ACTA SCIENTIARUM POLONORUM *Acta Sci. Pol., Technol. Aliment. 10(2) 2011, 209-221*
ISSN 1644-0730 (print) ISSN 1889-9594 (online)

DETERMINATION OF THEORETICAL RETENTION TIMES FOR PEPTIDES ANALYZED BY REVERSED--PHASE HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY*

Jerzy Dziuba, Piotr Minkiewicz, Damir Mogut
University of Warmia and Mazury in Olsztyn

Background. Peptides are important components of foods mainly due to their biological activity. The basic method of their identification is reversed phase high-performance liquid chromatography coupled with electrospray-ionization mass spectrometry (RP-HPLC--ESI-MS). Retention time (t_R) prediction *in silico* is very helpful in analysis of multicomponent peptide mixtures. One of problems associated with RP-HPLC-ESI-MS is deterioration of mass spectra quality by trifluoroacetic acid (TFA). This problem can be avoided through the use of chromatographic columns designed for work with low TFA concentrations in mobile phase. The objective of this study was to determine the correlations between peptide retention times predicted with the use of a program available on-line and experimental retention times obtained using the column working with low TFA concentrations.
Material and methods. The set of synthetic peptides and bovine α-lactalbumin fragments (18 peptides) was used in the experiment. Theoretical retention times were calculated using Sequence Specific Retention Calculator (SSRC) program. The experimental retention times were measured via RP-HPLC-ESI-MS method using column working with low TFA content. The dependence between theoretical and experimental t_R was expressed via empirical equations.
Results. The best correlation between theoretical and experimental retention times of peptides containing at least four amino acid residues has been obtained when third order polynomial (R^2 = 0.9536). Prediction quality for di- and tripeptides was significantly lower. The method described can be applied for cysteine-containing peptides although our sample preparation procedure did not include modification of this amino acid, taken into attention by SSRC program.

*This study was financed by the Ministry of Science and Higher Education under research project No. NN 312 286 333.

© Copyright by Wydawnictwo Uniwersytetu Przyrodniczego w Poznaniu

Corresponding author – Adres do korespondencji: Prof. dr hab. Jerzy Dziuba, Department of Food Biochemistry of University of Warmia and Mazury in Olsztyn, Plac Cieszyński 1, 10-726 Olsztyn-Kortowo, Poland, e-mail: jerzy.dziuba@uwm.edu.pl

Conclusions. The results of this study validate the usefulness of a third degree polynomial as a simple function describing the correlation between peptide retention times predicted by an on-line application and experimental retention times. The above function can effectively predict retention times in situations when experimental conditions differ from the computational environment (various columns, mobile phase composition, use or resignation from chemical modifications during sample preparation, various HPLC equipments). On-line available t_R predicting application with correction based on user's own data may be a useful tool in food peptidomics.

Key words: peptides, RP-HPLC, prediction of retention times, high-performance liquid chromatography, polynomial functions

INTRODUCTION

Recent achievements in proteomics and peptidomics have expanded our knowledge base of proteins and the products of their hydrolysis. *In silico*, *in vitro* and *in vivo* experiments have demonstrated that many products of specific protein proteolysis may have a positive or a negative effect on the human body [Minkiewicz et al. 2008 a, Dziuba et al. 2009, Iwaniak and Dziuba 2009]. Peptides are also involved in food nanotechnology [Ozimek et al. 2010].

Reversed-phase high-performance liquid chromatography with mass spectrometry (RP-HPLC-MS) is the chief method applied in peptidomic research and one of the key techniques in proteomic studies [Boonen et al. 2008, Minkiewicz et al. 2008 a, Bączek and Kaliszan 2009, Schmidt et al. 2009]. The most popular method of analyte ionization in LC-MS is electrospray ionization (ESI). In this technique, the quality of the obtained mass spectra is adversely affected by trifluoroacetic acid (TFA). In reversed-phase chromatography, the standard TFA concentration is 0.1% (v/v). This quantity of TFA deteriorates the quality of ESI-MS spectra. TFA forms strong ion pairs with analytes. These ion pairs cannot be broken with satisfactory yield in ESI analysis conditions. TFA causes also spray instability. [Apffel et al. 1995, García 2005]. The above problem can be avoided through the use of chromatographic columns designed for low TFA concentrations (0.01% v/v), such as the Jupiter Proteo 90 Å column [Phenomenex 2009].

Recent years have witnessed the growing popularity of a strategy termed hypothesis--driven proteomics [Schmidt et al. 2009]. This strategy includes the identification, quantitative determination, investigation of the kinetics of hydrolysis and modification of peptides selected based on theoretical presumptions. Such peptides include e.g. markers of allergenic proteins in food products.

The prediction of peptide retention times is a useful feature of the discussed strategy [Shinoda et al. 2008, Bączek and Kaliszan 2009]. Retention time predictions may also be used in experiments involving chemometric analyses of peptide mixtures, such as analyses of milk and soy protein hydrolysates [Dziuba et al. 2004]. An extensive body of research has been conducted into predicting peptide retention times and the effect of various columns and analytical environments on retention times. The results of such experiments were published by, among others, Meek [1980], Casal et al. [1996], Palmblad et al. [2002, 2009], Petritis et al. [2003, 2006], Krokhin et al. [2004 a, b], Bączek et al. [2005], Kawakami et al. [2005], Wang et al. [2005 a, b], Tarasova et al. [2009], Bączek and Kaliszan [2009], Vu et al. [2010], Bodzioch et al. [2010]. The Se-

quence Specific Retention Calculator application which predicts retention times based on amino acid sequences is available on-line [Krokhin 2006, Spicer et al. 2007]. Yet none of the existing publications cite data relating to the prediction of retention times for peptides analyzed with the use of special columns designed for low TFA concentrations.

The objective of this study was to determine the correlations between peptide retention times predicted with the use of a program available on-line and experimental retention times obtained using the Jupiter Proteo HPLC column with pore diameter 90Å, designed for solutions with low trifluoroacetic acid concentrations.

MATERIAL AND METHODS

The chromatographic analysis was carried out with the involvement of Sigma-Aldrich peptides (the amino acid sequence is given as a one-letter code): dipeptide AA (cat. no A9502), dipeptide AL (cat. no A1878), tripeptide AAA (cat. no A9627), tripeptide EVF (cat. no G3751), tripeptide LLL (cat. no L0879), human β-casomorphin YPFVEPI (cat. no C0783), β-lipotropin fragment 39-45 KKDSGPY (cat. no L2759), R^8-vasotocin-GKR CYIQNCPRGGKR (V1258), Y^{11}-somatostatin AGCKNFFWKTYTSC (cat. no S8508), calcineurin substrate DLDVPIPGRFDRRVSVAAE (cat. No C5207), standard HPLC peptide mixture: GY, VYV, YGGFM, YGGFL, DRVYIHPF (cat. No H2016). Bovine α-lactalbumin (α-La) was also supplied by Sigma (cat. no L6010). Protein hydrolysis was performed using Promega trypsin (EC 3.4.21.4; number 12 and S01.151, respectively, in the BIOPEP database at http://www.uwm.edu.pl/biochemia [Minkiewicz et al. 2008 b] and the MEROPS database at http://merops.sanger.ac.uk/ [Rawlings 2009, Rawlings et al. 2010]): Sequencing Grade Modified Trypsin, frozen with a resuspension buffer (Promega, cat. no V5113), and the following Sigma-Aldrich reagents: urea (cat. no U6504), Trizma® base solution (cat. no T1699), dithiothreitol (DTT; cat. no 43817), calcium chloride, 25% ammonia p.a. (P.O.Ch. sp. z o.o.) and sodium azide (P.P.H. Stanlab s.c.; cat. no 00 498/7) were used. Water deionized using Synergy UV (Millipore) device, was applied for HPLC with Lab-Scan acetonitrile (ACN; HPLC Gradient Grade, cat. no C73C11X) and J.T. Baker trifluoroacetic acid (TFA; cat. no 9470-01).

An *in silico* simulation of α-La proteolysis was performed using the PeptideMass program available on-line at http://expasy.org/tools/peptide-mass.html [Gasteiger et al. 2005] based on protein sequences from the UniProt database [Jain et al. 2009] (No. P00711). Theoretical retention times of standard peptides and the anticipated products of protein hydrolysis were computed on-line using the Sequence Specific Retention Calculator application at http://hs2.proteome.ca/SSRCalc/SSRCalc.html [Krokhin 2006, Spicer et al. 2007], relying on an algorithm proposed by Krokhin et al. [2004 a, b]. Parameter a, i.e. the retention time of the injection peak (the first peak containing non-adsorbing substances), was 2.02, and parameter b was determined at 0.94, corresponding to the rate of changes in the concentrations of acetonitrile – 0.66% per minute (solvent B, Table 1). The third parameter was the pore diameter. 100 Å was selected from among the available pore diameters in the application as it most closely correspondent to the pore diameter in the applied column (90 Å).

α-La hydrolysis was performed according to the instructions supplied by Promega [2003]. For initial protein denaturation, 150 μl buffer containing 6 M urea, 50 mM Tris

Table 1. Gradient of solvent B applied in the analysis

Time, min	Solvent B concentration, %	Stage
0	0	analysis
60	40	
65	100	column rinsing
70	100	
71	0	column equilibration
80	0	

Table 2. Analyzed peptide data

Peptide sequence[1]	Position in α-La[2] chain	Calculated $[M + H]^+$ [Da]	Measured $[M + H]^+$ [Da]	t_{Rpred} min	t_{Rexp} min	S[3] min	n[4]
AA (2)	–	160.17	161.2	3.40	3.68	0.05	6
AAA (3)	–	231.25	232.2	2.70	4.81	0.18	6
GY (2)	–	238.24	239.1	2.60	12.61	0.30	6
KKDSGPY (7)	–	793.86	794.6	4.70	13.37	0.31	6
AL (2)	–	202.25	203.1	3.00	15.56	0.39	6
CYIQNCPRGGKR (12)	–	1 392.61	1 393.2	8.50	19.24	0.31	3
DDQNPHSSNICNISCDK (17)	82-98	1 888.77	1 889.8	12.30	21.51	0.18	3
CEVFR (5)	25-29	652.30	653.3	16.80	23.49	0.68	3
IWCK (4)	78-81	548.28	549.3	16.30	26.20	1.36	3
DRVYIHPF (8)	–	1 046.18	1 046.9	25.40	32.68	0.30	6
LDQWLCEK (8)	134-141	1 033.49	1 034.5	25.40	32.70	0.24	3
YGGFM (5)	–	573.66	574.2	25.10	33.87	0.28	6
VGINYWLAHK (10)	118-127	1 199.65	1 200.6	28.70	35.09	0.11	3
DLDVPIPGRFDRRVSVAAE (19)	–	1 056.74[5] $[M + 2H]^{2+}$	1 057.4[5] $[M + 2H]^{2+}$	30.20	35.43	0.19	6
AGCKNFFWKTYTSC (14)	–	1653.88	1 654.3	27.50	37.64	0.26	6
YGGFL (5)	–	555.62	556.3	26.70	37.68	0.32	6
YPFVEPI (7)	–	864.00	864.7	31.80	41.72	0.25	6
FLDDDLTDDIMCVK (14)	99-112	1 641.73	1 642.7	31.50	44.20	0.04	3

The data for peptides EGF, LLL and VYF, for which the predicted retention times were shorter than the retention times for non-adsorbing substances, were disregarded.

[1]The number of amino acid residues in the peptide chain is shown in parentheses.

[2]The position in the chain is shown only for α-La fragments.

[3]S – standard deviation.

[4]n – number of measurements.

[5]Peptide molecular mass was higher than the maximum mass-to-charge ratio that could be measured with the applied instrument. For this reason, the mass of the double-ionized peptide was measured.

and 2 mM DTT were added to 1 mg protein. The sample was heated at 95°C for 20 min. After cooling, 850 μl buffer containing 50 mM Tris and 1 mM $CaCl_2$ were added to the sample. 20 μg trypsin was dissolved in 200 μg resuspension buffer and left for 15 min at 37°C. Protein and enzyme solutions were mixed and incubated for 13 h at 37°C. 5 μg TFA was added to inactivate the enzyme. The precipitate was centrifuged (10 min, 10 000 × G). The supernatant was analyzed by chromatography. The samples of Sigma--Aldrich standard peptides were dissolved in DI water. Peptide concentration was 1 mg·ml^{-1}.

The analysis was performed by mass spectrometry using electrospray ionization with an ion trap (ESI-IT-MS) supplied by Varian, a chromatography kit comprising two 212-LC pumps, a ProStar Model 410 autosampler with a thermostatic chamber, a Degassit degasser supplied by MetaChem Tech. Inc. and the LC/MS 12-2 nitrogen generator (Domnick Hunter). The results were stored and processed with the use of the MS WorkStation v. 6.9 software. Chromatographic separation was performed using the Jupiter Proteo 250 × 2 mm column with 90 Å pore diameter and 4 μm particle diameter, designed for use with low TFA concentrations. The analysis was carried out with the use of 0.01% TFA/DI water solution (solvent A) and 0.01% TFA/acetonitrile solution (solvent B). Total flow reached 200 μL·min^{-1}. Data collection time was 60 min. The gradient of solvent B is presented in Table 1.

Peptides were identified based on the mass-to-charge ratio. The mass spectrometer was set for positive polarization with ionization current of 600V. The capillary potential was 100V, needle potential – 5000V and disk target potential – 6000V. The spraying and drying gas was nitrogen at the pressure of 55 psi and 30 psi, respectively. The flow of helium, applied as the damping gas, was 1.5 ml·min^{-1}. The analyzed masses were in the range of 100 to 2000 m·z^{-1}. The results were processed in the Excel™ application (Microsoft). The number of replications for each peptide is presented in Table 2.

RESULTS AND DISCUSSION

Figure 1 a and b present an unprocessed peak corresponding to peptide with sequence AL (Table 2), smoothed with an algorithm proposed by Savitzky and Golay [1964] with the use of 3 and 11 points, respectively. The peaks produced by smoothing with 11 points have single vertices, thus supporting an unambiguous and repeated determination of retention times. As regards one standard peptide, chromatographic peaks with two vertices (Fig. 1 a) were artifacts. Successive calculations were performed using retention times corresponding to the vertices of peaks smoothed with the Savitzky-Golay algorithm, with the use of 11 points.

Figure 2 presents chromatograms of standard peptides (a) and the hydrolysis products of α-lactalbumin (b). Their retention times are shown in Table 2 which also indicates the mass of the analyzed peptides and standard deviation for retention times from several replications. The position of peptides produced from α-lactalbumin in the protein chain is also indicated. The group of peptides determined in the α-La hydrolysate differed from that reported by Maynard et al. [1997] and Wehbi et al. [2006]. The above differences could have resulted from the adoption of a different hydrolysis procedure than applied in this experiment. The group of peptide markers specific to a given protein (proteotypic peptides) is determined by the hydrolysis procedure, the detection method

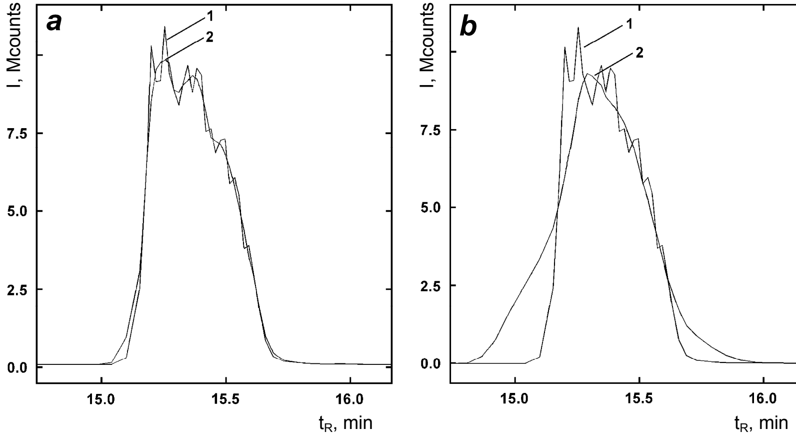

Fig. 1. Smoothing a chromatographic peak by the method proposed by Savitzky and Golay [1964]: a – a peak smoothed with the use of 3 points, b – a peak smoothed with the use of 11 points; numbers 1 and 2 in the figure drawings denote peaks before and after smoothing, respectively

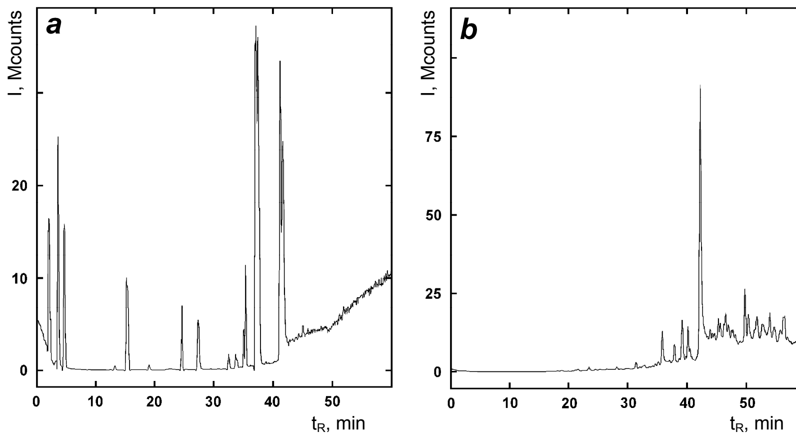

Fig. 2. Chromatograms of standard peptides (a) and the hydrolysis products of α-lactalbumin (b)

and the applied mass spectrometer [Mallick et al. 2007]. Delected α-La fragments (25-29 and 78-81) are fragments of that protein's epitopes described by Maynard et al. [1997] (fragments 17-58 and 59-93 or 59-94 respectively).

The data presented in Table 2 were used to calculate regression equations determining the correlation between peptide retention times computed in the SSRC application and the experimental retention times. The coefficients of determination (R^2) for first, second, third and fourth degree polynomial equations were 0.9339, 0.9342, 0.9386 and 0.9386, respectively. In the analyzed group, third and fourth degree polynomials

were characterised by the slightly better fit, than the linear or second degree equation. The third degree polynomial (equation 1) offers a simpler solution, and it can be recommended for practical application.

$$t_{Rexp} = 0.0014t_{Rpred}^{3} - 0.0767t_{Rpred}^{2} + 2.1876t_{Rpred} + 3.7872 \quad (1)$$

$$R^2 = 0.9386$$

where:

t_{Rpred} – predicted retention time calculated by the Sequence Specific Retention Calculator (SSRC) application,
t_{Rexp} – experimental retention time,
R^2 – coefficient of determination.

The curve described by the above equation is presented in Figure 3. The curve's goodness of fit to peptides of various length can be evaluated with parameter U, determined based on the following equation:

$$U = |t_{Rexp} - t_{Rcor}|/t_{Rcor} \quad (2)$$

where:

t_{Rcor} – predicted retention time calculated by the SSRC application and adjusted based on equation 1,
t_{Rexp} – experimental retention time.

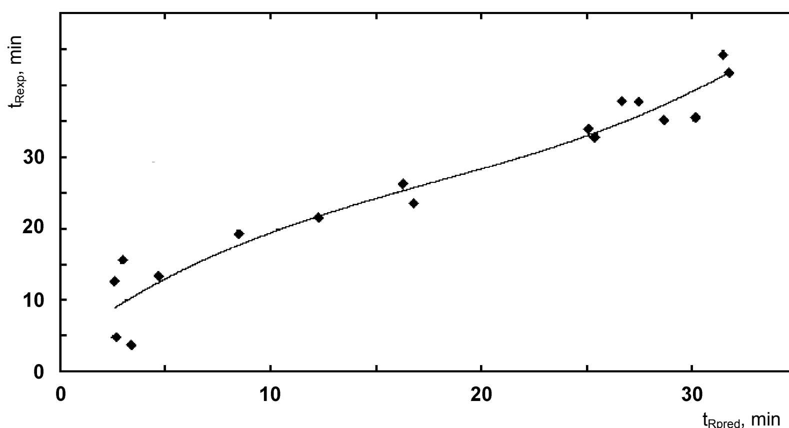

Fig. 3. Diagram of correlations between peptide retention times calculated by the SSRCalc application and experimental retention times obtained with the Jupiter Proteo HPLC column (equation 1)

The values of parameter U for peptides with varied chain length are presented in Figure 4. This figure drawing indicates that the values of U for dipeptides and tripeptides are much higher than for compounds containing a higher number of amino acid residues. The average values of parameter U reached 0.53 ±0.11 (n = 4) and 0.05 ±0.03 (n = 14) for peptides containing two and three amino acid residues, respectively, and for longer peptides (n is the number of peptides). The difference is statistically significant

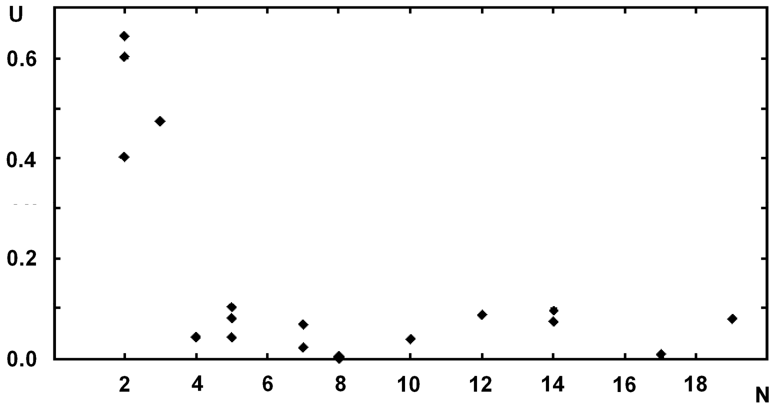

Fig. 4. Correlation between parameter U (equation 2) and the remaining amino acid residues in a peptide molecule. N – number of amino acid residues

at $p = 0.01$. The predicted retention times for three tripeptides were shorter than the retention time noted for a non-retained compound. Those peptides were disregarded in the curve mapping process.

Some peptides summarized in the Table 2 contain cysteine. Calculation procedure described by Krokhin et al. [2004 a] has been designed for the peptides modified by reaction of cysteine sulfhydryl groups with iodoacetamide. In our experiment sulfhydryl groups of cysteine were not modified. The equation proposed take into attention all possible causes of retention time shift due to changes in experimental procedure such as application of different columns, mobile phase composition, as well as presence or absence of modification of sulfhydryl groups.

The location of points corresponding to the predicted and the experimental retention times of each compound demonstrates that the analyzed equation effectively illustrates the behavior of compounds containing four and more amino acid residues, while dipeptides and tripeptides are described less effectively. Krokhin et al. [2004 a] described similar limitations for the original algorithm used to develop the SSRC program. This limitation cannot be eliminated by adjusting the predicted retention time with the use of a polynomial. The frequent presence of dipeptides and tripeptides in protein sequences renders them useless in proteomic studies, nonetheless, they are an important consideration in nutritional sciences owing to the ease (in comparison with longer peptides) of their absorption from the gastrointestinal tract into the bloodstream, and the demonstrated biological activity [Minkiewicz et al. 2008 a].

The elimination of dipeptide and tripeptide data led to a slight increase in the value of determination coefficients R^2 for equations fitting theoretical and experimental data. A third degree polynomial describing peptides with more than three amino acid residues is presented below:

$$t_{Rexp} = 0.0011t_{Rpred}^3 - 0.0562t_{Rpred}^2 + 1.7428t_{Rpred} + 6.7354 \quad (3)$$

$$R^2 = 0.9536$$

where:

t_{Rpred} – predicted retention time calculated by the SSRC application,
t_{Rexp} – experimental retention time,
R^2 – coefficient of determination.

The coefficients of determination for the linear equation, second degree and fourth degree polynomial equations are 0.9470; 0.9495 and 0.9537, respectively. The diagram of the correlation described by equation 3 is presented in Figure 5.

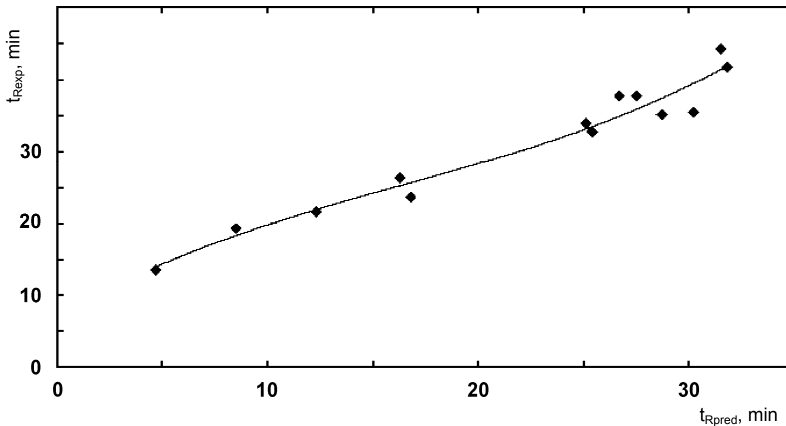

Fig. 5. Diagram of correlations between peptide retention times calculated by the SSRCalc application and experimental retention times obtained with the Jupiter Proteo HPLC column, disregarding dipeptide and tripeptide data (equation 3)

To date, polynomials have not been used as empirical equations describing the correlations between the predicted and experimental retention times of peptides or the correlations between peptide retention times obtained under various analytical conditions. A linear correlation is sometimes observed between retention times produced by different chromatographic columns, as discussed by Tarasova et al. [2009]. In some cases, a polynomial is more effective in describing the correlation between experimental and predicted retention times (expressed as the coefficient of determination or correlation) than a linear equation. Deviations from the linear correlation between experimental and predicted retention times are frequently noted. The situations in which changes of the chromatographic column or the analytical environment modify the peptide outflow sequence have been described by Casal et al. [1996], Minkiewicz et al. [2000], Bączek et al. [2005] and Kawakami et al. [2005].

The resulting equations may be applied to calculate adjustments in the process of predicting the retention times of peptides separated under different analytical conditions than the computational environment of the SSRC application [Krokhin 2006, Spicer et al. 2007]. The algorithm's greatest disadvantage is its relative ineffectiveness in describing peptides with two or three amino acid residues.

CONCLUSIONS

The results of this study validate the usefulness of a third degree polynomial as a simple function describing the correlation between peptide retention times predicted by an on-line application and experimental retention times. The above function can effectively predict retention times in situations when experimental conditions differ from the computational environment. The described correlation may be recommended as a useful tool for predicting the retention times of peptides containing at least four amino acid residues. On-line available t_R predicting application with correction based on user's own data may be useful tool in food peptidomics.

REFERENCES

Apffel A., Fischer S., Goldberg G., Goodley P.C., Kuhlmann F.E., 1995. Enhanced sensitivity for peptide mapping with electrospray liquid chromatography-mass spectrometry in the presence of signal suppression due to trifluoroacetic acid-containing mobile phases. J. Chromatogr. A, 712, 177-190.

Bączek T., Kaliszan R., 2009. Predictions of peptide's retention times in reversed-phase liquid chromatography as a new supportive tool to improve protein identification in proteomics. Proteomics 9, 835-847.

Bączek T., Wiczling P., Marszałł M., Van der Heyden Y., Kaliszan R., 2005. Prediction of peptide retention at different HPLC conditions from multiple linear regression models. J. Proteome Res. 4, 555-563.

Bodzioch K., Durand A., Kaliszan R., Bączek T., Van der Heyden Y., 2010. Advanced QSRR modeling of peptides behavior in RPLC. Talanta 81, 1711-1718.

Boonen K., Landuyt B., Baggerman G., Husson S.J., Huybrechts J., Schoofs L., 2008. Peptidomics: the integrated approach of MS, hyphenated techniques and bioinformatics for neuropeptide analysis. J. Separ. Sci. 31, 427-445.

Casal V., Martin-Alvarez P.J., Herraiz T., 1996. Comparative prediction of the retention behaviour of small peptides in several reversed-phase high-performance liquid chromatography columns by using partial least squares and multiple linear regression. Anal. Chim. Acta 326, 77-84.

Dziuba M., Dziuba B., Iwaniak A., 2009. Milk proteins as precursors of bioactive peptides. Acta Sci. Pol., Technol. Aliment. 8 (1), 71-90.

Dziuba J., Nałęcz D., Minkiewicz P., Dziuba B., 2004. Identification and determination of milk and soybean protein preparations using enzymatic hydrolysis followed by chromatography and chemometrical data analysis. Anal. Chim. Acta 521, 17-24.

García M.C., 2005, The effect of the mobile phase additives on sensitivity in the analysis of peptides and proteins by high-performance liquid chromatography-electrospray mass spectrometry. J. Chromatogr. B, 825, 111-123.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A., 2005. Protein identification and analysis tools on the ExPASy server. In: The proteomics protocols handbook. Ed. J.M. Walker. Humana Press. Totowa, USA, 571-607.

Iwaniak A., Dziuba B., 2009. Motifs with potential physiological activity in food proteins – BIOPEP database. Acta Sci. Pol., Technol. Aliment. 8 (3), 59-85.

Jain E., Bairoch A., Duvaud S., Phan I., Redaschi N., Suzek B.E., Martin M.J., McGarvey P., Gasteiger E., 2009. Infrastructure for the life sciences: design and implementation of the UniProt website. BMC Bioinform. art. no 10, 136.

Kawakami T., Tateishi K., Yamano Y., Ishikawa T., Kuroki K., Nishimura T., 2005. Protein identification from product ion spectra of peptides validated by correlation between measured

and predicted elution times in liquid chromatography/mass spectrometry. Proteomics 5, 856--864.

Krokhin O.V., 2006. Sequence Specific Retention Calculator – a novel algorithm for peptide retention prediction in ion-pair RP-HPLC: application to 300 Å and 100 Å pore size C18 sorbents. Anal. Chem. 78, 7785-7795.

Krokhin O.V., Craig R., Spicer V., Ens W., Standing K.G., Beavis R.C., Wilkins J.A., 2004 a. An improved model for prediction of retention times of tryptic peptides in ion pair reversed--phase HPLC. Mol. Cell. Proteom. 3, 908-919.

Krokhin O.V., Ying S., Craig R., Spicer V., Ens W., Standing K.G., Beavis R.C., Wilkins J.A., 2004 b. New sequence-specific correction factors for prediction of peptide retention in RP--HPLC: application to protein identification by off-line HPLC-MALDI-MS. In: Proceedings of the 52-th ASMS Conference on Mass Spectrometry and Allied Topics. Nashville, TN, USA, May 23-27, 2004, TPZ 503, http://hs2.proteome.ca/SSRCalc/OKrokhin_ASMS07.pdf.

Mallick P., Schirle M., Chen S.S., Flory M.R., Martin D., Ranish J., Raught B., Schmitt R., Werner T., Kustr B., Aebersold R., 2007. Computational prediction of proteotypic peptides for quantitative proteomics. Nature Biotechnol. 25, 125-131.

Maynard F., Jost R., Wal J.-M., 1997. Human IgE binding capacity of tryptic peptides from bovine alpha-lactalbumin. Int. Arch. Allergy Immunol. 113, 478-488.

Meek J.L., 1980. Prediction of peptide retention times in high-pressure liquid chromatography on the basis of amino acid composition. Proc. Natl. Acad. Sci. USA 77, 1632-1636.

Minkiewicz P., Dziuba J., Darewicz M., Iwaniak A., Dziuba M., Nałęcz D., 2008 a. Food peptidomics. Food Technol. Biotechnol. 46, 1-10.

Minkiewicz P., Dziuba J., Iwaniak A., Dziuba M., Darewicz M.,, 2008 b. BIOPEP database and other programs for processing bioactive peptide sequences. J. AOAC Int. 91, 965-980.

Minkiewicz P., Slangen C.J., Dziuba J., Visser S., Mioduszewska H., 2000. Identification of peptides obtained via hydrolysis of bovine casein by chymosin using HPLC and mass spectrometry. Milchwissenschaft 55, 14-17.

Ozimek L., Pospiech E., Narine S., 2010. Nanotechnologies in food and meat processing. Acta Sci. Pol., Technol. Aliment. 9 (4), 401-412.

Palmblad M., Ramström M., Markides K.E., Håkansson P., Bergquist J., 2002. Prediction of chromatographic retention and protein identification in liquid chromatography/mass spectrometry. Anal. Chem. 74, 5826-5830.

Palmblad M., Van der Burgt Y.E.M., Dalebout H., Derks R.J.E., Schoenmaker B., Deelder A.M., 2009. Improving mass measurement accuracy in mass spectrometry based proteomics by combining open source tools for chromatographic alignment and internal calibration. J. Proteom. 72, 722-724.

Petritis K., Kangas L.J., Ferguson P.L., Anderson G.A., Paša-Tolić L., Lipton M.S., Auberry K.J., Strittmatter E.F., Shen Y.Y., Zhao R., Smith R.D., 2003. Use of artificial neural networks for the accurate prediction of peptide liquid chromatography elution times in proteome analysis. Anal. Chem. 75, 1039-1048.

Petritis K., Kangas L.J., Yan B., Monroe M., Strittmatter E.F., Qian W.-J., Adkins J.N., Moore R.J., Xu Y., Lipton M.S., Camp D.G., Smith R.D., 2006. Improved peptide elution time prediction for reversed-phase liquid chromatography-MS by incorporating peptide sequence information. Anal. Chem. 78, 5026-5039.

Phenomenex, 2009. Jupiter TM. Application note. http://www.instrument.com.cn/Quotation /Manual/103878.pdf.

Promega, 2003. Sequencing Grade Modified Trypsin. Application note. http://www.promega.com /tbs/9piv511/9piv511.pdf.

Rawlings N.D., 2009. A large and accurate collection of peptidase cleavages in the MEROPS database. Database, 1, Art. No. bap015.

Rawlings N.D., Barrett A.J., Bateman A., 2010. MEROPS: the peptidase database. Nucleic Acids Res. 38, D227-D233.

Savitzky A., Golay M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36, 1627-1638.
Schmidt A., Claassen M., Aebersold R., 2009. Directed mass spectrometry: towards hypothesis-driven proteomics. Curr. Opin. Chem. Biol. 13, 510-517.
Shinoda K., Sugimoto M., Tomita M., Ishihama Y., 2008. Informatics for peptide retention properties in proteomic LC-MS. Proteomics 8, 787-798.
Spicer V., Yamchuk A., Cortens J., Sousa S., Ens W., Standing K.G., Wilkins J.A., Krokhin O.V., 2007. Sequence-specific retention calculator. A family of peptide retention time prediction algorithms in reversed-phase HPLC: applicability to various chromatographic conditions and columns. Anal. Chem. 79, 8762-8768.
Tarasova I.A., Guryča V., Pridatchenko M.L., Gorshkov A.V., Kieffer-Jaquinod S., Evreinov V.V., Masselon C.D., Gorshkov M.V., 2009. Standardization of retention time data for AMT tag proteomics database generation. J. Chromatogr. B, 877, 433-440.
Vu H., Spicer V., Gotfrid A., Krokhin O.V., 2010. A model for predicting slopes S in the basic equation for the linear-solvent-strength theory of peptide separation by reversed-phase high-performance liquid chromatography. J. Chromatogr. A, 1217, 489-497.
Wang Y., Gu X., Zhang J., Zhang X.M., 2005 a. Prediction of peptide retention in RP-LC. Chromatographia 62, 385-392.
Wang Y., Zhang J., Gu X., Zhang X.-M., 2005 b. Protein identification assisted by the prediction of retention time in liquid chromatography/tandem mass spectrometry. J. Chromatogr. B, 826, 122-128.
Wehbi Z., Pérez M.-D., Dalgalarrondo M., Sánchez L., Calvo M., Chobert J.-M., Haertlé T., 2006. Study of ethanol-induced conformational changes of holo and apo α-lactalbumin by spectroscopy and limited proteolysis. Mol. Nutr. Food Res. 50, 34-43.

WYZNACZANIE TEORETYCZNEGO CZASU RETENCJI PEPTYDÓW ANALIZOWANYCH METODĄ WYSOKOSPRAWNEJ CHROMATOGRAFII CIECZOWEJ Z ODWRÓCONYMI FAZAMI

Wstęp. Peptydy są ważnymi składnikami żywności ze względu na ich aktywność biologiczną. Główną metodą ich identyfikacji jest wysokosprawna chromatografia cieczowa z odwróconymi fazami w połączeniu ze spektrometrią mas z jonizacją za pomocą elektrorozpylania (RP-HPLC-ESI-MS). Przewidywanie czasów retencji (t_R) *in silico* jest bardzo przydatne w analizie wieloskładnikowych mieszanin peptydów. Jednym z problemów związanych z zastosowaniem RP-HPLC-ESI-MS jest obniżenie jakości widm masowych spowodowane obecnością kwasu trifluorooctowego (TFA). Wymieniony problem można ominąć, używając kolumn chromatograficznych zaprojektowanych do pracy z małą zawartością TFA w fazie ruchomej. Celem pracy było wyznaczenie zależności między czasami retencji peptydów przewidywanymi za pomocą programu dostępnego on-line a rzeczywistymi czasami retencji uzyskanymi z wykorzystaniem kolumny chromatograficznej przystosowanej do pracy z roztworami o małym stężeniu TFA.
Materiał i metody. W eksperymencie zostały użyte peptydy syntetyczne oraz fragmenty krowiej laktoalbuminy-α (18 peptydów). Teoretyczne czasy retencji peptydów były obliczane za pomocą dostępnego w Internecie programu Sequence Specific Retention Calculator (SSRC). Eksperymentalne czasy retencji były mierzone z użyciem metody RP--HPLC-ESI-MS z zastosowaniem kolumny przystosowanej do pracy z małą zawartością TFA. Zależność między teoretycznymi i eksperymentalnymi czasami retencji została wyrażona za pomocą równań empirycznych.

Wyniki. Najlepszą korelację między teoretycznymi i doświadczalnie wyznaczonymi czasami retencji peptydów zawierających co najmniej cztery reszty aminokwasowe uzyskano, stosując wielomian trzeciego stopnia ($R^2 = 0{,}9536$). Jakość przewidywania czasów retencji dla di- i tripeptydów była znacznie gorsza. Opisana metoda może być stosowana do przewidywania czasów retencji peptydów zawierających reszty cysteiny, mimo że użyta w eksperymencie procedura przygotowania próbek pomija modyfikację tych reszt, uwzględnianą przez program SSRC.
Wnioski. Rezultaty opisane w pracy wskazują na przydatność wielomianu jako prostej funkcji opisującej zależność między czasami retencji peptydów przewidywanymi za pomocą programu dostępnego w Sieci a eksperymentalnymi czasami retencji. Taka funkcja może mieć zastosowanie do przewidywania czasów retencji, jeśli warunki analizy różnią się od tych, które uwzględnia program (zastosowanie różnych kolumn lub składu fazy ruchomej, użycie bądź rezygnacja z chemicznej modyfikacji peptydów podczas przygotowania próbki; zastosowanie różnych zestawów HPLC). Program przewidujący czasy retencji, dostępny w Sieci z poprawką opartą na własnych danych użytkownika, może być użytecznym narzędziem w peptydomice żywności.

Słowa kluczowe: peptydy, RP-HPLC, przewidywanie czasów retencji, wysokosprawna chromatografia cieczowa, funkcje wielomianowe

Received – Przyjęto: 6.01.2011

Accepted for print – Zaakceptowano do druku: 11.02.2011

For citation – Do cytowania: Dziuba J., Minkiewicz P., Mogut D., 2011. Determination of theoretical retention times for peptides analyzed by reversed-phase high-performance liquid chromatography. Acta Sci. Pol., Technol. Aliment. 10(2), 209-221.