A Proteomic Analysis of Human Hemodialysis Fluid*

Henrik Molina‡§, Jakob Bunkenborg‡§¶, G. Hanumanthu Reddy‖, Babylakshmi Muthusamy‖, Paul J. Scheel**, and Akhilesh Pandey‡ ‡‡

The vascular compartment is an easily accessible compartment that provides an opportunity to measure analytes for diagnostic, prognostic, or therapeutic indications. Both serum and plasma have been analyzed extensively by proteomic approaches in an effort to catalog all proteins and polypeptides. Limitations of such approaches in obtaining a comprehensive catalog of proteins include the fact that a handful of proteins constitute over 90% of plasma protein content and that the renal glomeruli filter out proteins and polypeptides that are smaller than 66 kDa from blood. We chose to study hemodialysis fluid because it contains a higher concentration of small proteins and polypeptides and is also simultaneously depleted of the most abundant proteins present in blood. Using gel electrophoresis in combination with LC-MS/MS, we identified 292 proteins of which greater than 70% had not been previously identified from serum or plasma. More than half of the proteins identified from the hemodialysis fluid were smaller than 40 kDa. We also found 50 N-terminally acetylated peptides that allowed us to unambiguously map the N termini of mature forms of the corresponding proteins. Several identified proteins, including cytokines, were only present as predicted transcripts in data bases and thus represent novel proteins. The proteins identified in this study could serve as biomarkers in serum using more sensitive methods such as ELISA-specific antibodies. Molecular & Cellular Proteomics 4:637–650, 2005.

A comprehensive analysis of human serum and plasma has proven to be difficult, especially for low molecular weight and low abundance proteins, because of the wide range of concentrations with the 10 most abundant proteins constituting almost 90% of the serum proteome by mass (1). The dynamic range issue is exacerbated for low molecular weight species because the kidneys filter away molecules with molecular mass of less than 66 kDa (2). The abundant proteins such as albumin, immunoglobulins, and transferrin hamper identification of less abundant proteins. Although several methods including ultracentrifugation (3), immunodepletion, solvent extraction/precipitation (4), and size exclusion chromatography (5) have been tried for removal of abundant components, it is still difficult to completely eliminate the interference from residual amounts of these abundant proteins.

Hemodialysis fluid has previously been used as a source of polypeptides (6) due to its higher concentration of low molecular weight components, but no comprehensive list of constituents present in the hemofiltrate has yet been published. We chose to analyze hemodialysis fluid because it is greatly reduced in the protein content, from 70 g/liter in the plasma to ~70 mg/liter. This is because the filtration cutoff used during hemodialysis results in selective depletion of proteins greater than ~60 kDa. It has been shown that the concentration of albumin in the hemodialysis fluid is 5,000-fold lower compared with its normal concentration in serum (7).

Our strategy involved one-dimensional gel electrophoresis separation and in-gel digestion of proteins followed by LC-MS/MS for identifying proteins in the hemodialysis fluid. Using this approach, we identified 292 proteins of which 205 had never been previously reported in the serum or plasma. Western blot analysis of a subset of these proteins revealed that they were also present in normal serum indicating that the sensitivity of detection might be the major reason why the majority of these proteins have never been identified previously in serum or plasma. We were also able to identify the N termini of a number of proteins based on peptides sequences that were acetylated at their N termini. A number of semitryptic peptides that were identified in this study were most likely derived from in vivo proteolysis. We were also able to identify a number of novel proteins including several cytokines in this analysis. The lack of a major overlap between the list of proteins identified in this study and previously reported proteins in serum or plasma reflects the difficulty of identifying these components using current proteomic methods. As demonstrated by our study, it is likely that more sensitive methods such as Western blotting or ELISA will be able to detect these proteins in serum or plasma.

EXPERIMENTAL PROCEDURES
Sample Collection and Preparation and Reagents—The hemodialysis fluid was obtained from a 68-year-old white male with acute renal
failure following a coronary artery bypass surgery. The patient was not known to have any other diseases. Vascular access was obtained utilizing a dual lumen catheter in the femoral position. Continuous venovenous hemodialysis was performed with a Prisma® dialysis machine (Cobe Renal Intensive Care, Lakewood, CO). The dialysis membrane was an acrylonitrile and sodium methallyl sulfonate copolymer, AN69® (Cobe Renal Intensive Care). Sterile dialysate was prepared in 5.0-liter aliquots (140 megcl/liter sodium, 111 megcl/liter chloride, 3.5meq/liter calcium, 3 meq/liter lactate, 1 meq/liter magnesium, 32 meq/liter bicarbonate, 2.0 meq/liter potassium), and the dialysis was carried out with a blood flow of 180 ml/min and a dialysate flow of 1.0 liter/hr. The spent dialysate was collected in 5.0-liter aliquots and analyzed after concentrating using a 3-kDa-cutoff filter (Centricron YM3, Millipore). The protein content in the hemodialysis fluid was measured using a modified Lowry protein assay kit (Bio-Rad). The concentration of hemodialysis fluid was run on precast NuPage 4–12% bis-tris*1 and 10–20% Tricine minigels (Invitrogen). The gels were silver-stained (8), and between 20 and 25 visible protein bands were excised. Protein gel bands were reduced with dithiothreitol (Fluka, Buch, Switzerland) and alkylated with iodoacetamide (Sigma) before digestion with trypsin (Promega, Madison, WI) in 100 mM NH₄HCO₃ (Fluka). Solvents for liquid chromatography included heptafluorobutyric acid (Sigma), glacial acetic acid (Fisher Scientific), and HPLC-grade acetonitrile (U. T. Baker Inc.).

Trypsin Digestion and LC-MS/MS Analysis—The excised gel slices were digested with trypsin as follows. The gel bands were washed twice in 0.1 M NH₄HCO₃, washed twice in 50% acetonitrile, and subsequently cut into 2 × 2-mm pieces. The gel pieces were shrunk with 100% acetonitrile, and proteins were reduced by addition of 0.1 M dithiothreitol followed by an incubation step at 56 °C for 45 min. The washing procedure described above was repeated, and proteins were alkylated by adding 55 mM iodoacetamide and incubating for 30 min at room temperature in the dark. After an additional wash and shrinkage, 10 ng/μl trypsin in 0.1 M NH₄HCO₃ sufficient to cover the gel pieces was added followed by an incubation on ice for 20 min. When the gel pieces were completely rehydrated, any excess trypsin solution was removed and replaced by 0.1 M NH₄HCO₃, and samples were incubated overnight at 37 °C. The digestion was stopped by adding 10 μl of glacial acetic acid, and the supernatant containing the trypptic peptides was harvested. An extraction step was carried out to recover the peptides from the gel slices by adding 50% acetonitrile and incubating at room temperature for 30 min. The supernatant was harvested again and pooled. The pooled peptide extracts were dried down to ~10 μl and subjected to LC-MS/MS analysis as follows. Samples were injected onto a 5-cm C18 trap column (inner diameter, 75 μm) packed with YMC ODS-A 5–15-μm beads (Kanematsu USA Inc., New York, NY) using an autosampler (1100 microwell plate autosampler, Agilent Technologies, Palo Alto, CA). The peptides were eluted from the trap column onto an analytical 10-cm C18 column (inner diameter, 75 μm) packed with Vydac MS218 5-μm beads (Vydac, Columbia, MD) with a gradient increasing from 10% solvent B, 90% solvent A (solvent A: 0.4% acetic acid, 0.005% heptfluorobutyric acid; solvent B: 90% acetonitrile, 0.4% acetic acid, 0.005% heptfluorobutyric acid) to 45% solvent B, 55% solvent A in 30 min. A flow of 4 μl/min during loading and 300 nl/min during elution was delivered by a nanoflow pump (Agilent Technologies 1100 nano-OP). The LC setup was connected to either a quadrupole-time-of-flight mass spectrometer (QTOF API-US, Micromass, Manchester, UK) or an ion trap mass spectrometer (LC/MSD Trap XCT, Agilent Technologies) using nanoelectrospray sources from Proxeon (Odense, Denmark).

Western Blotting—For Western blotting experiments, serum from a healthy person was first depleted of albumin, IgG, IgA, transferrin, haptoglobin, and antitrypsin using an Agilent Technologies multiple affinity removal kit. 20 μl of serum were diluted to 100 μl in Buffer A (Buffer A and Buffer B as supplied in the multiple affinity removal kit by Agilent Technologies) prior to loading onto the column at 250 μl/min in Buffer A. The elution from the column was monitored at 280 nm, and the depleted serum was collected in the interval from 2 to 4 min. Cleaning of the column was carried out by eluting the bound proteins as follows. 10 min after the injection, Buffer A was exchanged with Buffer B, and the flow rate was simultaneously increased to 1,000 μl/min. Following this, Buffer B at 1,000 μl/min was continued for 18 min after which the conditions were returned to the initial loading conditions. After conditioning the column under loading conditions for 5 min, the system was ready for a new depletion process.

Depleted serum was resolved by SDS-PAGE, and the proteins were transferred onto a nitrocellulose membrane. The membrane was blocked at 4 °C overnight with 5% BSA in phosphate-buffered saline containing 0.1% Tween 20. The membrane was incubated with the relevant antibodies for 2 h, washed, and incubated with horseradish peroxidase-conjugated secondary antibody for 1 h. The proteins were visualized using enhanced chemiluminescence detection according to the manufacturer’s instructions (Amersham Biosciences). The sources of primary antibodies were as follows: Cathepsin D, connective tissue growth factor, Galectin 3, and Lipocalin 2 (R&D Systems, Inc, Minneapolis, MN); Nucleophosmin 1 (Zymed Laboratories Inc.); α-defensins 1–3 (BD Biosciences); Cathepsin H (Serotec, Oxford, UK); and Cofilin 2 (Upstate Biotechnology, Lake Placid, NY). Conjugated secondary anti-mouse and anti-rabbit were from Amersham Biosciences, and anti-goat antibody was from Santa Cruz Biotechnology (Santa Cruz, CA).

Data Base Search and Analysis—The mass spectrometry data files from individual LC-MS/MS experiments were merged and then searched against the RefSeq data base (human: 27,975 entries; July 28, 2004) using Mascot (Matrix Sciences Ltd., London, UK). The search parameters were as follows: mass accuracy of the monoisotopic precursor and peptide fragments was set to 1.5 and 0.5 Da, respectively, for the data acquired on the ion trap mass spectrometer and 0.2 and 0.2 Da, respectively, for the data acquired on the quadrupole-time-of-flight mass spectrometer. The following variable modifications were permitted: oxidation of methionine, histidine, and tryptophan residues; N-terminal acetylation of proteins; and cyclization of N-terminal glutamine. Two missed tryptic cleavages were allowed. Additional searches using data acquired on the quadrupole time-of-flight mass spectrometer were performed with the following criteria: 1) semitryptic constraints, 2) N-acetylhexosamine modification of asparagine residues, and 3) hydroxylation of proline residues. For validation purposes, the retrieved peptide sequences were divided into three groups as follows: 1) peptides with a low score (quadrupole-time-of-flight data, <25; ion trap data, <30), 2) peptides with intermediate to high scores (quadrupole-time-of-flight data, >25; ion trap data, >30), and 3) proteins identified with only one intermediate or high scoring peptide. All of the low scoring peptides were discarded. All intermediate and high scoring peptides as well as single peptide hits used to identify proteins were manually validated if the following criteria were met: 1) several consecutive y-ions although absence of y-ions after proline and glycine, 2) the existence of lower a- and b-ions, 3) none or few unassigned fragments ions, and 4) a charge state of the precursor ion and fragment ions that are in accordance with basic amino acids in the assigned peptide sequence. Supplemental Fig. 6 pro-

---

*1 The abbreviations used are: bis-tris, 2-[bis(2-hydroxyethyl)-amino]-2-(hydroxymethyl)propane-1,3-diol; Tricine, N-[2-hydroxy-1,1-bis(hydroxyethyl)glycine; FABP, fatty acid-binding protein; HIV-1, human immunodeficiency virus, type 1.
Our strategy to characterize the hemodialysis fluid proteome by mass spectrometry using a gel LC approach was as follows. The hemodialysis fluid was first desalted and concentrated using a 3-kDa-cutoff filter. For fractionating the proteins and polypeptides in the hemodialysis fluid, the desalted sample was resolved by SDS-PAGE and silver-stained, and the bands were excised, digested with trypsin, and analyzed by LC-MS/MS. Supplemental Table 1 lists all the proteins that were identified in this study along with the number of peptides that matched each protein, molecular weight, and assignments of the approved Human Genome Organization (HUGO) gene symbols, wherever available. Supplemental Table 1 provides a count of how many gene products were actually identified from this sample as the count on the basis of proteins can be misleading because of the presence of a large number of protein isoforms.

**RESULTS AND DISCUSSION**

Of the proteins identified in this study, 85% were smaller than 60 kDa, and more than half were <30 kDa. This is expected because of the hemodialysis process in which blood is filtered through a membrane that has a molecular mass cutoff of ~60 kDa, which should only allow proteins less than 60 kDa into the hemodialysis fluid. Fig. 1 shows a histogram of the number of proteins identified plotted against the corresponding molecular masses. Most of the proteins are found between 15 and 30 kDa. Superimposed on the histogram is a similar plot from the combined studies by Anderson et al. (9) and Chan et al. (10). Those two studies resulted in more than 2,100 proteins of which ~50% were below 65 kDa, and only 16% were below 30 kDa. Another distinguishing feature of the hemodialysis fluid data set is the absence of the 40–80-kDa hump that is conspicuous in the serum/plasma proteome data set. A subset of the serum/plasma proteins were identified by Tirumalai et al. (3), who used serum that had been filtered through a 30-kDa filter prior to analysis. In terms of the number of proteins identified, this data set (340 proteins) is comparable to our hemodialysis data set (292 proteins), although significant differences are observed. Comparing the molecular weight distribution of the proteins identified in these two studies, it is clear that a considerably greater fraction of the proteins found in the hemodialysis fluid (56% of proteins with molecular mass < 30 kDa) had a lower molecular weight than in the analysis by Tirumalai et al. (18% of proteins with molecular mass < 30 kDa).

We compared the known serum/plasma proteins with those identified in this study to determine the degree of overlap based on the genes that encoded the expressed proteins and their isoforms. We first mapped all the protein entries to HUGO-approved gene symbols wherever possible (www.gene.ucl.ac.uk/nomenclature), which allowed us to eliminate redundancy and facilitated this comparison. Although the serum/plasma catalog contained more than 7 times as many protein entries, it was remarkable that 205 of 292 proteins found in the hemodialysis fluid were not reported in the serum plasma proteome. Supplemental Fig. 1 shows a distribution of proteins identified in this study according to their molecular function. Table I lists the proteins that were identified from hemodialysis fluid but not previously reported in serum/plasma. Not surprisingly, several of these proteins are of low molecular mass, validating our use of hemodialysis fluid to identify smaller proteins and polypeptides. Fig. 2 shows a Venn diagram indicating the relative numbers of overlapping and non-overlapping gene products.

It can be argued that this difference observed between serum and hemodialysis fluid is attributable to the underlying condition in the patient undergoing dialysis. In that case, the proteins specific to our list of hemodialysis fluid would not be observed in serum from normal dialysis. It should be noted that the components that are normally filtered away by the kidneys are excreted in the urine. Therefore, we also compared our results with two proteomic analyses of the human urinary proteome (11, 12) with a total of 358 proteins. Of the proteins specifically found in hemodialysis fluid, 29 of those have been identified from urine of normal individuals. This indicates that a significant fraction of the proteins are indeed found in normal serum/plasma but have not been identified earlier because of the technical limitations mentioned earlier.

**Western Blot Analysis to Confirm Expression in Normal Serum**

To test whether the expression of the proteins found in hemodialysis fluid was also detectable in normal serum, we
## Proteomic Analysis of Human Hemodialysis Fluid

### TABLE I

| Protein name                                                                 | Accession number | Gene symbol | Molecular (Daltons) mass |
|-------------------------------------------------------------------------------|------------------|-------------|--------------------------|
| 1. Thymosin, β 10                                                             | NP_066926.1      | TMSB10      | 5,023                    |
| 2. Thymosin-like 2                                                             | NP_877594.1      | TMSL2       | 5,036                    |
| 3. Thymosin-like 6                                                             | NP_852093.1      | TMSL6       | 5,039                    |
| 4. Thymosin, β 4                                                              | NP_066932.1      | TMSB4X      | 5,050                    |
| 5. Thymosin-like 4                                                             | NP_877595.1      | TMSL4       | 5,109                    |
| 6. Metallothionein 1H                                                          | NP_005942.1      | MT1H        | 7,175                    |
| 7. Metallothionein 1G                                                          | NP_005941.1      | MT1G        | 7,205                    |
| 8. Metallothionein 1F                                                          | NP_005940.1      | MT1F        | 7,221                    |
| 9. Metallothionein 1A                                                          | NP_789846.1      | MT1A        | 7,268                    |
| 10. Metallothionein 1K                                                         | NP_005937.1      | MT1K        | 7,272                    |
| 11. Metallothionein 1B                                                         | NP_005938.1      | MT1B        | 7,307                    |
| 12. Antioxidant protein 1                                                       | NP_004036.1      | ATOX1       | 7,997                    |
| 13. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 14. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 15. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 16. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 17. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 18. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 19. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 20. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 21. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 22. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 23. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 24. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 25. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 26. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 27. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 28. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 29. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 30. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 31. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 32. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 33. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 34. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 35. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 36. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 37. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 38. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 39. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 40. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 41. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 42. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 43. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 44. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 45. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 46. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 47. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 48. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 49. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 50. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 51. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 52. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 53. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 54. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 55. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |
| 56. Metallothionein 1H                                                         | NP_006820.1      | C10orf16    | 7,850                    |

Previously identified in urine
| Protein name | Accession number | Gene symbol | Molecular (Daltons) mass |
|--------------|------------------|-------------|------------------------|
| Niemann-Pick disease, type C2 precursor | NP_006423.1 | NPC2 | 16,559 |
| Similar to Ig κ variable region | XP_376073 | LOC400967 | 16,693 |
| Ubiquitin-conjugating enzyme E2D 2 isoform 1 | NP_003330.1 | UBE2D2 | 16,724 |
| Cystatin M precursor | NP_001314.1 | CST6 | 16,785 |
| Gliadin maturation factor, β | NP_004115.1 | GMFB | 16,874 |
| Ubiquitin-conjugating enzyme E2N | NP_003339.1 | UBE2N | 17,127 |
| Pancreatic ribonuclease precursor | NP_002924.1 | RNASE1 | 17,633 |
| Ubiquitin-conjugating enzyme E2L 2 isoform 1 | NP_003338.1 | UBE2L3 | 17,850 |
| Similar to peptidyl-Pro cis-trans isomerase | XP_376073 | LOC400967 | 16,693 |
| Ubiquitin-conjugating enzyme E2N | NP_002945.1 | RPS27A | 17,953 |
| Pancreatic ribonuclease precursor | NP_002900.2 | REG1A | 18,199 |
| Cystatin M precursor | NP_001314.1 | CST6 | 16,785 |
| Glia maturation factor, /H9252 | NP_003330.1 | UBE2D2 | 16,724 |
| Ubiquitin-conjugating enzyme E2D 2 isoform 1 | NP_003338.1 | UBE2L3 | 17,850 |
| Similar to peptidyl-Pro cis-trans isomerase | XP_292085.1 | LOC341457 | 17,901 |
| Ubiquitin-conjugating enzyme E2L 3 isoform 1 | NP_003339.1 | UBE2L3 | 17,850 |
| Similar to peptidyl-Pro cis-trans isomerase | XP_376073 | LOC400967 | 16,693 |
| Ubiquitin-conjugating enzyme E2N | NP_002945.1 | RPS27A | 17,953 |
| Pancreatic ribonuclease precursor | NP_002900.2 | REG1A | 18,199 |
| Ubiquitin-conjugating enzyme E2L 2 isoform 1 | NP_003338.1 | UBE2L3 | 17,850 |
| Similar to peptidyl-Pro cis-trans isomerase | XP_292085.1 | LOC341457 | 17,901 |
| Ubiquitin-conjugating enzyme E2N | NP_002945.1 | RPS27A | 17,953 |
| Pancreatic ribonuclease precursor | NP_002900.2 | REG1A | 18,199 |
| Ubiquitin-conjugating enzyme E2L 2 isoform 1 | NP_003338.1 | UBE2L3 | 17,850 |
| Similar to peptidyl-Pro cis-trans isomerase | XP_292085.1 | LOC341457 | 17,901 |
| Ubiquitin-conjugating enzyme E2N | NP_002945.1 | RPS27A | 17,953 |
| Pancreatic ribonuclease precursor | NP_002900.2 | REG1A | 18,199 |
| Ubiquitin-conjugating enzyme E2L 2 isoform 1 | NP_003338.1 | UBE2L3 | 17,850 |
| Similar to peptidyl-Pro cis-trans isomerase | XP_292085.1 | LOC341457 | 17,901 |
| Ubiquitin-conjugating enzyme E2N | NP_002945.1 | RPS27A | 17,953 |
| Pancreatic ribonuclease precursor | NP_002900.2 | REG1A | 18,199 |
| Ubiquitin-conjugating enzyme E2L 2 isoform 1 | NP_003338.1 | UBE2L3 | 17,850 |
| Similar to peptidyl-Pro cis-trans isomerase | XP_292085.1 | LOC341457 | 17,901 |
| Ubiquitin-conjugating enzyme E2N | NP_002945.1 | RPS27A | 17,953 |
| Pancreatic ribonuclease precursor | NP_002900.2 | REG1A | 18,199 |
| Ubiquitin-conjugating enzyme E2L 2 isoform 1 | NP_003338.1 | UBE2L3 | 17,850 |
| Similar to peptidyl-Pro cis-trans isomerase | XP_292085.1 | LOC341457 | 17,901 |
| Ubiquitin-conjugating enzyme E2N | NP_002945.1 | RPS27A | 17,953 |
| Pancreatic ribonuclease precursor | NP_002900.2 | REG1A | 18,199 |
| Ubiquitin-conjugating enzyme E2L 2 isoform 1 | NP_003338.1 | UBE2L3 | 17,850 |
| Similar to peptidyl-Pro cis-trans isomerase | XP_292085.1 | LOC341457 | 17,901 |
| Ubiquitin-conjugating enzyme E2N | NP_002945.1 | RPS27A | 17,953 |
| Pancreatic ribonuclease precursor | NP_002900.2 | REG1A | 18,199 |
| Ubiquitin-conjugating enzyme E2L 2 isoform 1 | NP_003338.1 | UBE2L3 | 17,850 |
| Similar to peptidyl-Pro cis-trans isomerase | XP_292085.1 | LOC341457 | 17,901 |
TABLE I—continued

| Protein name                                                                 | Accession number | Gene symbol | Molecular (Daltons) mass | Comment                                      |
|------------------------------------------------------------------------------|------------------|-------------|--------------------------|----------------------------------------------|
| Triosephosphate isomerase 1                                                 | NP_000356.1      | TPI1        | 26,938                   | Previously identified in urine               |
| Arginine-rich, mutated in early stage tumors                                 | NP_006001.2      | ARG1        | 27,402                   |                                               |
| 6-Phosphogluconolactonase                                                   | NP_036220.1      | TIGR        | 27,815                   |                                               |
| Electron-transfer-flavoprotein, β polypeptide                               | NP_01976.1       | ETFB        | 27,826                   |                                               |
| Insulin-like growth factor-binding protein 4                                 | NP_01543.1       | IGFBP4      | 27,923                   |                                               |
| Hypothetical protein BCO01573                                                | NP_620164.1      | LOC134147   | 28,030                   |                                               |
| 3′(2′), 5′-bisphosphate nucleotidase 1                                      | NP_006076.3      | BPNT1       | 28,395                   |                                               |
| WNT1 inducible signaling pathway protein 2 precursor                        | NP_003872.1      | WISP2       | 28,460                   |                                               |
| Thiopurine S-methyltransferase                                              | NP_001543.1      | IGFBP4      | 28,617                   |                                               |
| Hypothetical protein BC001573                                               | NP_620164.1      | LOC134147   | 28,826                   |                                               |
| 3′/H11032, 5′/H11032-bisphosphate nucleotidase 1                            | NP_000358.3      | TPMT        | 29,617                   |                                               |
| WNT1 inducible signaling pathway protein 2 precursor                        | NP_003872.1      | WISP2       | 29,617                   |                                               |
| Dimethylarginine dimethylaminohydrolase 2                                   | NP_036220.1      | PGLS        | 29,923                   |                                               |
| Hypothetical protein BCO01573                                               | NP_620164.1      | LOC134147   | 29,929                   |                                               |
| Carbonyl reductase 1                                                        | NP_001748.1      | CBR1        | 30,444                   |                                               |
| Dimethylarginine dimethylaminohydrolase 1                                   | NP_036220.1      | PGLS        | 30,617                   |                                               |
| Mitochondrial short-chain enoyl-coenzyme A hydratase 1 precursor            | NP_000358.3      | TPMT        | 31,807                   |                                               |
| Thioredoxin 3,4-dioxygenase                                                  | NP_036220.1      | PGLS        | 32,122                   |                                               |
| Lactamase, β 2                                                              | NP_057111.1      | LACTB2      | 32,522                   |                                               |
| 3-Mercaptopyruvate sulfurtransferase                                         | NP_066949.1      | MRST        | 32,603                   |                                               |
| CGI-150 protein                                                             | NP_057164.2      | C17orf25    | 32,826                   |                                               |
| Thiosulfate sulfurtransferase                                                | NP_003303.2      | TST         | 33,408                   |                                               |
| Biliverdin reductase A                                                      | NP_007073.2      | BLVRA       | 33,692                   |                                               |
| Regucalcin                                                                  | NP_004674.1      | RGN         | 33,802                   |                                               |
| Secreted phosphoprotein 1                                                   | NP_00573.1       | CRK         | 33,908                   |                                               |
| Mitochondrial malate dehydrogenase precursor                                | NP_005909.2      | MDH2        | 34,812                   |                                               |
| Aldo-keto reductase family 1, member B1                                      | NP_001619.1      | AKR1B1      | 35,827                   | Previously identified in urine               |
| Hypothetical protein FLJ11151                                               | NP_060610.1      | FLJ11151    | 35,942                   |                                               |
| Glycerinaldehyde-3-phosphate dehydrogenase                                  | NP_002037.2      | GAPD        | 36,201                   | Previously identified in urine               |
| Poly(rC) binding protein 3                                                  | NP_065389.1      | PCBP3       | 36,201                   |                                               |
| 3-Hydroxyanthranilate 3,4-dioxygenase                                        | NP_036337.1      | HAAO        | 36,201                   |                                               |
| Isopenetyl-diphosphate δ isomerase                                          | NP_004499.2      | ID1         | 36,201                   |                                               |
| Lactamase, β 2                                                              | NP_057111.1      | LACTB2      | 36,201                   |                                               |
| 3-Mercaptopyruvate sulfurtransferase                                         | NP_066949.1      | MRST        | 36,201                   |                                               |
| CGI-150 protein                                                             | NP_057164.2      | C17orf25    | 36,201                   |                                               |
| Thiosulfate sulfurtransferase                                                | NP_003303.2      | TST         | 36,201                   |                                               |
| Biliverdin reductase A                                                      | NP_007073.2      | BLVRA       | 36,201                   |                                               |
| Regucalcin                                                                  | NP_004674.1      | RGN         | 36,201                   |                                               |
| Secreted phosphoprotein 1                                                   | NP_00573.1       | CRK         | 36,201                   |                                               |
| Mitochondrial malate dehydrogenase precursor                                | NP_005909.2      | MDH2        | 36,201                   |                                               |
| Aldo-keto reductase family 1, member B1                                      | NP_001619.1      | AKR1B1      | 36,201                   |                                               |
| Hypothetical protein FLJ11151                                               | NP_060610.1      | FLJ11151    | 36,201                   |                                               |
| Glycerinaldehyde-3-phosphate dehydrogenase                                  | NP_002037.2      | GAPD        | 36,201                   |                                               |
| Poly(rC) binding protein 3                                                  | NP_065389.1      | PCBP3       | 36,201                   |                                               |
| 3-Hydroxymethyl-3-methylglutaryl-CoA lyase-like 1                            | NP_061909.1      | HMGCLL1     | 36,201                   |                                               |
| Aldo-keto reductase family 1, member A1                                      | NP_006057.1      | AKR1A1      | 36,201                   |                                               |
| Hypothetical protein BCO01573                                               | NP_057707.3      | BM-009      | 36,201                   |                                               |
| Cathepsin H isoform b precursor                                             | NP_683880.1      | CTSH         | 36,201                   |                                               |
| Aldose 1-epimerase                                                          | NP_620156.1      | GALM        | 36,201                   |                                               |
| Capping protein (actin filament), gelsolin-like                             | NP_001738.1      | CAPG         | 36,201                   |                                               |
| Hypothetical protein FLJ20446                                               | NP_060295.1      | ADPRHL2     | 36,201                   |                                               |
| Protein tyrosine kinase 9-like                                              | NP_009215.1      | PTK9L       | 36,201                   |                                               |
| H2A histone family, member Y isoform 2                                      | NP_004884.1      | H2AFY       | 36,201                   |                                               |
| Tropomodulin 3 (ubiquitous)                                                 | NP_053582.1      | TMOOD3      | 36,201                   |                                               |
| Class I alcohol dehydrogenase, α subunit                                    | NP_006057.1      | AKR1A1      | 36,201                   |                                               |
| Aldo-keto reductase family 1, member C1                                      | NP_001344.2      | AKR1C1      | 36,201                   |                                               |
| α-Actin-associated LIM protein                                              | NP_035291.1      | PDLIM3      | 36,201                   |                                               |
| Connective tissue growth factor                                              | NP_001892.1      | CTGF        | 36,201                   |                                               |
| α-1 actin precursor                                                         | NP_001091.1      | ACTA1       | 36,201                   |                                               |
| Pepsinogen 5, group I (pepsinogen A)                                        | NP_055039.1      | PGA5        | 36,201                   |                                               |
| Class II alcohol dehydrogenase 4 π subunit                                  | NP_000661.1      | ADH4        | 36,201                   |                                               |
| Similar to peroxisomal acyl-coenzyme A thioester hydrolase 2a              | XP_090885.5      | C14orf42    | 36,201                   |                                               |
| Serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 9   | NP_004146.1      | SERPINB9    | 43,004                   |                                               |
| UV excision repair protein RAD23 homolog B                                  | NP_002865.1      | RAD23B      | 43,004                   |                                               |
| Chitinase 3-like 2                                                          | NP_003991.1      | CHI3L2      | 43,004                   |                                               |
| Alanine-glyoxylate aminotransferase                                         | NP_00021.1       | AGXT        | 43,004                   |                                               |
obtained antibodies that work in Western blotting experiments against a subset of proteins. As shown in Fig. 3, we were able to observe the expression of Galectin 3, Cofilin 2, Cathepsin H, H9251-defensins 1 and 3, Cathepsin D, Nucleophosmin 1, connective tissue growth factor, and Lipocalin 2 in normal serum. Thus, although we cannot formally rule out the possibility that some of the proteins are found in the serum because of the underlying disease of the patient undergoing hemodialysis, we think that the majority of the proteins identified in this study are normal components of serum/plasma and have not been previously identified due to their lower concentrations in the blood because of clearance by the kidneys.

Secreted Proteins Identified from the Hemodialysis Fluid

We identified several proteins that were either known or predicted to be secreted proteins. Below we will discuss a subset of these proteins that are likely to be especially interesting from a biological perspective. Many of them have not been previously identified in serum/plasma. All accession numbers are from the RefSeq data base.

**A Novel Predicted Osteoblast Protein (FAM3C; NP_055703)—** This protein was initially identified from genomic data bases using structure-based methods in search of novel “four-helix bundle” cytokines. Four genes were iden-
tified in this family, although functional prediction has only been done for a related protein encoded by a related gene, FAM3B, which inhibits basal insulin secretion. This novel protein encoded by the FAM3C gene is expressed in all the tissues examined but was named as a predicted osteoblast protein as it was initially observed in osteoblasts. It contains an N-terminal signal peptide and no transmembrane domain, which suggests that it is a secreted protein (13).

**Fatty Acid-binding Protein 3** (FABP3) (NP_004093)—This was originally identified by screening a human adult muscle /H9261 gt11 expression library with an antibody to muscle FABP (14). The same molecule was also identified as mammery-derived growth inhibitor based on its activity as a growth inhibitor in lactating bovine mammary gland. This protein plays a role in the intracellular transport of long-chain fatty acids and their acyl-CoA esters. FABP3 is a candidate tumor suppressor gene involved in breast cancer (15).

**Antrum Mucosa Protein 18 kDa** (NP_062563)—This molecule, also designated as CA11, was initially isolated using differential display in human gastric cancer tissue. The expression of CA11 was observed to be down-regulated in gastric cancer tissue as compared with the normal gastric mucosa. Northern blot and RACE indicated that it is predominantly expressed in stomach and at low levels in the uterus and placenta (16). It has been suggested that the loss of its expression in gastric tissues may play an important role in gastric carcinogenesis (17). This protein is localized to the secretory granules of mucosal epithelium lining the stomach lumen (18).

**Resistin** (NP_065148)—Resistin (resistance to insulin) belongs to a family of proteins that is involved in inflammatory processes and in regulating metabolism. Human resistin (also referred to as FIZZ3) was identified by searching sequence databases with a related mouse protein called FIZZ1 (19). The expression of resistin is induced during adipogenesis, and it is normally secreted by adipocytes. Elevated resistin levels are seen in serum in both genetic and diet-induced obesity suggesting that it could potentially link obesity to diabetes (20).

**Dermcidin** (NP_444513)—Dermcidin is a novel human antimicrobial peptide secreted by the sweat glands. Screening of a subtracted cDNA library of primary melanoma and benign melanocytic nevus tissues with cDNA arrays led to isolation of dermcidin. It has been shown to possess antimicrobial activity against *Escherichia coli*, *Enterococcus faecalis*, *Staphylococcus aureus*, and *Candida albicans* (21). It may also play a role

---

**TABLE II**

A list of semitryptic peptides identified from the hemodialysis fluid

| Protein name                                      | Accession number | Peptide sequence                  |
|--------------------------------------------------|------------------|-----------------------------------|
| β globin                                         | NP_000509        | KVLGAFSDGLAHLDNLK                 |
|                                                  |                  | DNLKGTFLS_ELC ASPR                |
|                                                  |                  | AYOKWAVANL_AAPR                   |
|                                                  |                  | DGLGLYSLS_HIANVER                |
|                                                  |                  | VKNRTLTSWGLDSGK                   |
|                                                  |                  | SGKLYTLVLTD_PAPSR                 |
|                                                  |                  | SGKLYTLVLTD_PAPSR                 |
| Gelsolin                                         | NP_000168        | DGLGLSYLSSHIANVER                 |
| Prostatic binding protein                        | NP_002558        | VKNRTLTSWGLDSGK                   |
| Albumin precursor                                | NP_000468        | SGKLYTLVLTD_PAPSR                 |
| Aldo-keto reductase family 1                     | NP_006057        | SGKLYTLVLTD_PAPSR                 |
| δ globin                                         | NP_000510        | SGKLYTLVLTD_PAPSR                 |
| Peptidylprolyl isomerase A                       | NP_066953        | SGKLYTLVLTD_PAPSR                 |
| α 2 globin                                       | NP_000508        | SGKLYTLVLTD_PAPSR                 |
| Ubiquitin and ribosomal protein S27a precursor   | NP_002945        | SGKLYTLVLTD_PAPSR                 |

---

**FIG. 2.** A Venn diagram showing the degree of overlap based on the non-redundant set of genes that encoded the proteins identified in this study and those previously reported in serum/plasma (9, 10).

**FIG. 3.** Western blot analysis to confirm expression in normal serum. The figure shows a Western blot analysis of serum from a healthy volunteer that was depleted of albumin, IgG, IgA, transferrin, haptoglobin, and antitrypsin. The antibodies used are directed against the proteins whose names are indicated above each panel.
HIV-1 were identified. One of them was plasma (27). All the family members are known to be expressed in bone marrow, leukocytes, and neutrophils (24). α-Defensins have been shown to inhibit the replication of HIV-1 (25).

α-Defensin 3 (NP_005208)—Several proteins secreted by activated CD8⁺ T cells from long term non-progressors with HIV-1 were identified. One of them was α-defensin 3 encoded by the DEFA3 gene (25). α-Defensins 1, 2, and 3 collectively account for much of the anti-HIV-1 activity of CD8 antiviral factor that is not attributable to β-chemokines (24). It is known to be expressed in bone marrow, leukocytes, and neutrophils (24, 26). All the family members are known to be present in plasma (27).

Thymosin β 10 (NP_066926)—Thymosin β 10 was isolated from a kidney cDNA library and is an actin-sequestering protein (28). Thymosin β 10 has been shown to be a putative progression marker for human cutaneous melanoma (28). It is likely that this protein is released into the plasma after lysis of cells.

Chromosome 19 Open Reading Frame 10 (NP_061980)—A novel secreted protein was identified in a murine system using an expression cloning strategy (29). We have identified the human ortholog of this murine secreted protein. The function of this protein is not known, although its sequence suggests that it is likely to be a cytokine.

Semitryptic Versus Full Tryptic Data Base Searching

All of the 292 proteins in this study were identified by searching data bases using tryptic constraints for peptides. However, one would also expect to observe fragments derived after proteolysis in vivo. Thus, we tested this by searching a subset of our data having the highest mass accuracy (data from quadrupole-time-of-flight mass spectrometer) with semitryptic constraints. Searching our data with semitryptic constraints. Searching our data with semitryptic constraints resulted in a higher total score for nearly all of the proteins identified. However, for most of the entries, the higher score was a result of the contribution of low scoring peptides that did not pass our validation criteria. Semitryptic high scoring peptides that passed our manual validation are listed in Table II. Although some of the peptides presented in Table II could results from in-source fragmentation of labile proline-containing peptides (30), we expect that the majority arise from proteolytic cleavage events that occurred in vivo. Fig. 4A shows an example of an MS/MS spectrum of a semitryptic peptide derived from Gelsolin.

Proteolytic Activity in Hemodialysis Fluid

Theoretically the hemodialysis fluid should not contain proteins greater in size than 60 kDa. Nevertheless it is possible for larger proteins to be detected in the hemodialysis fluid if they undergo proteolytic cleavage in vivo. Thus, we decided to examine the distribution of peptides in greater detail. Of the proteins greater than 90 kDa, more than half exhibited a grouping of several peptides in either the N or C terminus of the protein, suggesting the presence of a fragment or isoform. Smaller fragments of proteins can be the result of proteolytic activity in plasma. Fig. 5 shows five examples of such groupings within a larger protein. For most of the proteins shown in Fig. 5, all peptides are clustered in regions that constitute only 4–8% of the whole protein sequence. For the 260-kDa protein fibronectin 1 and the 189-kDa protein complement component 3, a clustering of peptides was found in the middle of the proteins. This observation could be explained if these proteins were cleaved twice resulting in the observed fragment. For plasminogen (93 kDa) and desmocollin 1 (100 kDa), a clustering of peptides was observed in the N-terminal part of the protein, whereas a C-terminal clustering was observed in the case of Perlecian (480 kDa). Furthermore all of the above mentioned proteins were identified from bands that were at least a third of their expected molecular mass. For instance, fibronectin 1 was identified from the 15–19-kDa gel band, and desmocollin 1 was identified from the 6–15-kDa gel band. It is also worth noting that although several proteolytic cleavage sites are located in the C-terminal region of plasminogen (31), our data suggest that a cleavage occurred in the N terminus of the protein to generate Angiostatin (32), an important player in the regulation of angiogenesis (33). The probability of clustering of five peptides within a region that is 5% of the total protein length is 10⁻⁶ strongly suggesting that we are observing true in vivo proteolytic events. Of the above described proteins, all were retrieved from protein bands migrating in the ~6–15-kDa region, except for complement component 3 that was retrieved from the ~60–80-kDa region. For the latter this observation is in support of a single cleavage rather than a multicleavage event and is in agreement with literature where C3 convertase has been reported to cleave complement component 3 at residue 748, generating C3a anaphylatoxin (34).

Post-translational Modifications

Many proteins and peptides found in plasma and hemodialysis fluid have undergone a number of changes since the formation of nascent polypeptide chains. These changes have to be taken into account when interpreting the mass spectrometry data partly to get a better description of the biological sample and partly to avoid false-positive identifications. Basically the modifications can be grouped into enzyme-catalyzed changes and spontaneously occurring modifications. A large number of peptides were oxidized at methionine and tryptophan residues or exhibited pyroglutamine formation.
by cyclization by an N-terminal glutamine residue. It is difficult to infer any biological significance from these spontaneously occurring modifications, and they will not be discussed in detail. We searched the data from the quadrupole time-of-flight mass spectrometer to identify some of the commonly occurring post-translational modifications.

**FIG. 4.** Tandem mass spectra of a semitryptic peptide, LSSHIANVER, from Gelsolin (A), an N-terminally acetylated peptide, Ac-ADKPDMGEIASFDK, from thymosin β 10 (B), an N-glycosylated peptide, NLEKNHexNAcSTKQELAALEK, from prosaposin (C), a peptide containing hydroxyproline, GSAGP^Hyp^GATGFP^Hyp^GAAGR, derived from Collagen I (D). The corresponding peptide sequence is indicated in each case. Ac refers to acetylation, Hyp refers to hydroxyproline, and HexNAc refers to N-acetylhexosamine.

**FIG. 5.** A distribution of peptides for five large proteins. The proteins shown in the figure were identified from bands corresponding to molecular masses that were significantly lower than the calculated molecular masses. The peptides corresponding to the proteins are indicated by short thick lines. comp., component.
Acetylation

The N terminus of most proteins in vivo is processed by aminopeptidases and N-acetyltransferases (35). Aminopeptidases generally cleave the initiator methionine if the penultimte residue has a small radius of gyration (e.g., Gly, Ala, Ser, Cys, Thr, Pro, or Val). N-terminal acetylation is a very common co- and post-translational process where an acetyl group is transferred from acetyl-CoA to the N-terminal α-amino acid of a protein (35). In our study, the acetylated N termini of 43 proteins were detected. The proteins and the corresponding gene symbols are listed in Table III.

### Table III

A list of N-terminally acetylated peptides and the corresponding proteins

| Protein name | Accession number | Gene symbol | Peptide sequence | Instrument |
|--------------|------------------|-------------|------------------|------------|
| Thymosin, β 10 | NP_066926.1 | TMSB10 | Ac-ADKPDMGEIASFDK | Ion trap |
| Thymosin-like 2 | NP_877594.1 | TMSL2 | Ac-SDKPDMAEIEK | Q-TOF |
| Thymosin, β 4 | NP_066932.1 | TMSB4X | Ac-SDKPDMoxAEIEFDK | Q-TOF |
| Thymosin-like 4 | NP_877595.1 | TMSL4 | Ac-SDKPDMAEIEK | Q-TOF |
| Metallothionein 2A | NP_005944.1 | MT2A | Ac-MDPNCSCAGDSCTCAGSCK | Q-TOF |
| Metallothionein 1X | NP_005943.1 | MT1X | Ac-MoxDPNCSCGVSCTCAGSCK | Q-TOF |
| Metallothionein 1G | NP_005941.1 | MT1G | Ac-MDPNCSCAGVSCTCAGSCK | Q-TOF |
| Metallothionein 1F | NP_005940.1 | MT1F | Ac-MoxDPNCSCGVSCTCAGSCK | Q-TOF |
| Metallothionein 1B | NP_005938.1 | MT1B | Ac-MDPNCSCGVSCTCAGSCK | Q-TOF |
| Heat shock 10-kDa protein 1 | NP_002148.1 | HSPE1 | Ac-AGAQFRK | Q-TOF |
| Cystatin B | NP_000091.1 | CSTB | Ac-MMCGAPSATQPATAETQHIADQVR | Q-TOF |
| Death-associated protein | NP_004385.1 | DAP | Ac-SSPPEGKLETK | Ion trap |
| Glutaredoxin (thioltransferase) | NP_002055.1 | GLRX | Ac-AQEFVNCK | Ion trap |
| Metallothionein 1E | NP_783316.1 | MT1E | Ac-MDPNCSCAGDSCTCAGSCK | Q-TOF |
| Fatty acid-binding protein 4, adipocyte | NP_001433.1 | FABP4 | Ac-VNPTVFDDAIADVGEPLGR* | Q-TOF |
| Fatty acid-binding protein 3 | NP_004093.1 | FABP3 | Ac-VNPTVFDDAIADVGEPLGR | Q-TOF |
| Profilin 1 | NP_005013.1 | PPN1 | Ac-AGWNAAYDNMLMAGDTQCCDAQVYK | Q-TOF |
| β-Galactosidase binding lectin precursor | NP_002296.1 | LGALS1 | Ac-AGGLASNLNLKGPCECLR | Q-TOF |
| Fatty acid-binding protein 5 (psoriasis-associated) | NP_001435.1 | FABP5 | Ac-ATVQOLEGR | Q-TOF |
| Coactosin-like 1 | NP_006972.1 | COTL1 | Ac-ATKIDKACR | Ion trap |
| Glia maturation factor, β | NP_004115.1 | GMFB | Ac-SESLWCVDAEDVLEKR | Ion trap |
| Peptidylprolyl isomerase A (cyclophilin A) | NP_006953.1 | PPIA | Ac-VNPTVFDDAIADVGEPLGR | Ion trap |
| Cofilin 1 (non-muscle) | NP_00498.1 | CFL1 | Ac-AGVSVDGGIK | Q-TOF |
| Cofilin 2 | NP_086733.1 | CFL2 | Ac-AGVSVDGGIK | Ion trap |
| Ubiquitin-conjugating enzyme E2 G 1 isoform 1 | NP_003333.1 | UBE2G1 | Ac-MDPNCSCAGDSCTCAGSCK | Q-TOF |
| ADP-ribosylation factor 1 | NP_001649.1 | ARF1 | Ac-AGVSLHNLK | Q-TOF |
| Transgelin 2 | NP_003555.1 | TAGLN2 | Ac-ANRGPAYGLSR | Ion trap |
| Transgelin | NP_003177.1 | TAGLN | Ac-ANRGPAYGLSR | Ion trap |
| Chloride intracellular channel 1 | NP_001279.2 | CLIC1 | Ac-ATKIDKACR | Ion trap |
| Carbonic anhydrase I | NP_001729.1 | CA1 | Ac-ASDPWGGYDKNGPEQWSK | Q-TOF |
| Carbonic anhydrase II | NP_000058.1 | CA2 | Ac-ASDPWGGYDK | Ion trap |
| Dimethylarginine dimethylaminohydrolase 1 | NP_003529.1 | DDAH1 | Ac-AGLHHPAAFG | Q-TOF |
| Regucalcin | NP_004674.1 | RGN | Ac-AGVFLLSLTP | Ion trap |
| Hypothetical protein FLJ11151 | NP_006081.1 | FLJ11151 | Ac-SAAEAGHFWR | Q-TOF |
| Aldo-keto reductase family 1, member A1 | NP_006057.1 | AKR1A1 | Ac-AGVSLHNLK | Q-TOF |
| β actin | NP_001092.1 | ACTB | Ac-AGVFLLSLTP | Ion trap |
| Phosphoglycerate kinase 1 | NP_00282.1 | PGK1 | Ac-ASDPWGGYDK | Ion trap |
| LIM protein (similar to rat protein kinase C-binding enigma) | NP_006448.1 | LIM | Ac-ASDPWGGYDK | Ion trap |
| LIM domain binding 3 | NP_009009 | LDB3 | Ac-SSVLGLPGPGPWGFT | Q-TOF |
| Tolloid-like 1 | NP_063829.3 | TLL1 | Ac-GLGTLSPR | Q-TOF |
| Chromosome 6 open reading frame 115 | XP_371848 | C6orf115 | Ac-MoxNVDHEVNLVEEHR | Q-TOF |

---

**Molecular & Cellular Proteomics** 4.5 647

Proteomic Analysis of Human Hemodialysis Fluid
acetylated peptides are listed in Table III. We were able to find two different N termini for peptidylprolyl isomerase A; in one instance, the initiator methionine was acetylated, and in another, the initiator methionine was removed and the next amino acid, valine, was acetylated. Fig. 4B shows an MS/MS spectrum of an acetylated peptide derived from thymosin β 10 protein. Our data on N-terminal acetylation is in good agreement with the preference of methionine aminopeptidases that the ultimate amino acid is small or unmodified. Of the 43 acetylated proteins, one N-acetylated peptide was found in the middle of a predicted protein (XP_371848), suggesting that this could be a wrongly predicted protein. Indeed a Blast analysis of orthologous proteins confirms our new assignment of the translational initiation site (see Supplemental Fig. 2), which is in agreement with sequence conservation in five different species beginning at this methionine. Supplemental Fig. 3 shows the MS/MS spectra of all peptides containing an acetylated residue.

**N-Glycosylation of Peptides**

Three peptides were found to be covalently modified with an N-acetylhexosamine moiety, which corresponds to a mass increase of 203 Da (Table III and Supplemental Fig. 4). The CID-induced fragmentation spectrum of the triply charged ion with \( m/z \) 711.396 is shown in Fig. 4C. The N-acetylhexosamine gives rise to an intense oxonium ion at \( m/z \) 204, and further fragmentation of the oxonium ion gives rise to ions with \( m/z \) 186, 168, 138, and 126 (see Fig. 4C, inset). Interpretation of the mass spectrum revealed that the second asparagine in the peptide was covalently linked to an N-acetylhexosamine moiety. This modified asparagine occurs in the glycosylation consensus sequence NX(S/T) where X can be any amino acid. This is an unusual modification because the common N-glycan is cotranslationally transferred to the asparagine residue en bloc as the oligosaccharide GlcNAc₂Man₅Glc₃ and seldom trimmed beyond the trimannosyl-chitobiose core. It is interesting that only three peptides are found with this modification and come from a protein called prosaposin. Because N-glycosylation takes place only on asparagines in the glycosylation sequon NX(S/T), there will always be an adjacent serine of threonine. It is difficult to find peptides where O-GlcNAc is at least not a theoretically possible explanation (except in the rare instances where X is arginine or lysine, and trypsin cleavage is not hindered by the glycan moiety). However, there is a doubly charged ion in the spectrum that shows, even though it is small, that asparagine is the site of modification.

Prosaposin is a 524-amino acid glycoprotein that gives rise to four saposins that are predominantly localized in late endosomal/lysosomal compartment. A large number of endo- and exoglycosidases are also present in the lysosome, and it is possible that the extensive trimming of the glycan down to a single N-acetylgalactosamine takes place there. Whether this trimming of the glycan structure has any biological significance is not known, but it should be noted that two cases of metachromatic leukodystrophy have been reported where the mutations in the glycosylation sequon, either N215H (36) or T217I (37) corresponding to glycosylation sequon in the identified peptide TNSTFVQALVEHVKEECDR, led to a dysfunctional saposin B protein emphasizing the importance of glycosylation.

**Proline Hydroxylation**

A number of structural molecules in the extracellular matrix are known to undergo extensive post-translational modifications. Hydroxylation of prolines by prolyl hydroxylase is a common modification of collagen that confers structural stability to the collagen triple helix (38). We identified one proline hydroxylation site in fibrinogen, two sites in collagen α 2 type I, and six sites in collagen α 1 type I (see Fig. 4D for the MS/MS spectrum of a peptide containing two hydroxylated proline residues). Table IV lists all the hydroxylation sites identified in this study, and the MS/MS spectra are shown in Supplemental Fig. 5.

**Conclusions**

Using one-dimensional gel electrophoresis and LC-MS/MS, we have identified 292 proteins from hemodialysis fluid of which more than half were proteins smaller than 30 kDa.
Analysis of the modified peptides led to identification of 43 N-terminally acetylated proteins and three proteins hydroxylated on prolines. We also found three peptides from prosaposin to be modified with a single HexNAc at two different glycosylation sequons. We were able to map the identified peptides onto larger proteins, which showed groupings of peptides within limited regions. A comparison of our results with previously published studies that examined serum and plasma proteomes showed that two-thirds of the proteins identified in this study had not been identified previously as components of serum or plasma. We feel that this is mainly due to two major contributing factors: the first is the greater dynamic range of protein concentrations in serum/plasma samples, and the second is enrichment of the lower molecular weight proteins in the hemodialysis fluid. The proteins identified in this study will allow further investigations into their detection in serum/plasma and possible use as biomarkers of disease states.

This study presents the first comprehensive list of hemofiltrate proteins and in-depth analysis of the post-translational modifications. This proteomic survey is by no means exhaustive, and there are probably many proteins in the low molecular weight region we have not identified. This is not an uncommon phenomenon. In the Anderson et al. (9) list where literature and three proteomic studies have been compared, 196 of 1,275 proteins were reported in more than one study, and only 46 were reported in all four studies. It is currently not possible to extrapolate from the hemofiltrate to plasma or urine constituents, but it represents a little step toward mapping the enormous unknown of the human body fluid proteomes in health and disease.

Acknowledgments—We thank John Kloss for help with data base searching and Raghuhan Reddy for creating helpful scripts. We thank Sun Microsystems for providing us a computer cluster under the Academic Equipment Grant Program.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

†† Supported by a grant from the Danish Natural Sciences Research Council.

‡‡ Supported by Beckman Young Investigator and Sidney Kimmel Scholar awards. Serves as Chief Scientific Advisor to the Institute of Bioinformatics. The terms of this arrangement are being managed by The Johns Hopkins University in accordance with its conflict of interest policies. To whom correspondence should be addressed. Tel.: 410-502-6662; Fax: 410-502-7544; E-mail: pandey@jhmi.edu.
FIZZ1, a novel cysteine-rich secreted protein associated with pulmonary inflammation, defines a new gene family. *EMBO J.* **19**, 4046–4055

20. Banerjee, R. R., Rangwala, S. M., Shapiro, J. S., Rich, A. S., Rhoades, B., Qi, Y., Wang, J., Rajala, M. W., Pocai, A., Scherer, P. E., Steppan, C. M., Ahima, R. S., Obici, S., Rossetti, L., and Lazar, M. A. (2004) Regulation of fasted blood glucose by resistin. *Science* **303**, 1195–1198

21. Cunningham, T. J., Hodge, L., Speicher, D., Reim, D., Tyler-Polsz, C., Levitt, P., Eagleson, K., Kennedy, S., and Wang, Y. (1998) Identification of a survival-promoting peptide in medium conditioned by oxidatively stressed cell lines of nervous system origin. *J. Neurosci.* **18**, 7047–7060

22. Porter, D., Weremowicz, S., Chin, K., Seth, P., Keshaviah, A., Lahti-Do- menici, J., Bae, Y. K., Monitto, C. L., Merlos-Suarez, A., Chan, J., Hulette, C. M., Richardson, A., Morton, C. C., Marks, J., Duyao, M., Hruban, R., Gabrieelson, E., Gelman, R., and Polyak, K. (2003) A neural survival factor is a candidate oncogene in breast cancer. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 10931–10936

23. Daher, K. A., Lehrer, R. I., Ganz, T., and Kronenberg, M. (1988) Isolation and characterization of human defensin cDNA clones. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 7327–7331

24. Date, Y., Nakazato, M., Shiomi, K., Toshimori, H., Kangawa, K., Matsuo, H., and Matsukura, S. (1994) Localization of human neutrophil peptide (HNP) and its messenger RNA in neutrophil series. *Ann. Hematol.* **69**, 73–77

25. Zhang, L., Yu, W., He, T., Yu, J., Caffrey, R. E., Dalmasso, E. A., Fu, S., Pham, T., Mei, J., Ho, J. J., Zhang, W., Lopez, P., and Ho, D. D. (2002) Contribution of human α-defensin 1, 2, and 3 to the anti-HIV-1 activity of CD8 antiviral factor. *Science* **298**, 995–1000

26. Lanzmeier, R., Michaelson, D., Liu, L., and Ganz, T. (1993) The structure of neutrophil defensin genes. *FEBS Lett.* **326**, 299–300

27. Ashitani, J., Nakazato, M., Mukae, H., Taniguchi, H., Date, Y., and Matsukura, S. (2000) Recombinant granulocyte colony-stimulating factor induces production of a human neutrophil peptides in lung cancer patients with neutropenia. *Regul. Pept.* **95**, 87–92

28. McCreary, V., Kartha, S., Bell, G. I., and Toback, F. G. (1988) Sequence of a human kidney cDNA clone encoding thymosin β 10. *Biochem. Biophys. Res. Commun.* **152**, 862–866

29. Tulin, E. E., Onoda, N., Nakata, Y., Maeda, M., Hasegawa, M., Nomura, H., and Kitamura, T. (2001) SF20/IL-25, a novel bone marrow stroma-derived growth factor that binds to mouse thymic shared antigen-1 and supports lymphoid cell proliferation. *J. Immunol.* **167**, 6338–6347

30. Olsen, J. V., Ong, S. E., and Mann, M. (2004) Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell. Proteomics* **3**, 608–614

31. Ohyama, S., Harada, T., Chikanishi, T., Miura, Y., and Hasumi, K. (2004) Nonlysine-analog plasminogen modulators promote autoproteolytic generation of plasmin(ogen) fragments with angiostatin-like activity. *Eur. J. Biochem.* **271**, 809–820

32. Patterson, B. C., and Sang, Q. A. (1997) Angiostatin-converting enzyme activities of human matrilysin (MMP-7) and gelatinase B/type IV collagenase (MMP-9). *J. Biol. Chem.* **272**, 28823–28825

33. O’Reilly, M. S., Holmgren, L., Shing, Y., Chen, C., Rosenthal, R. A., Moses, M., Lane, W. S., Cao, Y., Sage, E. H., and Folkman, J. (1994) Angiostatin: a novel angiogenesis inhibitor that mediates the suppression of metastases by a Lewis lung carcinoma. *Cell* **79**, 315

34. Hugli, T. E. (1975) Human anaphylatoxin (C3a) from the third component of complement. Primary structure. *J. Biol. Chem.* **250**, 8293–8301

35. Polevoda, B., and Sherman, F. (2003) N-terminal acetyltransferases and sequence requirements for N-terminal acetylation of eukaryotic proteins. *J. Mol. Biol.* **325**, 595–622

36. Wrobe, D., Henseler, M., Huettler, S., Pascual Pascual, S. I., Chabas, A., and Sandhoff, K. (2000) A non-glycosylated and functionally deficient mutant (N215H) of the sphingolipid activator protein B (SAP-B) in a novel case of metachromatic leukodystrophy (MLD). *J. Inherit. Metab. Dis.* **23**, 63–76

37. Kretz, K. A., Carson, G. S., Morimoto, S., Kishimoto, Y., Fluharty, A. L., and O’Brien, J. S. (1990) Characterization of a mutation in a family with saposin B deficiency: a glycosylation site defect. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 2541–2544

38. Pihlajaniemi, T., Myllyla, R., and Kivirikko, K. I. (1991) Prolyl-4-hydroxylase and its role in collagen synthesis. *J. Hepatol.* **13**, Suppl. 3, S2–S7