A centrifugal ultrafiltration strategy for isolating the low-molecular weight (≤ 25K) component of human plasma proteome

David W. Greening[1,2] and Richard J. Simpson[2]

[1]Department of Biochemistry and Molecular Biology, The University of Melbourne

[2]Joint Proteomics Laboratory, Ludwig Institute for Cancer Research & The Walter and Eliza Hall Institute of Medical Research, Royal Melbourne Hospital, Parkville, Victoria, Australia

Correspondence:

Prof. Richard J. Simpson

Ludwig Institute for Cancer Research,

PO Box 2008, Royal Melbourne Hospital,

Parkville, Vic 3050, Australia;

Fax: +61-3-9341-3192

E-mail: richard.simpson@ludwig.edu.au

Abbreviations: false-discovery rate, FDR; low-molecular weight fraction, LMF; low-molecular weight, low-M_r; nominal molecular weight cut-off, NMWC.

Abstract

The low-molecular weight fraction (LMF) of the human plasma proteome is an invaluable source of biological information, especially in the context of identifying plasma-based biomarkers of disease. In this study, a separation and enrichment strategy based on centrifugal ultrafiltration was developed for the LMF (i.e., $\leq$ 25K) of plasma routinely prepared from normal, healthy volunteers. Four commercially-available filter membranes of similar nominal molecular weight cut-off (NMWC), but differing membrane chemistries and filter orientations (Microcon®, Millipore; Centrisart®, Sartorius; Amicon Ultra®, Millipore; Vivaspin®, Sartorius), were evaluated. Of these filtration devices, only the Sartorius Vivaspin® tangential membrane, NMWC 20K was effective in the non-retention of M_r >50K, and recovery and enrichment of low-M_r components from human plasma. This filter membrane device was further optimized with respect to plasma buffer composition, centrifugal force, duration and temperature. Optimal ultrafiltration conditions were obtained using 100 µL of normal plasma in 10% acetonitrile, and a centrifugation force of 4,000 × g for 35 min at 20°C. In this LMF, 44 proteins (from 266 unique peptides) were identified using a combination of 1D-SDS-PAGE / nanoLC-MS/MS and a stringent level of identification (FDR <1%). We report the identification of several proteins (e.g., protein KIAA0649 (*Q9Y4D3*), rheumatoid factor D5, serine protease inhibitor A3, and transmembrane adapter protein PAG)

previously not reported in extant high-confidence Human Proteome Organization (HUPO) Plasma Proteome Project datasets. When compared with the low-M_r human plasma/serum proteome datasets of Zhou et al. (Electrophoresis, 2004. 25, 1289-98), Gundry et al. (Proteomics Clin. Appl., 2007. 1,73-88) and Villanueva et al. (Anal Chem, 2004. 76,1560-70), 64% of our identifications (28 proteins) were novel; these include cofilin-1, PPIase A, and the SH3 domain-binding glutamic acid-rich-like protein 3. In addition to intact proteins, many peptide fragments from high-abundance proteins (e.g., fibrinogen, clusterin, Factor XIIIa, transferrin, kinogen-1, and inter-alpha-trypsin inhibitor), presumably derived by *ex vivo* proteolysis, were observed.

1   Introduction

With 20–25% of all proteins encoded by the genome predicted to be secreted [1], human plasma is one of the most informative and important proteomes from a clinical perspective. Hence, the plasma proteome can be considered to be a valuable window of normal and pathophysiological states. For example, the low-molecular-weight (low-M_r, $\leq 25K$) plasma or serum proteome has been the focus of recent attempts to identify novel diagnostic and prognostic cancer biomarkers [2, 3] (reviewed in [4]). Because low-M_r proteins/peptides exhibit far greater tumour and vascular permeability due to their molecular size range [5-8] they are also useful therapeutic targets.

A key challenge in identifying constituents in the low-molecular weight fraction (LMF) (i.e., $M_r \leq 25K$) of plasma is the issue of the high dynamic concentration range of protein

abundance (10 orders of magnitude between the least abundant (1–5 pg/mL - e.g., interleukins, cytokines etc) and most-abundant (35–70×10^9 pg/mL - e.g., albumin, IgG etc) [9]. Several strategies have been developed to circumnavigate this problem. These include multidimensional enrichment and separation strategies [10], and depletion of high-abundance [11] and high-molecular weight proteins [12]. Collectively, these measures permit identification of low-abundance protein/peptide species.

In 2005, the HUPO Plasma Proteome Project generated a core set of high-confidence 889 proteins (derived from a 3,020 low-confidence dataset) proteins [13, 14]. Interestingly, the low-M_r (i.e., $\leq 25K$) component of the blood proteome was under-represented in these studies [4]. This finding was confirmed by Zhou et al. [15] who identified an aggregate of 210 low-M_r proteins and peptides from plasma, of which only 62 were identified in the 3,020 HUPO protein dataset (2.0%), and 26 in the 889 HUPO dataset (2.9%) [13]. Strategies used to overcome the dynamic range of protein concentrations and enrich for low-M_r proteins and peptides of low-abundance include fractionation of plasma into sub-proteomes (i.e., LMF, glycoproteome etc), and/or to deplete one or more of the abundant proteins through immuno-depletion [16, 17]. With the latter strategy, it has been reported that a range of proteins/peptides may co-purify with highly abundant proteins targeted for depletion, such as albumin [18]. Using ELISA-based assays it has been reported that the levels of low-abundance proteins, such as cytokines MIP-1β, IL-4, -6, -8, -10, and -18, tumour-necrosis factor and growth-related oncogene, GRO can be significantly reduced following albumin depletion [19, 20].

Several different isolation approaches utilizing ultrafiltration have been reported for enriching the plasma/serum LMF [21-24]. Centrifugal ultrafiltration is a facile method which utilizes both centrifugation force and a semi-permeable membrane to retain high-M_r solutes [25]. This method allows low-M_r solutes to pass through the membrane (filtrate) of nominal molecular weight cut-off (NMWC) while high-M_r solutes are retained (retentate). For example, Zheng et al. [23] utilized centrifugal ultrafiltration (NMWC 10K) to identify endogenous peptides from serum (300 unique peptides with 2 ppm or better mass accuracy). However, few peptides/proteins between 3,000 and 10,000 Da were represented in these data. In another study, Tirumalai et al. [21] explored the enrichment of the low-M_r serum proteome using Centriplus® centrifugal concentrator membranes (NMWC 30K) before a 2D-LC–MS/MS approach. Although 341 proteins were identified, surprisingly, no albumin and albumin-derived peptides were reported. Recently, centrifugal ultrafiltration (NMWC 50K) combined with solution-phase isoelectric focusing and LC-MS/MS resulted in identification of 1,394 unique peptides (262 proteins) from 100 μL of serum [22]. Further, Jung and co-workers [26] utilized various organic solvent precipitation procedures and buffer systems, in combination with centrifugal ultrafiltration (NMWC 30K), to enrich the plasma LMF. While both studies contain extensive coverage of high-M_r abundant proteins, such as albumin, transferrin, α-1-antitrypsin, and the apoproteins (A-I, A-II), low-M_r proteins were underrepresented. The occurrence of abundant, high-M_r plasma proteins using centrifugal ultrafiltration may be associated with the heterogeneity in the pore sizes of filter membranes which influences the ability of filter membranes to effectively fractionate proteins of a narrow molecular weight range [27]. It should be emphasized

that – (i) the nominal pore-size (NMWC) (and M_r cut-off) of any filter membrane is only an average and that a normal distribution of smaller and larger pores exist, and – (ii) high concentration protein mixtures (e.g., plasma - 40-60 mg/mL), can result in membrane polarization or clogging [22, 27].

Here, we describe a rapid, facile strategy for the enrichment of the LMF of human plasma. We report optimized conditions for usage of Sartorius Vivaspin® tangential centrifugal ultrafiltration membranes that influence the transmembrane pressure and permeability. Plasma buffer composition, and centrifugation force, duration and temperature were evaluated.

2   Materials and methods

2.1   Plasma collection, preparation and storage

Whole blood was collected from healthy, volunteer blood donors (n=5) at the Australian Red Cross Blood Service (ARCBS), Melbourne, according to standard procedures (described therein [28]). All included subjects provided written informed consent for their blood donation, ethics of which was approved by the ARCBS, Melbourne. Briefly, whole blood from each donor (450 ± 50 mL) was routinely collected into an enclosed, single-use standard bag system containing 70 mL of citrate–phosphate–dextrose (PL-146, Fenwal Division, Baxter Healthcare Corp., Deerfield, IL, USA). Blood was obtained from the cubital vein by clean venipuncture at a single time-point. Plasma was obtained

by centrifugation of whole blood for 10 min at 4,200 × g, 22ºC and collection of the plasma layer by a manual plasma extractor. The plasma from each separate donor (n=5) was pooled (Serial #9285490, A+ type), rapidly frozen (-30ºC) and transported on dry ice. The plasma pack was thawed vertically for 6 h at room temperature (RT) and a Complete protease inhibitor cocktail (Roche, Mannheim, Germany) added. Plasma was centrifuged in 200 mL lots at 5,000 × g for 15 min after which aliquots (35 mL, 10 mL, and 1 mL) were placed in polypropylene (50 mL, 15 mL, and 1.5 mL) tubes and refrozen at -80ºC. Prior to use plasma samples were thawed at RT to minimize cryoprecipitate formation [29, 30], and immediately used. The protein concentration of the initial plasma sample was 67 mg/mL as determined by bicinchoninic (BCA) assay, using BSA as a standard as described [31]. The LMF recoveries using various centrifugal ultrafiltration membrane devices (in comparison to the plasma source material) were also determined by BCA protein assay.

2.2 Centrifugal ultrafiltration

The centrifugal filter membranes were prepared according to manufacturer's instructions. Membranes were rinsed firstly in deionized water (Milli-Q water, HPLC grade, 18MΩ from an A10-Synthesis water polishing system, Millipore) followed by each buffer additive (see below). Various filtration devices were analyzed; Microcon® YM-30 (30K) (Millipore, Bedford, MA), Centrisart® 20K (Sartorius, Hannover, Germany), Amicon Ultra-4® 30K (Millipore, Bedford, MA), and Vivaspin-4® 20K (Sartorius, Hannover, Germany). For preliminary membrane evaluation studies, 100 µL of plasma was diluted

with 900 μL 5% (v/v) aqueous ACN and centrifuged at 2,000 × g for 30 min. For

optimization studies using the Vivaspin-4® 20K device, plasma was diluted with 900 μL

aqueous 25 mM NH_4HCO_3, pH 7.8 or various concentrations of aqueous ACN (0-30%),

and allowed to stand at RT for 2 min. Each plasma sample was centrifuged at 14,000 × g

for 2 min at RT. The supernatant was applied to the prepared centrifugal filters and

samples were centrifuged using an M4 swing bucket rotor, optimized for g force and

duration, to deplete high-M_r proteins (*Section 3.1*). For SDS-PAGE analyzes, filtrates

were lyophilized to dryness and re-suspended in 20 μL of Laemmli non-reducing sample

buffer (NRSB: 0.2M Tris-HCl, 40% (v/v) aqueous glycerol, 4% SDS, trace bromophenol

blue), heated at 95ºC for 5 min then cooled prior to gel loading.

2.3   SDS-PAGE analyzes

One dimensional SDS-PAGE was performed using an Invitrogen Novex Mini-Cell

electrophoresis unit utilizing NuPAGE Bis-Tris precast gels (4–12% gradient

acrylamide), as described [28]. Under non-reducing conditions NuPAGE MES SDS

running buffer (#NP0002) was used as the buffer system at a constant voltage of 150V

for 80 min.

2.4   Protein visualization

Proteins were visualized by staining gels with SilverSNAP® Stain Kit II gel stain (Pierce,

Rockford, Ill.) according to the manufacturer's instructions. Briefly, gels were washed

(×2) for 5 min in deionized water, fixed (×2) (30% (v/v) aqueous ethanol containing 10% (v/v) acetic acid) for 15 min, washed for 10 min first in 10% ethanol, and then in deionized water for 10 min. Gels were incubated in SilverSNAP® Sensitizer solution for 1 min, stained in SilverSNAP® silver solution for 30 min, washed (×2) in deionized water for 1 min, developed in SilverSNAP® develop solution for approximately 3-5 min, and the gels fixed in 5% (v/v) acetic acid. SilverSNAP® stained gels were imaged with a Personal Densitometer SI (Molecular Dynamics) with 100 μm pixel size. For samples prepared for mass spectrometry (MS)-based protein identification, the gel was incubated with Coomassie R-250 (Imperial Protein Stain, Pierce Biotechnology). Gels were rinsed (×3) and washed (×3) for 5 min in deionized water followed by staining in Imperial Coomassie solution for 3 h on an orbital shaker. Gels were destained in deionized water for at least 3 h before scanning with a Personal Densitometer SI (Molecular Dynamics) with 100 μm pixel size. All gel images were processed using ImageQuant™ software (Molecular Dynamics).

2.5 Nano-LC-MS/MS analysis

The LMF of plasma was prepared using the Vivaspin-4® 20K (Sartorius, Hannover, Germany) cartridge, with the filtrate subjected to nano-LC-MS/MS analysis as follows. Equivalent amounts of total protein (30 μg) from a single LMF preparation (4,000 × *g* for 35 min at 20°C) were separated by SDS-PAGE as described. Gel lanes were excised (1.5-mm gel bands, n=30) and extensively washed with deionized water. Excised gel bands were individually digested with trypsin (0.05 μg) and generated peptides were

extracted and then concentrated to ~10 µL by centrifugal lyophilization (Savant, U.S.A.) for nano-electrospray-Ion Trap (nESI-IT) tandem mass spectrometry (MS/MS) (LTQ-Orbitrap, Thermo Fisher Scientific, MA, U.S.A.). Extracted peptides (~10 µL in 1% (v/v) formic acid) were transferred into 96-well polypropylene plates (ABgene™ Thermo-Fast, Thermo Fisher Scientific) for automated injection and fractionation by capillary reversed-phase-HPLC (Model 1200, Agilent, Germany) using a nanoAcquity™ (C18) 150 × 1.0 mm I.D. RP-capillary column (nanoAcquity-C18, 1.8 µm, Waters Corp, MA. U.S.A) developed with a linear 60-min gradient from 0-100% B with a flow rate of 0.5 µL/min at 45°C. Solvent A was 0.1% (v/v) aqueous formic acid and Solvent B was 0.1% aqueous formic acid/60% (v/v) ACN. The capillary HPLC was coupled on-line to the mass spectrometer and subjected to MS/MS as described above.

Positive ion mode was used for data-dependent acquisition on the LTQ-Orbitrap. Survey MS scans were acquired with the resolution set to a value of 30,000. Each scan was recalibrated in real time by co-injecting an internal standard from ambient air into the C-trap [32]. Up to five of the most intense ions per cycle were fragmented and analyzed in the linear trap. Target ions already selected for MS/MS were dynamically excluded for 180 s to optimize peptide coverage.

The parameters used to generate the peak lists, using *extract-msn*, were as follows: minimum mass 700; maximum mass 5,000; grouping tolerance 0.01 Da; intermediate scans 200; minimum group count 1; 10 peaks minimum and TIC of 100. Peak lists for

each LC-MS/MS run were merged into a single MGF file for MASCOT searches. Charge state of the selected ions was automatically determined from the survey scan.

2.6 Data processing and analysis

Protein/peptide identifications were obtained using the MASCOT search algorithm (v2.2, Matrix Science, U.K.) [33]. All acquired MS/MS spectra were searched against the generated LudwigNR_subset database (created January 2008; URL:ftp://ftp.ch.embnet.org/pub/databases/nr_prot/) [34]. Database search parameters were as follows: fixed modification, carboxymethylation of cysteine (+58 Da); variable modifications, NH_2-terminal acetylation (+42 Da), methionine oxidation (+16 Da). Peptide mass tolerance was ± 20 ppm and #13C defined as 1 with the allowance for up to three missed tryptic cleavage sites. In this study, for a peptide to be considered a potentially positive identification, the acceptance criteria required that the ProteinScore to be ≥ 42 (<1% FDR) (described therein [34]). Additionally, manual verification of the spectra (Java™ spectrum applet) was performed in accordance with previously established guidelines for inclusion of true peptide identifications [35]. Proteins were then submitted to *BLAST* http://www.ncbi.nlm.nih.gov/BLAST/) with a cut-off of 95% to remove redundancy. Different resources were used in order to classify identified proteins based on several predictive algorithms (SignalP (3.0), TMHMM (2.0)), and Gene Ontology (GO) and the UniProt database [36-38].

3 Results & discussion

3.1 Evaluation of centrifugal ultrafiltration for enrichment of the LMF

Four commercially-available centrifugal ultrafiltration devices with differing membrane chemistries and membrane orientation (flat-based or tangential) were evaluated: *membrane A* - Microcon® YM-30, regenerated cellulose membrane (flat-based); *membrane B* - Centrisart®, polyethersulfone membrane (flat-based); *membrane C* - Amicon Ultra-4®, polyethersulfone membrane (tangential); and *membrane D* - Vivaspin-4®, cellulose triacetate membrane (tangential). The separation efficiency of these four filtration devices were assessed by 1D-SDS-PAGE using a starting volume of 100 μL of human plasma. It can be seen in Figure 1 that significant amounts of high-M_r proteins are not retained by ultrafiltration membrane devices A, B and C, and appear in the corresponding filtrates ($\leq$ 30K). In contrast, for membrane D all high-M_r proteins were retained by the membrane using identical conditions.

The recovery of the LMF for each filter was quantitated using the BCA protein assay as described (*Section 2.2*). The assay indicated that ~0.2 – 2.9 mg of the LMF from each membrane filtrate was obtained from ~6.7 mg of proteinaceous material contained in a human plasma sample using each of the membrane filter units (relative to a standard, BSA) (Table 1). The corresponding retentate recoveries from each corresponding filter membrane are also provided (ranging from 3.4 – 6.1 mg). The filtrate LMF recovery from membrane devices A-C were compromised by the presence of high-M_r proteins as demonstrated (Figure 1). In this current analysis, *membrane D* was selected for further

optimization of low-M_r protein/peptide enrichment due to both the selectivity in fractionating the plasma LMF, and recovery of the plasma filtrate.

With our experience using these filtration devices for a variety of protein purification and fractionation procedures, we believe that it is the specific membrane chemistry and orientation (flat or angular) of the filter unit which directly influences the transmembrane pressure and permeability of each membrane, and subsequently, the selectivity of protein separation through the pores in the membrane. Vertical or angular membrane configuration reduces concentration polarization (membrane fouling) and allows continuous flow rates even with high proteinaceous solutions such as plasma. It is for this reason why *membrane* D, with its low-protein binding cellulose triacetate membrane, and angular filter orientation, is so effective in the reproducible fractionation of LMW components in human plasma. We have further attempted to optimize the other filter membrane units (A-C) over a variety of operating conditions using human plasma, although with low separation effectiveness.

We optimized the Vivaspin-4® 20K ultrafiltration device with respect to centrifugation force (1,000 – 5,000 × g), operating temperature, and sample buffer additives (e.g., acetonitrile, ammonium bicarbonate etc). It can be seen in Figure 1 (Panel D, lane F, filtrate) that the optimal conditions for protein fractionation occur at 10% acetonitrile, centrifuged 4,000 × g for 35 min at 20°C. Under these conditions there is slight leakage of ~50K protein(s), but not enough to confound the analysis of the enriched LMF. (The

presence of small amount of ~50K protein could be obviated by reducing the sample load in the filtrate).

3.2   Application of centrifugal ultrafiltration to study the LMF of plasma

To demonstrate the applicability of this method, we analyzed the LMF of normal human plasma using the optimized Vivaspin-4® 20K ultrafiltration conditions. The LMF filtrate (NMWC 20K) from plasma was collected (950 µL), lyophilized to dryness and re-suspended in 20 µL of Laemmli non-reducing sample buffer and applied to a 1D-SDS-PAGE gel. Following electrophoresis the gel lane was excised into 1.5-mm gel bands (n=30), the individual fractions trypsinized and the extracted peptides subjected to nanoLC-MS/MS, as previously described [34]. Using this approach, 44 low-M_r plasma proteins (from 266 unique peptides) were identified (Table 2). (The Protein Score, amino acid sequence, peptide charge state and reported ion and homology scores (http://www.matrixscience.com/) are given in Supplementary Table 1). Of these, 24 proteins showed a good correlation with the ≤ 25K theoretical molecular weight (based on UniProt annotation). The remaining proteins identified were derived from proteolytic fragments of abundant plasma proteins such as albumin and IgG (albumin, 18% sequence coverage).

The efficiency of the method is demonstrated by the identification of the low-abundance classically-secreted proteins cystatin-C (plasma concentration 0.62–1.02 µg/mL [39])

14

(Figure 2), CXCL7, serine protease inhibitor A3, and cystatin-M. In all, 48% of the proteins identified in the plasma LMF (Table 2) are known secreted proteins.

3.3   Comparison with plasma/serum proteome studies

The LMF was compared with published plasma proteomic datasets including - (i) a detailed human blood plasma protein reference set by Mann and colleagues [40], - (ii) a high-confidence HUPO plasma proteome collaborative study by Hanash and colleagues [13], and - (iii) a high-confidence HUPO reference plasma proteome reported by Greening et al. [34] (Figure 3A). In the human plasma proteome study by Schenk et al. [40], these authors utilized a series of depletion (HSA / HSA, transferrin, haptoglobin, alpha-1-antitrypsin, IgA and IgG) and fractionation procedures (1D-SDS, off-gel electrophoresis (OGE), 2-D gel electrophoresis). The later two studies analyzed data previously reported by Omenn et al. [14], in which a wide variety of methods, including multiple LC-MS/MS instruments, MALDI-MS, and FT-ICR-MS; depletion of abundant proteins; fractionation of intact proteins on 2-D gels or with LC or IEF methods; protein enrichment or labeling methods; immunoassays or antibody arrays; and direct (SELDI) MS were utilized. In this study, 13 proteins were co-identified in the global plasma proteome analyzes including Apo-AI, fibrinopeptide A, kininogen-1, prostaglandin-H2 D-isomerase, and transthyretin (Figure 3A, Table 2). Several other co-identified proteins (e.g., fibronectin and platelet basic protein) have reported functions throughout the coagulation system including cell adhesion, cell motility, and wound healing. We report the identification of several proteins (e.g., desmocollin-3, hypothetical protein (*Q6P5S8*),

myosin-reactive immunoglobulin kappa-chain variable region, protein KIAA0649 (*Q9Y4D3*), rheumatoid factor D5, serine protease inhibitor A3, and transmembrane adapter protein (PAG) not reported in previous global studies of the human plasma proteome (Figure 3A, Table 2).

Our plasma LMF dataset was further compared with several low-M_r human plasma/serum studies, including the serum peptidome by Tempst and colleagues (molecular mass range of 1-15K) [41], proteins and peptides bound and co-purified with albumin during depletion (i.e., the albuminome) [18], and the low-M_r proteins/peptides associated with abundant serum proteins (i.e., albumin, transferrin, IgA, IgM, apolipoprotein A-I) [42] (Figure 3B, Table 2). Apolipoprotein A-I, fibrinogen alpha chain, and transthyretin were co-identified between each of these datasets, while ITI heavy chain H4, kininogen-1, and platelet basic protein were co-identified in two of these studies (Table 2). Interestingly, 28 proteins we identified in our LMF study (64%) were not previously reported in each of these low-M_r studies (Table 2). These include lysozyme *C*, plasma retinol-binding protein, PPIase A, profilin-1, SH3 domain- binding protein 1, and transgelin-2.

The efficacy of our plasma LMF study was revealed by the identification, for the first time, of 12 plasma proteins (25% of identifications) not observed in previous plasma proteome studies (Table 2). Of these, Ig kappa chain V-I region AG, Ig kappa chain V-I region DEE, myosin-reactive immunoglobulin kappa chain variable region (fragment), peptidyl-prolyl cis-trans isomerase A (PPIase A) and rheumatoid factor D5 light chain (fragment) were ≤ 25K theoretical molecular weight (based on UniProt annotation).

In summary, optimal conditions were developed for the Vivaspin® tangential membrane filtration device for the purpose of fractionating the low-M_r component (LMF) of human plasma. In this study 12 low-M_r proteins not previously reported in high-confidence global plasma proteome analyzes, proteomic studies investigating interacting proteins with the high-M_r plasma proteins, or LMF/peptidome studies, demonstrating the efficacy of the method. This facile, one-step strategy for enriching the low-M_r plasma proteome provides a rapid (less than 1 h fractionation procedure), cost-effective (does not require sophisticated separation technology) and reproducible method for the enrichment of the LMF of human plasma.

Acknowledgements

Supplementary Information

Supplementary Table 1 - Inferred proteins based on identification of peptides from the human plasma LMF.

References

1. Chen, Y., P. Yu, J. Luo, and Y. Jiang, *Secreted protein prediction system combining CJ-SPHMM, TMHMM, and PSORT.* Mamm Genome, 2003. **14**(12): p. 859-65.

2. Lopez, M.F., A. Mikulskis, S. Kuzdzal, D.A. Bennett, J. Kelly, E. Golenko, J. DiCesare, E. Denoyer, W.F. Patton, R. Ediger, L. Sapp, T. Ziegert, C. Lynch, S. Kramer, G.R. Whiteley, M.R. Wall, D.P. Mannion, G. Della Cioppa, J.S. Rakitan, and G.M. Wolfe, *High-resolution serum proteomic profiling of Alzheimer disease samples reveals disease-specific, carrier-protein-bound mass signatures.* Clin Chem, 2005. **51**(10): p. 1946-54.

3. Liotta, L.A. and E.F. Petricoin, *Serum peptidome for cancer detection: spinning biologic trash into diagnostic gold.* J Clin Invest, 2006. **116**(1): p. 26-30.

4. Drake, R.R., L. Cazares, and J.O. Semmes, *Mining the low molecular weight proteome of blood.* Proteomics Clin. Appl., 2007. **1**: p. 758-768.

5. Jain, R.K., *Transport of molecules across tumor vasculature.* Cancer Metastasis Rev, 1987. **6**(4): p. 559-93.

6. Dvorak, H.F., J.A. Nagy, J.T. Dvorak, and A.M. Dvorak, *Identification and characterization of the blood vessels of solid tumors that are leaky to circulating macromolecules.* Am J Pathol, 1988. **133**(1): p. 95-109.

7. Yuan, F., M. Dellian, D. Fukumura, M. Leunig, D.A. Berk, V.P. Torchilin, and R.K. Jain, *Vascular permeability in a human tumor xenograft: molecular size dependence and cutoff size.* Cancer Res, 1995. **55**(17): p. 3752-6.

8.   Dellian, M., F. Yuan, V.S. Trubetskoy, V.P. Torchilin, and R.K. Jain, *Vascular permeability in a human tumour xenograft: molecular charge dependence.* Br J Cancer, 2000. **82**(9): p. 1513-8.

9.   Shen, Y., J. Kim, E.F. Strittmatter, J.M. Jacobs, D.G. Camp, 2nd, R. Fang, N. Tolie, R.J. Moore, and R.D. Smith, *Characterization of the human blood plasma proteome.* Proteomics, 2005. **5**(15): p. 4034-45.

10.   Tang, H.Y., N. Ali-Khan, L.A. Echan, N. Levenkova, J.J. Rux, and D.W. Speicher, *A novel four-dimensional strategy combining protein and peptide separation methods enables detection of low-abundance proteins in human plasma and serum proteomes.* Proteomics, 2005. **5**(13): p. 3329-42.

11.   Greenough, C., R.E. Jenkins, N.R. Kitteringham, M. Pirmohamed, B.K. Park, and S.R. Pennington, *A method for the rapid depletion of albumin and immunoglobulin from human plasma.* Proteomics, 2004.

12.   Bjorhall, K., T. Miliotis, and P. Davidsson, *Comparison of different depletion strategies for improved resolution in proteomic analysis of human serum samples.* Proteomics, 2004. **5**(1): p. 307-317.

13.   States, D.J., G.S. Omenn, T.W. Blackwell, D. Fermin, J. Eng, D.W. Speicher, and S.M. Hanash, *Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study.* Nat Biotechnol, 2006. **24**(3): p. 333-8.

14.   Omenn, G.S., D.J. States, M. Adamski, T.W. Blackwell, R. Menon, H. Hermjakob, R. Apweiler, B.B. Haab, R.J. Simpson, J.S. Eddes, E.A. Kapp, R.L. Moritz, D.W. Chan, A.J. Rai, A. Admon, R. Aebersold, J. Eng, W.S. Hancock,

S.A. Hefta, H. Meyer, Y.K. Paik, J.S. Yoo, P. Ping, J. Pounds, J. Adkins, X. Qian, R. Wang, V. Wasinger, C.Y. Wu, X. Zhao, R. Zeng, A. Archakov, A. Tsugita, I. Beer, A. Pandey, M. Pisano, P. Andrews, H. Tammen, D.W. Speicher, and S.M. Hanash, *Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database.* Proteomics, 2005. **5**(13): p. 3226-45.

15. Zhou, M., D.A. Lucas, K.C. Chan, H.J. Issaq, E.F. Petricoin, 3rd, L.A. Liotta, T.D. Veenstra, and T.P. Conrads, *An investigation into the human serum "interactome".* Electrophoresis, 2004. **25**(9): p. 1289-98.

16. Rothemund, D.L., V.L. Locke, A. Liew, T.M. Thomas, V. Wasinger, and D.B. Rylatt, *Depletion of the highly abundant protein albumin from human plasma using the Gradiflow.* Proteomics, 2003. **3**(3): p. 279-87.

17. Pieper, R., Q. Su, C.L. Gatlin, S.T. Huang, N.L. Anderson, and S. Steiner, *Multi-component immunoaffinity subtraction chromatography: an innovative step towards a comprehensive survey of the human plasma proteome.* Proteomics, 2003. **3**(4): p. 422-32.

18. Gundry, R.L., Q. Fu, C.A. Jelinek, J.E. Van Eyk, and R.J. Cotter, *Investigation of an albumin-enriched fraction of human serum and its albuminome.* Proteomics Clin. Appl., 2007. **1**(1): p. 73-88.

19. Granger, J., J. Siddiqui, S. Copeland, and D. Remick, *Albumin depletion of human plasma also removes low abundance proteins including the cytokines.* Proteomics, 2005. **5**(18): p. 4713-8.

20. Lowenthal, M.S., A.I. Mehta, K. Frogale, R.W. Bandle, R.P. Araujo, B.L. Hood, T.D. Veenstra, T.P. Conrads, P. Goldsmith, D. Fishman, E.F. Petricoin, 3rd, and L.A. Liotta, *Analysis of albumin-associated peptides and proteins from ovarian cancer patients.* Clin Chem, 2005. **51**(10): p. 1933-45.

21. Tirumalai, R.S., K.C. Chan, D.A. Prieto, H.J. Issaq, T.P. Conrads, and T.D. Veenstra, *Characterization of the low molecular weight human serum proteome.* Mol Cell Proteomics, 2003. **2**(10): p. 1096-103.

22. Harper, R.G., S.R. Workman, S. Schuetzner, A.T. Timperman, and J.N. Sutton, *Low-molecular-weight human serum proteome using ultrafiltration, isoelectric focusing, and mass spectrometry.* Electrophoresis, 2004(25): p. 1299-1306.

23. Zheng, X., H. Baker, and W.S. Hancock, *Analysis of the low molecular weight serum peptidome using ultrafiltration and a hybrid ion trap-Fourier transform mass spectrometer.* J Chromatogr A, 2006. **1120**(1-2): p. 173-84.

24. Hu, L., X. Li, X. Jiang, H. Zhou, X. Jiang, L. Kong, M. Ye, and H. Zou, *Comprehensive peptidome analysis of mouse livers by size exclusion chromatography prefractionation and nanoLC-MS/MS identification.* J Proteome Res, 2007. **6**(2): p. 801-8.

25. Luque-Garcia, J.L. and T.A. Neubert, *Sample preparation for serum/plasma profiling and biomarker identification by mass spectrometry.* J Chromatogr A, 2007. **1153**(1-2): p. 259-76.

26. Jung, W.W., S. Phark, S. Oh, J.Y. Khim, J. Lee, M.H. Nam, J.B. Seo, S.Y. Park, E. Jo, S. Choi, Z. Zheng, J.Y. Lee, M. Lee, E. Lee, and D. Sul, *Analysis of low*

molecular weight plasma proteins using ultrafiltration and large gel two-dimensional electrophoresis. Proteomics, 2009. **9**(7): p. 1827-40.

27.   Georgiou, H.M., G.E. Rice, and M.S. Baker, *Proteomic analysis of human plasma: failure of centrifugal ultrafiltration to remove albumin and other high molecular weight proteins.* Proteomics, 2001. **1**(12): p. 1503-6.

28.   Greening, D.W., K. Glenister, R. Sparrow, and R.J. Simpson, *Enrichment of human platelet membrane-cytoskeletal proteins for proteomic analysis*, in *Proteomic analysis of membrane proteins: Methods and protocols. Methods in Molecular Medicine Series.*, M. Pierce, Editor. 2009, Humana Press. p. 245-258.

29.   Pieters, M., J.C. Jerling, and J.W. Weisel, *Effect of freeze-drying, freezing and frozen storage of blood plasma on fibrin network characteristics.* Thromb Res, 2002. **107**(5): p. 263-9.

30.   Lewis, M.R., P.W. Callas, N.S. Jenny, and R.P. Tracy, *Longitudinal stability of coagulation, fibrinolysis, and inflammation factors in stored plasma samples.* Thromb Haemost, 2001. **86**(6): p. 1495-500.

31.   Smith, P.K., R.I. Krohn, G.T. Hermanson, A.K. Mallia, F.H. Gartner, M.D. Provenzano, E.K. Fujimoto, N.M. Goeke, B.J. Olson, and D.C. Klenk, *Measurement of protein using bicinchoninic acid.* Anal Biochem, 1985. **150**(1): p. 76-85.

32.   Olsen, J.V., L.M. de Godoy, G. Li, B. Macek, P. Mortensen, R. Pesch, A. Makarov, O. Lange, S. Horning, and M. Mann, *Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap.* Mol Cell Proteomics, 2005. **4**(12): p. 2010-21.

33.   Perkins, D.N., D.J. Pappin, D.M. Creasy, and J.S. Cottrell, *Probability-based protein identification by searching sequence databases using mass spectrometry data.* Electrophoresis, 1999. **20**(18): p. 3551-67.

34.   Greening, D.W., K.M. Glenister, E.A. Kapp, R.L. Moritz, R.L. Sparrow, G.W. Lynch, and R.J. Simpson, *Comparison of human platelet-membrane cytoskeletal proteins with the plasma proteome: Towards understanding the platelet-plasma nexus.* Proteomics Clin. Appl., 2008. **2**: p. 63-77.

35.   Kapp, E.A., F. Schutz, G.E. Reid, J.S. Eddes, R.L. Moritz, R.A. O'Hair, T.P. Speed, and R.J. Simpson, *Mining a tandem mass spectrometry database to determine the trends and global factors influencing peptide fragmentation.* Anal Chem, 2003. **75**(22): p. 6251-64.

36.   Nielsen, H., J. Engelbrecht, S. Brunak, and G. von Heijne, *Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites.* Protein Eng, 1997. **10**(1): p. 1-6.

37.   Harris, M.A., J. Clark, A. Ireland, J. Lomax, M. Ashburner, R. Foulger, K. Eilbeck, S. Lewis, B. Marshall, C. Mungall, J. Richter, G.M. Rubin, J.A. Blake, C. Bult, M. Dolan, H. Drabkin, J.T. Eppig, D.P. Hill, L. Ni, M. Ringwald, R. Balakrishnan, J.M. Cherry, K.R. Christie, M.C. Costanzo, S.S. Dwight, S. Engel, D.G. Fisk, J.E. Hirschman, E.L. Hong, R.S. Nash, A. Sethuraman, C.L. Theesfeld, D. Botstein, K. Dolinski, B. Feierbach, T. Berardini, S. Mundodi, S.Y. Rhee, R. Apweiler, D. Barrell, E. Camon, E. Dimmer, V. Lee, R. Chisholm, P. Gaudet, W. Kibbe, R. Kishore, E.M. Schwarz, P. Sternberg, M. Gwinn, L. Hannick, J. Wortman, M. Berriman, V. Wood, N. de la Cruz, P. Tonellato, P. Jaiswal, T.

Seigfried, and R. White, *The Gene Ontology (GO) database and informatics resource.* Nucleic Acids Res, 2004. **32**(Database issue): p. D258-61.

38. Apweiler, R., A. Bairoch, C.H. Wu, W.C. Barker, B. Boeckmann, S. Ferro, E. Gasteiger, H. Huang, R. Lopez, M. Magrane, M.J. Martin, D.A. Natale, C. O'Donovan, N. Redaschi, and L.S. Yeh, *UniProt: the Universal Protein knowledgebase.* Nucleic Acids Res, 2004. **32**(Database issue): p. D115-9.

39. Croda-Todd, M.T., X.J. Soto-Montano, P.A. Hernandez-Cancino, and E. Juarez-Aguilar, *Adult cystatin C reference intervals determined by nephelometric immunoassay.* Clin Biochem, 2007. **40**(13-14): p. 1084-7.

40. Schenk, S., G.J. Schoenhals, G. de Souza, and M. Mann, *A high confidence, manually validated human blood plasma protein reference set.* BMC Med Genomics, 2008. **1**: p. 41.

41. Villanueva, J., J. Philip, D. Entenberg, C.A. Chaparro, M.K. Tanwar, E.C. Holland, and P. Tempst, *Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry.* Anal Chem, 2004. **76**(6): p. 1560-70.

42. Zhou, B.P., J. Deng, W. Xia, J. Xu, Y.M. Li, M. Gunduz, and M.C. Hung, *Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition.* Nat Cell Biol, 2004. **6**(10): p. 931-40.

43. Villanueva, J., D.R. Shaffer, J. Philip, C.A. Chaparro, H. Erdjument-Bromage, A.B. Olshen, M. Fleisher, H. Lilja, E. Brogi, J. Boyd, M. Sanchez-Carbayo, E.C. Holland, C. Cordon-Cardo, H.I. Scher, and P. Tempst, *Differential exoprotease*

activities confer tumor-specific serum peptidome patterns. J Clin Invest, 2006. **116**(1): p. 271-84.

Figure Legends

Figure 1 - Evaluation of four centrifugal ultrafiltration devices for enriching the plasma LMF

Four membrane filter devices were compared for their ability to fractionate the plasma LMF: *membrane A* - Microcon® YM-30, regenerated cellulose membrane (flat-based); *membrane B* - Centrisart®, polyethersulfone membrane (flat-based); *membrane C* - Amicon Ultra-4®, polyethersulfone membrane (tangential); and *membrane D* - Vivaspin-4®, cellulose triacetate membrane (tangential). The separation efficiency of these four filtration devices were assessed by 1D-SDS-PAGE using a starting volume of 100 µL. Sample load on gel, 30 µg total protein. Proteins were visualized by silver staining using SilverSNAP® III. Lane P, unfractionated plasma; lane R, retentate (retained volume) lyophilized and reconstituted to 20 µL in Laemmli non-reducing sample buffer; lane F, filtrate (volume passed through membrane) lyophilized and reconstituted in non-reducing sample buffer (*Section 2.2*). The ≤ 30K fraction of the gel is indicated by a dashed line (--). Analyzes shown are representative of triplicate experiments.

Figure 2 - Full MS scan and MS/MS spectrum of peptide R.ALDFAVGEYNK.A from cystatin-C (cystatin-3) (neuroendocrine basic polypeptide)

In the LMF, many peptides were singly charged species, often presenting issues during identification and sequence annotation. Specifically, 10 proteins correlated with a single peptide (22%), 9 from two peptides (20%), and 5 from three peptides (11%). An example of a singly charged peptide was identified in cystatin-C, where full MS scan (the

insert is a magnified view of the precursor ion showing relative molecular mass of 1226.61 Da) is shown (A). MS/MS spectrum of this singly charged peptide (+1) is indicated (B). The ProteinScore for the protein and IonScore for the peptide are provided. The identified peptide sequence is <u>ALDFAVGEYNK</u> with terminal peptides R and A.

Figure 3 - Comparison of plasma LMF with human plasma/serum proteome studies

The plasma LMF (44 proteins) was compared with the global plasma proteome datasets (A) and low-M_r plasma/serum studies and proteins/peptides associated with abundant serum proteins (i.e., albuminome) (B). For the high-confidence HUPO Plasma Proteome Project dataset of States et al. [13], 16 proteins were co-identified, in comparison to plasma proteome dataset of Schenk et al. [40], in which 30 proteins were co-identified. In the high-confidence plasma proteome dataset of Greening et al. [34] 17 proteins were co-identified in comparison to the plasma LMF. The plasma LMF dataset was further compared with several low-M_r plasma/serum studies, including the serum peptidome study by Villanueva et al. [43] (6 proteins co-identified), the low-M_r interactome study by Zhou et al. [15] (11 proteins co-identified), and the investigation of interacting low-M_r proteins/peptides with albumin by Gundry et al. [18] (8 proteins co-identified). All proteins were compared between each of these datasets based on a common identifier (Entrez GeneID).

Table 1 – Plasma LMF recoveries using various centrifugal ultrafiltration devices.

Centrifugal ultrafiltration membrane	Protein recoveries[a]		
	Filtrate (mg, %)	Retentate (mg, %)	Overall recovery (mg, %)
Membrane A Microcon®, Millipore (30K)	0.20 (3.0%)	5.6 (83.6%)	5.8 (86.6%)
Membrane B Centrisart®, Sartorius (20K)	2.9 (43.3%)	3.4 (50.7%)	6.3 (94.0%)
Membrane C Amicon Ultra®, Millipore (30K)	1.4 (20.9%)	4.8 (71.6%)	6.2 (92.5%)
Membrane D Vivaspin®, Sartorius (20K)	0.3 (4.0%)	6.1 (91.0%)	6.3 (94.0%)

a   100 µL plasma (67 mg/mL; 6.7 mg) was diluted to 1000 µL (various buffers utilized, refer Materials and methods) and loaded onto each prepared filtration device. Amount of protein recovered in the filtrate and retentate is expressed as a percentage of the initial plasma protein concentration loaded onto each filtration device. The volume of the filtrates and retentates have been adjusted to a total volume of 1000 µL, refer Materials and methods; values shown in parentheses represent the percentage of the initial plasma protein volume loaded onto each filtration device. Overall recovery represents the summation of protein recovery in both the filtrate and retentate for each membrane; values shown in parentheses represent the summation of filtrate and retentate recoveries expressed as a percentile. Each value is representative of experiments performed in triplicate.

Table 2 - The plasma LMF. Proteins identified from the human plasma LMF, with corresponding subcellular localization and unique peptides identified for each protein. Each protein was classified according to SignalP V2.0, with 20 proteins classically secreted according to their signal peptide – these have been corrected for loss of signal peptide (as indicated[†]). Further a comparison with other human plasma/serum proteomic studies is provided. Shaded proteins indicate calculated molecular weights ≤ 25 kDa. For detailed peptide information, refer Supplementary Table 1.

Accession number[a]	Protein description[b]	Unique peptides*[c]	Calculated molecular weight (kDa)[d]	Subcellular localization[e]	HUPO PPP States, 2006 889 Proteins[f]	Plasma Proteome Schenk, 2008 697 Proteins[g]	HUPO PPP Greening, 2008 647 Proteins[h]	Serum LMF Villanueva, 2004 16 Proteins[i]	Albuminome Gundry, 2007 35 Proteins[j]	Serum Interactome Zhou, 2004 210 Proteins[k]
P02647	Apolipoprotein A-I (Apo-AI) (ApoA-I)	5	30.7 (28.7)[†]	Secreted.	Y	Y	Y	Y	Y	Y
P19341	Beta-2-microglobulin (Fragment).	2	4.8	Secreted.						
P61769	Beta-2-microglobulin	6	13.6 (11.4)[†]	Secreted.		Y	Y			
Q9H8T5	CDNA FLJ13244 fis, clone OVARC1000679, highly similar to Homo sapiens myosin-IXa mRNA. (Fragment).	1	61.1	Unknown						
P23528	Cofilin-1 (Cofilin, non-muscle isoform).	3	18.4	Nucleus matrix. Cytoplasm, cytoskeleton.		Y	Y			
P0C0L5	Complement C4-B precursor (Basic complement C4)	12	192.7	Secreted.		Y				

Accession	Protein	Count	MW	Subcellular location						
P00746	Complement factor D precursor (EC 3.4.21.46) (C3 convertase activator)	2	26.9 (24.8)†	Secreted.			Y			Y
P01034	Cystatin-C precursor (Cystatin-3)	6	15.7 (13.0)†	Secreted.		Y	Y			Y
Q15828	Cystatin-M precursor	1	16.4 (13.4)†	Secreted.		Y				
Q14574	Desmocollin-3 precursor (Desmocollin-4)	3	99.9 (96.9)†	Cell membrane; Single-pass type I membrane protein. Cell junction, desmosome. Note=Innermost portion of the desmosomal plaque.						
P15924	Desmoplakin (DP)	3	331.7		Y	Y	Y		Y	
P02671	Fibrinogen alpha chain	29	94.9 (92.9)†	Secreted.	Y	Y	Y	Y	Y	Y
P02751	Fibronectin	1	262.5	Secreted, extracellular space, extracellular matrix.	Y	Y	Y			Y
Q6P5S8	Hypothetical protein.	14	25.7 (22.7)†	Unknown						
P01857	Ig gamma-1 chain C region.	1	36.0	Secreted.		Y				Y
P01593	Ig kappa chain V-I region AG.	2	11.9	Unknown						Y
P01597	Ig kappa chain V-I region DEE.	2	11.6	Unknown						
P01619	Ig kappa chain V-III region B6.	4	11.6	Unknown	Y	Y				

Accession	Protein	#	Mass (kDa)	Location						
P01625	Ig kappa chain V-IV region Len.	2	12.6	Unknown	Y	Y				
P01717	Ig lambda chain V-IV region Hil.	1	11.4	Unknown		Y				
Q569I9	IGKC protein.	12	25.3	Unknown		Y				Y
Q6NS95	IGL@ protein.	8	25.1 (23.0)†	Unknown						
Q8N355	IGL@ protein.	8	24.7 (22.6)†	Unknown	Y	Y				Y
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4	8	103.3 (100.2)†	Secreted.	Y	Y	Y	Y	Y	
P01042	Kininogen-1 precursor (Alpha-2-thiol proteinase inhibitor)	2	71.9 (69.9)†	Secreted, extracellular space.	Y	Y	Y	Y	Y	
P61626	Lysozyme C	2	16.5 (14.5)†	Unknown	Y	Y	Y			
Q9UL85	Myosin-reactive immunoglobulin kappa chain variable region (Fragment).	4	11.7	Unknown						
O00151	PDZ and LIM domain protein 1 (Elfin	5	36.0	Cytoplasm (By similarity). Cytoplasm, cytoskeleton (By similarity).		Y				
P62937	Peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8) (PPIase A)	1	17.9	Cytoplasm.						
Q9NWQ8	Phosphoprotein associated with glycosphingolipid-enriched microdomains 1	1	46.9	Cell membrane; Single-pass type III membrane protein.						

Accession	Protein			Localization							
P02753	Plasma retinol-binding protein precursor (PRBP) (RBP)	19	22.9 (20.9)†	Secreted.	Y	Y	Y				
P02775	Platelet basic protein precursor (PBP) (Small inducible cytokine B7) (CXCL7)	10	13.8 (10.2)†	Secreted.	Y	Y	Y	Y		Y	
P07737	Profilin-I	12	15.0	Unknown		Y	Y				
P41222	Prostaglandin-H2 D-isomerase precursor (EC 5.3.99.2) (Lipocalin-type prostaglandin-D synthase)	2	21.01 (18.6)†	Rough endoplasmic reticulum. Nucleus membrane. Golgi apparatus. Cytoplasm, perinuclear region. Secreted.	Y	Y	Y				
Q9Y4D3	Protein KIAA0649 (Fragment).	1	126.8	Nucleus, nucleolus.							
A0N5G5	Rheumatoid factor D5 light chain (Fragment).	3	12.7	Unknown							
P01011	Serine protease inhibitor A3 (Serpin A3) (Alpha-1-antichymotrypsin)	1	47.6 (45.1)†	Secreted.							
P02787	Serotransferrin precursor (Transferrin)	4	77.0 (74.9)†	Secreted.	Y	Y	Y		Y		
P02768	Serum albumin.	41	69.3 (67.3)†	Secreted.	Y	Y	Y				Y
Q9H299	SH3 domain-binding glutamic acid-rich-like protein 3 (SH3 domain-binding protein 1)	3	10.4	Cytoplasm. Nucleus.		Y					
O75368	SH3 domain-binding glutamic acid-rich-like protein.	1	12.7	Unknown		Y					

P37802	Transgelin-2 (SM22-alpha homolog).	6	22.3	Unknown		Y				
P62988	Ubiquitin.	2	8.5	Cytoplasm. Nucleus.		Y				
P02766	Transthyretin (Prealbumin)	10	15.8 (13.7)†	Secreted.	Y	Y	Y	Y	Y	Y