Off-Line High-pH Reversed-Phase Fractionation for In-Depth Phosphoproteomics

Batth, Tanveer Singh; Francavilla, Chiara; Olsen, Jesper Velgaard

Published in:
Journal of Proteome Research

DOI:
10.1021/pr500893m

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Batth, T. S., Francavilla, C., & Olsen, J. V. (2014). Off-Line High-pH Reversed-Phase Fractionation for In-Depth Phosphoproteomics. DOI: 10.1021/pr500893m
Off-Line High-pH Reversed-Phase Fractionation for In-Depth Phosphoproteomics

Tanveer S. Batth, Chiara Francavilla, and Jesper V. Olsen*

Proteomics Program, Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Science, University of Copenhagen, Blegdamsvej 3B, DK-2200 Copenhagen, Denmark

ABSTRACT: Protein phosphorylation is an important post-translational modification (PTM) involved in embryonic development, adult homeostasis, and disease. Over the past decade, several advances have been made in liquid chromatography–tandem mass spectrometry (LC−MS/MS)-based technologies to identify thousands of phosphorylation sites. However, in-depth phosphoproteomics often require off-line enrichment and fractionation techniques. In this study, we provide a detailed analysis of the physicochemical characteristics of phosphopeptides, which have been fractionated by off-line high-pH chromatography (HpH) before subsequent titanium dioxide (TiO2) enrichment and LC−MS/MS analysis. Our results demonstrate that HpH is superior to standard strong-cation exchange (SCX) fractionation in the total number of phosphopeptides detected when analyzing the same number of fractions by identical LC−MS/MS gradients. From 14 HpH fractions, we routinely identified over 30 000 unique phosphopeptide variants, which is more than twice the number of that obtained from SCX fractionation. HpH chromatography displayed an exceptional ability to fractionate singly phosphorylated peptides, with minor benefits for doubly phosphorylated peptides over that with SCX. Further optimizations in the pooling and concatenation strategy increased the total number of multiphosphorylated peptides detected after HpH fractionation. In conclusion, we provide a basic framework and resource for performing in-depth phosphoproteome studies utilizing off-line basic reversed-phase fractionation. Raw data is available at ProteomeXchange (PXD001404).

KEYWORDS: Phosphoproteomics, high-pH reversed-phase, phosphorylation, peptides, enrichment, fractionation, titanium dioxide, Orbitrap

INTRODUCTION

Great advances have been made in the field of biomedical science due to the availability of large-scale proteomics data sets arising from high-performance liquid chromatography–tandem mass spectrometry (LC−MS/MS) experiments. Along with the capability to identify and quantify peptides and proteins by shotgun proteomics, it is also possible to analyze and localize post-translational modifications (PTMs) to their respective sites.1 Protein phosphorylation, which mainly modifies Ser, Thr, and Tyr residues in eukaryotes, is of particularly great interest due to its pivotal role in the regulation of several cellular responses and is therefore one of the most widely studied PTMs. Site-specific and dynamic phosphorylation is believed to orchestrate virtually all intracellular signaling cascades, thereby

Received: August 27, 2014
Published: October 22, 2014
developed to address some of these issues by reducing sample complexity. Several off- and on-line peptide fractionation methods, such as ion exchange chromatography (IEC), hydrophilic interaction chromatography (HILIC), and electrostatic repulsion–hydrophilic interaction chromatography (ERLIC), have been described to increase the depth of the proteome as well as the number of PTMs analyzed.5–8 Off-line separation methods are commonly utilized in large-scale PTM analyses due to the ease of implementation and the ability to use large amounts of starting material. I on exchange chromatography remains the method of choice for peptide fractionation due to the availability of several different column types and its well-characterized separation properties that are orthogonal to those of reversed-phase chromatography. Charged-based separation, such as strong cation exchange (SCX) chromatography, is well-suited for enrichment of phosphopeptides from a population of unmodified tryptic peptides due to the negative charges of the phosphoryl group.9 For these reasons, SCX (as well as HILIC) chromatography in combination with TiO2/IMAC has, so far, been the most widely used off-line fractionation strategy applied in large-scale phosphoproteomics studies.10–15 However, fractionation methods based on hydrophobic interactions, such as reversed-phase chromatography, are more efficient for the separation of peptides compared to that of charge-based separations like SCX. Off-line peptide fractionation based on a basic reversed-phase (HpH) approach prior to LC–MS/MS analysis has shown great promise in recent years.16,17 This is due to its high-resolving power based on hydrophobic interactions and the use of concatenation strategies of pooling fractions from different parts of the gradient when running at high pH.17 This is highly orthogonal with the downstream phosphopeptide enrichment and LC–MS/MS analysis using low-pH reversed-phase chromatography in-line with the mass spectrometer. However, so far, only a limited number of studies have assessed the impact on global phosphoproteome coverage with prior fractionation based on hydrophobic interactions.18 This study provides the first insights into phosphopeptide fractionation by HpH; however, this was performed using a sophisticated on-line 3D fractionation platform inaccessible to most users.18 Here, we show a thorough analysis of phosphoproteome coverage with standard off-line high-pH reversed-phase fractionation (HpH) prior to phosphopeptide enrichment and MS analysis.

### EXPERIMENTAL PROCEDURES

**NIH-3T3 Culture and Lysis**

NIH-3T3 cells were grown in DMEM (Gibco, Invitrogen, Denmark) with 50 U/mL penicillin, streptomycin, and 4 mM glutamine and with the addition of 10% newborn calf serum (NBCS, Gibco, Invitrogen, Denmark). Cell cultures were grown to 60–80% confluence on 245 mm² polystyrene Petri dishes and then serum starved overnight in DMEM without serum. The following day, cells were serum-stimulated for 10 min with DMEM containing 10% NBCS for 10 min. Immediately following stimulation, the media was decanted, and the cells were washed with ice cold 1X PBS. Briefly, cold RIPA buffer (1% NP-40, 0.1% Na-deoxycholate, 150 mM NaCl, 1 mM EDTA, and 50 mM Tris, pH 7.6) supplemented with 50 mM sodium fluoride (NaF), 10 mM sodium orthovanadate, 50 mM β-glycerol phosphate, and 1 protease inhibitor cocktail tablet (Roche, Basel, Switzerland) was added to each 245 mm² plate, which was placed on a shaker at 4 °C for 20–30 min. Cells were scrapped and collected along with remaining buffer into Falcon tubes. Tubes were then centrifuged for 25 min at 7197g (4 °C). Protein in the supernatant was acetone-precipitated with four excess volumes of ice-cold acetone and stored at −20 °C until further processing.

**Protein Digestion**

Following acetone precipitation, the precipitates were centrifuged for 5 min at 2000g. Acetone was discarded, and the resulting pellets were resuspended in 5 mL of denaturing buffer consisting of 6 M urea, 2 M thiourea, 10 mM HEPES, pH 8.0. Approximate protein concentration was determined using the Bradford assay (Bio-Rad, Hercules, CA, USA), and protein disulfides were reduced with diithiothreitol (DTT) at 1 mM final concentration for 60 min at room temperature (RT). Following DTT reduction, free cysteine thiols were alkylated with chloroacetamide (5 mM final) for 60 min at RT in the dark. Lys-C protease (Wako Chemicals, Richmond, VA, USA) at a ratio of 1:100 to the total protein amount was added and allowed to incubate at 30 °C for 4 h. Urea concentration was diluted to 2 M using 50 mM ammonium bicarbonate, and the samples were incubated overnight at 37 °C with trypsin (1:50 w/w) (Life Technologies, Carlsbad, CA, USA).

**C18 Sep-Pak Cleanup of Peptides**

Following digestion, the resulting peptide mixtures were acidified with trifluoroacetic acid (TFA) to 2% final concentration and centrifuged at 4000g for 5 min, and supernatants were transferred to reversed-phase C18 Sep-Pak cartridges (Waters, Milford, MA, USA) for desalting and concentration. Sep-Pak cartridges were prepared by sequential washing with methanol, acetonitrile, and 0.1% TFA prior to loading of the peptide mixtures. Gravity was used for washing and loading of the samples. Sep-Pak cartridges were stored at 4 °C until further processing.

Peptides were eluted from Sep-Pak with 40% (4 mL) and 60% (2 mL) acetonitrile sequentially. For samples eluted for HpH fractionation, the elutes were dried in a SpeedVac centrifuge for 40 min at 45 °C, and peptide concentration was determined by Nanodrop (Thermo Fisher Scientific, Bremen, Germany) measurement at A_{260nm}.

**Basic Reversed-Phase (HpH) High-Pressure Liquid Chromatography (HPLC) Fractionation**

A few milligrams (2 to 3) of peptide mixture from biological replicates (each one resulting from a single 245 mm² cell culture plate) was fractionated using a Waters XBridge BEH130 C18 3.5 μm 4.6 × 250 mm column on a Ultimate 3000 HPLC (Dionex, Sunnyvale, CA, USA) operating at 1 mL/min. A rhodyne MXII pump (IDEX Corporation, Rohnert Park, CA, USA) was coupled to the HPLC and used to inject sample into the column. For initial experiments, buffer A consisted of 10 mM ammonium formate and buffer B consisted of 10 mM ammonium formate with 90% acetonitrile; both buffers were adjusted to pH 10 with ammonium hydroxide as described previously.19 For experiments conducted without ammonium formate, 10 mM ammonium hydroxide was used as the only additive to the mobile phases.

A similar injection protocol and gradient were used for all fractionation experiments; all fractions were collected using a Dionex AFC-3000 fraction collector in a 96 deep well plate at 1 min intervals. Samples were initially loaded onto the column at 1 mL/min for 4 min, after which the fractionation gradient...
commenced as follows: 1% B to 25% B in 50 min, 60% B in 4 min, and ramped to 70% B in 2 min. At this point, fraction collection was halted, and the gradient was held at 70% B for 5 min before being ramped back to 1% B, where the column was then washed and equilibrated.

The total number of fractions concatenated was set to 14 (Scheme 1) throughout all experiments. For the optimized method, the initial fractions were pooled and represented as 1 fraction and concatenated with other fractions as described. Prior to concatenation, ammonium hydroxide and/or ammonium formate were evaporated in a SpeedVac operating at 60 °C. It should be noted that lower temperatures are preferable when analyzing labile modifications such as O-glycosylation.

**Strong Cation Exchange (SCX) Fractionation**

SCX fractionation was performed as described. Briefly, 2 to 3 mg of peptides was fractionated on a 1 mL Resource S column (GE Healthcare, Sweden) using an AKTA FPLC system (GE Healthcare, Sweden). Peptides eluted from Sep-Pak were directly injected onto the column. Peptides were separated by increasing salt concentration at 1 mL/min from 100% SCX buffer A (5 mM KH$_2$PO$_4$, pH 2.7, 30% ACN) to 30% SCX buffer B (5 mM KH$_2$PO$_4$, pH 2.7, 350 mM KCl, 30% ACN) over the course of 30 min; B was then increased to 100%, where it was held for 6 min. Fractions were collected at 2 min intervals (2 mL each) and pooled at various intervals to bring total fraction number to 11 (Scheme 1).

**Phosphopeptide Enrichment**

For all experiments, phosphopeptides from each HpH and SCX fraction were enriched in batch mode with metal oxide affinity enrichment (MOAC) using titanium dioxide beads ($\mu$m Titansphere, GL Sciences, Japan). TiO$_2$ beads were preincubated in 2,5-dihydroxybenzoic acid (20 mg/mL) in 80% ACN and 1% TFA (5 $\mu$L/mg of beads) for 20 min. Prior to addition of the beads, all fractions were brought to 80% acetonitrile and 5% TFA in a final volume of 10 mL. Two milligrams (in 10 $\mu$L of DHB solution) of beads was added to each fraction, which

---

**Scheme 1. Overview of the Workflows Presented in This Study**

“Tryptic peptide mixtures are first fractionated using either a strong cation exchange or high-pH reversed-phase approach followed by titanium dioxide phosphopeptide enrichment. This is followed by a cleanup and loading step prior to LC–MS analysis.”
was then incubated for 30 min while rotating. After incubation, fractions were centrifuged at 4000g for 5 min, and supernatant was removed.

Beads were washed with 30% ACN and 1% TFA and loaded onto C8 STAGE-tip (3 M Empore, St. Paul, MN, USA) and centrifuged at 500g. Beads were washed on-tip with 50% ACN and 1% TFA followed by 80% ACN and 1% TFA, after which the phosphorylated peptides were eluted first with 20 μL of 5% NH₄OH followed by 20 μL of 10% NH₄OH with 25% ACN. Eluted peptides were concentrated in a SpeedVac, loaded onto C18 STAGE-tips, and primed for LC–MS analysis.

**LC–MS/MS Analysis**

All phosphopeptide samples were analyzed with Easy-nLC 1000 (Proxeon, Odense, Denmark) coupled to Thermo Fisher Q-Exactive Orbitrap or Q-Exactive Orbitrap Plus. Peptides were separated on in-house packed column (75 μm i.d. × 15 cm length) with 1.9 μm C18 beads (Dr. Maisch, Germany). Separation was achieved using a linear gradient of increasing buffer B (80% ACN and 0.1% formic acid) and decreasing buffer a (0.1% formic acid) at 250 nL/min. Buffer B was increased to 30% B in 70 min and ramped to 40% B in 5 min followed by a quick ramp to 80% B, where it was held for 5 min before a quick ramp back to 5% B, where it was held and the column was re-equilibrated.

Q-Exactive mass spectrometer was operated in positive polarity mode with capillary temperature of 275 °C. Full MS survey scan resolution was set to 70 000 with an automatic gain control (AGC) target value of 1 × 10⁶ for a scan range of 350–1750 m/z. A data-dependent top 10 method was operated during which higher-energy collisional dissociation (HCD)²⁴ spectra were obtained at 17 500 MS² resolution with AGC target of 1 × 10⁶ and maximum ion injection time (IT) of 48 ms, 4 m/z isolation width, and normalized collisional energy (NCE) of 25. Former precursor ions targeted for HCD were dynamically excluded of 15 s. Q-Exactive mass spectrometer was operated in positive polarity mode with capillary temperature of 275 °C.

**Data Analysis**

All files were analyzed using the MaxQuant software suite 1.4.1.4 (www.maxquant.org) with the Andromeda search engine.²⁵ For NIH-3T3 cell line sample files, the HCD MS/MS spectra were searched against an in silico tryptic digest of *Mus musculus* proteins from the UniProt sequence database (v. April 2014) containing 81 824 sequences, including spliced isoforms, that is split between 24 488 reviewed protein entries (Swiss-Prot IDs) and 57 336 unreviewed protein entries (TrEMBL IDs). All MS/MS spectra were searched with the following MaxQuant parameters for phosphopeptide identification: acetyl (protein N-terminus), methionine oxidation, phosphorylation (STY), and glutamine cyclization to pyroglutamate were searched as variable modifications; cysteine carbamidomethylation was set as fixed modification; max 3 missed cleavages; and precursors were initially matched to 4.5 ppm tolerance and 20 ppm for fragment spectra. Peptide spectrum matches, proteins, and sites were automatically filtered to a 1% false discovery rate based on Andromeda score, peptide length, and individual peptide mass errors. Modified peptides required a minimum peptide length of at least seven amino acids (AA) and a minimum Andromeda score of 40 to be considered a positive hit. Analysis of phosphopeptide variants was performed using the MaxQuant Evidence table. Confidently assigned phosphorylation sites with localization probabilities greater than 75% were defined as Class I sites, and were used as site localization cutoff from the Phospho (STY)Sites table generated by MaxQuant analysis, and the Evidence table was used for phosphopeptide variant analysis. All raw files from this study have been deposited to the ProteomeXchange Consortium with data set identifier PXD001404 via the PRIDE partner repository.

Functional analysis for gene ontology (GO) enrichment was performed using DAVID.²⁷,²⁸ Corresponding UniProt identifiers to proteins of Class I sites for each experiment were uploaded as a gene list. Background based on an LC–MS experiment of the NIH-3T3 proteome run was utilized. A cutoff of *p* < 0.05 was utilized for all GO categories in our analysis. Sequence analysis of phosphorylation sites was performed using icelogo 1.2.⁹ Peptide sequences were uploaded as positive and negative set, and the percent difference was selected as the scoring system, with *p* values less than 0.01.

### RESULTS AND DISCUSSION

**Comparison of High-pH Reversed-Phase (HpH) and Strong Cation Exchange (SCX) for Phosphoproteomics**

To determine the efficacy of basic reversed-phase (HpH) fractionation for large-scale phosphoproteomics, we analyzed tryptic digests of whole-cell lysates from mouse NIH-3T3 cells that were separated by HpH followed by TiO₂ enrichment of phosphopeptides from concatenated fractions. The resulting phosphopeptide fractions were separated on-line by nanoflow liquid chromatography and analyzed by high-resolution tandem mass spectrometry on a Q-Exactive instrument. We carefully evaluated the number of the identified phosphopeptides and their properties, and, on the basis of this, we optimized the HpH fractionation and concatenation strategy. For comparative purposes, strong cation exchange (SCX) provided a good reference point, as it has become a widely established protocol for fractionation of peptides and phosphopeptides (Scheme 1). In all experiments, we fractionated total digests first and enriched phosphopeptides afterward. We reasoned that the decreased complexity of the sample mixtures in fractions would allow for greater enrichment efficiency of phosphopeptides. This was confirmed in the comparison of HpH and SCX fractionation, where we observed phosphopeptide enrichment efficiencies of more than 90% and up to 97% in some experiments (Table 1 and Supporting Information Table 1). For both HpH and SCX, we maintained a total of 14 fractions. In the HpH chromatography, we collected a total of 56 fractions that were pooled and concatenated to 14 (Figure 1), and for the optimized SCX protocol, we collected 11 fractions.

| Fractionation | 1 P | 2 P | ≥3 P | Total Phospho | Non-Phospho | % Enrichment |
|---------------|-----|-----|------|--------------|-------------|--------------|
| HpH           | 12609 (±2655) | 4199 (±1005) | 759 (±165) | 17 566 (±3737) | 892 (±256) | 95 (±2) |
| SCX           | 1651 (±330) | 3360 (±1077) | 1295 (±519) | 6215 (±1759) | 350 (±287) | 94 (±6) |
from which double TiO$_2$ (MOAC) enrichment was performed on 3 fractions for a total of 14 (Figure 1). All SCX and HpH fractions were analyzed using the same 70 min LC gradient and the same MS settings (see Experimental Procedures).

The number of phosphopeptides from HpH was generally higher across fractions compared to that from SCX, where we observed that a few fractions carried the main proportion of phosphopeptides (Supporting Information Figure 1). Using four biological replicates from each experiment, we found that HpH had an exceptional ability to separate singly phosphorylated peptides, as we found an average of 12,609 (±2655) singly phosphorylated peptides in the HpH-separated samples compared to 15,611 (±330) when using SCX, a difference of almost 8-fold (Table 1 and Supporting Information Table 1). HpH fractionation also revealed a slight advantage for the enrichment of doubly phosphorylated peptides, as we found an average of 4199 (±1005) compared to 3360 (±1077) in the SCX experiments. However, SCX displayed a higher propensity toward peptides with three or more phosphoryl groups due to its advantage of separation based on charge (Table 1 and Supporting Information Table 1). The analysis of localized class I phosphorylation sites (pSTY) showed a total of 15,557 phosphorylation sites enriched after the HpH fractionations and only 6827 from SCX. This indicates that separating peptides with SCX produced less than half of the number of class I sites compared to that from HpH fractionation. Excellent overlap of these sites was observed, as 86% of the class I sites identified from SCX experiments were also identified in HpH (Figure 2A and Supporting Information Table 2), demonstrating that HpH is indeed superior for in-depth phosphoproteomics. Altogether, these results indicate a severe undersampling of singly phosphorylated peptides with SCX fractionation, suggesting that the greater resolving power via hydrophobic interactions of HpH is required for efficient separation and detection of phosphopeptides.

**Optimization of HpH for Phosphoproteomics**

As described previously, at higher pH values, multiphosphorylated peptides attain a high number of negative charges, thus reducing interactions with the reversed-phase material and leading to the loss of a high percentage of multiphosphorylated peptides in the initial flow through. We observed a similar trend with our off-line fractionation setup (Supporting Information Figure 1). To maximize the total number of phosphopeptides identified and to increase the amount of multiphosphorylated peptides identified with HpH, we made a few informed changes to our strategy (HpH Opt). We increased the quality of the HCD fragment spectra and performed double incubation of specific HpH fractions or pools of fractions (Figure 1B). With optimized parameters, we were able to increase the total number phosphopeptide variants to almost 40,000, an increase of more than 2-fold compared to that from the initial HpH strategy and 6-fold to that from the SCX experiments (Tables 1 and 2). Furthermore, we were able to dramatically increase the number of class I phosphorylation sites by almost 4-fold compared to that from SCX and 2-fold compared to that from previous HpH experiments (Figure 2B,C).

Combination of optimized MS parameters with double TiO$_2$ incubation (HpH Opt TiO$_2$) resulted in identification of 32,446 phosphopeptides, an increase of 84% from the average of previous HpH experiments (Table 2). Additionally, class I sites increased by 66% to 18,250 from one experiment over the average number of sites from initial HpH experiments (Supporting Information Table 2). Using this double TiO$_2$ enrichment strategy, we were able to significantly increase the number of phosphopeptides with three or more phosphoryl groups to more than 2000 (Figure 2D). This number of multiphosphorylated peptides was also substantially higher compared to the average of the SCX experiments. These results demonstrate that high coverage of multiphosphorylated peptides is possible with HpH and illustrates the capacity of HpH to considerably increase the depth of the phosphoproteome.

In phosphoproteomics, we are faced with the challenge of localizing phospho groups in the identified peptide sequence.
with single amino acid accuracy. To successfully achieve this, high sequence coverage in peptide fragmentation spectra is a necessity. Higher quality HCD spectra allow for better y and b ion series coverage of the peptide, including accurate localization of the phosphoryl group. In the case of Orbitrap instruments, as presented in this study, higher scanning time is required for higher quality spectra, doubling the resolution results with a doubling of the total scan time. Consequently, the percentage of peptides identified increases at the cost of scan speed. As described by Kelstrup et al.,\textsuperscript{30} optimization of parameters for sensitive and fast analysis depends on a variety of factors such as sample complexity, starting material, and run time. For samples of high complexity and high sample amounts, fast scanning methods are preferred in order to sequence as many peptides as possible for high coverage of the proteome.\textsuperscript{30,31} For samples that are less complex (i.e., of low abundance, such as fractionated and enriched), scan speed is not a limiting factor due to low number of co-eluting peptides.

In this case, sensitive scanning methods are optimal for identification and localization of phospho groups (Figure 3A). Despite the high number of phosphopeptides identified in each fraction in the first four HpH experiments, the percentage of MS/MS identified (peptide spectrum matches post FDR filtering at 1%) was only on the order of 6% on average (Supporting Information Table 3). This indicated that the general quality of the acquired HCD spectra was too low. We reasoned that increasing the HCD spectrum quality by doubling the Orbitrap analyzer resolution from 17,500 to 30,000 (at fwhm) and by utilizing sub 2 μm/μm precursor isolation windows (1.3 μm/μm) with higher c-trap injection times (108 ms) would increase identification rates (Figure 3A). Most importantly, with higher injection times (IT), peptide ions are accumulated for a longer period, thus increasing the MS signal intensities of the fragment ions. Additionally, narrow isolation would reduce the amount of co-fragmenting interfering species, thus reducing the fragment spectra

Table 2. Evaluation of Phosphopeptides Sequenced with Optimized Parameters

| fractionation | biological replicate | 1 P | 2 P | 3 P | total phospho | non-phospho | % enrichment |
|---------------|----------------------|-----|-----|-----|---------------|-------------|-------------|
| HpH Opt       | 1                    | 31,082 | 7411 | 820  | 39,283        | 4833        | 89.05       |
|               | 2                    | 30,344 | 5962 | 767  | 37,073        | 4765        | 88.61       |
| HpH Opt Double TiO2 | 1            | 20,785 | 9198 | 2462 | 32,445        | 3197        | 91.03       |

Figure 2. (A) Analysis of overlap of class I phosphorylation sites between SCX and HpH experiments. (B) Class I site overlap between HpH (n = 4) and HpH Opt (n = 3) experiments performed on Q-Exactive Plus with optimized parameters. (C) Number of class 1 sites arising from the three different strategies. (D) Number of multiply phosphorylated peptides identified by SCX, HpH, and HpH Opt with double TiO2 incubation.
complexity, which can hinder identification and localization. The strategy was validated, as the average percentage of MS/MS identified from HpH optimized experiments was 21.6%, an increase of almost 4-fold compared to that with the previous MS acquisition method (Supporting Information Table 3). Besides the higher number of phosphopeptide variants and class I sites detected (Table 2 and Supporting Information Table 2), we observed that the Andromeda scores for phosphopeptides were significantly higher using the optimized parameters (Figure 3B). The average Andromeda score of 142.8 (more than 60% increase) indicated that sequence assignment can be achieved with higher confidence. Higher scores for optimized parameters were independent of the fractionation method, as SCX and HpH had similar average Andromeda scores, 85.7 and 88.8, respectively, despite higher number of phosphopeptides sequenced using HpH. These results demonstrate the importance of having high-quality fragment spectra for analysis of fractionated phosphopeptides.

Although these results validated our initial assessment and optimization of the HpH protocol, further experiments over time showed a sharp decline in each measurable category (i.e., the number of phosphopeptides identified) (Figure 4A). As a note, the original HpH and HpH Opt TiO2 experiments were performed on the same fractionation column before degrada-
As mentioned by Percy et al., removal of ammonium formate phosphorylation sites in peptides without missed cleavages. Bars are represented as log10-transformed the separation efficiency and coverage of phosphorylation sites (Table 2). In fact, we observed a significant increase in the total number of phosphopeptides detected using this simple buffer system. These surprising results led us to further assess the impact of ammonium formate in the running buffers. Analysis of unique phosphopeptide variants displayed a linear increase in the number of total phosphopeptides over the fractions, indicating a small fraction-to-fraction variation (Figure 4B). These results demonstrate that a similar separation of phosphopeptides is achieved without the addition of ammonium formate at higher pH. Although the column manufacturer reports stability at high pH and high column temperatures, compatibility with buffer systems for separation of peptides is equally important. Therefore, we recommend future high-pH reversed-phase experiments to be performed in buffer systems for separation of peptides; however, the opposite was observed (Supporting Information Table 4), where the mean of the m/z distribution of phosphopeptides was roughly 80 m/z units higher (727.59 m/z vs 646.95 m/z). We reasoned that this could be attributed to generally longer peptide lengths of phosphopeptides and to the presence of the phosphoryl group (+79.96 Da). Analyses of peptide length showed that the phosphopeptides were, on average, five amino acid residues longer, with an average length of 17 for those without any missed cleavage sites (Supporting Information Figure 2). As a result, the m/z distribution was expected to be lower than that of unmodified peptides; however, the opposite was observed (Supporting Information Figure 3), where the mean of the m/z distribution of phosphopeptides was roughly 80 m/z units higher (727.59 m/z vs 646.95 m/z). We reasoned that this could be attributed to generally longer peptide lengths of phosphopeptides and to the presence of the phosphoryl group (+79.96 Da). Analyses of peptide length showed that the phosphopeptides were, on average, five amino acid residues longer, with an average length of 17 for those without any missed cleavage sites (Supporting Information Table 4). Further analysis revealed that more than half of the unique phosphopeptide sequences (56%) carried missed cleavages, compared to less than one-fourth for unmodified peptides (23%). Sequence motif analysis using iceLogo of phosphopeptides carrying missed cleavages (positive set) showed a significant enrichment of Arg/Lys residues in −1,
−3 positions N-terminal to the phosphorylation site (Figure 5). Surprisingly, we observed a significant enrichment of tyrosine phosphorylation sites (p < 10 × 10^{−9}) from phosphopeptides without missed cleavages compared to that from missed cleaved phosphopeptides (Figure 5B), suggesting that tyrosine phosphorylation may not form similar salt bridges with proximal arginine and lysine residues, which can prevent efficient trypsin cleavage. These results were consistent with earlier findings, whereby using synthetic peptides it was reported that phosphorylated tyrosines proximal to arginine or lysine residues were much more amiable to tryptic digestion compared to those peptides carrying phosphoserine or phosphothreonine. Phosphotyrosines are particularly interesting because tyrosine kinases are key initiators of signaling cascades.

**Phosphoproteome Coverage by HpH and Its Impact on Cell Signaling Analyses**

Although serine and threonine generally make up most of the phosphorylation sites detected in phosphoproteomics workflows (~98% in our studies) (Supporting Information Figure 4), tyrosine phosphorylation is a key component of signal transduction and cellular responses. To overcome the low abundance of tyrosine phosphorylation, phosphotyrosine-specific antibodies are typically utilized for enrichment and immunoprecipitation of tyrosine phosphorylated peptides. In our HpH experiments of the NIH-3T3 phosphoproteome, we observed a very large coverage and overrepresentation of phosphotyrosine class I sites when compared with that from SCX (Table 3 and Supporting Information Figure 4). We found that the majority (72%) of phosphotyrosine containing peptides were singly phosphorylated and 22% were double phosphorylated (data not shown). The ability of HpH fractionation to resolve eight times as many singly phosphorylated peptides compared to that of SCX indicates a correlation with the increased coverage of phosphotyrosine peptides. Close to 500 localized tyrosine phosphorylation sites were observed in the optimized HpH experiments (Table 3), representing 357 phosphoproteins. Gene ontology (GO) analysis revealed enrichment of transmembrane receptor protein tyrosine kinase signaling pathways such as ErbB, VEGF, and other growth factor signaling pathways and other biological processes linked to signaling cascades (data not shown). Cell compartment (CC) analysis of the tyrosine phosphorylated proteins showed enrichment for plasma membrane proteins, suggesting that HpH fractionation permits coverage of membrane tyrosine phosphorylation without affinity enrichment. It should be noted that phosphotyrosine coverage is expected to be even greater under specific conditions such as short growth factor stimulation, which specifically stimulates receptor tyrosine kinases. Nevertheless, our results obtained with a 10 min stimulation of growing cells with serum demonstrates the power of HpH for in-depth analysis of the tyrosine phosphoproteome.

To explore the extent and coverage of cell signaling pathways covered by off-line HpH compared to that with SCX fractionation in large-scale phosphoproteome experiments, we performed gene ontology (GO) enrichment analysis using DAVID of the SCX and HpH data sets. As expected, we observed a greater number of enriched categories of biological processes, compartmental categorization and molecular function (data not shown) above the cutoff (p < 0.05), and higher number of counts per category in experiments conducted with HpH compared to that with SCX. Surprisingly, some categories were underrepresented in the SCX experiments but overrepresented in HpH experiments. For instance, analysis of compartmental categorization revealed overrepresentation of cytosolic phosphoproteins in HpH experiments but not in SCX (Chart 1). We believe that the increased coverage is a result of large number of singly phosphorylated peptides derived from proteins in the cytosol identified by HpH fractionation. Understanding the role of phosphorylation in different cellular organelles requires deep coverage of the phosphoproteome, and our data sets demonstrate that this is now possible with HpH fractionation. We imagine that HpH fractionation in combination with quantitative proteomics technologies such as stable isotope labeling by amino acids in culture (SILAC) will be particularly useful for analyzing global and dynamic phosphorylation site changes in cell signaling pathways.

### Table 3. Number of Class I Phosphotyrosine Sites Covered with the Different Fractionation Strategies

| Fractionation | Biological replicate | Class I pY | Unique experiment | Unique fractionation | Total |
|---------------|---------------------|------------|-------------------|---------------------|-------|
| HpH           | 1                   | 148        | 196               | 520                 | 526   |
|               | 2                   | 87         |                   |                     |       |
|               | 3                   | 75         |                   |                     |       |
|               | 4                   | 71         |                   |                     |       |
| HpH Opt       | 1                   | 322        | 496               |                     |       |
|               | 2                   | 328        |                   |                     |       |
| HpH Opt TiO₂  | 1                   | 205        |                   |                     |       |
| SCX           | 1                   | 26         | 39                | 39                  |       |
|               | 2                   | 16         |                   |                     |       |
|               | 3                   | 14         |                   |                     |       |
|               | 4                   | 22         |                   |                     |       |

Reduction of the high dynamic range of protein abundance is crucial for deep coverage of the proteome and especially the phosphoproteome, where the dynamic range is even more pronounced. This issue is usually addressed in two ways: either through introduction of more powerful and sensitive mass spectrometric instrumentation or by decreasing sample complexity via fractionation. Although great advances in instrument sensitivity and sequencing speed have been made, fractionation is still required for in-depth coverage of phosphoproteomes. Off-line high-pH reversed-phase fractionation is exceptional in its resolving power for the separation of peptides and is highly orthogonal to low-pH reversed-phase separation used in tandem with MS analysis. Deep coverage can be crucial when attempting to understand complex biological systems. As our results indicate, greater coverage has implications in cell signaling analyses. Robust and reproducible sample preparation and MS methods are essential, as global phosphoproteomics is the method of choice for unbiased analysis of cell signaling networks and is becoming routine in many non-MS expert laboratories.

Our results show that phosphopeptide separation based on high-pH reversed-phase chromatography is a powerful fractionation method for global phosphoproteomics and that it is superior to traditional SCX-based strategies. Removal of ammonium formate from high-pH buffers can increase column performance over time, and double TiO₂ enrichment allows for high coverage of multiphosphorylated peptides from pooled samples.
HpH fractions. Similarly, our results indicate that MS parameters must be optimized based on the samples analyzed to increase the general quality of the HCD spectra, contributing to greater phosphorylation site localization. We hope that this study will serve as a useful resource and stimulate the wide adoption of high-pH-based fractionation in phosphoproteomics studies.

**ASSOCIATED CONTENT**

* Supporting Information

Table 1: Phosphopeptide variants for each individual biological replicate arising from HpH and SCX. Table 2: Class I phosphorylation sites for each biological replicate of SCX, HpH, and HpH Opt. Table 3: (A) Percent of HCD spectra (MS/MS) identified and MS/MS spectra identified (SIL) for which an isotope pattern was detected in each replicate. (B) Average identified MS/MS spectra for each individual experiment. Table 4: (A) Number of unmodified peptides with and without missed tryptic cleavage sites from a typical large-scale proteomics data set. (B) Number of phosphopeptides with and without missed tryptic cleavages. The average peptide sequence lengths (number of amino acid residues) are also indicated. Figure 1: Unique phosphopeptide variants are plotted based on total counts (A) and with varying number of phosphoryl groups (B–D). For comparison purposes, three biological replicates performed together in series were used. Figure 2: Peptide charge state as percentages for unmodified and phosphopeptides. Figure 3: Peptide precursor m/z ratios in bins of 50 units are plotted for unmodified and phosphophorylated peptides. Figure 4: Phosphorylation site distributions of serine, threonine, and tyrosine based on all class I sites from this study. Excel workbooks for phosphopeptide variants for each experiment and phosphorylation sites. This material is available free of charge via the Internet at http://pubs.acs.org.

**AUTHOR INFORMATION**

* Corresponding Author

E-mail: jesper.olsen@cpr.ku.dk. Tel.: +45 35 32 50 00. Fax: +45 35 32 50 01.

**NOTES**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

The Novo Nordisk Foundation (NNF) Center for Protein Research is supported by an NNF donation. The work was supported by the Seventh Framework Program of the European Union (262067-PRIME-XS). The authors thank Mats Wikström for HPLC assistance as well as Christian Kelstrup and Rosa Rakownikow Jersie-Christensen for helpful discussions. The authors additionally thank Dorte Breinholdt Bekker-Jensen for assistance with figures and helpful discussions.

**REFERENCES**

(1) Olsen, J. V.; Mann, M. Status of large-scale analysis of post-translational modifications by mass spectrometry. *Mol. Cell. Proteomics* 2013, 12, 3444−3452.

(2) Lemmon, M. A.; Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* 2010, 141, 1117−1134.

(3) de Graaf, E. L.; Giansanti, P.; Altelaar, A. F. M.; Heck, A. J. R. Single step enrichment by Ti<sup>4+</sup>-IMAC and label free quantitation enables in-depth monitoring of phosphorylation dynamics with high reproducibility and temporal resolution. *Mol. Cell. Proteomics* 2014, 13, 2426−2434.

(4) Lundby, A.; Andersen, M. N.; Steffensen, A. B.; Horn, H.; Kelstrup, C. D.; Francavilla, C.; Jensen, L. J.; Schmitt, N.; Thomsen, M. B.; Olsen, J. V. In vivo phosphoproteomics analysis reveals the cardiac targets of β-adrenergic receptor signaling. *Sci. Signaling* 2013, 6, rs11.

(5) Washburn, M. P.; Wolters, D.; Yates, J. R. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* 2001, 19, 242−247.

(6) Villen, J.; Gygi, S. P. The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nat. Protoc.* 2008, 3, 1630−1638.

(7) Di Palma, S.; Boersma, P. J.; Heck, A. J. R.; Mohammed, S. Zwitterionic hydrophilic interaction liquid chromatography (ZIC-HILIC and ZIC-cHILIC) provide high resolution separation and increase sensitivity in proteome analysis. *Anal. Chem.* 2011, 83, 3440−3447.

"p values for each category are −log<sub>10</sub>-transformed and plotted. The number of counts for each significantly enriched category from the corresponding experiments are indicated."
Cell A tissue-specific atlas of mouse protein phosphorylation and in response to TGF-β: dissection of early phosphoproteome and ensuing proteome changes
Choudhary, C.; Dijke, P. t.; Mann, M.; Sharma, K. Time-resolved phosphoproteomics applied to the yeast pheromone signaling pathway.
[635]
Servick, K. Online TiO2-based phosphoproteomics applied to study endogenous phosphorylation in K-hexamer DNA
[648].
Mol. Cell. Proteomics 2010, 9, 7310–7321.
(8) Hao, P.; Guo, T.; Li, X.; Adav, S. S.; Yang, J.; Wei, M.; Sze, S. K. Novel application of electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) in shotgun proteomics: comprehensive profiling of rat kidney proteome. J. Proteome Res. 2010, 9, 3520–3526.
(9) Beausoleil, S. A.; Jedrychowski, M.; Schwartz, D.; Elias, J. E.; Villén, J.; Li, J.; Cohn, M. A.; Cantley, L. C.; Gygi, S. P. Large-scale characterization of HeLa cell nuclear phosphoproteins. Proc. Natl. Acad. Sci. U.S.A. 2004, 101, 12130–12135.
(10) D’Souza, R. C. J.; Knittle, A. M.; Nagaraj, N.; Dinther, M. V.; Choudhary, C.; Dijke, P. t.; Mann, M.; Sharma, K. Time-resolved dissection of early phosphorylated and ensuing proteome changes in response to TGF-β. Sci. Signal. 2014, 7, rs5.
(11) Gruter, A.; Olsen, J. V.; Mohammed, S.; Mortensen, P.; Faegerman, N. J.; Mann, M.; Jensen, O. N. Quantitative phosphoproteomics applied to the yeast phenome signaling pathway. Mol. Cell. Proteomics 2005, 4, 310–327.
(12) Huttlin, E. L.; Jedrychowski, M. P.; Elias, J. E.; Goswami, T.; Rad, R.; Beausoleil, S. A.; Villén, J.; Haas, W.; Sowa, M. E.; Gygi, S. P. A tissue-specific atlas of mouse protein phosphorylation and expression. Cell 2010, 143, 1174–1189.
(13) Olsen, J. V.; Blagoev, B.; Gnud, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell 2006, 127, 635–648.
(14) Pinkse, M. W. H.; Mohammed, S.; Gouw, J. W.; van Breukelen, B.; Vos, H. R.; Heck, A. J. R. Highly robust, automated, and sensitive online TiO2-based phosphoproteomics applied to study endogenous phosphorylation in Drosophila melanogaster. J. Proteome Res. 2008, 7, 687–697.
(15) McNulty, D. E.; Annan, R. S. Hydrophilic interaction chromatography reduces the complexity of the phosphoproteome and improves global phosphopeptide isolation and detection. Mol. Cell. Proteomics 2008, 7, 971–980.
(16) Gilart, M.; Olivova, P.; Daly, A. E.; Geblher, J. C. Orthogonality of separation in two-dimensional liquid chromatography. Anal. Chem. 2005, 77, 6426–6434.
(17) Wang, Y.; Yang, F.; Gritsenko, M. A.; Wang, Y.; Clauss, T.; Liu, T.; Shen, Y.; Monroe, M. E.; Lopez-Ferrer, D.; Reno, T.; Moore, R. J.; Klems, R. L.; Camp, D. G.; II; Smith, R. D. Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. Proteomics 2011, 11, 2019–2026.
(18) Ficarro, S. B.; Zhang, Y.; Carrasco-Alfonso, M. J.; Garg, B.; Adelman, G.; Webber, J. T.; Luckey, C. J.; Marto, J. A. Online nanoflow multidimensional fractionation for high efficiency phosphopeptide analysis. Mol. Cell. Proteomics 2011, 10, 775–786.
(19) Udeshi, N. D.; Svinkina, T.; Mertins, P.; Kuhn, E.; Mani, D. R.; Qiao, J. W.; Carr, S. A. Refined preparation and use of anti-diglycine remnant (K-ε-GG) antibody enables routine quantification of 10,000s of ubiquitination sites in single proteomics experiments. Mol. Cell. Proteomics 2012, 12, 825–831.
(20) Francavilla, C.; Hekmat, O.; Blagoev, B.; Olsen, J. V. SILAC-based temporal phosphoproteomics. Methods Mol. Biol. 2014, 1188, 125–148.
(21) Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. D. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. Mol. Cell. Proteomics 2005, 4, 873–886.
(22) Pinkse, M. W. H.; Uitto, P. M.; Hills, J.; Uitto, J.; Blagoev, B.; Beck, A. J. R. Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC–ESI–MS/MS and titanium oxide precolumns. Anal. Chem. 2004, 76, 3935–3943.
(23) Rapp, P.; Mann, M.; Ishihama, Y. Protocol for microencapsulation, purification, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. 2007, 2, 1896–1906.
(24) Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Hornung, S.; Mann, M. Higher-energy C-trap dissociation for peptide modification analysis. Nat. Methods 2007, 4, 709–712.