Accessing Low-Abundance Proteins in Serum and Plasma With a Novel, Simple Enrichment and Depletion Method

Aran Paulus, Steve Freeby, Katrina Academia, Vanitha Thulasiraman, Tim Wehr, Ning Liu, Steve Roth, and Kate Smith

Life Science Group, Bio-Rad Laboratories, Inc., Hercules, CA 94547

Introduction

Biomarker discovery is one of the driving forces of proteomics programs in the life sciences. The utility of biomarkers as the biophysical or biochemical measurements and indicators of disease status or response to therapy is established in clinical research. However, an increased number of validated markers is needed, and advances in genomics and proteomics research have opened the door to finding them. The hope is high that proteomics will be the tool to identify biomarkers, since it is believed that the responses of an organism to a disease or therapy should be established first in its protein content. The proteome, which is highly dynamic and more complex than the genome due to alternative splicing events and posttranslational modifications, offers a rich source of potential biomarkers.

In clinical proteomics, protein biomarkers are utilized either to detect or to monitor the progression of a disease or a response to a pharmaceutical treatment. Typically, protein changes for a given disease can be measured in tissue samples or body fluids; however, the invasive nature of biopsies makes tissue samples harder to obtain and, therefore, less freely available. As a result, body fluids such as serum, plasma, cerebrospinal fluid (CSF), and urine are the primary sample sources in clinical proteomics. Of those, plasma and serum are used the most because they are relatively easily and routinely obtained from patients. In addition, since blood is in contact with the whole organism, proteins leaked from all tissues should be present in blood samples. This presents both an opportunity and a challenge: the Human Plasma Proteome Project has thus far identified in plasma the expression products encoded by 3,778 distinct genes and a total of 7,518 proteins and isoforms (www.plasmaproteomedatabase.org, Muthusamy et al. 2005).

The proteins present in serum and plasma span a concentration range of 11 orders of magnitude, with the 20 most abundant proteins representing 97–99% of the total protein mass (Anderson and Anderson 2002). No single analytical method is capable of resolving all plasma or serum proteins, and no detection method can cover more than 4 or 5 orders of magnitude. Due to limitations in resolving power, and because of the high abundance of a relatively small number of proteins, most analytical schemes for serum and plasma involve the depletion of high-abundance proteins to reduce both the complexity and dynamic range of the samples.

Immunodepletion, the selective binding of specific antibodies against the proteins of highest abundance, has been used successfully to remove the most abundant 6, 12, or 20 serum and plasma proteins (Echan et al. 2005, Zolotarjova 2005, Huang et al. 2005). This approach involves binding the selected antibodies to a chromatographic support. When serum or plasma proteins are in contact with the antibody-decorated beads, the high-abundance proteins are retained and the low-abundance proteins are eluted for use in downstream analysis schemes (Figure 1, panel A). Although this approach works quite well, its disadvantages are the relatively high cost of antibodies, dilution of the sample, and a never-ending catch-up game: after the 6 most abundant proteins are removed, the sample then contains a different set of high-abundance proteins. For further depletion, additional antibody sets for removing 12, 20, or more high-abundance proteins are available. However, this approach of raising antibodies against the next set of high-abundance proteins is a never-ending race if one believes the best chance for biomarker discovery is in the low-abundance (ng/ml) region of protein concentrations.

ProteoMiner™ technology uses a combinatorial library of hexapeptides bound to a chromatographic support, an alternative approach that should overcome most, if not all, of the disadvantages of immunodepletion while still effectively depleting high-abundance proteins. Combinatorial synthesis allows the creation of a large library of products by using a “split-couple-recombine” synthesis method on solid supports with a relatively small number of building blocks. For example, if the “split-couple-recombine” steps (Pinella et al. 2003) are repeated six times with the 20 amino acids used in proteins as building blocks, hexapeptides are formed, with each bead attached to a unique peptide sequence.

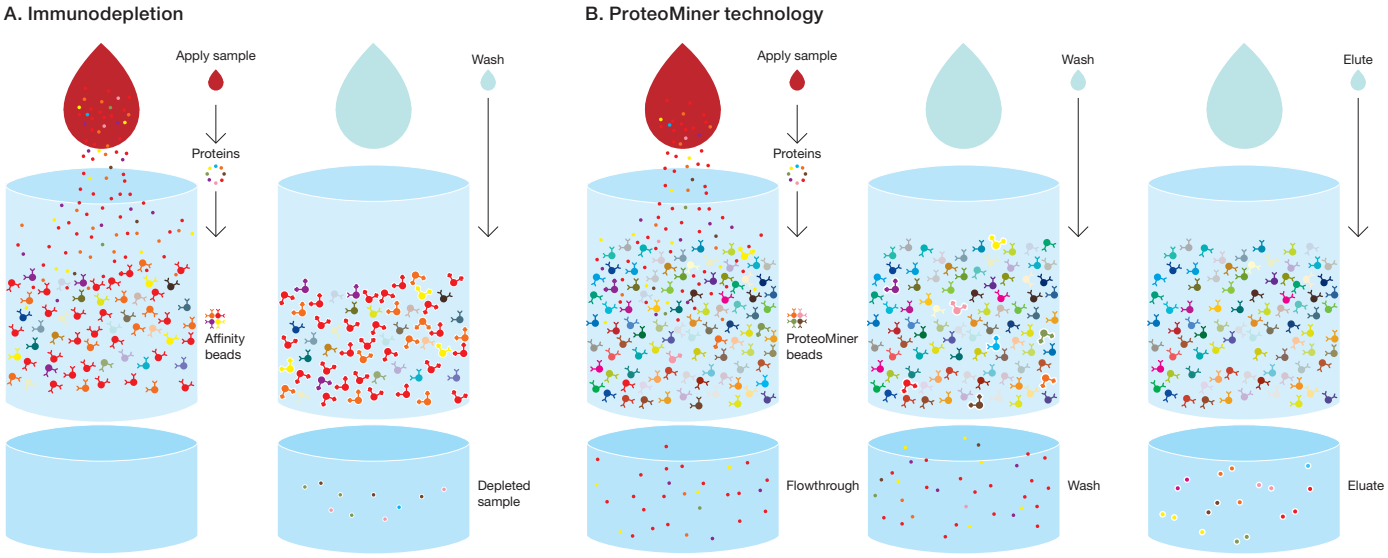

Fig. 1. Depletion of plasma or serum samples using immunodepletion (A) or the ProteoMiner protein enrichment kit (B). Note that the eluate from the ProteoMiner kits contains a greater diversity of protein species.

A library of 20^6 = 64 x 10^6 total unique beads is thus created, and these hexapeptides act as unique binders to proteins. Unlike immunodepletion, in which the capacity of the bound antibodies limits the sample amount to typically less than 100 µl, large sample volumes of 1 ml and more can be run over ProteoMiner beads. Since there are a limited number of binding sites per protein, high-abundance proteins quickly reach the bead capacity, and excess proteins are washed out. The bound proteins can be eluted using acidic CHAPS buffer (Figure 1, panel B), and the elution volume can be limited to 100 or 300 µl, effectively concentrating the low-abundance proteins. This is also in contrast to immunodepletion, where the sample volume following elution is increased and, therefore, the protein concentration is decreased.

This tech note describes the use of the ProteoMiner protein enrichment kit with either serum or plasma samples in both surface-enhanced laser desorption/ionization (SELDI) and two-dimensional gel electrophoresis (2DGE) proteomics workflows. We demonstrate that use of the kit upstream of SELDI and 2DGE analysis increases peak and spot counts, respectively, and show that proteins of low to medium abundance are enriched by this method.

Methods

All experiments were run in triplicate unless otherwise noted.

Sample Preparation Using ProteoMiner Kits

The ProteoMiner protein enrichment kit contains 100 µl bead volume of ProteoMiner beads conveniently prepackaged in easy-to-use mini spin columns as well as all the buffers required for binding, washing, and elution. After removing the storage solution by centrifugation, the beads were washed with deionized water followed by phosphate-buffered saline (PBS). Human serum or plasma samples (1 ml, Bioreclamation Inc.) were then applied to the column, and, to ensure effective binding, the samples were slowly rotated with the ProteoMiner beads for 2 hr prior to washing away unbound proteins with PBS buffer. To elute the bound low-abundance proteins, the ProteoMiner beads were treated 1–3 times with 100 µl of an acidic urea/CHAPS buffer (5% acetic acid, 8 M urea, 2% CHAPS) that is directly compatible with downstream SELDI and 2DGE experiments. Protein quantitation was performed using the *RC DC*™ protein assay.

One- and Two-Dimensional (1- and 2-D) Electrophoresis, Image Analysis, and Protein Identification

For SDS-PAGE, serum samples (1 ml) were treated with 100 µl ProteoMiner beads. From the eluted fraction, 250, 100, and 50 µg total protein samples were loaded onto Criterion™ 8–16% Tris-HCl gels, separated for 1 hr at 200 V, and stained with Flamingo™ fluorescent gel stain.

For 2DGE experiments, 100 µg protein was loaded onto an 11 cm ReadyStrip™ IPG strip, pH 5–8. Isoelectric focusing (IEF) was performed using a PROTEAN® IEF cell at 250 V for 30 min followed by 8,000 V until 45,000 V-hr were reached. After transfer of the IPG strip onto Criterion 8–16% Tris-HCl gels, the second dimension was run for 1 hr at 200 V prior to staining with Flamingo fluorescent gel stain.

Gels were imaged using the Molecular Imager® PharosFX™ system and analyzed with PDQuest™ 2-D analysis software, version 8.0. Selected spots were excised with the EXQuest™ spot cutter and digested with trypsin prior to protein identification using an Agilent 1100 nanoflow liquid chromatography (LC) system coupled to an LTQ ion-trap mass spectrometer (MS, Thermo Scientific). Protein identification with tandem mass spectrometry was done only once per excised spot.

SELDI Analysis

For this study, a ProteinChip® SELDI system, Enterprise Edition and ProteinChip CM10 (carboxymethyl weak cation exchange) were used. The arrays were equilibrated twice with 5 µl 100 mM sodium acetate, pH 4 buffer. After equilibration, the liquid was removed and 0.5 µl ProteoMiner kit-treated serum or plasma sample was mixed with 4.5 µl 100 mM sodium acetate, pH 4 buffer and spotted onto the array. After a 30 min incubation with shaking, each spot was washed three times with 5 µl binding buffer for 5 min, followed by a quick rinse with deionized water. After air drying, sinapinic acid dissolved in an acetonitrile:trifluoroacetic acid:water mixture (49.5:0.5:50) was added twice in 1 µl increments and allowed to air dry. All arrays were analyzed with an ion acceleration potential of 20 kV and a detector voltage of 2.8 kV. Data processing steps included baseline subtraction and external calibration using a mixture of known peptide and protein calibrants. Peak detection (with signal-to-noise set to >3) and peak clustering were performed automatically using ProteinChip data manager software, version 3.2.

Results

Plasma and serum samples were fractionated using the ProteoMiner protein enrichment kit. The fractions obtained were first analyzed by protein assay and SDS-PAGE to evaluate the mass distribution and mass balance of the samples during and after fractionation. Next, the samples were analyzed using the ProteinChip SELDI system and 2DGE to demonstrate that use of ProteoMiner technology increases peak and spot counts by these methods. In addition, spots were selected from the 2-D gels and analyzed by MS to demonstrate that fractionation with the ProteoMiner kit improves resolution on 2-D gels and leads to the enrichment of proteins of low to medium abundance.

SDS-PAGE Analysis and Yield Analysis

SDS-PAGE was performed to determine the mass distribution of the different fractions obtained by treatment with the ProteoMiner protein enrichment kit. Various protein loads were analyzed, and the results obtained using 75 µg plasma protein are shown in Figure 2. The most abundant band (66 kD) in the plasma, flowthrough, and wash fractions is most likely albumin and was markedly reduced in the eluted fraction (red ellipse). In addition, a series of bands appeared in the range below 50 kD in the elution fraction that did not appear in the control plasma sample. These bands represent proteins that have been concentrated in the elution fraction. To determine the effect of fractionation using ProteoMiner technology on the protein mass distribution in the different fractions, we measured the protein mass in the starting sample, flowthrough, wash, and elution fraction. From the 65.00 mg total protein present in the original 1 ml plasma sample, 58.83 mg (or 90.5%) was found in the flowthrough, 3.75 mg (5.8%) was in the 3 ml wash solution, and 2.15 mg (3.3%) was in the 300 µl of elution buffer. Nearly 96% of the total protein mass of plasma was removed (in the flowthrough and wash fractions) using the ProteoMiner protein enrichment kit.

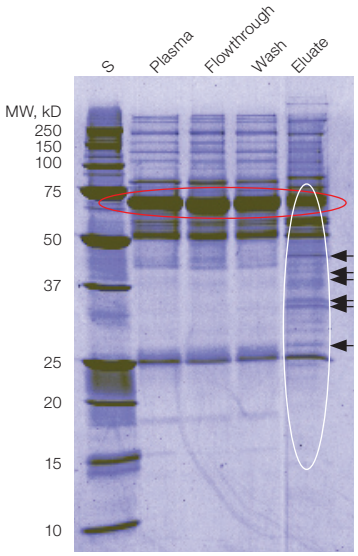

Fig. 2. SDS-PAGE analysis of ProteoMiner kit-treated plasma. A plasma sample (1 ml) was treated with the ProteoMiner protein enrichment kit, and 75 µg protein each from the original plasma sample, flowthrough, wash, and elution fractions were analyzed by SDS-PAGE. S, Precision Plus Protein™ standard. The red ellipse shows the 66 kD albumin band, and the white ellipse highlights bands that were enriched in the elution fraction. Arrowheads indicate the six bands used to analyze lot-to-lot reproducibility.

SELDI Analysis

Since the SELDI workflow is tailored to biomarker discovery, a depletion strategy combining ProteoMiner technology with the ProteinChip SELDI system generates a fast and high-throughput analysis approach that also quantitates low-abundance proteins.

Figure 3 shows a significant increase in peak detection in the mass range of 2,000–10,000 Da that occurs when a plasma sample is treated with the ProteoMiner kit. Whereas the untreated plasma sample yielded 170 peaks, the ProteoMiner kit-treated plasma yielded 263, or 55% more, peaks. In similar analyses of serum samples (data not shown), 214 peaks were found in the control sample as compared to 284 in the ProteoMiner kit-treated sample, a 32% increase.

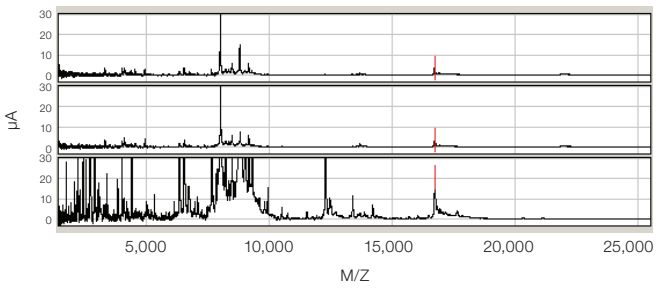

Fig. 3. SELDI mass spectrometry analysis of plasma samples treated with the ProteoMiner protein enrichment kit. One peak (red lines) was selected to show the enrichment experienced after the sample was treated with the ProteoMiner kit. The peak was very small before treatment and much larger after treatment, demonstrating enrichment of the corresponding protein. From top to bottom: untreated plasma, flowthrough, and eluate.

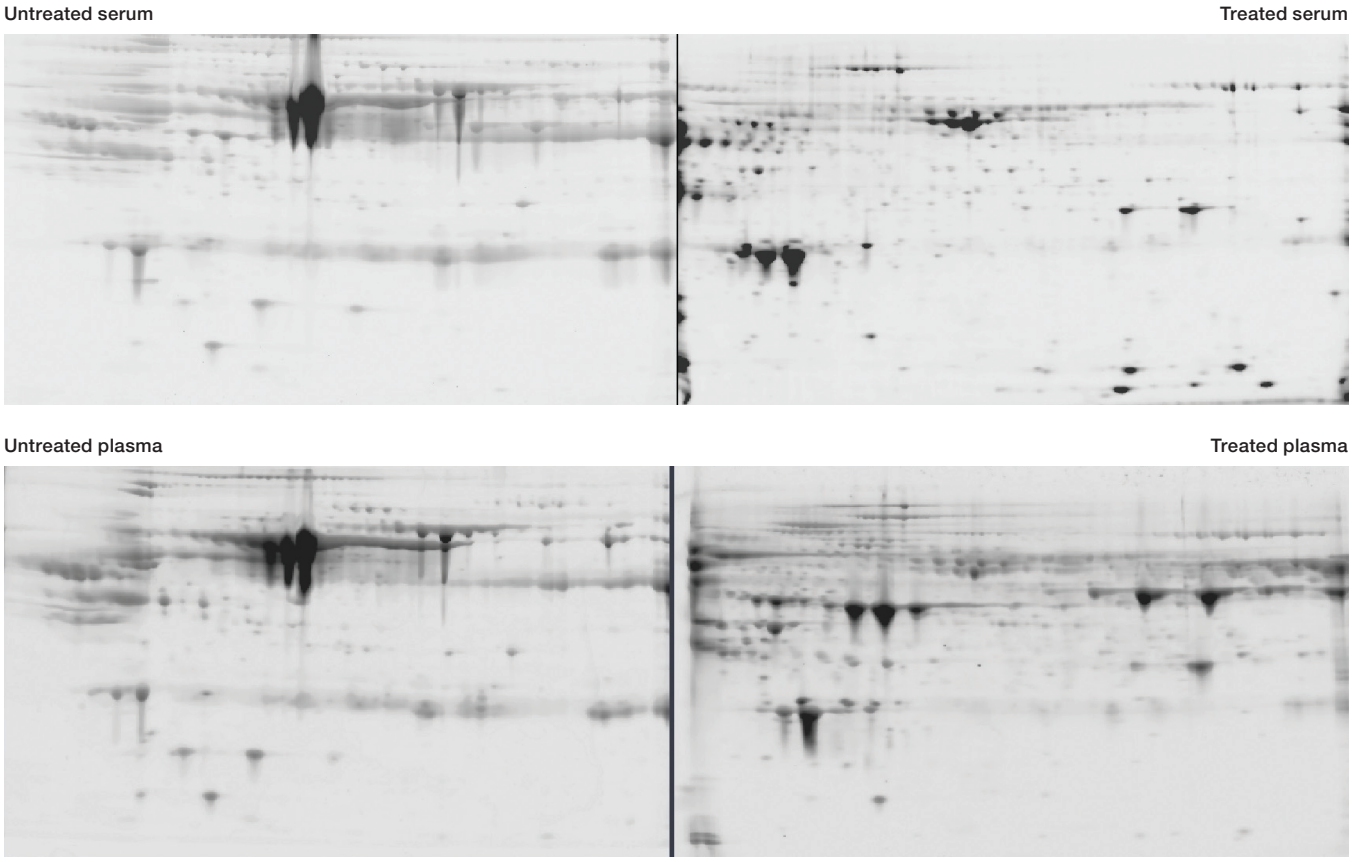

Fig. 4. 2DGE analysis of plasma and serum samples treated with the ProteoMiner protein enrichment kit. Plasma and serum samples were treated with the ProteoMiner kit, and 100 µg protein each from the original sample and elution fractions was analyzed by 2DGE using a pH 5–8 IPG strip. Note the improvement in resolution achieved with treatment of both sample types.

2DGE Analysis

The effects of ProteoMiner protein enrichment kit on 2DGE analysis of serum and plasma samples is illustrated in Figure 4. In this experiment, 1 ml serum or plasma was treated with the ProteoMiner kit, and from the resulting eluate, 100 µg protein was loaded onto an 11 cm, pH 5–8 IPG strip for IEF and subsequent 2-D analysis. Table 1 shows the peak count for each gel.

Treatment with the ProteoMiner kit increased the spot count by up to a third (Table 1), and the gels obtained after treatment showed superior resolution (Figure 4). The streaking that is typical of serum and plasma was reduced, and spots were more defined. Moreover, because the high-abundance protein spots were markedly reduced, more unique spots were visible.

Table 1. Spot count following 2DGE analysis of plasma and serum samples treated with the ProteoMiner protein enrichment kit.

	Treatment	
Sample Type	Untreated	ProteoMiner Kit
Serum	223	296
Plasma	223	338

Reproducibility

In a typical biomarker discovery study, a large number of samples are run and quantitative protein expression patterns are compared. Therefore, the reproducibility of each step in the analysis influences the confidence in any observed differences. In the case of binding a complex protein sample such as plasma or serum to an equally or more complex stationary phase (for example, ProteoMiner beads), reproducibility is a crucial element.

We used two different methods to test the lot-to-lot reproducibility of the ProteoMiner protein enrichment kit. In the first method, we used SDS-PAGE to compare the protein profiles of the elution fractions obtained with each lot. Specifically, we analyzed the presence of six bands between 20 and 50 kD (Figure 2). All four lots tested showed these bands.

In the second method, we used 2DGE to compare the protein profiles of the elution fractions. Figure 5 shows nine different 2-D gels obtained from three different lots of ProteoMiner beads, with each lot represented by three different mini spin columns. With the exception of the two most abundant spots in the middle of the right side in each image, which exhibited rather severe vertical streaking, the general spot pattern was the same in all gels, supporting the reproducibility of the method. Careful spot analysis revealed 41 matching spots in all images. Only a handful of very faint spots could not be seen in all panels.

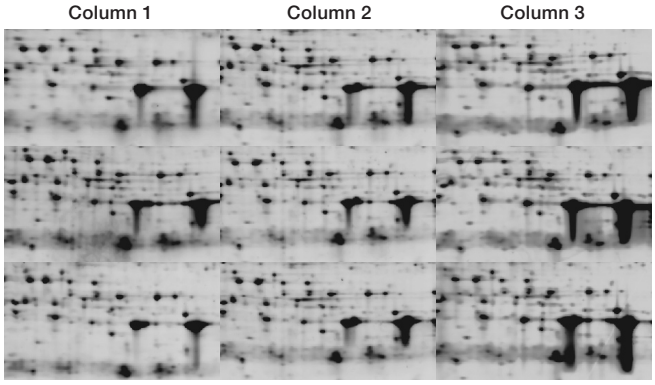

Fig. 5. Lot-to-lot reproducibility of ProteoMiner kits. Three columns from separate lots of ProteoMiner beads (each row) were used to fractionate 1 ml samples of the same plasma stock, and the elution fractions were analyzed by 2DGE. To facilitate comparison, only details from the pH 6.0–7.5 and 20–50 kD region from each gel are shown.

These results, demonstrate that the ProteoMiner kits generate highly reproducible results.

Protein Identification

The purpose of any biomarker discovery study is to first single out the protein spots or peaks demonstrating quantitative differences and then to identify the proteins composing those spots or peaks. Following 2DGE, this is usually done by in-gel digestion of the spot with trypsin followed by either peptide mass fingerprinting on a matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) MS or by nano liquid chromatography (nanoLC) separation of the peptide fragments followed by MS-MS in an ion-trap MS. We used the LC-MS-MS approach to demonstrate that the additional peaks visible after a 2-D separation also result in the identification of medium- to low-abundance proteins.

A comparison of the gel images of an untreated plasma sample and a ProteoMiner kit-treated sample revealed the following: 1) spots that appeared in both images were more intense in the gels from the ProteoMiner kit-treated sample compared with the untreated plasma (spot 1, Figure 6), and 2) the ProteoMiner kit-treated sample yielded more spots: examples of three spot that appeared in the treated sample but not in the untreated sample (spots 14–16) are shown in Figure 6.

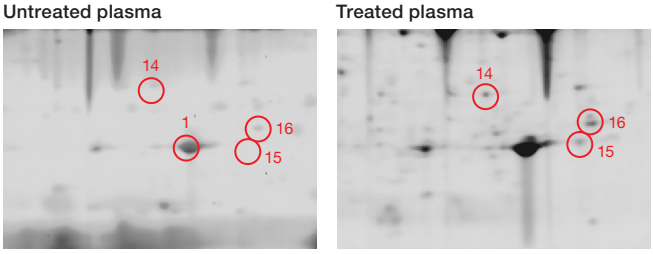

Fig. 6. Treatment of plasma sample with the ProteoMiner protein enrichment kit-treated reveals new protein spots. Red circles indicate regions excised for identification by MS.

We excised the circled areas in both gels, digested the spots, and subjected them to MS analysis. Spot 1 was identified as ZA protein in both cases; however, in untreated plasma, the protein was found with a unified score of 54, coverage of 29%, and six peptides. In the ProteoMiner kit-treated plasma, the same ZA protein was identified with a higher unified score of 76, coverage of 36%, and eight peptides. Thus, treatment with the ProteoMiner kit enabled identification with increased confidence.

The three spots labeled 14, 15, and 16 in Figure 6 were all identified as apolipoprotein E. Given the different positions of the spots in the gel, it is likely that the spots represent different isoforms, breakdown products, or posttranslational modifications of apolipoprotein E (Corder et al. 1993). Spots were excised from the same location in gels from the untreated serum sample, and as expected, MS analysis of these did not result in identification of any proteins.

Finally, we analyzed all of the protein spots unique to the ProteoMiner kit-treated plasma sample (Figure 7). All spots circled in Figure 7 were excised, digested, and subjected to MS analysis for identification, and the following proteins were identified (Table 2): plasma glutathione peroxidase precursor (spot 8), apolipoprotein E (spots 10, 14, 15, and 16), and clusterin (apolipoprotein J, spot 11). These proteins are known to exist at low to medium abundance in plasma (Anderson and Anderson 2002). In addition, a number of other proteins with medium to high abundance and poor MW or pI match were found: H factor (complement, spots 2 and 5), complement component 4B proprotein (spots 3, 4, 7, and 13), factor H homologue (spot 6), fibrinogen fragment D (spot 9), and CD5 antigen-like (spot 12). Poor matches to pI and MW are fairly common in serum and plasma analysis, since plasma and serum samples typically contain large amounts of proteolytic fragments of proteins that are only in their uncleaved form in the databases used for identification. Taken together, these results demonstrate that treatment with the ProteoMiner protein enrichment kit leads to more confident protein identification and the enrichment of low- to medium-abundance proteins in plasma samples.

Table 2. Proteins identified in plasma samples treated with the ProteoMiner protein enrichment kit. Spot numbers correlate with spots shown in Figure 7.

Spot	Probability	Peptide Coverage	Database Sequences	Database MW	pI	Protein Identity	Accession
1	5.6×10^{-05}	21.37	4	1,537	8.33	ZA protein	AAB36122.1
2	1.1×10^{-09}	17.88	3	37,638	8.33	H factor (complement)-like 1	NP_002104
3	5.2×10^{-10}	1.38	2	192,678	6.00	Complement component 4B proprotein	NP_000583.1
4	1.1×10^{-12}	8.08	11	192,678	6.00	Complement component 4B proprotein	NP_000583.1
5	7.4×10^{-08}	6.06	1	37,638	8.33	H factor (complement)-like 1	NP_002104
6	3.9×10^{-10}	6.06	1	37,627	7.17	Factor H homologue	AAA35947.1
7	4.8×10^{-11}	1.83	1	192,678	6.00	Complement component 4B proprotein	NP_000583.1
8	4.9×10^{-13}	28.32	7	25,490	8.33	Plasma glutathione peroxidase precursor	P22352
9	3.4×10^{-13}	52.98	12	36,158	6.00	Fibrinogen fragment D	1FZA
10	1.2×10^{-11}	20.50	5	36,133	4.25	Apolipoprotein E	NP_000032.1
11	2.0×10^{-06}	5.35	3	52,462	6.00	Clusterin; complement-associated protein SP-40	NP_001822.1
12	2.8×10^{-13}	46.11	10	38,064	4.25	CD5 antigen-like	NP_005885.21
13	8.7×10^{-13}	5.79	9	192,678	6.00	Complement component 4B proprotein	NP_000583.1
14	1.7×10^{-12}	26.50	7	36,133	4.25	Apolipoprotein E	NP_000032.1
15	8.1×10^{-13}	32.18	10	36,133	4.25	Apolipoprotein E	NP_000032.1
16	2.7×10^{-11}	27.13	7	36,133	4.25	Apolipoprotein E	NP_000032.1

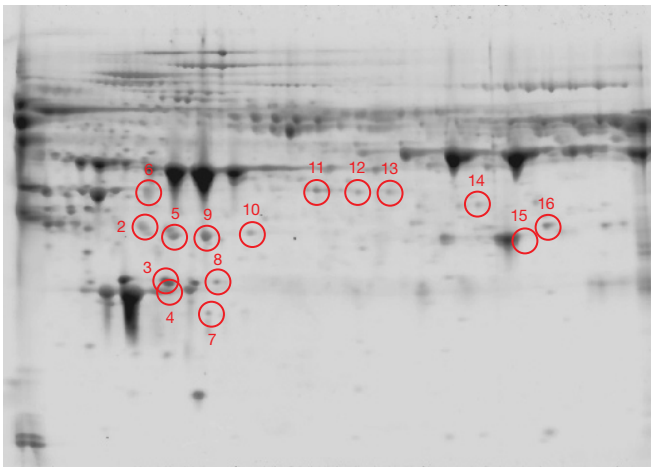

Fig. 7. Identification of protein spots unique to the Proteominer kit treated sample. Spots unique to the treated sample relative to an untreated control are circled and numbered for identification (Table 2).

Conclusions

The ProteoMiner protein enrichment kit represents an effective and novel sample preparation method that can both concentrate low-abundance proteins as well as shrink the dynamic range of protein concentrations in a sample. The procedure uses ProteoMiner beads conveniently packaged in mini spin columns and is fast, convenient, and applicable to serum and plasma samples, which exhibit an extremely wide dynamic range and contain a dozen or so high-abundance proteins. Therefore, the value of using ProteoMiner beads in biomarker discovery projects — where most significant biomarkers are assumed to be in the medium-to-low-abundance range — is clear. Downstream separation schemes can involve SELDI, 1-D, and 2DGE as well as LC methods (not shown). We demonstrate here that the use of the ProteoMiner protein enrichment kit in a SELDI- or 2DGE-based proteomics workflow can increase peak or spot counts from plasma and serum samples by 30–50%, thereby allowing detection and MS identification of medium- to low-abundance proteins.

References

Anderson NL and Anderson NG (2002). The human plasma proteome: History, character, and diagnostic prospects. Mol Cell Proteomics 1, 845-867.

Corder EH et al. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 261, 921-923.

Echan LA et al. (2005). Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. Proteomics 5, 3292-3303.

Huang L et al. (2005). Immunoaffinity separation of plasma proteins by IgY microbeads: Meeting the needs of proteomic sample preparation and analysis. Proteomics 5, 3314-3328.

Muthusamy B et al. (2005). Plasma proteome database as a resource for proteomics research. Proteomics 5, 3531-3536.

Pinilla C et al. (2003). Advances in the use of synthetic combinatorial chemistry: Mixture-based libraries. Nat Med 9, 118-122.

Zolotarjova N et al. (2005). Differences among techniques for high-abundant protein depletion. Proteomics 5, 3304-3313.

Information in this tech note was current as of the date of writing (2007) and not necessarily the date this version (Rev B, 2009) was published.

BIO-RAD

Bio-Rad Laboratories, Inc.

Life Science Group

Web site *www.bio-rad.com* ***USA*** *800 4BIORAD* ***Australia*** *61 02 9914 2800* ***Austria*** *01 877 89 01* ***Belgium*** *09 385 55 11* ***Brazil*** *55 21 3237 9400* ***Canada*** *905 364 3435* ***China*** *86 21 6426 0808* ***Czech Republic*** *420 241 430 532* ***Denmark*** *44 52 10 00* ***Finland*** *09 804 22 00* ***France*** *01 47 95 69 65* ***Germany*** *089 318 84 0* ***Greece*** *30 210 777 4396* ***Hong Kong*** *852 2789 3300* ***Hungary*** *36 1 455 8800* ***India*** *91 124 4029300* ***Israel*** *03 963 6050* ***Italy*** *39 02 216091* ***Japan*** *03 6361 7000* ***Korea*** *82 2 3473 4460* ***Mexico*** *52 555 488 7670* ***The Netherlands*** *0318 540666* ***New Zealand*** *0508 805 500* ***Norway*** *23 38 41 30* ***Poland*** *48 22 331 99 99* ***Portugal*** *351 21 472 7700* ***Russia*** *7 495 721 14 04* ***Singapore*** *65 6415 3188* ***South Africa*** *27 861 246 723* ***Spain*** *34 91 590 5200* ***Sweden*** *08 555 12700* ***Switzerland*** *061 717 95 55* ***Taiwan*** *886 2 2578 7189* ***United Kingdom*** *020 8328 2000*