

# Protein aggregation capture on microparticles enables multi-purpose proteomics sample preparation

Tanveer S. Batth<sup>1</sup>, Maxim A.X. Tollenaere<sup>2‡</sup>, Patrick L. R  ther<sup>1</sup>, Alba Gonzalez-Franquesa<sup>1,3</sup>, Bhargav S. Prabhakar<sup>1</sup>, Simon Bekker-Jensen<sup>2</sup>, Atul S. Deshmukh<sup>1</sup>, Jesper V. Olsen<sup>\*,1</sup>

1) The Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Denmark

2) Cellular Stress Signaling Group, Center for Healthy Aging, University of Copenhagen, Denmark

3) The Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, Denmark

\*Corresponding author

‡ Current address: LEO Pharma A/S, Industriparken 55, 2750 Ballerup, Denmark

Email: jesper.olsen@cpr.ku.dk

## KEYWORDS

Proteomics, sample preparation, mass spectrometry, phosphoproteomics, secretomics, microparticles, magnetic beads

## ABSTRACT

Universal proteomics sample preparation is challenging due to the high heterogeneity of biological samples. Here we describe a novel mechanism that exploits the inherent instability of denatured proteins for non-specific immobilization on microparticles by protein aggregation capture. To demonstrate the general applicability of this mechanism, we analyzed phosphoproteomes, tissue proteomes, and interaction proteomes as well as dilute secretomes. The findings presents a practical, sensitive and cost-effective proteomics sample preparation method.

# INTRODUCTION

Dedicated sample preparation for shotgun proteomics is essential for removing impurities and interfering species which may affect peptide chromatography, ionization during the electrospray process, and sequencing by mass spectrometers. To represent the in-vivo state of the global proteome including membrane-bound proteins, it is of high importance to ensure complete lysis of cells and tissues prior to protease digestion. This typically requires strong detergents that are difficult to remove afterwards, however crucial in order to avoid signal interference during MS analysis. Considerable developments have been made based on a variety of different biochemical principles which utilize filters, traps, or protein precipitation techniques which address different sample types (1-3). However, a primary challenge remaining is the development of a universal sample preparation method that has the potential to scale across different sample amounts, which typically range from ng to mg of starting material. Moreover, such a method needs to be compatible with different lysis buffers, biological material (i.e. cell lines, tissues), robust, reproducible, cost effective, and perhaps above all; practical. Although several methods have been developed to individually address different proteomics sample preparation challenges, a simple solution spanning all sample types remains elusive. Here we report a mechanism, termed protein aggregation capture (PAC), which utilizes the phenomenon of non-specifically immobilizing precipitated and aggregated proteins on any type of sub-micron particles irrespective of their surface chemistry. We explore the fundamental process underlying this phenomenon behind methods such as SP3 and determine the optimal parameters leading to effective sample preparation for shotgun proteomics analysis by mass spectrometry of different sample types. Our developments demonstrate the potential for low cost, simple, robust and sensitive sample preparation procedures for proteomics analysis, which can be easily implemented in any setting with great potential for full automation.

## RESULTS

Our hypothesis was based on a series of reports and observations that lead us to conclude the mechanism of non-specific aggregation, initially on magnetic beads with carboxyl group surface chemistries. Carboxyl coated magnetic beads have been reported for sensitive proteomics sample preparation as an alternative to other approaches such as FASP with limited starting material (4). The binding mechanism was attributed to hydrophilic interactions (HILIC)(5) with the carboxyl surface groups and the method was termed “SP3”, recent improvements of the protocol, such as pH control have rendered it more practical (6-8). As HILIC principles dictate preferential polar and ionic interactions under non-aqueous conditions, protein interaction to the carboxyl surface of the beads was hypothesized to be induced by the addition of acetonitrile to the protein lysate. However, we observed that stringent binding of proteins to the microspheres could not be completely reversed under aqueous conditions even with extended washing (Figure 1A, supplementary Figure 1A). Proteins however could be released in solubilization buffers such as lithium dodecyl sulfate (Figure 1A). We therefore wondered whether protein immobilization could additionally be driven by aggregation of insoluble proteins on magnetic microspheres. To test this, we treated native protein lysates either by incubation at room temperature (25°C) where proteins should stay in their native state, or induced aggregation by three different known mechanisms, these being organic solvent (acetonitrile; 70% final), high temperature (80°C) for 5 minutes), or high salt (2.5 M ammonium sulfate), followed by the addition of magnetic carboxyl microspheres. Immobilization of aggregated and insoluble proteins was only observed under the three conditions known to induce aggregation indicating that protein aggregation was essential to the underlying mechanism of protein capture (Figure 1B). Importantly, the induced protein aggregation was very effective - especially using acetonitrile - as judged by the little protein amounts remaining in the supernatants.

We subsequently investigated the role of microsphere surface chemistry on protein immobilization and found no impact on protein aggregation irrespective of microsphere surface chemistry including those containing hydrophobic C<sub>18</sub> surfaces (Figure 2A). To rule out the role of the magnetic properties of the microspheres leading to immobilization, we tested protein aggregation on porous 3µm C<sub>18</sub> hydrophobic beads, which are typically utilized for packing reversed-phase nano-columns and found similar immobilization mechanisms (Supplementary Figure 1B). We further examined whether coated smooth surface microspheres were essential for protein immobilization by inducing protein aggregation (with acetonitrile) on fine iron powder microparticles (grain size 5-9µm) and observed aggregation in a similar manner (Supplementary Figure 1C), however due to the poor solubility of carbonyl powder (in water or water/organic solvent mix) the recovery was found to be less reproducible resulting in low protein aggregation and recovery particular as the volume of protein containing solutions were scaled higher (data not shown). Furthermore, we tested the order of addition of beads and solvent and found no noticeable difference on the effect of protein aggregation capture on beads (Supplementary figure 1D). We next inquired whether protein aggregation on microspheres was a function of microparticle surface area by gauging protein aggregation at very low microsphere concentrations relative to a constant concentration of protein lysate at 0.25 µg/ul (Figure 2B). Although very low amounts of beads were sufficient to aggregate proteins from solution (Figure 2B), we found the structural integrity of visible protein-bead precipitate to be unstable when the bead to protein ratio was less than 1:4, leading to dispersion of small aggregated pieces in solution upon mild disruption. Conversely, solutions with low protein concentrations (<0.075 µg/µl) were found to require higher bead to protein ratios for efficient capture and recovery (Supplementary Figure 1D, E). These results indicate that solutions with higher protein concentrations can aggregate on minute amounts of microparticles, however lower protein concentration require relatively higher amounts of microparticles in order to effectively capture aggregated proteins. Collectively the data suggest that microparticle surface (irrespective of

surface chemistry) acts as a nucleation site or carrier to induce a immobilization cascade of insoluble protein aggregates, which ultimately serve to tightly maintain a microparticle-protein structure (Figure 2C).

We assessed the impact on trypsin digestion efficiency of immobilized proteins on carboxyl coated microparticles and compared it to a commonly-used chaotropic agent based in-solution digestion protocol (9). Analyzing all samples by single shot nanoflow liquid chromatography tandem mass spectrometry (LC-MS/MS) we found significantly reduced number of missed tryptic cleavages between the two methods at different Lys-C and trypsin ratios (Figure 3A, supplementary figure 2A, supplementary table 1). These findings imply the possibility to considerably reduce proteomics sample preparation costs as proteases typically constitute one of the largest expense of the workflow prior to MS analysis. As previously determined by missed cleavages across different digestion protocols (10), our results indicate that 10-20x reduction in trypsin and Lys-C usage (1:500 - 1:1000 ratio) leads to comparable missed cleavage rates (below 30%) to the standard 1:50 trypsin-to-protein ratio utilized in most proteomics studies (Figure 3A, supplementary figure 2A).

Post-translational modifications (PTMs) such as site-specific phosphorylation can rapidly modulate the function of proteins by changing their enzymatic activity, subcellular localization, turnover, and interaction partners (11). It is therefore important to develop proteomics methods that enable global analysis of phosphoproteomes in a robust, reproducible and sensitive manner. We examined whether the analysis of protein phosphorylation status is affected by the on-bead protein aggregation capture workflow on serum stimulated HeLa cells.

Phosphopeptides were enriched by magnetic Ti-IMAC beads and the eluates analyzed by LC-MS/MS in turn. The results demonstrate no impact in the number of identified phosphorylated peptide variants and phosphorylation sites (Figure 3B,C, supplementary table 2). Importantly, >1000 more phosphorylated peptides and 779 localized sites were identified on average using microsphere protein aggregation followed by protease digestion compared to standard in-

solution digestion. Moreover, high degree of overlap for localized sites was found between the two methods (Figure 3C) and no bias in the phosphopeptide enrichment was observed between the two experiments as we achieved an enrichment efficiency >99% for all replicates (Supplementary figure 2B). Surprisingly, we did not observe a major difference between missed cleavage rates for phosphopeptides between the two methods (Supplementary figure 2C).

We next assessed the potential of utilizing magnetic microparticles for proteomics analysis of organs and tissues. This can be particularly challenging as it often requires harsh solubilization buffers for efficient protein extraction from hard and soft tissues. To test aggregation on microparticles we utilized skeletal muscle tissue samples from *Mus musculus*. After homogenization and solubilization in 4% SDS lysis buffer, we examined protein recovery and digestion by using either the established filter-aided sample preparation (FASP) protocol or via aggregation of proteins on magnetic microspheres with sulfonic acid surface chemistry. Significantly higher peptide recovery (>2 fold) after Lys-C/trypsin digestion was observed using microspheres from initial starting material of 1.8 mg (as determined by tryptophan assay)(12) per replicate (Figure 4A). Incidentally, sample preparation with magnetic microspheres resulted in a cleaner peptide mixture as no polymer peaks were observed in the mass spectrometry analysis, leading to higher number of identified proteins and unique peptides in a single shot LC-MS/MS analysis (Supplementary Figure 2D, E, supplementary table 3).

Proteins rarely operate alone in the cell and their function is usually dependent on the protein complexes they are part of (13). One potential application of magnetic microparticles is for the rapid analysis of protein-protein interactions (PPIs). Antibody-based pulldown in combination with mass spectrometry is a popular approach for elucidating protein-protein interactions (PPIs). We inquired whether a simplified microparticle based methods such as SP3 for analyzing PPIs provides depth of coverage comparable to standard in-gel protocols. To test this we utilized a SILAC (14) based setup for determining interactors of ZFP36 (Tristetraprolin or TTP), an RNA binding protein (Supplementary Figure 3A). On average we identified 100 more

proteins (438) when eluted proteins were aggregated on beads (with HILIC surface groups) based on single shot MS analysis compared to a conventional in-gel digestion workflow (330) with 5 fractions per replicate (Figure 4B, supplementary Figure 3B, supplementary table 4). Crucially, we found good overlap between the two groups for IL-1b regulated TTP-interactors (Figure 4C, supplementary table 4). We identified known interactors of TTP upon IL-1b stimulation such as 14-3-3 subunits and RNA-binding factors which regulate stability such as UPF1 and PABC1/4 (15)(Supplementary figure 3C), as well as potential novel interactors which displayed interesting interaction dynamics with TTP upon IL-1b stimulation (Supplementary figure 3D).

MS-based analysis of cellular secretome holds enormous promise for the investigation of cellular communication. However, the analysis of cellular secretome presents several challenges as proteins are usually secreted in low concentrations making their detection in culture media (which are rich in salts and other compounds) difficult (16). We sought to determine the applicability of microparticles for enriching secreted proteins in the background of cell culture media contaminants. To this end we benchmarked urea (in-solution) and filter (FASP) based methods as reported previously (17, 18), against PAC on microparticles for secreted proteins. As described above (Supplementary figure 1E, F), high concentration of microspheres (>300 µg/ml) were required in order to provide sufficient surface area for immobilization of dilute aggregated proteins. Protein aggregation on microspheres consistently identified the largest number of proteins and unique peptides resulting in the highest sequence coverage (Figure 4D, supplementary figure 4A, B, supplementary table 5). Although the low peptide recovery with FASP protocol led to the fewest number of protein identifications, the microsphere and FASP methods produced the cleanest peptide sample as determined by spectroscopy analysis (Supplementary Figure 4C, D). Using previously described computational workflow to predict potentially secreted proteins (18), we found majority (>50%) of the detected proteins have been previously characterized as secreted proteins through the classical

secretory pathway (via signal peptide) or other non-conventional pathways (Figure 4E)(19). Proteins secreted through the classical pathway were found at higher abundance relative in the media (Figure 4F). This includes low abundant cytokines such as Csf3 and Cxcl10 which were only identified using by aggregating proteins on magnetic microbeads, demonstrating sensitivity of the protocol. Thus, microparticle aggregation can enable simultaneous quantification of 100's of secreted proteins in the background of complex cell culture media.

## DISCUSSION

We have described the protein aggregation capture mechanism behind microsphere based protein immobilization. Understanding the mechanism of PAC in detail has led to optimized protocols that outperform competing methods for preparing samples for proteomic analysis of tissue, enriched subproteomes such as phosphoproteomes, low abundant immunoprecipitated and secreted proteins. The rates of missed cleavages were low (<30%) when proteins aggregated on-beads were digested with lys-c/trypsin. While the rates of missed cleavages were significantly higher for phosphopeptides (66-67%) using the PAC protocol, we saw no difference in the rate using the other methods tested. This can be explained by the fact that phosphate groups are known to impair protease digestion efficiency (20, 21). In the secretome analyses reported here, serum-containing media was replaced with serum-free media prior to stimulation which avoided the usual challenge of large dynamic range of high amounts of albumin that can obstruct detection of secreted proteins. Alternatively, newly synthesized secreted proteins could be enriched by bio-orthogonal amino acid incorporation in combination with protein aggregation on beads (22). We demonstrate that PAC is scalable from very low to high starting amounts of material, which has advantages for reducing cost, simplicity, and time. The workflow is unbiased for downstream sequencing of peptides and would be compatible with alternative novel protein sequencing technologies such as those that



utilize fluorophores labeling and sequencing of peptide mixtures (23). The PAC approach could potentially fail if samples contain certain components (such as acids) that could prevent efficient aggregation on microparticle surfaces. Other large biomolecules such as DNA/RNA can also co-precipitate with proteins if not adequately removed. We hope awareness of this mechanism will lead to further novel developments and applications. Future developments could utilize microparticle surface functional group specificities for peptide level enrichment/fractionation following non-specific aggregation at the protein level.

# EXPERIMENTAL PROCEDURES

## Reagents

Chemicals were purchased from Sigma-Aldrich unless otherwise specified. Chemicals were purchased from Sigma-Aldrich unless otherwise specified. 1  $\mu\text{m}$  diameter Sera-mag carboxyl magnetic beads (cas # 45152105050250 and cas # 65152105050250) were purchased from GE-Healthcare. 0.5  $\mu\text{m}$  diameter SIMAG-Sulfon (cas # 1202), SiMAG-Q (cas # 1206), and SiMAG-Octadecyl (cas # 1301) magnetic beads were all purchased from Chemicell GmbH. 5-10  $\mu\text{m}$  average diameter HILIC, TiO<sub>2</sub>, and Ti-IMAC magnetic beads were purchased from ReSyn Biosciences. Carbonyl-iron powder with 5-9  $\mu\text{m}$  diameter grain size was purchased from Sigma-Aldrich (cas # 44890).

## Cell culture

Human bone osteosarcoma epithelial (U2OS) and human epithelial cervix carcinoma (HeLa) adherent cells were grown in DMEM media (Gibco) supplemented with fetal bovine serum (Gibco) at 10% final. The media also contained penicillin (Invitrogen) at 50 U/mL and streptomycin (Invitrogen) at 100  $\mu\text{g}/\text{mL}$ . Cells were grown in a humidified incubator at 37°C with 5% CO<sub>2</sub>. In all cases, cells were grown to 80-90% confluency before harvesting with different lysis buffers in Nunc petridishes (100 or 150mm diameter).

To generate stably expressing GFP-TTP cells under a doxycycline inducible promoter, ZFP/TTP was gateway cloned into a pCDNA4/TO/GFP expression vector by gateway cloning (Thermo Fisher), and co-transfected with pcDNA6/TR (Thermo Fisher) into U2OS cells. Cells were selected with zeocin and blasticidin for 14 days, after which individual clones were picked and screened for GFP-TTP expression. For SILAC labelling, cells were cultured in media containing either L-arginine and L-lysine (Light), L-arginine [<sup>13</sup>C6] and L-lysine [<sup>2</sup>H4] (Medium) or L-arginine [<sup>13</sup>C6-<sup>15</sup>N4] and L-lysine [<sup>13</sup>C6-<sup>15</sup>N2] (Heavy; Cambridge Isotope Laboratories).

RAW264.7 macrophage cells were derived from *Mus musculus* and grown in 10% in DMEM media with 10% FBS in 150mm diameter Nunc petridishes. The media was removed and cells were washed with PBS prior to addition of phenol-red free DMEM media without serum, penicillin, and streptomycin. Cells were stimulated with lipopolysaccharids (LPS) with 1µg/ml for 4 hours. 400 µl of the media was removed and processed for secretome analysis and filtered through 0.22µm filter (Sartorius #16532) prior to further processing.

### **Cell lysis and sample preparation**

Cells lysis as presented in this study was performed with either one of the three buffers: 1) 6M guanidine hydrochloride in 100 mM tris hydrochloride (Life technologies) at pH 8.5, 2) 1% SDS in 100 mM 100 mM Tris Hydrochloride (pH 8.5) or 3) 0.1% NP-40 in 1X phosphate buffered saline solution (pH 7.4) containing β-glycerol phosphate (50 mM), sodium orthovanadate (10 mM), and protease inhibitor cocktail (Roche). In all cases, supernatant from adherent cell plates was removed and the cells were rinsed with ice cold 1X PBS prior to the addition of the lysis buffer.

Guanidine hydrochloride buffer was pre-heated to 99°C prior to the addition to the cell plates. After the addition of guanidine or SDS buffer, cells were manually collected and heated at 99°C for 10 minutes followed by sonication using a probe to shear RNA and DNA. Proteins were immediately reduced and alkylated with the 10 mM tris(2-carboxyethyl)phosphine (TCEP) and 11 mM 2-chloroacetamide (CAA) for samples lysed with guanidine hydrochloride or SDS lysis buffer prior to further processing. For cells lysed using 0.1% NP-40 buffer, 1µl of benzonase (≥250U/ul) was added to the lysis solution for 1 hour on ice. The lysis solution was centrifuged at 5000 x G for 10 minutes and the supernatant was transferred to a new tube.

GFP-TTP immunoprecipitations were performed using GFP-Trap magnetic agarose beads (Chromotec) according to manufacturer's instructions. Cell lysis and immunoprecipitations were carried out using low salt EBC lysis buffer (150 mM NaCl; 50 mM TRIS pH 7.5; 1 mM EDTA; 0,5% NP-40) for 1 hour followed by 5 washes with the same buffer.

Proteins were eluted by boiling in 2% SDS, 10% glycerol, 5% 2-mercaptoethanol, 0.1 M Tris-HCl (pH 6.8).

### **On-bead protein aggregation**

Aggregation was induced by the addition of acetonitrile (unless stated otherwise) and magnetic microparticles were added to solution followed by mixing in order to uniformly mix the bead solution. The solution was allowed to settle for 10 minutes and beads were separated using a magnet for 60 seconds. Magnetic microspheres were retained by magnet and the supernatant was removed by vacuum suction. In the case of analysis by protein gel electrophoresis (SDS-PAGE), the supernatant was transferred to new tubes. Beads were washed using acetonitrile once followed by one wash with 70% ethanol. See extended methods for experiment specific protocols. Samples and washes were prepared for analysis by protein gel electrophoresis (SDS-PAGE) by the addition of 4x LDS sample buffer (Thermo Fisher Scientific) to 1x final, and DTT (100 mM). Samples were heated for 10 minutes at 80°C. For eluting bead bound protein aggregates, LDS buffer (containing DTT) was added to bead containing solutions and the mixture was heated for 10 minutes at 80°C. Heated beads in LDS buffer were separated by magnet and the supernatant was analyzed by SDS-PAGE or transferred to a new tube and stored at -20°C until SDS-PAGE analysis. Samples were loaded on NuPAGE 4-12% Bis-Tris protein gel (Thermo Fisher Scientific) and ran with 200 volts for 40 minutes. Gels were stained for 15 minutes using instant Blue (Expedeon) and destained overnight with Milli-Q water and scanned on EPSON V750 PRO.

### **Lys-C and Trypsin digestion and peptide cleanup**

Proteins were aggregated on microspheres and washed as described above. For on-bead digestion, 50 mM HEPES buffer (pH 8.5) was added to submerge microspheres. Proteins were reduced and alkylated with the 5mM tris(2-carboxyethyl)phosphine (TCEP) and 5.5mM 2-chloroacetamide (CAA) for 30 minutes if not treated immediately after lysis. Lys-c (Wako

chemicals) was added at ratio of 1:200 (to protein) and allowed to react for 1 hour at 37°C followed by the addition of trypsin at a ratio of 1:100 (unless specified otherwise). Trypsin digestion was allowed to occur overnight at 37°C. Beads were separated by magnet and the supernatant was transferred to new tube and acidified.

In-solution digestion with guanidine hydrochloride buffer was carried out under similar reduction and alkylation conditions. Lys-c was added to solution and allowed to react for 1 hour at 37°C. The concentration of guanidine hydrochloride concentration was reduced to >1M prior to the addition of trypsin for overnight digestion. Solution was acidified by with 1% trifluoroacetic acid (TFA) and centrifuged for 5 minutes at 5000 x G and the supernatant transferred to new tubes.

Peptide mixtures were clarified with solid phase extraction. Briefly, hydrophobic C18 sep-pak (Waters Corporation) were prepared by washing with acetonitrile and 0.1% TFA, followed by loading of the acidified peptide mixtures by gravity. Sep-paks were washed with 0.1% TFA and peptides were eluted using 50% Acetonitrile (0.05% TFA). Organic solvent was evaporated and peptides concentrated using a speedvac prior to MS analysis.

### **Protein extraction from mouse skeletal tissue**

We used skeletal muscle which were isolated for previously published study (Schonke et al 2018, Proteomics). Frozen gastrocnemius muscles were crushed using mortar and pestle. Powdered muscle was homogenized using Ultra Turrax T8 homogenizer (IKA Labortechnik) in 4% SDS buffer (100 mM Tris-HCl, pH 7.4). Protein lysates were boiled at 100 °C for 5 mins. Lysates were sonicated using a tip and centrifuged at 16000g for 10 mins followed by reduction and alkylation as described above, the supernatant was then processed using FASP or PAC or frozen until further analysis.

### **Filter-aided sample preparation (FASP)**

Urea powder was added to 400µL of filtered cell culture supernatant for a final 2M concentration, and pH for digestion adjusted with 40µL Tris 1M pH 8.5. FASP protocol was

adapted from as previously described<sup>1</sup>. Samples were heated for 10 min at 56°C and centrifuged (7000g, 10min). Following centrifugation steps were performed applying the same conditions. Ultracel-30 membrane filters (Millipore #MRCF0R030) were cleaned with 10% acetonitrile and 15% methanol, filters were centrifuged and equilibrated with 200µL urea buffer (2M, 0.1M Tris, pH8.5), and centrifuged again. Samples were added into the filters and the filters were centrifuged and washed two times with urea buffer. Reduction was performed by 1µL of 0.5M TCEP in 100µL urea buffer. The device was centrifuged and alkylation was performed by 1µL of 550mM CAA in 100µL urea buffer for 30 minutes in the dark. Filters were centrifuged and 200µL urea buffer was added prior to another centrifugation. Subsequently, 4µL of 0.5 ug/µL lys-c in 40µL urea buffer was added for 3h at 37°C with gentle orbital shaking. 4µL of 0.5µg/µL trypsin was added for an overnight digestion in the wet chamber at 37°C with gentle orbital shaking. 1.5mL eppendorf tubes were cleaned with absolute methanol and air dried, prior to inserting the filter device, which was then centrifuged. Subsequently 40µL of milli-Q water was added followed by centrifugation. The enzymatic digestion was stopped by acidifying the sample to pH<2.5 with TFA. StageTipping was performed right after.

### **In-solution urea digestion sample preparation**

All following chemicals have the same references and concentrations as in the FASP sample preparation. Urea powder was added to 400µL of filtered (0.22uM) cell culture supernatant for a final 2M concentration, and pH for digestion adjusted with 40µL Tris 1M pH8.5. 100% v/v. TCEP was added and the tubes were incubated for 30 minutes. Subsequently, samples were incubated with CAA for 20 minutes in the dark. The digestion step included addition of lys-c , incubation during 3h, followed by the addition trypsin (0.5ug/µL), and incubation overnight at room temperature. The enzymatic digestion was stopped by acidifying the sample to pH<2.5 with TFA. Samples were desalted and concentrated using Stage-Tips.

### **In-gel digestion**

In-gel protein digestion and downstream processing was performed as described earlier (Lundby and Olsen 2011, see references in extended supplementary methods).

### **Enrichment of phosphorylated peptides**

Adherent HeLa cells were grown as described above. Cells were washed and serum starved (DMEM without FBS) for 4 hours followed by 10 minute stimulation with FBS (10%). Cells were rapidly washed and lysed using guanidine hydrochloride buffer as described above. Protein concentration was estimated using tryptophan assay. Lys-C and trypsin digestion was carried out as described above. Peptides were clarified using SPE as described above with the exception that the peptide mixture was not concentrated using a speedvac. Small aliquot representing 5% was removed for determining peptide concentration using nanodrop which was estimated to roughly 200µg.

Ultra high phosphopeptide enrichment efficiency was achieved using Ti-IMAC magnetic beads (ReSyn Biosciences) with slight modification to the manufacturer protocol (see extended supplementary methods). Phosphopeptide containing solution were loaded onto C18 STAGE-tips where the phosphopeptides were loaded and washed. The STAGE-tips were stored at 4°C until elution and analysis by MS.

### **Mass spectrometry and liquid chromatography**

Samples were injected on a 15 cm nanocolumn (75µM inner diameter) packed with 1.9µM C<sub>18</sub> beads (Dr. Maisch GmbH) using an Easy-LC 1200 (Thermo Fisher Scientific). Peptides were separated and eluted from the column with an increasing gradient of buffer B (80% acetonitrile, 0.1% formic acid) at a flow rate of 250 nL/minute.

All samples were analyzed on a Q-Exactive HF-X (Thermo Fisher Scientific) mass spectrometer coupled to EASY-nLC 1200. With the exception of two replicates of in-gel TTP pulldown and one replicate from PAC TTP pulldown experiments were analyzed on a Lumos

(Thermo Fisher Scientific) mass spectrometer with similar scan settings. The mass spectrometer was operated in positive mode with TopN method.

## Data analysis

All mass spectrometric data are available via ProteomeXchange with identifier PXD011677. Raw files generated from LC/MS/MS experiments were analyzed using MaxQuant (1.6.1.1) software (Cox and Mann 2008). Samples generated from human cell lines (HeLa and U2OS) were searched against the reviewed Swiss-Prot human proteome (proteome ID: UP000005640, release date March 2018) with 21006 entries. Samples generated from mouse cell lines (Raw264.7) and tissue were searched against the *Mus musculus* reviewed Swiss-Prot proteome (proteome ID: UP000000589, release date October 2018) with 22325 entries. The protease specificity was set to “Trypsin/P” with maximum number of missed cleavages set to 2 with the exception for the analysis of protease digestion efficiency experiment, where it was set to 9 with “semi-tryptic” search. All searches were performed with carbamidomethyl of cysteines as a fixed modification while methionine oxidation and protein n-terminus acetylation were set as variable. Phosphorylation of serine, threonine, and tyrosine were set as variable modification for analysis of phosphopeptide enriched samples using Ti-IMAC. Maximum number of modifications was set to 5 for all analysis. Mass tolerance of 20 parts per million (ppm) was set to the first search of precursor ions followed by 4.5 ppm for main search after mass calibration. 20 ppm mass tolerance was set for fragment ion series. Minimum peptide length of 7 amino acids was required for all identifications and modified peptides required a minimum Andromeda score of 40 be considered for identification. A false discovery rate (FDR) of 1% was utilized for peptide spectral matches, peptides, and proteins. Proteins had to be identified by minimum of 2 peptides in order to be counted. Match between runs feature was utilized only for the analysis of phosphopeptide enriched samples.



## Experimental Design and Statistical Rationale

Number of replicates were denoted by “n = x” for all results where statistical analysis was performed and marked in the figures. Figure legends provided further clarification of statistical tests and criteria for determining significance in each case. Analysis of protease efficiency were performed in duplicates for the two digestion methods (in solution vs PAC) with 25 different protease conditions for each, resulting in the analysis of 100 samples. One replicate from PAC digestion with no Lys-C and Trypsin at 1:50 ratio was discarded due to experimental error. Same biological source was utilized in order to limit the variation only to the sample preparation methods. For phosphoproteomics analysis (as presented in Figure 3B,C,D), same HeLa protein extract was utilized in order to limit the variation to the sample preparation conditions. The protein extract was equally aliquoted 8 times for quadruplicate analysis of digestion methods (in solution vs PAC) leading to 8 different samples. Each sample was independently prepared (in solution or PAC) followed by independent enrichment of phosphopeptides from each sample after protease digestion. Skeletal tissue protein extract was aliquoted 6 times for triplicate analysis of peptide recovery and proteomics analysis between FASP and PAC. Each replicate was prepared and analyzed separately. SILAC analysis of ZFP36 (Tristetraprolin or TTP) interactors was performed in quadruplicates by growing cells in light, medium, and heavy states in 4 different cell culture plates (for a total of 12 different plates) and mixed to produce 4 separate samples from which ZFP36 interactors were determined. Elution from GFP-Trap beads were evenly split into half for either in-gel or PAC analysis. Quantile normalization was utilized to synchronize protein intensities and SILAC ratio distributions across replicates and experiments (see extended methods, Amaratunga et al. and Bolstad et al). Secretome analysis of RAW264.7 macrophage cells was performed in quadruplicates by growing the cells in 4 separate cell culture petri dishes. The supernatant from each replicate was prepared independently for proteomics analysis by FASP, in solution, or PAC method.

## ACKNOWLEDGEMENTS

We would like to thank Prof. Christian Kelstrup for constructive discussion and feedback. Work at The Novo Nordisk Foundation Center for Protein Research (CPR) is funded in part by a generous donation from the Novo Nordisk Foundation (Grant number NNF14CC0001). The proteomics technology developments applied was part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 686547 (MSmed). We would like to thank the PRO-MS Danish National Mass Spectrometry Platform for Functional Proteomics and the CPR Mass Spectrometry Platform for instrument support and assistance. T.S.B is funded by the HOPE project grant from the Novo Nordisk Foundation (Grant number NNF17SA0027704). P.L.R. is supported by the Marie Skłodowska-Curie European Training Network (ETN) "TEMPERA", a project funded by the European Union's EU Framework Programme for Research and Innovation Horizon 2020 (Grant Agreement number 722606). Work in the Bekker-Jensen lab was supported by grants from the Lundbeck Foundation, The NEYE Foundation and The Danish Medical Research Council.

## DATA AVAILABILITY

All mass spectrometric data are available via ProteomeXchange with identifier PXD011677.

Reviewer account details:

Username: reviewer94929@ebi.ac.uk

Password: 26B84GDh

Annotated spectra of post-translationally modified peptides can be found using the MS-Viewer here (The search key for the saved data set is **lj9dbrbqem**):

[http://msviewer.ucsf.edu/prospector/cgi-bin/mssearch.cgi?report\\_title=MS-Viewer&search\\_key=lj9dbrbqem&search\\_name=msviewer](http://msviewer.ucsf.edu/prospector/cgi-bin/mssearch.cgi?report_title=MS-Viewer&search_key=lj9dbrbqem&search_name=msviewer)

## AUTHOR CONTRIBUTIONS

T.S.B and J.V.O conceived and designed the project. T.S.B developed the workflows and carried out the experiments. M.A.X.T generated the stable U2OS cell lines expressing GFP-TTP

and performed the experiments and GFP-TTP pulldowns. M.A.X.T and S.B-J provided the Raw264.7 cell lines used for secretomics analysis and assisted in the data analysis of SILAC GFP-TTP pulldown experiments. P.L.R carried out the experiment for determining protease digestion efficiency and performed the data analysis. A.G.F and A.S.D performed the secretome experiments using FASP and in-solution workflows and assisted in the downstream data analysis of secreted proteins. A.S.D provided the *Mus musculus* skeletal muscle tissue and assisted with the experiments. B.S.P performed the SDS-PAGE experiments. T.S.B wrote the initial draft of the manuscript and all authors contributed to the manuscript.

## CONFLICTS OF INTEREST

Authors declare no competing interests.

## REFERENCES

1. Wiśniewski JR, Zougman A, Nagaraj N, Mann M. 2009. Universal sample preparation method for proteome analysis. *Nat Methods* 6:359.
2. Shevchenko A, Tomas H, Havli J, Olsen JV, Mann M. 2007. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* 1:2856–2860.
3. Zougman A, Selby PJ, Banks RE. 2014. Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. *Proteomics* 14:1006–1000.
4. Hughes CS, Foehr S, Garfield DA, Furlong EE, Steinmetz LM, Krijgsveld J. 2014. Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol Syst Biol* 10:757–757.
5. Alpert AJ. 1990. Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. *J Chromatogr A* 499:177–196.
6. Sielaff M, Kuharev J, Bohn T, Hahlbrock J, Bopp T, Tenzer S, Distler U. 2017. Evaluation of FASP, SP3, and iST Protocols for Proteomic Sample Preparation in the Low Microgram Range. *J Proteome Res* 16:4060–4072.
7. Moggridge S, Sorensen PH, Morin GB, Hughes CS. 2018. Extending the Compatibility of the SP3 Paramagnetic Bead Processing Approach for Proteomics. *J Proteome Res* 17:1730–1740.
8. Hughes, C.S. et al., 2019. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nature protocols*, 14(1), pp.68–85.

9. Poulsen JW, Madsen CT, Young C, Poulsen FM, Nielsen ML. 2013. Using Guanidine-Hydrochloride for Fast and Efficient Protein Digestion and Single-step Affinity-purification Mass Spectrometry. *J Proteome Res* 12:1020–1030.
10. Ludwig KR, Schroll MM, Hummon AB. 2018. Comparison of In-Solution, FASP, and S-Trap Based Digestion Methods for Bottom-Up Proteomic Studies. *J Proteome Res* 17:2480–2490.
11. Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, Mortensen P, Mann M. 2006. Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. *Cell* 127:635–648.
12. Wiśniewski JR, Gaugaz FZ. 2015. Fast and Sensitive Total Protein and Peptide Assays for Proteomic Analysis. *Anal Chem* 87:4110–4116.
13. Huttlin EL, Ting L, Bruckner RJ, Gebreab F, Gygi MP, Szpyt J, Tam S, Zarraga G, Colby G, Baltier K, Dong R, Guarani V, Vaites LP, Ordureau A, Rad R, Erickson BK, Wühr M, Chick J, Zhai B, Kolippakkam D, Mintseris J, Obar RA, Harris T, Artavanis-Tsakonas S, Sowa ME, De Camilli P, Paulo JA, Harper JW, Gygi SP. 2015. The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell* 162:425–440.
14. Ong S-E, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. 2002. Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Mol Cell Proteomics* 1:376–386.
15. Wu X, Brewer G. 2012. The regulation of mRNA stability in mammalian cells: 2.0. *Gene* 500:10–21.
16. Chevallet M, Diemer H, Van Dorssealer A, Villiers C, Rabilloud T. 2007. Toward a better analysis of secreted proteins: the example of the myeloid cells secretome. *Proteomics* 7:1757–1770.

17. Meissner F, Scheltema RA, Mollenkopf H-J, Mann M. 2013. Direct Proteomic Quantification of the Secretome of Activated Immune Cells. *Science* 340:475–478.
18. Deshmukh AS, Cox J, Jensen LJ, Meissner F, Mann M. 2015. Secretome Analysis of Lipid-Induced Insulin Resistance in Skeletal Muscle Cells by a Combined Experimental and Bioinformatics Workflow. *J Proteome Res* 14:4885–4895.
19. Butler GS, Overall CM. 2009. Proteomic identification of multitasking proteins in unexpected locations complicates drug targeting. *Nat Rev Drug Discov* 8:935–948.
20. Batth TS, Francavilla C, Olsen JV. 2014 Off-line high-pH reversed-phase fractionation for in-depth phosphoproteomics. *J. Proteome Res.* 13, 6176–6186.
21. Dickhut C, Feldmann I, Lambert J, Zahedi RP. 2014. Impact of digestion conditions on phosphoproteomics. *J. Proteome Res.* 13, 2761–2770.
22. Eichelbaum K, Winter M, Diaz MB, Herzig S, Krijgsveld J. 2012. Selective enrichment of newly synthesized proteins for quantitative secretome analysis. *Nat. Biotechnol* 30, 984–990.
23. Swaminathan J, Boulgakov AA, Hernandez ET, Bardo AM, Bachman JL, Marotta J, Johnson AM, Anslyn EV, Marcotte EM. 2018. Highly parallel single-molecule identification of proteins in zeptomole-scale mixtures. *Nat Biotechnol*. doi:10.1038/nbt.4278.

## FIGURE LEGENDS

### **Figure 1. Protein aggregation driven immobilization on microparticles**

A) The hypothesis of HILIC based bead interactions was tested by inducing bead – protein interaction on 20µg of HeLa protein lysate (in 1% SDS) with the addition of acetonitrile (70% final concentration) and carboxyl coated microparticles (20µg of Sera-mag beads) separated by magnet. The resulting supernatant was analyzed by SDS gel electrophoresis. Beads were sequentially washed three times with the indicated buffers. All washes were analyzed by SDS-PAGE for protein elution by beads including washes by milli-Q water and the lanes indicated in red. After bead washing, LDS buffer was added to the beads (see materials section) and analyzed. B) U2OS protein lysates (in 0.1% NP-40) were treated with different conditions as indicated in green. Carboxyl coated magnetic beads were added to the lysates after each treatment and the resulting supernatant analyzed by SDS-PAGE.

### **Figure 2. Elucidating the mechanism of protein aggregation capture on microparticles**

A) Acetonitrile was added to HeLa lysate (in 1% SDS) to a final concentration of 70% and equal amounts of microparticles with different surface chemistries were added to the lysates and the supernatant removed. LDS buffer was added to the microparticles and the resulting supernatant analyzed by SDS-gel after removal by magnet. B) Aggregation of equal amount of HeLa lysate (20µg at 0.25µg/µl after addition of acetonitrile) was induced in a similar fashion as indicated above and carboxyl coated microparticles were added to the lysate at different amounts as indicated in the figure. The supernatant was removed and the LDS buffer was added to the different samples and analyzed by SDS-PAGE after separation of microparticles by magnet. C) The hypothesized model for protein aggregation capture (PAC) on microparticles is illustrated based on the above observations.

**Figure 3. Effects of protease digestion and post-translation modification (PTM) analysis of proteins immobilized on microparticles.**

A) Average (based on duplicates) percentage of peptides containing missed cleavages at arginine or lysine after digestion with trypsin and lys-c proteases in different combinations and ratios are displayed by heatmap. Missed cleavage rates were investigated on lysates prepared by protein aggregation on microspheres or in-solution digestion. B) Average number of unique phosphopeptide variants were counted (after removal of contaminant or reverse hits as defined by MaxQuant analysis) for the different experiments. Phosphopeptides were tallied after enrichment from lysates prepared with in-solution or PAC digestion. C) Average number of phosphorylation sites with high localization probabilities (as defined by site localization probability  $\geq 0.75$ ) are presented between the two different methods. D) Overlap of localized phosphorylation sites (localization probably  $\geq 0.75$ ) between the two experiments. The site had to be identified in two of the four replicates in both experiments for it to be valid.

**Figure 4. Exploring the boundaries of PAC for proteomics analysis of different sample types.**

A) Average peptide recovery after protease digestion of equal amounts (~1.8 mg) of mouse skeletal tissue prepared using the FASP or PAC protocol measured by nanodrop absorbance at 280/260nm. B) The average number of proteins with SILAC ratios were counted between the two different experiments after removal of contaminating proteins and reverses hits. C) Overlap of statistically regulated protein between two preparation methods as determined by t-test with permutation-based FDR of 0.05 with 250 randomizations and s0 of 0.1 (Supplementary table 4). D) Number of proteins containing LFQ intensities from the different secretome workflows. E) Number of secreted proteins between the different experiments (see Methods). F) Median Log10 transformed LFQ intensities of the proteins identified in PAC are plotted against those



identified in the base proteome of Raw264.7 cells. Classical and non-classical secreted proteins are highlighted. \*Error bars represent standard deviation in all cases.

# FIGURES

Figure 1

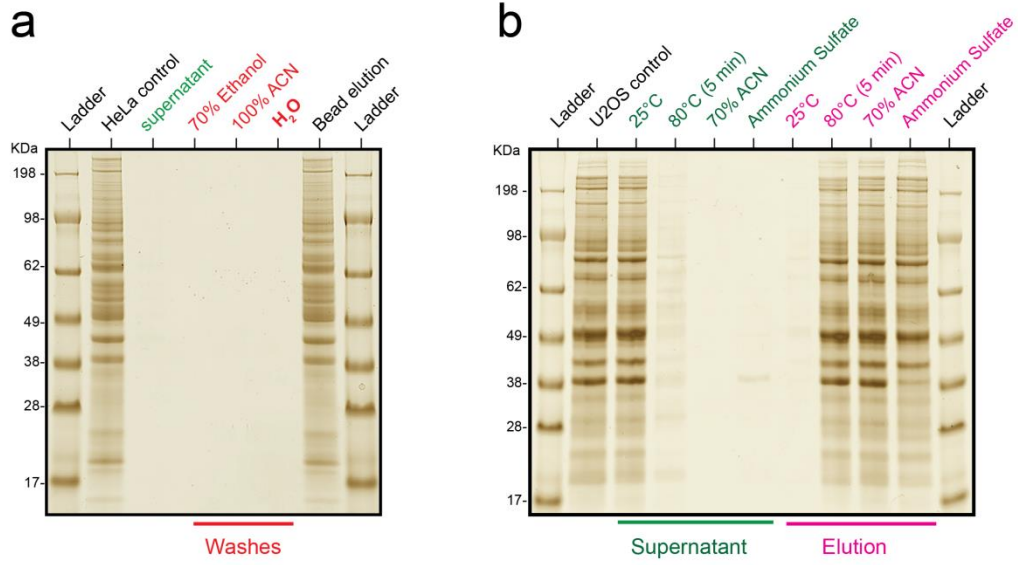


Figure 2

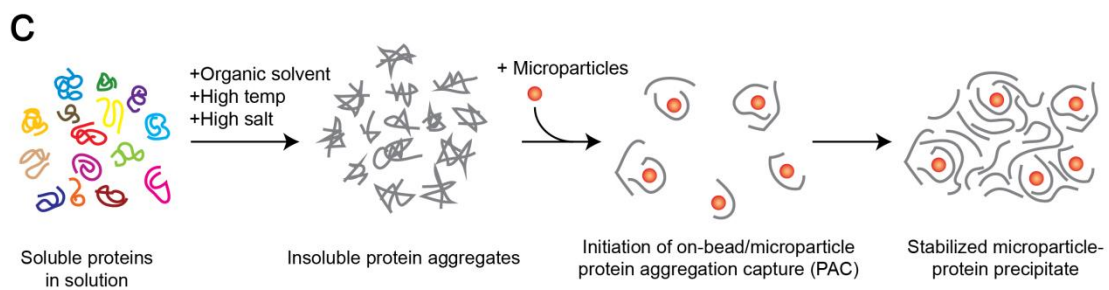
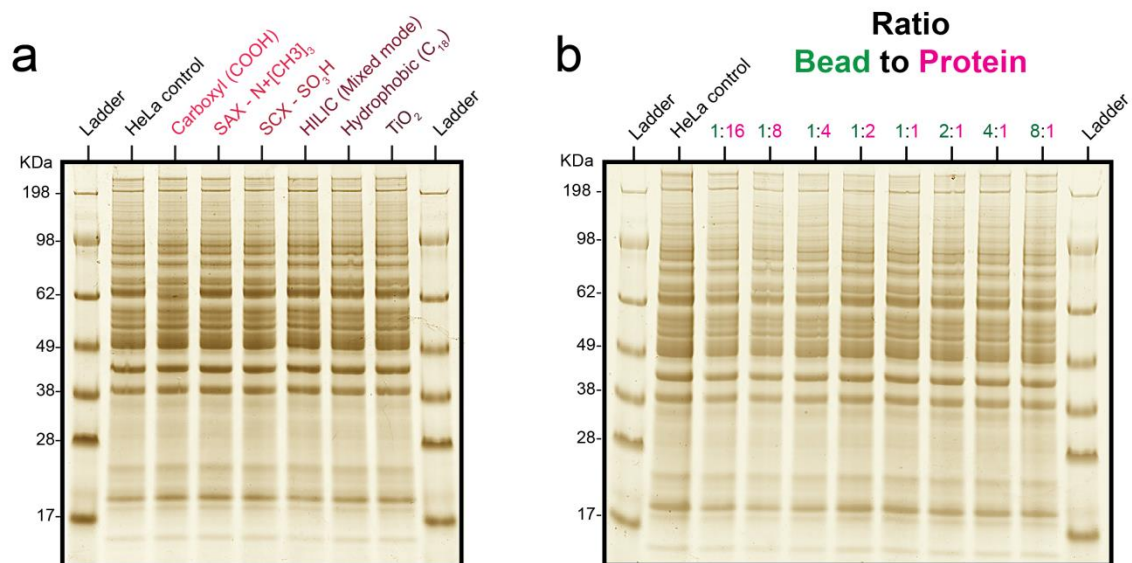


Figure 3

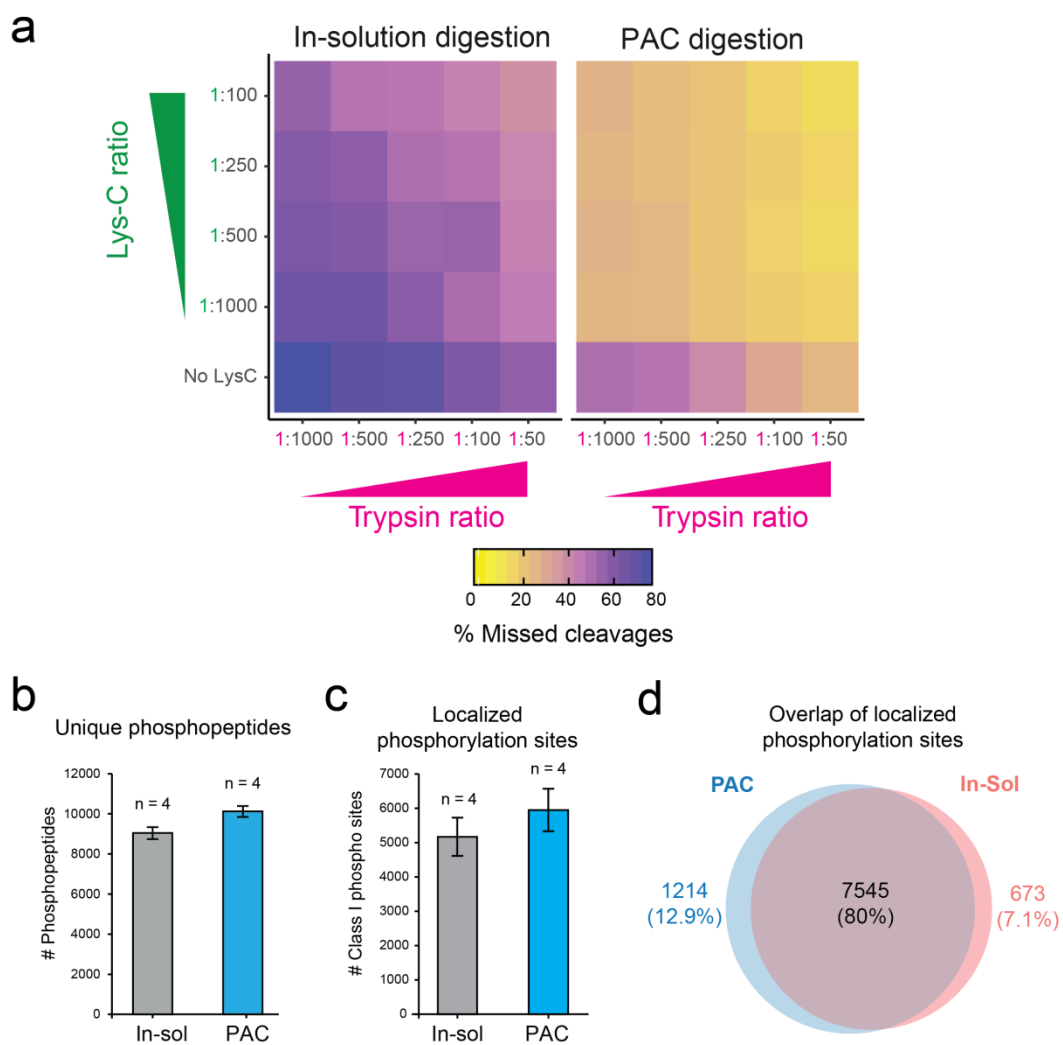


Figure 4

