Journal of proteome research

Article

pubs.acs.org/jpr

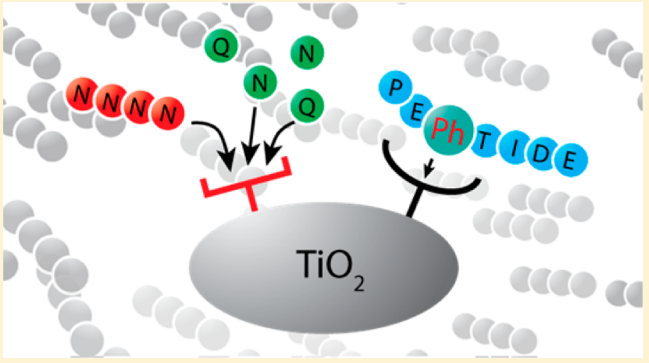
Displacement of N/Q-rich Peptides on TiO_2 Beads Enhances the Depth and Coverage of Yeast Phosphoproteome Analyses

Evgeny Kanshin,[†,‡] Stephen W. Michnick,*[,†] and Pierre Thibault*[,†,‡,§]

[†]Department of Biochemistry and [‡]Institute for Research in Immunology and Cancer, and [§]Department of Chemistry, Université de Montréal, Montréal, Quebec, Canada

Supporting Information

ABSTRACT: Phosphorylation is a reversible protein modification that regulates major cellular processes such as cell division, growth, and differentiation through highly dynamic and complex signaling pathways. Large-scale phosphoproteomics analyses have been greatly facilitated using affinity chromatography such as metal oxide affinity chromatography (e.g., TiO_2), which in combination with mass spectrometry has enabled unbiased detection and quantification of thousands of phosphorylation sites in a single experiment. However, global phosphoproteome analyses do not provide comparable enrichment yields for different model organisms. While the proportion of phosphopeptides exceed 90% in mammalian cells using TiO_2, similar levels have been notoriously difficult to achieve for yeast or dictylostelium cells. In a systematic study of TiO_2 using cell extracts from different organisms, we determined that phosphopeptides are coenriched with peptides containing repetitive stretches of glutamine and asparagine residues. The proportion of these nonspecific binders can reach up to 50% in cell extracts from budding yeast and thus limit the depth and comprehensiveness of phosphoproteomics analyses. To address this limitation, we developed an effective method that used decoy amino acids to reduce the extent of nonspecific peptide binding and improve the recovery and detection of low abundance phosphopeptides that remained undetected by conventional TiO_2 enrichment protocols.



KEYWORDS: *phosphoproteomics, TiO_2, yeast, N/Q-rich proteins, mass-spectrometry*

INTRODUCTION

Protein phosphorylation is one of the most ubiquitous post-translational modification found in eukaryotes and plays an important role in the regulation of different cellular functions such as protein stability, enzyme activity, DNA damage repair, cell proliferation, and cell signaling in response to intracellular and extracellular stimuli.[1] It is also implicated in many signal transduction pathways where sequential phosphorylation and dephosphorylation events are necessary for the progression through the cell cycle.[2] This modification is reversibly controlled by protein kinases and phosphatases that add and remove phosphate group on target substrates. Genes that encode kinases and phosphatases represent 2−3% of open reading frames of the human genome.[3−5] Misregulated kinase activity has been implicated in numerous human diseases including cancer,[6] metabolic disorders,[7] inflammation[8] and autoimmune diseases.[9] Consequently, the identification of site-specific changes in phosphorylation associated with aberrant kinase activity is of prime importance to understand the complex regulation of this modification in human diseases.

While a large proportion of protein phosphorylation in eukaryotes is found on hydroxylated amino acids such as serine, threonine, and tyrosine, this modification has also been reported on at least six other residues (histidine, cysteine, arginine, lysine, aspartate, and glutamate), though to a significantly lower extent.[10] Current estimates suggest that approximately 100 000 sites could be present in human proteins and that the extent to which a site is phosphorylated is typically less than 30%.[11,12] Typical methods for studying protein phosphorylation such as metabolic labeling with radioactive phosphate, phosphospecific antibodies or *in vitro* kinase assays are time and resource intensive and often necessitate prior knowledge of the modified residue. In this context, recent advances in sample preparation, affinity chromatography, sensitive mass spectrometry (MS) instrumentation and computational methods have greatly facilitated the large scale analysis of protein phosphorylation. MS-based phosphoproteomics is now recognized as an enabling technology providing both qualitative and quantitative information on thousands of phosphorylation sites from microgram amounts of protein extracts. The remarkable developments in both phosphopeptide enrichment methods and high throughput MS instrumentation that we witnessed over the past decade have led to a significant expansion in the

Received: March 4, 2013
Published: April 23, 2013

inventory of phosphosites that now represents more than 104 000 sites curated from the literature.[13]

Several phosphopeptide enrichment methods are currently available and protocols are regularly evolving to improve recoveries.[10,14−17] The most frequently used methods include immobilized metal affinity chromatography (IMAC)[18−20] and metal oxide affinity chromatography (MOAC).[10,14−16] IMAC is traditionally performed using Fe(III) or Ga(III) immobilized via nitrilotriacetic/iminodiacetic acid onto a solid phase support. Ion pair formation between the immobilized cation and the negatively charged phosphate group enabled the selective capture of phosphopeptides. To mitigate the nonspecific retention of acidic peptides to the IMAC resin, binding and washing steps are often performed at pH ∼3 below the pK_a of carboxylic acid groups.[21] Elution is subsequently accomplished with inorganic phosphate or ammonium hydroxide buffers at basic pH. Esterification of carboxylic acid was also proposed to reduce nonspecific binding,[22] though side reaction products (partial hydrolysis of peptides, deamidation of asparagine and glutamine residues) can increase sample complexity and further complicate data analysis.

MOAC affinity media make use of metal oxides such as TiO_2, ZrO_2, Nb_2O_5, and Al_2O_3 to form complexes with the negatively charged phosphate group. At present, TiO_2 is by far the most common MOAC medium used in phosphoproteomics experiments. Similar to that of IMAC, phosphopeptides are enriched onto TiO_2 beads at acidic pH and eluted at basic pH. To decrease nonspecific binding of acidic peptides, MOAC protocols typically use buffer modifiers containing saturated solution of organic acids, such as DHB, glycolic acid, and lactic acid.[23,24] IMAC and MOAC have complementary selectivity as IMAC typically yield a higher proportion of multiplyphophorylated peptides while MOAC preferentially enriches singly phosphorylated peptides.[25] It is noteworthy that IMAC and TiO_2 can be combined in a sequential fashion to enrich both monophosphorylated and multiply phosphorylated peptides from cell digest.[26]

Remarkable improvements in affinity enrichment protocols using either IMAC or MOAC have been made over the past 10 years, and the proportion of identified phosphopeptides from tryptic digests now often exceed 80% for mammalian cells. However, none of the current affinity media provide 100% selectivity toward the enrichment of phosphorylated peptides. In an effort to improve the selectivity of MOAC affinity media, we conducted systematic phosphoproteomics studies on yeast, drosophila, and human cells and found that phosphopeptides recoveries varied significantly across species. The examination of nonspecific peptide binders revealed that TiO_2 affinity medium can retain peptides containing multiple asparagine and glutamine residues (N/Q-rich). The presence of these N/Q-rich peptides is observed for different cell types, but is significantly more abundant in yeast cells, and adversely affects phosphopeptide recoveries. Binding of N/Q-rich peptides is specific to TiO_2 beads and is not observed using IMAC. We also report an effective method to remove N/Q-rich peptides and enhance TiO_2 selectivity, thereby facilitating the identification of trace level phosphopeptides not typically observed using traditional protocols.

MATERIALS AND METHODS

Acetonitrile (ACN) and HPLC grade water were purchased from Thermo Fisher Scientific. Other reagents and chemical were obtained from Sigma-Aldrich unless otherwise stated.

Cell Culture

S. Cerevisiae strain YAL6B (a generous gift of Ole Jensen, University of Southern Denmark) was grown in liquid YPD medium (1% yeast extract, 2% bacto-peptone, 2% glucose; all from Bioshop). Yeast cells were grown by rotation at 500 rpm at 30 °C until they reached an A_{600} absorbance between 0.6 and 0.8. *Drosophila melanogaster* Schneider S2 cells were cultured in serum-free and protein-free insect medium at 28 °C until 95% confluence. HeLa and HEK293 were cultured in high glucose DMEM (Hyclone SH30081.02) supplemented with 10% fetal bovine serum, 1% L-glutamine and pyruvate until 95% confluence. Cells were harvested by centrifugation for 5 min at 2000× *g* at 4 °C, washed in ice-cold phosphate buffered saline (PBS), snap frozen in liquid nitrogen and stored at −80 °C until further analysis.

Cell Lysis and Protein Extraction

S2, HeLa and HEK293 cells were lysed by sonication (3 × 5 s pulses) directly in the corresponding lysis buffer (see below) on ice. Since yeast cells are more difficult to disrupt, they were lysed in a freezer mill apparatus (SPEX Sample Prep) that allows breaking the cells in liquid nitrogen, which preserves proteins and their modifications. A total of 8 cycles, comprising 2 min grind and 2 min cool down, were used to disrupt the cells. Efficiency of cell lysis was monitored by observation of the cells under a microscope. Frozen powders were stored at −80 °C for further analysis. We used several different protocols for protein extraction and digestion. Our preferred method was based on using 1% sodium deoxycholate (SDC) in 50 mM ammonium bicarbonate (ABC) as lysis buffer. Samples were boiled for 5 min, which resulted in protein denaturation and inactivation of cellular proteases and phosphatases;[27,28] there was thus no need to add protease/phosphatase inhibitors that could subsequently interfere with either tryptic digestion or phosphopeptide purification. SDC also increases trypsin activity and can be easily removed from the peptide sample by acidification. Alternatively samples were boiled for 5 min in 0.1% RapiGest (Waters) solution in 50 mM ABC or solubilized in 8 M urea 50 mM TRIS, pH = 8 with HALT phosphatase inhibitors (Thermo Fisher Scientific).

Trypsin Digestion

Samples were centrifuged 13000× *g* for 10 min at 4 °C to remove insoluble debris and protein concentrations were measured by BCA protein assay (Thermo Fisher Scientific). Proteins were reduced with 5 mM dithiothreitol (DTT) for 30 min at 37 °C and alkylated with 15 mM iodoacetamide (IAA) in the dark for 30 min at room temperature. Excess of IAA was quenched with 5 mM DTT in the dark for 15 min at room temperature. Urea concentration was reduced by 6 times dilution with 1 mM $CaCl_2$ in 20 mM TRIS, pH = 8. Proteins were digested by sequencing grade modified trypsin (Promega) at 37 °C overnight (enzyme/protein ratio: 1/50). The peptide samples were acidified to 1% FA and insoluble material was removed by centrifugation at 20000× *g* for 10 min. Peptides were desalted on SepPak tC18 cartridges (Waters) according to the manufacturer protocol; eluates were snap frozen in liquid nitrogen, dried on speedvac and stored at −80 °C for further analysis.

Phosphopeptide Isolation

Peptide samples were subjected to the TiO_2 enrichment protocol as described previously[29,30] except when noted. Sample loading, washing, and elution were performed in

homemade spin columns assembled in a similar fashion to that of StageTips[31] and were comprised of 200 uL pipette tips with a frit made of SDB-XC membrane (3M) and filled with 5 mg of TiO_2 beads (Canadian Life Science). SDB-XC material has similar hydrophobic properties as C_{18} and allows combining phosphopeptide enrichment and desalting steps. Centrifugation speed was set to 2000× *g*. Before peptide loading columns were equilibrated with 100 μL of loading buffer (250 mM lactic acid in 70% ACN 3% TFA). Peptides (750 μg of digest) were solubilized in 100 μL of loading buffer and applied on TiO_2 column. Each column was washed with 100 μL of loading buffer followed by two 100 μL portions of washing buffer (70% ACN 3% TFA) to remove nonspecific binding. Subsequent washing with 50 μL of 1% FA was used to equilibrate SDB-XC frit material, then phosphopeptides were eluted from TiO_2 with 2 × 50 uL portions of 500 mM K_2HPO_4, pH 7 and retained on SDB-XC. Peptides were desalted with 50 μL of 1% FA and subsequently eluted from SDB-XC with 50% ACN 0.5% FA. Eluates were dried on speedvac and stored at −80 °C. Prior to MS analysis peptides were solubilized in 4% FA. In order to improve enrichment efficiency by removal of poly-N/Q peptides we introduced additional washing steps (2 × 100 uL) just before the last wash with 70% ACN 3% TFA. We used the following amide modifiers (solutions in 70% ACN, 3% TFA): 125 mM of each asparagine and glutamine (Bioshop), 125 mM lactamide, 125 mM malonamide and 125 mM of *N,N′*-tetramethyldiamide of tartaric acid. For the titration of TiO_2 phase with peptides, we used mixtures of light and heavy isotopic forms of synthetic phosphopeptides (MS PhosphoMix Standards) to evaluate absolute recovery of phosphopeptides. These were spiked into the sample before (heavy forms) and after (light forms) the enrichment to a level of 200 fmoles. Phosphopeptide enrichment on PhosSelect IMAC beads was performed essentially as described previously,[32] except that SDB-XC material was used to make a frit in the spin column.

NanoLC–ESI–MS/MS

A nanoflow HPLC system (Eksigent, Thermo Fisher Scientific) was used for online reversed phase chromatographic separation of peptide samples prior to nanoESI and MS detection. The peptides were loaded onto 5 mm long trap column (inner diameter 300 μm) and separated on 18 cm long fused silica capillary analytical column (inner diameter 150 μm), both packed with 3 μm 200A Magic AQ C18 reverse phase material (Michrom Bioresources Inc.). The peptides were eluted at 600 nL/min by an increasing concentration of ACN (5–40% of buffer B (100% ACN 0.2% FA) in either 51 or 100 min). At the end of the gradient, the column was washed with 80% buffer B and re-equilibrated with 5% buffer B. Peptides were analyzed utilizing automated data-dependent acquisition on a LTQ-Orbitrap XL or Elite mass spectrometers (Thermo Fisher Scientific). Each MS scan was acquired at a resolution of 240 000 (Elite) or 60000 (XL) (at 400 *m/z*) with lock mass option enabled and was followed by up to 12 MS/MS scans using collision induced dissociation (CID). Dynamic exclusion was defined by a list size of 500 features and exclusion duration of 30 s. The automatic gain control (AGC) target value was 1×10^6 for MS scans in the Orbitrap and 5×10^4 for the MS/MS scans in the LTQ.

Data Processing and Analysis

MS data were analyzed using the MaxQuant[33,34] software version 1.3.0.5 and searched against the corresponding SwissProt part of the Uniprot database (www.uniprot.org, *S. Cerevisiae*, 6629 entries; *D. Melanogaster*, 17533 entries; or *H. Sapiens*, 68079 entries). A list of 248 common laboratory contaminants provided by MaxQuant was also added to the database. The enzyme specificity was set to trypsin with the maximum number of missed cleavages set to 2. The precursor mass tolerance was set to 20 ppm for the first search (used for nonlinear mass recalibration) and was set to 6 ppm for the main search, fragment mass deviation was set to 0.5 Th. Oxidized methionines, phosphorylation of serine, thereonine and tyrosine residues and N-terminal protein acetylation were searched as variable modifications. Carbamidomethylation of cysteines was searched as a fixed modification. The false discovery rate (FDR) was calculated based on searching a decoy database and was set to <1% for both peptide and protein identifications, the minimum peptide length was set to 6. To match identifications across different conditions the "match between runs" option in MaxQuant was enabled with a retention time window of one minute. Gene Ontology enrichment analyses were performed by using the DAVID functional annotation tool.[35,36] Motif enrichment analyses were obtained using motif-x.[37,38]

RESULTS AND DISCUSSION

The use of TiO_2 to affinity purify phosphopeptides from protein digests has gained significant popularity since the introduction of buffer additives such as 2,5-dihydroxybenzoic acid (DHB) that enabled the reduction of nonspecific binding of acidic peptides.[23] Further improvement of isolation protocols used hydrophilic hydroxylated modifiers such as lactic acid to enhance selectivity and capacity of TiO_2 toward phosphorylated peptides without compromising reverse phase LC performance.[24] While enrichment levels of more than 90% are typically obtained on mammalian cell extracts using optimized isolation protocols, variable yields can be observed for protein digests of different cell model systems. For example, previous phosphoproteomics experiments performed on mouse, drosophila and dictyostelium cell extracts provided enrichment levels of 89, 72, and 17%, respectively.[39,40] To investigate the source of this variability and determine the nature of nonspecific binders, we evaluated different protein isolation protocols and purified phosphopeptides using the same enrichment method. These experiments are described below together with modifications to existing protocols to improve phosphopeptide recovery yields.

Table 1. Number of Identified Phosphopeptides using Different Protein Isolation Protocols

	Cell type	Rapigest	SDC	Urea
Peptides (all)	HEK293	2138	2429	2309
	HeLa	2090	2327	2309
	S2	1903	2383	2223
	Yeast	2290	2383	1985
Phosphopeptides	HEK293	1940	1832	1892
	HeLa	1856	1597	1776
	S2	1536	1433	1631
	Yeast	1448	1249	1018
Phosphosites (conf. >0.75)	HEK293	1714	1637	1643
	HeLa	1622	1409	1557
	S2	1327	1264	1352
	Yeast	1217	1065	902

dx.doi.org/10.1021/pr400198e | J. Proteome Res. 2013, 12, 2905–2913

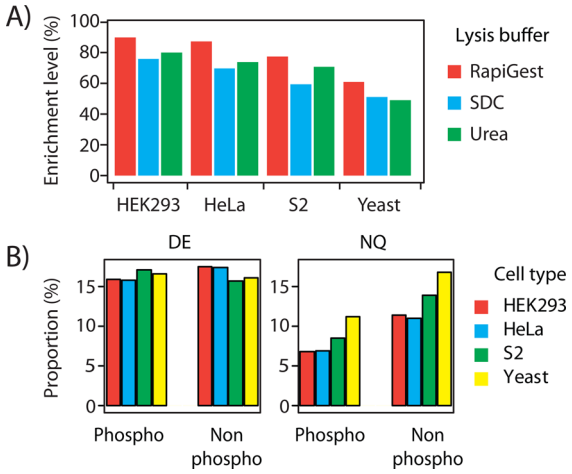


Figure 1. Enrichment of phosphopeptides using TiO_2 for different cell model systems and protein solubilization protocols. (A) Distribution of phosphopeptide enrichment levels for digests of human (HEK293, HeLa), Drosophila (S2) and yeast cells following the solubilization of protein extracts using Rapigest, SDC and urea. (B) Proportion of acidic (DE) and amide (NQ) amino acids in identified phosphopeptides and nonphosphorylated peptides from SDC.

Peptides Containing Poly Asparagine and Glutamine Are Enriched on TiO_2 Beads

We used three different protein isolation protocols namely RapiGest, sodium deoxycholate (SDC) and urea to solubilize and digest extracts of human (HEK293 and HeLa), *Drosophila melanogaster* (S2) and *S. cerevisiae* (yeast) cells. Each tryptic digest (750 μg of cell extract) was enriched on TiO_2 stage tips using 250 mM lactic acid (70% ACN, 3% TFA) prior to phosphopeptide elution with phosphate buffer. Samples corresponding to one-third of the phosphopeptide extracts were analyzed by LC−MS/MS on a LTQ-Orbitrap Elite and data were processed using MaxQuant. Approximately 2,000 peptides were identified irrespective of the method used, though a higher number of phosphopeptides was generally observed using Rapigest (Table 1). For this solubilization protocol, the relative proportion of phosphopeptide identified for human, S2 and yeast cells was 90, 81 and 63%, respectively (Figure 1). These experiments also indicated that phosphopeptide enrichment levels were systematically lower for yeast (51−63%) compared to either S2 (60−81%) or human (69−91%) cells (Figure 1A). For convenience, peptide identifications obtained for each cell extract are reported in Table S1 and high confidence phosphosites are listed in Table S2, Supporting Information. To determine the nature of nonspecific binders, we next examined the distribution of aspartate and glutamate residues for both phosphopeptides and nonphosphopeptides and found that their relative proportions were similar, ranging between 15 and 17%. This is illustrated in Figure 1B for TiO_2-enriched extracts following SDC protein solubilization. Similar results were also obtained when protein extracts were solubilized using RapiGest or urea (data not shown). Acidic amino acid residues were not enriched in nonphosphorylated peptides compared to phosphopeptides, demonstrating that lactic acid used in our protocol is effective at removing these. Closer examination of these analyses revealed that nonphosphorylated peptides contained a higher proportion of asparagine (N) and glutamine (Q) residues compared to phosphopeptides (Figure 1B). Furthermore, the relative enrichment of these N and Q residues in nonphosphorylated peptides differed among species, ranging from 11% in human cells to 17% in yeast cells.

To our knowledge, nonspecific binding of N/Q-rich peptides to TiO_2 has not been reported before. To determine if these nonspecific binders depend on the type of affinity medium, we compared the distribution of amino acids in both phosphopeptides and nonphosphorylated peptides for yeast extracts

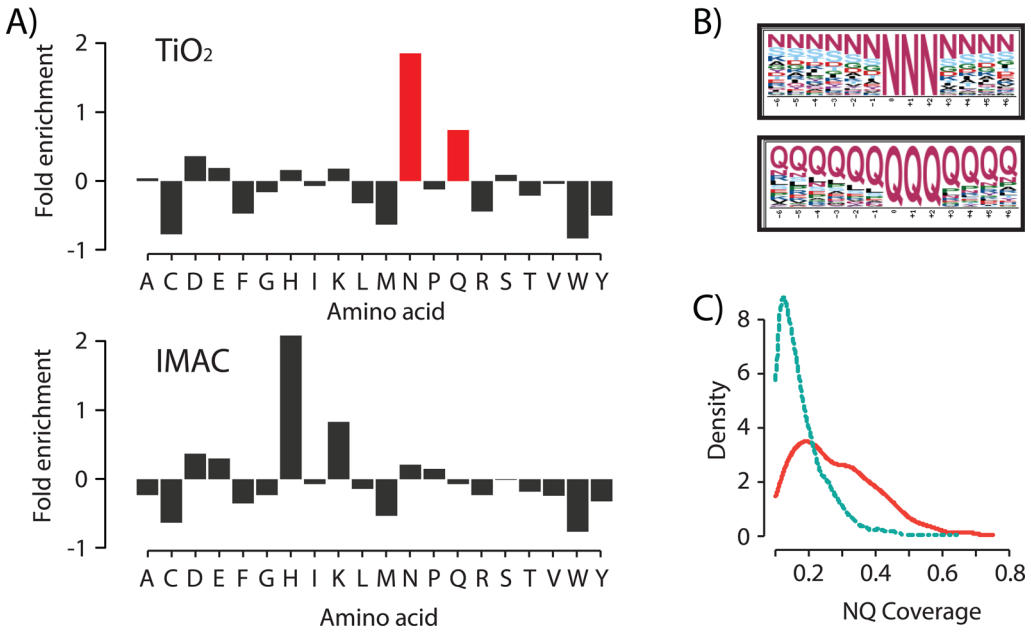


Figure 2. Distribution of amino acids on peptides found in TiO_2_ and IMAC-enriched extracts of yeast cells. (A) Enrichment of specific amino acids between nonphosphorylated peptides retained on either TiO_2 or IMAC and proteome. (B) Motif analysis of peptide sequences displaying clustered asparagines and glutamine residues. (C) Coverage of peptide sequences with asparagines and glutamine residues observed for phosphorylated (green) and nonphosphorylated (red) peptides retained on TiO_2.

dx.doi.org/10.1021/pr400198e | J. Proteome Res. 2013, 12, 2905−2913

Table 2. Examples of N/Q-rich Peptides Identified in Yeast Extracts

peptide sequence	score	N+Q	ORF	gene
LLNNDRNHQNN(de)NNNNNNNNNNNNNNNNNNNNNIINK	141	28	YBL084C	CDC27
FQALQQQQQQQQNQQQQNQQPQQQQQQQQNPK	300	26	YPL016W	SWI1
FQALQQQQQQQQNQQQQNQQPQQQQ(de)QQQQNPK	234	25		
SNFMPQQQQQGGQPLIMNDQQQPVMSEQQQQQQQQQQQQ	231	23	YNL016W	PUB1
VQQQQQQQQQQQQQQQQQQQQQQQQR	392	23	YDR145W	TAF12
VQQQQQQQQQQQQQQ(de)QQQQQQQQQR	260	22		
QLQLQQQQQQQQQQQQQQQQPSDVDNGNAAASEENYPK	275	21	YHR030C	SLT2
QLQLQQQQ(de)QQQQQQQQQQQQQPSDVDNGNAAASEENYPK	221	20		
NNMMLQLQEQQQQQQQLQQQHQQLDQEDNNGPLLIK	241	19	YBR212W	NGR1
NSYHGYYNNNNNNNNNNNNN(de)NNSNATNSNSAEK	165	18	YGL014W	PUF4
NSYHGYYNNNNNNNNNNNNNNNSNATNSNSAEK	122	19		
QQAAQAYNNTSNSNSSNPASIPTENVPNSSQQQQQQQQQQTR	144	19	YDR145W	TAF12
QQQQQEQLQQNQQQEEQQKAQLQEQNQR	199	19	YKL088W	CAB3
QQQQQEQLQ(de)QNQQQEEQQKAQLQEQNQR	173	18		

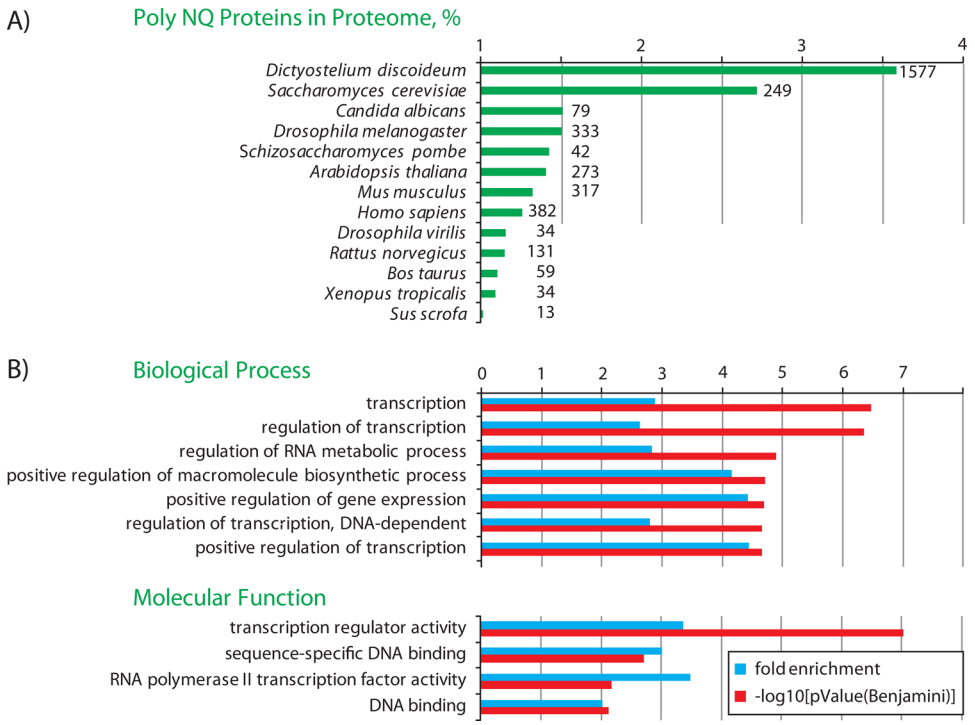


Figure 3. Proteome content and gene ontology (GO) terms associated with N/Q-rich proteins. (A) Relative proportion of N/Q-rich proteins in the proteome of different species. (B) Most significant GO terms for biological process and molecular function for N/Q-rich proteins found in TiO_2-enriched extracts of yeast cells.

purified on TiO_2 and IMAC beads. In each case, 750 μg of yeast tryptic digests were purified on affinity media prior to LC−MS/MS analyses. We identified a total of 2874 and 1221 phosphopeptides for digests purified on TiO_2 and IMAC, respectively. The comparison of the relative enrichment of individual amino acids in nonphosphorylated peptides retained on affinity media indicated that a higher fold enrichment of basic and acidic residues (e.g., histidine, lysine, aspartate and glutamate) was observed for nonphosphorylated peptides isolated with IMAC, whereas asparagine and glutamine where specifically enriched in the corresponding peptides using TiO_2 beads (Figure 2A). The isolation of N/Q-rich peptides on TiO_2 beads thus appears to be a characteristic property of this affinity medium and further supports the different selectivity observed between IMAC and TiO_2 beads.

Interestingly, we also noted that N and Q residues were not randomly distributed within the backbone of nonphosphorylated peptides, rather they formed homogeneous clusters that extended up to 25 residues. Motif logo analyses revealed stretches of N and Q residues of more than 5 residues in the corresponding peptides (Figure 2B). Also, the density distributions for coverage of peptide amino acid sequence with asparagine and glutamine residues are different for phosphorylated and nonphosphorylated peptides retained on TiO_2 (Figure 2C). The proportion of N/Q residues within the sequence backbone ranged from 14% for phosphopeptides to 22% for nonphosphorylated peptides isolated with TiO_2 beads (Figure 2C). Selected examples of N/Q-rich peptides identified in yeast extracts are presented in Table 2 and display extended clusters of either N or Q residues.

dx.doi.org/10.1021/pr400198e | J. Proteome Res. 2013, 12, 2905−2913

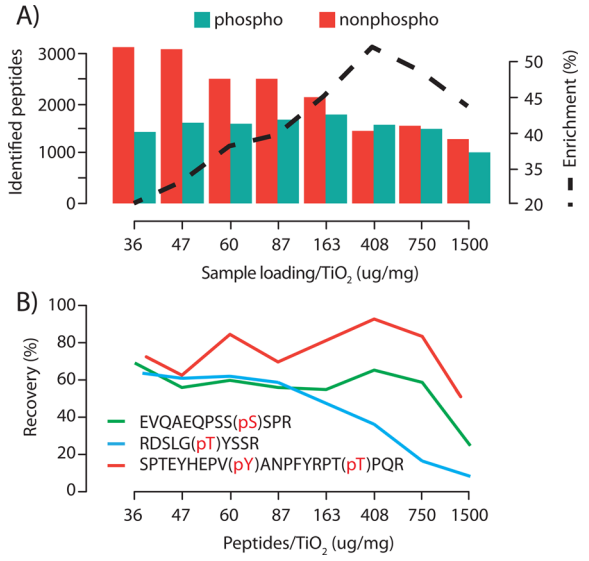


Figure 4. Identification and recovery of phosphopeptides on TiO_2 beads for different sample loadings. (A) Proportion and distribution of phosphopeptides identified in yeast digests according to sample loadings. (B) Recovery of phosphopeptides from isotopically labeled phosphopeptides spiked at 400 fmoles in the digest of yeast extracts.

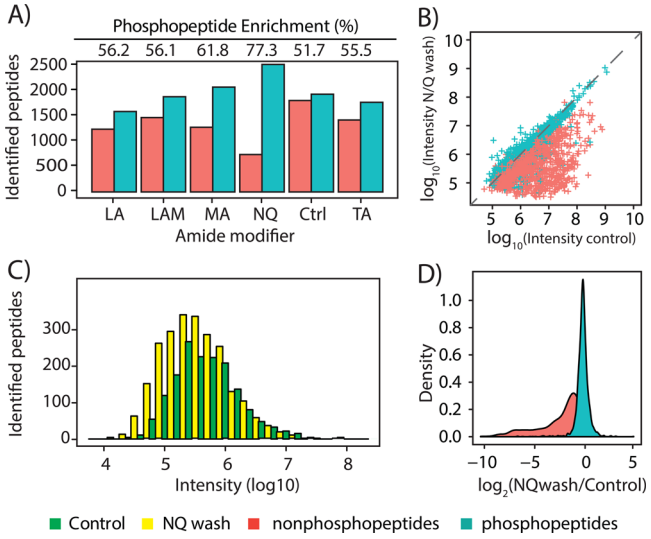


Figure 5. Comparison of phosphopeptide enrichment levels for different amide-containing wash buffers. (A) Enrichment level and distribution of phosphopeptides identified in yeast digests using different small molecule amide modifiers (125 mM each): lactic acid (LA), lactamide (LAM), malonamide (MA), mixture of asparagine and glutamine (NQ), *N,N′*-tetramethylamide of tartaric acid (TA); and standard protocol without any modifiers (Ctrl). (B) Scatter plot of Log_{10} peptide intensity for TiO_2-enriched yeast digest with and without NQ wash. NQ wash resulted in a large decrease of nonphosphorylated peptides while phosphopeptides remained unaffected (option "match between runs" in MaxQuant was used). (C) Intensity distribution of phosphopeptides with and without NQ wash. (D) Distributions of log_2 intensity ratios of phosphopeptides (green) and nonphosphorylated peptides (red) with and without NQ wash.

Distribution of N/Q-rich Proteins Differs Across Species and Accounts for Differential Phosphopeptide Recoveries

The significant difference in enrichment efficiency observed across cell types was intriguing in view of the fact that the same phosphopeptide enrichment protocol was used in all cases. We reasoned that these differences could be associated with the number and abundance of N/Q-rich proteins in different species. Accordingly, we determined the number of proteins from Uniprot that contained compositional bias with poly-N or poly-Q sequences for more than 10 different species including those described previously (Figure 3A). Among all species examined, we noted that the proportion of N/Q-rich proteins is the highest for *Dictyostelium discoideum* (slime mold) where these structural features corresponded to approximately 3.6% (1577 proteins) of its proteome. This was followed by *S. Cerevisiae*, *Candida albicans* and *D. melanogaster* where N/Q-rich proteins represent 2.7% (249 proteins), 1.5% (79 proteins), and 1.5% (333 proteins) of their respective proteomes. In contrast, the proportion of these N/Q-rich proteins in the mouse and human proteomes correspond to approximately 1.3%. The higher number of proteins with N/Q-rich domains in the proteomes of *D. discoideum* and *S. Cerevisiae* account at least in part for the lower enrichment of phosphopeptides in the cells extract of these species. The relative abundance of these proteins in the corresponding extracts could also provide additional insights into the prevalence of nonspecific N/Q-rich peptides enriched on TiO_2 affinity medium.

It is noteworthy that a group of prion proteins in yeast contain N/Q-rich domains (e.g., YPL226W, SUP35, URE2) though some of these proteins are thought to be non-pathogenic, and to propagate epigenetic information through a loss of function upon aggregation.[41] In humans, polyQ diseases are a family of neurodegenerative disorders caused by expansion of a CAG trinucleotide repeat of certain genes, and are associated with Huntington disease, spinocerebellar ataxias and X-linked spinal-bulbar muscular atrophy.[42] Expanded polyQ and N/Q repeats affect protein folding and aggregation. The unsuspected occurrence of N/Q-rich peptides in TiO_2 extracts prompted us to examine gene ontology (GO) terms associated with the corresponding proteins (Figure 3B). We found that proteins containing N/Q stretches were involved in biological processes enriched in transcription (p-value: 3.98×10^{-7}), regulation of RNA metabolic processes (p-value: 1.58×10^{-5}), molecular functions associated with transcription regulator activity (p-value: 8.91×10^{-8}) and DNA binding (p-value: 6.31×10^{-3}). In retrospect these observations are consistent with known functions of these proteins. Of particular note, protein with N/Q-rich stretches belong to the general class of low complexity (LC)-sequences containing proteins that have been shown to be principal components of cellular bodies called RNA granules, including Cajal bodies, stress granules, and P-bodies. These proteins have been demonstrated to undergo liquid−liquid phase transitions to hydrogels *in vitro* and *in vivo*, perhaps organizing RNA processing machineries in distinct ways and locations in the cell.[43−45] These structures are of great interest but difficult to isolate and it is thus interesting to suggest that TiO_2 bead affinity chromatography could be used as a component of an isolation protocol.

Enhancement of Phosphoproteome Analyses Using Amide Modifiers

A high abundance of N/Q-rich proteins in yeast and dictyostelium cells could account for a large proportion of binding of N/Q-rich peptides to TiO_2 affinity medium. We reasoned that increasing sample loading on TiO_2 beads should favor the competition between N/Q-rich peptides and phosphopeptides for active sites. Enhanced binding of

dx.doi.org/10.1021/pr400198e | J. Proteome Res. 2013, 12, 2905−2913

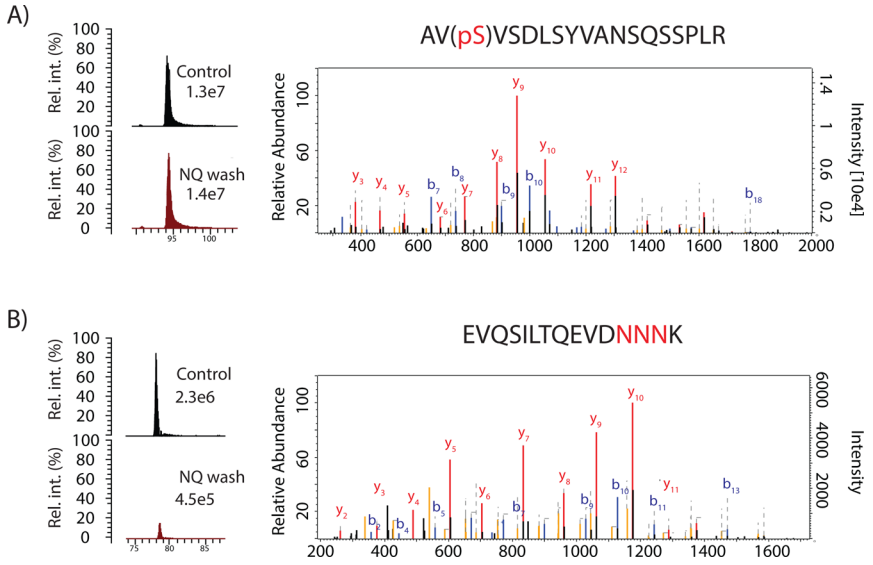


Figure 6. Identification and abundance profile of phosphopeptides and N/Q rich peptides with and without NQ wash. Extracted ion chromatograms for TiO_2 extracts with and without NQ wash, and MS/MS spectrum for (A) AV**pS**VSDLSYVANSQSSPLR from Acetyl-CoA carboxylase, and (B) EVQSILTQEVDNNNK from Inosine triphosphate pyrophosphatase. The reduced binding of N/Q rich peptides on TiO_2 affinity beads enhanced the recovery of low abundance peptides not typically detected using traditional enrichment protocol.

phosphopeptides is expected if we assume that phosphopeptides bind more strongly and displace N/Q-rich peptides on the TiO_2 phase. To evaluate this possibility, we performed LC−MS/MS experiments on extracts obtained from increasing amounts of yeast tryptic digest loaded onto TiO_2 stage tips (Figure 4). We noted that the number of phosphopeptides increased from 1360 to 1690 for a sample loading up to 163 μg per mg of TiO_2, and gradually decreased beyond this point. In contrast, the number of nonphosphorylated peptides progressively decreased from 2970 to 1222 for sample loading of 36 to 1500 μg per mg of TiO_2. An optimal phosphopeptide enrichment level of 54% was obtained for a sample loading of 408 μg per mg of TiO_2.

To determine the binding efficiency and recovery of phosphopeptides from TiO_2 beads, we used mixtures of light and heavy isotopic forms of synthetic phosphopeptides (400 fmoles) that were spiked before (heavy forms) and after (light forms) TiO_2 enrichment (Figure 4B). As observed, an optimal recovery (67−90%) was obtained for 2 out of 3 phosphopeptides at 408 μg/mg of TiO_2, while the recovery of the basic phosphopeptide RDSLGpTYSSR progressively decreased over the same sample loading range. These results suggest that phosphopeptides and N/Q-rich peptides might interact with different binding sites on the TiO_2 surface and that optimization of peptide/TiO_2 ratio cannot be solely used to mitigate nonspecific binding.

Next, we evaluated alternate approaches to increase the selectivity of TiO_2 beads for phosphoproteomics analyses of yeast cells using competitive binding with different buffer modifiers. In view of the common amide functionality of asparagine and glutamine residues, we surmised that this group could mediate binding to TiO_2 beads. Preliminary experiments were performed using increasing concentration of the amino acids asparagine and glutamine in the wash buffer following the removal of nonspecific binders using 3% TFA in 70% ACN (Supplementary Figure S1, Supporting Information). The phosphopeptide enrichment increased progressively with the concentration of amino acids and reached an optimal value of ~80% for a concentration of 125 mM. We also tested other amide-containing compounds including polyasparagine tri and tetrapeptides (Supplementary Figure S2, Supporting Information), lactamide, diamide of malonic and tartaric acids (Supplementary Figure S3, Supporting Information) as buffer modifiers during the wash steps. Peptide identifications obtained for each amide modifier are presented in Table S3, Supporting Information. The number of identifications and the phosphopeptide enrichment levels obtained for the different wash solutions are shown in Figure 5A and indicate that a maximum phosphopeptide recovery of 75% is obtained using 125 mM of asparagine and glutamine in the wash buffer (N/Q wash). A comparison of identifications obtained using the N/Q wash is provided as Supplementary Figure S4, Supporting Information, and a complete list of peptide identifications is reported in Table S4, Supporting Information.

The enhanced identification of phosphopeptides was also accompanied by a large reduction in the number and abundance of nonphosphorylated peptides (Figure 5B). On average, we noted a 5-fold reduction in the intensity of nonphosphorylated peptides with no noticeable change in phosphopeptide intensities when using the N/Q wash (Figure 5B,D). Nonspecific peptide binders that comprised long stretches of N/Q residues were most affected and showed the highest decrease in abundance. The reduced proportion of N/Q-rich peptides resulted in a 30% increase in phosphopeptide coverage and most of these additional identifications corresponded to low abundance phosphopeptides of intensities less than 300,000 counts (Figure 5C). This protocol was found to be reproducible and we typically obtained phosphopeptide enrichment yield of 80% on yeast extracts analyzed using different LC−MS/MS systems over a time period of 9 months (Supplementary Table S5, Supporting Information). Figure 6 shows representative examples of peptides identified in the yeast proteome with and without the additional wash with the amide modifier. The N/Q wash enabled a 5-fold reduction in abundance of the nonspecific binding peptide EVQSILTQEVDNNNK from Inosine triphosphate pyrophosphatase

dx.doi.org/10.1021/pr400198e | J. Proteome Res. 2013, 12, 2905−2913

(P47119) while the intensity of the phosphopeptide AV**pS**VSDLSYVANSQSSPLR from Acetyl-CoA carboxylase (Q00955) remained relatively constant (Figure 6A and B).

CONCLUSION

The systematic evaluation of TiO_2-enriched extracts from mammalian, fly and yeast cells revealed that phosphopeptide recoveries can be compromised by the presence of N/Q-rich peptides that bind nonspecifically to this affinity media. This was particularly true for yeast cells where the proportion of phosphopeptides in TiO_2 extracts was consistently lower (51−63%) compared to either S2 (60−81%) or human (69−91%) cells, irrespective of the protein solubilization method used. The occurrence of nonspecific binders in TiO_2 extracts was related to the relative proportion of N/Q-rich proteins in the proteomes of the corresponding organisms, which represented 2.7, 1.5 and 1.3% for yeast, fly and human cells, respectively. The presence of N/Q-rich peptides on TiO_2 beads was specific to this affinity medium, as IMAC beads did not show a propensity to bind to these peptides. Although not explored in the present study, we anticipate that the affinity of TiO_2 toward N/Q-rich peptides could be used advantageously to enrich for peptides from RNA granules, yeast prion proteins or proteins associated with polyQ diseases in humans.

Titration experiments on TiO_2 beads using increasing amounts of tryptic digest of yeast cells suggested that phosphopeptides and N/Q-rich peptides might interact with different binding sites and that under optimal peptide/TiO_2 ratio the phosphopeptide enrichment levels only reached 54%. To mitigate nonspecific binding of N/Q peptides on TiO_2 beads, we evaluated different amide-containing modifiers in the wash buffers, and found that a mix of asparagine and glutamine amino acids provided enhanced phosphopeptide recoveries. By using a wash solution containing 125 mM of asparagine and of glutamine, we observed a 5-fold reduction in the intensity of nonphosphorylated peptides with no noticeable change in phosphopeptide intensities. The decreased abundance of nonspecific peptide binders including those containing long stretches of N/Q residues enhanced the detection of low abundance phosphopeptides and provided a 30% increase in phosphopeptide enrichment.

ASSOCIATED CONTENT

Supporting Information

Supplementary figures and tables. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*S.W.M.: Phone 514 343-5849, Fax 514 343-2015, E-mail stephen.michnick@umontreal.ca. P.T.: Phone: (514) 343-6910, Fax: (514) 343-6843. E-mail: pierre.thibault@umontreal.ca.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

IRIC receives infrastructure support funds from the Fonds de la Recherche en Santé du Québec (FRSQ). This work was carried out with the financial support of operating grants from the Canadian Institute for Health Research (MOP-GMX-152556 to S.W.M. and P.T.), National Science and Engineering Research Council (NSERC) (194582 to S.W.M.) and the Canada Research Chair programs. We also thank the Eric Bonneil for technical assistance.

ABBREVIATIONS

ACN, acetonitrile; CAD, collision activated dissociation; DHB, 2,5-dihydroxybenzoic acid; DTT, dithiothreitol; IMAC, immobilized metal ion affinity chromatography; LC−MS/MS, liquid chromatography−tandem mass spectrometry; LTQ, linear trap quadrupole; MOAC, Metal oxide affinity chromatography; PBS, phosphate buffer saline; PTM, post-translational modification; SDC, Sodium dodecyl cholate; TFA, trifluoroacetic acid; TiO_2, titanium dioxide.

REFERENCES

(1) Manning, G.; Plowman, G. D.; Hunter, T.; Sudarsanam, S. Evolution of protein kinase signaling from yeast to man. *Trends Biochem. Sci.* **2002**, *27* (10), 514−20.

(2) Nurse, P. A long twentieth century of the cell cycle and beyond. *Cell* **2000**, *100* (1), 71−8.

(3) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The protein kinase complement of the human genome. *Science* **2002**, *298* (5600), 1912−34.

(4) Roy, J.; Cyert, M. S. Cracking the phosphatase code: docking interactions determine substrate specificity. *Sci. Signal.* **2009**, *2* (100), re9.

(5) Ubersax, J. A.; Ferrell, J. E., Jr. Mechanisms of specificity in protein phosphorylation. *Nat. Rev. Mol. Cell Biol.* **2007**, *8* (7), 530−41.

(6) Lim, Y. P. Mining the tumor phosphoproteome for cancer markers. *Clin. Cancer Res.* **2005**, *11* (9), 3163−9.

(7) Hojlund, K.; Beck-Nielsen, H. Impaired glycogen synthase activity and mitochondrial dysfunction in skeletal muscle: markers or mediators of insulin resistance in type 2 diabetes? *Curr. Diabetes Rev.* **2006**, *2* (4), 375−95.

(8) Gaestel, M.; Kotlyarov, A.; Kracht, M. Targeting innate immunity protein kinase signalling in inflammation. *Nat. Rev. Drug Discovery* **2009**, *8* (6), 480−499.

(9) Mavropoulos, A.; Orfanidou, T.; Liaskos, C.; Smyk, D. S.; Billinis, C.; Blank, M.; Rigopoulou, E. I.; Bogdanos, D. P. p38 mitogen-activated protein kinase (p38 MAPK)-mediated autoimmunity: Lessons to learn from ANCA vasculitis and pemphigus vulgaris. *Autoimmun. Rev.* **2012**, *12* (5), 580−90.

(10) Reinders, J.; Sickmann, A. State-of-the-art in phosphoproteomics. *Proteomics* **2005**, *5* (16), 4052−61.

(11) Kalume, D. E.; Molina, H.; Pandey, A. Tackling the phosphoproteome: tools and strategies. *Curr. Opin. Chem. Biol.* **2003**, *7* (1), 64−9.

(12) Zhang, H.; Zha, X.; Tan, Y.; Hornbeck, P. V.; Mastrangelo, A. J.; Alessi, D. R.; Polakiewicz, R. D.; Comb, M. J. Phosphoprotein analysis using antibodies broadly reactive against phosphorylated motifs. *J. Biol. Chem.* **2002**, *277* (42), 39379−87.

(13) Hornbeck, P. V.; Kornhauser, J. M.; Tkachev, S.; Zhang, B.; Skrzypek, E.; Murray, B.; Latham, V.; Sullivan, M. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* **2012**, *40* (Database issue), D261−70.

(14) D'Ambrosio, C.; Salzano, A. M.; Arena, S.; Renzone, G.; Scaloni, A. Analytical methodologies for the detection and structural characterization of phosphorylated proteins. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2007**, *849* (1−2), 163−80.

(15) Dunn, J. D.; Reid, G. E.; Bruening, M. L. Techniques for phosphopeptide enrichment prior to analysis by mass spectrometry. *Mass Spectrom. Rev.* **2010**, *29* (1), 29−54.

(16) Morandell, S.; Stasyk, T.; Grosstessner-Hain, K.; Roitinger, E.; Mechtler, K.; Bonn, G. K.; Huber, L. A. Phosphoproteomics strategies for the functional analysis of signal transduction. *Proteomics* **2006**, *6* (14), 4047−56.

(17) Schmelzle, K.; White, F. M. Phosphoproteomic approaches to elucidate cellular signaling networks. *Curr. Opin. Biotechnol.* **2006**, *17* (4), 406−14.
(18) Andersson, L.; Porath, J. Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. *Anal. Biochem.* **1986**, *154* (1), 250−4.
(19) Nuhse, T.; Yu, K.; Salomon, A. Isolation of phosphopeptides by immobilized metal ion affinity chromatography. *Curr Protoc. Mol. Biol.* **2007**, Chapter 18, Unit 18 13.
(20) Thingholm, T. E.; Jensen, O. N. Enrichment and characterization of phosphopeptides by immobilized metal affinity chromatography (IMAC) and mass spectrometry. *Methods Mol. Biol.* **2009**, *527*, 47−56.
(21) Kokubu, M.; Ishihama, Y.; Sato, T.; Nagasu, T.; Oda, Y. Specificity of immobilized metal affinity-based IMAC/C18 tip enrichment of phosphopeptides for protein phosphorylation analysis. *Anal. Chem.* **2005**, *77* (16), 5144−54.
(22) Ficarro, S. B.; McCleland, M. L.; Stukenberg, P. T.; Burke, D. J.; Ross, M. M.; Shabanowitz, J.; Hunt, D. F.; White, F. M. Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nat. Biotechnol.* **2002**, *20* (3), 301−5.
(23) Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell. Proteomics* **2005**, *4* (7), 873−86.
(24) Sugiyama, N.; Masuda, T.; Shinoda, K.; Nakamura, A.; Tomita, M.; Ishihama, Y. Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications. *Mol. Cell. Proteomics* **2007**, *6* (6), 1103−9.
(25) Jensen, S. S.; Larsen, M. R. Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques. *Rapid Commun. Mass Spectrom.* **2007**, *21* (22), 3635−45.
(26) Thingholm, T. E.; Jensen, O. N.; Robinson, P. J.; Larsen, M. R. SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. *Mol. Cell. Proteomics* **2008**, *7* (4), 661−71.
(27) Proc, J. L.; Kuzyk, M. A.; Hardie, D. B.; Yang, J.; Smith, D. S.; Jackson, A. M.; Parker, C. E.; Borchers, C. H. A quantitative study of the effects of chaotropic agents, surfactants, and solvents on the digestion efficiency of human plasma proteins by trypsin. *J. Proteome Res.* **2010**, *9* (10), 5422−37.
(28) Zhou, J.; Zhou, T.; Cao, R.; Liu, Z.; Shen, J.; Chen, P.; Wang, X.; Liang, S. Evaluation of the application of sodium deoxycholate to proteomic analysis of rat hippocampal plasma membrane. *J. Proteome Res.* **2006**, *5* (10), 2547−53.
(29) Trost, M.; English, L.; Lemieux, S.; Courcelles, M.; Desjardins, M.; Thibault, P. The phagosomal proteome in interferon-gamma-activated macrophages. *Immunity* **2009**, *30* (1), 143−54.
(30) Thingholm, T. E.; Jorgensen, T. J.; Jensen, O. N.; Larsen, M. R. Highly selective enrichment of phosphorylated peptides using titanium dioxide. *Nat. Protoc.* **2006**, *1* (4), 1929−35.
(31) Rappsilber, J.; Mann, M.; Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2007**, *2* (8), 1896−906.
(32) Dephoure, N.; Gygi, S. P. A solid phase extraction-based platform for rapid phosphoproteomic analysis. *Methods* **2011**, *54* (4), 379−86.
(33) Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26* (12), 1367−72.
(34) Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **2011**, *10* (4), 1794−805.
(35) Huang da, W.; Sherman, B. T.; Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2009**, *4* (1), 44−57.
(36) Huang da, W.; Sherman, B. T.; Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **2009**, *37* (1), 1−13.
(37) Chou, M. F.; Schwartz, D. Biological sequence motif discovery using motif-x. *Curr. Protoc. Bioinformatics* **2011**, Chapter 13, Unit 13, 15−24.
(38) Schwartz, D.; Gygi, S. P. An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat. Biotechnol.* **2005**, *23* (11), 1391−8.
(39) Boulais, J.; Trost, M.; Landry, C. R.; Dieckmann, R.; Levy, E. D.; Soldati, T.; Michnick, S. W.; Thibault, P.; Desjardins, M. Molecular characterization of the evolution of phagosomes. *Mol. Syst. Biol.* **2010**, *6*, 423.
(40) Marcantonio, M.; Trost, M.; Courcelles, M.; Desjardins, M.; Thibault, P. Combined enzymatic and data mining approaches for comprehensive phosphoproteome analyses: application to cell signaling events of interferon-gamma-stimulated macrophages. *Mol. Cell. Proteomics* **2008**, *7* (4), 645−60.
(41) Wickner, R. B.; Edskes, H. K.; Shewmaker, F.; Nakayashiki, T. Prions of fungi: inherited structures and biological roles. *Nat. Rev. Microbiol.* **2007**, *5* (8), 611−8.
(42) Orr, H. T.; Zoghbi, H. Y. Trinucleotide repeat disorders. *Annu. Rev. Neurosci.* **2007**, *30*, 575−621.
(43) Han, T. W.; Kato, M.; Xie, S.; Wu, L. C.; Mirzaei, H.; Pei, J.; Chen, M.; Xie, Y.; Allen, J.; Xiao, G.; Mcknight, S. L. Cell-free Formation of RNA Granules: Bound RNAs Identify Features and Components of Cellular Assemblies. *Cell* **2012**, *149* (4), 768−79.
(44) Kato, M.; Han, T. W.; Xie, S.; Shi, K.; Du, X.; Wu, L. C.; Mirzaei, H.; Goldsmith, E. J.; Longgood, J.; Pei, J.; Grishin, N. V.; Frantz, D. E.; Schneider, J. W.; Chen, S.; Li, L.; Sawaya, M. R.; Eisenberg, D.; Tycko, R.; Mcknight, S. L. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **2012**, *149* (4), 753−67.
(45) Weber, S. C.; Brangwynne, C. P. Getting RNA and protein in phase. *Cell* **2012**, *149* (6), 1188−91.