Fully Automatic Separation and Identification of Phosphopeptides by Continuous pH-Gradient Anion Exchange Online Coupled with Reversed-Phase Liquid Chromatography Mass Spectrometry

ARTICLE *in* JOURNAL OF PROTEOME RESEARCH · DECEMBER 2008

Impact Factor: 4.25 · DOI: 10.1021/pr800381w · Source: PubMed

research articles Journal of proteome research

Fully Automatic Separation and Identification of Phosphopeptides by Continuous pH-Gradient Anion Exchange Online Coupled with Reversed-Phase Liquid Chromatography Mass Spectrometry

Jie Dai, Lian-Shui Wang, Yi-Bo Wu, Quan-Hu Sheng, Jia-Rui Wu, Chia-Hui Shieh,* and Rong Zeng*

Key Laboratory of Systems Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

Received May 26, 2008

Most current technologies for the enrichment of phosphopeptides rely on a tandem combination of different chromatography modes. Here, a fully automatic two-dimensional liquid chromatography mass spectrometry method was developed for global phosphopeptide identification. The peptide mixtures were loaded on a strong anion exchange (SAX) column under basic pH conditions and eluted with a continuous gradient to pH 2.0. This SAX system could be coupled online with reversed-phase liquid chromatography mass spectrometry (RP-LC-MS/MS). For peptide digests from a standard protein mixture spiked with synthesized phosphopeptides, most of the nonphosphorylated peptides were eluted in more basic pH than phosphopeptides, and the phosphopeptides were focused to acidic pH ranges and gradually eluted according to the phosphorylated states of peptides. Compared with the pH step elution method, the continuous gradient method displayed better repeatability and less peptide cross-overlap between fractions. This system provided a robust and fully automatic approach to large-scale phosphoproteomic profiling. For protein tryptic digests from HeLa cells, 1833 nonredundant phosphorylation sites were identified based on this two-phase separation. Compared with the method combining cation exchange and titanium dioxide, this anion-exchange based system preferred to identify more acidic and multiphosphorylated peptides. It also covered a more complete series of phosphorylation states of peptides.

Keywords: Phosphoproteomics • Phosphopeptides • Two-dimensional • Anion exchange chromatography • pH continuous online gradient • Mass spectrometry

Introduction

Genome sequencing has presented the genetic background for an increasing number of species; thus, the focus of biological science has switched to disclose protein activities and gene regulation. The phosphorylation of tyrosine, serine and threonine triggers or mediates a number of cellular processes, including metabolism, cell cycle, apoptosis and gene transcription.[1−4] The phosphorylation occurs in a reversible manner that switches on or off signal pathway or gene transcription, controlled by protein kinase and phosphotase. The genome sequencing and annotation disclosed hundreds of protein kinases and phosphotases, which regulate protein phosphorylation in humans or mice.[5−7] Searching substrates of the kinase/phsosphotase has become one of the main missions to elucidate protein phosphorylation events.

Many more technologies have recently emerged to enrich phosphopeptides to eliminate the mass spectral signal suppression of phosphopeptides caused by coelution with nonphosphopeptides in liquid chromatography. One strategy is to use affinity purification, including immunopurification with antibody,[8,9] immobilized metal affinity chromatography (IMAC),[10−12] and titanium dioxide (TiO_2).[13−15] However, for complex peptide mixtures, the single step purification by affinity materials was not effective enough to isolate all phosphopeptides. Thus, additional studies used ion exchange chromatography to split complex peptide mixtures to multiple fractions, and each fraction was further purified with affinity materials.[16−19] In these combinations, strong cation exchange (SCX) was frequently used in offline mode prior to affinity purification. Additional reversed-phase liquid chromatography mass spectrometry was combined offline with affinity chromatography. In our recent work, we constructed an offline multiple dimensional liquid chromatography mass spectrometry (Yin-yang MDLC-MS/MS) method for comprehensive protein expression and phosphorylation profiling based on pH elution.[20] It was observed that the anion exchange could isolate phosphopeptides from complex systems more effectively than cation exchange. Recently, Motoyama et al. also found the anion exchange can enhance the phosphopeptide identification when mixed with cation exchange in MudPit mode using salt step elution.[21] Han et al. also found anion exchange can enrich phosphopeptides using offline salt elution.[22]

* To whom correspondence should be addressed. E-mail: (R.Z.) zr@sibs.ac.cn; (C.-H.S.) jhxie@sibs.ac.cn.

10.1021/pr800381w CCC: $40.75

Published on Web 11/18/2008

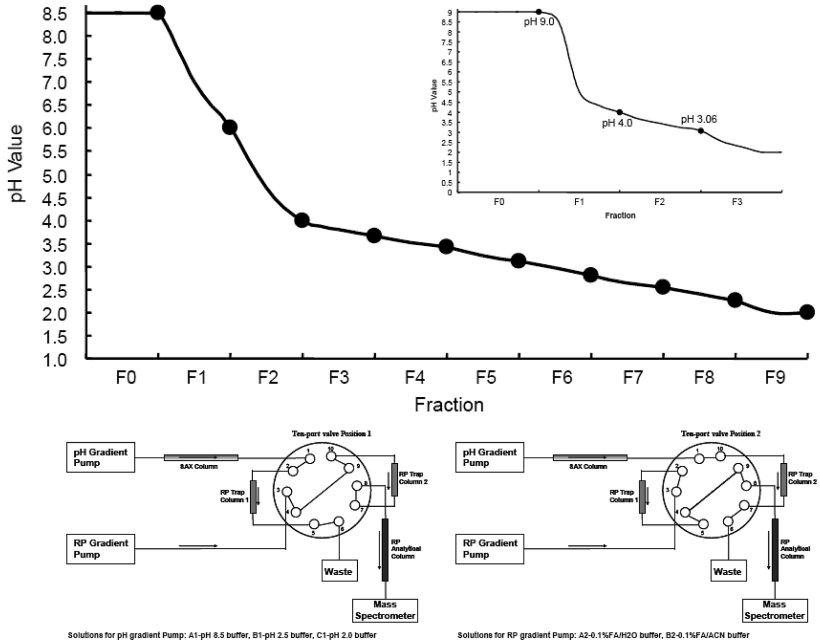

Figure 1. The configuration of the pCOG-SAX-RP-LC-MS/MS system, which is composed of two pumps, one SAX column, two trap reversed-phase columns and a separating reversed-phase column. Ten fractions were designed for complex proteome from HeLa cells. Four fractions were applied for standard protein mixture spiked with synthesized peptides (inset small picture). F0 is flow-through fraction at loading condition.

In our present work, we developed a pH continuous online gradient (pCOG) in a strong anion exchange column. The strong anion exchange column was coupled online with reversed-phase liquid chromatography mass spectrometry. This system implemented fully automatic separation, enrichment and identification of phosphopeptides. For protein tryptic digests from human HeLa cells, the pCOG-SAX-RP-LC-MS/MS identified 1833 nonredundant phosphorylated sites. This method was further compared with SCX-TiO2-RP-LC-MS/MS mode.

Experimental Section

Chemicals. All the water was prepared using a Milli-Q system (Millipore, Bedford, MA); dithiothreitol (DTT), ammonium bicarbonate and iodoacetamide (IAA) were purchased from Bio-Rad (Hercules, CA); citric acid, urea and sodium chloride (NaCl), were obtained from Sigma (St. Louis, MO); trypsin was purchased from Promega (Madison, WI); and formic acid (FA) and acetonitrile (ACN) were obtained from Aldrich (Milwaukee, WI). Sodium orthovanadate (Na_3VO_4) and sodium fluoride (NaF) were from Sigma (St. Louis, MO); ethylene diamine tetraacetic acid (EDTA), ethylene glycol-bis-[2-aminoethyl ether]-*N*,*N*,*N*′,*N*′ tetraaceticacid (EGTA) and phenyl methyl sulfonyl fluoride (PMSF) were purchased from Amresco (Solon, OH). All the chemicals were of analytical grade except acetonitrile, which was HPLC grade.

Trypsin Digestion of Standard Protein Mixture. Seven proteins (fetal calf fetuin, bovine β-casein, horse cytochrome C, chicken egg lysozyme, bovine serum albumin, human hemoglobin and horse heart myoglobin) were mixed with equal amounts. A total of 500 μg of these standard proteins were diluted to 200 μL with 50 mM ammonium bicarbonate. After that, 2 μL of 1 M DTT was added in the mixture. The mixture was incubated at 37 °C for 150 min and then 10 μL of 1 M IAA was added and incubated for an additional 40 min at room temperature in darkness. The protein mixture was then precipitated overnight with 50% acetone, 50% ethanol, and 0.1% acetic acid. After centrifugation, the precipitate was redissolved with 50 mM ammonium bicarbonate buffer at pH 8.3, and incubated overnight with trypsin (25:1) at 37 °C. The tryptic peptide mixtures were collected and lyophilized. Finally, 6 synthesized peptides (IGEGTYGVVYK, IGEGTY*GVVYK, IGEGT*YGVVYK, EEVASEPEEAASPTTPK, EEVAS*EPEEAASPTTPK and EEVAS*EPEEAASPT*TPK, * indicates phosphorylation) were properly added to standard peptides mixture for additional analysis.

Trypsin Digestion of HeLa Cell Protein Mixture. Briefly, 500 μg of HeLa cell proteins was diluted to 200 μL with 50 mM ammonium bicarbonate. After reduction and alkylation of cysteines by DTT and IAA, the protein mixture was then precipitated overnight. After centrifugation, the precipitate was redissolved with 50 mM ammonium bicarbonate and digested with trypsin (25:1) at 37 °C overnight. The tryptic peptide mixtures were collected and lyophilized for additional analysis.

Continuous pH Elution SAX-RP-LC-MS/MS Analysis. A Surveyor liquid chromatography system (ThermoFisher, San Jose, CA) consisting of a degasser, 2 MS Pumps and an autosampler was applied. The following separation conditions and columns were utilized: (1) a SAX column (10 μm, 320 μm $\times$ 100 mm, Column Technology Inc., CA); (2) two C18 trap columns (5 μm, 300 μm $\times$ 5 mm, Agilent Technologies); (3) an analytical C18 column (5 μm, RP, 75 μm $\times$ 150 mm, Column Technology Inc., CA) and (4) 15 μm (internal diameter) non-coated SilicaTip PicoTip nanospray emitter (New Objective, Woburn, MA).

The solvents for standard proteins were 3 different pH solutions: A, pH 9.0 (NH_4OH and FA); B:, pH 2.5 (2 mM Citric acid adjusted by formic acid); and C, pH2.0 (2 mM Citric acid adjusted by formic acid). The solvents for HeLa cell were A, pH 8.5 (5 mM Citric acid adjusted by NH_4OH); B, pH 2.5 (5

mM Citric acid adjusted by formic acid); and C, pH2.0 (5 mM Citric acid adjusted by formic acid).

One picomole standard peptide mixture was dissolved in 80 μL of pH 9.0 buffer A, and then loaded into SAX column at a 3 μL/min flow rate after the split. The pH continuous gradient was constituted by mixture of buffer A (pH 9.0) and buffer B (pH 2.5). Then, the final gradient was a mixture of buffer B (pH 2.5) and buffer C (pH 2.0) and maintained at 100% buffer C (Figure 1, inset). When trap column 1 was performed by reversed-phase gradient, trap column 2 was loaded by peptides eluted from SAX column by the following pH gradient, and vice versa.[23] After each experiment, the strong anion exchange column was re-equilibrated by using pH 8.5 buffer.

The 500 μg HeLa cell peptide mixtures were dissolved in 80 μL of pH 8.5 buffer A (5 mM Citric acid adjusted by NH_4OH) containing 20% acetonitrile, and then loaded into the SAX column. The peptides bound in SAX column were then separated by pH continuous gradient with buffer A (pH 8.5) and buffer B (pH 2.5), and the last two gradients were formed by mixing Buffer B (pH 2.5) with Buffer C (pH 2.0) and maintained at 100% Buffer C (Figure 1). The HPLC solvents for reversed-phase were 0.1% formic acid (v/v) aqueous (A) and 0.1% formic acid (v/v) acetonitrile (B). The RP gradient ran to 35% mobile phase B in 95 min for 7 standard protein mixtures and 165 min for HeLa cells. The flow rate was 140 μL/min before the splitting and 250 nL/min after the splitting. The entire chromatography process was fully automatic and controlled by program without any manual attendance.

pH Step Elution SAX-RP-LC-MS/MS Analysis. The pH step elution procedure on SAX column was based on our previous work.[20] Briefly, 1 pmol standard peptide mixture was dissolved in 80 μL of pH 9.0 buffer A (NH_4OH and FA), and then loaded into SAX column at a 3 μL/min flow rate after the split. The trap column 1 was loaded by flow through peptides mixture from SAX and set for the reversed-phase HPLC and MS/MS analysis. Then 80 μL of pH 4.0 buffers were injected from the autosampler with buffer A to SAX to elute peptides into trap column 2. When trap column 2 was running the reversed-phase gradient, 80 μL of pH 3.06 buffers was injected to elute peptides from SAX into trap column 1. Finally, the remaining bound peptides of SAX were eluted with 80 μL of pH 2.0 buffer into trap column 2, and prepared for additional reversed-phase analysis. The reversed-phase HPLC and mass spectrometry conditions were the same as the standard peptides mixture identification with continuous pH elution.

SCX-TiO_2-RP-LC-MS/MS Analysis. The SCX and TiO_2 separation was according to Beausoleil et al.[17] Briefly, 500 μg of HeLa cell peptide mixtures was redissolved in 20 μL of SCX buffer A (5 mM KH_2PO_4, pH 2.65/30% acetonitrile). Strong cation exchange chromatography was carried out on a SCX column (320 μm $\times$ 100 mm, Column Technology, Inc., CA) at a flow rate of 6 μL/min. After a 4 min flow through by buffer A (5 mM KH_2PO_4, pH 2.65/30% acetonitrile), a linear gradient from 0% to 21% buffer B (5 mM KH_2PO_4, pH 2.65/30% acetonitrile/350 mM KCl) over 32 min followed. Then, several washing steps with 100% buffer B lasted for 4 min. Finally, we collected 10 fractions (4-min intervals) and lyophilized them for additional offline TiO_2 (320 μm $\times$ 50 mm, Column Technology, Inc., CA) enrichment of phosphopeptide. Each fraction from SCX was dissolved with 5% formic acid and 20% acetonitrile and loaded into TiO_2, and then the bound peptides, including phosphopeptides in TiO_2 column, were eluted by pH 10.5 buffer with 20% acetonitrile. After being collected and lyophilized, the eluted peptide mixtures were prepared for reversed-phase HPLC and MS/MS analysis. The reversed-phase HPLC and mass spectrometry conditions were set as the continuous pH elution SAX-RP-LC-MS/MS.

Mass Spectrometry Identification. A linear ion trap/Orbitrap (LTQ-Orbitrap) hybrid mass spectrometer (ThermoFisher, San Jose, CA) equipped with an NSI nanospray source was operated in data dependent mode to automatically switch between in MS and MS/MS acquisition with ion transfer capillary of 200 °C and NSI voltage of 1.85 kV. Normalized collision energy was 35.0. The mass spectrometer was set so that one full MS scan was acquired in the Orbitrap parallel to (or following) 10 MS^2 scans in the linear ion trap on the 10 most intense ions from the full MS spectrum with the following Dynamic Exclusion settings: repeat count 2, repeat duration 30 s, exclusion duration 90 s. The resolving power of the Orbitrap mass analyzer was set at 100 000 ($m/\Delta m$ 50% at m/z 400) for the precursor ion scans. The MS^2 spectra containing neutral loss ions related to phosphorylation were used to do the pseudo MS^3.[24]

DATA Analysis and Validation. The acquired MS^2 and pseudo MS^3 data were automatically combined in a single spectrum during data collection and produced dta based on Bioworks 3.2. The dta files were searched against the Human International Protein Index protein sequence database containing 68 020 protein sequences downloaded as FASTA-formatted sequences from EBI-IPI (version 3.28, ftp://ftp.ebi.ac.uk/IPI). To calculate confidence levels and false positive rates, a decoy database containing the reverse sequences of the 68 020 proteins was appended to the original human database, and the SEQUEST program in Bioworks 3.2 was used to find the matching sequences from the combined database on a computer cluster running under the Linux operating system.

Database searches were performed by applying the precursor ion m/z tolerances of $\pm$500 ppm and fragment ion m/z tolerances of $\pm$1 Da. Cysteine residues were statically modified by 57.02146 Da due to carboxyamidomethylation of the sample. Dynamic modifications were permitted to allow for the detection of oxidized methoinine residues (+15.99492), and phosphorylated serine, threonine and tyrosine residues (+79.96633). The maximum number of internal cleavage site was set to 2. All output results were filtered and combined together using in-house software named BuildSummary to delete the redundant data. The peptides identified by Orbitrap were filtered by precursor ion m/$z \leq$ 10 ppm. When deltaCN of peptides $\geq$0.1, only peptides with Xcorr 1+ Charged $\geq$1.8, 2+ Charged $\geq$2.0, 3+ Charged $\geq$2.5, remained. The final FDR were 0.89%.[25] In the BuildSummary tool, if a scan was matched to phosphopeptides with the same peptide and same number of phosphorylated sites, but with different positions of phosphorylated sites, the positions were considered "ambiguous". If the same peptide(s) were assigned to multiple proteins, the multiple proteins were clustered to a "protein group" and only one representative protein was selected to count the identified number.

Results and Discussion

Construction of a Fully Automatic 2D-LC-MS/MS Based on pH Continuous Online Gradient Strong Anion Exchange (pCOG-SAX). Current 2D-LC-MS/MS methods have been developed based mainly on salt-eluting strong cation exchange (SCX) coupled with reversed-phase liquid chromatography mass spectrometry (RP-LC-MS/MS). There are two

main modes, offline and online. In both modes, the salt removal is a major concern since the salt cannot be directly connected to the mass spectrometry. In the offline mode, continuous salt gradient was applied in SCX column and fractions were collected according to the salt gradient ranges, followed by further desalting before secondary loading on RP-LC-MS/MS.[26] In the online mode, there are still two kinds of systems. The first was MudPit system with one LC pump and using rapid salt step elution from the sampler.[27] Our previous work applied alternative pH step elution to form an integrated multiple dimensional liquid chromatography mass spectrometry (IMDL-MS/MS), also based on one LC pump and a biphasic column integrating SCX and RP materials.[28] The second system also used salt step elution while composed of two pumps and two trap columns in a 10-port valve.[23,29,30] This system allowed salt to be removed completely by using trap columns in order not to interfere with the mass spectrometric detection. Theoretically, the continuous gradient on SCX can be implemented based on a two-pump configuration since one pump can run gradient in SCX while the other runs RP gradient. However, this mode is impractical to couple online with MS, since the salt gradient will introduce a large amount of salt in the LC-MS/MS system and crush the electrospray. Recently, we implemented the continuous online gradient in the SCX-RP-LC-MS/MS based on pH elution in the two-pump system,[23] which showed less peptide cross-overlap among fractions and better sensitivity than salt step online elution. The pH buffer instead of salt elution in ion exchange was not only effective for elution of peptides but also completely compatible with mass spectrometry. On the other hand, the pH step elution in our recent Yin-yang MDLC-MS/MS had shown that phosphopeptides could be separated from nonphosphorylated peptides in strong anion exchange (SAX). In this work, a fully automatic 2D-LC-MS/MS applying pH continuous online gradient strong anion exchange (pCOG-SAX) was successfully constructed. Figure 1 shows the configuration of this system. There were two liquid chromatography pumps, one strong anion exchange (SAX) column and two reversed-phase (RP) C18 trap columns on a 10-port valve, and one RP C18 nanoflow column and nanospray tips connected to mass spectrometry. The first pump ran pH continuous gradient and the second pump ran reversed-phase gradient. The pH gradient was from basic to acidic as a continuous mode, connected to a SAX column to elute peptides continuously. When peptides were loaded at basic pH, the unbound peptides flowed through the SAX and were bounded by the trap column 1. Then, SAX was switched to trap column 2, and the pH gradient was started to the targeted value to elute the peptides to trap column 2. In the meanwhile, the peptides in the trap column 1 were eluted with acetonitrile gradient and connected to separation reversed-phase column and mass spectrometry. Accordingly, when trap column 1 was loaded by peptides eluted from SAX column, trap column 2 was eluted by reversed-phase gradient and connected to mass spectrometry. The procedures were switched between the SAX and two trap columns to form fractions according to pH gradient ranges. NH_4OH, citric acid and formic acid were used to form pH gradient from basic to acidic. Herein, the binding of peptides was weakened by decrease of pH value, not by increase of salt concentration. Thus, only a very low concentration of pH buffer (5 mM citric acid) was used. Hence, no further desalting step was needed in the traps and separating reversed-phase column, even after continuous elution of buffer for 140 min. As a result, the pH continuous gradient can be online coupled with further reversed-phase LC-MS/MS. For standard protein mixture, four fractions were obtained by pH ranges from 9.0 to 2.0 (Figure 1 inset). The acidic pH was finally down to 2.0 in order to probe whether extremely acidic peptides were bound in the SAX column. For complex protein mixture, 10 fractions were split and optimized according to the pH curve as shown in Figure 1. The loading pH was 8.5 in order to remove more nonphosphopeptides at the loading step.

Binding Features of Nonphosphorylated and Phosphorylated Peptides of pCOG-SAX-RP-LC-MS/MS. For peptide mixture, the signals of phosphopeptides are generally suppressed by nonphosphopeptides due to the negative feature of phosphate group. Therefore, the isolation and enrichment of phosphopeptides from nonphosphopeptides are of vital importance for detectability of phosphopeptides. To evaluate the performance of pCOG-SAX-RP-LC-MS/MS, we mixed 7 standard proteins (see Experimental Section) and digested them with trypsin. The other 6 synthesized peptides, IGEGTYGVVYK, IGEGT*YGVVYK, IGEGTY*GVVYK, EEVASEPEEAASPTTPK, EEVAS*EPEEAASPTTPK, and EEVASEPEEAASPT*TPK (* indicates phosphorylated site) were mixed with the protein digests (Table S-1 in Supporting Information). All of the 6 synthesized peptides, including 4 phosphopetides, were identified. In addition, the monophosphopeptide and its nonphosphorylated isoform of β-casein were also identified. Figure 2a shows the total ion chromatogram of four fractions in this pCOG-SAX-RP-LC-MS/MS for standard protein mixture and synthesized peptides. The fraction 1 was flow-through peptides at the loading condition at pH 9.0. None of the phosphopeptides were eluted in this fraction, while most of the nonphosphorylated peptides were obtained here. A synthesized nonphosphorylated peptide IGEGTYGVVYK was found in this flow-through fraction, which has p*I* at 6.0. However, its phosphopeptide isoforms, IGEGT*YGVVYK and IGEGTY*GVVYK, were bound in the SAX column, and eluted in fraction 2 when pH was down to 4.0. Though the nonphosphorylated peptide cannot be bound by the SAX column, the phosphorylation rendered the peptide acidic features and made it bound by the SAX. In fraction 3, with gradient from pH 4.0 to 3.1, a singly phosporylated peptide, EEVAS*EPEEAASPTTPK, as well as two acidic nonphosphorylated peptides, EEVASEPEEAASPTTPK (pI 3.98) and FQSEEQQQTEDELQDK (p*I* 3.77), were eluted. In the final fraction with pH 3.1 to 2.0, a doubly phosphorylated peptide, EEVAS*EPEEAASPT*TPK, and the monophosphopeptide of β-casein, FQS*EEQQQTEDELQDK, were both found. These observations showed that if a peptide is acidic itself, it would be easily bound by the SAX, while its phosphorylated form would be bound more tightly and eluted in more acidic pH than the nonphosphorylated form. Also, the multiple-phophorylated peptides would be eluted in more acidic condition, compared with singly phosphorylated peptides. Though EEVAS*EPEEAASPTTPK and EEVASEPEEAASPTTPK were both found in fraction 3 since only four fractions were split here, it is believed that they can be resolved if more fractions were split.

Continuous Gradient Elution Had Better Repeatability and Less Peptide Cross-Overlap than Step Gradient Elution. To compare the performance of continuous gradient and step gradient, pH step online gradient coupled with SAX-RP-LC-MS/MS (pSOG-SAX-RP-LC-MS/MS) was also applied. As Figure 2b shows, most of the nonphosphorylated peptides came out in flow-through and early basic fractions, and phosphopeptides mostly appeared in acidic condition, consistent with our previous report using pH step offline elution.[20] Figure 2c,d

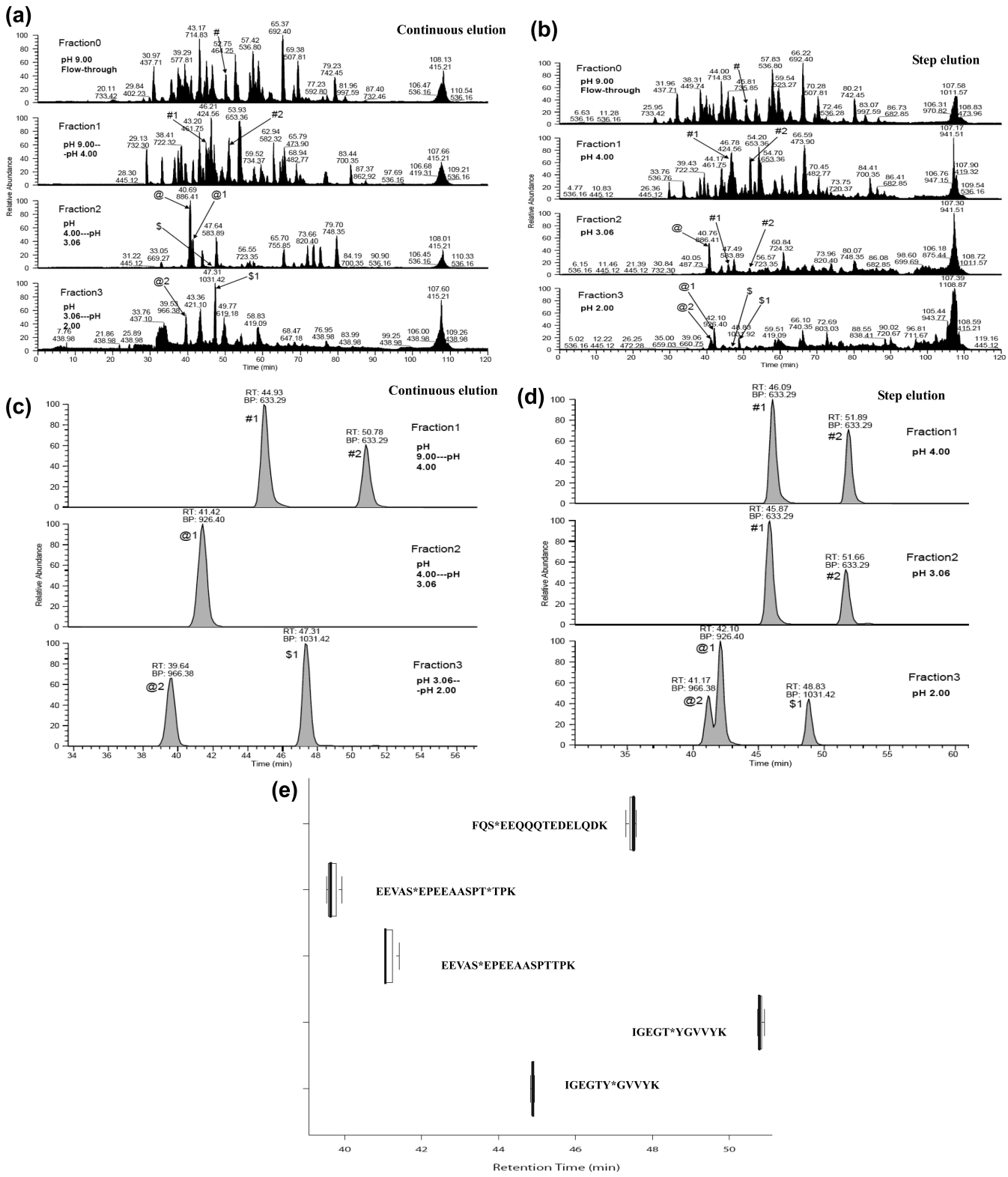

Figure 2. Total Ion Chromatograms (TIC) of four fractions for standard protein mixture spiked with synthesized peptides in pH continuous gradient elution (a) and pH step gradient elution (b). Base peaks of phosphopeptides in pH continuous gradient elution (c) and pH step gradient elution (d). Repeatability of retention times of five synthesized phosphopeptides (e). Notes: #, IGEGTYGVVYK; #1, IGEGTY*GVVYK; #2, IGEGT*YGVVYK; @, EEVASEPEEAASPTTPK; @1, EEVAS*EPEEAASPTTPK; @2, EEVAS*EPEEAASPT*TPK; $, FQSEEQQQTEDELQDK; $1, FQS*EEQQQTEDELQDK.

shows the base peaks of all five known phosphopeptides in two elution modes. In pCOG-SAX-RP-LC-MS/MS, all phosphopeptides were eluted in single fraction as a single peak. However, in pSOG-SAX-RP-LC-MS/MS, some of the nonphsophorylated peptides as well as phosphopeptides could be detected in multiple fractions. This phenomenon was called peptide cross-overlap, which occurred frequently in step elution. For example, in the pSOG mode, IGEGT*YGVVYK and

Table 1. Distributions of Phosphopeptides and Nonphosphopeptides in Each Fraction of pCOG-SAX-RP-LC-MS/MS, Split by pH Ranges

fraction	pH range	unique phosphopeptides	phosphopeptide hits	unique nonphosphopeptides	nonphosphopeptides hits
F0	8.50	15	15	1866	6183
F1	8.50–6.01	92	139	1230	3109
F2	6.01–4.00	297	571	691	2690
F3	4.00–3.65	368	854	550	2820
F4	3.65–3.40	309	960	405	1276
F5	3.40–3.12	254	744	427	1174
F6	3.12–2.80	172	479	324	912
F7	2.80–2.50	131	366	380	1001
F8	2.50–2.25	123	464	486	1209
F9	2.25–2.00	63	436	484	1133
Total		1561	5038	4087	21507

IGEGTY*GVVYK were detected in both fraction 2 and fraction 3 (Figure 2d), while in pCOG mode, they were only detected in fraction 2 (Figure 2c). Figure S1 in Supporting Information shows that the repeatability of the pCOG-SAX-RP-LC-MS/MS method is much better than the pH step elution. In the step elution, the elution buffers were injected from autosampler generally less than 80 μL. The rapid elution of peptides with such a small volume of buffers makes it difficult to wash all peptides completely. Compared with step elution, the continuous gradient can elute peptides completely with several hundred microliters of buffer during 140 min pH gradient. This elution makes peptides less diffuse among the fractions, thus, enhancing the sensitivity and repeatability of identification. In pCOG-SAX-RP-LC-MS/MS, the phosphopeptides were identified reproducibly according to the elution time (Table S-1 in Supporting Information). The variations of elution time of phosphopeptides among three duplicates were within 0.21 min (Figure 2e).

SAX-RP-LC-MS/MS for Phosphopeptides from Complex Proteome. To verify the performance of the above platform in a complex system, the protein tryptic digests from human HeLa cells were analyzed based on the pCOG-SAX-RP-LC-MS/MS coupled with LTQ-Orbitrap mass spectrometry. To reach high resolution of the separation, a total of 10 fractions were split by the pH ranges from pH 8.5 to 2.0 in SAX. Finally, 1561 unique phosphorylated peptides and 4087 unique nonphosphopeptides were identified. The ratio of phosphopeptide versus total peptides was 27.46%. Among the phosphopeptides, 763 were singly phosphorylated, 538 were doubly phosphorylated and 260 were triply phosphorylated. Accordingly, 1833 phosphorylation sites were identified, including 1497 distinct and 336 ambiguous sites. The phosphopeptides were listed in Table S-2 in Supporting Information. Of all 1497 distinct phosphorylation sites, 1221 (81.6%) were phosphorylated serine (Ser), 231 (15.4%) were phosphorylated threonine (Thr), and 45 (3.0%) were phosphorylated tyrosine (Tyr). Table 1 shows the distributions of phosphopeptides and nonphosphopeptides among fractions in pCOG-SAX-RP-LC-MS/MS method. Although 4530 nonphosphopeptides were detected, most of their hits (68.83%) were detected in the first 4 fractions (F0−F3) from pH 8.5 to 3.65. Below pH 3.65, the number of nonphosphopeptides decreased dramatically. On the other hand, up to 68.60% phosphopeptide hits were found in fractions with pH under 3.65 (F4−F9). Figure 3a presents the ratio of phosphopeptide hits (vs total phosphopeptide hits) and nonphosphopeptide hits (vs total nonphosphopeptide hits) in each fraction. The tendencies of phosphopeptides and nonphosphopeptides were reverse. The crossing point was at F2 when pH covered 6.01−4.0. These different tendencies facilitated the natural isolation of phosphopeptides from nonphosphopeptides in a single anion exchange phase. Figure 3b shows the distribution of singly and multiply phosphorylated peptides in SAX fractions. The acidic pH elution obviously increased the detectability of multiply phosphorylated peptides. Though not all nonphosphopeptides were eluted in flow-through fraction, the bound nonphosphopeptides could be washed in early fractions and phosphopeptides prefered to be eluted in later fractions. In this pCOG-SAX-RP-LC-MS/MS work, 237 groups of peptide isoforms with common peptide sequences but different numbers of phosphorylated sites were detected (as shown in Table S-3 in Supporting Information). In the 237 groups of peptide isoforms, almost all phosphopeptides were eluted in more

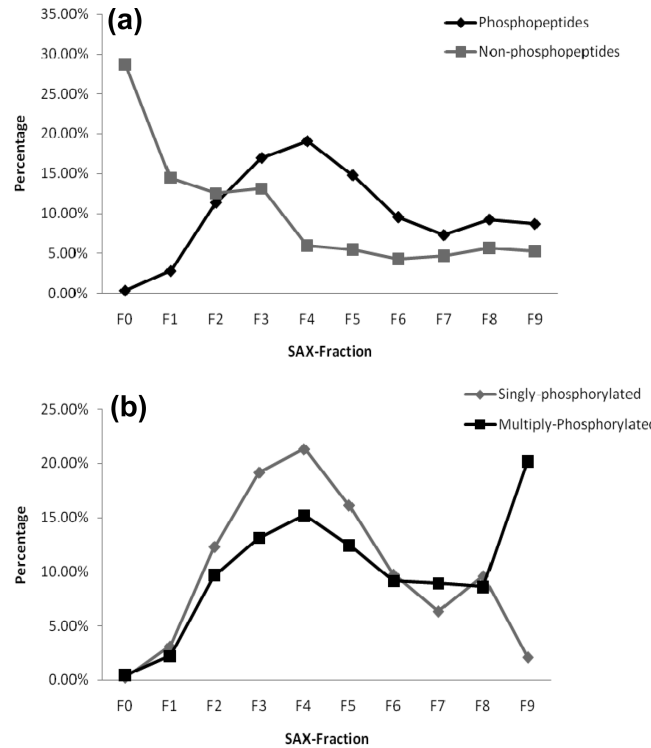

Figure 3. Ratios of phosphopeptide hits (vs total phosphopeptide hits) and nonphosphopeptide hits (vs total nonphosphopeptide hits) in fractions by pCOG-SAX-RP-LC-MS/MS for Hela cells (a). Distributions of singly phosphorylated and multiply phosphorylated peptides among fractions by pCOG-SAX-RP-LC-MS/MS (b).

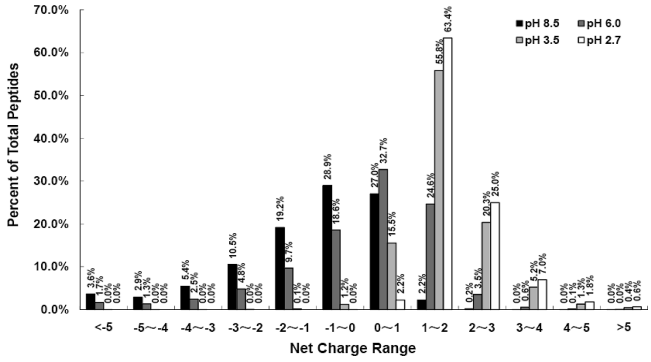

Figure 4. Net charge distribution of tryptic peptides theoretically digested from proteins in human IPI database under different pH environments (pH 8.5, 6.0, 3.5 and 2.7).

acidic pH fractions than their nonphosphopeptides counterparts. Peptides, with more phosphorylated sites were found in more acidic fractions than their less phosphorylated isoforms. For example, peptide LFDEEEDSSEKLFDDSDER has three Ser residues, the monophosphopeptide was in fraction 4 (pH 3.65−3.4), diphosphopeptide was in fraction 7 (pH 2.8−2.5) and triphosphopeptide was in fraction 9 (pH 2.25−2.0). Only in a few cases, the nonphosphopeptides were detected in multiple fractions. These results might be caused by the strong binding of high-abundance nonphosphopeptides in columns. The wide-coverage of complete series of peptide isoforms facilitated the comprehensive understanding of phosphorylation events.

Binding Features of SAX and SCX for Peptides. The performance of peptides in ion exchange column was based on net charge features. The net charge of a peptide can be calculated by Bjellqvist algorithm under different pH environments.[31] Figure 4 shows the net charge distribution of theoretical tryptic peptides in the human IPI database (IPI 3.28), under different pH environments. To simplify the data, no miss cleavage was allowed in the theoretical tryptic digestion. At pH 8.5 for sample loading in SAX, 29.2% theoretical peptides have positive net charge between 0 and ∼+2, which will not be bound in the SAX column, while about 71% peptides with negative charges will be bound in the SAX column. Since a single phosphorylation will deliver 2 negative charges to peptides at pH 8.5, almost all of the phosphopeptides will remain at pH 8.5 in SAX column. This process can deplete a part of nonphosphopeptides as well as decrease the sample complexity, just in the first loading step.

At the pH 6.0, the phosphate group will carry −1 charge, and the phosphopeptide binding will be weakened, especially for the singly phosphorylated peptides without other residual COOH. When at pH 3.5, 98.7% nonphosphopeptides are positive charged and eluted out, except for some phosphopeptides still in the SAX column. When at pH 2.0, the phosphate group will also lose its negative charge and all of the phosphopeptides should be washed out. Therefore, the elution of peptides from basic pH to acidic pH isolated the phosphopeptides from the nonphosphopeptides. For SAX mode, salt elution could be also used for phosphopeptides.[22] However, as indicated by the net charge distribution of theoretical peptides in Figure 4, approximately 70% and 40% peptides will be negative charged at pH 8.5 and pH 6.0. The phosphorylation will deliver more negative charges, resulting in very strong binding in the SAX column, especially for multiply phosphorylated peptides. The way to weaken their binding is to decrease the negative charges by lowering the pH value of the elution buffer.

Beausoleil et al. reported the technology using static pH at 2.7 while changing the salt concentration in SCX for phosphopeptide separation.[16,17] However, the authors also mentioned the strong cation exchange only obtained a part of phosphopeptides, since the phosphorylation will make some peptides too acidic to be bound by the SCX in acidic pH range. The following work further proved that the acidic peptides would flow through in the SCX.[32] At pH 2.7, almost all of the nonphosphorylated peptides are positive charged, and should theoretically be bound in SCX. However, most nonphophorylated peptides have net charge ranging from 0 to ∼+3. Therefore, one or two phosphorylation will make these peptides less positively charged, even zero or negative charged. The zero and negative charged phosphopeptides cannot be bound by SCX and will be lost at loading step. The less positive phosphopeptides will be still mixed with the nonphosphorylated peptides in SCX separation. Thus, the SCX-based approaches for phosphoprteomics generally combined three kinds of separation phases, such as SCX-IMAC-RP and SCX-TiO_2-RP. In this present work, the different elution features of nonphosphopeptide and phosphopeptide in SAX largely reduce the coelution issue. Thus, combined with the secondary fractionation by reversed-phase, this system is sufficient to isolate phosphopeptides from nonphosphopeptides. This two-phase separation based on SAX and RP combination is ready to be combined and online coupled with mass spectrometry. Furthermore, the pH elution, instead of a salt elution, facilitates the continuous online gradient, resulting in better repeatability and robustness.

pCOG-SAX-RP-LC-MS/MS Is Complementary to SCX-TiO_2-RP-LC-MS/MS. As mentioned above, SCX was generally used as fractionation tool to reduce the peptide complexity, and further affinity enrichment of TiO_2 was combined. Here, SCX was applied as offline elution by salt gradient, at static acidic condition as pH 2.65, and then each fraction was collected and conducted further offline TiO_2 enrichment of phosphopeptides. Each eluted fraction was analyzed by RP-LC-MS/MS (see Experimental Section). This procedure combined three separation phases including two offline and one online reversed phase coupled with mass spectrometry. By this SCX-TiO_2-RP-LC-MS/MS mode for protein digests from HeLa cells, 701 unique phosphorylated peptides and 722 nonredundant sites were identified, in which 475 were singly phosphorylated, 176 were doubly phosphorylated and 50 were triply phosphorylated. Figure 5a demonstrates the distribution of phosphopeptide with different numbers of phosphorylated sites in pCOG-SAX-RP-LC-MS/MS and SCX-TiO_2-RP-LC-MS/MS methods. Generally, pCOG-SAX-RP-LC-MS/MS method identified more multiple phosphorylated peptides than SCX-TiO_2-RP-LC-MS/MS, which may be contributed by the acidic elution down to pH 2.0. Figure 5b shows the theoretical p*I* distribution of identified phosphorylated peptides (calculated for the original peptide backbone), in both methods. It is interesting that these two methods are somehow complementary. The pCOG-SAX-RP-LC-MS/MS detected more acidic peptides and SCX-TiO_2-RP-LC-MS/MS identified more basic peptides. The peptide p*I* distribution in our SCX-TiO_2-RP-LC-MS/MS method was consistent with previous work,[19] which confirmed the complementary features of peptide identification by these two methods. Finally, phosphorylated events occurring on 57 overlapped peptides were identified by both platforms. Table S-4 in Supporting Information presents the overlapped peptides and phosphorylated states identified in each platform. For 26

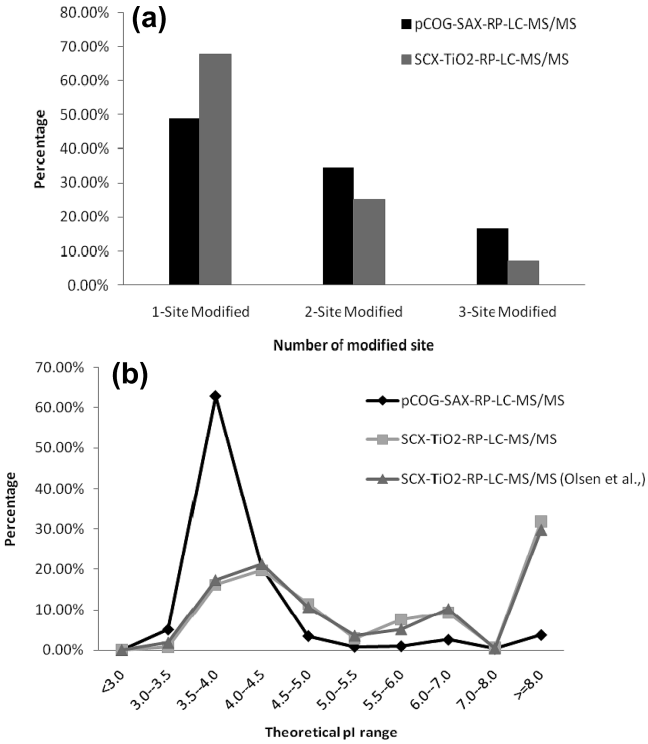

Figure 5. Percentage of phosphopeptides with different number of phosphorylated sites in pCOG-SAX-RP-LC-MS/MS and SCX-TiO_2-RP-LC-MS/MS methods for HeLa cells (a). Distribution of theoretical p*I* range of peptides found with phosphorylation by pCOG-SAX-RP-LC-MS/MS, this present SCX-TiO_2-RP-LC-MS/MS and reported work based on SCX-TiO_2-RP-LC-MS/MS[19] (b).

peptides, these two methods both covered identical phosphorylated isoforms, and in another 12 peptides, pCOG-SAX-RP-LC-MS/MS method identified the high-state phosphorylation and SCX-TiO_2-RP-LC-MS/MS obtained low-state phosphorylation. For the other 18 peptides, SCX-TiO_2-RP-LC-MS/MS only identified the low-state phosphorylation, while pCOG-SAX-RP-LC-MS/MS detected high-state as well as low-state phosphorylation, obtaining as a result a more complete series of phosphopeptides than SCX-TiO_2-RP-LC-MS/MS. For example, HIVSNDSSDSDDESHEPK has five potential phosphorylated Ser residues. Three kinds of phosphopeptides, singly, doubly and triply phosphorylated, were all detected and identified in pCOG-SAX-RP-LC-MS/MS. In SCX-TiO_2-RP-LC-MS/MS, only singly and doubly phosphorylated peptides were identified. This phosphopeptide was also found in reported work either using TiO_2 or IMAC enrichment. In previous work for HeLa cell,[19] only singly phosphorylated peptides were detected by SCX-TiO_2-based enrichment. Zahedi et al. found the triply phosphopeptides in human platelets using the IMAC based enrichment, but did not find singly and doubly phosphorylated peptides.[33] Previous comparisons indicated that TiO_2 preferred to enrich more singly phosphoryalted peptides and IMAC had better performance on enrichment of multiply phosphorylated peptides.[34] Herein, compared with affinity enrichment, the pCOG-SAX-RP-LC-MS/MS showed less bias detection and obtained a more complete series of phosphorylated states of peptides. On the other hand, the bottleneck for identification of phosphopeptides is the complexity of the sample. Thus, subcellular fractionation and more replicates will be helpful to identify more phosphopeptides.[19]

Conclusion

In this work, stong anion exhcnage and reversed-phase chromatography were online connected with mass spectrometry. As we know, it is a novel system using one-step loading and two-phase separations, which achieves robust separation and detection of phosphopeptides in complex cell proteome. The high-coverage detection of peptide phosphorylated isoforms makes it very promising in the quantitative analysis of the dynamics of protein phosphorylation in cell signaling. The combination of this method with SCX-TiO_2-RP-LC-MS/MS might give the most comprehensive picture of phosphopeptides. In summary, this fully automatic separation and identification of phosphopeptides provides a robust approach for large-scale phosphoproteomics.

Abbreviations: pCOG, pH continuous online gradient; SAX, strong anion exchange; SCX, strong cation exchange; IMAC, immobilized metal affinity chromatography; LC, liquid chromatography; RP-LC, reversed-phase liquid chromatography; ESI, electrospray ionization; MS/MS, tandem mass spectrometry; TiO_2, titanium dioxide; pSOG, pH step online gradient.

Acknowledgment. This work was supported by the National Natural Science Foundation (30425021, 30521005), Basic Research Foundation (2006CB910700), CAS Project (KSCX2-YW-R-106), High-technology Project (2007AA02Z334) and Proteomage Project of EU.

Supporting Information Available: Identification and reproducibility of 5 known phosphopeptides in protein mixture; the base peaks of fraction 2 of three replicates by pH continuous elution and pH step elution; complete data set of phosphopeptides identified by SAX method, and pCOG-SAX-RP-LC-MS/MS method, and complete data set of overlaped phosphopeptides identified in this study. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Cohen, P. The regulation of protein function by multisite phosphorylation--a 25 year update. *Trends Biochem. Sci.* **2000**, *25*, 596–601.
(2) Pawson, T.; Scott, J. D. Protein phosphorylation in signaling--50 years and counting. *Trends Biochem. Sci.* **2005**, *30*, 286–290.
(3) Cohen, P. The origins of protein phosphorylation. *Nat. Cell Biol.* **2002**, *4*, E127–130.
(4) Schmelzle, K.; White, F. M. Phosphoproteomic approaches to elucidate cellular signaling networks. *Curr. Opin. Biotechnol.* **2006**, *17*, 406–414.
(5) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The protein kinase complement of the human genome. *Science* **2002**, *298*, 1912–1934.
(6) Caenepeel, S.; Charydczak, G.; Sudarsanam, S.; Hunter, T.; Manning, G. The mouse kinome: discovery and comparative genomics of all mouse protein kinases. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 11707–11712.
(7) Forrest, A. R.; Ravasi, T.; Taylor, D.; Huber, T.; Hume, D. A.; et al. Phosphoregulators: protein kinases and protein phosphatases of mouse. *Genome Res.* **2003**, *13*, 1443–1454.
(8) Gronborg, M.; Kristiansen, T. Z.; Stensballe, A.; Andersen, J. S.; Ohara, O.; Mann, M.; Jensen, O. N.; Pandey, A. A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phospho-specific antibodies: identification of a novel protein, Frigg, as a protein kinase substrate. *Mol. Cell. Proteomics* **2002**, *1*, 517–527.
(9) Rush, J.; Moritz, A.; Lee, K. A.; Guo, A.; Goss, V. L.; Spek, E. J.; Zhang, H.; Zha, X. M.; Polakiewicz, R. D.; Comb, M. J. Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat. Biotechnol.* **2005**, *23*, 94–101.
(10) Nuhse, T. S.; Peck, S. C. Peptide-based phosphoproteomics with immobilized metal ion chromatography. *Methods Mol. Biol.* **2006**, *323*, 431–436.

(11) Ndassa, Y. M.; Orsi, C.; Marto, J. A.; Chen, S.; Ross, M. M. Improved immobilized metal affinity chromatography for large-scale phosphoproteomics applications. *J. Proteome Res.* **2006**, *5*, 2789–2799.
(12) Haydon, C. E.; Eyers, P. A.; Aveline-Wolf, L. D.; Resing, K. A.; Maller, J. L.; Ahn, N. G. Identification of novel phosphorylation sites on Xenopus laevis Aurora A and analysis of phosphopeptide enrichment by immobilized metal-affinity chromatography. *Mol. Cell. Proteomics* **2003**, *2*, 1055–1067.
(13) Pinkse, M. W.; Uitto, P. M.; Hilhorst, M. J.; Ooms, B.; Heck, A. J. 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.
(14) 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*, 873–886.
(15) Pocsfalvi, G.; Cuccurullo, M.; Schlosser, G.; Scacco, S.; Papa, S.; Malorni, A. Phosphorylation of B14.5a subunit from bovine heart complex I identified by titanium dioxide selective enrichment and shotgun proteomics. *Mol. Cell. Proteomics* **2007**, *6*, 231–237.
(16) Ballif, B. A.; Villen, J.; Beausoleil, S. A.; Schwartz, D.; Gygi, S. P. Phosphoproteomic analysis of the developing mouse brain. *Mol. Cell. Proteomics* **2004**, *3*, 1093–1101.
(17) Beausoleil, S. A.; Jedrychowski, M.; Schwartz, D.; Elias, J. E.; Villen, 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.
(18) Trinidad, J. C.; Specht, C. G.; Thalhammer, A.; Schoepfer, R.; Burlingame, A. L. Comprehensive identification of phosphorylation sites in postsynaptic density preparations. *Mol. Cell. Proteomics* **2006**, *5*, 914–922.
(19) Olsen, J. V.; Blagoev, B.; Gnad, 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.
(20) Dai, J.; Jin, W. H.; Sheng, Q. H.; Shieh, C. H.; Wu, J. R.; Zeng, R. Protein phosphorylation and expression profiling by yin-yang multidimensional liquid chromatography (yin-yang MDLC) mass spectrometry. *J. Proteome Res.* **2007**, *6*, 250–262.
(21) Motoyama, A.; Xu, T.; Ruse, C. I.; Wohlschlegel, J. A.; Yates, J. R., III. Anion and cation mixed-bed ion exchange for enhanced multidimensional separations of peptides and phosphopeptides. *Anal. Chem.* **2007**, *79*, 3623–3634.
(22) Han, G.; Ye, M.; Zhou, H.; Jiang, X.; Feng, S.; Jiang, X.; Tian, R.; Wan, D.; Zou, H.; Gu, J. Large-scale phosphoproteome analysis of human liver tissue by enrichment and fractionation of phosphopeptides with strong anion exchange chromatography. *Proteomics* **2008**, *8*, 1346–1361.
(23) Zhou, H.; Dai, J.; Sheng, Q. H.; Li, R. X.; Shieh, C. H.; Guttman, A.; Zeng, R. A fully automated 2-D LC-MS method utilizing online continuous pH and RP gradients for global proteome analysis. *Electrophoresis* **2007**, *28*, 4311–4319.
(24) Schroeder, M. J.; Shabanowitz, J.; Schwartz, J. C.; Hunt, D. F.; Coon, J. J. A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry. *Anal. Chem.* **2004**, *76*, 3590–3598.
(25) Zhai, B.; Villén, J.; Beausoleil, S. A.; Mintseris, J.; Gygi, S. P. Phosphoproteome analysis of *Drosophila melanogaster* embryos. *J. Proteome Res.* **2008**, *7*, 1675–1682.
(26) Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J. Proteome Res.* **2003**, *2*, 43–50.
(27) Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R., III. Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **1999**, *17*, 676–682.
(28) Dai, J.; Shieh, C. H.; Sheng, Q. H.; Zhou, H.; Zeng, R. Proteomic analysis with integrated multiple dimensional liquid chromatography/mass spectrometry based on elution of ion exchange column using pH steps. *Anal. Chem.* **2005**, *77*, 5793–5799.
(29) Chelius, D.; Hühmer, A. F.; Shieh, C. H.; Lehmberg, E.; Traina, J. A.; Slattery, T. K.; Pungor, E., Jr. Analysis of the adenovirus type 5 proteome by liquid chromatography and tandem mass spectrometry methods. *J. Proteome Res.* **2002**, *1*, 501–513.
(30) Jiang, X. S.; Zhou, H.; Zhang, L.; Sheng, Q. H.; Li, S. J.; Li, L.; Hao, P.; Li, Y. X.; Xia, Q. C.; Wu, J. R.; Zeng, R. A high-throughput approach for subcellular proteome: identification of rat liver proteins using subcellular fractionation coupled with two-dimensional liquid chromatography tandem mass spectrometry and bioinformatic analysis. *Mol. Cell. Proteomics* **2004**, *3*, 441–455.
(31) Bjellqvist, B.; Hughes, G. J.; Pasquali, Ch.; Paquet, N.; Ravier, F.; Sanchez, J.-Ch.; Frutiger, S.; Hochstrasser, D. F. The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis* **1993**, *14*, 1023–1031.
(32) Lim, K. B.; Kassel, D. B. Phosphopeptides enrichment using on-line two-dimensional strong cation exchange followed by reversed-phase liquid chromatography/mass spectrometry. *Anal. Biochem.* **2006**, *354*, 213–219.
(33) Zahedi, R. P.; Lewandrowski, U.; Wiesner, J.; Wortelkamp, S.; Moebius, J.; Schütz, C.; Walter, U.; Gambaryan, S.; Sickmann, A. Phosphoproteome of resting human platelets. *J. Proteome Res.* **2008**, *7*, 526–534.
(34) Bodenmiller, B.; Mueller, L. N.; Mueller, M.; Domon, B.; Aebersold, R. Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nat. Methods* **2007**, *4*, 231–237.

PR800381W