Characterization of a two-step enrichment approach for quantitative cancer phosphoproteomics

Xuemei Zeng, Steven Mullett, Zhiyuan Sun, Mai Sun, Nathan Yates

Biomedical Mass Spectrometry Center, University of Pittsburgh Schools of the Health Sciences, Pittsburgh, PA, USA

Phosphoproteomics and cancer: opportunities and challenges

- Dynamic regulation of protein phosphorylation is a key mechanism for regulation of protein activity in a variety of signaling pathways
- Perturbation of more than a hundred different protein kinases has been implicated in human cancer
- It is estimated that protein kinases constitute 1/4 to 1/3 of targets in drug discovery program
- Systemic characterization of protein phosphorylation is made possible by the advent of proteomic technologies, usually involving mass spectrometry (MS), and can be used to facilitate identification of aberrantly activated signaling pathways governing oncogenesis, downstream effectors of mutant kinases, therapeutic targets of kinase inhibitors, and unintended targets of kinase inhibitors, etc.
- Enrichment of phosphopeptides prior to MS analysis is essential due to the low stoichiometry and low ionization efficiency of phosphopeptides
- Several strategies have been developed for selective enrichment of phosphopeptides, including methods involving chemical derivatization of phosphoryl groups and methods utilizing affinity chromatography (metal chelation, ion exchange, immuno-affinity capture, etc.)
- Phosphopeptide enrichment strategies with single modality often have limited selectivity
- A two-step enrichment strategy combining calcium phosphate precipitation followed by affinity capture (metal chelation or TiO2) was reported[1] recently to provide highly selective enrichment of phosphopeptides
- Feasibility of the two-step enrichment method to quantitative proteomics remains to be evaluated.

Objective

- To adopt and optimize a two-step enrichment strategy combining calcium phosphate precipitation and titanium dioxide (TiO2) affinity capture to enrich phosphopeptides with high selectivity
- To evaluate reproducibility and linearity of the phosphopeptide enrichment workflow
- To carry out a proof-of-principle study to evaluate the potential of applying the developed workflow for quantitative assessment of phosphorylation status under different conditions

Sample preparation workflow

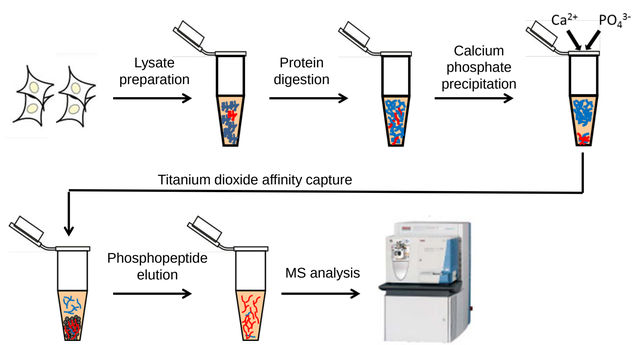


Figure 1: schematic illustration of a two-step phosphopeptide enrichment workflow from complex protein mixtures. Lysate samples are prepared from biological specimens or cultured cells, followed by protein digestion into peptide fragments. Calcium chloride and disodium phosphate are added to peptide solution which leads to calcium phosphate precipitation and co-precipitation of phosphopeptides. Pellet containing the enriched phosphopeptides is dissolved, desalted, and then subjected to titanium dioxide affinity capture to further enrich phosphopeptides. Phosphopeptides eluted from TiO2 are analyzed by LC-MS/MS.

Evaluation of sample preparation

Sample preparation: six replicate phosphopeptide enrichments were carried out using 400ug each of mouse brain homogenate to assess sample preparation reproducibility. Different amounts of bovine beta-casein, a known phosphorylated protein, were spiked-in to each sample to assess the linearity of the assay

LC-MS/MS analysis: the enriched phosphopeptides were analyzed on a LTQ/Orbitrap (Thermo Fisher Scientific Inc.) hybrid mass spectrometer online coupled with nanoflow liquid chromatography system (EASY-nLC™, Thermo Fisher Scientific Inc.)

Data analysis: the MS/MS spectra were searched using Andromeda[2] search engine against mouse UniProt database. Peptide precursor peak area were extracted using MaxQuant[3] and were used as surrogate metric for relative abundance of peptides.

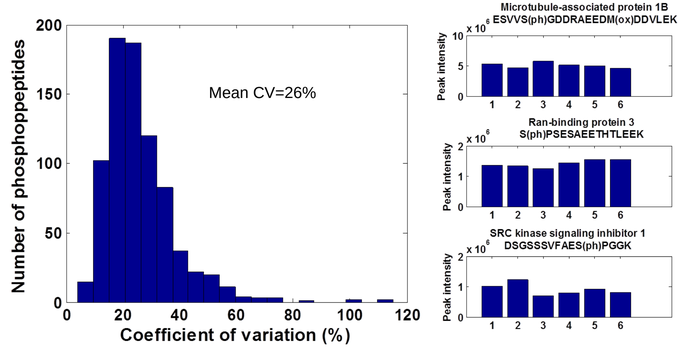


Figure 2: reproducibility of phosphopeptide enrichment workflow. Left panel showed the histogram distribution of the coefficient of variation and the right panel showed examples for the relative abundance of representative phosphopeptides across 6 replicate samples.

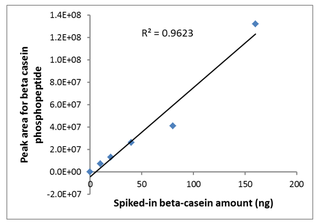


Figure 3: scatter plot distribution of beta-casein phosphopeptide (FQSEEQQQTEDELQDK) peak area vs. spiked-in amount with overlaid linear regression line. Results support good linearity of the assay.

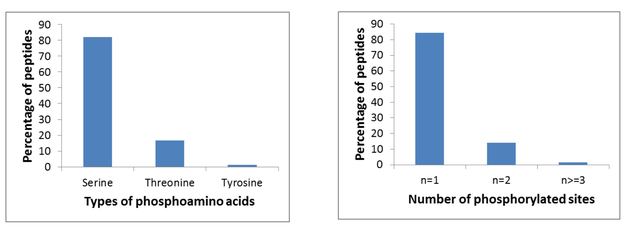


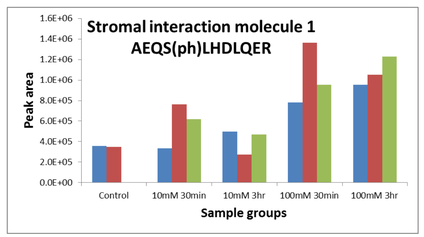
Figure 4: (A) Distribution of three different phosphorylated amino-acid residues (Ser/Thr/Tyr); (B) Distribution of the number of phosphorylated sites per peptide. A total of 2478 peptides were identified from the combined 6 replicate enrichments, with 1764 peptides originated from 729 proteins are phosphorylated (a total of 2092 phosphorylated sites).

Proof-of-principle study

As a proof-of-principle study to demonstrate the potential of using the developed workflow to decipher dynamic regulation of cancer phosphoproteome, we applied it for quantitative assessment of the phosphorylation status of C6 glioma cells, an immortalized glioblastoma derived from *rattus Norvegicus*, in response to rotenone, a pesticide which inhibits complex I of electron transport chain in mitochondria and induces free radical generation, DNA damage, and eventually trigger cell death via apoptosis. Phosphopeptides were enriched from C6 cells without treatment (n=2), treated with 10mM rotenone for 30 minutes (n=3), 10mM for 3 hours (n=3), 100mM rotenone for 30 minutes (n=3), or 100mM for 3 hours (n=3), followed by LC-MS/MS analysis. Peak area for identified phosphopeptides were extracted using MaxQuant as previously described and statistical tests were used to identify features with significant association with treatment dose (0, 10mM or 100mM) and/or treatment length (0min, 30min, or 3 hours).

TABLE 1. List of phosphopeptides with significant association with treatment dose and/or length (*P* value for Pearson's correlation less than 0.005). Highlighted in red are proteins involved in apoptosis.

Protein	ModiPep	P value		Average peak area				
		dose	time	untreated	10mM 30min	10mM 3hr	100mM 30min	100mM 3hr
Protein Srrm1	HRPS(ph)SPAT(ph)PPPK	0.0661	0.0024	7.9E+06	3.5E+06	0.0E+00	2.3E+06	1.6E+06
Eukaryotic translation elongation factor 1 beta 2	YGPVSVADTTGSGAADAKDDDDIDLFGS(ph)DDEEESEDAK	0.0102	0.0033	2.4E+07	8.4E+06	9.1E+06	1.0E+07	5.9E+06
Protein Limch1	SINHQIES(ph)PGER	0.0002	0.1064	4.4E+06	1.8E+06	2.0E+06	0.0E+00	8.6E+05
Protein Ranbp2	SHET(ph)DGGSAHGDEEDDGPHFEPVVPLPDK	0.3602	0.0042	5.6E+06	4.2E+06	0.0E+00	3.5E+06	1.9E+06
Protein Xpc	AVKEEALS(ph)DDGDDFR	0.4183	0.0046	1.1E+06	7.2E+05	2.6E+05	7.7E+05	6.3E+05
Protein Smarcc2	TLTDEVNS(ph)PDADRR	0.0048	0.2821	3.3E+06	2.2E+06	3.1E+06	1.5E+06	6.4E+05
Microtubule-associated protein 1 A, isoform CRA_c	DTEQTEPEQREPT(ph)PYPDER	0.4440	0.0025	3.2E+06	2.0E+06	1.9E+06	3.0E+06	1.8E+06
Insulin-like growth factor 2 receptor, isoform CRA_b	LVSFHDDS(ph)DEDLLHI	0.0013	0.2444	1.9E+07	8.9E+06	1.3E+07	6.7E+06	9.3E+06
Nestin, isoform CRA_b	RKS(ph)IDTQEPLWSTEVAR	0.0025	0.7356	1.1E+07	6.3E+06	8.5E+06	0.0E+00	5.1E+06
Stathmin	RAS(ph)GQAFELILS(ph)PR	0.0024	0.1925	4.4E+05	2.9E+05	3.7E+05	1.6E+05	4.0E+04
Aquaporin-1	VWT(ph)SGQVEEYDLDADDINSR	0.0023	0.0855	2.9E+06	5.1E+05	1.3E+06	3.2E+05	2.5E+05
PDZ and LIM domain protein 4	S(ph)SIGSISLEDNR	0.0006	0.1885	1.7E+06	1.7E+06	1.2E+06	0.0E+00	0.0E+00
Mitogen-activated protein kinase 1	VADPDHDHTGFLT(ph)EYVATR	0.0039	0.0332	4.7E+06	5.7E+05	1.7E+06	8.9E+05	2.4E+05
Stromal interaction molecule 1	AEQS(ph)LHDLQER	0.0003	0.3437	3.5E+05	5.7E+05	4.2E+05	1.0E+06	1.1E+06
RNA-binding protein 8A	MREDYDS(ph)VEQDGDEPGPQR	0.0043	0.0175	0.0E+00	6.8E+05	2.2E+06	2.5E+06	2.2E+06
Eukaryotic translation initiation factor 3 subunit B	AKPAAQSEEETAAS(ph)PAAS(ph)PTPQSAQEPSAPGK	0.0030	0.3809	2.2E+07	1.1E+07	2.0E+07	1.0E+07	7.5E+06
Protein Rrp1	EGGS(ph)ETEASSADPSR	0.0848	0.0013	1.3E+06	9.3E+05	7.0E+05	9.4E+05	7.3E+05
Transgelin-2	NFS(ph)DNQLQEGK	0.0045	0.4080	2.2E+06	5.0E+06	3.0E+06	6.6E+06	6.7E+06
Cation-dependent mannose-6-phosphate receptor	GVGDDQLGEES(ph)EERDDHLLPM(ox)	0.0011	0.5675	2.6E+06	1.2E+06	2.2E+06	0.0E+00	5.9E+05
General transcription factor IIF subunit 1	HRT(ph)LTAEEAEEEWER	0.0028	0.2389	1.5E+06	2.7E+05	9.0E+05	1.6E+05	2.5E+05
Protein phosphatase 1 regulatory subunit 1B	IAESHLQTISNLSENQAS(ph)EEEDELGELR	0.0036	0.0155	2.5E+06	1.0E+06	1.2E+06	1.0E+06	1.0E+06
Caspase-12	QLSLQFPS(ph)DDEEDELQK	0.0048	0.2049	3.7E+06	1.4E+06	1.9E+06	4.8E+05	1.2E+06
Thioredoxin-like protein 1	IKQHLENDPGS(ph)NEDTDIPK	0.0022	0.0491	1.4E+06	6.8E+05	6.6E+05	3.3E+05	3.1E+05



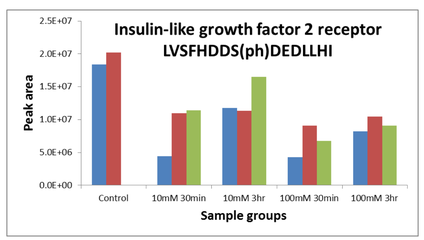


Figure 5: bar graph representation for relative abundance of two representative phosphopeptides.

Conclusions

- A quantitative phosphoproteomic workflow utilizing a two-step enrichment strategy coupled with high accuracy LC-MS/MS analysis is presented
- The developed workflow has good reproducibility and linearity, and is robust for quantification of global phosphorylation status
- The workflow is more appropriate for phosphopeptides with mono-phosphorylated serine or threonine
- Application of the described workflow to assess phosphoproteomic changes due to inhibition of mitochondrial function using C6 glioma cells identified some phosphorylated proteins of interest; however the validity remains to be validated

References

1. Zhang X, Ye J, Jensen ON, Roepstorff P. Highly Efficient Phosphopeptide Enrichment by Calcium Phosphate Precipitation Combined with Subsequent IMAC Enrichment. Mol Cell Proteomics. 2007 6:2032-42.
2. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M.(2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res.* 10(4):1794-805.
3. Cox J and Mann M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol.* 26(12):1367-72.

Acknowledgement

• This project used the Biomedical Mass Spectrometry Center and UPCI Cancer Biomarker Facility that are supported in part by award P30CA047904.