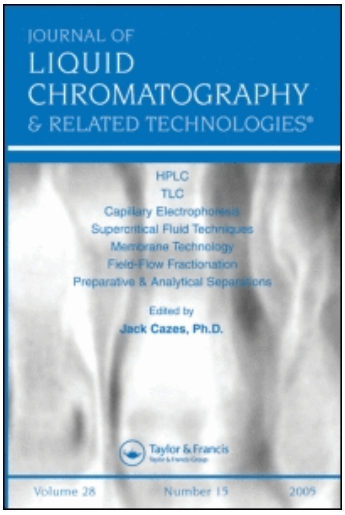

Journal of Liquid Chromatography & Related Technologies
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597273

EFFECT OF PHOSPHORYLATION ON THE RETENTION BEHAVIOR OF PEPTIDES IN ION PAIRING REVERSED-PHASE HPLC BASED ON A PREDICTION MODEL

Yanjun Yu[a]; Suxia Yuan[a]; Rongxin Su[a]; Wei Qi[a]; Zhimin He[a]
[a] State Key Laboratory of Chemical Engineering, School of Chemical Engineering and Technology, Tianjin University, Tianjin, P. R. China

Online publication date: 24 March 2010

To cite this Article Yu, Yanjun , Yuan, Suxia , Su, Rongxin , Qi, Wei and He, Zhimin(2010) 'EFFECT OF PHOSPHORYLATION ON THE RETENTION BEHAVIOR OF PEPTIDES IN ION PAIRING REVERSED-PHASE HPLC BASED ON A PREDICTION MODEL', Journal of Liquid Chromatography & Related Technologies, 33: 6, 733 — 747
To link to this Article: DOI: 10.1080/10826071003683945
URL: http://dx.doi.org/10.1080/10826071003683945

Journal of Liquid Chromatography & Related Technologies, 33:733–747, 2010

ISSN: 1082-6076 print/1520-572X online
DOI: 10.1080/10826071003683945

EFFECT OF PHOSPHORYLATION ON THE RETENTION BEHAVIOR OF PEPTIDES IN ION PAIRING REVERSED-PHASE HPLC BASED ON A PREDICTION MODEL

Yanjun Yu, Suxia Yuan, Rongxin Su, Wei Qi, and Zhimin He

State Key Laboratory of Chemical Engineering, School of Chemical Engineering and Technology, Tianjin University, Tianjin, P. R. China

□ *A prediction model of the peptide retention time was developed on the basis of 358 non-phosphorylated peptides, which had various amino acid residues in the C-terminal from the casein pancreatic hydrolysates. This model was applied to predict the retention times (RT) of another 43 peptides on C4 reversed phase columns with trifluoroacetic acid (TFA) as the ion-pairing reagent, with a relatively high R^2 value (0.969). Furthermore, the experimental RTs of 32 phosphopeptides were compared with the predictive RTs of their non-phosphorylated cognates. Mono- and poly-phosphopeptides seemed to elute after or before the non-phosphorylated predictive cognates, respectively. The positive charges of peptides were partly masked by the ion-pairing reagent TFA. Single and multiple phosphorylation might generally lead to an increase or reduction in the overall hydrophobicity of peptides, respectively.*

Keywords mass spectrometry, phosphopeptide, phosphorylation, prediction retention time

INTRODUCTION

Reversible protein phosphorylation is one of the most significant and prevalent intracellular post modifications, and has been regarded as an immense body of research. It plays a key role in numerous regulations of biological processes such as cell proliferation, differentiation, apoptosis migration, and metabolism. Protein phosphorylation is believed to contribute to many diseases including diabetes, neurological and autoimmune disorders, and cancers.[1–3] Reversible phosphorylation in eukaryotic cells is widely used to regulate protein-protein interactions, so as to transmit and integrate signals received from their environment.[4,5] It was estimated

Address correspondence to Rongxin Su, State Key Laboratory of Chemical Engineering, School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, P. R. China. E-mail: surx@tju.edu.cn

that about 30–50% proteome consists of phosphoproteins.[6] Therefore, protein phosphorylation has been promoted as a separate research system in life sciences, especially in proteomics.

Although protein phosphorylation is of major interest in many worldwide laboratories, the identification of the phosphorylated sites in protein sequence is still a great challenge for life scientists. Until now, protein phosphorylation has been typically studied by means of radioactive 3P labeling, antibody recognition, Edman deposition, electrophoresis, staining techniques, and immunological methods.[7,8]

Recently, the application of mass spectrometry (MS) accelerated the proteomics, particularly phosphoproteomics. Its combination with reversed phase high performance liquid chromatography (HPLC) has provided the most powerful and versatile tool to characterize protein phosphorylation.[9,10] To date, numerous phosphorylated proteins have been detected via MS; however, more information about protein phosphorylation stoichiometry is required to perform unambiguous analyses.

The retention information of phosphopeptides from protein hydrolysates during the HPLC separation can enhance the confidence of phosphoprotein identification in mass spectrometry based proteomics.[11–13] Hoffmann et al.[14] observed that synthetic monophosphorylated peptides, with TFA as ion-pairing reagent, eluted earlier than its unmodified counterparts, and that anathe retention times of di-phosphopeptides were sometimes greater than unphosphorylated and monophosphorylated analogs. Tholey et al.[15] demonstrated that most synthetic peptides with a single phosphorylation site eluted before their unmodified cognates under the 0.05% TFA system. At the same time, β-CN (33–48) and β-CN (1–25) from casein tryptic hydrolysates showed a significant increase in retention time after single and quadruple dephosphorylation, respectively. Recently, Kim et al.[16] observed that peptide phosphorylation generally led to a decrease in elution time, but a few peptides displayed increased elution times due to phosphorylation. In contrast, Steen et al.[17] obtained the opposite results when 0.2% formic acid was used as the ion-pairing reagent. All the monophosphopeptides eluted later than the unphosphorylated cognates due to the reduction in the overall hydrophilicities after modification. Consequently, the effect of phosphorylation, especially multiple phosphorylation, on the retention behavior of peptides is still not well understood.

Insights into how peptide phosphorylation affects its retention time are difficult, because it is generally hard to obtain a substantial amount of phosphopeptides and their unphosphorylated cognates simultaneously. Preparation of these peptides requires enzymatic phosphorylation or dephosphorylation,[16] which are both time consuming and expensive. The prediction models for peptide retention in reverse phase liquid chromatography (RPLC) might offer a solution to this difficulty by

providing the predictive retention time of unphosphorylated peptides. To date, various prediction models for retention time of peptides in liquid chromatography have been proposed, and the retention behavior of peptides during RPLC is believed to be mainly affected by the amino acid composition, peptide length, and neighbor effects.[18–20]

The objectives of this study are to develop a prediction model for the retention time of unphosphorylated peptides during RPLC by taking into account the retention coefficients and length of peptides, and to evaluate the effect of phosphorylation on peptide retention behavior by using the established model to compare the difference between the measured retention time of phosphopeptides and the predictive value of their unphosphorylated cognates.

EXPERIMENTAL

Materials

Bovine casein of technical grade, porcine pancreatin, and pepsin were purchased from Sigma (St. Louis, MO 63178, USA). Bovine hemoglobin of electrophoresis grade was obtained from Tianjin Blood Research Center (Tianjin, China). Acetonitrile (ACN) and trifluoroacetic acid (TFA), of HPLC grade, were bought from Merck (Darmstadt, Germany). Ultrapure water was obtained from an ELGA water purification unit (ELGA Ltd., Bucks, England). All other common reagents and solvents were of analytical grade and sourced commercially.

Preparation of Protein Digests for HPLC-MS Analysis

The peptides used in this work were derived from bovine casein pancreatic hydrolysates[21] and bovine hemoglobin peptic hydrolysates.[22] Bovine casein (10 g/L) was digested by pancreatin (0.1 g/L) at pH 8.0 and 37°C, whereas bovine hemoglobin (10 g/L) was hydrolyzed by pepsin (0.5 g/L) at pH 4.5 and 37°C. During hydrolysis, aliquots of peptide mixtures were taken out at given times. To inactivate the enzyme, casein hydrolysates were heated in boiling water for 10 min, while hemoglobin hydrolysates were immediately modulated up to a value of pH 10.0 by ammonium hydroxide. The samples were lyophilized and stored at −20°C.

Analysis of Peptide Mixtures by On-Line RPLC-ESI-MS/MS

All samples were separated and identified by an on-line liquid chromatography–tandem mass spectrometry (LC–MS/MS) system, equipped with a

Surveyor LC pump and a Surveyor autosampler attached to a Surveyor PDA detector, and an LCQ ion trap mass spectrometer (Thermo Finnigan, San Jose, CA). Liquid chromatography was performed on a 4.6×250 mm column (5 μm, 300 Å, 214TP54 C4, Vydac). Samples (10 μL) were eluted with buffer A (water containing 0.1% TFA and 2% ACN) and buffer B (ACN containing 0.085% TFA and 5% water) at a flow rate of 1 mL/min. The programme started with isocratic elution 100% A for 5 min, followed by linear gradient elution 100–50% A for 50 min, and then 50% A for 10 min.

The MS and tandem MS experiments were performed with an electrospray interface operated in the positive ion mode, and in a data dependent scan mode to automatically switch between MS and MS/MS acquisition controlled by the Xcalibur software. Full scans were performed between m/z 300 and 2000. A full scan MS spectrum was acquired followed by tandem mass spectra using collision induced dissociation (CID) of the three most intense precursor ions present in the MS scan. Electrospray conditions were as follows: capillary temperature, 300°C; ESI voltage, +5.0 kV; and capillary voltage, 2.1 kV.

Peak Assignments and Peptide Identification

All the MS/MS spectra were searched using SALSA and SEQUEST algorithm (Thermo Finnigan, Bioworks 3.1 version). Because phosphoserine and phosphothreonine commonly undergo a β-elimination reaction and readily lose phosphoric acid during the CID process in the ion-trap mass spectrometer, the SALSA algorithm[23] was applied to search MS/MS spectra for fragment ions formed by the neutral loss of phosphoric acid (i.e., 24.5, 32.7, 49, and 98 Da from $(M+4H)^{4+}$, $(M+3H)^{3+}$, $(M+2H)^{2+}$, and $(M+H)^{+}$ ions, respectively). Subsequently, they were searched against bovine database for peptide sequence and identification of site (s) of phosphorylation via SEQUEST algorithm.[24] Search parameters included differential mass modification (+80 Da) on serine and threonine due to their possible phosphorylation. To increase the accuracy of the prediction model, only those peptides achieving Xcorr ≥ 1.9 for singly, 2.2 for doubly, and 3.75 for triply charged peptides were considered.

Prediction Model Development

The prediction model for retention times of peptides during RPLC was developed using MATLAB Release 7.1 (The Mathworks, Natick, MA).

RESULTS AND DISCUSSION

Data Modification and Alignment

The prediction model of the retention time was often based on a gradient of 1% ACN per minute,[18] which was different from our experimental condition. To be unified, the modified retention time (RT) was calculated as the following:

$$RT = g \times (RT_0 - t_{iso}) \tag{1}$$

where RT_0 was the retention time of the peptide under our experimental conditions, t_{iso} was the elution time before the gradient of ACN (in our case, 5 min), g corresponded to the shift from our experimental gradient of ACN (in our case, 0.93 percent ACN per minute) to the gradient of 1% ACN per minute.

To determine the dead volume of the RPLC system, the retention times of peptides were initially evaluated from the bovine hemoglobin peptic digestion (a relatively simple mixture). According to Guo et al.[25] the predictive retention time of peptides (τ) equals the sum of the retention coefficients of the amino acid residues and end groups $\left(\sum R_c\right)$, plus the time of elution of the non-retained compounds (t_{nr}), expressed as the following:

$$\tau = \sum_{i=1}^{20} R_{c_i} n_i + t_{nr} \tag{2}$$

Here, R_{c_i} is the retention coefficients for the 20 amino acids, n_i is the number of each amino acid, and the value of t_{nr} can be estimated as the sum of the dead time of the column plus the dwell time of the HPLC system.

As shown in Figure 1, the value of t_o (14.81) was obtained for bovine hemoglobin peptic hydrolysates, together with the value of R^2 (0.936), by using multiple least regression. To develop the prediction model, the peptides eluting earlier than t_{nr} cannot be reserved. Therefore, the peptide samples with RT_0 greater than 20.92 $(14.81/0.93 + 5)$ should be discarded.

As a result, 358 unphosphorylated peptides were identified and selected from bovine casein enzymatic hydrolysates, and were divided randomly into the training data of 315 peptides and the testing data of 43 peptides.

Model Development

The retention time of peptide in RPLC (RT) depends mainly on its overall hydrophobicity (H). Krokhin et al.[18] depicted out the remarkable

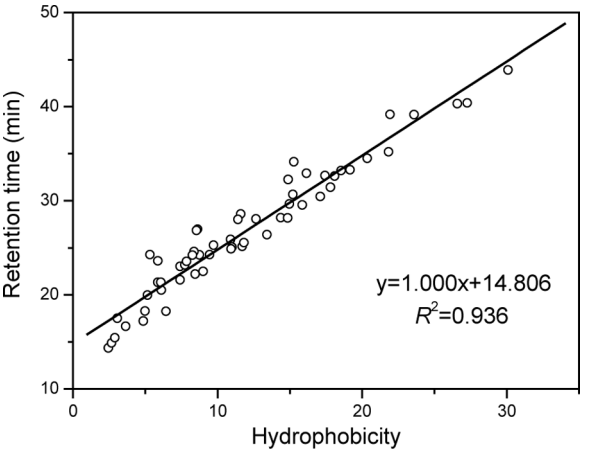

FIGURE 1 The multiple linear regression of prediction model for retention times of peptides from bovine hemoglobin peptic hydrolysates.

linear relationship between RT and H with a high value of R^2, expressed as the following:

$$RT = sH + b \tag{3}$$

where s was the slope, b was the intercept, and the starting value of the hydrophobicity (H) was equal to the sum of retention coefficients for each amino acid residue:

$$H = \sum Rc \tag{4}$$

When the slope s was 1, the retention time can be calculated by the following equation:

$$RT = \sum Rc + b \tag{5}$$

Similar to Eq. (2), Eq. (5) can be regressed by using MLR to determine the retention coefficient of each amino acid residue. In our work, all the peptides containing Cys residues in their sequences were excluded due to the scarcity of this residue in protein hydrolysates. As shown in Table 1 and Figure 2, the relative retention coefficient of each amino acid residue in the experiment was similar to previous studies.[18,20,25,26] However, the R^2 value of 0.834 was relatively low, which resulted in unavoidable optimization for the model.

TABLE 1 Comparison of Retention Coefficients for Individual Amino Acid Residues from Different Models

Amino acid residue	Retention coefficients					
	Present work 1[a]	Present work 2[b]	Krokhin[18]	Guo[25]	Meek[26]	Palmbad[20]
Trp	4.4	12.3	11.0	8.8	18.1	24.69
Phe	4.4	12.2	10.5	8.1	13.9	14.13
Leu	3.1	10.1	9.6	8.1	10.0	12.04
Ile	2.8	9.0	8.4	7.4	11.8	14.24
Met	1.1	5.6	5.8	5.5	7.1	5.19
Tyr	1.3	5.6	4.0	4.5	8.2	14.67
Val	1.1	5.1	5.0	5.0	3.3	12.88
Pro	0.3	3.2	0.2	2.0	3.3	5.1
Gly	0.0	2.1	−0.9	−0.2	−0.5	1.5
Glu	−0.3	1.7	0.0	1.1	−2.5	5.4
Ala	−0.3	1.6	0.8	2.0	−0.1	2.18
Thr	−0.4	1.4	0.4	0.6	1.5	1.94
Asp	−0.4	1.4	−0.5	0.2	−2.8	0.2
Ser	−0.5	0.7	−0.8	−0.2	−3.7	−3.7
Arg	−0.1	0.2	−1.3	−0.6	−4.5	−4.0
His	−0.4	0.2	−1.3	−2.1	0.8	3.0
Asn	−0.8	0.1	−1.2	−0.6	−1.6	−2.9
Gln	−1.3	−0.7	−0.9	0.0	−7.5	5.4
Lys	−0.8	−1.2	−1.9	−2.1	−3.2	−3.5
R^2	0.83	0.96	0.75	0.67	0.77	0.58

[a]According to equation 5.
[b]According to equation 6.

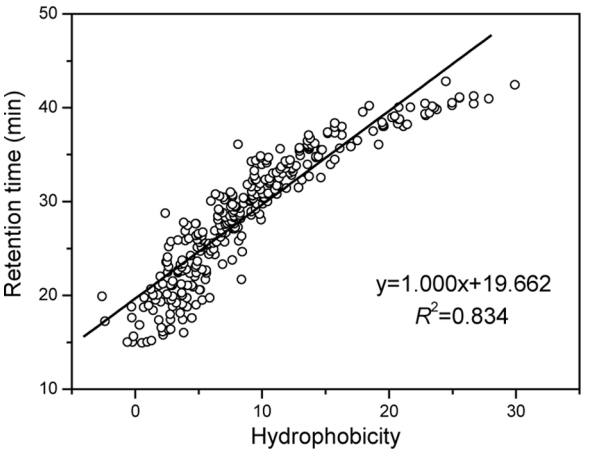

FIGURE 2 The multiple linear regression of prediction model for retention times of peptides from bovine casein enzymatic hydrolysates.

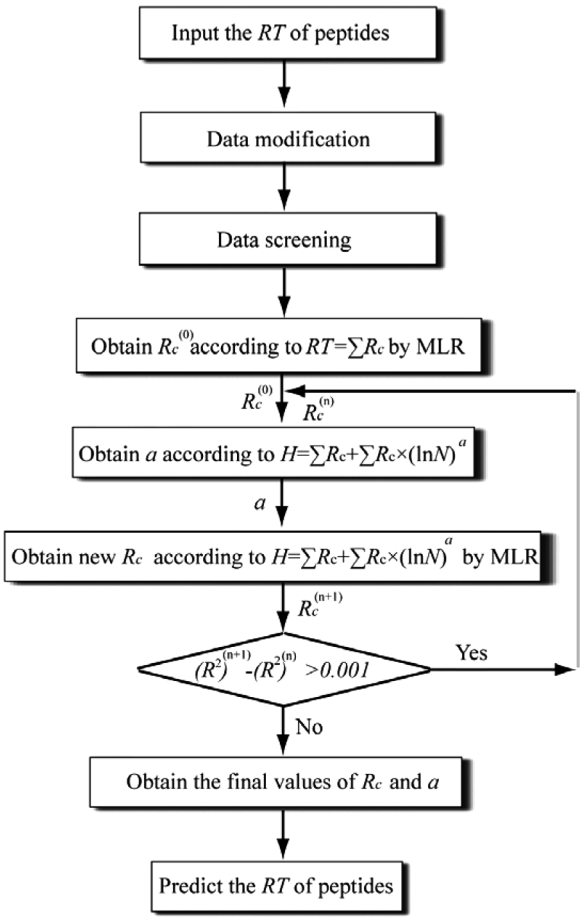

FIGURE 3 Scheme for the prediction model development of retention time of peptides during reversed-phase liquid chromatography.

Model Optimization

The retention behavior of peptides in RPLC was found to be governed by the amino acid composition, peptide length, the position of the amino acid residues, pI, nearest neighbor effect of charged side chains, and secondary structures. Taking these factors into account, Krokhin et al.[27] developed a sequence specific retention calculator (SSRCalc) algorithm with the high R^2 value of 0.98. Among these factors, the peptide length plays the second most significant role on the model just after the retention coefficients. Mant et al.[28] suggested that the peptide including more than 20 amino acid residues could not be well predicted by a simple model just concerning the retention coefficients.

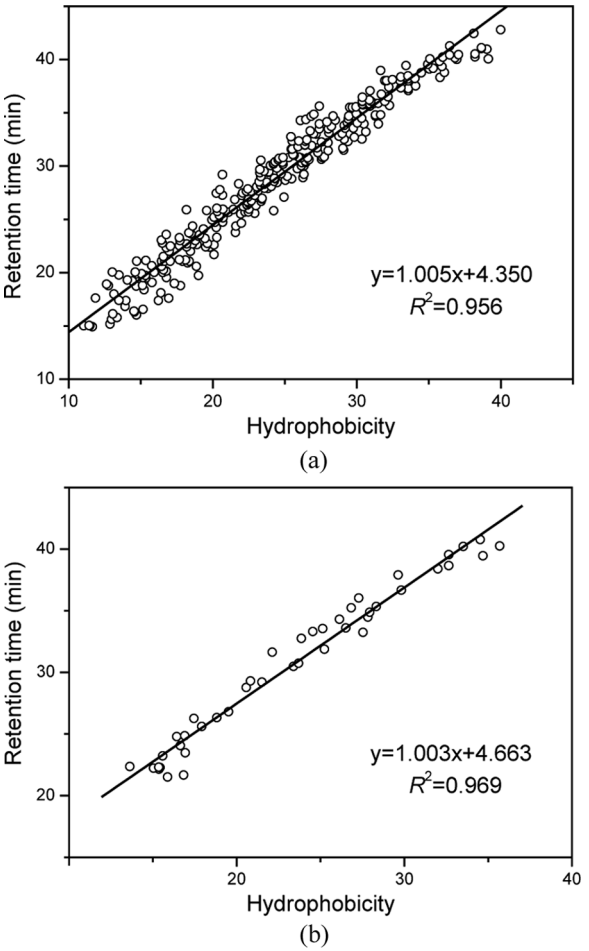

FIGURE 4 The measured retention time *versus* hydrophobicity, which was calculated by concerning retention coefficients and length of peptides, plots for the training peptides (a) and the testing peptides (b).

Guo et al.[25] found that the predicted retention had a linear relationship with the product of peptide hydrophobicity $\left(\sum R_c\right)$ and the logarithm of the residue number. Considering the retention coefficient (R_c) and peptide length (N), we optimized these two key factors at the same time, calculated as,

$$H = \sum R_c + \sum R_c \times \ln(N)^a \tag{6}$$

where a was the correction parameter for the effect of peptide length. As shown in Figure 3, the arithmetic program was developed and accomplished using MATLAB. The iterative regression was stopped when the increase in

the R^2 value was lower than 0.001 for the plot of retention time versus the hydrophobicities of peptides. In Figure 4a, a high R^2 value of 0.956 after the regression is shown. Thus, the significance of peptide length is again exhibited in the shift of the R^2 value from 0.835 to 0.956.

The model was further validated by the random test of a set consisting of 43 peptides from casein hydrolysates (Figure 4b). The value of R^2 of 0.969 was somewhat higher than that of the training data, but the obtained value for the slope (1.003) is similar to that of training data (1.005). Consequently, this also indicates the reliability of this model. In this work, the residues with the most positive retention coefficients were the strongly hydrophobic ones (W, F, L, I), then followed by the weakly hydrophobic ones (Y, M, V). On the other hand, the Lys residue showed the most negative retention. Such results are in good agreement with other previous studies, as reported in Table 1.

EFFECT OF PHOSPHORYLATION ON RETENTION BEHAVIOR OF PEPTIDE

In this work, we discarded the larger casein phosphopeptides (CPPs) with more than 45 amino acid residues, the CPPs with the low Xcorr score, and the CPPs, which eluted earlier than the t_{nr} (non-retained elution time). As a result, 32 identified CPPs were selected for the analysis of the effect of phosphorylation on peptide retention behavior (Table 2). Among these CPPs, 15 (46.9%), 3 (9.4%), 7 (21.9%), and 7 (21.9%) were singly, doubly, triply, and quadruply phosphorylated, respectively.

Subsequently, the retention behavior of the phosphopeptide in RPLC was examined. The predictive retention time of the unphosphorylated cognate was calculated and compared to the CPPs by using the established prediction model (Figure 4). The ΔRT was calculated by subtracting the predictive RT of the unmodified peptide from the experimental RT of its phosphorylated cognate.

As shown in Table 2 and Figure 5, the singly phosphorylated peptides eluted after their unphosphorylated analogues, whereas most of the poly-phosphopeptides eluted earlier than their unmodified counterparts. It has been reported that a phosphor motif might possess two negative charges at $pH < 7$, which would complement some positive charge of peptides due to the N-terminal residue and the unshielded basic residues of Lys, Arg, and His.

Trifluoroacetic acid (TFA) is a strong ion-pairing reagent for peptides during RPLC. TFA can mask certain positive charges of peptides. In the electrospray-ionization process, all the negative phosphor moieties are neutralized and the phosphopeptides carried two or three positive charges,

TABLE 2 Phosphopeptides Released from Bovine Casein Enzymatic Hydrolysates

No.	Protein fragment	Peptide sequence[a]	Sp[b]	ΔRT^c	RT^d	Z^e	N_b^f
1	α_{S1}-CN 104–119 (1P)	YKVPQLEIVPNS*AEER	1	1.16	24.96	2	2
2	α_{S1}-CN 104–121 (1P)	YKVPQLEIVPNS*AEERLH	1	0.03	26.32	2	3
3	α_{S1}-CN 104–122 (1P)	YKVPQLEIVPNS*AEERLHS	1	0.32	26.23	3	3
4	α_{S1}-CN 104–123 (1P)	YKVPQLEIVPNS*AEERLHSM	1	1.49	28.45	2	3
5	α_{S1}-CN 104–124 (1P)	YKVPQLEIVPNS*AEERLHSMK	1	1.27	27.23	3	4
6	α_{S1}-CN 104–142 (1P)	YKVPQLEIVPNS*AEERLHSMKEGIHAQQKEPMIGVNQEL	1	0.25	29.64	3	6
7	α_{S1}-CN 104–143 (1P)	YKVPQLEIVPNS*AEERLHSMKEGIHAQQKEPMIGVNQELA	1	0.07	29.43	3	6
8	α_{S1}-CN 104–149 (1P)	YKVPQLEIVPNS*AEERLHSMKEGIHAQQKEPMIGVNQELAYFYPEL	1	1.71	35.14	3	6
9	β-CN 29–52 (1P)	INKKIEKFQS*EEQQQTEDELQDKIHPFAQTQ	1	2.97	34.36	3	5
10	β-CN 29–56 (1P)	KIEKFQS*EEQQQTEDELQDKIHPF	1	2.60	27.44	2	4
11	β-CN 30–40 (1P)	KIEKFQS*EEQQQTEDELQDKIHPFAQTQ	1	2.37	25.79	3	4
12	β-CN 30–51 (1P)	IEKFQS*EEQQQTEDELQDKIHP	1	1.08	23.72	3	3
13	β-CN 30–52 (1P)	IEKFQS*EEQQQTEDELQDKIHPF	1	1.50	27.23	3	3
14	β-CN 33–51 (1P)	FQS*EEQQQTEDELQDKIHP	1	2.11	22.80	2	2
15	β-CN 33–52 (1P)	FQS*EEQQQTEDELQDKIHPF	1	2.63	26.72	3	2
16	α_{S1}-CN 37–58 (2P)	VNELSKDIGS*ES*TEDQAMEDIK	2	−1.08	22.61	2	2
17	α_{S1}-CN 37–59 (3P)	VNELS*KDIGS*ES*TEDQAMEDIKQ	2	−0.94	21.97	2	2
18	α_{S1}-CN 43–58 (2P)	DIGS*ES*TEDQAMEDIK	2	−0.75	20.30	2	1
19	α_{S2}-CN 115–149 (3P)	NAVPITPTLNREQLS*TS*EENSKKTVDMES*TEVFTK	3	−1.44	24.44	3	4
20	α_{S2}-CN 115–150 (3P)	NAVPITPTLNREQLS*TS*EENSKKTVDMES*TEVFTKK	3	−1.03	24.69	3	5
21	α_{S2}-CN 115–151 (3P)	NAVPITPTLNREQLS*TS*EENSKKTVDMES*TEVFTKKT	3	−0.91	24.36	3	5
22	α_{S2}-CN 115–152 (3P)	NAVPITPTLNREQLS*TS*EENSKKTVDMES*TEVFTKKTK	3	−0.91	23.65	3	6
23	α_{S1}-CN 35–52 (3P)	EKVNELS*KDIGS*ES*TEDQ	3	−4.02	15.46	2	2

(*Continued*)

TABLE 2 Continued

No.	Protein fragment	Peptide sequence[a]	*Sp*[b]	ΔRT[c]	RT[d]	Z[e]	N_b^f
24	α_{S1}-CN 35–58 (3P)	EKVNELS*KDIGS*ES*TEDQAMEDIK	3	−1.19	21.87	3	3
25	α_{S1}-CN 35–59 (3P)	EKVNELS*KDIGS*ES*TEDQAMEDIKQ	3	−0.45	21.87	2	3
26	α_{S2}-CN 1–19 (4P)	KNTMEHVS*S*S*EESIIS*QET	4	−0.90	19.42	2	2
27	α_{S2}-CN 1–24 (4P)	KNTMEHVS*S*S*EESIIS*QETYKQEK	4	0.87	20.14	2	4
28	β-CN 1–24 (4P)	RELEELNVPGEIVES*LS*S*S*EESIT	4	−3.91	26.72	2	1
29	β-CN 1–25 (4P)	RELEELNVPGEIVES*LS*S*S*EESITR	4	−3.82	26.29	2	2
30	β-CN 1–28 (4P)	RELEELNVPGEIVES*LS*S*S*EESITRINK	4	−3.27	27.25	2	3
31	β-CN 1–29 (4P)	RELEELNVPGEIVES*LS*S*S*EESITRINKK	4	−2.27	27.32	3	4
32	β-CN 1–32 (4P)	RELEELNVPGEIVES*LS*S*S*EESITRINKKIEK	4	−2.47	28.05	3	5

[a]S*corresponds to phosphoserine.
[b]The number of phosphorylated residues.
[c]The differences in retention time (experimental RT (phosphopeptides) − predictive RT (unphosphorylated cognates)).
[d]The experimental RT of phosphopeptides.
[e]The charge state of phosphopeptides in the electrospray-ionization process.
[f]The number of basic residues (Lys, Arg, His).

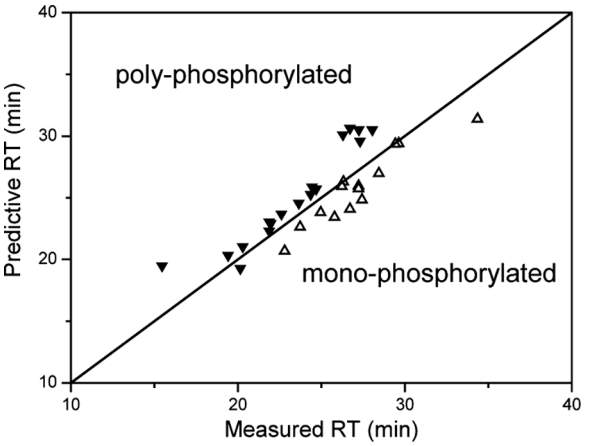

FIGURE 5 Comparison between the experimental retention time (*RT*) of phosphopeptides and the predictive *RT* of their unphosphorylated cognates.

as shown in Table 2. It was assumed in our study that the unmodified cognates of the phosphopeptides might also have 2–3 positive charges in the elutant solution including TFA, though they possessed 1–6 basic residues in the sequence. In the case of monophosphopeptides, the presence of one phosphor moiety with two negative charges might enhance the overall hydrophobicity of peptides due to charge neutralization (i.e., the reduction in the net charge). As a result, single phosphorylation increased the retention of peptides in ion-pairing RPLC (Table 2). On the other hand, the excessive complementation because of multiple phosphorylation might reduce the overall hydrophobicity of the peptide, and, thus, decrease its retention time in the RPLC. In addition, the more phosphorylation sites would result in the more negative charges in the peptide and the greater decrease in peptide retention. In this work, the quadruply phosphorylated peptides, except α_{s2}-CN (1–24) 4P, had the most remarkable negative differences among the phosphopeptides.

CONCLUSIONS

The effect of phosphorylation on the retention behavior of peptides in ion-pairing reversed phase HPLC was investigated via a new method. A prediction model was developed on the basis of 358 non-phosphorylated peptides from the casein hydrolysates, with the retention coefficient and the length of peptide taken into account. Such a model can be used to accurately predict the retention time (*RT*) of unmodified peptides on C4 reversed phase (300 Å pore size) columns with TFA as the ion-pairing reagent. Since TFA is a strong ion-pairing reagent for peptides during

RPLC, the positive charges of peptides would be partly masked by it. The monophosphorylated peptides eluted after their unphosphorylated analogues, while most of the poly-phosphopeptides eluted earlier than their unmodified counterparts. The single and multiple phosphorylation generally led to an increase or reduction in the overall hydrophobicity of peptides, respectively.

ACKNOWLEDGMENTS

The authors thank the financial supports from the Natural Science Foundation of China (No. 20806057) and the Ministry of Education (No. 200800561004). This research is also supported by the National High Technology Research and Development Program of China (863 Program) (No. 2008AA10Z318), and the Program for Changjiang Scholars and Innovative Research Team in University of China (IRT0641).

REFERENCES

1. Blume-Jensen, P.; Hunter, T. Oncogenic kinase signalling. Nature **2001**, *411*, 355–365.
2. Hermiston, M.L.; Xu, Z.; Weiss, A. CD45: A critical regulator of signaling thresholds in immune cells. Annu. Rev. Immunol. **2003**, *21*, 107–137.
3. Lee, H.G.; Perry, G.; Moreira, P.I.; Garrett, M.R.; Liu, Q.; Zhu, X.W.; Takeda, A.; Nunomura, A.; Smith, M.A. Tau phosphorylation in Alzheimer's disease: pathogen or protector? Trends Mol. Med. **2005**, *11*, 164–169.
4. Graves, J.D.; Krebs, E.G. Protein phosphorylation and signal transduction. Pharmacol. Ther. **1999**, *82*, 111–121.
5. Hunter, T. Signaling - 2000 and beyond. Cell **2000**, *100*, 113–127.
6. Manning, G.; Whyte, D.B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The protein kinase complement of the human genome. Science **2002**, *298*, 1912–1934.
7. McCormick, D.J.; Holmes, M.W.; Muddiman, D.C.; Madden, B.J. Mapping sites of protein phosphorylation by mass spectrometry utilizing a chemical enzymatic approach: characterization of products from alpha-S1Casein phosphopeptides. J. Proteome Res. **2005**, *4*, 424–434.
8. Yan, J.X.; Packer, N.H.; Gooley, A.A.; Williams, K.L. Protein phosphorylation: technologies for the identification of phosphoamino acids. J. Chromatogr. A **1998**, *808*, 23–41.
9. McLachlin, D.T.; Chait, B.T. Analysis of phosphorylated proteins and peptides by mass spectrometry. Curr. Opin. Chem. Biol. **2001**, *5*, 591–602.
10. Beausoleil, S.A.; Jedrychowski, M.; Schwartz, D.; Elias, J.E.; Villen, J.; Li, J.X.; Cohn, M.A.; Cantley, L.C.; Gygi, S.P. Large-scale characterization of HeLa cell nuclear phosphoproteins. Proc. Natl. Acad. Sci. USA **2004**, *101*, 12130–12135.
11. Aebersold, R.; Goodlett, D.R. Mass spectrometry in proteomics. Chem. Rev. **2001**, *101*, 269–295.
12. Cantin, G.T.; Yates, J.R. Strategies for shotgun identification of post-translational modifications by mass spectrometry. J. Chromatogr. A **2004**, *1053*, 7–14.
13. Garcia, B.A.; Shabanowitz, J.; Hunt, D.F. Analysis of protein phosphorylation by mass spectrometry. Methods **2005**, *35*, 256–264.
14. Hoffmann, R.; Segal, M.; Otvos, L. Separation of sets of mono- and diphosphorylated peptides by reversed-phase high performance liquid chromatography. Anal. Chim. Acta **1997**, *352*, 327–333.
15. Tholey, A.; Toll, H.; Huber, C.G. Separation and detection of phosphorylated and nonphosphorylated peptides in liquid chromatography–mass spectrometry using monolithic columns and acidic or alkaline mobile phases. Anal. Chem. **2005**, *77*, 4618–4625.

16. Kim, J.; Petritis, K.; Shen, W.; Camp, D.G.; Moore, R.J.; Smith, R.D. Phosphopeptide elution times in reversed-phase liquid chromatography. J. Chromatogr. A **2007**, *1172*, 9–18.
17. Steen, H.; Jebanathirajah, J.A.; Rush, J.; Morrice, N.; Kirschner, M.W. Phosphorylation analysis by mass spectrometry – Myths, facts, and the consequences for qualitative and quantitative measurements. Mol. Cell Proteomics **2006**, *5*, 172–181.
18. Krokhin, O.V.; Craig, R.; Spicer, V.; Ens, W.; Standing, K.G.; Beavis, R.C.; Wilkins, J.A. An improved model for prediction of retention times of tryptic peptides in ion pair reversed-phase HPLC – Its application to protein peptide mapping by off-line HPLC-MALDI MS. Mol. Cell Proteomics, **2004**, *3*, 908–919.
19. Palmblad, M.; Ramstrom, M.; Bailey, C.G.; McCutchen-Maloney, S.L.; Bergquist, J.; Zeller, L.C. Protein identification by liquid chromatography–mass spectrometry using retention time prediction. J. Chromatogr. B **2004**, *803*, 131–135.
20. Palmblad, M.; Ramstrom, M.; Markides, K.E.; Hakansson, P.; Bergquist, J. Prediction of chromatographic retention and protein identification in liquid chromatography/mass spectrometry. Anal. Chem, **2002**, *74*, 5826–5830.
21. Su, R.X.; Qi, W.; He, Z.M.; Yuan, S.X.; Zhang, Y.B. Pancreatic hydrolysis of bovine casein: Identification and release kinetics of phosphopeptides. Food Chem. **2007**, *104*, 276–286.
22. Su, R.X.; Qi, W.; He, Z.M. Time-dependent nature in peptic hydrolysis of native bovine hemoglobin. Eur. Food Res. Technol. **2007**, *225*, 637–647.
23. Liebler, D.C.; Hansen, B.T.; Davey, S.W.; Tiscareno, L.; Mason, D.E. Peptide sequence motif analysis of tandem MS data with the SALSA algorithm. Anal. Chem. **2002**, *74*, 203–210.
24. Eng, J.K.; McCormack, A.L.; Yates, J.R. An approach to correlate tandem mass spectral data of peptides with amino-acid-sequences in a protein database. J. Am. Soc. Mass Spectrom. **1994**, *5*, 976–989.
25. Guo, D.; T., M.C.; Taneja, A.K.; Parker, J.M.R.; Hodges, R.S. Prediction of peptide retention times in reversed-phase high-performance liquid chromatography I. Determination of retention coefficients of amino acid residues of model sythetic peptides. J. Chromatogr. **1986**, *359*, 499–517.
26. Meek, J.L. Prediction of peptide retention times in high-pressure liquid chromatography on the basis of amino acid composition. Proc. Natl. Acad. Sci. USA **1980**, *77*, 1632–1636.
27. Krokhin, O.V. Sequence-specific retention calculator. Algorithm for peptide retention prediction in ion-pair RP-HPLC: Application to 300- and 100-angstrom pore size C_{18} sorbents. Anal. Chem. **2006**, *78*, 7785–7795.
28. Mant, C.T.; Burke, T.W.L.; Black, J.A.; Hodges, R.S. Effect of peptide chain length on peptide retention behavior in reversed-phase chromatography. J. Chromatogr. **1988**, *458*, 193–205.