Fast and Deep Phosphoproteome Analysis with the Orbitrap Astral Mass Spectrometer

Noah M. Lancaster[1,2,17], Pavel Sinitcyn[3,17], Patrick Forny[4,17], 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], Julie Rojas[11], Audrey P Gasch[11,12,13], Michael S. Westphall[2,5], Hamish Stewart[6], Christian Hock[6], Eugen Damoc[6], David J. Pagliarini[4,14,15], Vlad Zabrouskov[16], and Joshua J. Coon[1,2,3,5*]

[1]Department of Chemistry, University of Wisconsin-Madison, Madison, WI, USA
[2]Department of Biomolecular Chemistry, University of Wisconsin-Madison, Madison, WI, USA
[3]Morgridge Institute for Research, Madison, WI, USA
[4]Department of Cell Biology and Physiology, Washington University School of Medicine, St. Louis, MO, USA
[5]National Center for Quantitative Biology of Complex Systems, Madison, WI, USA
[6]Thermo Fisher Scientific (Bremen) GmbH, Bremen, Germany
[7]Cellular and Molecular Biology Graduate Program, University of Wisconsin-Madison, Madison, WI, USA
[8]Department of Nutritional Sciences, University of Wisconsin-Madison, Madison, WI, USA
[9]Department of Medicine, University of Wisconsin-Madison, School of Medicine and Public Health, Madison, WI, USA
[10]University of Wisconsin Carbone Cancer Center, University of Wisconsin-Madison, School of Medicine and Public Health, Madison, WI, USA
[11]Center for Genomic Science Innovation, University of Wisconsin-Madison, Madison, WI, USA
[12]Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI, USA
[13]Laboratory of Genetics, University of Wisconsin-Madison, Madison, WI, USA
[14]Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO, USA
[15]Department of Genetics, Washington University School of Medicine, St. Louis, MO, USA
[16]Thermo Fisher Scientific, San Jose, CA, USA
[17]These authors contributed equally: Noah M. Lancaster, Pavel Sinitcyn, Patrick Forny
*address correspondence to coon@wisc.edu

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 was benchmarked to other state-of-the-art MS platforms using both synthetic peptide standards and with EGF-stimulated HeLa cells. We applied this approach to generate a phosphoproteome multi-tissue 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, 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 itself elevate the difficultly.[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 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 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**). 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 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 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 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.[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 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 **Supplemental Figure 1A** for the base peak chromatogram). The resultant data was searched both with Spectronaut and

163 a developmental build of CHIMERYS (see Methods). From the Spectronaut search, we
164 localized 29,190 phosphorylation sites (**Figure 2B,** see **Supplementary Figure 2A** for
165 comparison of Spectronaut and CHIMERYS, and **Supplementary Figure 2B** and **2C**
166 phosphopeptide multiplicity and localization probability distributions, respectively). We
167 note that the use of a wider DIA window (4 *m/z*) produced 14% fewer localized
168 phosphorylation sites. Next, we examined the impact of loading amount on performance
169 using a dilution series ranging from 1 µg to 50 ng loaded on column (**Figure 2C**). We note
170 that similar results are obtained down to 250 ng loads. The sensitivity of the system is
171 evident in that a 50-ng load produced ~60% of the detected sites observed with a 1 µg
172 load. To determine whether the Astral scan speed would allow for increased throughput,
173 we next evaluated a range of active gradient lengths (**Figure 2D, Supplementary Figure**
174 **1A**). A four-fold reduction in gradient length (30 to 7 minutes) results in just a ~20%
175 reduction in localized phosphorylation sites. Additional method parameters are justified in
176 the **Supplemental Text**.
177
178 To determine the impact of gradient length on reproducibility, we performed two additional
179 analyses. First, we examined identification reproducibility across triplicate injections
180 (**Figure 2E**). The 7- ,15-, and 30-minute active gradient methods resulted in 56.7%,
181 61.8%, and 62.7% of phosphosites detected across all three injection replicates,
182 respectively. Second, we assessed quantitative precision by determining percent relative
183 standard deviation (RSD) of phosphorylation sites detected across all three replicates
184 (**Figure 2F**). The 15- and 30-minute active gradients achieved median RSDs of 10.9 and
185 12.6%. Not unexpectedly, the 7-minute active gradient method has considerably reduced
186 precision (21.4% median RSD), most likely due to insufficient sampling across the
187 narrower elution peaks (**Supplementary Figure 1B-D**).
188
189 Since these data were collected with a novel mass analyzer and new instrument, we next
190 sought to ensure that the phosphosites were identified and localized accurately. To test
191 this, we first performed an entrapment search with a combined human and maize protein
192 database.[46] In this experiment, which is commonly used to evaluate internal false
193 discovery rate (FDR), maize false identifications allow for the calculation of an external
194 FDR. The results of this analysis are shown in **Figure 2G**, where internal and external
195 FDR is plotted for all precursors and phosphopeptide precursors. For lower q-values the
196 internal and external FDRs disagree; however, these curves converge at approximately
197 the 1% FDR threshold, suggesting that the identifications reported here are reliable. We
198 conclude that, as the DIA software tools continue to evolve and are adopted for the new
199 data type, this discordance will be diminished.
200
201 In the case of phosphoproteomics, peptide identification is not the final result – the site of

202 phosphorylation is ideally localized to a specific residue with confidence. To evaluate the
203 quality of phosphosite localization, we searched these data with phosphorylated proline
204 as a variable modification where any detected proline phosphorylation is false.[47] S, T, and
205 Y all have a significant amount of highly confident localized phosphosites whereas the P
206 does not (**Figure 2H, Supplemental Figure 2H**). To further explore localization
207 confidence, we analyzed a set of synthetic phosphopeptide standards that had been
208 spiked into a complex mixture yeast tryptic phosphopeptides at various concentrations.
209 Note this same 225 synthetic phosphopeptide mix has been used by others and thereby
210 provides a means of both assessment and comparison[41,48]; specifically, the experiment
211 allows for calculation of localization error rate based on ground truth knowledge. From
212 these data we conclude that, depending on localization probability cutoff, the error rate
213 ranges between 1 and 5 percent, a similar trend as the previous study (**Figure 2I,**
214 **Supplemental Figure 3A**).[41] **Supplemental Figure 3B** demonstrates that we retain the
215 majority of correct precursors even at more stringent localization probability cutoffs.
216 Taken together, these results confirm that a library-free DIA search, using MS/MS spectra
217 from an Astral analyzer, can reliably detect and localize phosphorylation sites.
218
219 The aforementioned phosphopeptide standards also provided an opportunity to evaluate
220 quantitative linearity.[42,48] To accomplish this, phosphopeptide standards were diluted
221 from 10,000 to 39 attomole (per standard, on-column) into a constant complex mixture of
222 yeast tryptic phosphopeptides. Using the 15-minute DIA Orbitrap Astral method, we
223 analyzed the various spike-in samples and then calculated the linear fit between the
224 observed MS intensities and the load amount. **Supplemental Figure 3D** displays three
225 randomly chosen curves while **Figure 2J** confirms that the majority of the phosphosite
226 calibration curves have R^2 values greater than 0.95, indicating good quantitative linearity.
227 **Supplemental Figure 3C** confirms that even at loads as low as 39 attomole, half of the
228 phosphorylation sites were detected across all three injections. Finally, we note that the
229 phosphopeptide standard intensity distributions observed across the dilution series
230 overlap with the yeast phosphopeptide intensity distribution (**Supplemental Figure 3E**).
231 Together, these results demonstrate that our method exhibits reliable detection and
232 quantification over the concentration range typically observed in phosphoproteomic
233 analysis.
234

235 Comparison of Orbitrap Astral to other phosphoproteomics platforms

236
237 To compare the performance of the Orbitrap Astral to other phosphoproteomics platforms
238 we replicated the HeLa EGF stimulation experiment as previously reported by Skowronek
239 et al. using the timsTOF Pro hybrid mass spectrometer with dia-PASEF acquisition.[33]
240 Briefly, we cultured HeLa cells in triplicate followed by 15 minutes of EGF stimulation, cell

241 lysis, protein extraction, trypsin digestion, and phosphopeptide enrichment. The resulting
242 phosphopeptide samples were then analyzed with a 15-minute active gradient, closely
243 replicating the total acquisition time for the Skowronek et al. publication, with detection
244 using either the Orbitrap Astral or Orbitrap Ascend™ Tribrid mass spectrometers. To
245 enable a direct comparison, we searched our raw files and those from the Skowronek et
246 al. study with the same protein database and Spectronaut library-free method. **Figure 3A**
247 shows the phosphoproteomic depth and completeness across triplicates for all three
248 datasets. The Orbitrap Astral method provided the deepest phosphoproteomic analysis
249 and yielded near 60% phosphosite completeness across replicates from approximately
250 three-fold more localized sites as compared to either the timsTOF Pro or Orbitrap Ascend
251 instruments. Note the Orbitrap Ascend is not optimally operated in DIA mode given its
252 slower scan speed; however, we chose to use a DIA method for direct comparison.
253 Interestingly, the Orbitrap Ascend generated the highest level of phosphosite overlap of
254 all three instruments at 64%.
255
256 To examine the dynamic range of detected phosphorylation sites across platforms we
257 plotted the intensity distributions of phosphorylation sites detected in all three replicates
258 using only the EGF-stimulated data. **Figure 3B** and **3C** present the Orbitrap Astral
259 intensity distributions with overlapping sites indicated for either the timsTOF Pro or
260 Orbitrap Ascend datasets, respectively. Both plots reveal that the unique phosphosites in
261 the Orbitrap Astral dataset are biased toward lower intensities. In contrast, unique
262 phosphosites for the timsTOF Pro dataset are distributed more evenly across the intensity
263 range (**Figure 3D**). These observations suggest that our Orbitrap Astral exhibits improved
264 sensitivity, resulting in a larger dynamic range. Note the Orbitrap Astral and Ascend
265 shows higher overlap as compared to the timsTOF Pro (**Figure 3E**); we suppose this
266 difference is likely due to variation in EGF treatment and sample preparation.
267
268 To see how this performance variation translates to biological discovery we performed a
269 differential expression analysis between control and EGF-stimulated samples for all three
270 datasets (**Figure 3F**). Each datapoint in these plots presents a phosphorylation site with
271 the differentially expressed ones highlighted in blue. Phosphorylation sites occurring
272 directly on the EGFR are labeled, many of which are detected across all platforms. To
273 further compare we performed a pathway enrichment analysis from each dataset (**Figure**
274 **3G**) – all three platforms had the EGF/EGFR signaling pathway as the top enriched term.
275 We conclude that despite differences in treatment, sample preparation, and
276 instrumentation, all technologies could arrive at the same biological conclusions.
277
278 ## Deep phosphoproteome analysis of the mouse
279

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 4A** 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 (**Supplementary Table 3**). Note to provide complementary protein abundance we separately analyzed non-enriched peptides from each tissue (**Supplementary Figure 4, Supplementary Table 4**).

Depending upon the source tissue, the number of identified phosphorylation sites ranged from approximately 20,000 to 40,000 sites (**Figure 4B**). 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 (**Supplementary Figure 5C**). Also shown in **Figure 4B** 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 (median 55.3%). Tissue-specificity was also plotted (**Figure 4C** 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 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

319 improvements in sensitivity and reproducibility are likely needed to detect them. This
320 hypothesis is further evinced by the results shown in **Figure 4D-E**, where we examined
321 the detected peak intensities for each site as a function of how many tissues in which it
322 was observed. From these data, both studies show concordance between number of
323 tissues detected and overall signal strength (*i.e.*, abundance).
324
325 Furthermore, we compared our comprehensive phosphoprotein atlas of the mouse with
326 the most recent multi-tissue study conducted by Giansanti and co-workers.[51] Similar to
327 our sample preparation, the phosphopeptides from multiple tissues were offline
328 fractionated and further concatenated into 4 samples. These samples were measured
329 using a 90-minute gradient DDA on the Q Exactive Orbitrap HF MS system. Despite a 6-
330 fold shorter measurement time, the Orbitrap Astral detected 2-3 times more phosphosites
331 per tissue, with improved agreement with our dataset compared to Huttlin et al. (median
332 64.2% versus 55.3%; **Supplementary Figure 6A**). Interestingly, Giansanti's study shows
333 a similar trend towards a higher number of sites that are present across all tissues
334 (**Supplementary Figure 6B**). The portion of pY in this study is 1.1% (**Supplementary**
335 **Figure 6C**), which additionally highlights the capability of combining Astral sensitivity with
336 DIA acquisition to enable the detection of 4.6% pY in our data.
337
338 ## Phosphorylation sites in the context of protein sequence and structure
339
340 Phosphorylation sites within proteins exhibit a specificity determined by both the
341 sequence and structural motifs. In our investigation, we sought to unravel this specificity
342 by clustering phosphorylation sites and their flanking regions. To achieve this, we
343 computed all pairwise comparisons of phosphosite sequences (site and five amino acids
344 in both C and N-terminal direction) and projected this comparison onto a 2D plane using
345 the t-SNE algorithm (**Supplementary Figure 7A**). Notably, our analysis highlighted the
346 importance of the phosphorylation amino acid as a major factor driving separation
347 (**Supplementary Figure 7B**). That is clusters 1-8 are Ser-based, while 9 is Tyr-based,
348 and 10-12 are Thr-based. Intriguingly, upon repeating the analysis while excluding the
349 phosphorylated amino acid from binary comparisons (**Figure 5A**), we identified
350 predominant clusters organized by specific motifs (**Figure 5B**), such as S/T-P (cluster 1),
351 and RXX-S/T (cluster 9). Proline-directed protein kinases demonstrate a preference for
352 phosphorylating serine or threonine residues immediately preceding a proline residue in
353 proteins. Prominent examples of such kinases include extracellular signal-regulated
354 kinases (ERKs) and cyclin-dependent kinases (CDKs). Leveraging annotations from the
355 PhosphoSitePlus database[52], our analysis revealed that the majority of ERK/CDK
356 annotated phosphorylation sites belong to the S/T-P cluster (**Supplementary Figure 7C**).
357 Note there is limited representation of Y in clusters 1 and 9, even though the

358 phosphorylated residue was not considered in this analysis (**Supplementary Figure 7D**).
359 These data confirm known biology that Y is not a substrate for P and R-directed protein
360 kinases.[53]
361
362 To explore structural motifs, we mapped the 81,120 phosphorylation sites detected here
363 onto protein structures predicted by AlphaFold.[34] The UniProt structural library
364 (assembled using AlphaFold) contains structures with varying degrees of certainty as
365 measured by a per residue confidence score (pLDDT) where 0 is very low and 100 is very
366 high confidence. Note that lower confidence scores are indicative of either flexible or
367 disordered protein regions.[54] Upon plotting the confidence score distribution for
368 phosphorylated S, T, and Y sites, we observed a striking prevalence for these sites to be
369 in low confidence regions (**Figure 5C**), especially when compared to all S, T, and Y
370 residues or even all amino acids. Phosphorylation sites that were detected in all tissues
371 had an even slightly higher preference for these low confidence regions. However, all
372 amino acids observed in non-enriched proteomics experiments did not follow the same
373 trend, emphasizing the specificity of this effect for phosphorylation sites. It is noteworthy
374 that when plotted individually, S, T, and Y phosphorylation sites all exhibit the same trend
375 (**Supplementary Figure 7E** and **7F**). These global results confirm previous work that
376 suggests phosphorylation sites are enriched in intrinsically disordered regions.[35,36] In fact,
377 some efforts have used intrinsically disordered regions to refine phosphorylation site
378 prediction models.[37] Finally, although phosphorylation is more likely to occur in less
379 structured regions, examination of the surrounding environment shows the confidence
380 scores are increasing with distance from the phosphorylated residue (**Supplementary**
381 **Figure 7G**). These data suggest further research on the role of structure and
382 phosphorylation site is warranted.
383

384 ## Kinase predictions based on flanking regions of phosphorylation sites

385
386 Further leveraging our extensive dataset of phosphorylation sites across various tissues,
387 we applied a recently published kinase prediction tool[38], which was developed by using
388 synthetic peptide libraries to profile the substrate sequence specificity of 303
389 serine/threonine kinases (**Figure 5D**). As expected, a systematic analysis of all
390 phosphorylation sites in our dataset revealed a trend where many phosphorylation sites
391 were assigned to a limited number of putative kinases; however, half of all
392 phosphorylation sites were predicted to be targeted by 24 or more kinases (**Figure 5E**).
393 Furthermore, this approach enabled us to predict kinases for sites that showed tissue
394 enrichment (**Figure 5F**). By comparing enriched phosphorylation motifs between tissue
395 pairs, we identified kinases with established activities in certain tissues (**Supplementary**
396 **Figure 8**). For instance, comparing brain and liver enriched motifs revealed kinases such

397 as GSK3A[55], CAMK subfamily 2 members[56–58], and PCKA/PCKB[59,60], which are known
398 to modulate the brain phosphoproteome and have been implicated in neurological
399 disease states (**Figure 5G**). While these kinases have activities in other tissues, our
400 dataset emphasized that the phosphorylation sites predicted to be clients of these kinases
401 are strongly overrepresented in brain as compared to liver tissue. Additionally, we
402 identified PHKG2 as the kinase with the strongest frequency factor for liver tissue,
403 consistent with this enzyme's role as a liver isoform critical for glycogen breakdown.[61] Our
404 analysis suggests additional kinases with probable tissue-specific activities, highlighting
405 the potential of our resource to facilitate the discovery of such tissue-specific kinase
406 functions.
407

408 ## Novel phosphorylation sites on mitochondrial proteins

409

410 We next sought to investigate how our comprehensive analysis of phosphorylation
411 patterns in the mouse proteome can be further applied in the biochemical and biological
412 context. We first compared our dataset to the mouse phosphorylation dataset from
413 PhosphoSitePlus[52], revealing a capture of 38,454 previously unidentified phosphorylation
414 sites, constituting 53% of the total identified sites within our dataset (**Figure 6A**). Upon
415 categorizing phosphorylation sites, we observed a significant representation of
416 mitochondrial proteins, with 55% of known mitochondrial proteins harboring at least one
417 phosphorylation site (**Figure 6B**). This observation aligns with the growing appreciation
418 for protein phosphorylation in regulating mitochondrial function.[62] An exemplary
419 mitochondrial pathway whose protein phosphorylation sites showed the most dynamic
420 pattern across all tissues was the tricarboxylic acid cycle (**Supplementary Figure 9A**)
421 with the majority of sites being detected for the first time in this study (**Supplementary**
422 **Figure 9B**), setting the stage for further exploration into the understudied regulation of
423 this key metabolic pathway.[63]
424

425 The occurrence of phosphorylation sites was variable across tissues in the entire dataset
426 (**Supplementary Figure 9C**) as well as in the mitochondrial subset of our data (**Figure**
427 **6C**), indicating that tissue-specific phosphorylation patterns exist within the murine
428 mitochondrial phosphoproteome. To investigate this in more detail, we calculated the
429 intersections of all mitochondrial phosphorylation sites across tissues and found the liver
430 to harbor the most unique phosphorylation sites (**Figure 6D**). Moreover, these unique
431 liver phosphorylation sites showed the largest proportion of mitochondrial matrix proteins,
432 suggesting that phosphoregulation of proteins in this sub-compartment is particularly
433 important in this tissue. We identified carbamoyl phosphate synthetase I (CPS1) to harbor
434 the most phosphorylation sites in this group (**Figure 6E**).

435 Notably, 40 phosphorylation sites were detected on CPS1, which is the first and rate-
436 limiting enzyme of the urea cycle. None of these phosphorylation sites have been
437 functionally characterized, and 22 of these sites have not been reported previously
438 (**Figure 6F**), underscoring the significant discovery potential of our dataset. Deficiency of
439 human CPS1 results in hyperammonemia ranging from neonatally lethal to
440 environmentally induced adult-onset disease in affected individuals.[64] The
441 phosphorylation site S913, located within the L2 integrating domain, overlaps with a
442 known severe disease-causing missense variant in CPS1 (p.S913L)[65], indicating the
443 potential regulatory significance of these sites in CPS1.
444
445 Our dataset also provides enhanced phosphotyrosine coverage without specific
446 enrichment methods (**Figure 4B**), revealing novel sites that may be functionally relevant.
447 N-acetyl-L-glutamate (NAG) is an essential allosteric activator of CPS1. By binding to the
448 C-terminal L4 domain of CPS1, NAG induces long-range conformational changes,
449 impacting distant catalytic activities.[66] We identify Y1450 as a conserved phosphorylated
450 residue in this domain, which is located proximal to the key R1453 residue that is required
451 for NAG binding and is the location of two patient variants (c.4357C>T and c.4358G>A
452 resulting in p.R1453W and p.R1453Q, respectively)[67] (**Figure 6G**). Previous work
453 suggested that a distinct post-translational modification of Y1450 (nitration) may impact
454 CPS1 activity by obstructing NAG binding.[68] Structural modeling demonstrates shifts in
455 this Y1450-R1453-containing helix of CPS1 upon NAG binding, highlighting the
456 confirmational change of the R1453 residue and its proximity to the Y1450 residue
457 (**Figure 6H**).
458

459 ## Most dynamic phosphoregulation in brain tissue

460
461 The brain is characterized by a diverse array of expressed kinases and phosphatases
462 and its proteins have a higher tendency to be phosphorylated as compared to other
463 tissues.[49] Our data substantiates this observation, as brain tissue showed the highest
464 prevalence of unique phosphorylation sites (*i.e.*, sites defined as exclusive to a singular
465 tissue, **Figure 7A**). Moreover, we investigated all pan-tissue expression of
466 phosphorylation sites (sites detected across all twelve tissues) by a z-score analysis,
467 enabling comparison of each phosphosite to the average organismal values, and found
468 brain tissue to exhibit the strongest abundance changes of phosphorylation sites
469 (**Supplementary Figure 10A**). We then applied the same z-score-based principle on the
470 entire dataset (including imputed values for undetected sites in some tissues) to
471 impartially identify sites undergoing significant changes and observed a normal
472 distribution in z-score patterns for all tissues, except the brain, which displayed a
473 pronounced enrichment of phosphorylation sites with elevated z-scores (**Supplementary**

474 **Figure 10B, Supplementary Table 6**). By applying the same absolute z-score threshold
475 of 2, we found 9,541 phosphorylation site outliers in the brain (**Figure 7B**), accounting for
476 23% of all outliers across all twelve tissues. Further prioritizing these brain sites based on
477 z-score ranking within the mitochondrial phosphoproteome, we identified serine 298 of
478 Optic atrophy protein 1 (OPA1) as a top hit (**Figure 7C**).
479
480 OPA1 facilitates multiple functions in mitochondria including membrane fusion, cristae
481 biogenesis, mitochondrial DNA maintenance, and respiration.[69] Importantly, mutations in
482 OPA1 are the most common cause of dominant optic atrophy (DOA). In our data, we
483 detected multiple phosphorylation sites of OPA1 across the twelve tissues
484 (**Supplementary Figure 10C**), none of which are functionally characterized as
485 documented in PhosphoSitePlus.[52] These sites are localized within the GTPase and PH
486 domain of OPA1, which are also common areas for missense variants in patients with
487 DOA.[70] Additionally, many of the detected phosphorylation sites have not been previously
488 reported in mouse, among which is the novel S298 site within the first GTP-binding
489 domain (G1/P-loop) of the GTPase (**Figure 7D**). The consensus motif GxxxxGKS/T of
490 G1/P-loop is highly conserved across species and other proteins of the dynamin
491 superfamily (**Figure 7E**). Further, the Q and S residue within this motif are important for
492 the assembly stimulated GTPase activity as shown for dynamin.[71] Notably, S298 has
493 been identified as a site of missense mutations in DOA patients (c.893G>A, p.S298N[72] ;
494 c.892A>G, p.S298G[73]), suggesting its importance for OPA1 function, too. Furthermore, a
495 S298A mutation reduces GTPase activity and dimerization *in vitro*[74]; the S298N mutation
496 abolishes respiratory growth in *S.cerevisiae* with mitochondrial DNA depletion and altered
497 mitochondrial morphology.[75] Our structural modelling of the OPA1's GTPase domain
498 shows that S298 is close to the nucleotide-interacting G300 of the G1/P-loop and the
499 nucleotide, suggesting that phosphorylation may affect nucleotide binding or hydrolysis
500 (**Figure 7F**). Taken together, current evidence clearly underscores the functional
501 importance of S298 for OPA1 activity, but the effect and relevance of phosphorylation we
502 have identified on S298 requires further study.
503
504 **DISCUSSION**
505 Here we demonstrated that the Orbitrap Astral mass spectrometer provides a platform for
506 fast and deep phosphoproteome analysis. Key to empowering this application is the ability
507 of the system's high MS/MS duty cycle to enable DIA data collection with narrow isolation
508 windows (2 Th). This capability, combined with high mass accuracy, mass resolution, and
509 sensitive ion detection of the Astral analyzer allow confident phosphosite localization from
510 low abundance peptides as evinced by the performance of the system for low overall
511 phosphopeptide loads. Altogether this approach allowed for the collection of the deepest
512 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, ubquitinylation, 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.

COMPETING INTERESTS

The authors declare the following 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.

ACKNOWLEDGEMENTS

We are grateful for support from the National Institutes of Health (grants P41GM108538 and R35GM118110 to JJC, R01DK098672 to DJP, R35GM150899 to AG, and R35GM147014 to JF), the National Science Foundation grant 2010789, and the Department of Energy grant number DE-SC0018409. 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.

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.

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 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 were cultured in Dulbecco's Modified Essential Medium (DMEM) (Gibco, 11995-065) with 10% FBS (HyClone, 89133-098) and 1% penicillin/streptomycin (Thermo

591 Fisher Scientific, 15-140-122) in 37°C/5% CO2 incubator. Cells were resuspended in PBS
592 (Gibco, 02-0119-0500) with or without 100 ng/mL human Endothelial Growth Factor
593 (Thermo, AF-100-15) for 15 minutes at room temperature. Cells were gently resuspended
594 every 5 minutes during incubation. After incubation, cells were washed twice with PBS
595 and stored in -80°C until use. Protein extraction was performed as described above for
596 the HEK293T samples. Protein concentrations were estimated via protein BCA (Pierce,
597 23235).
598
599 **Yeast Cell Preparation**
600 *Saccharomyces cerevisiae S288C-derivative strain, BY4741,* was cultured in triplicate for
601 ~5 generations into log phase in rich YPD medium at 30°C. Cells were centrifuged at
602 3000 RPM, 3 min, rinse in water and snap-freeze in liquid nitrogen. Frozen cell pellets
603 were resuspended in 8M urea Sigma-Aldrich, U5378) in 100 mM Tris, pH 8 (Invitrogen, 1
604 M Tris pH 8.0, 0.2 µm filtered, AM9856) and vortexed with glass beads (425-600 µm,
605 Sigma-Aldrich, G8772-500G) to lyse (2 minutes of total vortexing with 30s vortexing
606 followed by 30s on ice). Protein concentrations were estimated via protein BCA (Pierce,
607 23235).
608
609 **Mouse Tissue Preparation**
610 All experiments were performed in accordance with the National Institute of Health Guide
611 for the Care and Use of Laboratory Animals and were approved by the Animal Care and
612 Use Committee at the University of Wisconsin-Madison. Six-week-old male C57BL6/J
613 mice (n=3) were euthanized by cervical dislocation and tissues were immediately
614 collected, and flash frozen in liquid nitrogen. A total of twelve tissues (pancreas, small
615 intestine, spleen, liver, kidney, testes, heart, lung, subcutaneous white adipose tissue
616 (WAT), brown adipose tissue (BAT), gastrocnemius, and brain) were collected. All tissues
617 were stored at -80°C prior to cryo-pulverization. For each tissue, samples from three mice
618 were pulverized together into a fine power under liquid nitrogen. For each pulverized
619 tissue, ~60mg frozen wet weight was resuspended in 4 mL of 5.4M guanidine
620 hydrocholoride (Sigma Life Science, 8M, pH 8.5, G7294-100mL) in 100 mM Tris, pH 8
621 (Invitrogen, 1 M Tris pH 8.0, 0.2 µm filtered, AM9856) with Pierce™ Phosphatase Inhibitor
622 Mini Tablets (A32957) with one tablet/10mL. Tissue samples were vortexed and
623 sonicated for 20 min in a bath sonicator (chilled to 4°C) to homogenize. A probe sonicator
624 was used briefly to sonicate samples on ice as needed. The protein concentration was
625 estimated via protein BCA (Pierce, 23235) and samples were diluted with the guanidine
626 hydrochloride buffer above to give a protein concentration of ~2mg/mL prior to digestion.
627
628 **Protein digestion**
629 After extracting proteins from human cells, yeast cells, or mouse tissues, methanol
630 (Optima LC/MS grade, Fisher Scientific) was added to 90% (v/v) to precipitate protein

631 and samples were vortexed prior to centrifugation at 4000g for 15 minutes. The
632 supernatant was removed, and the pellet was resuspended in 8M urea (Sigma-Aldrich,
633 U5378), 100mM Tris (Invitrogen, 1 M Tris pH 8.0, 0.2 µm filtered, AM9856), 10mM TCEP
634 (Sigma-Aldrich, C4706-2G), 40mM 2-chloroacetamide (Sigma Aldrich, ≥98%, C0267-
635 100G) pH 8 at ~1.5mg protein/mL. Lysyl Endopeptidase (LysC, 100369-826, VWR) was
636 added at a ratio of 1:50 enzyme:protein and gently rocked at ambient temperature for 4
637 hours, followed by the dilution of the solution to 2M urea with 100mM Tris, pH 8
638 (Invitrogen, 1 M Tris pH 8.0, 0.2 µm filtered, AM9856). Promega Sequencing Grade
639 Modified Trypsin (V5113) was added at a ratio of 1:50 enzyme:protein and incubated
640 overnight. Following overnight digestion, the solution was acidified to <pH 2 with 10%
641 trifluoroacetic acid (Sigma-Aldrich, HPLC grade, >99.9%) to quench the digestion. The
642 sample was then centrifuged at 4000g for 10 minutes to remove particulate matter prior
643 to desalting with a Strata-X 33 µm polymeric reversed phase SPE cartridge. Peptides
644 were dried via a SpeedVac (Thermo Scientific) and stored at -80°C until phosphopeptide
645 enrichment or, in the case of the mouse proteomics experiments, until fractionation.
646
647 **Phosphopeptide enrichment**
648 Phosphopeptides were enriched from digested peptides using MagReSyn Ti-IMAC HP
649 beads (ReSyn Biosciences, MR-THP005). A volume of 100µL beads were used per 1mg
650 of peptides. Input peptide masses of 2-3mg were utilized for human and yeast
651 enrichments, and input peptide masses ranging from ~0.3 mg to 1.2 mg were utilized for
652 the different mouse tissues. Beads were washed three times with 1 mL 80% acetonitrile
653 (Optima LC-MS grade, Fisher Scientific)/6% trifluoracetic acid (Sigma-Aldrich, HPLC
654 grade, >99.9%) prior to resuspending the sample in 1 mL 80% acetonitrile (Optima LC-
655 MS grade, Fisher Scientific)/6% trifluoracetic acid (Sigma-Aldrich, HPLC grade, >99.9%)
656 and vortexing the sample with the beads for 1 hour. After the 1-hour of vortexing, the
657 beads were washed three times with 1mL 80% acetonitrile (Optima LC-MS grade, Fisher
658 Scientific) /6% trifluoracetic acid (Sigma-Aldrich, HPLC grade, >99.9%), once with 1 mL
659 80% acetonitrile, once with 1 mL 80% acetonitrile (Optima LC-MS grade, Fisher Scientific)
660 /0.5 M glycolic acid (Sigma-Aldrich, 99%, 124737-500G), and three times with 1 mL 80%
661 acetonitrile (Optima LC-MS grade, Fisher Scientific). The phosphopeptides were eluted
662 from the beads with the addition of 300 µL 50% acetonitrile (Optima LC-MS grade, Fisher
663 Scientific)/1% ammonium hydroxide (28% in H_2O, ≥99.99% trace metals basis, Sigma-
664 Aldrich), followed by a second elution with another 300 µL 50% acetonitrile/1%
665 ammonium hydroxide. The samples were acidified via addition of 15 µL 10%
666 trifluoroacetic acid (Sigma-Aldrich, HPLC grade, >99.9%). The samples were then dried
667 down in a SpeedVac (Thermo Scientific) prior to being resuspended in 0.2%
668 trifluoroacetic acid (Sigma-Aldrich, HPLC grade, >99.9%) and desalted as previously
669 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, 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 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.

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. A summary of all the phosphopeptide standards used in the analysis is provided in **Supplementary Table 1**.

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.[39] 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 the 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 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.

748

749 **LC-MS Operation for Benchmarking on the Orbitrap Ascend**

750 Nanoflow capillary columns (75µm I.D., 360µm O.D.) with pulled nanoESI emitters were
751 packed to 40cm at high pressures with C18 1.7 µm diameter, 130 Å pore size BEH C18
752 particles (Waters) as previously described.[39] Samples were analyzed with a Vanquish
753 Neo UHPLC (Thermo Scientific) coupled to an Orbitrap Ascend mass spectrometer
754 (Thermo Scientific) using a NanoSpray Flex source (Thermo Scientific) incorporating a
755 homebuilt column heating compartment. The column was heated to 50°C. The flow rate
756 was set to 400nL/min at the start of the method to decrease delay time and turned to
757 300nL/min at the start of the active gradient. Initial conditions of 2%B were ramped to
758 14% from 0 to 5 minutes. The active gradient was generally set to 14% to 54%B with
759 curve type 6 beginning at 5.2 minutes, with the exact %B settings adjusted for each active
760 gradient length to evenly distribute peptide signal across the gradient. The column was
761 washed for 5 minutes at 100%B and 400 nL/min at the end of the gradient, followed by
762 fast equilibration on the Vanquish Neo LC with an upper pressure limit of 1100 bar.

763

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

773

774 **LC-MS Operation for Proteomics with Orbitrap Eclipse Analysis**

775 The chromatography setup described above for the Orbitrap Ascend analysis was utilized
776 for the Orbitrap Eclipse experiments. A source voltage of 2000 V was used for all
777 experiments. Mobile phase A and B were 0.2% formic acid (Optima LC-MS grade, Fisher
778 Scientific) in water (Optima LC-MS grade, Fisher Scientific) and 0.2% formic acid/80%
779 acetonitrile (Optima LC-MS grade, Fisher Scientific), respectively. The column was
780 heated to 50°C and a flow rate of 300 nL/min was used. Initial conditions of 0%B were
781 ramped to 6%B from 0 to 1 minute. The active gradient was set to 6% to 52%B with curve
782 type 6 from 1 to 73 minutes. The column was washed for 5 minutes at 100%B, followed
783 by fast equilibration on the Vanquish Neo LC with an upper pressure limit of 1150 bar.

784

785 For DDA experiments on the Orbitrap Eclipse MS, MS^1 spectra were collected in the
786 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.[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. Note that the entrapment search in Figure 2G and the phosphoproline variable modification search in Figure 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.

826
827 HEK293T phosphoproteomics DIA data was also processed in Proteome Discoverer
828 3.1.0.638 with CHIMERYS. Numbers of localized phosphosites were reported from the
829 Modification Sites table after filtering for rows with 'Phospho' in the Modification Name
830 column. These searches were performed in a developmental version of Proteome
831 Discoverer 3.1 to show that the phosphoproteomic depth achieved with our methods is
832 not restricted to Spectronaut searches. However, this developmental version does not yet
833 incorporate quantitative values for phosphorylation sites detected with DIA methods.
834 Furthermore, from internal conversations, we know the results of this development build
835 should be treated as preliminary and that the algorithms being utilized are still in active
836 development. Consequently, we chose to perform our analyses using Spectronaut, which
837 provides identification, localization, and quantification for DIA phosphoproteomics.
838
839 **Localization Error Rate Calculation**
840 To calculate the localization error rate, a reference sheet of all the possible precursors
841 that could be detected from the phosphopeptide standards was constructed with the
842 phosphorylation sites specified. Precursors with sequences that could also arise from
843 the yeast proteome (based on an *in silico* digest) were removed from consideration.
844 Then, the error rate was calculated as $\text{Error Rate } (\%) = \frac{\text{False Sites}}{\text{False Sites} + \text{True Sites}} \times 100$, where
845 'False Sites' is the number of phosphopeptide standard precursors detected with
846 phosphorylation states not indicated in the reference sheet, and 'True Sites' is the
847 number of phosphopeptide standard precursors detected with phosphorylation states
848 indicated in the reference sheet. This calculation was performed for each
849 phosphopeptide standard raw data file, and the average error rates are reported as a
850 function of the localization probability cutoff.
851
852 **Phosphopeptide Standard Quantification Calculations**
853 The phosphopeptide standard reference sheet described above was modified so that
854 each row corresponds to a phosphorylation site, so that the 'PTM.Quantity' value in the
855 PTM Site Report could be utilized for quantification assessment. As described above,
856 phosphopeptides with sequences that could arise from the yeast proteome were
857 removed from consideration. Linear regression and R^2 calculations were performed
858 using the 'pearsonr' and 'linregress' functions within the 'scipy.stats' module for
859 phosphosites detected across at least three concentration points.[76]
860
861 **EGF-stimulated HeLa Differential Expression and Pathway Analysis**
862 Differential expression analysis for the EGF treatment experiment was performed on all
863 phosphosites that were detected across triplicates in either the EGF-treatment or control

group. Missing value imputation was performed in Perseus[77] 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.

Pathway enrichment analysis was performed using Enrichr.[78] 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.[79] 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.[77]

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

903 Annotation of human variants in CPS1 and OPA1 were identified using The Human
904 Gene Mutation Database at the Institute of Medical Genetics in Cardiff (HGMD) using
905 the public site entries (retrieval date 2023-11-18 [80]).
906
907 **Multiple Sequence Alignment**
908 Clustal Omega[81] hosted at the EMBL-EBI webserver was used for protein sequence
909 alignment using the following UniProt identifiers for CPS1: P08686, F7EZ, P07756,
910 Q8C196, F1ML89, A0A8C0NKH0; for OPA1: P58281, P58281-2, O60313, P32266; and
911 other dynamin family members: Q811U4, Q80U63, Q8K1M6 (retrieved from UniProt
912 2023-11-18).
913
914 **Structural modelling prediction**
915 The PyMOL Molecular Graphics System, Version 2.5.7 Schrödinger, LLC. was used for
916 predicting structural models for CPS1 (RCSB PDB: 5DOT (Apo), 5DOU (NAG)[66]) and
917 OPA1 (RCSB PDB structure 6JTG [74]). For modeling purposes, the human crystal
918 structures (highly homologous to mouse) were used.
919
920 **Code Availability**
921 All code will be available at the time of publication.
922
923 **Data Availability**
924 All data will be available at the time of publication.
925

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) 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 Communications in Mass Spectrometry* **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. *Proceedings of the National Academy of Sciences* **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*. *Molecular & Cellular 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. *Proceedings of the National Academy of Sciences* **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. *Proceedings of the National Academy of Sciences* **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. *Molecular & Cellular 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. *Proceedings of the National Academy of Sciences* **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 Reports* **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. *Molecular & Cellular 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.

1064 (35) Bludau, I.; Willems, S.; Zeng, W.-F.; Strauss, M. T.; Hansen, F. M.; Tanzer, M.
1065 C.; Karayel, O.; Schulman, B. A.; Mann, M. The Structural Context of
1066 Posttranslational Modifications at a Proteome-Wide Scale. *PLoS Biol* **2022**, *20* (5),
1067 e3001636. https://doi.org/10.1371/journal.pbio.3001636.
1068 (36) Tyanova, S.; Cox, J.; Olsen, J.; Mann, M.; Frishman, D. Phosphorylation
1069 Variation during the Cell Cycle Scales with Structural Propensities of Proteins. *PLoS*
1070 *Comput Biol* **2013**, *9* (1), e1002842. https://doi.org/10.1371/journal.pcbi.1002842.
1071 (37) Iakoucheva, L. M.; Radivojac, P.; Brown, C. J.; O'Connor, T. R.; Sikes, J. G.;
1072 Obradovic, Z.; Dunker, A. K. The Importance of Intrinsic Disorder for Protein
1073 Phosphorylation. *Nucleic Acids Res* **2004**, *32* (3), 1037–1049.
1074 https://doi.org/10.1093/nar/gkh253.
1075 (38) Johnson, J. L.; Yaron, T. M.; Huntsman, E. M.; Kerelsky, A.; Song, J.; Regev, A.;
1076 Lin, T.-Y.; Liberatore, K.; Cizin, D. M.; Cohen, B. M.; Vasan, N.; Ma, Y.; Krismer, K.;
1077 Robles, J. T.; van de Kooij, B.; van Vlimmeren, A. E.; Andrée-Busch, N.; Käufer, N.
1078 F.; Dorovkov, M. V.; Ryazanov, A. G.; Takagi, Y.; Kastenhuber, E. R.; Goncalves,
1079 M. D.; Hopkins, B. D.; Elemento, O.; Taatjes, D. J.; Maucuer, A.; Yamashita, A.;
1080 Degterev, A.; Uduman, M.; Lu, J.; Landry, S. D.; Zhang, B.; Cossentino, I.; Linding,
1081 R.; Blenis, J.; Hornbeck, P. V.; Turk, B. E.; Yaffe, M. B.; Cantley, L. C. An Atlas of
1082 Substrate Specificities for the Human Serine/Threonine Kinome. *Nature* **2023**, *613*
1083 (7945), 759–766. https://doi.org/10.1038/s41586-022-05575-3.
1084 (39) Shishkova, E.; Hebert, A. S.; Westphall, M. S.; Coon, J. J. Ultra-High Pressure
1085 (>30,000 Psi) Packing of Capillary Columns Enhancing Depth of Shotgun Proteomic
1086 Analyses. *Anal. Chem.* **2018**, *90* (19), 11503–11508.
1087 https://doi.org/10.1021/acs.analchem.8b02766.
1088 (40) Cox, J.; Mann, M. MaxQuant Enables High Peptide Identification Rates,
1089 Individualized p.p.b.-Range Mass Accuracies and Proteome-Wide Protein
1090 Quantification. *Nat Biotechnol* **2008**, *26* (12), 1367–1372.
1091 https://doi.org/10.1038/nbt.1511.
1092 (41) Bekker-Jensen, D. B.; Bernhardt, O. M.; Hogrebe, A.; Martinez-Val, A.; Verbeke,
1093 L.; Gandhi, T.; Kelstrup, C. D.; Reiter, L.; Olsen, J. V. Rapid and Site-Specific Deep
1094 Phosphoproteome Profiling by Data-Independent Acquisition without the Need for
1095 Spectral Libraries. *Nat Commun* **2020**, *11* (1), 787. https://doi.org/10.1038/s41467-
1096 020-14609-1.
1097 (42) Kitata, R. B.; Choong, W.-K.; Tsai, C.-F.; Lin, P.-Y.; Chen, B.-S.; Chang, Y.-C.;
1098 Nesvizhskii, A. I.; Sung, T.-Y.; Chen, Y.-J. A Data-Independent Acquisition-Based
1099 Global Phosphoproteomics System Enables Deep Profiling. *Nat Commun* **2021**, *12*
1100 (1), 2539. https://doi.org/10.1038/s41467-021-22759-z.
1101 (43) Srinivasan, A.; Sing, J. C.; Gingras, A.-C.; Röst, H. L. Improving
1102 Phosphoproteomics Profiling Using Data-Independent Mass Spectrometry. *J.*
1103 *Proteome Res.* **2022**, *21* (8), 1789–1799.
1104 https://doi.org/10.1021/acs.jproteome.2c00172.
1105 (44) Oliinyk, D.; Meier, F. Ion Mobility-Resolved Phosphoproteomics with Dia-PASEF
1106 and Short Gradients. *PROTEOMICS* **2023**, *23* (7–8), 2200032.
1107 https://doi.org/10.1002/pmic.202200032.
1108 (45) Lou, R.; Cao, Y.; Li, S.; Lang, X.; Li, Y.; Zhang, Y.; Shui, W. Benchmarking
1109 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*. *Molecular & Cellular 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 Research* **2021**, *81* (9), 2495–2509. https://doi.org/10.1158/0008-5472.CAN-20-3804.

(51) 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.

(52) 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.

(53) 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.

(54) 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.

(55) 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.

(56) 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.

(57) 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* **2019**, *39* (28), 5424–5439. https://doi.org/10.1523/JNEUROSCI.1341-18.2019.

(58) 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.

(59) 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.

(60) 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.

(61) 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.

1201 (62) Niemi, N. M.; Pagliarini, D. J. The Extensive and Functionally Uncharacterized
1202 Mitochondrial Phosphoproteome. *J Biol Chem* **2021**, *297* (1), 100880.
1203 https://doi.org/10.1016/j.jbc.2021.100880.
1204 (63) Arnold, P. K.; Finley, L. W. S. Regulation and Function of the Mammalian
1205 Tricarboxylic Acid Cycle. *J Biol Chem* **2023**, *299* (2), 102838.
1206 https://doi.org/10.1016/j.jbc.2022.102838.
1207 (64) Häberle, J.; Burlina, A.; Chakrapani, A.; Dixon, M.; Karall, D.; Lindner, M.;
1208 Mandel, H.; Martinelli, D.; Pintos-Morell, G.; Santer, R.; Skouma, A.; Servais, A.; Tal,
1209 G.; Rubio, V.; Huemer, M.; Dionisi-Vici, C. Suggested Guidelines for the Diagnosis
1210 and Management of Urea Cycle Disorders: First Revision. *J Inherit Metab Dis* **2019**,
1211 *42* (6), 1192–1230. https://doi.org/10.1002/jimd.12100.
1212 (65) Häberle, J.; Shchelochkov, O. A.; Wang, J.; Katsonis, P.; Hall, L.; Reiss, S.;
1213 Eeds, A.; Willis, A.; Yadav, M.; Summar, S.; Urea Cycle Disorders Consortium;
1214 Lichtarge, O.; Rubio, V.; Wong, L.-J.; Summar, M. Molecular Defects in Human
1215 Carbamoy Phosphate Synthetase I: Mutational Spectrum, Diagnostic and Protein
1216 Structure Considerations. *Hum Mutat* **2011**, *32* (6), 579–589.
1217 https://doi.org/10.1002/humu.21406.
1218 (66) de Cima, S.; Polo, L. M.; Díez-Fernández, C.; Martínez, A. I.; Cervera, J.; Fita, I.;
1219 Rubio, V. Structure of Human Carbamoyl Phosphate Synthetase: Deciphering the
1220 on/off Switch of Human Ureagenesis. *Sci Rep* **2015**, *5*, 16950.
1221 https://doi.org/10.1038/srep16950.
1222 (67) Pekkala, S.; Martínez, A. I.; Barcelona, B.; Yefimenko, I.; Finckh, U.; Rubio, V.;
1223 Cervera, J. Understanding Carbamoyl-Phosphate Synthetase I (CPS1) Deficiency
1224 by Using Expression Studies and Structure-Based Analysis. *Hum Mutat* **2010**, *31*
1225 (7), 801–808. https://doi.org/10.1002/humu.21272.
1226 (68) Takakusa, H.; Mohar, I.; Kavanagh, T. J.; Kelly, E. J.; Kaspera, R.; Nelson, S. D.
1227 Protein Tyrosine Nitration of Mitochondrial Carbamoyl Phosphate Synthetase 1 and
1228 Its Functional Consequences. *Biochem Biophys Res Commun* **2012**, *420* (1), 54–
1229 60. https://doi.org/10.1016/j.bbrc.2012.02.114.
1230 (69) Del Dotto, V.; Fogazza, M.; Carelli, V.; Rugolo, M.; Zanna, C. Eight Human OPA1
1231 Isoforms, Long and Short: What Are They For? *Biochim Biophys Acta Bioenerg*
1232 **2018**, *1859* (4), 263–269. https://doi.org/10.1016/j.bbabio.2018.01.005.
1233 (70) Ferré, M.; Caignard, A.; Milea, D.; Leruez, S.; Cassereau, J.; Chevrollier, A.;
1234 Amati-Bonneau, P.; Verny, C.; Bonneau, D.; Procaccio, V.; Reynier, P. Improved
1235 Locus-Specific Database for OPA1 Mutations Allows Inclusion of Advanced Clinical
1236 Data. *Hum Mutat* **2015**, *36* (1), 20–25. https://doi.org/10.1002/humu.22703.
1237 (71) Chappie, J. S.; Acharya, S.; Leonard, M.; Schmid, S. L.; Dyda, F. G Domain
1238 Dimerization Controls Dynamin's Assembly-Stimulated GTPase Activity. *Nature*
1239 **2010**, *465* (7297), 435–440. https://doi.org/10.1038/nature09032.
1240 (72) Santarelli, R.; Rossi, R.; Scimemi, P.; Cama, E.; Valentino, M. L.; La Morgia, C.;
1241 Caporali, L.; Liguori, R.; Magnavita, V.; Monteleone, A.; Biscaro, A.; Arslan, E.;
1242 Carelli, V. OPA1-Related Auditory Neuropathy: Site of Lesion and Outcome of
1243 Cochlear Implantation. *Brain* **2015**, *138* (Pt 3), 563–576.
1244 https://doi.org/10.1093/brain/awu378.
1245 (73) Loo, J. L.; Singhal, S.; Rukmini, A. V.; Tow, S.; Amati-Bonneau, P.; Procaccio,
1246 V.; Bonneau, D.; Gooley, J. J.; Reynier, P.; Ferré, M.; Milea, D. Multiethnic

1247 Involvement in Autosomal-Dominant Optic Atrophy in Singapore. *Eye (Lond)* **2017**,
1248 *31* (3), 475–480. https://doi.org/10.1038/eye.2016.255.
1249 (74) Yu, C.; Zhao, J.; Yan, L.; Qi, Y.; Guo, X.; Lou, Z.; Hu, J.; Rao, Z. Structural
1250 Insights into G Domain Dimerization and Pathogenic Mutation of OPA1. *J Cell Biol*
1251 **2020**, *219* (7), e201907098. https://doi.org/10.1083/jcb.201907098.
1252 (75) Del Dotto, V.; Fogazza, M.; Musiani, F.; Maresca, A.; Aleo, S. J.; Caporali, L.; La
1253 Morgia, C.; Nolli, C.; Lodi, T.; Goffrini, P.; Chan, D.; Carelli, V.; Rugolo, M.; Baruffini,
1254 E.; Zanna, C. Deciphering OPA1 Mutations Pathogenicity by Combined Analysis of
1255 Human, Mouse and Yeast Cell Models. *Biochim Biophys Acta Mol Basis Dis* **2018**,
1256 *1864* (10), 3496–3514. https://doi.org/10.1016/j.bbadis.2018.08.004.
1257 (76) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.;
1258 Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt,
1259 S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.;
1260 Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.;
1261 Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.;
1262 Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P.; SciPy 1.0
1263 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python.
1264 *Nat Methods* **2020**, *17* (3), 261–272. https://doi.org/10.1038/s41592-019-0686-2.
1265 (77) Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M. Y.; Geiger, T.; Mann,
1266 M.; Cox, J. The Perseus Computational Platform for Comprehensive Analysis of
1267 (Prote)Omics Data. *Nat Methods* **2016**, *13* (9), 731–740.
1268 https://doi.org/10.1038/nmeth.3901.
1269 (78) Chen, E. Y.; Tan, C. M.; Kou, Y.; Duan, Q.; Wang, Z.; Meirelles, G. V.; Clark, N.
1270 R.; Ma'ayan, A. Enrichr: Interactive and Collaborative HTML5 Gene List Enrichment
1271 Analysis Tool. *BMC Bioinformatics* **2013**, *14*, 128. https://doi.org/10.1186/1471-
1272 2105-14-128.
1273 (79) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.;
1274 Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.;
1275 Cournapeau, D.; Brucher, M.; Perrot, M.; Duschesnay, E. Scikit-Learn: Machine
1276 Learning in Python. *JMLR* **2011**, *12*, 2825–2830.
1277 (80) Stenson, P. D.; Mort, M.; Ball, E. V.; Chapman, M.; Evans, K.; Azevedo, L.;
1278 Hayden, M.; Heywood, S.; Millar, D. S.; Phillips, A. D.; Cooper, D. N. The Human
1279 Gene Mutation Database (HGMD®): Optimizing Its Use in a Clinical Diagnostic or
1280 Research Setting. *Hum Genet* **2020**, *139* (10), 1197–1207.
1281 https://doi.org/10.1007/s00439-020-02199-3.
1282 (81) Madeira, F.; Pearce, M.; Tivey, A. R. N.; Basutkar, P.; Lee, J.; Edbali, O.;
1283 Madhusoodanan, N.; Kolesnikov, A.; Lopez, R. Search and Sequence Analysis
1284 Tools Services from EMBL-EBI in 2022. *Nucleic Acids Res* **2022**, *50* (W1), W276–
1285 W279. https://doi.org/10.1093/nar/gkac240.
1286 (82) Rath, S.; Sharma, R.; Gupta, R.; Ast, T.; Chan, C.; Durham, T. J.; Goodman, R.
1287 P.; Grabarek, Z.; Haas, M. E.; Hung, W. H. W.; Joshi, P. R.; Jourdain, A. A.; Kim, S.
1288 H.; Kotrys, A. V.; Lam, S. S.; McCoy, J. G.; Meisel, J. D.; Miranda, M.; Panda, A.;
1289 Patgiri, A.; Rogers, R.; Sadre, S.; Shah, H.; Skinner, O. S.; To, T.-L.; Walker, M. A.;
1290 Wang, H.; Ward, P. S.; Wengrod, J.; Yuan, C.-C.; Calvo, S. E.; Mootha, V. K.
1291 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.

(83) 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.

MAIN FIGURES

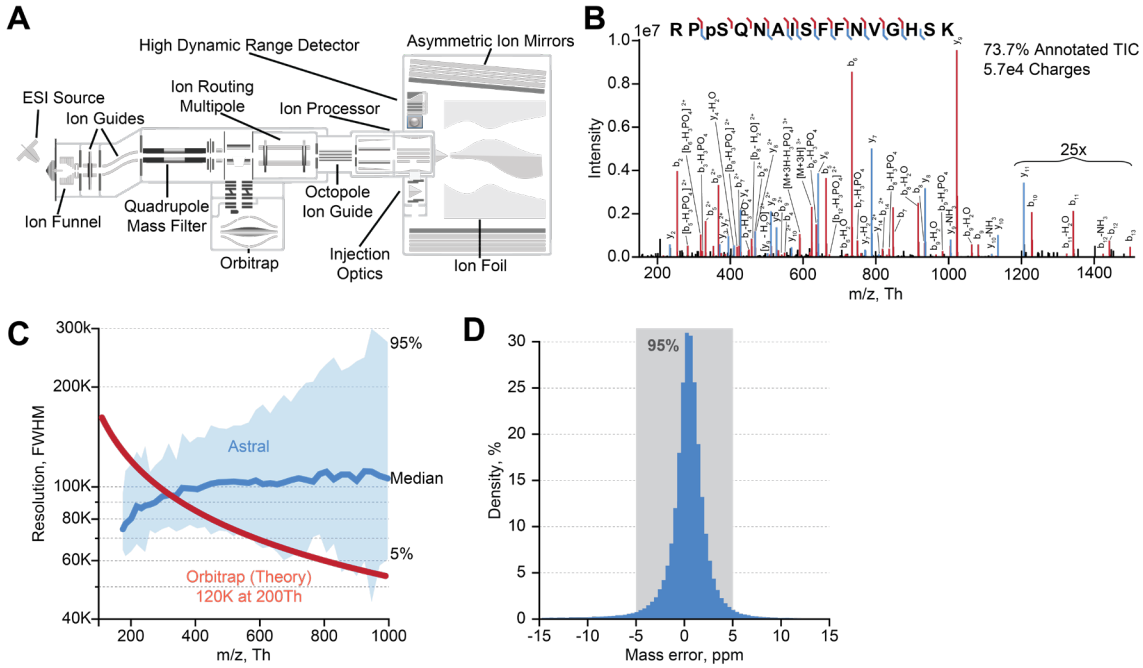


Figure 1. Overview of Orbitrap Astral MS and its key figures of merit. (A) Orbitrap Astral instrument schematic, highlighting the quadrupole, Orbitrap, and Astral analyzers. (B) Tandem mass spectrum of representative phosphopeptide collected using the Astral analyzer. (C) Distribution of Astral analyzer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

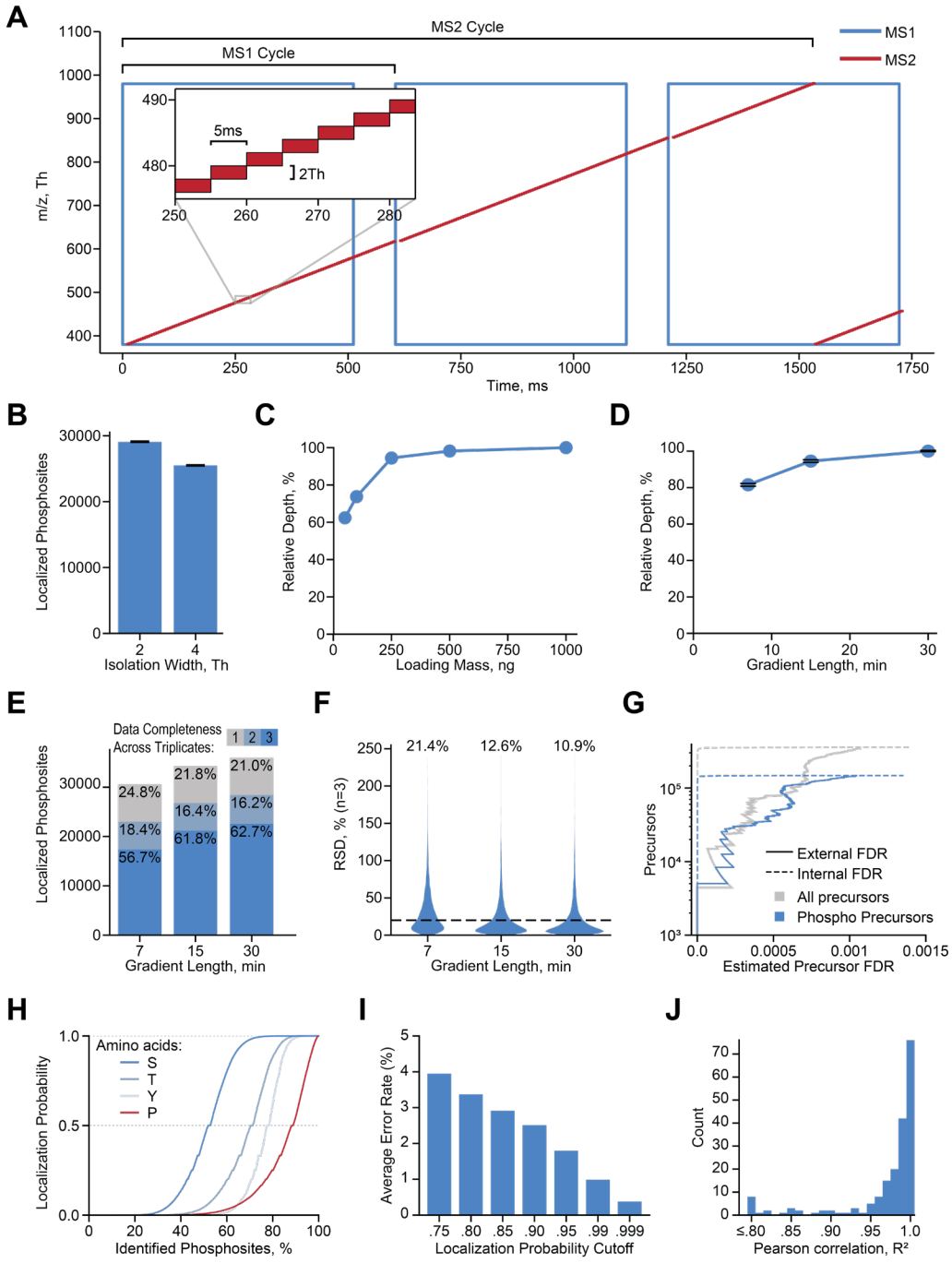


Figure 2. Overview of DIA phosphoproteomics on the Orbitrap Astral MS. (A) Illustration of DIA acquisition scheme. (B) Number of localized phosphorylation sites identified using DIA method with various isolation widths. (C) Evaluation of phosphopeptide loading mass on performance. Note data was collected 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. (I) The average site localization error rate as a function of localization probability cutoff is shown. 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 are shown for phosphopeptides standards detected in at least three concentrations across a five-point dilution series into a constant yeast phosphopeptide background.

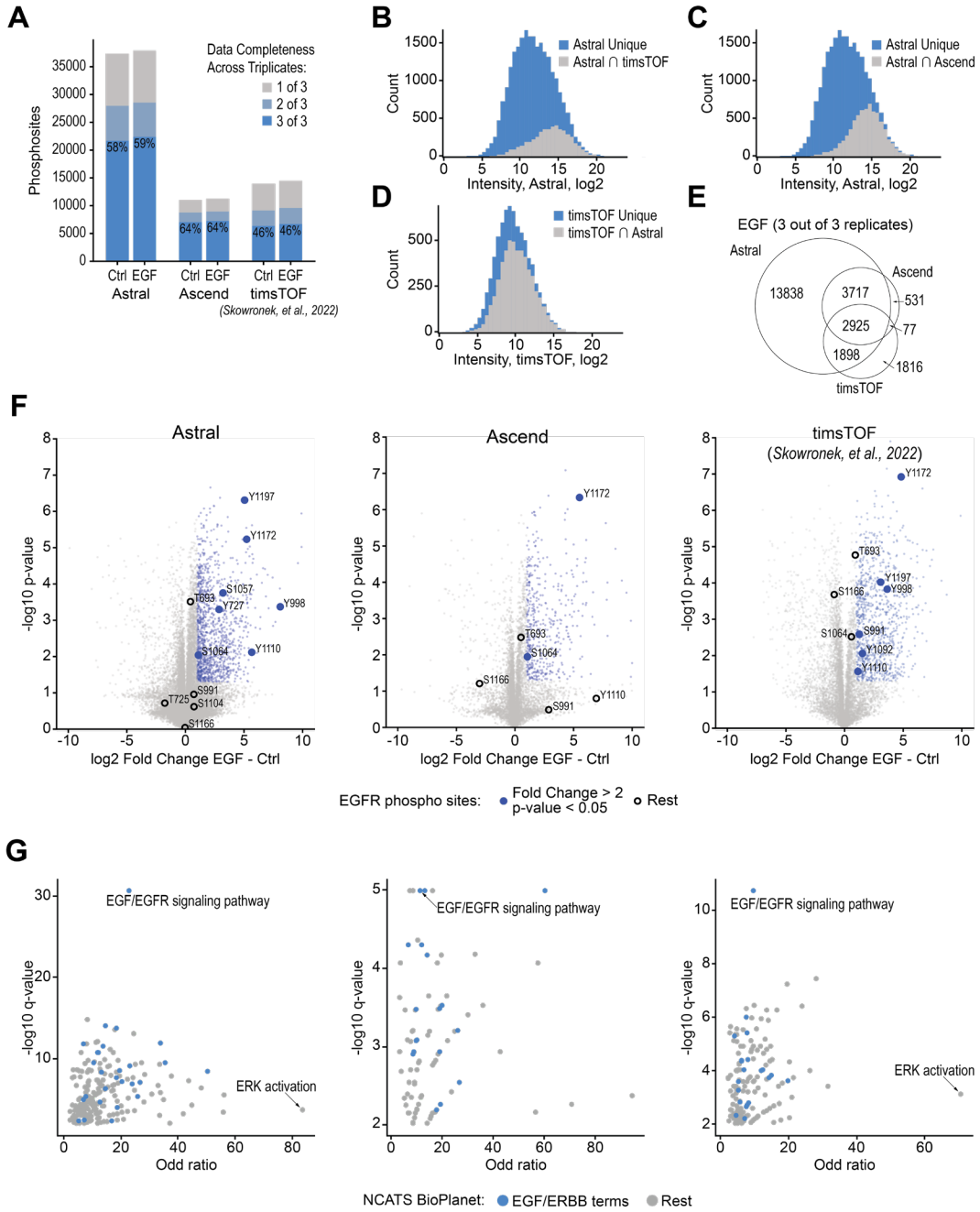


Figure 3. Biological validation of phosphoproteomics platforms using EGF stimulation. (A) Reproducibility of detected phosphosites across biological triplicates for three phosphoproteomics platforms – Orbitrap Astral, Orbitrap Ascend, and timsTOF Pro.[33] Intensity distribution of phosphosites consistently 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) Vulcano plots of phosphosites between EGF stimulation and control across platforms with phosphosites meeting the differential expression criteria (fold-change >2,

1339 p-value < 0.05) indicated in blue and EGFR phosphosites labeled. (G) Pathway
1340 enrichment analysis using the NCATS BioPlanet terms was performed for the three
1341 platforms with EGF/ERBB-associated terms indicated in blue.
1342

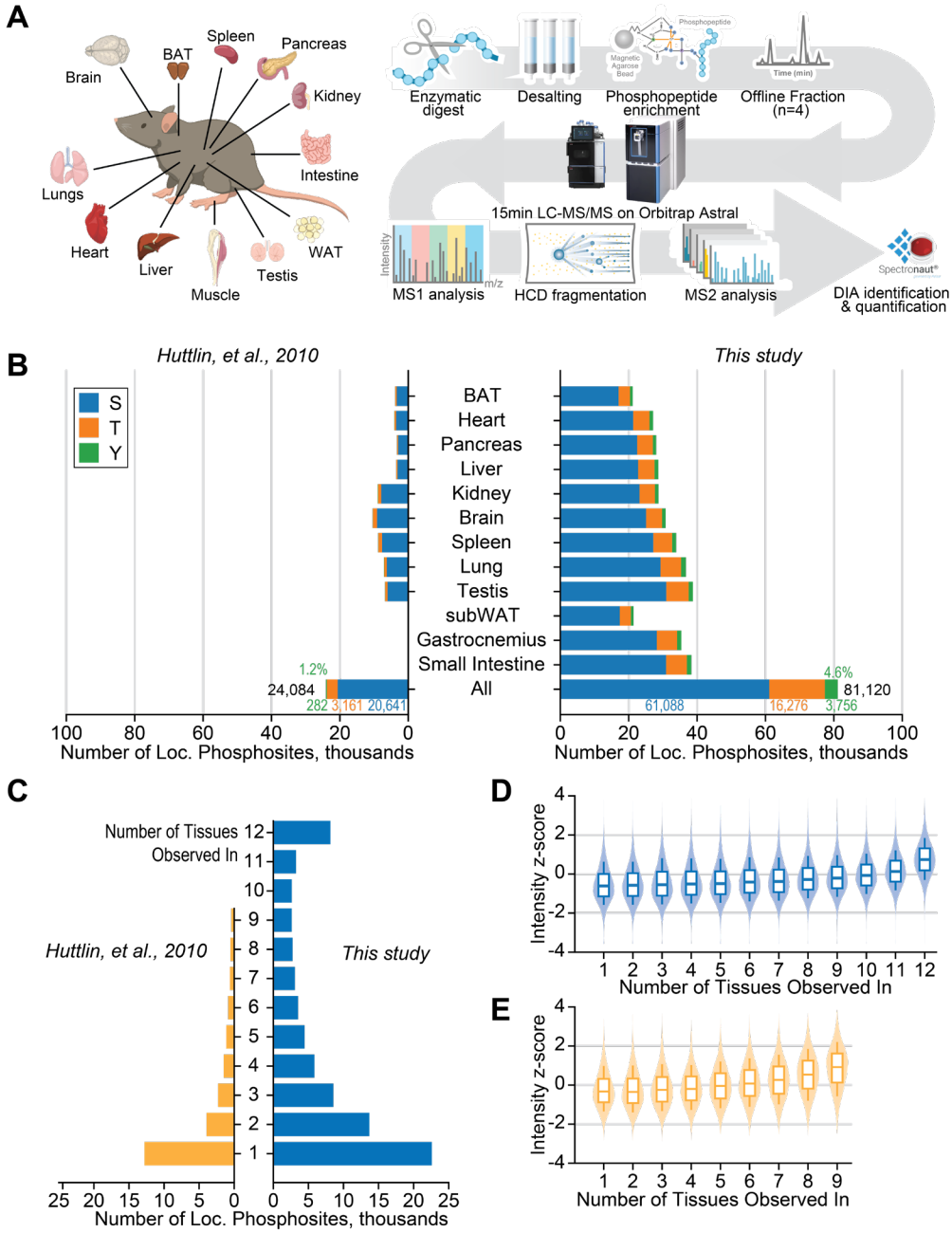


Figure 4. Mouse Phosphorylation Atlas Workflow and Results. (A) Mouse Phosphorylation Atlas Workflow. (B) Result for Mouse Tissue Phosphoproteomic Analysis. Numbers of unique phosphorylation sites are shown for each tissue and the total unique sites with the fraction of S,T, and Y localizations indicated. The Huttlin et al. results 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.

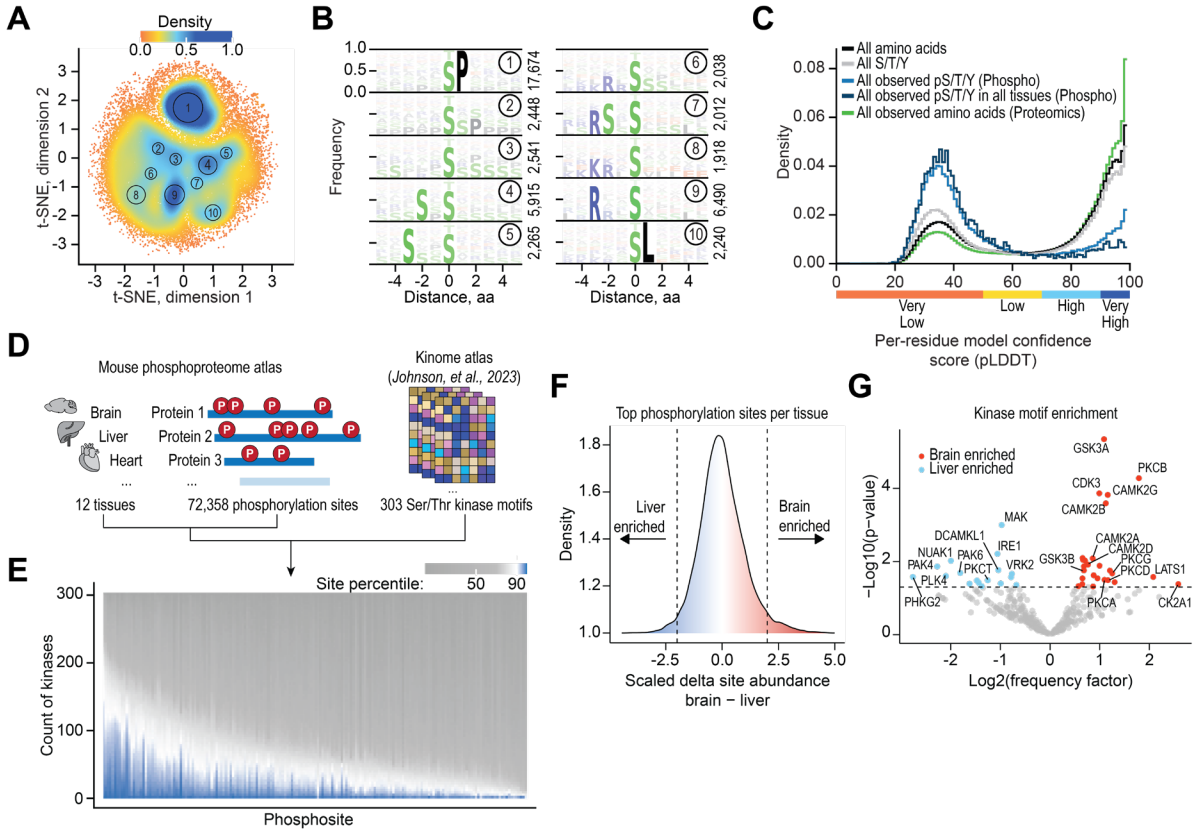


Figure 5. Sequence, Structural and Kinase Specificity Context of Phosphosites. (A) Two-dimensional representation of all phosphosites and their 5-amino acid flanking sequences, excluding the central amino acid from the comparison. Each cluster has been manually selected to emphasize the densest regions. (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. (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.

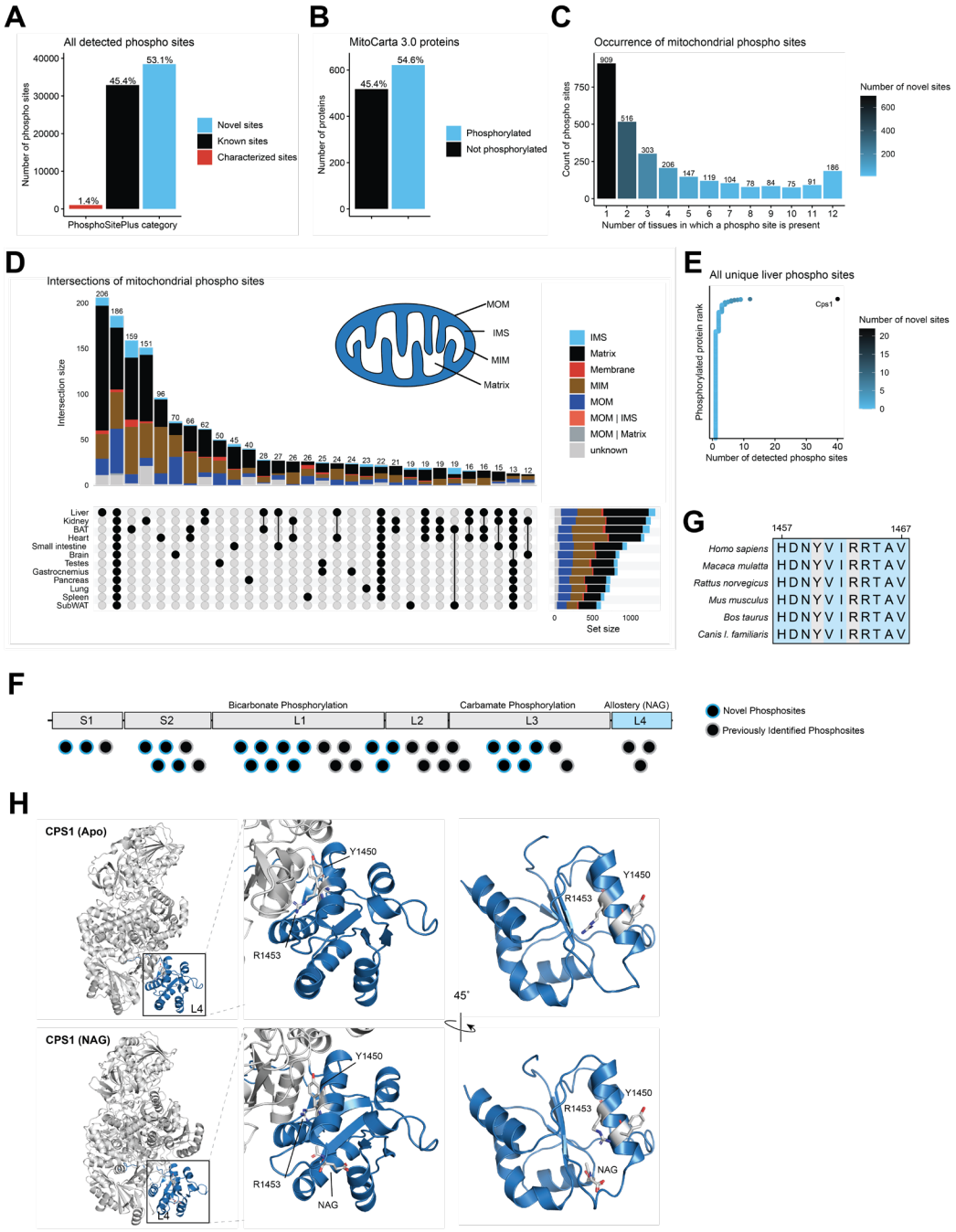


Figure 6. Mitochondrial phosphoproteomics reveals novel liver-specific phosphorylation site. (A) Bar plot of all detected phosphorylation sites in our study stratified into categories that are directly derived from the PhosphoSitePlus database (downloaded August 22, 2023). (B) Mouse MitoCarta 3.0[82] 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

1380 indicated on the x-axis. Novel sites are according to the PhosphoSitePlus database. (F)
1381 Schematic representation of the CPS1 peptide chain with all detected phosphorylation
1382 sites indicated as novel or previously identified. (G) Sequence alignment of CPS1
1383 orthologues using Clustal Omega[81]. Patient variant residue and phosphorylation site
1384 residue of interest are in grey. (H) Structural modeling based on structures (PDB: 5DOT
1385 (Apo), 5DOU (NAG)) from RCSB PDB[66].
1386

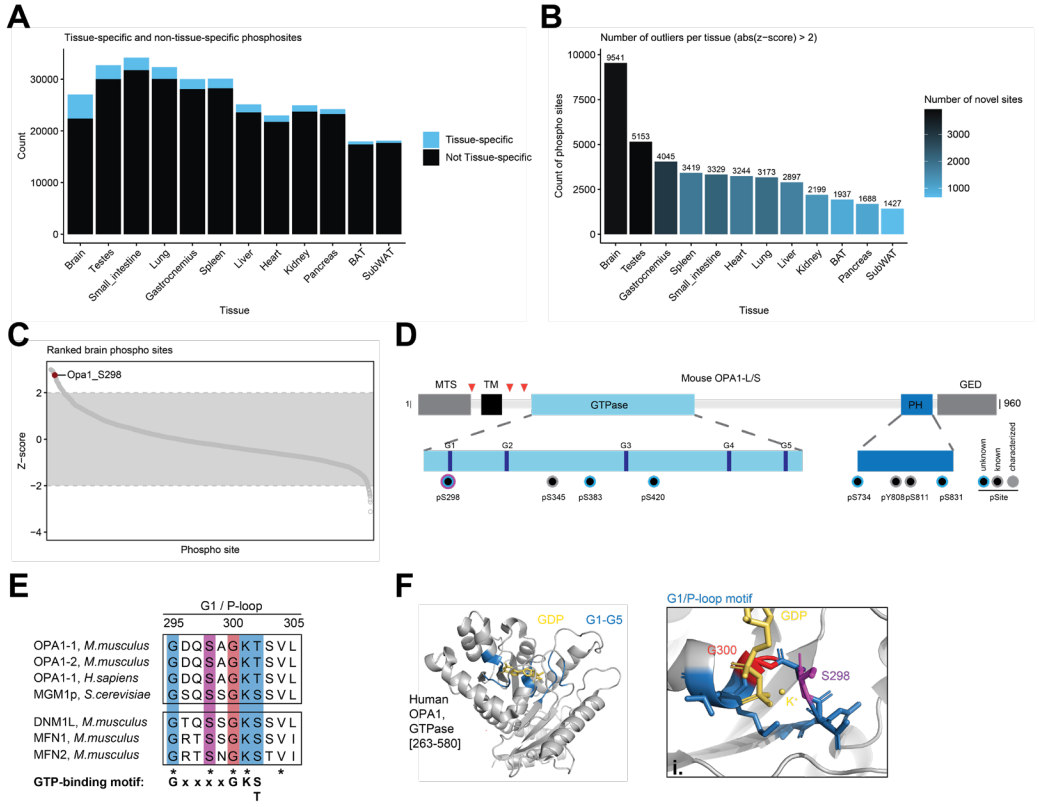
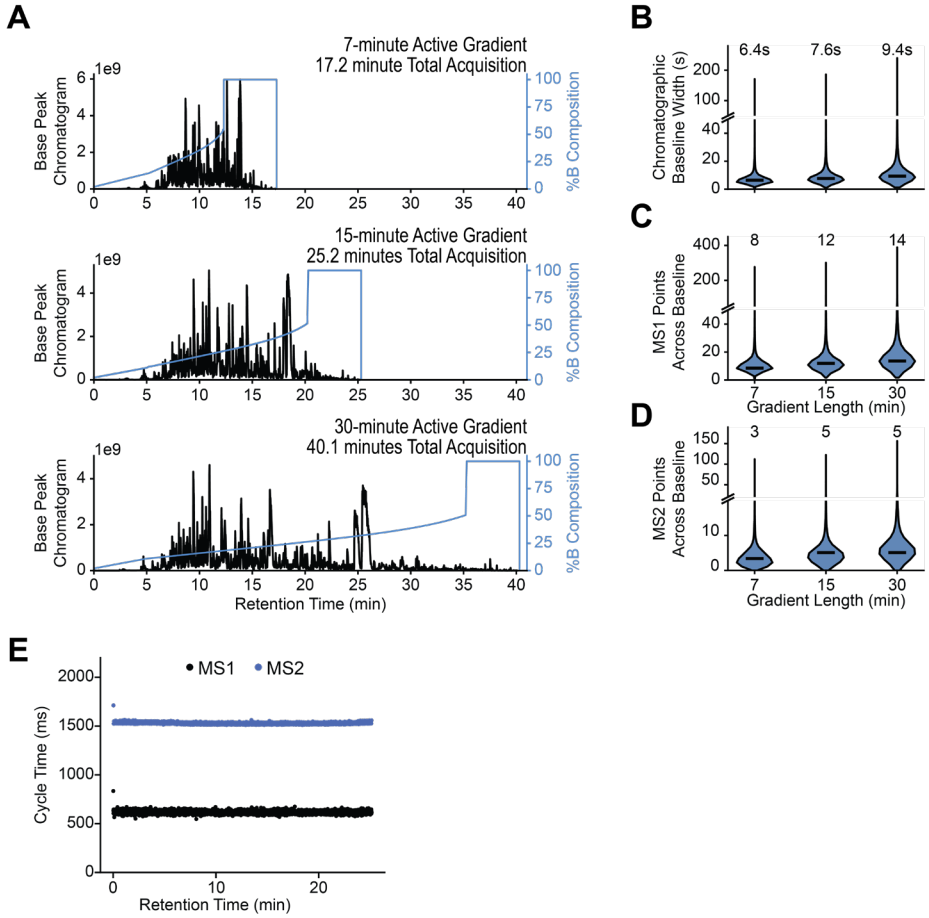
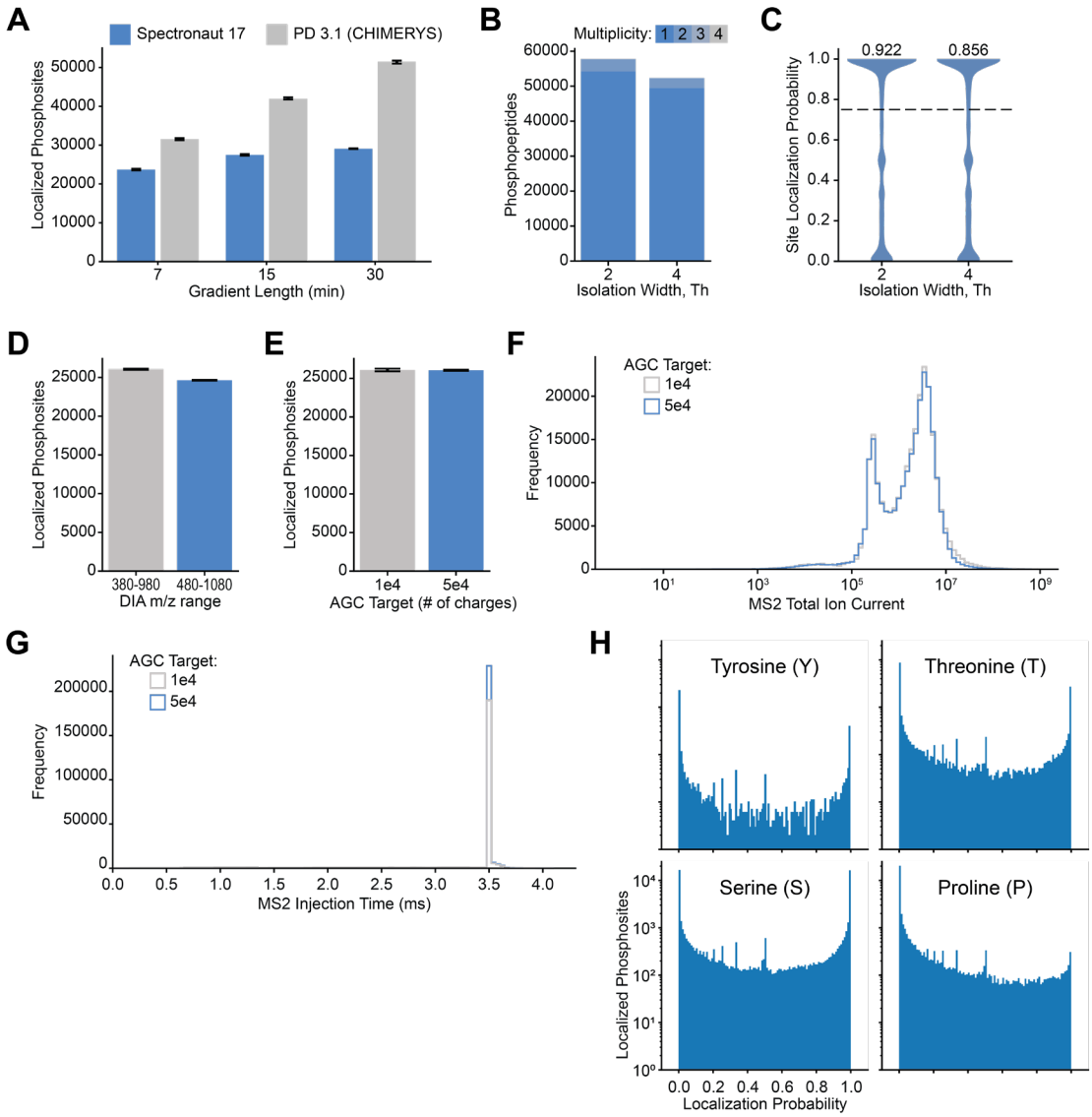


Figure 7. Unique phosphorylation sites in brain tissue. (A) Bar plot indicating all sites (black) that were detected per tissue, highlighting the sites that are unique in each tissue (blue). (B) Number of outliers per tissue based on z-score analysis with an absolute z-score cut-off of 2. (C) Ranked 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[74]. 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.

SUPPLEMENTARY INFORMATION

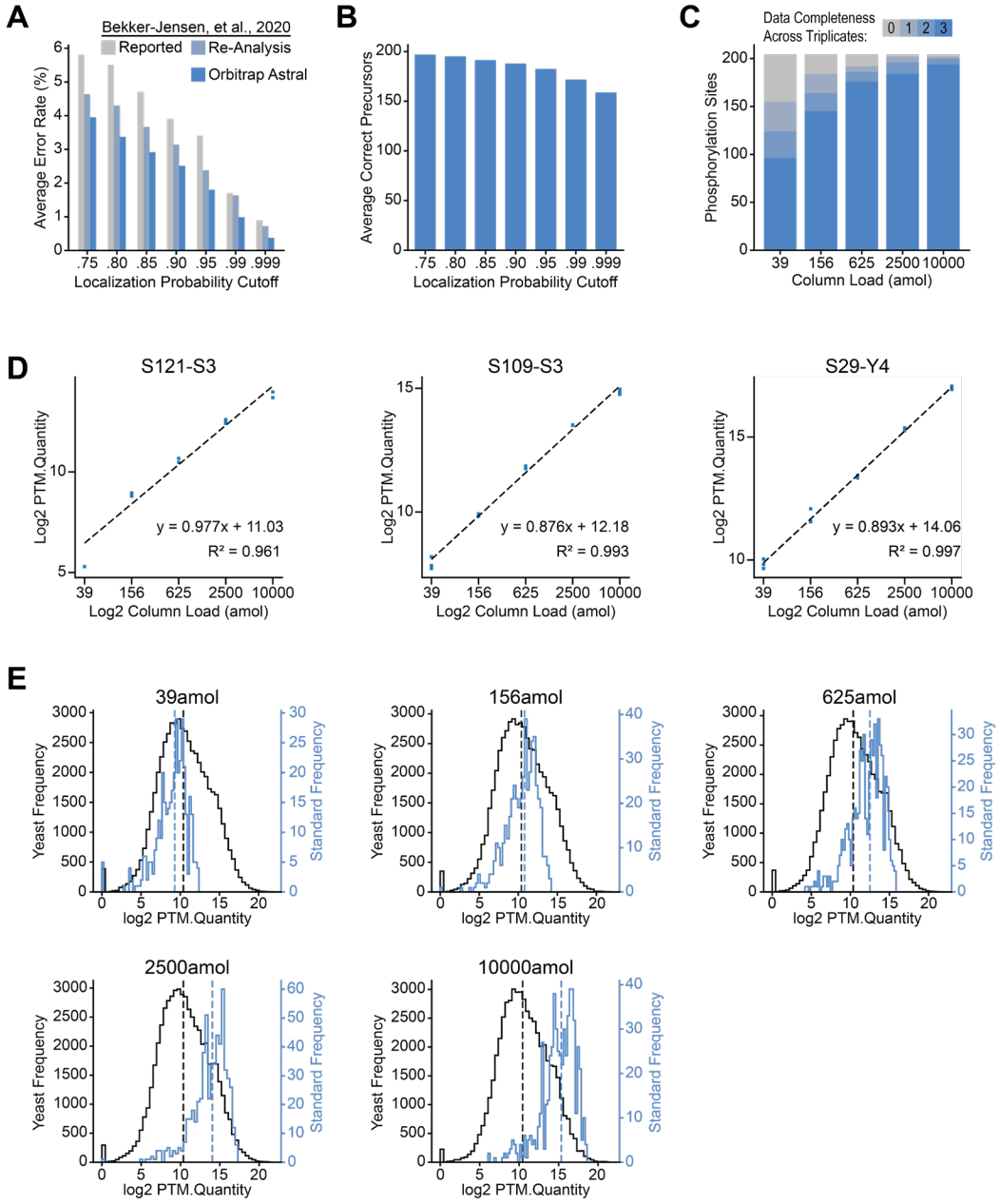


Supplementary Figure 1. Chromatographic Properties of HEK293T Phosphoproteomics Datasets (A) Representative chromatograms for different gradient lengths. The base peak chromatograms for the three different active gradient lengths are shown with the gradient composition curve plotted. The distribution of (B) chromatographic baseline widths, (C) MS1 points across the baseline peak, and (D) 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). (E) The cycle times for the MS1 and MS2 scan sequences are shown as a function of RT.



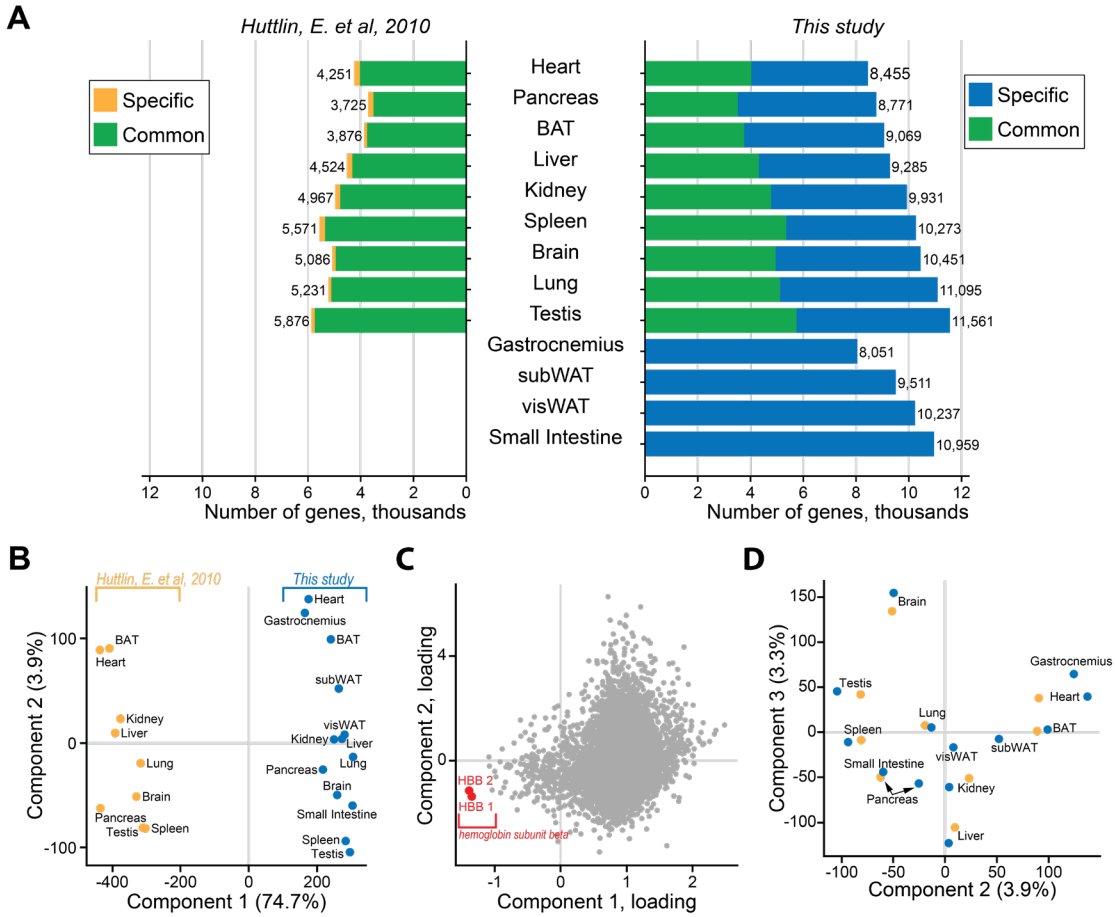
Supplementary Figure 2. Evaluation of Phosphoproteomic Method and Processing Parameters. (A) The average number of localized phosphosites across triplicate injections for three different gradient lengths is shown for searches in Spectronaut (blue) and CHIMERYS in Proteome Discoverer 3.1 (gray). The error bars represent the maximum and minimum value across the triplicates. (B) The number of phosphopeptides and the multiplicity distributions are shown for the data in **Figure 2B**. (C) The site localization probability distributions are shown for the data in **Figure 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 MS2 total ion current distributions are indicated for the two AGC targets in **Supplementary Figure 2E**. (G) The distributions of MS2 injection times are indicated for the two AGC targets in

1438 **Supplementary Figure 2E**. (H) Histograms of localization probabilities for S, T, Y, and P
1439 residues are shown as an alternative visualization for the phosphoproline decoy search
1440 results in **Figure 2H**.
1441

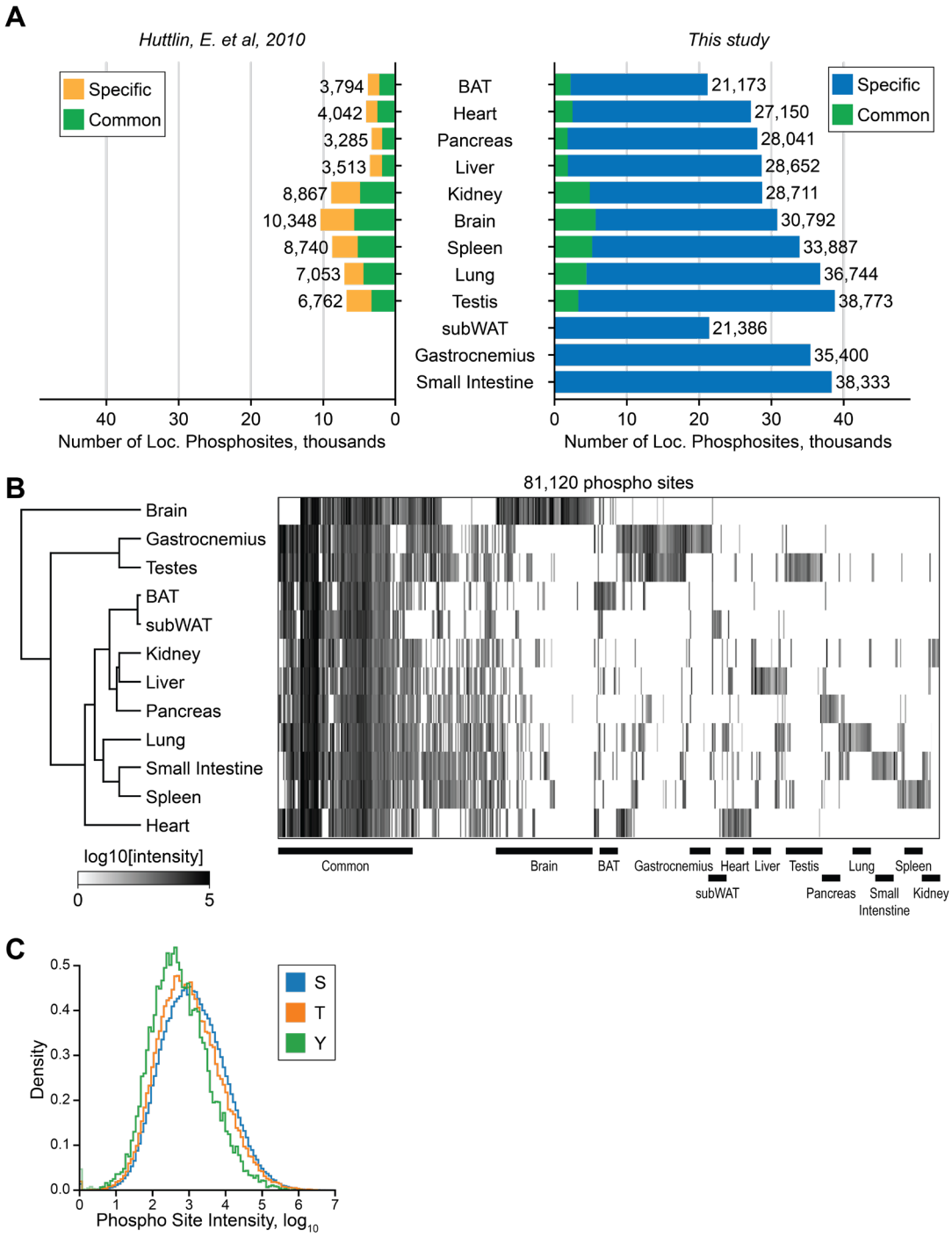


Supplementary Figure 3. **Validation of Site Localization and Quantification using Synthetic Phosphopeptide Standards**. (A) The localization error rates shown in **Figure 2I** are compared to the values previously reported by Bekker-Jensen *et al*[41] alongside the error rates calculated following re-analysis 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 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 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)

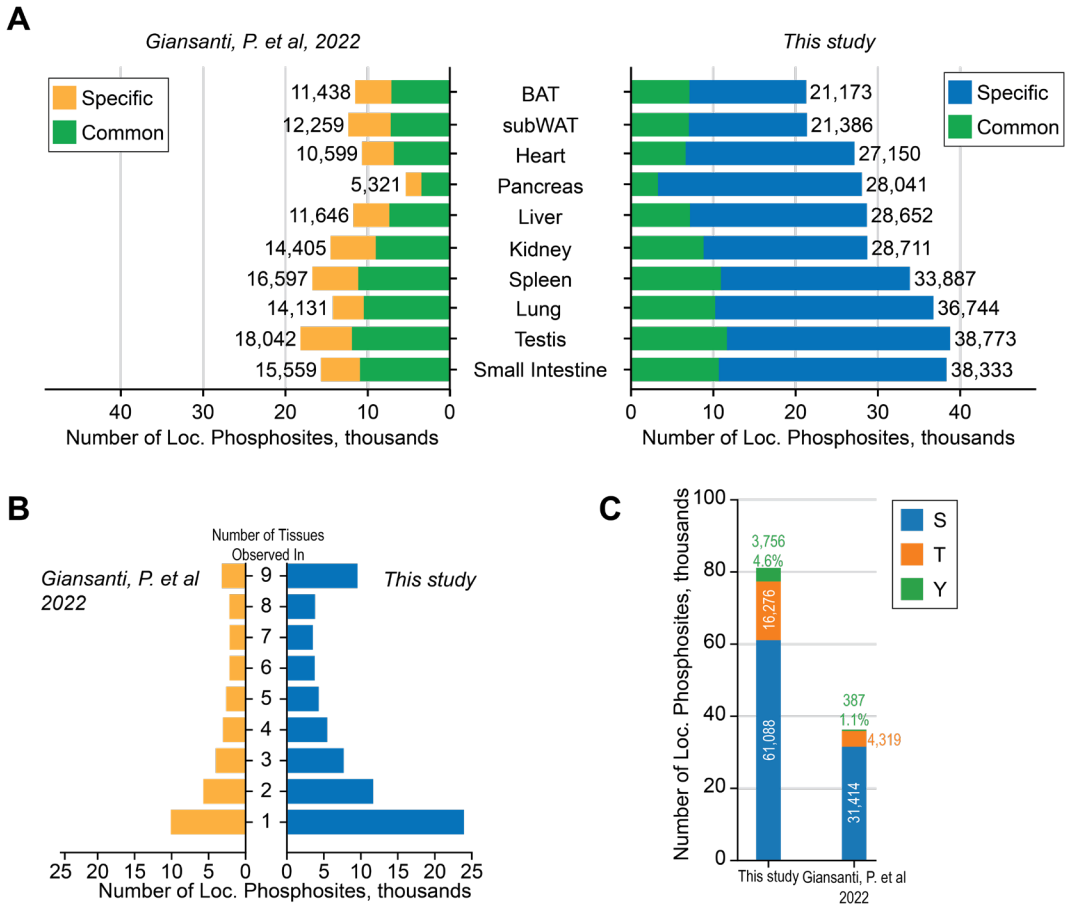
1455 for the synthetic phosphopeptide standards and yeast phosphopeptides are shown in blue
1456 and black, respectively, for each of the dilution series points. The median intensity for the
1457 phosphopeptide standards and yeast phosphopeptides is indicated on each plot with a
1458 blue and black dotted line, respectively.
1459



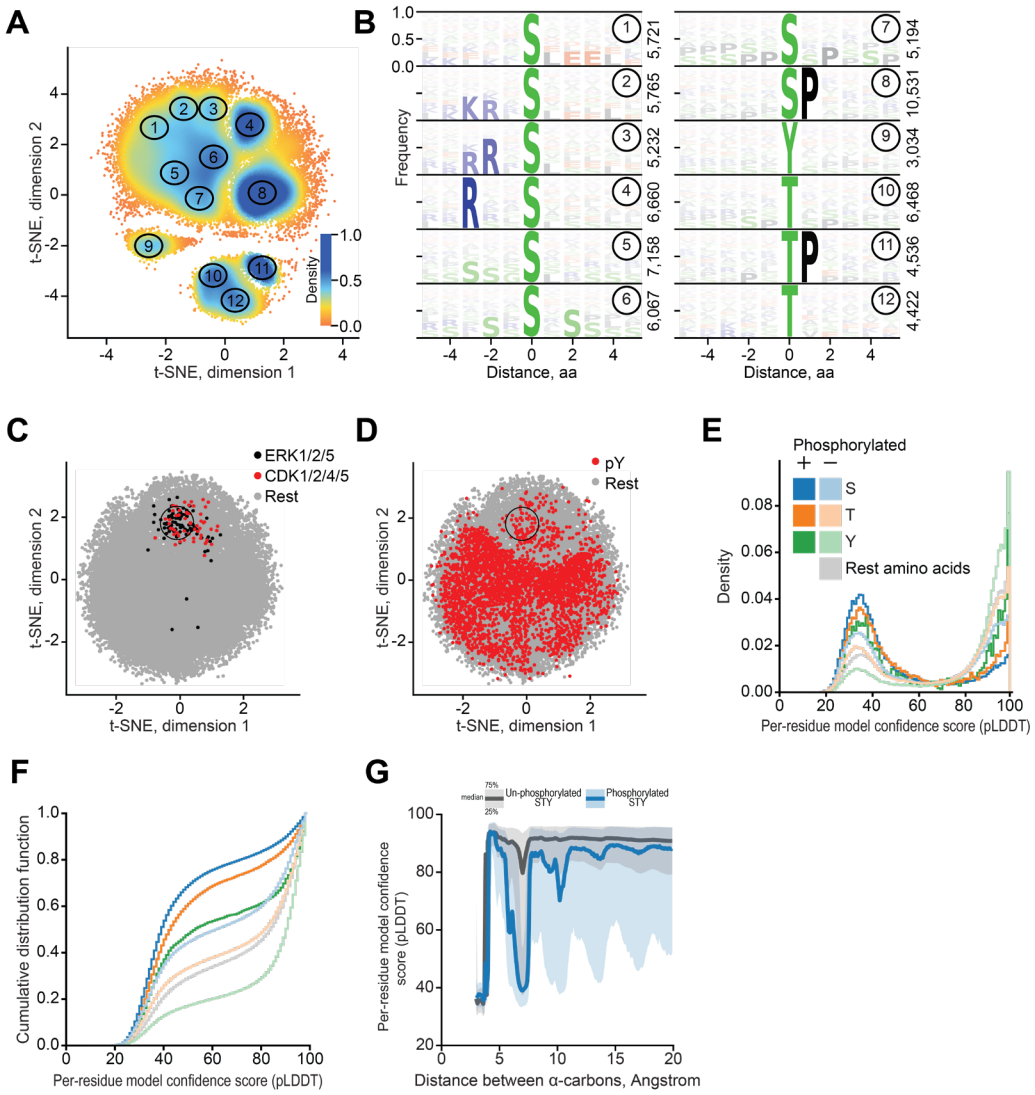
Supplementary Figure 4. Proteomic Profiling of Mouse Tissues. (A) Results from Mouse Tissue Proteomic Analysis. The numbers of protein groups are indicated for each tissue, with common protein groups highlighted in green, those unique to Huttlin et al. in yellow, and those unique to this study in blue.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



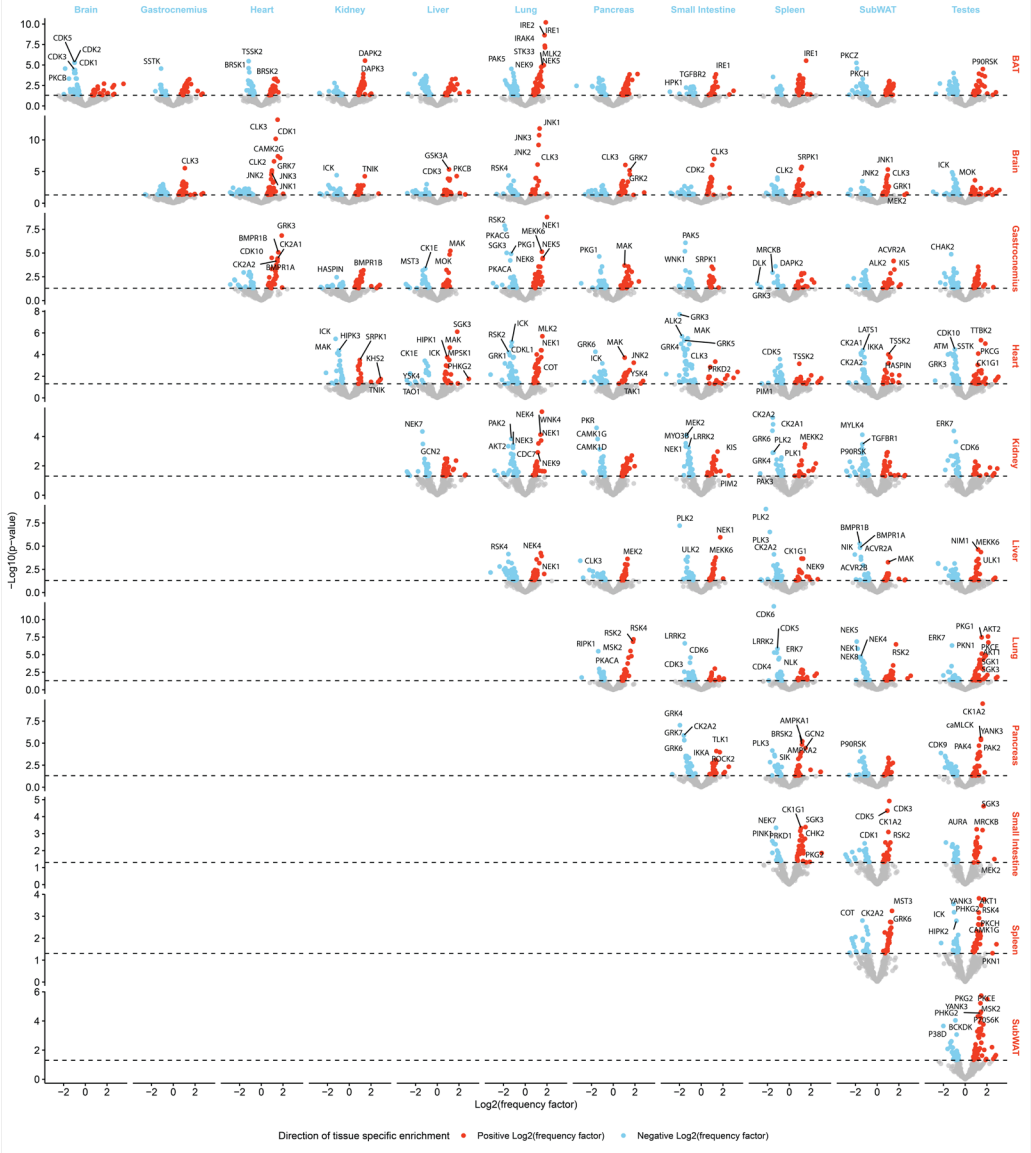
Supplementary Figure 5. Mouse Phosphorylation Atlas Comparison to Huttlin et al. (A) Overlap of phosphorylation sites detected in this study and by Huttlin et al. (B) Heatmap showing the relative intensities of phosphosites detected across tissues. Note that there are many phosphorylation sites with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



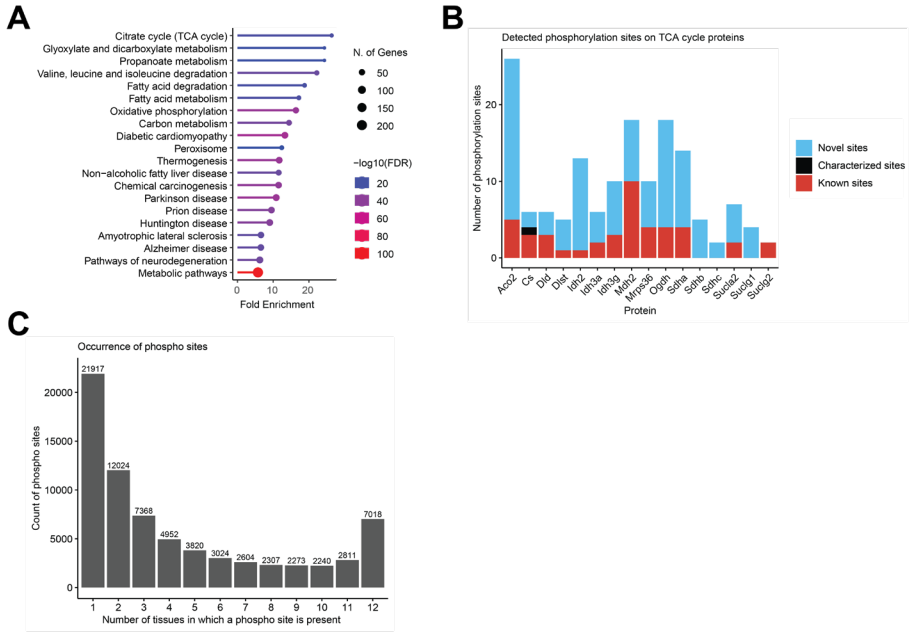
Supplementary Figure 6. Mouse Phosphorylation Atlas Comparison to Giansanti et al. (A) Overlap of phosphorylation sites detected in this study and by Giansanti et al. (B) Tissue Specificity of Detected Phosphorylation sites, excluding the pancreas, brain, and gastrocnemius tissues from comparison. The 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. The percentage of the total sites occurring on phosphotyrosine residues is indicated for both studies.



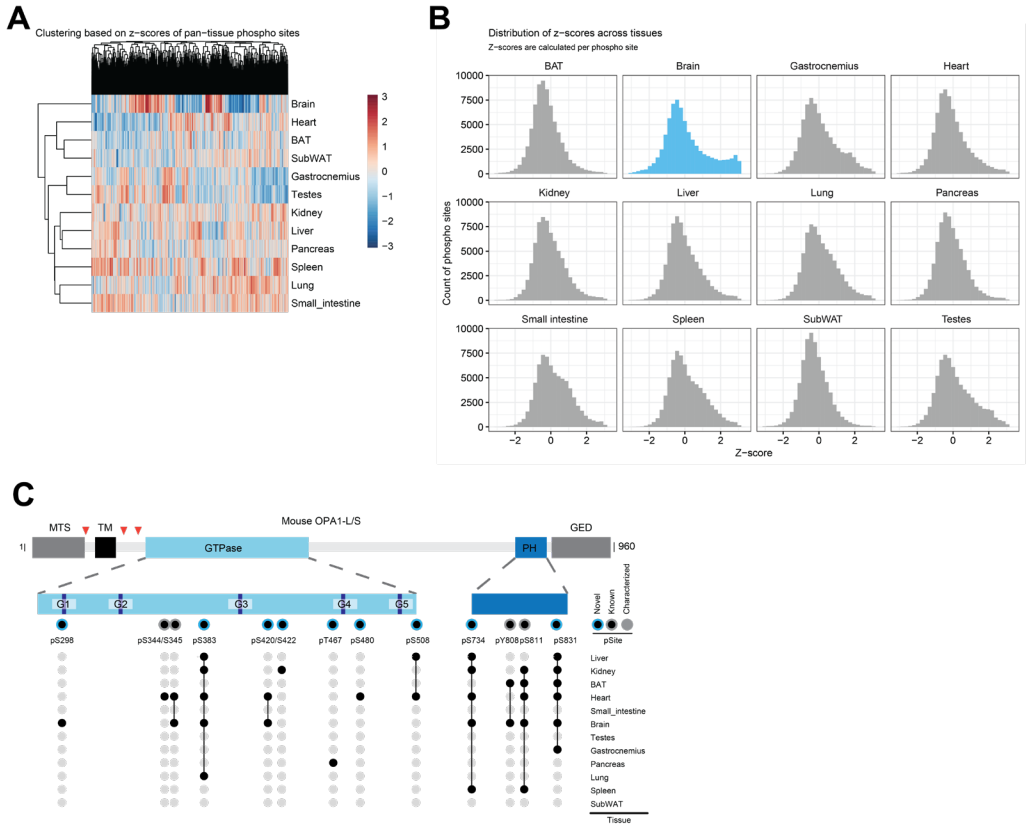
Supplementary Figure 7. Sequence and Structural Context of Phosphosites. (A) Two-dimensional representation of all phosphosites and their 5-amino acid flanking sequences, including the central amino acid in the comparison. Each cluster has been manually selected to highlight the densest regions. (B) 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. (G) Confidence score distribution based on the distance between alpha-carbons surrounding phosphorylation site and phosphorylation state – phosphorylated and un-phosphorylated S/T/Y.



Supplementary Figure 8. 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 labelled in red on the right side of the plot. The blue dots correspond to the tissues plotted across the top of the graph. Source data is provided in **Supplementary Table 5**.



Supplementary Figure 9. 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.[83] 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.



Supplementary Figure 10. 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.

1530 **Supplementary Table 1.** Phosphopeptide Standards Summary and Dilution Series Data
1531
1532 **Supplementary Table 2.** Summary of all detected phosphosites for the EGF-stimulation
1533 experiments.
1534
1535 **Supplementary Table 3.** Summary of all detected phosphosites from the
1536 phosphoproteomics experiment for mouse tissues.
1537
1538 **Supplementary Table 4.** Summary of all protein groups detected in the proteomics
1539 experiment for mouse tissues.
1540
1541 **Supplementary Table 5.** Summary of kinase enrichment analysis for mouse tissues.
1542
1543 **Supplementary Table 6.** Expression enrichment and mitochondrial annotation of
1544 phosphosites.