One-Hour Phosphoproteome Analysis with the Orbitrap Astral Mass Spectrometer

Noah M. Lancaster[1,2,14], Pavel Sinitcyn[3,14], Patrick Forny[4,14], Trenton M. Peters-Clarke[1,2], Caroline Fecher[4], Andrew J. Smith[4], Evgenia Shishkova[2,5], Tabiwang N. Arrey[6], Anna Pashkova[6], Margaret Lea Robinson[2], Nicholas Arp[3,7], Jing Fan[3,7,8], Juli Hansen[9], Andrea Galmozzi[2,9,10], Lia R. Serrano[1,2], Michael S. Westphall[2,5], Hamish Stewart[6], Christian Hock[6], Eugen Damoc[6], David J. Pagliarini[4,11,12], Vlad Zabrouskov[13], and Joshua J. Coon[1,2,3,5*]

[1]Department of Chemistry, University of Wisconsin-Madison, Madison, WI

[2]Department of Biomolecular Chemistry, University of Wisconsin – Madison, Madison, WI

[3]Morgridge Institute for Research, Madison, WI

[4]Department of Cell Biology and Physiology, Washington University School of Medicine, St. Louis, MO 63110, USA.

[5]National Center for Quantitative Biology of Complex Systems, Madison, WI

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

[7]Cellular and Molecular Biology Graduate Program, University of Wisconsin-Madison, Madison, WI

[8]Department of Nutritional Sciences, University of Wisconsin-Madison, Madison, WI

[9]Department of Medicine, University of Wisconsin-Madison, School of Medicine and Public Health, Madison, WI, United States.

[10]University of Wisconsin Carbone Cancer Center, University of Wisconsin-Madison, School of Medicine and Public Health, Madison, WI, United States.

[11]Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110, USA

[12]Department of Genetics, Washington University School of Medicine, St. Louis, MO 63110, USA.

[13]Thermo Fisher Scientific, San Jose, CA

[14]These authors contributed equally: Noah M. Lancaster, Pavel Sinitcyn, Patrick Forny

*address correspondence to coon@wisc.edu

ABSTRACT

As a key mechanism for cellular signal transduction, protein phosphorylation plays critical roles in myriad cellular processes. This modification, however, is highly dynamic and occurs at sub-stoichiometric levels. Mass spectrometry is an essential tool for studying this modification on a global scale; however, the technology's impact suffers from four main limitations: need for site localization, dynamic range, reproducibility, and throughput. Here we describe the use of a novel mass spectrometer (Orbitrap Astral) coupled with data-independent acquisition (DIA) to achieve detection of close to 40,000 unique phosphorylation sites within one hour of analysis. We applied this approach to generate a phosphoproteome atlas of the mouse. Altogether, we detected 81,120 unique phosphorylation sites within 12 hours of measurement. With this unique dataset, we examine the sequence and structural 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.[3–14] 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.[15–29] Still, current phosphoproteomic analyses are neither routine nor straightforward when compared to conventional shotgun proteomics (i.e., 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.[16,22,30,31] Additionally, sensitivity and dynamic range of the mass analyzer can also be important for detection of low level, but critical, product ions.[12,25,29] Next, the often sub-stoichiometric amounts of phosphorylated protein itself elevate the difficultly.[2,14,18–20] 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.[32] This tremendous dynamic range necessitates both enrichment of phosphorylated peptides and chromatographic fractionation ahead of conventional capillary liquid chromatography tandem MS (nLC-MS/MS).[2,18,19]

Phosphorylation site quantification is often essential to elucidating biological insight.[18] 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.[19] 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 spectrometer 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.[33,34] 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 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. Next, 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, we conclude that most phosphorylation events are directed towards unstructured regions of proteins. Finally, we provide examples of key phosphorylation event discoveries relevant to metabolism.

RESULTS AND DISCUSSION

Overview of the Orbitrap Astral mass spectrometer

Owing to its extremely 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 1A**). Ions isolated by the quadrupole are first accumulated in the ion routing multipole before being passed to the ion processor, a conjoined collision cell and pulsed extraction ion trap[34], and are optionally fragmented with collisions before ejection into the Astral analyzer. Parallel processing of ion packets in different regions of the ion processor (and the ion routing multipole) allows the ion processor to operate above 200 Hz, while automatic gain control (AGC), prevents overload of the trap and increases its dynamic range by up to 100-fold.

Injected ion packet trajectories are shaped by the injection optics, an assembly of deflectors and lenses, to match the acceptance of the downstream mirror system. Ions oscillate between the opposing ion mirrors, and due to the slight ~2° injection angle, drift

down the length of the mirror system. The ion mirrors are asymmetric, tilted with respect to another, causing the ions to be deflected back slightly with each oscillation, reducing and eventually reversing the drift velocity and directing ions back through the analyzer to the high dynamic range (HDR) detector. The ion foil, a set of specially shaped compensation electrodes mounted above and below the ion beam, compensates for the distortion in ion oscillation frequency caused by the non-parallel mirrors, as well as improving the focus of ions at the detector surface. The asymmetric mirrors provide a >30-meter folded track for ions, a sufficient path length for achieving high resolving power. Due to a lack of grids and trajectory focusing in all dimensions, ion transmission may exceed 90%.

After striking the conversion dynode of the HDR detector, a scintillator converts electrons generated by the ion impact to photons, which are then detected by a photomultiplier tube (PMT). The signal is split into two measurement channels with a 10x difference in gain between them, proportionally increasing the dynamic range of the detection. This system can detect single ions, producing a signal-to-noise ratio approximately an order of magnitude higher than a typical Orbitrap MS/MS acquisition.

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.[33]

Rapid phosphopeptide Analysis with the Orbitrap Astral mass spectrometer

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[35], 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 1B** represents an example single-scan tandem mass spectrum that was identified following a traditional MaxQuant database search. 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 thirty-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 unique sites of phosphorylation. Note these MS/MS scans were collected within ~6.3 ms on average, enabling extremely 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 1C**) and mass accuracy (**Figure 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.[33] 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.[36–41] 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 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.

With a 250-ng injection of tryptic phosphopeptides, a 30-minute nLC separation, the Orbitrap Astral MS with the DIA method described above. The resultant data was searched both with Spectronaut and CHIMERYS (see Methods). From the Spectronaut search, we localized 29,190 phosphorylation sites (**Figure 2B,** see **Supplemental Figure 1** for comparison of Spectronaut and CHIMERYS performance). We note that the use of a wider DIA window (4 *m/z*) produced 14% fewer localized phosphorylation sites, despite decreasing cycle time. 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 2C**). We note that similar results are obtained down to 250 nanogram 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 2D, Supplemental Figure 2**). Impressively, a four-fold reduction in gradient length (30 to 7 minutes) results in just a ~20% reduction in localized phosphorylation sites.

To determine the impact of gradient length on reproducibility, we performed two additional analyses. First, we examined identification reproducibility across triplicate injections (**Figure 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 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 (**Supplemental Figure 3**).

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 first performed an entrapment search with a combined human and maize protein database.[42] In this experiment, which is commonly used to evaluate internal FDR, maize false identifications allow for the calculation of an external FDR. The results of this analysis are shown in **Figure 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.[43] S, T, and Y all have a significant amount of highly confident localized phosphosites whereas the P does not (**Figure 2H).** These results confirm that a library-free DIA search, using MS/MS spectra from an Astral analyzer, can reliably detect and localize phosphorylation sites.

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.[44] **Figure 3A** 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 analyzed non-enriched peptides from each tissue (**Supplementary Figure 4**).

Depending upon the source tissue, the number of identified phosphorylation sites ranged from approximately 20,000 to 40,000 sites (**Figure 3B**). 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.[32,45] We suppose that the increased sensitivity and depth afforded by this new analyzer permits the detection of these low-expression phosphorylation events (**Supplementary Figure 5C**) and represents an exciting advance. Also shown in **Figure 3B** 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 (**Supplementary Figure 5A**). Despite the age and strain differences between mice, and the dynamic nature of phosphorylation, we see good overlap in detection of sites. Tissue-specificity was also plotted (**Figure 3C** and **Supplementary Figure 5B**) 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 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 by the results shown in **Figure 3D-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).

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 (**Supplementary Figure 6A**). Notably, our analysis highlighted the significance of the phosphorylation amino acid as a major factor driving separation (**Supplementary Figure 6B**). 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 4A**), we identified

predominant clusters organized by specific motifs (**Figure 4B**), 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[46], our analysis revealed that the majority of ERK/CDK annotated phosphorylation sites belong to the S/T-P cluster (**Supplementary Figure 6C**). Note there is limited representation of Y in clusters 1 and 9, even though the phosphorylated residue was not considered in this analysis (**Supplementary Figure 6D**). These data confirm known biology that Y is not a substrate for P and R-directed protein kinases.[47]

To explore structural motifs, we mapped the 81,120 phosphorylation sites detected here onto protein structures predicted by AlphaFold.[48] 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 regions.[49] 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 4C**), 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. Note that when plotted individually, S, T, and Y phosphorylation sites follow the same trend (**Supplementary Figure 6E** and **6F**). These global results confirm previous work that suggests phosphorylation sites are enriched in intrinsically disordered regions.[50,51] In fact, some efforts have used intrinsically disordered regions to refine phosphorylation site prediction models.[52] 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 4D**). These data suggest further research on the role of structure and phosphorylation site is warranted.

Novel phosphorylation sites on mitochondrial proteins

We next sought to investigate how our comprehensive analysis of phosphorylation patterns in the mouse proteome can be applied in the biochemical and biological context. We first compared our dataset to the PhosphoSitePlus[46], revealing a capture of 38,454 previously unidentified phosphorylation sites, constituting 53% of the total identified sites within our dataset (**Figure 5A**). 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 5B**). This observation aligns

with the growing appreciation for protein phosphorylation in regulating mitochondrial function.[53] An exemplary mitochondrial pathway whose protein phosphorylations showed the most dynamic pattern across all tissues was the tricarboxylic acid cycle (**Supplementary Figure 7A**) with the majority of sites being detected for the first time in this study (**Supplementary Figure 7B**), setting the stage for further exploration into the understudied regulation of this key metabolic pathway.[54]

The occurrence of phosphorylation sites was variable across tissues in the entire dataset (**Supplementary Figure 7C**) as well as in the mitochondrial subset of our data (**Figure 5C**), 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 5D**). Moreover, these unique liver phosphorylation sites showed the largest proportion of mitochondrial matrix proteins, suggesting that phospho regulation of proteins in this sub-compartment is particularly important in this tissue. We identified carbamoyl phosphate synthetase I (CPS1) as the strongest driver of this observed matrix protein enrichment among liver unique sites, as it harbored the most phosphorylation sites in this group (**Figure 5E**).

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 5F**), 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.[55] The phosphorylation site S913, located within the L2 integrating domain, overlaps with a known severe disease-causing missense variant in CPS1 (p.S913L)[56], indicating the potential regulatory significance of these sites in CPS1.

Our dataset also provides enhanced phosphotyrosine coverage without specific enrichment methods (**Figure 3B**), 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. 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)[57] (**Figure 5G**). Previous work suggested that a distinct post-translational modification of Y1450 (nitration) may impact CPS1 activity by obstructing NAG binding.[58] 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 5H**). Therefore, phosphorylation of the Y1450 residue may distort the NAG binding site in addition to affecting the interaction of this helix with key NAG binding residues.

Phosphorylation in brain tissue is highly abundant

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.[44] 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 6A**). Moreover, we investigated all pan-tissue expression of phosphorylation sites (sites detected across all twelve tissues) by a z-score analysis and found brain tissue to exhibit the strongest abundance changes of phosphorylation sites (**Supplementary Figure 8A**). 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 (**Supplementary Figure 8B**). By applying the same absolute z-score threshold of 2, we found 9,541 phosphorylation site outliers in the brain (**Figure 6B**), 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 6C**).

OPA1 facilitates multiple functions in mitochondria including membrane fusion, cristae biogenesis, mitochondrial DNA maintenance, and respiration.[59] 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 (**Supplementary Figure 8C**), none of which are functionally characterized as documented in PhosphoSitePlus.[46] These sites are localized within the GTPase and PH domain of OPA1, which are also common areas for missense variants in patients with DOA.[60] 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 6D**). The consensus motif GxxxxGKS/T of G1/P-loop is highly conserved across species and other proteins of the dynamin superfamily (**Figure 6E**). Surprisingly, serine at position 4 is not considered part of this motif despite its high conservation. Notably, S298 has been identified as a site of missense mutations in DOA patients (c.893G>A, p.S298N[61] ; c.892A>G, p.S298G[62]), suggesting its importance for OPA1 function. Furthermore, a S298A mutation reduces

GTPase activity and dimerization *in vitro*[63]; the S298N mutation abolishes respiratory growth in *S.cerevisiae* with mitochondrial DNA depletion and altered mitochondrial morphology.[64] 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 6F**). 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.

CONFLICT OF INTEREST

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.

ACKNOWLEDGEMENTS. We are grateful for support from the National Institutes of Health (grants P41GM108538 and R35GM118110 to JJC, R01DK098672 to DJP, and R35GM150899 to AG). PS is supported by a Morgridge Interdisciplinary Postdoctoral Fellowship. TMPC and LRS acknowledge support from the National Human Genome Research Institution through a training grant to the Genomic Science Training Program (NIH T32HG002760). MLR acknowledges support from the UW-Madison Biotechnology Training Program (NIH T32GM135066). TMPC also acknowledges the ACS Division of Analytical Chemistry and Agilent for support through a graduate fellowship. PF acknowledges the support of postdoctoral fellowships by the European Molecular Biology Organization (ALTF 263-2022) and the Swiss National Science Foundation (P500PB_211038). DJP acknowledges the support of funds from the BJC Investigators Program.

METHODS

HEK293T Cell Preparation

HEK293T cells (ATCC) 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. Culture media was replaced every 24 hours. Cells were expanded to 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 107 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.

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. 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. 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 either human 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. 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%) prior to being desalted as previously described. 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, 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 280nm 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.

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 40cm at high pressures with C18 1.7 µm diameter, 130 Å pore size BEH C18 particles (Waters) as previously described.[35] Samples were analyzed with a Vanquish Neo UHPLC 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 Column Oven PRSO-V2 (Sonation Lab Solutions) and the flow rate was set to 400nL/min at the start of the method to decrease delay time and turned to 300nL/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 0.6s 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 for 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.

LC-MS Operation for Proteomics with Orbitrap Eclipse Analysis

Nanoflow capillary columns (75µm I.D., 360µm 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.[35] Fractionated mouse protein samples were analyzed with a Vanquish Neo UHPLC (Thermo Scientific) coupled to an Orbitrap Eclipse mass spectrometer (Thermo Scientific) using a NanoSpray Flex source (Thermo Scientific) incorporating a homebuilt column heating compartment. 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 25ppm 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.

DDA Data Processing

Proteomics DDA data was processed in MaxQuant 2.4.9.0 with default parameters.[65] 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. A HEK239T spectral library was generated in Spectronaut from the eight HEK2923T phosphopeptide fractions analyzed with a 30-minute active gradient 2 Th DIA method and the library was used to search all HEK293T phosphoproteomics experiments. 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.

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.

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.[66] 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.[67]

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 [68]).

Multiple Sequence Alignment

Clustal Omega[69] 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)[70]) and OPA1 (RCSB PDB structure 6JTG [63]).

Data analysis

Data analysis and figure creation was performed in Python 3. The code will be accessible at the time of publication.

Data Availability

The data will be accessible at the time of publication.

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). *International Journal of Molecular Medicine* **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. *Molecular BioSystems* **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 Science* **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. *Archives of Biochemistry and Biophysics* **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) Hart, S. R.; Waterfield, M. D.; Burlingame, A. L.; Cramer, R. Factors Governing the Solubilization of Phosphopeptides Retained on Ferric NTA IMAC Beads and Their Analysis by MALDI TOFMS. *J. Am. Soc. Mass Spectrom.* **2002**, *13* (9), 1042–1051. https://doi.org/10.1016/S1044-0305(02)00432-4.
(9) 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.
(10) 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.
(11) 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 Communications in Mass Spectrometry* **2010**, *24* (2), 219–231. https://doi.org/10.1002/rcm.4377.
(12) 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.
(13) 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.
(14) 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.
(15) 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.
(16) 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. *Proceedings of the National Academy of Sciences* **2004**, *101* (26), 9528–9533. https://doi.org/10.1073/pnas.0402700101.
(17) Olsen, J. V.; Mann, M. Status of Large-Scale Analysis of Post-Translational Modifications by Mass Spectrometry*. *Molecular & Cellular Proteomics* **2013**, *12* (12), 3444–3452. https://doi.org/10.1074/mcp.O113.034181.
(18) Nilsson, C. L. Advances in Quantitative Phosphoproteomics. *Anal. Chem.* **2012**, *84* (2), 735–746. https://doi.org/10.1021/ac202877y.
(19) 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.
(20) 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.
(21) 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.
(22) 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. *Proceedings of the National Academy of Sciences* **2007**, *104* (7), 2193–2198. https://doi.org/10.1073/pnas.0607084104.
(23) 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. *Proceedings of the National Academy of Sciences* **2009**, *106* (4), 995–1000. https://doi.org/10.1073/pnas.0811964106.
(24) Lemeer, S.; Heck, A. J. The Phosphoproteomics Data Explosion. *Current Opinion in Chemical Biology* **2009**, *13* (4), 414–420. https://doi.org/10.1016/j.cbpa.2009.06.022.
(25) 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.
(26) 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.
(27) 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.
(28) 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.
(29) Savitski, M. M.; Lemeer, S.; Boesche, M.; Lang, M.; Mathieson, T.; Bantscheff, M.; Kuster, B. Confident Phosphorylation Site Localization Using the Mascot Delta Score. *Molecular & Cellular Proteomics* **2011**, *10* (2), S1–S12. https://doi.org/10.1074/mcp.M110.003830.
(30) Molina, H.; Horn, D. M.; Tang, N.; Mathivanan, S.; Pandey, A. Global Proteomic Profiling of Phosphopeptides Using Electron Transfer Dissociation Tandem Mass Spectrometry. *Proceedings of the National Academy of Sciences* **2007**, *104* (7), 2199–2204. https://doi.org/10.1073/pnas.0611217104.
(31) 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.
(32) 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 Reports* **2014**, *8* (5), 1583–1594. https://doi.org/10.1016/j.celrep.2014.07.036.

(33) 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.

(34) 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.

(35) 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.

(36) 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.

(37) 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.

(38) 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.

(39) 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. *Molecular & Cellular Proteomics* **2022**, *21* (9), 100279. https://doi.org/10.1016/j.mcpro.2022.100279.

(40) 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.

(41) 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.

(42) 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*. *Molecular & Cellular Proteomics* **2017**, *16* (12), 2296–2309. https://doi.org/10.1074/mcp.RA117.000314.
(43) 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.
(44) 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.
(45) 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 Research* **2021**, *81* (9), 2495–2509. https://doi.org/10.1158/0008-5472.CAN-20-3804.
(46) 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.
(47) 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.
(48) 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.
(49) 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.
(50) 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.
(51) 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.
(52) 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.
(53) 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.

(54) 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.
(55) 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.
(56) 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.
(57) 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.
(58) 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.
(59) 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.
(60) 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.
(61) 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* **2015**, *138* (Pt 3), 563–576. https://doi.org/10.1093/brain/awu378.
(62) 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)* **2017**, *31* (3), 475–480. https://doi.org/10.1038/eye.2016.255.
(63) 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.
(64) 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.

(65) 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.
(66) 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.
(67) 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.
(68) 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.
(69) 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.
(70) 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.
(71) 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.
(72) Ge, S. X.; Jung, D.; Yao, R. ShinyGO: A Graphical Gene-Set Enrichment Tool for Animals and Plants. *Bioinformatics* **2020**, *36* (8), 2628–2629. https://doi.org/10.1093/bioinformatics/btz931.

Figure 1

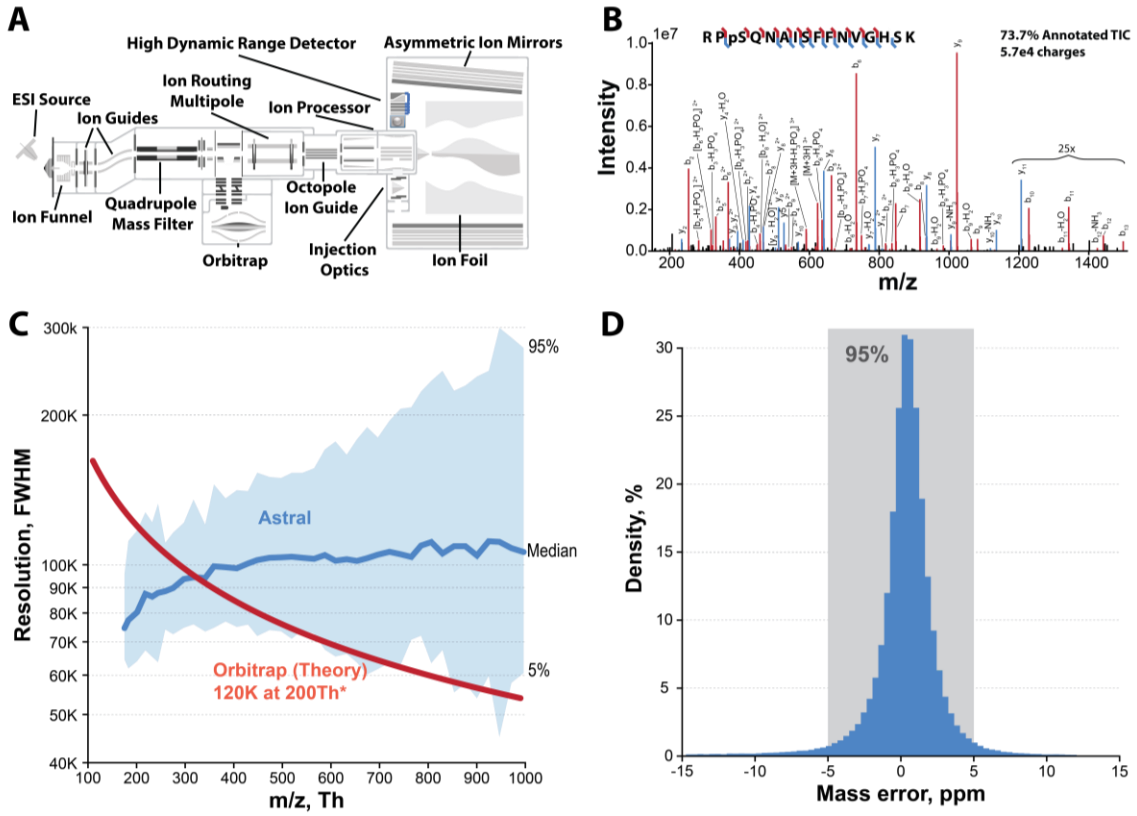


Figure 1. Overview of Orbitrap Astral MS and its key figures of merit. (A) Orbitrap Astral instrument schematic, highlighting the quadrupole, Orbitrap, and Astral analyzers. (B) Tandem mass spectrum of representative phosphopeptide collected using the Astral analyzer. (C) Distribution of Astral analyzer resolving power as a function of mass from phosphopeptide product ion spectra collected here. (D) Phosphopeptide product ion mass measurement error from Astral analyzer.

Figure 2

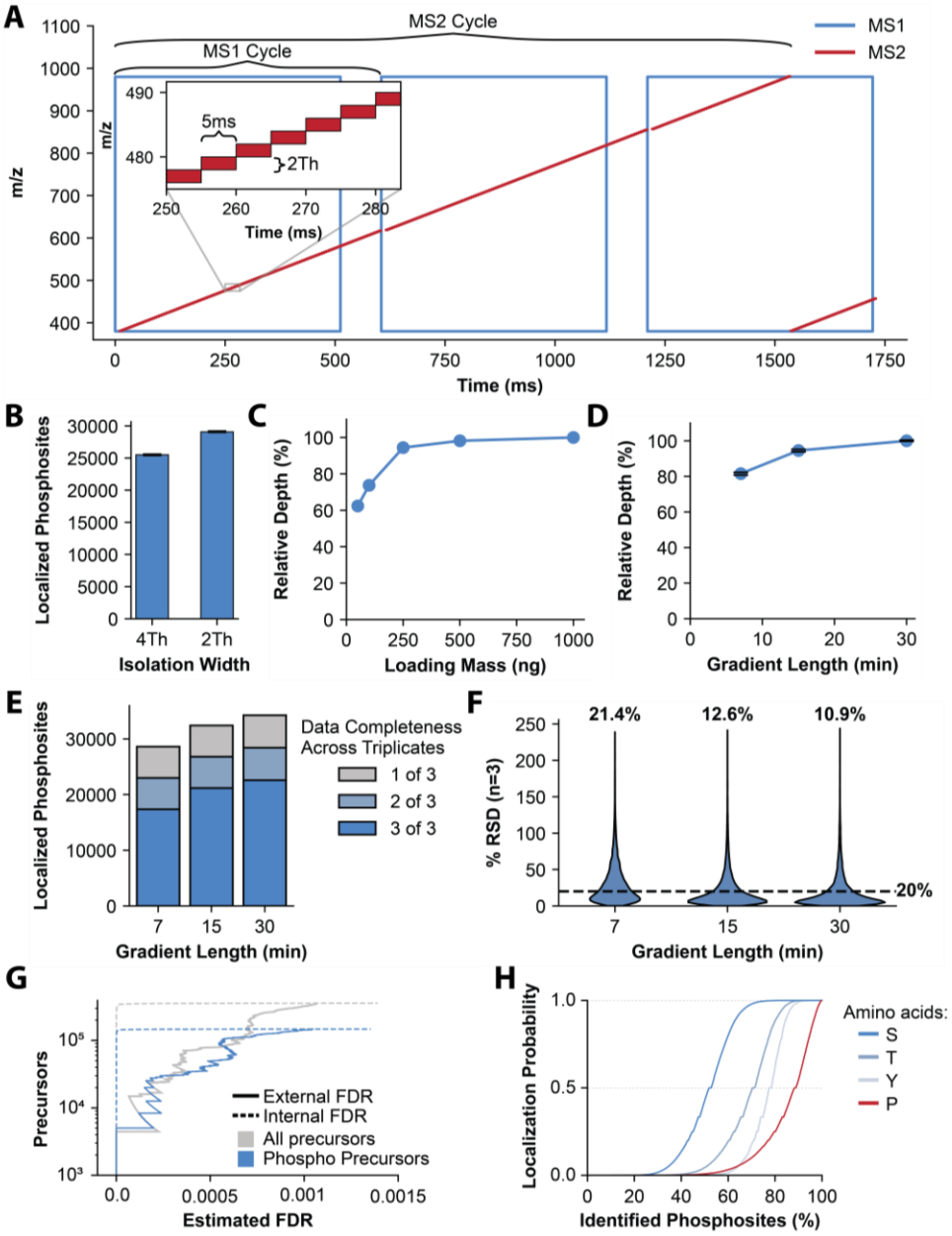


Figure 2. Overview of DIA phosphoproteomics on the Orbitrap Astral MS. (A) Illustration of DIA acquisition scheme. (B) Number of localized phosphorylation sites identified using DIA method with various isolation widths. (C) Evaluation of phosphopeptide loading mass on performance. Note data was collected using a 30-minute active gradient and 2 *m/z* DIA isolation width. (D) Evaluation of gradient length on performance. (E) Evaluation of phosphosite identification reproducibility (note this plot uses results from (D)). (F) Evaluation of phosphosite quantitative precision. Here 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.

Figure 3

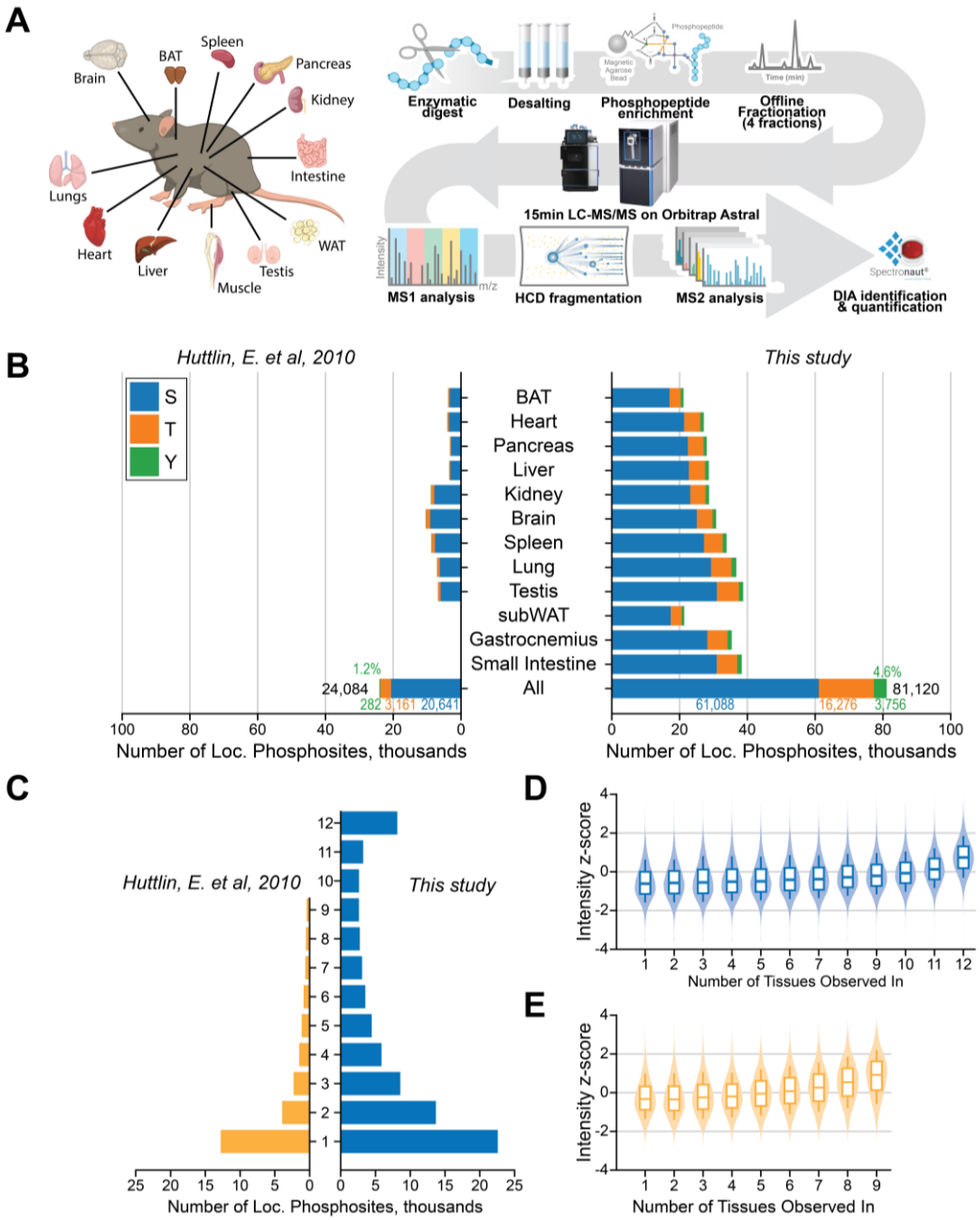


Figure 3. Mouse Phosphorylation Atlas Workflow and Results. (A) Mouse Phosphorylation Atlas Workflow. (B) Result for Mouse Tissue Phosphoproteomic Analysis. Numbers of unique phosphorylation site are shown for each tissue and the total unique sites with the fraction of S,T, and Y localizations indicated. The Huttlin et al. results 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. (D) and (E) Intensity distributions for phosphorylation sites detected in a given number of tissues for this study and by Huttlin et al., respectively.

Figure 4

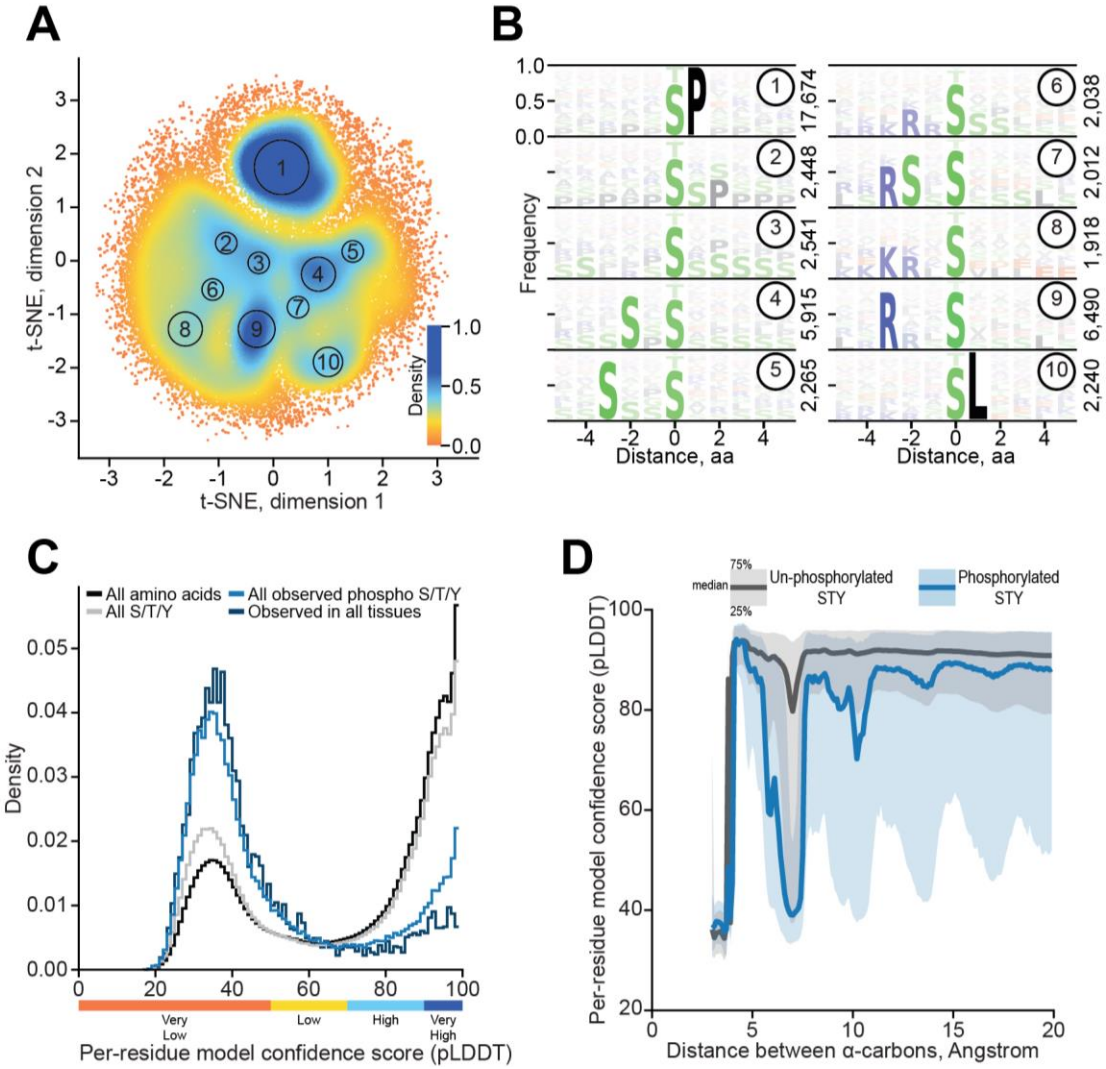


Figure 4. Sequence and Structural Context of Phosphosites. (A) Two-dimensional representation of all phosphosites and their 5-amino acid flanking sequences, excluding the central amino acid from the comparison. Each cluster has been manually selected to emphasize the densest regions. (B) Sequence logo 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) Confidence score distribution based on the distance between alpha-carbons surrounding phosphorylation site and phosphorylation state – phosphorylated and un-phosphorylated S/T/Y.

Figure 5

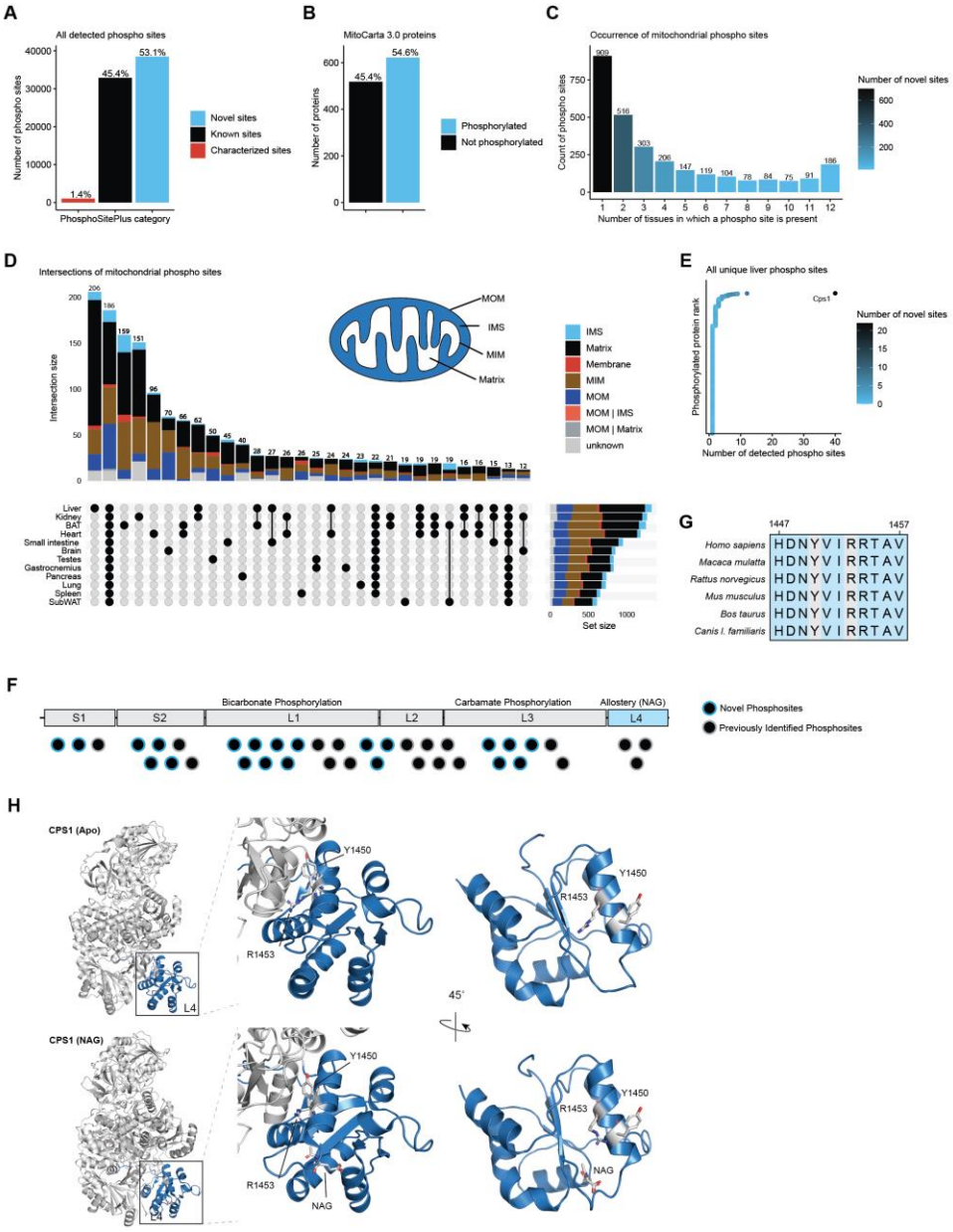


Figure 5. Mitochondrial phosphoproteomics reveals novel liver-specific phosphorylation site. (A) Bar plot of all detected phosphorylation sites in our study stratified into the categories from the PhosphoSitePlus database (downloaded August 22, 2023). (B) Mouse MitoCarta 3.0[71] proteins with at least 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[69]. 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[70].

Figure 6

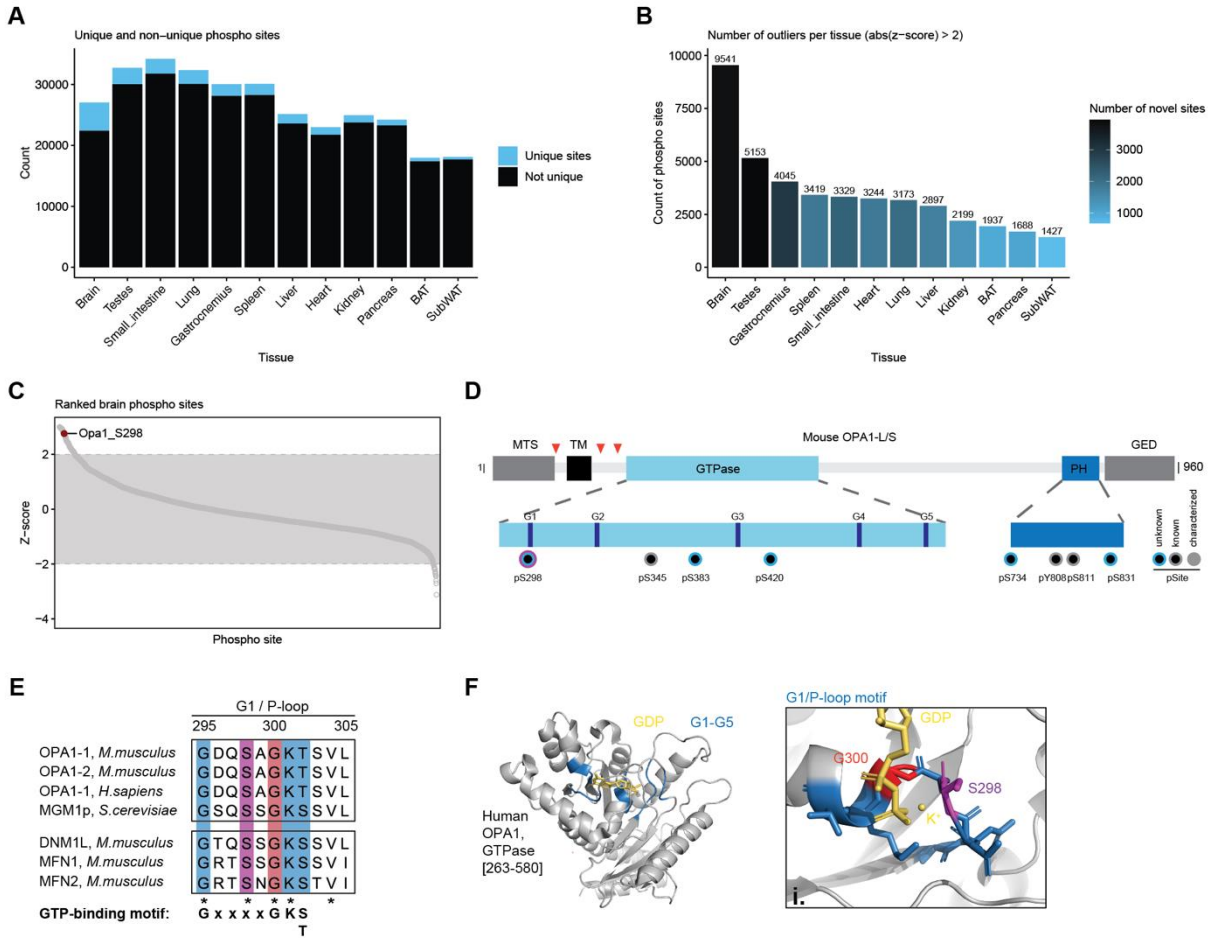
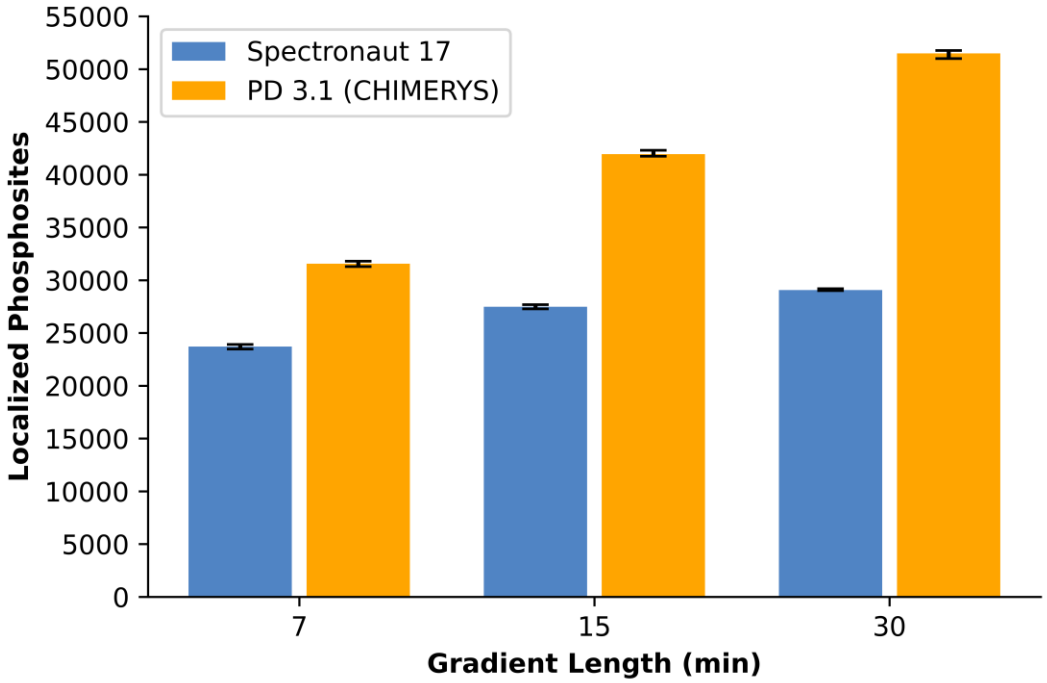


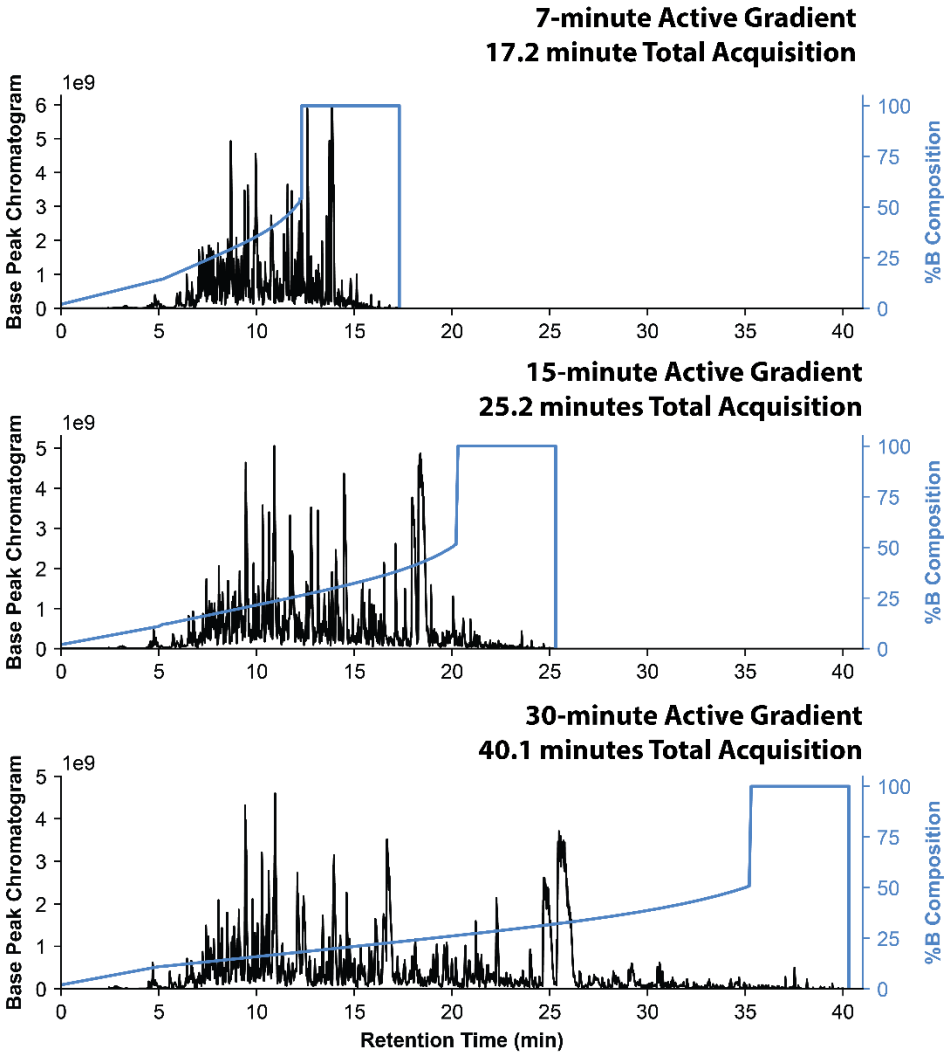
Figure 6. Unique phosphorylation sites in brain tissue. (A) Bar plot indicating all sites (black) that were detected per tissue, highlighting the sites that are unique in each tissue (blue). (B) Number of outliers per tissue based on z-score analysis with an absolute z-score cut-off of 2. (C) Ranked 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 Supplementary Figure 8C 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[63]. 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

Supplemental Figure 1



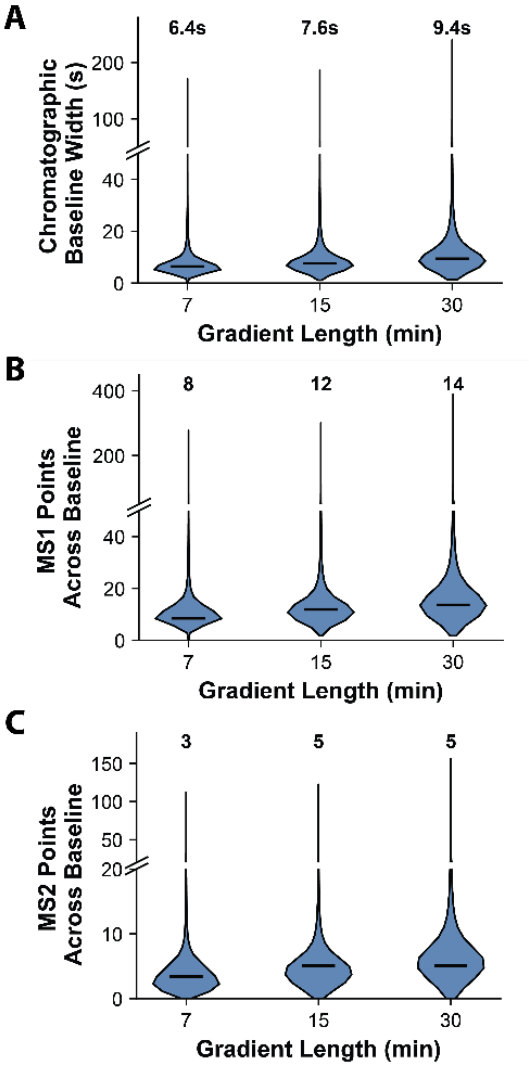
Supplemental Figure 1. DIA Phosphoproteomics Search with Spectronaut and CHIMERYS. The average number of localized phosphosites across triplicate injections for three different gradient lengths is shown for searches in Spectronaut (blue) and CHIMERYS in Proteome Discoverer 3.1 (orange). The error bars represent the maximum and minimum value across the triplicates.

Supplemental Figure 2



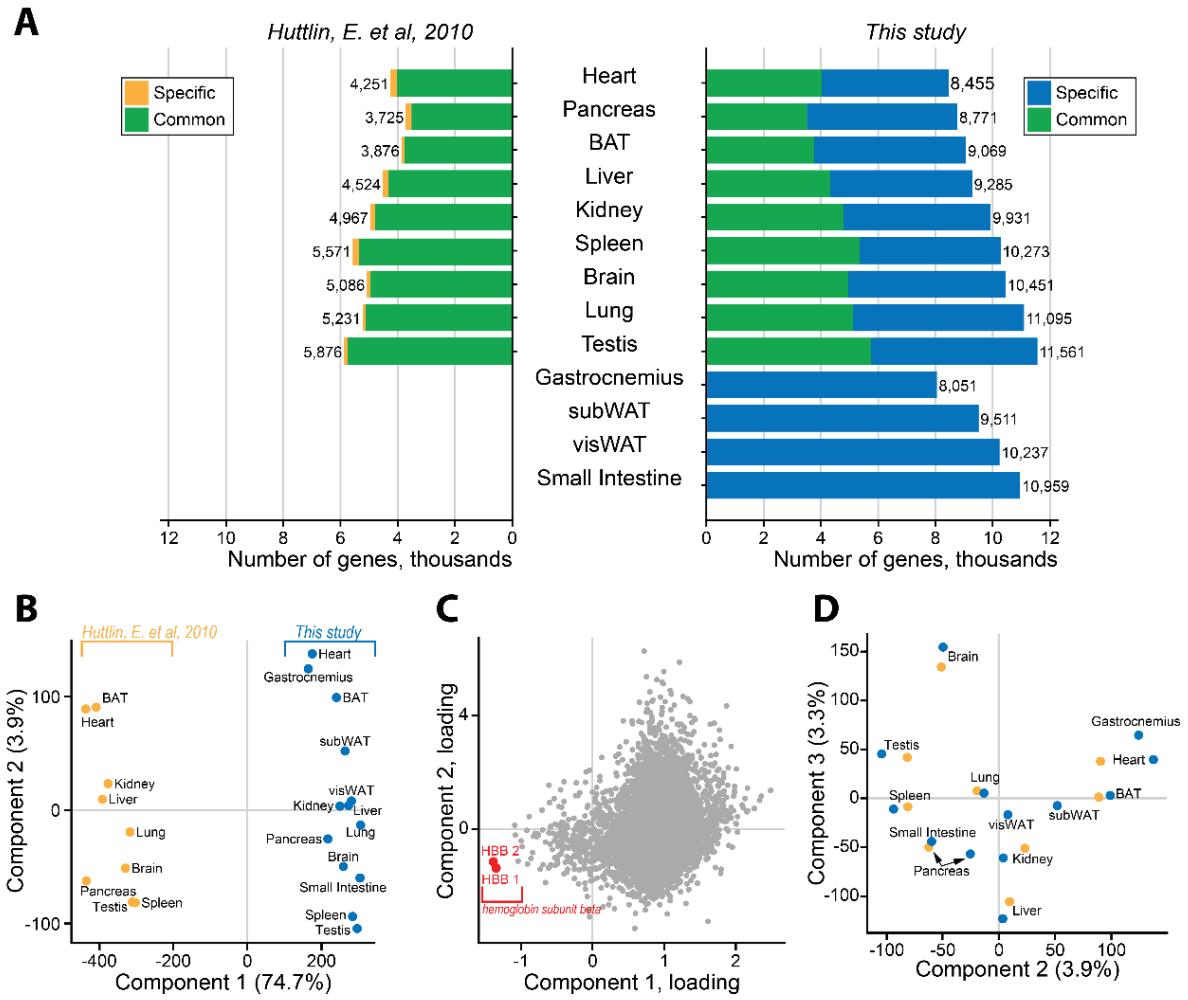
Supplemental Figure 2. Representative chromatograms for different gradient lengths. The base peak chromatograms for the three different active gradient lengths are shown with the gradient composition curve plotted.

Supplemental Figure 3



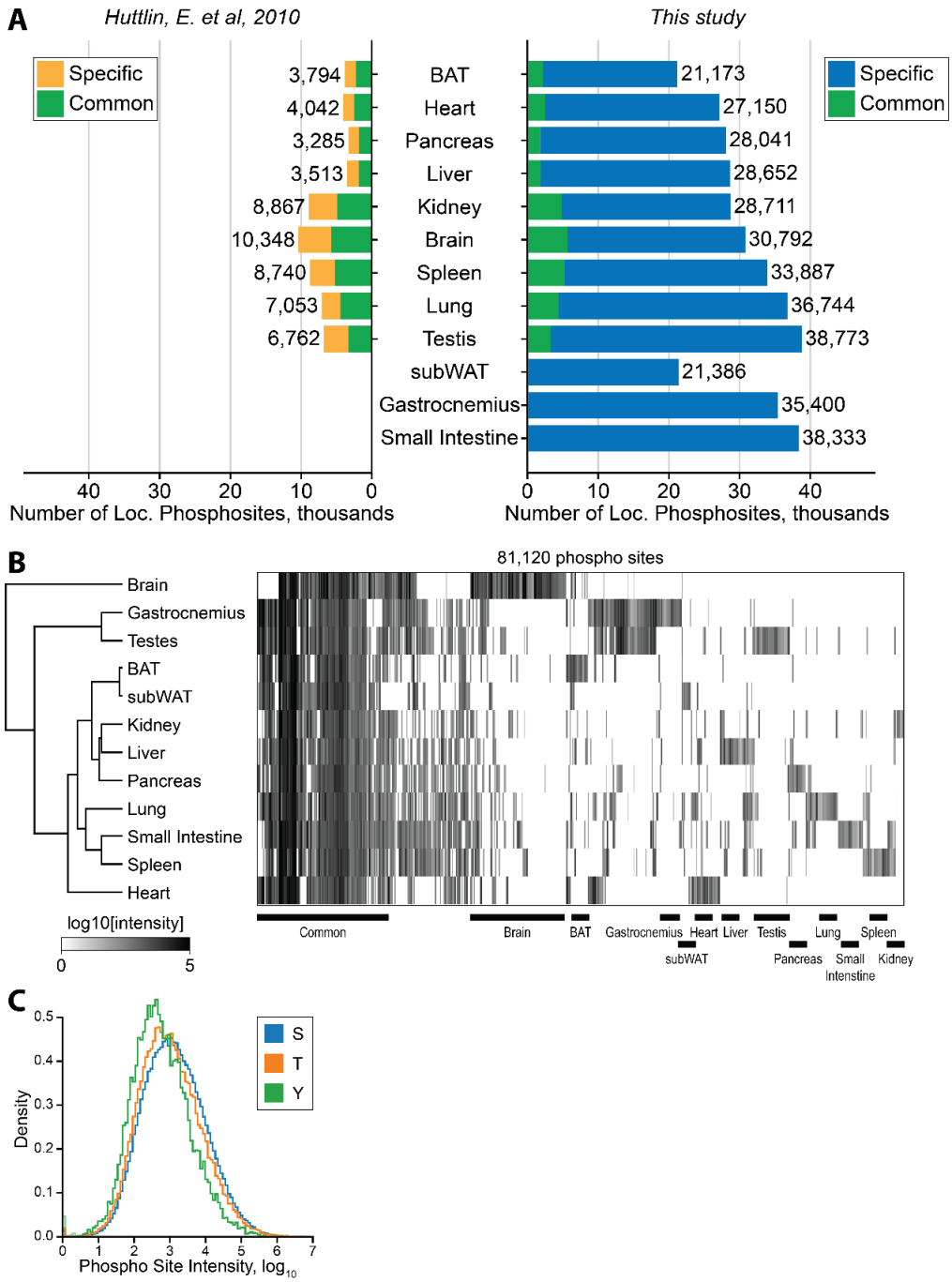
Supplemental Figure 3. As compared with the 15- or 30-minute gradient methods, the 7-minute method has fewer MS and MS/MS scans per chromatographic peak. The distribution of (a) chromatographic baseline widths, (b) MS1 points across the baseline peak, and (c) MS2 points across the baseline peaks are 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, MS1 points across FWHM, and MS2 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).

Supplemental Figure 4



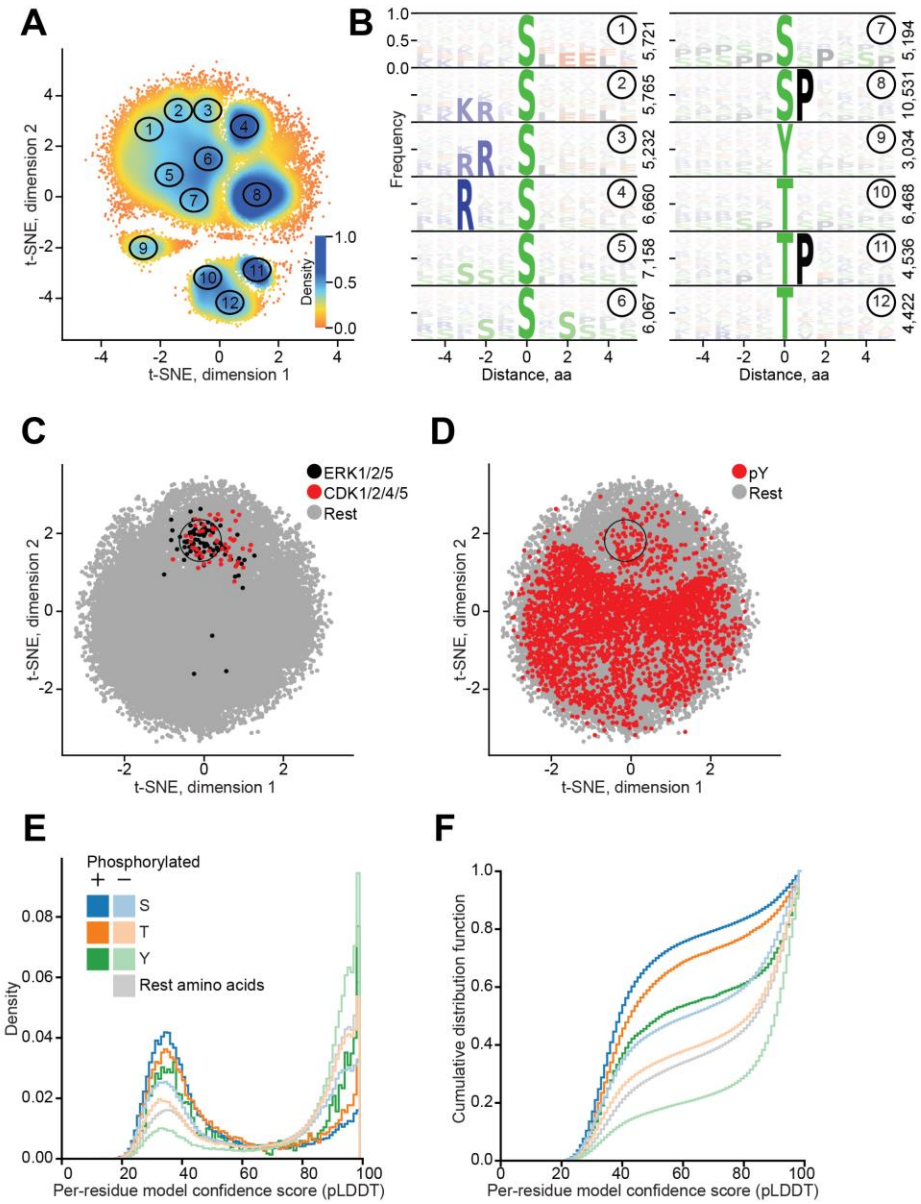
Supplemental Figure 4. Proteomic Profiling of Mouse Tissues. (A) Results from Mouse Tissue Proteomic Analysis. The numbers of protein groups are indicated for each tissue, with common protein groups highlighted in green, those unique to Huttlin et al. in yellow, and those unique to this study in blue. 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.

Supplemental Figure 5



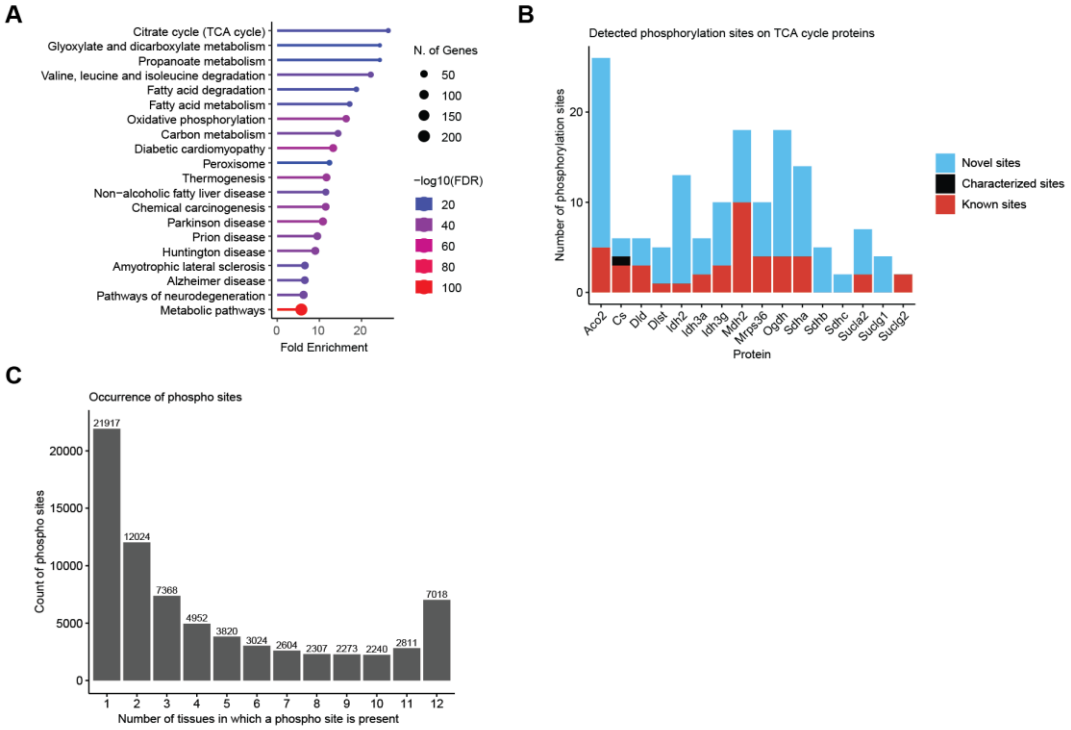
Supplemental Figure 5. Mouse Phosphorylation Atlas. (A) Overlap of phosphorylation sites detected in this study and by Huttlin et al. (B) Heatmap showing the relative intensities of phosphosites detected across tissues. Note that there are many phosphorylation sites with similar intensities across all 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.

Supplemental Figure 6



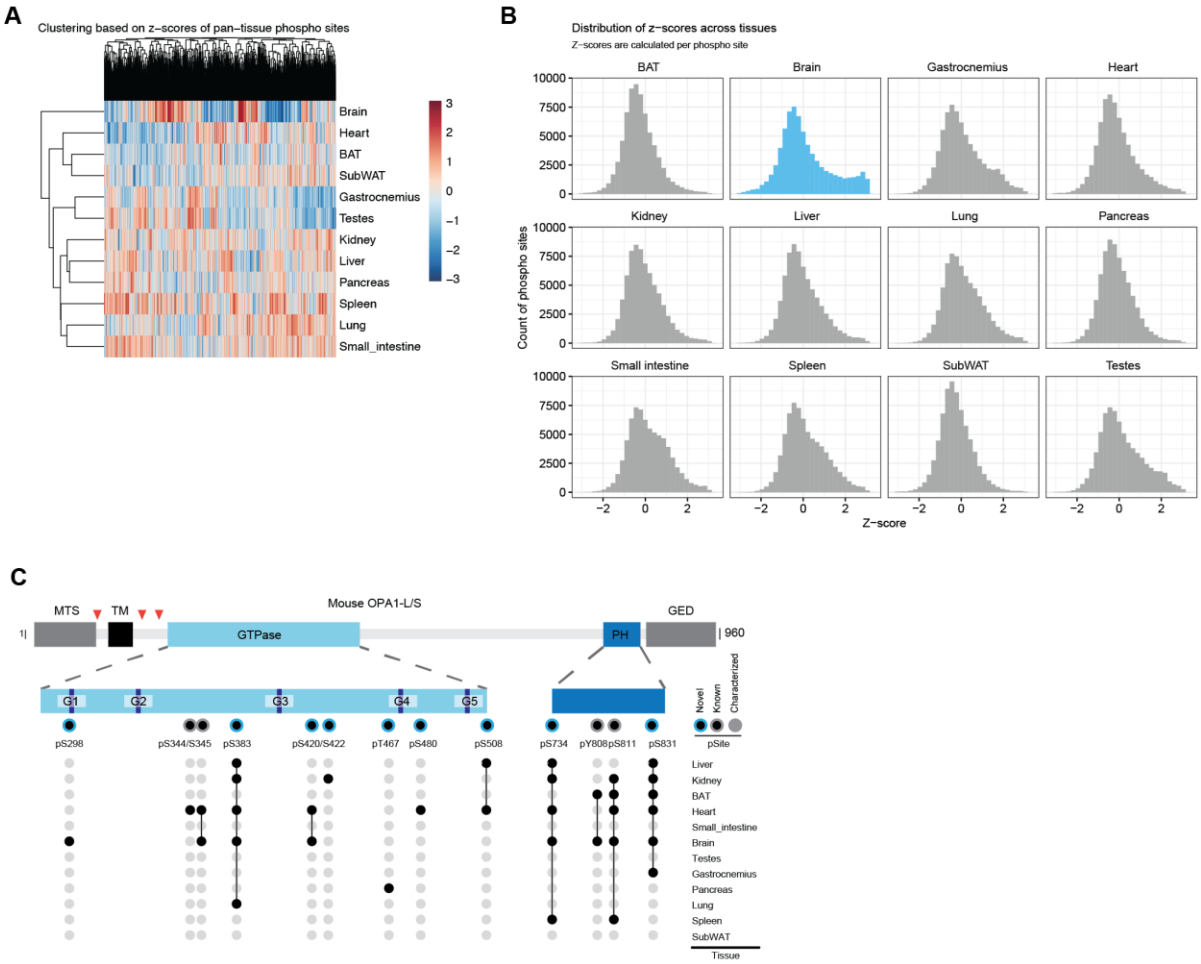
Supplemental Figure 6. Sequence and Structural Context of Phosphosites. (A) Two-dimensional representation of all phosphosites and their 5-amino acid flanking sequences, including the central amino acid in the comparison. Each cluster has been manually selected to highlight the densest regions. (B) Sequence logo plot for all clusters depicted in (A). (C) Similar to Figure 5A with highlights of ERK and CDK kinases based on the PhosphoSitePlus database. (D) Similar to Figure 5A 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.

Supplemental Figure 7



Supplementary Figure 7. 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.[72] 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.

Supplemental Figure 8



Supplementary Figure 8. 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 5D with additional phosphorylation sites not observed in brain tissue. The occurrence of OPA1 phosphorylation site within all twelve dataset is shown below.