High Dynamic Range Characterization of the Trauma Patient Plasma Proteome*§

Tao Liu‡, Wei-Jun Qian‡, Marina A. Gritsenko‡, Wenzhong Xiao§, Lyle L. Moldawer¶, Amit Kaushal§, Matthew E. Monroe‡, Susan M. Varnum‡, Ronald J. Moore‡, Samuel O. Purvine‡, Ronald V. Maier¶, Ronald W. Davis§, Ronald G. Tompkins**, David G. Camp Il‡, Richard D. Smith‡ ‡‡, and the Inflammation and the Host Response to Injury Large Scale Collaborative Research Program§§

Although human plasma represents an attractive sample for disease biomarker discovery, the extreme complexity and large dynamic range in protein concentrations present significant challenges for characterization, candidate biomarker discovery, and validation. Herein we describe a strategy that combines immunoaffinity subtraction and subsequent chemical fractionation based on cysteinyl peptide and N-glycopeptide captures with two-dimensional LC-MS/MS to increase the dynamic range of analysis for plasma. Application of this "divide-and-conquer" strategy to trauma patient plasma significantly improved the overall dynamic range of detection and resulted in confident identification of 22,267 unique peptides from four different peptide populations (cysteinyl peptides, non-cysteinyl peptides, N-glycopeptides, and non-glycopeptides) that covered 3654 different proteins with 1494 proteins identified by multiple peptides. Numerous low abundance proteins were identified, exemplified by 78 "classic" cytokines and cytokine receptors and by 136 human cell differentiation molecules. Additionally a total of 2910 different N-glycopeptides that correspond to 662 N-glycoproteins and 1553 N-glycosylation sites were identified. A panel of the proteins identified in this study is known to be involved in inflammation and immune responses. This study established an extensive reference protein database for trauma patients that provides a foundation for future high throughput quantitative plasma proteomic studies designed to elucidate the mechanisms that underlie systemic inflammatory responses. Molecular & Cellular Proteomics 5:1899–1913, 2006.

The search for disease-specific biomarkers from various human biofluids (e.g. plasma/serum (1–3), cerebrospinal fluid (4), bronchoalveolar lavage fluid (5), synovial fluid (6), nipple aspirate fluid (7), saliva (8), and urine (9)) is gaining increasing attention due to significant advances in genomic and proteomic technologies and their potential for discovering novel disease biomarkers. An advantage of using human plasma for biomarker discovery is that plasma samples are readily obtained from patients. More importantly, because almost all cells in the body communicate with blood either directly or indirectly and tissue-specific proteins are released into the bloodstream upon cell damage and cell death, the onset or presence of most disease states can potentially be determined by measuring either the altered presence or the differential abundances of proteins in the blood plasma.

Although progress has been made toward quantitatively studying the plasma proteome (10–13), significant effort still remains focused on broadly characterizing the plasma protein constituents. These constituents include a handful of major proteins such as albumin that dominate the plasma protein content as well as other tissue leakage proteins that are generally present only at extremely low levels. As a result, robust and specific detection of the broad spectrum of proteins that span such a wide dynamic range in concentration challenges current analytical technologies. Various separation approaches (e.g. prefractonation of intact plasma proteins (14–17), utilization of solution isoelectric focusing (18), and ultrahigh efficiency separations (19, 20) at the peptide level and multidimensional fractionation on both the protein and peptide levels (21)) have been applied to address this challenge. Efforts have also been made to characterize subsets of the plasma proteome (e.g. the low molecular weight subproteome (22) and the N-glycoproteome (23)). Additionally the removal of albumin, immunoglobulins, and a group of high abundance proteins in plasma/serum has been demonstrated to be very effective for attaining more comprehensive proteome coverage (20, 23–26). Recently developed multicomponent immunoaffinity subtraction systems, based on mammalian (27, 28) and avian (29) antibodies, are now available for highly efficient, reproducible, and reversible depletion of high abundance proteins in plasma/serum samples. The common
concern of potentially losing a subset of nontargeted proteins with these methods by way of protein-protein interactions (e.g. the “sponge effects” of albumin (30)) can be addressed by studying the “interactomes” (29).

Herein we report on an alternative strategy for profiling the human plasma proteome that combines immunopurification subtraction and subsequent chemical fractionation via solid-phase capture of cysteinyl peptides and N-glycopeptides with two-dimensional (2D) LC-MS/MS to extend the dynamic range of analysis and increase proteome coverage. The advantages afforded by dividing a complex proteome into several subproteomes include 1) reduced complexity of the subproteome samples (particularly enriched cysteinyl peptide and N-glycopeptide samples) that allows more low abundance proteins to be identified, 2) the complementary nature of different subproteome fractions that significantly improves overall proteome coverage, and 3) the relative simplicity, efficiency, and reproducibility afforded by these fractionation methods that make them amenable to automation. Application of this strategy to characterize blood plasma obtained from human trauma patients resulted in broad proteome coverage that included many low abundance proteins (e.g. cytokines) and proteins involved in both inflammation, a hallmark of many human diseases (31, 32), and the immune response. Previous inflammation studies have relied heavily on gene expression analysis of blood leukocyte populations obtained from trauma patients (33, 34). An extensive reference plasma proteome database for trauma has been established from this study, providing the basis for subsequent comparative quantitative proteomic studies aimed at revealing the underlying mechanisms of inflammatory response in trauma patients.

EXPERIMENTAL PROCEDURES

Immunopurification Subtraction—The human blood plasma samples were supplied by the Department of Surgery at the University of Florida College of Medicine, which serves as the Sample Collection and Coordination Site for a multicentered clinical study (Inflammation and the Host Response to Injury). Approval for the conduct of this programmatic research was obtained from the Institutional Review Boards of the University of Florida College of Medicine and the Pacific Northwest National Laboratory in accordance with federal regulations. Plasma samples were pooled from six severe trauma patients and one healthy subject, and the protein concentration was determined by Coomassie protein assay (Pierce) to be 60 mg/ml. Twelve plasma proteins (albumin, IgG, α1-antitrypsin, IgA, IgM, transferrin, haptoglobin, α1-acid glycoprotein, α2-macroglobulin, apolipoprotein A-I, apolipoprotein A-II, and fibrinogen) that constitute up to 96% of the total protein mass of human plasma were removed in a single step by using the prepacked 2-ml Seppro™ MIXED12 affinity LC column (GenWay Biotech, San Diego, CA) on an Agilent 1100 series HPLC system (Agilent, Palo Alto, CA) according to the manufacturer’s instructions. A total of 2625 µl (25 µl × 105 injections) of plasma was subjected to MIXED12 depletion. The flow-through fractions were pooled and concentrated under reduced pressure to one-tenth of the original volume and then desalted by using a prepacked PD-10 column (Amersham Biosciences) equilibrated with 50 mM NH4HCO3. The desalted sample was further concentrated, and the total protein amount was determined by Coomassie protein assay to be 8.85 mg.

Plasma Protein Digestion—One-third of the proteins from the depleted plasma sample were directly subjected to trypsic digestion prior to cysteinyl peptide enrichment (CPE). The proteins were denatured and reduced in 50 mM Tris buffer (pH 8.2), 8 mM urea, 10 mM DTT for 1 h at 37 °C. The resulting protein mixture was diluted 10-fold with 20 mM Tris buffer (pH 8.2), and then sequencing grade modified porcine trypsin (Promega, Madison, WI) was added at a trypsin:protein ratio of 1:50 (w/w). The sample was incubated overnight at 37 °C. The following day, the trypsic digest sample was loaded on a 1-ml solid phase extraction (SPE) C18 column (Supelco, Bellefonte, PA) and washed with 4 ml of 0.1% TFA, 5% ACN. Peptides were eluted from the SPE column with 1 ml of 0.1% TFA, 80% ACN and lyophilized. Peptide samples were stored at −80 °C until CPE treatment.

Fractionation via Capture of Cysteinyl Peptides—Unless otherwise noted, all solutions used in this step were degassed to prevent oxidation of the thiol content. The trypsic digest was reduced with 5 mM DTT in 80 µl of 50 mM Tris buffer (pH 7.5), 1 mM EDTA (coupling buffer) for 30 min at 37 °C after which the samples were diluted to 400 µl by adding coupling buffer and then split into four 100-µl aliquots. Thiopropyl-Sepharose 6B thiol affinity resin (4 × 100 µl; Amersham Biosciences) was prepared from dried powder according to the manufacturer’s instruction. Briefly the dried powder was rehydrated in water for 15 min and washed by 50 bed volumes of water followed by 50 bed volumes of coupling buffer in a Handee Mini-Spin column (Pierce). The reduced peptide sample was then incubated with the resin for 1 h at room temperature with gentle mixing, and the unbound portion (non-cysteinyl peptides) was collected by spinning the column at low speed. The resin was washed 6 times in the spin column sequentially with each of the following solutions: 0.5 ml of 50 mM Tris buffer (pH 8.0), 1 mM EDTA (washing buffer); 2 mM NaCl; 80% ACN, 0.1% TFA (solution; and washing buffer. To release the captured cysteinyl peptides, 100 µl of 20 mM DTT freshly prepared in washing buffer was added to the resin and incubated for 30 min at room temperature. The resin was further washed with 100 µl of 80% ACN. The pH of the combined eluates was adjusted to 8.0, and the sample was alkylated with 80 mM iodoacetamide for 30 min at room temperature and in the dark. Both the eluted cysteinyl peptide samples and the unbound non-cysteinyl peptide samples were desalted by using an SPE C18 column and then lyophilized.

Fractionation via Capture of N-Linked Glycopeptides—Hydrazide resin (Bio-Rad) and a method similar to that reported previously (23, 35) were used for capturing glycoproteins. Two-thirds of the concentrated MIXED12 flow-through fraction was diluted 10-fold in coupling buffer (100 mM sodium acetate and 150 mM NaCl, pH 5.5) and oxidized in 15 mM sodium periodate at room temperature for 1 h in the dark with constant shaking. The sodium periodate was subsequently removed by using a prepacked PD-10 column equilibrated with coupling buffer. Four 1-ml portions of hydrazide resin were each washed five times with coupling buffer; the oxidized protein sample was split into four equal aliquots, and then each aliquot was added to a portion of resin and incubated overnight at room temperature. The following
day, the non-glycoproteins were removed by washing the resin briefly three times with 100% methanol and then three times with 8 M urea in 0.4 M NH₄HCO₃. The glycoproteins were denatured and reduced in 8 M urea and 10 mM DTT at 37 °C for 1 h. Protein cysteiny1 residues were alkylated with 20 mM iodoacetamide for 90 min at room temperature. After washing sequentially with 8 M urea and 50 mM NH₄HCO₃, the resin was resuspended as a 20% slurry in 50 mM NH₄HCO₃ after which sequencing grade trypsin (Promega) was added at a 1:100 (w/w) trypsin:protein ratio (based on the initial plasma protein concentration of 60 mg/ml). The sample was then digested on-resin overnight at 37 °C. After centrifuging at 3000 × g for 5 min, the trypsin-released peptides (non-glycopeptides) were collected from the supernatants and lyophilized. The resin was further washed extensively with the following three different solutions: 2 M NaCl, 100% methanol, and 50 mM NH₄HCO₃. The resin was resuspended as a 50% slurry in 50 mM NH₄HCO₃ and the N-glycopeptides were released by incubating the resin with peptide N-glycosidase F (PNGase F; New England Biolabs, Beverly, MA) at a ratio of 1 µl of PNGase F/100 µl of plasma for 4 h at 37 °C. The released deglycosylated peptides were cleaned by using an SPE C₁₈ column before being lyophilized.

Strong Cation Exchange (SCX) Peptide Fractionation—All of the enriched cysteiny1 peptides and deglycosylated peptides and 300 µg each of non-cysteiny1 peptides and non-glycopeptides were individually reconstituted with 300 µl of 10 mM ammonium formate (pH 3.0), 25% ACN and fractionated by SCX chromatography on a Polysulfoethyl A 200 × 2.1-mm column (PolyLC, Columbia, MD) that was preceded by a 10 × 2.1-mm guard column. The separations were performed on an Agilent 1100 series HPLC system (Agilent) at a flow rate of 200 µl/min and with mobile phases that consisted of 10 mM ammonium formate (pH 3.0), 25% ACN (A) and 500 mM ammonium formate (pH 6.8), 25% ACN (B). After loading 300 µl of sample onto the column, the gradient was maintained at 100% A for 10 min. Peptides were separated by using a gradient from 0 to 50% B over 40 min followed by a gradient of 50–100% B over 10 min. The gradient was then held at 100% B for 10 min. A total of 30 fractions were collected for each peptide population, and each fraction was dried under vacuum. The fractions for each population were dissolved in 30 µl of 25 mM NH₄HCO₃, and 10 µl of each fraction were analyzed by capillary LC-MS/MS.

Reversed-phase Capillary LC-MS/MS Analyses—A custom-built high pressure capillary LC system (36) coupled on line to a linear ion trap mass spectrometer (LTQ; ThermoElectron) via an in-house-manufactured electrospray ionization interface was used to analyze peptide samples. The reversed-phase capillary column was prepared by slurry-packing 3-µm Jupiter C₁₈ bonded particles (Phenomenex, Torrance, CA) into a 65-cm-long, 150-µm-inner diameter × 360-µm-out diameter fused silica capillary (Polymicro Technologies, Phoenix, AZ) that incorporated a retaining stainless steel screen in an HPLC union (Valco Instruments Co., Houston, TX). The mobile phases, which consisted of 0.2% acetic acid and 0.05% TFA in water (A) and 0.1% TFA in 90% ACN, 10% water (B), were degassed on line by using a vacuum degasser (Jones Chromatography Inc., Lakewood, CO). After loading 10 µl of peptides onto the column, the mobile phase was held at 100% A for 20 min. An exponential gradient elution was achieved by increasing the mobile phase composition in a stainless steel mixing chamber from 0 to 70% B over 150 min. To identify the eluting peptides, the linear ion trap mass spectrometer was operated in a data-dependent MS/MS mode (m/z 400–2000) in which each full MS scan was followed by 10 MS/MS scans. The 10 most intense precursor ions were dynamically selected in order of highest to lowest intensity and then subjected to collision-induced dissociation; a normalized collision energy setting of 35% and a dynamic exclusion duration of 1 min were used. The temperature of the heated capillary and the ESI voltage were 200 °C and 2.2 kV, respectively.

Data Analysis—The SEQUEST (37) algorithm (ThermoFinnigan) was used to independently search all MS/MS spectra against the human International Protein Index (IPI) database (Version 3.05 that consists of 49,161 protein entries; available on line at www.ebi.ac.uk/IPI) and the reversed human IPI protein database. Tandem MS peaks were generated by extract_msn.exe, part of the SEQUEST software package. Dynamic carboxamidomethylation of cysteine and oxidation of methionine were used to identify cysteiny1 peptides, non-cysteiny1 peptides, and non-glycopeptides; another PNGase F-catalyzed conversion of asparagine to aspartic acid at the site of carbohydrate attachment was added to identify the formerly N-linked glycopeptides.

Because false positive peptide/protein identifications are a common concern in proteomic investigations, we developed a set of criteria based on the reversed database approach for filtering the raw data to limit false positive identifications to <5%. The reversed human protein database was created by reversing the order of the amino acid sequences for each protein, and the percentage of false positive peptide identifications was estimated as described previously (38) by dividing the number of peptides identified from the reversed database search by the number of peptides identified from the normal database search. Only the cysteine-containing and the NXS(T)/T motif-containing peptides were used to estimate false positive identifications for cysteiny1 peptides and formerly N-linked glycopeptides, respectively. (An in silico assessment of the number of cysteine-containing and NXS(T)/T motif-containing peptides among all tryptic peptides shows the number of these kinds of peptides in the normal and reversed databases to be similar; thus, the number of random false positive identifications is assumed to be similar for both databases.)

Criteria that would yield an overall confidence of over 95% at the unique peptide level were established for filtering raw peptide identifications. Table I summarizes the cross-correlation score (Xcorr) values for the different peptide populations; a delta correlation (ΔCn) value of 0.10 was used to determine these values. Two additional ΔCn cutoff values of 0.05 and 0.16 were applied to reduce false negative identifications while maintaining a 95% level of confidence for peptide assignments. For ΔCn ≥ 0.05, the minimum acceptable Xcorr value was raised to achieve a comparable percentage of false positive rate identifications, and similarly for ΔCn ≥ 0.16, the minimum acceptable Xcorr value was reduced. In an attempt to remove redundant protein entries, Protein Prophet software was used as a clustering tool to group similar or related protein entries into a “protein group” (39). All peptides that passed the filtering criteria were assigned an identical probability score of 1 and entered into the software program (solely for clustering analysis) to generate a final list of nonredundant proteins or protein groups. One protein identification was randomly selected to represent each corresponding protein group that contains member database entries.

Gene Ontology (GO) component, function, and process terms extracted from text-based annotation files downloaded from the European Bioinformatics Institute ftp site (ftp.ebi.ac.uk/pub/databases/GO/goa/HUMAN) were used to categorize the identified proteins. Further biological interpretation in the context of gene ontology was performed with the aid of GOstat (40). Protein transmembrane helices were determined by TMHMM, a prediction model based on a hidden Markov model (www.cbs.dtu.dk/services/TMHMM/).

Pathway and network analysis was performed by using the Ingenuity Pathways Knowledge Base (IPKB) (41). Canonical signal transduction and metabolic pathways were examined among the proteins identified. Queries were also made to obtain cellular localization information of the detected proteins from the IPKB. The significant functional networks among the extracellular proteins were computed as described elsewhere (41).
ELISA—Specific plasma protein levels in the pooled trauma patient plasma sample were determined in duplicate using Quantikine ELISA kits (R&D Systems, Minneapolis, MN) following the manufacturer’s instructions. Each sample was analyzed at two different dilutions to determine the optimal dilution. A plasma sample from a healthy subject obtained from Golden West Biologicals (Temecula, CA) was also assayed for comparison.

RESULTS

High Dynamic Range Proteomic Profiling Strategy—Fig. 1 schematically illustrates the sample processing and fractionation used to characterize the trauma patient plasma proteome. High abundance proteins (H) were first removed using immunoaffinity subtraction. The resulting less abundant proteins (L) were split and submitted individually for solid-phase cysteiny peptide and N-glycopeptide captures. Non-cysteiny peptides and non-glycopeptides generated at the same time were also collected. All four different peptide populations were then fractionated by SCX chromatography, and each fraction was analyzed by capillary LC-MS/MS.

Peptide and Protein Identifications—30 SCX fractions of each peptide population were analyzed by LC-MS/MS, and the resulting spectra were searched against the human IPI protein database by using SEQUEST. Only those peptide identifications that passed the filtering criteria in Table I were considered as confident identifications. The numbers of confident peptide and protein identifications obtained from the four different populations are summarized in Table II. A total of 22,267 unique peptides (81.8% fully tryptic) were identified from the four different peptide populations, corresponding to 3654 nonredundant proteins, including 1494 multiprotein proteins (Supplemental Table 1) and 2160 single peptide proteins (Supplemental Table 2). Note that the non-cysteiny peptide sample contributed the largest set of peptide identifications (9951), and the N-glycopeptide fraction contributed the smallest set of peptide identifications (2910).

Of the 3654 proteins identified from the pooled trauma patient plasma, 1341 proteins (36.7%) were present in more than one peptide population, and 330 proteins (9.0%) were observed in all four peptide populations (Fig. 2B), demonstrat-
ing the complementary nature of the four different types of peptide identifications for obtaining higher proteome coverage. 662 distinct N-glycoproteins were identified from the 2910 N-glycopeptides. The overlap between peptides identified in the four peptide populations (Fig. 2A) shows that the cysteinyl peptide and N-glycopeptide fractions share only a small percentage of identifications with the other peptide fractions. Although the non-cysteinyl peptide and non-glycopeptide fractions exhibit much greater overlap in peptide identifications, these fractions still provided additional proteome coverage (e.g. 523 proteins only from non-glycopeptides and 729 proteins only from non-cysteinyl peptides). Interestingly the same sequences for 12.5% of the N-glycopeptides were also found in 94 cysteinyl peptides, 58 non-glycopeptides, and 260 non-cysteinyl peptides, indicating that these N-glycosylation sites are only partially modified and that both N-glycosylated and unmodified peptides from different fractions are detected.

Confidence of Identifications —To illustrate the stringency of our approach for limiting peptide false discovery rate (FDR) to <5%, the filtering criteria developed in this study were compared with two sets of published criteria (21, 42) used previously to filter plasma/serum MS/MS data. The three different sets of criteria were applied to filter the SEQUEST identifications for all four peptide populations, and FDRs resulting from each set of criteria were assessed using the reversed database approach (38). The results of this evaluation are summarized in Table III. Our filtering criteria generated the smallest number of peptide and protein identifications, consistent with the significantly lower percentage of false positive identifications (~4%). Using the criteria recommended by the Human Proteome Organization (HUPO) (42), peptide identifi-
cations increased 1.4-fold, and protein identifications increased 2.2-fold; this yielded a lower percentage of proteins with multiple peptide “hits” and increased the overall FDR at the peptide level to ~25%. The application of the criteria recently reported by Hood et al. (21) resulted in 18,958 non-redundant protein identifications; however, FDR at the pep-

| Filter                | Peptides identified | Proteins identified | Multipepptide proteins | Average peptides per protein | FDR |
|-----------------------|--------------------|---------------------|------------------------|-----------------------------|-----|
| Reversed database     | 22,267             | 3,654               | 1,494 (40.9%)          | 6.1                         | 4   |
| HUPO                  | 30,524             | 7,928               | 2,850 (35.9%)          | 3.9                         | 25  |
| Hood et al. (21)      | 66,839             | 18,958              | 11,653 (61.5%)         | 3.5                         | 66  |

* Non-redundant protein identifications.

b Calculated at unique peptide level based on reversed database evaluation (38).

![Reference concentrations for 80 selected proteins identified from trauma patient plasma.](image)

**Figure 3.** Reference concentrations for 80 selected proteins identified from trauma patient plasma. The protein concentration is plotted on a log scale spanning 11 orders of magnitude. The reference concentrations were obtained from Refs. 1 and 11, and all data were based on healthy human subjects. It should be noted that the actual detected concentrations for the trauma sample may be significantly different from these reference concentrations. The peptide population from which those proteins were identified was labeled using the following scheme: triangles, cysteinyl peptides; squares, non-cysteinyl peptides; circles, N-glycopeptides; diamonds, non-glycopeptides. Cys, cysteinyl peptides; Non-Cys, non-cysteinyl peptides; N-Glyco, N-glycopeptides; Non-Glyco, non-glycopeptides; PEDF, pigment epithelium-derived factor; ACE, angiotensin-converting enzyme; ICAM, vascular cell adhesion molecule; ICAM, intercellular adhesion molecule; TIMP, tissue inhibitor of metalloprotease; HCC, hemofiltrate CC chemokine; EGF, epidermal growth factor; MMP, matrix metalloprotease; CXCL16, CXC chemokine ligand 16; IFN, interferon; CTACK, cutaneous T-cell-attracting chemokine; IL-1RA, interleukin 1 receptor antagonist; 6CKine, 6-cysteine chemokine; Flt-3L, fms-like tyrosine kinase 3 ligand.
tide level was ~66%. The common feature between our criteria and the HUPO criteria is that only fully and partially tryptic peptides are considered; however, the HUPO criteria use much lower Xcorr cutoff values for the partially tryptic peptides. The criteria used by Hood et al. (21) are similar to two other previously reported sets of criteria (22, 25) in which chymotryptic and elastic peptides as well as those peptides that had high Xcorr scores but no proteolytic end were also considered. Clearly our filtering criteria are significantly more stringent when compared with the two sets of previously published criteria. Furthermore the cysteinyl peptide and N-glycopeptide identifications were specifically enriched through reliable solid-phase capture, which further increases the overall confidence level for our data.

The confidence of our current filtering criteria was also compared with the independent statistical model Peptide Prophet program (43) by analyzing the same subset of data-sets applying both approaches. Histograms of Xcorr distributions in peptides identified by reversed database approach showed that low FDR can be achieved using filtering criteria listed in Table I (Supplemental Fig. 1). Similarly high confidence peptide identification can be obtained at a certain Peptide probability ($p_{\text{comp}}$) level (Supplemental Fig. 2). 6279 unique peptides were identified from the reversed database approach compared with 6341 unique peptides from Peptide Prophet ($p_{\text{comp}}$ = 0.95) among which 5973 (~95%) peptides were found in common. Note that the estimated FDR predicted by Peptide Prophet is 0.3% ($p_{\text{comp}}$ ≥ 0.95), further supporting the stringency of the current reversed database filtering criteria. After processing of both sets of peptides by Protein Prophet, a total of 1301 proteins and 1254 proteins were identified for the reversed database approach and Peptide Prophet, respectively. The high percentage of overlap at the peptide level and the comparable numbers of protein identifications reflect the overall high quality of the currently reported data.

**Dynamic Range of Detection and the Detection of Low Abundance Proteins**—Fig. 3 plots reference protein concentrations under normal conditions for 80 selected proteins identified in this study as well as their peptide distribution within the four peptide populations. It is obvious that most of the high abundance proteins and proteins having a midrange concentration (e.g. from mg/ml to ng/ml level) were identified from multiple peptide populations, whereas many of the low abundance proteins (e.g. cytokines) were identified from only one peptide population. It is speculated that some protein concentrations in trauma patients may be significantly different from those under normal conditions as a result of inflamm-

### Table IV

Concentrations of selected proteins in the pooled trauma patient sample and a normal plasma sample measured by ELISA

| Protein   | Concentration | Standard curve | r² value |
|-----------|---------------|----------------|----------|
| Trauma PDGF-B | 2,807         | Not detectable | 0.9972   |
| Normal PDGF-B | Not detectable | Not detectable | 0.9936   |
| Trauma IL-1RA | 19,666        | 1,179          |          |
| Normal IL-1RA | Not detectable | Not detectable |          |
| Trauma VEGF R1 | 556           | 232            | 0.9987   |
| Normal VEGF R1 | Not detectable | Not detectable | 0.9975   |
| Trauma M-CSF | 1,042         | 157            |          |
| Normal M-CSF | Not detectable | Not detectable |          |
| Trauma CCL21 | 1,165         | 1,115          | 0.9873   |
| Normal CCL21 | Not detectable | Not detectable |          |
| Trauma TNF R1 | 6,014         | 945            | 0.9989   |

**Fig. 4.** Representative MS/MS spectra of low abundance proteins identified in this study. The top panel shows a deglycosylated peptide identified from M-CSF (1.0 ng/ml), and the bottom panel shows a cysteinyl peptide identified from small inducible cytokine A21 (1.1 ng/ml).
### TABLE V

**Categorization of 78 classic cytokines and cytokine receptors identified in this study**

Cytokine nomenclature and categorization are from R&D Systems and Ref. 57. LT-H925, lymphotoxin beta; TRAIL, TNF-related apoptosis-inducing ligand; DR6, death receptor 6; GM-CSF, granulocyte macrophage colony-stimulating factor; IL-1RA, interleukin 1 receptor antagonist; IL-1R AP, interleukin 1 receptor accessory protein; IL-18 BP, interleukin 18-binding protein; HGF, hepatocyte growth factor; MSP, macrophage-stimulating protein; Met, hepatocyte growth factor receptor; PIGF, placenta growth factor; EGF, epidermal growth factor; OB, leptin; OSM, oncostatin M; CNTF, ciliary neurotrophic factor; SCF, stem cell factor; Flt-3L, fms-like tyrosine kinase 3 ligand; c-kit, stem cell factor receptor; MK, midkine; PTN, pleiotrophin; RPTP, receptor-type tyrosine-protein phosphatase; BMP, bone morphogenetic protein; GDF, growth/differentiation factor; Inh C, inhibin C chain; Inh E, inhibin E chain; GFR, glia-derived neurotrophic factor receptor; Act, activin; NK cells, natural killer cells; PMNs, polymorphonuclear leukocytes; HCC-1, HCC-4, 6Ckine, Eot-2, CTACK, CXC chemokine ligand 16; MIF, macrophage migration inhibitory factor; GH R, growth hormone receptor; IFN, interferon; TLR, toll-like receptor.

| Family               | Features                                                                 | Cytokines                                                                 | Receptors                                                                 |
|----------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------|
| **TNF**              | Contain highly conserved carboxyl-terminal domains; can induce receptor trimerization influencing signaling pathways | TNF RI (p55), TNF RII (p75), LT-β R, Fas (CD95), TRAIL R3, CD27, TNF R21 (DR6) |                                                                           |
| γc                   | Receptors contain a common γ chain (γc)                                   | IL-2                                                                      |                                                                           |
| IL-4 and IL-13       | Bind to shared heteromultimeric receptor complexes                         | IL-13 Rε1                                                                 |                                                                           |
| βc                   | Receptors contain a common β chain (βc)                                   | IL-5 Rα, GM-CSF Rα                                                        |                                                                           |
| **IL-1**             | A superfamily of proteins comprised of IL-1 and IL-18 members             | IL-1 RI, IL-1 RII, IL-1R AP, IL-18 BP                                     |                                                                           |
| **IGF**              | Share sequence homology with the insulin family of proteins               | IGF1a, IGF2                                                               | IGF RI                                                                   |
| **HGF and MSP**      | Contain a 4-kringle domain and a pseudo-serine protease domain that lacks enzyme activity |                                                                           |                                                                           |
| **FGF**              | Heparin-binding polypeptides                                              |                                                                           |                                                                           |
| Neurotrophic factors | Induce signal transduction through ligand-induced dimerization and activation of trk receptors |                                                                           |                                                                           |
| PDGF, VEGF, PIGF     | Dimeric angiogenic factors containing an 8-cysteine motif                 | PDGF-B, VEGF-D                                                            |                                                                           |
| **EGF**              | Contain at least one extracellular EGF structural unit (conserved 6-cysteine motif that forms three disulfide bonds) |                                                                           |                                                                           |
| Gp130                | Receptors of the IL-6 superfamily are homologous to or contain the gp130 subunit as the common signaling component | Ob                                                                         | Ob R, OSM R, IL-6 Rα, IL-6 Rβ (gp130), CNTF Rα                           |
| SCF, Flt-3L, M-CSF   | Contain a 4-helix bundle structure in the extracellular domain and 4 conserved cysteines; receptors are tyrosine kinases | Flt-3L, M-CSF                                                             |                                                                           |
| MK and PTN           | Products of retinoic acid-responsive genes; developmentally regulated molecules |                                                                           |                                                                           |
| **TGF-β**            | Contain a highly conserved 7-cysteine domain that forms a characteristic cysteine knot | TGF-β1, TGF-β3, BMP-1, BMP-3, GDF-2, Inh β C, Inh β E                       |                                                                           |
| **CC chemokine**     | Target population include: lymphoid cells (T-cells and NK cells), monocytes and PMNs | HCC-1, MIP-16, HCC-4, 6Ckine, Eot-2, CTACK,                                |                                                                           |
| **CXC chemokine**    | Target population includes predominantly PMNs but also T- and B-cells      | PF4, CXCL16                                                               |                                                                           |
| **Unique**           |                                                                           | IL-32, MIF                                                                | GH R, IFN-α/β Rα, IFN-α/β R8, IL-10 Rα, IL-10 Rβ, IL-17 R, IL-23 R, IL-31 Rα, TLR4 |
monoclonal antibodies and used to characterize the cell surface immunophenotypes of different leukocyte subpopulations (e.g. B-cells, helper T-cells, cytotoxic T-cells, and natural killer cells). In 90.4% of the identified CD antigens each had at least one predicted transmembrane domain, whereas among all other proteins, only 15.5% had predicted transmembrane domain(s). This finding is consistent with the fact that the majority of CD antigens are believed to be membrane-associated molecules.

**GO and Pathway Analysis of the Detected Proteins**—Fig. 5 shows the categories of proteins identified from this study in terms of cellular location based on gene ontology analysis. Comparison of cellular components for N-glycoproteins and the other proteins identified shows major differences. The majority of N-glycoproteins (Fig. 5A) are predicted to be extracellular/secreted proteins (38.8%) and membrane-associated proteins (48.8%), whereas all other proteins (Fig. 5B) are predicted to distribute more evenly across all cellular locations. None of the N-glycoproteins identified are from the nucleus, cytoplasm, mitochondrion, ribosome, proteasome, and cytoskeleton; this is consistent with the biological functions of N-linked glycoproteins (23). The high percentage of intracellular proteins in this dataset indicates that large numbers of proteins present in plasma may result from different levels of cellular leakage.

Further GOstat analyses that compared the distribution of GO terms of identified proteins with the entire human IPI database revealed over- and under-represented molecular functions and biological processes (data not shown). Over-represented molecular function categories included hematopoietin/interferon class cytokine receptor activity, insulin-like growth factor (IGF) binding, VEGF receptor activity, metallopeptidase activity, protease inhibitor, extracellular matrix structural constituent, lipid binding and transporter activity, polysaccharide binding, receptor protein kinase activity, and oxidoreductase activity. In the GO comparison of biological processes, proteins involved in response to wound, regulation of body fluids, complement activation, and proteolysis categories appeared over-represented among the proteins identified. These findings reflect certain distinguishing features of the trauma patient plasma proteome, e.g. the presence of...
many inflammation- and immune response-related proteins, exemplified by the proteins listed in Table VI. These proteins include acute phase reactants, cytokines and growth factors, complement proteins and coagulation factors, hormones, extracellular matrix proteins, cell adhesion molecules, and secreted proteases and protease inhibitors in addition to other proteins and immunoglobulins.

Pathway analysis revealed the significant representation of specific signaling pathways, e.g. NF-κB signaling (inflammation and immune regulation), apoptosis signaling, ERK (extracellular signal-regulated kinase)/MAPK (mitogen-activated protein kinase) signaling, and Wnt/β-catenin signaling (data not shown). As an example, Fig. 6A shows a global representation of the extracellular proteins that are involved in immune response comprising a network of 193 proteins and their interactions. (A larger version of this image may be viewed online (Supplemental Fig. 3).) In total, 113 of the 193 (58.5%) known players were identified. Fig. 6B further illustrates the coverage of specific regions of this network and highlights the IGF and IGF-binding proteins, laminins, and matrix metalloproteinases.

### DISCUSSION
Specific biomarkers for diagnosis/prognosis of disease and for monitoring disease progression and response to therapy...
have been clinically applied to screen patient tissues and blood samples as well as used to develop therapeutics and segment the population for specific treatments (44). Proteomics is increasingly being used in this field to describe and enumerate the systematic changes in the protein constituency of a cell, to generate lists of proteins that change in expres-

Fig. 6. Network representation of the extracellular pathways that are involved in immune response. A, the network consists of 193 genes with 113 identified in this study. B, selected regions of the network highlight several groups of proteins. Group 1, IGF and IGF-binding proteins (IGFBP). Group 2, laminins (LAM). Group 3, matrix metalloproteinases (MMP). A larger version of this image may be viewed online (Supplemental Fig. 3).
sion as a cause or consequence of disease, and more importantly to characterize the information flow through the intra- and extracellular molecular protein networks that interconnect organ and circulatory systems. These networks are expected to provide new targets for therapeutics and to reveal the dynamic biological changes that give rise to new candidate biomarkers (45). Because of its constant perfusion through tissues within the body, blood plasma is anticipated to contain ample information regarding these networks and therefore provides a basis for candidate disease biomarker discovery. However, several intrinsic features of plasma, such as an enormous dynamic range in concentrations of protein of interest and extreme sample complexity and heterogeneity, hamper effective proteomic analysis.

Our strategy for analyzing blood plasma addresses these issues by combining multicomponent immunoaffinity subtraction and multiple chemical fractionations (Fig. 1) with 2D LC-MS/MS. The single step depletion of 12 high abundance proteins on an automated LC system significantly increases the dynamic range of detection and reduces sample heterogeneity (due to the simultaneous removal of the highly variable IgG, IgA, and IgM populations). The high efficiency CPE step further reduces sample complexity and in turn enables detection of low abundance proteins (46, 47). Simultaneous analysis of the non-cysteinyl peptides, generated as a “byproduct” during CPE, significantly increases proteome coverage (47, 48). The N-glycopeptide enrichment step affords yet another effective way of reducing plasma sample complexity (23, 35, 49). N-Glycosylation is particularly prevalent in proteins that are secreted and located on the extracellular side of the plasma membrane and in proteins that are contained in various body fluids (e.g. blood plasma) (50). Because the N-glycosylation sites generally fall into a consensus N\((X)(S/T)\) sequence motif, where \(X\) represents any amino acid residue except proline (51), the availability of \(N\)-glycopeptides from each protein is limited; this provides the basis for reducing sample complexity. The performance of the immunoaffinity subtraction MIXED12 column has been reported previously to be specific, efficient, and reproducible (29). In our study, the 12 immunoabsorption-targeted proteins were specifically and effectively removed, but peptides from these proteins were still detected by MS following the depletion step (data not shown) presumably due to the initial high concentrations of these proteins. Binding of other nontarget proteins to the MIXED12 column did occur, however, in a slight and fairly reproducible fashion (data not shown). Therefore, this column should be applicable to quantitative studies. The methods for cysteinyl peptide and \(N\)-glycopeptide enrichment can be integrated seamlessly into the quantitative strategies by using stable isotope labeling (35, 46) (e.g. \(^{18}\)O labeling) and are amenable to automation. The major challenges in proteomic studies on inflammation are that, whereas cancer biomarkers are usually binomial (i.e. they are absent in normal and present in disease and probably do not vary extensively over short time periods), the inflammatory biomarkers are likely not a matter of presence or absence but more of a change in concentration compared with normal and likely to change rapidly over time, addressing the need for high throughput temporal analysis in this field. Although the current study is qualitative, the analytical strategy proposed here is able to be directly linked to quantitative automation. Taking full advantage of the broad proteome coverage and in-depth detection dynamic range, samples from temporal studies on inflammation prepared by the current strategy, with or without stable isotope labeling, can be analyzed by LC-MS-based approaches for high throughput biomarker discovery by using the accurate mass and time tag approach (52).

Application of this “divide-and-conquer” strategy significantly improved overall dynamic range of detection (estimated to be \(>10^7\)) for the characterization of the blood plasma proteome. In a previous report a nonredundant human plasma protein list was developed by combining proteins from four separate sources (53). The 46 proteins found in common among the four datasets were all relatively high abundance classic plasma proteins. In the current study, some of the 330 proteins observed in all four of the peptide populations appear to be classic plasma proteins (e.g. complement components, coagulation factors, and protease inhibitors); however, there are also many proteins such as cathepsins D and L that are anticipated to be present at relatively low concentrations in plasma (<10 ng/ml at normal conditions (54, 55)). These results suggest the overall high dynamic range of detection applying the current strategy that allows many low abundance proteins to be consistently detected in all four peptide populations. Additionally improvements in overall proteome coverage were achieved by combining identifications from these four different, but complementary, peptide populations. Only very small overlap was observed between the cysteinyl peptide and non-cysteinyl peptide fractions, between the \(N\)-glycopeptide and non-glycopeptide fractions, and between the cysteinyl peptide and \(N\)-glycopeptide fractions (Fig. 2A). A total of 3654 non-redundant proteins (Table II) were confidently identified that included 662 \(N\)-glycoproteins (Fig. 2B) as well as numerous cytokines, cytokine receptors, CD antigens, and proteins involved in both inflammation and immune response processes. Approximately 63% of these nonredundant proteins were identified solely from one peptide population and included many low abundance proteins such as IL-2 from non-glycopeptides; PDGF B chain from glycopeptides; CCL21 from cysteinyl peptides, and calcitonin from non-cysteinyl peptides.

A major challenge in the analysis of MS data is the accurate assessment of the extent of false positive identifications. Various criteria have been developed to filter raw MS/MS identifications; however, further statistical evaluation is essential to ensure that high confidence protein identifications can be derived from such analyses. For example, an initial large scale HUPO plasma proteome collaborative study assembled a list
of 3020 proteins identified with two or more peptides using data acquired on different instruments from 18 different laboratories (42). Recently the list has been reduced to 889 proteins (containing both multipeptide and single peptide protein identifications) identified with a confidence level of at least 95% using a rigorous statistical approach taking into account the length of coding regions in genes and multiple hypothesis-testing techniques (56). Our filtering criteria developed based on the reversed database searching are much more stringent compared with the early HUPO criteria (Table III). Their stringency is also supported by the comparable results from the reversed database approach and Peptide/Protein Prophet program in terms of generating high confidence protein identifications. Furthermore reanalyzing the data from one of our early plasma profiling studies (19) using the HUPO criteria yielded 1073 proteins; the length-dependent statistical analysis yielded ~2-fold reduction in protein identifications (433 proteins with confidence >95%) (56). Similarly for the data presented in this study, the reversed database analysis also resulted in ~2-fold fewer protein identifications compared with those identifications obtained if the HUPO criteria were applied (3654 versus 7928 proteins using previous HUPO criteria, Table III), suggesting an approximate comparable level of confidence for protein identifications obtained between the reversed database criteria and the recently published length-dependent statistical analysis (56). These comparisons between independent statistical approaches reflect the overall high quality of the currently reported data obtained by using the novel combined approach; however, the single peptide protein identifications will have a higher FDR compared with the multipeptide proteins; therefore, these single peptide proteins are listed separately in Supplemental Table 2.

The relatively low abundance cytokines present in plasma mediate not only host responses to invading organisms, tumors, and trauma but also maintain our capacity of daily survival in our germ-laden environment (57). The detection of cytokines in disease states (e.g. inflammation) may provide very useful diagnostic and/or therapeutic tools; for example, IL-1 receptor antagonist has been shown to play a role in systemic host responses and improve survival in septic shock (58, 59), and some cytokine receptors are also being evaluated for the anti-inflammation effects (60). Our strategy enabled the detection of members of all major cytokine families (Table V), demonstrating the applicability of this strategy for discovering cytokine inflammation biomarkers in quantitative studies.

An area of contention surrounding biomarker discovery is whether a single protein marker or a panel of biomarkers should be used for disease diagnosis and therapeutic treatment. An increasingly common view is that the use of a single biomarker lacks the required sensitivity and specificity when applied to a heterogeneous population; however, these limitations may be overcome by utilizing panels of biomarkers (61). As in cancer, the dysfunctional or malignant cell growth may result from the changes in multiple members of the deranged protein signal transduction pathways. Therefore, an understanding of the pathways and networks that involve plasma proteins released from the cells would facilitate the development of a disease biomarker panel for clinical applications. The pathway analysis reveals that our dataset indeed provides extensive coverage for important signaling pathways (e.g. NF-κB signaling pathway) and protein networks involved in inflammatory and innate immune responses. Such coverage suggests the potential for simultaneously monitoring the temporal changes of many protein players for a specific pathway/network when the current strategy is coupled with quantitative methodologies (e.g. stable isotope 18O labeling).

Acknowledgment—We thank Jimmy Eng at the Fred Hutchinson Cancer Research Center for assistance with Peptide and Protein Prophet analyses.

* Portions of this work were supported by the NIGMS, National Institutes of Health, Large Scale Collaborative Research Grant U54 GM-62119-02 and the National Institutes of Health National Center for Research Resources Grant RR18522. Work was performed in the Environmental Molecular Science Laboratory, a United States Department of Energy (DOE) national scientific user facility located on the campus of Pacific Northwest National Laboratory (PNNL) in Richland, WA. PNNL is a multiprogram national laboratory operated by Battelle Memorial Institute for the DOE under Contract DE-AC05-76RL01830. 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.

The on-line version of this article (available at http://www.mcponline.org) contains supplemental material.

To whom correspondence should be addressed: Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, P. O. Box 999, MSIN: K8-98, Richland, WA 99352. E-mail: rds@pnl.gov.

Additional participating investigators in the Large Scale Collaborative Research Program entitled “Inflammation and the Host Response to Injury”: Henry V. Baker, Paul E. Bankey, Timothy R. Billiar, Bernard H. Brownstein, Steve E. Calvano, Celeste Campbell-Finnerty, George Casella, Irshad H. Chaudry, Mashkoor Choudhry, J. Perren Cobb, Asl De, Constance Elson, Bradley Freeman, Richard L. Gammell, Nicole S. Gibran, Brian G. Harbrecht, Douglas L. Hayden, David N. Herndon, Jureta W. Horton, William Hubbard, John Lee Hunt, Jeffrey L. Johnson, Matthew B. Klein, James A. Lederer, Tanya Logvinenko, Stephen F. Lowry, John A. Mannick, Phillip H. Mason, Grace P. McDonald-Smith, Bruce McKinley, Carol L. Miller-Graziano, Michael Mindrinos, Joseph P. Minei, Ernest E. Moore, Fredrick A. Moore, Avery B. Nathens, Grant E. O’Keefe, Laurence G. Rahme, Daniel G. Remick, David Schoenfeld, Michael B. Shapiro, Martin Schwacha, Geoffrey M. Silver, John Storey, Mehmet Toner, H. Shaw Warren, and Michael A. West.

REFERENCES

1. Anderson, N. L., and Anderson, N. G. (2002) The human plasma proteome: history, character, and diagnostic prospects. Mol. Cell. Proteomics 1, 845–867
2. Chen, R., Pan, S., Brentnall, T. A., and Aebersold, R. (2005) Proteomic profiling of pancreatic cancer for biomarker discovery. Mol. Cell. Proteomics 4, 525–533
3. Jacobs, J. M., Adkins, J. N., Qian, W. J., Liu, T., Shen, Y., Camp, D. G., II, and Smith, R. D. (2005) Utilizing human blood plasma for proteomic biomarker discovery. J. Proteome Res. 4, 1073–1085

4. Hu, Y., Malone, J. P., Fagan, A. M., Townsend, R. R., and Holtzman, D. M. (2005) Comparative proteomic analysis of intra- and interindividual variation in human cerebrospinal fluid. Mol. Cell. Proteomics 4, 2000–2009

5. Wattiez, R., and Falmagne, P. (2005) Proteomics of bronchoalveolar lavage fluid. J. Chromatogr. B: Anal. Technol. Biomed. Life Sci. 815, 169–178

6. Liao, H., Wu, J., Kuhn, E., Chin, W., Chang, B., Jones, M. D., O’Neil, S., Clauser, K. R., Karl, J., Hasler, F., Roubenoff, R., Zolg, W., and Guild, B. C. (2004) Use of mass spectrometry to identify protein biomarkers of disease severity in the synovial fluid and serum of patients with rheumatoid arthritis. Arthritis Rheum. 50, 3792–3803

7. Varnum, S. M., Covington, C. C., Woodbury, R. L., Petritis, K., Kangas, L. J., Abdulllah, M. S., Pounds, J. G., Smith, R. D., and Zangar, R. C. (2003) Proteomic characterization of nipple aspirate fluid: identification of potential biomarkers of breast cancer. Breast Cancer Res. Treat. 80, 87–97

8. Fingleton, B., Menon, R., Carter, K. J., Overstreet, P. D., Hachey, D. L., Matrisian, L. M., and McIntyre, J. O. (2004) Proteinase activity in human and murine saliva as a biomarker for proteinase inhibitor efficacy. Clin. Cancer Res. 10, 7865–7874

9. Meinke, K. H., Issaq, H. J., and Veenstra, T. D. (2004) Direct ampholyte-free reversed-phase high-performance liquid chromatographic prefractionation and linear ion trap mass spectrometry identification. J. Am. Soc. Mass Spectrom. 15, 2070–2080

10. Sheng, S., Chen, D., and Van Eyk, J. E. (2006) Multi-dimensional liquid chromatography separation of intact proteins by chromatographic fractionation and linear ion trap mass spectrometry identification. J. Proteome Res. 4, 169–178

11. Qian, W. J., Jacobs, J. M., Camp, D. G., II, Monroe, M. E., Moore, R. J., and Smith, R. D. (2005) The human plasma N-glycoprotein analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry. J. Proteome Res. 4, 700–709

12. Shen, Y., Kim, J., Strittmatter, E. F., Jacobs, J. M., Camp, D. G., II, Fang, R., Tolie, N., Moore, R. J., and Smith, R. D. (2005) Characterization of the human blood plasma proteome. Proteomics 5, 4034–4045

13. Hood, B. L., Zhou, M., Chan, K. C., Lucas, D. A., Kim, G. J., Issaq, H. J., Veenastra, T. D., and Conrads, T. P. (2005) Investigation of the mouse serum proteome. J. Proteome Res. 4, 1561–1568

14. Triunafal, R. S., Chan, K. C., Prieto, D. A., Issaq, H. J., Conrads, T. P., and Veenastra, T. D. (2003) Characterization of the low molecular weight human serum proteome. Mol. Cell. Proteomics 2, 1096–1103

15. Liu, T., Qian, W. J., Gritsenko, M. A., Camp, D. G., II, Monroe, M. E., Moore, R. J., and Smith, R. D. (2005) The human plasma N-glycoprotein analysis by immunoaffinity subtraction, hydrazide chemistry, and mass spectrometry. J. Proteome Res. 4, 1073–1085

16. Pieper, R., Gatin, C. L., Makusky, A. J., Russo, P. S., Schatz, C. R., Miller, S. S., Su, Q., McGrath, A. M., Estock, M. A., Parmar, P. P., Zhao, M., Huang, S. T., Zhou, J., Wang, F., Esquer-Blasco, R., Anderson, N. L., Taylor, J., and Steiner, S. (2003) The human serum proteome: display of nearly 3700 chromatographically separated protein spots on two-dimensional electrophoresis gels and identification of 325 distinct proteins. Proteomics 3, 1345–1358

17. Adkins, J. N., Varnum, S. M., Aubbery, K. J., Moore, R. J., Angell, N. H., Smith, R. D., Springer, D. L., and Pounds, J. G. (2002) Toward a human blood serum proteome analysis: by multidimensional separation coupled with mass spectrometry. Mol. Cell. Proteomics 1, 947–955

18. Chromy, B. A., Gonzales, A. D., Perkins, J., Choi, M. W., Corzett, M. H., Chang, B. C., Corzett, C. H., and McCutchen-Maloney, S. L. (2004) Proteomic analysis of human serum by two-dimensional differential gel electrophoresis after depletion of high-abundant proteins. J. Proteome Res. 3, 1120–1127

19. Pieper, R., Su, Q., Gatin, C. L., Huang, S. T., Anderson, N. L., and Steiner, S. (2003) Multi-component immunoaffinity subtraction chromatography: an innovative step towards a comprehensive survey of the human plasma proteome. Proteomics 3, 422–432

20. Zolotarova, N., Martosella, J., Nicol, G., Bailey, J., Boyes, E. B., and Baker, H. V. (2005) Quantitative techniques for high-abundant protein depletion. Proteomics 5, 3304–3313

21. Huang, L., Harvie, G., Feitelson, J. S., Grammatick, K., Herold, D. A., Allen, D. L., Amupumla, R., Nagler, A. R., Pisano, M. R., Zhang, W. W., and Fang, X. (2005) Immunoaffinity separation of plasma proteins by IgY microbeads: meeting the needs of proteomic sample preparation and analysis. Proteomics 5, 3314–3328

22. Uotta, I. A., Ferrari, M., and Petricoin, E. (2003) Clinical proteomics: written in blood. Nature 425, 905.
based evaluation of peptide and protein identifications from tandem mass spectrometry and SEQUEST analysis: the human proteome. J. Proteome Res. 4, 53–62
14. Nesvizhskii, A. I., Keller, A., Kolkner, E., and Aebbersold, R. (2002) A statistical model for identifying proteins by tandem mass spectrometry. Anal. Chem. 75, 4646–4658
15. Beissbarth, T., and Speed, T. P. (2004) GOstat: find statistically overrepresented Gene Ontologies within a group of genes. Bioinformatics 20, 1464–1465
16. Calvano, S. E., Xiao, W., Richards, D. R., Felciano, R. M., Baker, H. V., Cho, R. J., Chen, R. O., Brownstein, B. H., Cobb, J. P., Tschoeke, S. K., Miller-Graziano, C., Moldawer, L. L., Mindrinos, M. N., Davis, R. W., Tompkins, R. G., and Lowry, S. F. (2005) A network-based analysis of systemic inflammation in humans. Nature 437, 1032–1037
17. Omenn, G. S., States, D. J., Adamski, M., Blackwell, T. W., Menon, R., Hermjakob, H., Apweiler, R., Haab, B. B., Simpson, R. J., Eades, J. S., Kapp, E. A., Moritz, R. L., Chan, D. W., Rai, A. J., Admon, A., Aebbersold, R., Eng, J., Hancock, W. S., Hefta, S. A., Meyer, H., Paik, Y. K., Yoo, J. S., Ping, P., Pounds, J., Adkins, J., Qian, X., Wang, R., Wasinger, V., Wu, C. Y., Zhao, X., Zeng, R., Archakov, A., Tsugita, A., Beer, I., Pandey, A., Pisano, M., Andrews, P., Tammen, H., Speicher, D. W., and Hanash, S. M. (2005) 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 5, 3226–3245
18. Keller, A., Nesvizhskii, A. I., Kolkner, E., and Aebbersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal. Chem. 74, 5383–5392
19. Scarios, O., and Fisler, R. (2005) Biomarker technology round-up: from discovery to clinical applications, a broad set of tools is required to translate from the lab to the clinic. BioTechniques 38, S30–32
20. Peticon, E., Wulfkuhle, J., Espina, V., and Liotta, L. A. (2004) Clinical proteomics: revolutionizing disease detection and patient tailoring therapy. J. Proteome Res. 3, 209–217
21. Liu, T., Qian, W. J., Strittmatter, E. F., Campbell, D. G., Anderson, G. A., Thrall, B. D., and Smith, R. D. (2004) High throughput comparative proteome analysis using a quantitative cysteinyl-peptide enrichment technology. Anal. Chem. 76, 5345–5353
22. Liu, T., Qian, W. J., Chen, W. U., Jacobs, J. M., Moore, R. J., Anderson, D. J., Gritsenko, M. A., Monroe, M. E., Thrall, B. D., Camp, D. G., II, and Smith, R. D. (2005) Improved proteome coverage by using high efficiency cysteinyl peptide enrichment: the human mammary epithelial cell proteome. Proteomics 5, 1263–1273
23. Wang, H., Qian, W. J., Chin, M. H., Petyuk, V. A., Barry, R. C., Liu, T., Gritsenko, M. A., Mottaz, H. M., Moore, R. J., Thrall, B. D., Camp, D. G., II, Khan, A. H., Smith, D. J., and Smith, R. D. (2006) Characterization of the mouse brain proteome using global proteomic analysis complemented with cysteinyl-peptide enrichment. J. Proteome Res. 5, 361–369
24. Zhang, H., Yi, E. C., Li, X. J., Mallick, P., Kelly-Spratt, K. S., Masselon, C. D., Camp, D. G., Il, Smith, R. D., Kemp, C. J., and Aebbersold, R. (2005) High throughput quantitative analysis of serum proteins using glycopeptide capture and liquid chromatography mass spectrometry. Mol. Cell. Proteomics 4, 144–155
25. Roth, J. (2002) Protein N-glycosylation along the secretory pathway: relationship to organelle topography and function, protein quality control, and cell interactions. Chem. Rev. 102, 285–303
26. Rauea, E. (1983) Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes. Biochim. J. 209, 331–336
27. Smith, R. D., Anderson, G. A., Lipton, M. S., Pasa-Tolic, L., Shen, Y., Conrads, T. P., Veenstra, T. D., and Udseth, H. R. (2002) An accurate mass tag strategy for quantitative and high throughput proteome measurements. Proteomics 2, 513–523
28. Anderson, N. L., Polanski, M., Pieper, R., Gatlin, T., Tirumalai, R. S., Conrads, T. P., Veenstra, T. D., Adkins, J. N., Pounds, J. G., Fagan, R., and Lobley, A. (2004) The human plasma proteome: a nonredundant list developed by combination of four separate sources. Mol. Cell. Proteomics 3, 311–316
29. Hara, T., Nishida, H., and Nakajima, R. (1994) Highly sensitive detection of proteins separated by capillary zone electrophoresis using on-line chemiluminescence detection. Anal. Sci. 10, 823–825
30. Lang, T. H., Willinger, U., and Holzer, G. (2004) Soluble cathepsin-L: a marker of bone resorption and bone density? J. Lab. Clin. Med. 144, 163–166
31. States, D. J., Omenn, G. S., Blackwell, T. W., Fermin, D., Eng, J., Speicher, D. W., and Hanash, S. M. (2006) Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. Nat. Biotechnol. 24, 333–338
32. Thompson, A. W., and Lodish, H. F. (2003) The Cytokine Handbook, 4th Ed., Academic Press, San Diego, CA
33. Fischer, E., Van Zee, K. J., Marano, M. A., Rock, C. S., Kenney, J. S., Poutsiaka, D. D., Dinarello, C. A., Lowry, S. F., and Moldawer, L. L. (1992) Interleukin-1 receptor antagonist circulates in experimental inflammation and in human disease. Blood 79, 2196–2200
34. Fischer, E., Marano, M. A., Van Zee, K. J., Rock, C. S., Hawes, A. S., Thompson, W. A., DeForge, L., Kenney, J. S., Remick, D. G., Blood, D. C., Thompson, R. C., Lowry, S. F., and Moldawer, L. L. (1992) Interleukin-1 receptor blockade improves survival and hemodynamic performance in Escherichia coli septic shock, but fails to alter host responses to sublethal endotoxemia. J. Clin. Investig. 89, 1551–1557
35. Van Zee, K. J., Kohn, T., Fischer, E., Rock, C. S., Moldawer, L. L., and Lowry, S. F. (1992) Tumor necrosis factor soluble receptors circulate during experimental and clinical inflammation and can protect against excessive tumor necrosis factor alpha in vitro and in vivo. Proc. Natl. Acad. Sci. U. S. A. 89, 4845–4849
36. Verma, M., Maruvada, P., and Srivastava, S. (2004) Epigenetics and cancer. Crit. Rev. Clin. Lab. Sci. 41, 585–607