

Lung alveolar proteomics of bronchoalveolar lavage from a pulmonary alveolar proteinosis patient using high-resolution FTICR mass spectrometry

Yu Bai · Dmitry Galetskiy · Eugen Damoc ·
Jan Ripper · Markus Woischnik · Matthias Griese ·
Zhiqiang Liu · Shuying Liu · Michael Przybylski

Received: 28 March 2007 / Revised: 17 May 2007 / Accepted: 25 May 2007 / Published online: 20 June 2007
© Springer-Verlag 2007

Abstract High-resolution Fourier transform ion cyclotron resonance (FTICR) mass spectrometry was developed and applied to the proteome analysis of bronchoalveolar lavage fluid (BALF) from a patient with pulmonary alveolar proteinosis. With use of 1-D and 2-D gel electrophoresis, surfactant protein A (SP-A) and other surfactant-related lung alveolar proteins were efficiently separated and identified by matrix-assisted laser desorption/ionization FTICR mass spectrometry. Low molecular mass BALF proteins were separated using a gradient 2-D gel. An efficient extraction/precipitation system was developed and used for the enrichment of surfactant proteins. The result of the BALF proteome analysis show the presence of several isoforms of SP-A, in which an N-non-glycosylated form and several proline hydroxylations were identified. Furthermore, a number of protein spots were found to contain a mixture of proteins unresolved by 2-D gel electrophoresis, illustrating the feasibility of high-resolution mass spectrometry to

provide identifications of proteins that remain unseparated in 2-D gels even upon extended pH gradients.

Keywords Pulmonary alveolar proteinosis · Lung surfactant proteins · Bronchoalveolar lavage fluid · Fourier transform ion cyclotron resonance mass spectrometry · Proteome analysis

Abbreviations

FTICR-MS	Fourier transform ion cyclotron resonance mass spectrometry
BALF	bronchoalveolar lavage fluid
PAP	pulmonary alveolar proteinosis
GM-CSF	granulocyte-macrophage colony stimulating factor

Yu Bai and Dmitry Galetskiy both contributed equally to this work.

Y. Bai · D. Galetskiy · E. Damoc · M. Przybylski (✉)
Laboratory of Analytical Chemistry and Biopolymer Structure
Analysis, Department of Chemistry, University of Constance,
Box M 731, 78457 Constance, Germany
e-mail: michael.przybylski@uni-konstanz.de

Y. Bai · Z. Liu · S. Liu
Laboratory of New Drugs, Changchun Institute of Applied
Chemistry, Chinese Academy of Sciences,
159 Remin Street,
130022 Changchun, China

J. Ripper · M. Woischnik · M. Griese
Department of Pediatric Pneumology, Dr. von Haunersches
Kinderspital, University of Munich,
Lindwurmstr. 4,
80337 Munich, Germany

Introduction

Pulmonary alveolar proteinosis (PAP) is a syndrome characterized by intra-alveolar accumulation of surfactant proteins A, B, C and D (SP-A, SP-B, SP-C and SP-D), and precursors of SP-B [1–5]. The idiopathic form in adults is caused by auto-antibodies directed against granulocyte-macrophage colony stimulating factor [6]. In neonates and children several mechanisms leading to intra-alveolar surfactant accumulation have been shown. These include mutations of the genes of SP-B [7] and SP-C [8, 9]. The subject of this study was a child with SP-C gene mutation, which is associated with combined histological patterns of nonspecific interstitial pneumonia and PAP [10].

Over the last two decades, fiberoptic bronchoscopy with bronchoalveolar lavage has become a useful method of

obtaining epithelial fluid that covers the airways and alveoli. Proteome analysis of bronchoalveolar lavage fluid (BALF) from lung disease patients, such as interstitial lung disease, asthma, chronic obstructive pulmonary disease, acute respiratory distress syndrome and PAP, has been a powerful tool for the identification of potential markers of lung diseases [11–17]. The application of high-resolution Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) in combination with gel-electrophoretic separation provided the efficient identification of proteins with medium and low abundance [18]. The accurate mass determination obtained by FTICR-MS allows the use of substantially lower tolerance thresholds during a database search, which greatly improves the selectivity of protein identification that provides only low sequence coverage [19–23]. In a recent study two nonglycosylated SP-A degradation products were identified by FTICR-MS proteome analysis [18]. The previous data suggested proteolytic degradation of both lung surfactant proteins SP-A and SP-D associated with lung diseases [24–26]; degraded SP-A had lost calcium-dependent lectin properties, i.e., loss of binding to mannose and ability to agglutinate bacteria [27].

In this study 1-D gel electrophoresis and 2-D gel electrophoresis were performed in the proteome analysis of BALF from a pulmonary proteinosis patient. We have developed an effective combination of a prepreparation step using a three-phase extraction/precipitation system for enrichment of surfactant proteins with 1-D electrophoretic separation, and subsequent protein identification by FTICR-MS. The high resolution and mass accuracy of FTICR-MS proteome analysis enabled the identification of a series of protein bands corresponding to surfactant proteins as well as serum proteins following electrophoretic separation. In particular, several SP-A forms including nonglycosylated intact SP-A were identified. Furthermore, the results of this study show that FTICR-MS provided direct, unequivocal identification of protein mixtures from gel bands that were unresolved by electrophoretic separation.

Materials and methods

Materials and preparation of BALF

Bronchoalveolar fluid was obtained using standardized protocols as previously described [10]. All samples were obtained according to approved protocols by the hospital's review board, and upon informed consent of patients and parents. Immediately after the lavage procedure, fresh BALF was centrifuged (200,000 g for 10 min). Protein concentration was determined according to the protocol of Bradford [28]. The cell-free supernatant was prepared for

electrophoresis using three different methods with all solvents and reagents of highest available purity:

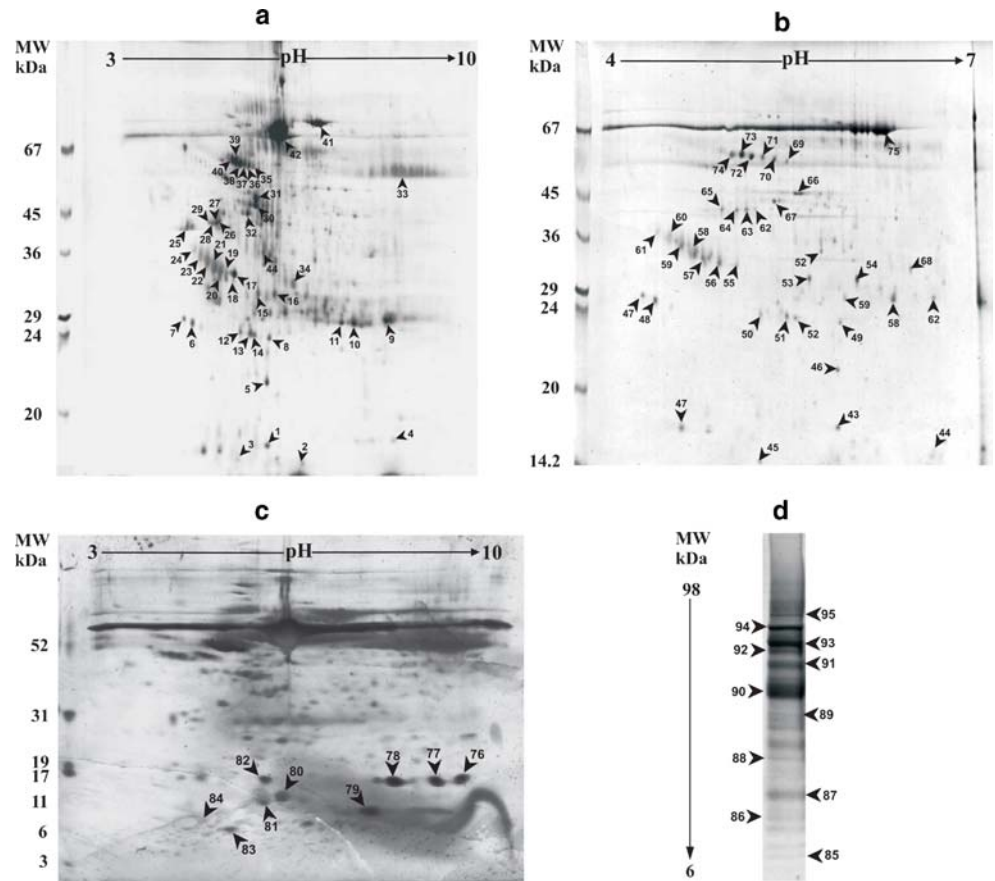
- A: BALF containing 250 μ g protein was lyophilized, dissolved in 100 μ l Milli-Q water and precipitated with a mixture of chloroform/methanol (1:4, v:v). The protein mixture was vortexed for 2 min and centrifuged for 5 min at 5,741 g, the supernatant was removed and the protein pellet was used for 2-D gel electrophoresis.
- B: BALF containing 80 μ g protein was prepared as described for method 1, desalted and concentrated with Microcon YM-3 membranes (Millipore, Bedford, USA) and finally vacuum-dried before gradient 2-D gel electrophoresis.
- C: A 20-ml BALF sample was extracted with a mixture of chloroform/methanol/10 mM HCl (2:2:1, v:v:v), vortexed for 10 min and centrifuged for 10 min at 2,756 g. The aqueous phase was extracted again with chloroform/methanol/10 mM HCl (8.6:1.4:1, v:v:v), vortexed for 10 min and centrifuged for 10 min at 2,756 g. The proteins precipitated in the interphase between the upper aqueous and the lower organic phase were collected and separated by 1-D gel electrophoresis.

Electrophoretic separation

2-D electrophoretic separation with Laemmli gel was carried out with a Multiphor horizontal electrophoresis system (Amersham Biosciences, Uppsala, Sweden) using 17-cm immobilized pH gradient (IPG) strips (pH 3–10 or 4–7 linear), with the sample applied overnight using the in-gel rehydration method. For isoelectric focusing (IEF) the sample was dissolved in rehydration solution—7 M urea, 2 M thiourea, 4% 3-[(3-cholamidopropyl)dimethylammonio]propanesulfonic acid, 0.3% dithiothreitol (DTT), 2% Servalyte 3–10 or 4–7 and trace amounts of bromophenol blue). The rehydrated strip was run in the first dimension for about 30 kVh at 20 °C. After the IEF step, the IPG strip was equilibrated for 15 min in 6 M urea, 30% glycerol, 2% w/v sodium dodecyl sulfate (SDS), 0.05 M tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl; pH 8.8), 1% w/v DTT and a trace of bromophenol blue, then for 15 min in the same solution except the DTT solution was replaced by 4.5% (w/v) iodoacetamide. The second-dimensional separations were carried out with a Bio-Rad Protean II xi vertical electrophoresis system using 12% SDS polyacrylamide gel electrophoresis (PAGE) self-made gels (1.5 mm \times 200 mm \times 200 mm). The strip placed on the vertical gel was overlaid with 1% w/v agarose in SDS running buffer (25 mM Tris-HCl, 192 mM glycine and 0.1% w/v SDS) and subjected to electrophoresis at 25 mA per gel for 30 min and 40 mA per

Fig. 1 Gel-electrophoretic separations of bronchoalveolar lavage fluid (BALF) proteins of an alveolar proteinosis patient. Protein spots were excised from the gel, digested with trypsin and identified by matrix-assisted laser desorption/ionization (MALDI) Fourier transform ion cyclotron resonance (FTICR) mass spectrometry (MS) (see “Materials and methods”). The gel spots labeled with numbers correspond to the proteins identified by FTICR-MS (Tables 1–3).

a 2-D gel electrophoresis of 250 μ g BALF protein, pH range 3–10, visualized by Coomassie blue staining. **b** 2-D gel electrophoresis of 250 μ g BALF protein, pH range 4–7, visualized by Coomassie blue staining. **c** Separation of low molecular mass proteins by 2-D gel electrophoresis of 80 μ g BALF proteins using gradient 12–14% gel, visualized by silver staining. **d** Surfactant-enriched protein fraction of 20 ml BALF separated by 1-D gel electrophoresis, visualized by Coomassie blue staining



gel until the tracking dye reached the anodic end of the gel. Proteins were visualized by high-sensitive colloidal Coomassie blue [29] and were scanned using a Bio-Rad GS-710 calibrated imaging densitometer.

2-D electrophoretic separation with a gradient [bis(2-hydroxyethyl)amino]tris(hydroxymethyl)methane (Bis-Tris) gel was carried out on IPG strips, pH 3–10. The second-dimensional separation was performed on a gradient of 12–14% Bis-Tris gel, and proteins were visualized by silver staining [30].

1-D gel electrophoresis was carried out with a NuPAGE Novex Bis-Tris gel system from Invitrogen, in particular the 10% Bis-Tris gel (1.0 mm, ten wells; cat. no. NP0301BOX) with 2-morpholinoethanesulfonic acid running buffer (20 \times concentrate; cat. no. NP0002), lithium dodecyl sulfate (LDS) sample buffer (4 \times concentrate; cat. no. NP0008), and sample reducing agent (10 \times concentrate; cat. no. NP0009). The sample was lyophilized, resolubilized in 25 μ l 1 \times sample buffer (65% H₂O, 25% LDS sample buffer, 10% reducing agent), and heated for 10 min at 70 $^{\circ}$ C. Samples were then applied to the wells in the gel. The gel was run in a NuPAGE XCell minicell chamber at a constant voltage of 200 V for 35–40 min, after which the gels were taken out and stained with colloidal Coomassie blue.

In-gel digestion and peptide extraction

Gel bands were excised accurately to minimize excess gel material into 1-mm cubes, dehydrated with acetonitrile, and dried in a Speed Vac centrifuge. For Coomassie blue stained gel, the gel pieces were destained using acetonitrile/H₂O (3:2) and 50 mM ammonium bicarbonate; for silver-stained gel, gel pieces were destained in 30 mM K₃[Fe(CN)₆]/100 mM Na₂S₂O₃ (1:1, v:v) [31] for 10 min and washed with Milli-Q water. The gel pieces were then dehydrated with acetonitrile and dried in vacuo for approximately 15 min. Then 40 μ l trypsin solution (12.5 ng μ l⁻¹ trypsin in 50 mM NH₄HCO₃) was added and incubated for 45 min at 4 $^{\circ}$ C. After pulling off the reaction solution, 10 μ l of buffer without protease was added and the reaction continued for 20 h at 37 $^{\circ}$ C. The solution was then lyophilized and desalted using C18-ZipTip (Millipore, USA).

Matrix-assisted laser desorption/ionization–FTICR-MS

Mass-spectrometric analysis was performed with a Bruker APEX II FTICR instrument equipped with an actively shielded 7 T superconducting magnet, a cylindrical

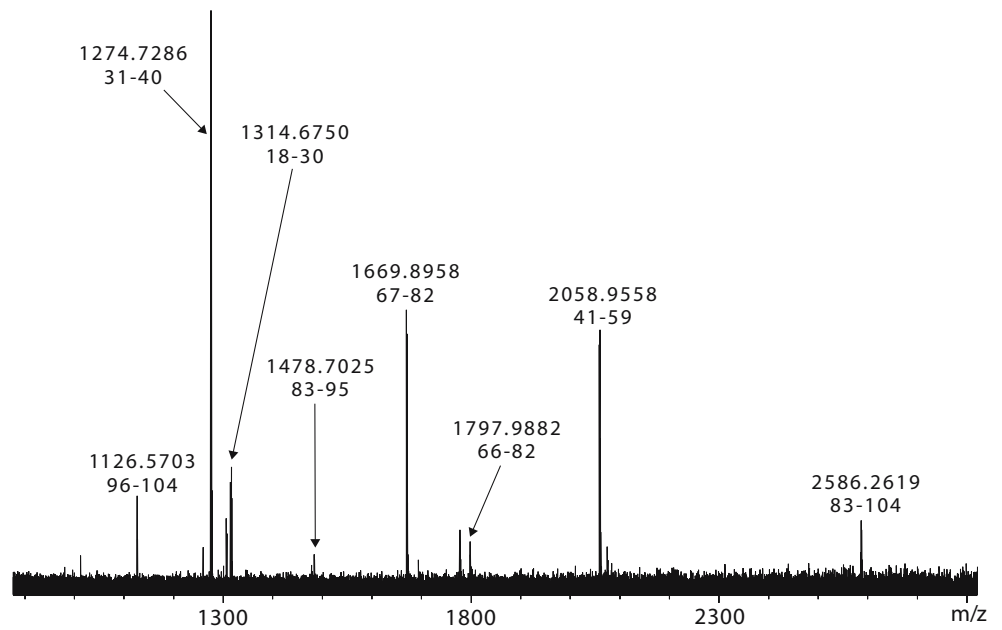
Table 1 Protein isoforms from bronchoalveolar lavage fluid (BALF) of a patient with pulmonary alveolar proteinosis (PAP) identified by matrix-assisted laser desorption/ionization (MALDI) Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) mass fingerprinting

Protein	Spot no.	$M_{r \text{ gel}}$ (kDa)	pI_{gel}	Sequence coverage (%)	Accession no.
Transthyretin	1, 43	17	5.8	66	P02766
TAP2 (fragment)	2, 44	14	6.3	22	Q03519
Galectin-1	3, 45	14	5.1	48	P09382
Cyclophilin	4	18	8.0	42	Q71V99
Ferritin light chain	5, 46	21	5.4	37	P02792
Immunoglobulin J chain	6, 47	28	4.1	53	P01591
Immunoglobulin J chain	7, 48	28	4.0	53	P01591
Glutathione S-transferase A1	8, 49	24	5.5	26	P08263
IG λ light chain variable region	9	28	7.8	46	Q96SB0
IG λ light chain variable region	10	28	7.2	46	Q96SB0
Triosephosphate isomerase	11	28	6.9	38	P60174
Serum albumin (fragment)	11	28	6.9	13	P02768
Serum albumin (fragment)	11	28	6.9	11	P02768
γ -Actin (fragment)	12, 50	25	5.1	30	P63261
γ -Actin (fragment)	13, 51	25	5.4	49	P63261
Apolipoprotein A-1	14, 52	24	5.5	37	P02647
Glutathione S-transferase A1	14, 52	24	5.5	26	P08263
Cathepsin D heavy chain	15, 53	26	5.3	62	P07858
Cathepsin D heavy chain	16, 54	26	5.3	36	P07858
Annexin A-5	17	34	4.9	24	P08758
Surfactant protein A	18, 55	32	5.0	47	P07714
Surfactant protein A	19, 56	33	4.9	47	P07714
Surfactant protein A	20, 57	34	4.8	47	P07714
Surfactant protein A	21, 58	35	4.7	47	P07714
Surfactant protein A	22, 59	36	4.6	47	P07714
Surfactant protein A	23, 60	37	4.5	47	P07714
Surfactant protein A	24, 61	38	4.3	47	P07714
Pepsinogen C	25	41	4.0	9	P20142
Vimentin	26, 62	43	4.7	32	P08670
Vimentin	27, 63	44	4.6	32	P08670
Vimentin	28, 64	44	4.5	32	P08670
Vimentin	29, 65	44	4.4	32	P08670
β -Actin	30, 66	48	5.4	39	P60709
β -Actin	31, 67	49	5.4	18	P60709
β -Actin	32	46	5.2	39	P60709
IG γ -1 chain C region	33	36	8.5	34	P01857
Serum albumin (fragment)	34, 68	33	6.2	5	P02768
α -1-Antitrypsin	35, 69	50.5	5.4	40	P01009
α -1-Antitrypsin	36, 70	51	5.3	40	P01009
α -1-Antitrypsin	37, 71	51.5	5.2	40	P01009
α -1-Antitrypsin	38, 72	52	5.1	40	P01009
α -1-Antitrypsin	39, 73	52.5	5.0	40	P01009
α -1-Antitrypsin	40, 74	53	4.9	40	P01009
Transferrin	41	78	6.7	38	P02787
Serum albumin	42, 75	72	5.8	32	P02768

infinity ion cyclotron resonance analyzer cell, and an external Scout 100 fully automated X-Y target stage matrix-assisted laser desorption/ionization (MALDI) source with pulsed collision gas (Bruker Daltonik, Bremen, Germany). The pulsed nitrogen laser was operated at 337 nm, and ions were directly desorbed into a hexapole ion guide situated 1 mm from the laser target [20]. The

device for pulsing collision gas in direct proximity to the laser target provides cooling of the ions, which have a kinetic energy spread of several electronvolts when produced by the MALDI process. These ions are trapped in the hexapole, where positive potentials at the laser target and at an extraction plate help trap ions along the longitudinal axis. After a predefined trapping time, the

Fig. 2 MALDI-FTICR-MS identification of hemoglobin β chain, spot 78 (Fig. 3; pI 7.8, M_r 16 kDa). Eight peptide masses were matched by a database search



voltage of the extraction plate is reversed and the trapped ions are extracted for transmission to the ion cyclotron resonance cell. Accumulation of ions from multiple laser shots in the hexapole before mass-spectrometric analysis increases sensitivity [20, 23]. Ions generated by 20 laser shots were accumulated in the hexapole for 0.5–1 s at 30 V and extracted at –15 V into the analyzer cell. The ion trapping potential was 1.0–1.2 V. Mass spectra were acquired in the mass range from m/z 500 to m/z 3,500 and processed using sine bell apodization. The mass spectrometer was calibrated externally within an m/z range between 500 and 4,000 using a standard peptide mixture. A 50 mg ml⁻¹ solution of 2,5-dihydroxybenzoic acid (Aldrich, Germany) in acetonitrile/0.1 % aqueous trifluoroacetic acid (2:1) was used as the matrix. A mixture of 0.5 μ l matrix

solution and 0.5 μ l sample solution was prepared, applied on the stainless steel target and allowed to dry at room temperature.

Database search

Monoisotopic masses of all singly charged ions from the MALDI-FTICR mass spectra were directly used for database search procedures using Mascot peptide mass fingerprinting search engine [32] or ProFound [33]. The database employed was NCBIInr, a compilation of several databases, including Swiss-Prot, PIR, PRF, PDB and GenBank CDS translation. Criteria for protein identification were three or more proteolytic peptides identified with mass accuracies of less than 8 ppm.

Table 2 Low molecular mass BALF proteins of a PAP patient separated on a 12–14% gradient 2-D gel and identified by FTICR-MS

Protein	Spot no.	$M_{r, \text{gel}}$ (kDa)	pI _{gel}	Sequence coverage (%)	Accession no.
Hemoglobin α chain	76	17	9	49	P69905
Profilin-1	76	17	9	31	P07737
Hemoglobin β chain	76	17	9	29	P68871
Hemoglobin α chain	77	16	8.6	49	P69905
Hemoglobin β chain	78	16	7.8	55	P68871
Calgranulin A	79	11	7.5	55	P05109
Calgranulin B	80	14	6.1	38	P06702
Calgranulin B	81	13	5.7	36	P06702
Transthyretin	82	17	5.8	24	P02766
Serum albumin (fragment)	83	7	5.2	10	P02768
Clara cell phospholipid binding protein	84	8	4.7	42	P11684

Table 3 BALF proteins in a surfactant-enriched fraction of from PAP patient identified by MALDI-FTICR-MS after 1-D electrophoretic separation

Protein	Band no.	$M_{r, \text{gel}}$ (kDa)	Sequence coverage (%)	Accession no.
Ubiquitin	85	6	36	P62988
Calgranulin A	86	12	34	P05109
Calgranulin B	86	12	68	P06702
α -1-Acid glycoprotein-1 (fragment)	86	12	15	P02763
Calgranulin B	87	16	69	P06702
Hemoglobin β chain	87	16	38	P68871
Calcyphosin	88	21	27	Q13938
Surfactant protein A	89	27	36	P07714
Surfactant protein A	90	35	47	P07714
α -1-Acid glycoprotein-1	91	41	39	P02763
α -1-Antitrypsin	92	44	40	P01009
α -1-Antitrypsin	93	50	40	P01009
α -1-Antitrypsin	94	69	32	P01009
Serum albumin	94	69	24	P02768
Galectin-3-binding protein	95	80	16	Q08380

Results and discussion

Gel-electrophoretic separation of lung alveolar proteins from a pulmonary proteinosis patient

BALF containing 250 μg protein (see “Materials and preparation of BALF,” method 1) was reduced, alkylated and separated by 2-D PAGE using linear pH gradient 3–10 strips, and visualized by Coomassie blue staining (Fig. 1a). In order to increase the resolution of 2-D separations, an IPG strip with pH gradient 4–7 was selected as a second 2-D gel with the same sample (Fig. 1b). The resolution was clearly increased by using the narrow-range IPG strip. The protein

spots were excised, destained and digested with trypsin, and the proteolytic peptide mixtures were analyzed by MALDI-FTICR-MS. Monoisotopic masses of singly charged ions from the MALDI-MS data were subjected to a database search using Mascot peptide fingerprinting and the ProFound search engines. With the MALDI-FTICR-MS procedure employed, mass determination accuracies were approximately 5 ppm at a mass resolution of 100,000. The data of the tryptic peptides identified by FTICR-MS are summarized in Table 1.

Compared with constant-percentage gels, gradient separation gels have become an efficient approach for separation of a wide range of proteins with different molecular masses. Low molecular mass proteins are often

Fig. 3 MALDI-FTICR mass spectrum of an isoform of surfactant protein A (spot 20 in Fig. 1a; pI 4.8, M_r 34 kDa). Peptides containing one, two and four hydroxyprolines are labeled

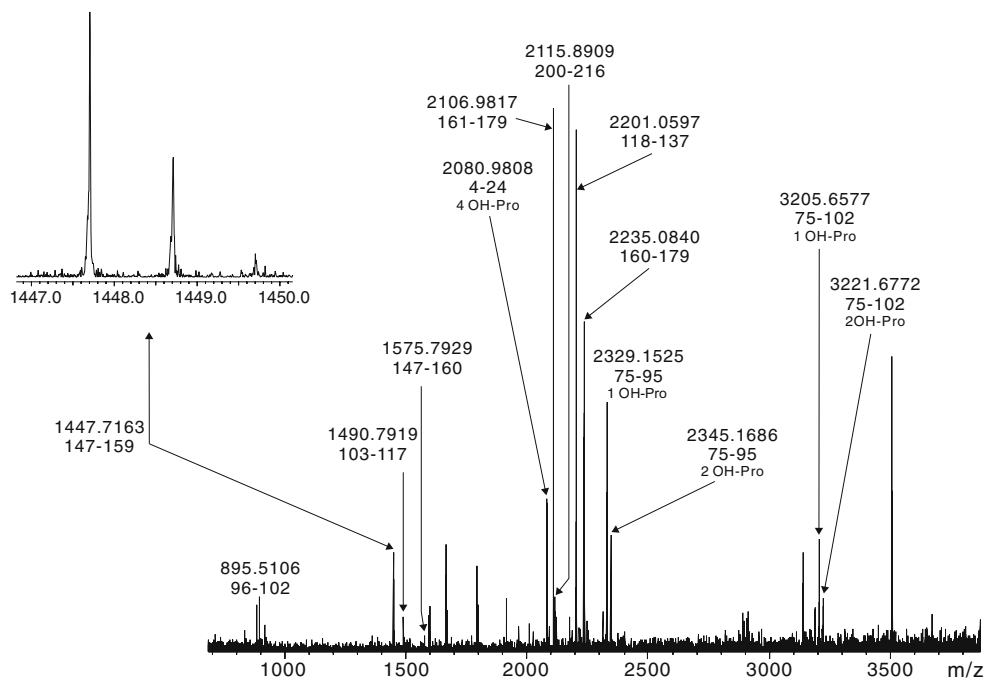


Table 4 Modifications of surfactant protein A identified by FTICR-MS

[M+H] ⁺	Sequence	Tryptic peptide	Modifications
2,007.96	4–24	DVCVGSPIGTPGSHGLPGR	Cys6 reduced ^a , 3 OH-Pro
2,066.96	4–24	DVCVGSPIGTPGSHGLPGR	Cys6 alkylated ^b , 3 OH-Pro
2,023.95	4–24	DVCVGSPIGTPGSHGLPGR	Cys6 reduced ^a , 4 OH-Pro
2,080.97	4–24	DVCVGSPIGTPGSHGLPGR	Cys6 alkylated ^b , 4 OH-Pro
2,329.16	75–95	GPPGLPAHLDEELQATLHDFR	1 OH-Pro
2,345.15	75–95	GPPGLPAHLDEELQATLHDFR	2 OH-Pro
3,205.65	75–102	GPPGLPAHLDEELQATLHDFRHQILQTR	1 OH-Pro
3,221.65	75–102	GPPGLPAHLDEELQATLHDFRHQILQTR	2 OH-Pro
895.51	96–102	HQILQTR	
1,490.79	103–117	GALSLQGSIMTVGEK	
2,144.01	118–137	VFSSNGQSITFDAIQEACAR	Cys135 reduced ^a
2,201.05	118–137	VFSSNGQSITFDAIQEACAR	Cys135 alkylated ^b
1,447.71	147–159	NPENEAIASFVK	
1,575.80	147–160	NPENEAIASFVKK	
2,235.07	160–179	KYNTYAYVGLTEGSPGDFR	
2,106.98	161–179	YNTYAYVGLTEGSPGDFR	
1,735.78	180–193	YSDGTPVNYTNWYR	Nonglycosylated Asn187 ^c
1,736.76	180–193	YSDGTPVDYTNWYR	Deglycosylated Asn187 ^d
2,115.89	200–216	GKEQCVEMYTDGQWNR	Cys204 alkylated ^b

^a Reduced with dithiothreitol

^b Reduced with dithiothreitol and alkylated with iodoacetamide

^c Present only in 27-kDa form

^d Present after treatment with PNGase F

difficult to separate by 2-D gel electrophoresis. In order to separate the proteins from a PAP patient in the low molecular mass region, the sample was prepared using a Microcon YM-3 membrane (see “[Materials and preparation of BALF](#),” method 2) and gradient 12–14% 2-D electrophoretic separation was performed. Desalting and concentration on the membrane due to 2-D electrophoretic separation applied in our previous work [18] showed minimal loss of protein and good gel performance using 5–100 µg protein. BALF containing 80 µg protein was reduced, alkylated, separated by 2-D gel electrophoresis and visualized by silver staining (Fig. 1c). The marked protein spots (Fig. 1c) were destained, digested with trypsin, and analyzed by FTICR-MS. Figure 2 shows the FTICR mass spectrum and the ProFound search result for protein spot 78 (pI 7.8, M_r 16 kDa). The identification was obtained by eight peptides with a mass accuracy of less than 5 ppm. Hemoglobin α chain, hemoglobin β chain, profilin-1, calgranulin A, calgranulin B, transthyretin, serum albumin fragment and Clara cell phospholipid binding protein were efficiently separated by gradient 2-D gel electrophoresis and unambiguously identified by FTICR-MS (Table 2).

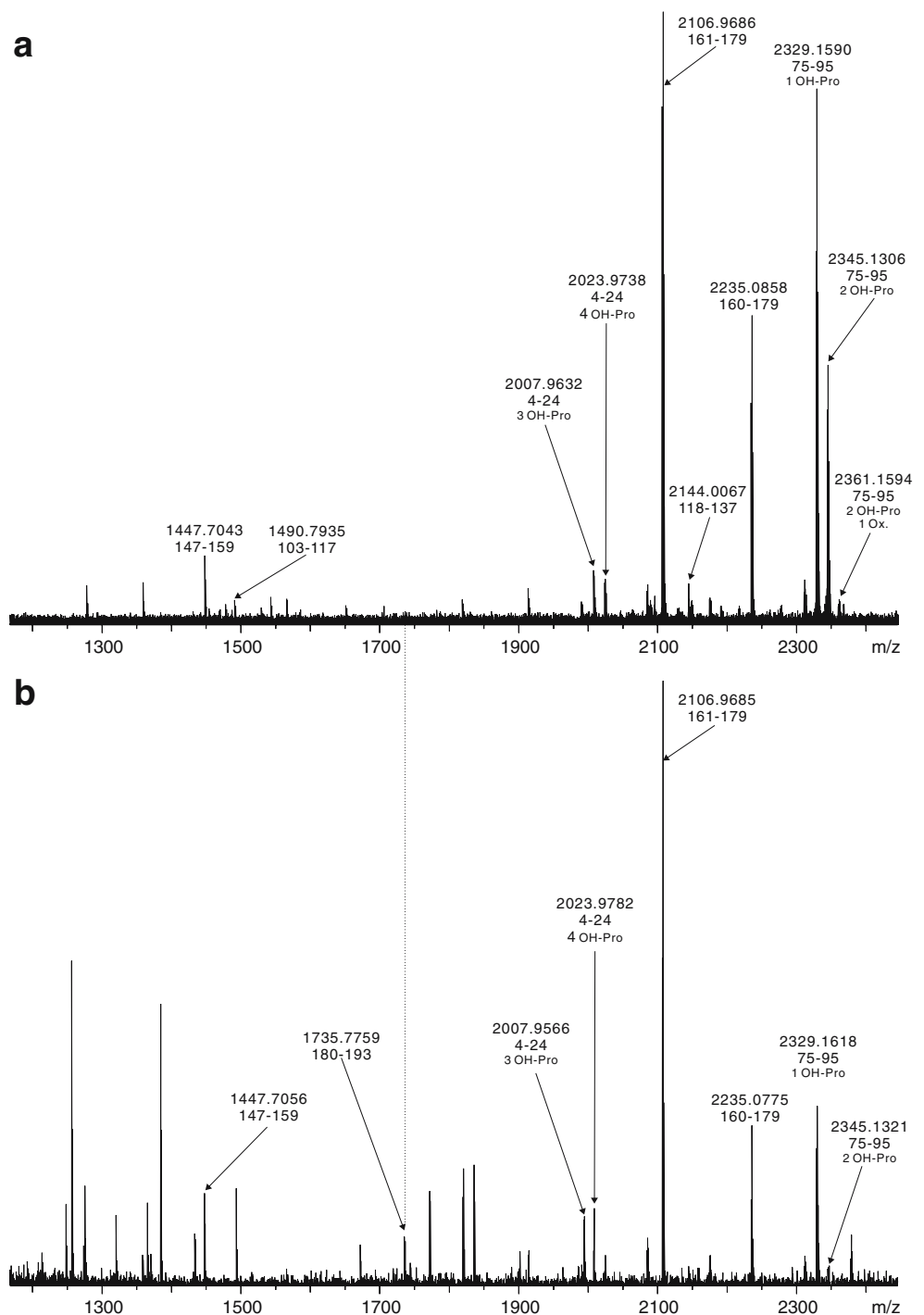
For the isolation of surfactant proteins, a three-phase extraction/precipitation system was developed and used for BALF from PAP patients (see “[Materials and preparation of BALF](#),” method 3). Surfactant proteins precipitate partially

under these conditions, while the hydrophilic proteins dissolved in the aqueous phase and the hydrophobic proteins dissolved in the organic phase. This method was very effective for the isolation of low-abundance lipophilic proteins with intermediate hydrophobicity not soluble in the aqueous phase after dissolution of the lung surfactant lipid-protein complex and removing of phospholipids and not soluble in the organic phase, such as nonglycosylated SP-A, which could not be directly identified after 2-D gel electrophoresis. Following extraction, these proteins were separated on a 10% Bis-Tris gel. The protein bands (Fig. 1d) were destained, digested and analyzed by FTICR-MS. The proteins identified by MALDI-FTICR-MS are summarized in Table 3.

Characterization of structural modifications in SP-A

Seven isoforms of intact SP-A in a mass range from 32 to 38 kDa were identified from 2-D gels at gradient pH 3–10 and 4–7. Several modifications by hydroxyproline were directly identified. Figure 3 shows the MALDI-FTICR mass spectrum and identification of SP-A from spot 20 (pI 4.8, M_r 34 kDa). Thirteen peptides matched SP-A in the database search with accuracies of the measured peptide masses identified as SP-A fragments of less than 5 ppm. The ions at m/z 2,329.1525 and m/z 2,345.1686 were identified as a sequence (75–95) corresponding to the

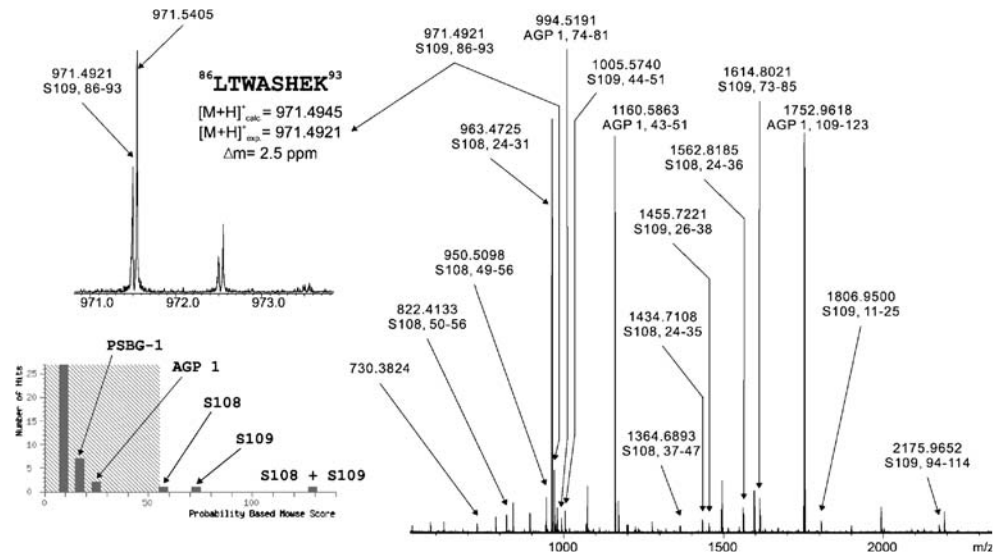
Fig. 4 MALDI-FTICR-MS of two isoforms of surfactant protein A from a pulmonary alveolar proteinosis patient after separation of surfactant-enriched BALF protein fraction by 1-D gel electrophoresis. **a** Glycosylated intact surfactant protein A identified in the 35-kDa protein band (band 90 in Fig. 1c). **b** Nonglycosylated surfactant protein A identified in the 27-kDa protein (band 89 in Fig. 1c). The tryptic peptide fragment at m/z 1,735.7759 corresponds to the nonglycosylated sequence (180–193), containing Asp187



collagen-like domain of SP-A containing one and two hydroxyproline residues, in agreement with previous report showed that Pro77 was fully hydroxylated, while Pro80 was partially hydroxylated [34]. These identifications were confirmed by the ions at m/z 3,205.6577 and m/z 3,221.6772, due to the sequence (75–102) containing hydroxylated Pro77 and Pro80. Furthermore, the ion at m/z 2080.9808 was due to the peptide (4–24) in which all four proline residues at positions 10, 13, 16 and 22 were found

hydroxylated. The ions at m/z 895.5106, m/z 1,490.7919 and m/z 2,201.0597 were identified as peptides (96–102), (103–117) and (118–137), respectively, corresponding to the triple coiled coil neck region of SP-A. Peptides at m/z 1,447.7163, m/z 1,575.7929, m/z 2,106.9817, m/z 2,235.0840 and m/z 2,115.8909 covered the sequences (147–159), (147–160), (161–179), (160–179) and (200–216), respectively, of the carbohydrate recognition domain. The comparison of the sequence of known gene products of

Fig. 5 MALDI-FTICR mass spectrum and Mascot identification of calgranulin A (S108) calgranulin B (S109) and α -1-acid glycoprotein-1 (AGP 1) in the 12-kDa band (band 86 in Fig. 1c). An additional two peptides (m/z 730.3824 and m/z 971.5405) are assigned to pregnancy-specific β -1-glycoprotein (PSBG-1)

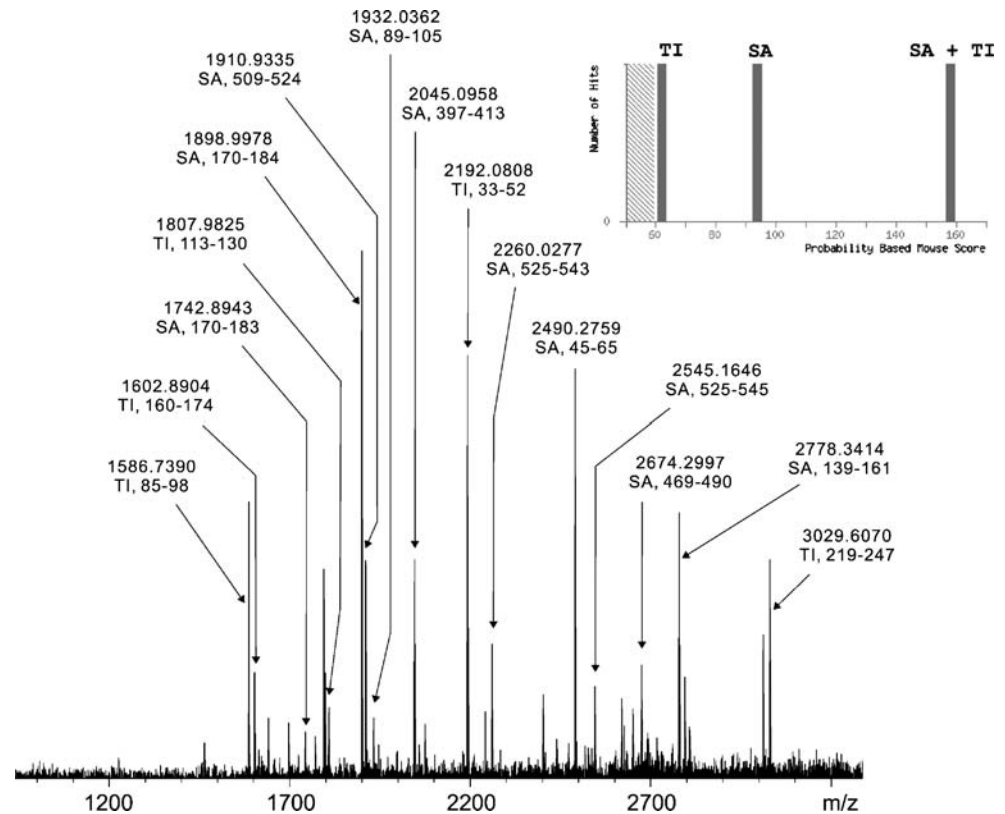


SP-A within the sequence (200–216) revealed the presence of Arg199, thus providing a cleavage site by trypsin owing to the single nucleotide polymorphism Arg199Trp. This mutation is an important disease marker. Trp199 was reported in idiopathic pulmonary fibrosis patients [35]. A change from the large and basic arginine to the large and aromatic tryptophan alters protein behaviour owing to truncation by oxidation and may serve as a marker to identify subgroups of patients of risk. The absence of this specific mutation in the patient can exclude SP-A degradation

caused by Arg199Trp mutation. Tryptic fragments of SP-A characterized by PAP patients are summarized in Table 4.

We have shown previously [18] that SP-A from a BALF sample from a patient was N-glycosylated at Asn187. The heterogeneity of the SP-A isoforms observed by 2-D gel electrophoresis was explained by microheterogeneity of glycan structures. The peptide fragment at m/z 1,736.77 appearing after treatment with PNGase F was assigned to the Asn187-deamidated peptide 180 YSDGTPVDYTN-WYR 193 ; and with use of FTICR-MS, this peptide was

Fig. 6 MALDI-FTICR-MS identification of serum albumin (SA) and triosephosphate isomerase (TI) after separation by 2-D gel electrophoresis (spot 11 in Fig. 1a). Identified sequence parts of albumin in the gel spot at pI 6.9, M_r 28 kDa were assigned to two different fragments (Table 1): N-terminal fragment (residues 45–184) and C-terminal fragment (residues 387–545)



clearly resolved from the nonglycosylated peptide at m/z 1,735.77 [36]. The capability of FTICR-MS to differentiate between nonglycosylated and deamidated sites [22, 37–40] was used to identify nonglycosylated forms of intact SP-A in PAP patient samples. Intact SP-A was identified in the most intense band (M_r 35 kDa) after 1-D electrophoretic separation of the surfactant-enriched BALF protein fraction, illustrating the three-phase extraction/precipitation for the isolation of SP-A. Figure 4 shows the FTICR mass spectrum of two forms of SP-A (M_r 35 and 27 kDa) after reduction and in-gel tryptic digestion. In addition to the expected peptides, a new peak in the second form at m/z 1,735.7759 corresponded to the nonglycosylated peptide fragment $^{180}\text{YSDGTPVNYTNWYR}^{193}$ ($\Delta m=3$ ppm) (Fig. 4b). Two nonglycosylated proteolytic fragments were found previously in this patient [18]. The presence of nonglycosylated intact SP-A and SP-A fragments suggests that glycosylation of SP-A is an important feature for the stability of protein, and thus non-glycosylated protein may be proteolytically degraded and removed from the lung surface. Additionally, FTICR-MS data clearly show that truncated SP-A found in the patient is a result of the lack in protein glycosylation and is not caused by degradation of an active intact SP-A.

Identification of protein mixtures from gel bands using high-resolution FTICR-MS

Owing to its high resolution and mass accuracy, FTICR-MS provides direct identification of proteins in complex biological mixtures, such as in tryptic digests from unseparated gel bands. Mass spectra of up to 128 scans were averaged, and in-source accumulations of the ions from up to 20 laser shots were used under collisional cooling of ions to increase the sensitivity without significant loss of mass accuracy. This procedure was found to be effective for the analysis of low-abundance proteins and provided the direct identification of protein mixtures. Several protein mixtures were unambiguously identified in this manner from unresolved gel bands (Tables 1–3).

As an example, the 12-kDa gel band (Fig. 1d, band 86) was excised, destained and digested with trypsin. The tryptic peptides were extracted and identified by FTICR-MS using the monoisotopic peaks for the Mascot database engine at a mass accuracy of 8 ppm (Fig. 5). Calgranulins A and B were identified in a spot by six peptides for calgranulin A (mass accuracy 4 ppm) by and six peptides for calgranulin B (mass accuracy 2 ppm). In addition, three peptides (m/z 994.5191, m/z 1,160.5863 and m/z 1,752.9618; mass accuracy 2 ppm) matched the N-terminal fragment of α -1-acid glycoprotein-1; two other peptides (m/z 730.3824 and m/z 971.5405; mass accuracy 1 ppm) were assigned to pregnancy-specific β -1-glycoprotein. The identification of the fragment of pregnan-

cy-specific β -1-glycoprotein is not included in Table 3 because of the identification criteria (see “Materials and methods”). At the mass resolution (approximately 113,000) the two adjacent peaks with a mass difference of 50 ppm could be separated, one of which (m/z 971.4921) was identified as a tryptic fragment of calgranulin B (residues 86–93). High-resolution mass-spectrometric data showing the presence of small and highly modified proteins and protein fragments in the complex mixture are essential for characterization of lung diseases. In particular, calgranulins are important disease markers. High levels of calgranulin A were observed in lung diseases such as sarcoidosis, idiopathic pulmonary fibrosis and hypersensitivity pneumonitis [11]; hence, its specific identification is of particularly importance in disease samples. In a second protein band (spot 87) a mixture of calgranulin B and hemoglobin β chain was identified, while protein band 101 was identified as a mixture of serum albumin and α -1-antitrypsin.

The resolution of a 2-D electrophoresis gel is frequently limited when wide pH range IPG strips are used. Even with narrow pH range IPG strips, proteins with similar physical-chemical properties may be difficult to separate. Therefore, the application of FTICR-MS, with high resolution and high mass accuracy, greatly facilitates the unequivocal identification of protein mixtures. A further example for protein mixture identification using FTICR-MS is protein spot 11 in Fig. 1a (Fig. 6). This spot was identified as a mixture of serum albumin and triosephosphate isomerase. Five peaks matched to triosephosphate isomerase (mass accuracy 4 ppm), while ten additional peaks matched serum albumin (mass accuracy 3 ppm). Since the molecular mass of this spot is 28 kDa, identified sequence parts of albumin (residues 45–545, M_r 57 kDa) were assigned to two different fragments (Table 1): N-terminal fragment (residues 45–184) and C-terminal fragment (residues 387–545). Furthermore, protein spots 14 (Fig. 1a) and 52 (Fig. 1b) were identified to contain a mixture of apolipoprotein A-1 and glutathione *S*-transferase A; and three proteins were identified in spot 76 (Fig. 1c): hemoglobin α chain, profilin-1 and hemoglobin β chain.

Conclusions

BALF, which contains a complex mixture of proteins, was analyzed by gel-electrophoretic methods followed by subsequent proteome analysis using high-resolution FTICR-MS combined with a database search to identify a series of specific proteins. With use of 2-D gel electrophoresis at pH 3–10 and pH 4–7 SP-A and other surfactant-related lung alveolar proteins were identified, including several structure modifications. Several hydroxyproline modifications in the collagen-like domain of SP-A were

directly identified. In most cases, the accurate mass determination allowed the use of low mass tolerance thresholds for database searches, and enabled the identification of proteins with a minimum number of peptides. With use of 12–14% gradient 2-D gels, the low molecular mass proteins were separated and identified by FTICR-MS. A three-phase extraction/precipitation surfactant-enrichment system was shown in combination with 1-D gel electrophoresis to be an effective method to separate and identify surfactant proteins. Because of its high resolution and mass accuracy, FTICR-MS provided rapid, direct and unambiguous identification of proteins and protein mixtures from 2-D gel bands, thus presenting an efficient approach for proteome analysis of biological samples that are difficult to separate or unfeasible for separation from biological material.

Acknowledgements Y.B. acknowledges a scholarship from the Max Planck Society. We gratefully acknowledge the expert assistance of Nikolay Youhnovski with FTICR-MS procedures. This work was supported by grants from the Deutsche Forschungsgemeinschaft (Alveolar Proteomics, to M.G. and M.P.; and Biopolymer-MS, to M.P.) and the Fonds der Chemischen Industrie.

References

- Seymour JF, Presneill JJ (2002) *Am J Respir Crit Care Med* 166:215
- Mason RJ, Greene K, Voelker DR (1998) *Am J Physiol* 275:L1
- Shah PL, Hansell D, Lawson PR, Reid KB, Morgan C (2000) *Thorax* 55:67
- Maygarden SJ, Iacocca MV, Funkhouser WK, Novotny DB (2001) *Diagn Cytopathol* 24:389
- Wang BM, Stern EJ, Schmidt RA, Pierson DJ (1997) *Chest* 111:460
- Kitamura T, Tanaka N, Watanabe J, Uchida, Kanegasaki S, Yamada Y, Nakata K (1999) *J Exp Med* 190:875
- Nogee LM, Garnier G, Dietz HC, Singer L, Murphy AM, deMello DE, Colten HR (1994) *J Clin Invest* 93:1860
- Nogee LM, Dunbar AE 3rd, Wert S, Askin F, Hamvas A, Whitsett JA (2002) *Chest* 121:20S
- Nogee LM, Dunbar AE 3rd, Wert SE, Askin F, Hamvas A, Whitsett JA (2001) *N Engl J Med* 344:573
- Brasch F, Griese M, Tredano M, Johnen G, Ochs M, Rieger C, Mulugeta S, Muller KM, Bahuau M, Beers MF (2004) *Eur Respir J* 24:30
- Wattiez R, Hermans C, Cruyt C, Bernard A, Falmagne P (2000) *Electrophoresis* 21:2703
- Wattiez R, Hermans C, Bernard A, Lesur O, Falmagne P (1999) *Electrophoresis* 20:1634
- Wattiez R, Falmagne P (2005) *J Chromatogr B* 815:169
- Plymoth A, Lofdahl CG, Ekberg-Jansson A, Dahlback M, Lindberg H, Fehniger TE, Marko-Varga G (2003) *Proteomics* 3:962
- Bowler RP, Duda B, Chan ED, Enghild JJ, Ware LB, Matthay MA, Duncan MW (2004) *Am J Physiol Lung Cell Mol Physiol* 286:L1095
- Noel-Georis I, Bernard A, Falmagne P, Wattiez R (2001) *Dis Markers* 17:271
- Signor L, Tigani B, Beckmann N, Falchetto R, Stoeckli M (2004) *Proteomics* 4:2101
- Bai Y, Galetskiy D, Damoc E, Paschen C, Liu Z, Griese M, Liu S, Przybylski M (2004) *Proteomics* 4:2300
- Damoc E, Youhnovski N, Crettaz D, Tissot JD, Przybylski M (2003) *Proteomics* 3:1425
- Baykut G, Jertz R, Witt M (2000) *Rapid Commun Mass Spectrom* 14:1238
- Conrads TP, Anderson GA, Veenstra TD, Pasa-Tolic L, Smith RD (2000) *Anal Chem* 72:3349
- Albach C, Damoc E, Denzinger T, Schachner M, Przybylski M, Schmitz B (2004) *Anal Bioanal Chem* 378:1129
- Witt M, Fuchser J, Baykut G (2003) *J Am Soc Mass Spectrom* 14:553
- von Bredow C, Birrer P, Griese M (2001) *Eur Respir J* 17:716
- von Bredow C, Wiesener A, Griese M (2003) *Lung* 181:79
- Griese M, von Bredow C, Birrer P (2001) *Electrophoresis* 22:165
- Griese M, Wiesener A, Lottspeich F, von Bredow C (2003) *Biochim Biophys Acta* 1638:157
- Bradford MM (1976) *Anal Biochem* 72:248
- Neuhoff V, Arold N, Taube D, Ehrhardt W (1988) *Electrophoresis* 9:255
- Heukeshoven J, Dernick R (1985) *Electrophoresis* 6:103
- Gharahdaghi F, Weinberg CR, Meagher DA, Imai BS, Mische SM (1999) *Electrophoresis* 20:601
- Mascot (2007) Matrix Science, London. http://www.matrixscience.com/search_form_select.html
- ProFound (2007) The Rockefeller University, New York. <http://prowl.rockefeller.edu/prowl/cgi/profound.exe>
- Berg T, Leth-Larsen R, Holmskov U, Hojrup P (2000) *Biochim Biophys Acta* 1543:159
- Selman M, Lin HM, Montano M, Jenkins AL, Estrada A, Lin Z, Wang G, DiAngelo SL, Guo X, Umstead TM, Lang CM, Pardo A, Phelps DS, Floros J (2003) *Hum Genet* 113:542
- Küster B, Wheeler SF, Hunter AP, Dwek RA, Harvey DJ (1997) *Anal Biochem* 250:82
- Otto VI, Damoc E, Cueni LN, Schurpf T, Frei R, Ali S, Callewaert N, Moise A, Leary JA, Folkers G, Przybylski M (2006) *Glycobiology* 16:1033
- Murray S, Nilsson CL, Hare JT, Emmett MR, Korostelev A, Ongley H, Marshall AG, Chapman MS (2006) *J Virol* 80:6171
- Itoh S, Kawasaki N, Hashii N, Harazono A, Matsuishi Y, Hayakawa T, Kawanishi T (2006) *J Chromatogr A* 1103:296
- Dalpathado DS, Irungu J, Go EP, Butnev VY, Norton K, Bousfield GR, Desaire H (2006) *Biochemistry* 45:8665