

High Specificity Depletion of Twenty High-Abundance Proteins from Human Plasma Facilitates Examination of Plasma Proteome Differences Between Individuals

G. B. I. Scott, A. S. Crawford, H. A. Chapman, J. Wildsmith, S. L. Cockrill, K. B. Ray, R. J. Mehig, D. Chen, M. D. Schuchard, C. D. Melm

Sigma-Aldrich
Life Science and Technology Center
2909 Laclede Ave., St. Louis, MO 63103



Abstract

Identification of disease biomarkers has significantly increased study of the human plasma proteome. Unfortunately, disease biomarkers typically appear at low concentrations. Another challenge associated with the plasma proteome is the large dynamic range of individual protein concentrations (ten orders of magnitude). Therefore, identification of low-copy-number proteins of interest is difficult due to the confounding presence of higher-abundance proteins. An antibody-based resin has been developed which depletes 20 high-abundance proteins from human plasma. This depletion technology uses a mixture of small single-chained recombinant antibody ligands along with conventional affinity purified polyclonal antibodies. Depletion of 99% of these 20 high-abundance proteins removes 97–98% of total plasma protein. Following 2DE or LC separation, greater visualization of proteins that typically comigrate with, or are masked by, the high-abundance proteins was observed. Depletion also allows the remaining proteins to be loaded at a 20–50 times higher level for improved identification of lower-abundance proteins. This increase was found to be three times higher than when six proteins were removed. This increased protein load should help fuel the discovery of proteins of biological or pharmaceutical significance. We have demonstrated the utility of this technology in discerning differences in the plasma proteome between multiple individuals from different gender and ethnic backgrounds. Analysis of both the depleted plasma samples and the eluted bound protein fraction showed increased low-abundance protein visualization and high specificity for proteins removed from the plasma samples.

Introduction

- The study of the human plasma proteome is an area of great interest, especially for the pharmaceutical potential of identifying disease biomarkers. Many proteins of pharmaceutical interest appear at low concentrations in the plasma and are therefore difficult to detect.
- Identification of potential biomarkers is especially difficult due to the presence of higher-abundance proteins. Depletion of high-abundance proteins allows for visualization of proteins that comigrate with, and are masked by, the high-abundance proteins on 1DE or 2DE gels. Plasma proteins can then be loaded onto the gels or IPG strips at higher levels for improved visualization of low-copy-number proteins.
- An affinity resin has been developed for removal of 20 high-abundance proteins from 8 μ L of plasma. Depletion of these 20 high-abundance proteins removes approximately 97% of total plasma protein, and permits loading of 30– to 50-fold more of each individual protein for improved visualization of lower-copy-number proteins.
- Here we have evaluated three sources of human plasma samples by 2DE following depletion to identify proteomic differences.
- Nonspecifically bound proteins were also evaluated.

Methods

High-Abundance Protein Depletion

Twenty (20) high-abundance proteins were depleted from plasma using the ProteoPrep[®] 20 Plasma Immunodepletion Technology. A large 3.7-mL prototype spin column for depleting 100 μ L of plasma was used. The plasmas used included a pooled plasma source (Cat. No. P9523), Asian American Female, and Asian Indian Male. Concentration of multiple depletions was carried out using 5000-NMWL filters (Cat. No. M0286 or Millipore Amicon Ultra 4). Protein concentration of whole and depleted plasma was determined using the Bradford Assay (Cat. No. B6916) using BSA as the standard.

Two-Dimensional Electrophoresis (2DE)

Whole citrated plasma samples or depleted plasma (using PROT20) were employed. Salts were removed by buffer exchange using a 5000-NMWL filter. The samples were then used to dissolve the IPG strip rehydration reagent (Protein Extraction Reagent Type 4, Cat. No. C0356) and reduced and alkylated using PROTRA (Tributylphosphine and Iodoacetamide). IPG strips (Cat. No. I3531, 11 cm, pH 4–7) were rehydrated with the samples and focused overnight (85,000 Vhr). The strips were equilibrated for 15 min with IPG Equilibration Buffer (Cat. No. I7281) and loaded onto 8–16% SDS-PAGE gels with IPG wells. The gels were electrophoresed at 180 V for 1.25 h. The marker lanes contain SigmaMarker Wide Range (Cat. No. S8445). The second dimension gel was fixed and stained with EZBlue. The gels were imaged using a UMAX PowerLook 2100XL at 400 dpi. Gel images were then analyzed using Phoretix 2D Expression software from Nonlinear Dynamics.

Acetone Precipitation (for Bound Proteins)

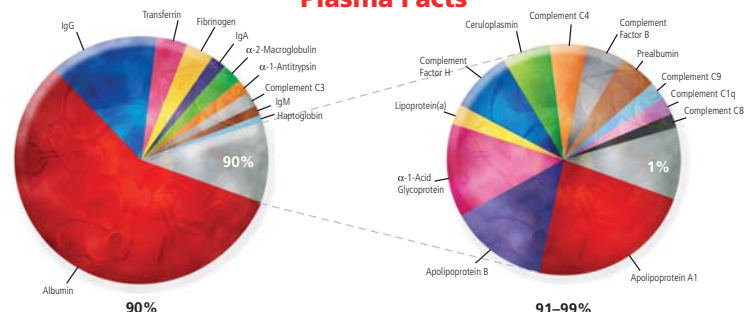
To the PROT20-bound fraction was added five volumes of 100% acetone and the samples incubated at -20°C overnight. The protein was pelleted by centrifugation and washed 3 times with 50% acetone (-20°C) with centrifugation. The washed protein pellet was air dried at room temperature. The pellet was dissolved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

Trypsin Digestion (for Bound Proteins)

The sample was 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 μ L 0.1% TFA. Samples of the digests (5 μ L) were injected onto a 15 cm \times 2.1 mm C18 RPLC Supelco Discovery column (5- μ m bead size) and then into the LTQ mass spectrometer. The LC run was 100 μ L/min for 200 min.

Results

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

Proteome Comparison Between Three Plasma Sources

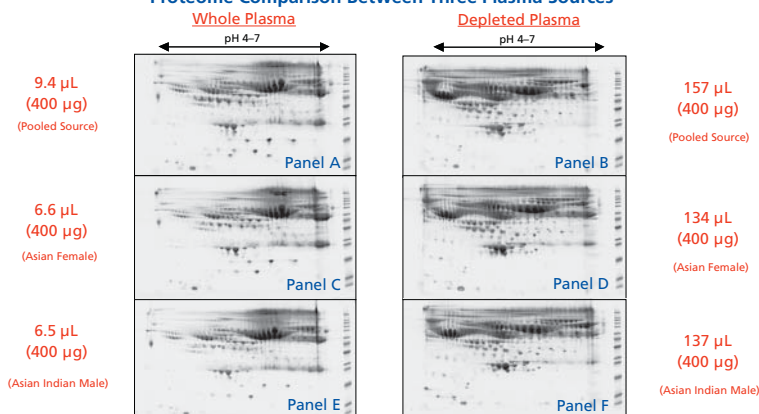


Figure 1. Depletion of 20 proteins allows for increased loading of the lower-abundance proteins.

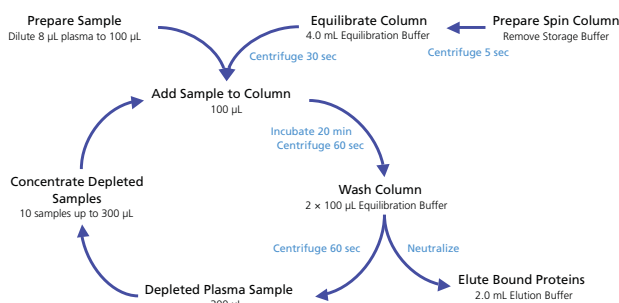
- Depletion of 20 proteins unmasks many areas of the gel.
- Unmasking allows for visualization of other comigrating proteins and reduces potential contamination of spots, which can interfere with mass spectrometric identification.
- Depletion of 20 proteins allowed for an increase of approximately 20-fold in the load of low-abundance proteins over whole plasma (undepleted).

A 400- μ g sample of whole pooled plasma and PROT20 depleted plasmas (pooled, Asian female, and Asian Indian male) were concentrated using 5000-NMWL spin filters as described in the Methods section. Two-dimensional electrophoresis was carried out on all four samples in duplicate as described in the Methods section. Protein concentration was determined by Bradford Assay. The original plasma volume for 400 μ g of each sample was 9.4 μ L, 6.6 μ L, and 6.5 μ L for whole plasma and 157 μ L, 134 μ L, and 137 μ L for PROT20 depleted plasma from the pooled, Asian female, and Asian Indian male plasma samples, respectively. The gels were imaged as described in the Methods section.

ProteoPrep[®] 20 Plasma Immunodepletion Kit (PROT20)

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

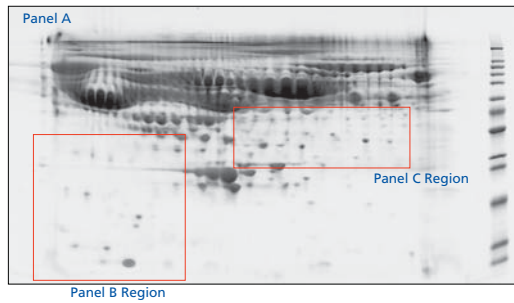
ProteoPrep[®] 20 Workflow



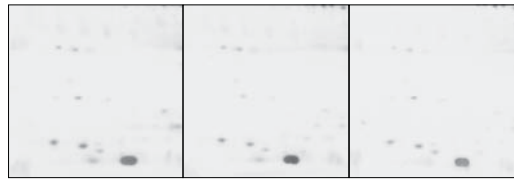
Proteome Differences Observed Between Three Plasma Sources

Nonspecifically Bound Proteins

Depleted Plasma, Pooled Source



Whole Plasma



Depleted Plasma

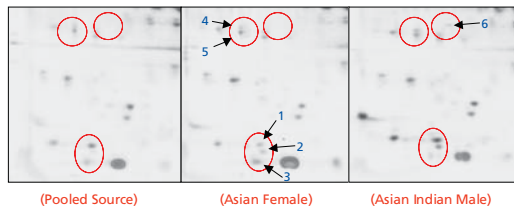


Figure 2, Panel C

Whole Plasma



Depleted Plasma

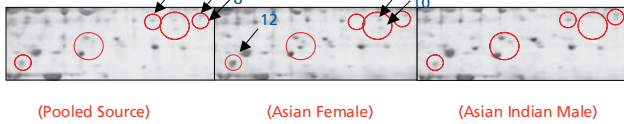


Figure 2. Proteome differences observed between three plasma sources.

- Two regions of the gels were highlighted in which expression differences were observed.
- Several protein spots in each region show significant (> twofold) differences in expression levels. These are differences which would not have been detected in the whole plasma.

Panel A illustrates the regions of the whole plasma and depleted plasma gels, which are highlighted in Panels B and C.

Protein Identification	XC Score	Number of Peptides	Coverage
haptoglobin-related protein	78	8	16
similar to hypothetical protein	50	5	16
similar to hypothetical protein	40	4	20
similar to hypothetical protein	40	4	5
pregnancy-zone protein	20	2	2
A-gamma globin, hemoglobin gamma-a chain	10	1	7
alpha 2 globin	10	1	11
alpha-2-HS-glycoprotein	10	1	3
angiotensin I converting enzyme isoform 2 precursor	8	1	2
apolipoprotein L1 isoform a precursor	10	1	6
arginyl-tRNA synthetase	10	1	2
ATP-binding cassette, sub-family F, member 1	10	1	1
beta globin	10	1	6
bradykinin receptor B2	4	1	5
cadherin EGF LAG seven-pass G-type receptor 1	10	1	1
cholinergic receptor, nicotinic, alpha polypeptide 5	2	1	3
chromosome 14 open reading frame 93	10	1	3
chromosome 15 open reading frame 16	8	1	2
clusterin, complement-associated protein SP-40	10	1	3
complement component 2 precursor, C3/C5 convertase	2	1	2
cyclin-dependent kinase 8	10	1	3
D111ap1e-like	8	1	2
golgi phosphoprotein 4	10	1	2
golgi-specific brefeldin A resistance factor 1	10	1	1
HIRA interacting protein 3	8	1	2
histidine-rich glycoprotein precursor	10	1	3
hypothetical protein DKFZp54780714	6	1	2
hypothetical protein FLJ35867	6	1	1
hypothetical protein FLJ36144	8	1	2
hypothetical protein MGC40084	10	1	3
hypothetical protein XP_353145	10	1	12
integrin alpha 7 precursor	10	1	1
isocitrate dehydrogenase 3 (NAD+) alpha precursor	10	1	3
kallikrein 14 preproprotein	6	1	5
keratin 15	8	1	2
KIAA0409 protein	10	1	4
KIAA1354 protein	8	1	3
kinesin family member 14	6	1	1
lymphocyte-specific protein tyrosine kinase, oncogene LCK	10	1	3
membrane component, chromosome 11, surface marker 1	8	1	2
mitochondrial ribosomal protein S17	10	1	17
mitogen-activated protein kinase kinase 5 isoform A	8	1	3
nebulin	4	1	0
NEFA-interacting nuclear protein NIP30	8	1	4
neuroepithelial cell transforming gene 1	10	1	2
nuclear receptor subfamily 1, group H, member 3	4	1	3
nucleoporin 214kDa	10	1	1
phospholipid transfer protein isoform a precursor	8	1	4
plectin 1, intermediate filament binding protein 500kDa	8	1	0
PREX1 protein	8	1	1
Propionyl-Coenzyme A carboxylase, alpha polypeptide precursor	10	1	2
proteasome alpha 4 subunit, proteasome component C9	8	1	4
proteoglycan 4	8	1	1
rhomboid-related protein 2	8	1	6
rhylin 2	8	1	2
ryanodine receptor 2	6	1	0
secretogogin precursor	10	1	8
serologically defined colon cancer antigen 1	10	1	0
similar to 1-beta glycin	10	1	0
similar to C-terminal binding protein 2 isoform 2	8	1	3
similar to glutamate receptor	10	1	9
similar to KIAA0542 protein	6	1	1
similar to KIAA1272 protein	10	1	1
similar to Prostaglandin F2 receptor negative regulator precursor	8	1	1
similar to protease (prosome, macropain) 265 subunit, ATPase 1	10	1	3
Smad nuclear interacting protein	8	1	2
sphingosine kinase 1	10	1	4
synaptonemal complex protein 3	10	1	3
testis-specific leucine zipper protein nurit	8	1	4
TGF-beta-induced factor 2 (TALE family homeobox)	10	1	6
transforming, acidic coiled-coil containing protein 3	6	1	1
v-myb myeloblastosis viral oncogene homolog	8	1	2
v-raf-1 murine leukemia viral oncogene homolog 1	10	1	3
zinc finger protein 262, cell death inhibiting RNA	8	1	1
zinc finger protein 274 isoform b, KRAB zinc finger protein HFB101	8	1	2

Table 1. Bound proteins from PROT20 were evaluated using LC-MS/MS.

- Five (5) nonspecifically bound proteins were identified with two or more peptides (shaded area).
- Seventy (70) nonspecifically bound proteins were identified with one peptide.

Bound proteins eluted from the PROT20 resin were acetone-precipitated and trypsin-digested as indicated in the Methods section. A sample was injected onto a C18 column and separated with a 200-min acetonitrile gradient and the peptides identified with a LTQ LC-MS/MS (Thermo). The HUPO filter (Xcorr values ≥ 1.9 , 2.2, and 3.75 for singly, doubly, and triply charged ions with deltaCN value ≥ 0.1 and Rsp ≥ 4) was used for determination of protein identifications.

Conclusions

- Depletion of 20 high-abundance proteins from human plasma greatly improves the ability to visualize lower-abundance proteins.
- This novel antibody resin displays high depletion capability (average 99%) for the 20 proteins from human plasma.
- Nonspecifically bound proteins are known to associate with several of these depleted high-abundance proteins. Several of these nonspecifically bound proteins have been identified. Further identification of these proteins is ongoing.
- Depletion of 20 high-abundance proteins permits greater loading capacity and visualization of low-abundance proteins for electrophoretic and/or chromatographic separation prior to mass spectrometry.

References

- Anderson, N. L. and Anderson, N. G. The Human Plasma Proteome. *Mol. Cell. Proteomics* **2002**, *1*, 845.
- Adkins, J. N. et al. Toward a Human Blood Serum Proteome. *Mol. Cell. Proteomics* **2002**, *1*, 947.
- Rengarajan, K. et al. Removal of Albumin from Multiple Human Serum Samples. *BioTechniques* **1996**, *20*, 30.
- 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.

Expression Differences Between the Plasma Sources

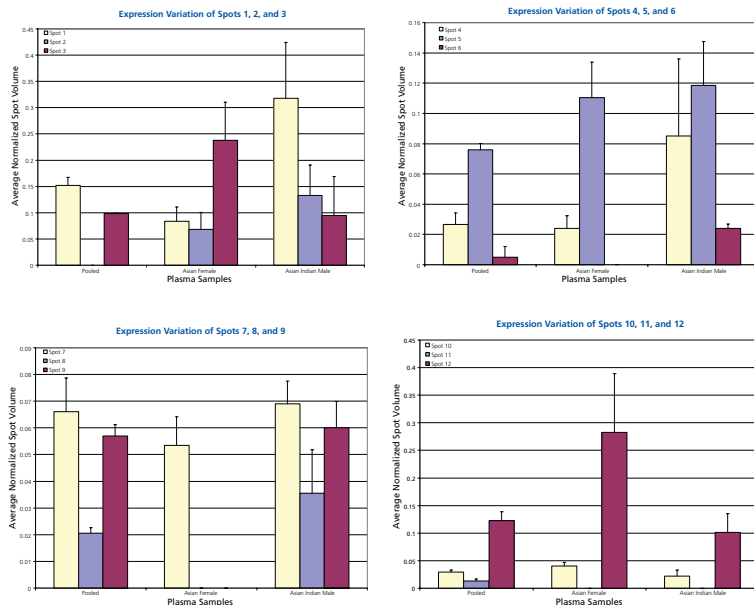


Figure 3. Graphical representation of proteome differences between three plasma sources.

- Among the 12 spots highlighted, significant (> twofold) differences are seen between the three plasma samples analyzed.
- These are differences which would not have been detected in the whole plasma.

The expression level differences were detected and determined using the Phoretix 2D Expression software (Nonlinear Dynamics).