Optimization of Instrument Parameters for Efficient Phosphopeptide Identification and Localization by Data-dependent Analysis Using Orbitrap Tribrid Mass Spectrometers

Seamus R. Morrone, Michael R. Hoopmann, David D. Shteynberg, Ulrike Kusebauch, and Robert L. Moritz*

Institute for Systems Biology, 401 Terry Ave N, Seattle, WA, 98109, USA

* Corresponding author:

Dr. Robert L. Moritz

Email: rmoritz@systemsbiology.org

ORCID ID
Seamus R. Morrone 0000-0001-6454-8970
Michael R. Hoopmann 0000-0001-7029-7792
David D. Shteynberg 0000-0001-7982-0498
Ulrike Kusebauch 0000-0001-6162-7577
Robert L. Moritz 0000-0002-3216-9447

Keywords
Method optimization, protein phosphorylation, peptide phosphorylation, phosphosite localization, phosphopeptide enrichment, false discovery rate, false localization rate, mass spectrometry, data-dependent analysis, HeLa
Boost 30%
Abstract
The analysis of protein phosphorylation site identification by mass spectrometry-based methods continues to improve with increased efficiency at multiple points of the pipeline including affinity isolation, sample handling with automated protein digestion and phosphopeptide extraction routines, and more sensitive phosphopeptide detection capabilities. The role mass spectrometry parameters play in the quantity and quality of results is understudied, with some reports presenting improvements without sufficient details of how these parameters were derived, nor how they were fully optimized. Here, we systematically scrutinize and optimize parameters of two Orbitrap Tribrid mass spectrometers by varying instrument parameters at both the MS1 and MS2 levels using Titanium IMAC-enriched phosphopeptide samples to illustrate how the recovery of phosphopeptides, and the confidence of phosphosite localizations, vary accordingly. These optimized values are then used to compare different IMAC beads for phosphopeptide enrichment, and optimal bead combinations used to further improve phosphopeptide detection. These results have implications for planning detailed phosphoprotein identification and quantitation experiments to provide confidence in detection, site localization and reproducibility.
Introduction
Phosphorylation is an important post-translational modification (PTM) in all domains of life. In eukaryotes phosphorylation generates phosphoamino acids that do not resemble any natural amino acid; instead, they act as new chemical entities and provide a means of diversifying the chemical nature of, for example, proteins. Due to the reversible nature of this modification, phosphoproteins can act as a switch, regulator, or signal on the protein itself or to other proteins in the cell. Phosphorylation of serine, threonine and tyrosine plays a key role in nearly every cellular process and is one of the most extensively studied PTMs \(^1\). With over 75% of the human proteome undergoing phosphorylation at some time during the organism’s life \(^2\), the role phosphorylation plays as a biomarker of disease states or phenotypic variation is well known \(^3\), and drugs which target kinases, the proteins responsible for phosphorylation and central players in many types of cancers, are major areas of pharmaceutical development \(^4\) \(^5\).

Various methods now exist to enrich phosphopeptides from digested cell or tissue lysate, with the most popular being immobilized metal affinity chromatography (IMAC), and others that include immunoaffinity precipitation, non-immunoaffinity-precipitation, and metal oxide affinity chromatography (MOAC) \(^6\) \(^7\). In recent years, automation has become an important component of proteomics workflows, and mass spectrometry (MS) has emerged as the primary method to study protein phosphorylation events at large scale \(^2\) \(^8\). Phosphopeptide enrichment has been adapted for high-throughput automation \(^6\) \(^9\) \(^10\) such as liquid handler robots that are designed for capturing magnetic bead particles functionalized for phosphorylation moieties on peptides from a complex sample, enriching, de-salting, and eluting these peptides so that they are ready for MS analysis. The use of liquid chromatography-mass spectrometry (LC-MS) by data-dependent analysis (DDA) is the method of choice for identifying phosphopeptides and their phosphate site attachment, and some studies report increases from several hundred to potentially tens of thousands of unique phosphopeptides being identified on even microgram quantities of unfractionated phospho-enriched samples \(^9\) \(^11\) \(^12\).

Although frequently used by many publications, unoptimized MS analysis and data interpretation can lead to loss of phosphosite information and incorrect interpretation \(^13\). Downstream of sample preparation, LC settings such as gradient length and MS parameter settings, especially within DDA, such as automatic gain control (AGC), maximum injection time (IT), normalized collision energy (NCE), and the number of precursor ions selected for fragmentation (TopN), can play different roles for different sample types. As an unfractionated phospho-enriched sample is less complex than a whole-cell lysate, LC-MS parameters may not be translatable between the two conditions. Aside from this, the nature of post-translational modifications requires confidence in the localization of the modification, which in turn requires different fine-tuning of parameters compared to samples not enriched for modifications. This is especially the case for phosphorylation, where often neighboring residues undergo modification and must be resolved from each other (e.g. in multi-phosphorylated proteins the majority of sites are within four amino acids of each other \(^14\)). While previous studies have investigated different fragmentation modes for phosphopeptide identification and localization \(^15\), the roles of individual instrument scan parameters have been underexplored. This leads to the question of which values for certain parameters are optimal for phosphopeptide-enriched samples, balancing the...
demand for high coverage of the phosphoproteome with high fragment intensity for confident localization of the site.

Here, we investigate and report the optimization of mass spectrometry settings using a Thermo-Fisher Orbitrap Fusion Lumos on automated phosphopeptide-enriched samples by IMAC, and show how settings such as IT and AGC on both the MS1 and MS2 level, NCE, gradient length, and TopN affect the quantity and quality of the spectral information returned and outcome of identification. These optimized settings were then tested on two different IMAC bead versions and combinations of two different metal types as a means of comparison. The bead combination yielding the greatest amount of phosphopeptides was then used to compare and refine the parameters on a Thermo-Fisher Orbitrap Eclipse, allowing for the identification and high localization confidence of over ten thousand phosphopeptides in minimal lysate starting material. The results of this work will be of interest for examining the trade-offs between quantity (number of phosphopeptides recovered) and quality (localization confidence) of phosphopeptide data with the variation of these parameters, and have implications on data quality for the type of experiment being run and the goal of the analysis.
Methods

Materials

Acetonitrile (ACN; Cat. No. A955-4, CAS #75-05-8), Tris (2-carboxyethyl) phosphine hydrochloride (TCEP; Cat. No. 20490, CAS #51805-45-9), isopropanol (Cat. No. A464-4, CAS #67-63-0), trifluoroacetic acid (TFA; Cat. No. 28903, CAS #76-05-1), and formic acid (FA; Cat. No. A117-50, CAS #64-18-6) were purchased from Fisher Scientific. Iodoacetamide (IAA; Cat. No. 16125, CAS #144-48-9), glycolic acid (Cat. No. 124737, CAS #79-14-1), and ammonium hydroxide solution (Cat. No. 05002, CAS #1336-21-6) were purchased from Sigma-Aldrich. Ethanol (Cat. No. 2701, CAS #64-17-5) was purchased from Decon Laboratories, Inc. Carboxylate beads (Cat. Nos. 24152105050250 and 65152105050250) were purchased from GE Healthcare, while immobilized metal-affinity beads (Ti-IMAC: Cat. No. MR-TIM002; Ti-IMAC-HP: Cat. No. MR-THP002; Zr-IMAC: Cat. No. MR-ZRM002; Zr-IMAC-HP: Cat. No. MR-ZHP002) were from ReSyn, South Africa. All other materials and instruments are specified by their manufacturer in the text.

HeLa cell lysis, quantification, reduction, and alkylation

HeLa S3 cells (ATCC Cat No. CCL-2.2) were cultured in Dulbecco’s Modified Eagle’s Medium supplemented with 10% fetal bovine serum at 37 °C in 5% CO₂ to confluency. Cells were rinsed with cold PBS three times, snap-frozen on ethanol/dry ice in pellets (1.1-1.2e7 cells/pellet) and stored at -80 °C. Pellets were lysed in urea buffer (8 M urea, 0.1 M Tris pH 8.2) by three rounds of flash-freezing in dry-ice/ethanol followed by vortexing. The lysate was sonicated for 10 rounds of 30 s in a cup-horn sonicator (full power) to shear any DNA, then centrifuged for 15 min at 23,000 x g at 4 °C. BCA assay (Pierce) was performed to determine the protein quantity. Lysate was reduced with 5 mM TCEP for 30 min at 37 °C in a benchtop incubator at 850 RPM, followed by cysteine-blocking with 15 mM IAA for 30 min at 20 °C in the dark at 850 RPM, and then quenched with 5 mM TCEP for 30 min at 20 °C at 850 RPM.

Protein-Aggregation Capture and Digestion

Lysate was diluted into 80% ethanol to a final concentration of 0.4 μg/μL. Automated Protein-Aggregation Capture was performed on a Thermo-Fisher/Qiagen BioSprint-96 according to the following method: plate 1 - magnetic head comb, plate 2 - 1.6 μg/μL carboxylate beads (1:1 mixture of hydrophilic and hydrophobic beads [mass/mass]) in 500 μL H₂O per well, plate 3 - 500 μL 80% ethanol per well, plate 4 - 0.4 μg/μL lysate in 500 μL 80% ethanol per well, plate 5 to 7 - 500 μL 80% ethanol per well, plate 8 - 300 μL 50 mM ammonium bicarbonate per well. The program was set to collect the beads from plate 2, wash for 5 min at medium speed in plate 3, bind protein for 15 min at medium speed in plate 4, wash for 5 min per plate at medium speed for plates 5 through 7, and mix beads for 10 min at medium speed in plate 8. The beads were then left in plate 8, transferred to 1.5 mL microcentrifuge tubes, the wells washed with 100 μL 50 mM ammonium bicarbonate, and trypsin at a ratio of 1:100 (w/w) trypsin:protein lysate was added to the tubes. Digestion was carried out for 18 hr at 37 °C on a tube rotator. Digestion was quenched with addition of formic acid to 1%, beads were resolved on a magnet, and supernatant was dried in a SpeedVac.
**Phosphopeptide-enrichment and Cleanup**

Dried peptides were resuspended in “Loading Buffer” (80% ACN, 5% TFA, 1 M glycolic acid) to 0.4 μg/μL. Automated phosphopeptide enrichment was carried out using a BioSprint-96 according to the following method: plate 1 - magnetic head comb, plate 2 - 0.8 μg/μL Ti-IMAC beads in 500 μL ACN per well, plate 3 - 500 μL Loading Buffer per well, plate 4 - 0.4 μg/μL peptides in 500 μL Loading Buffer per well, plate 5 - 500 μL Loading Buffer per well, plate 6 - 500 μL Wash Buffer 1 (80% ACN, 1% TFA) per well, plate 7 - 500 μL Wash Buffer 2 (10% ACN, 0.2% TFA) per well, plate 8 - 400 μL Elution Buffer (2% ammonium hydroxide) per well. The program was set to collect the beads from plate 2, wash for 5 min at medium speed in plate 3, bind peptides for 15 min at medium speed in plate 4, wash for 5 min per plate at medium speed for plates 5 through 7, and elute phosphopeptides for 15 min at medium speed in plate 8. Beads were then transferred back to plate 2. Eluate was collected, acidified to 5% TFA, and a second round of automated enrichment was carried out with the same sample and fresh beads and buffers (double extraction). Both acidified elution volumes were combined, and samples were subjected to a clean-up step using Atlas Cerex columns on a positive-pressure manifold. Columns were washed with 250 μL ACN followed by 250 μL 0.1% TFA, eluted samples from phosphopeptide enrichment were applied, followed by 2 x 1 mL washing steps of 0.1% TFA, and peptides were eluted in three steps by adding 450 μL 15% ACN/0.1% TFA, 450 μL 30% ACN/0.1% TFA, and 450 μL 70% ACN/0.1% TFA. Eluate was then dried down in a SpeedVac.

**Liquid Chromatography using an Easy nano-LC 1000 system**

All samples were resuspended in 2% ACN/0.1% FA and loaded into an Easy nanoLC 1000 (ThermoFisher USA) system (EasyLC). An Easy-Spray PepMap RSLC C18 column (ThermoFisher USA), inner diameter 75 μm, length 50 cm and 2 μm dp was used for peptide separation in combination with an Acclaim PepMap 100 nanoViper C18 column (ThermoFisher USA) of inner diameter 75 μm, length of 2 cm and 3 μm dp in-line pre-column. Mobile phase A was 0.1% FA in water, while mobile phase B was 0.1% FA in ACN. A gradient from 5% B to 35% B was applied for 120 min (or 60min or 240 min as specified in the text), followed by a ramp-up to 80% B over 30 s and held for 9 min as a wash step, then reduced back to 5% over 30 s and held for 20 min for re-equilibration.

**Liquid Chromatography using an EvoSep One system**

For biological replicates, different cell pellets were used for each replicate (i.e., four separate cell pellets, 200 μg peptides from each pellet); for technical replicates, one cell pellet was enriched and 4x200 μg were pooled and then applied to four EvoTips). All samples were resuspended in 0.1% FA and loaded on Evo-Tips which had been washed with 0.1% FA in ACN, soaked in isopropanol, and pre-equilibrated with 0.1% FA. Tips were then washed with 0.1% FA and stored with 0.1% FA until injection into the EvoSep system. A PepSep column, 25 cm in length, inner diameter 150 μm, and 1.9 μm dp packed with ReproSil C18 beads was used for the separation, connected to a PepSep stainless-steel emitter. Mobile phase A was 0.1% FA in water, while mobile phase B was 0.1% FA in ACN. A modified 88 min pre-programmed Evosep+ method (~15 samples per day) was used to improve hydrophobic peptide elution.

**Liquid Chromatography using a Vanquish Neo nano-HPLC system**
An Easy-Spray PepMap RSLC C18 column (ThermoFisher USA), inner diameter 75 µm and length 50 cm, with 2 µm dp was used for the separation, with a PepMap C18 column (ThermoFisher USA) of inner diameter 300 µm and length of 5 mm with 5 µm dp in-line pre-column. Mobile phase A was 0.1% FA in water, while mobile phase B was 0.1% FA in ACN. A gradient from 5% B to 35% B was applied for 120 min (or 60 min or 240 min as specified in the text), followed by a ramp-up to 80% B over 30 s and held for 9 min as a wash step, then a pre-programmed re-equilibration was performed for three column volumes (about 15 min).

**Mass Spectrometry using an Orbitrap Fusion Lumos**

A sample amount corresponding to 200 µg of initial lysate as measured in the previous section was injected for each run into the LC system (EasyLC or EvoSep) attached to an Orbitrap Fusion Lumos (ThermoFisher USA) without the Advanced Peak Determination (APD) algorithm turned on.

All samples were analyzed with a spray voltage of 1900 V in positive mode with the RF lens at 30%, using the Orbitrap with an MS1 resolution of 120,000, with maximum IT and AGC as specified in the text. TopN was varied from 5 to 30 for the experiments specified in the text, and a value of 15 was used for all other experiments. MS2 resolution for the Orbitrap was 30,000 (FWHM) at m/z 200, with maximum IT and AGC as specified in the text. A normalized collision energy of 28% was used with HCD except in the experiments where it was varied between 24-32%, with an isolation window of 1.6 m/z. Dynamic exclusion was 30 s.

**Mass Spectrometry using an Orbitrap Eclipse**

A sample amount corresponding to 200 µg of initial lysate as measured and described above was injected for each run into the LC system (EasyLC or Vanquish Neo) attached to an Orbitrap Eclipse (ThermoFisher USA USA) with the Advanced Peak Determination (APD) algorithm turned on.

All samples were analyzed with a spray voltage of 1900 V in positive mode with the RF lens at 30%, using the Orbitrap with an MS1 resolution of 120,000, with maximum IT and AGC as specified in the text. TopN was varied from 5 to 30 for the experiments specified in the text, and a value of 15 was used for all other experiments. MS2 resolution for the Orbitrap was 30,000 (FWHM) at m/z 200, with maximum IT and AGC as specified in the text. A normalized collision energy of 28% was used with HCD except in the experiments where it was varied between 24-32%, with an isolation window of 1.6 m/z. Dynamic exclusion was 30 s.

**Data Analysis**

Thermo .RAW files were converted to mzML format using msConvert (ProteoWizard version 3.0.19225-a1ce12329) with “peakPicking true 1-” and “zeroSamples removeExtra” filters. mzML files were searched using Comet version 2021.01 rev. 0 with the reviewed UniProt Homo sapiens proteome (UP000005640) containing 20,371 entries, downloaded October 12, 2021. The database was appended with the common Repository of Adventitious Proteins (cRAP) database, and DeBrujin randomized decoys (k=2) to the extended database were generated with the Trans-Proteomic Pipeline v.6.0.0 (OmegaBlock). Two different
independent sets of DeBruijn decoys were generated for validation purposes, labeled with DECOY0 and DECOY1 prefixes in the fasta database. In all, the fasta database contained 50% target (including cRAP), 25% DECOY0 and 25% DECOY1 sequences. During TPP mixture modeling DECOY0 PSMs were used in the semi-supervised learning of the probabilities and discarded by being assigned probability 0. The DECOY1 PSMs are independent decoys that can remain with non-zero probabilities and can be used to independently estimate decoy-based error rates, given the expectation that among non-DECOY0 random matches in the results 33% will match to DECOY1 and the rest to target (and cRAP) sequences. The search was performed with a peptide mass tolerance of 20 ppm and a fragment binning tolerance of 0.02, trypsin as the search enzyme and semi-tryptic digestion with two allowed missed cleavages, with carbamidomethyl of cysteines as a fixed modification and methionine oxidation and serine/threonine/tyrosine phosphorylation as variable modifications. Comet results were then processed with PeptideProphet with Z-accurate mass binning, using the Expect Score as the discriminant, and decoy (DECOY0) hits to pin down the negative distribution with a non-parametric model. The PeptideProphet results were further processed with iProphet, and localization of phosphorylation sites were analyzed with PTMProphet set to high MS2 accuracy, using an iProphet probability value of 0.9 (< 0.01 FDR) and a mean best probability of 0.75 for confident localization. Peptides were label-free quantified in both the MS1 level by XPRESS and the MS2 level by QuanTic in the TPP suite with a tolerance of 0.02 Da.
Results

Differences between methods for non-phosphopeptides and phosphopeptides
The primary objective of this study is to explore how various parameters within a mass spectrometry DDA method affect the quantity and quality of phosphopeptide identification and use the optimized parameters to extract the maximum identifications of phosphopeptides in a standard lysate-enrichment preparation. As a starting point, we compared methods with and without phosphopeptide identification in mind. The role which different parameters, specifically AGC and IT at the MS1 and MS2 level, gradient length, TopN, NCE, and phosphopeptide-capture type, play in increasing the identification of phosphopeptides is explored below.

We established a baseline comparison on an Orbitrap Fusion Lumos Tribrid (Lumos) mass spectrometer between a standard shotgun method optimized in our lab on whole-cell lysate (referred to as standard method or SM) and with a published method designed for phosphoproteomics by the Olsen et al and the Hess et al groups (referred to as the “phospho-based method” or PBM). Both methods are contrasted in Supplementary Table 1, with instrument parameter differences in MS1/MS2 of resolution, AGC, IT, and NCE indicated. When using both methods to analyze an unenriched HeLa lysate digest, little difference was observed in the number of identified spectra as well as unique peptides (Fig 1A). In contrast, analyzing a HeLa lysate digest enriched for phosphopeptides by these two methods illustrated an advantage when identifying phosphopeptides with the PBM over the SM: we observed a 12% increase in unique phosphopeptides, and a larger percentage of confidently localized phosphopeptides (54% versus 70% with SM or PBM, respectively), leading to a 45% increase in unique localized phosphopeptides (Fig 1B). These results indicate that, to take advantage of a high rate of return of identified phosphopeptides from phospho-enriched samples, the instrument parameters for optimal data collection need to be selected using evidence-based approaches. We further explored the method parameters that differentiate the PBM from the SM and analyzed the effects of various instrument parameters on the success of phosphopeptide identification and localization. We then adapted the optimized PBM to an Orbitrap Eclipse Tribrid (Eclipse) mass spectrometer to demonstrate the notable differences from the previous model (Lumos).

The effect of IT on MS1 and MS2 on phospho-enriched samples
In the Orbitrap Tribrid series of instruments, AGC controls the ion population within the Orbitrap, allowing the user to specify a maximum amount to limit space-charge effects; maximum IT allows the user to control the amount of time in which ions populate the C-trap before being transferred to the Orbitrap. AGC and IT are coupled: if maximum AGC is reached before maximum IT, or vice versa, then the scan is completed and stored. Therefore, to see the effects of IT separate from AGC, the invariant parameter must be set to an arbitrarily high value to avoid being triggered. Uncoupling IT from AGC can be accomplished by setting the AGC to its maximum value (5e6), allowing for the majority of MS1 or MS2 scans to reach their maximum IT; as shown by Supplementary Fig 1A where MS1 scans associated with identified precursors reach the maximum IT in the majority of cases (indicating the maximum AGC was reached first) excluding the final timepoint of 40 ms, while for testing various MS2 maximum IT all MS2 scans associated with identified peptides reached the maximum specified (Supplementary Fig 1B).
This indicates that each MS1 (or MS2 in those experiments focusing on it) scan of interest is influenced almost exclusively by the injection time and not AGC.

For MS1 scans, setting maximum injection time between 2.5 ms and 20 ms has a very small effect on the number of identified unique phosphopeptides (Fig 1C), while doubling again to 40 ms showed a decrease of 15% for unique phosphopeptides, indicating that a longer injection time has a negative effect on the number of phosphopeptide identifications. This is borne out by the fact that, on average, the 40 ms IT resulted in a net loss of 1 scan/sec compared to the 20 ms IT (Supplementary Fig 1C). Localization confidence was unaffected by injection time (Fig 1C), as an accumulation of precursor ions in a survey scan would not be expected to affect the confidence of localization of a phosphosite that depends on MS2-level information (see below). However, there was a strong trend for longer injection times to capture more low-intensity precursors for subsequent selection and fragmentation: a distribution of identified precursor intensities showed at least a two-fold increase in the fraction with an intensity value equal to or less than 1e6 (below the halfway mark of dynamic range log-distribution) for a 40 ms injection time versus a 2.5 ms injection time (Fig 1D). This indicates that, to optimize the identification of low-intensity (and, correspondingly, low-abundance) peptides, a longer injection time should be chosen. Altogether, MS1 IT had little to no effect on the overall numbers of identified phosphopeptides at shorter fill times and a small negative effect at the longer fill time sampled here, but increased the likelihood to select lower intensity phosphopeptides that were then identified from their corresponding MS2 spectra.

For MS2 scan events, comparing from lowest injection time used (15 ms) to the highest (240 ms), two competing trends can be observed: first, the number of unique phosphopeptides increases with increased injection time, until 120 ms, then begins to decrease with increasing injection time (Fig 1E); second, the quality of the spectra as measured by confidence in localization of the phosphorylated peptide shows an increasing trend as injection time increases (Fig 1E). An explanation of these trends is that increasing injection time increases signal abundance during MS2 analysis, improving the spectral quality for the ability to identify the phosphopeptide and localize the site(s) of phosphorylation. However, the increased injection time will reduce the number of scan events possible during a run, eventually leading to fewer unique phosphopeptides selected and ultimately identified. Nonetheless, the stronger signal achieved for the highest injection times allows improved phosphosite localization among the phosphopeptides that are selected for MS2 analysis. Precursor intensities show a similar trend with unique phosphopeptides at low intensities, i.e., lower-intensity precursors (1e6) were identified more in 60-120 ms than lower or higher injection times (about 1.8x as many between the 60-120 ms point and the 15 ms point) (Fig 1F). One reason for this may be that lower-intensity precursors can be more confidently matched to fragment ions with increasing fill times (due to a higher fragment ion accumulation), but at higher maximum IT there will be fewer precursors selected due to fewer scan events, with DDA favoring more intense precursors to select for fragmentation. These data suggest optimal MS2 IT should be dictated by the goals of the experiment: e.g., as a means of survey of general peptide phosphorylation in a sample, such as determining whether certain peptides are phosphorylated in one condition versus another, a lower MS2 injection time should be used to eliminate candidates while also
optimizing the maximum total number of phosphopeptides. However, if biomarker characterization is the primary goal, a higher MS2 IT to maximize phosphosite localization confidence should be used.

Considering MS1 and MS2 injection time optimization together, as a balance between the most phosphopeptides recovered (indicated by the maximum of the functions plotted in the figures) and high localization confidence as well as capture of low-intensity precursors, which sometimes act in opposition, a 20 ms injection time for MS1 and a 60 ms injection time for MS2 were chosen for optimized values in further experiments.

The effect of AGC on MS1 and MS2
Uncoupling AGC from IT is more difficult, as the range is orders of magnitude higher than IT and so takes longer to accumulate ions at the higher end, leading to very long injection times. MS1 accumulates ions much more quickly than MS2 due to the fairly dense initial ion cloud. Aside from this, there is only one MS1 scan event per cycle, while (in the current experimental setup) there are 15 MS2 scans which multiply the total cycle time; hence, MS1 AGC is easier to uncouple from IT to see its effects on phosphopeptide recovery, while determining phosphopeptide recovery on MS2 is confounded by the higher number of scans. Although in practice it is unlikely that similar AGC values for MS1 and MS2 are used due to the differences in ion cloud densities, the ranges tested here for MS2 are similar to MS1 to test the extremes of AGC on phosphopeptide recovery. To uncouple MS1 AGC from IT, maximum IT was set to 200 ms and resulted in the majority of MS1 IT associated with precursors being less than the maximum and hence being triggered by maximum AGC (Supplementary Fig 2A); a second dataset was collected at the optimum injection time of 20 ms as well with a finer AGC variation as a means of testing more practical values (Supplementary Fig 2B). Uncoupling AGC from IT for MS2 is far more difficult; setting the maximum IT to an inconveniently high value (1s; thus, a total cycle of 15 scans occupies some 15 seconds) still resulted in the majority of MS2 scans associated with an identified peptide maxing out in injection time before reaching the target AGC value (Supplementary Fig 2C). In addition, a longer MS2 injection time results in a total scan time of at least an order of magnitude larger due to 15 MS2 scans per cycle. A second dataset for MS2 was also collected at the optimum MS2 IT value previously chosen (60 ms) and at a finer AGC gradation (Supplementary Fig 2D).

The effect of increasing AGC on MS1 with regards to peptide identification reaches a maximum at 1e5 (Fig 2A) that stays fairly flat for all higher AGC values tested. Unexpectedly, the fraction of unique phosphopeptides confidently localized decreased from an initial high of 0.9 at the lowest AGC value tested (1e3) to a minimum around 0.54 at 1e5 and higher values (Fig 2A). This can be explained as so: at low AGC, a very small portion of spectra are identified (~3500 at 1e3 vs ~20,000 at higher values); those that are identified most likely represent highly abundant peptides and thus have high fragment ion populations to assist in confident localization of the phosphosite. This is confirmed by the distribution of precursor intensities, which show the precursors from the low-AGC dataset at a much higher intensity distribution than the high-AGC precursors (Fig 2B). Therefore, in this case confident localization is the result of the instrument selecting highly abundant precursors leading to high localization confidence from increased
The effect of increasing AGC on MS1 does not appear to be large when using the optimal maximum IT. There is a corresponding slight maximum in unique phosphopeptides identified at an AGC of 3e6, but this only represents a 4% increase from the lowest observed at 1e6 (Supplementary Fig 2B). The decrease in localization confidence represents a little over 1% decrease from the maximum at 2e6 (71.4% for unique phosphopeptides) to the minimum at 4e6 (70.0%) (Supplementary Fig 2B). Taken together, a high AGC value in combination with a reasonable injection time (at least 1e6 and 60 ms, respectively) results in a large number of identified phosphopeptides without sacrificing significant cycle time.

Evaluating the effect of AGC on MS2 is more difficult, given the impact of longer associated cycle times with increasing AGC. With the high IT dataset, the number of unique phosphopeptides showed a sharp increase and peak around 1e4 followed by an exponential decline (Fig 2C). The confidence in localization, however, plateaued at the highest values of AGC (1e5-1e6) (Fig 2C). As higher MS2 AGC results in increased accumulation of the precursor ion packet, the MS2 fragment ions are concomitantly increased, resulting in higher localization confidence from a richer, more intense MS2 spectra. Comparing the dataset collected at the optimum IT of 60 ms showed no major differences between low and high AGC in terms of unique phosphopeptides and the same is true for the confidence in localization (Supplementary Fig 2D). This indicates that 60 ms is sufficient to reach at least 5e4 ion intensity, and the number of precursors of higher intensity is not significant enough to change any confidence in localization or number of unique phosphopeptides identified.

As higher AGC values lead to a very long injection time, setting a maximum IT threshold is the better option for triggering the scan event. For MS1, a high AGC coupled with a 20 ms IT balances ion accumulation with speed, while for MS2, 60 ms is sufficient to accumulate at least 5e4 ions. Therefore, an AGC value of 3e6 for MS1 and 1e6 for MS2 are the optimal values for maximizing phosphopeptide identification and phosphosite localization.

The effect of normalized collision energy on phosphopeptide identification
The normalized collision energy (NCE) for HCD peptide fragmentation was tested to determine the effect on phosphopeptide identification (Supplementary Fig 2E). NCE values tested ranged from 24-32%. Lower NCE may be insufficient to suitably fragment isolated precursor ions, and higher NCE may lead to over-fragmentation and loss of the phosphopeptide moiety. The highest number of PSMs and phosphopeptides were identified at 28% NCE, and the localization confidence varied little between the different energies used. With the five values tested, 28% NCE resulted in the most optimal return of identification and localization.

The effect of TopN and gradient length on phosphopeptide identification
In order to evaluate the effects of gradient length of the number of unique peptides identified, three varying lengths of a 5-35% B gradient were used during separation by liquid chromatography: 60 min, 120 min, and 240 min. In parallel to this, the number of precursor ions selected for fragmentation (TopN) was set for either 5, 10, 12, 15, 20, or 30. Results show that, in terms of localization, the 120 min gradient yielded the most unique phosphopeptides.
There was a diminishing return in the number of identified phosphopeptides with increasing gradient, as a 5-35%B gradient is the portion of organic solvent where the bulk of peptides elute, one might expect a substantial increase in numbers reported from 60 min to 120 min and from 120 min to 240 min (a “boost” when comparing the ratios of 120 min to 60 min [120vs60] or 240 min to 120 min [240vs120]). However, unique phosphopeptides showed only modest gains in going from 60 min to 120 min, and a decrease in doubling time again (Supplementary Fig 2G). A TopN of 20 yielded the highest number of phosphopeptide results in both the 120 min and 240 min gradient, while in the 60 min gradient a TopN of 10 was optimal. This can be explained by narrower elution profiles in the shorter gradient. With narrower chromatographic peak widths and a high TopN, precursors may elute between survey scans before they can be selected for fragmentation. A smaller TopN reduces cycle time and will increase the likelihood of novel precursors being selected. Taking the 120 min gradient as the optimal gradient, a Top15 precursor selection resulted in a compromise between cycle time, sensitivity due to chromatographic peak broadening, and selection of low-abundance and high-abundance phosphopeptides, as measured by precursor peak area (Supplementary Fig 2H).

Evaluating optimized MS parameters

To test these optimized parameters and to investigate the complementarity of different types of IMAC microparticles (as well as testing manufacturing differences) for the capture of phosphopeptides, we used either standard or high-performance (HP) beads (hereafter referred to by base bead type) and performed a double extraction enrichment: first, we used Ti-IMAC or Zr-IMAC, then a second round was performed on the same sample with either the same microparticle type used in the first round or the alternate metal ion (i.e., Zr-IMAC instead of Ti-IMAC or vice versa). In total we tested four types of microparticles including Ti-IMAC, Zr-IMAC, Ti-IMAC-HP and Zr-IMAC-HP to enrich phosphorylated peptides from HeLa cell lysate. The samples were then analyzed with the optimized LC-MS parameters with the EasyLC and Orbitrap Fusion Lumos. All bead types perform similarly in terms of unique phosphopeptides identified, though in all cases the HP version of the chelated metal ions performs better (6-16% increase in high-confidence localized phosphopeptides) (Fig 3A). This can be seen by combining the results from all metals of the same base bead type (both Ti- and Zr-IMAC or both Ti- and Zr-IMAC-HP) and comparing them (Fig 3B). Comparing unique phosphopeptides within the same base bead type, there was at least 80% overlap between any two experiments, and among all of them there was at least 50% complete overlap (Supplementary Tables 2 and 3).

In terms of chemical differences between phosphopeptides captured from different bead types, there are no major differences in distributions of pl, GRAVY score, or precursor mass for the peptides, indicating that all are equally capable of enriching similar sets of peptides from a similar matrix with regards to charge state, hydrophobicity and length (Supplementary Fig 3). Comparing the same metal types used between different base bead types, at least three-quarters of the unique phosphopeptides in the standard beads are found in IMAC-HP beads, while the fractional unique phosphopeptide overlaps are slightly lower comparing IMAC-HP to standard beads (63-73%) due to the overall higher number of phosphopeptides found in the IMAC-HP beads (Supplementary Table 4). These results suggest that there is an advantage of using the HP beads over the standard IMAC beads.
Using an EvoSep One nanoLC coupled to an Orbitrap Fusion Lumos mass spectrometer provides an increase in sensitivity

The EvoSep One nanoHPLC system uses a disposable trapping tip, optimized gradients, and a sample compression method to increase the throughput of sample analysis by LC-MS [29]. This raises the question if these differences affect phosphopeptide identification compared to conventional nanoLC systems. In order to evaluate the use of the EvoSep system with the optimized Orbitrap parameters, an 88-min 15SPD Evosep+ pre-programmed gradient was used with the same optimized mass spectrometry method parameters from the previous section and using a combination of Ti-IMAC-HP beads followed by Zr-IMAC-HP. Compared to the EasyLC, the EvoSep One resulted in over 1,000 more unique phosphopeptides with high localization confidence as well as increased reproducibility in total identification of phosphopeptides observed across both technical and biological replicates (Fig 3C; values from the same bead-types of 3A,B is 7190). This exceptional gain with the use of a much shorter gradient and sample loading time compared to a conventional nano-LC has been noted elsewhere [10].

Orbitrap Eclipse with Advanced Peak Determination (APD) provides an increase in sensitivity

Using the optimal combination of IMAC beads for phosphopeptide enrichment (Ti-IMAC-HP, followed by Zr-IMAC-HP), the same parameters were explored on the Orbitrap Eclipse with the Advanced Peak Determination (APD) algorithm with a Vanquish Neo nano-HPLC [30,31] (Supplementary Fig 4A, Fig 4). Many of the parameters showed the same general trends as on the Orbitrap Fusion Lumos and thus will not be repeated in every detail. MS1-AGC (Fig 4A), MS1-IT (Fig 4B), and MS2-AGC (Fig 4C) were found to be optimal at different values, indicating that a change in acquisition rate, one of the prime differences between the Orbitrap Lumos and Orbitrap Eclipse, plays a subtle role in parameter optimization. In addition, a MS2 selection of Top20 or a total 3-s cycle time provided the highest return in number of phosphopeptides and high confidence of localization (Supplementary Fig 4B,C). Combining these parameters resulted in a 15% increase of high-confidence-localization phosphopeptides from what was observed on the EvoSep One-Orbitrap Lumos (Fig 4E).

These results show how careful optimization of instrument parameters can improve phosphopeptide identification and localization downstream of sample preparation methods. As a comparison to previous DDA studies with similar protein capture and phosphopeptide enrichment procedures, we have included a reanalysis of the best protein-aggregation capture dataset from Batth et al [11] to indicate how instrument optimization can improve phosphopeptide identification at both the overall numbers identified as well as confidence in localization (Fig 4E). Although this previously published study differs on many variables such as gradient length and LC-MS instrumentation, it stands as a benchmark for the substantial increase in phosphopeptide recovery and identification it demonstrated over previous studies. With proper optimization, we have an improvement in high-confidence localization by about 66%. 
Discussion

In this report we examined the effects of MS1 and MS2 parameter settings on the quantity and quality of phosphopeptide data recovered from a complex lysate sample between two different Orbitrap mass spectrometers, the Orbitrap Fusion Lumos and Orbitrap Eclipse, as well as the role different LC gradients, systems, and phosphopeptide capture materials have on phosphopeptide identification and localization. In terms of quantity, shorter injection times in MS1 and MS2 can lead to a larger number of phosphopeptides identified, while there is a trade-off in terms of localization of the phosphosite (quality). With regards to the effects of different parameters on the quality of the data, the strongest trend appears to be the effect of confidence of localization on a phosphosite with increasing MS2 IT. Another apparent trend is the accumulation of low-intensity precursors with increasing MS1 IT and AGC. These results suggest these parameters should be varied according to the nature of the experimental question, such that a shorter MS1/MS2 injection time may lead to a larger overall set of phosphopeptides, but a longer time will lead to higher localization confidence. Thus, for experiments designed as surveys or preliminary identification of phosphopeptides, shorter injection times will generally lead to higher rates of identification; for those in which detailed site information is required, longer injection times and higher AGC will be of more use (Table 1).

As for the use of different IMAC microparticle types, with the optimized MS parameters, the four types of affinity beads tested (Ti-IMAC, Zr-IMAC, Ti-IMAC-HP, and Zr-IMAC-HP) appear to recover roughly the same numbers and chemo-physical types of phosphopeptides with similar localization confidences. The HP beads have a slight advantage over the standard IMAC beads in terms of the number of identified phosphopeptides, and all beads capture a small portion of peptides unique to the particular bead (see Supplementary Tables 2-4). A more detailed study under various capture conditions may further differentiate these microparticle types.

Gradient lengths have a diminishing rate of return in terms of processing time versus unique phosphopeptides identified, and the middle-ground of 120 min explored here yielded the best compromise. When sample quantity is plentiful, multiple short gradients can provide better rates of identification, but with small amounts of important samples a longer gradient with less replicates may be best. One alternative is to use the EvoSep One system, designed as a high throughput LC. Here, the system run with shorter gradient times (88min versus 150 min for total sample run time) resulted in more phosphopeptides with higher localization confidence than with the EasyLC on a 120 min gradient. It should be noted that the newest generation of a Thermo-Fisher nanoLC-system, the Vanquish Neo, provided similar detection rates as the EvoSep One system.

As these parameters were refined on Orbitrap type mass spectrometers and explored mainly with a 120-min gradient length, these parameters may need additional optimization for the individual setup with different instruments or gradients, but the parameters described here will provide a good starting point. Investigations of other mass spectrometry instruments and how they compare to the Orbitrap systems used here in terms of dynamic range of phosphopeptides captured and numbers recovered will be of interest for the field of phosphoproteomics.
In summary, through careful and separate evaluation of instrument parameters that pertain to Orbitrap instruments, the optimization of HPLC conditions, and the use of these optimized settings provides a substantial improvement in phosphopeptide identification and localization over general, default settings that many users proceed with in their own analysis. Comparing the standard method with the final optimized phospho-based method, these increases include 43% in total phosphopeptide identifications and 111% in confident phosphosite identification from a HeLa S3 cell digest respectively.

Data Availability
All mass spectrometry raw data and search parameters and conditions have been deposited with the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD039921. Reviewer account details: Username: reviewer_pxd039921@ebi.ac.uk Password: PhyvU2Z9

Acknowledgements
We thank Jennifer Villazon and Timm Prokop for their excellent technical assistance. This work was partially funded by grants from the National Institutes of Health, National Institute of General Medical Sciences R01GM087221, the National Heart, Lung, and Blood Institute R01HL133135, the Office of the Director S10OD026936, the National Institute on Aging U19AG023122, and the National Science Foundation award 1920268. We thank Nicolai Bache, Evosep Biosystems, Denmark and Oleksander Boychenko, Thermo-Fisher Scientific, Germany for early access to some instrumentation used in this work.
Figures

A, Identified spectra (blue) and unique peptides (orange) from unfractionated total HeLa cell lysate analyzed with the standard or the phospho-base method protocol. B, Unlocalized and high-confidence localized unique phosphopeptides from phospho-enriched HeLa cell lysate analyzed with the standard or the phospho-base method protocol. The numbers indicate the percentage change. C, The total number of unique phosphopeptides (blue) and fraction confidently localized (orange) plotted as a function of MS1 IT. D, The distribution of precursor intensities of unique confidently localized phosphopeptides identified as a normalized fraction of the total identified, colored according to MS1 injection time. E, The total number of unique phosphopeptides (blue) and fraction confidently localized (orange) identified as a function of MS2 IT. F, The distribution of precursor intensities of unique confidently localized phosphopeptides identified as a normalized fraction of the total identified, colored according to MS2 injection time. All numbers are from a peptide FDR < 0.01, localization mean-best probability of 0.75 in PTMProphet.

Figure 1: Whole lysate optimized method versus phosphopeptide enrichment method and the effect of injection time (IT) on MS1 and MS2.
Figure 2: The effect of Automatic Gain Control (AGC) on MS1 and MS2.

A, the total number of unique phosphopeptides (blue) and fraction confidently localized (orange) plotted as a function of MS1 AGC.

B, the distribution of precursor intensities of unique confidently localized phosphopeptides identified as a normalized fraction of the total identified, colored according to MS1 AGC.

C, the total number of unique phosphopeptides (blue) and fraction confidently localized (orange) plotted as a function of MS2 AGC. All numbers are from a peptide FDR < 0.01, localization mean-best probability of 0.75 in PTMProphet. au = arbitrary units.
Figure 3: Differences in phosphopeptide recovery using different IMAC microparticles and LC systems.

A, Total number of unique phosphopeptides (blue bars) and fraction of confidently localized phosphopeptides (orange dots) graphed according to the bead type indicated. B, Total number of unlocalized (blue) or high confidence localized (orange) unique phosphopeptides identified according to the base bead type. Numbers represent percentage drop between unlocalized and high-confidence localization. C, Number of unique phosphopeptides confidently localized by either biological or technical replication using the EvoSep LC system. In figure 3A, the first bead type listed is the bead type used during the first round of extraction, the second bead type is the one used during the second round of extraction. In figure 3B, Ti/Zr-IMAC indicates any combination of Ti-IMAC and/or Zr-IMAC, while Ti/Zr-IMAC-HP indicates any combination of Ti-IMAC-HP and/or Zr-IMAC-HP, for extraction. Error bars in 3B and 3C are standard errors from the four corresponding experiments shown in 3A. All numbers are from a peptide FDR < 0.01, localization mean-best probability of 0.75 in PTMProphet.
Figure 4: Optimization of MS parameters on the Vanquish-Eclipse.
A, Total number of unique phosphopeptides (blue) and fraction confidently localized (orange) identified as a function of MS1 AGC. B, Total number of unique phosphopeptides (blue) and fraction confidently localized phosphopeptides (orange) identified as a function of MS1 IT. C, Total number of unique phosphopeptides (blue) and fraction confidently localized phosphopeptides (orange) identified as a function of MS2 AGC. D, Total number of unique phosphopeptides (blue) and fraction confidently localized phosphopeptides (orange) identified as a function of MS2 IT. E, Total number of unlocalized (blue) or confidently localized (orange) unique phosphopeptides from the optimized protocol using the instruments indicated as a comparison to Bath et al (2019). All numbers are from a peptide FDR < 0.01, localization mean-best probability of 0.75 in PTMProphet. au = arbitrary units.
Table 1: Optimized parameters for maximizing either total phosphopeptide identifications (Max_ID) or high-confidence localization (Max_Local).

<table>
<thead>
<tr>
<th>Method Nickname</th>
<th>Max_ID</th>
<th>Max_Local</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MS1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max IT (ms)</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>AGC Target</td>
<td>1.0E+05</td>
<td>1.0E+05</td>
</tr>
<tr>
<td>Number of dependent scans</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td><strong>MS2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max IT (ms)</td>
<td>60</td>
<td>120</td>
</tr>
<tr>
<td>AGC Target</td>
<td>1.0E+05</td>
<td>1.0E+06</td>
</tr>
<tr>
<td>Collision Energy (%)</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>Gradient Length</td>
<td>120</td>
<td>60</td>
</tr>
</tbody>
</table>
Supplementary Figure 1: Injection time scan lengths. A, Percentage of MS1 scans associated with an identified precursor which reached the maximum injection time as a function of the set maximum MS1 injection time when AGC is set to the maximum (5e6). B, Percentage of MS2 scans associated with an identified peptide which reached the maximum injection time as a function of the set maximum MS2 injection time when AGC is set to the maximum (5e6). C, Average number of MS1 scans per second expressed as a function of MS1 injection time.
Supplementary Figure 2: AGC scan lengths and supplementary collected data on the Orbitrap Fusion Lumos. A, Percentage of MS1 scans associated with an identified precursor which did not reach the maximum injection time as a function of the set maximum MS1 AGC. B, Total number of unique phosphopeptides (blue) and fraction confidently localized (orange) identified as a function of MS1 AGC with maximum IT fixed at low value. C, the percentage of MS2 scans associated with an identified peptide which did not reach the maximum injection time as a function of the set maximum MS2 AGC. D, the total number of unique phosphopeptides (blue) and fraction confidently localized (orange) identified as a function of MS2 AGC with maximum IT fixed at low value (60 ms). E, the total number of unique phosphopeptides (blue) and fraction confidently localized (orange) identified as a function of normalized collision energy. F, the total unique phosphopeptides identified at a given TopN with a 60-min (blue), 120-min (orange), or 240-min (gray) gradient. G, the ratio of unique phosphopeptides confidently localized as a function of TopN with a 60-min (blue), 120-min (orange), or 240-min (gray) gradient, normalized against the 60-min gradient. H, the binned distribution of peak area of confidently localized phosphopeptides in the 120-min gradient colored according to TopN, expressed as a fraction of the total number of confidently localized phosphopeptides. All numbers are from a peptide FDR < 0.01, localization mean-best probability of 0.75 in PTMProphet. au = arbitrary units.
Supplementary Figure 3: Physiochemical similarity of phosphopeptides captured by different IMAC microparticle types

Shown are the binned distributions of isoelectric points (pI), GRAVY scores, and precursor masses of all unique confidently localized phosphopeptides identified in the different IMAC capture experiments. Bars are expressed as a fraction of the total phosphopeptides identified, colored according to the different bead types used for capture. All numbers are from a peptide FDR < 0.01 and localization mean-best probability of 0.75 in PTMProphet.
Supplementary Figure 4: Additional parameter exploration on the Eclipse. A, the number of unlocalized (blue) or high confidence localized (orange) unique phosphopeptides identified according to whether the automatic peak detection (APD) algorithm was off (left) or on (right). B, the total unique phosphopeptides identified at a given TopN with a 60-min (blue), 120-min (orange), or 240-min (gray) gradient. C, the fraction of confidently localized phosphopeptides identified at a given TopN with a 60-min (blue), 120-min (orange), or 240-min (gray) gradient.
Supplementary Table 1: Comparison of Standard and Phospho-Based Mass Spectrometry Methods. The parameters which differ between the two methods are shown along with their values.

<table>
<thead>
<tr>
<th>Method Nickname</th>
<th>Standard</th>
<th>Phospho-Based</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MS1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution</td>
<td>60K</td>
<td>120K</td>
</tr>
<tr>
<td>Scan range (m/z)</td>
<td>375-1375</td>
<td>375-1550</td>
</tr>
<tr>
<td>Max IT (ms)</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>AGC Target</td>
<td>5.00E+05</td>
<td>3.00E+06</td>
</tr>
<tr>
<td>Number of dependent scans</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td><strong>MS2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collision Energy (%)</td>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td>Resolution</td>
<td>15K</td>
<td>30K</td>
</tr>
<tr>
<td>Max IT (ms)</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td>AGC Target</td>
<td>4.00E+04</td>
<td>1.00E+06</td>
</tr>
</tbody>
</table>
Supplementary Table 2: Comparison of overlap of confidently localized unique phosphopeptides between standard IMAC beads using different metal types.

<table>
<thead>
<tr>
<th>Bead Type</th>
<th>Total Unique Phosphopeptides</th>
<th>Complete Overlap</th>
<th>At least 2</th>
<th>At least 1</th>
<th>No overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ti-IMAC/Ti-IMAC</td>
<td>5528</td>
<td>0.65</td>
<td>0.79</td>
<td>0.89</td>
<td>0.11</td>
</tr>
<tr>
<td>Ti-IMAC/Zr-IMAC</td>
<td>5772</td>
<td>0.62</td>
<td>0.80</td>
<td>0.91</td>
<td>0.09</td>
</tr>
<tr>
<td>Zr-IMAC/Ti-IMAC</td>
<td>5712</td>
<td>0.63</td>
<td>0.81</td>
<td>0.92</td>
<td>0.08</td>
</tr>
<tr>
<td>Zr-IMAC/Zr-IMAC</td>
<td>6365</td>
<td>0.56</td>
<td>0.74</td>
<td>0.87</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Number of confidently localized unique phosphopeptides, overlap of all four experiments, overlap between the specified bead-type and two other beads, the overlaps between the specified bead-type and one other bead, and the unique phosphopeptides to that bead type are all specified. For example, from a total of 5528 unique confidently-localized phosphopeptides identified in the Ti-IMAC/Ti-IMAC experiment using Ti-IMAC for both extraction rounds (Ti-IMAC/Ti-IMAC), 3575 of those were also found in all the other three experiments (65%), 4354 were found in at least two of the other experiments (79%), 4917 were found in at least one of the other experiments (89%), and 611 were unique to only the Ti-IMAC/Ti-IMAC experiment (11%).
Supplementary Table 3: Comparison of overlap of confidently localized unique phosphopeptides between IMAC-HP beads using different metal types.

<table>
<thead>
<tr>
<th>Bead Type</th>
<th>Total Unique Phosphopeptides</th>
<th>Complete Overlap</th>
<th>At least 2 At least 1 overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ti-IMAC-HP/Ti-IMAC-HP</td>
<td>7154</td>
<td>0.64</td>
<td>0.81</td>
</tr>
<tr>
<td>Ti-IMAC-HP/Zr-IMAC-HP</td>
<td>7449</td>
<td>0.61</td>
<td>0.77</td>
</tr>
<tr>
<td>Zr-IMAC-HP/Ti-IMAC-HP</td>
<td>6784</td>
<td>0.67</td>
<td>0.83</td>
</tr>
<tr>
<td>Zr-IMAC-HP/Zr-IMAC-HP</td>
<td>7243</td>
<td>0.63</td>
<td>0.80</td>
</tr>
</tbody>
</table>

Supplementary Table 3: Comparison of overlap of confidently localized unique phosphopeptides between IMAC-HP beads using different metal types.

The total confidently localized unique phosphopeptides, overlaps of all four experiments, overlaps between the specified bead-type and two other beads, the overlaps between the specified bead-type and one other bead, and the confidently localized unique phosphopeptides to that bead type are all specified. For example, from a total of 7154 unique confidently-localized phosphopeptides identified in the Ti-IMAC-HP/Ti-IMAC-HP experiment, 4545 of those were also found in all the other three experiments (64%), 5809 were found in at least two of the other experiments (81%), 6582 were found in at least one of the other experiments (92%), and 572 were unique to only the Ti-IMAC-HP/Ti-IMAC-HP experiment (8%).
Supplementary Table 4: Comparison of overlap of confidently localized unique phosphopeptides between IMAC and IMAC-HP using the same metal capture type.

As an explanation, in the Ti-IMAC/Ti-IMAC experiment, 5528 unique confidently localized phosphopeptides were identified, in the Ti-IMAC-HP/Ti-IMAC-HP experiment 7154 were identified, and 4485 were found in both experiments (81% of 5528, 63% of 7154).

<table>
<thead>
<tr>
<th>Bead Comparison</th>
<th>Total IMAC</th>
<th>Total IMAC-HP</th>
<th>IMAC Overlap</th>
<th>IMAC-HP Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ti-IMAC/Ti-IMAC vs Ti-IMAC-HP/Ti-IMAC-HP</td>
<td>5528</td>
<td>7154</td>
<td>0.81</td>
<td>0.63</td>
</tr>
<tr>
<td>Ti-IMAC/Zr-IMAC vs Ti-IMAC-HP/Zr-IMAC-HP</td>
<td>5772</td>
<td>7449</td>
<td>0.85</td>
<td>0.66</td>
</tr>
<tr>
<td>Zr-IMAC/Ti-IMAC vs Zr-IMAC-HP/Ti-IMAC-HP</td>
<td>5712</td>
<td>6784</td>
<td>0.79</td>
<td>0.67</td>
</tr>
<tr>
<td>Zr-IMAC/Zr-IMAC vs Zr-IMAC-HP/Zr-IMAC-HP</td>
<td>6365</td>
<td>7243</td>
<td>0.83</td>
<td>0.73</td>
</tr>
</tbody>
</table>
References


