*Electronic Supplemetary Information*

# Coupling Functionalized Cobalt Ferrite Nanoparticle Enrichment with Online LC/MS/MS for Top-down Phosphoproteomics

Bifan Chen[a, ⊥], Leekyoung Hwang[a, ⊥], William Ochowicz[a], Ziqing Lin[b], Tania M. Guardado-Alvarez[a], Wenxuan Cai[b], Lichen Xiu[a], Kunal Dani[a], Cyrus Colah[b], Song Jin[a], and Ying Ge[a,b,c*]

[a] Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin, USA

[b] Department of Cell and Regenerative Biology, University of Wisconsin-Madison, Madison, Wisconsin, USA

[c] Human Proteomics Program, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, Wisconsin, USA

[⊥] These authors contributed equally to this work.

*To whom correspondence may be addressed: Dr. Song Jin, 1101 University Avenue, Madison, WI 53706, USA. E-mail: jin@chem.wisc.edu; Tel: 608-262-1562; Fax: 608-262-0453. Dr. Ying Ge, 1300 University Ave., SMI 130, Madison, WI 53706, USA. E-mail: ge2@wisc.edu; Tel: 608-263-9212; Fax: 608-265-5512.

## Part I. Materials and methods

### Chemicals and reagents

All chemicals and reagents were used as received without further purification unless otherwise noted. N-(tert-Butoxycarbonyl)-L-tyrosine methyl ester (Boc-Tyr-OMe), di-(2-picolyl)amine (DPA), paraformaldehyde, trifluoroacetic acid (TFA), glutaric anhydride, N-ethyldiisopropylamine (EDIPA), Iron (III) acetylacetonate (acac) (97%), benzyl ether (98%), 1,2-hexadecanediol (90%), oleic acid (OA) (99%), oleylamine (OE) (>70%), zinc chloride ($ZnCl_2$), anhydrous N,N-dimethylforamide (DMF), and anhydrous dichloromethane were purchased from Sigma-Aldrich (St. Louis, MO, USA). ***N,N,N′,N′*-Tetramethyl-*O*-(1*H*-benzotriazol-1-yl) uronium hexafluorophosphate** (HBTU) were obtained from TCI America (Portland, OR, USA). (3-aminopropyl)trimethoxysilane (APTMS) and 2-methoxy (polyethyleneoxy) propyl trimethoxysilane (average molecular weight =550-600, hereafter referred as Si-PEG) were purchased from Gelest (Morrisville, PA, USA). Bovine serum albumin (BSA), β-casein, and pepsin were purchased from Sigma-Aldrich. For the enrichment of β-casein for further top-down MS analysis, β-casein (MS grade) was purchased from Protea (Morgantown, WV, USA). All proteins were used without further purification or modification. 50 mM HEPES (4-(2-hydroxyethyl)-piperazineethanesulfonic acid) buffer was made by directly diluting 0.5 M HEPES buffer (pH=7.7; home-made with HEPES obtained from Fluka) with 18 MΩ nanopure de-ionized water from Milli-Q water (Millipore, Corp., Billerica, MA, USA). 100 mM sodium phosphate buffer was prepared by diluting 0.5 M sodium phosphate (pH 7.3; home-made with sodium phosphate dibasic obtained from Sigma-Aldrich). Bradford protein assay reagent was purchased from Bio-Rad (Hercules, CA, USA). 12.5% mini-gel (10 comb well, 10.0 cm × 10.0 cm) for SDS-

Polyacrylamide Gel Electrophoresis (SDS-PAGE) was home-made. 12.5% Criterion precast gel (18 comb well, 13.3 cm × 8.7 cm (W×L)) was purchased from Bio-Rad (Hercules, CA, USA). Pierce™ Concentrators PES (polyethersulfone)-0.5 mL centrifugal filter with a molecular-weight cutoff of 10k was purchased from ThermoFisher Scientific (Rockford, IL, USA). Sypro Ruby and Pro-Q Diamond staining solutions were purchased from ThermoFisher Scientific (Rockford, IL, USA). Phos-tag gel staining solution was purchased from GeneCopoeia (Rockville, MD, USA). 3,4-hexanedione was perchased from Sigma-Aldrich.

**Synthesis of functionalized $CoFe_2O_4$ nanoparticles (NPs)**

**$CoFe_2O_4$-OA/OE NPs.** The cobalt ferrite NPs were prepared by a seed-mediated growth method following the procedure reported by Sun *et.al*.[1] Briefly, $Co(acac)_2$ (1 mmol), $Fe(acac)_3$ (2 mmol), 1,2-hexadecanediol (10 mmol), OA (6 mmol), and OE (6 mmol) were mixed in benzyl ether (20 mL) to synthesize $CoFe_2O_4$ seed NPs. The mixture was heated to 200 °C for 2 h with stirring under $N_2$ (g), and then heated to reflux at ~300 °C for 1 h. After the mixture was cooled to room temperature, EtOH (40 mL) was added to the mixture. A black precipitate was obtained by centrifugation, which was subsequently redissolved in hexane in the presence of OA (50 μL) and OE (50 μL). Undissolved residue was removed by centrifugation (5000 rpm, 20 min). The supernatant was precipitated with EtOH, centrifuged (5000 rpm, 20-30 min) to remove the solvent, and then redispersed in hexane. ~80 mg of the $CoFe_2O_4$ seed NPs was further grown to a bigger size of $CoFe_2O_4$ NPs by adding $Co(acac)_2$ (1 mmol), $Fe(acac)_3$ (2 mmol), 1,2-hexadecanediol (10 mmol), OA (2 mmol), and OE (2 mmol) in benzyl ether (20 mL). The mixture was heated to 100 °C for 30 min and then to 200 °C for 1h with stirring under $N_2$ (g), and then heated to reflux at ~300 °C for 30 min. After the mixture was cooled to room temperature, EtOH (40 mL) was added

to the mixture. As described in the synthesis of $CoFe_2O_4$ seed NPs, the workup procedures were repeated to obtain $CoFe_2O_4$-OA/OE NPs with a size of ~10 nm.

**$CoFe_2O_4$-$NH_3$/PEG NPs and $CoFe_2O_4$-GAPT NPs.** The procedures of NP functionalization by trialkoxysilane molecules and GAPT ligand molecules have been described in the supporting material of our previous communication with minor modifications.[2] Briefly, under ambient conditions, 0.5% (v/v) trialkoxy silane solution in hexane containing APTMS and Si-PEG in a ~2:1 molar ratio was added to the $CoFe_2O_4$ NPs dispersed in hexane (~ 50 mg in 120 mL). The mixture was stirred and refluxed at 70-72 °C for 24 h, during which a black-brown precipitate was formed. The precipitate was washed three times with hexane and three times with EtOH to remove excess silane molecules in solution. Finally, the NPs were dried for the next reaction.

GAPT (0.546 mmol) was dissolved in anhydrous DMF (~6 mL), and then HBTU (0.82 mmol) and EDIPA (3.28 mmol) were added to the solution. The $CoFe_2O_4$-$NH_3$/PEG NPs were redispersed in DMF (~2 mL). EDIPA (0.7 mmol) was added to this NP solution and it was sonicated for 1 min. Then the DMF solution containing GAPT, HBTU, and EDIPA was added, and the whole mixture was stirred at room temperature for 24 h. The resulting black-brown precipitate was washed two times with DMF, one time with DMF/MeOH (v/v 3:1), one time with DMF/MeOH (v/v 1:1), and one time with MeOH to remove unreacted GAPT, HBTU, and EDIPA. The resulting $CoFe_2O_4$-GAPT NPs were redispersed in EtOH at a concentration of ~50 mg/mL.

**$CoFe_2O_4$-GAPT-Zn NPs.** A 120-150 μL aliquot of $CoFe_2O_4$-GAPT NP stock solution (at ~50 mg/mL) was transferred to a 2 mL vial. The NPs (~6.0-7.5 mg) were washed with de-ionized water three times to remove EtOH. After 1.5 mL of 10 mM $ZnCl_2$ (aq) was added to the $CoFe_2O_4$-GAPT NPs, the mixture was agitated for 3.5 h at room temperature. The $CoFe_2O_4$-GAPT-Zn NPs were precipitated by a magnet (centrifugation was also used to collect the NPs if needed) and then

sonicated in de-ionized water to re-disperse. The final product, $CoFe_2O_4$-GAPT-Zn NPs, was used for enriching phosphoproteins.

**Characterizations of the as-synthesized and functionalized NPs.** Transmission electron microscopy (TEM) samples were prepared by pipetting one drop of as-synthesized $CoFe_2O_4$-OA/OE NP solution onto a copper TEM grid with carbon film. TEM was conducted on Philips CM200 Ultra Twin instrument operated at 200 kV and images were collected using a Gatan CCD image system with digital micrograph software program. Powder X-ray diffraction (PXRD) data were collected on as-synthesized $CoFe_2O_4$-OA/OE NPs deposited on a glass substrate using a Bruker D8 Advance using Cu Kα radiation (λ= 1.5418 Å), and the background from the glass substrate was subtracted. Transmission Fourier transform infrared (FT-IR) spectroscopy measurements were recorded on a Bruker Equinox 55 FT-IR spectrometer in the range of 4000 $cm^{-1}$ to 600 $cm^{-1}$ at about 2 $cm^{-1}$ resolution by using potassium bromide (KBr) pellet. Thermogravimetric analysis (TGA) was carried out using a TGA Q500 thermal analysis system. All TGA measurements were performed under a $N_2$ atmosphere at a constant heating rate of 20 °C/min from 100 °C to 600 °C. All samples were first heated to 100 °C and held at that temperature for 3 min to remove adsorbed water. Data were analyzed using the Thermal Advantage (TA) universal analysis software program. Dynamic light scattering (DLS) was carried out using a Malvern Zetasizer Nano. All DLS measurements of individual samples were performed three times after 120 sec equilibration.

**Protein sample preparation.** Each of the standard proteins BSA, β-casein, and pepsin was dissolved in 50 mM HEPES buffer (pH 7.7, with 150 mM NaCl) at a concentration of about ~15-20 mg/mL to be used as stock solutions. For the test of enrichment performance of $CoFe_2O_4$-

GAPT-Zn NPs, simple mixtures containing 300 μg of BSA, 100 μg of β-casein , and 100 μg of pepsin in 1.0 mL of 50 mM HEPES buffer (pH 7.7 ± 0.1, 150 mM NaCl) were prepared.

Swine hearts were excised from healthy Yorkshire domestic pigs, snap frozen in liquid $N_2$, and stored under -80 °C before use as approved by the University of Wisconsin Animal Care and Use Committee. First, the frozen pig tissue samples (~1 g) were cut into small pieces and immediately washed twice in cold 25 mM HEPES buffer (pH 7.4). The tissue samples were homogenized 5-6 times using a Polytron electric homogenizer (Model PRO200, PRO Scientific Inc., Oxford, CT, USA) for 5-7 secs on ice in buffer containing 25 mM HEPES (pH 7.4), 50 mM NaF, 0.5 mM phenylmethanesulfonyl fluoride (PMSF), 2.5 mM EDTA, and protease inhibitor cocktail (Roche, Switzerland). After centrifugation, the remaining pellet was briefly homogenized two times in the same buffer followed by 30 min of centrifugation at 16,000 ×g at 4 °C. The supernatants were collected, the concentration of the tissue lysate was determined by Bradford protein assay. The samples were stored at -80 °C for later study.

The swine heart tissue extract was diluted in loading buffer consisting of 50 mM HEPES buffer (pH 7.7) with 150 mM NaCl to achieve a final protein concentration of ~0.8-0.9 mg/mL (based on Bradford protein assay). Various concentrations of β-casein solution were spiked into the complex tissue extract solution.

**Verification of enrichment by SDS-PAGE analysis.** For standard protein mixtures, the combined elution fractions were concentrated to ~100 μL and then the protein solutions before and after enrichment were separated by SDS-PAGE (12.5% mini-gel). For swine heart tissue extract containing varying amounts of the spiked-in β-casein, 10 μg of protein mixtures before and after enrichment were separated on a Criterion Tris-HCl (10.5-14%) precast gel (Bio-Rad). The gels were first stained with Phos-tag or Pro-Q diamond solution for visualizing phosphoporteins and

then Sypro Ruby solution for total proteins. The gel images were taken by a Typhoon 9200 imager (GE Healthcare, Bio-Science, Piscataway, NJ, USA) with excitation at 532 nm and emission filter of 580 nm or a ChemiDoc™ MP Imaging System (Bio-Rad, Hercules, CA, USA).

**Western blot.** Equal amount of samples from loading mixture, flow through, and elution was loaded and sepeartated by SDS-PAGE, and was subsequently transferred to PVDF membranes (Millipore, Billerica, MA, USA). Membranes were then blocked with protein-Free Blocking buffer (Thermo Scientific) and bloted with antibodies against cTnI phosphorylated at Ser 22/23 (Cell Signaling Technology, Beverly, MA, USA).

**Reverse phase chromatography (RPC).** Reverse phase chromatography was performed with an ACQUITY M-Class UPLC system (Waters; Milford, MA, USA). Mobile phase A (MPA) contained 0.1% formic acid in water, and mobile phase B (MPB) contained 0.1% formic acid in 50/50 ethanol/acetonitrile. For each injection, 5 µL of sample was loaded on a PLRP-S (5 µm, 1000Å) capillary trap column (nanoLCMS LLC; Gold River, CA, USA) for online desalting with 95% MPA at 6 µL/min for 4 minutes. A home-packed 200 mm × 250 µm PLRP-S (5 µm, 1000Å; Agilent Technology, Santa Clara, CA, USA) column was used at a constant 4 µL/min flow rate. The RPC gradient consisted of the following concentrations of MPA: 95% for 2 mins, 20% at 60 min, 15% at 61 min, and back to 95% at 70 min. Each run was 75 min long.

**Top-down MS and MS/MS analysis.** Samples eluted from RPC separation were electrosprayed into a maXis II ETD Q-TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) for online LC/MS and LC/MS/MS experiments. End plate offset and capillary voltage were set at 500 and 4000 V, respectively. The nebulizer was set to 0.5 bar, and the dry gas flow rate was 4.0 L/min at 220 °C. The quadrupole low mass was set to 600 *m/z* during MS and 200 *m/z* during MS/MS. Mass range was set to 200-3000 *m/z* and spectra were acquired at 1 Hz for LC/MS

runs. For the top 3 data-dependent LC/MS/MS CID runs, spectra were acquired across 200-3000 *m/z* at 2 – 6 Hz with active exclusion after 4 spectra. Targeted LC/MS/MS CID or ETD was performed at 1 Hz after determining the elution time frame from the targeted proteins. The ETD reagent was the radical anion of 3,4-hexanedione (114 *m/z*). All data were collected with OtofControl 3.4 (Bruker Daltonics). Data were analyzed and processed in DataAnalysis 4.3 (Bruker Daltonics). Maximum Entropy algorithm (Bruker Daltonics) was used to deconvolute all mass spectra with resolution set to 80000. The SNAP algorithm was applied to determine the monoisotopic mass of all detected ions.

**Protein identification and fragmentation mapping.** Fragmentation ion lists consisting of monoisotopic mass, intensity and charge were generated from DataAnalysis 4.3, and were subsequently converted to MSAlign files. Alignment-based MS-Align+ algorithm[3,4] for intact protein identification based on protein spectrum matches, was used to search against the Uniprot-Swissprot Sus scrofa (Pig) database, which was released on May 14$^{th}$ 2016 and contains 26167 protein sequences. Fragment mass tolerance was set to 15 ppm. All identifications were validated with statistically significant *P* and *E* values (<0.01) and satisfactory numbers of assigned fragment (>10).

**Calculation of relative aboundance of phosphorylated β-casein proteoforms.** We divided the ion intensity of all β-casein proteoforms by the total ion intensity across the same elution time frame.

**Part II. Supplementary data on the characterization of $CoFe_2O_4$ NPs**

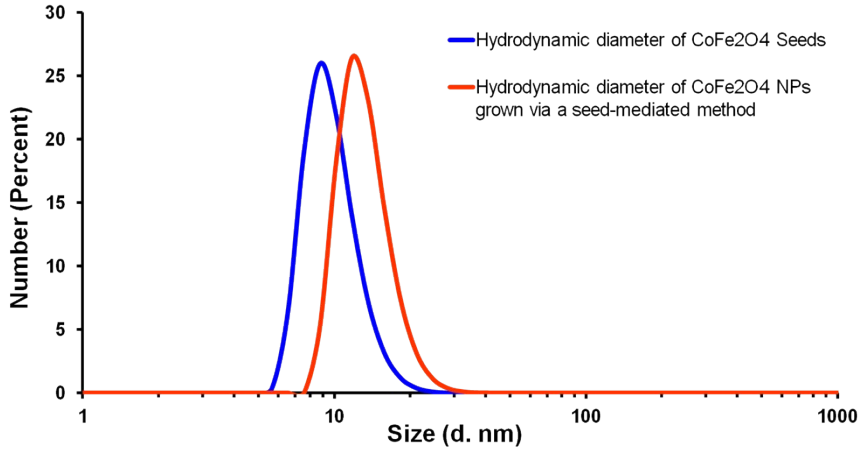


**Figure S1**. Particle-size (average hydrodynamic diameter) distribution obtained by dynamic light scattering (DLS). DLS shows the as-synthesized $CoFe_2O_4$ seeds (blue, 9.07 ± 2.31 nm) and $CoFe_2O_4$ NPs (red, 11.77 ± 2.93 nm) produced via a seed-mediated growth method.

**Part III. Supplementary data on paralleled SDS-PAGE analysis of phosphoproteins enriched using $CoFe_2O_4$ NPs**

We tested the phosphoprotein enrichment from a standard protein mixture containing bovine serum albumin (BSA, a non-phosphoprotein), pepsin (phosphoprotein with one phosphorylation site), and β-casein (phosphoprotein with five phosphorylation sites) in 3:1:1 mass ratio. After SDS-PAGE analysis of the pre- and post-enrichment solutions, the gel was first stained by Phos-tag staining to visualize phosphoproteins, and then by Sypro Ruby for total protein staining. Dark bands of β-casein and pepsin which has only a single phosphoserine in elution lane indicates the effective capture and enrichment as seen in both total protein staining (Figure S2a) and phosphoprotein staining (Figure S2b). Additionally, the band of BSA is faint in the same elution lane, showing minimal non-specific binding. This result confirms the selectivity of the GAPT-Zn ligand complex on the surface of the $CoFe_2O_4$ NPs toward phosphate groups.

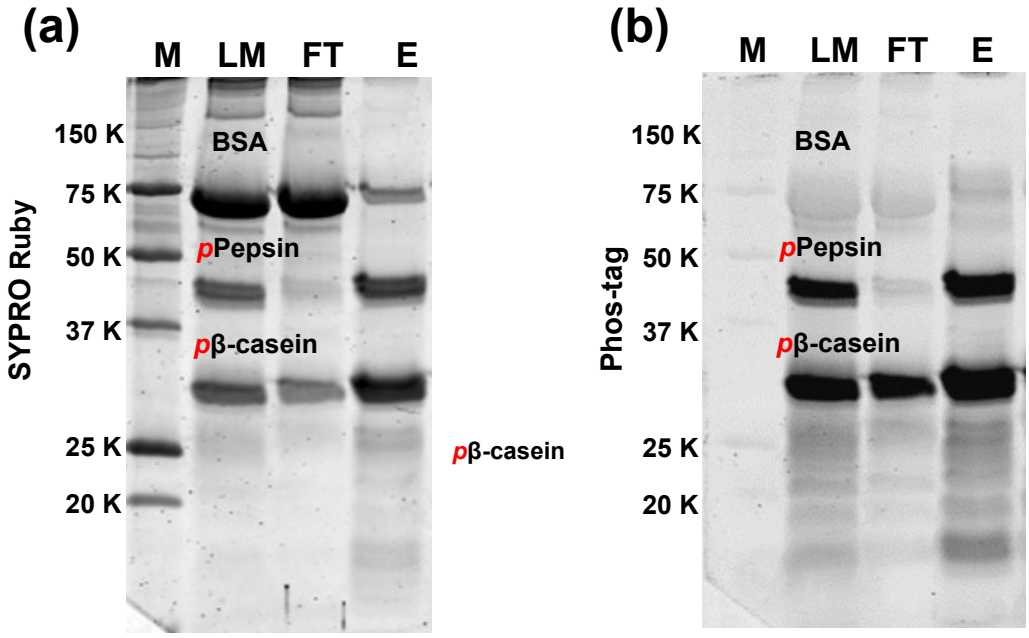


**Figure S2.** SDS-PAGE analysis with (a) Sypro Ruby and (b) Phos-tag gel stain solutions confirmed the highly specific enrichment of phosphoproteins (β-casein and pepsin) from a standard protein mixture containing non-phosphoprotein (BSA). M: marker; LM: loading mixture; FT: flow through; E: elution.

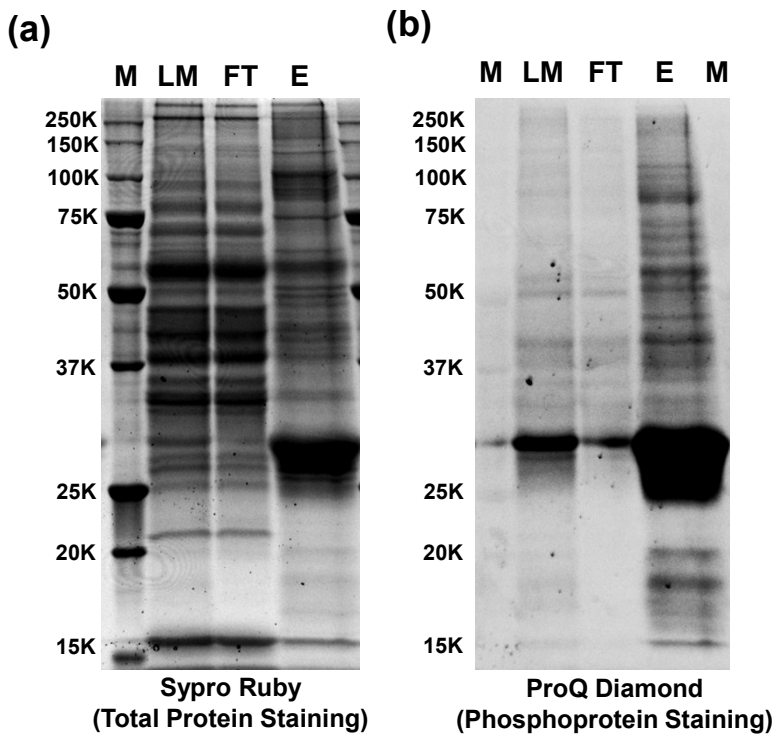


**Figure S3.** SDS-PAGE analysis of 5% β-casein spike-in swine heart tissue extract with (a) Sypro Ruby staining and (b) ProQ Diamond staining corresponding to Figure 4. M: marker; LM: loading mixture; FT: flow through; E: elution.

**Esimating the percentage (w/w) of the spiked-in β-casein in swine heart tissue extract.** 10 mg/mL of β-casein (HPLC grade; average molecular weight 24 kDa) was prepared as a stock solution. Varying aliquots of β-casein stock solution were added into swine heart tissue extract, and then the total protein amounts was estimated by Bradford protein assay. The percentage (w/w) of the spiked-in β-casein was calculated as follows: % of spiked-in β-casein = [the amount of spiked-in β-casein /the total amount of the proteins in the mixture after adding the β-casein into swine heart tissue extract] × 100%.

**Table S1**. Final protein concentrations of the spike-in β-casein in 1 mL of the swine heart tissue extract used for phosphoprotein enrichment and the percentage (w/w) of the spiked-in β-casein in swine heart tissue extract. I, II, III, and IV corresponds to the lanes in Figure 3a, b.

| | I | II | III | IV |
|---|---|---|---|---|
| **Concentration (μg/ μL) of spiked-in β-casein** | 0.03 | 0.05 | 0.07 | 0.1 |
| **Concentration (pmol/ μL) of spiked-in β-casein** | 1.3 | 2.1 | 2.9 | 4.2 |
| **Total protein quantity of swine heart tissue extract (LM) (μg)** | 885 | 910 | 764 | 996 |
| **% (w/w) of β-casein in LM** | 3 | 5 | 9 | 10 |

**Part IV. Supplementary data on top-down MS**

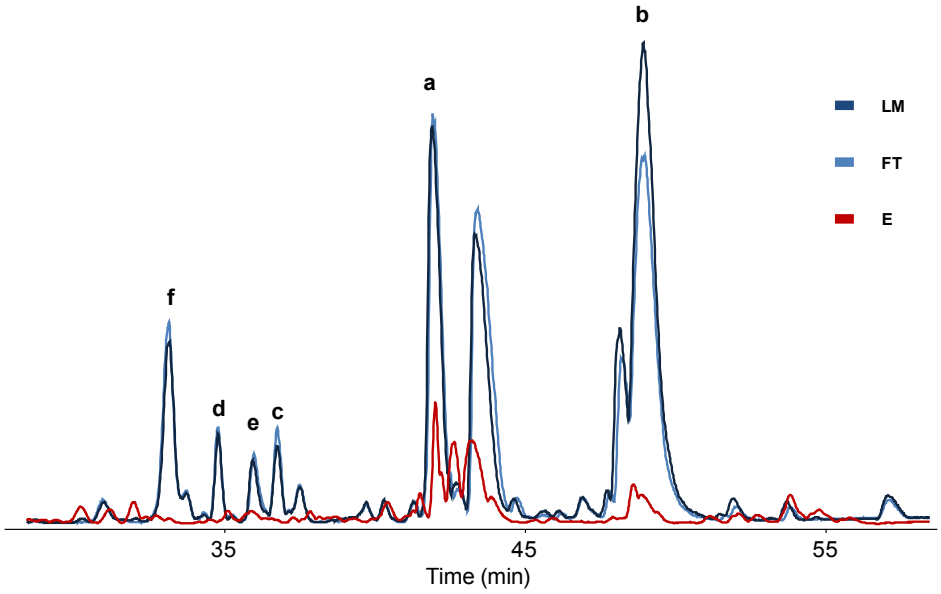


**Figure S4.** The overlapped base peak chromatograms (BPC) of LM, FT, and E. The loading mixture (LM; pre-enrichment solution), flow through (FT; washing solution), and elution (E; post-enrichment solution) of swine heart tissue extract containing 50 μg of spike-in β-casein were first separated by reverse phase chromatography. Equal amount of LM, FT, and E solutions were injected into the LC/MS. The BPC of LM (dark blue) and FT (light blue) was almost identical with each other but not with BPC of E (red). This demonstrates that the majority of proteins (presumably nonphosphoproteins) from LM were washed off into FT. The abundant a, b, c, d, e, and f peaks in LM and FT were identified and shown in Figure S5 and S6.

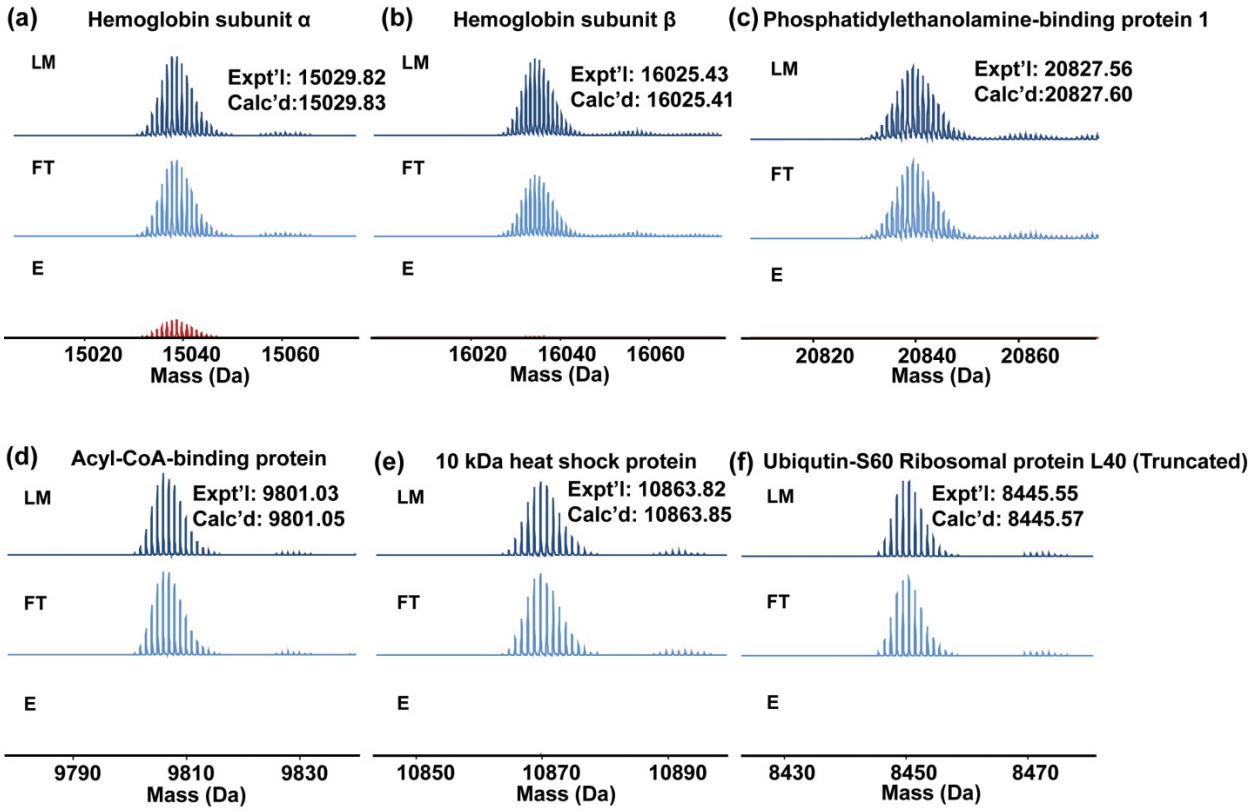


**Figure S5**. Depletion of abundant non-phosphoproteins including blood proteins among others after enrichment, shown by comparison of normalized deconvoluated spectra from loading mixture (LM), flow through (FT), and elution (E). (a) Deconvoluted spectra of Hemoglobin subunit α from 41.8 to 42.3 min; (b) Deconvoluted spectra of Hemoglobin subunit β from 48.0-48.4 min; (c) Deconvoluted spectra of Phosphatidylethanolamine-binding protein 1 from 35.8-36.2 min; (d) Deconvoluted spectra of Acyl-CoA-binding protein from 34.7-35.0 min; (e) Deconvoluted spectra of 10 kDa heat shock protein from 36.7-37.0 min; (f) Deconvoluted spectra of Ubiquitin-S60 Ribosomal protein L40 (truncated) from 33.0-33.5 min. a, b, c, d, e, and f correspond to the labeled most abundant peaks in Figure S4. All proteins were identified by MSAlign+ search algorithm from LC/MS/MS with CID fragmentation (Figure S6).

**(a)** All proteins / sp|P01965|HBA_PIG Hemoglobin subunit alpha OS=Sus scrofa GN=HBA PE=1 SV=1 / Peptide #1

## Protein-Spectrum-Match for Spectrum #47

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| PrSM ID: | 32 | Scan(s): | 2382 | Precursor charge: | 17 | Precursor m/z: | 885.121 |
| Precursor mass: | 15029.935 | Adjusted precursor mass: | 15029.831 | Duplicate score: | 131 | Unique score: | 81 |
| E-value: | 1.2E-56 | P-value: | 5.8E-57 | Spectral FDR: | n/a | Protein mass: | 15029.831 |

No N-terminal modification(canonical)

V⌊L⌊S⌊A A D K A N⌉V K⌊A⌉A⌉W⌉G K⌊V⌉G G Q A⌉G A⌉H⌉G⌉A⌉E⌉A L E R M F L G⌉F⌉P T T K
T Y F⌉P H F N⌉L⌊S⌉H⌉G S D Q V K A H G Q K⌊V⌊A D⌉A L T K⌊A V G H⌊L⌊D⌊D⌊L⌊P⌊G⌉A⌊L⌊
S⌊A⌊L⌊S⌊D⌊L⌊H A H K L R V D⌉P V N F K L⌊L⌊S⌊H⌊C L⌉L⌊V⌊T L A A⌊H⌊H⌉P D D F N⌉P S⌊
V⌊H⌊A⌊S L D⌊K⌊F⌊L⌊A N⌊V⌊S T⌊V⌊L⌊T⌊S⌊K Y R

All proteins / tr|F1RII7|F1RII7_PIG Hemoglobin subunit beta OS=Sus scrofa GN=HBB PE=1 SV=1 / Peptide #5

**(b)** **Protein-Spectrum-Match for Spectrum #107**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| PrSM ID: | 75 | Scan(s): | 2800 | Precursor charge: | 22 | Precursor m/z: | 729.442 |
| Precursor mass: | 16025.563 | Adjusted precursor mass: | 16025.412 | Duplicate score: | 67 | Unique score: | 46 |
| E-value: | 2.2E-37 | P-value: | 2.2E-37 | Spectral FDR: | n/a | Protein mass: | 16156.452 |

N-Terminal Methionine Excision(canonical)

M]V H L S⌉A E E K E A⌊V⌉L⌊G⌉L⌊W⌉G K V N⌊V⌉D⌊E⌉V G⌉G E A L G R⌉L⌊L⌊V⌉V⌊Y⌉P W T Q
R F F E⌉S F G D L S N A⌊D⌊A⌊V⌊M⌊G N P K V K A H G K K V L Q S F S⌊D⌊G L⌊K H L D
N L K G T F A K L S E L H C D Q L H V D⌊P E N F R L L G N V I V V V L A R R L G
H D F N⌊P D⌊V⌊Q⌊A⌊A⌊F Q K⌊V⌊V⌊A⌊G⌊V⌊A⌊N⌊A⌊L⌊A⌊H K⌊Y H

All proteins / tr|F1RKG8|F1RKG8_PIG Uncharacterized protein OS=Sus scrofa GN=PEBP1 PE=1 SV=1 / Peptide #11

**(c)** **Protein-Spectrum-Match for Spectrum #16**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| PrSM ID: | 9 | Scan(s): | 1945 | Precursor charge: | 26 | Precursor m/z: | 802.076 |
| Precursor mass: | 20827.789 | Adjusted precursor mass: | 20827.597 | Duplicate score: | 45 | Unique score: | 33 |
| E-value: | 6.2E-25 | P-value: | 2.1E-25 | Spectral FDR: | n/a | Protein mass: | 20958.638 |

N-Terminal Methionine Excision(canonical)

M]P V D⌉L⌉G K W⌉S⌉G⌉P⌉L⌉S⌉L⌉Q⌉E⌉V⌉D E R P Q H⌉P L Q V K Y G G A E V D E L G K V
L T P T Q V K S R P T S I T⌉W D⌉G⌉L N⌉P D K⌉L⌉Y T⌉L⌉V L T D P D A P S R K D P K
Y R E W H H F L V V N M K G N D I S S G T V L S D Y V G S G P P K G T G L H R Y
V W L V Y E Q D G P L K C D E P I L S N R S G D H R G K F K V A S F R K K Y Q L
G A P V A⌊G T⌊C⌊Y⌊Q⌊A⌊E⌊W⌊D D⌊Y⌊V⌊P⌊K L Y E Q L S G K

**(d)**

All proteins / sp|P12026|ACBP_PIG Acyl-CoA-binding protein OS=Sus scrofa GN=DBI PE=1 SV=2 / Peptide #14

## Protein-Spectrum-Match for Spectrum #10

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| PrSM ID: | 6 | Scan(s): | 1857 | Precursor charge: | 13 | Precursor m/z: | 754.941 |
| Precursor mass: | 9801.142 | Adjusted precursor mass: | 9801.054 | Duplicate score: | 40 | Unique score: | 30 |
| E-value: | 2.8E-23 | P-value: | 1.4E-23 | Spectral FDR: | n/a | Protein mass: | 9890.084 |

N-Terminal Methionine Excision(canonical) and N-Terminal Acetylation

M]S[42.01] Q|A|E|F E K|A A E E V K N L K T K P A D D|E|M|L|F|I|Y|S H Y K Q A T|V|G D I|
N T E R P G I L D|L|K|G K A K|W D A W N G L K G T S K E D|A M|K|A Y I|N|K|V|E E|
L K K K Y G I

**(e)**

All proteins / tr|F1SMZ6|F1SMZ6_PIG Uncharacterized protein OS=Sus scrofa GN=HSPE1 PE=1 SV=1 / Peptide #22

## Protein-Spectrum-Match for Spectrum #22

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| PrSM ID: | 15 | Scan(s): | 2006 | Precursor charge: | 14 | Precursor m/z: | 777.003 |
| Precursor mass: | 10863.945 | Adjusted precursor mass: | 10863.847 | Duplicate score: | 43 | Unique score: | 29 |
| E-value: | 7.3E-20 | P-value: | 7.3E-20 | Spectral FDR: | n/a | Protein mass: | 10952.877 |

N-Terminal Methionine Excision(canonical) and N-Terminal Acetylation

M]A[42.01] G Q A F R K F L|P L F D R V L V E R S A A E T V T K G G I M|L|P E K S Q G K
V L Q|A T|V|V|A|V|G S G S K G K G G E I Q|P|V|S|V K V G D K V L|L|P E Y G G T R
V V L D|D K|D Y|F L F R D G D I L G K Y V D

**(f)**

All proteins / sp|P63053|RL40_PIG Ubiquitin-60S ribosomal protein L40 OS=Sus scrofa GN=UBA52 PE=1 SV=2 / Peptide #7

## Protein-Spectrum-Match for Spectrum #3

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| PrSM ID: | 2 | Scan(s): | 1763 | Precursor charge: | 10 | Precursor m/z: | 845.572 |
| Precursor mass: | 8445.646 | Adjusted precursor mass: | 8445.574 | Duplicate score: | 76 | Unique score: | 50 |
| E-value: | 4.9E-35 | P-value: | 3.5E-37 | Spectral FDR: | n/a | Protein mass: | 14718.958 |

Removal of protein suffix
No N-terminal modification(canonical)

M|Q|I|F|V K T|L|T G K|T|I|T|L|E|V|E|P|S D|T|I E|N V|K|A K I Q D|K|E|G I|P|P D|Q
Q R L I F|A G|K|Q|L|E|D|G R T L S D|Y N|I|Q|K|E|S|T|L|H|L V L R L R[G G I I E P
S L R Q L A Q K Y N C D K M I C R K C Y A R L H P R A V N C R K K K C G H T N N
L R P K K K V K

**Figure S6**. Identification of the highly abundant non-phosphoproteins (corresponding to figure S5) based on LC/MS/MS data with MS-Align+ search algorithm. (a) Hemoglobin subunit α (UnitProtKB/Swiss-Prot, P01965|HBA_PIG); (b) Hemoglobin subunit β (UnitProtKB/Swiss-Prot, F1RII7_PIG); (c) Phosphatidylethanolamine-binding protein 1 (UnitProtKB/Swiss-Prot, F1RKG8_PIG); (d) Acyl-CoA-binding protein (UnitProtKB/Swiss-Prot, P12026|ACBP_PIG); (e) 10 kDa heat shock protein (UnitProtKB/Swiss-Prot, F1SMZ6_PIG); (f) Ubiquitin-S60 Ribosomal protein L40 (truncated) (UnitProtKB/Swiss-Prot, P63053|RL40_PIG).

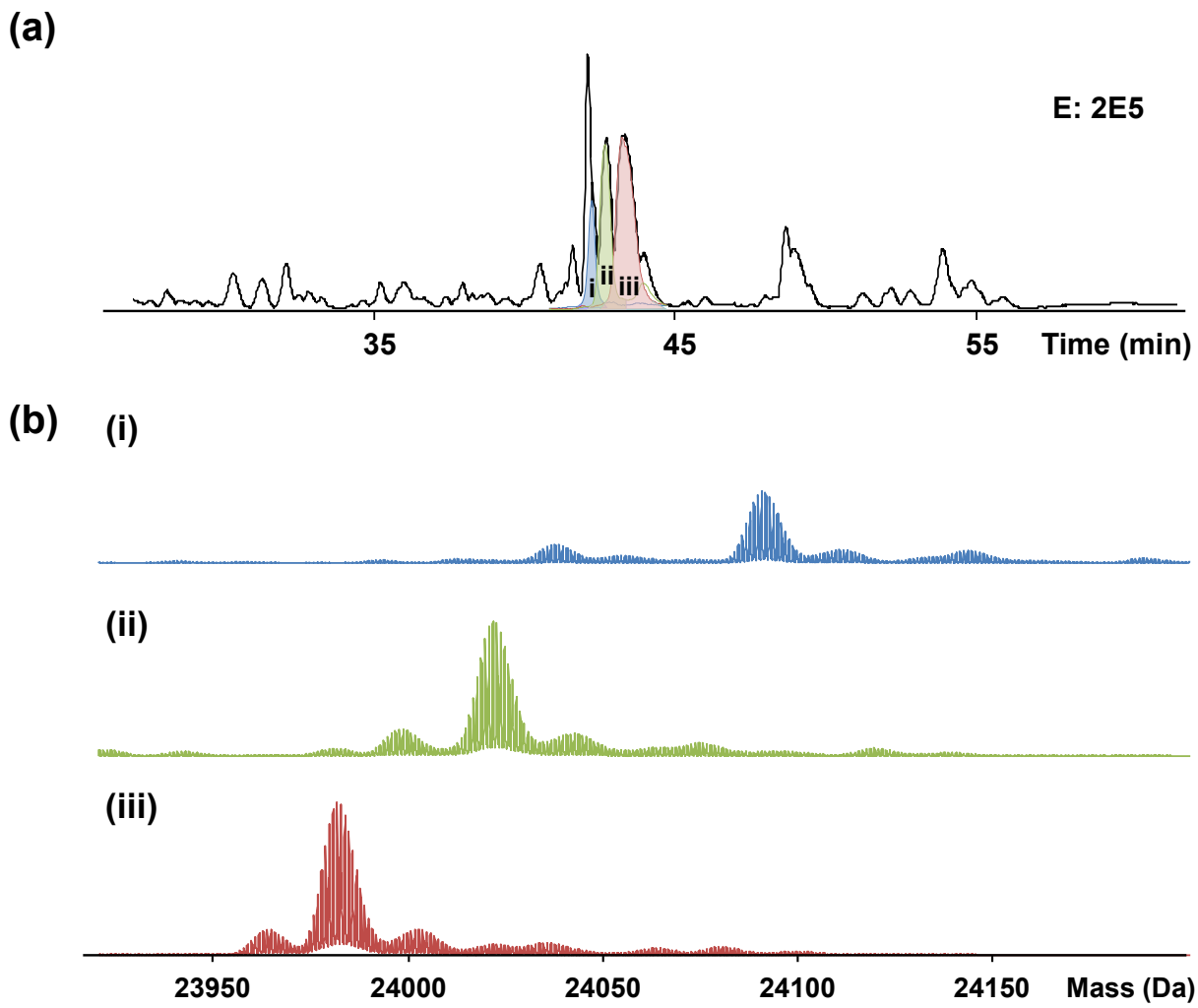


**Figure S7**. Top-down LC/MS analysis of multiple β-casein variants in elution after enrichment with swine heart tissue extract with 0.05 μg/μL of spike-in β-casein. **(a)** Base peak chromatogram of the elution fraction. (i), variant A1 (ii), and variant A2 (iii). **(b)** Normalized deconvoluted mass spectra of β-casein isoforms; Variant B (i) showed a mass difference of 109 Da from A2 variant, which corresponds with a previously reported mass difference by two amino acids at position 67 and 122, having His 67 and Ser 122 instead of Pro 67 and Arg 122 in the variant A2 . Variant A1 (ii) revealed a mass difference of 40 Da from A2 variant (iii), which corresponds with a single amino acid difference from Pro 67 (in the variant A2) to His 67.[5]

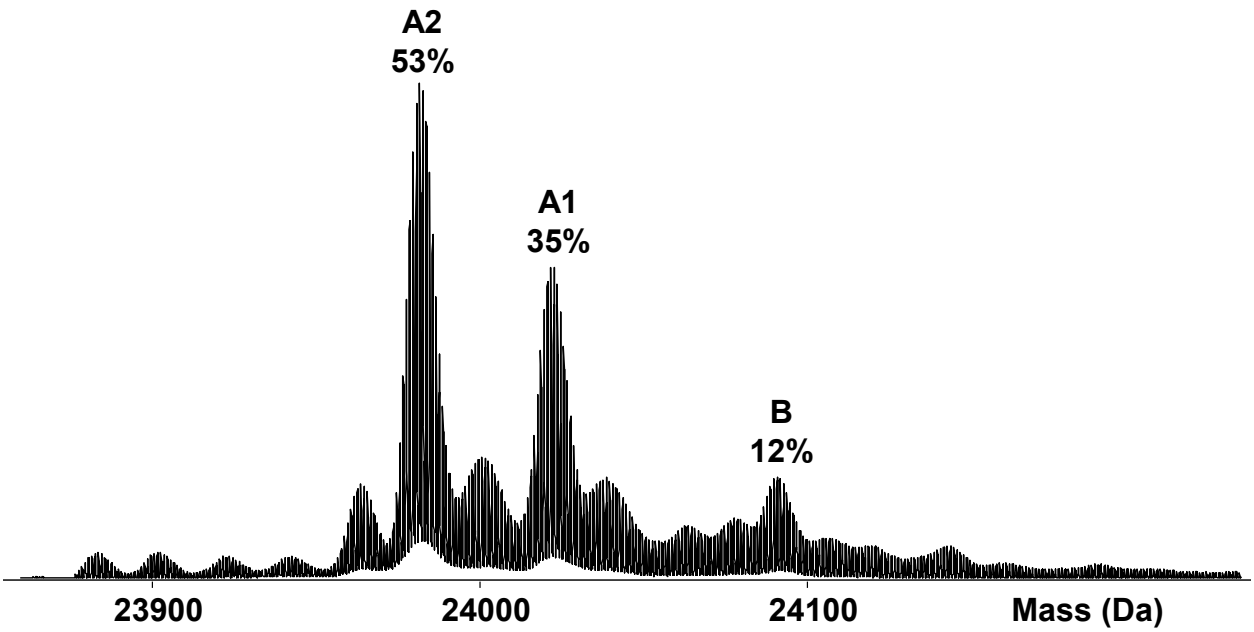


**Figure S8**. Deconvoluted spectrum of β-casein proteoforms from 41.9 – 44.5 min in the top-down LC/MS run of elution sample indicating the relative abundance among variant A2, A1, and B is 53%, 35%, and 12% respectively.

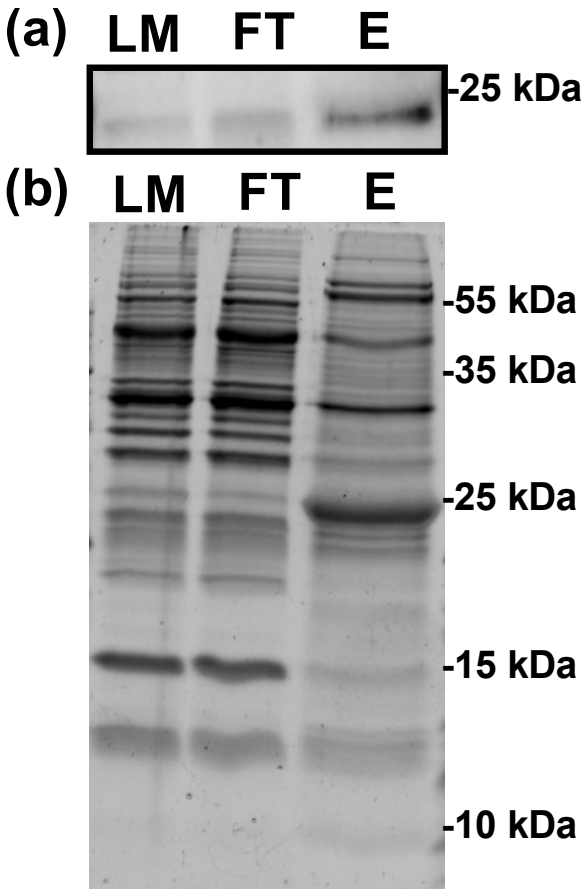


**Figure S9**. Western blots analysis of phosphorylated cTnI demonstrating the enrichment of phosphoproteins in the elution. (a) Representative Western blot for cTnI phosphorylated at Ser22/23 with equal amount of proteins loaded in loading mixture (LM), flow through (FT), and elution (E). (b) Total proteins stained by Sypro Ruby with equal loading in LM, FT, and E.

**(a)** All proteins / tr|F1RHJ2|F1RHJ2_PIG Uncharacterized protein OS=Sus scrofa GN=HDGF PE=1 SV=1 / Peptide #2

## Protein-Spectrum-Match for Spectrum #1

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| PrSM ID: | 0 | Scan(s): | 1 | Precursor charge: | 29 | Precursor m/z: | 940.249 |
| Precursor mass: | 27238.02 | Adjusted precursor mass: | 27237.855 | Duplicate score: | 27 | Unique score: | 21 |
| E-value: | 4.6E-13 | P-value: | 8.8E-14 | Spectral FDR: | n/a | Protein mass: | 27088.011 |

N-Terminal Methionine Excision(canonical) and N-Terminal Acetylation
A modification (or multiple modifications) on internal residue

M]S(42.01) R S N R]Q]K]E Y K C G D L V F A K M K G Y P H W P A R I D E M P E A A V K S
T A N K Y Q V F F F G T H E T A F L G P K D L F P Y E E S K E K F G K P N K R K
G F S E G L W E I E N N P T V K A S G Y Q S S Q K K S C V E E P E P K A D]A A E
G]D G]D(238.97) K K G N A E G S S D E E G K L V I D E P T K E K N E K G A L K R R A G D
L L E D S P K R P K]E A E]D P]E G E E K]E V A T]L E G E R P L P V E]A E]K]N S T
P S]E P S S G R G P P P E E E E E E E E E E E E E E E A T K E]D A]E]A P G I R D H
E S L

**(b)** All proteins / tr|F1RHJ2|F1RHJ2_PIG Uncharacterized protein OS=Sus scrofa GN=HDGF PE=1 SV=1 / Peptide #1

## Protein-Spectrum-Match for Spectrum #1

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| PrSM ID: | 0 | Scan(s): | 1 | Precursor charge: | 29 | Precursor m/z: | 940.249 |
| Precursor mass: | 27238.02 | Adjusted precursor mass: | 27237.915 | Duplicate score: | 68 | Unique score: | 45 |
| E-value: | 1.2E-21 | P-value: | 2.3E-22 | Spectral FDR: | n/a | Protein mass: | 27088.011 |

N-Terminal Methionine Excision(canonical) and N-Terminal Acetylation
A modification (or multiple modifications) on internal residue

M]S(42.01) R S N R Q K E]Y K C G D L V]F(238.93) A K M K G Y P H W P]A R I D E M P E A A V K S
T A N K Y Q V F F F G T H E T A F L G P K D L F P Y E E S K E K F G K P N K R K
G F S E G L W E I E N N P T V K A S G Y Q S S Q K K S C V E E P E P K A D A A E
G D G D K K G N A E G S S D E E G K L V I D E P T K E K N E K G A L K R R A G D
L L E D S P K R P K E A E D]P E]G E E K E]V A]T]L]E]G E R P L]P V E A E K N S T]
P S E]P S S G R G P P P E]E]E]E]E]E]E]E]E]E]E]E]A]T]K]E]D]A]E]A]P G I]R D]H
E S L

**Figure S10**. Identification of the 27 kDa multiply phosphorylated protein by LC/MS/MS (a) ETD and (b) CID with MS-Align+ algorithm[3]. Uncharacterized protein (UnitProtKB/Swiss-Prot, F1RHJ2_PIG) was best matched and identified with N-terminal Met loss and two possible modifications (acetylation and multiple phosphorylations). The Basic Local Alignment Search Tool (BLAST) found 93.8% sequence similarity between the identified protein and Hepatoma-derived growth factor (Homo sapiens) (P51858|HDGF_HUMAN), suggesting its identity.

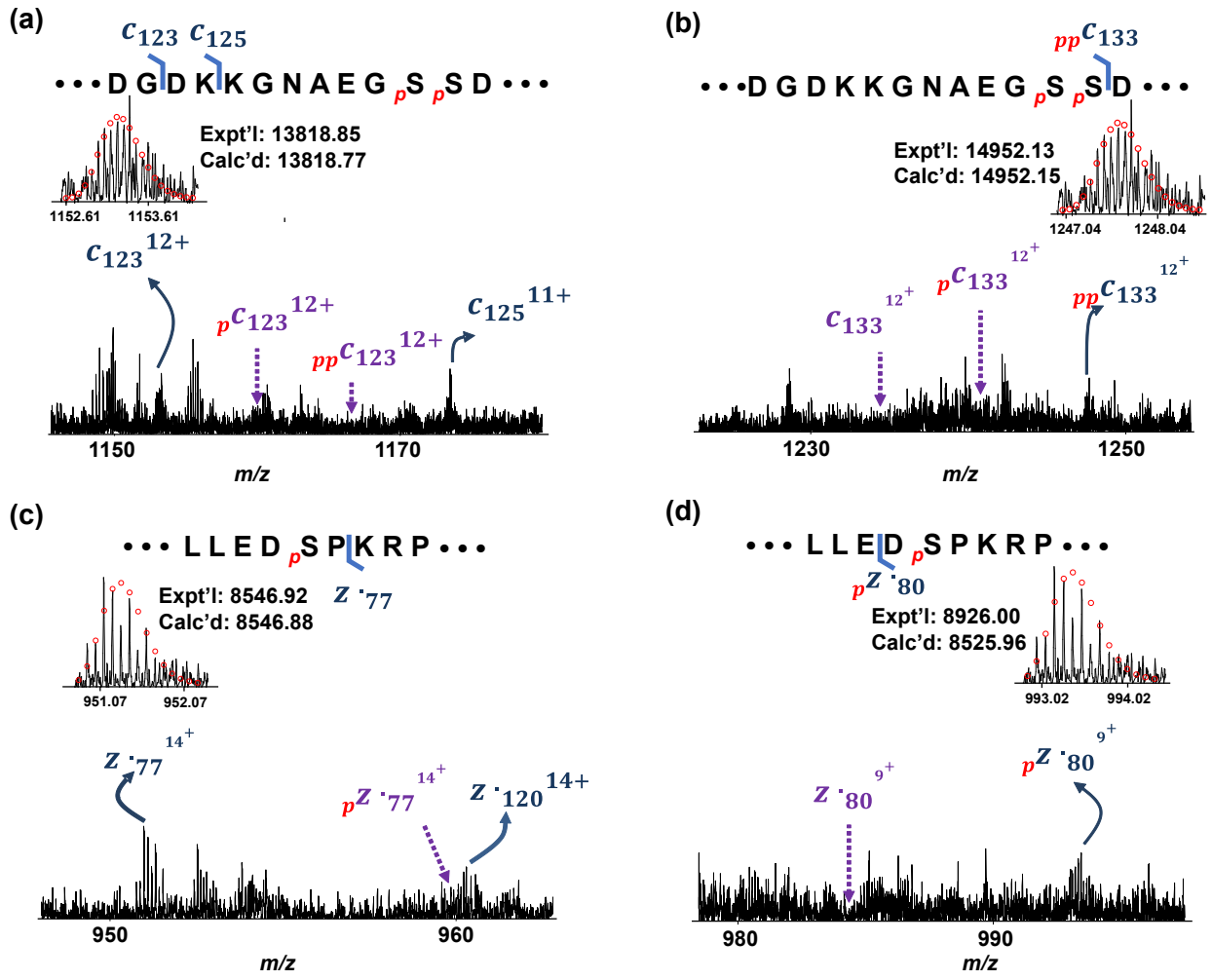


**Figure S11**. MS/MS mapping of the phosphorylation sites of the identified swine hepatoma-derived growth factor after NP enrichment. (a-d) Representative MS/MS spectra of ***c*** and ***z•*** type ions from ECD spectra of triply-phosphorylated hepatoma-derived growth factor using online LC/MS/MS. The phosphorylation sites were localized to Ser132, Ser133, and Ser165. The assignment of fragment ions was made based on the sequence of uncharacterized protein (UnitProtKB/Swiss-Prot, F1RHJ2_PIG) with N-terminal Met excision and acetylation near the N-terminal. Blue lines indicate existing ions, and purple dash lines indicate non-existing ions.

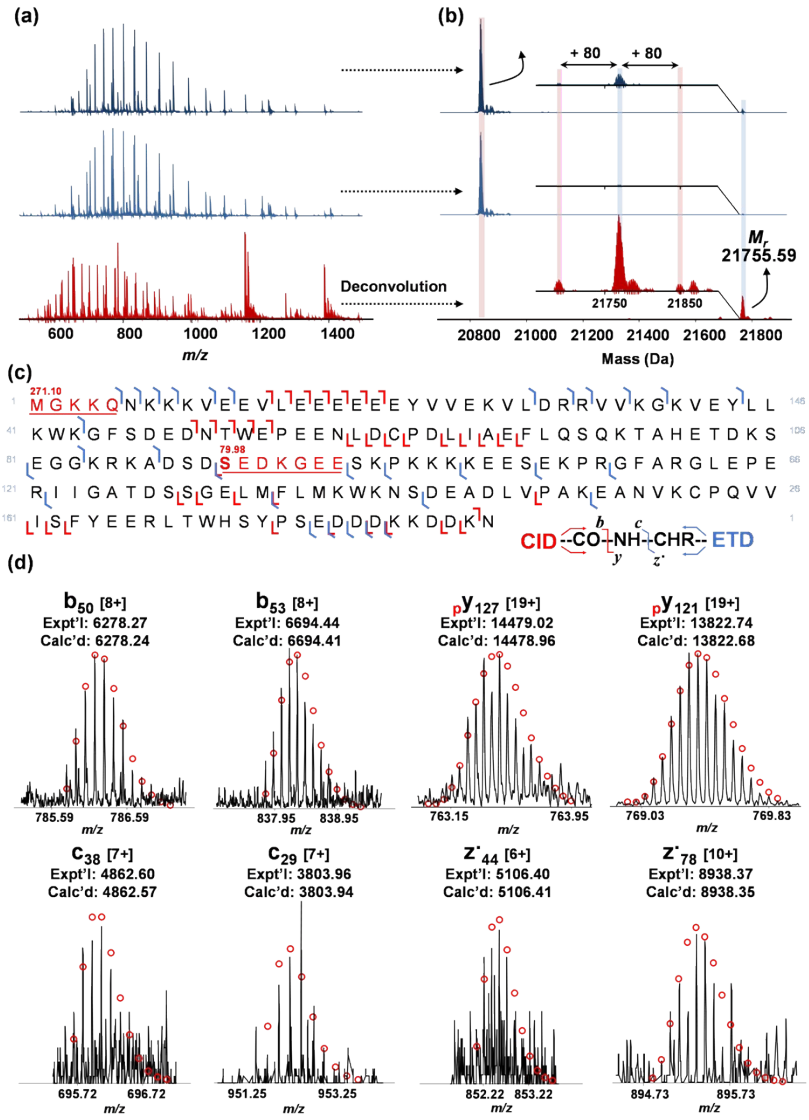


**Figure S12.** Representative MS/MS analysis of endogenous phosphoproteins (22 kDa) enriched by functionalized $CoFe_2O_4$ NPs from swine heart tissue extract**.** (a) Average spectra and (b) deconvoluted spectra of loading mixture (LM, dark blue), flow through (FT, light blue), and elution (E, red) from 35.7 min to 35.9 min. (c) Fragment ion map of ETD and CID product ions from triply phosphorylated precursor ion at 31+ charge state of chromobox protein homolog 1 through online MS/MS analysis with CID and ETD. Red "S" (79.98) indicates potential phosphorylation sites. Red number above the underlined red sequence reveals modifications with their additional mass.

**(a)** All proteins / tr|F1RWH1|F1RWH1_PIG Uncharacterized protein OS=Sus scrofa GN=CBX1 PE=1 SV=2 / Peptide #1

## Protein-Spectrum-Match for Spectrum #2

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| PrSM ID: | 1 | Scan(s): | 1 | Precursor charge: | 31 | Precursor m/z: | 702.802 |
| Precursor mass: | 21755.65 | Adjusted precursor mass: | 21755.65 | Duplicate score: | 26 | Unique score: | 22 |
| E-value: | 2.8E-15 | P-value: | 4.6E-18 | Spectral FDR: | n/a | Protein mass: | 21404.572 |

Unusual N-terminal modification
A modification (or multiple modifications) on internal residue

271.10
M G K K Q N K K]K V]E]E V L E E E E E E Y V V E K V L]D R]R V V]K G]K V E Y]L L
79.99
K W K]G F S D E D N T W E P E E N L D C P D L I A E F L Q S Q K T A H E T D K S
E G G K R K A D S D S E D K G E E S K P K K K K E E S E K P R]G F A R G L E P E
R]I I G A T D S]S G E L M]F L M K W K N S]D E A D L]V P A K]E A N V K C P Q V V
I S F Y E E R L T W H S Y P S]E]D]D]D K K D D K N

**(b)** All proteins / tr|F1RWH1|F1RWH1_PIG Uncharacterized protein OS=Sus scrofa GN=CBX1 PE=1 SV=2 / Peptide #2

## Protein-Spectrum-Match for Spectrum #2

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| PrSM ID: | 1 | Scan(s): | 1 | Precursor charge: | 31 | Precursor m/z: | 702.802 |
| Precursor mass: | 21755.64 | Adjusted precursor mass: | 21755.64 | Duplicate score: | 88 | Unique score: | 37 |
| E-value: | 1.4E-12 | P-value: | 1.4E-15 | Spectral FDR: | n/a | Protein mass: | 21404.572 |

Unusual N-terminal modification
A modification (or multiple modifications) on internal residue

271.10
M G K K Q N K K K V E E V L E]E]E]E]E]E]Y V V E K V L D R R V V K G K V E Y L L
79.97
K W K G F S D E D]N]T]W]E]P E E]N]L]D]C]P D]L]I]A]E]F L Q S Q K T A H E T D K S
E G G K R K A D S D S E D K G E E S K P K K K K E E S E K P R G F A R G L E P E
R I I G A T D S]S]G E L]M]F L M K W K N S D E A D L V]P A K E A N V K C P Q V V]
I]S F Y E E R L T W H S Y]P S E]D]D]D]K K D]D]K N

**Figure S13**. Identification of the 22 kDa phosphorylated protein by (a) ETD LC/MS/MS and (b) CID LC/MS/MS with MS-Align+ algorithm. Uncharacterized protein (UnitProtKB/Swiss-Prot, F1RWH1_PIG) was best matched and identified with N-terminal sequence variation and a possible phosphorylation. The Basic Local Alignment Search Tool (BLAST) found 100% sequence similarity between the identified protein and Chromobox protein homolog 1 (Homo sapiens) (P83916|CBX1_HUMAN), indicating its identity.

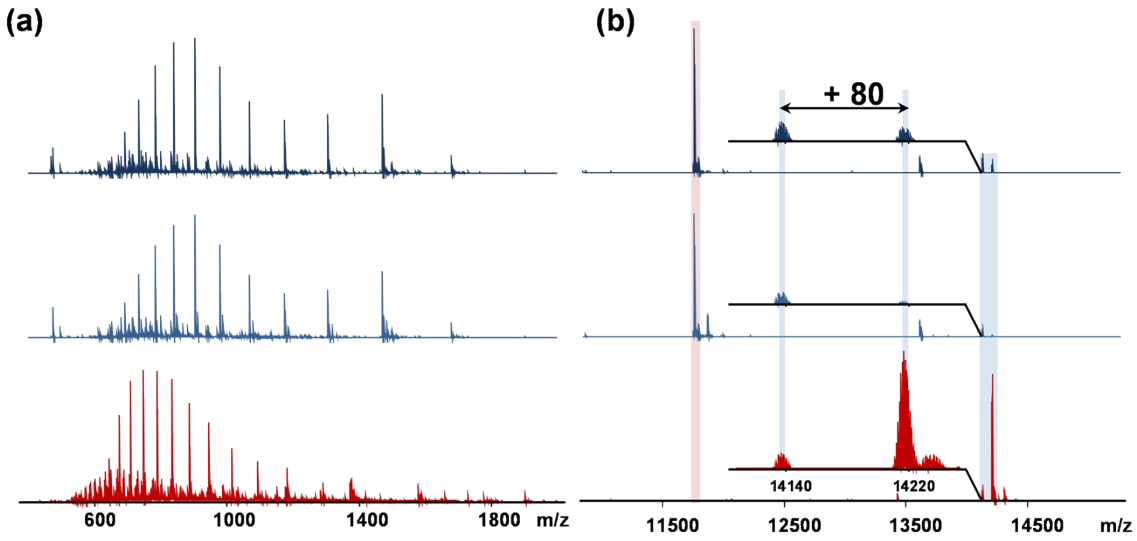


**Figure S14.** Representative MS/(MS) analysis of endogenous phosphoproteins (14 kDa) enriched by functionalized $CoFe_2O_4$ NPs from swine heart tissue extract**.** (a) Mass spectra and (b) deconvoluted spectra of loading mixture (LM, dark blue), flow through (FT, light blue), and elution (E, red) from 37.8 min to 38.1 min.

All proteins / tr|F1RNX2|F1RNX2_PIG Programmed cell death protein 5 OS=Sus scrofa GN=PDCD5 PE=1 SV=1 / Peptide #2

## Protein-Spectrum-Match for Spectrum #1

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| PrSM ID: | 0 | Scan(s): | 1 | Precursor charge: | 19 | Precursor m/z: | 748.861 |
| Precursor mass: | 14209.22 | Adjusted precursor mass: | 14209.22 | Duplicate score: | 35 | Unique score: | 21 |
| E-value: | 2.4E-6 | P-value: | 1.7E-6 | Spectral FDR: | n/a | Protein mass: | 14218.324 |

N-Terminal Methionine Excision(canonical) and N-Terminal Acetylation
A modification (or multiple modifications) on internal residue

M]A(42.01) E⌊E⌋E L E A L R K Q R L A E L Q A K H G D⌋P G D A⌋A⌊Q Q E A K H R E A E M R
N S I L A Q⌋V⌊L D Q S A R A R L S N L A L V K P E K T K A V E N Y⌋L⌋I⌋Q⌊M⌊A R Y
G Q L S G K V S E Q G L I E I L E K V S Q Q T E K K T T V K F N R R K V(79.93) M D S D
E D D⌋D⌋Y

**Figure S15**. Identification of the 14 kDa phosphorylated protein by CID LC/MS/MS with MS-Align+ algorithm. Programmed cell death protein 5 (UnitProtKB/Swiss-Prot, F1RNX2_PIG) was best matched and identified with N-terminal Met excision and a possible phosphorylation.

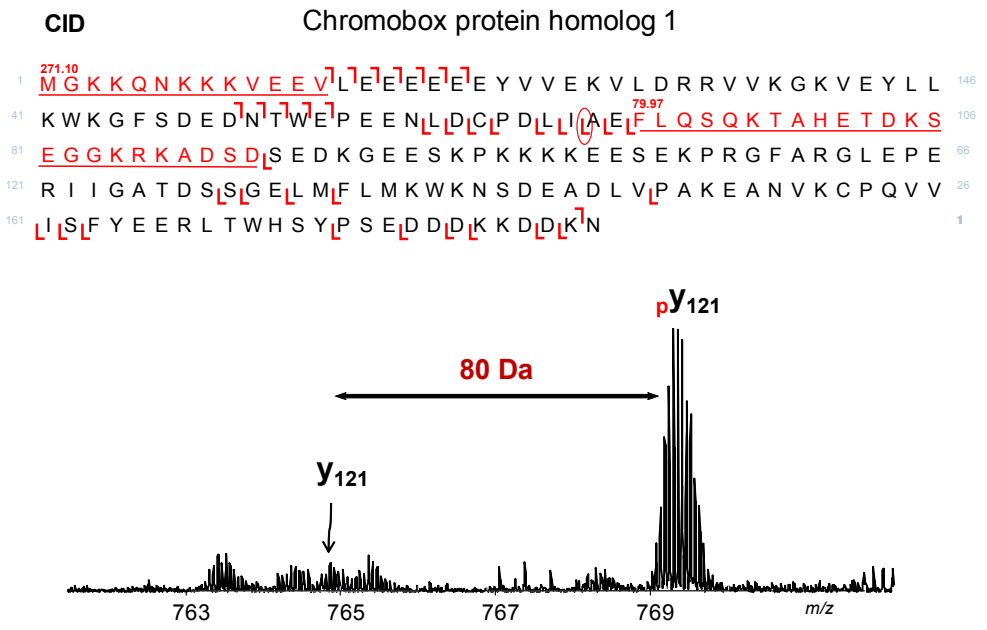


**Figure S16**. Fragmentation map and representing ions of the 22 kDa phosphoryalted protein chromobox protein homolog 1. The highly abundant phosphorylated y121 ion and its minimal unphosphorylated counterpart indicate preservation of phosphorylation during collisionally induced dissociation (CID).

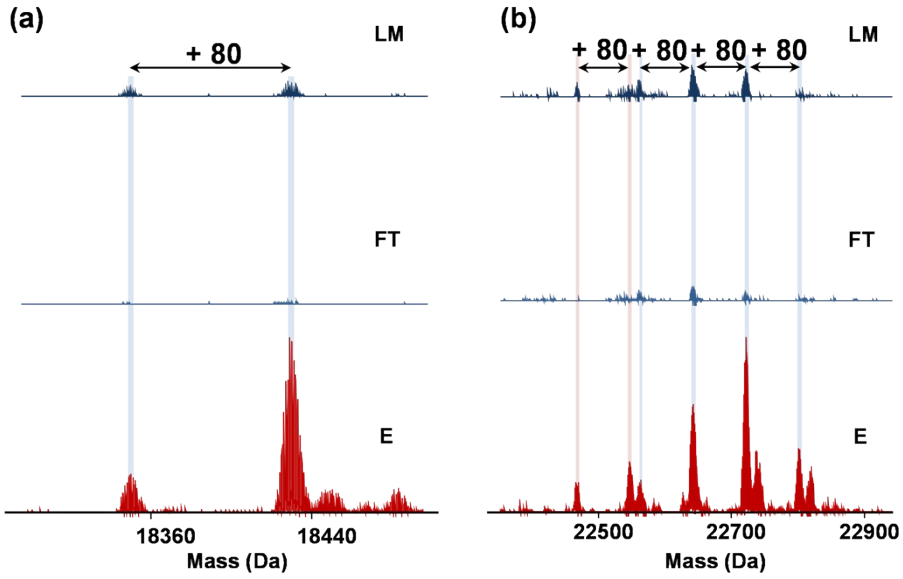


**Figure S17.** Other examples of endogenous phosphoproteins that have been enriched using NPs, demonstrated by comparing normalized deconvoluted average MS spectra of loading mixture (LM), flow through (FT), and elution (E) from (a) 37.2-37.5 min and (b) 32.7-33.0 min.

**Part V. Supporting Information References**


(S1). Sun, S.; Zeng, H.; Robinson, D. B.; Raoux, S.; Rice, P. M.; Wang, S. X.; Li, G. Monodisperse MFe2O4 (M = Fe, Co, Mn) Nanoparticles. *J. Am. Chem. Soc.* **2003,** *126*, 273.

(S2). Hwang, L.; Ayaz-Guner, S.; Gregorich, Z. R.; Cai, W.; Valeja, S. G.; Jin, S.; Ge, Y. Specific Enrichment of Phosphoproteins Using Functionalized Multivalent Nanoparticles. *J. Am. Chem. Soc.* **2015,** *137*, 2432.

(S3). Liu, X. W.; Sirotkin, Y.; Shen, Y. F.; Anderson, G.; Tsai, Y. S.; Ting, Y. S.; Goodlett, D. R.; Smith, R. D.; Bafna, V.; Pevzner, P. A. Protein Identification Using Top-Down. *Mol. Cell. Proteomics* **2012,** *11*, M111.008524.

(S4). Cai, W. X.; Guner, H.; Gregorich, Z. R.; Chen, A. J.; Ayaz-Guner, S.; Peng, Y.; Valeja, S. G.; Liu, X. W.; Ge, Y. MASH Suite Pro: A Comprehensive Software Tool for Top-Down Proteomics. *Mol. Cell. Proteomics* **2016,** *15*, 703.

(S5). Wu, S.; Yang, F.; Zhao, R.; Tolic, N.; Robinson, E. W.; Camp, D. G.; Smith, R. D.; Pasa-Tolic, L. Integrated workflow for characterizing intact phosphoproteins from complex mixtures. *Anal. Chem.* **2009,** *81*, 4210.