

EXPERIMENTAL RESEARCH

IDENTIFICATION OF HUMAN BLOOD PLASMA PROTEINS USING SPIKE-IN PEPTIDES IN SHOTGUN PROTEOMICS

*A.T. Kopylov, O.V. Tikhonova, T.E. Farafonova, N.A. Petushkova, Yu.V. Miroshnichenko, V.G. Zgoda**

Institute of Biomedical Chemistry, 10 Pogodinskaya str., Moscow, 119121 Russia; *e-mail: victor.zgoda@gmail.com

LC-MS/MS allows identification of thousands of proteins in the complex proteomes. However, a significant part of a proteome remains inaccessible for identification due to the absence or poor quality of MS/MS spectra. The method described herein allows identifying the desired proteins of human blood plasma by comparing aligned chromatographic data of digested by trypsin sample and the same sample with spiked-in synthetic peptides. Identification of human blood plasma proteins is archived by assigning tandem mass spectra of spiked-in peptides to the corresponding aligned chromatographic peaks of proteolytic peptides. Using the described approach we have identified 19 low abundant proteins in human blood plasma, which corresponded to 19 synthetic peptides used in the study. SRM verification of the identifications with isotopically labelled standards (SIS) confirmed the presence in the plasma of above 17 proteins.

Key words: mass spectrometry; protein identification; shotgun proteomics; spike-in peptides**DOI:** 10.18097/BMCRM00093

Abbreviations: AMT – accurate mass and time; FA – formic acid; HCD – higher energy collisional dissociation; NaDC – sodium deoxycholate; RTs – retention times; SIS – stable isotope-labelled synthetic; SWATH-MS – sequential window acquisition of all theoretical fragment-ion spectra mass spectrometry; PMF – peptide mass fingerprint; MS – mass-spectra; MS/MS - tandem MS

INTRODUCTION

Mass spectrometry is the main technique used in proteomics for analysis of biological samples [1]. Peptides obtained by proteolytic cleavage of a protein or a mixture of proteins are analyzed using the bottom-up approach. In the classical version of bottom-up proteomics, protein identification is carried out using PMF [2,3] and MS/MS techniques based on the mass spectra of proteolytic peptides (MS1) or their fragments (MS2), respectively [4]. PMF is mostly used for identifying purified proteins or simple protein mixtures [5].

MS/MS is used for the analysis of complex protein mixtures to identify as many proteins as possible in a sample (from several hundred [6] to several thousand [7] in a single experiment) in an approach known as shotgun proteomics. In this case, tandem mass spectra (MS2 spectra) obtained after multidimensional chromatographic separation of a mixture of proteolytic peptides and scans of precursor ions (MS1) and their fragments (MS2) are analyzed [1]. Nevertheless, a significant part of a proteome in biological materials, including proteins present at low concentrations, remains inaccessible for MS/MS analysis [8]. As a consequence, over 70% of all recorded MS2 spectra remain unidentified or contain false identifications [10]. Such large-scale losses of information can be avoided by processing MS1 spectra only. This approach is much more sensitive but suffers from a high rate of false positive identifications [9,11]. Nevertheless, over the past ten years different research groups developed mass spectrometry methods with and without MS/MS in parallel over the past ten years.

Methods based on the AMT (Accurate Mass and Time) tag strategy provide another example; unlike PMF, it can analyze more complex protein mixtures and identify proteins from MS1 spectra and retention times (RTs) of peptides, bypassing the fragmentation stage [15,16]. However, the effectiveness

of AMT methods is limited by: (1) the effect of MS2 data used in AMT tags databases for protein search [17], (2) the labour intensity of creating AMT tags databases, and (3) the dependence of chromatography data (in this case RTs) on the conditions of a particular experiment, which complicates the use of AMT tags databases across laboratories [18]. Models for predicting theoretical RTs have been proposed at different times for normalizing experimental RTs, but none of them has become universally suitable for generating error-free RTs [19] or adapting to different experimental conditions [20,21].

Another strategy called SWATH-MS (Sequential Window Acquisition of all Theoretical fragment-ion spectra Mass Spectrometry) combines the high throughput of shotgun proteomics with the accuracy and reproducibility of quantitative SRM analysis [12,13]. The feasibility of SWATH-MS is achieved by cyclic sequential scanning of fragment ions (MS2) of all precursors in the range between 400 and 1200 *m/z*, allowing large datasets to be processed with high accuracy in a dynamic range of up to four orders of magnitude [12,13]. The key drawbacks of SWATH-MS include interference of fragment ions and the need for creating specialized spectrum libraries, the number of identified proteins being dependent on the quality of the libraries [13,14].

The general idea of the study is based on the observation that identifications of peptides can be archived using MS/MS spectra of the same feature obtained in other LC-MS runs after all runs alignment by the Progenesis software. Thus, the goal of this work was to identify of desired proteins in human blood plasma using 19 synthetic proteotypic peptides for guaranteed recording of their MS2 spectra. Mass spectrometry was performed by LC-MS/MS and the high-resolution MS data were processed using the Mascot and Progenesis LC-MS software. Identification of human blood plasma proteins was archived by assigning tandem mass spectra of spiked in peptides to the corresponding



aligned chromatographic peaks of proteolytic peptides, which had MS2 spectra of spike-in synthetic peptides. Analysis of MS1 and MS2 data allowed to identify 19 proteins in human blood plasma, which corresponded to 19 spike-in synthetic peptides. SRM verification of the identifications with SIS standards confirmed the presence in the plasma of 17 above proteins.

MATERIALS AND METHODS

Materials

Formic acid (FA), thiourea and HPLC grade water were purchased from «Acros Organics» (USA). Acetonitrile (ACN) was from «Merck» (Germany); ammonium bicarbonate was from «Pierce» (USA); porcine trypsin modified (sequencing grade) was from «Promega» (USA).

Internal Standard Production

The SIS peptides desired were obtained using the solid-phase peptide synthesis on the Overture («Protein Technologies», USA) synthesizer according to the published method [26]. The isotope-labelled leucine ($^{13}\text{C}_6^{15}\text{N}$), arginine ($^{13}\text{C}_6^{15}\text{N}_4$) or lysine ($^{13}\text{C}_6^{15}\text{N}_2$) were used for incorporation into peptide standards instead of the unlabeled leucine, arginine or lysine, respectively. Concentrations of the synthesized peptides were measured by the method of amino acids analysis with fluorescent signal detection of amino acids derived after acidic hydrolysis of peptides.

Biomaterial

Venous blood samples were collected from 3 volunteers in disposable plastic tubes (10 mL) containing $\text{K}_3\text{-EDTA}$ as an anticoagulant, followed by centrifugation in an Armed CH80-2S centrifuge (Germany) at 3000 rpm for 10 min to obtain plasma and sediment blood cell were stored at -80°C . The plasma supernatant was filtered through $0.22\ \mu\text{m}$ cellulose-acetate filters («Whatman», USA) aliquoted and stored at -80°C until the proteomic analysis. The protein concentration was determined using the Micro BCA protein assay («Thermo Scientific», USA). Protein digestion was performed according to the protocol described in detail by Zgoda et al. [12].

LC-MS/MS analysis

Samples were analyzed using the UltiMate 3000 nano-flow HPLC system («Thermo Scientific») connected to an Orbitrap Q-Exactive HF mass spectrometer («Thermo Scientific»). Peptides separation was carried out on a Zorbax 300SB-C18 column, ($150\ \text{mm} \times 75\ \mu\text{m}$, $3.5\ \mu\text{m}$ particle size; «Agilent Technologies», USA), in a linear gradient from 3% to 35% of mobile phase B (80% ACN, 0.1% FA) over 70 min, then from 35% to 99% of mobile phase B over 5 min at $0.3\ \mu\text{L}/\text{min}$ followed by a washing the column at 99% of mobile phase B for 10 minutes and post-analytical equilibration at 3% solvent B for 5 minutes.

Mass spectra were acquired in the positive ionization mode with a resolution of 70000 (at m/z 400) for MS and 15000 (m/z 400) for MS/MS scans. Survey MS scan was followed by MS/MS acquisition of top-ten most abundant precursors. Higher-energy collisional dissociation (HCD) at 28 CE was used to generate fragment ions. The signal threshold for precursor ions

was set to 100000 and ions were isolated within $2\ m/z$ window. Tandem mass spectra were acquired within a range of the fixed first mass ($m/z = 130$) and charge-dependent last mass. Singly charged ions and ion with undefined charge state were excluded from MS/MS triggering. Dynamic exclusion for 10 s was applied if precursor ion was targeted at least 3 times.

LC-SRM analysis

Chromatographic separation of peptides was carried out using the UPLC Agilent 1290 («Agilent Technologies») system composed of a micro-flow pump and an autosampler. Ten microliters of sample were loaded onto an Eclipse Plus SBC-18 column ($2.1 \times 100\ \text{mm}$, $1.8\ \mu\text{m}$, 100 Å; «Agilent Technologies»). Separation was performed in a linear gradient of mobile phase A (0.1% (v/v) FA) and mobile phase B (80% (v/v) HPLC grade ACN/water with 0.1% (v/v) FA). Peptides were loaded at starting conditions of 3% B and increasing to 32% of B over 50 min, then from 32% to 53% of B over 3 min followed by washing the column at 90% of B for 5 minutes). The column was equilibrated at the initial gradient condition for 5 minutes before the next sample run. Mass spectra were acquired on a G6495 triple quadrupole mass spectrometer («Agilent Technologies») equipped with the Jet Stream ionization source. The following parameters were used for the Agilent Jet Stream ionization source: the temperature of the drying gas – 280°C , the pressure in the nebulizer – 18 psi, the flow of the drying gas – $14\ \text{l}/\text{min}$, and the voltage on the capillary – 3000V.

Data analysis

Raw data files were processed in the Progenesis LC-MS software (version 4.1; «Nonlinear Dynamics Ltd.», UK). Protein identification was performed using the Mascot software (version 2.4.1). The following search parameters were used: database, UniProt KB (version 2012_11) limited to *Homo sapiens* taxonomy; cutting enzyme: trypsin; precursor ions tolerance $\pm 10\ \text{ppm}$; fragment ions tolerance $\pm 0.05\ \text{Da}$; one permissible missed cleavage sites was allowed; fixed modifications: cysteine

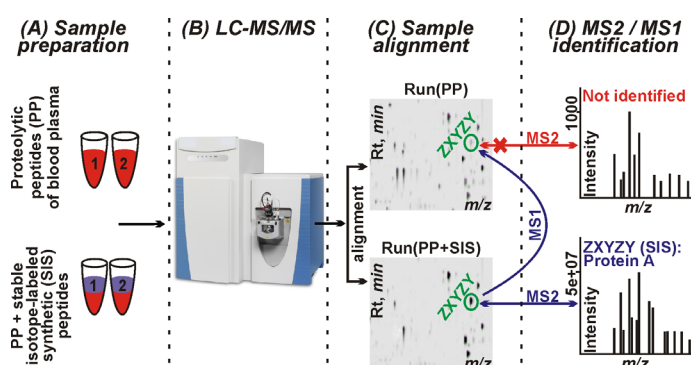


Figure 1. Diagram of an experiment workflow allowing co-alignment of samples of proteolytic peptides of human blood plasma with and without spike-in synthetic peptides. The spike-in peptides were aligned with the natural peptides by the RTs and m/z of their MS1 features in order to identify proteins by high quality MS2 spectra of spike-in synthetic peptides. Key stages: (A) sample preparation; (B) LC-MS/MS analysis by high resolution mass spectrometer; (C) the samples alignment using Progenesis software; (D) search for protein identifications in Mascot software; mapping of MS2 synthetic peptide identifications to MS1 features of aligned samples and the identification of human blood plasma proteins corresponding to spike-in peptides in a sample free of these peptides.

Table 1. List of identified of human plasma proteins/peptides.

#	UniProt ID	Name of protein (gene name)	Peptide	Precursor m/z ²	PP/SP ³
1.	Q9UIL8	PHD finger protein 11 (PHF11)	EAGLLNYLLEEILDK ¹	866.97++	+/+
2.	P00742	Coagulation factor X (F10)	ETYDFDIAVLR	671.34++	-/+
3.	P11279	Lysosome-associated membrane glycoprotein 1 (LAMP1)	TVESITDIR	517.28++	-/+
4.	Q9UBN4	Short transient receptor potential channel 4 (TRPC4)	NFSLFDLTTLIHPR	558.64+++	-/+
5.	Q96CX2	BTB/POZ domain-containing protein KCTD12 (KCTD12)	DLQLVLPDYFPER	802.92++	-/+
6.	Q96IY4	Carboxypeptidase B2 (CPB2)	DTGTYGFLPER	684.84++	-/+
7.	Q15762	CD226 antigen (CD226)	DLFTESWDTQK	685.35++	-/+
8.	Q9UNA4	DNA polymerase iota (POLI)	DLQTFSPK	468.25++	-/+
9.	O95164	Ubiquitin-like protein 3 (UBL3)	ETLPEPNSQGQR	678.33++	-/+
10.	O75343	Guanylate cyclase soluble subunit beta-2 (GU-CY1B2)	FINSQFVLK	548.31++	-/+
11.	P08709	Coagulation factor VII (F7)	FSLVSGWGQLLDR	739.39++	-/+
12.	P22830	Ferrochelatase, mitochondrial (FECH)	GDPYPQEVSATVQK	759.88++	-/+
13.	P23368	NAD-dependent malic enzyme, mitochondrial (ME2)	HISDSVFLEAAK	439.57+++	-/+
14.	Q9C0E2	Exportin-4 (XPO4)	SPPLNFLSSPVQR	721.39++	-/+
15.	Q9Y6X0	SET-binding protein (SETBP1)	TMPNLQPISALPTK	504.28+++	-/+
16.	Q9UL03	Integrator complex subunit 6 (INTS6)	YELEPSPLTQFILER	917.98++	-/+
17.	Q9BWV3	Cytidine and dCMP deaminase domain-containing protein 1 (CDADC1)	YIIHAEQNALTFR	525.95+++	-/+
18.	Q96KW9	Protein SPACA7 (SPACA7)	TVIPGSPTEIPFSSK	–	-/+
19.	P61960	Ubiquitin-fold modifier 1 (UFM1)	FAAEFEK	–	-/+

Note. ¹ peptide identified in samples of human blood plasma. ² SRM m/z for precursor ions of the endogenous peptides. ³ Identification of peptide/protein in human plasma [before/after (SP)] alignment of the MS1 features in the PP and SP samples.

carbamidomethylation; variable modifications, oxidation of methionine. The criterion for positive identification was considered with a score >13; significance threshold, $p < 0.05$; FDR < 1%. For the SRM data analysis and quantification of the human blood plasma proteins, the Skyline software (version 3.7) was used [27].

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [28] partner repository with the dataset identifier PXD007580 and 10.6019/PXD007580.

RESULTS AND DISCUSSION

Figure 1 demonstrates the schematic workflow of the designed experiment. In our study, we used 19 tryptic peptides which corresponded to 19 proteins (Table 1). According to Nextprot records, most of the chosen peptides and corresponded proteins have been identified in human plasma or serum [PMID: 27457493, 29101746].

At the first step of the workflow, 19 peptides were spiked in digested plasma samples. All samples were analyzed by LC-

MS/MS in triplicates as described in the «Material and methods» section. The obtained raw files were processed using Progenesis. Progenesis LC-MS was used for sequential alignment of experimental runs of human blood plasma proteolytic peptide mixtures samples (PP); PP runs and runs of the same samples with spike-in synthetic peptides (SP). The MS/MS spectra of recorded features were uploaded into the Mascot search engine in the .mgf format, and protein identification searching was performed. After co-alignment of the MS (MS1) features found in PP and SP, the analysis of PP yielded proteolytic peptides/proteins identical with or corresponding to the spike-in peptides, whose MS1 features were associated with identified MS/MS spectra. Totally 126 proteins were identified in PP runs only, and 184 proteins were identified in SP; 117 proteins were common for both samples. Among 184 proteins identified in SP, we found all 19 spiked-in peptides and all of them were assigned to the corresponding features (peptides) of the PP (Table 1).

As can be seen from Table 1, only one peptide (TGIVSGFGR₁) from the list was identified in the sample without the synthetic peptides (PP). This can indicate that remaining 18 ions of these proteolytic peptides were either (1) not selected

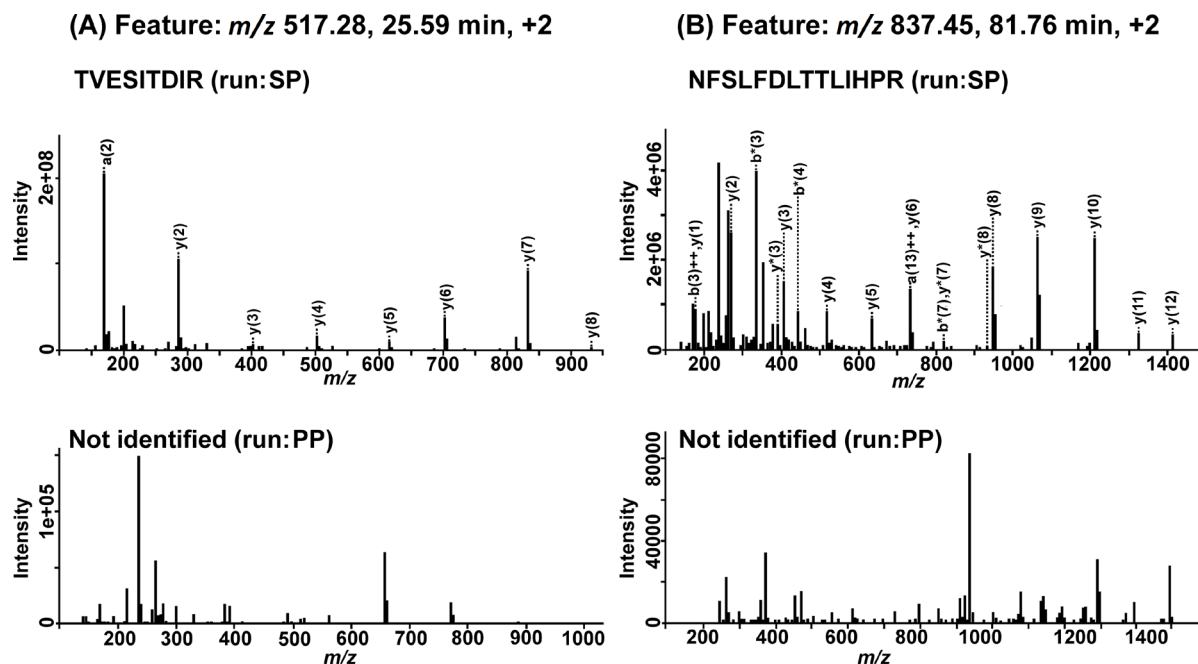


Figure 2. An example of MS2 spectra corresponding to the same MS1 features of samples of mixtures of proteolytic peptides of human blood plasma (PP) co-aligned with samples of the same mixtures with spike-in synthetic peptides (SP). (A) An identified (for SP) MS2 spectrum of peptide TVESITDIR and an unidentified (for PP) MS2 spectrum corresponding to a MS1 feature with m/z 517.28. (B) An identified (for SP) MS2 spectrum of peptide NFSLFDLTTLIHPR and an unidentified (for PP) MS2 spectrum corresponding to a MS1 feature with m/z 837.45.

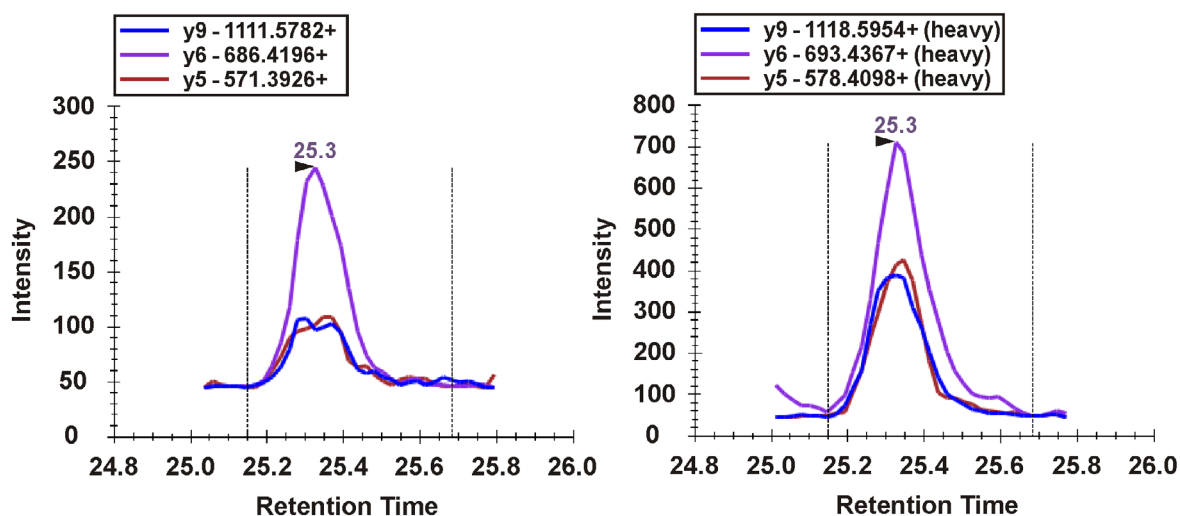


Figure 3. SRM identification of the protein Coagulation factor X (P00742) in human blood plasma. Shown ion chromatograms of native and stable isotope-labeled synthetic (heavy) peptide ETYDFDIIVLR from the Skyline software.

for HCD fragmentation; or (2) their MS2 spectra were not good enough to identify the peptide. Figure 2 shows that for the same MS1 features (peptides) of co-aligned PP and SP, both identified (for SP) spectra of peptides TVESITDIR (m/z 517.28) and NFSLFDLTTLIHPR (m/z 837.45) had high-intensity peaks and unidentified (for PP) MS2 spectra had low-intensity peaks.

Therefore, identification of human blood plasma proteins (peptides) using the corresponding synthetic peptides was based on data of accurate alignment by RTs, m/z in MS1 features, and the MS2 identifications of these spike-in peptides. As an example, results of co-alignment of MS1 features in PP and SP runs for the ion intensity map regions with matching RTs and m/z are shown in Figure 1 and Supplementary Materials.

Verification of protein identifications was carried out using a triple quadrupole mass spectrometer in the SRM mode. SRM is characterized by high sensitivity, reproducibility of measurements and a linear dynamic range of five to six orders of magnitude for quantitative analysis of proteins [29]. This targeted approach requires development of a method and synthesis of isotopically labeled peptide standards specific to each protein [30]. For the validation purpose, we synthesized SIS with the same sequences of peptides used for the identification of the proteins in LC-MS/MS experiments (see Table 1).

The detection and quantification of proteins by SRM were carried out in the same plasma samples and SIS peptides were added to the plasma samples. The data obtained were analyzed

using the criteria, which had been previously determined for the SRM method: first, a complete coincidence of chromatographic profiles of endogenous and isotope-labeled standard; secondly, the presence of signals from all three fragmented transitions; the difference between the relative intensities of fragmented transitions for the endogenous peptide and isotope-labeled peptide did not exceed 25% [31].

The examples given in Supplementary Materials in Figure 2 show SRM profiles of peptides accepted in the course of quality control and fit identification criteria. Figure 3 shows traces of positive identification of the protein Coagulation factor X (P00742) where natural and SIS peptides are matched perfectly.

As a result of SRM traces quality control, we have detected the signals for 17 of 19 desired proteins in the plasma samples (see Table 1). Quantitative estimations for 17 detected proteins were obtained by SRM in a dynamic range of 2.5 orders of magnitude from 10 µg/ml (Coagulation factor X, P00742) to 50 ng/ml (Cytidine and dCMP deaminase domain-containing protein 1, Q9BWV3) using SIS peptides added to the plasma samples. Thus, we obtained evidence of the presence of 17 proteins among 19 proteins identified in LC-MS/MS experiments.

In summary, LC-MS/MS shotgun analysis of blood plasma proteolytic peptides mixtures that were spiked with synthetic peptides yields MS2 spectra of greater intensity for the most abundant ions of spike-in peptides relatively to proteolytic peptides. This allows the reliable identification of human blood plasma proteins.

COMPLIANCE WITH ETHICAL STANDARDS

All blood donors gave informed consent to participate in this study. Human-related procedures were performed according to the guidelines of the local ethical committee (Institute of Biomedical Problems of RAS).

ACKNOWLEDGMENTS

We acknowledge the IBMC «Human Proteome» Core Facility for assistance with the generation of mass-spectrometry data.

FUNDING

This study was supported by the Program of the Presidium of the Russian Academy of Sciences («Proteomic and Metabolomic Profile of Healthy Human»).

SUPPLEMENTARY

Supplementary Materials are available in online version at <http://dx.doi.org/10.18097/BMCRM00093>.

REFERENCES

1. Domon, B., & Aebersold, R. (2006). Mass spectrometry and protein analysis. *Science*, **312**(5771), 212-217. DOI: 10.1126/science.1124619
2. James, P., Quadroni, M., Carafoli, E., & Gonnet, G. (1993). Protein identification by mass profile fingerprinting. *Biochemical and biophysical research communications*, **195**(1), 58-64. DOI: 10.1006/bbrc.1993.2009
3. Pappin, D. J., Hojrup, P., & Bleasby, A. J. (1993). Rapid identification of proteins by peptide-mass fingerprinting. *Current biology*, **3**(6), 327-332. DOI: 10.1016/0960-9822(93)90195-T
4. Henzel, W. J., Watanabe, C., & Stults, J. T. (2003). Protein identification: the origins of peptide mass fingerprinting. *Journal of the American Society for Mass Spectrometry*, **14**(9), 931-942. DOI: 10.1016/S1044-0305(03)00214-9
5. Camerini, S., & Mauri, P. (2015). The role of protein and peptide separation before mass spectrometry analysis in clinical proteomics. *Journal of chromatography A*, **1381**, 1-12. DOI: 10.1016/j.chroma.2014.12.035
6. Miroshnichenko, Y. V., Petushkova, N. A., Teryaeva, N. B., Lisitsa, A. V., Zgoda, V. G., Belyaev, A. Y., & Potapov, A. A. (2015). Identification of Central Nervous System Proteins in Human Blood Serum and Plasma. *Bulletin of experimental biology and medicine*, **160**(1), 35-39. DOI: 10.1007/s10517-015-3092-5
7. Michalski, A., Damoc, E., Hauschild, J. P., Lange, O., Wieghaus, A., Makarov, A., ... & Horning, S. (2011). Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Molecular & Cellular Proteomics*, **10**(9), M111-011015. DOI: 10.1074/mcp.M111.011015
8. Miroshnichenko, Y. V., Petushkova, N. A., Moskaleva, N. E., Teryaeva, N. B., Zgoda, V. G., Ilgisonis, E. V., & Belyaev, A. Y. (2015). The possibility of using the PlasmaDeepDive™ MRM-panel in clinical diagnostics. *Biochemistry (Moscow) Supplement Series B: Biomedical Chemistry*, **9**(3), 283-289. DOI: 10.1134/S1990750815030051
9. Michalski, A., Cox, J., & Mann, M. (2011). More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. *Journal of proteome research*, **10**(4), 1785-1793. DOI: 10.1021/pr101060v
10. Laskay, U. A., Lobas, A. A., Srzentić, K., Gorshkov, M. V., & Tsybin, Y. O. (2013). Proteome digestion specificity analysis for rational design of extended bottom-up and middle-down proteomics experiments. *Journal of proteome research*, **12**(12), 5558-5569. DOI: 10.1021/pr400522h
11. Geiger, T., Cox, J., & Mann, M. (2010). Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation. *Molecular & Cellular Proteomics*, **9**(10), 2252-2261. DOI: 10.1074/mcp.M110.001537
12. Sajic, T., Liu, Y., & Aebersold, R. (2015). Using data-independent, high-resolution mass spectrometry in protein biomarker research: perspectives and clinical applications. *PROTEOMICS-Clinical Applications*, **9**(3-4), 307-321. DOI: 10.1002/prca.201400117
13. Gillet, L. C., Navarro, P., Tate, S., Röst, H., Selevsek, N., Reiter, L., ... & Aebersold, R. (2012). Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Molecular & Cellular Proteomics*, **11**(6), O111-016717. DOI: 10.1074/mcp.O111.016717
14. Schubert, O. T., Gillet, L. C., Collins, B. C., Navarro, P., Rosenberger, G., Wolski, W. E., ... & Aebersold, R. (2015). Building high-quality assay libraries for targeted analysis of SWATH MS data. *Nature protocols*, **10**(3), 426. DOI: 10.1038/nprot.2015.015
15. Conrads, T. P., Anderson, G. A., Veenstra, T. D., Paša-Tolić, L., & Smith, R. D. (2000). Utility of accurate mass tags for proteome-wide protein identification. *Analytical chemistry*, **72**(14), 3349-3354. DOI: 10.1021/ac0002386
16. Palmblad, M., Ramström, M., Markides, K. E., Håkansson, P., & Bergquist, J. (2002). Prediction of chromatographic retention and protein identification in liquid chromatography/mass spectrometry. *Analytical chemistry*, **74**(22), 5826-5830. DOI: 10.1021/ac0256890
17. Stanley, J. R., Adkins, J. N., Slys, G. W., Monroe, M. E., Purvine, S. O., Karpievitch, Y. V., ... & Dabney, A. R. (2011). A statistical method for assessing peptide identification confidence in accurate mass and time tag proteomics. *Analytical chemistry*, **83**(16), 6135-6140. DOI: 10.1021/ac2009806
18. Pridatchenko, M. L., Tarasova, I. A., Guryca, V., Kononikhin, A. S., Adams, C., Tolmachev, D. A., ... & Zubarev, R. A. (2009). Use of models of biomacromolecule separation in AMT database generation for shotgun proteomics. *Biochemistry (Moscow)*, **74**(11), 1195. DOI: 10.1134/S0006297909110030
19. Bączek, T., & Kalisz, R. (2009). Predictions of peptides' retention times in reversed-phase liquid chromatography as a new supportive tool to improve protein identification in proteomics. *Proteomics*, **9**(4), 835-847. DOI: 10.1002/pmic.200800544
20. Kunda, P. B., Benavente, F., Catalá-Clariana, S., Giménez, E., Barbosa, J., & Sanz-Nebot, V. (2012). Identification of bioactive peptides in a functional yogurt by micro liquid chromatography time-of-flight mass spectrometry assisted by retention time prediction. *Journal of Chromatography A*, **1229**, 121-128. DOI: 10.1016/j.chroma.2011.12.093
21. Moruz, L., Staes, A., Foster, J. M., Hatzou, M., Timmerman, E., Martens, L., & Käll, L. (2012). Chromatographic retention time prediction for posttranslationally modified peptides. *Proteomics*, **12**(8), 1151-1159. DOI: 10.1002/pmic.201100386
22. Lobas, A. A., Verenchikov, A. N., Goloborodko, A. A., Levitsky, L. I., & Gorshkov, M. V. (2013). Combination of Edman degradation of peptides with liquid chromatography/mass spectrometry workflow for peptide identification in bottom-up proteomics. *Rapid Communications in Mass Spectrometry*, **27**(3), 391-400. DOI: 10.1002/rcm.6462
23. Goloborodko, A. A., Mayerhofer, C., Zubarev, A. R., Tarasova, I. A., Gorshkov, A. V., Zubarev, R. A., & Gorshkov, M. V. (2010). Empirical approach to false discovery rate estimation in shotgun proteomics. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry*, **24**(4), 454-462. DOI: 10.1002/rcm.4417

24. Bateman, N. W., Goulding, S. P., Shulman, N. J., Gadok, A. K., Szumlanski, K. K., MacCoss, M. J., & Wu, C. C. (2014). Maximizing peptide identification events in proteomic workflows using data-dependent acquisition (DDA). *Molecular & Cellular Proteomics*, **13**(1), 329-338. DOI: 10.1074/mcp.M112.026500
25. Ting, Y. S., Egerton, J. D., Payne, S. H., Kim, S., MacLean, B., Käll, L., ... & MacCoss, M. J. (2015). Peptide-centric proteome analysis: an alternative strategy for the analysis of tandem mass spectrometry data. *Molecular & Cellular Proteomics*, **14**(9), 2301-2307. DOI: 10.1074/mcp.O114.047035
26. Hood, C. A., Fuentes, G., Patel, H., Page, K., Menakuru, M., & Park, J. H. (2008). Fast conventional Fmoc solid-phase peptide synthesis with HCTU. *Journal of peptide science: an official publication of the European Peptide Society*, **14**(1), 97-101. DOI: 10.1002/psc.921
27. MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., ... & MacCoss, M. J. (2010). Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*, **26**(7), 966-968. DOI: 10.1093/bioinformatics/btq054
28. Vizcaino, J. A., Csordas, A., Del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., ... & Xu, Q. W. (2015). 2016 update of the PRIDE database and its related tools. *Nucleic acids research*, **44**(D1), D447-D456. DOI: 10.1093/nar/gkv1145
29. Lange, V., Picotti, P., Domon, B., & Aebersold, R. (2008). Selected reaction monitoring for quantitative proteomics: a tutorial. *Molecular systems biology*, **4**(1), 222. DOI: 10.1038/msb.2008.61
30. Gianazza, E., Tremoli, E., & Banfi, C. (2014). The selected reaction monitoring/multiple reaction monitoring-based mass spectrometry approach for the accurate quantitation of proteins: clinical applications in the cardiovascular diseases. *Expert review of proteomics*, **11**(6), 771-788. DOI: 10.1586/14789450.2014.947966
31. Carr, S. A., Abbatiello, S. E., Ackermann, B. L., Borchers, C., Domon, B., Deutsch, E. W., ... & Liebler, D. C. (2014). Targeted peptide measurements in biology and medicine: best practices for mass spectrometry-based assay development using a fit-for-purpose approach. *Molecular & Cellular Proteomics*, **13**(3), 907-917. DOI: 10.1074/mcp.M113.036095

Received: 22.03.2019
 Revised: 15.04.2019
 Accepted: 29.04.2019

ИДЕНТИФИКАЦИЯ БЕЛКОВ ПЛАЗМЫ КРОВИ ЧЕЛОВЕКА С ИСПОЛЬЗОВАНИЕМ ВНУТРЕННИХ ПЕПТИДНЫХ СТАНДАРТОВ В ПАНОРАМНОМ ПРОТЕОМНОМ АНАЛИЗЕ

*А.Т. Копылов, О.В. Тихонова, Т.Е. Фарафонова, Н.А. Петушкова, Ю.В. Мирошниченко, В.Г. Згода**

Научно-исследовательский институт биомедицинской химии имени В.Н. Ореховича,
 119121, Москва, Погодинская ул. 10, *e-mail: victor.zgoda@gmail.com

Использование метода панорамной жидкостной хроматографии-масс-спектрометрии (ЖХ-МС/МС) позволяет идентифицировать в сложных протеомах до нескольких тысяч белков. Однако значительная часть протеома остается недоступной для идентификации из-за отсутствия или плохого качества МС/МС спектров. Описываемый в данной работе метод повышает вероятность идентификации белков плазмы крови человека путем выравнивания хроматографических данных образца смеси протеолитических пептидов и этого же образца, разбавленного синтетическими пептидами. Такая идентификация происходит в результате сопоставления тандемных масс-спектров синтетических пептидов с соответствующими выровненными хроматографическими пиками протеолитических пептидов. Используя данный подход, мы выявили 19 низко представленных белков плазмы крови человека, соответствующих 19 синтетическим пептидам, выбранным для исследования. Проверка идентификаций методом мониторинга диссоциативных переходов с изотопно-мечеными стандартами подтвердила наличие в плазме 17 белков.

Ключевые слова: масс-спектрометрия; идентификация белков; панорамный протеомный анализ; внутренний пептидный стандарт

ФИНИНСИРОВАНИЕ

Работа выполнена в рамках Программы президиума РАН «Протеомно-метаболический профиль здорового человека».

ДОПОЛНИТЕЛЬНЫЕ МАТЕРИАЛЫ

К данной статье приложены дополнительные материалы, свободно доступные в электронной версии на сайте журнала (<http://dx.doi.org/10.18097/BMCRM00093>).

Поступила: 22.03.2019; после доработки: 15.04.2019; принята к публикации: 29.04.2019.