Mass spectrometric monitoring of exhaled breath condensate proteome of a patient after lung transplantation

V. S. Kurova,[a] E. C. Anaev,[b] A. S. Kononikhin,[a, c] I. A. Popov,[a] K. Yu. Fedorchenko,[a] E. N. Nikolaev,[a] S. D. Varfolomeev[a], and A. G. Chuchalin[b]*

[a]*N. M. Emanuel Institute of Biochemical Physics Russian Academy of Sciences, 4 ul. Kosygina, 119334 Moscow, Russian Federation. Fax: +7 (499) 137 4101. E-mail: vkurova@yahoo.com*

[b]*Research Institute of Pulmonology, Federal Medicobiological Academy of Russia, 32 ul. 11 Parkovaya, 105077 Moscow, Russian Federation. Fax: +7 (495) 465 5264*

[c]*Institute for Energy Problems of Chemical Phisics, Russian Academy of Sciences, 38/12 Leninsky prosp., 119334 Moscow, Russian Federation*

Analysis of samples of exhaled breath condensate (EBC) is a promising noninvasive method for the control of the status of the human respiratory system. In the present work, the proteome of the EBC samples received from a patient with lung dystrophy at different times before and after bilateral lung transplantation is analyzed by ion cyclotron resonance mass spectrometry. Qualitative protein composition of EBC samples obtained during the first month correlates with the clinical data on the acceptance of the transplanted lungs (allograft). Fifteen months after the lung surgery, the protein spectrum was similar to the normal composition of EBC proteins. This result agrees with the medical conclusion about normal lung functioning. The results suggest that the mass spectrometric monitoring of the protein spectra of EBC may be a tool for noninvasive pulmonological diagnostics.

Key words: ion cyclotron resonance mass spectrometry, proteomic analysis, exhaled breath condensate, diagnostics of respiratory diseases.

Recent instrumental advances in mass spectrometry (MS) and chromatography allow perfection of the methods of medical diagnostics based on the studies of the structure and function of biomacromolecules. The sensitivity of the most recent mass spectrometers allows for the detection of femtomole concentration of proteins and peptides with the dynamic range of concentrations of several orders of magnitude. The accuracy of mass determination of charged species with the state-of-the-art methods reaches one millionth part of the ion mass. These characteristics are pertinent, in particular, to the Fourier transform ion cyclotron resonance mass spectrometers (FT-ICR MS). Coupling of high-performance liquid chromatography with FT-ICR mass spectrometry ensures the high performance of the analysis and allows the quantitative analysis of the components of complex mixtures.

In the present study, the potential of FT-ICR chromato-mass spectrometry was employed for identification of the protein composition of the exhaled breath condensate (EBC) of a person who had undergone lung surgery. The analysis of EBC is a promising method for the noninvasive diagnosis of human respiratory system, but the preparation of samples for the analysis significantly differs from that used for blood or urine. The method of EBC collection and analysis of protein composition by chromato-mass spectrometry has previously been practized with material obtained from healthy non-smoking donors. Below we present the results of the qualitative analysis of EBC protein composition of a patient with bullous emphysema who underwent bilateral lung transplantation. Thorough medical monitoring of the patient in pre- and post-surgery periods allowed us to obtain a unique set of samples of exhaled breath condensate.

Experimental

Collection of exhaled breath condensate. The condensate of exhaled air was collected from a patient under artificial lung ventilation the day before surgery, two times on the first day after surgery and daily afterwards during spontaneous respiration. The condensate of the air exhaled for 10 min was collected in a teflon receiver cooled down to −10 °C using the EcoScreen condenser (VIASYSHealthcare — Erich Jaeger, Hohberg, Germany). After thawing, the condensates were transferred into low temperature-resistant polypropylene test-tubes with a low protein-absorbing surface. The samples were frozen and stored at −85 °C. For further analysis of the protein composition, definite amount

Published in Russian in *Izvestiya Akademii Nauk. Seriya Khimicheskaya,* No. 1, pp. 284—288, January, 2010.

1066-5285/10/5901-0292

of the sample was lyophilized and prepared for digestion with trypsin followed by mass spectrometry.

Identification of the EBC proteins. Lyophilized mixtures of the proteins of the EBC were digested with trypsin. Peptides formed were fractionated by high performance nanoflow chromatography (Agilent 1100, Agilent Technologies, USA) in a gradient mode. HPLC was performed by gradient elution: the mobile phase A was 0.1% formic acid in water, and the mobile phase B (80% acetonitrile, 20% water, 0.1% formic acid) was changed from 3% to 50% over 40 min; the elution rate was 0.3 μL/min. Peptides were identified by the exact masses of the protonated ions and their fragments obtained by collision-induced dissociation in the ion trap in combination with ion cyclotron resonance mass spectrometry (7-Tesla Finnigan LTQ FT mass spectrometer, Thermo Electron, Bremen, Germany) with electrospray ionization. The spectra were analyzed using Bioworks Browser 3.1 SR1 software (Thermo Electron, Bremen, Germany) producing the list of exact masses of the peptides and their fragments. The mass data obtained were used for searching and identification of proteins in the NCBInr (non redundant) database using the Mascot software (Matrix Science, London, United Kingdom; version 2.0.04).

Results and Discussion

Collection of the exhaled breath condensate

The choice of the material of the study was motivated by the availability of the unique set of samples systematically collected under medical control. Thus, a comparative analysis of the protein spectrum of EBC can demonstrate the dynamics of the change in the state of the respiratory system of the patient during treatment.

When collecting EBC samples,[1] one should focus on the exclusion of contamination of the samples by saliva and nasal mucosa proteins. Donors were requested to thoroughly rinse their mouths with distilled water and breathed with their noses clipped so that the air entered the lungs through the mouth only. The design of the device for condensate provides for the retention of saliva, so that only the air exhaled by the patient reaches the collector where it is condensed at low temperature. Salivary contamination of the samples can be revealed by the presence of a saliva-specific protein alpha-amylase in the sample spectrum. The samples that were found to contain alpha-amylase were excluded from the study.

According to the previously obtained data,[1] the concentration of proteins in EBC of healthy people does not exceed 1 μg/mL. Taking in the account that the recommended duration of EBC collection (10 min) allows the collection of not more than 3 mL of the sample material, a loss of proteins in the processes of sample collection and analysis should be reduced to a maximum degree. For this purpose, the collection of EBC was performed using teflon or polypropylene collector, and the thawed condensate was transferred from the collector to the test tubes with low absorption surface. All other manipulations with the EBC samples were performed *in situ* up to withdrawal of a portion for chromato - mass spectrometric analysis.

Identification of proteins in the exhaled air condensate

Identification of proteins according to mass spectra was carried out using the Mascot search software, which performs search and identification of proteins from primary sequence databases of proteins and the corresponding genes. From all the currently available softwares, the search engine Mascot is the most powerful tool for the comparison of experimental and theoretical data in the identification of proteins and peptides. Its principle of operation is based on the MOWSE scoring algorithm (MOlecular Weight SEarch), developed in 1993 by Darryl Pappin.[2]

The first stage of identification is to compare the theoretical masses of peptides that can be obtained under certain conditions by fragmentation of proteins from the considered database with the set of experimental mass spectrometry data. The Mascot dialog window allows one to specify a database for the search, the protein fragmentation method (chemical or enzymatic hydrolysis), and potential modifications of the amino acid residues (predefined or theoretically possible under these conditions). One can also optimize the search by selecting a biological taxon, indicating the expected mass of the required protein, and determining the accuracy with which the experimental and theoretical data should match. Next, after downloading a list of experimental masses and charge states of the ions, one can execute the search. The Mascot software allows one to access various databases, the most versatile of which are MSDB, NCBI, SwissProt, and dbEST.

The search algorithm is described in detail on the site of the Mascot developer www.matrixscience.com. It allows the identification based on the calculated masses of the peptides specific to each protein. The match of each calculated value of the mass of the peptide with the experimental value is characterized by a certain error. Instead of just counting the number of matching peptides, MOWSE-algorithm uses the empirically derived factors to determine the statistical weight of each peptide match. A frequency factor matrix F is calculated at the time of the database access and is created in such a way that each row corresponded to an interval of 100 Da of the peptide masses and each column corresponded to the interval of 10 kDa of intact protein masses. Each sequence processing from the database produces the corresponding matrix element f_{ij}, allowing one to find the statistical distribution of the peptide masses as a function of the protein masses. Elements of each column are normalized to the largest value in column $|f_{ij}|_{\max \text{ in column}(j)}$. The mass matrix M is calculated by the formula

$$m_{i,j} = f_{i,j}/|f_{i,j}|_{\max \text{ in column}(j)}. \quad (1)$$

After comparing the experimental and theoretical data, one assigns a score (a value describing the level of reliability of the identification results *Score*) for each element according to formula:

$$Score = 50000/(M_{prot} \cdot \Pi_n m_{i,j}), \quad (2)$$

where M_{prot} is the molecular weight of each matching protein and Π_n is the product of n elements of the matrix, which is calculated from the Mowse mass matrix M for each match of the experimental and theoretical masses $m_{i,j}$. As a result, the score is the absolute value, which shows how accidental the observed mass is.

The algorithm of search using the peptide masses (Peptide Mass Fingerprint) was implemented before that based on tandem mass spectrometry data (MS/MS-Search), where the search pattern for both types of analysis is the same. In the case of tandem mass spectrometry, the mass of intact protein in the expression for the score is replaced by the mass of the peptide under consideration, and its fragments obtained by physical methods directly in the mass spectrometer are treated as peptides. Mascot presents the results in tabular form where proteins are placed according to their score. Each protein is accompanied by a list of peptides whose theoretical masses match the experimental data within a given error. For the analysis of tandem mass spectrometry data, Mascot calculates the theoretical peptide fragmentation spectra and compares them with the experimentally obtained masses of the fragments.

The results of Mascot identification are illustrated by the analysis of the mass spectrum of the tryptic hydrolysate of a sample obtained on the tenth day after the surgery. A doubly charged ion with m/z 968.057198 observed in the spectrum represents the mass of the uncharged spacies 1934.099844. The desmoplakin III peptide found in the NCBI library with the sequence TLVTQNSGVEALIHAILR has the mass difference from the experimental value of 0.01 Da. Table 1 shows the theoretical masses of the fragments of the peptide TLVTQNSGVEALIHAILR obtained by collision-induced dissociation in the ion trap, which coincide with the fragments of the m/z 968.057198 cation with the accuracy to within 0.5 Da. Table 1 shows the fragments of the so-called b- and y-series, which are formed upon cleavage of the peptide bond. The b-series ions represent the protonated N-terminal parts of the peptide (to the left of the cleaved peptide bond), whereas the y-series ions are the respective protonated C-terminal (right) parts. Table 1 shows also the masses of the doubly charged ions of these fragments and the masses of the fragments with the loss of ammonia (b* and y*) or water (b^0 and y^0) molecules. Using the NCBI library, Mascot identified the fragment by its ion mass and the masses of its fragments as the peptide TLVTQNSGVEALIHAILR

Table1. Theoretical masses of the fragment TLVTQNSGVEALIHAILR of desmoplakin III matching the experimental masses of the fragments produced by the collision - induced dissociation of a doubly-charged cation with m/z 968.057198, the accuracy $\leq$ 0.5 Da

Fragment number	b	b^{++}	b*	b^0	AA	y	y^{++}	y*	y^{*++}	y^0	y^{0++}	Fragment number
1					T							18
2					L							17
3	314.21				V		860.99					16
4				397.24	T	1621.90	811.45	1604.88				15
5			526.29		Q	1520.85		1503.83	752.42		751.93	14
6					N	1392.80		1375.77		1374.79		13
7					S	1278.75						12
8		401.21			G	1191.72						11
9	900.48				V	1134.70						10
10					E	1035.63			509.81		509.31	9
11	1100.56		1083.53		A	906.59						8
12	1213.64		1196.62		L	835.55						7
13	1326.73	663.87		1308.72	I	722.47						6
14	1463.79		1446.76		H	609.38						5
15	1534.82				A	472.32						4
16	1647.91				I	401.29						3
17	1760.99				L	288.20						2
18					R							1

Note. The ions of the b- and y-series formed upon the peptide bond cleavage correspond to the protonated N-terminal (to the left of the cleaved bond) fragment and C-terminal (to the right of the cleaved bond) fragments of the peptide; b^{++} and y^{++} are the masses of doubly charged ions of the fragments of the b- and y-series; b* and y* are the masses of the fragments of the b- and y-series following the loss of the ammonia molecule; b^0 and y^0 are the masses of the fragments of the b- and y-series following the loss of the water molecule. AA is an amino acid in the peptide.

of desmoplakin III with the score of 75. It should be noted that in this case Mascot suggested a single amino acid sequence. This is not always the case, and peptides with similar sequences may belong to completely different proteins. In such cases one should consider various proteins offered by Mascot, taking into account the information about the sample, localization of the proteins in the organism, etc.

It has previously been shown with the standard mixtures of proteins that the sensitivity of nanoflow chromatography-mass spectrometry allows reliable identification of as little as 40 fmol of protein in a sample[1] introduced into the chromatograph. As it turned out, this sensitivity is sufficient to detect over twenty different proteins in individual samples of EBC of healthy people[1].

In the sample collected before surgery, only keratin 9, dermcidin, lysozyme, and ubiquitin were identified with certainty but with a low score. Such a limited range of proteins in the condensate correlates with the patient´s lung dysfunction, which affects the dynamics of exhalation and reduces the number of non-volatile compounds, including proteins, in the EBC.

The samples collected on the first day after surgery also revealed less diverse spectrum of proteins as compared to that of healthy donors.[1,3] We found proteins characteristic of the majority of EBC samples collected from healthy non-smokers[1,3]: keratin pairs 1/10, 2/9, and 5/14 and dermcidin.

In the mixtures of the samples of the fourth and eighteenth days after the surgery, we identified both «normal» proteins and non-keratin cell components: desmoplakin and profilaggrin (Table 2). Profilaggrin (filaggrin) is a calcium-binding protein of the cytoskeleton.[4] The identification of desmoplakin, a component of desmosomes, suggested the increased degradation of the respiratory epithelium.[5] Also found were the fragments of a serine protease inhibitor and bleomycin hydrolase (Table 3). Identification of such a set of proteins is consistent with the clinical data on the status of the transplanted lungs in the early postoperative period and with the therapy of the patient with cytostatics similar to bleomycin. In the individual condensate obtained on the tenth day after the operation, except for the components of desmosomes (desmoglein, desmoplakin III), we also detected a fragment of annexin A1, a protein participating in the suppression of inflammatory processes. Annexin 1 inhibits the activity of phospholipase A2 and cyclooxygenases-1 and -2 retarding the

Table 2. Peptides identified in the tryptic hydrolysates of the exhaled breath condensates of the patients that are not characteristic of healthy donors

m/z (I_{rel}(%))	Protein, identifying peptides	Peptide score	Peptide name
	The analysis of the sample collected on the tenth day after the surgery (with deprotonated keratin 9 as the reference)		
792.86 $[M+2H]^{2+}$ (2)	GPYESGSGHSSGLGHR	99	Hornerin
649.97 $[M+3H]^{3+}$ (1)	QSLGHGQHGSGSGQSPSPSR	30	
812.57 $[MH]^{+}$ (5)	LVQLLVK	33	Desmoplakin III
618.85 $[M+2H]^{2+}$ (2)	VSVELTNSLFK	44	
968.06 $[M+2H]^{2+}$ (6)	TLVTQNSGVEALIHAILR	75	
767.37 $[M+2H]^{2+}$ (1)	QEPSDSPMFIINR	63	Desmoglein I
818.43 $[M+2H]^{2+}$ (1)	YQGTILSIDDNLQR	90	
1023.06 $[M+2H]^{2+}$ (3)	ISGVGIDQPPYGIFVINQK	78	
1121.14 $[M+2H]^{2+}$ (0.3)	ASAISVTVLNVIEGPVFRPGSK	89	
851.95 $[M+2H]^{2+}$ (3)	GLGTDEDTLIEILASR	70	Annexin I
	The analysis of samples collected on the third and eighteenth days after surgery (with monoprotonated desmoplakin I as the reference)		
793.95 $[MH]^{+}$ (100)	ALLQAILQTEDMLK	82	Desmoplakin I
645.71 $[M+3H]^{3+}$ (3)	TLVTQNSGVEALIHAILR	54	Desmoplakin III
792.86 $[M+2H]^{2+}$ (5)	GPYESGSGHSSGLGHR	127	Hornerin
874.39 $[M+2H]^{2+}$ (1)	GSGSGQSPSSGQHGTGFGR	74	
649.97 $[M+3H]^{3+}$ (2)	QSLGHGQHGSGSGQSPSPSR	31	
776.86 $[M+2H]^{2+}$ (0.6)	GYSPTHREEEYGK	28	Profilaggrin
840.90 $[M+2H]^{2+}$ (0.7)	KGYSPTHREEEYGK	69	
912.00 $[M+2H]^{2+}$ (0.8)	NIFFSPLSLSAALGMVR	54	Serine proteinase inhibitor
653.37 $[M+2H]^{2+}$ (2)	IGPITPLEFYR	39	Bleomycin hydrolase

Table 3. The dynamics of protein composition of the exhaled air condensate

Proteins	1 Day before surgery	After surgery			
		1st day	4th and 18th day*	10th day	15 months
Dermcidin	+	+	+	+	+
Keratin 1	–	+	+	+	+
Keratin 2	–	+	+	+	+
Keratin 4	–	–	–	+	–
Keratin 5	–	+	+	+	+
Keratin 6	–	+	–	+	–
Keratin 8	–	–	+	–	–
Keratin 9	+	–	+	+	+
Keratin10	–	+	+	+	+
Keratin 14	–	+	+	+	+
Keratin 16	–	–	–	+	–
Ubiquitine	+	–	–	+	+
Lysozyme	+	–	–	–	–
Lipocalin	–	–	–	–	+
Cytochrome C	–	+	–	–	+
Desmoglein	–	+	–	+	—
Cystatin A	–	–	–	+	+
Hornerin	–	–	+	+	–
Desmoplakin	–	–	+	+	–
Annexin 1	–	–	—	+	–
Filaggrin	–	–	+	+	–
Serine (cysteine) protease inhibitor	–	–	–	+	–
Bleomycine hydrolase	–	–	–	+	–
Prolactin-induced protein	–	–	–	+	–
Glyceraldehyde 3-phosphate dehydrogenase	–	–	–	+	–
Arginase 1	–	–	–	+	–
Peroxiredoxin	–	–	–	+	–

* Mixture of samples.

biosynthesis of prostaglandins and leukotrienes.[6,7] In addition, annexin 1 promotes inflammatory apoptosis cells.[8] The appearance of annexin A1 in the sample suggests inflammation of the airways of the patient, which is consistent with the results of clinical examination in this period.

In the condensate obtained 15 months after the operation, only «normal» proteins were identified, characteristic of people without respiratory pathology: keratins 1/10, 2/9, and 5/14, dermcidin, cystatin A, and ubiquitin.

The dynamics of changes in the protein composition of the EBC during the postoperative rehabilitation is presented in Table 3. It can be concluded that the results of the analysis of protein composition of the EBC are consistent with the results of clinical condition of the patient who underwent lung transplantation, and, thus, are characteristic of the process of allograft acceptance.

References

1. V. S. Kurova, E. C. Anaev, A. S. Kononikhin, K. Y. Fedorchenko, I. A. Popov, T. L. Kalupov, D. O. Bratanov, E. N. Nikolaev, S. D. Varfolomeev, *Clin. Chem. Lab. Med.*, 2009, 47, 706.
2. D. J. Pappin, P. N. Hojrup, A. J. Bleasby, *Curr. Biol.,* 1993, **3**, 327.
3. V. S. Kurova, I. N. Kurochkin, G. R. Kalamkarov, A. E. Bugrova, K. Y. Fedortchenko, S. D. Varfolomeev, *Biotechnol Adv.*, 2009, **27**, 945.
4. N. G. Markova, L. N. Marekov, C. C. Chipev, S. Q. Gan, W. W. Idler, P. M. Steinert, *Mol. Cell Biol.*, 1993, **13**, 613.
5. U. S. Chentsov, *Introduction to cell biology*, IKC "Academkniga", Moscow, 2005, 495 pp.
6. F. D´Acquisto, A. Merghani, E. Lecona, G. Rosignoli, K. Raza, C. D. Buckley, R. J. Flower, M. Perretti, *Blood*, 2007, **109,** 1095.
7. F. D´Acquisto, M. Perretti, R. J. Flower, *Br. J. Pharmacol.,* 2008, **155**, 152.
8. L. Parente, E. Solito, *Inflamm. Res.*, 2004, **53**, 125.

Received July 30, 2009,
In revised form October 26, 2009