Rembert Pieper
Christine L. Gatlin
Anthony J. Makusky
Paul S. Russo
Courtney R. Schatz
Stanton S. Miller
Qin Su*
Andrew M. McGrath
Marla A. Estock
Prashanth P. Parmar
Ming Zhao
Shih-Ting Huang
Jeff Zhou
Fang Wang
Ricardo Esquer-Blasco
N. Leigh Anderson**
John Taylor
Sandra Steiner

Large Scale Biology Corporation, Germantown, MD, USA

The human serum proteome: Display of nearly 3700 chromatographically separated protein spots on two-dimensional electrophoresis gels and identification of 325 distinct proteins

Plasma, the soluble component of the human blood, is believed to harbor thousands of distinct proteins, which originate from a variety of cells and tissues through either active secretion or leakage from blood cells or tissues. The dynamic range of plasma protein concentrations comprises at least nine orders of magnitude. Proteins involved in coagulation, immune defense, small molecule transport, and protease inhibition, many of them present in high abundance in this body fluid, have been functionally characterized and associated with disease processes. For example, protein sequence mutations in coagulation factors cause various serious disease states. Diagnosing and monitoring such diseases in blood plasma of affected individuals has typically been conducted by use of enzyme-linked immunosorbent assays, which using a specific antibody quantitatively measure only the affected protein in the tested plasma samples. The discovery of protein biomarkers in plasma for diseases with no known correlations to genetic mutations is challenging. It requires a highly parallel display and quantitation strategy for proteins. We fractionated blood serum proteins prior to display on two-dimensional electrophoresis (2-DE) gels using immunoaffinity chromatography to remove the most abundant serum proteins, followed by sequential anion-exchange and size-exclusion chromatography. Serum proteins from 74 fractions were displayed on 2-DE gels. This approach succeeded in resolving approximately 3700 distinct protein spots, many of them post-translationally modified variants of plasma proteins. About 1800 distinct serum protein spots were identified by mass spectrometry. They collapsed into 325 distinct proteins, after sequence homology and similarity searches were carried out to eliminate redundant protein annotations. Although a relatively insensitive dye, Coomassie Brillant Blue G-250, was used to visualize protein spots, several proteins known to be present in serum in $<$ 10 ng/mL concentrations were identified such as interleukin-6, cathepsins, and peptide hormones. Considering that our strategy allows highly parallel protein quantitation on 2-DE gels, it holds promise to accelerate the discovery of novel serum protein biomarkers.

Keywords: Blood plasma / Mass spectrometry / Multidimensional liquid chromatography / Protein biomarker / Serum proteome / Two-dimensional gel electrophoresis PRO 0449

1 Introduction

A resurgence of interest in the human plasma proteome has occurred in recent years because of the central role plasma plays in clinical diagnostics. Protein concentrations in plasma are tightly controlled to balance their physiological functions in areas such as immunity, coagulation, small molecule transport, inflammation, and lipid metabolism. Lack of function and out-of-balance concentrations of plasma proteins can cause or result from disease processes. For example, both quantitative and qualitative deficiencies of clotting factor VIII are known to lead to the phenotype of bleeding disorder hemophilia A [1]. Thrombotic microangiopathies comprise a group of diseases for which the molecular causes lie either in muta-

Correspondence: Rembert Pieper, Large Scale Biology Corporation, 20451 Seneca Meadows Parkway, Germantown, MD 20876, USA
E-mail: rembert.pieper@lsbc.com
Fax: +1-301-354-1300

Abbreviations: **ACTH**, adrenocorticotropic hormone; **AEC**, anion-exchange chromatography; **2-DLC**, two-dimensional liquid chromatography; **IASC**, immunoaffinity subtraction chromatography; **Ig**, immunoglobulin; **pAbs**, polyclonal antibodies; **SEC**, size-exclusion chromatography

* Present address: National Cancer Institute, NIH, Room 3B02, Bldg. 36, 36 Covent Dr., Bethesda, MD, 20892, USA
** Present address: Plasma Proteome Institute, P.O. Box 53450, Washington, DC 20008, USA

 0173-0835/03/0707–1345 $17.50+.50/0

tions or in the inactivation of the von Willebrand Factor-cleaving protease due to the presence of auto-antibodies [2, 3]. Eventually, this adversely affects the coagulation pathway. The molecular causes of this disease group present a good example for the possibility to either track a disorder at the genetic level (mutations in the protease) or at the protein level (presence of truncated von Willebrand Factor-cleaving protease with a lower molecular mass or protease auto-antibodies). The majority of diseases are thought not to be linked to a single structural gene mutation. There is strong evidence that both genetic and epigenetic factors contribute to the causes of many disease processes. To investigate epigenetic factors, human serum and plasma are easily accessible sample sources and available at various time points in a progressing or regressing disease. Plasma protein changes can present many clues to gain insight into the anatomical and molecular origin of a disease.

Clinical tests for disease diagnoses in plasma have been developed, although at a slow pace as reviewed recently [4]. Individual proteins have been measured as biomarkers in ELISAs, such as γ-glutamyl transferase [5] to reveal liver malfunction, or troponin T [6], myoglobin [7], and creatine kinase MB [8] to diagnose myocardial infarction or prostate-specific antigen [9] for prostate cancer diagnosis. Parallel quantitative display of proteins is believed to be the most promising strategy for biomarker discovery. One core technology, 2-DE, was established 25 years ago and applied to display blood plasma proteins [10–12]. The second core technology, mass spectrometry (MS), has revolutionized the proteomics field more recently [13–15]. MS has been applied to identify many plasma proteins from 2-DE gel-purified spots with a yield of about 70 unique protein annotations [16–18]. The powerful combination of liquid chromatography (LC) and MS has succeeded in identifying *ca.* 500 plasma hemofiltrate peptides [19] and *ca.* 500 uniquely annotated proteins after tryptic digestion of serum proteins and 2-D LC peptide separation [20]. Furthermore, surface-enhanced laser desorption ionization mass spectrometry (SELDI) has been implemented to advance proteomic biomarker discovery in serum in a recent study on ovarian cancer. Instead of identifying individual protein markers, polypeptide abundance patterns were produced employing a novel bioinformatic data analysis procedure [21].

The limited dynamic range of proteins covered in any high-throughput proteomic approach is a challenge common to all of the aforementioned technologies. Both plasma and serum show tremendous variations in individual protein abundances, *e.g.*, albumin is 10^9-fold more abundant in serum than troponin T. Fractionation approaches should expand the dynamic range of protein measurements in serum. They may be based on electrophoretic techniques, such as narrow-range pH gradient 2-DE methods [22] and free-flow electrophoresis [23], or on LC methods [24, 25]. We applied a recently developed immunoaffinity subtraction procedure [26] in conjunction with one or two sequential LC methods followed by 2-DE to fractionate and separate medium- and low-abundance serum proteins. The gel-displayed protein spots were subjected to highly automated protein identification procedures using both MALDI-MS and LC-MS/MS methods. About 1800 distinct protein features including many post-translational variants of serum proteins separated by 2-DE were MS-identified.

2 Materials and methods

2.1 Materials

Polyclonal antisera were obtained from various manufacturers. Sigma-Aldrich (St. Louis, MO, USA): immunoglobulin G (IgG) fraction of anti-albumin (rabbit), IgG fraction of anti-haptoglobin (rabbit), IgG fraction of anti-transferrin (goat), IgG fraction of anti-transthyretin (rabbit), IgG fraction of anti-α-1-antitrypsin (rabbit). Kent Laboratories (Bellingham, WA, USA): antiserum to α-2-macroglobulin (goat), antiserum to α-1-acid glycoprotein (goat), antiserum to hemopexin (goat). POROS® A 20, POROS® G 20 and POROS® HQ 20 resins were purchased from Applied Biosystems (Foster City, CA, USA). The Superdex™ 200 prep grade resin and the 2-DE carrier ampholytes (pH range 8–10.5) were from Amersham Biosciences (Piscataway, NJ, USA). Sequencing grade porcine trypsin was purchased from Promega (Madison, WI, USA). Angiotensin II, $ACTH_{18–39}$ (adrenocorticotropic hormone fragment 18–39) and Glu^1-fibrinopeptide B were purchased from Sigma-Aldrich.

2.2 Human serum

Human blood was obtained by venipuncture from two healthy male donors (ages 40 and 80). The blood samples were allowed to clot over 2 h at 20°C. The clotted material was removed by centrifugation spinning at $3000 \times g$ for 15 min. The supernatant sera obtained from the blood samples of both donors were combined in equal volumes. Sodium azide (0.01%) was added before freezing 1.5 mL serum aliquots at −20°C. They were thawed and recentrifuged at $15\,000 \times g$ for 15 min prior to use. Insoluble matter was discarded and the supernatant applied to chromatographic analysis.

2.3 Immunoaffinity subtraction chromatography

Using the Biocad-Vision™ workstation (Applied Biosystems), a multistep chromatographic procedure was used to purify, immobilize, and chemically cross-link antibodies specific for eight abundant human serum proteins – albumin, haptoglobin, transferrin, transthyretin, α-1-antitrypsin, α-1-acid glycoprotein, hemopexin, and α-2-macroglobulin. Antisera specific for each of the aforementioned antigens were loaded selectively purifying polyclonal antibodies (pAbs) on the appropriate protein antigen affinity columns. pAbs were acid-eluted at pH 2.1, exchanged into neutral buffer by gel filtration and immobilized on protein A or G resins. For rabbit antisera, a protein A-derivatized POROS A column was used to trap the antibodies. A protein G-derivatized POROS G column was prepared to trap antibodies from goat antisera followed by covalent cross-linking of antibodies to the matrix as described earlier [26]. Thus, a series of recyclable antibody-coupled columns, each specific for IgG and one of the eight aforementioned abundant serum proteins, were generated. The resin slurries were combined in volume ratios proportional to the relative amount of each protein in serum. A 7.8 mL mixed-bed immunoaffinity subtraction chromatography (IASC) column was generated. This mixed-bed IASC column was evaluated with respect to its approximate binding specificities and capacities for the nine aforementioned serum proteins applying whole serum as described earlier [26]. 100 μL serum aliquots were loaded in buffer A (25 mM sodium phosphate, 500 mM sodium chloride, 0.01% sodium azide, 2 mM EDTA, and 1 mM benzamidine, pH 7.2) and column-flowthrough fractions were collected at a flow rate of 1.5 mL/min. 4 mL of the acidic elution buffer B (5% acetic acid, 150 mM sodium chloride, pH 2.1) were injected eluting the affinity-bound proteins and recycling the column. In a series of sequential IASC runs, 20 mL serum was processed pooling all flowthrough fractions, which were concentrated on YM5 membranes (M_r cutoff of 5 kDa) in a Stirred Cell (Millipore, Billerica, MA, USA).

2.4 Anion-exchange chromatography

The immunoaffinity-subtracted serum protein concentrate, stored at −20°C before use for anion-exchange chromatography (AEC), was equilibrated in 12 mL buffer C (25 mM Tris-HCl, 25 mM NaCl, 2 mM EDTA, 1 mM benzamidine, and 0.1 mg/mL leupeptin, pH 7.6). In each of three chromatography runs, 4 mL of the protein concentrate was applied to a POROS HQ column (3.9 mL resin). At a flow rate of 2.5 mL/min, a linear gradient elution was run from 25 mM to 375 mM NaCl in buffer C over 60 mL followed by a steeper gradient elution from 375 mM to 1.5 M NaCl (30 mL). 22 AEC protein fractions were collected in each LC run. Equivalent fractions from the three LC runs were combined and spin-concentrated to volumes of 1 mL in Ultrafree®-4 (5K) membrane units (Millipore) at 3500 rpm. Protein amounts in all AEC fractions were measured using the bicinchoninic acid (BCA) assay (Pierce, Rockford, IL, USA). To generate samples for 2-DE from fractions after the AEC separation, 25% of each concentrate was set aside. After combination of some fraction aliquots with low protein content, 8 (out of 22) final AEC serum protein samples were obtained and prepared for 2-DE as described in Section 2.6.

2.5 Size-exclusion chromatography (SEC)

From the remaining 75% of the volume of the AEC fractions, six final fraction pools were obtained combining some adjacent fractions with low protein content. All six protein samples were further concentrated in Ultrafree-4 (5K) units to final volumes of 500 μL prior to SEC. The samples were loaded onto the Superdex 200 column (1.6 × 100 cm, 200 mL) sequentially. Chromatography runs were carried out in buffer A (with 150 mM NaCl) at a flow rate of 0.75 mL/min collecting 18 protein fractions. Two or three fractions with lower protein amounts eluting in adjacent fractions were pooled resulting in 11 final protein samples from each of the six SEC runs. Thus, 66 samples were prepared for 2-DE.

2.6 Two-dimensional gel electrophoresis

Protein fractions eluted from either the AEC or the SEC column were subjected to a 100-fold buffer exchange into buffer D (25 mM ammonium bicarbonate, 0.5 mM sodium EDTA, and 0.5 mM benzamidine) in Ultrafree-4 (5K) units. Protein concentrates of 200–300 μL were pipetted into microvials, lyophilized for 15–24 h and solubilized in the 2-DE isoelectric focusing buffer: 2% CHAPS, 9 M urea, 62.5 mM DTT, 2% pH 8–10.5 carrier ampholytes. 2-DE was performed using the high-throughput ProGEx™ system (Large Scale Biology Corporation) as described before [27]. Briefly, solubilized samples (5–20 μL) were loaded manually onto 4% T IEF tube gels with the pH range of 4–7. Proteins were focused in the first electrophoretic dimension for 25 000 Vh. Employing the proprietary Angelique™ computer-controlled gradient-casting system, SDS slab gels were prepared in batches of 28 gels at a time featuring a linear gradient ranging from 8% T to 15% T (top to bottom). Each focused IEF gel was placed on top of a polymerized slab gel and held in place with 1% agarose. Second-dimensional slab gels were resolved in the M_r range between *ca.* 200 and 10 kDa over 2 h in sets of 25 in cooled DALT™ tanks (1300 Vh,

20°C). Electrophoresed gels were fixed overnight and stained in a Coomassie Brilliant Blue G-250 staining solution for 3 days [27].

2.7 Sample preparation for MS and MALDI-TOF analysis

2-DE gels were scanned using the Kepler® software package assigning positional locations to each spot on each gel. Spot location data was stored in a relational database and retrieved by a proprietary spot-cutter. All visible CBB-stained gel spots were systematically cut out and collected into bar-coded 96-well microtiter plates for further processing. Sample preparation of gel plugs included destaining, reduction, alkylation and trypsin digestion using a TECAN Genesis Workstation 200 (Tecan, Durham, NC, USA) as described previously [27]. After digestion with trypsin, peptides were extracted from the gel plugs and spotted onto MALDI target plates using the 96-tip CyBi-Well robot (CyBio, Wobern, MA, USA). A fraction of the sample volumes was deposited onto a 384-format Bruker 600 μm AnchorChip MALDI target followed by α-cyano-4-hydroxy-cinnamic acid matrix. Samples plus matrix were allowed to dry, followed by a wash with 1% TFA. The remainder of the samples was prepared for LC-MS/MS analysis using a Packard Multiprobe II EX liquid handling system (Perkin Elmer, Boston, MA, USA), transferred to 96-well microtiter plates (220 μL) and brought to a volume of 10 μL. MALDI targets were automatically run on a Bruker Biflex or Autoflex mass spectrometer. Both instrument models were equipped with delayed ion extraction, pulsed nitrogen lasers (10 Hz Biflex, 20 Hz Autoflex), dual microchannel plates, and 2 GHz transient digitizers. All mass spectra represented signal averaging of 120 laser shots. The performance of the mass spectrometers had sufficient mass resolution to produce isotopic multiplets for each ion species below *m/z* 3000. Spectra were internally calibrated using two spiked peptides (angiotensin II and $ACTH_{18–39}$) and database-searched with a mass tolerance of 50 ppm.

2.8 LC-MS/MS analysis

Samples that did not get positive identifications by MALDI were subjected to LC-MS/MS analysis using Finnegan LCQ mass spectrometers. A micro-electrospray interface similar to an interface described previously [28] was employed. Briefly, the interface utilized a PEEK micro-tee (Upchurch Scientific, Oak Harbor, WA, USA) into one stem of which was inserted a 0.025" platinum-iridium wire (Surepure Chemetals, Florham Park, NJ, USA) to supply the electrical connection. The spray voltage was 1.8 kV. A 30 μm ID PicoTip spray needle (New Objectives, Cambridge, MA, USA) was inserted into another arm of the tee and aligned with the MS orifice. A 10 cm microcapillary column packed with 5 μm reversed-phase C18-Zorbax material (Microtech Scientific, Vista, CA, USA) was plumbed into the last arm of the tee. A 20 μL/min flow from a Microtech UltraPlus II 3-pump solvent delivery system (Microtech Scientific) was reduced using a splitting tee to achieve a column flow rate of 400 nL/min. Samples were injected from an Endurance autosampler (Spark-Holland, The Netherlands) onto a trapping cartridge (Cap-Trap, Michrom BioResources, Auburn, CA, USA) with pump C. Seven minute reversed-phase gradients from pumps A and B eluted peptides off the trap and the capillary LC column into the MS. Spectra were acquired in automated MS/MS mode with a relative collision energy (RCE) preset to 35%. To maximize data acquisition efficiency, the additional parameters of dynamic exclusion, isotopic exclusion, and top-3-ions were incorporated into the auto-MS/MS procedure. The scan range for MS mode was set at *m/z* 375–1400. A parent ion default charge state of +2 was used to calculate the scan range for acquiring MS/MS data.

2.9 MS data analysis

MS data was automatically registered, analyzed, and searched with the appropriate public protein/genome databases using RADARS, a separate relational database provided by Harvard Biosciences (Holliston, MA, USA). For MALDI peptide mapping, Mascot (Matrix Science, London, UK) and Profound (Harvard Biosciences) search engines were employed. Identifications were noted in the Kepler® relational database, when one of the following situations occurred: (i) both Profound and Mascot search results were above the 95th percentile of significance showing the same protein identification (scores of Profound $\geq$ 1.65 and Mascot $\geq$ 50); (ii) one of the two search engines delivered results above the 95th percentile of significance, whereas the other search engine was below it (scores as low as 1.0 for Profound and 35 for Mascot), but with the same protein identification as the top hit; (iii) one of the two search engines delivered results above the 95th percentile of significance with no corroborative result from the other search engine, however, where the manually observed spectrum had a peptide fingerprint quality positively identifying the protein. Mascot was used for peptide sequence searching of LC-MS/MS data. Scores above the 95th percentile (Mascot $\geq$ 50) were noted in the Kepler® database.

3 Results

3.1 Separation of serum proteins using immunoaffinity subtraction, anion-exchange and size-exclusion chromatography combined with 2-DE

In the first-dimension chromatography, we utilized a recently developed immunoaffinity-based method [26] removing nine very abundant plasma proteins – albumin, immunoglobulin G, haptoglobin, transferrin, transthyretin, α-1-antitrypsin, α-1-acid glycoprotein, α-2-macroglobulin, and hemopexin – from serum samples. After depletion of the abundant proteins in one chromatography cycle, approximately 620 unique protein features were observed in a CBB-stained 2-DE gel (see Fig. 1) spanning a dynamic range for protein detection of three to four orders of magnitude. This technique was suitable to moderately enrich lower-abundance proteins in serum and plasma. It increased the protein measurement sensitivity in a CBB-stained gel to a level of ~10 μg protein/mL serum and permitted, for example, the detection of serum amyloid P component and retinol-binding protein, as indicated in Fig. 1. The concentrations of these two proteins amount to *ca.* 10–15 μg/mL [29] and 30–60 μg/mL [30] in serum, respectively. Proteins such as interleukins and various enzymes (*e.g.*, L-lactate dehydrogenase), which are routinely detected in serum *via* ELISA measurements and present in human serum in amounts five to nine orders of magnitude lower than albumin (35–45 mg/mL) and IgG (10–15 mg/mL), were not visualized.

In an effort to achieve a more comprehensive coverage of the serum proteome, human serum was prefractionated by either a 2-D LC (2-DLC) procedure preparing eight protein fractions for 2-DE, or by an approach using three sequential chromatography techniques (3-DLC), which yielded 66 fractions, as illustrated in Fig. 2. Immunoaffinity subtraction of serum aliquots was carried out in recycling mode in a series of LC runs, because the IASC antibody column was limited in its serum protein-binding capacity to *ca.* 10 mg/LC run. Twenty mL of serum were processed enriching low- and medium-abundance plasma proteins in the column-flowthrough for further fractionation. After concentration, the yield was *ca.* 112 mg protein. This sample was applied to a quarternary amine anion exchanger and proteins were eluted with an increasing sodium chloride gradient (see Fig. 3A). The fractions were divided into 75% further chromatographically fractionated in the 3-DLC procedure and 25% applied to 2-DE after combining adjacent AEC fractions to eight final fraction pools. Each of these fraction pools yielded a reproducible protein pattern as illustrated in the gel images F1 to F8 of Fig. 4. 2-DLC improved the spot

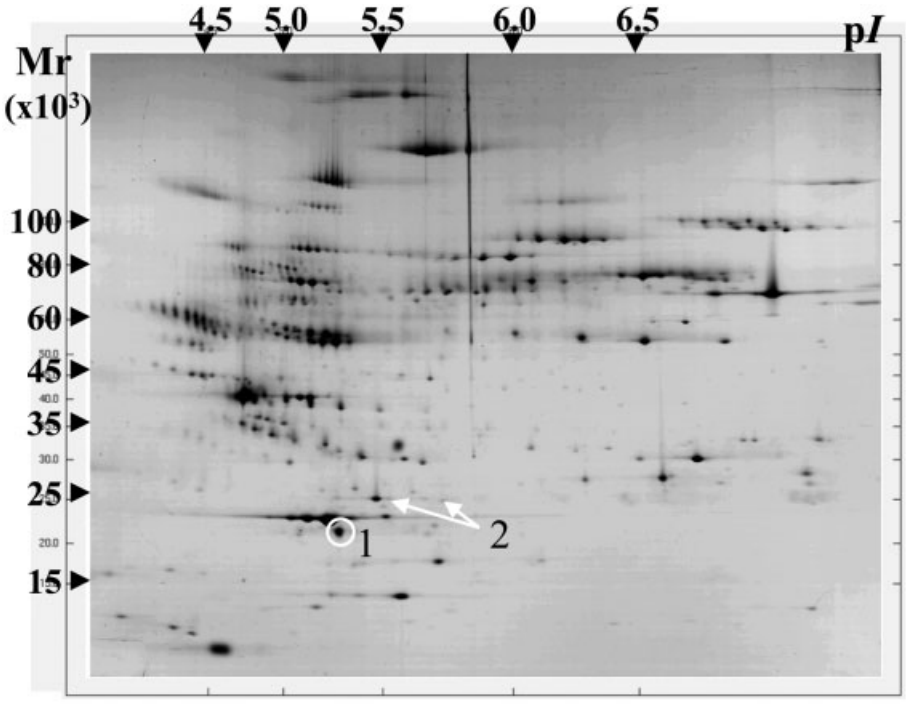

Figure 1. Human serum protein pattern after removal of several abundant proteins by IASC as visualized in a CBB-stained 2-DE gel. Following IASC, which subtracted *ca.* 85% of total serum protein, the fraction of lower-abundance serum proteins was subjected to 2-DE. 2-DE gel run conditions are described in the text. 130 μg protein was loaded onto the IEF first-dimension gel. Following second-dimension separation, protein spots were visualized with Coomassie Brilliant Blue G-250 dye (CBB). Spot 1 was identified as retinol-binding protein, spots 2 as serum amyloid P. The approximate M_r and p*I* scales were derived from protein calibration curves in a 2-DE gel run of the same batch. The calibrants consisted of a set of *ca.* 200 2-DE-separated and MS-identified rat liver proteins with M_r and p*I* values (in 2-DE gels) known to match the respective theoretical values well.

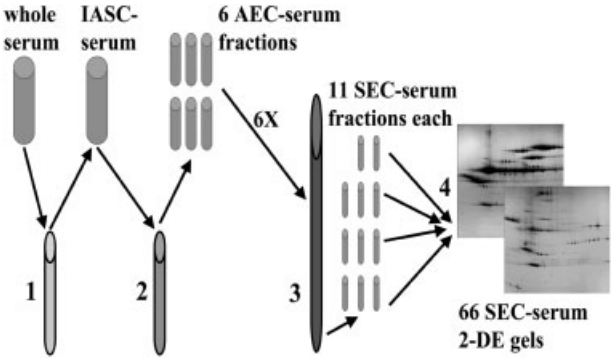

Figure 2. Scheme for the 3-DLC fractionation of the human serum proteome. 1. IASC: aliquots of human serum were separated into immunoaffinity-subtracted proteins and a fraction of lower-abundance unretained proteins in recycling mode. 2. AEC: unretained proteins (IASC-serum) were fractionated *via* anion exchange. 3. SEC: each of the six AEC-serum fractions was subjected to SEC obtaining 11 fractions in each run. 4. 2-DE: each of the 66 SEC-serum fractions was concentrated and aliquots were applied to the first (IEF) and second (SDS-PAGE) electrophoretic dimension resulting in 66 distinct protein patterns.

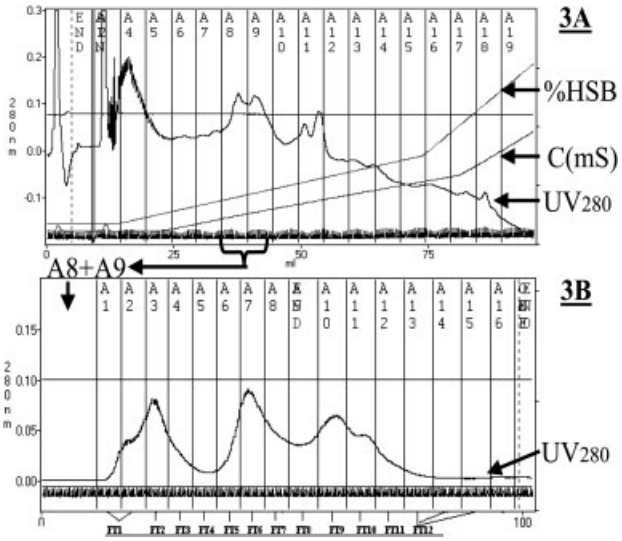

Figure 3. Selected chromatograms for serum protein separations *via* (A) strong anion-exchange (AEC) and (B) SEC. To separate proteins *via* AEC, an increasing gradient of NaCl – here monitored in %HSB (high salt buffer C with 1.5 M NaCl) and in the conductivity trace C (measured in mS) – was applied. Fractions A4–A15: 0.025–0.375 M NaCl; fractions A16–A20: 0.375–1.5 M NaCl. In the SEC chromatogram, a concentrated protein sample derived from fractions A8 and A9 of the AEC step was applied and proteins were separated according to their native molecular weights. The UV_{280} trace was monitored from 60 to 160 mL column elution volume. The A3, A7, and A10 peak areas correspond to M_r's of approximately 600, 150, and 80 kDa, respectively.

resolution compared to IASC prefractionation alone and allowed the detection of *ca.* 3600 spots combined from all eight gels. Due to fraction-to-fraction overlaps, 2100 of the protein spots were estimated to be unique, which included many proteins varying only in their post-translational modifications. These modifications of the same gene product are displayed in 2-DE gels as more or less effectively separated 2-DE spots, which are part of a spot train extended particularly in the p*I* dimension. The enhanced dynamic range for protein detection in serum was reflected in the enrichment of many lower-abundance protein spots. For example, cholinesterase and glutathione peroxidase, enzymes measured in plasma in concentrations below 10 μg/mL [30, 31], were visualized as indicated in gel F6 (Fig. 4B). Considering that protein binding to the AEC matrix occurs through negatively charged protein amino acid side chains and sialic acid groups of glycoproteins, one would assume a strong tendency of proteins eluting in low-salt AEC fractions to cluster in the basic region and of proteins eluting in high-salt AEC fractions to cluster in the acidic region of a 2-DE gel. However, only a minor clustering trend was observed, arguably caused by the participation of serum proteins in stable complexes, which alter the surface charges of individual protein components, and the contribution of charge-unrelated binding effects between the chromatographic matrix and some protein surface structures. Consequently, anion-exchange chromatography is a protein separation method complementary to 2-DE.

Even after serum fractionation including IASC and AEC, dense spot patterns were still visible, particularly in the M_r range between 50 and 80 kDa (see gels F3–F7, Fig. 4). To further increase the dynamic range of proteins detectable in serum, SEC separations were carried out with six pooled protein samples derived from the AEC fractionation. The chromatogram in Fig. 3B illustrates the protein separation of one of six AEC serum samples (A8+A9) in the molecular mass range between 10 and 500 kDa. Protein losses of up to 60% (loading 5–12 mg protein in each LC run) were measured after SEC, likely due to the large Superdex column volume (200 mL). Recovered proteins distributed over a total of 66 fractions were displayed in an equivalent number of 2-DE gels. 3-DLC, as demonstrated in four selected CBB-stained gel images (Figs. 5–8), was effective in further resolving proteins. Approximately 16 000 spots were registered by image processing of the 66 CBB-stained gels. Due to spot overlaps across fractions in the AEC- and the SEC-dimensions, the number for unique protein spots was estimated to be 3500. Following MS analysis, the 2-DE locations of spots yielding identifications were retrieved and permitted the assembly of numerous 2-DE serum protein maps, as

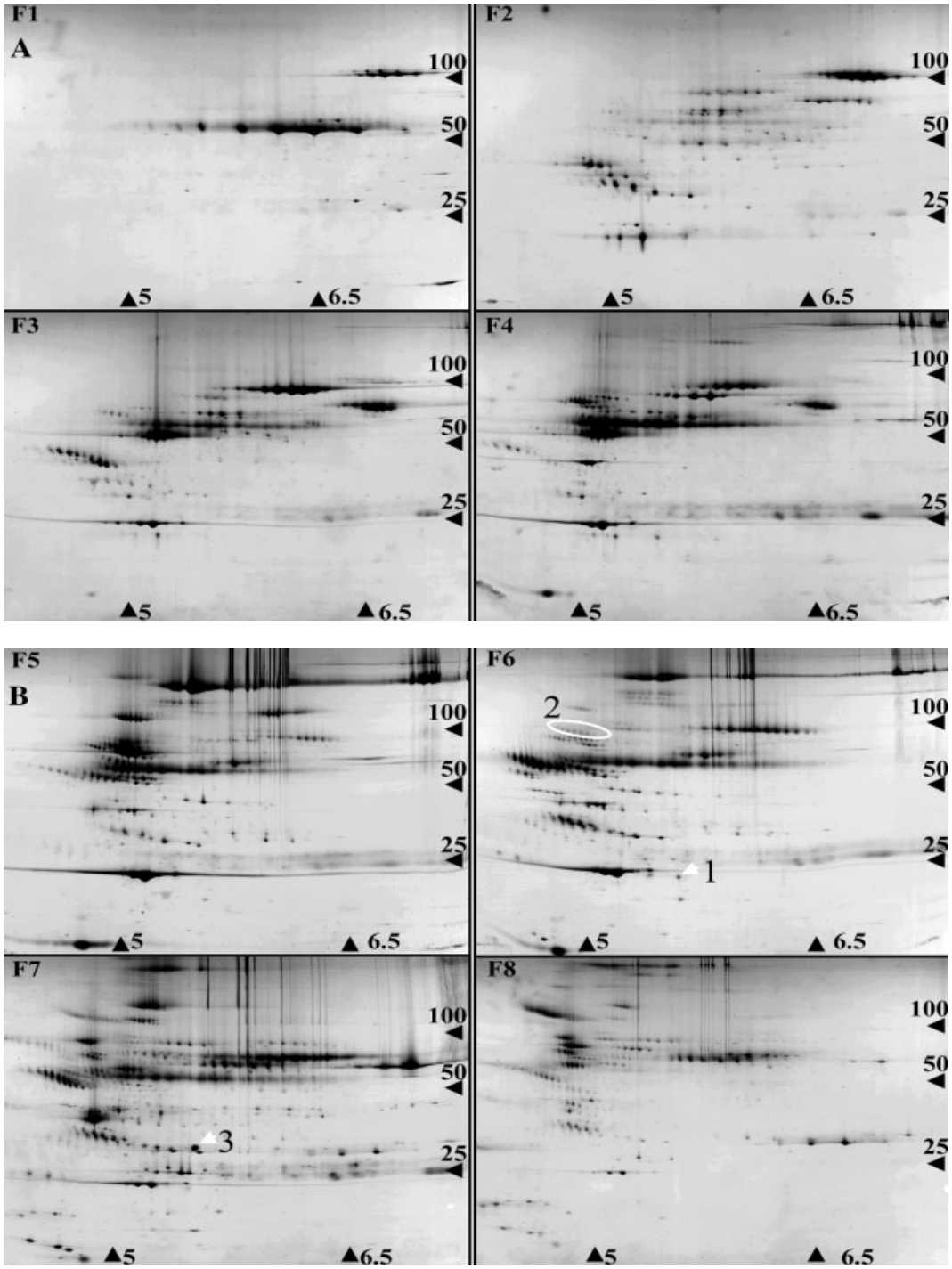

Figure 4. Proteins displayed on eight CBB-stained 2-DE gels following human serum protein fractionation including IASC and AEC (2-DLC). 2-DE run parameters and the method applied for p*I* and M_r approximations are described in the text. p*I*'s are indicated at the bottom of each gel (5 and 6.5) and M_r's on the right side of each gel (25, 50, and 100 kDa). The gels F1 to F4 (in 4A) and F5 to F8 (in 4B) correspond to the order of fractions eluted from the POROS-HQ column using a two-step linear NaCl gradient in 25 mM Tris, pH 7.6, as illustrated in chromatogram (A) of Fig. 3. F1: 0.025 M NaCl; F2: 0.1 M NaCl; F3: 0.15 M NaCl; F4: 0.25 M NaCl; F5: 0.35 M NaCl; F6: 0.45 M NaCl; F7: 0.6 M NaCl; F8: 1.5 M NaCl. Spot 1 was identified as plasma glutathione peroxidase and spot train 2 as cholinesterase (both in gel F6). Retinol-binding protein and transthyretin were identified from one spot, spot 3 in gel F7.

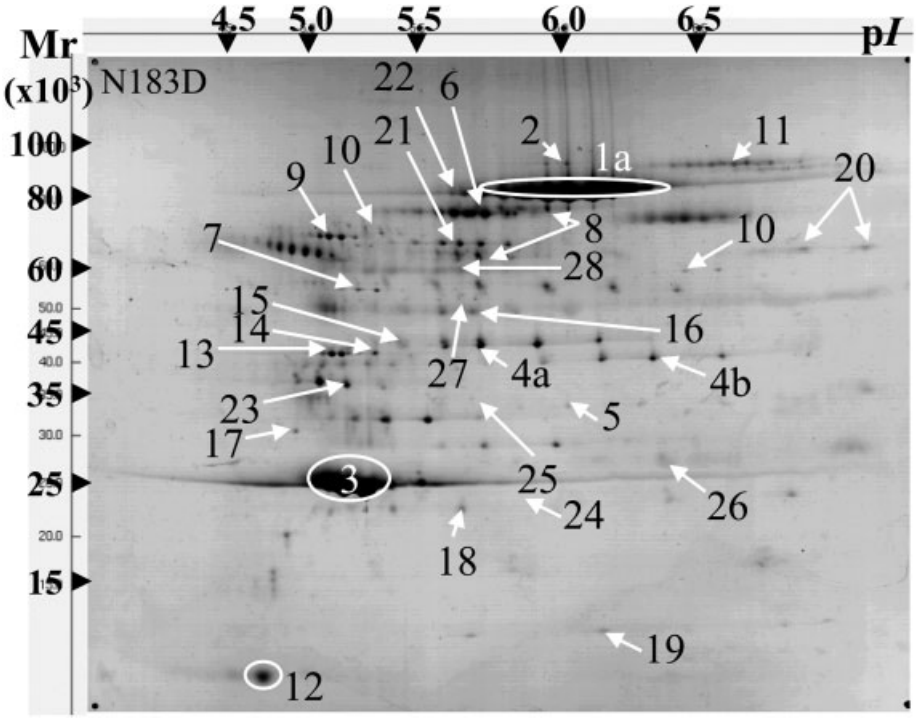

Figure 5. 2-DE spot positions of MS-identified proteins in serum following 3-DLC fractionation. The CBB-stained gel N183D corresponds to a fraction eluted with 25 mM Tris, pH 7.6, and 0.1 M NaCl from the POROS HQ column (AEC), which – upon fractionation by SEC – eluted in the M_r range of 95–110 kDa. The spot numbers match the ones listed in Table 1.

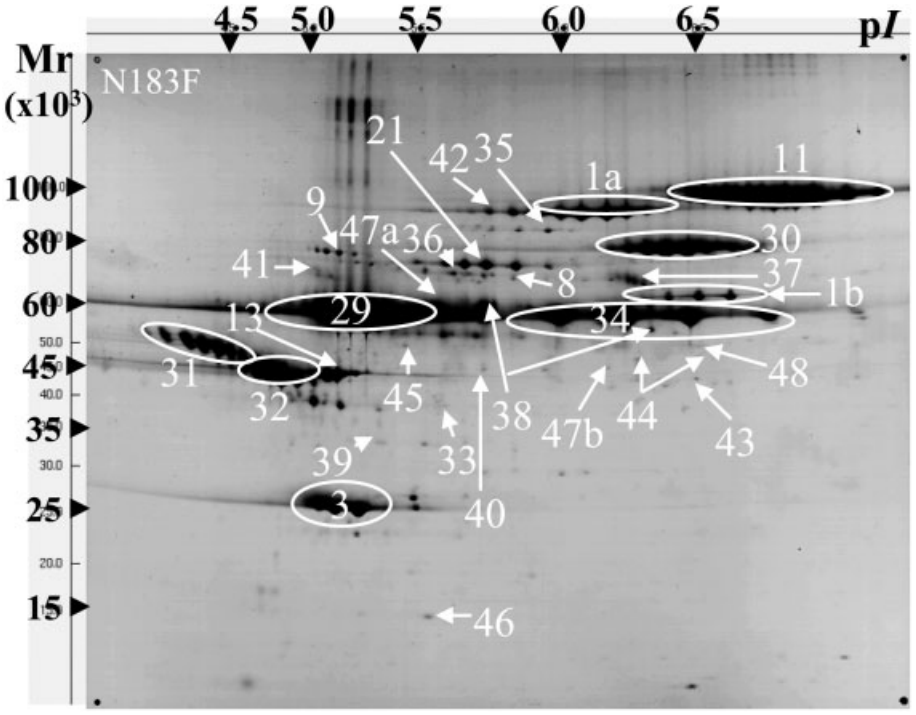

Figure 6. 2-DE spot positions of MS-identified proteins in serum following 3-DLC fractionation. The CBB-stained gel N183F corresponds to a fraction eluted with 25 mM Tris, pH 7.6, and 0.1 M NaCl from the POROS HQ column (AEC), which – upon fractionation by SEC – eluted in the M_r range of 75–85 kDa. The spot numbers match the ones listed in Table 1.

exemplified by the images in Figs. 5–8. Proteins of medium abundance in serum were highly enriched in distinct 3-DLC fractions and, despite extensive fractionation, some of their spot trains were still not perfectly resolved. This was observed, *e.g.*, for vitamin D-binding protein, plasminogen, and β-2-glycoprotein 1 in gel N183F (Fig. 6, spot trains numbered 29, 11 and 34, respectively). However, many lower-abundance proteins in serum were effectively resolved including, if applicable, their differentially glycosylated protein variants.

Many glycosylated plasma proteins were displayed as spot trains with higher M_r values than calculated from the polypeptide lengths. An example is cholinesterase (spot train 2 in gel F6, Fig. 4B), which has nine *N*-linked carbohydrate chains and whose spots band around 90 kDa. The enzyme's polypeptide-based M_r is 68.4 kDa. For other proteins, proteolytic fragments were displayed at lower M_r's in 2-DE gels than predicted from the respective full-length protein sequences. Several of them are well-characterized plasma proteins expressed as pro-molecules and cleaved site-specifically during or after secretion into the blood, *e.g.*, complement component C3. We detected C3α, C3α fragments and C3β with the denatured molecular masses (second dimension of 2-DE) predicted from the known complement C3 cleavage and disulfide reduction sites: C3β (~70 kDa) in spot 51b in gels N214B (Fig. 7) and N214H (Fig. 8); C3α (~130 kDa) in spot 51a in gel N214B; C3α fragments (± 40 kDa) in several spots denoted 51a in gel N214H. C3α and C3β form disulfide bonds with each other after complement C3 cleavage and C3α undergoes further proteolysis [32]. Under native conditions, the protein subunits remain associated in form of multimeric complexes: a complex of C3α and C3β proteins, possibly tetrameric, was observed to migrate in a high M_r range (> 550 kDa) during SEC fractionation, whereas a complex composed of smaller C3α fragments and C3β eluted in the M_r range of 240–290 kDa from the SEC column.

M_r values of proteins, identified in 2-DE gel spots and not part of the group of plasma proteins in circulation (see Table 1, Addendum), did occasionally not match the predicted molecular masses either, particularly cell membrane-bound proteins (category 4, Table 1). Apparently, their extracellular domains (or parts thereof) were proteolytically cleaved from blood and endothelial cell surfaces and released into the blood plasma resulting in molecular masses lower than those derived from full-length protein sequences. Examples are the α- (284 kDa) and β-chain (289 kDa) components of spectrin, whose subunits form a large erythrocyte cell membrane protein complex. Several α- and β-chain fragments were gel-displayed and identified in the M_r range between 25 and 60 kDa. Due to multicomponent complex formation and proteolytic fragmentation of proteins in serum, nondenaturing SEC proved to be complementary to the SDS-PAGE dimension of 2-DE protein display. The gel in Fig. 7 visualizes denatured proteins between 20 kDa and 90 kDa, which, prior to 2-DE application, eluted *via* SEC in an M_r range above 550 kDa. Covalent and noncovalent associations of plasma proteins forming high M_r complexes are known to be essential with respect to protein function or retention in the blood during glomerular filtration. Remarkably, even the denaturing and reducing conditions applied during 2-DE were not strong enough to dissociate two proteins: retinol-binding protein and transthyretin, binding partners in a strong protein complex described before

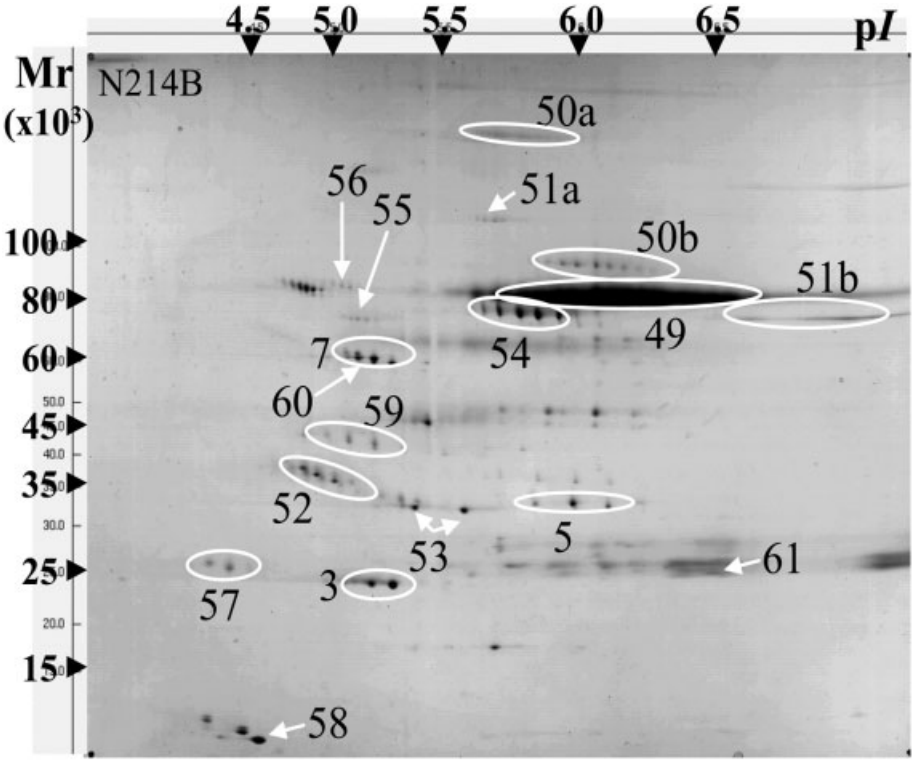

Figure 7. 2-DE spot positions of MS-identified proteins in serum following 3-DLC fractionation. The CBB-stained gel N214B corresponds to a fraction eluted with 25 mM Tris, pH 7.6, between 0.275 and 0.375 M NaCl from the POROS HQ column (AEC), which – upon fractionation by SEC – eluted in the M_r range higher than 550 kDa. The spot numbers match the ones listed in Table 1.

[33]. In a spot with a molecular mass of *ca.* 35 kDa (spot 3 in gel F7, Fig. 4B), both proteins were identified by MALDI-MS fingerprinting. Proteins known to be present in normal plasma in concentrations of less than 1 μg/mL were identified in 2-DE gels: C-reactive protein (< 1 μg/mL), metallothionein-II (< 5 ng/mL), and even interleukin-6 (< 10 pg/mL) [34]. This reflects a significant gain in the dynamic range of serum proteins visualized in CBB-stained gels following 3-DLC and 2-DE. Due to the well-documented difficulties to focus proteins *via* IEF with p*I*'s above 7.5, our 2-DE approach excluded the display of basic proteins. However, compared to tissue extracts, such proteins are less numerous in serum, as glycosylation typical for secreted proteins shifts their p*I* values towards the acidic range.

3.2 Identification of separated human serum proteins

With nearly 20 000 protein spots displayed on 2-DE gels (combined from eight gels after the 2-DLC and 66 gels after the 3-DLC fractionation experiments), MALDI-TOF peptide fingerprinting as well as LC-MS/MS peptide sequencing were applied on a large scale. The complexity of peptides in trypsin-digested samples was low as expected from the extensive chromatographic separation and protein resolution on 2-DE gels. Generally, a mathematical procedure obtaining high-confidence scores with two search algorithms (Mascot and Profound) for MALDI-spectra was employed to define a protein as identified. In *ca.* 4% of all MALDI-MS identifications, a protein was confirmed as identified resulting from a high-confidence score by one search algorithm and a corroborative MALDI-MS peptide fingerprint. If a MALDI-MS result was inconclusive, the sample was cued for LC-MS/MS analysis, which had a higher success rate for protein identification. With respect to nonredundant protein annotations, 63.5% of the proteins were identified only by MALDI-MS, 22.5% by LC-MS/MS, and 14% by MS analysis using both techniques.

Overall, high-confidence scores for approximately 500 protein annotations were obtained. Applying a modified version of the BLAST sequence alignment tool, redundant annotations setting the filters at a similarity score of 95% and higher and the homology score at 99% and higher were removed. Following 2-DLC fractionation, about 1100 unique protein spots were identified and collapsed into 157 nonredundant database annotations. Following 3-DLC fractionation, approximately 1700 spots were identified, which collapsed into 295 unique protein annotations. Thus, 5–7 spots on average were associated with each identified protein accounting for various M_r- and p*I*-altering modifications of many proteins in serum likely including glycosylation, acetylation, phosphorylation, and proteolytic cleavage. Merging the 2-DLC and 3-DLC data, approximately 1800 protein spots were successfully MS-identified related to 325 nonredundant protein annotations.

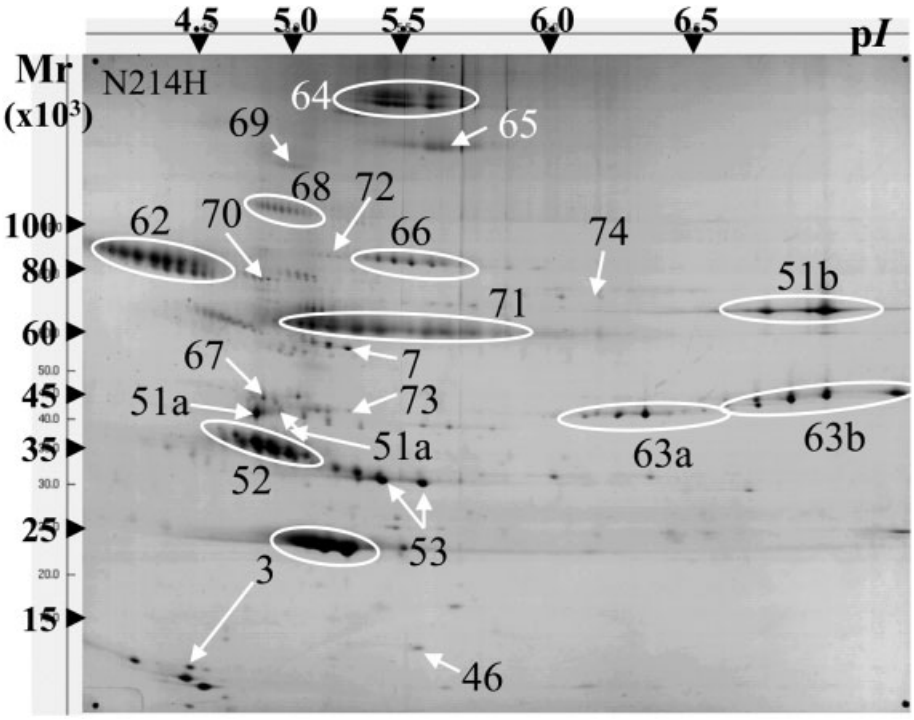

Figure 8. 2-DE spot positions of MS-identified proteins in serum following 3-DLC fractionation. The CBB-stained gel N214H corresponds to a fraction eluted with 25 mM Tris, pH 7.6, between 0.275 and 0.375 M NaCl from the POROS HQ column (AEC), which – upon fractionation by SEC – eluted in the M_r range of 240–290 kDa. The spot numbers match the ones listed in Table 1. Spots 51a and 51b were identified by LC-MS/MS as complement component C3 (P01024). The pro-molecule C3α was visualized as a faint spot train in gel N214B (51a, M_r ~130 kDa, Fig. 7), whereas C3α-fragments of lower M_r (±40 kDa) were displayed in this gel (51a). Pro-molecule C3β did not appear to be further cleaved (spots 51b, M_r ~70 kDa, in both gels).

Table 1 (Addendum) provides an overview of all proteins divided in seven categories, which, in a broader sense, represent anatomical or cellular designations of the proteins. Mascot and/or Profound scores and sequence coverage data are listed for each MALDI-identified protein, whereas a Mascot score is listed for each LC-MS/MS identification. Each protein is described by a common name and a database accession number. Proteins known to exist either in multiple isoforms (*e.g.*, actins) or polymorphisms (immunoglobulin chains) as well as specific proteolytically activated factors (*e.g.*, complement factor 1 heavy and light chains) and separately annotated in the databases were retained in the table despite the fact that the similarity and homology searches defined them as redundant. Similar to guidelines used by Putnam [30] and Anderson [4] to classify proteins in plasma, we associated the identified proteins with the following categories as illustrated in Fig. 9: 1. Classical circulating plasma proteins (126 entries); 2. Proteins localized in other extracellular fluids (29 entries); 3. Vesicular proteins with secretion signal sequences, known or assumed to be released into plasma (25 entries); 4. Cell membrane proteins (21 entries); 5. Intracellular proteins apparently leaking into the plasma due to cellular damage or lysis (113 entries); 6. Proteins not categorized due to insufficient information (10 entries); 7. Microbial proteins (1 entry).

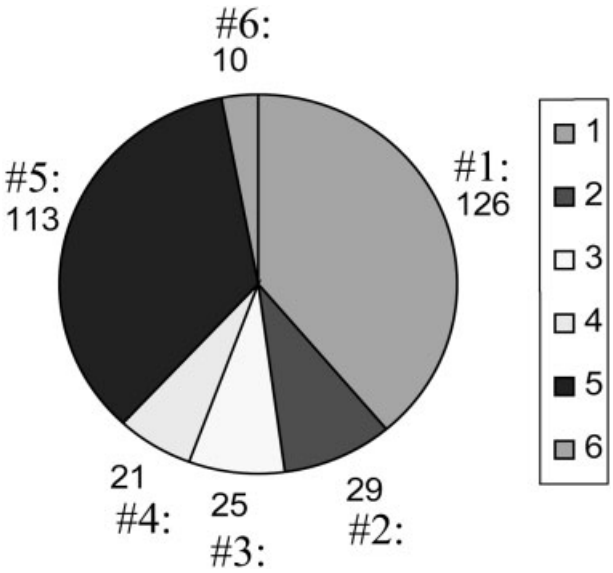

Figure 9. Categories of proteins identified in human serum. The sizes of the pie segments (with adjacent numbers) are proportional to the number of nonredundant protein annotations for the following serum protein categories. #1. Classical plasma proteins in circulation; #2. Proteins in the extracellular matrix or secreted into body fluids other than plasma; #3. Vesicular proteins (including endoplasmic reticulum, lysosomes, peroxisomes, Golgi apparatus) also – presumably or knowingly – exported into extracellular fluids; #4. Cell surface membrane proteins; #5. Intracellular proteins, presumably leaking from cells and tissues into blood plasma; #6. Uncategorized (proteins for which cellular designations are unknown).

4 Discussion

This report describes the largest effort, published to date, characterizing the human serum proteome using a 2-DE display approach combined with protein spot identification employing MALDI and LC-MS/MS methodologies. It is important to distinguish between our strategy, which addressed the characterization of a proteome on the protein level and is amenable to protein spot-based quantitation, and other strategies, whose first steps are protein digestion followed by partial resolution of resulting peptides by analytical LC. The latter approach is not straightforwardly compatible with quantitation of peptide fragments – which in theory represent their proteins of origin in proportional quantities – from chromatographic peaks due to the remaining peptide complexity. It cannot be easily adapted to quantitative evaluation of peptide spectra during the mass spectrometric analysis either. In addition, the peptide LC-MS/MS methods only infer, rather than confirm, the association of multiple peptides expected together in one protein. However, as recently demonstrated by Adkins *et al.* [20], serum proteome analysis solely based on a peptide 2-DLC fractionation approach combined with MS/MS is very sensitive with sufficient resolution to identify 490 distinct proteins of nonredundant annotations. Among the 490 proteins were some of low abundance ($<$ 1 ng/mL), such as human growth hormone, interleukin-12a, and prostate-specific antigen.

While quantitative measurements of proteins were not the focus of this report, 2-DE-displayed proteins can be subjected to spot-based quantitation in a highly parallel manner. Furthermore, from the 74 gels analyzed in this study, numerous distinct post-translational variants of identified proteins and yet unidentified proteins were visualized. The 3-DLC separation strategy yielded 3500 unique spots and allowed identification of 1700 of them. As mentioned in Section 3.1, a similar level of detection sensitivity (less than 1 ng protein/mL serum) was reached with our approach comparable to the recently described LC-MS/MS studies on human serum [20, 35]. We identified approximately 100 proteins to our knowledge not previously described in serum. Most of them appear to result from tissue and cell leakage. The overlap of presumable cell leakage proteins identified in fractionated serum in our study compared to the data of Adkins *et al.* is surprisingly low. Only six of the 134 proteins in the cellular designation

categories 4 and 5 were also identified by Adkins *et al.* This could reflect protein differences in the donor serum samples, varying serum preparation methods, the substantial technical differences in the fractionation (analysis on the protein *versus* peptide level), differences in the MS analysis procedures and/or data consolidation of protein annotations. With respect to serum preparation from donated blood, it is plausible that intracellular or cell membrane proteins, which originate in erythrocytes, leukocytes, and platelets are released into the plasma (and therefore into serum) as a result of nonphysiological cell lysis, either during blood clotting or centrifugal separation procedures. Further studies are necessary to support any of these assumptions.

It is of particular interest, that our strategy permitted the visualization and identification of proteins known to be involved in disease processes. This includes the C-reactive protein and interleukin-6, diagnostic biomarkers for inflammation and coronary heart disease [36]; metallothionein-I and -II, potential prognostic markers for the response to chemotherapy [37]; prostate-specific membrane antigen, a diagnostic biomarker for prostate cancer [38]; L-lactate dehydrogenase, a general marker for irreversible cell damage [39]; creatine kinase M, a diagnostic leakage marker for myocardial infarction [40] and cathepsins, suggested to be predictive for tumor growth and invasion processes [41]. In addition, peptide hormones such as follitropin and parathyroid hormone, apoptosis-related enzymes such as caspase 10 and growth factors such as interleukin-7 and ciliary neurotrophic factor were identified from 2-DE spots. Table 2 (Addendum) lists proteins used in practice or evaluated as disease biomarkers as well as proteins reported to be expressed tissue-specifically. Quantitation of these proteins in serum in a differential display approach could provide important links to damage or malfunction of the tissues in which they are specifically expressed, providing leads for disease biomarkers.

This work has been supported in part by a Small Business Innovative Research grant in the IMAT program by the National Cancer Institute, National Institutes of Health, Bethesda, MD, USA (Grant No. 5 R44 CA082038-03).

Received February 6, 2003
Revised March 12, 2003
Accepted March 13, 2003

5 References

[1] Hoyer, L. W., *Hum. Pathol.* 1987, *18*, 153–161.
[2] Furlan, M., Robles, R., Lamie, B., *Blood* 1996, *87*, 4223–4234.
[3] Tsai, H. M., *Blood* 1996, *87*, 4235–4244.
[4] Anderson, N. L., Anderson, N. G., *Mol. Cell. Proteomics* 2002, *1*, 845–867.
[5] Schmidt, E., Schmidt, F. W., *Prog. Liver Dis.* 1982, *7*, 411–428.
[6] Katus, H. A., Remppis, A., Looser, S., Hallermeier, K. *et al.*, *J. Mol. Cell. Cardiol.* 1989, *21*, 1349–1353.
[7] Drexel, H., Dworzak, E., Kirchmair, W., Milz, M. M. *et al.*, *Am. Heart J.* 1983, *105*, 642–651.
[8] Neumeier, D., Knedel, M., Würzburg, U., Hennrich, N., Lang, H., *Klin. Wochenschr.* 1975, *53*, 329–333.
[9] Kuriyama, M., Wang, M. C., Papsidero, L. D., Killian, C. S., Shimano, T. *et al.*, *Cancer Res.* 1980, *40*, 4658–4662.
[10] O'Farrell, P. H., *J. Biol. Chem.* 1975, *250*, 4007–4021.
[11] Klose, J., *Humangenetik* 1975, *26*, 231–243.
[12] Anderson, N. L., Anderson, N. G., *Proc. Natl. Acad. Sci. USA* 1977, *74*, 5421–5425.
[13] YatesIII, J. R., *J. Mass Spectrom.* 1998, *33*, 1–19.
[14] Aebersold, R., Patterson, S. D., *Electrophoresis* 1995, *16*, 1791–1814.
[15] Mann, M., Hendrickson, R. C., Pandey, A., *Annu. Rev. Biochem.* 2001, *70*, 437–473.
[16] Hughes, G. J., Frutiger, S., Paquet, N., Ravier, F., Pasquali, C., Sanchez, J. C. *et al.*, *Electrophoresis* 1992, *13*, 707–714.
[17] Golaz, O., Hughes, G. J., Frutiger, S., Paquet, N. *et al.*, *Electrophoresis* 1993, *14*, 1223–1231.
[18] Sanchez, J. C., Appel, R. D., Golaz, O., Pasquali, C. *et al.*, *Electrophoresis* 1995, *16*, 1131–1151.
[19] Richter, R., Schulz-Knappe, P., Schrader, M., Standker, L. *et al.*, *J. Chromatogr. B* 1999, *726*, 25–35.
[20] Adkins, J. N., Varnum, S. M., Auberry, K. J., Moore, R. J. *et al.*, *Mol. Cell. Proteomics* 2002, *1*, 947–955.
[21] Petricoin, E. F., Ardekani, A. M., Hitt, B. A., Levine, P. J. *et al.*, *Lancet* 2002, *359*, 572–577.
[22] Görg, A., Obermaier, C., Boguth, G., Weiss, W., *Electrophoresis* 1999, *4–5*, 712–717.
[23] Eckerskorn, C., Weber, G., Weber, P. J., Schneider, U., in: From Genome to Proteome, 5th Siena Meeting 2002, Plenary Session 1: Technical Aspects.
[24] Wall, D. B., Kachman, M. T., Gong, S., Hinderer, R. *et al.*, *Anal. Chem.* 2000, *72*, 1099–1111.
[25] Badock, V., Steinhusen, U., Bommert, K., Otto, A., *Electrophoresis* 2001, *22*, 2856–2864.
[26] Pieper, R., Su, Q., Gatlin, C. L., Huang, S. T. *et al.*, *Proteomics* 2003, *3*, 422–432.
[27] Steiner, S., Gatlin, C. L., Lennon, J. L., McGrath, A. M. *et al.*, *Toxicol. Lett.* 2001, *120*, 369–377.
[28] Gatlin, C. L., Kleemann, G. R., Hays, L. G., Link, A. J., Yates, IIIJ.R., *Anal. Biochem.* 1998, *263*, 93–101.
[29] Adachi, T., Mogi, M., Harada, M., Kojima, K., *J. Chromatgr. B* 1996, *682*, 47–54.
[30] Putnam, F. W., *The Plasma Proteins*, Academic Press, London 1984, *Vol. IV.*
[31] Zachara, B. A., Trafikowska, U., Adamowicz, A., Nartowicz, E., Manitius, J., *J. Trace Elem. Med. Biol.* 2001, *15*, 161–166.
[32] Sahu, A., Lambris, J. D., *Immunol. Rev.* 2001, *180*, 35–48.
[33] Rostom, A. A., Sunde, M., Richardson, S. J., Schreiber, G. *et al.*, *Proteins* 1998, *Suppl. 2*, 3–11.
[34] Akintola, D. F., Sampson, B., Burrin, J., Price, C. *et al.*, *Clin. Chem.* 1997, *43*, 845–847.
[35] Wu, S. L., Amato, H., Biringer, R., Choudhary, G. *et al.*, *J. Proteome Res.* 2002, *1*, 459–465.
[36] Blake, G. J., Ridker, P. M., *J. Intern. Med.* 2002, *252*, 283–294.

[37] Siu, L. L., Banerjee, D., Khurana, R. J., Pan, X. *et al.*, *Clin. Cancer Res.* 1998, *4*, 559–565.

[38] Israeli, R. S., Powell, C. T., Fair, W. R., Heston, W. D., *Cancer Res.* 1993, *53*, 227–230.

[39] Tyson, C. A., Green, C. E., in: Rauckmann, E. J., Padilla, G. M. (Eds.), *The Isolated Hepatocyte*, Academic Press, Orlando, FL 1987, pp. 119–158.

[40] Apple, F. S., *Am. J. Clin. Pathol.* 1992, *97*, 217–226.

[41] Kos, J., Werle, B., Lah, T., Brunner, N., *Int. J. Biol. Markers* 2000, *15*, 84–89.

[42] Bhattacharya, M., Barlow, J. J., *Int. Adv. Surg. Oncol.* 1979, *2*, 155–176.

[43] Berchem, G., Glondu, M., Gleizes, M., Brouillet, J. P. *et al.*, *Oncogene* 2002, *21*, 5951–5955.

[44] Staack, A., Koenig, F., Daniltchenko, D., Hauptmann, S. *et al.*, *Urology* 2002, *59*, 308–312.

6 Addendum

Table 1. Mass-spectrometric analysis of the human serum proteome

Cellular designation	Annotation	Protein name	Spot No.	MALDI-MS			LC-MS/MS Mascot
				Mascot	Profound	Seq Cov	
Category 1							
Plasma	gi_223099	IG α-1 chain Bur					198
Plasma	gi_223130	Fibrinopeptide β B 1-118	47a, 47b				91
Plasma	gi_223961	Complement component C4C					180
Plasma	gi_223962	Complement component C4C variant		90	2.1	8	146
Plasma	gi_1335055	Complement factor 1, heavy chain	16				85
Plasma	gi_1335056	Complement factor 1, light chain					86
Plasma	gi_15826311	Apolipoprotein E4					63
Plasma	gi_1942471	Apolipoprotein E3					135
Plasma	SP_000187	Mannan-binding lectin serine protease 2		65	1.5	10	
Plasma	SP_014791	Apolipoprotein L1		37	2.3	18.5	323
Plasma	SP_043866	CD5 antigen-like, IgM-associated peptide		117	2.3	37	
Plasma	SP_075882	Attractin		37	2.4	10.5	131
Plasma	SP_P00450	Ceruloplasmin		48	2.4	11	
Plasma	SP_P00488	Coagulation factor XIII A chain		49	2.3	15.3	58
Plasma	SP_P00734	Prothrombin		38	2.2	14	
Plasma	SP_P00736	Complement component C1r	66	86	2.4	16.2	273
Plasma	SP_P00737	Haptoglobin-1	59	49	2.2	24	
Plasma	SP_P00738	Haptoglobin-2	59				240
Plasma	SP_P00740	Coagulation factor IX		37	1.7	19	
Plasma	SP_P00742	Coagulation factor X		51	1.9	13.3	
Plasma	SP_P00747	Plasminogen	11	47	2.4	11.4	275
Plasma	SP_P00748	Coagulation factor XII	17				54
Plasma	SP_P00751	Complement factor B	1a, 1b	45	2.3	16.1	521
Plasma	SP_P01008	Antithrombin III		50	2.3	23	403
Plasma	SP_P01009	α-1 Antitrypsin	60	49	2.4	30	379
Plasma	SP_P01011	α-1 Antichymotrypsin		45	2.3	27.9	416
Plasma	SP_P01019	Angiotensinogen		40	2.4	25.2	365
Plasma	SP_P01023	α-2 Macroglobulin	50a, 50b	40	2.4	10	804
Plasma	SP_P01024	Complement component C3	51a, 51b	39	2.1	7	67
Plasma	SP_P01028	Complement component C4	72	50	1.8	8	273
Plasma	SP_P01031	Complement component C5					79
Plasma	SP_P01042	Kininogen		41	2.4	20.2	361
Plasma	SP_P01591	Ig J chain	57				102
Plasma	SP_P01616	Ig κ chain V-II region MIL					63
Plasma	SP_P01617	Ig κ chain V-II region TEW					88
Plasma	SP_P01620	Ig κ chain V-III region SIE					67
Plasma	SP_P01623	Ig κ chain V-III region WOL					106
Plasma	SP_P01834	Ig κ-chain C region	26				186
Plasma	SP_P01842	Ig λ chain C region					105
Plasma	SP_P01871	Ig μ chain C region	49	47	2.3	20	

Table 1. Continued

Cellular designation	Annotation	Protein name	Spot No.	MALDI-MS			LC-MS/MS Mascot
				Mascot	Profound	Seq Cov	
Plasma	SP_P01876	Ig α-1 chain C region	71	47	2.1	26	
Plasma	SP_P01877	Ig α-2 chain C region	71	43	2	13.5	213
Plasma	SP_P02647	Apilipoprotein A-I	3	49	2.4	29.6	683
Plasma	SP_P02649	Apolipoprotein E	53	42	2.4	26	674
Plasma	SP_P02652	Apolipoprotein A-II	12				99
Plasma	SP_P02655	Apolipoprotein C-II	58				95
Plasma	SP_P02656	Apolipoprotein C-III					92
Plasma	SP_P02735	Serum amyloid A protein	19	94	2	56	
Plasma	SP_P02741	C-reactive protein		36	2.2	24	
Plasma	SP_P02743	Serum amyloid P component		36	1.9	30	
Plasma	SP_P02748	Complement component C9	55	36	2.3	18	
Plasma	SP_P02749	β-2 glycoprotein I	34				232
Plasma	SP_P02750	Leucine-rich α-2 glycoprotein	31	44	2.4	26.9	277
Plasma	SP_P02751	Fibronectin	64				358
Plasma	SP_P02753	Plasma retinol-binding protein					201
Plasma	SP_P02760	α-1 Microglobulin		37	2.3	21.6	175
Plasma	SP_P02765	α-2-HS glycoprotein		59	2	17.7	
Plasma	SP_P02766	Transthyretin	46	48	2.1	49	
Plasma	SP_P02768	Albumin		49	2.2	11.8	449
Plasma	SP_P02774	Vitamin D-binding protein	29	48	2.4	16.7	368
Plasma	SP_P02787	Transferrin	30	47	2.1	16.3	591
Plasma	SP_P02790	Hemopexin	36				102
Plasma	SP_P03952	Plasma kallikrein					176
Plasma	SP_P04003	Complement C4b-binding protein, alpha chain	54				226
Plasma	SP_P04004	Vitronectin		39	2.4	17	
Plasma	SP_P04070	Vitamin K-dependent protein C		36	2.9	16.3	
Plasma	SP_P04114	Apolipoprotein B-100					451
Plasma	SP_P04180	Phosphatidylcholine-sterol acyltransferase		58	2.3	25	
Plasma	SP_P04196	Histidine-rich glycoprotein	37	47	2.4	18.9	329
Plasma	SP_P04217	α-1-B glycoprotein	9	45	2.4	20.7	254
Plasma	SP_P04220	Ig μ heavy chain disease protein BOT					128
Plasma	SP_P05090	Apolipoprotein D					76
Plasma	SP_P05121	Plasminogen activator inhibitor-1		45	1.7	23.4	
Plasma	SP_P05154	Plasma serine protease inhibitor	44				335
Plasma	SP_P05155	Plasma protease C1 inhibitor		117	2.1	20	
Plasma	SP_P05156	Complement factor I	23				223
Plasma	SP_P05160	Coagulation factor XIII, B chain	6	44	2.3	15.6	108
Plasma	SP_P05164	Myeloperoxidase		68	1.7	18	
Plasma	SP_P05452	Tetranectin	18				142
Plasma	SP_P05543	Thyroxine-binding globulin		39	2.3	23.1	267
Plasma	SP_P05546	Heparin cofactor II		39	2.2	17	
Plasma	SP_P06309	Ig κ-chain V-II region GM607	61				64
Plasma	SP_P06310	Ig κ-chain V-II region RPMI					62
Plasma	SP_P06396	Gelsolin	35	136	2.4	32.5	419
Plasma	SP_P06681	Complement factor C2	8	47	2.2	12	396
Plasma	SP_P06727	Apolipoprotein A-IV	13	184	2.4	52	779
Plasma	SP_P07225	Vitamin K-dependent protein S		91	2.2	17	
Plasma	SP_P07357	Complement component C8, α chain		43	1.7	13.2	170
Plasma	SP_P07358	Complement component C8, β chain	20	75	2	18.1	217
Plasma	SP_P07360	Complement component C8, γ chain		97	2.4	50	
Plasma	SP_P08185	Corticosteroid-binding globulin		44	2.4	22.2	278
Plasma	SP_P08603	Complement factor H	65	38	2.3	11	
Plasma	SP_P08697	α-2 Antiplasmin	41	44	2.4	32.8	216
Plasma	SP_P09871	Complement factor C1s		46	2.3	15	
Plasma	SP_P10643	Complement component C7		52	1.6	14.1	110
Plasma	SP_P10909	Clusterin	52	42	2.3	19	

Table 1. Continued

Cellular designation	Annotation	Protein name	Spot No.	MALDI-MS			LC-MS/MS Mascot
				Mascot	Profound	Seq Cov	
Plasma	SP_P13671	Complement component C6		75	2.2	11	199
Plasma	SP_P15169	Carboxypeptidase N, catalytic chain	63a, 63b	50	2.4	22	
Plasma	SP_P19823	Inter-α-trypsin inihibitor heavy chain H2		41	2	14	
Plasma	SP_P19827	Inter-α-trypsin inhibitor heavy chain H1		42	2.4	15	
Plasma	SP_P20472	Parvalbumin		70	1.8	44	
Plasma	SP_P20742	Pregnancy zone protein					54
Plasma	SP_P20851	Complement C4b-binding protein, β chain		67	1.7	31	250
Plasma	SP_P22352	Plasma glutathione peroxidase		79	2.2	32	208
Plasma	SP_P22792	Carboxypeptidase N, regulatory chain	62	48	2.4	21	
Plasma	SP_P25311	Zinc-α-2 glycoprotein	32	47	2.3	22.7	475
Plasma	SP_P27169	Serum paraoxonase	67	37	2.3	23	
Plasma	SP_P29622	Kallistatin	38				164
Plasma	SP_P35542	Serum amyloid A-4 protein					76
Plasma	SP_P35858	Insulin-like growth factor-binding protein complex acid labile chain	10	44	2.4	19.8	230
Plasma	SP_P36955	Pigment epithelium-derived factor		81	2.4	22	422
Plasma	SP_P36980	Complement factor H-related protein 2	39				104
Plasma	SP_P43652	Afamin		48	2.3	15.2	220
Plasma	SP_P48740	Complement-activating component of Ra-reactive factor		90	2.2	20	
Plasma	SP_P80108	Phophatidylinositol-glycan specific phospholipase D 1	68	48	2.4	15	
Plasma	SP_P80748	Ig λ chain V-III region LOI					106
Plasma	SP_Q03591	Complement factor H-related protein 1	4a, 4b				160
Plasma	SP_Q04756	Hepatocyte growth factor activator					119
Plasma	SP_Q06033	Inter-α-trypsin inhibitor heavy chain H3		89	2.4	19	
Plasma	SP_Q13790	Apolipoprotein F			1.7	12.1	
Plasma	SP_Q14624	Inter-α-trypsin inhibitor heavy chain H4	7	48	2.3	12.3	
Plasma	SP_Q15166	Serum paraoxonase 3		84	2.4	29	
Plasma	SP_Q15848	Adiponectin		36	2.1	46.3	
Plasma	SP_Q96IY4	Plasma carboxypeptidase B2	40				53
Plasma	SP_Q9BWW9	Apolipoprotein L5		59			
Plasma	SP_Q9U/K55	Protein Z-dependent protease inhibitor					115
Category 2							
Extracellular	gi_1436678	Lactotransferrin					152
Extracellular	gi_8118083	Secreted attractin					66
Extracellular	SP_O75636	Ficolin 3	5	44	2.4	21.4	
Extracellular	SP_O95479	GDH/6PGL endoplasmic bifunctional protein		47	2.3	18.3	
Extracellular	SP_O95967	Fibulin-4			1.6	16.5	
Extracellular	SP_P00739	Hapotglobin-related protein					94
Extracellular	SP_P04278	Sex hormone-binding globulin	15	77	2.4	30	
Extracellular	SP_P04279	Semenogelin I protein		172	2.3	32	
Extracellular	SP_P07996	Thrombospondin 1		73	1.9	15	
Extracellular	SP_P08123	Collagen α-2 (I) chain		54			59
Extracellular	SP_P08670	Vimentin		62	2.3	28	
Extracellular	SP_P09486	Osteonectin		36	1.8	20	
Extracellular	SP_P22105	Tenascin-X	69				238
Extracellular	SP_P23142	Fibulin-1		38	1.9	9	
Extracellular	SP_P24043	Laminin α-2 chain		123	2.3	53	69
Extracellular	SP_P27658	Collagen, alpha 1 (VIII) chain		52	1.1	13	
Extracellular	SP_P31025	von Ebner's gland protein					75
Extracellular	SP_P43251	Biotinidase		70	2.3	14.1	
Extracellular	SP_P51884	Lumican		69	2.2	26	
Extracellular	SP_P53778	Mitogen-activated protein kinase 12		53	1.8	22	
Extracellular	SP_P81605	Dermcidin					95
Extracellular	SP_Q02383	Semenogelin II		63	1.4	24	

Table 1. Continued

Cellular designation	Annotation	Protein name	Spot No.	MALDI-MS			LC-MS/MS Mascot
				Mascot	Profound	Seq Cov	
Extracellular	SP_Q03692	Collagen α 1 (X) chain		51	1.9	19.8	
Extracellular	SP_Q15113	Procollagen C-proteinase enhancer protein	48	72	1.6	25.8	72
Extracellular	SP_Q15149	Plectin 1		65	2	4	
Extracellular	SP_Q15582	TGF-β-induced protein IG-H3					142
Extracellular	SP_Q16610	Extracellular matrix protein 1		38	1.9	18.5	76
Extracellular	SP_Q96KN2	Glutamate-carboxypeptidase-like protein 2					115
Category 3							
Vesicular/secreted	SP_O00462	β-Mannosidase		42	1.8	13.8	
Vesicular/secreted	SP_O00469	Procollagen-lysine 2-oxoglutarate 5-dioxygenase 2		70	1.1	11	
Vesicular/secreted	SP_O00750	Phosphatidylinsitol 3-kinase C2, β polypeptide		64	1.1	6	
Vesicular/secreted	SP_P01225	Follitropin β chain		59	1.5	41	
Vesicular/secreted	SP_P01270	Parathyroid hormone		55	1.3	38	
Vesicular/secreted	SP_P04040	Catalase		47	2.4	23	228
Vesicular/secreted	SP_P04066	Tissue α-L fucosidase			3	14.8	
Vesicular/secreted	SP_P05231	Interleukin-6		51	2.4	42	
Vesicular/secreted	SP_P07237	Protein disulfide isomerase		89	2.1	25	
Vesicular/secreted	SP_P07339	Cathepsin D	45				103
Vesicular/secreted	SP_P07711	Cathepsin L		41	2.2	26.7	
Vesicular/secreted	SP_P09172	Dopamine β-monooxygenase		47	2.3	19.7	91
Vesicular/secreted	SP_P11021	78 kDa Glucose-regulated protein		60	1.5	17	
Vesicular/secreted	SP_P13232	Interleukin-7		53	1.5	33	
Vesicular/secreted	SP_P14625	Endoplasmin		132	2.4	13	
Vesicular/secreted	SP_P15144	Aminopeptidase N					77
Vesicular/secreted	SP_P25774	Cathepsin S		64	1.7	22	
Vesicular/secreted	SP_P37268	Farnesyl-diphosphate farnesyltransferase		58	1.7	25	
Vesicular/secreted	SP_P53634	Depeptidyl peptidase I (cathepsin C)		31	2	27	
Vesicular/secreted	SP_Q13439	Golgi autoantigen, golgin subfamily member A4		79	2.3	12	
Vesicular/secreted	SP_Q92820	γ-glutamyl hydrolase	14	58	1.1		
Vesicular/secreted	SP_Q92896	E-selectin ligand 1		62			
Vesicular/secreted	SP_Q99518	Dimethylaniline monooxygenase 2		62	1.2	24	
Vesicular/secreted	SP_Q9UHG3	Prenylcysteine lyase		51	2.2	17	
Vesicular/secreted	SP_Q9Y4L1	Hypoxia-upregulated protein 1, 150 kDa oxygen-regulated		57	1.5	14	
Category 4							
Cell membrane	gi_13376091	Emilin-like protein EndoGlyx-1					70
Cell membrane	SP_O15269	Serine palmitoyltransferase 1		62	1.6	32	
Cell membrane	SP_O43491	Band 4.1-like protein 2		74	1.4	10	
Cell membrane	SP_P75330	Hyaluronan-mediated motility receptor		51	1.8	15	
Cell membrane	SP_P08571	Monocyte differentiation antigen CD14			1.8	3.4	92
Cell membrane	SP_P08575	Leukocyte common antigen		64	1	7	
Cell membrane	SP_P10316	HLA class I histocompatibility antigen AW69, α chain	33				65
Cell membrane	SP_P18206	Vinculin	2	52			195
Cell membrane	SP_P18428	Lipopolysaccharide-binding protein		60	2.3	23	
Cell membrane	SP_P23634	Calcium-transporting ATPase, isoform 4		54	1.7	12	
Cell membrane	SP_P27216	Annexin A13		59	2.1	22	
Cell membrane	SP_P33151	Vascular endothelial cadherin		55	2.1	13	
Cell membrane	SP_P35916	Vascular endothelial growth factor receptor 3		68	1.7	11.2	
Cell membrane	SP_P40123	Adenylyl cyclase-associated protein 2		36	2.1	16	
Cell membrane	SP_Q04609	Prostrate-specific membrane antigen (folate hydrolase 1)		68	1.7	12	

Table 1. Continued

Cellular designation	Annotation	Protein name	Spot No.	MALDI-MS			LC-MS/MS Mascot
				Mascot	Profound	Seq Cov	
Cell membrane	SP_Q04656	Copper-transporting ATPase 1	56	41	1.9	11	
Cell membrane	SP_Q07075	Glutamyl aminopeptidase EAP		64	1.6	13.6	
Cell membrane	SP_Q13813	Spectrin, non-erythroid α chain		80	2.3	12	
Cell membrane	SP_Q93034	Vasopressin-activated calcium-mobilizing receptor		84	1.3	18	
Cell membrane	SP_Q9H254	Spectrin, non-erythroid β chain		66	1.8	7	
Cell membrane	SP_Q9YSI	Protocadherin α-11		64	1.9	9	
Category 5							
Intracellular	SP_P07476	Involucrin		77	1.8	10	
Intracellular	SP_O14829	Serine/threonine protein phosphatase with EF-hands-1		51	1.4	19.3	
Intracellular	SP_O14920	Inhibitor of nuclear factor kappa B kinase, β subunit		54	1.1	13	
Intracellular	SP_O43237	Dynein light intermediate chain 2		42	2.2	24	
Intracellular	SP_O43781	Dual-specificity tyrosine-phosphorylation regulated kinase 3		66	1.4	13	
Intracellular	SP_O43903	Growth-arrest-specific protein 2		70	2	25	
Intracellular	SP_O60610	Diaphanous protein homolog 1		65	1.4	10	
Intracellular	SP_O60225	SSX5 protein		77	1.4	24	
Intracellular	SP_O60262	Guanine nucleotide-binding protein G(I) γ-7 subunit		52	1.9	65	
Intracellular	SP_O60858	Leukemia-associated protein 5		52	2.3	19	
Intracellular	SP_O75533	Splicing factor 3B subunit 1		70	1.8	12.3	
Intracellular	SP_O76074	cGMP-specific 3',5'-cyclic phosphodiesterase		58	1.3	16	
Intracellular	SP_O95613	Kendrin		73	2.2	9	
Intracellular	SP_P00915	Carbonic anhydrase I		125	2.4	52	
Intracellular	SP_P00918	Carbonic anhydrase II		56	2.2	23	
Intracellular	SP_P02023	Hemoglobin beta chain		49	2.3	19	
Intracellular	SP_P02568	α-Actin, skeletal muscle		65	2.4	26	
Intracellular	SP_P02570	β-Actin, cytoplasmic 1	73	48	2.3	36	192
Intracellular	SP_P02571	γ-Actin, cytoplasmic 2	73	40	2.4	29	132
Intracellular	SP_P02795	Metallothionein II		57	1.5	70	
Intracellular	SP_P34062	Proteasome subunit alpha type 6		47	1.7	24.8	
Intracellular	SP_P03996	α-Actin, aortic smooth muscle					76
Intracellular	SP_P05109	Calgranulin A		54	1.9	70	
Intracellular	SP_P05165	Propionyl-CoA carboxylase, α chain		75	1.3	25	
Intracellular	SP_P06276	Cholinesterase					130
Intracellular	SP_P06576	ATP synthase, β chain		53			
Intracellular	SP_P06732	Creatine kinase M		52	0.8	12	
Intracellular	SP_P06753	Tropomyosin α-3 chain, skeletal muscle		54	2.4	14.4	
Intracellular	SP_P07195	L-Lactate dehydrogenase B chain	25				98
Intracellular	SP_P07196	Neurofilament triplet L protein		66	2.8	12	
Intracellular	SP_P07226	Tropomyosin α-4 chain, fibroblast type		82	1.6	23	
Intracellular	SP_P07451	Carbonic anhydrase III		40	1.9	23.8	
Intracellular	SP_P07738	Bisphosphoglycerate mutase		87	2.3	51	
Intracellular	SP_P07951	Tropomyosin β chain		51	2.4	21.8	
Intracellular	SP_P08758	Annexin V		63	2.1	32	
Intracellular	SP_P09104	γ-Enolase		63	1.7	15	
Intracellular	SP_P09455	Cellular retinol-binding protein		55	16	25	
Intracellular	SP_P09493	Tropomyosin α-1 chain, skeletal muscle		58	2.3	22	
Intracellular	SP_P10644	cAMP-dependent protein kinase type I-α regulated chain		38	1.8	17.6	
Intracellular	SP_P12882	Myosin heavy chain, skeletal muscle		72	1.3	4	
Intracellular	SP_P12883	Myosin heavy chain, cardiac muscle beta isoform		69	2.4	6	

Table 1. Continued

Cellular designation	Annotation	Protein name	Spot No.	MALDI-MS			LC-MS/MS Mascot
				Mascot	Profound	Seq Cov	
Intracellular	SP_P12955	Proline dipeptidase	27	39	1.8	23.9	78
Intracellular	SP_P13533	Myosin heavy chain, cardiac muscle alpha isoform		87	2.4	8	
Intracellular	SP_P13796	L-plastin		66	1.7	14	
Intracellular	SP_P13797	T-plastin	22	56	2	17.1	
Intracellular	SP_P14136	Glial fibrillary acidic protein		50	1.2	17	
Intracellular	SP_P14619	cGMP-dependent protein kinase 1, β isoenzyme		61	2.4	17.3	
Intracellular	SP_P16383	GC-rich sequence DNA-binding factor		59	1	14	
Intracellular	SP_P16930	Fumarylacetoacetase	43				114
Intracellular	SP_P17480	Nucleolar transcription factor 1		83	0.9	13	
Intracellular	SP_P17661	Desmin	74	53	1.9	19.1	
Intracellular	SP_P19484	Transcription factor EB	42				52
Intracellular	SP_P20700	Lamin B1		68	1.3	15	
Intracellular	SP_P21266	Glutathione *S*-transferase Mu3		73	2.1	40	
Intracellular	SP_P29992	Guanine nucleotide-binding protein G(Y)α-subunit		64	2.4	17.8	
Intracellular	SP_P26441	Ciliary neurotrophic factor		76	2.3	40	
Intracellular	SP_P29312	14-3-3 protein ζ/δ		56	2.2	40	
Intracellular	SP_P31151	S100 calcium-binding protein A7		76	2.1	43	
Intracellular	SP_P31947	14-3-3 protein σ		64	1	31	
Intracellular	SP_P31350	Ribonucleotide-diphosphate reductase M2 chain		50	1.1	20	
Intracellular	SP_P32119	Peroxiredoxin 2			1.8	19.7	98
Intracellular	SP_P34932	Heat shock 70 kDa protein		54	1.6	12	
Intracellular	SP_P35237	Placental thrombin inhibitor		54	1.6	9.6	
Intracellular	SP_P35573	Glycogen-debranching enzyme		65	1.3	15	
Intracellular	SP_P35609	α-Actinin 2		61	1.6	13	
Intracellular	SP_P35998	26S protease regulatory subunit 7		44	1.8	16.2	
Intracellular	SP_P40939	78 kDa gastrin-binding protein		52	1.1	20	
Intracellular	SP_P41182	B-cell lymphoma 6 protein		71	1.2	16	
Intracellular	SP_P43034	Platelet-activating factor acetylhydrolase Ib, α subunit		65	1.2	17	
Intracellular	SP_P45974	Ubiquitin carboxyl-terminal hydrolase 5		42	2.2	15	
Intracellular	SP_P46926	Glucosamine-6-phosphate isomerase		69	1.1	27	
Intracellular	SP_P48595	Bomapin (protease inhibitor 10)			1.8	22.7	
Intracellular	SP_P48637	Glutathione synthetase		69	2.3	26	
Intracellular	SP_P48729	Casein kinase I, α isoform		62	1	14	
Intracellular	SP_P49888	Estrogen sulfotransferase			2.2	25.9	
Intracellular	SP_P50148	Guanine nucleotide-binding protein G(q), α subunit		64	2.3	25.8	
Intracellular	SP_P50990	T-complex protein 1, theta subunit		54	2.3	24.1	
Intracellular	SP_P51451	Tyrosine-protein kinase BLK		62	1.9	17	
Intracellular	SP_P52732	Kinesin-related motor protein Eg5		66	1.8	16	
Intracellular	SP_P52272	Heterogeneous ribonuclear protein M		74	1.2	10	
Intracellular	SP_P53396	ATP-citrate (pro-S-)-lyase		60			
Intracellular	SP_P54274	Telomeric repeat-binding factor 1		47	1.9	19	
Intracellular	SP_P55265	Double-stranded RNA-specific adenosine deaminase		56	2.2	16.3	
Intracellular	SP_P58304	Homeobox protein CHX10		56	1.6	24.9	
Intracellular	SP_P80297	Metallothionein-IL		58	2.4	57	
Intracellular	SP_P82094	TATA element modulatory factor		70	2.1	11	
Intracellular	SP_Q02252	Methylmalonate-semialdehyde dehydrogenase		51	1.1	14	
Intracellular	SP_Q02952	A-kinase anchor protein 12		55	2.5	6.6	
Intracellular	SP_Q05193	Dynamin 1		66	1.3	8	

Table 1. Continued

Cellular designation	Annotation	Protein name	Spot No.	MALDI-MS			LC-MS/MS Mascot
				Mascot	Profound	Seq Cov	
Intracellular	SP_Q12874	Splicing factor 3A subunit 3		36	2.4	12.4	
Intracellular	SP_Q13200	26S proteasome subunit non-ATPase regulated subunit 2		56	2.3	12.6	
Intracellular	SP_Q13322	Growth factor receptor-bound protein 10		75	1.3	11	
Intracellular	SP_Q13330	Metastasis-associated protein MTA 1		66	1.7	18	
Intracellular	SP_Q13451	FK506-binding protein 5		60	1.4	19	
Intracellular	SP_Q13573	Nuclear protein SKIP		78	1	18	
Intracellular	SP_Q13610	Periodic tryptophan protein 1 homolog		61	1.1	18	
Intracellular	SP_Q13882	Tyrosine protein kinase 6	28	68	1.7	27	
Intracellular	SP_Q13976	cGMP-dependent protein kinase 1, α isozyme		76	2.4	13.6	
Intracellular	SP_Q14203	Dynactin 1, 150 kDa isoform		65	2.9	1.8	
Intracellular	SP_Q15436	Protein transport protein Sec23A		57	1.4	15	
Intracellular	SP_Q92696	RAB geranylgeranyltransferase α subunit		50	1.8	23.3	
Intracellular	SP_Q92851	Caspase 10		61	1.4	14	
Intracellular	SP_Q99575	Ribonuclease P/MRP protein subunit POP1		68	1.8	18	
Intracellular	SP_Q99661	Mitotic centromere-associated kinesin		61			
Intracellular	SP_Q99856	Dead ringer-like 1 protein	70	53	1.5	14.8	
Intracellular	SP_Q99996	A kinase anchor protein 9		75	2	5	
Intracellular	SP_Q9C0C2	182 kDa tankyrase 1-binding protein		37	1.7	8	
Intracellular	SP_Q9H223	EH domain-containing protein 4		36	2.2	20.7	
Intracellular	SP_Q9ULV0	Myosin Vb (fragment)		72	1.3	14	
Intracellular	SP_Q9Y4E8	Ubiquitin carboxyl-terminal hydrolase 15		60	1.6	1.9	
Intracellular	SP_Q9Y600	Cysteine sulfinic acid decarboxylase		61	2	19.7	
Intracellular	SP_Q9Y6M4	Casein kinase I, γ 3 isoform		70	0.9	21	
Intracellular	SP_Q9Y6K8	Adenylate kinase isoenzyme 5	24	63	1.5	36	
Category 6							
NA	gi_9836652	Brain-selective and mapped . . . CMAP in cystatin cluster					180
NA	gi_12845793	Putative protein (AK010386), *Mus musculus*					52
NA	SP_P28370	Possible global transcription factor SNF2L1		59	1.3	9	
NA	SP_P55854	Ubiquitin-like protein SMT3A		63	2.3	47	
NA	SP_P78395	Melanoma antigen preferentially expressed in tumors		54	1.1	12	
NA	SP_Q13618	Cullin homolog 3		64	1.1	12	
NA	SP_Q8WXK1	Ankyrin repeat and SOCS box-containing protein 15		56	1.6	13	
NA	SP_Q96PD5	Peptidoglycan recognition protein L	21				144
NA	SP_Q9NYH9	Hepatocellular carcinoma-associated antigen 66		68	1.7	13	
NA	SP_Q9UK41	VPS28 homolog		55	1.8	29	
Category 7							
Microbial	SP_P24305	Outer membrane porin protein 32					132

Following removal of redundant annotations (not included), the final yield was 325 unique protein annotations (gene products), listed together with one common protein name. Proteins are listed according to the categories described in Fig. 9. It should be noted that several of the identified proteins have been associated with more than one cellular designation. In such cases, one category was chosen arbitrarily. NA, categorization not applied. Category 7: not depicted in Fig. 9 with one identified bacterial protein. Annotations as well as the information on extra- and intracellular designations derive from the SWISS-PROT (SP_) and NCBI (gi_) databases. For MALDI-MS scores obtained through Profound searches, the peptide sequence coverage (seq cov) is listed in % of the entire protein sequence.

Table 2. Selected human serum proteins with potential value as tissue-specific or disease-associated biomarkers

Annotation	Protein name	Biomarker utility or tissue-specefic expression of protein
SP_O00469	Lysyl-hydroxylase 2	Highly expressed in muscle and pancreas tissues
SP_O14829	Serine/threonine-protein phosphatase 7	Expressed in the retina, retinoblastoma cells, and fetal brain
SP_O60262	Guanine-nucleotide-binding protein G(I) γ-subunit	Expression in pancreatic cancer and pancreatic carcinoma cell lines downregulated
SP_O75330	Hyaluronan-mediated motility receptor	Expressed in normal breast tissue, localized intracellularly in breast cancer cells
SP_O75636	Ficolin-3	Expressed in lung tissue, highly abundant in serum of Lupus erythematosus patients
SP_P08571	Monocyte differentiation antigen CD14	Highly expressed on the cell surface of monocytes
SP_P13796	L-plastin	Specifically expressed in the spleen and lymph node-containing organs
SP_P14136	Glial fibrillary acidic protein	Specifically expressed by cells of astroglial lineage in the brain
SP_P18206	Vinculin	Specifically expressed in muscle tissue
SP_P27216	Annexin A13	Specifically expressed in intestine tissue
SP_P29312	14-3-3 protein ζ/δ	Mostly expressed in nerve terminals of neurons
SP_P33151	Vascular endothelial-cadherin	Mostly expressed in endothelial tissues and the brain
SP_P35609	α-Actinin 2	Specifically expressed in skeletal and cardiac muscle tissues
SP_P51451	Tyrosine-protein kinase BLK	Specifically expressed in B-lymphocytes
SP_P78395	Melanoma antigen preferentially expressed in tumors	cell surface tumor antigen
SP_Q13882	Tyrosine-protein kinase 6	Highly expressed in colon tissue, expressed in some breast tumors but not in normal breast
SP_Q14203	Dynactin 1	Specifically expressed in the brain
SP_Q9H223	EH-domain containing protein 4	Highly expressed in pancreas and heart tissues
SP_Q9Y6K8	Adenylate kinase isoenzyme 5	Specifically expressed in the brain
SP_P02741	C-reactive protein	Acute phase reactant, biomarker for myocardial infarction and coronary heart disease risk
SP_P04066	Tissue-α-L-fucosidase	Potential biomarker for ovarian cancer [42]
SP_P05231	Interleukin-6	Acute phase reactant, proinflammatory cytokine
SP_P06732	Creatin kinase M	Biomarker for irreversible cell damage and myocardial infarction
SP_P07195	L-Lactate dehydrogenase B-chain	Biomarker for irreversible cell damage and liver malfunction
SP_P07339	Cathepsin D	Potential tumor progression and metastasis biomarker [43]
SP_P07711	Cathepsin L	Potential tumor progression biomarker [44]
SP_P20742	Pregnancy zone protein	Potential biomarker for ovarian cancer [42]
SP_Q04609	Prostrate-specific membrane antigen	Highly expressed in prostate epithelium, membrane-bound form is prostate cancer biomarker

SWISS-PROT annotations with one common protein name (as in Table 1) are listed. Unless further specified [42–44], information on tissue specificity and biomarker utility is referenced in the text or was retrieved from protein descriptions in the SWISS-PROT database.