Published in final edited form as:

Anal Chem. 2011 October 15; 83(20): 7635–7644. doi:10.1021/ac201894j.

Rapid and reproducible single-stage phosphopeptide enrichment of complex peptide mixtures: Application to general and phosphotyrosine-specific phosphoproteomics experiments

Arminja N. Kettenbach and Scott A. Gerber*

Department of Genetics, Dartmouth Medical School, Lebanon, NH 03756, USANorris Cotton Cancer Center, Lebanon, NH 03756, USA

Abstract

Reversible protein phosphorylation is an essential regulatory component of virtually every cellular process, and is frequently dysregulated in cancer. However, significant analytical barriers persist that hamper the routine application of phosphoproteomics in translational settings. Here, we present a straightforward and reproducible approach for the broadscale analysis of protein phosphorylation that relies on a single phosphopeptide enrichment step using titanium dioxide microspheres from whole cell lysate digests and compared it to the well-established SCX-TiO₂ workflow for phosphopeptide purification on a proteome-wide scale. We demonstrate the scalability of our approach from 200 micrograms to 5 milligrams of total NCI-H23 non-small cell lung adenocarcinoma cell lysate digest and determine its quantitative reproducibility by label-free analysis of phosphopeptide peak areas from replicate purifications (median CV: 20% RSD). Finally, we combine this approach with immunoaffinity phosphotyrosine enrichment, enabling the identification of 3168 unique non-redundant phosphotyrosine peptides in two LC-MS/MS runs from 8mg of HeLa peptides, each with 80% phosphotyrosine selectivity, at a peptide FDR of 0.2%. Taken together, we establish and validate a robust approach for proteome-wide phosphorylation analysis in a variety of scenarios that is easy to implement in biomedical research and translational settings.

Introduction

Protein phosphorylation is a ubiquitous post-translational modification implicated in nearly all cellular signal transduction processes, including cell proliferation and differentiation, cell cycle progression, and apoptosis. In the context of human health, many oncogenes regulate the activity and expression of protein kinases, or are kinases themselves¹; thus, knowledge regarding dysregulated kinase pathways in human cancers has the potential to reveal mechanistic underpinnings of cellular transformation and/or novel entry points for therapeutic intervention for cancer care^{2–4}.

Recent advances in phosphoproteomics techniques and mass spectrometry instrumentation have brought the global analysis of cellular phosphorylation closer to our reach⁵. Given that protein phosphorylation is often substoichiometric and many phosphorylated signaling molecules are expressed at low abundance, phosphorylation analysis remains a challenging and difficult task. Phosphopeptides also suffer from low ionization efficiency⁶, and their detection is further impaired by signal suppression by generally more abundant unphosphorylated peptides⁷. Therefore, selective enrichment of phosphopeptides away from the large pool of unphosphorylated peptides is essential for their efficient detection and

*To whom correspondence should be addressed: scott.a.gerber@dartmouth.edu.

identification by MS/MS sequencing. Techniques such as phosphopeptide immunoprecipitations^{8–10}, β -elimination and Michael addition chemistries¹¹, strong-cation/anion exchange chromatography (SCX/SAX)¹², hydrophilic interaction chromatography¹³, immobilized metal ion affinity chromatography¹⁴ (IMAC), and titanium dioxide microspheres^{15, 16} (TiO₂), and their combinations¹⁷ have been described for the enrichment of phosphorylated peptides from complex mixtures. Phosphopeptide immunoprecipitations are limited to peptide sequences corresponding to the epitope recognized by the antibody and have mainly been used for the purification of phosphotyrosine peptides⁹ and phosphopeptides corresponding to a specific kinase motif of interest^{8, 18}. Chemically modifying phosphorylated residues through β -elimination and Michael addition has been shown to generate unspecific side products¹⁹. SCX separates peptides depending on their solution-phase charge state at acidic pH, enriching phosphopeptides in early eluting fractions¹². However, later SCX fractions also contain phosphopeptides rich in basic residues, together with many unphosphorylated peptides. To overcome this, SCX is now often followed by IMAC^{20, 21} or TiO₂²² enrichment. IMAC enrichment is based on the affinity of phosphopeptides for metal ions (Fe³⁺ and Ga³⁺), although peptides containing multiple acidic residues reduce selectivity unless esterified²³. Improvements in phosphopeptide selectivity using TiO₂ microspheres have been described in reports using 2,5-dihydroxybenzoic acid (DHB)¹⁵, glutamate²⁴ or α -hydroxy aliphatic acids²⁵. In most of these cases, however, method development is performed on a small number of peptides, usually from protein standards such as α casein, instead of complex peptide mixtures that more accurately reflect the complexity of large-scale phosphoproteomics experiments in which a background of highly abundant unphosphorylated peptides complicates these analyses. Most TiO₂ enrichments are conducted off-line from LC-MS/MS, although recently, online setups for TiO₂ enrichment have been described²⁶ that reduce sample handling and increase reproducibility, at the cost of selectivity and throughput.

Here we describe an offline TiO₂-based phosphopeptide enrichment approach for complex mixtures using lactic acid as a co-solvent that provides high phosphopeptide selectivity, is widely scalable, and affords excellent qualitative and quantitative reproducibility. Using complex peptide mixtures representative of real-world phosphoproteomics experiments, we carefully titrated the amount of lactic acid, TFA in the presence of lactic acid, TiO₂ microspheres, and time for optimal phosphopeptide selectivity and maximum phosphopeptide number. We demonstrate that phosphopeptide selectivity is largely dependent on the relative basicity of the respective peptide and independent of the complexity of the peptide mixture. Based on this finding, we propose, test, and validate the hypothesis that a single stage of phosphopeptide purification from whole cell lysate digests recapitulates the results from a workflow in which peptides are separated post-digest and prior to phosphopeptide enrichment. We also investigated the scalability and reproducibility of these single-stage purifications, and report our findings. Finally, we combine this single-stage phosphopeptide enrichment with immunoaffinity purification of phosphotyrosine peptides to identify in two LC-MS/MS runs 3168 unique non-redundant phosphotyrosine peptides.

Experimental Procedures

See Supporting Information for details on experimental procedures.

Results and Discussion

Assessment of parameters for phosphopeptide purification using titanium dioxide (TiO₂)

TiO₂ is now commonly used to enrich phosphopeptides from complex peptide mixtures. To increase the selectivity of phosphopeptide recovery from TiO₂ microspheres, peptides are

resuspended in ACN in the presence of acidic modifiers such as TFA²⁷, DHB¹⁵ and lactic acid²⁵, and often when DHB²⁸ and lactic acid²⁵ are diluted in the presence of TFA, that are intended to prevent or displace non phosphate-mediated interactions with the microsphere surface. Most commonly, these additives were assessed using phosphopeptide standards or simple single-protein digests. To mimic conditions typically found in proteomics experiments, we developed two samples of high peptide complexity. Twenty milligrams of HeLa cell digest peptides were separated by SCX into 24 fractions up to a solution charge state of +4. Three early fractions (6, 7, and 8; solution charge state of +1; “EF”) were combined, three later fractions (18, 19, and 20; solution charge states of +2 and +3; “LF”) were combined, and each pool was divided into 100µg peptide aliquots (Figure 1A). In the “EF” sample, phosphopeptides were already moderately enriched by virtue of tryptic peptide charge reduction in SCX when phosphorylated¹², while in the “LF” sample, phosphopeptides were more diluted in non-phosphopeptides. We note that the median number of basic residues (arginines, lysines, histidines, and peptide N-termini at pH 2.65) in the “EF” sample was two, while in the “LF” sample the median number of basic residues was 3 (Supporting Figure 1A); the median number of phosphorylation sites per peptide in both samples was one. The number of acidic amino acids (which are proton-bound at pH 2.65) scaled with peptide length (Supporting Figure 1B and 1C). These samples were subsequently used in TiO₂ purifications under variable conditions followed by LC-Orbitrap-MS/MS, peptide spectral matching, and filtering to achieve a false discovery rate (FDR) < 1% at the peptide level.

We also sought to assess quantitative differences in peptide recovery under these variable conditions through the use of “label-free” peak area integration²⁹. We found that the collection of peptides (unphosphorylated peptides in particular) changed across the range of conditions tested, and therefore determined the average fold-change difference of peptides identified between consecutive conditions. Because our “EF” sample was already enriched in phosphopeptides via SCX, phosphopeptide selectivity was in most cases 90% or higher, which prevented us from finding a sufficient number of unphosphorylated peptides in this sample to produce meaningful quantification results. We were, however, able to quantify fold-changes for phosphopeptides in both sample types. All conditions were conducted as technical replicates.

We first determined the optimal concentration of TFA in the presence of a fixed amount (2M) of lactic acid based on this practice in previous reports. Peptides were resuspended in 50% ACN, 2M lactic acid with concentrations of 0%, 0.04%, 0.2%, 1%, and 2% TFA, and vortexed with 375µg TiO₂ microspheres for 45 minutes. We found a modest but significant decrease in total phosphopeptide identifications with increasing amounts of TFA additive in the presence of 2M lactic acid (Figure 1B). The reduction in total phosphopeptide identifications between 0% and 2% TFA was similar in relative extent for the two sample types (33% and 32% reduction for “EF” and “LF” samples in 2% TFA, respectively). Phosphopeptide selectivity for both samples was very high (over 96%) and essentially unaffected by TFA. The relative standard deviation varied from 0.1% to 1.9% for phosphopeptide number and 0.01% to 0.2% for phosphopeptide selectivity for the “EF” and 0.1% to 2.9% for phosphopeptide number and 0.2% to 0.8% for phosphopeptide selectivity for the “LF” sample analyzed in duplicate. We also observed that the average peptide quantification difference between each TFA condition was consistent with the difference in peptide identification rates between those conditions (Figure S-2A). For example, the largest difference in peptide identification rate for phosphopeptides in the “EF” sample was between 0.2% and 1% TFA – this also represented the largest average difference in peptide peak areas observed between these two conditions. Similarly, for the more basic “LF” peptides, the largest drop in peptide identifications was between using no TFA and using only 0.04% TFA, which was also the largest transition for their relative peak area

differences. We note that TFA exhibited a very minimal differential effect on the recovery of phosphopeptides relative to unphosphorylated peptides in the “LF” sample. We surmise that additional TFA added to the binding solution either reduces the pH of the binding solution to a level that drives phosphate to a more proton-bound (phosphoric acid) state, or directly competes with phosphate for binding to the microsphere surface, or both. We therefore concluded that any additional TFA was detrimental from lactic acid-mediated phosphopeptide purifications, and so excluded it from subsequent experiments.

We determined the optimal concentration of lactic acid in a similar manner: peptides were resuspended in 50% ACN with 0M, 0.2M, 0.5M, 1M, 2M, or 3M lactic acid, and vortexed with 375 μ g TiO₂ microspheres for 45 minutes. The relative standard deviation varied from 1.5% to 7.1% for phosphopeptide number and 0.2% to 1.2% for phosphopeptide selectivity for the “EF” and 0.6% to 3.3% for phosphopeptide number and 0.8% to 2.3% for phosphopeptide selectivity for the “LF” samples analyzed in duplicate. We found that the number of phosphopeptides increased with increasing lactic acid concentration for both samples, and was only slightly reduced at 3M for the “EF” sample (Figure 1C). The selectivity of phosphopeptide purification reached a plateau of 99% at 1M lactic acid for the “EF” sample, and 92% at 2M lactic acid for the “LF” sample. Interestingly, we observed a very strong effect of lactic acid on the quantitative recovery of peptides from these samples (Figure S-2B). While a small amount of lactic acid resulted in a modest increase in phosphopeptide recovery, increasing the amount of this additive beyond 0.2M resulted in slight decreases in the quantitative phosphopeptide recovery with each additional increment for both sample types. However, the effect of lactic acid on unphosphorylated peptides was much more dramatic. In the transition from no lactic acid to 0.2M lactic acid, the quantitative recovery of unphosphorylated peptides from the “LF” condition was reduced by 16-fold. We did not observe a concentration of lactic acid at which the quantitative rate of phosphopeptide recovery loss was greater than that of unphosphorylated cognates. We consider a lactic acid concentration of 2M as the best balance between the number of unique phosphopeptide identifications, selectivity, and quantitative recovery.

Next, we investigated the amount of TiO₂ microspheres necessary for efficient phosphopeptide enrichment from complex mixtures by varying the amount of TiO₂ in a purification in the presence of 2M lactic acid. The relative standard deviation varied from 1.5% – 8.2% for phosphopeptide number and 0.1% to 0.2% for phosphopeptide selectivity for the “EF” and 2.5% to 14.3% for phosphopeptide number and 0.1% to 3.5% for phosphopeptide selectivity for the “LF” sample analyzed in duplicate. The number of identified phosphopeptides reached a plateau at 250 μ g TiO₂ for the “EF” sample, and increased only slightly more for the “LF” sample with the addition of 750 μ g TiO₂ (Figure 1D). The phosphopeptide selectivity did not change significantly for the “EF” sample, and was slightly reduced (14% lower than the maximum) for the “LF” sample when 750 μ g TiO₂ was employed, which is in contrast to a previous report³⁰ that indicates that selectivity increases with increasing amount of resin up to an optimum ratio, after which it decreases again. We note here that this particular study employed a very high amount of TFA (2%) in saturated glutamic acid, something our titration of TFA with lactic acid suggests would result in poor phosphopeptide recovery in the first place. It is also possible that the number of confident phosphopeptide identifications leveled off due to undersampling on our LTQ-Orbitrap and 50-minute LC gradient. This hypothesis is supported by the corresponding quantitative data (Figure S-2C) that demonstrates a generally positive trend in the average phosphopeptide recovery with increasing amounts of resin. Here, phosphopeptides and unphosphorylated peptides appeared to be equally responsive to additional TiO₂ microspheres present during purification. The decrease in selectivity in the “LF” sample with 750 μ g TiO₂ microspheres suggests that most phosphopeptides are bound and that increasing amounts of unphosphorylated peptides were binding non-specifically due to the

increase in surface area – this is plausible in the context of the quantification data if one assumes that the increase in average quantitative recovery is restricted to only certain peptides (to those that have many acidic residues, for example).

Finally, we fixed the lactic acid concentration (2M) and TiO_2 amount (400 μg) at their optimum in order to establish how binding time influenced overall phosphopeptide identification rates and selectivity. The relative standard deviation of the replicate analysis varied from 0.5% – 1.3% for phosphopeptide number and 0.01% to 0.2% for phosphopeptide selectivity for the “EF” and 0.1% to 3.2% for phosphopeptide number and 0.07% to 0.6% for phosphopeptide selectivity for the “LF” sample. Already after 20min incubation time we observed maximal binding for the “EF” sample. For the “LF” sample, although the majority of phosphopeptides were bound after 20min, the yield continued to increase steadily with increasing amount of binding time (Figure 1E), suggesting that the presence of either more basic residues in these phosphopeptides, or the larger excess of unphosphorylated peptides in later SCX fractions, contributes substantially to the rate of equilibration of phosphate with resin. Phosphopeptide selectivity was comparable in all “EF” and “LF” samples, regardless of incubation time, as was the quantitative recovery of both phosphorylated and unphosphorylated peptides (Figure S-2D). Taken together, we found that 2M lactic acid in the absence of TFA with 400 μg of TiO_2 microspheres per 100 μg of peptide and an incubation time of 1 hour lead to improved recovery and identification rates for these peptide samples of high complexity.

Influence of sample complexity on phosphopeptide enrichment selectivity

To determine if the selectivity of phosphopeptide enrichment is influenced by the presence of other peptides with different chemical attributes during purification, or depends only on the relative basicity of the purified peptide itself, we combined equal amounts of the “EF” and “LF” samples and enriched for phosphopeptides from 100 μg of the combined sample in the presence of variable concentrations of lactic acid. Interestingly, we found that the overall selectivity of the combined sample was intermediate to the “EF” and “LF” only sample at every lactic acid concentration (Figure 1F). Next, we determined if the peptides identified in the combined sample were previously found in the “EF” or in the “LF” sample, and re-plotted phosphopeptide selectivity according to their origin (Figure 1G). Surprisingly, we found that in the combined sample, phosphopeptides and non-phosphopeptides from early or late SCX fractions were enriched to an almost identical degree as in the individual “EF” or “LF” samples. This suggests that phosphopeptide enrichment selectivity using TiO_2 microspheres in the presence of lactic acid depends primarily on the number of basic residues in the specific peptide being enriched, and not on other peptides present in the sample.

Single-stage workflow for large-scale phosphoproteomics

The standard phosphopeptide purification workflow for complex samples such as tissue culture cells or primary tissue consists of sample lysis, digestion into peptides, followed by separation of the peptides on a 9.4 mm inner diameter SCX column into multiple fractions²¹. Phosphopeptides are then purified from each of the individual fractions and analyzed by LC-MS/MS (Figure 2A). While this process is very productive, it is also very labor-intensive and subject to variability due to the high degree of sample handling and manipulation. Based on our results, we hypothesized that phosphopeptides could be efficiently purified from whole cell lysate directly in the presence of lactic acid prior to fractionation by SCX chromatography. This would allow for the final separation to be conducted using lower flow rate columns (2.1 mm inner diameter or smaller) due to a large reduction in peptide loading, followed by analysis via LC-MS/MS (Figure 2B).

To test this hypothesis, we prepared peptides from digests of NCI-H23 non-small cell lung adenocarcinoma (NSCLC) cells and subjected them to both workflows. We separated 5mg in a single injection on a 9.4 mm SCX column at 2.5 ml/min into 24 fractions, purified phosphopeptides from each fraction, and analyzed them by LC-MS/MS (Figure 2C). We then performed a single TiO₂ phosphopeptide purification directly on 5mg of whole cell digest peptides using lactic acid, scaled to an identical degree as those previously performed on individual fractions. The resulting phosphopeptides were separated on a 2.1 mm SCX column at 0.2 ml/min into 24 fractions, desalted and analyzed by LC-MS/MS (Figure 2D). Similar to the conventional approach, phosphopeptide selectivity was high in the early fractions and lower in the later fractions. Overall, we identified 19,647 phosphopeptides (24,005 total peptides) using the 24-stage TiO₂ approach, and 14,532 phosphopeptides (20,141 total peptides) using the alternative single-stage TiO₂ approach (Figure 2E). Interestingly, the number of unique phosphorylation sites identified by both approaches was comparable with 11,080 unique phosphorylation sites in the 24-TiO₂ and 9,651 unique phosphorylation sites in the single-stage purification workflow (87.5%). The overall phosphopeptide selectivity was moderately lower in the single- versus 24-stage TiO₂ approach (72.2% versus 81.8%) (Figure 2E).

A qualitative comparison of the phosphorylation sites identified by both workflows showed that 6,041 were in common, a level comparable to the overlap of replicate injections in undersampled shotgun sequencing experiments in general. In addition, comparison of the protein coverage showed that 80% of the total number of proteins were identified in both workflows (Figure 2F).

Taken together, these findings demonstrate that an alternative workflow with a single stage of phosphopeptide enrichment prior to SCX using titanium dioxide microspheres under optimized conditions can produce results comparable to the classic multi-stage approach, with reduced sample handling and manual labor.

Scalability of single-stage phosphopeptide purifications

Samples of significant biological or biomedical interest are often only available in limited amounts (tumor tissue, primary and stem cells). We challenged our single-stage TiO₂ strategy by isolating phosphopeptides in the presence of 2M lactic acid from 5mg, 1mg, and 0.2mg of NCI-H23 whole cell digest peptides. For feature comparison, the base peak chromatogram of fraction 8 of the SCX separation for all three samples (Figure 3A). As expected, the peak intensity and number of discernable features increased with increasing amount of input sample, while all three chromatograms contain largely the same set of core features. To quantify this observation, we compared the peak area of peptides identified in all three samples and found a high correlation of peptide peak area with increased sample amount (median R² 0.96, median slope 0.98) (Figure 3B and C). The number of identified peptides increased from the 0.2mg (6,198 phosphopeptides/9,585 total peptides), the 1mg sample (10,266 phosphopeptides/14,395 total peptides), to the 5mg sample (14,532 phosphopeptides/20,141 total peptides) (Figure 3D). Phosphopeptide selectivity was 65% in the 0.2mg sample, 71% in the 1mg, and 72% in the 5mg samples. The number of unique phosphorylation sites increased from 4,397 in the 0.2mg, 6,945 in the 1mg, to 9,651 in the 5mg sample (Figure 4E). The overlap of unique phosphorylation sites in the three samples was 2,454, which is 56% of the total number of unique phosphorylation sites identified in the 0.2mg sample (Figure 4F). We concluded from these experiments that the single-stage TiO₂ approach scales quantitatively and can be utilized in the identification of phosphorylation sites from samples of various amounts of input.

Reproducibility of single-stage phosphopeptide purifications

In conventional phosphopeptide purification approaches, TiO₂ incubations are performed individually, which generates the potential for increased variability in the amounts of resulting phosphopeptides. An important advantage of the single-stage TiO₂ approach is that only one purification is performed on each desalted whole cell lysate digest. To test the reproducibility of whole cell lysate digest TiO₂ treatments, we divided 4mg of Jurkat cell lysate digest into separate 1mg aliquots, purified phosphopeptides from each of them and analyzed a portion of the purified peptides directly by LC-MS/MS. We also analyzed four repeat injections of one of the samples to assess intermediate precision. Visual inspection of base peak chromatograms of replicate injections displayed features that were qualitatively similar (Figure 4A). We identified 6004 unique phosphopeptides in the union of these technical replicates (on average 1501 unique phosphopeptides per replicate), of which 1,145 occurred in only one of the injections, 515 in two, 395 in three, and 661 in all four replicate injections, leading to an overlap across all four injections of 44% (Figure 4B). Peptides identified in all four replicates exhibited a median relative standard deviation (RSD) of 18% (range: 1.9% – 171%) in peak area (Figure 4C).

Next, we examined the four individual TiO₂ samples, which also generated similar base peak chromatograms (Figure 4D). We sequenced a total of 6,055 unique phosphopeptides across the four separate samples (on average 1514 unique phosphopeptides per sample), of which 1,141 occurred in only one, 510 in two, 390 in three, and 681 in all four, reaching an overlap of 45% across the four samples (Figure 4E). Comparison of the peak area of the peptides common to all four runs resulted in a median RSD of 20% (range: 1.2% – 155%; Figure 4F). Taken together, these analyses demonstrate that individual TiO₂ purifications are nominally as reproducible as replicate injections on our LC-MS/MS platform, and can be readily adapted for label-free quantification.

Application of single stage TiO₂ enrichment to phosphotyrosine phosphoproteomic experiments

Tyrosine phosphorylation is essential for many cellular processes, and its dysregulation is implicated in human disease and cancer³¹. However, comprehensive mapping of tyrosine phosphorylation remains challenging due to the infrequent occurrence and substoichiometric levels of the modification. In large-scale phosphoproteomics experiments, including those described above with H23 cell lysate digests, phosphotyrosine peptides are only 1% of total number of identified phosphopeptides. To address this problem, enrichment of phosphotyrosine peptides is often achieved by immunoaffinity purification of proteins or peptides using antibodies against phosphorylated tyrosine residues^{9, 10, 32, 33}. While phosphotyrosine peptide immunoprecipitations (IP) have been successfully employed in large-scale proteomics studies, they require careful optimization of IP conditions³⁴, or a second step of phosphopeptide enrichment to remove non-specifically bound unphosphorylated peptides^{35, 36}. This problem worsens with increasing amount of starting material and antibody beads, making large-scale purifications challenging. We found that by first performing a single-stage TiO₂ enrichment on whole cell lysate digests to purify total cellular phosphopeptides away from the background of unphosphorylated peptides, and then conducting the phosphotyrosine IPs (Figure 6A), we are able to achieve reproducible results with a low background of non-phosphorylated peptides even from larger amounts of starting material (Figure 6B). Using this approach we were able to identify 3168 unique non-redundant phosphotyrosine peptides (0.2% FDR) (2110 and 2504 unique non-redundant phosphotyrosine peptide in each replicate) in replicate 120 minute gradient LC-MS/MS analyses from 8mg of stimulated HeLa peptides each with 80% selectivity of phosphotyrosine peptides over other phospho- and non-phosphopeptides. Using the combined datasets, we performed motif-analysis using Motif-X and found 14 statistically

significant motifs³⁷ (Figure 5C, Figure S3). As previously described³¹, most commonly the motifs consisted of an acidic amino acid upstream or downstream of the phosphorylated tyrosine. Due to the large size of our dataset, we not only found a preference for serine in the +1 position or proline in the +3 position³⁴, but were able to identify more complex motifs containing an acidic amino acid as well as a serine or proline (EXXpYXXP and pYDXP, or EXXpYS, and DXXpYS). We also found alanine as well as glutamine in the +1 position.

Finally, we performed pathway analysis on the phosphotyrosine dataset using commercial software and found insulin receptor signaling to be one of the most significantly enriched pathways (Figure 5D). Tyrosine phosphorylation is known to be essential for insulin signal transduction from the insulin receptor in the plasma membrane to its downstream targets and has been the topic of previous phosphotyrosine proteomic studies^{38–40}. We envision that this coupled approach will add considerably greater depth to such explorations of insulin signaling in future experiments⁴¹.

Conclusions

The broadscale analysis of protein phosphorylation remains a daunting task. In addition to problems associated with the biological context of phosphorylation (stoichiometry, dynamic range, lability, etc.), many analytical platforms for general protein shotgun sequencing are poorly suited to phosphoproteomics. Given the potential wealth of value that translational and clinical phosphoproteomics has to offer biomedicine, we sought to augment the traditional analytical paradigm in phosphoproteomics to improve the ruggedness and reproducibility of these platforms. Here, we have presented data demonstrating that under the proper conditions, titanium oxide-based phosphopeptide enrichment can be conducted with high yield, selectivity and with quantitative precision, directly from extremely complex mixtures such as human cancer cell lysates. We extend the general utility of our approach by demonstrating its capability to contribute to phosphotyrosine phosphoproteomics experiments by coupling single-stage enrichments to phosphotyrosine immunoprecipitations. We also note that the high quantitative reproducibility of single-stage purifications suggests it should be possible to perform comparative experiments between different samples by label-free methods. Alternatively, it may be possible to label these purified phosphopeptide pools with reagent-based labeling chemistries such as iTRAQ or TMT reagents. Direct iTRAQ labeling of cell lysate digests for global phosphoproteomics experiments is cost-prohibitive, owing to the large amount of peptide digest required for a productive experiment (~5mg or more). However, we estimate that the peptide yield from a single-stage purification is roughly 0.5% of the total input, which would place a 5mg input sample on par with a typical labeling reaction for iTRAQ. Further experiments are planned to validate these hypotheses.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We acknowledge funding from the American Cancer Society (IRG-82-003-24) and the National Institutes of Health (P20-RR018787) for the IDeA Program of the National Center for Research Resources (to S.A.G.).

References

1. Davies H, Hunter C, Smith R, Stephens P, Greenman C, Bignell G, Teague J, Butler A, Edkins S, Stevens C, Parker A, O'Meara S, Avis T, Barthorpe S, Brackenbury L, Buck G, Clements J, Cole J, Dicks E, Edwards K, Forbes S, Gorton M, Gray K, Halliday K, Harrison R, Hills K, Hinton J, Jones

- D, Kosmidou V, Laman R, Lugg R, Menzies A, Perry J, Petty R, Raine K, Shepherd R, Small A, Solomon H, Stephens Y, Tofts C, Varian J, Webb A, West S, Widaa S, Yates A, Brasseur F, Cooper CS, Flanagan AM, Green A, Knowles M, Leung SY, Looijenga LH, Malkowicz B, Pierotti MA, Teh BT, Yuen ST, Lakhani SR, Easton DF, Weber BL, Goldstraw P, Nicholson AG, Wooster R, Stratton MR, Futreal PA. *Cancer Res.* 2005; 65:7591–7595. [PubMed: 16140923]
2. Fedorov O, Muller S, Knapp S. *Nat Chem Biol.* 2010; 6:166–169. [PubMed: 20154661]
3. Kettenbach AN, Schweppe DK, Faherty BK, Pechenick D, Pletnev AA, Gerber SA. *Sci Signal.* 2011; 4:rs5. [PubMed: 21712546]
4. Janne PA, Gray N, Settleman J. *Nat Rev Drug Discov.* 2009; 8:709–723. [PubMed: 19629074]
5. Nita-Lazar A, Saito-Benz H, White FM. *Proteomics.* 2008; 8:4433–4443. [PubMed: 18846511]
6. Janek K, Wenschuh H, Bienert M, Krause E. *Rapid Commun Mass Spectrom.* 2001; 15:1593–1599. [PubMed: 11544598]
7. Kratzer R, Eckerskorn C, Karas M, Lottspeich F. *Electrophoresis.* 1998; 19:1910–1919. [PubMed: 9740051]
8. Moritz A, Li Y, Guo A, Villen J, Wang Y, MacNeill J, Kornhauser J, Sprott K, Zhou J, Possemato A, Ren JM, Hornbeck P, Cantley LC, Gygi SP, Rush J, Comb MJ. *Sci Signal.* 2010; 3:ra64. [PubMed: 20736484]
9. Rush J, Moritz A, Lee KA, Guo A, Goss VL, Spek EJ, Zhang H, Zha XM, Polakiewicz RD, Comb MJ. *Nat Biotechnol.* 2005; 23:94–101. [PubMed: 15592455]
10. Zhang Y, Wolf-Yadlin A, Ross PL, Pappin DJ, Rush J, Lauffenburger DA, White FM. *Mol Cell Proteomics.* 2005; 4:1240–1250. [PubMed: 15951569]
11. Oda Y, Nagasu T, Chait BT. *Nat Biotechnol.* 2001; 19:379–382. [PubMed: 11283599]
12. Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, Villen J, Li J, Cohn MA, Cantley LC, Gygi SP. *Proc Natl Acad Sci U S A.* 2004; 101:12130–12135. [PubMed: 15302935]
13. McNulty DE, Annan RS. *Mol Cell Proteomics.* 2008; 7:971–980. [PubMed: 18212344]
14. Gruhler A, Olsen JV, Mohammed S, Mortensen P, Faergeman NJ, Mann M, Jensen ON. *Mol Cell Proteomics.* 2005; 4:310–327. [PubMed: 15665377]
15. Larsen MR, Thingholm TE, Jensen ON, Roepstorff P, Jorgensen TJ. *Mol Cell Proteomics.* 2005; 4:873–886. [PubMed: 15858219]
16. Pinkse MWH, Uitto PM, Hilhorst MJ, Ooms B, Heck AJR. *Analytical Chemistry.* 2004; 76:3935–3943. [PubMed: 15253627]
17. Bodenmiller B, Mueller LN, Mueller M, Domon B, Aebersold R. *Nature Methods.* 2007; 4:231–237. [PubMed: 17293869]
18. Matsuoka S, Ballif BA, Smogorzewska A, McDonald ER, Hurov KE, Luo J, Bakalarski CE, Zhao ZM, Solimini N, Lerenthal Y, Shiloh Y, Gygi SP, Elledge SJ. *Science.* 2007; 316:1160–1166. [PubMed: 17525332]
19. Li W, Backlund PS, Boykins RA, Wang G, Chen HC. *Anal Biochem.* 2003; 323:94–102. [PubMed: 14622963]
20. Villen J, Beausoleil SA, Gerber SA, Gygi SP. *Proc Natl Acad Sci U S A.* 2007; 104:1488–1493. [PubMed: 17242355]
21. Villen J, Gygi SP. *Nat Protoc.* 2008; 3:1630–1638. [PubMed: 18833199]
22. Olsen JV, Vermeulen M, Santamaria A, Kumar C, Miller ML, Jensen LJ, Gnäd F, Cox J, Jensen TS, Nigg EA, Brunak S, Mann M. *Sci Signal.* 2010; 3:ra3. [PubMed: 20068231]
23. Ficarro SB, McClelland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, Hunt DF, White FM. *Nat Biotechnol.* 2002; 20:301–305. [PubMed: 11875433]
24. Yu LR, Zhu Z, Chan KC, Issaq HJ, Dimitrov DS, Veenstra TD. *J Proteome Res.* 2007; 6:4150–4162. [PubMed: 17924679]
25. Sugiyama N, Masuda T, Shinoda K, Nakamura A, Tomita M, Ishihama Y. *Mol Cell Proteomics.* 2007; 6:1103–1109. [PubMed: 17322306]
26. Pinkse MW, Mohammed S, Gouw JW, van Breukelen B, Vos HR, Heck AJ. *J Proteome Res.* 2008; 7:687–697. [PubMed: 18034456]
27. Prsic S, Dankwa S, Schwartz D, Chou MF, Locasale JW, Kang CM, Bemis G, Church GM, Steen H, Husson RN. *Proc Natl Acad Sci U S A.* 2010; 107:7521–7526. [PubMed: 20368441]

28. Thingholm TE, Jorgensen TJ, Jensen ON, Larsen MR. *Nat Protoc.* 2006; 1:1929–1935. [PubMed: 17487178]
29. Mueller LN, Rinner O, Schmidt A, Letarte S, Bodenmiller B, Brusniak MY, Vitek O, Aebersold R, Muller M. *Proteomics.* 2007; 7:3470–3480. [PubMed: 17726677]
30. Li QR, Ning ZB, Tang JS, Nie S, Zeng R. *J Proteome Res.* 2009; 8:5375–5381. [PubMed: 19761217]
31. Hunter T. *Curr Opin Cell Biol.* 2009; 21:140–146. [PubMed: 19269802]
32. Blagoev B, Ong SE, Kratchmarova I, Mann M. *Nat Biotechnol.* 2004; 22:1139–1145. [PubMed: 15314609]
33. Thelemann A, Petti F, Griffin G, Iwata K, Hunt T, Settinari T, Fenyo D, Gibson N, Haley JD. *Mol Cell Proteomics.* 2005; 4:356–376. [PubMed: 15657067]
34. Boersema PJ, Foong LY, Ding VM, Lemeer S, van Breukelen B, Philp R, Boekhorst J, Snel B, den Hertog J, Choo AB, Heck AJ. *Mol Cell Proteomics.* 2009; 9:84–99. [PubMed: 19770167]
35. Iliuk AB, Martin VA, Alicie BM, Geahlen RL, Tao WA. *Mol Cell Proteomics.* 2010; 9:2162–2172. [PubMed: 20562096]
36. Zheng H, Hu P, Quinn DF, Wang YK. *Mol Cell Proteomics.* 2005; 4:721–730. [PubMed: 15659558]
37. Schwartz D, Gygi SP. *Nat Biotechnol.* 2005; 23:1391–1398. [PubMed: 16273072]
38. Kruger M, Kratchmarova I, Blagoev B, Tseng YH, Kahn CR, Mann M. *Proc Natl Acad Sci U S A.* 2008; 105:2451–2456. [PubMed: 18268350]
39. Hanke S, Mann M. *Mol Cell Proteomics.* 2009; 8:519–534. [PubMed: 19001411]
40. Ibarrola N, Molina H, Iwahori A, Pandey A. *J Biol Chem.* 2004; 279:15805–15813. [PubMed: 14739304]
41. Sano H, Peck GR, Kettenbach AN, Gerber SA, Lienhard GE. *J Biol Chem.* 2011; 286:16541–16545. [PubMed: 21454697]

Figure 1. Parameter sweep for phosphopeptide purification using titanium dioxide (TiO₂) and influence of sample complexity on phosphopeptide enrichment selectivity

(A.) Workflow of TiO₂ enrichment. Cell lysate digests were separated on a SCX column, 24 fractions were collected and fraction #6–8 (acidic, early fractions, “EF”) and #18–20 (basic, late fractions, “LF”) were combined, enriched as described and analyzed by LC-MS/MS.

(B.) Unique phosphopeptide identifications and selectivity purified at different TFA concentrations. (C.) Unique phosphopeptide identifications and selectivity purified at different lactic acid concentrations. (D.) Unique phosphopeptide identifications and selectivity purified with different amounts of TiO₂. (E.) Unique phosphopeptide identifications and selectivity from incubations of variable time. (F.) Phosphopeptide selectivity when isolated from early, late, and combined (early + late) fractions the presence of 2M lactic acid in 50% ACN, 375µg TiO₂ for 45 min. (G.) Phosphopeptide selectivity of peptides isolated from the combined fractions and plotted according to their origin from the early or late fractions.

Figure 2. Single-stage purification workflow for large-scale phosphoproteomics (A.) Standard proteomics workflow for broadscale phosphorylation analysis. (B.) Single-stage proteomics workflow. Phosphopeptides are purified in a single, larger TiO₂ enrichment step; the resulting phosphopeptides are separated by SCX chromatography and analyzed by LC-MS/MS. (C.) and (D.) Histograms depicting number of phosphopeptides (orange) and total peptides (blue) identified in each SCX fraction using the 24x-TiO₂ (C.) and 1x-TiO₂ (D.) approach. Phosphopeptide selectivity of each fraction is indicated (black). (E.) Total number of identified peptides, phosphopeptides, selectivity, and unique phosphorylation sites using the 24x-TiO₂ and 1x-TiO₂ approach. Overlap of unique phosphorylation sites (F.) and proteins (G.) identified using the 24x-TiO₂ and 1x-TiO₂ approaches.

Figure 3. Scalability of single-stage phosphopeptide purifications

(A.) Representative base peak chromatograms (fraction 8) of the SCX separation from 0.2mg, 1mg, and 5mg whole cell lysate single-stage purifications. (B.) Correlation coefficient of linear regression analyses of peptide peak area from 0.2mg to 5mg single-stage purifications. (C.) Slope of the least square lines of the linear regression analysis of peptide peak area from 0.2mg to 5mg single-stage purifications. (D.) Number of phosphopeptides (orange) and total peptides (blue) identified in the 0.2mg, 1mg and 5mg single-stage phosphopeptide enrichments. Phosphopeptide selectivity is indicated (black). (E.) Number of unique phosphorylation sites identified in the 0.2mg, 1mg and 5mg single-stage phosphopeptide enrichments. (F.) Overlap of unique phosphorylation sites identified in the 0.2mg, 1mg and 5mg single-stage phosphopeptide enrichments.

Figure 4. Reproducibility of single-stage phosphopeptide enrichments

(A.) Base peak chromatograms of replicate injections of one single-stage purification (sample A) from Jurkat lysate digest. (B.) Venn diagram depicting phosphopeptide occurrence in the four replicate injections of sample A. (C.) Histogram of peak area RSD of phosphopeptides from four replicate injections. (D.) Base peak chromatograms of replicate single-stage purifications (sample A, B, C, and D) from Jurkat lysate digest. (E.) Venn diagram depicting phosphopeptide occurrence in the four replicate TiO₂ enrichments. (F.) Histogram of peak area RSD of phosphopeptides from four replicate TiO₂ enrichments.

Figure 5. Phosphotyrosine enrichments from single-stage TiO_2 purifications

(A). Workflow of phosphotyrosine enrichment from TiO_2 enrichment. Phosphopeptides are purified in a single, large TiO_2 enrichment step; from the resulting phosphopeptides, phosphotyrosine (pY) peptides are then isolated by immunoprecipitation (IP). (B.) Base peak chromatograms of the biological pY IP replicates. (C.) Significant motifs identified from the union of pY IP replicates. (D.) Insulin signaling pathway with pY sites identified in the replicate LC-MS/MS analyses indicated in red.