# Investigation of Non-Specific Protein: Protein Interactions Following Removal of Twenty High Abundance Proteins From Human Plasma

**G. Scott, C. Melm, A. Crawford, H. Chapman, J. Wildsmith, K. Ray, D. Chen, and M. Schuchard.**

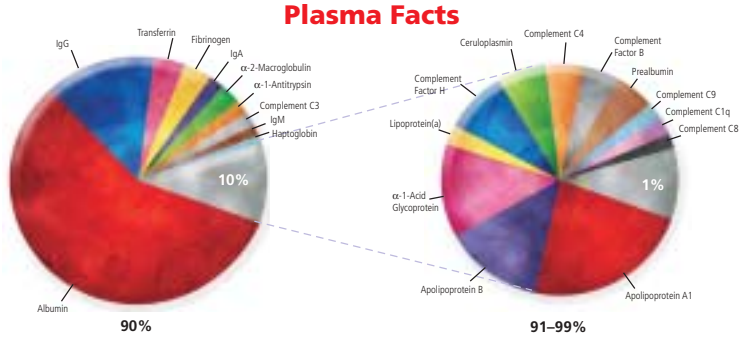
## Abstract

Biomarker discovery has increased interest in the study of the human plasma proteome. Given the wide dynamic range of plasma constituents, high abundance protein depletion has proven necessary to identify medium and low copy number proteins. Non-specific interactions associated with the high abundance fraction require investigation for complete discovery analysis. Human plasma was depleted of 20 high abundance proteins using the ProteoPrep® 20 technology in a prototype large (3.7 ml) format vacuum column. Bound proteins were removed and acetone precipitated following depletion. The bound protein pellet was tryptically digested and separated by 2D (SCX and RP) liquid chromatography. Protein identifications were made via ESI MS/MS. Here, low nonspecific binding to the ProteoPrep 20 resin is demonstrated. Also, the interactions of the nonspecifically bound proteins to those specifically targeted are discussed.

## Introduction

- Depletion of 20 high abundance proteins allows for observation of medium to low abundance proteins as a result of two effects: (1) unmasking of proteins that migrate and co-elute with the high abundance proteins, and (2) increasing the relative protein load of the depleted fraction.
- Protein biochemistry suggests that a number of proteins may be non-specifically associated with the high abundance proteins (e.g. albumin) and these species would thus also be all or partially removed during depletion. Initial data with the ProteoPrep 20 and other depletion technologies indicates the existence of protein-protein interactions, but the extent of these phenomena is unclear.
- Previously, an initial list of nonspecifically-bound proteins was determined using a prototype large format ProteoPrep 20 vacuum column and analyzed using 1D LC (RP separation) MS/MS.
- Here, a more extensive analysis is described using the same depletion technology and large vacuum column format coupled to 2D LC MS/MS. This study should help further define the protein interactions within human plasma as well as define the necessity to study the bound fraction.

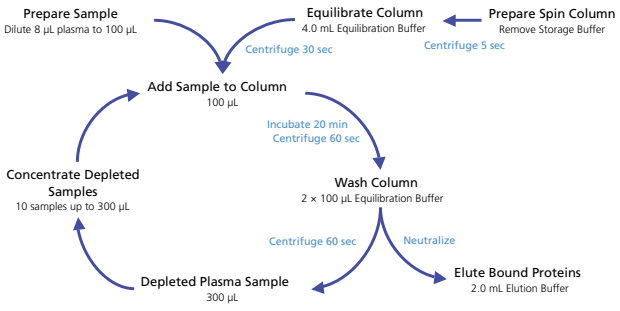
## Plasma Facts




- The ten most abundant proteins represent approximately 90% of the total protein mass in human plasma.
- The 22 most abundant proteins are said to represent approximately 99% of the total protein mass in human plasma.
- The ProteoPrep® 20 Plasma Immunodepletion column removes the 20 high-abundance plasma/serum proteins listed below. These 20 proteins represent approximately 97% of the total human plasma protein mass.

| | | | |
|---|---|---|---|
| Albumin | α-2-Macroglobulin | Apolipoprotein A1 | Complement C4 |
| IgGs | IgMs | Apolipoprotein A2 | Complement C1q |
| Transferrin | α-1-Antitrypsin | Apolipoprotein B | IgDs |
| Fibrinogen | Complement C3 | Acid-1-Glycoprotein | Prealbumin |
| IgAs | Haptoglobin | Ceruloplasmin | Plasminogen |

## ProteoPrep 20 Workflow




## ProteoPrep 20 Plasma Immunodepletion Kit (PROT20)

- Columns, three (each containing 0.3 mL of resin for depletion of 20 high-abundance proteins from 8 µL of human plasma)
- Equilibration Buffer (10× Concentrate)
- Elution Buffer (10× Concentrate)
- Kathon® (for long-term column storage)
- Collection Tubes
- Spin Filters (0.2 µm for plasma clarification)
- Spin Filters (5000 NMWL for concentration)
- Syringes (for column equilibration and elution)
- Luer Lock Caps

## Methods

### High Abundance Protein Depletion

Six (6) high abundance proteins were depleted from citrated plasma (Cat. No. P9523) using a commercially available spin column product (Agilent, Cat No. 5188-5230) according to supplied protocols. Twenty (20) high abundance proteins were depleted from plasma using a 3.7 ml prototype vacuum column containing the ProteoPrep® 20 Plasma Immunodepletion technology resin. Multiple depletions of plasma were carried out similar to that described in the workflow for use of the small columns in the ProteoPrep 20 kit. The depleted plasma from multiple depletions was pooled and concentrated using a 5,000 NMWL filter. The concentrate was then depleted a second time. Concentration and buffer exchange (50 mM ammonium bicarbonate, pH 8.3) of the depleted plasma was carried out using 5,000 NMWL filters (Amicon Ultra, Millipore). Bound proteins were extracted from the resin and acetone precipitated as described.

### Acetone Precipitation (Bound Fraction)

To the PROT20 bound fraction was added five volumes of 100% acetone and incubated at –20 °C overnight. The precipitated protein was pelleted by centrifugation and washed 3 times with 50 % acetone (–20 °C) with centrifugation. The washed protein pellet was air dried at room temperature. The pellet was dissolved with 50 mM ammonium bicarbonate, 9% acetonitrile, pH 8.2.

### Trypsin Digestion

Samples were reduced and alkylated (Cat. No. PROTRA) and digested with Trypsin (Cat. No. T6567) at a concentration of 1% (w/w) and allowed to incubate at 37 °C for 3 hrs. Trypsin was again added (1% w/w) and allowed to incubate at 37 °C overnight. The digests were dried down in a Speed Vac and dissolved with 50 µl 0.1% TFA.

### LC-MS/MS Identification (Bound Fraction)

A sample was injected onto a two dimensional Paradigm MS4N NanoFlow MDLC system (Microm Bioresources) having both SCX (1 × 10 mm, 5 µm, 200 A, poly SulfoEthylA) and RP (0.1 × 150 mm, 5 µm, 200 A, Magic C18) columns. An eleven step salt gradient (0–200 nM ammonium formate) was utilized on the SCX column at 30 min per step.

### LC MS/MS Identification (Depleted Fraction)

Digested samples (8 µl, 76 µg total protein) were separated by reverse phase on a 150 x 2.1 mm, 5 µm Supelco Discovery™ HS C18 column. The separation was performed over 3 hrs with a gradient of formic acidified water and acetonitrile. Tandem mass spectrometric data was aquired using a Thermo Finnegan LTQ via a Data Dependent Acquisition (DDA) method with dynamic exclusion. Data sets were searched using the SEQUEST algorithm on Bioworks 3.2.

### ELISA for High Abundance Plasma Proteins

The percent depletion of human plasma proteins was determined by ELISA. Whole citrated plasma and depleted plasma samples were directly coated onto 96 well ELISA plates overnight following dilution in carbonate buffer (Cat. No. C3041). The plates were washed with TBS-Tween 20 and then incubated with 20 primary antibodies in TBS-BSA for 2 hr. at 37 °C. The plates were washed and then incubated with HRP-conjugated secondary antibodies in TBS-BSA for 2 hr. at 37 °C. Finally, the plates were washed and developed with TMB substrate (Cat. No. T0440), stopped with an equal volume of 1 M HCl and the absorption measured at 450 nm.

## Results

**Non-specifically Bound Proteins**

| Protein Identification | XC Score | Peptides | % Coverage |
|---|---|---|---|
| Apolipoprotein E precursor (Apo E) | 30 | 3 | 15 |
| Apolipoprotein C-III precursor (Apo C-III) | 20 | 2 | 27 |
| Clusterin precursor (Apo J) | 20 | 2 | 8 |
| Fibronectin precursor | 20 | 2 | 2 |
| Hemoglobin beta chain | 20 | 2 | 16 |
| Homeobox protein Cux-2 | 18 | 2 | 2 |
| Bestrophin-2 | 10 | 1 | 5 |
| Collagen alpha 1 | 10 | 1 | 3 |
| Double-stranded RNA-specific editase 1 | 8 | 1 | 2 |
| Ephrin-B1 precursor | 10 | 1 | 6 |
| General transcription factor 3C polypeptide 3 | 10 | 1 | 2 |
| Integrin alpha-6 precursor | 10 | 1 | 2 |
| Myoferlin | 10 | 1 | 1 |
| Myosin light chain kinase 2, skeletal/cardiac muscle | 8 | 1 | 3 |
| Protocadherin Fat 2 precursor | 10 | 1 | 0 |
| RNA-binding protein Raly | 10 | 1 | 5 |
| Serum paraoxonase/arylesterase 1 | 10 | 1 | 5 |
| Sodium channel protein type VII alpha subunit | 10 | 1 | 1 |
| Target of Nesh-SH3 precursor | 10 | 1 | 1 |

**Specifically Bound Proteins**

| | Protein Identification | XC Score | Peptides | % Coverage |
|---|---|---|---|---|
| 1 | Serum albumin precursor | 208 | 21 | 28 |
| 2 | Ig gamma-1 chain C region | 60 | 6 | 29 |
| 2 | Ig gamma-2 chain C region | 10 | 1 | 5 |
| 2 | Ig gamma-3 chain C region | 18 | 2 | 10 |
| 2 | Ig gamma-4 chain C region | 10 | 1 | 5 |
| 3 | Serotransferrin precursor | 98 | 10 | 14 |
| 4 | Fibrinogen alpha chain precursor | 50 | 5 | 9 |
| 4 | Fibrinogen beta chain precursor | 70 | 7 | 18 |
| 4 | Fibrinogen gamma chain precursor | 70 | 7 | 24 |
| 5 | Ig alpha-1 chain C region | 40 | 4 | 14 |
| 5 | Ig alpha-2 chain C region | 10 | 1 | 4 |
| 6 | Alpha-2-macroglobulin precursor | 190 | 19 | 18 |
| 7 | Ig mu chain C region | 10 | 1 | 3 |
| 7 | Ig mu heavy chain disease protein | 20 | 2 | 6 |
| 8 | Alpha-1-antitrypsin precursor | 90 | 9 | 31 |
| 9 | Complement C3 precursor | 250 | 25 | 21 |
| 10 | Haptoglobin precursor | 10 | 1 | 2 |
| 10 | Haptoglobin-related protein precursor | 28 | 3 | 10 |
| 11 | Apolipoprotein A-I precursor | 90 | 9 | 38 |
| 12 | Apolipoprotein A-2 | | | |
| 13 | Apolipoprotein B-100 precursor | 50 | 5 | 2 |
| 14 | Alpha-1-acid glycoprotein 1 precursor | 40 | 4 | 23 |
| 15 | Ceruloplasmin precursor | 60 | 6 | 9 |
| 16 | Complement C4 precursor | 100 | 10 | 9 |
| 17 | Complement C1q subcomponent | 10 | 1 | 7 |
| 18 | Ig delta chain C region | 10 | 1 | 3 |
| 19 | Transthyretin precursor | 60 | 6 | 49 |
| 20 | Plasminogen precursor | 10 | 1 | 1 |

**Figure 1: Low nonspecific binding to PROT20 is demonstrated using 2D LC-MS/MS.**

- Of the 20 specifically bound proteins, 19 were positively identified by MS. Apolipoprotein-A2 was not specifically detected.
- Six (6) non-specifically bound proteins were identified with 2 or more peptides.
- Thirteen (13) non-specifically bound proteins were identified with 1 peptide.
- Several of the non-specifically bound proteins (e.g. Fibronectin, Clusterin and Serum paroxonase) were also identified in the depleted plasma (Figure 3).

Bound proteins eluted from a PROT20 high capacity column were acetone-precipitated and trypsin-digested as indicated in the Methods section. A 200 µg sample was injected onto a two dimensional LC system as described in the Methods section. Protein identifications were made from SEQUEST search and filtered as follows: Xcorr values 1.9, 2.4, and 2.9 for singly, doubly and triply charged ions, with deltaCN value ≥ 0.1, and protein probability of 99.9%.

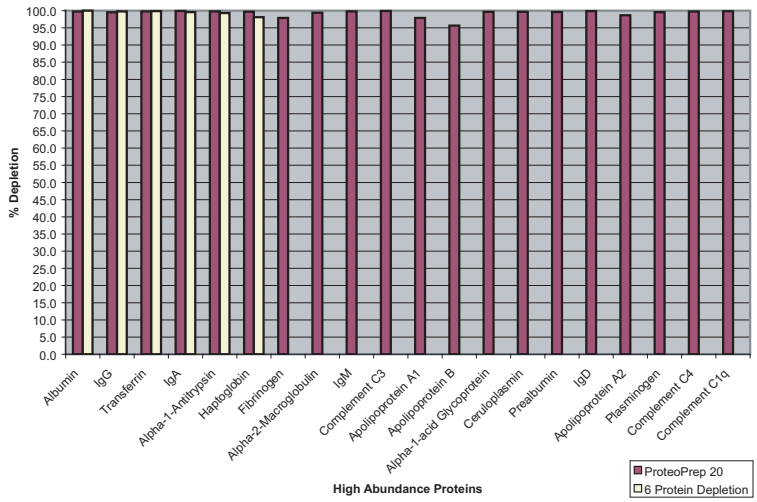



**Figure 2: The novel antibody resin (ProtePrep 20 technology) displays high binding capacity for 20 human plasma proteins.**

- Twenty high abundance proteins were removed with an average depletion of 99.3% when 2 x 100 µl plasma depletions were concentrated and depleted a second time.
- Six high abundance proteins were removed with an average depletion of 99.4% when 15 x 7 µl plasma were depleted using spin columns.
- The total amount of protein removed using these 2 technologies was approximately 94% (ProteoPrep 20) and 81% (6 protein depletion), as determined by Bradford Assay (Cat. No. B6916).

Two (2) initial depletions of 100 µl plasma were carried out using the 3.7 ml column as described in the workflow for use of the ProteoPrep 20 technology. The depleted plasma from both depletions was pooled and concentrated using a 5,000 NMWL filter. The concentrate was then depleted a second time. Fifteen (15) depletions of 7 µl plasma were carried out using the 6 protein depletion spin column as described in the Methods section. An ELISA assay was carried out on all 20 proteins as described in the Methods section.

**Panel A–Depleted Plasma**

| | PROT20 Depleted | | 6 Protein Depleted | | Whole Plasma | |
|---|---|---|---|---|---|---|
| Protein Identification | Peptide | % Coverage | Peptide | % Coverage | Peptide | % Coverage |
| Fibronectin precursor | 27 | 16 | 18 | 11 | 4 | 2 |
| Alpha-2-HS-glycoprotein precursor, Fetuin-A | 22 | 31 | 7 | 28 | 2 | 6 |
| Vitamin D-binding protein precursor | 18 | 36 | 15 | 31 | 7 | 17 |
| Apolipoprotein A-IV precursor | 18 | 43 | 9 | 28 | 1 | 4 |
| Inter-alpha-trypsin inhibitor heavy chain H2 precursor | 17 | 17 | 8 | 13 | 2 | 4 |
| Hemopexin precursor | 16 | 28 | 11 | 18 | 10 | 25 |
| Alpha-1B-glycoprotein precursor | 11 | 24 | 11 | 23 | 3 | 8 |
| Plasma protease C1 inhibitor precursor | 11 | 15 | 4 | 11 | 7 | 10 |
| Inter-alpha-trypsin inhibitor heavy chain H4 precursor | 10 | 13 | 5 | 7 | 1 | 1 |
| Afamin precursor | 10 | 14 | 2 | 4 | 1 | 3 |
| Beta-2-glycoprotein I precursor, Apo-H | 8 | 19 | 8 | 26 | 6 | 16 |
| Alpha-1-antichymotrypsin precursor | 7 | 16 | 4 | 12 | 2 | 7 |
| Angiotensinogen precursor | 5 | 16 | 5 | 17 | 1 | 4 |
| Carboxypeptidase N subunit 2 precursor | 4 | 7 | 3 | 7 | 2 | 9 |
| Vitronectin precursor | 4 | 12 | 2 | 6 | 3 | 7 |
| AMBP protein precursor | 4 | 10 | 1 | 3 | 1 | 3 |
| Complement factor B precursor | 17 | 19 | 10 | 14 | | |
| Kininogen-1 precursor | 14 | 18 | 12 | 18 | | |
| Gelsolin precursor | 14 | 20 | 7 | 11 | | |
| Complement factor H precursor | 13 | 11 | 5 | 4 | | |
| Prothrombin precursor | 13 | 21 | 3 | 6 | | |
| Inter-alpha-trypsin inhibitor heavy chain H1 precursor | 10 | 13 | 4 | 7 | | |
| Pigment epithelium-derived factor precursor | 8 | 16 | 6 | 11 | | |
| Histidine-rich glycoprotein precursor | 5 | 10 | 2 | 5 | | |
| Serum amyloid P-component precursor | 4 | 17 | 5 | 24 | | |
| Clusterin precursor | 4 | 10 | 3 | 11 | | |
| Heparin cofactor II precursor | 4 | 7 | 2 | 4 | | |
| Complement component C9 precursor | 3 | 6 | 1 | 2 | | |
| Serum paraoxonase/arylesterase 1 | 2 | 7 | 3 | 9 | | |
| C4b-binding protein alpha chain precursor | 2 | 2 | 2 | 2 | | |
| Complement component C8 gamma chain precursor | 2 | 13 | 1 | 7 | | |
| Complement factor I precursor | 2 | 4 | 1 | 2 | | |
| Adapter-related protein complex 4 beta 1 subunit | 1 | 2 | 3 | 5 | | |
| Sex hormone-binding globulin precursor | 1 | 6 | 1 | 6 | | |
| Plasma retinol-binding protein precursor | 1 | 5 | 1 | 5 | | |
| Complement C1s subcomponent precursor | 1 | 2 | 1 | 2 | | |
| Antithrombin-III precursor | 8 | 19 | | | | |
| N-acetylmuramoyl-L-alanine amidase precursor | 6 | 17 | | | | |
| Complement C5 precursor | 6 | 4 | | | | |
| Alpha-2-antiplasmin precursor | 4 | 11 | | | | |
| Zinc-alpha-2-glycoprotein precursor | 3 | 13 | | | | |
| Heat-shock protein beta-7 | 2 | 10 | | | | |
| Coagulation factor XII precursor | 2 | 5 | | | | |
| Complement C2 precursor | 2 | 3 | | | | |
| Tetranectin precursor | 1 | 6 | | | | |
| Activator 1 140 kDa subunit | 1 | 2 | | | | |
| Galectin-3 binding protein precursor | 1 | 2 | | | | |
| Kallistatin precursor | 1 | 4 | | | | |
| Complement component C6 precursor | 1 | 2 | | | | |
| Corticosteroid-binding globulin precursor | 1 | 3 | | | | |
| Leucine-rich alpha-2-glycoprotein precursor | 1 | 3 | | | | |
| Nuclear pore complex protein Nup98-Nup96 precursor | | | 2 | 1 | 2 | 1 |
| Apolipoprotein C-III precursor | | | 2 | 27 | 2 | 27 |
| Periplakin | | | 2 | 2 | | |
| Myosin Ic | | | 2 | 4 | | |
| Ninein | | | 1 | 1 | | |
| Probable G-protein coupled receptor 109B | | | 1 | 7 | | |
| Complement factor H-related protein 3 precursor | | | 1 | 3 | | |
| Ankyrin repeat and SOCS box protein 12 | | | 1 | 10 | | |
| Neuronal acetylcholine receptor protein | | | 1 | 4 | | |
| Type 2 lactosamine alpha-2,3-sialyltransferase | | | | | 3 | 7 |
| Zinc finger protein 205 | | | | | 1 | 6 |
| Tyrosine-protein kinase BTK | | | | | 1 | 3 |

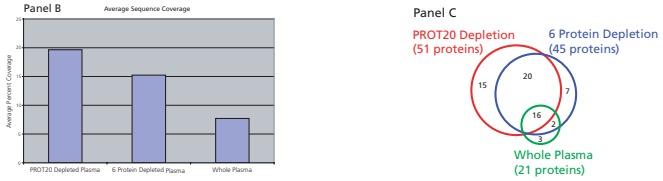



**Figure 3. Depleted Plasma—Depletion of more proteins increases the number of protein identifications and allows for increased sequence coverage.**

- The number of proteins identified was higher with the PROT20 depleted sample (Panels A and C).
- The sequence coverage was generally higher with the PROT20 depleted sample (Panel B).

PROT20 depleted plasma, 6 protein depleted plasma and whole plasma (380 µg) was trypsin digested overnight and dried using a vacuum centrifuge. The samples were prepared as described in the Methods section. Protein identifications were made via SEQUEST search and filtered as follows; Xcorr values ≥ 1.9, 2.4, and 2.9 for singly, doubly and triply charged ions, with deltaCN value ≥ 0.1, and protein probability of 99.9%. Panel A lists the lower abundance proteins identified in the 3 samples with the number of different peptides detected and their sequence coverage. The 20 higher abundance proteins are not listed. Panel B is a graph of the average sequence coverage for the 16 proteins in Panel A which were identified in all 3 samples. The Venn Diagram in Panel C illustrates the overlap of the 3 datasets.

## Conclusions

- Only a small number of non-specifically bound proteins are detected with two or more peptides in the bound fraction indicating that non-specific binding to the resin is low (Figure1).
- Apo E, Apo C-III and Apo J are likely associated with the lipoprotein complexes, brought down through depletion of Apolipoproteins A1, A2, and B.
- Hemoglobin and Fibronectin are likely identified in the bound fraction because of covalent attachment to Apolipoprotein B and Lipoprotein (a) respectively.
- The interaction of Homeobox protein Cux-2 with the specifically targeted proteins is unknown.
- Several of the non-specifically bound proteins are also detected in the depleted plasma.
- The novel PROT20 antibody resin displays high depletion capability (average 99%) for the 20 human plasma proteins (Figure 2).
- Depletion of 20 proteins generates more protein ID's and improved sequence coverage compared with the depletion of 6 proteins (Figure 3).

## References


1. Anderson, N. L. and Anderson, N. G. The Human Plasma Proteome. *Mol. Cell. Proteomics* **2002**, *1*, 845.
2. Adkins, J. N. et al. Toward a Human Blood Serum Proteome. *Mol. Cell. Proteomics* **2002**, *1*, 947.
3. Omenn, G. S. et al. 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*, 3226.
4. Gonos, E.S. and Trougalkos, I.P. Clusterin/Apolipoprotein J in Human Aging and Cancer. Int. J. Biochem. Cell Biol. **2002**, 34, 1430.
5. Ziouzenkova, O. et al. Oxidative Cross-linking of ApoB100 and Hemoglobin Results in Low Density Lipoprotein Modification in Blood. J. Biol. Chem. **1999**, 274, 18916.
6. Salonen, E-M. et al. Lipoprotein (a) Binds to Fibronectin and has Serine Proteinase Activity Capable of Cleaving it. EMBO J. **1989**, 8, 4035.
7. Gingras, H. et al. Biochemical Characterization of the Mammalian Cux2 Protein. Gene **2005**, 344, 273.


SIGMA-ALDRICH