphate synthetase I (CPS1) to harbor the most phosphorylation sites in this group (**Figure 2.14E**).

Notably, 40 phosphorylation sites were detected on CPS1, which is the first and rate-limiting enzyme of the urea cycle. None of these phosphorylation sites have been functionally characterized, and 22 of these sites have not been reported previously (**Figure 2.14F**), underscoring the significant discovery potential of our dataset. Deficiency of human CPS1 results in hyperammonemia ranging from neonatally lethal to environmentally induced adult-onset disease in affected individuals.[69] The phosphorylation site S913, located within the L2 integrating domain, overlaps with a known severe disease-causing missense variant in CPS1 (p.S913L)[70], indicating the potential regulatory significance of these sites in CPS1.

Our dataset also provides enhanced phosphotyrosine coverage without specific enrichment methods (**Figure 2.7B**), revealing novel sites that may be functionally relevant. N-acetyl-L-glutamate (NAG) is an essential allosteric activator of CPS1. By binding to the C-terminal L4 domain of CPS1, NAG induces long-range conformational changes, impacting distant catalytic activities.[67] We identify Y1450 as a conserved phosphorylated residue in this domain, which is located proximal to the key R1453 residue that is required for NAG binding and is the location of

two patient variants (c.4357C>T and c.4358G>A resulting in p.R1453W and p.R1453Q, respectively)[71] (**Figure 2.14G**). Previous work suggested that a distinct post-translational modification of Y1450 (nitration) may impact CPS1 activity by obstructing NAG binding.[72] Structural modeling demonstrates shifts in this Y1450-R1453-containing helix of CPS1 upon NAG binding, highlighting the confirmational change of the R1453 residue and its proximity to the Y1450 residue (**Figure 2.14H**).

Most dynamic phosphoregulation in brain tissue

The brain is characterized by a diverse array of expressed kinases and phosphatases and its proteins have a higher tendency to be phosphorylated as compared to other tissues.[49] Our data substantiates this observation, as brain tissue showed the highest prevalence of unique phosphorylation sites (*i.e.*, sites defined as exclusive to a singular tissue, **Figure 2.16A**). Moreover, we investigated all pan-tissue expression of phosphorylation sites (sites detected across all twelve tissues) by a z-score analysis, enabling comparison of each phosphosite to the average organismal values, and found brain tissue to exhibit the strongest abundance changes of phosphorylation sites (**Figure 2.17A**). We then applied the same z-score-based principle on the entire dataset (including imputed values for undetected sites in some tissues) to impartially identify sites undergoing significant changes and observed a normal distribution in z-score patterns for all tissues, except the brain, which displayed a pronounced enrichment of phosphorylation sites with elevated z-scores (**Figure 2.17B**). By applying the same absolute z-score threshold of 2, we found 9,541 phosphorylation site outliers in the brain (**Figure 2.16B**), accounting for 23% of all outliers across all twelve tissues. Further prioritizing these brain sites based on z-score ranking within the mitochondrial phosphoproteome, we identified serine 298 of Optic atrophy protein 1 (OPA1) as a top hit

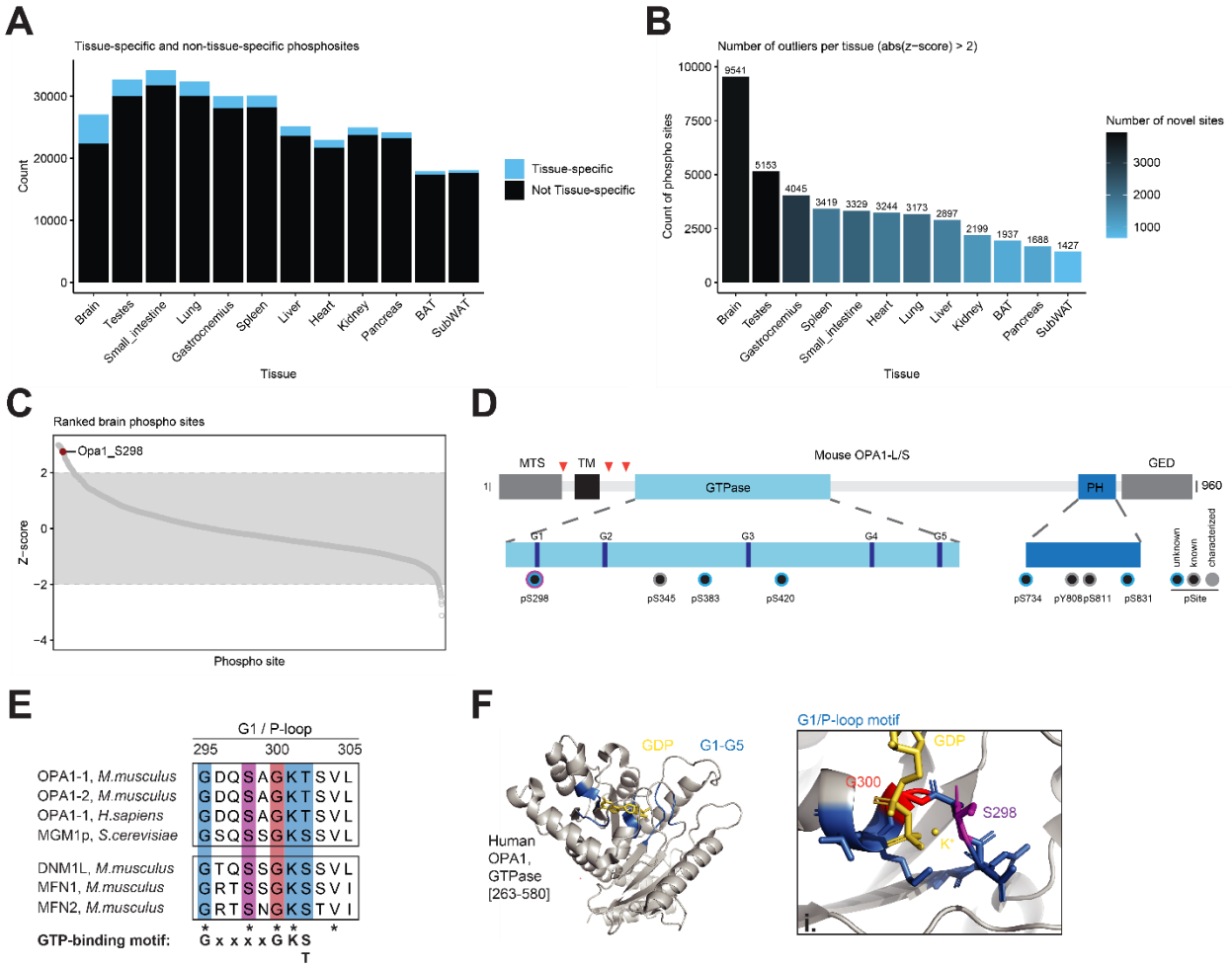

Figure 2.16. Unique phosphorylation sites in brain tissue. (A) Bar plot indicating all sites (black) that were detected per tissue, highlighting the sites that are unique in each tissue (blue). (B) Number of outliers per tissue based on z-score analysis with an absolute z-score cut-off of 2. (C) Ranked mitochondrial phosphorylation sites in brain based on z-score, highlighting the OPA1 top site. (D) Protein domain representation of mouse OPA1, including mitochondrial transition signal (MTS), transmembrane domain (TM), GTPase, PH, and GED domain. Red triangles represent sites of proteolytic cleavage generating the long and short proteo-forms of OPA1. Note, that all identified phosphosites are common to all OPA1 isoforms. Below, details of GTPase and PH domain with indicated phosphorylation sites (pS/Y) found in brain tissue (See **Figure 2.17C** for presence of OPA1 phosphorylation sites in other tissues).(E) Sequence alignment of first GTP-binding site/P-loop among mouse OPA1 isoforms and homologs in human and yeast, as well as related dynamin family members, DNM1L, MFN1, and MFN2, in mouse. Conserved residues are marked with *, including serine 298 in mouse OPA1. (F) Structure modeling of human OPA1 GTPase domain [263-580] based on RCSB PDB structure 6JTG[73]. Left, Modeling of the entire domain with highlighted GTP-binding domains (G1-5) in blue and bound GDP (yellow). Right, Detail (i.) of the G1/P-loop domain containing the discussed S298 (magenta) and the nucleotide-binding G300 (red). K+ (yellow) stabilizes the vicinity of the nucleotide. Image Description

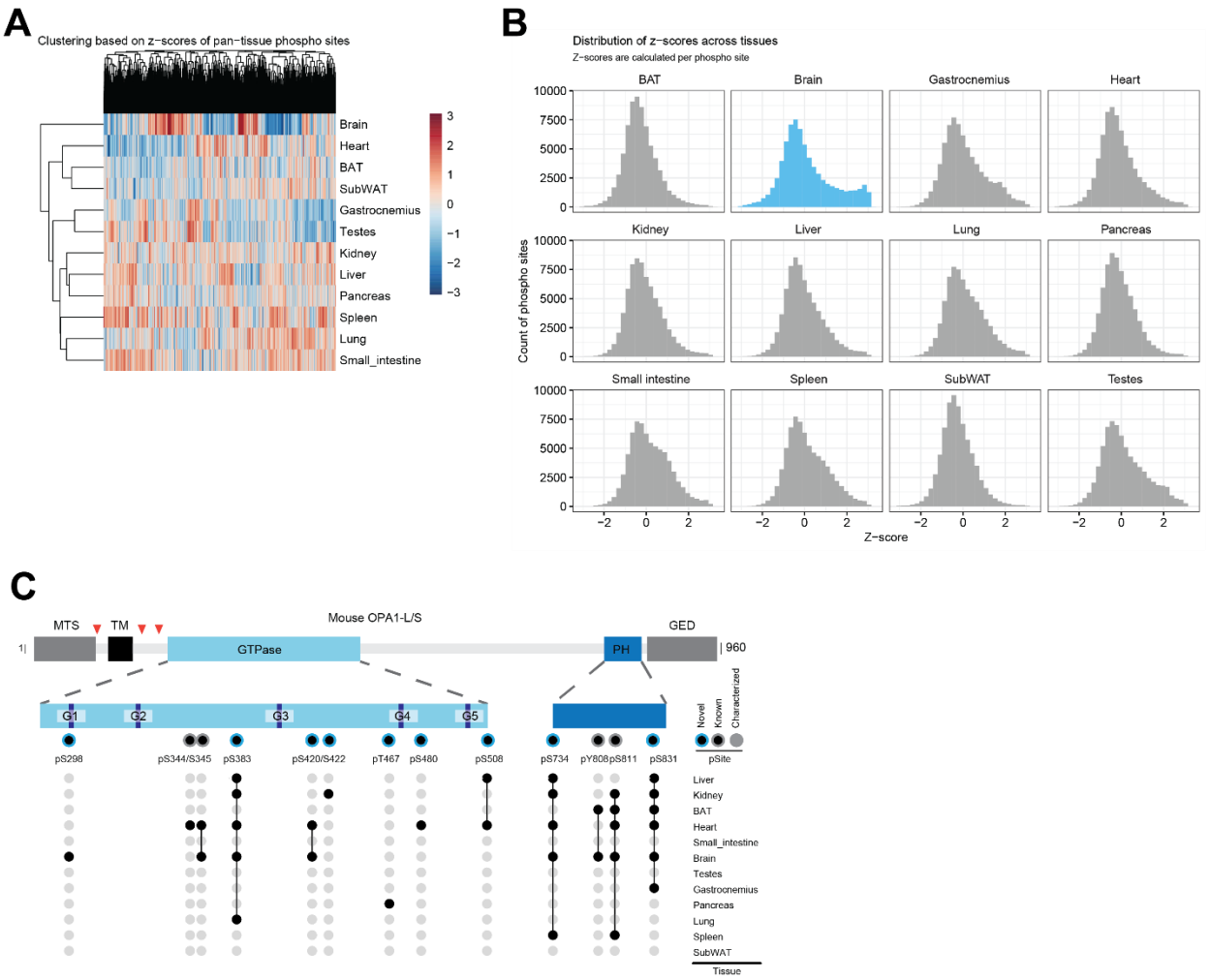

Figure 2.17. Z-score based outlier analysis and phosphorylation sites on Opa1. (A) Heatmap of z-scores calculated across all tissues. Only sites that occur in all tissues were considered. (B) Distribution of phosphorylation site z-scores in each tissue, using the entire imputed dataset. (C) Protein domain representation of mouse OPA1 as shown in **Figure 2.16D** with additional phosphorylation sites not observed in brain tissue. The occurrence of OPA1 phosphorylation site within all twelve datasets is shown below. Image Description

(**Figure 2.16C**).

OPA1 facilitates multiple functions in mitochondria including membrane fusion, cristae biogenesis, mitochondrial DNA maintenance, and respiration.[74] Importantly, mutations in OPA1 are the most common cause of dominant optic atrophy (DOA). In our data, we detected multiple phosphorylation sites of OPA1 across the twelve tissues (**Figure 2.17C**), none of which are functionally characterized as documented in PhosphoSitePlus.[53] These sites are localized within the GTPase and PH domain of OPA1, which are also common areas for missense variants in patients with DOA.[75] Additionally, many of the detected phosphorylation sites have not been previously reported in mouse, among which is the novel S298 site within the first GTP-binding domain (G1/P-loop) of the GTPase (**Figure 2.16D**).The consensus motif GxxxxGKS/T of G1/P-loop is highly conserved across species and other proteins of the dynamin superfamily (**Figure 2.16E**). Further, the Q and S residue within this motif are important for the assembly stimulated GTPase activity as shown for dynamin.[76] Notably, S298 has been identified as a site of missense mutations in DOA patients (c.893G>A, p.S298N[77] ; c.892A>G, p.S298G[78]), suggesting its importance for OPA1 function, too. Furthermore, a S298A mutation reduces GTPase activity and dimerization *in vitro*[73]; the S298N mutation abolishes respiratory growth in *S.cerevisiae* with mitochondrial DNA depletion and altered mitochondrial morphology.[79] Our structural modelling of the OPA1's GTPase domain shows that S298 is close to the nucleotide-interacting G300 of the G1/P-loop and the nucleotide, suggesting that phosphorylation may affect nucleotide binding or hydrolysis (**Figure 2.16F**). Taken together, current evidence clearly underscores the functional importance of S298 for OPA1 activity, but the effect and relevance of phosphorylation we have identified on S298 requires further study.

Discussion

Here we demonstrated that the Orbitrap Astral mass spectrometer provides a platform for fast and deep phosphoproteome analysis. Key to empowering this application is the ability of the system's high MS/MS duty cycle to enable DIA data collection with narrow isolation windows (2 Th). This capability, combined with high mass accuracy, mass resolution, and sensitive ion detection of the Astral analyzer allow confident phosphosite localization from low abundance peptides as evinced by the performance of the system for low overall phosphopeptide loads. Altogether this approach allowed for the collection of the deepest mouse phosphoproteomic atlas to date following only a half day of data collection. Nonetheless, the uniquely large mouse phosphoproteome atlas defined here provides an opportunity to explore the sequence and structural context of kinase activity.

From the phosphoproteome analysis we discovered that, despite its bacterial origins, over fifty percent of mitochondrial proteins carry at least one site of phosphorylation site. An interesting and potentially important example is the CPS1 protein on which we detected 22 novel sites. Notably, one of these sites, Y1450, is proximal to the NAG binding domain and near the location of known patient variants – suggesting potential clinical relevance.

As exemplified by the above example, this technology permitted detection of thousands of pY sites without the typical pY-specific antibody enrichments. Doubtless the sensitivity and depth afforded by the Orbitrap Astral instrument can allow direct access to these most critical, dynamic, and low-abundance phosphorylation events. Finally, we note that while all tissues exhibit a subset of unique phosphorylation events, the brain is distinguished and contains nearly 10,000 unique phosphorylation sites.

Protein phosphorylation analysis presents many challenges for mass spectrometric analysis; however, we demonstrate here that the Orbitrap Astral resolves many of these limitations and permits fast and deep phosphoproteome analysis. We suppose that these performance characteristics will be translated to the analysis of other post-translational modifications including acetylation, glycosylation, methylation, ubiquitination, etc. Furthermore, we expect use of extensive fractionation, more tissues, and multiple proteases would allow detection of even more phosphorylation sites and increase tissue-to-tissue phosphorylation site reproducibility.

Methods

HEK293T Cell Preparation. HEK293T cells (ATCC CRL-3216) were cultured in Dulbecco's modified Eagle's medium (DMEM) (Gibco, 11995-065) with 1% penicillin/streptomycin (Thermo Fisher Scientific, 15-140-122) and 10% fetal bovine serum (FBS) (HyClone, 89133-098) at 37°C and 5% CO2. Cell line authentication was performed by the commercial distributor. Culture media was replaced every 24 hours. Cells were expanded to the appropriate cell number, detached from tissue culture plate with 0.05% trypsin-EDTA (Gibco, 25300062), washed once with phosphate buffered saline (PBS) (Gibco, 02-0119-0500), cell number determined, and 20 x 10^7 cells pelleted. Cell pellet was immediately stored at -80°C until use. The cells were low passage number and tested negative for mycoplasma contamination. Frozen cell pellets were resuspended in 5.4M guanidine hydrocholoride (from Sigma Life Science, 8M, pH 8.5, G7294-100mL) in 100 mM Tris, pH 8 (Invitrogen, 1 M Tris pH 8.0, 0.2 µm filtered, AM9856) via vortexing, followed by heating in a sand bath for 5 minutes at 105℃ prior to brief (10-15s) sonication with a probe sonicator. The sample was diluted with the guanidine buffer above to give a ~1.5 mg/mL estimated protein concentration via NanoDrop (Thermo Scientific) prior to beginning digestion.

EGF-stimulated HeLa Cell Preparation. HeLa cells (ATCC CCL-2) were cultured in Dulbecco's Modified Essential Medium (DMEM) (Gibco, 11995-065) with 10% FBS (HyClone, 89133-098) and 1% penicillin/streptomycin (Thermo Fisher Scientific, 15-140-122) in 37°C/5% CO2 incubator. Cell line authentication was performed by the commercial distributor. Cells were resuspended in PBS (Gibco, 02-0119-0500) with or without 100 ng/mL human Endothelial Growth Factor (Thermo, AF-100-15) for 15 minutes at room temperature. Cells were gently resuspended

every 5 minutes during incubation. After incubation, cells were washed twice with PBS and stored in -80°C until use. Protein extraction was performed as described above for the HEK293T samples. Protein concentrations were estimated via protein BCA (Pierce, 23235). Three biological replicates each were prepared for the control and EGF treatment groups.

Yeast Cell Preparation. *Saccharomyces cerevisiae S288C-derivative strain, BY4741* (SRD GmbH, Y00000) was cultured in triplicate for ~5 generations into log phase in rich YPD medium at 30°C. Authentication was performed by the commercial distributor. Cells were centrifuged at 2000g, 3 min, rinsed in water, transferred to smaller tubes and centrifuged at 1600g, 3min, prior to snap-freezing in liquid nitrogen. Frozen cell pellets were resuspended in 8M urea Sigma-Aldrich, U5378) in 100 mM Tris, pH 8 (Invitrogen, 1 M Tris pH 8.0, 0.2 µm filtered, AM9856) and vortexed with glass beads (425-600 µm, Sigma-Aldrich, G8772-500G) to lyse (2 minutes of total vortexing with 30s vortexing followed by 30s on ice). Protein concentrations were estimated via protein BCA (Pierce, 23235).

Mouse Tissue Preparation. All experiments were performed in accordance with the National Institute of Health Guide for the Care and Use of Laboratory Animals and were approved by the Animal Care and Use Committee at the University of Wisconsin-Madison. Mice were kept on a 12-h light–dark cycle at 23°C and within a humidity range between 30% and 50%. Mice were fed standard chow diet (Teklad #2018). Six-week-old male C57BL6/J mice (n=3) were euthanized by cervical dislocation and tissues were immediately collected, and flash frozen in liquid nitrogen. A total of twelve tissues (pancreas, small intestine, spleen, liver, kidney, testes, heart, lung, subcutaneous white adipose tissue (WAT), brown adipose tissue (BAT), gastrocnemius, and brain) were collected. All tissues were stored at -80°C prior to cryo-pulverization. For each tissue,

samples from three mice were pulverized together into a fine power under liquid nitrogen (*i.e.,* one biological replicate per tissue). For each pulverized tissue, ~60mg frozen wet weight was resuspended in 4 mL of 5.4M guanidine hydrocholoride (Sigma Life Science, 8M, pH 8.5, G7294-100mL) in 100 mM Tris, pH 8 (Invitrogen, 1 M Tris pH 8.0, 0.2 µm filtered, AM9856) with Pierce™ Phosphatase Inhibitor Mini Tablets (A32957) with one tablet/10mL. Tissue samples were vortexed and sonicated for 20 min in a bath sonicator (chilled to 4°C) to homogenize. A probe sonicator was used briefly to sonicate samples on ice as needed. The protein concentration was estimated via protein BCA (Pierce, 23235) and samples were diluted with the guanidine hydrochloride buffer above to give a protein concentration of ~2mg/mL prior to digestion.

Protein digestion. After extracting proteins from human cells, yeast cells, or mouse tissues, methanol (Optima LC/MS grade, Fisher Scientific) was added to 90% (v/v) to precipitate protein and samples were vortexed prior to centrifugation at 4000g for 15 minutes. The supernatant was removed, and the pellet was resuspended in 8M urea (Sigma-Aldrich, U5378), 100mM Tris (Invitrogen, 1 M Tris pH 8.0, 0.2 µm filtered, AM9856), 10mM TCEP (Sigma-Aldrich, C4706-2G), 40mM 2-chloroacetamide (Sigma Aldrich, ≥98%, C0267-100G) pH 8 at ~1.5mg protein/mL. Lysyl Endopeptidase (LysC, 100369-826, VWR) was added at a ratio of 1:50 enzyme:protein and gently rocked at ambient temperature for 4 hours, followed by the dilution of the solution to 2M urea with 100mM Tris, pH 8 (Invitrogen, 1 M Tris pH 8.0, 0.2 µm filtered, AM9856). Promega Sequencing Grade Modified Trypsin (V5113) was added at a ratio of 1:50 enzyme:protein and incubated overnight. Following overnight digestion, the solution was acidified to <pH 2 with 10% trifluoroacetic acid (Sigma-Aldrich, HPLC grade, >99.9%) to quench the digestion. The sample was then centrifuged at 4000g for 10 minutes to remove particulate matter prior to desalting with

a Strata-X 33 μm polymeric reversed phase SPE cartridge. Peptides were dried via a SpeedVac (Thermo Scientific) and stored at -80°C until phosphopeptide enrichment or, in the case of the mouse proteomics experiments, until fractionation.

Phosphopeptide enrichment. Phosphopeptides were enriched from digested peptides using MagReSyn Ti-IMAC HP beads (ReSyn Biosciences, MR-THP005). A volume of 100μL beads were used per 1mg of peptides. Input peptide masses of 2-3 mg were utilized for human and yeast enrichments, and input peptide masses ranging from ~0.3 mg to 1.2 mg were utilized for the different mouse tissues. Beads were washed three times with 1 mL 80% acetonitrile (Optima LC-MS grade, Fisher Scientific)/6% trifluoracetic acid (Sigma-Aldrich, HPLC grade, >99.9%) prior to resuspending the sample in 1 mL 80% acetonitrile (Optima LC-MS grade, Fisher Scientific)/6% trifluoracetic acid (Sigma-Aldrich, HPLC grade, >99.9%) and vortexing the sample with the beads for 1 hour. After the 1-hour of vortexing, the beads were washed three times with 1mL 80% acetonitrile (Optima LC-MS grade, Fisher Scientific) /6% trifluoracetic acid (Sigma-Aldrich, HPLC grade, >99.9%), once with 1 mL 80% acetonitrile, once with 1 mL 80% acetonitrile (Optima LC-MS grade, Fisher Scientific) /0.5 M glycolic acid (Sigma-Aldrich, 99%, 124737-500G), and three times with 1 mL 80% acetonitrile (Optima LC-MS grade, Fisher Scientific). The phosphopeptides were eluted from the beads with the addition of 300 μL 50% acetonitrile (Optima LC-MS grade, Fisher Scientific)/1% ammonium hydroxide (28% in H_2O, ≥99.99% trace metals basis, Sigma-Aldrich), followed by a second elution with another 300 μL 50% acetonitrile/1% ammonium hydroxide. The samples were acidified via addition of 15 μL 10% trifluoroacetic acid (Sigma-Aldrich, HPLC grade, >99.9%). The samples were then dried down in a SpeedVac (Thermo Scientific) prior to being resuspended in 0.2% trifluoroacetic acid (Sigma-Aldrich, HPLC

grade, >99.9%) and desalted as described above. Desalted phosphopeptide samples were dried and resuspended in 0.1% formic acid (Fisher Scientific, LC-MS grade). The phosphopeptide concentration was estimated via NanoDrop (Thermo Scientific). The HEK293T phosphopeptides were pooled to generate a sample for method evaluation. Each mouse phosphopeptide sample was fractioned as described below.

High-pH Peptide Fractionation. High-pH fractionation of peptides was performed on an Agilent 1260 Infinity BioInert LC with an automated fraction collector. A 20-minute method was performed on a Waters XBridge, Peptide BEH C18, 3.5 µm, 130 Å, 4.6mm x 150 mm column with a flow rate of 0.8 mL/min. Mobile phase A and B were 10 mM ammonium formate (Sigma-Aldrich, >99/0%, LC-MS grade, 70221-100GF), pH 10 and 20% 10mM ammonium formate (pH 10)/80% methanol (Optima LC/MS grade, Fisher Scientific), respectively. The gradient went from 0 to 35%B from 0-2 minutes, 35 to 75%B from 2-8 minutes, 75% to 100%B from 8-13 minutes, followed by washing at 100%B from 13-15 minutes and equilibration at 0%B from 15-20 minutes. UV absorbance at 210 and 280 nm was recorded. For HEK293T spectral library generation and mouse proteomics samples, 16 fractions were collected from 5 to 18 minutes and concatenated into the final fractions by combining fraction 1 and 9, fraction 2 and 10, etc., resulting in 8 final fractions. For mouse phosphoproteomics samples, 8 fractions were collected from 5 to 18 minutes and concatenated into the final fractions by combining fraction 1 and 5, fraction 2 and 6, etc. Samples were dried down in a SpeedVac (Thermo Scientific) prior to being resuspended in 0.1% formic acid (Fisher Scientific, LC-MS grade) for LC-MS analysis.

Synthetic Phosphopeptide Dilution Series Preparation. Five sets of phosphopeptide standards were acquired: SpikeMix PTM-Kit 52 (JPT, SPT-PTM-POOL-Yphospho-1), SpikeMix

PTM-Kit 54 (JPT, SPT-PTM-POOL-STphospho-1), MS PhosphoMix 1 (Sigma, MSP1L-1VL), MS PhosphoMix 2 (Sigma, MSP2L-1VL), and MS PhosphoMix 3 (Sigma, MSP3L-1VL). The standards were reconstituted in 0.2% formic acid/20% acetonitrile/80% water via vortexing. The standards were pooled into an equimolar mixture of the 225 total phosphopeptide standards. The pooled equimolar mixture was then diluted and mixed with the yeast phosphopeptide sample to construct a dilution series comprised of five points of four-fold dilutions starting at 10000 amol, resulting in total quantities loaded onto the column of 10000, 2500, 625, 156.25, and 39.0625 amol per phosphopeptide standard along with a constant yeast phosphopeptide load of 250ng.

LC-MS Operation for Phosphoproteomics with Orbitrap Astral Analysis. Nanoflow capillary columns (75µm I.D., 360µm O.D.) with pulled nanoESI emitters were packed to 40 cm at high pressures with C18 1.7 µm diameter, 130 Å pore size BEH C18 particles (Waters) as previously described.[39] A slurry of C18 particles dissolved in chloroform was loaded into a custom-built packing setup with a high pressure pneumatic pump (Haskel) and ultrahigh-pressure capillary fittings (HiP) and packed into the column while slowly ramping up to 30,000 psi, holding at 30,000 psi for 4 hours, and then allowed to depressurize slowly. Samples were analyzed with a Vanquish Neo UHPLC (Thermo Scientific) coupled to an Orbitrap Astral mass spectrometer (Thermo Scientific) using a NanoSpray Flex source (Thermo Scientific). A source voltage of 2000 V was used for all experiments. Mobile phase A and B were 0.1% formic in water (Fisher Scientific, Optima LC-MS grade) and 0.1% formic acid/80% acetonitrile (Fisher Scientific, Optima LC-MS grade), respectively. The column was heated to 50°C with the Column Oven PRSO-V2 (Sonation Lab Solutions) and the flow rate was set to 400 nL/min at the start of the method to decrease delay time and turned to 300 nL/min at the start of the active gradient. Initial conditions of 2%B were

ramped to 14% from 0 to 5 minutes. The active gradient was generally set to 14% to 54%B with curve type 6 beginning at 5.2 minutes, with the exact %B settings adjusted for each active gradient length to evenly distribute peptide signal across the gradient. The column was washed for 5 minutes at 100%B and 400 nL/min at the end of the gradient, followed by fast equilibration on the Vanquish Neo LC with an upper pressure limit of 1100 bar.

For initial DDA experiments on the Orbitrap Astral MS, MS^1 spectra were collected in the Orbitrap every 0.6 s at a resolving power of 240,000 at *m/z* 200 over *m/z* 350-1350 with a normalized AGC target of 300% (3e6 charges) and a maximum injection time of 10ms. The MIPS filter was applied with Peptide mode and 'Relax Restrictions when too few Precursors are Found' set to True. Precursors were filtered to charges states 2-6. A Dynamic Exclusion filter was applied with 10s duration and 10ppm low and high mass tolerance and exclude isotopes set to True. An intensity filter was applied with a minimum precursor intensity of 5000 required for selection. MS^2 scans were collected in the Astral mass analyzer with an isolation window of 0.7 *m/z*, normalized collision energy of 27, a scan range of 150-2000 *m/z*, an AGC target of 100% (1e4 charges), and a maximum injection time of 10ms.

For DIA experiments on the Orbitrap Astral MS, MS^1 spectra were collected in the Orbitrap every 0.6s at a resolving power of 240,000 at *m/z* 200 over *m/z* 380-980. The MS^1 normalized AGC target was set to 300% (3e6 charges) with a maximum injection time of 10ms. DIA MS^2 scans were acquired in the Astral analyzer over a 380-980 *m/z* range with a normalized AGC target of 500% (5e4 charges) and a maximum injection time of 3.5ms and an HCD collision energy setting of 27% and a default charge state of +2. Window placement optimization was turned on. Isolation widths of 2 Th and active gradient lengths of 30 minutes were used with HEK293T

fraction analysis for HEK293T spectral library generation. Isolation bin widths of 2 Th and 4 Th with 1 Th overlap were compared for HEK293T phosphoproteomics method evaluation. Isolation widths of 2 Th were used for mouse phosphoproteomics experiments. Comparisons of different DIA *m/z* ranges and AGC targets are shown in **Figure 2.4D-G**.

Eight HEK293T phosphopeptide fractions were injected once for library generation. For HEK293T shotgun experiments, one injection replicate per loading mass was performed for the loading mass experiment. Three injection replicates per method were performed for the gradient length and isolation width experiments. For the phosphopeptide standard analysis, each concentration point was analyzed with three injection replicates. For the EGF treatment vs control HeLa phosphopeptide and proteomics experiments, each of the three biological replicates per condition was injected once. For the HeLa phosphopeptide method evaluation experiments, two injection duplicates were performed per method. For the mouse phosphopeptide analysis, each tissue sample was fractioned into 4 fractions and injected once.

LC-MS Operation for Benchmarking on the Orbitrap Ascend. Nanoflow capillary columns (75 µm I.D., 360 µm O.D.) with pulled nanoESI emitters were packed to 40cm at high pressures with C18 1.7 µm diameter, 130 Å pore size BEH C18 particles (Waters) as previously described.[39] See above for column packing description. Samples were analyzed with a Vanquish Neo UHPLC (Thermo Scientific) coupled to an Orbitrap Ascend mass spectrometer (Thermo Scientific) using a NanoSpray Flex source (Thermo Scientific) incorporating a homebuilt column heating compartment. The column was heated to 50°C. The flow rate was set to 400 nL/min at the start of the method to decrease delay time and turned to 30 0nL/min at the start of the active gradient. Initial conditions of 2%B were ramped to 14% from 0 to 5 minutes. The active gradient

was generally set to 14% to 54%B with curve type 6 beginning at 5.2 minutes, with the exact %B settings adjusted for each active gradient length to evenly distribute peptide signal across the gradient. The column was washed for 5 minutes at 100%B and 400 nL/min at the end of the gradient, followed by fast equilibration on the Vanquish Neo LC with an upper pressure limit of 1100 bar.

DIA experiments on the Orbitrap Ascend utilized similar method settings to those previously described by Bekker-Jensen *et al.*[41], but with the same *m/z* range as utilized on the Orbitrap Astral to enable direct comparisons. Briefly, MS^1 spectra were collected in the Orbitrap at a resolving power of 60,000 at *m/z* 200 over *m/z* 380-980 with an AGC target of 250% and maximum injection time of 123 ms. DIA MS^2 scans were collected in the Orbitrap with 15,000 at 200 *m/z* resolving power, a scan range of 150-2000 *m/z*, an AGC target of 200%, a maximum injection time of 27 ms, and a collision energy of 30%. A *m/z* range from 380-980 *m/z* was iterated through with 14 Th isolation windows with 1 Th overlap.

For the phosphopeptide standard analysis, each concentration point was analyzed with three injection replicates. For the EGF treatment vs control HeLa phosphopeptide and proteomics experiments, each of the three biological replicates per condition was injected once.

LC-MS Operation for Proteomics with Orbitrap Eclipse Analysis. The chromatography setup described above for the Orbitrap Ascend analysis was utilized for the Orbitrap Eclipse experiments. A source voltage of 2000 V was used for all experiments. Mobile phase A and B were 0.2% formic acid (Optima LC-MS grade, Fisher Scientific) in water (Optima LC-MS grade, Fisher Scientific) and 0.2% formic acid/80% acetonitrile (Optima LC-MS grade, Fisher Scientific), respectively. The column was heated to 50°C and a flow rate of 300 nL/min was

used. Initial conditions of 0%B were ramped to 6%B from 0 to 1 minute. The active gradient was set to 6% to 52%B with curve type 6 from 1 to 73 minutes. The column was washed for 5 minutes at 100%B, followed by fast equilibration on the Vanquish Neo LC with an upper pressure limit of 1150 bar.

For DDA experiments on the Orbitrap Eclipse MS, MS^1 spectra were collected in the Orbitrap analyzer every 0.6s at a resolving power of 240,000 at *m/z* 200 over *m/z* 300-1350 with a normalized AGC target of 250% and a maximum injection time of 50ms. The MIPS filter was applied with Peptide mode and 'Relax Restrictions when too few Precursors are Found' set to True. Precursors were filtered to charges states 2-5. A Dynamic Exclusion filter was applied with 10s duration and 25 ppm low and high mass tolerance and exclude isotopes set to True. MS^2 scans were collected in the ion trap with an isolation window of 0.5 *m/z*, normalized collision energy of 25, a scan range of 150-1350 *m/z*, an AGC target of 300%, and a maximum injection time of 14ms.

Each mouse tissue was fractioned into 8 fractions, each of which was injected once for DDA proteomics.

DDA Data Processing. Proteomics DDA data was processed in MaxQuant 2.4.9.0 with default parameters.[40] Phosphoproteomics DDA data was processed with the same version of MaxQuant with default parameters and Phospho (STY) enabled as a variable modification. The 'Phospho (STY)Site.txt' was used for analysis with a filter for localization probabilities greater than or equal to 0.75.

DIA Data Processing. Phosphoproteomics DIA data was processed in Spectronaut version 17.6.230428.55965 or 18.6.231227. A HEK293T spectral library was generated in Spectronaut (searching against the human proteome database (Swiss-Prot and TrEMBL) downloaded from

UniProt on 15-Jan-2023) from the eight HEK293T phosphopeptide fractions analyzed with a 30-minute active gradient 2 Th DIA method and the library was used to search the HEK293T phosphoproteomics method evaluation experiments shown in **Figure 2.2**. Note that the entrapment search in **Figure 2.2G** and the phosphoproline variable modification search in **Figure 2.2H** were performed as a library-free search. Mouse phosphoproteomics data was searched with the directDIA mode. Default Spectronaut search parameters were used with a variable Phospho (STY) modification included. The PTMSiteReport file was used for analysis of phosphorylation sites. To count the number of unique phosphorylation sites detected within an experiment, the PTMSiteReport was filtered for rows with 'Phospho (STY)' in the 'PTM.ModificationTitle' column and values greater than or equal to 0.75 in the 'PTM.SiteProbability' column. The data table was then sorted according to 'PTM.CollapseKey' and filtered for unique values in the 'PTM.Group'. At this stage, there can be rows with the same values of 'PTM.CollapseKey' that are not filtered by 'PTM.Group' due to PTM grouping differing across multiple files, so the dataset was filtered for unique values of 'PTM.CollapseKey' to determine the total number of unique phosphorylation sites measured in a single file. Mouse phosphoproteomics data was searched against the mouse proteome downloaded from UniProt from Swiss-Prot and TrEMBL (downloaded on 25-Aug-2023). For the mitochondrial biology investigation, a separate search was performed with just the Swiss-Prot database to facilitate the analysis.

HEK293T phosphoproteomics DIA data was also processed in Proteome Discoverer 3.1.0.638 with CHIMERYS. Numbers of localized phosphosites were reported from the Modification Sites table after filtering for rows with 'Phospho' in the Modification Name column. These searches were performed in a developmental version of Proteome Discoverer 3.1 to show

that the phosphoproteomic depth achieved with our methods is not restricted to Spectronaut searches. However, this developmental version does not yet incorporate quantitative values for phosphorylation sites detected with DIA methods. Furthermore, from internal conversations, we know the results of this development build should be treated as preliminary and that the algorithms being utilized are still in active development. Consequently, we chose to perform our analyses using Spectronaut, which provides identification, localization, and quantification for DIA phosphoproteomics.

Localization Error Rate Calculation. To calculate the localization error rate, a reference sheet of all the possible precursors that could be detected from the phosphopeptide standards was constructed with the phosphorylation sites specified. Precursors with sequences that could also arise from the yeast proteome (based on an *in silico* digest) were removed from consideration. Then, the error rate was calculated as $\text{Error Rate } (\%) = \frac{\text{False Sites}}{\text{False Sites} + \text{True Sites}} \times 100$, where 'False Sites' is the number of phosphopeptide standard precursors detected with phosphorylation states not indicated in the reference sheet, and 'True Sites' is the number of phosphopeptide standard precursors detected with phosphorylation states indicated in the reference sheet. This calculation was performed for each phosphopeptide standard raw data file, and the average error rates are reported as a function of the localization probability cutoff.

Phosphopeptide Standard Quantification Calculations. The phosphopeptide standard reference sheet described above was modified so that each row corresponds to a phosphorylation site, so that the 'PTM.Quantity' value in the PTM Site Report could be utilized for quantification assessment. As described above, phosphopeptides with sequences that could arise from the yeast proteome were removed from consideration. Linear regression and R^2 calculations were performed

using the ‘pearsonr’ and ‘linregress’ functions within the ‘scipy.stats’ module for phosphosites detected across at least three concentration points.[80]

EGF-stimulated HeLa Differential Expression and Pathway Analysis. Differential expression analysis for the EGF treatment experiment was performed on all phosphosites that were detected across triplicates in either the EGF-treatment or control group. Missing value imputation was performed in Perseus[81] assuming a normal distribution with ‘width = 0.3’ and ‘down shift = 1.8’. Enriched phosphosites were defined as those with an absolute fold change of more than 2 and a p-value less than 0.05 using a two-sided t-test without a multiple testing correction.

Pathway enrichment analysis was performed using Enrichr.[82] The enriched subset contains genes corresponding to phosphosites that are enriched as defined above, and the background was performed for the rest of genes with at least one detected phosphosite.

Sequence-based phosphosite clustering. Phosphosite sequences, along with their 5-amino acid flanking regions, were compared using the Blosum62 substitution matrix. The top 4 scores, representing a typical number of amino acids in kinase recognition motifs, were selected. These scores were subsequently averaged and transformed to a range between 0 (indicating identical sequences) and 1 (representing maximally different sequences). Subsequently, a similarity matrix between all sequence pairs was constructed and utilized to generate a 2D embedding space through t-SNE, implemented using the scikit-learn library.[83] Default parameters were used, except for metric="precomputed", init="random", n_iter=500, n_iter_without_progress=150, and random_state=42. Additionally, Perseus was employed for visualization purposes.[81]

Kinase Motif Enrichment. Initially, each phosphorylation site and its flanking region in

our dataset was used to search the kinase library website (https://kinase-library.phosphosite.org/site, accessed on February 22nd, 2024) for predictions of all 303 serine/threonine kinases for each site. For kinase motif enrichment analysis, only the top 15 kinases with regards to percentile score for each site were retained. To nominate tissue-enriched phosphorylation sites, we proceeded as follows: for each tissue-tissue comparison the log-transformed abundances of all shared phosphorylation sites were subtracted from each other. For selection of top phosphorylation sites per tissue two different approaches were applied: a z-score threshold of 2 was used, or the top and bottom 200 sites were selected. For the final analysis, the top and bottom 200 ranked sites were used. To calculate the frequency factor for each kinase the proportion of how often a kinase was predicted within the top phosphorylation sites versus the total of top phosphorylation sites was computed, divided by the same ratio for the unchanged sites, *i.e.*, not ranked among the top/bottom 200. A chi-squared test was calculated after applying Haldane's correction to the contingency table of the two proportions and the p-value was extracted.

Human variant information. Annotation of human variants in CPS1 and OPA1 were identified using The Human Gene Mutation Database at the Institute of Medical Genetics in Cardiff (HGMD) using the public site entries (retrieval date 2023-11-18 [84]).

Multiple Sequence Alignment. Clustal Omega[66] hosted at the EMBL-EBI webserver was used for protein sequence alignment using the following UniProt identifiers for CPS1: P08686, F7EZ, P07756, Q8C196, F1ML89, A0A8C0NKH0; for OPA1: P58281, P58281-2, O60313, P32266; and other dynamin family members: Q811U4, Q80U63, Q8K1M6 (retrieved from UniProt 2023-11-18).

Structural modelling prediction. The PyMOL Molecular Graphics System, Version 2.5.7

Schrödinger, LLC. was used for predicting structural models for CPS1 (RCSB PDB: 5DOT (Apo), 5DOU (NAG)[67]) and OPA1 (RCSB PDB structure 6JTG [73]). For modeling purposes, the human crystal structures (highly homologous to mouse) were used.

MS^1 Orbitrap Scan Settings. Because Astral MS^2 acquisition can be performed in parallel with the Orbitrap MS^1 scan, we selected the 240,000 resolving power at 200 *m/z* scan for high-resolution MS^1 scan acquisition. This scan takes just over 0.5 s[85], so we selected a 0.6 s cycle time (the instrument control software allows cycles times in 0.1 s increments) to maximize the Orbitrap MS^1 duty cycle.

MS^2 Maximum Inject Time. The timing sequence for Astral analysis has been previously described by Stewart *et al*.[31] Using a maximum injection time of 3.5 ms ensures that the ion accumulation time is close to maximally parallelized with other steps in the scan sequence. This setting enables a scan rate of ~200 Hz to be maintained throughout the runs. Below, the MS^1 and MS^2 cycle times are plotted for our 15-minute active gradient method with a 2 Th DIA window iterating over a 380-980 *m/z* range (a total of 300 windows). Cycle times of ~0.6 and ~1.5 s are observed for the MS^1 and MS^2 scans, respectively (**Figure 2.3E**). Maintaining short cycle times is essential as the columns employed in this study yield median chromatographic peak widths less than 8 seconds for our 15-minute method (as shown in **Figure 2.3B**).

DIA *m/z* Range. The *m/z* range iterated over during DIA analysis is a key method parameter. The size of this range limits the achievable cycle time (or the minimum isolation window width). In our experiments, we utilized a 380-980 *m/z* range, a common range used for DIA proteomics.[31,86,87] Enriched phosphopeptide samples could exhibit different *m/z* distributions than proteomics samples due to the mass shift from phosphorylation and increased missed

cleavages[28] so we compared the use of a 380-980 *m/z* range with a 480-1080 *m/z* range for the analysis of EGF-stimulated HeLa phosphopeptides. We observe that the use of the 480-1080 *m/z* range decreases the number of phosphosites by ~6% (**Figure 2.4D**), validating our use of a 380-980 *m/z* range in this study. However, we note that the optimal *m/z* range likely depends on the sample type being analyzed.

Astral AGC target. The AGC target for the Astral analyzer influences the number of charges that are targeted for each scan. We utilized a target of 5e4 (500%) for DIA phosphoproteomics. We compared this to a method using a target of 1e4 for analysis of the EGF-stimulated HeLa phosphopeptide samples. We note that there is no considerable difference in the depth achieved with these two methods (**Figure 2.4E**). We note that the MS^2 scan total ion current distributions are identical for both methods (**Figure 2.4F**). This observation is explained by looking at the MS^2 inject time distribution, with that both methods reach the maximum injection time setting of 3.5 ms for most scans (**Figure 2.4G**). The absence of an effect on phosphoproteomic depth then arises from the AGC target not being reached in either method for these particular analyses.

Author Contributions

NA and JF performed cell culture and collected HEK293T and HeLa cells. JH and AG collected mouse tissues. JR and APG performed yeast culture and collected yeast cells. TNA, AP, HS, CH, ED, and VZ developed the new instrument used in this study. NML, TMPC, TNA, AP, MLR, LRS, and ED performed mass spectrometry sample preparation and/or data acquisition. NML, PS, TMPC, and LRS performed mass spectrometry data processing. NML, PS, PF, CF, and AJS performed data analysis and assessment of the biological relevance of the study. NML, PS, PF, CF, AJS, ES, TNA, LRS, MSW, ED, VS, and JJC contributed to the figure content and design. NML, PS, PF, CF, AJS, ES, HS, DJP, and JCC wrote the manuscript. NML, PS, PF, TMPC, CF, AJS, ES, TNA, AP, HS, DJP, VZ, and JJC edited the manuscript. MSW, DJP, VZ, and JJC provided project supervision.

Competing Interests

JJC is a consultant for Thermo Fisher Scientific and on the scientific advisory board for Seer and 908 Devices. TNA, AP, HS, CH, ED, and VS are employees of Thermo Fisher Scientific. The remaining authors declare no competing interests.

References

(1) Ardito, F.; Giuliani, M.; Perrone, D.; Troiano, G.; Lo Muzio, L. The Crucial Role of Protein Phosphorylation in Cell Signaling and Its Use as Targeted Therapy (Review). *Int. J. Mol. Med.* **2017**, *40* (2), 271–280. https://doi.org/10.3892/ijmm.2017.3036.

(2) A. Solari, F.; Dell'Aica, M.; Sickmann, A.; P. Zahedi, R. Why Phosphoproteomics Is Still a Challenge. *Mol. Biosyst.* **2015**, *11* (6), 1487–1493. https://doi.org/10.1039/C5MB00024F.

(3) Neville, D. C. A.; Townsend, R. R.; Rozanas, C. R.; Verkman, A. S.; Price, E. M.; Gruis, D. B. Evidence for Phosphorylation of Serine 753 in CFTR Using a Novel Metal-Ion Affinity Resin and Matrix-Assisted Laser Desorption Mass Spectrometry. *Protein Sci.* **1997**, *6* (11), 2436–2445. https://doi.org/10.1002/pro.5560061117.

(4) Anguenot, R.; Yelle, S.; Nguyen-Quoc, B. Purification of Tomato Sucrose Synthase Phosphorylated Isoforms by Fe(III)-Immobilized Metal Affinity Chromatography. *Arch. Biochem. Biophys.* **1999**, *365* (1), 163–169. https://doi.org/10.1006/abbi.1999.1146.

(5) Posewitz, M. C.; Tempst, P. Immobilized Gallium(III) Affinity Chromatography of Phosphopeptides. *Anal. Chem.* **1999**, *71* (14), 2883–2892. https://doi.org/10.1021/ac981409y.

(6) Stensballe, A.; Andersen, S.; Jensen, O. N. Characterization of Phosphoproteins from Electrophoretic Gels by Nanoscale Fe(III) Affinity Chromatography with off-Line Mass Spectrometry Analysis. *PROTEOMICS* **2001**, *1* (2), 207–222. https://doi.org/10.1002/1615-9861(200102)1:2<207::AID-PROT207>3.0.CO;2-3.

(7) Ficarro, S. B.; McCleland, M. L.; Stukenberg, P. T.; Burke, D. J.; Ross, M. M.; Shabanowitz, J.; Hunt, D. F.; White, F. M. Phosphoproteome Analysis by Mass Spectrometry and Its Application to Saccharomyces Cerevisiae. *Nat. Biotechnol.* **2002**, *20* (3), 301–305. https://doi.org/10.1038/nbt0302-301.

(8) Pinkse, M. W. H.; Uitto, P. M.; Hilhorst, M. J.; Ooms, B.; Heck, A. J. R. Selective Isolation at the Femtomole Level of Phosphopeptides from Proteolytic Digests Using 2D-NanoLC-ESI-MS/MS and Titanium Oxide Precolumns. *Anal. Chem.* **2004**, *76* (14), 3935–3943. https://doi.org/10.1021/ac0498617.

(9) Sano, A.; Nakamura, H. Chemo-Affinity of Titania for the Column-Switching HPLC Analysis of Phosphopeptides. *Anal. Sci.* **2004**, *20* (3), 565–566. https://doi.org/10.2116/analsci.20.565.

(10) Aryal, U. K.; Ross, A. R. S. Enrichment and Analysis of Phosphopeptides under Different Experimental Conditions Using Titanium Dioxide Affinity Chromatography and Mass

Spectrometry. *Rapid Commun. Mass Spectrom.* **2010**, *24* (2), 219–231. https://doi.org/10.1002/rcm.4377.

(11) Olsen, J. V.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M. Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* **2006**, *127* (3), 635–648. https://doi.org/10.1016/j.cell.2006.09.026.

(12) Zhang, G.; Neubert, T. A. Use of Detergents to Increase Selectivity of Immunoprecipitation of Tyrosine Phosphorylated Peptides Prior to Identification by MALDI Quadrupole-TOF MS. *PROTEOMICS* **2006**, *6* (2), 571–578. https://doi.org/10.1002/pmic.200500267.

(13) Beltran, L.; Cutillas, P. R. Advances in Phosphopeptide Enrichment Techniques for Phosphoproteomics. *Amino Acids* **2012**, *43* (3), 1009–1024. https://doi.org/10.1007/s00726-012-1288-9.

(14) Schroeder, M. J.; Shabanowitz, J.; Schwartz, J. C.; Hunt, D. F.; Coon, J. J. A Neutral Loss Activation Method for Improved Phosphopeptide Sequence Analysis by Quadrupole Ion Trap Mass Spectrometry. *Anal. Chem.* **2004**, *76* (13), 3590–3598. https://doi.org/10.1021/ac0497104.

(15) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. Peptide and Protein Sequence Analysis by Electron Transfer Dissociation Mass Spectrometry. *Proc. Natl. Acad. Sci.* **2004**, *101* (26), 9528–9533. https://doi.org/10.1073/pnas.0402700101.

(16) Olsen, J. V.; Mann, M. Status of Large-Scale Analysis of Post-Translational Modifications by Mass Spectrometry*. *Mol. Cell. Proteomics* **2013**, *12* (12), 3444–3452. https://doi.org/10.1074/mcp.O113.034181.

(17) Nilsson, C. L. Advances in Quantitative Phosphoproteomics. *Anal. Chem.* **2012**, *84* (2), 735–746. https://doi.org/10.1021/ac202877y.

(18) Riley, N. M.; Coon, J. J. Phosphoproteomics in the Age of Rapid and Deep Proteome Profiling. *Anal. Chem.* **2016**, *88* (1), 74–94. https://doi.org/10.1021/acs.analchem.5b04123.

(19) Grimsrud, P. A.; Swaney, D. L.; Wenger, C. D.; Beauchene, N. A.; Coon, J. J. Phosphoproteomics for the Masses. *ACS Chem. Biol.* **2010**, *5* (1), 105–119. https://doi.org/10.1021/cb900277e.

(20) Muehlbauer, L. K.; Hebert, A. S.; Westphall, M. S.; Shishkova, E.; Coon, J. J. Global Phosphoproteome Analysis Using High-Field Asymmetric Waveform Ion Mobility Spectrometry on a Hybrid Orbitrap Mass Spectrometer. *Anal. Chem.* **2020**, *92* (24), 15959–15967. https://doi.org/10.1021/acs.analchem.0c03415.

(21) Chi, A.; Huttenhower, C.; Geer, L. Y.; Coon, J. J.; Syka, J. E. P.; Bai, D. L.; Shabanowitz, J.; Burke, D. J.; Troyanskaya, O. G.; Hunt, D. F. Analysis of Phosphorylation Sites on Proteins from Saccharomyces Cerevisiae by Electron Transfer Dissociation (ETD) Mass Spectrometry. *Proc. Natl. Acad. Sci.* **2007**, *104* (7), 2193–2198. https://doi.org/10.1073/pnas.0607084104.

(22) Swaney, D. L.; Wenger, C. D.; Thomson, J. A.; Coon, J. J. Human Embryonic Stem Cell Phosphoproteome Revealed by Electron Transfer Dissociation Tandem Mass Spectrometry. *Proc. Natl. Acad. Sci.* **2009**, *106* (4), 995–1000. https://doi.org/10.1073/pnas.0811964106.

(23) Beausoleil, S. A.; Villén, J.; Gerber, S. A.; Rush, J.; Gygi, S. P. A Probability-Based Approach for High-Throughput Protein Phosphorylation Analysis and Site Localization. *Nat. Biotechnol.* **2006**, *24* (10), 1285–1292. https://doi.org/10.1038/nbt1240.

(24) Taus, T.; Köcher, T.; Pichler, P.; Paschke, C.; Schmidt, A.; Henrich, C.; Mechtler, K. Universal and Confident Phosphorylation Site Localization Using phosphoRS. *J. Proteome Res.* **2011**, *10* (12), 5354–5362. https://doi.org/10.1021/pr200611n.

(25) Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J. Proteome Res.* **2011**, *10* (4), 1794–1805. https://doi.org/10.1021/pr101065j.

(26) Bailey, C. M.; Sweet, S. M. M.; Cunningham, D. L.; Zeller, M.; Heath, J. K.; Cooper, H. J. SLoMo: Automated Site Localization of Modifications from ETD/ECD Mass Spectra. *J. Proteome Res.* **2009**, *8* (4), 1965–1971. https://doi.org/10.1021/pr800917p.

(27) Savitski, M. M.; Lemeer, S.; Boesche, M.; Lang, M.; Mathieson, T.; Bantscheff, M.; Kuster, B. Confident Phosphorylation Site Localization Using the Mascot Delta Score. *Mol. Cell. Proteomics* **2011**, *10* (2), S1–S12. https://doi.org/10.1074/mcp.M110.003830.

(28) Molina, H.; Horn, D. M.; Tang, N.; Mathivanan, S.; Pandey, A. Global Proteomic Profiling of Phosphopeptides Using Electron Transfer Dissociation Tandem Mass Spectrometry. *Proc. Natl. Acad. Sci.* **2007**, *104* (7), 2199–2204. https://doi.org/10.1073/pnas.0611217104.

(29) Ferries, S.; Perkins, S.; Brownridge, P. J.; Campbell, A.; Eyers, P. A.; Jones, A. R.; Eyers, C. E. Evaluation of Parameters for Confident Phosphorylation Site Localization Using an Orbitrap Fusion Tribrid Mass Spectrometer. *J. Proteome Res.* **2017**, *16* (9), 3448–3459. https://doi.org/10.1021/acs.jproteome.7b00337.

(30) Sharma, K.; D'Souza, R. C. J.; Tyanova, S.; Schaab, C.; Wiśniewski, J. R.; Cox, J.; Mann, M. Ultradeep Human Phosphoproteome Reveals a Distinct Regulatory Nature of Tyr and Ser/Thr-Based Signaling. *Cell Rep.* **2014**, *8* (5), 1583–1594. https://doi.org/10.1016/j.celrep.2014.07.036.

(31) Stewart, H. I.; Grinfeld, D.; Giannakopulos, A.; Petzoldt, J.; Shanley, T.; Garland, M.; Denisov, E.; Peterson, A. C.; Damoc, E.; Zeller, M.; Arrey, T. N.; Pashkova, A.; Renuse, S.; Hakimi, A.; Kühn, A.; Biel, M.; Kreutzmann, A.; Hagedorn, B.; Colonius, I.; Schütz, A.; Stefes, A.; Dwivedi, A.; Mourad, D.; Hoek, M.; Reitemeier, B.; Cochems, P.; Kholomeev, A.; Ostermann, R.; Quiring, G.; Ochmann, M.; Möhring, S.; Wagner, A.; Petker, A.; Kanngiesser, S.; Wiedemeyer, M.; Balschun, W.; Hermanson, D.; Zabrouskov, V.; Makarov, A. A.; Hock, C. Parallelized Acquisition of Orbitrap and Astral Analyzers Enables High-Throughput Quantitative Analysis. *Anal. Chem.* **2023**, *95* (42), 15656–15664. https://doi.org/10.1021/acs.analchem.3c02856.

(32) Stewart, H.; Grinfeld, D.; Wagner, A.; Kholomeev, A.; Biel, M.; Giannakopulos, A.; Makarov, A.; Hock, C. A Conjoined Rectilinear Collision Cell and Pulsed Extraction Ion Trap with Auxiliary DC Electrodes. *J. Am. Soc. Mass Spectrom.* **2023**. https://doi.org/10.1021/jasms.3c00311.

(33) Skowronek, P.; Thielert, M.; Voytik, E.; Tanzer, M. C.; Hansen, F. M.; Willems, S.; Karayel, O.; Brunner, A.-D.; Meier, F.; Mann, M. Rapid and In-Depth Coverage of the (Phospho-)Proteome With Deep Libraries and Optimal Window Design for Dia-PASEF. *Mol. Cell. Proteomics* **2022**, *21* (9), 100279. https://doi.org/10.1016/j.mcpro.2022.100279.

(34) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.

(35) Bludau, I.; Willems, S.; Zeng, W.-F.; Strauss, M. T.; Hansen, F. M.; Tanzer, M. C.; Karayel, O.; Schulman, B. A.; Mann, M. The Structural Context of Posttranslational Modifications at a Proteome-Wide Scale. *PLoS Biol.* **2022**, *20* (5), e3001636. https://doi.org/10.1371/journal.pbio.3001636.

(36) Tyanova, S.; Cox, J.; Olsen, J.; Mann, M.; Frishman, D. Phosphorylation Variation during the Cell Cycle Scales with Structural Propensities of Proteins. *PLoS Comput. Biol.* **2013**, *9* (1), e1002842. https://doi.org/10.1371/journal.pcbi.1002842.

(37) Iakoucheva, L. M.; Radivojac, P.; Brown, C. J.; O'Connor, T. R.; Sikes, J. G.; Obradovic, Z.; Dunker, A. K. The Importance of Intrinsic Disorder for Protein Phosphorylation. *Nucleic Acids Res.* **2004**, *32* (3), 1037–1049. https://doi.org/10.1093/nar/gkh253.

(38) Johnson, J. L.; Yaron, T. M.; Huntsman, E. M.; Kerelsky, A.; Song, J.; Regev, A.; Lin, T.-Y.; Liberatore, K.; Cizin, D. M.; Cohen, B. M.; Vasan, N.; Ma, Y.; Krismer, K.; Robles, J. T.; van de Kooij, B.; van Vlimmeren, A. E.; Andrée-Busch, N.; Käufer, N. F.; Dorovkov, M.

V.; Ryazanov, A. G.; Takagi, Y.; Kastenhuber, E. R.; Goncalves, M. D.; Hopkins, B. D.; Elemento, O.; Taatjes, D. J.; Maucuer, A.; Yamashita, A.; Degterev, A.; Uduman, M.; Lu, J.; Landry, S. D.; Zhang, B.; Cossentino, I.; Linding, R.; Blenis, J.; Hornbeck, P. V.; Turk, B. E.; Yaffe, M. B.; Cantley, L. C. An Atlas of Substrate Specificities for the Human Serine/Threonine Kinome. *Nature* **2023**, *613* (7945), 759–766. https://doi.org/10.1038/s41586-022-05575-3.

(39) Shishkova, E.; Hebert, A. S.; Westphall, M. S.; Coon, J. J. Ultra-High Pressure (>30,000 Psi) Packing of Capillary Columns Enhancing Depth of Shotgun Proteomic Analyses. *Anal. Chem.* **2018**, *90* (19), 11503–11508. https://doi.org/10.1021/acs.analchem.8b02766.

(40) Cox, J.; Mann, M. MaxQuant Enables High Peptide Identification Rates, Individualized p.p.b.-Range Mass Accuracies and Proteome-Wide Protein Quantification. *Nat. Biotechnol.* **2008**, *26* (12), 1367–1372. https://doi.org/10.1038/nbt.1511.

(41) Bekker-Jensen, D. B.; Bernhardt, O. M.; Hogrebe, A.; Martinez-Val, A.; Verbeke, L.; Gandhi, T.; Kelstrup, C. D.; Reiter, L.; Olsen, J. V. Rapid and Site-Specific Deep Phosphoproteome Profiling by Data-Independent Acquisition without the Need for Spectral Libraries. *Nat. Commun.* **2020**, *11* (1), 787. https://doi.org/10.1038/s41467-020-14609-1.

(42) Kitata, R. B.; Choong, W.-K.; Tsai, C.-F.; Lin, P.-Y.; Chen, B.-S.; Chang, Y.-C.; Nesvizhskii, A. I.; Sung, T.-Y.; Chen, Y.-J. A Data-Independent Acquisition-Based Global Phosphoproteomics System Enables Deep Profiling. *Nat. Commun.* **2021**, *12* (1), 2539. https://doi.org/10.1038/s41467-021-22759-z.

(43) Srinivasan, A.; Sing, J. C.; Gingras, A.-C.; Röst, H. L. Improving Phosphoproteomics Profiling Using Data-Independent Mass Spectrometry. *J. Proteome Res.* **2022**, *21* (8), 1789–1799. https://doi.org/10.1021/acs.jproteome.2c00172.

(44) Oliinyk, D.; Meier, F. Ion Mobility-Resolved Phosphoproteomics with Dia-PASEF and Short Gradients. *PROTEOMICS* **2023**, *23* (7–8), 2200032. https://doi.org/10.1002/pmic.202200032.

(45) Lou, R.; Cao, Y.; Li, S.; Lang, X.; Li, Y.; Zhang, Y.; Shui, W. Benchmarking Commonly Used Software Suites and Analysis Workflows for DIA Proteomics and Phosphoproteomics. *Nat. Commun.* **2023**, *14* (1), 94. https://doi.org/10.1038/s41467-022-35740-1.

(46) Bruderer, R.; Bernhardt, O. M.; Gandhi, T.; Xuan, Y.; Sondermann, J.; Schmidt, M.; Gomez-Varela, D.; Reiter, L. Optimization of Experimental Parameters in Data-Independent Mass Spectrometry Significantly Increases Depth and Reproducibility of Results*. *Mol. Cell. Proteomics* **2017**, *16* (12), 2296–2309. https://doi.org/10.1074/mcp.RA117.000314.

(47) Leijten, N. M.; Heck, A. J. R.; Lemeer, S. Histidine Phosphorylation in Human Cells; a Needle or Phantom in the Haystack? *Nat. Methods* **2022**, *19* (7), 827–828. https://doi.org/10.1038/s41592-022-01524-0.

(48) Vaca Jacome, A. S.; Peckner, R.; Shulman, N.; Krug, K.; DeRuff, K. C.; Officer, A.; Christianson, K. E.; MacLean, B.; MacCoss, M. J.; Carr, S. A.; Jaffe, J. D. Avant-Garde: An Automated Data-Driven DIA Data Curation Tool. *Nat. Methods* **2020**, *17* (12), 1237–1244. https://doi.org/10.1038/s41592-020-00986-4.

(49) Huttlin, E. L.; Jedrychowski, M. P.; Elias, J. E.; Goswami, T.; Rad, R.; Beausoleil, S. A.; Villén, J.; Haas, W.; Sowa, M. E.; Gygi, S. P. A Tissue-Specific Atlas of Mouse Protein Phosphorylation and Expression. *Cell* **2010**, *143* (7), 1174–1189. https://doi.org/10.1016/j.cell.2010.12.001.

(50) Stopfer, L. E.; Flower, C. T.; Gajadhar, A. S.; Patel, B.; Gallien, S.; Lopez-Ferrer, D.; White, F. M. High-Density, Targeted Monitoring of Tyrosine Phosphorylation Reveals Activated Signaling Networks in Human Tumors. *Cancer Res.* **2021**, *81* (9), 2495–2509. https://doi.org/10.1158/0008-5472.CAN-20-3804.

(51) Huttlin, E. L.; Jedrychowski, M. P.; Elias, J. E.; Goswami, T.; Rad, R.; Beausoleil, S. A.; Villén, J.; Haas, W.; Sowa, M. E.; Gygi, S. P. A Tissue-Specific Atlas of Mouse Protein Phosphorylation and Expression. *Cell* **2010**, *143* (7), 1174–1189. https://doi.org/10.1016/j.cell.2010.12.001.

(52) Giansanti, P.; Samaras, P.; Bian, Y.; Meng, C.; Coluccio, A.; Frejno, M.; Jakubowsky, H.; Dobiasch, S.; Hazarika, R. R.; Rechenberger, J.; Calzada-Wack, J.; Krumm, J.; Mueller, S.; Lee, C.-Y.; Wimberger, N.; Lautenbacher, L.; Hassan, Z.; Chang, Y.-C.; Falcomatà, C.; Bayer, F. P.; Bärthel, S.; Schmidt, T.; Rad, R.; Combs, S. E.; The, M.; Johannes, F.; Saur, D.; de Angelis, M. H.; Wilhelm, M.; Schneider, G.; Kuster, B. Mass Spectrometry-Based Draft of the Mouse Proteome. *Nat. Methods* **2022**, *19* (7), 803–811. https://doi.org/10.1038/s41592-022-01526-y.

(53) Hornbeck, P. V.; Zhang, B.; Murray, B.; Kornhauser, J. M.; Latham, V.; Skrzypek, E. PhosphoSitePlus, 2014: Mutations, PTMs and Recalibrations. *Nucleic Acids Res.* **2015**, *43* (Database issue), D512-520. https://doi.org/10.1093/nar/gku1267.

(54) Pinna, L. A.; Ruzzene, M. How Do Protein Kinases Recognize Their Substrates? *Biochim. Biophys. Acta* **1996**, *1314* (3), 191–225. https://doi.org/10.1016/s0167-4889(96)00083-3.

(55) Alderson, T. R.; Pritišanac, I.; Kolarić, Đ.; Moses, A. M.; Forman-Kay, J. D. Systematic Identification of Conditionally Folded Intrinsically Disordered Regions by AlphaFold2. *Proc. Natl. Acad. Sci. U. S. A.* **2023**, *120* (44), e2304302120. https://doi.org/10.1073/pnas.2304302120.

(56) Chico, L. K.; Van Eldik, L. J.; Watterson, D. M. Targeting Protein Kinases in Central Nervous System Disorders. *Nat. Rev. Drug Discov.* **2009**, *8* (11), 892–909. https://doi.org/10.1038/nrd2999.

(57) Küry, S.; van Woerden, G. M.; Besnard, T.; Proietti Onori, M.; Latypova, X.; Towne, M. C.; Cho, M. T.; Prescott, T. E.; Ploeg, M. A.; Sanders, S.; Stessman, H. A. F.; Pujol, A.; Distel, B.; Robak, L. A.; Bernstein, J. A.; Denommé-Pichon, A.-S.; Lesca, G.; Sellars, E. A.; Berg, J.; Carré, W.; Busk, Ø. L.; van Bon, B. W. M.; Waugh, J. L.; Deardorff, M.; Hoganson, G. E.; Bosanko, K. B.; Johnson, D. S.; Dabir, T.; Holla, Ø. L.; Sarkar, A.; Tveten, K.; de Bellescize, J.; Braathen, G. J.; Terhal, P. A.; Grange, D. K.; van Haeringen, A.; Lam, C.; Mirzaa, G.; Burton, J.; Bhoj, E. J.; Douglas, J.; Santani, A. B.; Nesbitt, A. I.; Helbig, K. L.; Andrews, M. V.; Begtrup, A.; Tang, S.; van Gassen, K. L. I.; Juusola, J.; Foss, K.; Enns, G. M.; Moog, U.; Hinderhofer, K.; Paramasivam, N.; Lincoln, S.; Kusako, B. H.; Lindenbaum, P.; Charpentier, E.; Nowak, C. B.; Cherot, E.; Simonet, T.; Ruivenkamp, C. A. L.; Hahn, S.; Brownstein, C. A.; Xia, F.; Schmitt, S.; Deb, W.; Bonneau, D.; Nizon, M.; Quinquis, D.; Chelly, J.; Rudolf, G.; Sanlaville, D.; Parent, P.; Gilbert-Dussardier, B.; Toutain, A.; Sutton, V. R.; Thies, J.; Peart-Vissers, L. E. L. M.; Boisseau, P.; Vincent, M.; Grabrucker, A. M.; Dubourg, C.; Undiagnosed Diseases Network; Tan, W.-H.; Verbeek, N. E.; Granzow, M.; Santen, G. W. E.; Shendure, J.; Isidor, B.; Pasquier, L.; Redon, R.; Yang, Y.; State, M. W.; Kleefstra, T.; Cogné, B.; GEM HUGO; Deciphering Developmental Disorders Study; Petrovski, S.; Retterer, K.; Eichler, E. E.; Rosenfeld, J. A.; Agrawal, P. B.; Bézieau, S.; Odent, S.; Elgersma, Y.; Mercier, S. De Novo Mutations in Protein Kinase Genes CAMK2A and CAMK2B Cause Intellectual Disability. *Am. J. Hum. Genet.* **2017**, *101* (5), 768–788. https://doi.org/10.1016/j.ajhg.2017.10.003.

(58) Kool, M. J.; Proietti Onori, M.; Borgesius, N. Z.; van de Bree, J. E.; Elgersma-Hooisma, M.; Nio, E.; Bezstarosti, K.; Buitendijk, G. H. S.; Aghadavoud Jolfaei, M.; Demmers, J. A. A.; Elgersma, Y.; van Woerden, G. M. CAMK2-Dependent Signaling in Neurons Is Essential for Survival. *J. Neurosci. Off. J. Soc. Neurosci.* **2019**, *39* (28), 5424–5439. https://doi.org/10.1523/JNEUROSCI.1341-18.2019.

(59) Rigter, P. M. F.; de Konink, C.; Dunn, M. J.; Proietti Onori, M.; Humberson, J. B.; Thomas, M.; Barnes, C.; Prada, C. E.; Weaver, K. N.; Ryan, T. D.; Caluseriu, O.; Conway, J.; Calamaro, E.; Fong, C.-T.; Wuyts, W.; Meuwissen, M.; Hordijk, E.; Jonkers, C. N.; Anderson, L.; Yuseinova, B.; Polonia, S.; Beysen, D.; Stark, Z.; Savva, E.; Poulton, C.; McKenzie, F.; Bhoj, E.; Bupp, C. P.; Bézieau, S.; Mercier, S.; Blevins, A.; Wentzensen, I. M.; Xia, F.; Rosenfeld, J. A.; Hsieh, T.-C.; Krawitz, P. M.; Elbracht, M.; Veenma, D. C. M.; Schulman, H.; Stratton, M. M.; Küry, S.; van Woerden, G. M. Role of CAMK2D in Neurodevelopment and Associated Conditions. *Am. J. Hum. Genet.* **2024**, *111* (2), 364–382. https://doi.org/10.1016/j.ajhg.2023.12.016.

(60) Lordén, G.; Wozniak, J. M.; Doré, K.; Dozier, L. E.; Cates-Gatto, C.; Patrick, G. N.; Gonzalez, D. J.; Roberts, A. J.; Tanzi, R. E.; Newton, A. C. Enhanced Activity of Alzheimer

Disease-Associated Variant of Protein Kinase Cα Drives Cognitive Decline in a Mouse Model. *Nat. Commun.* **2022**, *13* (1), 7200. https://doi.org/10.1038/s41467-022-34679-7.

(61) Callender, J. A.; Newton, A. C. Conventional Protein Kinase C in the Brain: 40 Years Later. *Neuronal Signal.* **2017**, *1* (2), NS20160005. https://doi.org/10.1042/NS20160005.

(62) Maichele, A. J.; Burwinkel, B.; Maire, I.; Søvik, O.; Kilimann, M. W. Mutations in the Testis/Liver Isoform of the Phosphorylase Kinase Gamma Subunit (PHKG2) Cause Autosomal Liver Glycogenosis in the Gsd Rat and in Humans. *Nat. Genet.* **1996**, *14* (3), 337–340. https://doi.org/10.1038/ng1196-337.

(63) Niemi, N. M.; Pagliarini, D. J. The Extensive and Functionally Uncharacterized Mitochondrial Phosphoproteome. *J. Biol. Chem.* **2021**, *297* (1), 100880. https://doi.org/10.1016/j.jbc.2021.100880.

(64) Arnold, P. K.; Finley, L. W. S. Regulation and Function of the Mammalian Tricarboxylic Acid Cycle. *J. Biol. Chem.* **2023**, *299* (2), 102838. https://doi.org/10.1016/j.jbc.2022.102838.

(65) Rath, S.; Sharma, R.; Gupta, R.; Ast, T.; Chan, C.; Durham, T. J.; Goodman, R. P.; Grabarek, Z.; Haas, M. E.; Hung, W. H. W.; Joshi, P. R.; Jourdain, A. A.; Kim, S. H.; Kotrys, A. V.; Lam, S. S.; McCoy, J. G.; Meisel, J. D.; Miranda, M.; Panda, A.; Patgiri, A.; Rogers, R.; Sadre, S.; Shah, H.; Skinner, O. S.; To, T.-L.; Walker, M. A.; Wang, H.; Ward, P. S.; Wengrod, J.; Yuan, C.-C.; Calvo, S. E.; Mootha, V. K. MitoCarta3.0: An Updated Mitochondrial Proteome Now with Sub-Organelle Localization and Pathway Annotations. *Nucleic Acids Res.* **2021**, *49* (D1), D1541–D1547. https://doi.org/10.1093/nar/gkaa1011.

(66) Madeira, F.; Pearce, M.; Tivey, A. R. N.; Basutkar, P.; Lee, J.; Edbali, O.; Madhusoodanan, N.; Kolesnikov, A.; Lopez, R. Search and Sequence Analysis Tools Services from EMBL-EBI in 2022. *Nucleic Acids Res.* **2022**, *50* (W1), W276–W279. https://doi.org/10.1093/nar/gkac240.

(67) de Cima, S.; Polo, L. M.; Díez-Fernández, C.; Martínez, A. I.; Cervera, J.; Fita, I.; Rubio, V. Structure of Human Carbamoyl Phosphate Synthetase: Deciphering the on/off Switch of Human Ureagenesis. *Sci. Rep.* **2015**, *5*, 16950. https://doi.org/10.1038/srep16950.

(68) Ge, S. X.; Jung, D.; Yao, R. ShinyGO: A Graphical Gene-Set Enrichment Tool for Animals and Plants. *Bioinforma. Oxf. Engl.* **2020**, *36* (8), 2628–2629. https://doi.org/10.1093/bioinformatics/btz931.

(69) Häberle, J.; Burlina, A.; Chakrapani, A.; Dixon, M.; Karall, D.; Lindner, M.; Mandel, H.; Martinelli, D.; Pintos-Morell, G.; Santer, R.; Skouma, A.; Servais, A.; Tal, G.; Rubio, V.; Huemer, M.; Dionisi-Vici, C. Suggested Guidelines for the Diagnosis and Management of Urea Cycle Disorders: First Revision. *J. Inherit. Metab. Dis.* **2019**, *42* (6), 1192–1230. https://doi.org/10.1002/jimd.12100.

(70) Häberle, J.; Shchelochkov, O. A.; Wang, J.; Katsonis, P.; Hall, L.; Reiss, S.; Eeds, A.; Willis, A.; Yadav, M.; Summar, S.; Urea Cycle Disorders Consortium; Lichtarge, O.; Rubio, V.; Wong, L.-J.; Summar, M. Molecular Defects in Human Carbamoy Phosphate Synthetase I: Mutational Spectrum, Diagnostic and Protein Structure Considerations. *Hum. Mutat.* **2011**, *32* (6), 579–589. https://doi.org/10.1002/humu.21406.

(71) Pekkala, S.; Martínez, A. I.; Barcelona, B.; Yefimenko, I.; Finckh, U.; Rubio, V.; Cervera, J. Understanding Carbamoyl-Phosphate Synthetase I (CPS1) Deficiency by Using Expression Studies and Structure-Based Analysis. *Hum. Mutat.* **2010**, *31* (7), 801–808. https://doi.org/10.1002/humu.21272.

(72) Takakusa, H.; Mohar, I.; Kavanagh, T. J.; Kelly, E. J.; Kaspera, R.; Nelson, S. D. Protein Tyrosine Nitration of Mitochondrial Carbamoyl Phosphate Synthetase 1 and Its Functional Consequences. *Biochem. Biophys. Res. Commun.* **2012**, *420* (1), 54–60. https://doi.org/10.1016/j.bbrc.2012.02.114.

(73) Yu, C.; Zhao, J.; Yan, L.; Qi, Y.; Guo, X.; Lou, Z.; Hu, J.; Rao, Z. Structural Insights into G Domain Dimerization and Pathogenic Mutation of OPA1. *J. Cell Biol.* **2020**, *219* (7), e201907098. https://doi.org/10.1083/jcb.201907098.

(74) Del Dotto, V.; Fogazza, M.; Carelli, V.; Rugolo, M.; Zanna, C. Eight Human OPA1 Isoforms, Long and Short: What Are They For? *Biochim. Biophys. Acta Bioenerg.* **2018**, *1859* (4), 263–269. https://doi.org/10.1016/j.bbabio.2018.01.005.

(75) Ferré, M.; Caignard, A.; Milea, D.; Leruez, S.; Cassereau, J.; Chevrollier, A.; Amati-Bonneau, P.; Verny, C.; Bonneau, D.; Procaccio, V.; Reynier, P. Improved Locus-Specific Database for OPA1 Mutations Allows Inclusion of Advanced Clinical Data. *Hum. Mutat.* **2015**, *36* (1), 20–25. https://doi.org/10.1002/humu.22703.

(76) Chappie, J. S.; Acharya, S.; Leonard, M.; Schmid, S. L.; Dyda, F. G Domain Dimerization Controls Dynamin's Assembly-Stimulated GTPase Activity. *Nature* **2010**, *465* (7297), 435–440. https://doi.org/10.1038/nature09032.

(77) Santarelli, R.; Rossi, R.; Scimemi, P.; Cama, E.; Valentino, M. L.; La Morgia, C.; Caporali, L.; Liguori, R.; Magnavita, V.; Monteleone, A.; Biscaro, A.; Arslan, E.; Carelli, V. OPA1-Related Auditory Neuropathy: Site of Lesion and Outcome of Cochlear Implantation. *Brain J. Neurol.* **2015**, *138* (Pt 3), 563–576. https://doi.org/10.1093/brain/awu378.

(78) Loo, J. L.; Singhal, S.; Rukmini, A. V.; Tow, S.; Amati-Bonneau, P.; Procaccio, V.; Bonneau, D.; Gooley, J. J.; Reynier, P.; Ferré, M.; Milea, D. Multiethnic Involvement in Autosomal-Dominant Optic Atrophy in Singapore. *Eye Lond. Engl.* **2017**, *31* (3), 475–480. https://doi.org/10.1038/eye.2016.255.

(79) Del Dotto, V.; Fogazza, M.; Musiani, F.; Maresca, A.; Aleo, S. J.; Caporali, L.; La Morgia, C.; Nolli, C.; Lodi, T.; Goffrini, P.; Chan, D.; Carelli, V.; Rugolo, M.; Baruffini, E.; Zanna, C. Deciphering OPA1 Mutations Pathogenicity by Combined Analysis of Human, Mouse and Yeast Cell Models. *Biochim. Biophys. Acta Mol. Basis Dis.* **2018**, *1864* (10), 3496–3514. https://doi.org/10.1016/j.bbadis.2018.08.004.

(80) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P.; SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17* (3), 261–272. https://doi.org/10.1038/s41592-019-0686-2.

(81) Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M. Y.; Geiger, T.; Mann, M.; Cox, J. The Perseus Computational Platform for Comprehensive Analysis of (Prote)Omics Data. *Nat. Methods* **2016**, *13* (9), 731–740. https://doi.org/10.1038/nmeth.3901.

(82) Chen, E. Y.; Tan, C. M.; Kou, Y.; Duan, Q.; Wang, Z.; Meirelles, G. V.; Clark, N. R.; Ma'ayan, A. Enrichr: Interactive and Collaborative HTML5 Gene List Enrichment Analysis Tool. *BMC Bioinformatics* **2013**, *14*, 128. https://doi.org/10.1186/1471-2105-14-128.

(83) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duschesnay, E. Scikit-Learn: Machine Learning in Python. *JMLR* **2011**, *12*, 2825–2830.

(84) Stenson, P. D.; Mort, M.; Ball, E. V.; Chapman, M.; Evans, K.; Azevedo, L.; Hayden, M.; Heywood, S.; Millar, D. S.; Phillips, A. D.; Cooper, D. N. The Human Gene Mutation Database (HGMD®): Optimizing Its Use in a Clinical Diagnostic or Research Setting. *Hum. Genet.* **2020**, *139* (10), 1197–1207. https://doi.org/10.1007/s00439-020-02199-3.

(85) *UWPR*. https://proteomicsresource.washington.edu/instruments/orbitrapexploris480.php (accessed 2024-04-30).

(86) Heil, L. R.; Damoc, E.; Arrey, T. N.; Pashkova, A.; Denisov, E.; Petzoldt, J.; Peterson, A. C.; Hsu, C.; Searle, B. C.; Shulman, N.; Riffle, M.; Connolly, B.; MacLean, B. X.; Remes, P. M.; Senko, M. W.; Stewart, H. I.; Hock, C.; Makarov, A. A.; Hermanson, D.; Zabrouskov, V.; Wu, C. C.; MacCoss, M. J. Evaluating the Performance of the Astral Mass Analyzer for Quantitative Proteomics Using Data-Independent Acquisition. *J. Proteome Res.* **2023**, *22* (10), 3290–3300. https://doi.org/10.1021/acs.jproteome.3c00357.

(87) Guzman, U. H.; Martinez-Val, A.; Ye, Z.; Damoc, E.; Arrey, T. N.; Pashkova, A.; Renuse, S.; Denisov, E.; Petzoldt, J.; Peterson, A. C.; Harking, F.; Østergaard, O.; Rydbirk, R.; Aznar, S.; Stewart, H.; Xuan, Y.; Hermanson, D.; Horning, S.; Hock, C.; Makarov, A.; Zabrouskov, V.; Olsen, J. V. Ultra-Fast Label-Free Quantification and Comprehensive Proteome Coverage with Narrow-Window Data-Independent Acquisition. *Nat. Biotechnol.* **2024**. https://doi.org/10.1038/s41587-023-02099-7.

Chapter 3

SynchroSep-MS: Parallel LC Separations for Multiplexed Proteomics

Lancaster, N. M.; Chen, L.-Y.; Zhao, B.; Anderson, B. J.; Probasco, M. D.; Demichev, V.; Polasky, D. A.; Nesvizhskii, A. I.; Overmyer, K. A.; Quarmby, S. T.; Coon, J. J. SynchroSep-MS: Parallel LC Separations for Multiplexed Proteomics. *J. Am. Soc. Mass Spectrom*. **2025**, 36 (9), 1979–1987. https://doi.org/10.1021/jasms.5c00207.

Abstract

Achieving high throughput remains a challenge in MS-based proteomics for large-scale applications. We introduce SynchroSep-MS, a novel method for parallelized, label-free proteome analysis that leverages the rapid acquisition speeds of modern mass spectrometers. This approach employs multiple liquid chromatography columns, each with an independent sample, simultaneously introduced to a single mass spectrometer inlet. A precisely controlled retention time offset between sample injections creates distinct elution profiles, facilitating unambiguous analyte assignment. We modified the DIA-NN workflow to effectively process this unique parallelized data, accounting for retention time offsets. Using a dual-column setup with mouse brain peptides, SynchroSep-MS detected approximately 16,700 unique protein groups, nearly doubling the peptide information obtained from a conventional single proteome analysis. The method demonstrated excellent precision and reproducibility (median protein %RSDs less than 4%) and high quantitative linearity (median R^2 greater than 0.96) with minimal matrix interference. SynchroSep-MS represents a new paradigm for data collection and the first example of label-free multiplexed proteome analysis via parallel LC separations, offering a direct strategy to accelerate throughput for demanding applications like large-scale clinical cohorts and single-cell analyses without compromising peak capacity or causing ionization suppression.

Introduction

Advances in mass spectrometry (MS)-based proteomics are largely driven by improvements in instrumentation, including more comprehensive and high-throughput proteomics methods; however, obtaining complete human proteomes has typically required extensive fractionation, multiple injections, and long analysis times.[1,2] Recent advances in instrumentation that allow for MS/MS acquisition rates exceeding 200 Hz, *i.e.*, Astral mass analyzer, have greatly accelerated the rate of and depth of proteome analysis. Specifically, these high MS/MS acquisition rates can be combined with data-independent acquisition (DIA) to enable detection of over 10,000 proteins from human samples in just tens of minutes.[3–11] With one sample at a time, two or three samples per hour can now be analyzed excellent proteomic depth. That said, many clinical and single cell applications, for example, demand the analysis of hundreds and even thousands of samples.

To satiate this throughput demand, many have now put focus on the chromatographic separation step. Indeed, active effort in the field seeks to dramatically accelerate the speed of chromatographic analysis through short gradients,[5,12–19] or parallelization of sample loading/column washing with gradient separations using multiple columns.[20–33] Shortening gradients, however, both erodes quantitative precision and accuracy and reduces proteomic depth. Multiplexing can be achieved through isotopic labeling, *e.g.,* tandem mass tags, but that approach has its own complications including cost, considerably more preparation labor, ratio distortion, and incompatibility with DIA.

We supposed a novel route to multiplexed, label-free DIA proteome analysis could be achieved by use of multiple LC columns, each with their own sample and integrated nano-

electrospray emitter, which could be analyzed in parallel owing to the speed of the new Orbitrap Astral MS platform. This approach would eliminate the loss in ionization efficiency that occurs when multiple samples are loaded onto a single column and would provide the higher ion flux needed so that the MS system could operate at top speeds. Here we describe an initial implementation of this approach which we call SynchroSep-MS.[34]

Experimental Section

Materials and reagents. Ultrapure water was provided by a Barnstead GenPure Pro system (Thermo Scientific). Trifluoroacetic acid (TFA, HPLC grade, >99.9%), chloroacetamide (≥98%, C0267-100G), urea (U5378-1kg), tris(2-carboxyethyl)phosphine hydrochloride (TCEP, C4706-2G, guanidine hydrochloride solution (6M, SRE0066), and fetal bovine serum (F2442) were obtained from Sigma-Aldrich. Formic acid (LC-MS grade), methanol (Optima LC/MS grade), acetonitrile (Optima LC-MS grade), Tris Buffer (1 M Tris pH 8.0, 0.2µm filtered, Invitrogen, AM9856), and Pierce Peptide Retention Time Calibration Mixture (PRTC, 88320) were obtained from Fisher Scientific. Lysyl Endopeptidase (LysC, 100369-826) was sourced from VWR. Sequencing Grade Modified Trypsin (V5113) was obtained from Promega. BCA Protein Assay Reagent A (23228), BCA Protein Assay B (23224), Bovin Serum Albumin Standards (2 mg/mL, 23209), IMDM (12440053), and penicillin–streptomycin (15140122) were obtained from Thermo Scientific.

Sample Preparation. All experiments were performed in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the Animal Care and Use Committee at the University of Wisconsin-Madison. Brains were harvested from C57BL/6J adult female mice after euthanasia and immediately frozen in liquid nitrogen. The brains were pulverized under liquid nitrogen into a fine powder.

Mouse brain samples were weighed on dry ice and guanidine buffer (5.4 M guanidine hydrochloride, 100mM Tris, pH 8) was added. Tissue lysis was performed via vortexing and sonicating in a chilled sonicator bath. The protein concentration was measured using the Pierce BCA protein assay kit.

Human HAP1 KO cells acquired from Horizon Discovery were cultured in IMDM with 1% penicillin–streptomycin and 10% fetal bovine serum at 37 °C and 5% CO_2, as previously described.[11,35] Lysis was performed in 6 M guanidine with sonication. The protein concentration was then measured via the Pierce BCA protein assay kit.

After measuring the protein content of samples via the BCA assay, methanol was added to 90% to precipitate proteins and the samples were centrifuged at 9000g for 5 minutes. The supernatant was removed, and the protein pellets were resolubilized in 8 M urea, 100 mM Tris, 10 mM TCEP, 40 mM chloroacetamide, pH 8 at a target protein concentration of 1.5 mg/mL. The samples were diluted to 2 M urea with 100 mM Tris, pH 8. LysC was added at a 1:50 enzyme:protein ratio and the samples were rocked gently for 4 hours at ambient conditions. Trypsin was then added at a 1:50 enzyme:protein ratio and the samples were rocked gently overnight at ambient conditions. Digestion was quenched with 10% TFA in water and the samples were centrifuged for 5 min at 9000g prior to desalting with Strata-X 33 μm polymeric reversed phase SPE cartridges (Phenomenex). Desalted samples were dried down in a SpeedVac (Thermo Scientific) and reconstituted in 0.2% formic acid in water. The peptide concentration was measured via NanoDrop (Thermo Scientific) prior to LC-MS/MS analysis.

Data Collection. To implement a dual column SynchroSep-MS setup, two nanoLCs were used. A Vanquish Neo UHPLC (Thermo Scientific) was configured on the mass spectrometer control computer. A contact closure line connected the Vanquish Neo to an Ultimate 3000 nLC (Thermo Scientific) such that the Vanquish Neo could send an output signal that was then received by the Ultimate 3000. The Ultimate 3000 was configured and controlled by a separate computer. Sample injection onto Column 1 was performed using the Vanquish Neo nLC. At the end of the

sample injection step in the Vanquish Neo, a relay output signal was programmed into the method. The Ultimate 3000 method was programmed to start sample loading only after it received a relay signal from the Vanquish Neo. Thus, the retention time offset could be consistently maintained. By adjusting the time delay of the relay signal in the Vanquish Neo method, the absolute retention time offset can be tuned.

For both nLCs, mobile phase A was 0.2% formic acid in water and mobile phase B was 80% ACN/20% water/0.2% formic acid. Columns were prepared by laser pulling emitters on silica capillaries (360 µm O.D., 75 µm I.D.) and packing to 40 cm with 1.7 µm BEH C18 particles, 130 Å (Waters) at ultrahigh pressure using a custom-built high pressure packing station.[36] For the Vanquish Neo, a gradient separation was performed at 0.3 µL/min starting by ramping from 0 to 6% B over 2 minutes, then ramping from 6 to 47%B for 30 minutes, washing at 100% B for ~11.5 minutes, and equilibrating at 0% B for ~25 minutes. Note that this equilibration stage was extended to assure mass spectrometry data collection was performed throughout the entire gradient separation on the second column and could likely be reduced to maximize throughput. A relay signal was programmed into the Vanquish Neo method at 0, 1, or 2 minutes for the retention time offset tuning dataset. For most of this study, the relay signal was sent at the 1-minute timepoint in the Vanquish Neo method. For the Ultimate 3000, sample injection was delayed until the relay signal was detected as an input. Upon the relay signal, sample injection was performed by drawing up 8 µL of transfer fluid (0.2% formic acid in water), followed by drawing up the sample and 2.4 µL of transfer fluid using a 5 µL sample loop. The injection valve was then switched to inject onto the column. A gradient was ramped from 0 to 7% B over 2 minutes, followed by ramping from 7% B to 43% B over 30 minutes. Then, the %B was set to 100%B over 0.1 minutes followed by ~11.7

minutes of washing, and 10 minutes of equilibration at 0% B. The gradient conditions are slightly different for each LC as the specific settings were tuned to provide well distributed total ion chromatograms for the two columns. For this method, reproducible timing of injections is critical to maintain consistent retention time offsets. In our experience, the Vanquish Neo system exhibits higher variability in the total sample loading times, likely due to the injection preparation steps and sample loop pressurization step implemented. While this variability would not negatively impact typical performance since the retention time start is synchronized across separations, the use of a Vanquish Neo to control the second column might lead to variation in the delay between injections onto the first and second columns. We thus chose to employ an Ultimate 3000 to perform separations as this system does not perform as many injection preparation steps or sample loop pressurization and takes consistent amounts of time for sample loading across injections.

Both emitters were aligned to the inlet capillary using a 3D-printed holder. A spray voltage of 2000V was applied to both columns via a liquid-liquid junction. The ion transfer tube temperature was set to 280°C. A DIA method was employed. 240,000 resolving power MS^1 scans were collected in the Orbitrap analyzer with an *m/z* range of 380-980, with RF Lens (%) of 40, maximum injection time of 5 ms, and normalized AGC target of 500%. Astral MS^2 scans were acquired with a DIA method iterating over a DIA *m/z* range of 380-980 with 4 Th windows (with 1 Th overlap). The scan range in the Astral analyzer was from 150-2000 *m/z*, and the normalized AGC target was set to 500% with a maximum injection time of 5 ms, and an RF Lens (%) of 40. The cycle time for MS^1 scans was set to 0.6 seconds.

Data analysis. Raw data files were converted to the .mzml format with MSConvert (v. 3.0.241192ef69)[37] using the peakPicking filter set to 'Vendor'. Dual injection files were searched

against a mouse protein database using the default settings for the 'DIA_SpecLib_Quant' workflow within FragPipe.[38,39] The intermediate output files *{file-name}*_rank{*rank number*}.pepXML files were parsed using Python, where rank denotes the ranks of PSM quality as implemented in MSFragger-DIA algorithm. Peptide Spectral Matches (PSMs) in .pepXML files were filtered for an expectation value less than 1e-5. Duplicate PSMs were filtered for only pairs with a retention time difference greater than 30 seconds. The earlier PSM from each pair was assumed to be from Column 1 and the later from Column 2. The RTs from each column were used to construct a locally weighted scatterplot smoothing (LOWESS) model with respect to the Column 1 RTs using the 'lowess' function within the Statsmodels library[40] and the 'interp1d' function within SciPy.[41] The ΔRT model was defined as the difference between these models. The Pierce PRTC standard raw data were manually examined in Skyline (v. 24.1.0.414) to define experimental ΔRT values.[42]

A prototype version of DIA-NN 2.2.0 was employed for searching the data.[43] The prototype version takes in two additional parameters through the command line, --rt-shift and --rt-gap. The --rt-shift parameter defines the retention time offset (ΔRT) that will be applied to the retention times in the spectral library for searching a specific file. The --rt-gap parameter specifies the retention time window width that is used to search the chromatogram around the specified retention time. A Column 1- only injection is used for creating a retention time-aligned predicted spectral library by searching in DIA-NN and aligning retention times using an R script. A workflow diagram for the data processing procedure is shown in **Figure 3.1**. For the dilution series experiments, normalization was turned off using the --no-norm command. For data analysis, the following columns in the report.parquet output by DIA-NN were filtered for values less than or equal to 0.01:

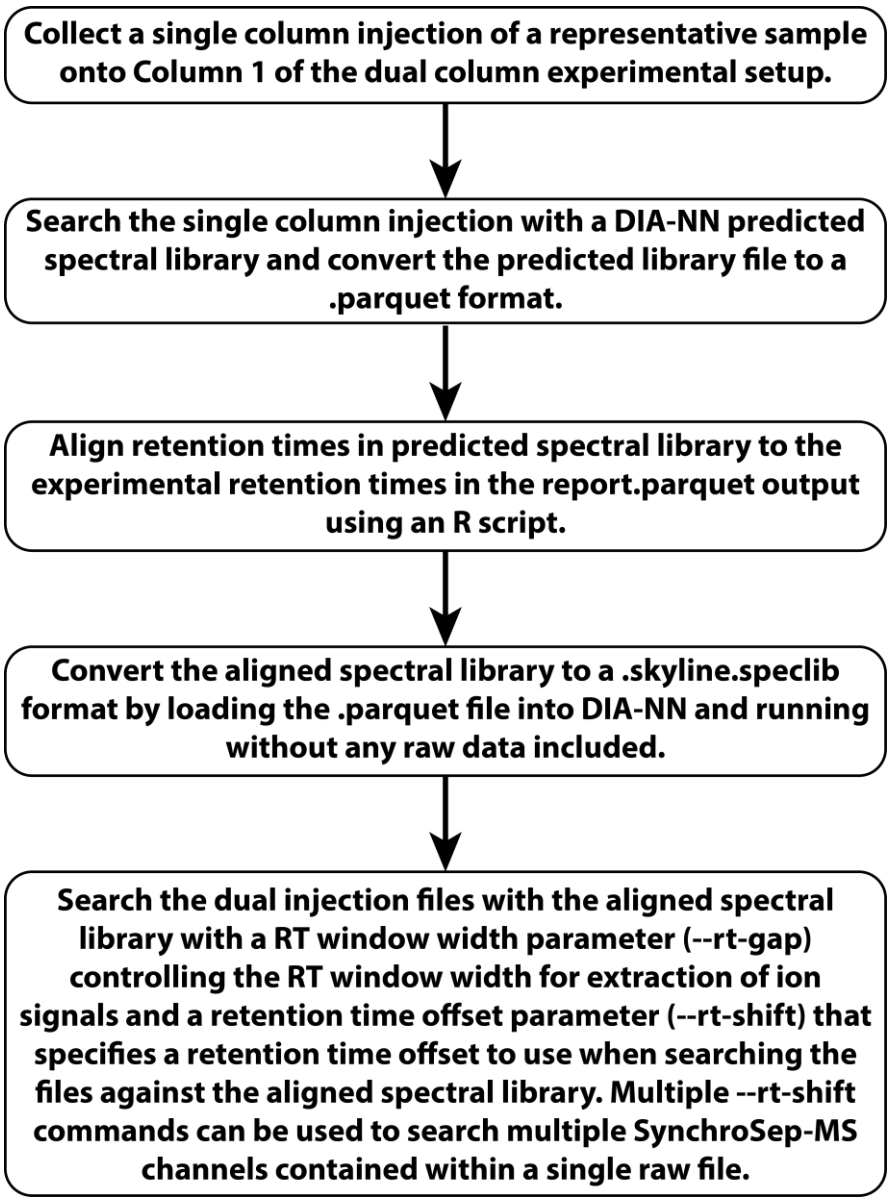

Figure 3.1. Processing SynchroSep-MS using Prototype DIA-NN Software. A schematic workflow is shown for processing the dual column SynchroSep-MS data with a modified version of the DIA-NN software. Image Description

Q.Value, Lib.Q.Value, Global.Q.Value, Lib.PG.Q.Value, Global.PG.Q.Value, PG.Q.Value, Lib.Peptidoform.Q.Value, and Global.Peptidoform.Q.Value.

Normal searches were performed using a mouse protein database downloaded from UniProt. The entrapment search was performed by combining a mouse and *C. elegans* protein database downloaded from UniProt.[44] Match-between-runs (MBR) and protein inference were turned off for a precursor-level FDR assessment. The 'Q.Value' column was filtered for values ≤0.01. For the entrapment search, the false discovery proportion (%) was calculated according to

$$\text{FDP (\%)} = \frac{N_\varepsilon}{N_\tau + N_\varepsilon} \times 100$$

Here, N_ε represents the number of *C. elegans* entrapment precursors, and N_τ represents the number of mouse target precursors. Precursors corresponding to both *C. elegans* and mouse protein identifiers were considered to be targets.

Figures were created in Python 3.11.5 using the matplotlib library.[45] Pearson correlations and linear regressions were calculated using 'pearsonr' and 'linregress' within SciPy.[41]

Results and Discussion

To explore the feasibility of SynchroSep-MS, we constructed two capillary LC columns and aligned their integrated emitters using a 3D-printed holder (**Figure 3.2A** and **Figure 3.2B**), placing them both in front of the atmospheric pressure inlet of an Orbitrap Astral MS system. Each column was connected to a nanoLC pump such that loading and gradient elution were controlled independently (**Figure 3.2C**). To distinguish which column/sample an analyte originates from, we developed an injection scheme (**Figure 3.2D**) that imparts a time delay between the two separations such that any given analyte present in both samples will be detected by the MS system across two elution profiles separated by a defined temporal gap (**Figure 3.2E** and **Figure 3.2F**).

Figure 3.3 presents initial results using the SynchroSep-MS technique to analyze tryptic peptides derived from mouse brain protein isolates. Here, identical samples were injected onto each column and analyzed using a 4-minute time delay and a DIA method iterating over *m/z* 380 – 980 with 4 *m/z* bins (Orbitrap Astral). For comparison, three chromatograms are shown – (**Figure 3.3A**) peptides loaded only on Column 1 and blank on Column 2, (**Figure 3.3B**) peptides loaded only on Column 2 and blank on Column 1, and (**Figure 3.3C**) peptides loaded on both columns. From **Figure 3.3C**, we observe that injection of peptides on both columns produces a chromatogram that appears as the superposition of the two single column chromatograms. Note the dual column injection resulted in a total ion intensity that is 94% of the sum of both single column injections. To determine how consistent the introduced retention time offset was across all peptides, we used the preliminary PSM files (.pepXML) generated by the MSFragger-DIA workflow[39] in FragPipe to construct a model of the retention time offset (ΔRT)**.** Specifically, duplicated PSMs were used to calculate the retention time offset throughout the separation, and

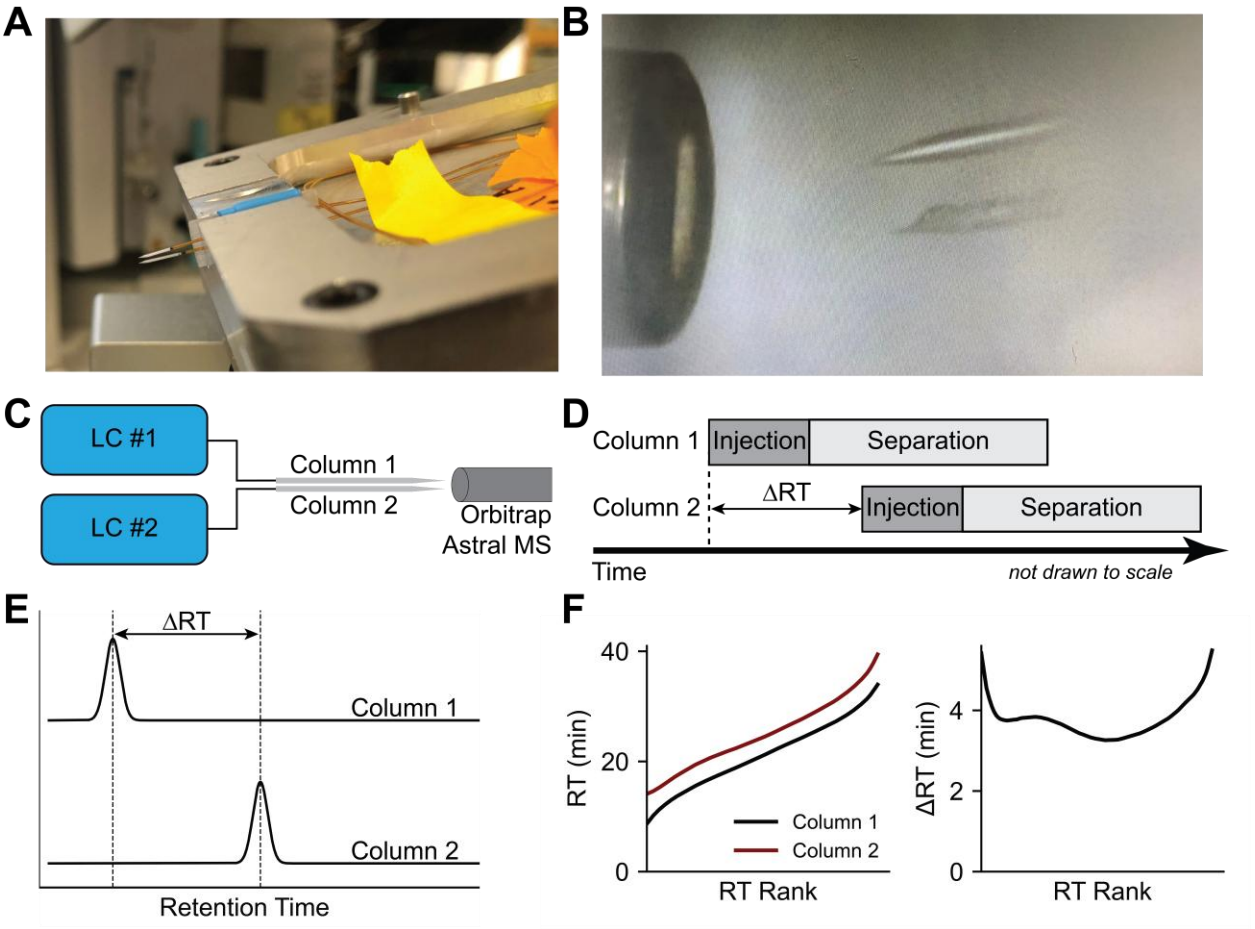

Figure 3.2. SynchroSep-MS Conceptual Overview. (A) Two packed capillary columns (75 µm I.D.) with integrated emitters installed in a custom-built column heater with a 3D-printed alignment piece. (B) Side view of emitters aligned to MS inlet. (C) Two (or more) LCs are connected to two (or more) columns, and the emitters are aligned to the mass spectrometer inlet. (D) An injection onto the first column is performed and a gradient separation is begun, followed by injecting a sample onto the second column and beginning a second gradient separation. The delay between sample injections can be controlled, allowing an adjustment of the resulting retention time offset (ΔRT). (E) The resulting data will have duplicate peaks in the chromatograms with a retention time offset of ΔRT, representing peaks from Column 1 and Column 2. (F) The retention time offset (ΔRT) can be modeled to determine the ΔRT throughout the gradient, which varies slightly throughout the gradient due to experimental variations in the columns and LC conditions. Image Description

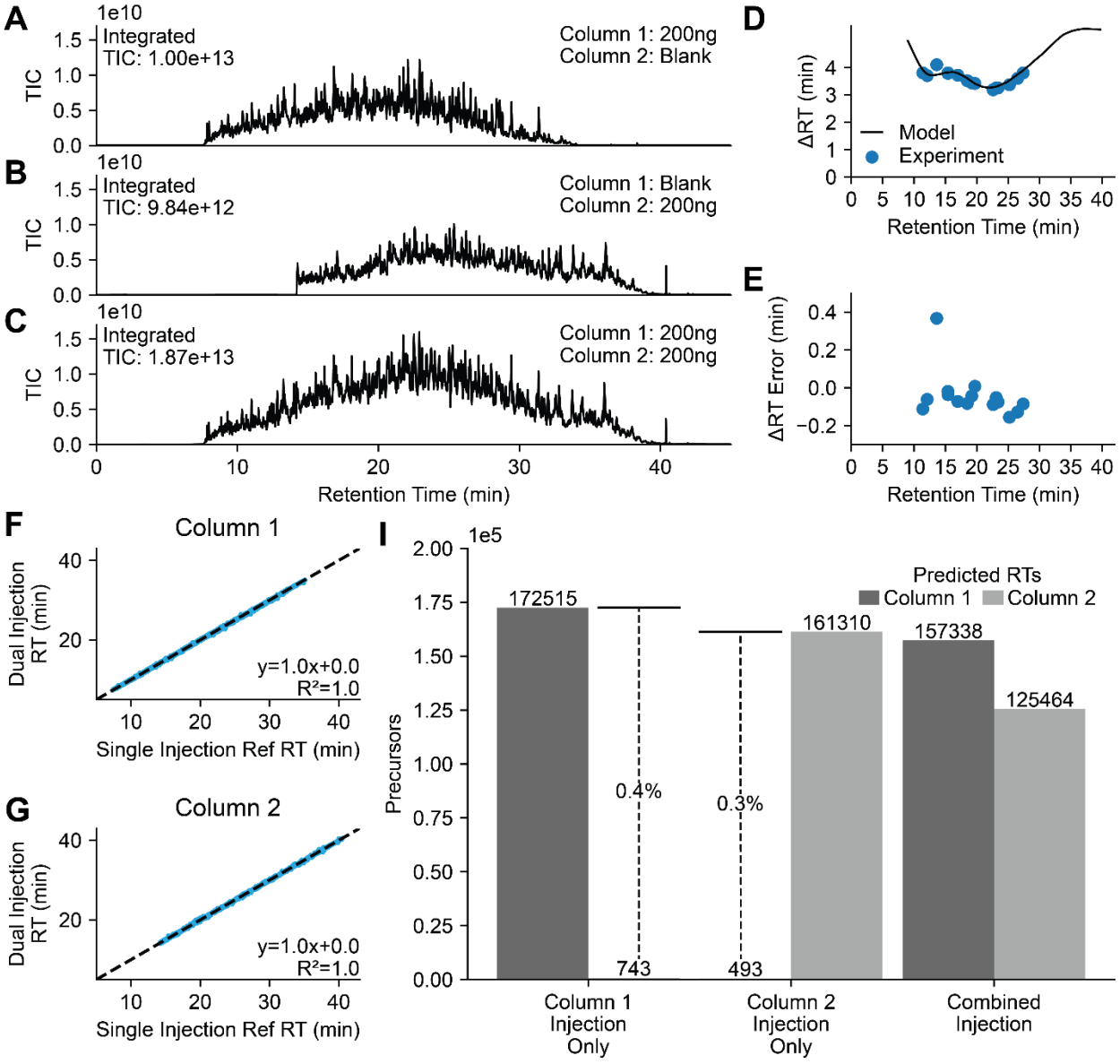

Figure 3.3. Assignment of Peaks to the Originating Column. (A) Representative chromatograms for 200 ng mouse brain peptide loads onto (A) Column 1-only, (B) Column 2-only, and (C) both columns at once are shown. Injections onto both columns results in a chromatogram that is essentially the sum of the individual injection chromatograms. (D) A LOWESS model of the retention time offset, ΔRT, is created based off of PSMs observed twice in a FragPipe-DIA search. Experimental ΔRT values observed for synthetic peptide standards (Pierce PRTC standard mix) are plotted as blue dots. The experimental ΔRT values are based off the observed retention time differences between PRTC peaks across 6 observations (3 single column injection each, 3 dual column injections). The LOWESS model was created from the data shown in (C). (E) The error in the ΔRT model is calculated based on the experimental values. Here, the model is accurate to ±0.4 min, which is small relative to the ~4-minute ΔRT shown in (B). Retention times of

precursors detected after searching the data shown in (A), (B), and (C) using the prototype DIA-NN software were plotted against each other for (F) Column 1 (172,515 total overlapping precursors used for the analysis) and (G) Column 2 (161,310 total overlapping precursors used for the analysis). Strong agreement between the RTs detected in the single-column injections and dual-column injections (as indicated by a linear regression slope of 1, y-intercept of 0, and R^2 of 1.0) indicate accurate peak assignments. (I) The single column injections for the chromatograms shown in (A) and (B) were searched against both the Column 1 and Column 2 predicted RTs, alongside the dual column injection in (C). False positive rates of 0.4% and 0.3% were observed for Column 1 and Column 2, respectively, indicating accurate column assignment. Image Description

then a local LOWESS regression model was fit to the data (**Figure 3.3D**). From these data, we conclude that the retention time offset was centered around 4 minutes; however, it varies throughout the chromatogram, likely due to experimental differences between chromatographic setups. That said, this variation was small relative to the absolute delay and this delay can be tuned experimentally (**Figure 3.4**). To further validate this conclusion, we injected a cocktail of 15 synthetic peptides onto both columns and repeated the experiment. Manual inspection of these results shows excellent agreement with the local regression model (**Figure 3.3E**).

From these data, we conclude that the SynchroSep-MS concept is viable, and, for full implementation, we next required an informatic workflow for global data processing. Accordingly, we modified the DIA-NN software to allow analysis of these results based on a retention time-aligned predicted library and the time offset (see **Experimental Section** and **Figure 3.1** for workflow details). Using the mouse brain peptide results described above, we validated the modified software's ability to accurately assign peptide precursors to the column from which it eluted. **Figure 3.3F** and **Figure 3.3G** plot the retention times of precursors identified in the single and dual injection data. The strong agreement confirms that the software is correctly assigning peptides to the originating column. **Figure 3.3I** shows the number of identified precursors assigned to each column across the three files. For example, when peptides are injected only on Column 1 (**Figure 3.3A**), 172,515 precursors were assigned to Column 1, while only 743 were assigned to Column 2. These 743 are, of course, false positives and constitute only 0.4% of the total precursors detected. Similar results were obtained when peptides were injected on Column 2 (**Figure 3.3B**). When samples were loaded onto both columns we detected the vast majority of peptides from the single column injections (91% and 78%) (**Figure 3.3C**). Validation of the retention time window

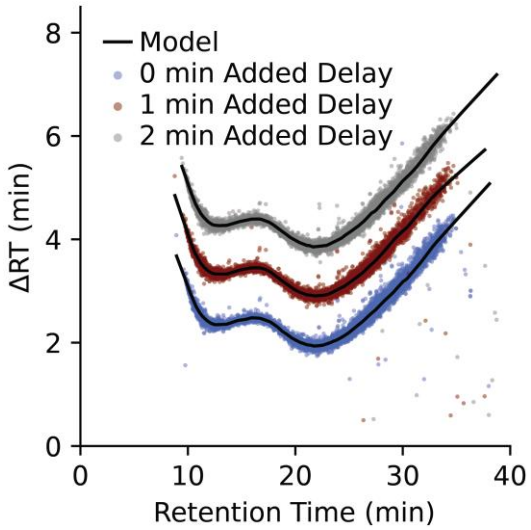

Figure 3.4. Tuning the Retention Time Offset. The ΔRT models for three analyses of a HAP1 digest with 0, 1 and 2 minutes added delay between the injections onto Column 1 and Column 2 are shown. Dots represent the ΔRT for features detected by MSFragger-DIA, and the lines represent the local regression model generated from these points. For this study, the '1 min Added Delay' method was selected as this method offered a balance between parallelization of the separations and a ΔRT magnitude large enough for confident column assignment. Note that the delay for the '0 min Added Delay' represents the minimum retention time offset achievable for this particular experimental setup, but this could be reduced by experimental adjustments. Image Description

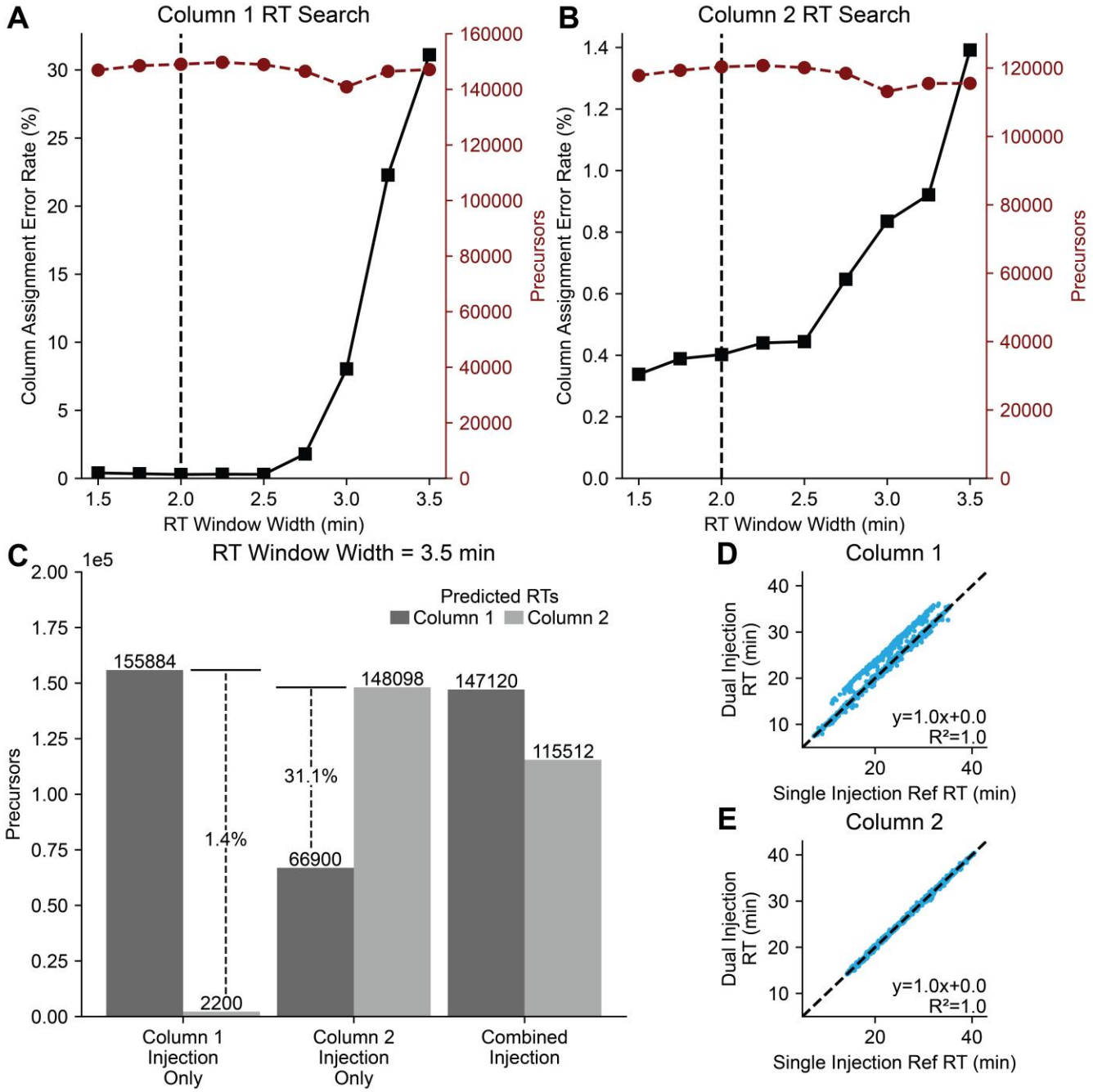

Figure 3.5. RT Window Width Scouting for DIA-NN Processing. The RT window width (which is controlled by the --rt-gap parameter in DIA-NN) dictates how widely around the predicted retention time DIA-NN should look when searching for peptide features. The error rate for assignments of precursors to an originating column can be assessed by searching data from a single-injection file with predicted retention times for both Column 1 and Column 2. For an injection of sample onto just Column 1 (with a blank on Column 2), any precursors detected with the predicted RTs for Column 2 would represent incorrect assignments. The error rate can then be calculated as the ratio of incorrect assignments/total assignments when using the predicted RTs for

Column 1. Note that the reported results are for analysis with a predicted library (*i.e.,* library-free), and we suppose that empirical DIA-based library (*e.g.,* made from bulk samples and used on multi-column analysis of single cells) could result in improved assignment accuracy. The calculated Column Assignment Error Rates (%) are shown for a Column 1-only injection (A) and a Column 2- only injection (B) alongside the total number of precursors detected when using the appropriate predicted retention times. The dashed line indicates the window width used for analysis in this study. As an example of what can go wrong if the RT window is too large, the results for a search using a 3.5-minute window width are shown in panels (C), (D), and (E). The single column injections for the chromatograms shown in **Figure 3.3A** and **Figure 3.3B** were searched against both the Column 1 and Column 2 predicted RTs, alongside the dual column injection in **Figure 3.3C**. False positive rates of 1.4% and 31.1% were observed for Column 1 and Column 2, respectively. The retention times of precursors detected for a dual-column and single-column injection for (E) Column 1 and (F) Column 2 are plotted against each other. Strong agreement between the RTs (as indicated by a linear regression slope of 1, y-intercept of 0, and R^2 of 1.0) in the single and dual column injections is observed for Column 2. However, the wider RT window leads to many precursors detected from Column 2 being assigned to Column 1, which is observed as a second cluster of precursors detected at an offset retention time relative to the reference retention time for the single column injection. Image Description

width for processing is demonstrated in **Figure 3.5**. We also performed an entrapment search of single- and dual-injection data on both columns against a combined *C. elegans*/Mouse protein database and observed consistent false discovery proportions across the different datasets, providing confidence that FDR control is not negatively impacted for this new type of data (**Figure 3.6**).

Figure 3.7A presents the number of proteins detected in the three experiments discussed above. To summarize: Column 1-only (~9,000), Column 2-only (~8,700), and Column 1 and 2 (*i.e.*, SynchroSep, ~8,700 from Column 1 and ~8,000 from Column 2 for a total of ~16,700). A more intuitive visual of these results is shown in a plot of proteins detected over time (**Figure 3.7B**) and here we see the additive nature of SynchroSep-MS and how nearly identical protein depth for each sample can be achieved with parallel analysis. The precursor-level results are shown in **Figure 3.8A** and **Figure 3.8B**. As expected, the protein identifications lost in the dual injections relative to the single injections correspond to lower intensity proteins (**Figure 3.9**). Next, we examined the quantitative reproducibility of the method. **Figure 3.7C** shows excellent precision for quantitation with median RSDs < 5% across >7000 protein groups (see **Figure 3.10** for zoom-in). Analysis at the precursor ion level shows similar trends (**Figure 3.8C**). These data confirm our guiding supposition that fast-scanning MS/MS technology is outpacing our relatively static approach to one LC column and on sample at a time.

One interesting observation in this dataset was that the depth achieved for Column 2 was more severely affected by the dual-column injection than Column 1 (**Figure 3.7A**). In this experimental setup, elution from Column 1 begins prior to and ends before elution of peptides from Column 2. Thus, during the first 5-6 minutes of elution from Column 1, there is no notable

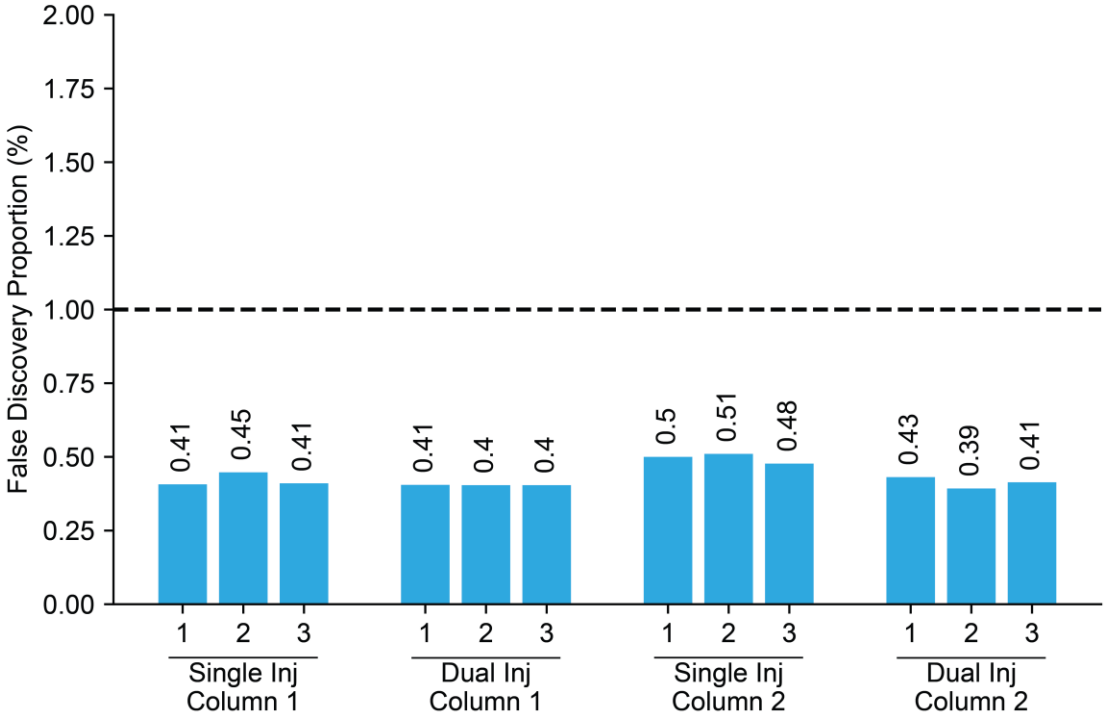

Figure 3.6. External FDR Assessment via Entrapment Analysis. The calculated false discovery proportion (%) values at the precursor level are shown for replicate injections on both columns. The dashed line indicates the 1% FDR level. 'Single Inj' refers to a datafile generated by mouse peptides injection onto only a single column with a blank loaded on the other. 'Dual Inj' refers to a data file generated by injecting mouse peptides onto both columns. The entrapment search was performed using the workflow described in **Figure 3.1** with a combined *C. Elegans*/mouse protein database. Image Description

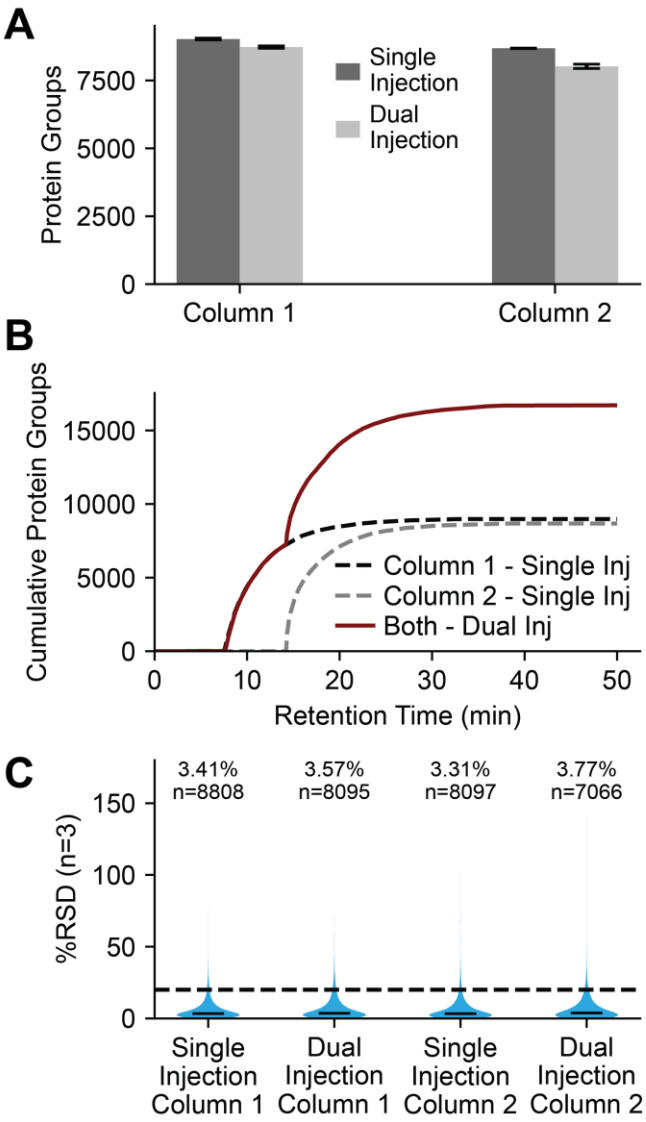

Figure 3.7. Protein Group Identifications. (A) The number of protein groups detected on Column 1 and Column 2 for a 200-ng mouse peptide load either as a single injection (with blank injected on the other column) or a dual column injection (200 ng mouse peptide load on the other column) are shown. Error bars represent the minimum/maximum values observed across triplicate injections. (B) The cumulative protein group identifications as a function of retention time are shown for a 200-ng load only on Column 1, a 200-ng load on Column 2, and 200-ng loads on both columns. For the injection on both columns, protein groups assigned to different columns are treated as unique identifications. (C) Protein groups detected across triplicate injections were used to calculate %RSD values for quantification for the dual- and single-injections onto Column 1 and Column 2. The median %RSD and number of protein groups included in each dataset are shown above each violin plot. The dashed line indicates %RSD = 20. Image Description

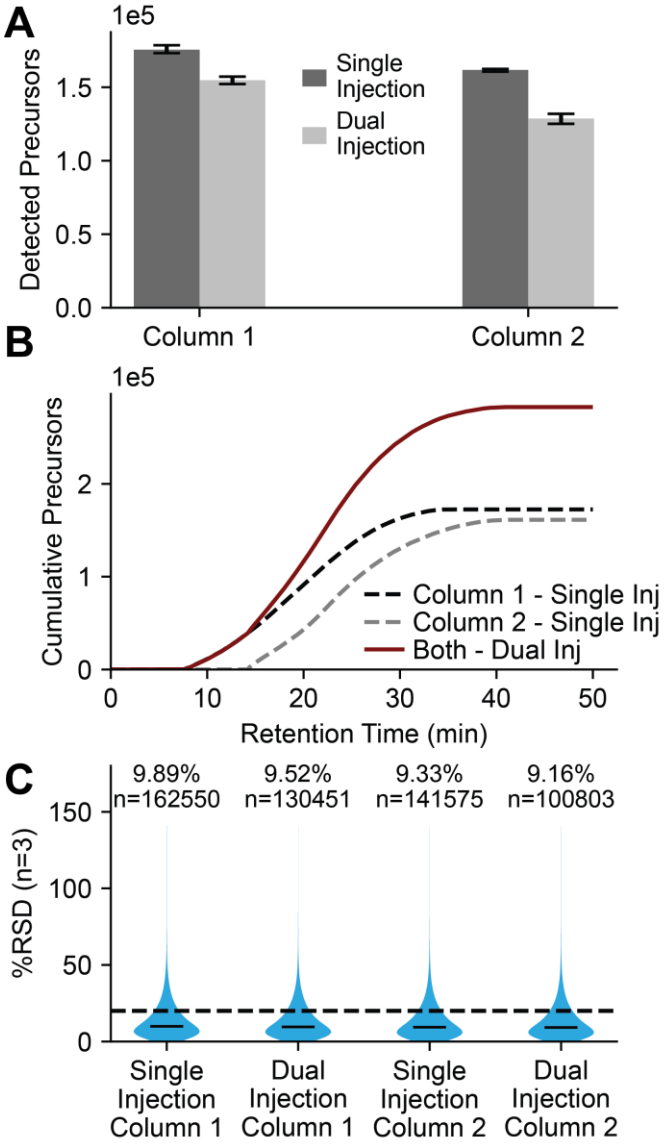

Figure 3.8. Precursor Identifications. (A) The number of precursors detected on Column 1 and Column 2 for a 200-ng mouse peptide load either as a single injection (with blank injected on the other column) or a dual column injection (200-ng mouse peptide load on the other column). Error bars represent the minimum/maximum values observed across triplicate injections. (B) The cumulative precursor identifications as a function of retention time are shown for a 200-ng load only on Column 1, a 200=ng load on Column 2, and 200-ng loads on both columns. For the injection on both columns, precursors assigned to different columns are treated as unique identifications. (C) Precursors detected across triplicate injections were used to calculate %RSD values for quantification for the dual- and single-injections onto Column 1 and Column 2. The median %RSD values and number of precursors included in each dataset are shown above each violin plot. The dashed line indicates %RSD = 20. Image Description

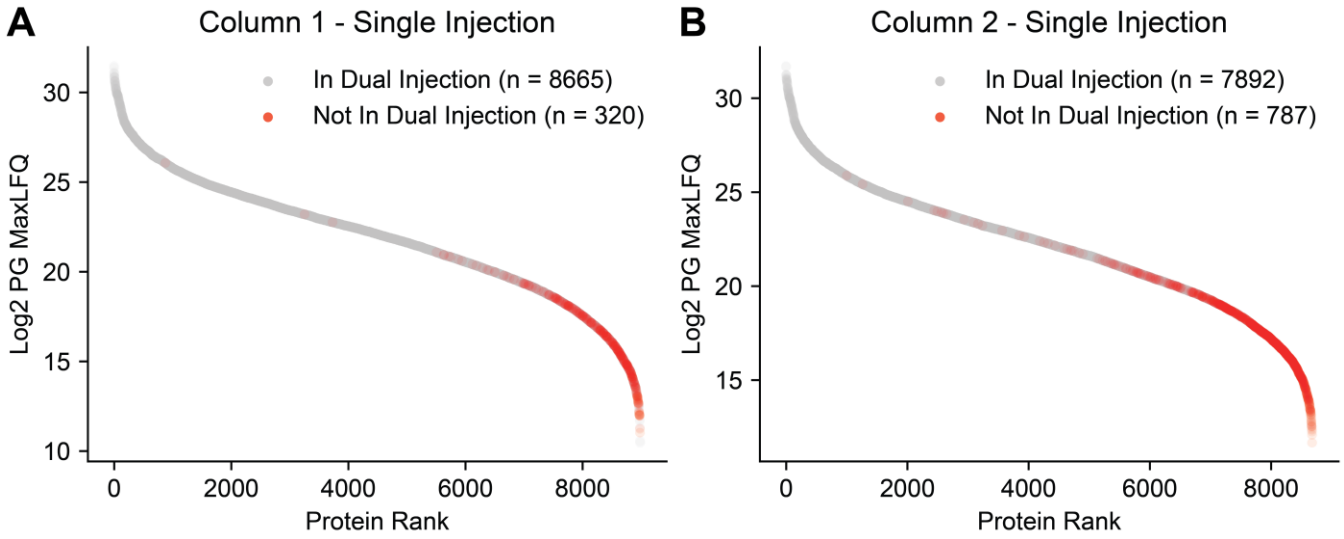

Figure 3.9. Dual vs. Single-Injection Dynamic Range. Protein group intensities for the single injections of mouse brain are plotted in ranked order for (A) Column 1 (corresponding to **Figure 3.3A**) and (B) Column 2 (corresponding to **Figure 3.3B**). Grey dots represent protein groups also detected for that column in the dual column injection (**Figure 3.3C**) and red dots represent proteins groups not detected in the dual column injection. Image Description

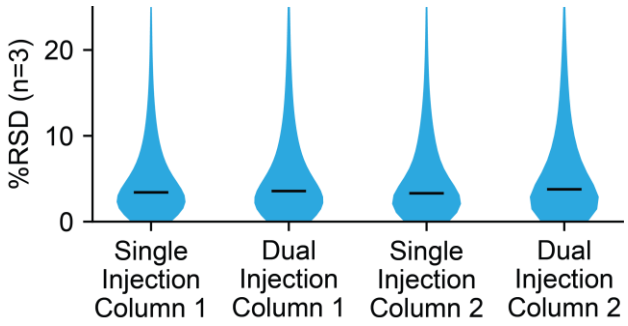

Figure 3.10. Protein Group RSD Zoom In. A zoomed in view of the protein group %RSD distributions shown in **Figure 3.7C** is shown. Note that this visualization cuts off the distribution at a %RSD of 25. Image Description

signal from Column 2 (**Figure 3.3A-C**). Because most precursor identifications are generated earlier in the gradient (**Figure 3.8B**), the identification performance of Column 1 might uniquely benefit from the absence of signal from Column 2 that would increase spectral complexity (and decrease the likelihood of identifications due to potential decreased dynamic range, signal-to-noise, etc.). In contrast, the identification rate for peptides eluting later from Column 2 might benefit from the lack of interference from Column 1 at the end of the gradient, although the lower rate of precursor identification observed at the end of the gradient for the single-column injections (**Figure 3.8B**) suggests that this might not compensate for the loss of identifications earlier in the gradient. To further investigate this phenomenon, we binned the precursors identified in the single-injection data by retention time and calculated the proportion of these precursors that were missing in the dual-injection identifications (**Figure 3.11**). With this analysis, we observe that Column 2 has a higher proportion of missing precursor identifications in the earlier part of the gradient, while Column 2 has a higher proportion of missing precursors in the later part of the gradient. Additionally, the magnitude of the trend seems to be stronger for Column 2 as suggested by the data in **Figure 3.8B**. This aspect of our SynchroSep-MS dataset helps to explain why Column 2 demonstrates a worse performance loss for dual-injections than Column 1 and suggests that increased spectral complexity caused by the addition of more columns for a higher degree of multiplexing would further diminish the identification performance. However, we hypothesize that the continued development of faster and more sensitive mass spectrometers will improve capacity to handle complex mixtures and decrease the magnitude of this phenomenon.

Ultimately, SynchroSep-MS provides an ideal and novel avenue for multiplexed and high-throughput proteome analysis. To achieve this goal, the approach must also provide reliable

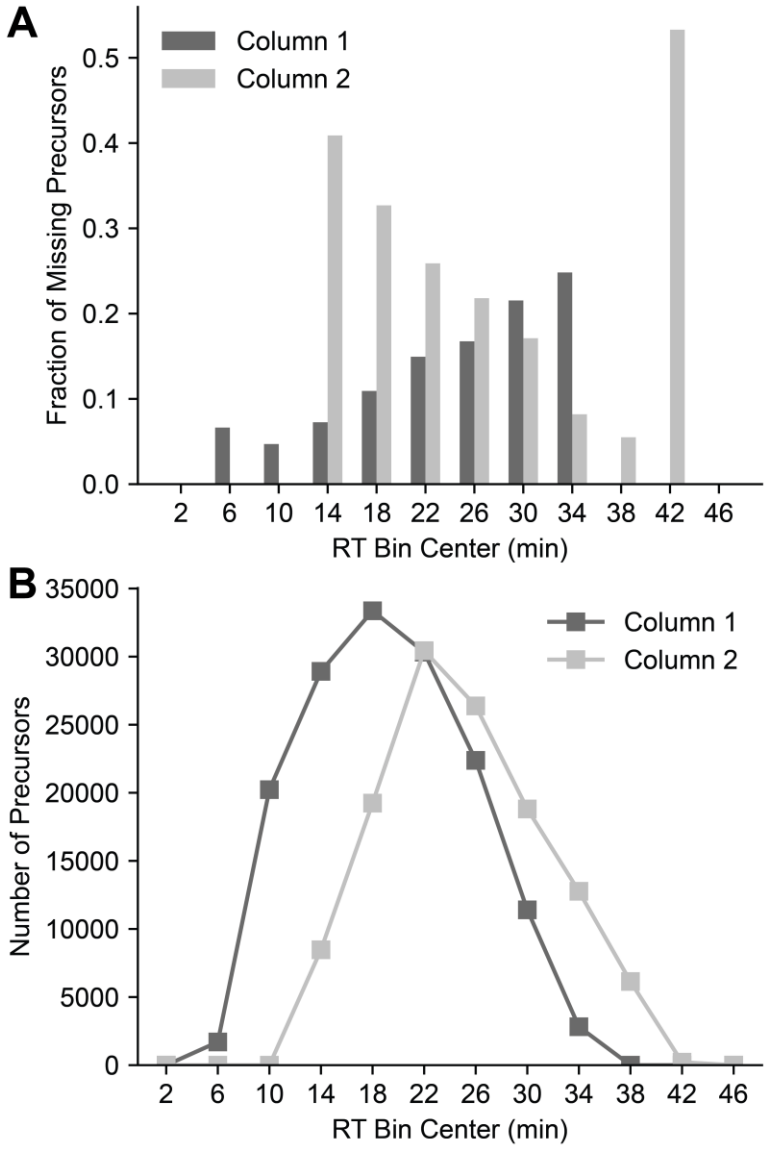

Figure 3.11. RT Binning Analysis of Dual Column SynchroSep-MS Data. (A) Precursors detected in the single-injection files for each column were binned by their detected retention time (RT). Then the fraction of these precursors that were missing in the dual-injection files was calculated and plotted as a function of the RT bin. (B) The number of precursors detected in the single-injection files were plotted as a function of the RT bin. Image Description

quantification. Accordingly, we performed two dilution series experiments in triplicate. The first reduced the amount of peptides loaded on each column equally across a range (Direct) and the second evaluated effect of interference between columns by varying the loading amounts in opposite directions (Inverted) (**Table 3.1** and **Table 3.2**).

Table 3.1. Mouse Peptide Direct Dilution Series Scheme

Column 1	Column 2
10 ng	10 ng
50 ng	50 ng
100 ng	100 ng
200 ng	200 ng

Table 3.2. Mouse Peptide Inverted Dilution Series Scheme

Column 1	Column 2
200 ng	10 ng
100 ng	50 ng
50 ng	100 ng
10 ng	200 ng

The Direct dilution series results are as expected – that is, as the load level is reduced the protein identifications also drop (**Figure 3.12A**, see **Figure 3.13A** for precursor level results). Importantly, similar protein identification trends are observed in the Inverted dilution series, except that they move in opposing directions (**Figure 3.12B**, see **Figure 3.13B** for precursor level results). We note that in both experiments, Column 1 outperforms Column 2, which is likely due to variation in separation quality and/or emitter positioning. To evaluate quantitative performance, we calculated R^2 values across these data and plotted the distributions for each dilution series in **Figure 3.12C**. The median R^2 value was over 0.95 for all experiments (and over ~0.9 at the precursor level, **Figure 3.13C**), suggesting excellent quantitative performance. Another

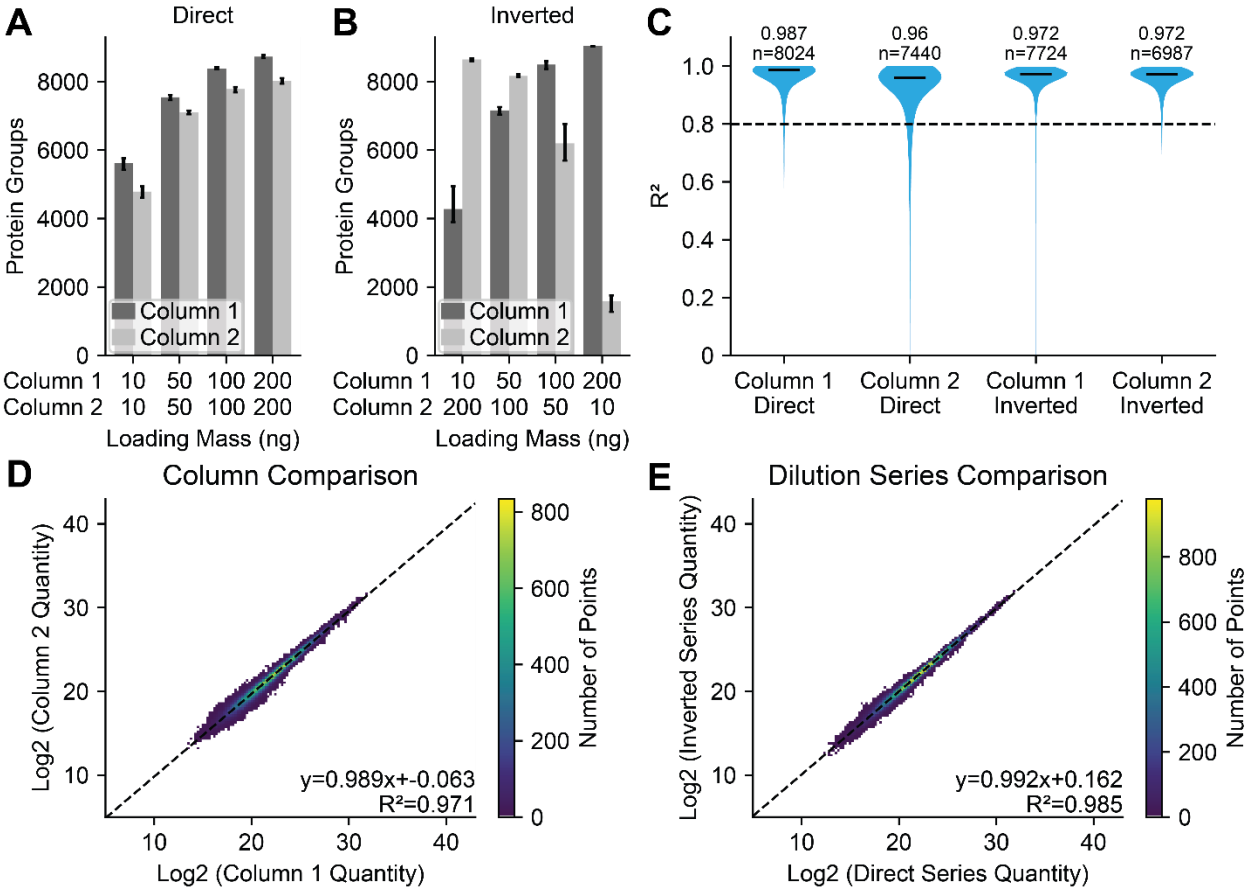

Figure 3.12. Performance of SynchroSep-MS Method Assessed with Dilution Series. The number of detected protein groups are shown for (A) the Direction Dilution Series (**Table 3.1**) and (B) the Inverted Dilution Series (**Table 3.2**).The direct dilution series consists of injections of the same loading mass onto both columns, whereas the inverted dilution series uses different loads to represent an extreme case of varied background matrix from the other column. Error bars represent the minimum/maximum values observed across triplicate injections. (C) The R^2 value distributions for Column 1 and Column 2 across the dilution series shown in **Table 3.1** and **Table 3.2** are shown. Values were only calculated for protein groups detected across at least three concentration points. The median R^2 value and number of protein groups included in each distribution are shown above each plot. The dashed line indicates an R^2 value of 0.8. (D) The log2-transformed quantities for protein groups detected across both Column 1 and Column 2 in the Direct Dilution series are plotted against each other. A slope less than 1 indicates that Column 1 protein quantities are generally higher. (E) The log2-transformed quantities for protein groups detected across both dilution series are plotted against each other. A slope less than 1 indicates that the direct dilution series protein quantities are generally higher. For this example, data on Column 1 is shown. Image Description

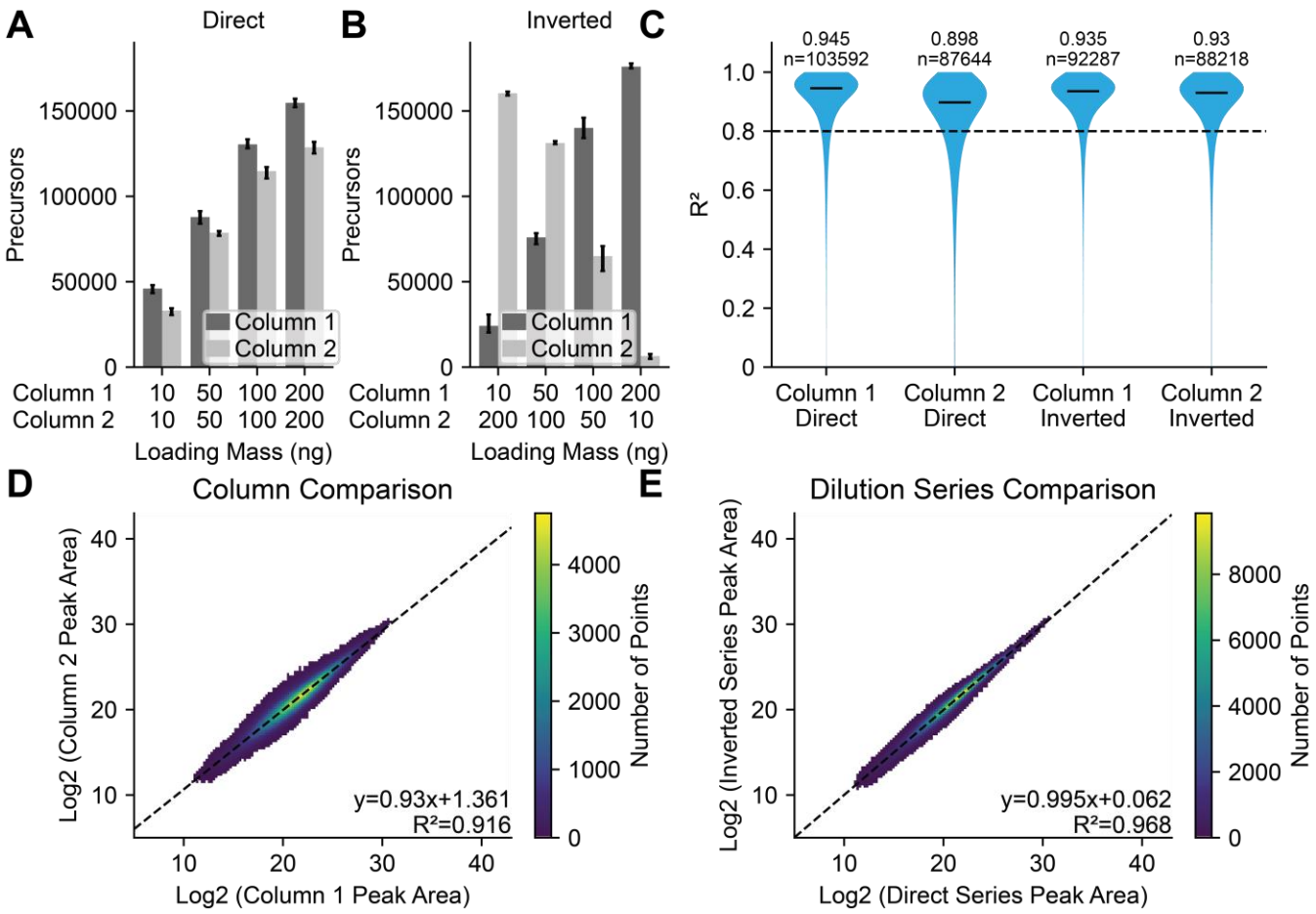

Figure 3.13. Performance of SynchroSep-MS Method Assessed with Dilution Series – Precursor Level. The number of precursors are shown for (A) the Direction Dilution Series (**Table 3.1**) and (B) the Inverted Dilution Series (**Table 3.2**). The direct dilution series consists of injections of the same loading mass onto both columns, whereas the inverted dilution series uses different loads to represent an extreme case of varied background matrix from the other column. Error bars represent the minimum/maximum values observed across triplicate injections. (C) The R^2 value distributions for Column 1 and Column 2 across the dilution series shown in **Table 3.1** and **Table 3.2** are shown. Values were only calculated for precursors detected across at least three concentration points. The median R^2 value and number of precursors included in each distribution are shown above each plot. The dashed line indicates an R^2 value of 0.8. (D) The log2-transformed peak areas for precursors detected across both Column 1 and Column 2 in the Direct Dilution series are plotted against each other. A slope less than 1 indicates that Column 1 peak areas are generally higher. (E) The log2-transformed peak areas for precursors detected across both dilution series are plotted against each other. A slope less than 1 indicates that the direct dilution series peak areas are generally higher. For this example, data on Column 1 is shown. Image Description

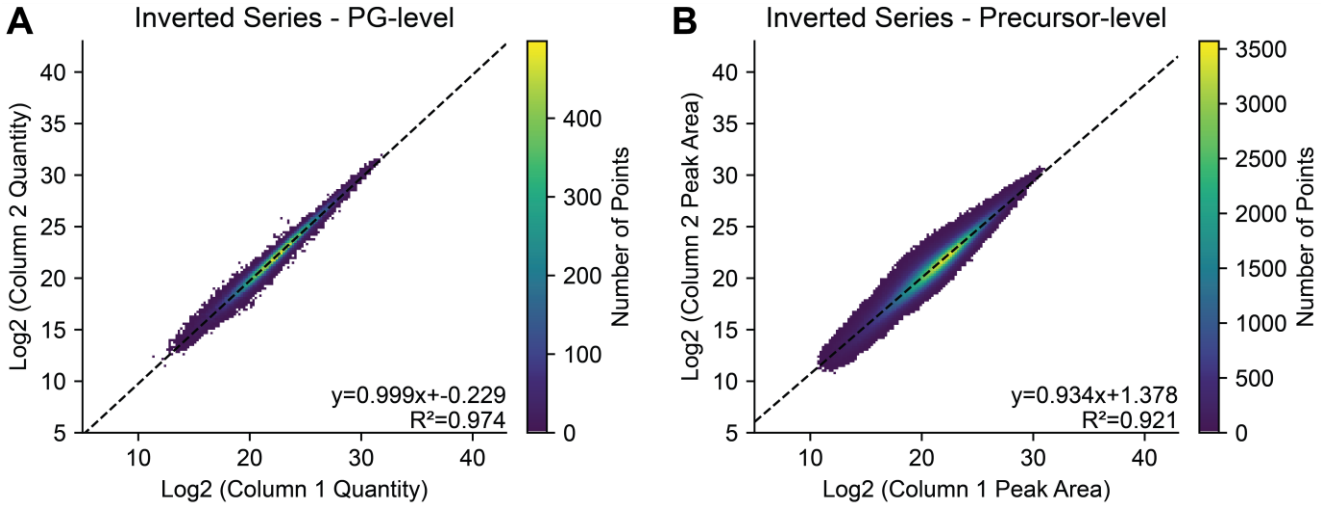

Figure 3.14. Column Intensity Comparisons for the Inverted Dilution Series. The log2-transformed quantities for (A) protein groups and (B) precursors detected across both Column 1 and Column 2 in the Direct Dilution series are plotted against each other. A slope less than 1 indicates that Column 1 protein quantities are generally higher. Image Description

performance requirement for SynchroSep-MS is that the measured quantities across columns are consistent. We assessed this figure of merit by plotting the quantities of proteins detected across Columns 1 and 2 in the Direct Dilution Series against each other and observed excellent correlation (**Figure 3.12D**). A slope of 0.989 and y-intercept of 0.063 indicate that the quantities are nearly identical with Column 1 having slightly higher abundances. The same analysis at the precursor level yields a slope of 0.93, indicating a similar trend (**Figure 3.13D)**. Similar results were observed when performing the same analysis for the Inverted Dilution series (**Figure 3.14**). This difference can easily be corrected by normalization approaches and likely be eliminated by optimization of the experimental setup (*i.e.*, emitter positioning). We also compared the abundances measured across the two dilution series schemes on Column 1 as shown in **Figure 3.12E.** Strong agreement at the protein quantity level (and the precursor level, **Figure 3.13E**) indicates that the method is able to selectively quantify analytes from both columns without notable interference.

Conclusions

Using SynchroSep-MS, we can simultaneously analyze two proteomes, acquiring nearly double the peptide information from a single proteome analysis, demonstrating that modern, fast-scanning mass analyzers are capable of handling the increased analysis complexity of this method. Our data also shows high quantitative precision and reproducibility for analysis of a complex mammalian proteome. Further dilution series experiments demonstrate good quantitative linearity and suggest that minimal matrix interference is present for the multiplexed analysis.

We conclude that the SynchroSep-MS method represents a new paradigm for data collection and the first example of label-free multiplexed proteome analysis. Although this proof-of-concept work demonstrates the method using two columns, we envision a multi-column setup with four or more columns that could significantly improve analysis throughput. This multiplexed setup would be especially attractive for large-scale clinical cohorts and single-cell analyses.

The main current limitation with large-scale implementation of SynchroSep-MS is the relative complexity of the LC pump setup. In our current efforts, each column has an independent pump. Moving forward, we envision splitting the flow of a single pump across the various columns while still maintaining the ability to load each independently. Elimination of these technical hurdles will allow a very direct path to accelerating the throughput of proteome analysis. Further, as MS analyzers continue to increase in acquisition speed, SynchroSep-MS offers a strategy to boost throughput without having to compress gradient lengths – which doubtless causes reduced peak capacity and ionization suppression of the low abundance analytes.

Author Contributions.

NML, KAO, STQ and JJC planned the project. NML prepared samples. NML and LYC collected data. NML, LYC, BZ, BJA, DAP, AN, and VD performed data analysis. DAP, AN, and VD provided software support. NML, KAO, STQ, JJC drafted the manuscript. All authors contributed to the revising and editing of the manuscript.

References

(1) Bekker-Jensen, D. B.; Kelstrup, C. D.; Batth, T. S.; Larsen, S. C.; Haldrup, C.; Bramsen, J. B.; Sørensen, K. D.; Høyer, S.; Ørntoft, T. F.; Andersen, C. L.; Nielsen, M. L.; Olsen, J. V. An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes. *Cell Syst.* **2017**, *4* (6), 587-599.e4. https://doi.org/10.1016/j.cels.2017.05.009.

(2) Sinitcyn, P.; Richards, A. L.; Weatheritt, R. J.; Brademan, D. R.; Marx, H.; Shishkova, E.; Meyer, J. G.; Hebert, A. S.; Westphall, M. S.; Blencowe, B. J.; Cox, J.; Coon, J. J. Global Detection of Human Variants and Isoforms by Deep Proteome Sequencing. *Nat. Biotechnol.* **2023**, *41* (12), 1776–1786. https://doi.org/10.1038/s41587-023-01714-x.

(3) Heil, L. R.; Damoc, E.; Arrey, T. N.; Pashkova, A.; Denisov, E.; Petzoldt, J.; Peterson, A. C.; Hsu, C.; Searle, B. C.; Shulman, N.; Riffle, M.; Connolly, B.; MacLean, B. X.; Remes, P. M.; Senko, M. W.; Stewart, H. I.; Hock, C.; Makarov, A. A.; Hermanson, D.; Zabrouskov, V.; Wu, C. C.; MacCoss, M. J. Evaluating the Performance of the Astral Mass Analyzer for Quantitative Proteomics Using Data-Independent Acquisition. *J. Proteome Res.* **2023**, *22* (10), 3290–3300. https://doi.org/10.1021/acs.jproteome.3c00357.

(4) Stewart, H. I.; Grinfeld, D.; Giannakopulos, A.; Petzoldt, J.; Shanley, T.; Garland, M.; Denisov, E.; Peterson, A. C.; Damoc, E.; Zeller, M.; Arrey, T. N.; Pashkova, A.; Renuse, S.; Hakimi, A.; Kühn, A.; Biel, M.; Kreutzmann, A.; Hagedorn, B.; Colonius, I.; Schütz, A.; Stefes, A.; Dwivedi, A.; Mourad, D.; Hoek, M.; Reitemeier, B.; Cochems, P.; Kholomeev, A.; Ostermann, R.; Quiring, G.; Ochmann, M.; Möhring, S.; Wagner, A.; Petker, A.; Kanngiesser, S.; Wiedemeyer, M.; Balschun, W.; Hermanson, D.; Zabrouskov, V.; Makarov, A. A.; Hock, C. Parallelized Acquisition of Orbitrap and Astral Analyzers Enables High-Throughput Quantitative Analysis. *Anal. Chem.* **2023**, *95* (42), 15656–15664. https://doi.org/10.1021/acs.analchem.3c02856.

(5) Guzman, U. H.; Martinez-Val, A.; Ye, Z.; Damoc, E.; Arrey, T. N.; Pashkova, A.; Renuse, S.; Denisov, E.; Petzoldt, J.; Peterson, A. C.; Harking, F.; Østergaard, O.; Rydbirk, R.; Aznar, S.; Stewart, H.; Xuan, Y.; Hermanson, D.; Horning, S.; Hock, C.; Makarov, A.; Zabrouskov, V.; Olsen, J. V. Ultra-Fast Label-Free Quantification and Comprehensive Proteome Coverage with Narrow-Window Data-Independent Acquisition. *Nat. Biotechnol.* **2024**. https://doi.org/10.1038/s41587-023-02099-7.

(6) Lancaster, N. M.; Sinitcyn, P.; Forny, P.; Peters-Clarke, T. M.; Fecher, C.; Smith, A. J.; Shishkova, E.; Arrey, T. N.; Pashkova, A.; Robinson, M. L.; Arp, N.; Fan, J.; Hansen, J.; Galmozzi, A.; Serrano, L. R.; Rojas, J.; Gasch, A. P.; Westphall, M. S.; Stewart, H.; Hock, C.; Damoc, E.; Pagliarini, D. J.; Zabrouskov, V.; Coon, J. J. Fast and Deep Phosphoproteome Analysis with the Orbitrap Astral Mass Spectrometer. *Nat. Commun.* **2024**, *15* (1), 7016. https://doi.org/10.1038/s41467-024-51274-0.

(7) Meier, F.; Beck, S.; Grassl, N.; Lubeck, M.; Park, M. A.; Raether, O.; Mann, M. Parallel Accumulation–Serial Fragmentation (PASEF): Multiplying Sequencing Speed and Sensitivity by Synchronized Scans in a Trapped Ion Mobility Device. *J. Proteome Res.* **2015**, *14* (12), 5378–5387. https://doi.org/10.1021/acs.jproteome.5b00932.

(8) Meier, F.; Brunner, A.-D.; Koch, S.; Koch, H.; Lubeck, M.; Krause, M.; Goedecke, N.; Decker, J.; Kosinski, T.; Park, M. A.; Bache, N.; Hoerning, O.; Cox, J.; Räther, O.; Mann, M. Online Parallel Accumulation–Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer*. *Mol. Cell. Proteomics* **2018**, *17* (12), 2534–2545. https://doi.org/10.1074/mcp.TIR118.000900.

(9) Meier, F.; Brunner, A.-D.; Frank, M.; Ha, A.; Bludau, I.; Voytik, E.; Kaspar-Schoenefeld, S.; Lubeck, M.; Raether, O.; Bache, N.; Aebersold, R.; Collins, B. C.; Röst, H. L.; Mann, M. diaPASEF: Parallel Accumulation–Serial Fragmentation Combined with Data-Independent Acquisition. *Nat. Methods* **2020**, *17* (12), 1229–1236. https://doi.org/10.1038/s41592-020-00998-0.

(10) Skowronek, P.; Thielert, M.; Voytik, E.; Tanzer, M. C.; Hansen, F. M.; Willems, S.; Karayel, O.; Brunner, A.-D.; Meier, F.; Mann, M. Rapid and In-Depth Coverage of the (Phospho-)Proteome With Deep Libraries and Optimal Window Design for Dia-PASEF. *Mol. Cell. Proteomics* **2022**, *21* (9). https://doi.org/10.1016/j.mcpro.2022.100279.

(11) Serrano, L. R.; Peters-Clarke, T. M.; Arrey, T. N.; Damoc, N. E.; Robinson, M. L.; Lancaster, N. M.; Shishkova, E.; Moss, C.; Pashkova, A.; Sinitcyn, P.; Brademan, D. R.; Quarmby, S. T.; Peterson, A. C.; Zeller, M.; Hermanson, D.; Stewart, H.; Hock, C.; Makarov, A.; Zabrouskov, V.; Coon, J. J. The One Hour Human Proteome. *Mol. Cell. Proteomics* **2024**, *0* (0). https://doi.org/10.1016/j.mcpro.2024.100760.

(12) Shen, Y.; Smith, R. D.; Unger, K. K.; Kumar, D.; Lubda, D. Ultrahigh-Throughput Proteomics Using Fast RPLC Separations with ESI-MS/MS. *Anal. Chem.* **2005**, *77* (20), 6692–6701. https://doi.org/10.1021/ac050876u.

(13) Shen, Y.; Strittmatter, E. F.; Zhang, R.; Metz, T. O.; Moore, R. J.; Li, F.; Udseth, H. R.; Smith, R. D.; Unger, K. K.; Kumar, D.; Lubda, D. Making Broad Proteome Protein Measurements in 1−5 Min Using High-Speed RPLC Separations and High-Accuracy Mass Measurements. *Anal. Chem.* **2005**, *77* (23), 7763–7773. https://doi.org/10.1021/ac051257o.

(14) Bache, N.; Geyer, P. E.; Bekker-Jensen, D. B.; Hoerning, O.; Falkenby, L.; Treit, P. V.; Doll, S.; Paron, I.; Müller, J. B.; Meier, F.; Olsen, J. V.; Vorm, O.; Mann, M. A Novel LC System Embeds Analytes in Pre-Formed Gradients for Rapid, Ultra-Robust Proteomics*. *Mol. Cell. Proteomics* **2018**, *17* (11), 2284–2296. https://doi.org/10.1074/mcp.TIR118.000853.

(15) Bian, Y.; Zheng, R.; Bayer, F. P.; Wong, C.; Chang, Y.-C.; Meng, C.; Zolg, D. P.; Reinecke, M.; Zecha, J.; Wiechmann, S.; Heinzlmeir, S.; Scherr, J.; Hemmer, B.; Baynham, M.;

Gingras, A.-C.; Boychenko, O.; Kuster, B. Robust, Reproducible and Quantitative Analysis of Thousands of Proteomes by Micro-Flow LC–MS/MS. *Nat. Commun.* **2020**, *11* (1), 157. https://doi.org/10.1038/s41467-019-13973-x.

(16) Bekker-Jensen, D. B.; Martínez-Val, A.; Steigerwald, S.; Rüther, P.; Fort, K. L.; Arrey, T. N.; Harder, A.; Makarov, A.; Olsen, J. V. A Compact Quadrupole-Orbitrap Mass Spectrometer with FAIMS Interface Improves Proteome Coverage in Short LC Gradients *. *Mol. Cell. Proteomics* **2020**, *19* (4), 716–729. https://doi.org/10.1074/mcp.TIR119.001906.

(17) Messner, C. B.; Demichev, V.; Bloomfield, N.; Yu, J. S. L.; White, M.; Kreidl, M.; Egger, A.-S.; Freiwald, A.; Ivosev, G.; Wasim, F.; Zelezniak, A.; Jürgens, L.; Suttorp, N.; Sander, L. E.; Kurth, F.; Lilley, K. S.; Mülleder, M.; Tate, S.; Ralser, M. Ultra-Fast Proteomics with Scanning SWATH. *Nat. Biotechnol.* **2021**, *39* (7), 846–854. https://doi.org/10.1038/s41587-021-00860-4.

(18) Ishikawa, M.; Konno, R.; Nakajima, D.; Gotoh, M.; Fukasawa, K.; Sato, H.; Nakamura, R.; Ohara, O.; Kawashima, Y. Optimization of Ultrafast Proteomics Using an LC-Quadrupole-Orbitrap Mass Spectrometer with Data-Independent Acquisition. *J. Proteome Res.* **2022**, *21* (9), 2085–2093. https://doi.org/10.1021/acs.jproteome.2c00121.

(19) Szyrwiel, L.; Gille, C.; Mülleder, M.; Demichev, V.; Ralser, M. Fast Proteomics with Dia-PASEF and Analytical Flow-Rate Chromatography. *PROTEOMICS* **2024**, *24* (1–2), 2300100. https://doi.org/10.1002/pmic.202300100.

(20) Bonneil, E.; Tessier, S.; Carrier, A.; Thibault, P. Multiplex Multidimensional nanoLC-MS System for Targeted Proteomic Analyses. *ELECTROPHORESIS* **2005**, *26* (24), 4575–4589. https://doi.org/10.1002/elps.200500603.

(21) Livesay, E. A.; Tang, K.; Taylor, B. K.; Buschbach, M. A.; Hopkins, D. F.; LaMarche, B. L.; Zhao, R.; Shen, Y.; Orton, D. J.; Moore, R. J.; Kelly, R. T.; Udseth, H. R.; Smith, R. D. Fully Automated Four-Column Capillary LC−MS System for Maximizing Throughput in Proteomic Analyses. *Anal. Chem.* **2008**, *80* (1), 294–302. https://doi.org/10.1021/ac701727r.

(22) Orton, D. J.; Wall, M. J.; Doucette, A. A. Dual LC–MS Platform for High-Throughput Proteome Analysis. *J. Proteome Res.* **2013**, *12* (12), 5963–5970. https://doi.org/10.1021/pr400738a.

(23) Lee, H.; Lee, J. H.; Kim, H.; Kim, S.-J.; Bae, J.; Kim, H. K.; Lee, S.-W. A Fully Automated Dual-Online Multifunctional Ultrahigh Pressure Liquid Chromatography System for High-Throughput Proteomics Analysis. *J. Chromatogr. A* **2014**, *1329*, 83–89. https://doi.org/10.1016/j.chroma.2013.12.084.

(24) Lee, H.; Mun, D.-G.; Bae, J.; Kim, H.; Yeon Oh, S.; Soo Park, Y.; Lee, J.-H.; Lee, S.-W. A Simple Dual Online Ultra-High Pressure Liquid Chromatography System (sDO-UHPLC) for

High Throughput Proteome Analysis. *Analyst* **2015**, *140* (16), 5700–5706. https://doi.org/10.1039/C5AN00639B.

(25) Webber, K. G. I.; Truong, T.; Johnston, S. M.; Zapata, S. E.; Liang, Y.; Davis, J. M.; Buttars, A. D.; Smith, F. B.; Jones, H. E.; Mahoney, A. C.; Carson, R. H.; Nwosu, A. J.; Heninger, J. L.; Liyu, A. V.; Nordin, G. P.; Zhu, Y.; Kelly, R. T. Label-Free Profiling of up to 200 Single-Cell Proteomes per Day Using a Dual-Column Nanoflow Liquid Chromatography Platform. *Anal. Chem.* **2022**, *94* (15), 6017–6025. https://doi.org/10.1021/acs.analchem.2c00646.

(26) Kreimer, S.; Haghani, A.; Binek, A.; Hauspurg, A.; Seyedmohammad, S.; Rivas, A.; Momenzadeh, A.; Meyer, J. G.; Raedschelders, K.; Van Eyk, J. E. Parallelization with Dual-Trap Single-Column Configuration Maximizes Throughput of Proteomic Analysis. *Anal. Chem.* **2022**, *94* (36), 12452–12460. https://doi.org/10.1021/acs.analchem.2c02609.

(27) Kreimer, S.; Binek, A.; Chazarin, B.; Cho, J. H.; Haghani, A.; Hutton, A.; Marbán, E.; Mastali, M.; Meyer, J. G.; Mesquita, T.; Song, Y.; Van Eyk, J.; Parker, S. High-Throughput Single-Cell Proteomic Analysis of Organ-Derived Heterogeneous Cell Populations by Nanoflow Dual-Trap Single-Column Liquid Chromatography. *Anal. Chem.* **2023**, *95* (24), 9145–9150. https://doi.org/10.1021/acs.analchem.3c00213.

(28) Staes, A.; Boucher, K.; Dufour, S.; Maia, T. M.; Timmerman, E.; Haver, D. V.; Pauwels, J.; Demol, H.; Vandenbussche, J.; Gevaert, K.; Impens, F.; Devos, S. High-Throughput Nanoflow Proteomics Using a Dual-Column Electrospray Source. *Anal. Chem.* **2024**, *96* (17), 6534–6539. https://doi.org/10.1021/acs.analchem.4c00845.

(29) Chen, L.; Zhang, Z.; Matsumoto, C.; Gao, Y. High-Throughput Proteomics Enabled by a Fully Automated Dual-Trap and Dual-Column LC-MS. *Anal. Chem.* **2024**, *96* (24), 9761–9766. https://doi.org/10.1021/acs.analchem.3c03182.

(30) Xie, X.; Truong, T.; Huang, S.; Johnston, S. M.; Hovanski, S.; Robinson, A.; Webber, K. G. I.; Lin, H.-J. L.; Mun, D.-G.; Pandey, A.; Kelly, R. T. Multicolumn Nanoflow Liquid Chromatography with Accelerated Offline Gradient Generation for Robust and Sensitive Single-Cell Proteome Profiling. *Anal. Chem.* **2024**, *96* (26), 10534–10542. https://doi.org/10.1021/acs.analchem.4c00878.

(31) Wang, H.; Hanash, S. M. Increased Throughput and Reduced Carryover of Mass Spectrometry-Based Proteomics Using a High-Efficiency Nonsplit Nanoflow Parallel Dual-Column Capillary HPLC System. *J. Proteome Res.* **2008**, *7* (7), 2743–2755. https://doi.org/10.1021/pr700876g.

(32) Hosp, F.; Scheltema, R. A.; Eberl, H. C.; Kulak, N. A.; Keilhauer, E. C.; Mayr, K.; Mann, M. A Double-Barrel Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) System to Quantify 96 Interactomes per Day*. *Mol. Cell. Proteomics* **2015**, *14* (7), 2030–2041. https://doi.org/10.1074/mcp.O115.049460.

(33) Shen, Y.; Tolić, N.; Zhao, R.; Paša-Tolić, L.; Li, L.; Berger, S. J.; Harkewicz, R.; Anderson, G. A.; Belov, M. E.; Smith, R. D. High-Throughput Proteomics Using High-Efficiency Multiple-Capillary Liquid Chromatography with On-Line High-Performance ESI FTICR Mass Spectrometry. *Anal. Chem.* **2001**, *73* (13), 3011–3021. https://doi.org/10.1021/ac001393n.

(34) Lancaster, N. M.; Zhao, B.; Chen, L.-Y.; Anderson, B. J.; Overmyer, K. A.; Quarmby, S. T.; Probasco, M. D.; Coon, J. J. SynchroSep-MS: Parallel LC Separations for Multiplexed Proteomics. In *Proceeding of the 73rd ASMS Conference on Mass Spectrometry and Allied Topics*; Baltimore, MD, 2025.

(35) Rensvold, J. W.; Shishkova, E.; Sverchkov, Y.; Miller, I. J.; Cetinkaya, A.; Pyle, A.; Manicki, M.; Brademan, D. R.; Alanay, Y.; Raiman, J.; Jochem, A.; Hutchins, P. D.; Peters, S. R.; Linke, V.; Overmyer, K. A.; Salome, A. Z.; Hebert, A. S.; Vincent, C. E.; Kwiecien, N. W.; Rush, M. J. P.; Westphall, M. S.; Craven, M.; Akarsu, N. A.; Taylor, R. W.; Coon, J. J.; Pagliarini, D. J. Defining Mitochondrial Protein Functions through Deep Multiomic Profiling. *Nature* **2022**, *606* (7913), 382–388. https://doi.org/10.1038/s41586-022-04765-3.

(36) Shishkova, E.; Hebert, A. S.; Westphall, M. S.; Coon, J. J. Ultra-High Pressure (>30,000 Psi) Packing of Capillary Columns Enhancing Depth of Shotgun Proteomic Analyses. *Anal. Chem.* **2018**, *90* (19), 11503–11508. https://doi.org/10.1021/acs.analchem.8b02766.

(37) Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P. A Cross-Platform Toolkit for Mass Spectrometry and Proteomics. *Nat. Biotechnol.* **2012**, *30* (10), 918–920. https://doi.org/10.1038/nbt.2377.

(38) Kong, A. T.; Leprevost, F. V.; Avtonomov, D. M.; Mellacheruvu, D.; Nesvizhskii, A. I. MSFragger: Ultrafast and Comprehensive Peptide Identification in Mass Spectrometry–Based Proteomics. *Nat. Methods* **2017**, *14* (5), 513–520. https://doi.org/10.1038/nmeth.4256.

(39) Yu, F.; Teo, G. C.; Kong, A. T.; Li, G. X.; Demichev, V.; Nesvizhskii, A. I. One-Stop Analysis of DIA Proteomics Data Using MSFragger-DIA and FragPipe Computational Platform. bioRxiv October 31, 2022, p 2022.10.28.514272. https://doi.org/10.1101/2022.10.28.514272.

(40) Seabold, S.; Perktold, J. Statsmodels: Econometric and Statistical Modeling with Python. *scipy* **2010**. https://doi.org/10.25080/Majora-92bf1922-011.

(41) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P.; SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17* (3), 261–272. https://doi.org/10.1038/s41592-019-0686-2.

(42) MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J. Skyline: An Open Source Document Editor for Creating and Analyzing Targeted Proteomics Experiments. *Bioinformatics* **2010**, *26* (7), 966–968. https://doi.org/10.1093/bioinformatics/btq054.

(43) Demichev, V.; Messner, C. B.; Vernardis, S. I.; Lilley, K. S.; Ralser, M. DIA-NN: Neural Networks and Interference Correction Enable Deep Proteome Coverage in High Throughput. *Nat. Methods* **2020**, *17* (1), 41–44. https://doi.org/10.1038/s41592-019-0638-x.

(44) The UniProt Consortium. UniProt: The Universal Protein Knowledgebase in 2025. *Nucleic Acids Res.* **2025**, *53* (D1), D609–D617. https://doi.org/10.1093/nar/gkae1010.

(45) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9* (3), 90–95. https://doi.org/10.1109/MCSE.2007.55.

Chapter 4

Regarding Emitter Positioning for Nanoflow Electrospray Ionization with a High-Capacity Inlet Capillary

Lancaster, N. M.; Quarmby, S. T.; Overmyer, K. A.; Coon, J. J. Regarding Emitter Positioning for Nanoflow Electrospray Ionization with a High-Capacity Inlet Capillary. *J. Am. Soc. Mass Spectrom.* **2026**, 37 (3), 754–759. https://doi.org/10.1021/jasms.5c00441.

Abstract

Nanoflow electrospray ionization is commonly used for proteomics due to its high sensitivity. Signal intensity, however, is dependent on optimal emitter positioning relative to the mass spectrometer inlet. Here, we characterize the effect of varied emitter positions on peptide signal intensity in all three dimensions using emitters and flows consistent with standard proteomic analyses. We observe improved signal robustness to x/y variations at increasing z distances and demonstrate that positioning within 1 to 2 mm of the optimal location will maintain consistent signal. Signal intensity behavior is consistent across the *m/z* range, suggesting emitter positions do not need to be fine-tuned for different analytes for proteomics analyses. These results provide insight for proteomics researchers using nanoflow LC-MS/MS.

Introduction

Owing to its high sensitivity, nanoflow electrospray liquid chromatography coupled with tandem mass spectrometry (nESI-LC-MS/MS) is widely used for protein sequence analysis.[1,2] Whereas high-flow electrospray sources typically have relatively fixed emitter alignments and are more robust to changes in emitter positioning, nESI sources frequently require manual alignment in three dimensions.[3,4] Since the number of ions reaching the MS system can vary depending on the location of the nESI emitter, positioning is a critical aspect of achieving robust and reproducible results.

Multiple reports illustrate the relationship between emitter positioning and electrospray signal intensity[2,4–10]; however, these studies typically examine a single dimension or evaluate emitters and flow regimes not typical of shotgun proteomic workflows. To understand these relationships, we investigate here the effect of emitter positioning on signal intensity when using fused-silica capillaries with integrated emitters at a flow rate of 300 nL/min, a common choice across the field.[11–17] Experiments were conducted using an Orbitrap Ascend mass spectrometer, a notable example of instruments commonly used in proteome analysis,[18] which features a nanoflow source and asymmetric inlet capillary (high-capacity transfer tube, HCTT). This work represents, to our knowledge, the first report in the literature characterizing how emitter positioning impacts signal intensity for nanoflow electrospray ionization into an inlet capillary without radial symmetry.

Since most nESI emitters are manually positioned, understanding the precision in positioning that is required could assist in reducing batch-to-batch variation that is common in larger proteomic experiments. Further, recent efforts to improve the throughput of proteome

analysis feature the use of multiple columns with multiple emitters aligned with the source at the same time.[19–23] By performing measurements in all three dimensions, this report provides insight for implementing such a multi-emitter setup.

Methods

Materials and Reagents: Ultrapure water was generated with a Barnstead GenPure Pro system (Thermo Scientific). Tris(2-carboxyethyl)phosphine hydrochloride (TCEP, C4706-2G), chloroacetamide (≥98%, C0267-100G), urea (U5378-1kg), and trifluoroacetic acid (TFA, HPLC grade, >99.9%) were obtained from Sigma-Aldrich. Tris Buffer (1 M Tris pH 8.0, 0.2 µm filtered, Invitrogen, AM9856), formic acid (LC-MS grade), acetonitrile (Optima LC/MS grade), and bovine serum albumin (BSA, Thermo Fisher Scientific, 23209, Albumin Standard) were obtained from Fisher Scientific. Trypsin (Sequencing Grade Modified Trypsin, V5113) was purchased from Promega. LysC (Lysyl Endopeptidase, 100369-826) was acquired from VWR.

Sample Preparation: BSA standard was diluted to 0.5 mg/mL with a final composition of 2 M Urea/100 mM Tris/2.5 mM TCEP/10 mM chloroacetamide. LysC was added at a 1:50 enzyme:BSA ratio and the sample was rocked at ambient conditions for 4 hours. Trypsin was added at a 1:50 enzyme:BSA ratio and the sample was rocked at ambient conditions overnight. Digestion was quenched with 10% TFA added to pH <2. The sample was centrifuged for 5 minutes at 9,000 g, desalted using a Strata-X 33 µm polymeric reversed phase SPE cartridge (Phenomenex), dried in a SpeedVac (Thermo Scientific), and resuspended in 0.2% formic acid. Peptide concentration was measured via NanoDrop (Thermo Scientific), and the sample was diluted to 0.5 mg/mL with a final solvent composition of 24% acetonitrile/0.2% formic acid.

Data Collection: Emitter positioning measurements were performed by infusing BSA peptides into an Orbitrap Ascend mass spectrometer (Thermo Scientific) with a Vanquish Neo UHPLC (Thermo Scientific). An emitter was prepared by laser pulling and etching a fused silica capillary (360 µm OD, 75 µm ID) with ~10 µm tip ID[24], connecting to the nanoLC, and aligned

to the inlet with the x/y/z micrometer on a NanoSpray Flex source (Thermo Scientific) and the source camera. The origin of the coordinate system used here was defined as the center of the inlet capillary opening. In particular, the z = 0 plane is orthogonal to the opening of the inlet capillary, and the x = 0 and y = 0 planes are parallel to the inlet capillary axis. Note that the xyz stage positions were indicated in intervals of 1 mm, so the uncertainty of the xyz positioning is ±0.5mm. Positioning results are reported accordingly. **Figure 4.1** shows top and side-view images of the starting position**.** The emitter angle in the y-dimension was measured to be ~17.4° using the ImageJ software.[25] Infusion was performed by injecting 20-22 µL using Direct Control within Xcalibur at a flow rate of 300 nL/min and mobile phase conditions identical to the sample solvent composition. The infusion method provided over an hour of stable signal (**Figure 4.2**). For each set of positioning experiments, an infusion was begun, and the signal was monitored until it had stabilized (~15 min). Then, the emitter position was varied for data collection for no more than approximately 30 minutes to ensure stable spray. For this study, this procedure was performed once to collect data for the 1D experiments shown in **Figure 4.7** and was performed once at each z position to generate the data in **Figure 4.12**. Measurements were taken at the starting positions throughout each data collection stage to confirm consistent signal intensity (**Figure 4.3**).

A spray voltage of 2000 V was applied with transfer tube temperature of 275°C. Orbitrap MS^1 scans were acquired at 60,000 resolving power with scan range 300–1500 *m/z*, 30% RF Lens setting, 250% Normalized AGC Target, and maximum injection time of 100 ms. Signal intensity was allowed to stabilize after adjusting positions. For each position, 100 scans were acquired in the Tune software.

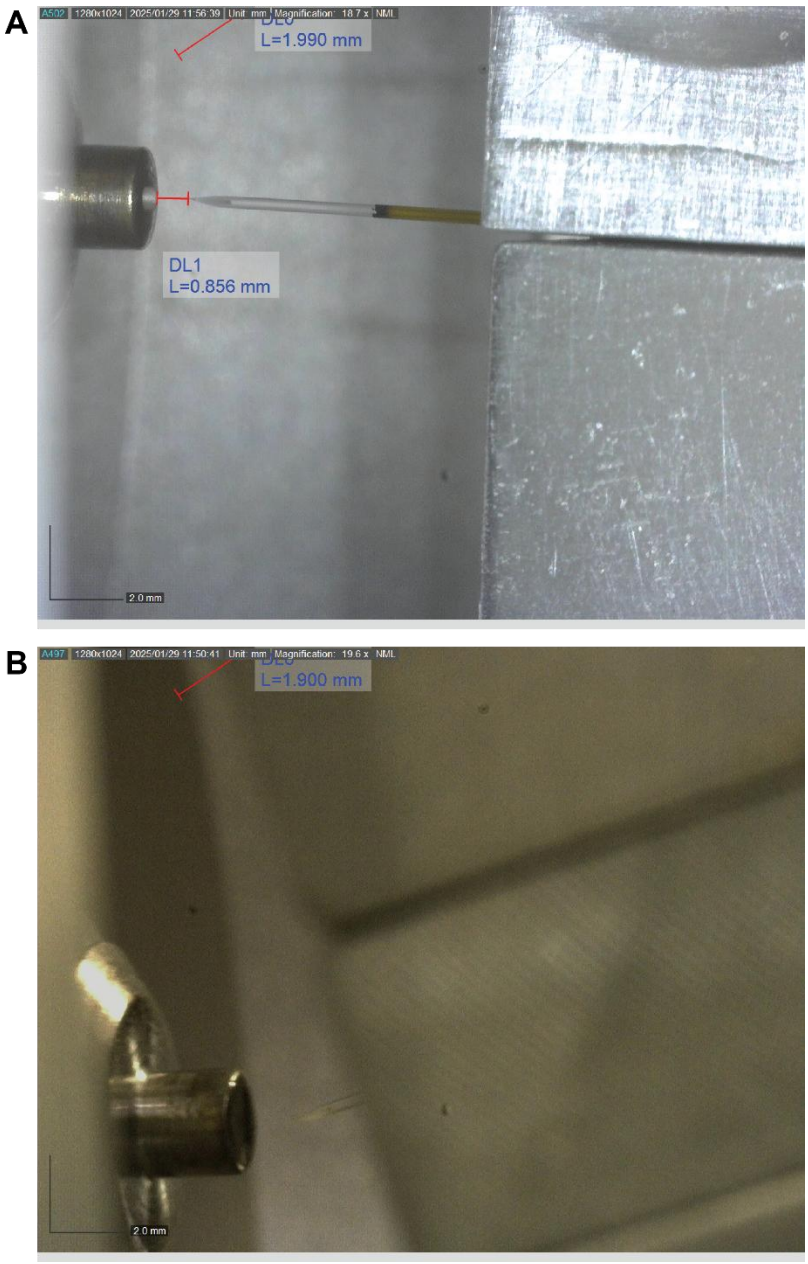

Figure 4.1. Emitter Position Images Recorded with Source Camera. (A) Top View of the emitter (x/z axes). The measured distance between the inlet capillary and the emitter tip was calibrated based on the known 360 µm O.D. of the silica capillary. Note that the 1.990 mm measurement is an artifact from the camera software. (B) Side view of the emitter (y/z axes). The angle of the emitter axis with respect to the inlet capillary axis was measured to be ~17.4° in ImageJ. Note that the 1.900 mm measurement is an artifact from the camera software. Image Description

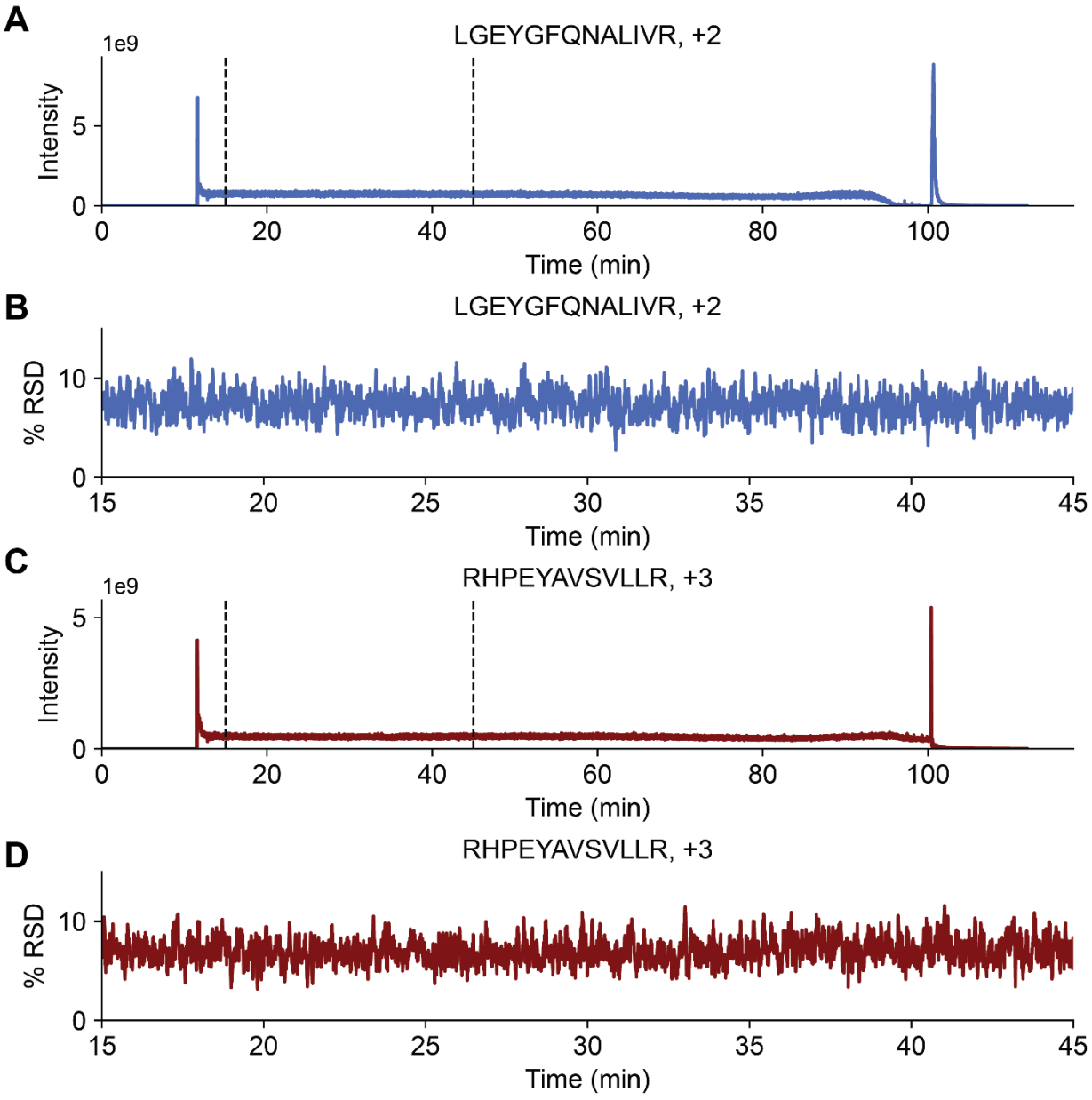

Figure 4.2. Signal Stability of Infused Peptides. (A) Signal stability over the infusion period for the LGEYGFQNALIVR^{2+} ion. Dashed lines indicate the typical time window when emitter positioning data was collected during the sample infusion. (B) Plot showing the rolling %RSD value for the LGEYGFQNALIVR^{2+} signal within the typical data collection window indicated in (A). (C) Signal stability over the infusion period for the RHPEYAVSVLLR^{3+} ion. Dashed lines indicate the typical time window when emitter positioning data was collected during the sample infusion. (D) Plot showing the rolling %RSD value for the RHPEYAVSVLLR^{3+} $^{+}$ signal within the typical data collection window indicated in (C). Rolling %RSD values were calculated over 20 scans. Image Description

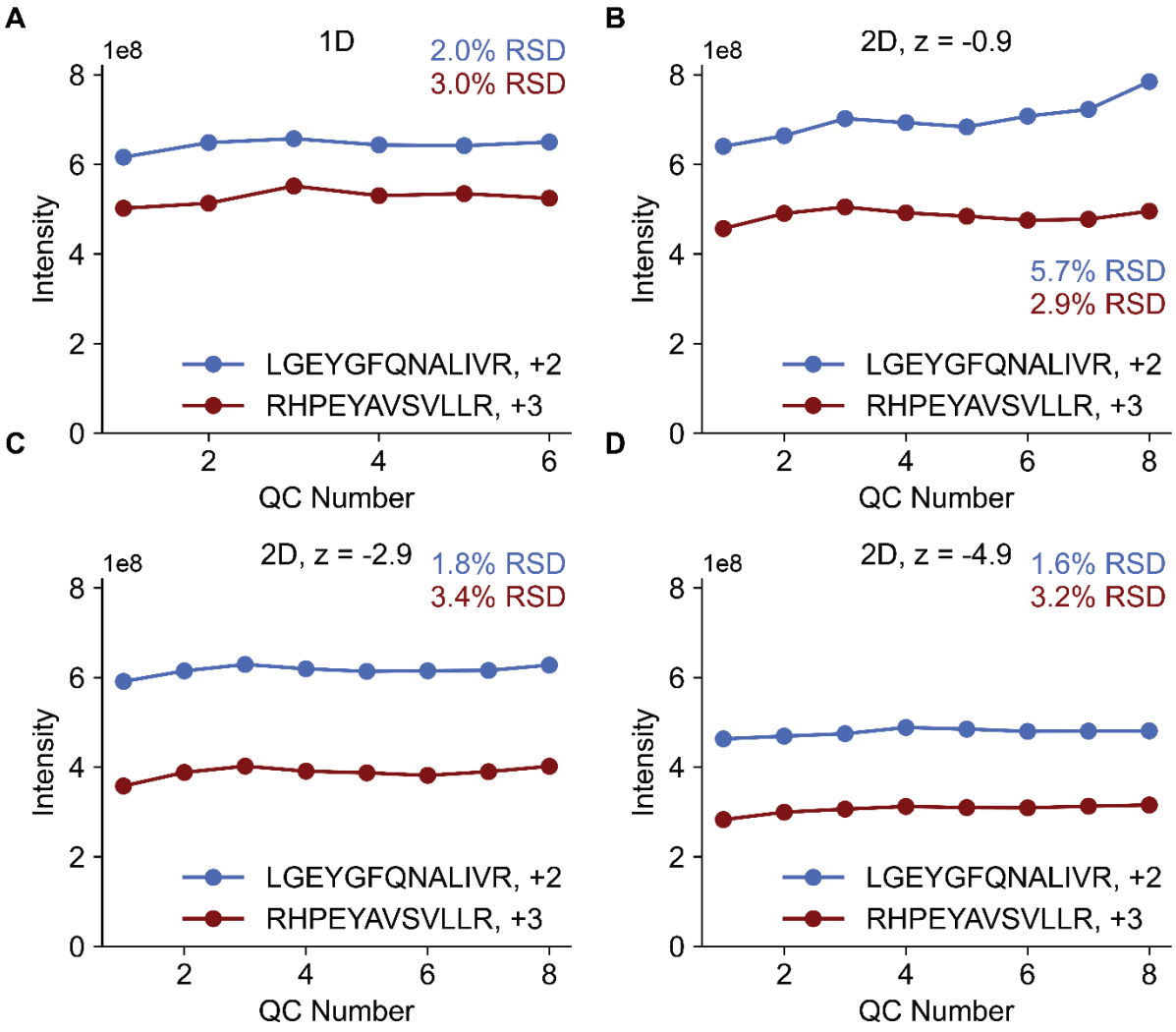

Figure 4.3. QC Measurements for Emitter Positioning Experiments. The position was returned to the starting position for each experiment periodically throughout data collection to confirm signal stability during the experiments. The signal intensity is shown for the (A) one-dimensional experiments shown in **Figure 4.7**, and the two-dimensional experiments shown in **Figure 4.12** and **Figure 4.13** at (B) z = -0.9, (C) z = -2.9, and (D) z = -4.9 mm. For all measurements, x = 0 mm and y = 0 mm. The relative standard deviations across QC measurements are shown on each plot with text in colors matching the corresponding peptide precursor data. Image Description

Data Analysis: An *in silico* BSA digest was performed using PeptideMass in Expasy[26] with two missed cleavages allowed. Outputs were used to select a doubly and triply charged precursor ($LGEYGFQNALIVR^{2+}$ and $RHPEYAVSVLLR^{3+}$) with high signal intensity and stability. Ion intensity was averaged across scans with a 15-ppm *m/z* tolerance using the 'pymsfilereader' library in Python. Calculation of approximate full-width half-maximum (FWHM) values for intensity distributions was performed by fitting the data with the 'UnivariateSpline' function within SciPy.[27] For the global comparisons across *m/z*, mass spectra for the two extreme z-positions (at x = 0 and y = 0) were averaged in Freestyle, followed by exporting peak lists and filtering peaks present in both spectra (within 5 ppm) for charge > 1 and S/N > 300. This filtering retained higher intensity peaks more likely to be reliable indicators for the effect of emitter positioning on signal intensity. Pearson correlations were calculated using 'pearsonr' within SciPy.[27]

Geometric Calculations: The Nanospray source used in this study exhibits a ~17.4° angle between the central axis of the emitter and the central axis of the inlet capillary in the y dimension. An illustration of this geometry is shown below (**Figure 4.4**), where θ represents the angle, d_z represents the distance of the emitter from the inlet in the z dimension, and Δy represents the offset between the emitter's y position and the y position at the intersection of the central axis of the emitter and the plane of the inlet capillary.

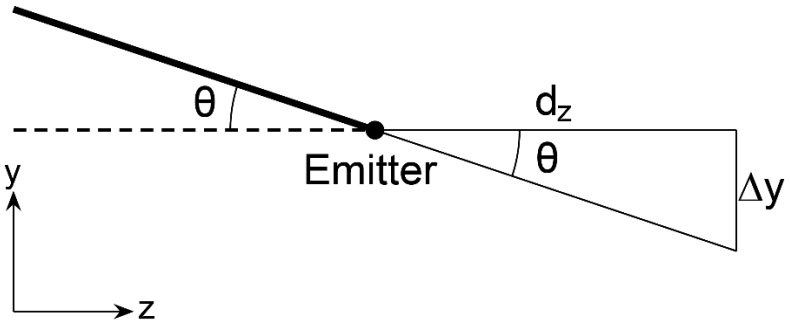

Figure 4.4. Geometric Diagram of Nano-Electrospray Emitter Configuration. Image Description

For an angle, θ, of 17.4° and an emitter distance, d_z, the offset, Δy, is defined by

$$d_z \tan\theta$$

Thus, for the z positions shown in **Figure 4.12** and **Figure 4.13**, the offsets are calculated as

Table 4.1. Predicted Offsets (Δy) at Different z Distances.

Absolute z Distance (mm)	Theoretical Δy (mm)
0.9	0.3
2.9	0.9
4.9	1.5

This feature of the source geometry would be expected to impact the emitter positioning results in the y dimension if the geometric configuration of the source was a major driver of ion trajectories. However, this trend is not clearly observed as shown in **Figure 4.5**.

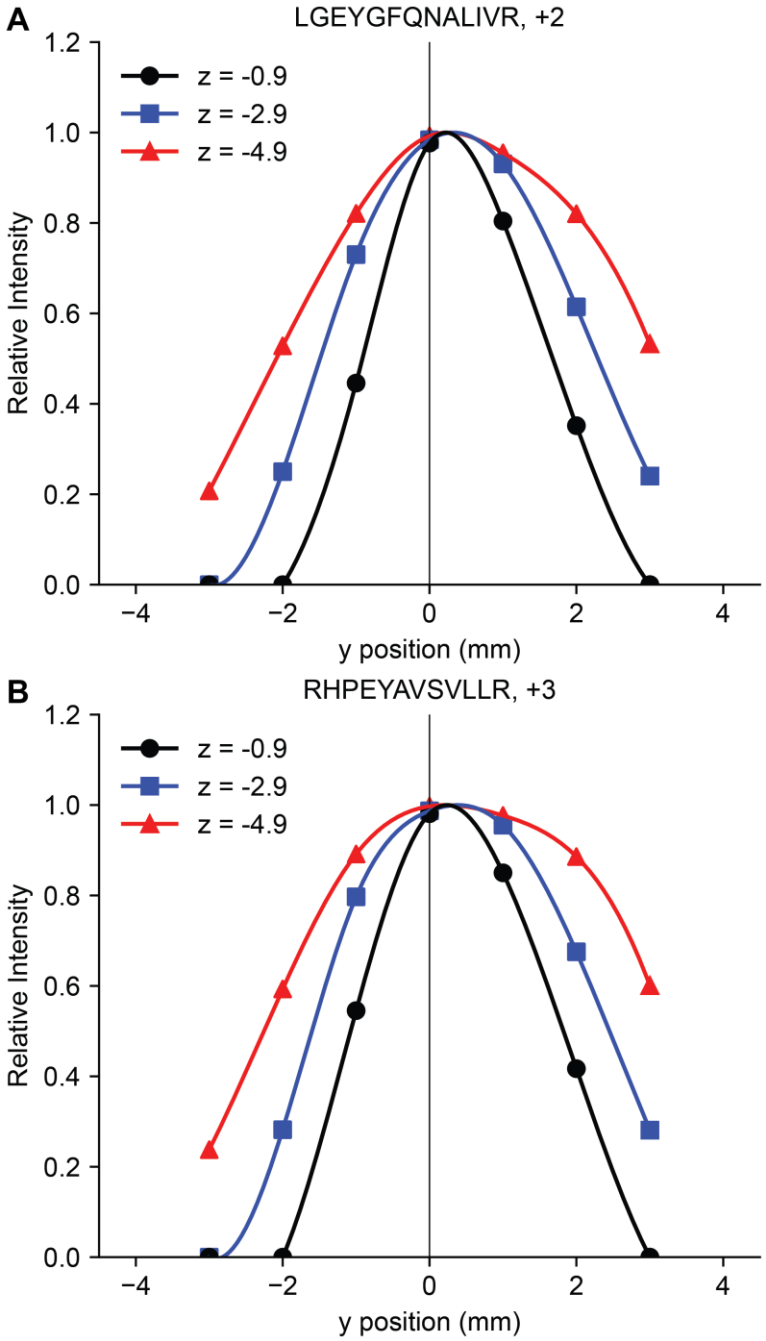

Figure 4.5. Dependence of y Intensity Distribution on z Position. For the data shown in **Figure 4.12** and **Figure 4.13**, the y position-dependent normalized intensity distributions (at x = 0) were fit to a spline interpolation function with SciPy. Data are shown at different z positions for (A) LGEYGFQNALIVR (+2) and (B) RHPEYAVSVLLR (+3). The y = 0 point is indicated with a black line. Dots represent the experimental datapoints and the lines represent the spline function. Image Description

Results and Discussion

We characterized signal intensity dependence on emitter position for nanoflow electrospray by infusing a mixture of peptides generated following trypsin digestion of purified BSA in an Orbitrap-quadrupole linear ion trap hybrid mass spectrometer (Orbitrap Ascend Tribrid). To ensure LC-MS applicability, we used an emitter size (~ 10 micron orifice) and flow rate (300 nL/min) identical to those typical of capillary LC-MS. [11–17] The resulting MS^1 analysis produced numerous ions across the *m/z* range (**Figure 4.6**). We selected two of these ions, a doubly protonated species (LGEYGFQNALIVR at *m/z* 740.4) and a triply protonated species (RHPEYAVSVLLR at *m/z* 480.6), for their strong signal intensity and stability over the time range of the experiments here (**Figure 4.2** and **Figure 4.3**). With the above setup, we tracked those *m/z* peaks as a function of emitter position in three dimensions and recorded intensity distributions (**Figure 4.7**). **Figure 4.7C** presents the ion signal across x-positions, where the x = 0 plane is aligned to the central axis of the inlet capillary and the dashed lines depict the MS inlet opening width in this dimension (0.6 mm, **Figure 4.8**).[28,29] Strikingly, we observe that the signal is relatively stable and high across two millimeters of the x-dimension. Further, greater than 50% of the ion signal is retained at distances up to ~five times the width of the inlet (*i.e.*, full width half max (FWHM)). **Figure 4.7D** presents the same concept but in the y-dimension, where good signal stability is also observed across a similar length of two millimeters; however, the larger inlet opening in the y-dimension (1.6 mm) does not appear to proportionally impact the width of the intensity distribution. The ion plume is roughly symmetrical in both x- and y-dimensions[8]; therefore, it is expected that the smaller inlet width in the x-dimension would have a higher probability of being in the high density region of

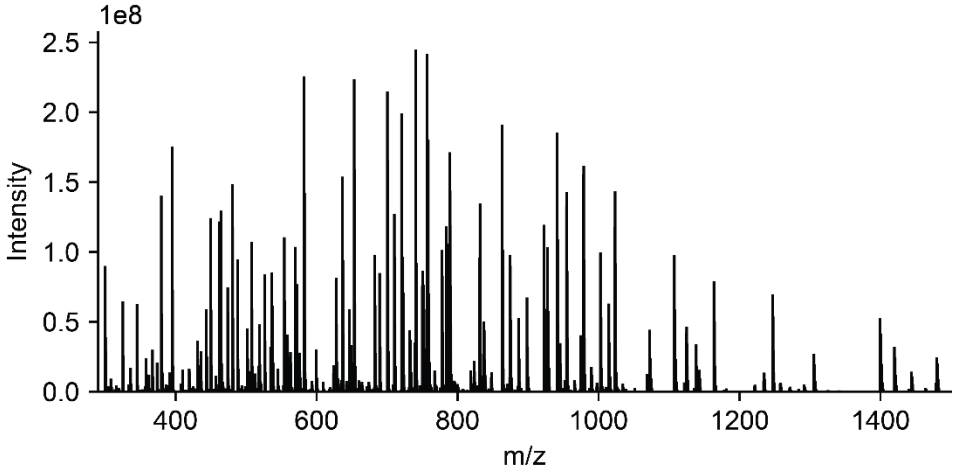

Figure 4.6. MS[1] Spectra of Infused BSA Tryptic Digest. The Orbitrap MS[1] spectra is shown for infusion of BSA peptides at x = 0 mm , y = 0 mm. z = -0.9 mm. Image Description

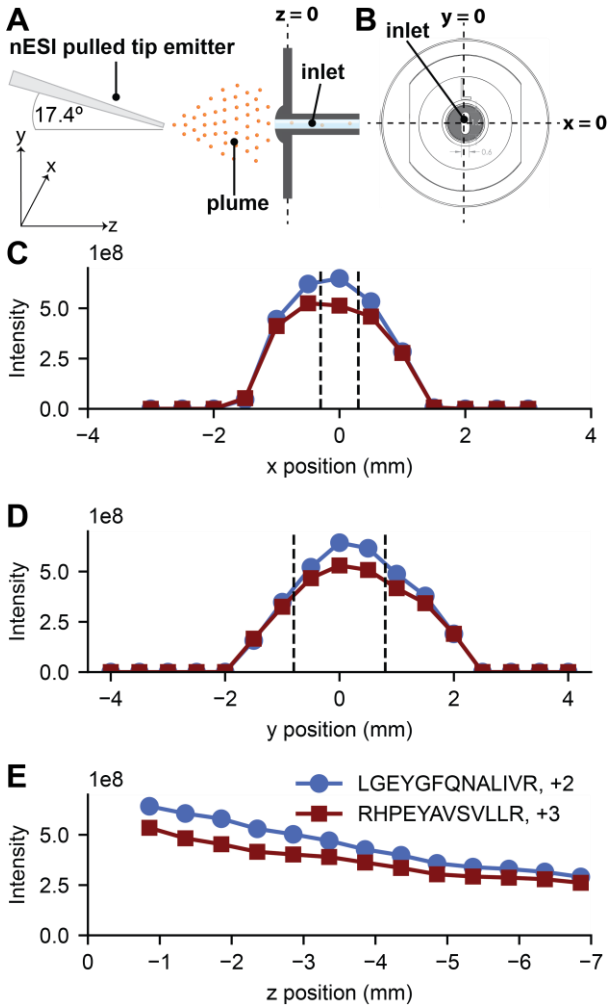

Figure 4.7. 1D Positioning Experiments. (A) The x/y/z coordinate system used for this study. (B) An x/y perspective of the inlet with coordinate definitions. (C) Intensity dependence on x-position for two selected peptides (at y=0 and z = -0.9). Dotted lines indicate the estimated position of the inlet capillary edges. (D) Intensity dependence on y-position for two selected peptides (at x = 0 and z = -0.9). Dotted lines indicate the estimated position of the inlet capillary edges. (E) Intensity dependence on z-position for two selected peptides (at x = 0 and y = 0). Image Description

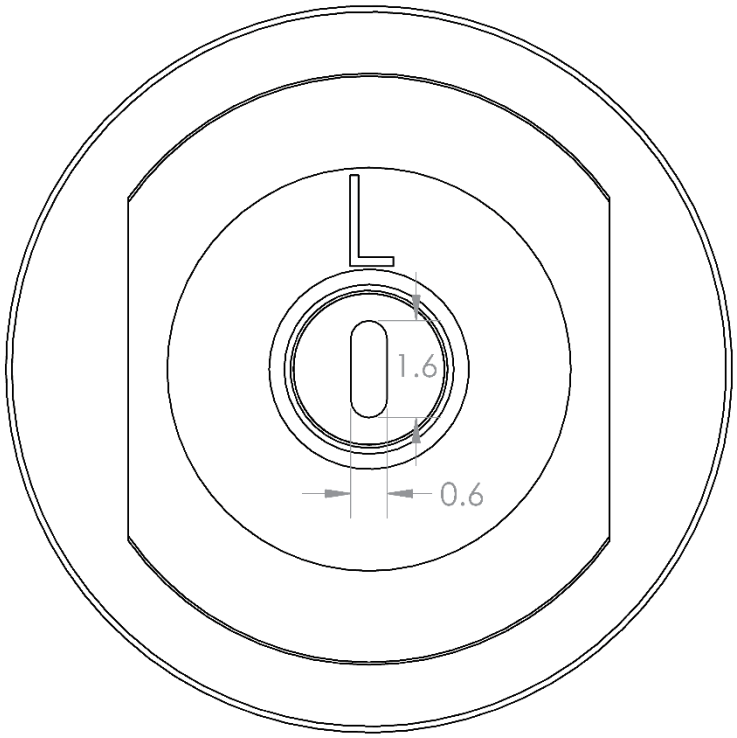

Figure 4.8. Inlet Capillary Drawing. A drawing of the high-capacity transfer tube (HCTT) is shown. Measurements are reported in mm with 0.6 mm being the width in the x dimension and 1.6 mm being the width in the y dimension. Image Description

the plume than in the y-dimension. Proportional to inlet opening, the x position of the emitter would consequently have less of an effect on the intensity than the y position. We note the intensity distributions appear to be slightly off-center; however, we find the centroids of the distributions (-0.1 and 0.3 mm, x- and y-respectively, **Figure 4.9**) were within the tolerance of the micrometer. **Figure 4.7E** presents the ion intensities as a function of emitter position in the z-dimension. Impressively, signal continues to be observed up to ~7 mm removed from the inlet at levels close to 50% of the highest. Not surprisingly, the highest signal is observed at the closest position, but, as noted above, that signal gradually declines with increasing distance. The relative intensities of these two peptide ions remain fairly constant at all observed z positions (**Figure 4.10**). Overall, identical trends were observed for both peptides over all dimensions. As these peptides are fairly high signal intensity, we wanted to confirm that the trends reported here were not biased by signal saturation, so we extracted signal intensities for low intensity features and compared the x, y, and z intensity distributions (**Figure 4.11**). This analysis confirms that these two peptides are representative of global trends for our tryptic peptide sample.

With these unidimensional observations complete, we next characterized the three-dimensional interplay of emitter position. For these experiments, we rastered the emitter across x, y, and z dimensions (**Figure 4.12** and **Figure 4.13**). A key observation from these data is that signal intensity becomes more robust to x- and y-positioning as distance from the inlet increases (z-position). This observation is likely explained by a widening and flattening of the ion plume spatial distribution with increasing distance as previously reported.[9] Specifically, with the emitter close to the inlet (z = -0.9 mm), the x, y tolerance to retain ~90% signal is ~1 mm, whereas at z = -4.9 mm the tolerance is ~2-3mm (**Figure 4.14**). To our knowledge, these data are the first to map the

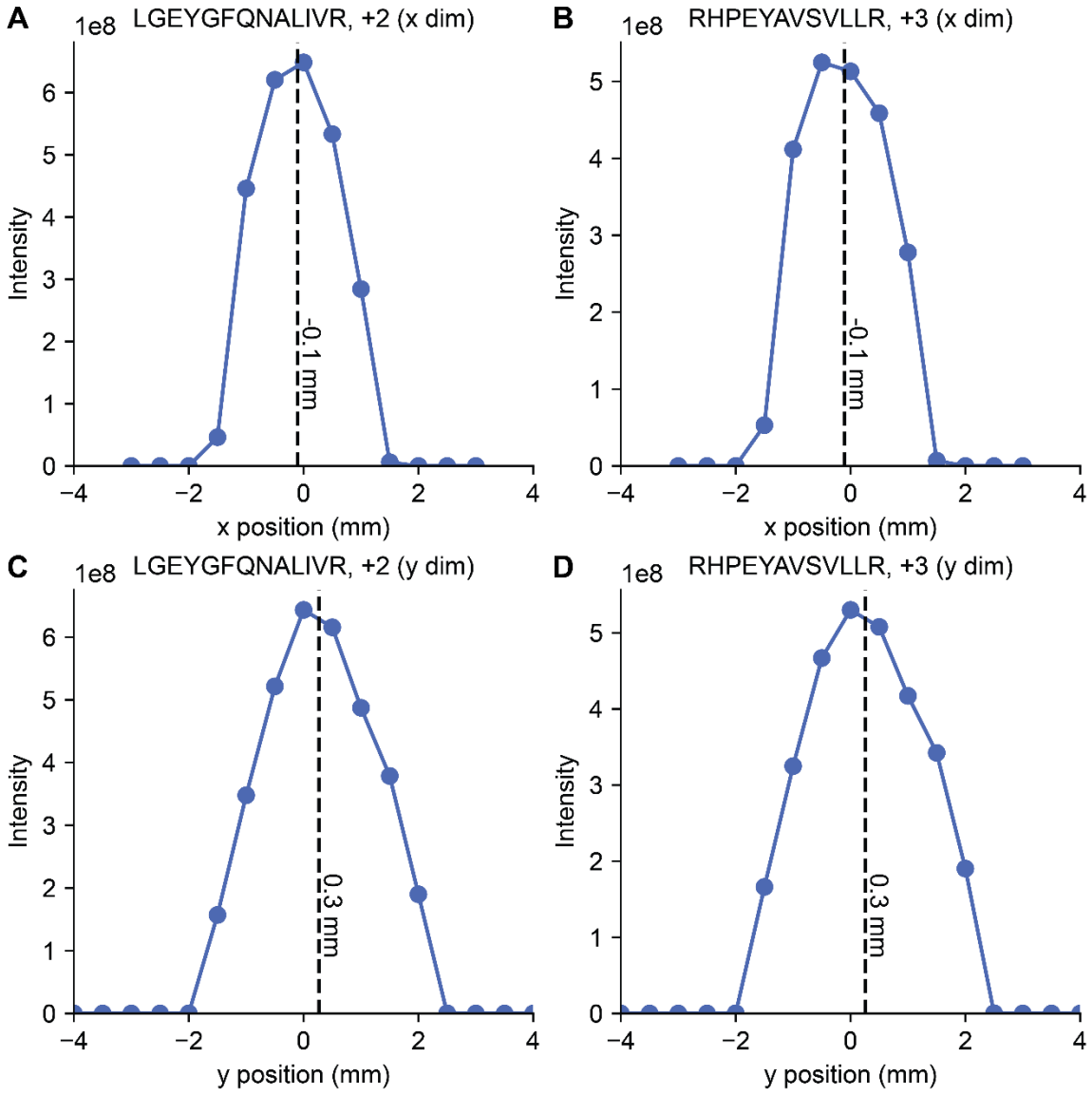

Figure 4.9. Estimating Distribution Centroids. For the data shown in **Figure 4.7C** and **Figure 4.7D**, the centroids of the distributions in the x- and y-dimensions (at z = -0.9) were determined. Data are shown for (A) LGEYGFQNALIVR (+2) for the x dimension, (B) RHPEYAVSVLLR (+3) for the x dimension, (C) LGEYGFQNALIVR (+2) for the y dimension, and (D) RHPEYAVSVLLR (+3) for the y dimension. Dotted black lines with text annotations indicate the position of the centroid calculated as $\bar{x} = \sum(x_i\ I_i)/\sum I_i$, where x_i is the x position for a datapoint and I_i is the corresponding intensity value at that position. Image Description

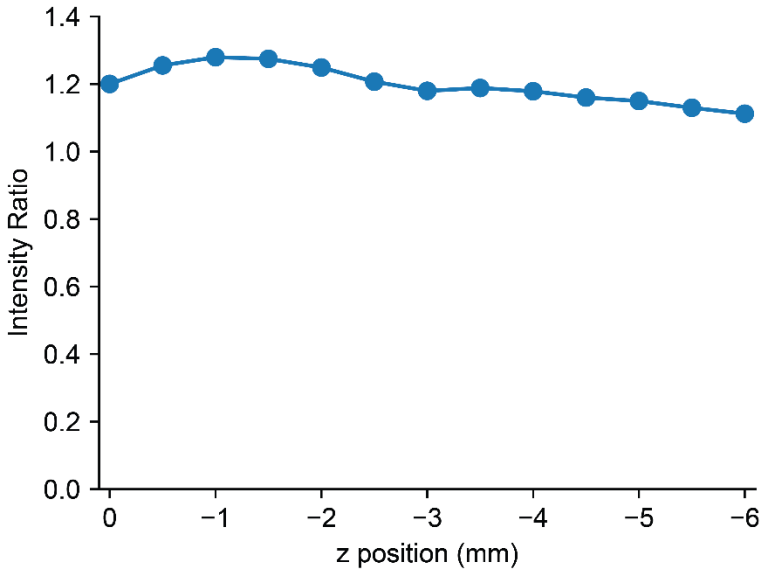

Figure 4.10. Dependence of Precursor Intensity Ratio on z Position. The ratio of intensities for LGEYGFQNALIVR (+2) and RHPEYAVSVLLR (+3) is calculated and plotted as function of z position using the data shown in **Figure 4.7E** (at x = 0 and y = 0). Image Description

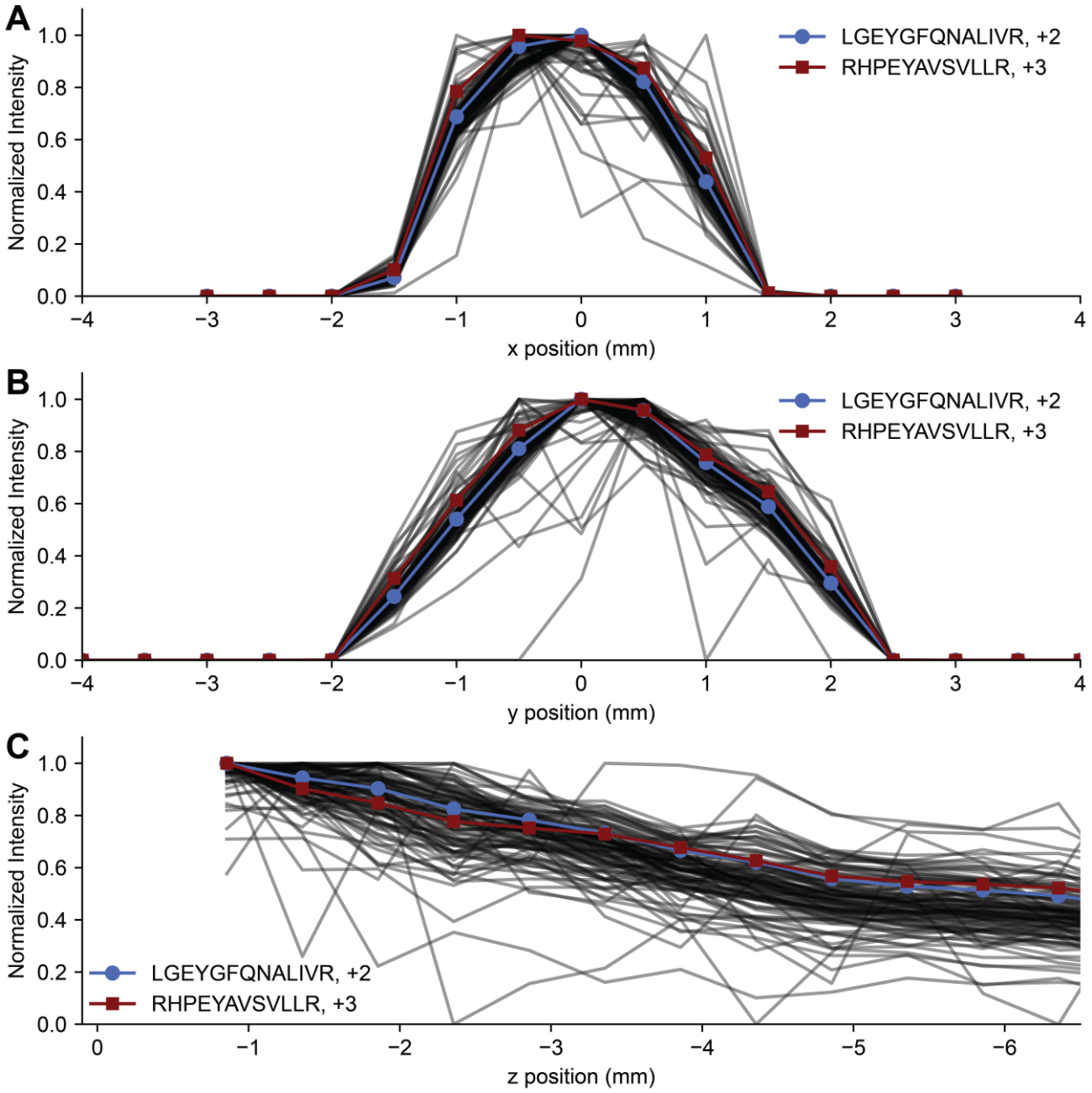

Figure 4.11. Dependence of Signal Intensity on Emitter Position for Low Level Precursor Ions. Mass spectra for the two extreme z-positions in **Figure 4.7E** (at x = 0 and y = 0) were filtered for *m/z* peaks present in both with S/N between 50 and 100 and charge >1 to represent lower-level intensity peaks. The normalized intensity dependence on position is shown for the (A) x dimension, (B), y dimension, and (C) z dimension. Normalized intensity profiles generated from the data shown in **Figure 4.7C-E** are overlaid to compare how these agree with the lower-level signals. Image Description

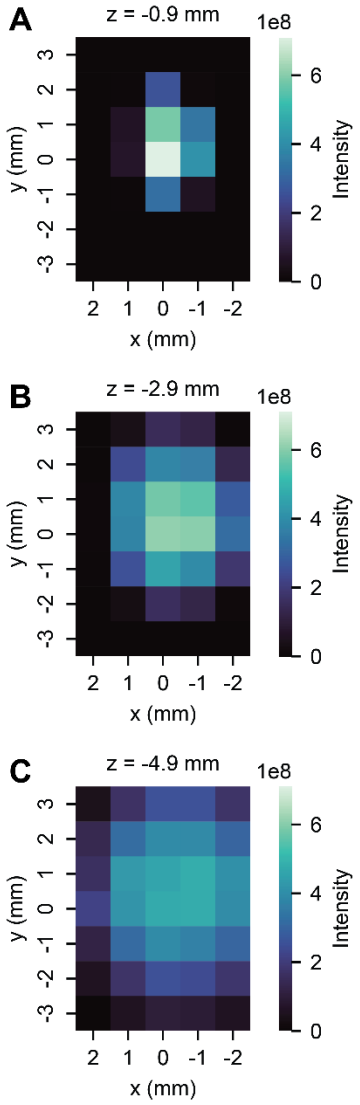

Figure 4.12. 2D Positioning Experiments. Heatmaps showing intensity of LGEYGFQNALIVR (+2) at different x/y positions at (A) z = -0.9 mm, (B) z = -2.9 mm, and (C) z = -4.9 mm. Image Description

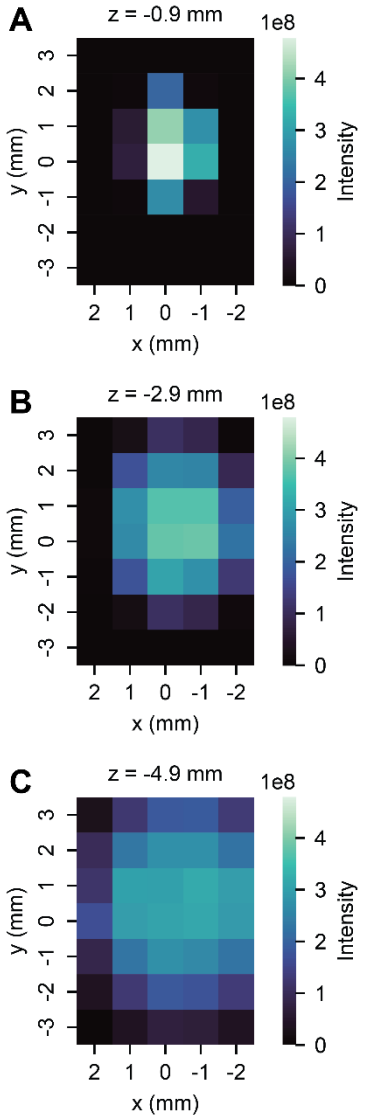

Figure 4.13. 2D Positioning Experiments for RHPEYAVSVLLR. Heatmaps showing the intensity of RHPEYAVSVLLR (+3) at different x/y positions at (A) z = -0.9 mm, (B) z = -2.9 mm, and (C) z = -4.9 mm. Image Description

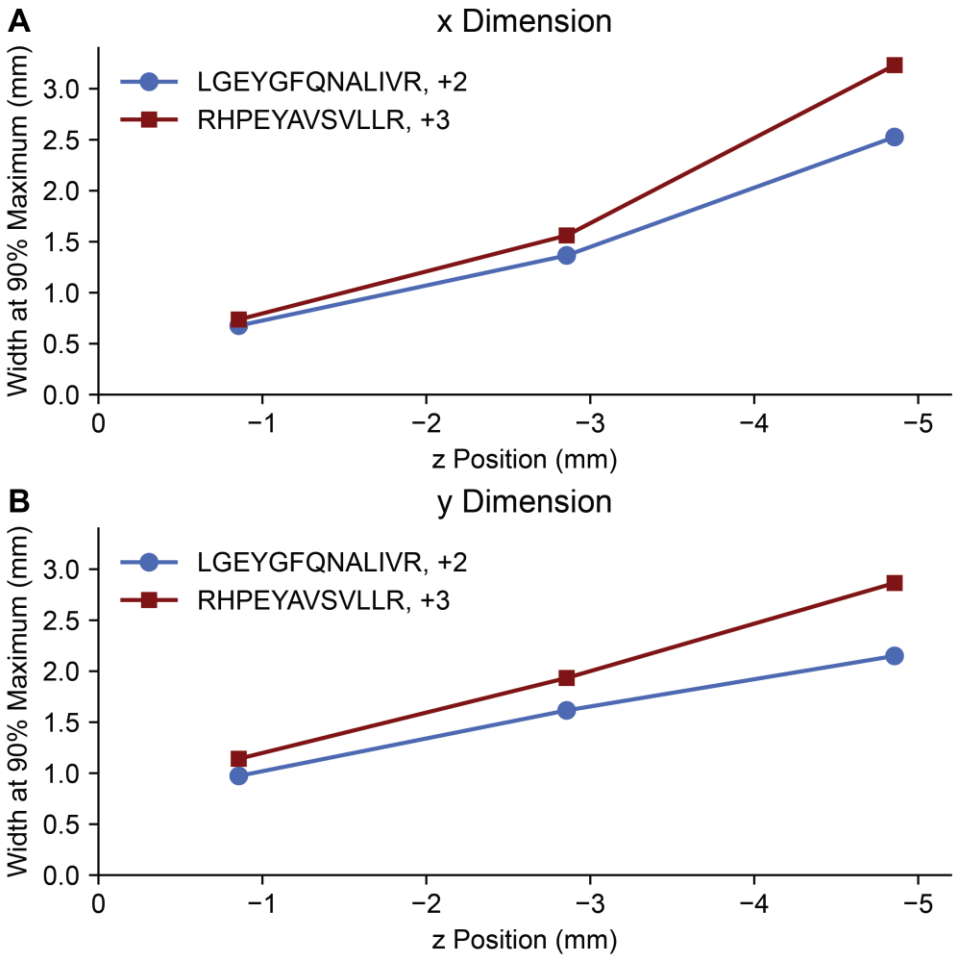

Figure 4.14. Width at 90% Maximum as Function of z Position. For the data shown in **Figure 4.12** and **Figure 4.13**, the width at 90% maximum of the intensity distribution in the x- and y-dimensions were approximated by fitting the data to a spline interpolation function with SciPy. The approximate width values for the (A) x- and (B) y-distributions are then shown for the two different precursors ions as a function of the z position. Image Description

three-dimensional space of the nESI plume in the context of the conditions that are typical in shotgun proteomics (*i.e.*, capillary LC-MS).[2,6,8–10,30–38] An especially exciting potential application of this knowledge is in the use of multiple emitters and/or parallel separations. In these cases, two or more emitters must be simultaneously positioned in front of the inlet, inhibiting the conventional alignment approaches where you put the emitter as centered and close as possible. Here, one desires to both achieve maximal ion signal while ensuring similar performance across the emitters. In such scenarios, our data suggest that moving 3-5 mm back in the z-dimension will ease the constraints on the x,y-positioning and reduce differences between the multiple emitters.

The source used in this study presented the emitter such that the central axis of the emitter is at an angle of ~17.4° relative to the y-axis (**Figure 4.7A** and **Figure 4.1B**). Other nanoflow sources feature angles ranging from 0° to 90°.[4,39–41] From a geometric perspective, this angle could shift the y-dimension ion intensity distribution to increasingly positive values as the emitter is positioned further away in the z-dimension (see **Geometric Calculations** in **Methods**, **Figure 4.4, Table 4.1**) due to the contributions of the initial momentum of the charged particles as they leave the emitter. However, we find the centers of the y-dimension intensity distributions are consistent at different z-positions suggesting the angle had little to no impact (**Figure 4.5**). Gas flow into the inlet capillary has a collimating effect on the electrospray plume and is the dominating factor controlling ion trajectories due to its influence on the aerodynamic flow of the charged particles.[31,42] Additionally, the electrical potential difference between the emitter and the inlet capillary could curve the ion plume toward the inlet. These two effects would dominate especially as desolvation decreases droplet size over time, increasing the mobility. These mechanisms likely explain why the geometric effect of the angled emitter axis relative to the y-axis is not observed.

Finally, to ensure there are no broader analyte specific effects at play, we looked at the effect of position across various *m/z* peaks (n = 243, charge > 1, *m/z* values ranging from ~350 to ~1250 *m/z*) stemming from the BSA tryptic digest. We plotted full-width half-maximum (FWHM) values for the x and y intensity distributions at the closest z position (**Figure 4.15A** and **Figure 4.15B**) and the intensity fold changes between the closest and furthest positions (*i.e.*, z=-0.9 and z=-6.9, **Figure 4.15C**) vs *m/z*. Although there is variation in measured FWHM values and fold-change across analytes, this variation is not correlated with *m/z*. Note that outliers in these plots correspond to lower intensity peaks more likely to exhibit unstable signal for our setup (**Figure 4.16**). We conclude that the effect of emitter positioning on signal intensity does not strongly depend on the analyte for peptide analyses, such that the same emitter position can be appropriate for most peptides.

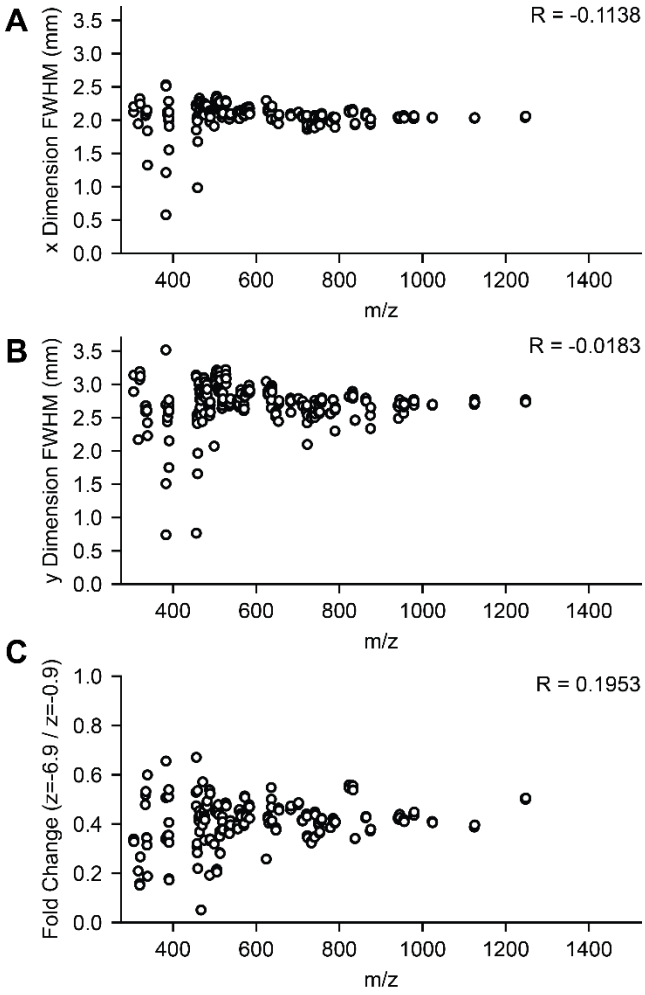

Figure 4.15. Dependence of Emitter Position Results on *m/z*. Mass spectra for the two extreme z-positions (at x = 0 and y = 0) were filtered for *m/z* peaks present in both with S/N>300 and charge >1. The full-width half-maximum (FWHM) was approximated via spline interpolation for the (A) x-distribution and (B) y-distribution at z = -0.9 mm, and (C) the fold change between the extreme z-positions was calculated (at x = 0, y = 0 mm). The Pearson correlation, R, of these values with *m/z* was calculated and is included at the top right of each panel. Image Description

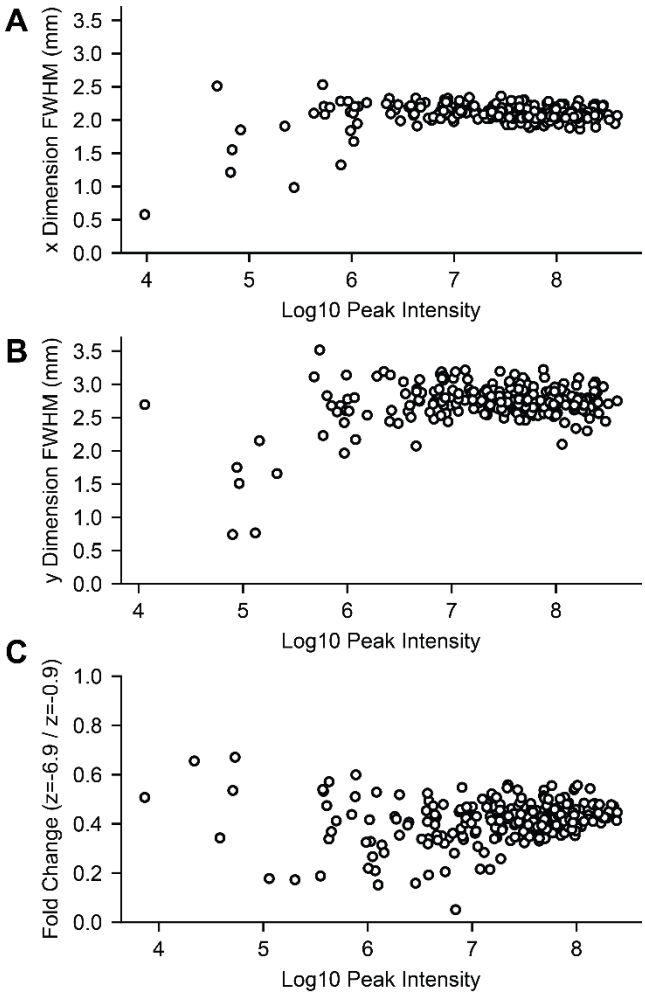

Figure 4.16. Dependence of Emitter Position Results on Peak Intensity. Mass spectra for the two extreme z-positions (at x = 0 and y = 0) were filtered for *m/z* peaks present in both spectra with S/N > 300 and charge > 1. The full-width half-maximum (FWHM) was approximated via spline interpolation for the (A) x-distribution (at z = -0.9 mm) and (B) y-distribution (at z = -0.9 mm), and (C) the fold change between the extreme z-positions (at x = 0, y = 0 mm) was calculated. These data show that low intensity peaks exhibit much greater variation. Image Description

Conclusions

We report the dependence of peptide signal intensity on emitter positioning for nanoflow electrospray. Specifically, within 1-2 mm in any dimension, one can achieve reasonably consistent and robust signal. Distinct intensity profiles we observed here for the x- and y-dimensions likely arise from the asymmetric shape of the ion capillary. We demonstrate improved robustness of signal intensity to x/y variation at increasing z distances, an observation that will be helpful for positioning multiple emitters, for example.[19–23]

We provide evidence that the effect of the emitter position on signal intensity does not strongly depend on the analyte for bottom-up proteomics analysis. This report provides insight into the role of emitter positioning on signal intensity for bottom-up proteomics and represents the first characterization of the effect of emitter position on electrospray signal intensity on an instrument with an inlet capillary lacking radial symmetry. Future areas of interest would be an examination of how trends vary across different flow rates or emitter types, as well as assessing how the exact intensity profiles vary across instruments with differences in their atmospheric interface.

Author Contributions

NML prepared samples, collected data, and performed analysis. All authors contributed to writing and editing of the manuscript.

References

(1) Emmett, M. R.; Caprioli, R. M. Micro-Electrospray Mass Spectrometry: Ultra-High-Sensitivity Analysis of Peptides and Proteins. J. Am. Soc. Mass Spectrom. 1994, 5 (7), 605–613. https://doi.org/10.1016/1044-0305(94)85001-1.

(2) Marginean, I.; Kelly, R. T.; Prior, D. C.; LaMarche, B. L.; Tang, K.; Smith, R. D. Analytical Characterization of the Electrospray Ion Source in the Nanoflow Regime. Anal. Chem. 2008, 80 (17), 6573–6579. https://doi.org/10.1021/ac800683s.

(3) Manisali, I.; Chen, D. D. Y.; Schneider, B. B. Electrospray Ionization Source Geometry for Mass Spectrometry: Past, Present, and Future. TrAC Trends Anal. Chem. 2006, 25 (3), 243–256. https://doi.org/10.1016/j.trac.2005.07.007.

(4) Kang, Y.; Schneider, B. B.; Bedford, L.; Covey, T. R. Design Characteristics to Eliminate the Need for Parameter Optimization in Nanoflow ESI-MS. J. Am. Soc. Mass Spectrom. 2019, 30 (11), 2347–2357. https://doi.org/10.1007/s13361-019-02301-8.

(5) Licklider, L.; Wang, X.-Q.; Desai, A.; Tai, Y.-C.; Lee, T. D. A Micromachined Chip-Based Electrospray Source for Mass Spectrometry. Anal. Chem. 2000, 72 (2), 367–375. https://doi.org/10.1021/ac990967p.

(6) Shepherd, S. O.; Green, A. W.; Resendiz, E. S.; Newton, K. R.; Kurulugama, R. T.; Prell, J. S. Effects of Nano-Electrospray Ionization Emitter Position on Unintentional In-Source Activation of Peptide and Protein Ions. J. Am. Soc. Mass Spectrom. 2024, 35 (3), 498–507. https://doi.org/10.1021/jasms.3c00371.

(7) Chen, L. C.; Tsutsui, S.; Naito, T.; Ninomiya, S.; Hiraoka, K. Electrospray Ionization Source with a Rear Extractor. J. Mass Spectrom. 2018, 53 (5), 400–407. https://doi.org/10.1002/jms.4072.

(8) Thompson, J. W.; Eschelbach, J. W.; Wilburn, R. T.; Jorgenson, J. W. Investigation of Electrospray Ionization and Electrostatic Focusing Devices Using a Three-Dimensional Electrospray Current Density Profiler. J. Am. Soc. Mass Spectrom. 2005, 16 (3), 312–323. https://doi.org/10.1016/j.jasms.2004.11.012.

(9) Page, J. S.; Kelly, R. T.; Tang, K.; Smith, R. D. Ionization and Transmission Efficiency in an Electrospray Ionization—Mass Spectrometry Interface. J. Am. Soc. Mass Spectrom. 2007, 18 (9), 1582–1590. https://doi.org/10.1016/j.jasms.2007.05.018.

(10) Janusson, E.; Hesketh, A. V.; Bamford, K. L.; Hatlelid, K.; Higgins, R.; McIndoe, J. S. Spatial Effects on Electrospray Ionization Response. Int. J. Mass Spectrom. 2015, 388, 1–8. https://doi.org/10.1016/j.ijms.2015.07.016.

(11) Jiang, Y.; Rex, D. A. B.; Schuster, D.; Neely, B. A.; Rosano, G. L.; Volkmar, N.; Momenzadeh, A.; Peters-Clarke, T. M.; Egbert, S. B.; Kreimer, S.; Doud, E. H.; Crook, O. M.; Yadav, A. K.; Vanuopadath, M.; Hegeman, A. D.; Mayta, M. L.; Duboff, A. G.; Riley, N. M.; Moritz, R. L.; Meyer, J. G. Comprehensive Overview of Bottom-Up Proteomics Using Mass Spectrometry. ACS Meas. Sci. Au 2024, 4 (4), 338–417. https://doi.org/10.1021/acsmeasuresciau.3c00068.

(12) Overmyer, K. A.; Shishkova, E.; Miller, I. J.; Balnis, J.; Bernstein, M. N.; Peters-Clarke, T. M.; Meyer, J. G.; Quan, Q.; Muehlbauer, L. K.; Trujillo, E. A.; He, Y.; Chopra, A.; Chieng, H. C.; Tiwari, A.; Judson, M. A.; Paulson, B.; Brademan, D. R.; Zhu, Y.; Serrano, L. R.; Linke, V.; Drake, L. A.; Adam, A. P.; Schwartz, B. S.; Singer, H. A.; Swanson, S.; Mosher, D. F.; Stewart, R.; Coon, J. J.; Jaitovich, A. Large-Scale Multi-Omic Analysis of COVID-19 Severity. Cell Syst. 2021, 12 (1), 23-40.e7. https://doi.org/10.1016/j.cels.2020.10.003.

(13) Serrano, L. R.; Peters-Clarke, T. M.; Arrey, T. N.; Damoc, N. E.; Robinson, M. L.; Lancaster, N. M.; Shishkova, E.; Moss, C.; Pashkova, A.; Sinitcyn, P.; Brademan, D. R.; Quarmby, S. T.; Peterson, A. C.; Zeller, M.; Hermanson, D.; Stewart, H.; Hock, C.; Makarov, A.; Zabrouskov, V.; Coon, J. J. The One Hour Human Proteome. Mol. Cell. Proteomics 2024, 0 (0). https://doi.org/10.1016/j.mcpro.2024.100760.

(14) Giansanti, P.; Samaras, P.; Bian, Y.; Meng, C.; Coluccio, A.; Frejno, M.; Jakubowsky, H.; Dobiasch, S.; Hazarika, R. R.; Rechenberger, J.; Calzada-Wack, J.; Krumm, J.; Mueller, S.; Lee, C.-Y.; Wimberger, N.; Lautenbacher, L.; Hassan, Z.; Chang, Y.-C.; Falcomatà, C.; Bayer, F. P.; Bärthel, S.; Schmidt, T.; Rad, R.; Combs, S. E.; The, M.; Johannes, F.; Saur, D.; de Angelis, M. H.; Wilhelm, M.; Schneider, G.; Kuster, B. Mass Spectrometry-Based Draft of the Mouse Proteome. Nat. Methods 2022, 19 (7), 803–811. https://doi.org/10.1038/s41592-022-01526-y.

(15) Zhong, X.; Li, Q.; Polacco, B. J.; Patil, T.; Marley, A.; Foussard, H.; Khare, P.; Vartak, R.; Xu, J.; DiBerto, J. F.; Roth, B. L.; Eckhardt, M.; von Zastrow, M.; Krogan, N. J.; Hüttenhain, R. A Proximity Proteomics Pipeline with Improved Reproducibility and Throughput. Mol. Syst. Biol. 2024, 20 (8), 952–971. https://doi.org/10.1038/s44320-024-00049-2.

(16) Jeong, J.; Hausmann, S.; Dong, H.; Szczepski, K.; Flores, N. M.; Garcia Gonzalez, A.; Shi, L.; Lu, X.; Lempiäinen, J.; Jakab, M.; Zeng, L.; Chasan, T.; Bareke, E.; Dong, R.; Carlson, E.; Padilla, R.; Husmann, D.; Thompson, J.; Shipman, G. A.; Zahn, E.; Barnes, C. A.; Khan, L. F.; Albertorio-Sáez, L. M.; Brill, E.; Kumary, V. U. S.; Marunde, M. R.; Maryanski, D. N.; Szany, C. C.; Venters, B. J.; Windham, C. L.; Nowakowski, M. E.; Czaban, I.; Jaremko, M.; Keogh, M.-C.; Le, K.; Soth, M. J.; Garcia, B. A.; Jaremko, Ł.; Majewski, J.; Mazur, P. K.; Gozani, O. NSD2 Inhibitors Rewire Chromatin to Treat Lung and Pancreatic Cancers. Nature 2026, 649 (8095), 205–215. https://doi.org/10.1038/s41586-025-09299-y.

(17) Li, J.; Van Vranken, J. G.; Pontano Vaites, L.; Schweppe, D. K.; Huttlin, E. L.; Etienne, C.; Nandhikonda, P.; Viner, R.; Robitaille, A. M.; Thompson, A. H.; Kuhn, K.; Pike, I.;

Bomgarden, R. D.; Rogers, J. C.; Gygi, S. P.; Paulo, J. A. TMTpro Reagents: A Set of Isobaric Labeling Mass Tags Enables Simultaneous Proteome-Wide Measurements across 16 Samples. Nat. Methods 2020, 17 (4), 399–404. https://doi.org/10.1038/s41592-020-0781-4.

(18) Peters-Clarke, T. M.; Coon, J. J.; Riley, N. M. Instrumentation at the Leading Edge of Proteomics. Anal. Chem. 2024, 96 (20), 7976–8010. https://doi.org/10.1021/acs.analchem.3c04497.

(19) Orton, D. J.; Wall, M. J.; Doucette, A. A. Dual LC–MS Platform for High-Throughput Proteome Analysis. J. Proteome Res. 2013, 12 (12), 5963–5970. https://doi.org/10.1021/pr400738a.

(20) Webber, K. G. I.; Truong, T.; Johnston, S. M.; Zapata, S. E.; Liang, Y.; Davis, J. M.; Buttars, A. D.; Smith, F. B.; Jones, H. E.; Mahoney, A. C.; Carson, R. H.; Nwosu, A. J.; Heninger, J. L.; Liyu, A. V.; Nordin, G. P.; Zhu, Y.; Kelly, R. T. Label-Free Profiling of up to 200 Single-Cell Proteomes per Day Using a Dual-Column Nanoflow Liquid Chromatography Platform. Anal. Chem. 2022, 94 (15), 6017–6025. https://doi.org/10.1021/acs.analchem.2c00646.

(21) Xie, X.; Truong, T.; Huang, S.; Johnston, S. M.; Hovanski, S.; Robinson, A.; Webber, K. G. I.; Lin, H.-J. L.; Mun, D.-G.; Pandey, A.; Kelly, R. T. Multicolumn Nanoflow Liquid Chromatography with Accelerated Offline Gradient Generation for Robust and Sensitive Single-Cell Proteome Profiling. Anal. Chem. 2024, 96 (26), 10534–10542. https://doi.org/10.1021/acs.analchem.4c00878.

(22) Hosp, F.; Scheltema, R. A.; Eberl, H. C.; Kulak, N. A.; Keilhauer, E. C.; Mayr, K.; Mann, M. A Double-Barrel Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) System to Quantify 96 Interactomes per Day*. Mol. Cell. Proteomics 2015, 14 (7), 2030–2041. https://doi.org/10.1074/mcp.O115.049460.

(23) Lancaster, N. M.; Chen, L.-Y.; Zhao, B.; Anderson, B. J.; Probasco, M. D.; Demichev, V.; Polasky, D. A.; Nesvizhskii, A. I.; Overmyer, K. A.; Quarmby, S. T.; Coon, J. J. SynchroSep-MS: Parallel LC Separations for Multiplexed Proteomics. J. Am. Soc. Mass Spectrom. 2025, 36 (9), 1979–1987. https://doi.org/10.1021/jasms.5c00207.

(24) Shishkova, E.; Hebert, A. S.; Westphall, M. S.; Coon, J. J. Ultra-High Pressure (>30,000 Psi) Packing of Capillary Columns Enhancing Depth of Shotgun Proteomic Analyses. Anal. Chem. 2018, 90 (19), 11503–11508. https://doi.org/10.1021/acs.analchem.8b02766.

(25) Schindelin, J.; Rueden, C. T.; Hiner, M. C.; Eliceiri, K. W. The ImageJ Ecosystem: An Open Platform for Biomedical Image Analysis. Mol. Reprod. Dev. 2015, 82 (7–8), 518–529. https://doi.org/10.1002/mrd.22489.

(26) Wilkins, M. R.; Lindskog, I.; Gasteiger, E.; Bairoch, A.; Sanchez, J.-C.; Hochstrasser, D. F.; Appel, R. D. Detailed Peptide Characterization Using PEPTIDEMASS – a World-Wide-Web-Accessible Tool. ELECTROPHORESIS 1997, 18 (3–4), 403–408. https://doi.org/10.1002/elps.1150180314.

(27) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P.; SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nat. Methods 2020, 17 (3), 261–272. https://doi.org/10.1038/s41592-019-0686-2.

(28) Silveira, J. A.; Schultz, G. A.; Rucker, K.; Lin, Y.; Tsai, M.; Belford, M.; Boeser, C.; Wouters, E. R. Mapping Nano-Electrospray Ionization Plumes on an Orbitrap Fusion Lumos Tribrid Mass Spectrometer Equipped with FAIMS Pro. https://assets.thermofisher.com/TFS-Assets/CMD/Flyers/fl-489255-asms23-mapping-nano-electrospray-ionization-plumes-on-an-orbitrap-fusion-lumos-tribrid-mass-spectrometer-fl489255-en.pdf.

(29) Ion transfer tube • Orbitrap Excedion Pro Operating Manual • Technical documentation. https://docs.thermofisher.com/r/Orbitrap-Excedion-Pro-Operating-Manual/2397502091v1en-US (accessed 2025-11-07).

(30) Petro, E. M.; Gallud, X.; Hampl, S. K.; Schroeder, M.; Geiger, C.; Lozano, P. C. Multiscale Modeling of Electrospray Ion Emission. J. Appl. Phys. 2022, 131 (19), 193301. https://doi.org/10.1063/5.0065615.

(31) Pauly, M.; Sroka, M.; Reiss, J.; Rinke, G.; Albarghash, A.; Vogelgesang, R.; Hahne, H.; Kuster, B.; Sesterhenn, J.; Kern, K.; Rauschenbach, S. A Hydrodynamically Optimized Nano-Electrospray Ionization Source and Vacuum Interface. Analyst 2014, 139 (8), 1856–1867. https://doi.org/10.1039/C3AN01836A.

(32) Thuppul, A.; Collins, A. L.; Wright, P. L.; Uchizono, N. M.; Wirz, R. E. Mass Flux and Current Density Distributions of Electrospray Plumes. J. Appl. Phys. 2021, 130 (10), 103301. https://doi.org/10.1063/5.0056761.

(33) Collins, A. L.; Wright, P. L.; Uchizono, N. M.; Wirz, R. E. High Angle Mass Flux of an Electrospray Plume. J. Electr. Propuls. 2022, 1 (1), 32. https://doi.org/10.1007/s44205-022-00031-w.

(34) Ramos-Tomás, R.; Villegas-Prados, D.; de Saavedra, B.; Cruz, J.; Sepúlveda, B. Impact of Tip Angle on the Divergence and Efficiency of Electrospray Thrusters. ACS Appl. Electron. Mater. 2024, 6 (10), 7319–7328. https://doi.org/10.1021/acsaelm.4c01224.

(35) Deng, W.; Gomez, A. Influence of Space Charge on the Scale-up of Multiplexed Electrosprays. J. Aerosol Sci. 2007, 38 (10), 1062–1078. https://doi.org/10.1016/j.jaerosci.2007.08.005.

(36) Gamero-Castaño, M. The Structure of Electrospray Beams in Vacuum. J. Fluid Mech. 2008, 604, 339–368. https://doi.org/10.1017/S0022112008001316.

(37) Girod, M.; Dagany, X.; Antoine, R.; Dugourd, P. Relation between Charge State Distributions of Peptide Anions and pH Changes in the Electrospray Plume. A Mass Spectrometry and Optical Spectroscopy Investigation. Int. J. Mass Spectrom. 2011, 308 (1), 41–48. https://doi.org/10.1016/j.ijms.2011.07.020.

(38) Wang, R.; Zenobi, R. Evolution of the Solvent Polarity in an Electrospray Plume. J. Am. Soc. Mass Spectrom. 2010, 21 (3), 378–385. https://doi.org/10.1016/j.jasms.2009.10.022.

(39) Beck, S.; Michalski, A.; Raether, O.; Lubeck, M.; Kaspar, S.; Goedecke, N.; Baessmann, C.; Hornburg, D.; Meier, F.; Paron, I.; Kulak, N. A.; Cox, J.; Mann, M. The Impact II, a Very High-Resolution Quadrupole Time-of-Flight Instrument (QTOF) for Deep Shotgun Proteomics *. Mol. Cell. Proteomics 2015, 14 (7), 2014–2029. https://doi.org/10.1074/mcp.M114.047407.

(40) Millar, A. L.; Langridge, J. I.; McKenna, T.; Pringle, S.; Young, P.; Giles, K. The Use of A Nanolockspray Electrospray Interface for Exact Mass Proteomics Studies. Waters Corp. 2022. https://www.waters.com/content/dam/waters/en/app-notes/2003/720000738/720000738-en.pdf.

(41) Agilent G1992A Nanospray Ion Source User Guide. Agil. Technol. 2012. https://www.agilent.com/cs/library/usermanuals/public/Copy%20of%20G1988-90000_Nanospray_UserGuide.pdf (Accessed January 12, 2026).

(42) Mayer, T.; Borsdorf, H. Ion Transfer from an Atmospheric Pressure Ion Funnel into a Mass Spectrometer with Different Interface Options: Simulation-Based Optimization of Ion Transmission Efficiency. Rapid Commun. Mass Spectrom. 2016, 30 (3), 372–378. https://doi.org/10.1002/rcm.7451.

Chapter 5

Conclusions and Outlook

Conclusions & Future Directions

The research presented in this dissertation focuses on the development and improvement of data acquisition methods for mass spectrometry-based bottom-up proteomics. Proteomics has potential to offer comprehensive insight into the states of biological systems to enable discovery and understanding of complex biological phenomena. For the realization of this potential, this dissertation presents progress towards proteomics methods that enable comprehensive characterization and quantification of proteins while maintaining high reproducibility and accelerating throughput to enable larger scale studies. This ambitious goal is achieved by leveraging new technology to develop improved proteomics methods and is described within three key studies that address these challenges: the development of a high performance phosphoproteomics method, introduction of a new label-free multiplexing method for higher throughput, and characterization of emitter positioning effects as a critical component of the experimental setup. The significance, limitations, and future directions of each of these studies are discussed below.

The research presented in **Chapter 2** provides an optimized method for phosphoproteomics using the newly introduced Orbitrap Astral mass spectrometer. Technical evaluation of the method demonstrated high phosphoproteomic depth, good reproducibility, good site localization accuracy, and high quantitative linearity. The method performance was then benchmarked against other instrument platforms through the analysis of control and EGF treatment groups of HeLa cells, yielding >200% more phosphorylation site detections than methods on other modern instrument platforms while leading to the same biological conclusions. These results demonstrate high method performance and validate the method's ability to generate the correct biological conclusions. The

method was then applied to generate a comprehensive atlas of protein phosphorylation across various mouse tissues. Compared to the most recent study analyzing phosphorylation across mouse tissues[1], our dataset detects >2-fold more phosphorylation sites while increasing the throughput of data collection by ~3-fold, representing a substantial improvement in the ability to comprehensively characterize protein phosphorylation events and serving as a valuable resource for the wider biochemistry community. This dataset confirmed previous findings regarding the tendency of phosphorylation events to occur on disordered regions of proteins[2–4], and downstream analysis highlighted multiple novel phosphorylation sites with potential human disease relevance. The phosphoproteomics method developed and the dataset presented in this work have been used and applied to multiple studies in our research group and have been used by the wider proteomics community[5–8], clearly indicating the value of the method and the dataset generated here.

While our method and the phosphoproteomics atlas has already found utility, there are several limitations of this study. For benchmarking of the phosphoproteomic method to the other instrumentation platforms, the comparison is very direct for the Orbitrap Ascend dataset, as this dataset was generated using an identical chromatographic setup/method and the aliquots of the same samples were analyzed. In contrast, the timsTOF Pro dataset was previously generated in another study using an earlier generation instrument using a different sample preparation, different chromatographic setup, and slightly longer active acquisition period. Another limitation of this study regards the generation of the mouse phosphoproteomic atlas. Due to material and instrumentation access limitations, the atlas only includes one biological replicate, constituting tissues pooled from three male mice, and would have been improved by the inclusion of more biological replicates as well as female mice samples. Despite these limitations, this study

represents a substantial innovation in the analysis of the phosphoproteome. Future directions for this work would be the continued application to biological studies as well as adaptation of this method to analyze other types of PTMs and the extension to lower input material analyses.

For modern proteome analysis using nanoflow liquid chromatography-mass spectrometry, the sample loading and column washing/equilibration steps, referred to as the 'LC overhead', constitute a notable fraction of the overall analysis time for an LC-MS/MS workflow. To accelerate the throughput of proteomics by bypassing the bottleneck of the LC overhead steps, **Chapter 3** introduces the SynchroSep-MS method. This new approach to sample multiplexing performs multiple gradient separations on multiple LC columns that are simultaneously eluting into the same mass spectrometer, resulting in improved throughput due to parallelized sample analysis without the need for chemical derivatization. As the mass spectrometer itself is unable to distinguish which column a peptide is coming from, data analysis is enabled by introducing a retention time offset between columns that allows for assigning eluting peptides to their originating column. This study demonstrates proof-of-concept for this method with a setup utilizing 2 columns and 2 LCs to analyze 2 samples at once with a single mass spectrometer, yielding an ~46% improvement in throughput. The use of the Orbitrap Astral and a narrow isolation window DIA scheme for high selectivity was employed to handle the increased spectral complexity observed with this method and a new data processing workflow was developed and validated using the DIA-NN software.[9] The method demonstrates retention of most protein identifications observed with a single column method while achieving higher throughput and maintaining similar quantitative performance. This study successfully provides proof-of-concept for the SynchroSep-MS method as a potential avenue for enhanced proteomics throughput.

The value of this study primarily consists in the proof-of-concept demonstration. For more routine application of this technique, further development is required. As a key step in making the method more accessible to the community, I note that the modifications in the DIA-NN software were incorporated in the currently available version of the DIA-NN software.[10] Further development of the experimental setup will be necessary as the setup used in this study was complicated and not very robust. Key improvements would include use of a dual-pump LC, instrument control with a single computer, and development of an electrospray source that makes installation and emitter alignment of multiple columns easier. We observed that the alignment of the emitters seemed to have a large effect on the relative performances of the columns. This challenge will be discussed later. Optimization of the retention time offset stability over extended analysis periods and refinement of the retention time modeling in the data processing workflow would enable more routine use and increase parallelization of sample analyses by enabling use of a smaller retention time offset. After optimizing the experimental setup, it is also of interest to perform more quantitative benchmarking through the use of the matrix-matched calibration curve approach[11] and apply the method to a biological study. Comparing these results to those generated with a traditional setup would serve as a validation of the utility of the method for biological analysis. Scaling up the method to use more columns would enable increased throughput gains. Handling the increased spectral complexity arising from the use of >2 columns would likely be enabled by further development of mass spectrometry instrumentation. The use of gas-phase ion mobility separations such as FAIMS or TIMS could decrease the complexity of ions entering the mass spectrometer. Overall, this method has promise for improving the throughput of proteome

analysis, but further simplifications in the implementation of the method would be needed to make it an attractive workflow.

As mentioned above, we found that emitter positioning was of particular importance to the performance of the SynchroSep-MS method. For modern mass spectrometry instrumentation, we found that the literature was lacking in reports of the effect of emitter position for nanoflow electrospray in all three dimensions while using experimental setups similar to those used in modern proteomics analysis. To both address challenges with the SynchroSep-MS method and to fill this gap in the literature, **Chapter 4** presents a detailed characterization of how emitter positioning affects signal intensity of tryptic peptides electrosprayed into the high-capacity inlet capillary common to many Thermo Fisher Scientific mass spectrometers which are broadly used in the proteomics community.[12] The results demonstrate that using emitter positions within 1 to 2 mm of the optima will maintain relatively consistent signal, serving as a positioning tolerance recommendation for researchers. Additionally, we found that the tolerance required to maintain relatively consistent signal increased in magnitude as the emitter was moved further away from the inlet in the direction of the inlet axis, suggesting that moving the emitters slightly further away from the inlet would minimize intensity differences between emitters when using the SynchroSep-MS approach or other experimental setups that feature multiple emitters aligned at once.[13–17] This study provides a set of experiments and analyses that would be valuable to carry out with other experimental setups, including those with different emitter types, different flow rates, and other mass spectrometers featuring different atmospheric pressure interfaces. The impact of using front-end ion mobility separations such as FAIMS on the emitter positioning results would also be

interesting. Finally, a direct measurement of emitter positioning results with a multi-emitter setup would provide more conclusive information to improve the SynchroSep-MS method.

As a combination of work presented in this dissertation, development of a phosphoproteomic method that takes advantage of the increased throughput of the SynchroSep-MS approach would be of interest. Adapting the SynchroSep-MS method for phosphoproteomics would present a challenge due to the need for higher selectivity for effective phosphorylation site localization, but this challenge could be alleviated through the use of ion mobility techniques to simplify the mass spectra.[18–20] As another future direction, the SynchroSep-MS method could be combined with previous multi-column work.[13–16,21–30] Construction of an experimental setup with two sets of columns (for instance, 2 columns in Set 1 and 2 columns in Set 2) aligned to the mass spectrometer at the same time with one set actively eluting into the mass spectrometer while the column washing, equilibration, and sample loading steps were performed on the other set would combine the SynchroSep-MS method with previous multi-column methods for greatly improved use of mass spectrometry instrument time.

Outlook

The work presented in this dissertation addresses some of the key challenges in the analysis of complex proteomics samples with an emphasis on the development of data acquisition approaches to improve throughput and depth while maintaining high reproducibility and accuracy. However, the field still faces outstanding challenges for data acquisition methods, primarily the development of higher throughput methods, enabling more sensitive analysis for analysis of small sample amounts, and improving quantitative performance and method reproducibility.

The main approach to improving method throughput is shortening the gradient length[31–35]; however, this approach by itself typically decreases the proteomic depth achieved by the method, which can be detrimental for studies where lower abundance proteins are of biological importance.[36,37] To maintain proteomics depth, the use of shorter gradients should be accompanied by the use of higher efficiency separations and/or faster and more sensitive mass spectrometers. Higher efficiency separations increase analysis throughput without sacrificing analyte resolution[38], but are not as helpful if the mass spectrometer is not fast and sensitive enough to sample the narrower elution peaks.[39] For data-independent acquisition, faster scanning instruments enable higher selectivity for MS^2 analysis by allowing the use of narrower isolation windows, which can compensate for the reduced separation efficiency of shorter gradient separations. The use of narrow isolation windows enabled by the fast-scanning speed of the Orbitrap Astral mass spectrometer was demonstrated to have a substantial effect on method performance for the phosphoproteomic method described in **Chapter 2**. Additionally, narrow window isolation with the Orbitrap Astral likely supports effective analysis with the increased spectral complexity observed with the SynchroSep-MS method described in **Chapter 3**. These examples illustrate the importance of time-of-flight mass analyzers in modern proteomics research, as these analyzers feature the fastest scan speeds with recent innovations yielding high resolution and mass accuracy while achieving high ion transmission.[40–47] Another way to enhance proteomics throughput while maintaining method performance is to alleviate bottlenecks in data acquisition such as the LC overhead. Thus, approaches like the SynchroSep-MS method and other multi-column methods[13–16,21–30] that parallelize LC steps enable a way to bypass this throughput limitation. The combination of methods that alleviate LC bottlenecks with fast, sensitive, and high-resolution mass spectrometers

for the development of higher throughput methods with deep proteome coverage will make comprehensive characterization of the proteome achievable for larger scale studies.

Sensitive analysis of low amounts of sample material represents another challenge for proteome analysis. For many biological studies, the amount of material that can be generated from biological samples is limited, and this limitation is especially severe when analyzing post-translational modifications (PTMs) such as phosphorylation due to the need for sample enrichment to detect low stoichiometry modifications. The work in **Chapter 2** demonstrates high detection depth with low sample loads of enriched phosphopeptides, demonstrating potential for high sensitivity phosphoproteomics, which will be further enhanced by the use of phosphopeptide enrichment workflows that require much less input material.[48–53] The characterization of emitter positioning effects in **Chapter 4** may also play a role in improving sensitivity by understanding the tolerances of emitter positioning needed to maximize ion signal while maintaining good reproducibility. Overall, improvements in method sensitivity will enable new analyses that were not previously possible.

Finally, quantitative performance and reproducibility are important for proteomics methods.[11,54–56] The development of methods that enhance quantitative accuracy and reproducibility results in improved biological analysis by enabling measurement of more subtle distinctions between biological conditions. Assessment of quantitative performance using approaches like a matrix-matched calibration curve enable reliable comparison of different methods[11] and will become especially important as new methods and instrumentation become available.[54–60] As higher throughput methods using shorter gradients become more common, quantitative assessment will be even more important, as shorter gradients can lead to worse

chromatographic peak sampling and increased co-elution and quantitative interference.[61] The methods developed in **Chapter 2** and **Chapter 3** both demonstrate assessments of quantitative performance. While good quantitative linearity was demonstrated using synthetic phosphopeptides in **Chapter 2**, the development of a matrix-matched calibration curve approach for phosphoproteomics would allow a more global assessment of method performance. For the SynchroSep-MS method in **Chapter 3**, a dilution series was used to confirm consistent quantitative performance across the single- and dual-column methods, but a matrix-matched calibration curve approach could provide more fine-grained insight into quantitative impacts observed when moving from the traditional single-column to the dual-column SynchroSep-MS method. For mass spectrometry analysis, higher signal intensity generally leads to greater precision. The characterizations of emitter positioning effects presented in **Chapter 4** will enable a more strategic approach to optimizing signal intensity. This work could also help improve reproducibility by guiding the selection of consistent emitter positions to use when the electrospray emitter must be repositioned multiple times throughout a study. Especially as modern proteomics methods began achieving comprehensive characterization of proteomes, a focus on quantitative performance and reproducibility will become increasingly important for future development efforts.

The proteomics field now offers many approaches for sample preparation, data acquisition, and data processing. The workflow that is selected should be aligned with the goals of the biological study. For example, a study aimed at generating a comprehensive set of observed protein modifications in a particular set of samples might select a workflow optimized for identification depth.[62] In contrast, a clinical proteomics study aimed at detecting biomarkers might employ a method optimized for quantitative accuracy and precision with high robustness to facilitate

sustained analysis of a large sample set.[63] Similarly, different studies might prioritize minimization of false positive versus false negative results. A discovery proteomics analysis performed in cell lines or a simple model organism, where downstream validation is easier, might tolerate an increased false positive rate compared to a proteomics study that will inform the selection of targets for a larger-scale clinical study. Careful consideration of the goals of a biological study will inform the selection of a proteomics workflow with the necessary performance criteria.

Overall, the ongoing development of new data acquisition methods expands the possibilities of measurement. In this dissertation, substantial progress is made in the development of higher throughput methods that achieve high proteomic depth, and groundwork is laid toward improving quantitative performance and reproducibility when performing comprehensive proteome characterization. The continued development of these proteomics methods will enable the discovery of new biochemical insight as highlighted in the quote at the beginning of this dissertation: "…improved analytical methods always yield new chemical results." [64,65]

References

(1) Giansanti, P.; Samaras, P.; Bian, Y.; Meng, C.; Coluccio, A.; Frejno, M.; Jakubowsky, H.; Dobiasch, S.; Hazarika, R. R.; Rechenberger, J.; Calzada-Wack, J.; Krumm, J.; Mueller, S.; Lee, C.-Y.; Wimberger, N.; Lautenbacher, L.; Hassan, Z.; Chang, Y.-C.; Falcomatà, C.; Bayer, F. P.; Bärthel, S.; Schmidt, T.; Rad, R.; Combs, S. E.; The, M.; Johannes, F.; Saur, D.; de Angelis, M. H.; Wilhelm, M.; Schneider, G.; Kuster, B. Mass Spectrometry-Based Draft of the Mouse Proteome. Nat. Methods 2022, 19 (7), 803–811. https://doi.org/10.1038/s41592-022-01526-y.

(2) Bludau, I.; Willems, S.; Zeng, W.-F.; Strauss, M. T.; Hansen, F. M.; Tanzer, M. C.; Karayel, O.; Schulman, B. A.; Mann, M. The Structural Context of Posttranslational Modifications at a Proteome-Wide Scale. PLoS Biol. 2022, 20 (5), e3001636. https://doi.org/10.1371/journal.pbio.3001636.

(3) Tyanova, S.; Cox, J.; Olsen, J.; Mann, M.; Frishman, D. Phosphorylation Variation during the Cell Cycle Scales with Structural Propensities of Proteins. PLoS Comput. Biol. 2013, 9 (1), e1002842. https://doi.org/10.1371/journal.pcbi.1002842.

(4) Iakoucheva, L. M.; Radivojac, P.; Brown, C. J.; O'Connor, T. R.; Sikes, J. G.; Obradovic, Z.; Dunker, A. K. The Importance of Intrinsic Disorder for Protein Phosphorylation. Nucleic Acids Res. 2004, 32 (3), 1037–1049. https://doi.org/10.1093/nar/gkh253.

(5) Garrido-Rodriguez, M.; Potel, C.; Burtscher, M. L.; Becher, I.; Rodriguez-Mier, P.; Müller-Dott, S.; Savitski, M. M.; Saez-Rodriguez, J. Benchmarking EGF Signaling Pathway Inference Using Phosphoproteomics and Kinase-Substrate Interactions. Nat. Commun. 2026, 17 (1), 2071. https://doi.org/10.1038/s41467-026-69332-0.

(6) Wen, B.; Hsu, C.; Shteynberg, D.; Zeng, W.-F.; Riffle, M.; Chang, A.; Mudge, M. C.; Nunn, B. L.; MacLean, B. X.; Berg, M. D.; Villén, J.; MacCoss, M. J.; Noble, W. S. Carafe Enables High Quality in Silico Spectral Library Generation for Data-Independent Acquisition Proteomics. Nat. Commun. 2025, 16 (1), 9815. https://doi.org/10.1038/s41467-025-64928-4.

(7) Park, J.; Kim, K.; Kang, U.-B.; Kim, S. DelPi Learns Generalizable Peptide–Signal Correspondence for Mass Spectrometry-Based Proteomics. bioRxiv January 7, 2026, p 2026.01.06.697814. https://doi.org/10.64898/2026.01.06.697814.

(8) Garrido-Rodriguez, M.; Potel, C.; Burtscher, M. L.; Becher, I.; Rodriguez-Mier, P.; Müller-Dott, S.; Savitski, M. M.; Saez-Rodriguez, J. Evaluating Signaling Pathway Inference from Kinase-Substrate Interactions and Phosphoproteomics Data. bioRxiv October 22, 2024, p 2024.10.21.619348. https://doi.org/10.1101/2024.10.21.619348.

(9) Demichev, V.; Messner, C. B.; Vernardis, S. I.; Lilley, K. S.; Ralser, M. DIA-NN: Neural Networks and Interference Correction Enable Deep Proteome Coverage in High Throughput. Nat. Methods 2020, 17 (1), 41–44. https://doi.org/10.1038/s41592-019-0638-x.

(10) vdemichev. DIA-NN 2.0-series Updates · vdemichev/DiaNN · Discussion #1366. GitHub. https://github.com/vdemichev/DiaNN/discussions/1366 (accessed 2026-03-03).

(11) Pino, L. K.; Searle, B. C.; Yang, H.-Y.; Hoofnagle, A. N.; Noble, W. S.; MacCoss, M. J. Matrix-Matched Calibration Curves for Assessing Analytical Figures of Merit in Quantitative Proteomics. J. Proteome Res. 2020, 19 (3), 1147–1153. https://doi.org/10.1021/acs.jproteome.9b00666.

(12) Peters-Clarke, T. M.; Coon, J. J.; Riley, N. M. Instrumentation at the Leading Edge of Proteomics. Anal. Chem. 2024, 96 (20), 7976–8010. https://doi.org/10.1021/acs.analchem.3c04497.

(13) Orton, D. J.; Wall, M. J.; Doucette, A. A. Dual LC–MS Platform for High-Throughput Proteome Analysis. J. Proteome Res. 2013, 12 (12), 5963–5970. https://doi.org/10.1021/pr400738a.

(14) Webber, K. G. I.; Truong, T.; Johnston, S. M.; Zapata, S. E.; Liang, Y.; Davis, J. M.; Buttars, A. D.; Smith, F. B.; Jones, H. E.; Mahoney, A. C.; Carson, R. H.; Nwosu, A. J.; Heninger, J. L.; Liyu, A. V.; Nordin, G. P.; Zhu, Y.; Kelly, R. T. Label-Free Profiling of up to 200 Single-Cell Proteomes per Day Using a Dual-Column Nanoflow Liquid Chromatography Platform. Anal. Chem. 2022, 94 (15), 6017–6025. https://doi.org/10.1021/acs.analchem.2c00646.

(15) Xie, X.; Truong, T.; Huang, S.; Johnston, S. M.; Hovanski, S.; Robinson, A.; Webber, K. G. I.; Lin, H.-J. L.; Mun, D.-G.; Pandey, A.; Kelly, R. T. Multicolumn Nanoflow Liquid Chromatography with Accelerated Offline Gradient Generation for Robust and Sensitive Single-Cell Proteome Profiling. Anal. Chem. 2024, 96 (26), 10534–10542. https://doi.org/10.1021/acs.analchem.4c00878.

(16) Hosp, F.; Scheltema, R. A.; Eberl, H. C.; Kulak, N. A.; Keilhauer, E. C.; Mayr, K.; Mann, M. A Double-Barrel Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) System to Quantify 96 Interactomes per Day*. Mol. Cell. Proteomics 2015, 14 (7), 2030–2041. https://doi.org/10.1074/mcp.O115.049460.

(17) Lancaster, N. M.; Chen, L.-Y.; Zhao, B.; Anderson, B. J.; Probasco, M. D.; Demichev, V.; Polasky, D. A.; Nesvizhskii, A. I.; Overmyer, K. A.; Quarmby, S. T.; Coon, J. J. SynchroSep-MS: Parallel LC Separations for Multiplexed Proteomics. J. Am. Soc. Mass Spectrom. 2025, 36 (9), 1979–1987. https://doi.org/10.1021/jasms.5c00207.

(18) Skowronek, P.; Thielert, M.; Voytik, E.; Tanzer, M. C.; Hansen, F. M.; Willems, S.; Karayel, O.; Brunner, A.-D.; Meier, F.; Mann, M. Rapid and In-Depth Coverage of the (Phospho-

)Proteome With Deep Libraries and Optimal Window Design for Dia-PASEF. Mol. Cell. Proteomics 2022, 21 (9). https://doi.org/10.1016/j.mcpro.2022.100279.

(19) Muehlbauer, L. K.; Hebert, A. S.; Westphall, M. S.; Shishkova, E.; Coon, J. J. Global Phosphoproteome Analysis Using High-Field Asymmetric Waveform Ion Mobility Spectrometry on a Hybrid Orbitrap Mass Spectrometer. Anal. Chem. 2020, 92 (24), 15959–15967. https://doi.org/10.1021/acs.analchem.0c03415.

(20) Oliinyk, D.; Meier, F. Ion Mobility-Resolved Phosphoproteomics with Dia-PASEF and Short Gradients. PROTEOMICS 2023, 23 (7–8), 2200032. https://doi.org/10.1002/pmic.202200032.

(21) Bonneil, E.; Tessier, S.; Carrier, A.; Thibault, P. Multiplex Multidimensional nanoLC-MS System for Targeted Proteomic Analyses. ELECTROPHORESIS 2005, 26 (24), 4575–4589. https://doi.org/10.1002/elps.200500603.

(22) Livesay, E. A.; Tang, K.; Taylor, B. K.; Buschbach, M. A.; Hopkins, D. F.; LaMarche, B. L.; Zhao, R.; Shen, Y.; Orton, D. J.; Moore, R. J.; Kelly, R. T.; Udseth, H. R.; Smith, R. D. Fully Automated Four-Column Capillary LC−MS System for Maximizing Throughput in Proteomic Analyses. Anal. Chem. 2008, 80 (1), 294–302. https://doi.org/10.1021/ac701727r.

(23) Lee, H.; Lee, J. H.; Kim, H.; Kim, S.-J.; Bae, J.; Kim, H. K.; Lee, S.-W. A Fully Automated Dual-Online Multifunctional Ultrahigh Pressure Liquid Chromatography System for High-Throughput Proteomics Analysis. J. Chromatogr. A 2014, 1329, 83–89. https://doi.org/10.1016/j.chroma.2013.12.084.

(24) Lee, H.; Mun, D.-G.; Bae, J.; Kim, H.; Yeon Oh, S.; Soo Park, Y.; Lee, J.-H.; Lee, S.-W. A Simple Dual Online Ultra-High Pressure Liquid Chromatography System (sDO-UHPLC) for High Throughput Proteome Analysis. Analyst 2015, 140 (16), 5700–5706. https://doi.org/10.1039/C5AN00639B.

(25) Kreimer, S.; Haghani, A.; Binek, A.; Hauspurg, A.; Seyedmohammad, S.; Rivas, A.; Momenzadeh, A.; Meyer, J. G.; Raedschelders, K.; Van Eyk, J. E. Parallelization with Dual-Trap Single-Column Configuration Maximizes Throughput of Proteomic Analysis. Anal. Chem. 2022, 94 (36), 12452–12460. https://doi.org/10.1021/acs.analchem.2c02609.

(26) Kreimer, S.; Binek, A.; Chazarin, B.; Cho, J. H.; Haghani, A.; Hutton, A.; Marbán, E.; Mastali, M.; Meyer, J. G.; Mesquita, T.; Song, Y.; Van Eyk, J.; Parker, S. High-Throughput Single-Cell Proteomic Analysis of Organ-Derived Heterogeneous Cell Populations by Nanoflow Dual-Trap Single-Column Liquid Chromatography. Anal. Chem. 2023, 95 (24), 9145–9150. https://doi.org/10.1021/acs.analchem.3c00213.

(27) Staes, A.; Boucher, K.; Dufour, S.; Maia, T. M.; Timmerman, E.; Haver, D. V.; Pauwels, J.; Demol, H.; Vandenbussche, J.; Gevaert, K.; Impens, F.; Devos, S. High-Throughput

Nanoflow Proteomics Using a Dual-Column Electrospray Source. Anal. Chem. 2024, 96 (17), 6534–6539. https://doi.org/10.1021/acs.analchem.4c00845.

(28) Chen, L.; Zhang, Z.; Matsumoto, C.; Gao, Y. High-Throughput Proteomics Enabled by a Fully Automated Dual-Trap and Dual-Column LC-MS. Anal. Chem. 2024, 96 (24), 9761–9766. https://doi.org/10.1021/acs.analchem.3c03182.

(29) Wang, H.; Hanash, S. M. Increased Throughput and Reduced Carryover of Mass Spectrometry-Based Proteomics Using a High-Efficiency Nonsplit Nanoflow Parallel Dual-Column Capillary HPLC System. J. Proteome Res. 2008, 7 (7), 2743–2755. https://doi.org/10.1021/pr700876g.

(30) Shen, Y.; Tolić, N.; Zhao, R.; Paša-Tolić, L.; Li, L.; Berger, S. J.; Harkewicz, R.; Anderson, G. A.; Belov, M. E.; Smith, R. D. High-Throughput Proteomics Using High-Efficiency Multiple-Capillary Liquid Chromatography with On-Line High-Performance ESI FTICR Mass Spectrometry. Anal. Chem. 2001, 73 (13), 3011–3021. https://doi.org/10.1021/ac001393n.

(31) Messner, C. B.; Demichev, V.; Bloomfield, N.; Yu, J. S. L.; White, M.; Kreidl, M.; Egger, A.-S.; Freiwald, A.; Ivosev, G.; Wasim, F.; Zelezniak, A.; Jürgens, L.; Suttorp, N.; Sander, L. E.; Kurth, F.; Lilley, K. S.; Mülleder, M.; Tate, S.; Ralser, M. Ultra-Fast Proteomics with Scanning SWATH. Nat. Biotechnol. 2021, 39 (7), 846–854. https://doi.org/10.1038/s41587-021-00860-4.

(32) Ishikawa, M.; Konno, R.; Nakajima, D.; Gotoh, M.; Fukasawa, K.; Sato, H.; Nakamura, R.; Ohara, O.; Kawashima, Y. Optimization of Ultrafast Proteomics Using an LC-Quadrupole-Orbitrap Mass Spectrometer with Data-Independent Acquisition. J. Proteome Res. 2022, 21 (9), 2085–2093. https://doi.org/10.1021/acs.jproteome.2c00121.

(33) Guzman, U. H.; Martinez-Val, A.; Ye, Z.; Damoc, E.; Arrey, T. N.; Pashkova, A.; Renuse, S.; Denisov, E.; Petzoldt, J.; Peterson, A. C.; Harking, F.; Østergaard, O.; Rydbirk, R.; Aznar, S.; Stewart, H.; Xuan, Y.; Hermanson, D.; Horning, S.; Hock, C.; Makarov, A.; Zabrouskov, V.; Olsen, J. V. Ultra-Fast Label-Free Quantification and Comprehensive Proteome Coverage with Narrow-Window Data-Independent Acquisition. Nat. Biotechnol. 2024. https://doi.org/10.1038/s41587-023-02099-7.

(34) Sinn, L. R.; Wang, Z.; Alvarez, C. P.; Chelur, A.; Batruch, I.; Pribil, P.; Ludwig, D.; Tate, S.; Castro-Perez, J.; Messner, C. B.; Demichev, V.; Ralser, M. Performance Characteristics of Zeno Trap Scanning DIA for Sensitive and Quantitative Proteomics at High Throughput. PROTEOMICS 2026, 26 (1), 68–81. https://doi.org/10.1002/pmic.70093.

(35) Guzman, U. H.; Rykær, M.; Hendriks, I. A.; Stewart, H.; Denisov, E.; Hagedorn, B.; Petzoldt, J.; Kreutzmann, A.; Mueller, Y.; Arrey, T. N.; Colonius, I.; Østergaard, O.; Koenig, C.; Kraegenbring, J.; Fort, K. L.; Couzijn, E.; Hauschild, J.-P.; Hermanson, D.; Zabrouskov, V.;

Hock, C.; Damoc, E.; Olsen, Jesper. V. Higher-Throughput Proteome Profiling Enabled by Parallelized Pre-Accumulation and Optimized Ion Processing in the Orbitrap Astral Zoom Mass Spectrometer. Mol. Cell. Proteomics 2026, 25 (2), 101504. https://doi.org/10.1016/j.mcpro.2025.101504.

(36) Zubarev, R. A. The Challenge of the Proteome Dynamic Range and Its Implications for In-Depth Proteomics. PROTEOMICS 2013, 13 (5), 723–726. https://doi.org/10.1002/pmic.201200451.

(37) Anderson, N. L.; Anderson, N. G. The Human Plasma Proteome: History, Character, and Diagnostic Prospects *. Mol. Cell. Proteomics 2002, 1 (11), 845–867. https://doi.org/10.1074/mcp.R200007-MCP200.

(38) Snyder, L. R.; Kirkland, J. J.; Dolan, J. W. Introduction. In Introduction to modern liquid chromatography; Wiley: Hoboken, N.J, 2010; pp 1–17. https://doi.org/10.1002/9780470508183.

(39) Michalski, A.; Cox, J.; Mann, M. More than 100,000 Detectable Peptide Species Elute in Single Shotgun Proteomics Runs but the Majority Is Inaccessible to Data-Dependent LC−MS/MS. J. Proteome Res. 2011, 10 (4), 1785–1793. https://doi.org/10.1021/pr101060v.

(40) Stewart, H. I.; Grinfeld, D.; Giannakopulos, A.; Petzoldt, J.; Shanley, T.; Garland, M.; Denisov, E.; Peterson, A. C.; Damoc, E.; Zeller, M.; Arrey, T. N.; Pashkova, A.; Renuse, S.; Hakimi, A.; Kühn, A.; Biel, M.; Kreutzmann, A.; Hagedorn, B.; Colonius, I.; Schütz, A.; Stefes, A.; Dwivedi, A.; Mourad, D.; Hoek, M.; Reitemeier, B.; Cochems, P.; Kholomeev, A.; Ostermann, R.; Quiring, G.; Ochmann, M.; Möhring, S.; Wagner, A.; Petker, A.; Kanngiesser, S.; Wiedemeyer, M.; Balschun, W.; Hermanson, D.; Zabrouskov, V.; Makarov, A. A.; Hock, C. Parallelized Acquisition of Orbitrap and Astral Analyzers Enables High-Throughput Quantitative Analysis. Anal. Chem. 2023, 95 (42), 15656–15664. https://doi.org/10.1021/acs.analchem.3c02856.

(41) Grinfeld, D.; Stewart, H.; Balschun, W.; Skoblin, M.; Hock, C.; Makarov, A. Multi-Reflection Astral Mass Spectrometer with Isochronous Drift in Elongated Ion Mirrors. Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip. 2024, 1060, 169017. https://doi.org/10.1016/j.nima.2023.169017.

(42) Stewart, H.; Grinfeld, D.; Hagedorn, B.; Ostermann, R.; Makarov, A.; Hock, C. Proof of Principle for Enhanced Resolution Multi-Pass Methods for the Astral Analyzer. Int. J. Mass Spectrom. 2024, 498, 117203. https://doi.org/10.1016/j.ijms.2024.117203.

(43) Stewart, H.; Grinfeld, D.; Petzoldt, J.; Hagedorn, B.; Skoblin, M.; Makarov, A.; Hock, C. Crowd Control of Ions in the Astral Analyzer. J. Mass Spectrom. 2024, 59 (4), e5006. https://doi.org/10.1002/jms.5006.

(44) Cooper-Shepherd, D. A.; Wildgoose, J.; Kozlov, B.; Johnson, W. J.; Tyldesley-Worster, R.; Palmer, M. E.; Hoyes, J. B.; McCullagh, M.; Jones, E.; Tonge, R.; Marsden-Edwards, E.; Nixon, P.; Verenchikov, A.; Langridge, J. I. Novel Hybrid Quadrupole-Multireflecting Time-of-Flight Mass Spectrometry System. J. Am. Soc. Mass Spectrom. 2023, 34 (2), 264–272. https://doi.org/10.1021/jasms.2c00281.

(45) Verenchikov, A. N.; Kirillov, S. N.; Vorobyev, A. V.; Makarov, V. V.; Yavor, M. I.; Tonge, R. P.; Langridge, J. I. Multi Reflecting TOF MS Approaching Resolution of 1,000,000 in a Wide Mass Range. Int. J. Mass Spectrom. 2025, 508, 117395. https://doi.org/10.1016/j.ijms.2024.117395.

(46) Verenchikov, A. N.; Wildgoose, J.; Kirillov, S. N.; Vorobyev, A. V.; Makarov, V. V.; Gethings, L. A.; Tonge, R. P.; Daly, M. E.; Johnson, W. J.; Langridge, J. I. A Novel Compact Multi-Reflecting Time-of-Flight Mass Spectrometer. J. Am. Soc. Mass Spectrom. 2026, 37 (3), 601–611. https://doi.org/10.1021/jasms.5c00321.

(47) Vorobyev, A. V.; Makarov, V. V.; Kozlov, S. P.; Verenchikov, A. N.; Ivanov, M. V.; Gorshkov, M. V. A Novel Ultrahigh-Resolution Y-Injection Multireflecting Time-of-Flight Mass Spectrometer for Bottom-Up Proteomics. Anal. Chem. 2025, 97 (45), 24946–24952. https://doi.org/10.1021/acs.analchem.5c04182.

(48) Humphrey, S. J.; Karayel, O.; James, D. E.; Mann, M. High-Throughput and High-Sensitivity Phosphoproteomics with the EasyPhos Platform. Nat. Protoc. 2018, 13 (9), 1897–1916. https://doi.org/10.1038/s41596-018-0014-9.

(49) Post, H.; Penning, R.; Fitzpatrick, M. A.; Garrigues, L. B.; Wu, W.; MacGillavry, H. D.; Hoogenraad, C. C.; Heck, A. J. R.; Altelaar, A. F. M. Robust, Sensitive, and Automated Phosphopeptide Enrichment Optimized for Low Sample Amounts Applied to Primary Hippocampal Neurons. J. Proteome Res. 2017, 16 (2), 728–737. https://doi.org/10.1021/acs.jproteome.6b00753.

(50) Low, T. Y.; Mohtar, M. A.; Lee, P. Y.; Omar, N.; Zhou, H.; Ye, M. Widening the Bottleneck of Phosphoproteomics: Evolving Strategies for Phosphopeptide Enrichment. Mass Spectrom. Rev. 2021, 40 (4), 309–333. https://doi.org/10.1002/mas.21636.

(51) Bortel, P.; Piga, I.; Koenig, C.; Gerner, C.; Martinez-Val, A.; Olsen, J. V. Systematic Optimization of Automated Phosphopeptide Enrichment for High-Sensitivity Phosphoproteomics. Mol. Cell. Proteomics 2024, 23 (5). https://doi.org/10.1016/j.mcpro.2024.100754.

(52) Oliinyk, D.; Will, A.; Schneidmadel, F. R.; Böhme, M.; Rinke, J.; Hochhaus, A.; Ernst, T.; Hahn, N.; Geis, C.; Lubeck, M.; Raether, O.; Humphrey, S. J.; Meier, F. µPhos: A Scalable and Sensitive Platform for High-Dimensional Phosphoproteomics. Mol. Syst. Biol. 2024, 20 (8), 972–995. https://doi.org/10.1038/s44320-024-00050-9.

(53) Oliinyk, D.; Heymann, T.; Henneberg, L.; Bardziukova, A.; Thielert, M.; Kjærgaard, J.; Zwiebel, M.; Vasconez, S.; Oeller, M.; Metousis, A.; Ugur, E.; Rodriguez, E. H.; Rosenberger, F. A.; Schallenberg, S.; Klauschen, F.; Mann, M. nanoPhos Enables Ultra-Sensitive and Cell-Type Resolved Spatial Phosphoproteomics. bioRxiv December 21, 2025, p 2025.05.29.656770. https://doi.org/10.1101/2025.05.29.656770.

(54) Heil, L. R.; Damoc, E.; Arrey, T. N.; Pashkova, A.; Denisov, E.; Petzoldt, J.; Peterson, A. C.; Hsu, C.; Searle, B. C.; Shulman, N.; Riffle, M.; Connolly, B.; MacLean, B. X.; Remes, P. M.; Senko, M. W.; Stewart, H. I.; Hock, C.; Makarov, A. A.; Hermanson, D.; Zabrouskov, V.; Wu, C. C.; MacCoss, M. J. Evaluating the Performance of the Astral Mass Analyzer for Quantitative Proteomics Using Data-Independent Acquisition. J. Proteome Res. 2023, 22 (10), 3290–3300. https://doi.org/10.1021/acs.jproteome.3c00357.

(55) Hsu, C.; Shulman, N.; Stewart, H. I.; Petzoldt, J.; Pashkova, A.; Plubell, D. L.; Denisov, E.; Hagedorn, B.; Damoc, N. E.; MacLean, B. X.; Remes, P. M.; Canterbury, J. D.; Makarov, A. A.; Hock, C.; Zabrouskov, V.; Wu, C. C.; MacCoss, M. J. Evaluation of a Prototype Orbitrap Astral Zoom Mass Spectrometer for Quantitative Proteomics─Beyond Identification Lists. J. Proteome Res. 2025, 24 (11), 5742–5754. https://doi.org/10.1021/acs.jproteome.5c00593.

(56) Beimers, W. F.; Overmyer, K. A.; Sinitcyn, P.; Lancaster, N. M.; Quarmby, S. T.; Coon, J. J. A Technical Evaluation of Plasma Proteomics Technologies. bioRxiv January 13, 2025, p 2025.01.08.632035. https://doi.org/10.1101/2025.01.08.632035.

(57) Heil, L. R.; Remes, P. M.; Canterbury, J. D.; Yip, P.; Barshop, W. D.; Wu, C. C.; MacCoss, M. J. Dynamic Data-Independent Acquisition Mass Spectrometry with Real-Time Retrospective Alignment. Anal. Chem. 2023, 95 (32), 11854–11858. https://doi.org/10.1021/acs.analchem.3c00903.

(58) Remes, P. M.; Jacob, C. C.; Heil, L. R.; Shulman, N.; MacLean, B. X.; MacCoss, M. J. Hybrid Quadrupole Mass Filter–Radial Ejection Linear Ion Trap and Intelligent Data Acquisition Enable Highly Multiplex Targeted Proteomics. J. Proteome Res. 2024, 23 (12), 5476–5486. https://doi.org/10.1021/acs.jproteome.4c00599.

(59) Wu, C. C.; Tsantilas, K. A.; Park, J.; Plubell, D.; Sanders, J. A.; Naicker, P.; Govender, I.; Buthelezi, S.; Stoychev, S.; Jordaan, J.; Merrihew, G.; Huang, E.; Parker, E. D.; Riffle, M.; Hoofnagle, A. N.; Noble, W. S.; Poston, K. L.; Montine, T. J.; MacCoss, M. J. Enrichment of Extracellular Vesicles Using Mag-Net for the Analysis of the Plasma Proteome. Nat. Commun. 2025, 16 (1), 5447. https://doi.org/10.1038/s41467-025-60595-7.

(60) Plubell, D. L.; Remes, P. M.; Wu, C. C.; Jacob, C. C.; Merrihew, G. E.; Hsu, C.; Shulman, N.; MacLean, B. X.; Heil, L.; Poston, K. L.; Montine, T. J.; MacCoss, M. J. Development of Highly Multiplex Targeted Proteomics Assays in Biofluids Using a Nominal Mass Ion Trap Mass Spectrometer. Mol. Cell. Proteomics 2026, 25 (2), 101506. https://doi.org/10.1016/j.mcpro.2026.101506.

(61) Serrano, L. R.; Peters-Clarke, T. M.; Arrey, T. N.; Damoc, N. E.; Robinson, M. L.; Lancaster, N. M.; Shishkova, E.; Moss, C.; Pashkova, A.; Sinitcyn, P.; Brademan, D. R.; Quarmby, S. T.; Peterson, A. C.; Zeller, M.; Hermanson, D.; Stewart, H.; Hock, C.; Makarov, A.; Zabrouskov, V.; Coon, J. J. The One Hour Human Proteome. Mol. Cell. Proteomics 2024, 0 (0). https://doi.org/10.1016/j.mcpro.2024.100760.

(62) Lancaster, N. M.; Sinitcyn, P.; Forny, P.; Peters-Clarke, T. M.; Fecher, C.; Smith, A. J.; Shishkova, E.; Arrey, T. N.; Pashkova, A.; Robinson, M. L.; Arp, N.; Fan, J.; Hansen, J.; Galmozzi, A.; Serrano, L. R.; Rojas, J.; Gasch, A. P.; Westphall, M. S.; Stewart, H.; Hock, C.; Damoc, E.; Pagliarini, D. J.; Zabrouskov, V.; Coon, J. J. Fast and Deep Phosphoproteome Analysis with the Orbitrap Astral Mass Spectrometer. Nat. Commun. 2024, 15 (1), 7016. https://doi.org/10.1038/s41467-024-51274-0.

(63) Messner, C. B.; Demichev, V.; Wendisch, D.; Michalick, L.; White, M.; Freiwald, A.; Textoris-Taube, K.; Vernardis, S. I.; Egger, A.-S.; Kreidl, M.; Ludwig, D.; Kilian, C.; Agostini, F.; Zelezniak, A.; Thibeault, C.; Pfeiffer, M.; Hippenstiel, S.; Hocke, A.; von Kalle, C.; Campbell, A.; Hayward, C.; Porteous, D. J.; Marioni, R. E.; Langenberg, C.; Lilley, K. S.; Kuebler, W. M.; Mülleder, M.; Drosten, C.; Suttorp, N.; Witzenrath, M.; Kurth, F.; Sander, L. E.; Ralser, M. Ultra-High-Throughput Clinical Proteomics Reveals Classifiers of COVID-19 Infection. Cell Syst. 2020, 11 (1), 11-24.e4. https://doi.org/10.1016/j.cels.2020.05.012.

(64) Szabadváry, F.; Belcher, R.; Gordon, L. History of Analytical Chemistry: International Series of Monographs in Analytical Chemistry; Elsevier Science: Burlington, 1966.

(65) Fresenius, H. Zur Erinnerung an R. Fresenius. Z. Für Anal. Chem. 1897, 36 (1), III–XVIII. https://doi.org/10.1007/BF01348148.

Appendix 1: Image Descriptions

Figure 2.1. Overview of Orbitrap Astral MS and its key figures of merit. A multi-panel figure is shown. Panel A shows an annotated schematic of the Orbitrap Astral mass spectrometer highlighting different key components in including features of the Astral analyzer like the ion processor, injection optics, ion foil, and asymmetric ion mirrors. Panel B shows an annotated phosphopeptide fragmentation spectrum, demonstrating high sequence coverage for the peptide and 73.7% of the TIC annotated for a spectrum containing 5.7e4 charges. Panel C compares the resolution vs m/z for the Astral and theoretical Orbitrap resolution. The Orbitrap resolution drops with increasing m/z, whereas the Astral resolution is relatively stable with m/z with the 5-95% confidence interval ranging from about 60k to 150k and reaching almost 300k at the higher m/z values. Panel D shows the distribution of mass error centered tightly around zero and 95% of the distribution within 5ppm error. See figure caption for detailed description.

Figure 2.2 Overview of DIA phosphoproteomics on the Orbitrap Astral MS. A multi-panel figure is shown. Panel A is a schematic representation of parallelized MS1 and rapid MS2 scans. Large blue boxes represent the MS1 scan range and time, and small red boxes represent the MS2 scan ranges and times. The MS1 scans are represented as acquiring every 0.6 s for a total acquisition time over a broad m/z range. The MS2 scans are represented as iterating over the same m/z range in 2 Thomson intervals with an interscan time of 5 milliseconds. A zoom-in shows the MS2 scan representation more clearly. Panel B shows a bar plot indicating that an isolation width of 2 Thomson produces more localized phosphosites than an isolation width of 4 Th. Both bars show tight error bars. Panel C shows a line plot indicating increasing relative depth with increasing loading mass with saturation reached at around 250 nanograms. The loading masses range from 50 to 1000 nanograms. Panel D shows a line plot indicating slight increases in Relative Depth (%) with increasing gradient length with three data points with tight error bars and ranging from 7 -30-minute gradient lengths. Panel E shows stacked bar plots indicating data completeness of phosphosites across triplicate analyses at three different gradient lengths. The data completeness increases with gradient length. Different colors are used to indicate the number of phosphosites meeting a particular Data Completeness Across Triplicates with labels indicating the percentage of each Data Completeness for each gradient length. Panel F shows violin plots of %RSD values showing that the %RSD median value decreases with increasing gradient length. Most of the distributions for all gradient lengths are below the 20% line. Panel G compares internal and external FDR measurements for all and phosphorylated precursors by showing line plots of the number of precursors as a function of estimated precursor FDR. The internal and external FDR lines converge at higher FDRs. The FDR curves for all precursors and phospho precursors show similar trends. Panel H shows cumulative distributions of localization probabilities for phosphorylation on S,T,Y, and P residues. The S residue curve is shifted towards the highest

probabilities followed by the T and Y curves. The P distribution is shifted toward lower localization probabilities. Panel I plots the average error rate versus localization probability cutoff, which decreases with increases cutoff. Panel J shows a histogram of Pearson R squared values with most values over 0.95. See figure caption for detailed description.

Figure 2.3 Chromatographic Properties of HEK293T Phosphoproteomics Datasets. A multi-panel figure is shown. Panel A shows representative chromatograms for the 7, 15 and 30-minute active gradient methods with blue curves showing the %B gradient settings. Panels B-D show violin plots distributions for chromatographic baseline widths, MS1 points across baseline, and MS2 points baseline at the three gradient lengths. Each of these distributions shift to higher values at the longer gradient times. Panel E shows a scatter plot of MS1 and MS2 cycle times across retention time. The MS1 cycle times are around 0.6 seconds and the MS2 cycle times are around 1.5 seconds. See figure caption for detailed description.

Figure 2.4. Evaluation of Phosphoproteomic Method and Processing Parameters. A multi-panel graph is shown. Panel A is a grouped bar chart comparing localized phosphosites across three gradient lengths for Spectronaut 17 and PD 3.. The number of sites increases with gradient length and CHIMERYS identifies substantially more sites. For each gradient length, the left bar is the Spectronaut results and the right bar is for PD 3.1. Panel B shows a stacked bar chart of total phosphopeptides for isolation widths 2 and 4, colored by peptide multiplicity. The multiplicity distributions are similar for each with 2 Thomson generating more phosphopeptide detections. For both bars, the majority of sites have a multiplicity of 1 with just 5-10% of the bar visually indicating multiplicity 2. It is not possible to visually distinguish multiplicity 3 and 4 values from this figure. Panel C is a violin plot showing the site localization probability distribution for isolation widths 2 and 4, with slightly higher values for the 2 Thomson width. A dashed line indicated that most sites are above the 0.75 threshold. Panels D and E are bar charts of localized phosphosites against DIA m/z range (380-980, 480-1080) and AGC Target (1e4, 5e4), respectively. The 380-980 m/z range slightly outperforms 480-1080. The AGC target values do not have a noticeable effect. Panel F is a line histogram displaying the frequency of MS2 Total Ion Current for two AGC targets, with a logarithmic x-axis. Both curves are basically identical. Panel G is a line histogram showing the frequency of MS2 Injection Time for two AGC targets with nearly identical distributions but the 5e4 AGC target has a slightly higher peak at 3.5 milliseconds. The distributions for Panels F and G across AGC target are nearly identical. Panel H contains four separate bar charts showing localized phosphosites versus localization probability for Tyrosine, Threonine, Serine, and Proline, with y-axes plotted on a logarithmic scale. The proline distribution is shifted toward lower localization probabilities. See figure caption for detailed description.

Figure 2.5. Validation of Site Localization and Quantification using Synthetic Phosphopeptide Standards. A multi-panel figure is shown. Panel A shows a bar plot of error rates as function of localization probability cutoff which decrease with increasing cutoff. The reported error rates for the Bekker-Jensen et al study are higher than a re-analysis of their data, and the error rates for the Orbitrap Astral are lower panel. At each probability cutoff, the leftmost bar is for the 'Reported' group, the middle bar is for the 'Re-Analysis' group, and the rightmost bar is for the 'Orbitrap Astral' group. Panel B shows a bar plot with only a slight drop in average correct precursors with increasing localization probability cutoff. Panel C shows a stacked bar plot showing a high degree of data completeness for the standards detected across triplicates at different columns loads. The completeness gets worse at the lower column loads with most of the sites demonstrating a completeness of 3 at the 10000 attomole load and this decreasing down to about 50% of the sites demonstrating a completeness of 3 at the 39 attomole column load. Panel D shows calibration curves for three of the synthetic phosphosites, showing R squared values >0.96 for each. Panel E shows the overlap of the yeast phosphopeptide background intensity distribution and the synthetic phosphopeptide intensity at different column loads for the synthetic standards, showing that the range of column loads covers the full range of intensity values observed in the yeast background. For this panel, the yeast distribution is indicated in black and the synthetic standard distribution is shown in blue. The black yeast distribution is consistent across all of the plots and is a smoother curve than the standard distribution. See figure caption for detailed description.

Figure 2.6. Biological validation of phosphoproteomics platforms using EGF stimulation. A multi-panel figure is shown comparing the results for analyses of EGF-treated and control HeLa cells across the Astral, Ascend, and timsTOF platforms. Panel A shows stacked bar plots of the number of phosphosites detected for each platform with the Astral detecting much more sites than the others and the Ascend having the highest proportion of sites detected across triplicates. Colors indicate the data completeness with % values annotated for the 3 of 3 completeness value. Panel B and C show overlapping intensity distributions for sites unique to Astral and overlapping with the timsTOF and Ascend datasets, showing that overlapping sites are biased to higher intensities. The large distributions in blue correspond to the Astral unique sites. Panel D shows intensity distributions for unique timsTOF sites and overlapping with Astral, showing the overlapping sites are fairly well distributed across intensities. The taller blue distribution corresponds to the timsTOF unique sites. Panel E shows a Venn diagram showing strong overlap of detected sites across platform with the Astral dataset having the most unique sites. Panel F shows three volcano plots for each platform looking at site differences between conditions. Similar results are obtained for the annotated EGFR sites. Datapoints with fold change >2 and p-value are considered significant and are colored blue. The datapoints that are labeled are colored solid blue if they are considered

significant and are colored with just a black outline if they are not significant. Panel G shows a pathway enrichment analysis (odd ratio vs -log10(p-value)) highlighting that each platform's dataset showed enrichment for both the EGF/EGFR signaling pathway and ERK activation. The colored blue dots indicate EGF/ERBB terms. See figure caption for detailed description.

Figure 2.7. Mouse Phosphorylation Atlas Workflow and Results. A multi panel figure is shown. Panel A shows graphics illustrating the tissues analyzed and the sample prep and data acquisition workflow. Panel B shows stacked bar plots of the number of phosphorylation sites detected in each tissue and overall for the Huttlin et al. study and this study. The proportion of S,T,Y sites are illustrated with different colors. The vast majority of sites occur on S, with a smaller proportion occurring on T, and the smallest proportion on Y. For the 'All' bars, these proportions are labeled. This study identified much more sites overall. Panel C shows histograms for the number of tissues a phosphosite was observed in with the Huttlin et al. paper showing most sites detected in 1 tissue. This study shows the distribution shifted toward more tissues with a notable bump for phosphosites detected in all 12 tissues. Panels D and E show intensity distributions as function of number of tissues observed in, with the intensity distributions shifting to higher intensity values as the number of tissues increased for this study and the Huttlin et al. study. respectively. See figure caption for detailed description.

Figure 2.8. Proteomic Profiling of Mouse Tissues. A multi-panel figure is shown. Panel A shows two stacked bar plots for the Huttlin et al. study and this study. The bar plots show that this study detected much more gene products and that almost all the Huttlin et al. study gene products overlap with this study and about half of gene products detected in this study for overlapping tissue being common between studies. Panels B-D show PCA analysis results of the two studies. Panel B shows clear separation between samples in this study and the Huttlin et al study along the Component 1 axis. Panel C shows the loading plot for Component 1 and 2 with HBB 1 and HBB 2 noted as outliers. Panel D shows that the tissues for each study group together closely when plotting Component 2 and 3. See figure caption for detailed description.

Figure 2.9. Mouse Phosphorylation Atlas Comparison to Huttlin et al. A multi-panel figure is shown. Panel A shows two stacked bar plots for the Huttlin et al. study and this study. The bar plots show that this study detected much more phosphosites and that almost half of the Huttlin et al. study phosphosites overlap with this study with small amounts of sites in this study overlapping with the Huttlin et al. study. Panel B shows a heatmap showing relative site intensities across different sites. Labels below the graph indicate regions of the heatmap where a particular tissue is divergent from the rest. Panel C shows the distribution of intensities for phosphorylated S, T, and

Y. The S distribution is shifted toward the highest values, followed by the T distribution. The Y distribution is shifted toward lower values.

Figure 2.10. Mouse Phosphorylation Atlas Comparison to Giansanti et al. A multi-panel figure is shown. Panel A shows two stacked bar plots for the Giansanti et al. study and this study. The bar plots show that this study detected about 2-fold more phosphosites and that a little more than half of the Giansanti et al. study phosphosites overlap with this study with one-fourth to one-third of sites in this study overlapping with the Giansanti et al. study. Panel B shows the distributions of how many tissues a site was observed in for both studies with similar shapes observed, but this study having a bit more sites detected in just 1 tissue. Panel C shows a stacked bar plot showing that this study detected about 2-fold more sites with a higher percentage of phosphorylated Y residues detected in this study. The majority of the sites are on S, with a smaller portion on T, and the smallest portion on Y. See figure caption for detailed description.

Figure 2.11. Sequence and Structural Context of Phosphosites. A multi-panel figure is shown. Panel A shows a t-SNE representation of phosphosites and flanking sequences with clusters indicated. Density is represented by color and clusters indicate regions of high density. Panel B shows site motifs for each cluster. Panel C shows the same t-SNE plot with ERK1/2/5 and CDK1/2/4/5 sites colored differently with a main cluster circled. Panel D shows pY sites in red with minimal overlap with the circled region. There is a small amount of overlap of the pY sites with the indicated sites in panel C. Panel E shows distribution of pLDDT scores for phosphorylated and other residues. Bimodal distributions are observed with phosphorylated residues showing enrichment in lower pLDDT scores. Panel F shows the same plot in Panel E as a cumulative distribution plot. Panel G plots the pLDDT score as a function of distance between alpha carbons for un-phosphorylated and phosphorylated STY. There are repeating patterns of dips in the pLDDT score, which is more notable in the phosphorylated curve. The unphosphorylated curve consistently shows higher pLDDT scores. See figure caption for detailed description.

Figure 2.12. Sequence, Structural and Kinase Specificity Context of Phosphosites. A multi-panel figure is shown. Panel A shows a t-SNE representation of phosphosites and flanking sequences with clusters indicated. Density is represented by color and clusters indicate regions of high density. Panel B shows site motifs for each cluster. Panel C shows distribution of pLDDT scores for phosphorylated and other residues. Bimodal distributions are observed with phosphorylated residues showing enrichment in lower pLDDT scores. Panel D shows a schematic representation of the kinase prediction workflow. Panel E shows a plot showing the number of kinases matched to phosphosites. Many phosphosites are just assigned to a few kinases but there are some that are assigned to many kinases. The site percentile shows high percentiles in blue and

lower in gray shades. Most of the high percentile plots are at the bottom left of the plot. Panel F shows a plot of the site abundance differences between brain and liver. Panel G shows a volcano plot of kinase enrichment between brain and liver with significant kinases annotated on the plot. Brain enriched datapoints are on the right of the figure and liver enriched datapoints are on the left. See figure caption for detailed description.

Figure 2.13. Kinase Enrichment Across Tissues. A large figure showing pairwise comparisons of kinase enrichment across all tissues. Volcano plots comparing the Log2(frequency factor) to the -log10(p-value) are shown for predicted kinases. Each plot is annotated with indicating kinases that show large differences in kinase enrichment between tissues. Blue dots on the left side of each volcano plot indicate enriched datapoints for the tissue labeled on the top side of the overall figure and red dots on the right side of each volcano plot indicate enriched datapoints for the tissue labeled on the right side of the overall figure. See figure caption for detailed description.

Figure 2.14. Mitochondrial phosphoproteomics reveals novel liver-specific phosphorylation site. A multi-panel figure is shown. Panel A shows a bar plot of phosphosites and their Phosphosite Plus Categories: Characterized, Known, and Novel. A small percentage of sites are 'Characterized' and over 50% of sites are 'Novel'. The bar on the left is ‘Characterized sites’, the middle bar is ‘Known sites’, and the bar on the right is ‘Novel sites’. Panel B shows a bar plot showing that ~500 proteins in the MitoCarta 3.0 database are not phosphorylated and >600 proteins are phosphorylated. The left bar is ‘Not phosphorylated’ and the right bar is ‘Phosphorylated’. Panel C shows the distribution of number of tissues a site was detected in for mitochondrial phospho sites. The bars are colored to indicate the number of novel sites. The majority of sites are detected in one tissue and the number of novel sites (indicated by the color of the bar) decreases with the number of tissues in which a phospho site is present in. Panel D shows a graphic of the mitochondria and an upset plot of mitochondrial sites across tissues with labels for different mitochondrial location classifications. The liver has the most unique mitochondrial phosphosites. Panel E shows a scatter plot of protein rank versus number of detected phospho sites with color indicating the number of novel sites. Cps1 is strong outlier with the highest rank and the greatest number of sites and the many novel sites. In general, the number of novel sites (indicated by datapoint color) increases from the left to the right of the graph. Panel F shows a representation of the Cps1 domains with previously identified and novel phosphosites indicated along the domain map. Panel G shows a sequence alignment of CPS1 across species. Panel H shows proteins structures for the Apo and NAG bound forms with the R1453 and Y1450 residues indicated. See figure caption for detailed description.

Figure 2.15. TCA cycle phosphorylation events and occurrence of phosphorylation sites across tissues. A multi-panel figure is shown. Panel A shows a GSEA analysis indicating enriched KEGG pathways. The citric acid cycle has the greatest fold enrichment. The -log10(FDR) is shown with coloring. There is a slight trend toward higher -log10(FDR) values going from the top to the bottom of the plot. Panel B shows a bar plot of Phosphosite Plus categories (Novel, Characterized, and Known sites) for TCA cycle proteins, showing many novel sites on these proteins. For most of these proteins, the novel sites are the majority of the detected sites. The novel site coloring is at the top of the bars, the characterized site coloring is in the middle (only observable for the Cs protein) and the known site coloring is on the bottom. Panel C shows the distribution of number of tissues in which phosphorylation sites were detected in. See figure caption for detailed description.

Figure 2.16. Unique phosphorylation sites in brain tissue. A multi-panel figure is shown. Panel A shows a stacked bar plot of phosphosites across tissues that are tissue-specific vs not tissue-specific. The brain has the highest number of tissue-specific phosphosites. Panel B shows the number of outliers per tissue as a bar plot with color indicating the number of novel sites. In general, the proportion of novel sites decreases from left to right. The brain has many novel sites and has the most outliers per tissue. Panel C shows a scatter plot showing the z-score for phosphosites. Opa1_S298 is indicated as a phosphosite on the higher end of z-scores. Panel D shows a domain representation of OPA1 with phosphosites indicated on it. Panel E shows a sequence alignment of OPA1 across species. Panel F shows the structure of OPA1 with the S298 site labeled in a zoom-in. See figure caption for detailed description.

Figure 2.17. Z-score based outlier analysis and phosphorylation sites on Opa1. A multi-panel figure is shown. Panel A shows a heatmap of phosphosite z-scores across tissues. The z-score values are represented with a color gradient. Panel B shows distributions of z-score for each tissue. The brain distribution is bimodal. Panel C shows a domain representation of OPA1 with novel phosphosites. For each phosphosite, an overlap plot indicates which tissue it was detected in. See figure caption for detailed description.

Figure 3.1. Processing SynchroSep-MS using Prototype DIA-NN Software. A flowchart describing how data processing is performed. The workflow involves collecting a single column injection of a representative sample on Column 1 of the dual column setup, searching this data with a DIA-NN predicted spectral library, converting the library file to a .parquet format, aligning retention times in the predicted library to experimental retention times in the report.parquet output using an R script. Then the aligned spectral library is converted to a .skyline.speclib format by loading the .parquet file into DIA-NN and running without any raw data included. Then this library

is used to search dual injection files by supplying a RT window width parameter (--rt-gap) controlling the RT window width for extraction of ion signals and a retention time offset parameter (--rt-shift) that specifies a retention time offset to use when searching the files. Multiple --rt-shift commands can be used to search for multiple SynchroSep-MS channels contained within a single raw file. See figure caption for detailed description.

A multi-panel figure illustrating the SynchroSep-MS concept. The first two panels show images of the two capillary columns in the same column heater and emitters aligned to the MS inlet at the same time. Panel C shows a schematic showing how each column is connected to a unique LC and aligned to a single mass spec. Panel D shows a conceptual representation of how an injection is performed and a separation starts on Column 1 and then the injection onto Column 2 is performed such that the separation windows overlap between columns. Panel E shows an image where the peak for Column 2 has a later retention time than Column 2. Panel F shows line plots illustrating how there is a retention time offset between the two columns. See figure caption for detailed description.

Figure 3.3. Assignment of Peaks to the Originating Column. A multi-panel figure is shown. Panels A-C show example chromatograms for a Column 1 single injection, Column 2 single injection, and dual column injection, respectively. Panel D plots the retention time offset as function of retention time which is a non-linear relationship but is typically about 4 minutes. The agreement between experimental data points and the model is shown. Panel E shows the error in the retention time offset model as a function of retention time. In general, the model accuracy is within about 0.3 minutes. Panel F and G show scatter plots of dual vs single-injection retention times for detected precursors on Column 1 and Column 2, respectively. Excellent agreement is observed. Panel I shows a bar plot showing the number of precursors assigned to Column 1 or 2 detected with a Column 1 injection, Column 2 injection, or combined injection. For each of these conditions, the Precursors with a Column 1 Predicted RT are indicated with the left bar, and Precursors with a Column 2 Predicted RT are indicated with the right bar. For the single column injections, less than 0.5% of detected precursors were assigned to the other column. For the dual column injection, Column 1 retains almost all detections as in the single injection and Column 2 retains fewer but still a large majority of precursor detections. See figure caption for detailed description.

Figure 3.4. Tuning the Retention Time Offset. A scatter plot showing the retention time offset across retention time for methods with different added delay times, showing that increased added delay leads to increased retention time offsets. The 0 min Added Delay data points are the lowest

curve, the 1 min Added Delay data points are the middle curve and the 2 min Added Delay data points are the top curve. See figure caption for detailed description.

Figure 3.5. RT Window Width Scouting for DIA-NN Processing. A multi-panel figure is shown. Panels A and B plot the column assignment error rates (square datapoints) and number of precursors detected (circular datapoints) as a function of RT window width for Column 1 RT search and Column 2 RT search, respectively. The number of precursors detected stays fairly consistent for both plots. The error rate rises quickly above 2.5 minutes window width, reaching >30% for Panel A and >1.4% for Panel B. Panel C shows a bar plot showing the number of precursors assigned to Column 1 or 2 detected with a Column 1 injection, Column 2 injection, or combined injection. For each of these conditions, the Precursors with a Column 1 Predicted RT are indicated with the left bar, and Precursors with a Column 2 Predicted Rt are indicated with the right bar. For the Column 1 only injection, about 1.5% of detected precursors were assigned to Column 2. For Column 2 only injection, about 31% of precursors were assigned to Column 1. For the dual column injection, Column 1 retains the almost all detections as in the single injection and Column 2 retains fewer but still a large majority of precursor detections. Panel D and E show scatter plots of dual vs single-injection retention times for detected precursors on Column 1 and Column 2, respectively. Excellent agreement is observed for Column 2, but Column 1 reveals a large amount of different retention times detected between the dual injection and single injection. See figure caption for detailed description.

Figure 3.6. External FDR Assessment via Entrapment Analysis. A bar chart showing the False Discovery Proportions across triplicate analyses for the Single Injection Column 1, Dual Injection Column 2, Single Injection Column 2 and Dual Injection Column 2. Each value is around 0.4-0.5%. See figure caption for detailed description.

Figure 3.7. Protein Group Identifications. A multi-panel figure is shown. Panel A shows the number of Protein Groups for Column 1 and Column 2 and for the Single and Dual Injection. Single Injections are shown as the left bar, and Dual Injections are shown as the right bar for the Column 1 and Column 2 data. Protein groups are similar between the Single and Dual injections, with a slight decrease for the Column 1 Dual injection, and a slightly larger decrease for the Column 2 Single Injection. Panel B is a line plot showing the cumulative protein group detections as a function of retention time for the Column 1 Single Injection, the Column 2 Single Injection and the Dual Column injection. The Column 2 curve is shifted to the right of the Column 1 curve, and the Dual Injection curve appears as the sum of the two curves. Panel C shows violin plots for %RSD values for protein groups for the Single Injection Column 1, Dual Injection Column 1,

Single Injection Column 2, and Dual Injection Column 2. The median is about 3.5% across all the datasets. See figure caption for detailed description.

Figure 3.8. Precursor Identifications. A multi-panel figure is shown. Panel A shows the number of precursors for Column 1 and Column 2 and for the Single and Dual Injection. Single Injections are shown as the left bar, and Dual Injections are shown as the right bar for the Column 1 and Column 2 data. Precursors are similar between the Single and Dual injections, with a slight decrease for the Column 1 Dual injection, and a slightly larger decrease for the Column 2 Single Injection. Panel B is a line plot showing the cumulative precursor detections as a function of retention time for the Column 1 Single Injection, the Column 2 Single Injection and the Dual Column injection. The Column 2 curve is shifted to the right of the Column 1 curve and the Dual Injection curve appears as the sum of the two curves. Panel C shows violin plots for %RSD values for precursors for the Single Injection Column 1, Dual Injection Column 1, Single Injection Column 2, and Dual Injection Column 2. The median is about 9.5% across all the datasets. See figure caption for detailed description.

Figure 3.9. Dual vs. Single-Injection Dynamic Range. A two panel figure showing scatter plots of Protein Group MaxLFQ quantity as a function of protein rank. The data points are colored based on whether they were detected in both the dual and single injection method or just the single injection. The data is similar for Column 1 and Column 2, showing that most of the protein groups unique to the single column injection exhibit much lower MaxLFQ quantities. See figure caption for detailed description.

Figure 3.10. Protein Group RSD Zoom In. A violin plot showing the distribution of %RSD values for Single Injection Column 1, Dual Injection Column 1, Single Injection Column 2, and Dual Injection Column 2. All distributions appear similar. See figure caption for detailed description.

Figure 3.11. RT Binning Analysis of Dual Column SynchroSep-MS Data. A two panel figure is shown. The top panel shows histograms of the fraction of mission precursors at different RT timepoints. At each RT Bin center, the Column 1 bar is on the left and the Column 2 bar is on the right. Column 1 has lower fraction of missing precursors at the beginning and Column 2 has lower fraction of missing precursors at the end. The bottom panel shows the distribution of detected precursors for Column 1 and Column 2 with respect to retention time. The Column 2 distribution is shifted to later retention times. See figure caption for detailed description.

Figure 3.12. Performance of SynchroSep-MS Method Assessed with Dilution Series. A multi-panel figure is shown. Panels A and B show the number of protein groups detected across the direct and inverted dilution series. For each loading mass point, the Column 1 data is shown as the left bar, and the Column 2 data is shown as the right bar. For both columns, the number of protein groups increases as the loading mass increases for that column. Panel C shows violin plots for R squared distribution for the Column 1 Direct, Column 2 Direct, Column 1 Inverted, and Column 2 Inverted datasets with median values >0.96 across the datasets. Panels D and E show the scatter plots for the quantities measured across columns or dilution series, respectively. Density of points is represented with a color gradient with the highest density along the linear fit line. Both plots have R squared values >0.97 and slopes close to 1, indicating strong agreement. See figure caption for detailed description.

Figure 3.13. Performance of SynchroSep-MS Method Assessed with Dilution Series – Precursor Level. A multi-panel figure is shown. Panels A and B show the number of precursors detected across the direct and inverted dilution series. For each loading mass point, the Column 1 data is shown as the left bar, and the Column 2 data is shown as the right bar. For both columns, the number of precursors increases as the loading mass increases for that column. Panel C shows violin plots for R squared distribution for the Column 1 Direct, Column 2 Direct, Column 1 Inverted, and Column 2 Inverted datasets with median values >0.96 across the datasets. Panels D and E show the scatter plots for the quantities measured across columns or dilution series, respectively. Density of points is represented with a color gradient with the highest density along the linear fit line. Both plots have R squared values >0.91 and slopes close to 1, indicating reasonable agreement. See figure caption for detailed description.

Figure 3.14. Column Intensity Comparisons for the Inverted Dilution Series. A two panel plot is shown. Scatter plots for the protein group and precursor quantities measured across columns for the inverted dilution series are shown, respectively. Density of points is represented with a color gradient with the highest density along the linear fit line. Both plots have R squared values >0.91 and slopes close to 1, indicating reasonable agreement. Note that agreement is higher at the PG-level than the precursor-level as would be expected. See figure caption for detailed description.

Figure 4.1. Emitter Position Images Recorded with Source Camera. A two panel figure with top and side view images of the emitter aligned to the mass spec inlet. A measurement indicates a distance of 0.856 mm between the emitter and the inlet in Panel A. Panel B shows a side view of the emitter. An angle is observed between the emitter axis and the inlet axis. See figure caption for more details.

Figure 4.2. Signal Stability of Infused Peptides. A multi-panel figure showing the intensity profiles for two peptides over time and a rolling %RSD over time for a zoomed in time range. Signal is stable for periods of >60 minutes with %RSD values <10%. See figure caption for detailed description.

Figure 4.3. QC Measurements for Emitter Positioning Experiments. A multi-panel figure shows line plots illustrating the signal intensity of two peptides across QC measurements. The panels represent measurements for the 1D experiments, 2D at z = -0.9, 2D at z = -2.9, and 2D at z = -4.9 mm experiments. %RSD values for the first peptide are 2.0, 5.7, 1.8, and 1.6, respectively. %RSD values for the second peptide are 3.0, 2.9, 3.4, and 3.2, respectively. Within each panel, the QC measurements are quite consistent. The +2 peptide intensity values are always the higher ones in the plots. See figure caption for detailed description.

Figure 4.4. Geometric Diagram of Nano-Electrospray Emitter Configuration. A geometric diagram showing the geometry of the emitter with respect to the inlet. The axis of the emitter is at an angle of theta with respect to the axis of the inlet, leading to a displacement in the y dimension that depends on the distance between the emitter and the inlet in the z dimension.

Figure 4.5. Dependence of y Intensity Distribution on z Position. A two panel figure shows the same plot for the two peptides analyzed in this study. Each panel is a line plot showing the relative intensity distribution in the y dimension at three different z positions. It is observed that the intensity distributions broaden with increasing z distance. The centers of the distributions are relatively consistent at increasing z distances. See figure caption for detailed description.

Figure 4.6. MS1 Spectra of Infused BSA Tryptic Digest. An MS1 spectra of infused BSA tryptic digest showing many peaks across different m/z values. See figure caption for detailed description.

Figure 4.7. 1D Positioning Experiments. A multi-panel figure is shown. The first two panels illustrate the geometry of the inlet and source with a 17.4% angle between the axis of the emitter and the inlet axis. Panels C-E show the intensity distributions for two peptides in the x,y, and z dimensions. See figure caption for detailed description.

Figure 4.8. Inlet Capillary Drawing. A mechanical drawing of the high-capacity transfer tube showing a stadium shape for the inlet with dimensions of 1.6 and 0.6 millimeters. See figure caption for detailed description.

Figure 4.9. Estimating Distribution Centroids. A multi-panel figure shows the intensity distributions for two peptides in the x- and y-dimensions. The centroids of the distributions are shown. For the x-dimension, both centroids are at -0.1 millimeters. For the y-dimension, both centroids are at 0.3 millimeters. See figure caption for detailed description.

Figure 4.10. Dependence of Precursor Intensity Ratio on z Position. A line plot showing the ratio of two peptide's signal intensity as a function of z position. The line plot is relatively stable with z-position at a value of around 1.2. See figure caption for detailed description.

Figure 4.11. Dependence of Signal Intensity on Emitter Position for Low Level Precursor Ions. A multi-panel figure shows the line plots for intensity profiles for the x-, y-, and z-dimensions. The profiles for two peptides are shown in blue and red overlaid on many black traces representing lower intensity profiles. The y-axis of each plot is depicts normalized intensity. See figure caption for detailed description.

Figure 4.12. 2D Positioning Experiments. A multi-panel figure shows heatmaps of signal intensity at different x and y positions. Each panel shows the results at a different z-position. The distributions are centered close to x = 0 and y = 0 and the width of the distributions broaden and flatten at further z distances. See figure caption for detailed description.

Figure 4.13. 2D Positioning Experiments for RHPEYAVSVLLR. A multi-panel figure shows heatmaps of signal intensity at different x and y positions. Each panel shows the results at a different z-position. The distributions are centered close to x = 0 and y = 0 and the width of the distributions broaden and flatten at further z distances. See figure caption for detailed description.

Figure 4.14. Width at 90% Maximum as Function of z Position. A two panel figure showing the Width of the Intensity Distributions at 90% maximum with increasing distance in the z distance for two peptides and for the x and y dimension. In general, the width increases as the distance between the inlet and emitter increases in the z dimension. See figure caption for detailed description.

Figure 4.15. Dependence of Emitter Position Results on *m/z*. A multi-panel figure shows scatter plots of x Dimension FWHM, y Dimension FWHM and Fold change between z = -6.9 and z = -0.9 versus m/z. There is not a strong dependence on m/z. The Pearson correlations are annotated as -0.1138, -0.0183, and 0.1953 for each panel, respectively. See figure caption for detailed description.

Figure 4.16. Dependence of Emitter Position Results on Peak Intensity. A multi-panel figure shows scatter plots of x Dimension FWHM, y Dimension FWHM and Fold change between z = -6.9 and z = -0.9 versus the Log10 Peak Intensity. In general, there is a greater spread of values at lower intensity values, and the values are relatively consistent at higher intensities. See figure caption for detailed description.