Published in final edited form as:

J Chromatogr A. 2007 November 16; 1172(1): 9–18.

Phosphopeptide elution times in reversed-phase liquid chromatography

Jeongkwon Kim[†], Konstantinos Petritis, Yufeng Shen, David G. Camp II, Ronald J. Moore, and Richard D. Smith^{*}

Environmental Molecular Sciences Laboratory, MSIN K8-98, Pacific Northwest National Laboratory, P.O. Box 999, Richland, WA, 99352, USA

Abstract

Elution time shifts between 33 different peptides and their corresponding phosphopeptides ranging from 4 amino acid residues to 35 amino acids in length were systematically investigated using high-resolution reversed-phase liquid chromatography (RPLC)-tandem mass spectrometry (MS/MS) analysis with trifluoroacetic acid as the ion pairing agent. Observed peptide elution time shifts for a single phosphorylation ranged from −5.28 min (for pYVPML) to +0.59 min (for HRDpSGLLDLGR). Peptides containing a phosphotyrosine residue displayed a significant decrease in elution time following phosphorylation compared to their similar-sized peptides with phosphoserine or phosphothreonine residues. While peptide phosphorylation generally led to a decrease in the observed elution time, five peptides displayed increased elution times as a result of phosphorylation. For large peptides (= 18 amino acids), the elution time shifts due to single phosphorylation were limited (ranging between −0.48 min and +0.03 min), while the elution time shifts for small peptides (< 18 amino acids) were characterized by a larger deviation (ranging between −5.28 min and +0.59 min). The predictive capability for the observed RPLC elution time change due to phosphorylation has been suggested, which will aid in assigning confident phosphopeptide identifications and their subsequent confirmation.

Keywords

phosphopeptide; phosphorylation; elution time; reversed-phase liquid chromatography; normalization; GRAVY; prediction capability

1. INTRODUCTION

Information gained by the chromatographic elution time for peptides represents a powerful parameter for addressing both peptide identification and for improving the overall confidence of the identification. In capillary electrophoresis, peptide electrophoretic mobility with [1] and without [2,3] posttranslational modifications has been successfully predicted. Strong cation exchange chromatography was utilized to enrich phosphopeptides that eluted earlier than non-phosphorylated tryptic peptides since phosphopeptides generally tend to have lower net positive charges than non-phosphorylated peptides [4]. Extensive efforts to predict or utilize

[†]Current address: Department of Chemistry, Chungnam National University, Daejeon, 305-764, South Korea

^{*}Address correspondence to: Dr. Richard D. Smith Environmental Molecular Sciences Laboratory Pacific Northwest National Laboratory P.O. Box 999, MSIN: K8-98, Richland, WA, 99352 USA Fax: (509) 376-7722 E-mail: rds@pnl.gov

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

the elution time information of peptides in reversed-phase liquid chromatography (RPLC) have been made [5-13].

Protein phosphorylation plays an important role in cellular signal transductions [14] and subsequent regulatory mechanisms [15] and the combination of electrospray ionization and mass spectrometry have contributed significantly to the analysis of protein phosphorylation events through the widely used application of combined liquid chromatography/mass spectrometry [16-19]. Phosphopeptides are traditionally believed to elute earlier than their corresponding non-phosphopeptides under the conditions of RPLC because the incorporation of the hydrophilic phosphate group increases the hydrophilicity of a peptide. Indeed, it has been observed that monophosphopeptides eluted earlier from the RPLC column than their non-phosphorylated counterparts in a diversity of conditions including a neutral [20] solvent system or an acidic solvent system (containing either an acidifier or ion-pairing agent such as formic acid) [21]; trifluoroacetic acid (TFA) [18,22,23]; pentafluoropropionic acid (PFPA) [23]; or heptafluorobutyric acid (HFBA) [23,24]. Graphite powder has been used to capture phosphopeptides, which were not retained on reversed-phase resin [25-27]. However, similar elution of the phosphopeptides and their non-phosphopeptides has also been observed under acidic RPLC conditions with formic acid [21], acetic acid [28], and TFA [8]. Finally, there have been cases where phosphopeptides were observed being systematically eluted following their non-phosphopeptide counterparts with formic acid [13,29].

Here we systematically investigate the changes in elution time due to peptide phosphorylation while using a high-resolution RPLC system with TFA as the ion pairing agent in combination with tandem mass spectrometry (MS/MS). Phosphopeptides and their corresponding non-phosphopeptides were prepared from standard phosphopeptides and enzymatically digested phosphoproteins such as chicken myelin basic protein (MBP), bovine <beta>-casein, and phosvitin.

2. EXPERIMENTAL

2.1 Sample preparation

Chicken MBP charge isomer C2 was provided by Dr. Zand of the University of Michigan (Ann Arbor, MI, USA). Bovine <beta>-casein (Catalog Number: C6905) and five standard peptides for normalization (ASHLGLAR (Catalog Number: A8651), APRTPGGRR (Catalog Number: A6583), pyroEPPGGSKVILF (Catalog Number: H2136), INLKALAALAKKIL-NH₂ (Catalog Number: M5280), and FLPLILGKLVKGLL-NH₂ (Catalog Number: M8780)) were purchased from Sigma (St. Louis, MO, USA) as shown in Table 1. Phosvitin was purchased from Pierce (Rockford, IL, USA). Sequencing grade modified trypsin was purchased from Promega (Madison, WI, USA). Endoproteinase Glu-C was obtained from Roche Diagnostics (Indianapolis, IN, USA). Standard phosphopeptides (NpYISKGSTFL, PFVKpYATLISN, DLDVPIPGRRFRRVpSVAAE, pYVNV, LRRApSLG, pYVPML, LKRApTLG-NH₂, RRApSPVA, and RRREEEpTEEEAA) were from AnaSpec (San Jose, CA, USA). Enzymatic digestions were performed overnight in 100 mM NH₄HCO₃ (pH 8.0) using an enzyme-to-protein ratio of 1:50 (w/w) at 37 °C. Tryptic digestion was performed on the individual proteins such as chicken MBP, bovine <beta>-casein, and phosvitin. Endoproteinase Glu-C digestion was performed for chicken MBP. Dephosphorylation was catalyzed using alkaline phosphatase (Boehringer-Mannheim, Mannheim, Germany) in 100 mM NH₄HCO₃ (pH 8.0) at a ratio of 1:1 (unit of phosphatase to µg of peptides) for 1 h at 37 °C.

Five sets of samples were prepared for the study; one mixture of standard phosphopeptides, one endoproteinase Glu-C digest of chicken MBP, and three tryptic digests of bovine <beta>-casein, chicken MBP, and phosvitin. Each set contained their corresponding non-phosphopeptides from dephosphorylation with phosphatase, except chicken MBP digests

which naturally contain partially dephosphorylated amino acids. The solution of the standard phosphopeptides and their corresponding non-phosphopeptides contained 1 µg/mL of each peptide. The digests of chicken MBP had a concentration of 200 µg/mL. The phosvitin digest had a concentration of 100 µg/mL for each phosphorylated and dephosphorylated peptide. The bovine <beta>-casein digest had a concentration of 200 µg/mL for each phosphorylated and dephosphorylated peptide. Every sample contained 2 µg/mL of each of the five standard peptides to be used for the normalization of the elution time. A volume of 10 µL of sample solution was injected for each LC/MS analysis, which would be ~20 pmol for each of the standard peptides used for the normalization, ~10 pmol for the standard phosphopeptides and their corresponding non-phosphopeptides, and ~100 pmol for the tryptic digests of chicken MBP, phosvitin, and bovine <beta>-casein,

2.2 RPLC-MS/MS and data analysis

The separations were performed as previously described [30,31]. Briefly, 80 cm fused-silica capillary columns (360 µm O.D. × 150 µm I.D.) (Polymicro Technologies, Phoenix, AZ, USA) packed with 5 µm Jupiter C18-bonded particles (300 Å pore size) (Phenomenex, Torrance, CA, USA) were used for the RPLC separations. Solvent A (0.2% acetic acid and 0.05% TFA in water) and Solvent B (0.1% TFA in 90% acetonitrile and 10% water) were delivered by two ISCO pumps (Model 100DM pump system, ISCO, Lincoln, NE, USA). Prior to loading a sample solution, the capillary column was conditioned with 100% Solvent A for 60 min. After loading a sample solution (10 µL) to the column, the column was conditioned with Solvent A for another 20 min before the gradient with Solvent B was started on the column. All sample solutions contained 1.0% phosphoric acid since it results in an improvement in the overall detection sensitivity for phosphopeptides as shown in our previous study [19].

The eluting solution from each RPLC separation was introduced into an LCQ ion-trap mass spectrometer (ThermoFinnigan, San Jose, CA, USA), where an in-house assembled ESI interface was used with a spray voltage of +2.5 kV. Data-dependent MS/MS mode was selected where three MS/MS scans for the most dominant peaks were performed after one MS survey scan. Dynamic exclusion was applied to avoid overlap of MS/MS scans on the same ion with a 3.00 min exclusion duration setting and an exclusion width of 4.00. The normalized collision energy was 45%.

All MS/MS spectra were confirmed manually to determine the exact sites of amino acid phosphorylation within each peptide. Since the MS/MS spectra from endoproteinase Glu-C digest of chicken MBP did not provide sufficient information to determine the phosphorylated amino acid residues, the information obtained from the tryptic digest of chicken MBP was used for determining the locations of phosphorylation. In some cases (peptides 24-1, 26-1 and 33-1 in Table 2), it was not possible with this approach to determine the exact location of the phosphorylated amino acid residue.

3. RESULTS AND DISCUSSION

3.1 Normalization of RPLC elution time

Different retention times are often observed in different chromatographic separations of the same sample. While the high-resolution RPLC separation minimizes random assignment of peak elution time, special treatment is necessary to minimize systematic variables, which occur from differences such as packing effectiveness, progressive degradation of column properties, and unstable room temperature. To compensate for these variables that may have affected the elution time of peptides from run to run separations, five standard peptides were used to normalize the elution time for all of the 152 RPLC-MS/MS experiments, which were performed using a total of 14 different columns.

As examples, table 1 shows the elution times from four different RPLC conditions for the five standard peptides that were utilized for the normalization. The first two experiments were performed on one column on two successive dates, while the third and the fourth experiments were performed on two different columns. Average elution times for the five standard peptides from the four different RPLC separations were employed as standards which were then applied to normalize the elution time of other RPLC separations as follows. Fig. 1 shows the elution time normalization procedure. First, using both the original basepeak (A) and the basepeak for each of the five standard peptides that was reconstructed (B), the elution time at the highest maximum was selected as the elution time for a peptide of interest. Second, the conversion factors (slopes and y-intercepts) for the four segmented elution time ranges (0 to Std-2, Std-2 to Std-3, Std-3 to Std-4, and Std-4 to 150 min) were determined by the comparison of the elution times of the five standard peptides from the original chromatogram and the average elution times (C) from Table 1. Third, the original chromatogram (A) was converted to the normalized chromatogram (D) by application of the conversion factors. Finally, the normalized elution times of phosphopeptides and their corresponding non-phosphopeptides were used for the comparison. A similar normalization method was once introduced by Sapirstein et al. to compare wheat protein RPLC chromatograms using five calibration peaks [32].

Elution time for a peptide of interest was confirmed by examining the MS/MS spectrum as shown in Fig. 2. An MS reconstructed chromatogram for the doubly charged HRDpSGLLDLGR (theoretical $m/z = 703.33$) with ± 0.5 m/z gives three abundant peaks at 21.27, 46.72, and 54.27 min. (Fig. 2 (B)). The MS/MS spectrum at 46.69 min confirmed that the MS peak at 46.72 min is from HRDpSGLLDLGR.

Table 2 summarizes the normalized elution times for each phosphopeptide and its corresponding non-phosphopeptide investigated in the current experiments. A total of 33 non-phosphopeptides and their corresponding phosphopeptides ranging from 4 amino acid residues to 35 amino acid residues in length were investigated, where four phosphopeptides (peptides 24, 26, 31, and 33) have both single and double phosphorylations, one phosphopeptide (peptide 22) has only double phosphorylation and one peptide (peptide 29) has only tetraphosphorylation. The maximum decrease in observed elution time for the samples investigated in this study (except the tetraphosphopeptide (peptide 29) which exhibited a decrease of -12.15 min) is -5.28 min from pYVPML. Interestingly, some phosphopeptides were retained longer on the reversed-phase column than their corresponding non-phosphopeptides. (e.g. peptides 16, 20, 23, and 31-1)

3.2 Dephosphorylation

Phosphatase was used to prepare samples of phosphopeptides with their corresponding non-phosphopeptides; a sample of standard phosphopeptides was split in half and one half enzymatically dephosphorylated with phosphatase and then combined with the other untreated half. Since chicken MBP has only partial phosphorylation, no dephosphorylation was performed for this protein. Since no non-phosphorylated peptide was observed for the tetraphosphopeptide (peptide 29 in Table 2) from bovine β -casein, dephosphorylation was performed for this protein even if the tryptic monophosphopeptide (peptide 19) showed a small amount ($< 3\%$) of non-phosphorylated peptide. Some low abundant phosphopeptides other than β -casein digested peptides were also detected as impurities (peptides 15, 17, 18, and 22) from the tryptic digest of bovine β -casein.

3.3 Effect of tyrosine phosphorylation

Fig. 3 shows the relationship between the normalized elution time of non-phosphorylated peptides and the number of amino acid residues of each of the peptides. The non-phosphorylated tyrosine-containing peptides (peptides 1, 3, 9, 12, 21, 24, and 26) have

significantly increased elution times, compared to the similar-sized non-phosphorylated threonine- or serine-containing peptides, which is due to the bulky phenolic side chain of the tyrosine residue. When this tyrosine residue is phosphorylated, the elution time of the peptide significantly decreased (Fig. 4 and 5), because phosphorylation on the protruding tyrosine residue contributes extensively to an overall increase of peptide hydrophilicity.

Most of the peptides with tyrosine phosphorylation (phosphotyrosine-containing peptides 1, 3, 9, 12, 21, and 24 in Table 2) display significant decreases in their chromatographic retention, compared to the peptides of the same length with serine or threonine phosphorylation as shown in Fig. 4. The phosphate group on tyrosine is believed to contribute more to the increase of hydrophilicity of the peptides than the phosphate group on serine or threonine since tyrosine makes a longer side chain than serine or threonine. For phosphopeptide 26, however, no significant decrease of the elution time was observed which may be due to the increased length of phosphopeptide 26 compared to the other phosphotyrosine-containing peptide and the fact that the phosphorylation is located in the middle of the sequence where the peptide itself is sufficiently long to overcome any effects of phosphorylation that might result in an increase in peptide hydrophilicity.

Recently, the early elution of tyrosine phosphorylated peptide compared to threonine phosphorylated peptide was observed in an acidic solvent system using formic acid, where the elution order from early to late was VADPDHDHTGFLTE_pYVATR < VADPDHDHTGFLTEYVATR < VADPDHDHTGFL_pTEYVATR [33]. The similar trend was also observed in the analysis of the Lys-C digest of stimulated epidermal growth factor receptor with RPLC solutions containing formic acid, where the elution order, determined by the specific monophosphorylated residue, was pY₁₀₆₈ phosphorylated < pY₁₀₄₅ phosphorylated < pS₁₀₅₇ phosphorylated < pS₁₀₄₆ phosphorylated peptides in the sequence of EDSFLQRY₁₀₄₅S₁₀₄₆SDPTGALTEDS₁₀₅₇IDDTFLPVPEY₁₀₆₈INQSVPK [34].

3.4 Effect of peptide hydrophobicity

The effect of peptide hydrophobicity on the elution time shift due to phosphorylation was evaluated using the grand average of hydropathicity (GRAVY) value [35] and hydrophobicity value [8].

3.4.1 GRAVY value—The GRAVY value for each of the non-phosphopeptides with five or higher amino acid residues was obtained from ExPASy - ProtParam tool (http://bioinformatics.org/sms2/protein_gravy.html) where positive and negative GRAVY values mean hydrophobic and hydrophilic, respectively [35].

Table 2 includes the GRAVY value for each non-phosphopeptide. The dependence of the elution times of the non-phosphopeptides upon GRAVY values is shown in Fig. 6, where no specific correlation was observed between the elution times and the GRAVY values, which means the retention of peptides in RPLC does not depend on the hydropathicity of peptides calculated based on GRAVY. Fig. 7 shows the relation between the GRAVY values and the normalized elution time shifts upon phosphorylation. A low correlation was observed where the squared correlation coefficient R^2 was calculated to be 0.5904, where the peptide 29 was omitted since it deviated too much from the straight line. When all of the phosphotyrosine-containing peptides (peptides 1, 3, 9, 12, 21, 24, and 26 in Table 2) were considered, the R^2 was 0.9645. The peptides with positive GRAVY values, peptides 1, 3, and 12, displayed significant correlation with phosphorylation, where the R^2 was 0.9934. It is not certain that the correlative decrease in retention for the peptides with higher GRAVY values results from either their positive GRAVY values, their tyrosine phosphorylation, or their N-terminal phosphorylation. In order to make a decisive conclusion, more peptides with positive GRAVY values would need to be evaluated.

3.4.2 Hydrophobicity value—Recently an improved model for calculation of peptide hydrophobicity was proposed by Krokhin and co-workers [8] for tryptic peptides. In order to apply the hydrophobicity calculation to the current study, the hydrophobicity value for each non-phosphopeptide was calculated from Sequence Specific Retention Calculator (<http://hs2.proteome.ca/SSRCalc/SSRCalc.html>). Table 2 includes the hydrophobicity value for each non-phosphopeptide. As shown in Fig. 8, a strong correlation was observed between the normalized elution time of non-phosphopeptides and their hydrophobicity values, where the R^2 was 0.9222. A better correlation was exhibited ($R^2 = 0.971$) when only the tryptic peptides in Table 2 were considered, which validates the calculation of the hydrophobicity values. No correlation was observed between the hydrophobicity value and the normalized elution time shifts upon phosphorylation, where the R^2 was 0.0115 (Graph not shown). Conclusively, the improved hydrophobicity calculation was beneficial for the elution time prediction of non-phosphopeptides, while not contributing to the prediction of the elution time shifts upon phosphorylation.

3.5 Effect of N-terminal phosphorylation

N-terminal phosphorylation is believed to contribute significantly to the decrease in retention since the phosphate groups on N-terminal amino acids will increase the hydrophilicity of the peptides. In the current study, four peptides (peptides 1, 3, 8, and 23 in Table 2) have N-terminal phosphorylation, among which peptides 1 and 3 were synthetic peptides and peptides 8 and 23 were tryptically digested peptides from chicken MBP. No missed cleavages by trypsin occur at the phosphorylation sites for peptides 8 and 23, demonstrating that the bond of R-pS or R-pT does not inhibit tryptic digestion.

When the elution time shifts of peptides 1, 3, 8, and 23 in Table 2, all of which have phosphorylation on the N-terminal amino acid residues, were compared, peptides 8 and 23 did not exhibit any significant decrease in elution time compared to the decrease for peptides 1 and 3 as shown in Fig. 4. The decrease of peptides 1 and 3 is believed to be primarily due to the short length of each peptide in combination with the N-terminal tyrosine residue that is phosphorylated. Dass and co-workers observed that phosphorylation on the N-terminal tyrosine residues of enkephalin-related peptides, such as YGGFM, YGGFL, YGGFMK, YGGFLK, YGGFMR, YGGFMRF, and YGGFMRGL, significantly decreased the retention of the phosphopeptides in TFA, PFPA, and HFBA mobile phase systems, where the existence of basic amino acid residues did not affect the decreasing trend [23]. The effect of N-terminal phosphorylation is believed to be dependent on the peptide length and the N-terminal amino acid residue of the peptide, where short peptides with an N-terminal phosphorylated tyrosine residue demonstrate significantly decreased elution times.

3.6 Effect on a peptide with multiple phosphorylations

A peptide with tetraphosphorylation (peptide 29 in Table 2) has shown a significant decrease in the elution time. (−12.15 min decrease), which is due to the large increase in hydrophilicity as a result of the tetraphosphorylation.

Diphosphorylated peptides generally eluted earlier than monophosphorylated peptides (peptides 24-1 vs. 24, 26-1 vs. 26, and 33-1 vs. 33), as expected by the presence of two hydrophilic phosphate groups. However, diphosphorylated peptide 31-1 eluted later than the corresponding monophosphorylated peptide 31 by +0.3 min. It is believed that the intramolecular hydrogen bonding interaction between the phosphate groups and the guanidinium groups of the nearby arginines provides a charge-shielding effect on the phosphate groups. This strong interaction between phosphate groups and arginine groups was investigated by Jackson and co-workers for the noncovalent complexes between arginine-rich VLRRRRKRNVN and SVSTDpTpSAE [36]. The similar phenomenon was observed by

Wettenhall et al. [37], in which a diphosphorylated peptide, RLpSpSLR, eluted later than the monophosphorylated form, RLpSSLR. Hoffmann and co-workers [38] also found that some diphosphorylated peptides eluted later than the corresponding monophosphorylated peptides, depending on the location of the phosphorylated residues. This shows that the specific location of the phosphorylation modification is one of the critical factors for characterizing the elution time change. Different locations of phosphorylation in the same sequence of a peptide often result in a different elution time for monophosphorylated peptides. From the monophosphorylated peptide of KSPVPKSPVQKKG, Otvos et al. [39] observed two distinct elution times depending upon which serine residue is phosphorylated.

3.7 Increases in elution time post-phosphorylation

In general, peptide phosphorylation was observed to decrease the retention of the phosphopeptide compared to its corresponding non-phosphopeptide since the phosphate group increases the overall hydrophilicity of the peptide. Certain peptides, however, exhibited increased retention following phosphorylation as shown by peptides 16, 20, 23, 26, and 31-1 in Table 2. This increase in retention may be explained by the intramolecular hydrogen bonding interaction between the oxygen in the phosphate and the hydrogen in the nearby basic residue or by the intramolecular shielding of the phosphate group by other amino acids [36].

Peptide 16 is believed to have a strong hydrogen bonding interaction between pS and the nearby R, which is strong enough to limit the tryptic digestion between (H)R-D(pS) bond, where DpSGLLDLGR was not observed while DSGLLDLGR was observed along with HRDGLLDLGR and HRDpSGLLDLGR. It is believed that the D located between R and pS is also involved in the hydrogen bonding with R because in other cases (“A” for peptide 11, “F” for peptide 20), no tryptic cleavage limitation was observed. Similar tryptic cleavage limitations were reported for the sequence of –KEpS– [40] and for the sequence of –RapS– [41]. This restriction is unique since the bond of R-pS or R-pT does not inhibit tryptic digestion as mentioned in **3.5 Effect of N-terminal phosphorylation**. This interaction between pS and the nearby R in peptide 16 reduces the effect from hydrophilic phosphate group and increases the hydrophobicity by increasing the molar mass. Peptide 20 is flanked by two bulky amino acids (F and W), which shield the phosphate group effectively. Peptide 23 is believed to have hydrogen bonding interaction between phosphate group on pS and Histidine. Peptide 26 is in the middle of the sequence, so the phosphate group is assumed to be shielded by other amino acids in the same peptide. For peptide 31-1, the intramolecular hydrogen bonding interaction between the phosphate group and nearby guanidinium groups provides a shielding effect on the phosphate groups.

3.8 Effect of different ion-pairing agents

In this investigation, TFA was used as the chromatographic acidic ion-pairing agent. Different acidic ion-pairing agents are anticipated to create changes in the elution time patterns (the amount of shift and possibly the order) since for the same mobile phase concentration, the hydrophobicity and pKa of the ion-pairing reagent will determine the amount of electrostatic interactions available in the column (assuming that the pH of the mobile phase is determined by the concentration of the ion-pairing reagent in the column)[42,43]. Indeed, for aliphatic acidic ion-pairing reagents, the longer the side chain, the more the column will behave as a dynamic cation-exchanger. In this case, phosphorylation not only increases the hydrophilicity of the peptides, but it also reduces the net positive charge if the peptide contains basic amino acid residues that are positively charged under standard RPLC conditions. This decreases the amount of interactions between the peptide and the dynamic cation-exchanger, thus further decreasing the retention of the peptide. In the case where “weak” ion pairing reagents are used such as formic acid, the increase of hydrophilicity can be overcompensated by charge neutralization thereby reducing the overall hydrophilicity and leading to an increase of the

retention time. This provides a possible explanation for why many phosphopeptides elute earlier than their corresponding non-phosphopeptides with strong acidic ion-pairing agents such as TFA [18,22,23], PFPA [23], and HFBA [23,24], while with formic acid, which is a relatively weak ion-pairing agent, increased retention time was commonly observed with phosphorylation [13,29]. Steen et al. [29] described that the phosphorylation of a peptide with TFA as an ion pairing agent decreases the elution time of the peptide because of the increase of the overall hydrophilicity of the peptide while with formic acid the number of phosphorylation sites must be higher than the number of basic amino acid residues to obtain a decreased elution time. Hoffmann et al. [38] found that TFA and HFBA behave differently in the elution pattern of phosphopeptides. Ohguro et al. [44] found that HFBA provided better separation for phosphopeptides from non-phosphopeptides than did either TFA or PFPA. Recently, Tholey et al. [24] also observed better separation with HFBA than for TFA for the separation of phosphopeptides and their corresponding non-phosphopeptides. Additional systematic research needs to be performed to examine the effect of different ion-pairing agents on different peptides.

3.9 Prediction of phosphopeptide retention in RPLC separation

While extensive efforts have been made to predict the retention of non-phosphorylated peptides in RPLC [5-9], little or no effort has been done for the prediction of phosphorylated peptides. The idea that phosphorylated peptides elute earlier than their corresponding non-phosphorylated peptides has prevailed without having quantitative information. Here we suggest a predictable capability for phosphorylated peptides.

As shown in Fig. 4, the elution time shift due to phosphorylation for large peptides (= 18 amino acids) is limited. It can be shown as follows:

$$-0.48 \text{ min} = \Delta t = 0.03 \text{ min} \quad [1]$$

where Δt is the normalized elution time shift (min) due to phosphorylation.

For small peptides (< 18 amino acids), the elution time shifts have various values as shown in Fig. 4. In order to characterize the relationship between the elution time shift and the number of amino acids in a phosphopeptide, a plot of $\Delta t/t_0$ vs. the number of amino acids (N) has been made as shown in Fig. 9, where Δt is the normalized elution time shift (min) from peptide phosphorylation and t_0 is the normalized elution time (min) of the corresponding non-phosphorylated peptide. Phosphotyrosine-containing peptides with less than 18 amino acids (peptides 1, 3, 9, 12, and 21) displayed a good correlation between $\Delta t/t_0$ and N with $R^2 = 0.9753$.

$$\frac{\Delta t}{t_0} = 0.0104 \times N - 0.1761 \quad [2]$$

or

$$\Delta t = (0.0104 \times N - 0.1761) \times t_0 \quad [3]$$

The peptides below the linear regression line (peptides 6 and 7) in Fig. 9 are believed to have a larger contribution of hydrophilicity from phosphate groups than phosphotyrosine-containing peptides have, while the peptides above the linear regression line are believed to have less contribution from phosphorylation.

Without a specific enrichment step, phosphopeptides in LC-MS/MS experiments are often not identified due to either their relatively low abundance in nature [4,45] or their abundant loss of the neutral phosphate group [46]. Providing RPLC elution time information for phosphopeptides with correlation to their corresponding non-phosphopeptides as shown in this study will facilitate identifying phosphopeptides with higher confidence. The phosphopeptide

can be specifically targeted for MS/MS or MS³ experiments to confirm the exact location of the modification [21,47].

4. CONCLUSIONS

Elution time shifts due to phosphorylation were systematically investigated in an RPLCMS/MS system with TFA as the ion-pairing agent. The majority of the phosphopeptides exhibited decreased retention following phosphorylation. The elution time decrease correlated closely with peptide length. Small peptides (< 18 amino acids) showed various ranges of elution time decrease, while larger peptides (= 18 amino acids) displayed only a limited decrease in elution time due to phosphorylation. Based on the currently investigated peptides (33 different peptides and phosphopeptide-counterparts), phosphopeptides with phosphorylated tyrosine residues showed larger correlative decreases in elution time than similar sized phosphopeptides with phosphorylated serine or threonine residues. A prediction model has been suggested for the small tyrosine phosphorylated peptides, in which the parameters of hydrophobicity and the length of the peptides have contributed to the model. The predictive capability of phosphopeptide elution time will help with identification and confirmation of phosphopeptides.

ACKNOWLEDGEMENTS

Authors would like to thank Dr. Zand for his kind donation of Chick MBP samples for fruitful discussions. Portions of this work were supported in part by the NIH (grant RR018522 to R. D. S) and the Environmental Molecular Sciences Laboratory (EMSL) for use of the instrumentation applied in this research. EMSL is a US Department of Energy (DOE) national scientific user facility located at PNNL in Richland, Washington. PNNL is a multi-program national laboratory operated by Battelle Memorial Institute for the DOE under Contract DE-AC05-76RL01830.

References

1. Kim J, Zand R, Lubman DM. Electrophoresis 2003;24:782. [PubMed: 12627438]
2. Grossman PD, Colburn JC, Lauer HH. Anal. Biochem 1989;179:28. [PubMed: 2757198]
3. Offord RE. Nature 1966;211:591. [PubMed: 5968723]
4. Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, Villen J, Li J, Cohn MA, Cantley LC, Gygi SP. Proc. Natl. Acad. Sci. U.S.A 2004;101:12130. [PubMed: 15302935]
5. Browne CA, Bennett HP, Solomon S. Anal. Biochem 1982;124:201. [PubMed: 7125223]
6. Palmblad M, Ramstrom M, Markides KE, Hakansson P, Bergquist J. Anal. Chem 2002;74:5826. [PubMed: 12463368]
7. Petritis K, Kangas LJ, Ferguson PL, Anderson GA, Pasa-Tolic L, Lipton MS, Auberry KJ, Strittmatter EF, Shen Y, Zhao R, Smith RD. Anal. Chem 2003;75:1039. [PubMed: 12641221]
8. Krokhin OV, Craig R, Spicer V, Ens W, Standing KG, Beavis RC, Wilkins JA. Mol. Cell. Proteomics 2004;3:908. [PubMed: 15238601]
9. Baczek T, Wiczling P, Marszall M, Heyden YV, Kalisz R. J. Proteome Res 2005;4:555. [PubMed: 15822934]
10. Shinoda K, Sugimoto M, Yachie N, Sugiyama N, Masuda T, Robert M, Soga T, Tomita M. J. Proteome Res 2006;5:3312. [PubMed: 17137332]
11. Petritis K, Kangas LJ, Yan B, Monroe ME, Strittmatter EF, Qian WJ, Adkins JN, Moore RJ, Xu Y, Lipton MS, Camp DG 2nd, Smith RD. Anal. Chem 2006;78:5026. [PubMed: 16841926]
12. Strittmatter EF, Kangas LJ, Petritis K, Mottaz HM, Anderson GA, Shen Y, Jacobs JM, Camp DG 2nd, Smith RD. J. Proteome Res 2004;3:760. [PubMed: 15359729]
13. Kawakami T, Tateishi K, Yamano Y, Ishikawa T, Kuroki K, Nishimura T. Proteomics 2005;5:856. [PubMed: 15668996]
14. Hunter T. Cell 2000;100:113. [PubMed: 10647936]
15. Kim JK, Mastronardi FG, Wood DD, Lubman DM, Zand R, Moscarello MA. Mol. Cell. Proteomics 2003;2:453. [PubMed: 12832457]

16. Ficarro SB, McClelland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, Hunt DF, White FM. *Nat. Biotechnol* 2002;20:301. [PubMed: 11875433]
17. Zappacosta F, Huddleston MJ, Karcher RL, Gelfand VI, Carr SA, Annan RS. *Anal. Chem* 2002;74:3221. [PubMed: 12141686]
18. Lee KA, Craven KB, Niemi GA, Hurley JB. *Protein Sci* 2002;11:862. [PubMed: 11910029]
19. Kim J, Camp DG 2nd, Smith RD. *J. Mass Spectrom* 2004;39:208. [PubMed: 14991691]
20. Lucas J, Henschen A. *J. Chromatogr* 1986;369:357. [PubMed: 3100570]
21. Tsay YG, Wang YH, Chiu CM, Shen BJ, Lee SC. *Anal. Biochem* 2000;287:55. [PubMed: 11078583]
22. Huddleston MJ, Annan RS, Bean MF, Carr SA. *J. Am. Soc. Mass Spectrom* 1993;4:710.
23. Dass C, Mahalakshmi P, Grandberry D. *J. Chromatogr. A* 1994;678:249. [PubMed: 7951986]
24. Tholey A, Toll H, Huber CG. *Anal. Chem* 2005;77:4618. [PubMed: 16013881]
25. Vacratsis PO, Phinney BS, Gage DA, Gallo KA. *Biochemistry* 2002;41:5613. [PubMed: 11969422]
26. Larsen MR, Graham ME, Robinson PJ, Roepstorff P. *Mol. Cell. Proteomics* 2004;3:456. [PubMed: 14762214]
27. Larsen MR, Cordwell SJ, Roepstorff P. *Proteomics* 2002;2:1277. [PubMed: 12362346]
28. Ruse CI, Willard B, Jin JP, Haas T, Kinter M, Bond M. *Anal. Chem* 2002;74:1658. [PubMed: 12033257]
29. Steen H, Jebanathirajah JA, Rush J, Morrice N, Kirschner MW. *Mol. Cell. Proteomics* 2006;5:172. [PubMed: 16204703]
30. Shen Y, Zhao R, Belov ME, Conrads TP, Anderson GA, Tang K, Pasa-Tolic L, Veenstra TD, Lipton MS, Udseth HR, Smith RD. *Anal. Chem* 2001;73:1766. [PubMed: 11338590]
31. Shen Y, Tolic N, Zhao R, Pasa-Tolic L, Li L, Berger SJ, Harkewicz R, Anderson GA, Belov ME, Smith RD. *Anal. Chem* 2001;73:3011. [PubMed: 11467548]
32. Sapirstein HD, Scanlon MG, Bushuk W. *J. Chromatogr. A* 1989;469:127.
33. Ballif BA, Roux PP, Gerber SA, MacKeigan JP, Blenis J, Gygi SP. *Proc. Natl. Acad. Sci. U.S.A* 2005;102:667. [PubMed: 15647351]
34. Wu SL, Kim J, Bandle RW, Liotta L, Petricoin E, Karger BL. *Mol. Cell. Proteomics* 2006;5:1610. [PubMed: 16799092]
35. Kyte J, Doolittle RF. *J. Mol. Biol* 1982;157:105. [PubMed: 7108955]
36. Jackson SN, Wang HY, Woods AS. *J. Proteome Res* 2005;4:2360. [PubMed: 16335986]
37. Wettenhall RE, Quinn MJ. *J. Chromatogr* 1984;336:51. [PubMed: 6526925]
38. Hoffmann R, Segal M, Otvos JL. *Anal. Chim. Acta* 1997;352:327.
39. Otvos L Jr, Tangoren IA, Wroblewski K, Hollosi M, Lee VM. *J. Chromatogr* 1990;512:265. [PubMed: 2229230]
40. Hashimoto E, Takeda M, Nishizuka Y, Hamana K, Iwai K. *J. Biol. Chem* 1976;251:6287. [PubMed: 185207]
41. Proud CG, Rylatt DB, Yeaman SJ, Cohen P. *FEBS Lett* 1977;80:435. [PubMed: 196939]
42. Petritis K, Brussaens S, Guenu S, Elfakir C, Dreux M. *J. Chromatogr. A* 2002;957:173. [PubMed: 12113341]
43. Petritis KN, Chaimbault P, Elfakir C, Dreux M. *J. Chromatogr. A* 1999;833:147.
44. Ohguro H, Palczewski K. *FEBS Lett* 1995;368:452. [PubMed: 7635197]
45. Tripet B, Cepeniene D, Kovacs JM, Mant CT, Krokhin OV, Hodges RS. *J. Chromatogr. A* 2007;1141:212. [PubMed: 17187811]
46. Put R, Daszykowski M, Baczek T, Vander Heyden Y. *J. Proteome. Res* 2006;5:1618. [PubMed: 16823969]
47. Mant CT, Hodges RS. *J. Chromatogr. A* 2006;1125:211. [PubMed: 16814308]
48. Krokhin OV, Ying S, Cortens JP, Ghosh D, Spicer V, Ens W, Standing KG, Beavis RC, Wilkins JA. *Anal. Chem* 2006;78:6265. [PubMed: 16944911]
49. Gorshkov AV, Tarasova IA, Evreinov VV, Savitski MM, Nielsen ML, Zubarev RA, Gorshkov MV. *Anal. Chem* 2006;78:7770. [PubMed: 17105170]

50. Coon JJ, Shabanowitz J, Hunt DF, Syka JE. J. Am. Soc. Mass Spectrom 2005;16:880. [PubMed: 15907703]
51. Tholey A, Reed J, Lehmann WD. J. Mass Spectrom 1999;34:117. [PubMed: 12440389]
52. Masselon C, Pasa-Tolic L, Tolic N, Anderson GA, Bogdanov B, Vilkov AN, Shen Y, Zhao R, Qian WJ, Lipton MS, Camp DG 2nd, Smith RD. Anal. Chem 2005;77:400. [PubMed: 15649034]

Fig. 1.

Normalization of elution time applying the segmented conversion of (A) the original basepeak, (B) reconstructed ion chromatogram for the five standard peptides, (C) conversion of (B) to the average elution time from Table 1, and (D) converted basepeak. The values in the parentheses are the average elution times for the five standard peptides that were used for the conversion. Four segmented ranges were used for the conversion. The slope and the y-intercept for each segment were calculated as shown in the figure and then applied to convert other peaks. A mixture of tryptic digest of 2.5 μ g chicken MBP and 20 ng of each standard peptide was loaded onto the column. The values in parentheses are the average elution times of the five standard peptides, used for the conversion. For the reconstructed ion chromatogram, scanning of ± 0.5 m/z of the abundant peaks of the standard peptides (singly charged state for Std-2 and Std-3 and doubly charge state for the other standard peptides) were made.

Fig. 2. Elution time confirmation for a peptide of interest. (A) Original basepeak, (B) reconstructed ion chromatogram for of the doubly charged HRDpSGLLDLGR, theoretical $m/z = 703.33$, and (C) tandem mass spectrum of the doubly charged HRDpSGLLDLGR (m/z 703.36) at 46.69 min. The normalized elution time of the peptide of interest is shown in (B). The peak with asterisk is from the result of neutral loss. The intensity around this peak was reduced by half since the peak was highly abundant. For the details of the sample, refer to Fig. 1.

Fig. 3.

Plot of the number of amino acids vs. the elution time of non-phosphorylated peptides. Circle or square around a peptide number means the peptide contains a potential phosphotyrosine or phosphothreonine residue, respectively. The other peptides contain potential phosphoserine residues. Potential phosphotyrosine-containing peptides (peptides with circles around the number) have shown higher elution times compared to similar sized peptides. Standard deviations are shown with average elution times. A dark circle (•) indicates that the peptide contains more than one phosphorylated amino acid.

Fig. 4. Effect of the number of amino acids to the normalized elution time shift due to phosphorylation. Circle or square around a peptide number means the peptide contains phosphotyrosine or phosphothreonine residue, respectively. The other peptides contain phosphoserine residues. Standard deviations are shown with average elution times. A dark circle (•) means that the peptide contains more than one phosphorylated amino acid.

Fig. 5.

Effect of the normalized elution time of the non-phosphorylated peptides (x-axis) to the normalized elution time shift due to phosphorylation. Circle or square around a peptide number means the peptide contains a phosphotyrosine or phosphothreonine residue, respectively. The other peptides contain phosphoserine residues. Standard deviations are shown with average elution times. A dark circle (•) means that the peptide contains more than one phosphorylated amino acid.

Fig. 6.

Dependence of the elution times of non-phosphopeptides upon their GRAVY values. No overall correlation is observed. Circle or square around a peptide number means the peptide contains phosphotyrosine or phosphothreonine residue, respectively. The other peptides contain phosphoserine residues. Standard deviations are shown with average elution times.

Fig. 7.

Relation between the GRAVY values and the normalized elution time shifts upon phosphorylation. Circle or square around a peptide number means the peptide contains phosphotyrosine or phosphothreonine residue, respectively. The other peptides contain phosphoserine residues. Standard deviations are shown with average elution times. Linear regression for the hydrophobic phosphotyrosine-containing peptides shows the R^2 as 0.9934, while linear regression for all the phosphotyrosine-containing peptides shows the R^2 as 0.9645.

Fig. 8. Dependence of the elution times of non-phosphopeptides upon their hydrophobicity values. Circle or square around a peptide number means the peptide contains phosphotyrosine or phosphothreonine residue, respectively. The other peptides contain phosphoserine residues. Standard deviations are shown with average elution times. Linear regression for all of the peptides and only for the tryptic peptides shows the R^2 values as 0.9222 and 0.971, respectively.

Fig. 9.

Dependence of $\Delta t/t_0$ on the number of amino acids (N), where Δt is the normalized elution time shift from phosphorylation and t_0 is the normalized elution time of non-phosphorylated peptide. The linear regression line is calculated based on the phosphotyrosine containing peptides with less than 18 amino acids (peptides 1, 3, 9, 12, and 21). Good correlation has been observed for the regression ($R^2 = 0.9753$). Circle or square around a peptide number means the peptide contains phosphotyrosine or phosphothreonine residue, respectively. The other peptides contain phosphoserine residues. Standard deviations are also shown. A dark circle (•) means that the peptide contains more than one phosphorylated amino acid.

Table 1
Elution times of standard peptides in different experimental conditions

Peptides	Average molar mass	elution time (min)			
		Exp. 1	Exp. 2	Exp. 3	Average
Sid-1	967.09	18.41	18.46	22.90	21.14
Sid-2	823.94	26.91	26.93	32.29	30.12
Sid-3	1125.32	44.45	44.46	50.45	48.07
Sid-4	1478.91	62.46	62.56	68.80	66.03
Sid-5	1523.00	70.95	70.75	78.06	74.89

Three different 80-cm packed fused-silica capillary columns (Columns A, B, and C) prepared on the same date were used for the elution time comparison of standard peptides. Exp. 1 and 2 were performed on the same column (Column A) on a different date (one day apart). The other experiments (Exp. 3 and Exp. 4) were performed on different columns (Column B and Column C, respectively). Exp. 2 and 3 were performed on the same date. Exp. 4 was performed one month earlier than the other experiments.

Table 2

Changes in elution times due to phosphorylation

Peptide ^{d)}	Source	N ^{b)}	GRAVY ^{c)}	H ^{d)}	n ^{e)}	Normalized elution time			
						t _R ^{f)}	Δt _R ^{g)}	Std. ^{h)}	Δt/t _R
1	pYNNV	4	0.9	8.45	5	29.73	25.55	-4.18	0.06
2	R.SbSFR.H	4	-0.825	7.49	3	21.30	19.58	-1.72	0.10
3	pYVPM.L	5	1.4	24.75	5	44.17	38.89	-5.28	0.12
4	LRRApSLG	7	-0.114	11.81	5	28.00	26.55	-1.45	0.12
5	LKRApTLG.NH ₂	7	-0.014	12.66	5	26.86	25.51	-1.35	0.11
6	RRApSPVA	7	-0.514	7.16	4	23.54	20.47	-3.07	0.04
7	R.SGpSPVARR	7	-0.3	5.93	3	21.13	18.41	-2.72	0.17
8	R.pTTPPMQAK.G	8	-1.15	7.9	3	24.64	23.23	-1.41	0.06
9	NpYISKGSTFL	10	-0.03	26.75	3	42.31	39.67	-2.64	0.09
10	G.SRSgSPVARR.-	10	-1.19	6.18	3	23.71	22.27	-1.44	0.16
11	R.ApSHVGSIPOR.S	10	-0.43	11.3	3	27.08	26.09	-0.99	0.10
12	PFVKpYATLISN	11	0.482	34.63	3	42.54	39.58	-2.96	0.04
13	K.LGGSGSRPGpSR.S	11	-0.982	4.89	4	21.86	21.06	-0.80	0.04
14	K.MASApSTTDHAR.H	11	-0.627	6.04	4	22.08	21.74	-0.34	0.19
15	K.TVDMEpSTEVFTK.K	12	-0.35	24.75	3	37.70	36.17	-1.53	0.08
16	R.HRDpSGLLDSLGR.F	12	-0.85	23.56	4	39.22	39.81	0.59	0.11
17	K.VPQLpEIVPNpSAEER.L	14	-0.536	26.92	3	39.58	38.68	-0.90	0.05
18	K.YKVPQLpEIVPNpSAEER.L	16	-0.794	29.36	3	41.44	41.06	-0.38	0.08
19	K.FQpSSEEQQTDELQDK.I	16	-2.331	16.93	3	30.92	30.80	-0.12	0.04
20	R.FpSWGGEGHKPGYSGK.F	16	-1.219	19.55	4	35.01	35.37	0.36	0.10
21	K.SQFNSSHSpYDIPNEWET	16	-1.494	30.42	3	41.40	40.96	-0.44	0.02
22 ^{j)}	K.DIGpSEpSTEDQAMEDIK.Q	16	-1.15	23.09	3	35.50	35.25	-0.25	0.07
23	R.pSQHGRPGIDNPVVHFFK. N	17	-1.153	25.91	3	42.42	42.89	0.47	0.16
24	K.SQFNSSHSpYDIPNEWET. L	18	-1.439	32.69	3	43.89	43.55	-0.34	0.03
24- 1 ^{k)}	K.SQFNSSHSpYDIPNEWET. L + 1 HPO ₃ ^{l)}	18	-1.439	32.69	3	43.89	43.20	-0.69	0.13
25	DLDPVPGFRFDRVpSVAAE	19	-0.242	34.6	3	42.62	42.57	-0.05	0.05
26	K.SQFNSSHSpYDIPNEWET. L.p	19	-1.163	37.92	3	51.03	51.06	0.03	0.08
26- 1 ^{k)}	K.SQFNSSHSpYDIPNEWET. L.p + 1 HPO ₃ ^{l)}	19	-1.163	37.92	3	51.03	50.78	-0.25	0.13
27	- .acetASQKRSpSRHGSKMAS ASTID.H ^{m)}	21	-1.119	9.89	3	27.27	26.81	-0.46	0.06
28	S.KIFKLGGSGSRPGSRSpS pVARR.-	23	-0.778	22.61	3	33.82	33.38	-0.44	0.04
29 ^{k)}	-REELNVPGIEVpSLpSpSp SEESITR.I	25	-0.596	38.17	3	55.18	43.03	-12.15	0.05
30	E.GQGTLISKIFKLGGSGSRPG SRSpSPVARR	27	-0.404	28.2	3	43.34	43.06	-0.28	0.05
31	E.GQGTLISKIFKLGGSGSRPG	29	-0.686	27.24	3	41.45	41.40	-0.05	0.06

	Peptide ^{a)}	Source	N ^{b)}	GRAVY ^{c)}	H ^{d)}	n ^{e)}	Normalized elution time			
							t _f ^{f)}	g ^{g)}	Average Δt ^{h)}	Δt/t ₀
31-1 k)	SRSGrSPVARR- E.GQGTLSKIFKLGSGSRPG pSRSGrSPVARR-	chicken MBP								
32	- .aceIASQKRSpSFRHGSKMAS ASTTDHARHGSPRHRD.S ^{m)}	Endoproteinase Glu-C digest of chicken MBP	29	-0.686	27.24	3	41.45	41.70	0.25	0.06
33	- .AEFGTEPDAKTSSSSSASS TApTSSSSSSASSPNR.K	Endoproteinase Glu-C digest of chicken MBP	32	-0.686	4.16	3	27.61	27.14	-0.47	0.06
33-1 k)	- .AEFGTEPDAKTSSSSSASS TApTSSSSSSASSPNR.K + 1 HPO ₃ ⁱ⁾	Tryptic digest of phosphitin	35	-0.851	11.64	3	27.91	27.43	-0.48	0.08
		Tryptic digest of phosphitin	35	-0.851	11.64	3	27.91	26.84	-1.07	0.05

a) "p" is placed in front of the phosphorylated amino acid. Amino acids before and after the peptide sequences are shown if possible.

b) "N" means number of amino acids.

c) "GRAVY" means Grand average of hydrophobicity.

d) "H" means hydrophobicity from Sequence Specific Retention Calculator (<http://hs2.proteome.ca/SSRCalc/SSRCalc.html>).

e) "n" means number of measurements.

f) t₀ means the normalized elution time of the non-phosphorylated peptide.

g) t means the normalized elution time of the phosphorylated peptide.

h) Δt = t - t₀

i) Standard deviation of Δt.

j) These proteins appeared as impurities in bovine <beta>-casein from Sigma.

k) The asterisk means that the peptides are multiply phosphorylated.

l) The exact location of the additional phosphorylation was not identified. The underlined amino acids are possible modification sites.

m) "acetA" means acetylated Alanine.