Clin Chem Lab Med 2009;47(6):706–712 © 2009 by Walter de Gruyter • Berlin • New York. DOI 10.1515/CCLM.2009.166

Proteomics of exhaled breath: methodological nuances and pitfalls

Viktoria S. Kurova*, Eldar C. Anaev, Alexey S. Kononikhin, Kristina Yu. Fedorchenko, Igor A. Popov, Timothey L. Kalupov, Dmitriy O. Bratanov, Eugenie N. Nikolaev and Sergey D. Varfolomeev

N.M. Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Moscow, Russia

Abstract

Background: The analysis of exhaled breath condensate (EBC) can be an alternative to traditional endoscopic sampling of lower respiratory tract secretions. This is a simple non-invasive method of diagnosing respiratory diseases, in particular, respiratory inflammatory processes.
Methods: Samples were collected with a special device-condenser (ECoScreen, VIASYS Healthcare, Germany), then treated with trypsin according to the proteomics protocol for standard protein mixtures and analyzed by nanoflow high-performance liquid chromatography tandem mass spectrometry (HPLC-MS/MS) with a 7-Tesla Finnigan LTQ-FT mass spectrometer (Thermo Electron, Germany). Mascot software (Matrixscience) was used for screening the database NCBInr for proteins corresponding to the peptide maps that were obtained.
Results: EBCs from 17 young healthy non-smoking donors were collected. Different methods for concentrating protein were compared in order to optimize EBC preparations for proteomic analysis. The procedure that was chosen allowed identification of proteins exhaled by healthy people. The major proteins in the condensates were cytoskeletal keratins. Another 12 proteins were identified in EBC from healthy non-smokers. Some keratins were found in the ambient air and may be considered exogenous components of exhaled air.
Conclusions: Knowledge of the normal proteome of exhaled breath allows one to look for biomarkers of different disease states in EBC. Proteins in ambient air can be identified in the respiratory tract and should be excluded from the analysis of the proteome of EBC. The results obtained allowed us to choose the most effective procedure of sample preparation when working with samples containing very low protein concentrations.
Clin Chem Lab Med 2009;47:706–12.

*Corresponding author: Dr. Viktoria Kurova, N.M. Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, 4/11 Kosygin Street, 119334 Moscow, Russian Federation
Fax: +7 495 137 41 01, E-mail: kurova@sky.chph.ras.ru
Received November 26, 2008; accepted April 21, 2009

Keywords: exhaled breath condensate (EBC); nanoflow HPLC-tandem mass spectrometry; proteomics.

Introduction

Exhaled breath contains a wide range of different substances including non-volatile salts, lipids and proteins (1–5). The collection of exhaled breath condensate (EBC) is a non-invasive method for obtaining samples from fluids lining the respiratory surfaces. Condensed exhaled breath is a promising source in which biomarkers of respiratory diseases may be identified (6).

Even though various methods and devices are now available, the efficiency of these methods for recovering airway mediators is not established (7). Although the American Thoracic Society and European Respiratory Society Task Force on Exhaled Breath Condensate recently published general recommendations on collection of samples, there are still some methodological pitfalls and unresolved questions (8). None of the works provide complete information about the normal condensate proteome, or propose an optimal method for preparing a probe for proteomic analysis. Using a standard proteomic approach, we optimized the analysis and identified different proteins, in addition to keratins, in condensates of 17 young healthy non-smoking donors. The proteins that are found in exhaled breath of healthy non-smoking people will allow comparison of the proteome of healthy people with that of individuals having diseases of the respiratory tract.

Materials and methods

EBC collection and processing

The Ethics Committee of N.M. Emanuel Institute of Biochemical Physics, Russian Academy of Science (IBCP RAS) approved our research on EBCs, and informed consent was obtained from the participants.

The most commonly used condensers are the EcoScreen (Erich Jaeger GmbH, Hoechberg, Germany), RTube (Respiratory Research Inc, Charlottesville, VA, USA) and in-house glass or Teflon devices (7, 9). Compared with RTube™, ECoScreen allows larger volumes of exhaled breath to be collected, and allows detection of proteins and lipids with greater sensitivity (7). It should be noted that breath condenser coatings can affect measurement of biomarkers in EBC. According to the literature, higher albumin recovery in vitro was obtained for a condenser system containing glass and silicone coating (9). Recovery of albumin looks more reproducible for polypropylene, aluminum and Teflon coatings (9) compared with glass or silicone coatings. In addition, protein absorption should be not significant at the temperature of collection (–10°C) using the ECoScreen.

Based on these data, we chose chemically inert and thermostable Teflon collection vials that were compatible with the ECoScreen system.

Each sample represents a breath condensate exhaled from tidal breath over a 10-min period. Samples were collected into Teflon tubes at –10°C with the ECoScreen condenser (VIASYS Healthcare, Hoechberg, Germany). The construction of this system prevents saliva from getting into a receiver.

Minute ventilation and respiratory rate were measured by a pneumotachograph (ECoVent, VIASYS Healthcare, Germany). Thawed condensate was transferred immediately into low temperature-resistant polypropylene test-tubes with a low protein-absorbing surface (Corning, cat. No. 430656). Total volumes of exhaled air and breath condensates are presented as mean ± SE of the mean.

One milliliter of each sample was transferred into an individual tube, frozen at –70°C and then lyophilized in the same tube. Residue of each sample was frozen at –70°C and stored for comparative studies.

Condensate collection with a protective filter

Two donors were asked to breathe using a protective filter. The protective filter prevented dust in the ambient air from getting into the condensate. The experiment was done in order to determine proteins present in the ambient air from those obtained from the lungs and respiratory tract. The results were compared with normal EBC collection.

Mass spectrometric analysis

All nanoflow liquid chromatography tandem mass spectrometry (LC-MS/MS) experiments were performed in duplicate on a nanoHPLC Agilent 1100 system (Agilent Technologies, Santa Clara, CA, USA) in combination with a 7-Tesla Finnigan LTQ-FT mass spectrometer (Thermo Electron, Bremen, Germany) equipped with a nanospray ion source (in-house system). HPLC was performed by gradient elution: the mobile phase B (80% acetonitrile, 20% water, 0.1% formic acid) was changed from 3% to 35% over 120 min; the elution rate was 0.3 μL/min. All MS/MS data were analyzed using Mascot (Matrix Science, London, UK; version 2.0.04). Proteins were identified against the NCBInr (National Center for Biotechnology Information) protein sequence database selected for Homo sapiens and taking into consideration the specificity of the protein hydrolysis with trypsin. Mass tolerance for parent ion mass spectrometry (MS) peaks were 10.0 ppm and 0.50 Da for MS/MS peaks. Significance of thresholds was <0.05. Variable modifications were not considered. The instrument set for Mascot search was specified as ''electrospray ionization (ESI)-Trap''.

The peptide Mascot score was considered significant if it indicated identity between the observed and theoretical peptide sequences. Proteins were considered to be correctly identified if at least one unique peptide was revealed with the Mascot probability-based score >70, or if more than two unique tryptic peptides were defined with significant scores.

Tryptic hydrolysis of standard protein mixture

A mixture of 48 proteins (0.5 pmol) from 6 to 80 kDa (MS standards UPS1, Sigma Aldrich, St. Louis, MO, USA) was hydrolyzed with 0.1 μg trypsin in 20 μL of 30 mM NH_4HCO_3 at 30°C for 20 h. The concentration of each protein was 0.025 pmol/μL. The reaction was stopped with addition of 2 μL of 1% formic acid. The resultant solution (hydrolysate) was analyzed using nanoflow LC-MS/MS in duplicate as described above. The sample volume that was injected was 4 μL.

EBC delipidation and protein precipitation (10)

Standard mixture A mixture (10 μL) of standard proteins (MS-CAL1, Sigma) was mixed in a 200-μL tube with 75 μL methanol followed by addition of 75 μL chloroform. The mixture contained 100 pmol of cytochrome C, 1 pmol of apomyoglobin, 1 pmol of aldolase and 1 pmol of bovine serum albumin (BSA). The mixture was centrifuged at 10,000 *g* for 5 min. Following aspiration of the organic phases, the tube was closed with porous film and the protein pellet was dried at ambient temperature, resuspended in 20 μL 50 mM NH_4HCO_3 and hydrolyzed with 0.02 μg of trypsin at 30°C for 16 h. The reaction was stopped with 0.1% formic acid. The resultant peptide mixture was analyzed using nanoflow high-performance liquid chromatography tandem mass spectrometry (HPLC-MS/MS) as described above, 1 μL ($\leq$50 fmol of each protein) of each probe was added to each mixture.

EBC Lyophilized samples of EBC were re-suspended in 10 μL of water and extracted with methanol and chloroform as described above.

Protein precipitation with sodium deoxicholate and trichloroacetic acid (11)

Into a 1.5-mL tube, 100 μL of 0.15% sodium deoxycholate (DOC) was added to 1 mL of EBC with the standard proteins (MS standards UPS1, Sigma), using an equimolar ratio (0.5 pmol). The sample was vortexed and chilled for 10 min at room temperature. Fifty microliter of 100% trichloroacetic acid (TCA) was added and the sample then vortexed and chilled for 15 min at –18°C. The mixture was centrifuged for 5 min at 10,000 × *g*. The supernatant was removed carefully. The remaining pellet was washed twice with 150 μL methanol/chloroform mixture (1/1 vol) and dried in the tube under a porous film at ambient temperature. Once dry, the pellet was resuspended in 20 μL 50 mM NH_4HCO_3 and hydrolyzed with 0.005 μg/μL of trypsin at 30°C for 20 h. The concentration of each protein was 0.025 pmol/μL. The reaction was stopped with 2 μL of 1% formic acid (up to 0.1% of acid in the sample). The resultant solution (hydrolysate) was analyzed using nanoflow LC-MS/MS. The sample volume that was injected was 4 μL.

Protein ultrafiltration on Microcon YM-3 membranes

A mixture of 1 mL of EBC with 2 μL of a standard mixture (MS standards UPS1, Sigma) with an equimolar ratio of proteins (0.5 pmol) was concentrated to 50 μL on a 3-kDa Microcon filter by centrifugation at 10,000 *g* according to the Microcon user's guide (www.millipore.com). Microcon YM-3 provides high recovery (above 90%) of proteins larger than 3 kDa (www.millipore.com/techpublications/tech1/PF185). The sample was diluted with 50 mM NH_4HCO_3 buffer (pH 8.1) and hydrolyzed with 0.005 μg/μL of trypsin at 30°C for 20 h. The concentration of each protein was $\leq$0.008 pmol/μL. The reaction was stopped with formic acid (up to 0.1% of acid in the sample). The resultant solution (hydrolysate) was analyzed using nanoflow LC-MS/MS. The volume of sample that was injected was 6 μL.

Recovery of peptides after concentrating on Millipore ZipTip C18

Mixtures of proteins with molecular weights from 12 to 67 kDa, at an equimolar ratio, were hydrolyzed with trypsin

(0.02 μg/μL, 50 mM NH_4HCO_3, 30°C 16 h). Solutions of the hydrolysate, each containing 0.5 pmol in 100 μL of a mixture of previously obtained EBC hydrolysate, were concentrated on the ZipTip cartridges with reversed phase C18 according to a ZipTip user guide (www.millipore.com). Peptides were rinsed in 10 μL of 70% methanol containing 0.1% formic acid. The samples were dried and re-dissolved in 10 μL of 0.1% formic acid. For use as the control, 0.5 pmol of standard hydrolysate was dissolved in 10 μL of 50% methanol containing 0.1% formic acid (without treatment on ZipTip), dried and re-dissolved in 10 μL of 0.1% formic acid. The resultant peptide mixture was analyzed using nanoflow LC-MS/MS as described above; 1 μL (50 fmol of each protein) of each probe was added to each mixture.

Proteomic analysis of EBC

Each of the 17 lyophilized EBC samples were dissolved in 40 μL of 50 mM NH_4HCO_3 at pH 8.0. One half of each probe was hydrolyzed with 0.02 μg trypsin at 30°C for 16 h. The reaction was stopped with 0.1% formic acid. The mixture was analyzed using nanoHPLC-MS/MS in duplicate. The volume injected was 4 μL. Empty probes (treatment with reaction buffer in an empty test-tube) and BSA (Serva Electrophoresis GmbH, Heidelberg, Germany) were used as controls of purity and the completeness of lysis by trypsin.

Electrophoresis

0.25 mL of EBC was lyophilized, re-dissolved in 20 μL of Laemmli Sample Buffer and separated electrophoretically on a 15% acrylamide gel (8×10 cm plates). The gels were stained with Silver Stain Plus Kit (Bio-Rad, USA). Sensitivity of Silver Stain Plus is annotated as approximately 0.5 ng/band. Precision Plus Protein™ standards (Bio-Rad) were used as controls. Amounts of standard proteins were 0.45–2.25 ng/band according to the Bio-Rad annotation.

Results

General features of EBC

The study population was represented by seven males [average age: 29±3 years old (SE)] and ten females [average age: 26±4 years old (SE)]. The mean volume of exhaled air over 10 min was 176±26 L (n=17). The mean volume of EBC over 10 min was 2.1±0.1 mL (n=17). EBC volumes comply with specifications of the device reported by the manufacturer.

Protein concentrations in EBC have been estimated based on silver-stained sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (figures not shown). Intensities of the EBC-bands are noticeably lower than those of the standards. Taking into account the sensitivity of silver stain (0.5 ng/band), one can conclude that 250 μL of the EBC-aliquot contains <1 μg of protein. Thus, protein concentrations in EBC are <1 μg/mL.

Optimization of the protein concentrating procedure

Analytical sensitivity Three protein concentrating procedures have been tested using standard aqueous mixtures with low protein concentrations, and in the presence or absence of the EBC matrix. Concentrated proteins were digested with trypsin and identified using nanoHPLC-ESI-MS/MS as described above. Analysis of a standard protein mixture before drying allowed us to identify 40 fmol of each protein with a high coverage of their primary sequences (Table 1). This amount was considered to be the limit of sensitivity for this method.

Freeze-drying The number of peptides identified following freeze-drying of proteins decreased by 50%–90% (Table 1, Figure 1). However, all protein standards were found with reliable scores (Table 1).

Similar results were obtained for equimolar mixtures of 48 recombinant human proteins (data not shown). Forty-three proteins were identified in the mixture before lyophilization, but this number decreased to 33 following lyophilization (Figure 2). The number of peptides with individual scores estimated using Mascot that were considered to be good enough for identification of proteins also decreased.

Influence of EBC matrix on protein identification Thirty-three proteins were identified in the lyophilized standard mixture without the EBC matrix and 37 proteins were identified after drying the same mixture with individual EBC. Thus, the EBC matrix does not reduce the number of the proteins that can be identified.

Precipitation and ultrafiltration Lyophilization of equimolar mixtures of recombinant human proteins with individual EBC was compared to trichloroacetic acid/sodium deoxycholate (TCA/DOC) precipitation and ultrafiltration (Figure 2). Following freeze-drying of the standard mixture with individual EBC, 37 proteins were identified. Thirty-three proteins were identified following precipitation with TCA/DOC. Following ultrafiltration, only two proteins were revealed demonstrating the unsuitability of this method.

Solid-phase extraction with ZipTipC18 Solid-phase extraction with ZipTipC18 could be very useful for

Table 1 Identification of proteins (40 fmol) in aqueous solution before and after lyophilization.

Standard proteins	MW, Da	Mascot score		Peptides identified	
		Before drying	After drying	Before drying	After drying
Cytochrome C (equine)	12,360.96	174	67	12	2
Myoglobin (equine)	16,951.27	254	120	12	6
Aldolase (rabbit muscle)	39,211.28	355	233	23	12
Albumin (bovine serum)	66,429.09	292	118	22	2

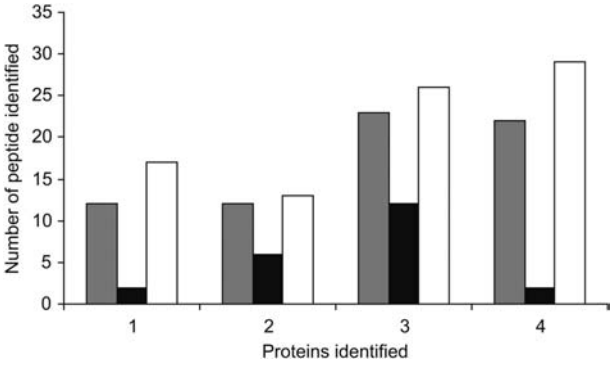

Figure 1 The number of identified peptides in non-prepared samples (gray), samples after lyophilization (black) and samples after precipitation with methanol/chloroform (white). 1, cytochrome C (equine); 2, myoglobin (equine); 3, aldolase (rabbit muscle); 4, albumin (bovine serum).

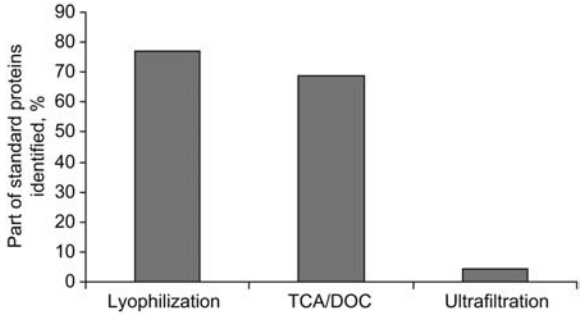

Figure 2 Number of proteins identified in a standard mixture of 48 recombinant proteins with the EBC matrix using different concentrating procedures.

sample preparation for nanoflow LC-ESI-MS/MS, especially in cases of pooled condensates which contain much higher concentrations of salts and lipids when compared with individual samples. However, introduction of an additional preparation stage can decrease the yield of identified peptides, thus decreasing the Mascot scores of proteins. To estimate peptide recovery after ZipTipC18, hydrolysates of standard proteins were analyzed. The number of peptides identified in three samples after ZipTipC18 was dramatically lower compared with non-extracted samples, and SD of the number of peptides associated with each protein were high (Table 2). Thus, this procedure is not recommended for EBC preparation.

EBC desalting and delipidation Methanol/chloroform precipitation was tested as a procedure for desalting and delipidation. The present method of using methanol/chloroform/water in a 7.5/7.5/1 ratio provided a higher number of identified peptides for all proteins in a standard mixture, compared with the hydrolyzed protein mixtures introduced without any preparation (Table 3 and Figure 2). This is likely due to denaturation of proteins in the organic solvents.

In spite of the effectiveness of this treatment for protein identification, we could not use it for non-concentrated samples because of high sample volumes and low protein concentrations. However, this procedure can be used after freeze-drying for improving hydrolysis.

Proteomic analysis of individual EBC of young healthy non-smoking donors

Proteomic investigation of EBC in 17 young healthy non-smoking volunteers revealed that the major proteins are cell keratins. However, the spectrum is polymorphous in different people. Pairs of cytoskeletal keratins 1/10 and 2/9 are invariant in most samples. No mutations in the sequences of these proteins from healthy volunteers were detected. We found that other keratins are substantially different for individual samples (Table 4).

In addition to keratins, dermcidin, prostaglandin H2 D-isomerase (PGDS2) and α_1-microglobulin/bikunin precursor (AMBP), ubiquitin and cystatin A also frequently occurred ($\geq$30% of donors, Table 4).

Analysis of the keratin background of ambient air and EBC collected with a protective filter

Keratins found in EBC belonged primarily to skin epithelia and hair follicles (12). Therefore, they are presumed to be exogenous components of EBC. In order to establish the nature of EBC keratins, we applied two approaches. First, we collected the EBC of donors who breathed with a protective filter as described above. This filter should separate inhaled air from dust as the main source of keratins. We compared the proteomes of EBC obtained with and without an anti-dust filter for three individuals. The results for one are presented in Table 5.

Sets of proteins identified in the condensates collected with and without a filter match, except for a slight discrepancy between the score value and the number of unique peptides identified. The latter is most likely due to the HPLC-MS/MS-analysis, rather than a method of collecting condensates. Therefore, our second approach investigated the proteome of ambient air. We collected and condensed the air of room. 1.5 mL and 5 mL of condensate were collected for 5 h and 20 h, respectively. The proteomes of samples are presented in Table 6.

Table 2 Identification of standard proteins after solid-phase extraction from aqueous solutions.

Number of identified peptides after ZipTipC18 (average mean $\pm$ SD, n = 4)	
Cytochrome C (equine)	4 $\pm$ 3
Apomyoglobin (equine)	Not identified
Aldolase (rabbit muscle)	4 $\pm$ 3
Albumin (bovine serum)	7 $\pm$ 3

Table 3 Identification of standard proteins after precipitation with methanol/chloroform.

Standard proteins	MW, Da	Mascot score	Peptides identified
Cytochrome C (equine)	12,360.96	424	17
Myoglobin (equine)	16,951.27	462	13
Aldolase (rabbit muscle)	39,211.28	1622	26
Albumin (bovine serum)	66,429.09	1798	29

Table 4 Occurrence of proteins in individual EBC.

Proteins	Number of donors
Keratins	
gi|17318569 Keratin 1	16
gi|47132620 Keratin 2	10
gi|18999435 Keratin 5	6
gi|5031839 Keratin 6	5
gi|55956899 Keratin 9	17
gi|40354192 Keratin 10	12
gi|12803709 Keratin 14	6
gi|24430192 Keratin 16	3
gi|4557701 Keratin 17	4
gi|114431246 Keratin 25	2
gi|908805 Keratin type II	6
Non-keratins	
gi|16751921 Dermcidin	16
gi|4502067 AMBP	9
gi|229532 Ubiquitin	6
gi|32171249 PGDS2	6
gi|1311047 Cystatin A	5
gi|29470 HSPG	2
gi|4504943 LAIR1	2
gi|30235 CSAct	1
gi|307110 LAMP2	1
gi|62898910 Kininogen 1	1
gi|3123580 Immunoglobulin light chain	1
gi|28592 Serum albumin	1

AMBP, α_1-microglobulin/bikunin precursor; PGDS2, prostaglandin H2 D-isomerase; HSPG, human basement membrane heparan sulfate proteoglycan core protein; LAIR1, leukocyte-associated immunoglobulin-like receptor 1; CSAct, cerebroside sulfate activator; LAMP2, lyosomal-associated membrane glycoprotein 2.

We used empty probes (treatment with a reaction buffer in an empty test-tube) and BSA as controls of purity and completeness of lysis with trypsin. We obtained good peptide maps for BSA. However, no peptides were found in the empty samples.

Table 5 Proteins identified in the EBC collected with and without a filter.

Proteins	Without a filter	With a filter
gi|17318569 Keratin 1	19	19
gi|40354192 Keratin 10	18	17
gi|55956899 Keratin 9	11	16
gi|47132620 Keratin 2	11	10
gi|908805 Keratin type II	2	5
gi|4557701 Keratin 17	1	1
gi|5031839 Keratin 6A	2	2
gi|15431310 Keratin 14	5	5
gi|18999435 Keratin 5	1	1
gi|1195531 Keratin 16	1	1
gi|16751921 Dermcidin preproprotein	5	5
gi|4502067 α_1-Microglobulin/bikunin precursor	3	3
gi|229532 Ubiquitin	3	2
gi|28592 Serum albumin	2	2
gi|32171249 Prostaglandin H2 D-isomerase	2	1
gi|29470 Human basement membrane heparan sulfate proteoglycan core protein	2	2
gi|307110 Lysosomal membrane glycoprotein-2	1	1

Table 6 Proteins in air condensates of the work room.

Protein identified	Peptides matched
A: 5 h (1.5 mL)	
gi|114667176 Similar to Keratin, type I cytoskeletal 14 score: 244	10
gi|17318569 Keratin 1 score: 176	8
gi|547754 Keratin, type II cytoskeletal 2 epidermal score: 136	3
B: 20 h (5 mL)	
gi|1346343 Keratin, type II cytoskeletal 1 score: 181	5
gi|122513 Hemoglobin subunit β1 score: 83	1

Discussion

We conducted a series of experiments using standard protein mixtures to highlight methodological factors that influence the mass-spectrometric identification of proteins in human EBC. We analyzed proteomes of EBC in healthy subjects. Before discussing our results it is important to note some special features of EBC and to explain our objectives. EBC is composed of more than 99% condensed water vapor (8, 13) and a very small fraction ($<1\%$) is represented by respiratory droplets. The formation of respiratory droplets is not regulated and their relative number and volume in EBC vary irrespective of the water volume. Thus, fluctuations in EBC volumes can influence the concentrations of non-volatile substances dissolved in respiratory droplets. According to the literature (13), the concentration of non-volatile analytes could be normalized by using dilutional indicators. In the present work, we have not normalized protein concentrations. We conducted a qualitative proteomic analysis and tried to recognize what proteins can be identified using mass-spectrometry in EBC collected using standardized conditions over 10 min. The period for collection of EBC was chosen as being optimal for comfort of the volunteers according to recommendations of the pulmonologists.

The main drawback is the low and variable protein concentrations in EBC. Since we cannot increase EBC volume, we tried to optimize sample preparation for proteomic analysis. First, we tried to minimize the number of sample preparation steps in order to increase protein recovery.

Comparison of different concentrating methods revealed that lyophilization was the most gentle and rapid procedure that could be standardized easily and integrated into a laboratory protocol for preparation of EBC. Lyophilization allowed us to identify at the fmol level all proteins in four-component mixture, and 77% of proteins in 48-component mixture (Table 1, Figures 1 and 2). TCA/DOC precipitation allowed us to identify a similar number of proteins (Figure 2), but it took more time and appeared to be more harmful due to usage of TCA.

Some proteins were not identified at all: insulin, epidermal growth factor, the alpha chain of hemoglobin, insulin-like growth factor II, platelet-derived growth factor B chain and ribosylhydronicotinamide dehydrogenase (NQO2). Such results can be due to two factors: incomplete hydrolysis of proteins or complicated ionization of peptides during MS-analysis of the trypsin hydrolysate. It is well-established that insulin-like growth factor II (14), like insulin, is resistant to trypsin hydrolysis under routine conditions because of covalent dimerization. Using electrophoresis and MS, insulin was found to be resistant to trypsin hydrolysis under routine conditions, and was digested only after a reduction/alkylation procedure (data not shown). In the case of NQO2, hydrolysis can be prevented by numerous carboxylic acids, and aromatic and hydrophobic residues close to cleavage sites that decrease the specificity of trypsin for protein substrate. In addition to proteins that are resistant to trypsinolysis, some proteins are expected to be lost during lyophilization leading to a decrease in the number of proteins identified in the mixtures.

When analyzing the Mascot results, we failed to discriminate between some homologues of keratin; examples include keratins K13, K15 and K10; K14 and K16; K4, K3 and K6; K7, K8 and K1. Some were identified due to their unique peptides (Table 4), while others could be neither identified nor excluded. Some of these keratins could have importance in diagnosing lung carcinoma or metastases, and in classification as suggested by the literature (15). At present, these keratins cannot be identified until their unique peptides in EBC are found.

The results we obtained lead us to speculate that there is a different level of expression of various keratins in the respiratory tract of different people and, most likely, keratin polymorphism of the epithelium. We identified only keratins 5, 6 and 14, which have been previously established as being expressed in the human airway (bronchi) (16). Thus, our results are consistent with those recently published (16).

We also found other proteins, apart from keratins, that frequently occurred in $\geq$30% of our study subjects (Table 4). Dermcidin, known to be a protein antibiotic from the sweat glands is probably an ''exogenous'' protein. Other proteins could play a diagnostic role. Some proteins that are normal participants of metabolism and appear in relatively healthy organisms may be able to change their expression level in certain pathologies. For example, cystatin A was initially characterized as an inhibitor of lysosomal cysteine proteases – cathepsins. Cathepsins are involved in processing and presentation of antigens, and are also involved in several pathological conditions such as inflammation and cancer. Cystatins can also induce synthesis of tumor necrosis factor and interleukin 10, and stimulate nitric oxide production (17).

PGDS2 participates in allergic and inflammatory reactions (18) and can be a good target for anti-allergy and anti-inflammatory drugs (19).

Finally, the AMBP secretes two separate proteins in blood. α_1-Microglobulin is found in blood and connective tissue of most organs. It is most abundant at interfaces between cells of the body and the environment, such as lungs, intestine, kidneys and placenta. α_1-Microglobulin inhibits the immunological functions of white blood cells in vitro, and its distribution is consistent with an anti-inflammatory and protective role in vivo (20). Bikunin excretion increases in inflammatory conditions and is often considered to be a positive acute phase protein. It was demonstrated that during inflammatory diseases, concentrations of serum high molecular weight bikunin products are dependent on their rate of utilization as well as on regulation of their biosynthesis (21).

Additionally, some proteins were found only in single samples and their appearance requires further investigation and interpretation.

A wide range of keratins identified in EBC allows us to conclude that these proteins occur in the respiratory tract due to exogenous sources and are probably retained by the respiratory system. The anti-dust filter did not restrict keratins identified in EBC collected in the laboratory. In addition, breathing with a filter was rather difficult and is not recommended for patients with labored breathing. Exogenous keratins can deposit continuously on the airway lining throughout human life. Exogenous proteins can be accumulated with dust indoors and inhaled by people. Thus, the keratin content of EBC is attributed both to living conditions and age of the individual. Proteomic investigations of ambient air may be an important component of proteomic studies of exhaled breath. We identified keratins in the condensates from ambient room air (Table 6). These data are also consistent with results of Hoffmann et al. (16). Such ''freely circulating'' proteins should be excluded from analysis of the proteome under investigation in order to truly establish the diagnostic significance and application of EBC analysis.

Conclusions

In this study, we attempted to optimize sample preparation of EBC for proteomic investigations. We compared different methods for protein concentration that have been used in previous studies for proteomic analysis of EBC. The results we obtained allowed us to choose the most effective procedure for sample preparation when working with EBC samples containing very low protein concentrations. The procedure we chose for proteomic analysis of EBC from healthy non-smoking volunteers proved to be rapid and effective.

Conflict of interest

The authors have declared no conflict of interest.

Acknowledgements

We thank Dr. Mike McEwan and Dr. Karsta Luettich for fruitful discussion of the manuscript. The work was supported by British American Tobacco Co. Ltd under contract 224-07.

References

1. Paredi P, Kharitonov SA, Barnes PJ. Analysis of expired air for oxidation products. Am J Respir Crit Care Med 2002;166:31–3.
2. Huszar E, Vass G, Vizi E, Csoma Z, Barát E, Molnár Világos G, et al. Adenosine in exhaled breath condensate in healthy volunteers and in patients with asthma. Eur Respir J 2002;20:1393–8.
3. Shahid SK, Kharitonov SA, Wilson NM, Bush A, Barnes PJ. Increased interleukin-4 and decreased interferon-gamma in exhaled breath condensate of children with asthma. Am J Respir Crit Care Med 2002;165:1290–3.
4. Scheideler L, Manke HG, Schwulera U, Inacker O, Hammerle H. Detection of non-volatile macromolecules in breath. A possible diagnostic tool? Am Rev Respir Dis 1993;148:778–84.
5. Wood LG, Gibson PG, Garg ML. Biomarkers of lipid peroxidation, airway inflammation and asthma. Eur Respir J 2003;21:177–86.
6. Hunt J. Exhaled breath condensate: an overview. Immunol Allergy Clin North Am 2007;27:587–96.
7. Soyer O, Dizdar E, Keskin O, Lilly C, Kalayci O. Comparison of two methods for exhaled breath condensate collection. Allergy 2006;61:1016–8.
8. Horváth I, Hunt J, Barnes P, Alving K, Antczak A, Baraldi E, et al. Exhaled breath condensate: methodological recommendations and unresolved questions. Eur Respir J 2005;26:523–48.
9. Rosias P, Robroeks C, Niemarkt H, Kester A, Vernooy J, Suykerbuyk J, et al. Breath condenser coatings affect measurement of biomarkers in exhaled breath condensate. Eur Respir J 2006;28:1036–41.
10. Brannian J, Zhao Y, Burbach J. Localization of lipid peroxidation-derived protein epitopes in the porcine corpus luteum. Biol Reprod 1997;57:1461–6.
11. Bollag D, Edelstein S. Protein methods. New York: Wiley-Liss. Inc., 1991;72–3.
12. Schweizer J, Bowden P, Coulombe P, Langbein L, Lane E, Magin T, et al. New consensus nomenclature for mammalian keratins. J Cell Biol 2006;174:169–74.
13. Effros R, Peterson B, Casaburi R, Su J, Dunning M, Torday J, et al. Epithelial lining fluid solute concentrations in chronic obstructive lung disease patients and normal subjects. J Appl Physiol 2005;99:1286–92.
14. Zumstein P, Luthi C, Humbel R. Amino acid sequence of a variant pro-form of insulin-like growth factor II. Proc Natl Acad Sci USA 1985;82:3169–72.
15. Moll R, Divo M, Langbein L. The human keratins: biology and pathology. Histochem Cell Biol 2008;129:705–33.
16. Hoffmann H, Tabaksblat L, Enghild J, Dahl R. Human skin keratins are the major proteins in exhaled breath condensate. Eur Respir J 2008;31:380–4.
17. Kopitar-Jerala N. The role of cystatins in cells of the immune system. FEBS Lett 2006;580:6295–301.
18. Psaty B, Furberg C. COX-2 inhibitors – lessons in drug safety. N Engl J Med 2005;352:1133–5.
19. Kanaoka Y, Fujimori K, Kikuno R, Sakaguchi Y, Urade Y, Hayaishi O. Structure and chromosomal localization of human and mouse genes for hematopoetic prostaglandin D synthase. Conservation of the ancestral genomic structure of sigma-class glutathione S-transferase. Eur J Biochem 2000;267:3315–22.
20. Akerström B, Lögdberg L, Berggard T, Osmark P, Lindqvist A. Alpha-1-microglobulin: a yellow-brown lipocalin. Biochim Biophys Acta 2000;1482:172–84.
21. Mizon C, Piva F, Queyrel V, Balduyck M, Hachulla E, Mizon J. Urinary bikunin determination provides insight into proteinase/proteinase inhibitor imbalance in patients with inflammatory diseases. Clin Chem Lab Med 2002;40:579–86.