ISSN 0006-2979, Biochemistry (Moscow), 2009, Vol. 74, No. 11, pp. 1195-1202. © Pleiades Publishing, Ltd., 2009.
Original Russian Text © M. L. Pridatchenko, I. A. Tarasova, V. Guryca, A. S. Kononikhin, C. Adams, D. A. Tolmachev, A. Yu. Agapov, V. V. Evreinov, I. A. Popov, E. N. Nikolaev, R. A. Zubarev, A. V. Gorshkov, C. D. Masselon, M. V. Gorshkov, 2009, published in Biokhimiya, 2009, Vol. 74, No. 11, pp. 1469-1478.
Originally published in Biochemistry (Moscow) On-Line Papers in Press, as Manuscript BM09-039, July 5, 2009.

Use of Models of Biomacromolecule Separation in AMT Database Generation for Shotgun Proteomics

M. L. Pridatchenko[1], I. A. Tarasova[1], V. Guryca[2], A. S. Kononikhin[1], C. Adams[3], D. A. Tolmachev[1], A. Yu. Agapov[1], V. V. Evreinov[4], I. A. Popov[5], E. N. Nikolaev[5], R. A. Zubarev[3], A. V. Gorshkov[4], C. D. Masselon[2], and M. V. Gorshkov[1]*

[1]Institute for Energy Problems of Chemical Physics, Russian Academy of Sciences, Leninsky pr. 38, 119334 Moscow, Russia; fax: (499) 137-8258; E-mail: mike.gorshkov@gmail.com
[2]CEA, Universite Joseph Fourier, 17 Avenue des Martyrs, Bat. C3, 38054 Grenoble Cedex 9, France; fax: +33 (43) 878-5051
[3]Uppsala University, Institute for Cell and Molecular Biology, Box 596, BMC, SE 75 124, Uppsala, Sweden; fax: +46 (18) 471-7209
[4]Semenov Institute of Chemical Physics, Russian Academy of Sciences, ul. Kosygina 4, 119991 Moscow, Russia; fax: (495) 137-8247
[5]Emmanuel Institute of Biochemical Physics, Russian Academy of Sciences, ul. Kosygina 4, 119334 Moscow, Russia; fax: (495) 137-4101

Received February 5, 2009
Revision received April 20, 2009

Abstract—Generation of a complex proteome database requires use of powerful analytical methods capable of following rapid changes in the proteome due to changing physiological and pathological states of the organism under study. One of the promising technologies with this regard is the use of so-called Accurate Mass and Time (AMT) tag peptide databases. Generation of an AMT database for a complex proteome requires combined efforts by many research groups and laboratories, but the chromatography data resulting from these efforts are tied to the particular experimental conditions and, in general, are not transferable from one platform to another. In this work, we consider an approach to solve this problem that is based on the generation of a universal scale for the chromatography data using a multiple-point normalization method. The method follows from the concept of linear correlation between chromatography data obtained over a wide range of separation parameters. The method is further tested for tryptic peptide mixtures with experimental data collected from mutual studies by different independent research groups using different separation protocols and mass spectrometry data processing tools.

DOI: 10.1134/S0006297909110030

Key words: high performance liquid chromatography, proteomics, mass spectrometry

A promising trend in the quantitative analysis and identification of proteins by chromatography-mass-spectrometry (HPLC/MS) is an approach based on the use of databases of accurate mass and chromatography retention times of peptide markers (or tag) of proteins. Databases of such markers have been called Accurate Mass and Time tag (AMT) [1-4]. Their use makes it possible to significantly increase the throughput of protein identification due to the possibility of carrying out the analysis without involving methods of tandem mass spectrometry (MS/MS).

However, a number of fundamental problems interfered with the development and wide application of AMT tags. Inaccuracies of genomic databases and low efficiency of tandem mass spectrometry used for peptide identification are "translated" into errors in generated AMT databases. Besides, chromatography data depend on specific experimental conditions of separation and used instrumental platforms. As a result, AMT databases compiled by a certain proteomic center cannot be used by other researchers having different experimental systems

Abbreviations: AMT, accurate mass and time; BioLCCC, liquid chromatography of biomacromolecules under critical conditions; LSS, Linear Solvent Strength theory proposed by Snyder; MPN, multi-point normalization; MS/MS, tandem mass spectrometry; SSRCalc, Sequence Specific Retention Calculator.

* To whom correspondence should be addressed.

and protocols. The solution of this problem occurs through the creation of a single scale of peptide retention times that are independent of experimental separation parameters.

There have been numerous attempts to introduce a unique procedure for normalization of chromatographic data. Thus, the procedure of experimental time retention was described [4] that reduced these times to the space [0.1] using linear transformation in which linear equation coefficients are determined empirically using a genetic algorithm based on the iteration optimization process [1, 2, 4]. Another approach to normalization of chromatographic retention times is also based on linear transformation in which coefficients are calculated using a chosen internal standard [5]. The real sample chromatogram is divided into time intervals on the basis of the standard peptide retention times, and linear equation coefficients are calculated for each time interval on the basis of experimental retention times of the internal standard peptides.

It should be noted that the use of so far proposed approaches to retention time standardization results in the situation when the developed chromatographic databases become attached to concrete experimental conditions and used instrumental systems. It is possible to escape this partially by bringing retention times in line with the peptide properties caused by their primary structures. To solve this problem, McIntosh et al. [3] proposed using as normalized retention times the peptide hydrophobicity values calculated using the SSRCalc (Sequence Specific Retention Calculator [6]) prediction algorithm, thus justifying such approach by existence of linear correlation between peptide hydrophobicity and their chromatographic retention times. It should be noted that the SSRCalc algorithm is based on multiparametric optimization of free parameters of the peptide separation model for particular HPLC conditions [7]. An alternative approach to retention time normalization was proposed in the work [8], in which HPLC data are "aligned" on the basis of logarithmic dependence of peptide retention times on the used organic solvent retention coefficient (solvent strength). This procedure is based on the Linear Solvent Strength (LSS) theory proposed by Snyder [9].

In this work the method of multi-point normalization (MPN), earlier proposed by the authors [10] and based on the concept of linearity of chromatographic data obtained under different separation conditions typical of proteomic studies, was used for standardization of peptide retention times. The MPN technique was tested on the example of complex mixtures of proteolytic peptides, experimental data for which were obtained by three independent research groups using different protocols of HPLC separation and mass-spectrometry data processing. This showed the possibility of generating a universal database for accurate mass and retention times suitable for combined use by independent research centers.

MATERIALS AND METHODS

The following commercial standards were used in this work: tryptic hydrolysates of cytochrome *c* and of a mixture of six proteins (bovine serum albumin, β-galactosidase, lysozyme, alcohol dehydrogenase, cytochrome *c*, and apo-transferrin) from Dionex/LCPacking (USA). The concept of chromatographic data linearity was validated using the cytochrome *c* protein hydrolysate in the Laboratory of Proteome Dynamics Investigation of the Ministry of Atomic Energy (Grenoble, France). Chromatography columns used in these studies are listed in Table 1. Experiments on separation and identification of peptides of six protein hydrolysates were carried out at the University of Uppsala (Sweden), the Laboratory of Proteome Dynamics Investigation of the Ministry of Atomic Energy (Grenoble, France), and the Institute of Biochemical Physics of the Russian Academy of Sciences (Moscow, Russia) using different instrumental systems and HPLC protocols listed in Table 2.

In all three laboratories hybrid ion cyclotron resonance mass spectrometers LTQ-FT (ThermoFisher, Germany) equipped with nanoelectrospray and gradient HPLC systems were used to obtain mass-spectrometry

Table 1. Chromatographic columns used in testing the concept of chromatographic data linearity

Column	Manufacturer	Phase	Inner diameter of column, μm	Column length, mm	Particle size, μm	Pore size, Å
I	LC Packing	PepMapC18	75	150	3	100
II	LC Packing	PepMapC18	75	250	3	100
III	Waters	AtlantiesC18	75	150	3	100
IV	LC Packing	PepMapC18	75	150	5	100
V	LC Packing	PLRP-S	75	150	5	300
VI	Merck	C18 Monolith	100	150	–	130

Table 2. Conditions of chromatographic separation used in analysis of standard proteolytic peptide mixture of six proteins (Dionex Inc.) in proteomic laboratories in Grenoble, Uppsala, and Moscow

	LC platform	LC column					LC gradient		Flow rate, nl/min	Specimen amount, fmol	Mobile phase	
		phase	particle size, μm	pore size, Å	inner diameter of column, μm	column length, mm	B, %	time, min			phase A	phase B
Grenoble	Ultimate 3000 (Dionex)	PepMap-C18	3	100	75	150	4-50	60	300	500	CH_3CN: H_2O: HCOOH (2:97.9: 0.1%, v/v)	CH_3CN: H_2O: HCOOH (80:19.92: 0.08%, v/v)
Moscow	Agilent 1100 (Agilent)	Reprosil-Pur-C18 AQ	3	120	75	120	0-35	120	300	250	H_2O: HCOOH (99.9:0.1%, v/v)	CH_3CN: H_2O: HCOOH (80:19.9: 0.1%, v/v)
Uppsala	Agilent 1100 (Agilent)	Reprosil-Pur-C18 AQ	3	120	75	150	2-45	91	200	500	CH_3CN: H_2O: CH_3COOH (2:97.5: 0.5%, v/v)	CH_3CN: H_2O: CH_3COOH (89.5:10: 0.5%, v/v)

data and peptide identifications. Peptide fragmentation was carried out using collision dissociation (CAD) (Moscow, Grenoble, Uppsala) and Electron Capture Dissociation (ECD) (Uppsala). The Mascot search engine was used for peptide identifications on the basis of MS/MS data with searching in the Swiss-Prot database. Only peptides identified in all three laboratories and having Mascot identity scores above 30 were chosen from the search results for testing the proposed procedure for retention time standardization. Experiments with the six-protein standard were successively interchanged with experiments with the cytochrome *c* peptides that were used for control of retention time reproducibility and chromatographic system calibration.

The BioLCCC model based on the concept of macromolecular chromatography under critical conditions and realized in the Theoretical Chromatograph software was used for calculation of theoretical peptide retention times [11-13] (http://theorchromo.ru). The Theoretical Chromatograph software predicts retention times of proteins and peptides separated on reverse-phase C18 under assigned experimental conditions (including column dimensions, gradient profile, and composition of solvent components) depending on their amino acid sequences.

As an alternative method for calculation of theoretical times of peptide retention, the SSRCalc based on an additive model of peptide separation and empirically determined hydrophobicity coefficients was used [6, 7] (http://hs2.proteome.ca/SSRCalc/SSRCalc32.html).

RESULTS AND DISCUSSION

Concept of linear correlation of chromatographic retention times. The earlier proposed approach [10] is based on a supposition concerning linear correlation of peptide retention times obtained under different separation conditions most widely used in proteomic studies. The use of this approach is restricted by conditions typical of proteomic studies using the HPLC-MS/MS techniques and suggests the reverse-phase C18 as adsorbent, pore size in the range from 90 to 300 Å, water and acetonitrile mixture as a binary solvent with pH values in the range 2.0 to 3.0, and linear gradient profiles with slope of 0.2 to 1.7% acetonitrile/min. Acetic or formic acid are usually used as ion-pair agents in HPLC-MS/MS. The concept of linear correlation of chromatographic data is not new when speaking about chromatography of low molecular weight compounds or short peptides and follows from LSS

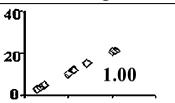
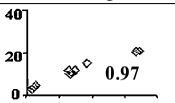
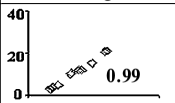
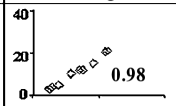
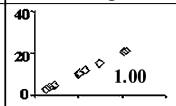
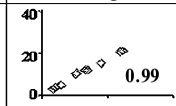
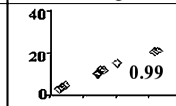
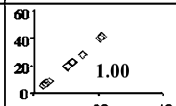
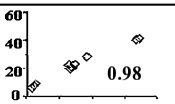
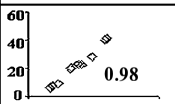
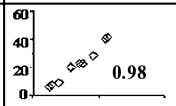
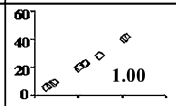
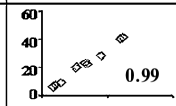
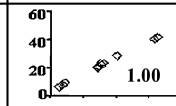
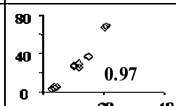
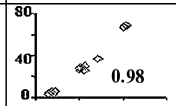
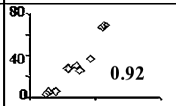
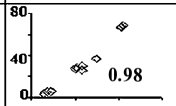
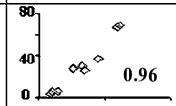
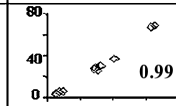
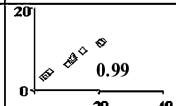
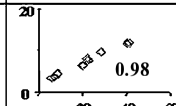
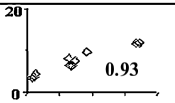
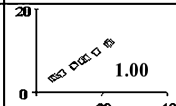
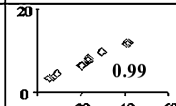
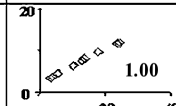
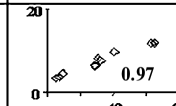
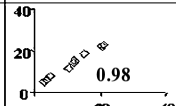
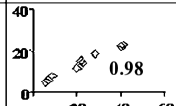
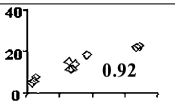
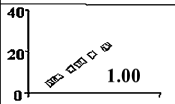
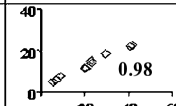
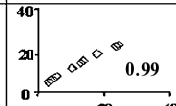
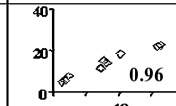
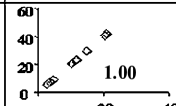
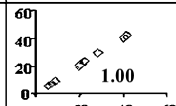
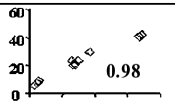
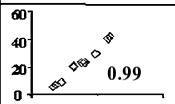
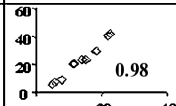
	2 × 150 mm, 100 Å, 0.5 ml/min, 0-60% ACN during 60 min	4.6 × 250 mm, 300 Å, 1 ml/min, 0-30% ACN during 60 min	3 × 100 mm, 110 Å, 0.8 ml/min, 0-30% ACN during 120 min	1 × 250 mm, 300 Å, 0.05 ml/min, 2-62% ACN during 30 min	0.3 × 150 mm, 90 Å, 0.004 ml/min, 2-62% ACN during 60 min	0.15 × 100 mm, 100 Å, 0.001 ml/min, 0-30% ACN during 60 min	0.075 × 150 mm, 100 Å, 0.0003 ml/min, 4-50% ACN during 60 min	0.075 × 120 mm, 120 Å, 0.0003 ml/min, 0-35% ACN during 120 min
2 × 150 mm, 100 Å, 0.5 ml/min, 0-60% ACN during 60 min								
4.6 × 250 mm, 300 Å, 1 ml/min, 0-30% ACN during 60 min								
3 × 100 mm, 110 Å, 0.8 ml/min, 0-30% ACN during 120 min								
1 × 250 mm, 300 Å, 0.05 ml/min, 2-62% ACN during 30 min								
0.3 × 150 mm, 90 Å, 0.004 ml/min, 2-62% ACN during 60 min								
0.15 × 100 mm, 100 Å, 0.001 ml/min, 0-30% ACN during 60 min								
0.075 × 150 mm, 100 Å, 0.0003 ml/min, 4-50% ACN during 60 min								
0.075 × 120 mm, 120 Å, 0.0003 ml/min, 0-35% ACN during 120 min								
0.075 × 150 mm, 120 Å, 0.0002 ml/min, 2-45% ACN during 90 min								

Fig. 1. Table of linear correlations of retention times (numbers on plots correspond to values of correlation R^2) for different HPLC parameters predicted by the BioLCCC theory. The range of theoretically considered conditions includes those usual for semi-preparative and analytical chromatography as well as for HPLC systems with micro- and nano-flows. Retention times calculated under conditions marked in an appropriate column of the table correspond to the X axis of each graph, while retention times calculated under conditions mentioned in lines correspond to the Y axis. ACN, acetonitrile.

theory [9]. At the same time, despite experimental evidence of linear correlation between chromatographic times of short peptide [10, 14], the problem is not so obvious in the case of long molecules, because in this case separation follows not only the adsorption mechanism but the exclusion one as well [12, 13].

To answer the question concerning possible functional dependence of chromatographic data obtained on different HPLC systems and under different separation conditions, we performed calculations of theoretical retention times using the BioLCCC model [11-13]. These calculations were made for typical separation protocols used under conditions of gradient HPLC and for peptides of cytochrome *c* protein hydrolysate. The choice of these protein peptides was deliberate: first, it is a sufficiently simple, well characterized, and commercially available peptide mixture, and second, by their adsorption property on reverse phase, peptides of this mixture envelop a practically complete range of physicochemical properties characteristic of tryptic peptides. Figure 1 shows results of calculation of retention times and their mutual correlations for arbitrarily chosen separation protocols. As a whole, the performed calculating experiment has shown the existence of linear correlation between chromatographic data.

It should be noted that results of calculation using the BioLCCC model show the dependence of separation selectivity on such conditions as column pore size and gradient slope, which in turn can be the reason for decrease in linear correlation from $R^2 = 1.00$ to 0.93. Nevertheless, the calculations predict the existence of a rather broad range of experimental conditions for peptide separation in gradient HPLC on the reverse phase for which linear correlation of chromatographic data is observed.

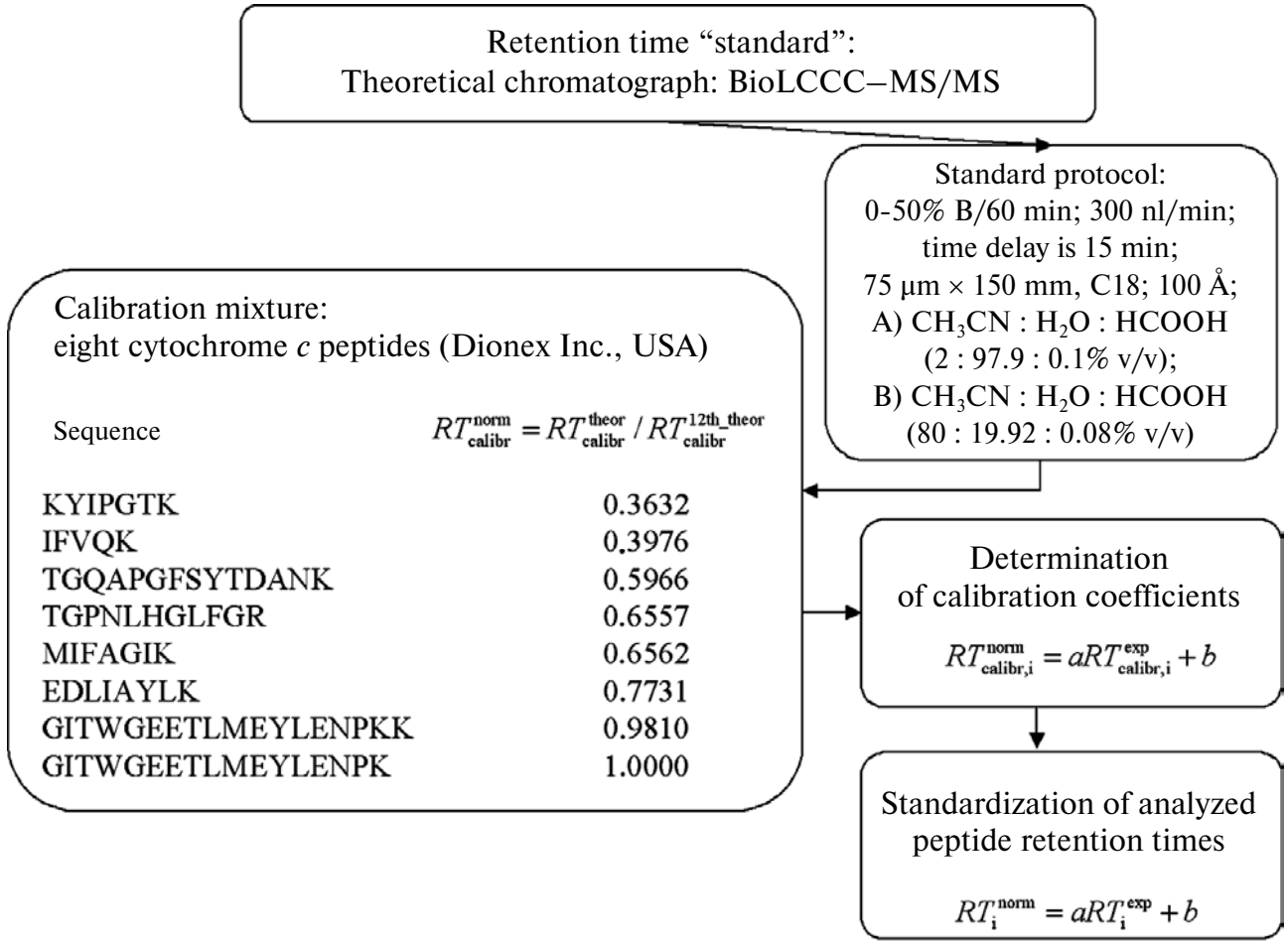
Tarasova et al. [10] carried out a number of experiments on determination of effects of gradient slope, rate of solvent flow, and column characteristics on chromatographic retention on the example of cytochrome *c* peptides. It was shown that for different types of reverse phases C18 (silica gel, monolith, and polymer) a high level of chromatographic data linear correlation is observed (on average, $R^2 = 0.985$-0.995). Table 3 in addition shows results of experiments in which the following separation conditions were changed: (i) gradient profile (at fixed column I and flow rate of 300 nl/min); (ii) flow rate (at fixed column I and gradient slope of 0.8% B/min); as well as (iii) column parameters upon separation in two different gradients, 1.7 and 0.3% B/min, and fixed flow of 300 nl/min. In this series of experiments the lowest correlation coefficient of $R^2 = 0.97$ was obtained for polymer phase RepMap PLRP-S (column V, Table 1).

Note that during experiments with cytochrome *c* peptides a pair of peptides was found which changed the order of their elution from the column upon transition from the sharp gradient (1.3% B/min) to the slowly sloping one (0.3% B/min). Of course, such peptides present a certain difficulty: linear transformation of retention times does not allow one to consider the inversion of the peptide elution order upon change in the separation protocol. However, such inversion in narrow range of retention times did not change the value of linear correlation coefficient R^2. This phenomenon is not affecting the accuracy of the chromatographic data calibration with the use of a linear equation.

BioLCCC model as basis for universal scale of retention times. To solve the problem of normalization of experimental data independently obtained by different research groups, we used in creation of the AMT databases the earlier developed method of multipoint normalization (MPN) [10]. Using predicted retention times for the chosen peptide standard and the parameter of chromatographic data linear correlation, it is possible to introduce a new scale of time in which experimental values of the

Table 3. Linear correlation coefficients for experimental retention times of cytochrome *c* peptides for different separation conditions, columns, and protocols. Coefficients R^2 indicate correlation of retention times under tested conditions with "standard" data obtained several months earlier on column I (Table 1) at 0-50% B gradient for 60 min

Gradient, % B/min	R_1^2	Solvent flow rate, nl/min	R_2^2	Column	R_3^2	
					1.7% B/min	0.3% B/min
1.7	0.992	200	0.998	I	0.994	0.988
1.4	0.998			II	0.994	0.992
1.2	0.998	300	0.998	III	0.978	0.966
0.8	0.997			IV	0.992	0.990
0.4	0.995	400	0.997	V	0.974	0.967
0.3	0.988			VI	0.990	0.988



Schematic representation of standardization of peptide retention times by MPN using the BioLCCC model for calculation of theoretical retention times

studied peptide retention times will have the same (for identical peptides) normalized times. The step-by-step description of the time standardization procedure used in this work is shown on Scheme.

The scale of normalized times was plotted as follows. The cytochrome *c* peptides were chosen as standard for which theoretical values of retention time under conditions of predetermined basic separation protocol were calculated. The choice of basic protocol is random. The protocol that is standard in microcolumn peptide chromatography was used as basic in this work (Scheme).

Then theoretical retention times for standard peptides were normalized in the [0.1] range by fixing the normalized time scale to the GITWGEETLMEYLENPK peptide having the longest retention time. In principle, the time scale reduction to the [0.1] fragment is not obligatory, but it seems more convenient with respect to the scale universality. Calibration coefficients *a* and *b* were determined by substitution into the linear equation of experimental and normalized theoretical values of standard peptide retention times (Scheme). With the use of these calibration coefficients all experimental times RT_i^{exp} for the studied peptides were converted into the new scale of normalized times using linear equation (Scheme).

In principle, the proposed procedure is analogous to a calibration procedure in mass spectrometry in which measured physical values bind via a corresponding calibration equation with masses expressed in atomic units and attached to the universal mass scale based on carbon ^{12}C.

Experimental validation of results of retention time normalization by the MPN method. The proposed normalization procedure was tested experimentally for tryptic peptide mixture of six proteins. Note that the mixture also includes cytochrome *c* peptides that were used for retention time normalization of all peptides of the mixture ("internal calibration"). Data obtained in different laboratories were combined in three groups: (i) peptides identified in Moscow and Grenoble (174 peptides); (ii) peptides identified in Moscow and Uppsala (103 peptides); and (iii) identifications common for all three laboratories (69 peptides). Figure 2 shows standard devia-

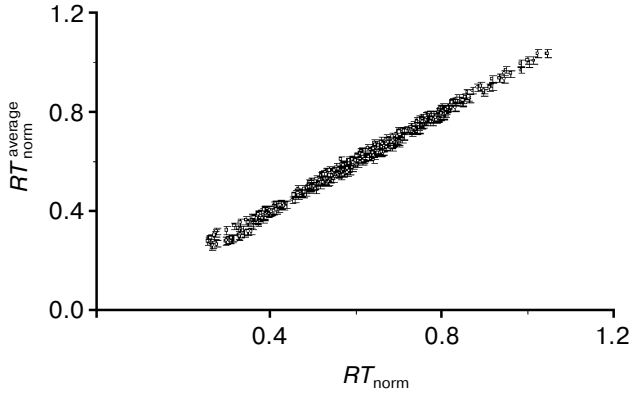


Fig. 2. Graph of experimental retention times of the six-protein standard peptides identified in laboratories of Grenoble, Uppsala, and Moscow and normalized by the MPN method using the cytochrome *c* internal standard. Mean values for measured and calibrated retention times were plotted on the Y axis of this graph $RT^{average}_{norm} = (1/3)\,\Sigma\,(RT^{Grenoble}_{norm} + RT^{Uppsala}_{norm} + RT^{Moscow}_{norm})$. The point size on the Y axis corresponds to standard deviation from the mean and on average it does not exceed 1.5%.

tions from mean value of retention times normalized using as "internal calibrant" cytochrome *c* peptides for data (i), (ii), and (iii), respectively. In this case standard deviations for obtained normalized chromatographic times of identified peptides were, respectively, 1.4, 1.6, and 1.5%. Thus, the use of the MPN method to reduce to a universal scale of retention times the chromatographic data obtained on different instrumental platforms and different parameters of proteolytic mixture separation has shown the agreement between normalized retention times for identical peptides with the accuracy of about 1.5%.

It should be noted that in proteomic studies there is no problem of choosing an internal calibrant: a natural choice can be peptide admixtures such as those from keratin, a satellite of practically any biological specimen, or peptides with the same sequence but present in different proteins. However, the use of "internal calibration" for retention time reduction to a universal scale is not always possible, which can be first of all due to difficulties of the identification of the same peptide within a complex mixture in different research centers using different equipment and HPLC protocols. An alternative is the retention time normalization using "external calibration" of the chromatographic system by peptides of a standard analyzed separately from the main specimen. In this work, retention times of peptides identified in Uppsala were normalized by external standard using cytochrome *c* peptides. Table 4 shows standard deviations of normalized retention times for different sets of data obtained with internal and external calibration using the MPN method and two different algorithms of retention time prediction – BioLCCC and SSRCalc.

As is seen in these data, the use of alternative models for retention time prediction makes it possible to obtain comparable data by the accuracy of normalized retention times. Thus, the accuracy of "internal calibration" using different models of retention calculation varied within the range of 0.9-1.5%. It should be noted that the SSRCalc algorithm is considered as the most precise algorithm for calculation of tryptic peptide retention times and is widely used in proteomics. However, the SSRCalc algorithm has been developed for a limited number of experimental

Table 4. Standard deviations (%) between different sets of data after retention time standardization by MPN using internal and external (BioLCCC*) calibrations on the basis of cytochrome *c* peptides as well as using alternative retention time calculators

		Uppsala			Moscow		Grenoble	
		BioLCCC	BioLCCC*	SSRCalc	BioLCCC	SSRCalc	BioLCCC	SSRCalc
Uppsala	BioLCCC		1.7	0.9	1.2	1.1	0.5	0.5
	BioLCCC*	1.7		2.4	2.7	2.1	1.8	1.9
	SSRCalc	0.9	2.4		1.5	1.1	1.0	1.0
Moscow	BioLCCC	1.2	2.7	1.5		0.9	1.4	1.3
	SSRCalc	1.1	2.1	1.1	0.9		1.2	1.1
Grenoble	BioLCCC	0.5	1.8	1.0	1.4	1.2		0.1
	SSRCalc	0.5	1.9	1.0	1.3	1.1	0.1	

Note: The retention time calculation using the SSRCalc calculator was performed for pore size 100 Å; coefficient taking into account dead volumes of the system $A = 7$; and coefficient associated with the gradient slope $B = 0.95$ (coefficients A and B were incorporated into the SSRCalc model by its developer, http://hs2.proteome.ca/SSRCalc/SSRCalc32.html). Coefficient A was optimized using experimental data from Grenoble and optimization was stopped when the standard deviation of retention times, normalized using BioLCCC and SSRCalc, reached 0.1%.

separation conditions and the change, for example, of pore size or gradient profile requires its readjustment [9]. The BioLCCC model is free of the above-mentioned shortcomings and allows retention time calculations for any HPLC conditions on reverse-phase C18 for arbitrarily preset gradient profiles and column parameters. In this case, due to a low number of phenomenological parameters of the model (20 amino acid surface-interaction energies), the BioLCCC model can be easily adapted for work with different phases.

The proposed procedure for normalization of peptide retention times enables standardization of experimental data obtained on reverse phase C18 under different gradient conditions, column, and mobile phase parameters. It was shown on the example of separation of tryptic peptide complex mixtures and using different experimental separation protocols and instrumental systems that in the case of internal calibration normalized retention times correlate for the same amino acid sequences with accuracy of the order of 1-1.6%. If external calibration standards are used, the precision of normalized retention times is 1.7-2.7%. The proposed approach to normalization of peptide retention times and their conversion to a universal time scale can be used for protein identification in shotgun proteomics and generation of AMT databases.

This work was supported by the Russian Foundation for Basic Research (grants 09-08-00633, 08-04-01339, and 08-04-91121-CRDF), INTAS (grant Genomics 05-10000004-7759), and Section of Chemical Sciences and Materials, Russian Academy of Sciences (SCSM program 4.2 "Creation of Efficient Methods for Chemical Analysis and Structural Investigation of Substances and Materials").

REFERENCES

1. Conrads, T. P., Anderson, G. A., Veenstra, T. D., Pasa-Tolic, L., and Smith, R. D. (2000) *Anal. Chem.*, **72**, 3349-3354.
2. Norbek, A. D., Monroe, M. E., Adkins, J. N., Anderson, K. K., Daly, D. S., and Smith, R. D. (2005) *Am. Soc. Mass Spectrom.*, **16**, 1239-1249.
3. May, D., Fitzgibbon, M., Liu, Y., Holzman, T., Eng, J., Kemp, C. J., Whiteaker, J., Paulovich, A., and McIntosh, M. (2007) *J. Proteome Res.*, **6**, 2685-2694.
4. Petritis, K., Kangas, L. J., Ferguson, P. L., Anderson, G. A., Pasa-Tolic, L., Lipton, M. S., Auberry, K. J., Strittmatter, E. F., Shen, Y., Zhao, R., and Smith, R. D. (2003) *Anal. Chem.*, **75**, 1039-1048.
5. Sapirstein, H. D., Scanlon, M. G., and Bushuk, W. (1989) *J. Chromatogr. A*, **469**, 127-135.
6. Krokhin, O. V., Craig, R., Spicer, V., Ens, V., Standing, K. G., Beavis, R. C., and Wilkins, J. A. (2004) *Mol. Cell. Proteom.*, **3**, 908-919.
7. Spiser, V., Yamchuk, A., Cortens, J., Sousa, S., Ens, W., Standing, K. G., Wilkins, J. A., and Krokhin, O. V. (2007) *Anal. Chem.*, **79**, 8762-8768.
8. Shinoda, K., Tomita, M., and Ishihama, Y. (2008) *Bioinformatics*, **24**, 1590-1595.
9. Snyder, L. R. (1983) *J. Chromatogr.*, **255**, 3-26.
10. Tarasova, I. A., Guryca, V., Pridatchenko, M. L., Gorshkov, A. V., Kieffer-Jaquinod, S., Evreinov, V. V., Masselon, C. D., and Gorshkov, M. V. (2009) *J. Chromatogr. B*, **877**, 433-440.
11. Gorshkov, A. V., Tarasova, I. A., Evreinov, V. V., Savitski, M. M., Nielsen, M. L., Zubarev, R. A., and Gorshkov, M. V. (2006) *Anal. Chem.*, **78**, 7770-7778.
12. Gorshkov, A. V., Evreinov, V. V., Tarasova, I. A., and Gorshkov, M. V. (2007) *Polymer Sci. Ser. B*, **49**, 93-107.
13. Tarasova, I. A., Gorshkov, A. V., Evreinov, V. V., Zubarev, R. A., and Gorshkov, M. V. (2008) *Polymer Sci. Ser. B*, **50**, 1-15.
14. Casal, V., Martin-Alvarez, P. J., and Herraiz, T. (1996) *Analyt. Chim. Acta*, **326**, 77-84.