Vanitha Thulasiraman[1]
Shanhua Lin[1]
Liliana Gheorghiu[2]
Julia Lathrop[2]
Lee Lomas[1]
David Hammond[2]
Egisto Boschetti[1]

[1]Ciphergen Biosystems, Fremont, CA, USA
[2]American Red Cross Biomedical R&D, Rockville, MD, USA

Reduction of the concentration difference of proteins in biological liquids using a library of combinatorial ligands

The discovery of polypeptides and proteins with relevance to a particular biological state is complicated by their vast number and concentration range in most biological mixtures. Depletion methodologies are frequently used to remove the most abundant species; however, this removal not only fails significantly to enrich trace proteins, it may also nonspecifically deplete them due to their interactions with the removed high-abundance proteins. Here we report a simple-to-use methodology that reduces the protein concentration range of a complex mixture like whole serum through the simultaneous dilution of high-abundance proteins and the concentration of low-abundance proteins. This methodology utilizes solid-phase ligand libraries of immense diversity, generated by "split, couple, recombine" combinatorial chemistry, that are used for affinity-based binding to the proteins of a given mixture. With a controlled sample-to-ligand ratio it is possible to modulate the relative concentration of proteins such that many peptides or proteins that are undetectable by classical analytical methods become easily accessible. The reduction in the dynamic range of unfractionated serum is specifically described along with treatment of other proteomes such as extracts from *Escherichia coli*, chicken egg white and cell culture supernatant. Mono- and bi-dimensional electrophoresis (1-DE and 2-DE respectively) and surface-enhanced laser desorption/ionization-mass spectrometry (SELDI-TOF-MS) technology demonstrate the reduction in protein concentration range. Combining this approach with additional fractionation methods further increased the number of detectable species.

Keywords: Dynamic range; Human serum; Ligand libraries; Low-abundance proteins; Plasma; Prefractionation; Proteomics DOI 10.1002/elps.200500147

1 Introduction

1.1 General aspects

Considerable effort has been expended toward a deeper understanding of protein composition and function in living organisms. This task is far more complex than genomic investigations because, beyond protein translation, there are a very large number of post-translational modifications (PTMs), enzymatic truncations, and generation of families of proteins, such as antibodies that are the result of the interaction of multiple gene products. Furthermore, not all proteins are present at the same time: they may be organ-, tissue-, cell- or organelle-specific, or expressed only at a given time in the cell cycle. More importantly, proteins are expressed over a vast concentration range. The human plasma proteome, for instance, has an estimated dynamic range of over ten orders of magnitude [1]. It comprises hundreds of thousands of different polypeptides; however, only few dozen of them make up 99% of the overall protein content while thousands of proteins constitute the remaining 1% [2, 3].

To date, the most commonly used methods for proteome analysis are bidimensional electrophoresis (2-DE) [4, 5] and mass spectrometry (MS) [6]. Both identify protein components of complex mixtures either by a combination of isoelectric point (p*I*) and gel mobility or by a very accurate mass analysis. Efforts to enhance the sensitivity and efficiency of both methods have been undertaken. Prefractionation or depletion of abundant proteins like albumin, transferrin, and immunoglobulins by targeted approaches such as immobilized dyes [7, 8] or immunoaffinity [9–11] facilitates the detection of a greater number of proteins, including more low-abundance species [2, 3, 12], due primarily to the removal of high-abundance species that previously obstructed trace protein detection. Immunodepletion of two abundant proteins from milk allowed detection of low-abundance proteins

Correspondence: Dr. Lee Lomas, Ciphergen Biosystems, 6611 Dumbarton Circle, Fremont, CA 94555, USA
E-mail: lomas@ciphergen.com
Fax: +1-510-505-2101

Abbreviations: CM10, carboxymethyl weak cationic exchange biochip; **IgG,** immunoglobulin G; **LDS,** lithium dodecyl sulfate; **SELDI,**Surface-enhanced laser desorption/ionization

identified by *N*-terminal sequencing [13]. Unfortunately, these methods cause the simultaneous depletion of other polypeptides (including low abundance) that are carried by depleted species in addition to being expensive, time-consuming, and not applicable to all sample types. Furthermore, the remaining polypeptides are not concentrated and may continue to fall outside the detection limit of analytical approaches.

The analysis of trace proteins has also been approached by other methods such as high-load 2-DE analysis under very narrow pH gradients [14, 15]. This method induces significant protein precipitation with consequent depletion of affected proteins. Ammonium sulfate precipitation has also been used to concentrate and promote detection of trace proteins [16], but low-molecular-weight species are lost due to poor precipitation.

Thus, the ability to detect low-abundance species remains a critical challenge in deciphering an entire proteome and correlating proteome changes with metabolic events for diagnostic purposes. The principle of a novel strategy based on the selective adsorption of proteins on a solid-phase combinatorial ligand library under capacity-limited binding conditions is described in Section 1.2, after which experimental data are reported and discussed.

1.2 Description of the principle

The spatial arrangement of amino acids within a protein defines its physicochemical properties, *e.g.*, p*I*, charge density and hydrophobicity index, and conformation. The latter determines the ability of a protein to interact *in vivo* with other molecules having complementary structures and forms the basis of protein separation by affinity chromatography where the interacting molecule (ligand) is chemically attached to a solid carrier [17]. The complementary proteins to the immobilized ligands are captured from complex mixtures up to the saturation of the available ligand. With sufficient diversity of ligands, it is theoretically possible to have a ligand to every protein in a complex mixture, ensuring that each is adsorbed. When a biological extract like serum is exposed to such a ligand library under specific capacity-restrained conditions, an abundant protein will quickly saturate all of its available high-affinity ligands and the vast majority of the same protein will remain unbound. In marked contrast, a trace protein will not saturate all its high-affinity ligands and the majority of the same protein will be bound. Thus, based on the saturation-overloading principle, a combinatorial solid-phase library will enrich for trace proteins relative to the abundant proteins. Following washing to remove unbound or weakly bound proteins, elution of the adsorbed proteins from the beads will result in a solution with a narrower dynamic range of protein concentrations while still representing all proteins present in the original material.

The library must meet three criteria: (i) a sufficient reproducible diversity of ligands must be present to bind each protein in the mixture; (ii) dissociation constants of the ligands and proteins must be compatible with the protein concentration; and (iii) the ligands and their support must be compatible with the unfractionated test sample and have a binding capacity high enough to capture sufficient protein to be detected by current methods. The technology is founded upon libraries of peptide ligands on which proteins can be adsorbed. Based on the pioneering work of Merrifield [18] on solid-phase synthesis using the "split, couple, recombine" method, libraries of peptide ligands are synthesized on resin beads [19–21]. Each bead has millions of copies of a single, unique ligand, and each bead potentially has a different ligand. Using just the 20 natural amino acids, a library of linear hexapeptides contains 20^6, or 64 million, different ligands. The addition of unnatural amino acids and D-enantiomers into branched, linear, or circular ligands generates a potential library diversity that is, for all practical purposes, unlimited, and may contain a ligand to every protein present in a biological sample. To date, using libraries of moderate diversity, ligands have been identified to a number of plasma proteins [22–24].

In the present work, a library of hexapeptides was mixed with a variety of biological samples. Once ligands bind their corresponding proteins, the beads are washed to eliminate unbound or weakly bound proteins. Adsorbed proteins are subsequently eluted by means of classical desorption methods used in chromatography. Figure 1 represents the typical process. The eluted protein mixture is then analyzed by standard methods such as 1- and 2-DE and/or MS.

2 Materials and methods

2.1 The solid-phase ligand library

The solid-phase combinatorial libraries of hexapeptides were synthesized *via* a spacer on polymethacrylate beads by Peptides International, Louisville, Kentucky by the split, couple, recombine method described earlier [19, 25]. By calculation using 20 different amino acids the number of different ligand structures was 20^6 or 64 million. Each bead with an average diameter of 65 μm carried about 50 pmoles of the same hexapeptide distributed throughout the core of the bead.

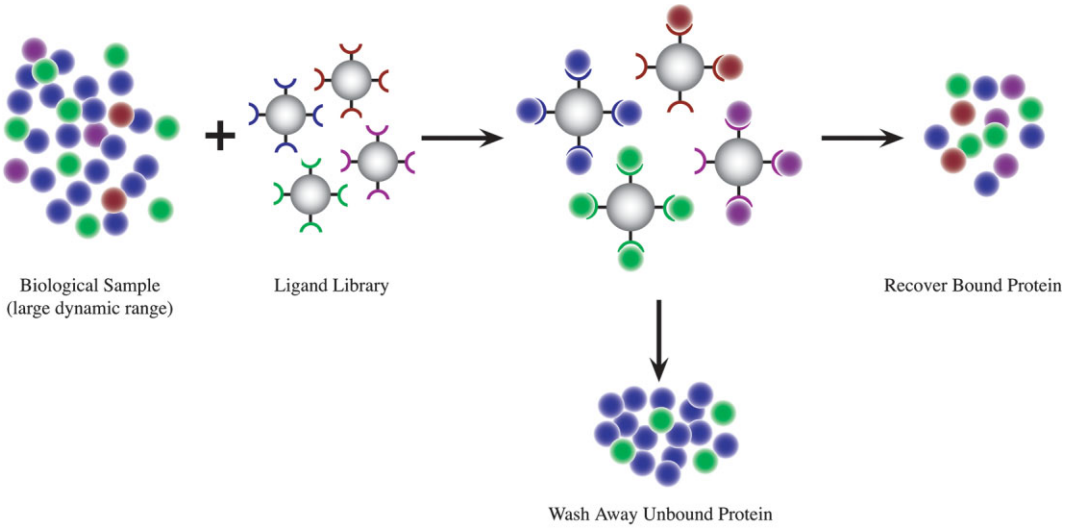

Figure 1. Schematic representation of the process for the reduction of concentration differences between proteins in a sample. First, the sample is contacted with solid-phase combinatorial ligand library at a given ratio. Second, nonadsorbed proteins are removed by washing; third, adsorbed proteins are recovered using standard affinity chromatography desorption methods, with subsequent detection by electrophoresis and surface-enhanced laser desorption/ionization (SELDI).

2.2 Sample processing with combinatorial ligand library

The crude biological sample was filtered through a 0.8 μm filter, and then mixed with a given amount of solid-phase ligand library and incubated for about 2 h. The beads were collected by centrifugation and washed with 20 bed volumes of a phosphate-buffered saline (PBS). The ratio of the volume of solid-phase ligand library to the volume of biological sample ranged from 1 to 150 according to the specific trials. Proteins captured by the solid-phase ligand library were either desorbed altogether using 6 M guanidine-HCl, pH 6, or sequentially using 1 M sodium chloride, 60% ethylene glycol in water, 400 mM glycine-HCl, pH 2.5, and 6 M guanidine-HCl. Recovered proteins were submitted to comparative analysis against native sample. Processed samples were human plasma and serum, whole extract from *Escherichia coli*, chicken egg white, and cell culture supernatant.

2.3 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

Electrophoresis of protein fractions was performed using 10-well precast 1-mm-thick 12% Bis-Tris polyacrylamide gel plates (Invitrogen, Carlsbad, CA, USA). Samples of appropriate protein concentration were diluted fourfold in lithium dodecyl sulfate (LDS) sample buffer. Thirty microliters of diluted solution was loaded *per* lane and electrophoresis migration was performed at 150 V for 90 min. Coomassie Brilliant Blue (CBB) staining and destaining were achieved using the method described by the supplier of reagents (Invitrogen).

2.4 2-DE

Desalted proteins were solubilized in the "2-DE sample buffer," containing 7 M urea, 2 M thiourea, 3% CHAPS, 0.5% ampholyte, pH 3–10, and bromophenol blue in traces. Volumes of 150 μL of the obtained protein solution (control sample), or of the fractions obtained *via* solid-phase ligand library, were then used for rehydrating 7-cm-long IPG 3–10 strips for 4 h. Isoelectric focusing (IEF) was carried out with a low initial voltage and then by applying a voltage gradient up to 5000 V, with a limiting current of 50 μA *per* strip. The total product time × voltage applied was 25 000 Vh for each strip, and the temperature was set at 20°C. For the second dimension, the IPG strips were equilibrated for 26 min in a solution of 6 M urea, 2% SDS, 20% glycerol, 375 mM Tris-HCl (pH 8.8), under gentle agitation. The IPG strips were then cemented with 0.5% agarose on a 12% SDS-PAGE in the cathode buffer (192 mM glycine, 0.1% SDS and Tris to pH 8.3). The anodic buffer was a solution of 375 mM Tris-HCl, pH 8.8. The electrophoretic run was performed by setting a current of 5 mA for each gel slab for 1 h, and then 10 mA *per* gel for 1 h and finally 20 mA *per* gel until the end of the run. Dur-

ing the entire run the temperature was set at 11°C. At the end of the run the gels were stained overnight with colloidal CBB G-250. Destaining was performed in 5% acetic acid for 2 h.

2.5 Fractionation by anion exchange chromatography: anion exchange fractionation of human serum using 96-well format

To each well in a 96-well filtration plate was added 100 μL of Q Ceramic Hyper D resin as a 50% suspension. The resin was equilibrated in 50 mM Tris-HCl pH 9.0 with 1 M urea and 0.22% CHAPS. To 25 μL of human serum was added 37.5 μL of 50 mM Tris-HCl pH 9.0 containing 9 M urea and 2% CHAPS, and incubated for 30 min at 4°C, followed by addition of another 37.5 μL of 50 mM Tris pH 9.0. Serum mix was added to the anion exchange resin in filter plate and mixed for 30 min on a micromix 5 (set at 20,7,60). Flow-through was collected in a 96-well microplate collection plate by applying a vacuum.

Elution of protein fractions was obtained by lowering the pH in a stepwise mode. The sequence of buffers used was 50 mM HEPES-acetate pH 7, 5, and 4 followed by a citrate buffer pH 3 and finally by a hydroorganic solution containing trifluoroacetic acid (TFA), isopropanol, and acetonitrile (ACN). Collected fractions were of 150 μL each.

2.6 SELDI-MS analysis

MS-compatible SPE arrays used throughout this study were CM10 (carboxymethyl weak cation exchange biochip), Q10 (strong anion exchanger), H50 (hydrophobic surface), and IMAC30 (metal ion chelating surface). Each spot of a ProteinChip® Array (from Ciphergen Biosystems, Fremont, CA, USA, named as biochip throughout the text) was equilibrated twice with 5 μL of the indicated array-specific binding buffer for 5 min. Each spot surface was then loaded with 6 μL of the sample previously sixfold diluted in the array binding buffer. After an incubation period of 30 min with shaking, each spot was washed three times with 5 μL of the binding buffer for 5 min to eliminate nonadsorbed proteins, followed by a quick rinse with deionized water.

All surfaces were dried and loaded twice with 1 μL of a saturated solution of Sinapinic acid in a mixture of ACN (49.5 vol.):TFA (0.5 vol.):deionized water (50 vol.), and dried again. All arrays were then analyzed by MS using a laser desorption, linear time-of-flight (TOF) mass reader in positive-ion mode (SELDI-TOF-MS), with an ion acceleration potential of 20 kV and a detector voltage of 2.8 kV. The molecular-weight range investigated by mass spectrometer was from 0 to 300 kDa. Time-lag focusing was optimized at 70 and 5 kDa for high and low mass range, respectively.

Processing of data obtained included baseline subtraction and external calibration using a mixture of known peptide and protein calibrants. Peak detection ($S/N > 3$) and peak clustering were performed automatically using Ciphergen ProteinChip Software 3.2. Unique peaks across different chromatographic fractions were only counted once even if they appeared in more than one fraction.

2.7 Preparation of *E. coli* whole lysate

E. coli Ultra BL21 (DE3) competent cell was purchased from Edge Biosystems (Gaithersburg, MD, USA). Stationary-phase cells were grown in Luria broth (LB) media (2% w/v tryptone, 1% w/v yeast extract, and 1% w/v NaCl) at 37°C under aerobic conditions for 16 h. Cells were harvested by centrifugation at 4000 rpm for 10 min. The cell pellet was resuspended in ice-cold lysis buffer (20 mM Tris-HCl, pH 8.2, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 1% Triton X-100, and on protease tablet from Boehringer Mannheim (Germany), and sonicated using five 30 s bursts with 60 s intermittent cooling periods on ice. The cell lysate was later centrifuged at 8000 rpm for 10 min. Protein concentration of the final supernatant was determined to be around 20 mg/mL using the Pierce BCA protein assay kit (Rockford, IL, USA) with BSA as a quantitative standard. *E. coli* whole lysate was aliquoted and quickly frozen in dry ice and stored at −70°C.

2.8 Preparation of cell culture medium

White cells were obtained from the buffy coat of freshly collected and centrifuged human blood. The cells were purified using isopycnic centrifugation and incubated for 12 h in serum-free media supplemented with 0.5% albumin. The supernatant was collected by centrifugation and stored frozen at −20°C prior to evaluation.

2.9 Identification of proteins: in-gel tryptic digestion

Excised gel bands from SDS-PAGE were treated to remove CBB stain and SDS by incubating successively with methanol/acetic acid, ammonium bicarbonate (pH 8), and ACN solutions. The gel pieces were dried in a Speed-Vac, then rehydrated with 10 μL of 25 mM ammonium bicarbonate (pH 8.0) containing 0.02 μg/μL modified trypsin (Roche Applied Science, Indianapolis, IN, USA).

The digests were incubated for 16 h at 37°C. One microliter of supernatant containing the peptides from the protease digestion was applied to an NP 20 ProteinChip Array. A 1 μL aliquot of saturated α-cyano-4-hydroxicinnamic acid (CHCA) in 50% ACN, 0.5% TFA was applied to each spot.

2.10 MS/MS analysis and protein identification

Positive-ion spectra were acquired on a QqTOF mass spectrometer (Micromass QTOF II, Manchester, UK) equipped with a Ciphergen ProteinChip Array Interface. Ions were created using a pulsed nitrogen laser, Laser Science, VSL 337 NDS (Franklin, MA, USA) operated at 30 pulses *per* second delivering an average pulse fluence of 130 μJ/mm^2. Nitrogen gas, at 10 mtorr of pressure, was used for collisional cooling of formed ions and argon was used as the collision gas for all low energy collision-induced dissociation (CID) experiments. The CID spectra were submitted for the MASCOT MS/MS ion search. All chemicals and biochemicals were of analytical grade from Sigma-Aldrich (St. Louis, MO, USA).

3 Results

3.1 Demonstration of dynamic range reduction in serum

The first demonstration of the dynamic range reduction of plasma proteins was performed using 1 mL of combinatorial ligand library. Based on average particle diameter of 65 μm, 1 mL of combinatorial library contains about 3–5 million different beads, with a potential total protein binding capacity of about 20 mg. Previous experience [22–24, 26] indicated this volume of resin to be sufficient for the identification of specific ligands of high affinity to plasma proteins.

Figure 2 represents 1-DE of the dynamic range reduction of 10 mL of pooled, unfractionated human plasma diluted 1:1 with citrate buffer, both with IgG present or depleted with Protein G, incubated with 1 mL of ligand library. Approximately equal amounts of protein (1 mg/mL) are loaded in each lane. A similar pattern of bands are detected in the starting material, unbound material after incubation, and wash (lanes 2 through 4 and 7 through 9), indicating that the relative concentration of the most abundant proteins does not change appreciably after saturation of their available ligands (unbound and wash). Several previously undetected bands are observed following elution with LDS-PAGE (lanes 5 and 6), indicating the enrichment of these proteins. The presence of highly abundant immunoglobulinsG (IgGs) does not interfere with the relative

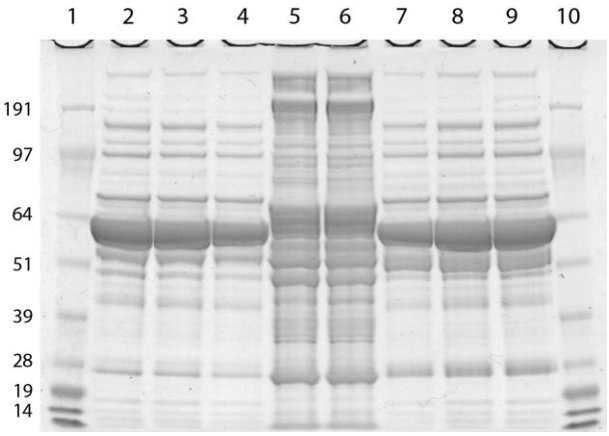

Figure 2. SDS-PAGE of initial plasma and the same after ligand library treatment. Plasma used was native (lanes 2, 3, and 4) or after depletion of IgG (lanes 7, 8, and 9). Lanes 1 and 10 represent molecular-weight standards; lanes 2, 3, and 4 represent neat plasma, flow-through, and wash of ligand library, respectively; lanes 5 and 6 represent the eluant of the ligand library from the native and IgG-depleted plasma, respectively; lanes 7, 8, and 9 are plasma, flow-through, and wash from the library of IgG-depleted plasma.

enrichment (compare lanes 5 and 6). Thus, the presence of the most abundant proteins including albumin and immunoglobulins does not interfere with detection of intermediate concentration proteins due to the highly specific affinity interactions among all of the proteins and their ligands.

3.2 Influence of bead/sample ratio on the "dynamic range reduction" process

To analyze the influence of both the ratio of beads to protein and the number of different ligands in the library on the reduction of dynamic range, two types of experiments have been performed with human serum. In the first experiment, the number of beads was kept constant at 3×106 in 1 mL while the volume of serum was varied from 1, 10, 50, and 150 mL. In a second experiment the beads/sample volume ratio was kept constant (1:10) while the volume of beads ranged from 0.18 mL (5×10^5 beads) to 16 mL ($60–70 \times 10^6$ beads). In both experiments the adsorbed proteins were eluted by single elution with 6 M guanidine-HCl, pH 6.0.

Following the general protocol described in Section 2, desorbed proteins were analyzed by 1- and 2-DE, and MS in comparison to nontreated serum.

Figure 3 shows analytical data from SELDI-TOF-MS and from 2-DE analysis. This study demonstrated a very clear increase in the number of peaks identified by SELDI-TOF-

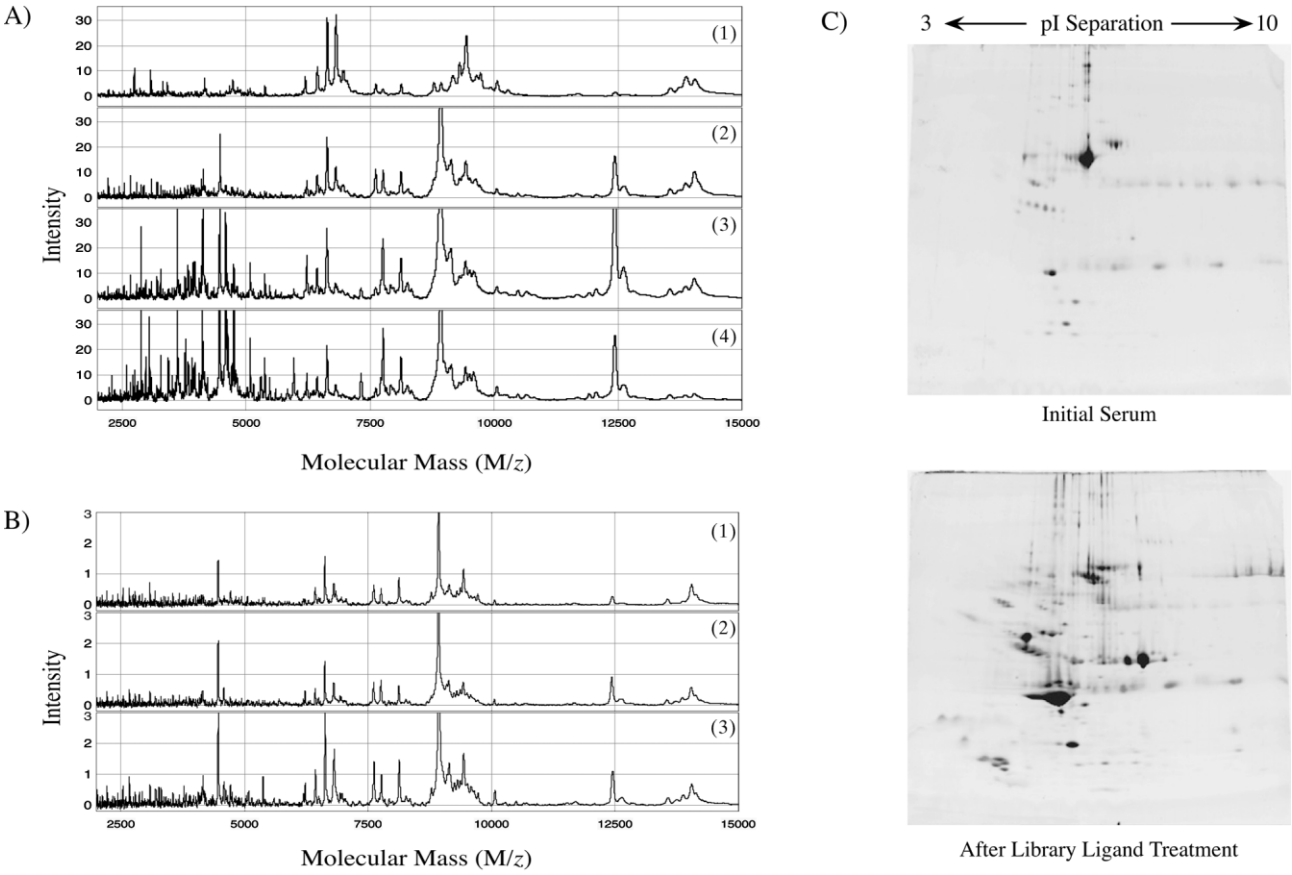

Figure 3. Analysis of serum proteins submitted treated with the ligand library. (A) represents the influence of sample volume on a fixed volume of beads (1 mL) as revealed by SELDI-TOF-MS. A1–4 represent data from bead-to-serum volume ratio of respectively 1:1, 1:10, 1:50, and 1:150. ProteinChip array used was IMAC30. (B) represents the influence of the number of beads on a fixed bead-to-serum volume ratio (1:10). B1–3 represent data from bead volumes of, respectively, 0.18, 1, and 16 mL. Analysis conditions were equivalent to those indicated above. (C) represents a typical 2-DE analysis of initial serum and treated serum using a bead-to-serum volume ratio of 1:10 (for details see Section 3.4).

MS, especially in the medium–low molecular masses (2–15 kDa). Detection of new species was enhanced by increasing volumes of serum:bead ratios (Fig. 3A), while the concentration of the most abundant species was decreased in all ratios. MS peak count variation as a function of resin:sample ratio is shown in Table 1. The number of proteins detected increased when the ratio increased from 1/10 to 1/150 when detected on either the metal chelate interaction or cation exchange surfaces. When the resin:serum ratio was increased to 1:150, the overall number of species detected on CM10 biochip (170 species) and IMAC30 biochip (341 species) increased 126 and 315%, respectively, over pretreatment samples.

In a second set of experiments, the resin:sample ratio was constant while absolute volumes were increased (Fig. 3B). Decreasing or increasing the volume of beads used while maintaining a constant bead:sample ratio did not have a significant impact on the number of protein species detected. For example, using 0.18 mL (about 5×10^5 ligands), 1 mL (3×10^6 ligands), or 16 mL (64×10^6

Table 1. Effect of bead to sample volume ratio on the number of detected species on a cation exchange (CM10) and chelating (IMAC30-Cu^{++}) array surfaces by SELDI-TOF-MS

Bead: sample ratio	Number of detected species	
	CM10 ProteinChip array	IMAC30 ProteinChip array
Neat serum	134	108
1:1	113	136
1:10	130	214
1:50	152	241
1:150	170	341

ligands) of beads while maintaining the bead:sample ratio at 1:10 resulted in combined CM and IMAC array detected species that differed by only 11%.

Figure 3C represents a 2-DE analysis of a nontreated serum (top) and the same serum after treatment with the ligand library at a ratio of 1/10. A large increase of detectable proteins was confirmed by gel spot counts: 487 *versus* 195 spots obtained with whole serum. A number of additional lower abundance proteins were identified, *e.g.*, vitronectin, C3a anaphylatoxin, C4a anaphylatoxin, serum amyloid A (SAA), apolipoprotein J (Apo-J), C4bP, ceruloplasmin, arylesterase, and interalpha-trypsin inhibitor-1 and -2 (ITIH1 and ITIH2).

Finally, to determine the reproducibility of the whole process five independent experiments were completed with 1 mL of beads and 10 mL of serum, each elution analyzed in triplicate, and a pooled CV of 16.8% for the entire method was obtained.

3.3 Interaction mechanism by sequential desorption

The adsorption of proteins on a peptide library is the result of several different interactions based on ion exchange, hydrophobic associations, hydrogen bonding, and other weak forces. To clarify this aspect of the technology and demonstrate fractionation by interaction type, adsorbed proteins were sequentially eluted with different agents. The first eluant (1 M NaCl) increased the ionic strength without pH modification; the second, 60% ethylene glycol, used hydrophobic displacement; the third, 0.4 M glycine pH 2.4, decreased the pH; and finally, a neutral chaotropic agent, 6 M GuHCl, associated with a high ionic strength, stripped off the remaining proteins. It should be noted that proteins present as multiprotein complexes were also dissociated from each other by the various elution conditions.

Figure 4 shows SDS-PAGE and SELDI-TOF-MS analysis of eluted fractions. Elution with high-salt, low-pH, and chaotropic buffers increased the number of detectable

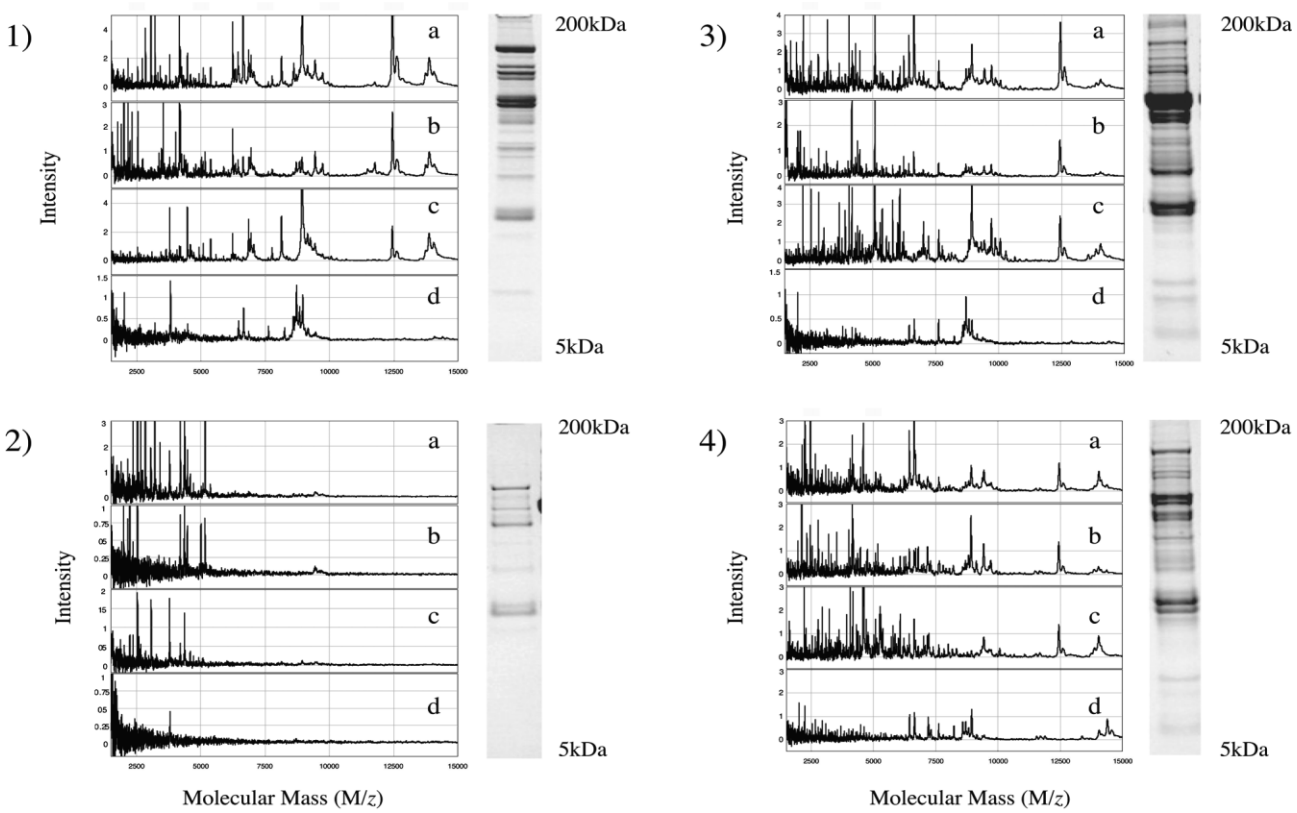

Figure 4. Electrophoresis and SELDI-TOF-MS results of fractions desorbed sequentially from the ligand library. Desorption agents were 1 M sodium chloride pH 7 (column 1), 60% ethylene glycol (column 2), 0.4 M glycine-HCl, pH 2.5, and 6 M guanidine-HCl, pH 6.0. Spectra "a" to "d" represent SELDI-TOF-MS analysis data from different chip surface, respectively, CM10 (cation exchange), Q10 (anion exchange), IMAC30 (copper chelator), and H50 (hydrophobic). NaCl extract was loaded on the gel 50% less and ethylene glycol extract was loaded twofold more when compared to glycine and guanidine extract.

proteins, particularly on IMAC30 and CM10 biochips. Hydrophobic elution using ethylene glycol did not result in a large number of species signals when compared to GuHCl elution, suggesting that strong interactions are present between proteins and ligand library involving hydrophobic associations; however, 60% ethylene glycol was not sufficient for the dissociation of these interactions (see Fig. 4 column 2 and especially box 2d). Overall, the experiment demonstrated that the protein–ligand interactions in the library involve various physicochemical forces as would be expected from the large structural diversity generated by the combinatorial assembly of amino acids.

3.4 Application of the method to other biological samples

The original method was extended to *E. coli* whole extract, chicken egg white, and cell culture supernatants. Crude samples were used as described in Section 2, and analysis performed by SDS-PAGE and SELDI-TOF-MS.

Figure 5A shows data from *E. coli* extract. The number of proteins detectable after treatment with the ligand library is much larger than from nontreated sample by both SDS-PAGE (> 20 kDa) and MS (< 20 kDa). Bands were cut out from the electrophoresis gel (Fig. 5B, lane 2) to identify proteins by in-gel digestion followed by MS/MS analysis. The protein identity for several of them is listed in Table 2. All these proteins are qualified as low abundance. Based on previous published work [27], ADP-L-glycero-B-mano-heptose-6-epimerase is present in around 220 copies *per* cell; the other five enzymes were not previously detected by 2-DE analysis out of the whole lysate because of their low concentrations; moreover, putative tagatose 6-phosphate kinase gatZ was previously reported only by DNA sequence.

Similar results were obtained with the chicken egg white (Fig. 5B). Proteins of high abundance, such as ovalbumin, conalbumin (ovotransferrin), and lysozyme, were significantly decreased while many others of high, medium, and low molecular mass appeared as shown by 1-DE and SELDI-TOF-MS. Two main groups of cationic proteins (masses between 9 and 13 kDa) appeared following fractionation on CM10 biochips.

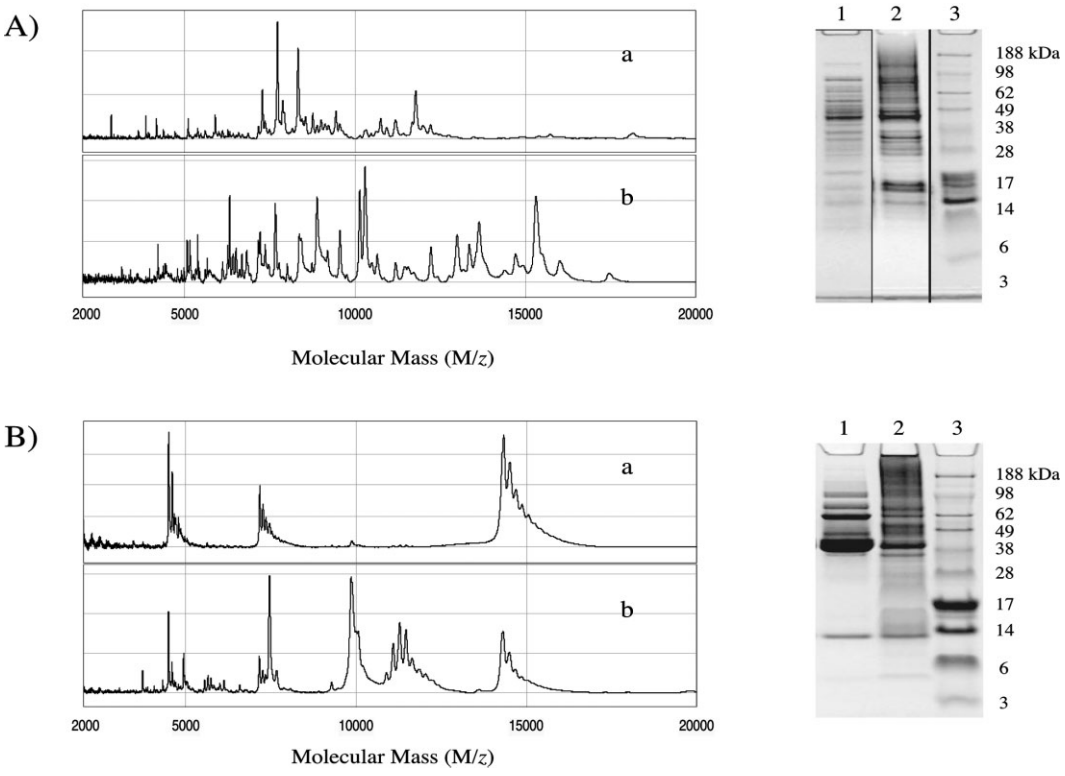

Figure 5. SELDI-TOF-MS and SDS-PAGE of *E. coli* extract (A) and chicken egg white (B). "a" represents MS spectra of both biological liquids before treatment; "b" represents the same analysis after treatment with the ligand library. All analysis was performed using CM10 ProteinChip arrays. Experiments were made using a bead-to-sample volume of 1:10. Lane 1, neat lysate of *E. coli* or chicken egg white; lane 2, extracts from combinatorial bead treatment; lane 3, *Mr* standards.

Table 2. Identification of proteins present in the *E. coli* lysate visible only after enrichment using the bead ligand library

Protein identity	p*I*	M_r
NADH-quinone oxidoreductase chain C/D	6.41	68 680
Putative tagatose 6-phosphate kinase gatZ	5.5	47 108
Glutamate-1-semialdehyde 2,1-aminomutase	4.74	45 356
Glycine acetyltransferase	5.88	43 117
Galactitol-1-phosphate-5-dehydrogenase	5.94	37 366
ADP-L-glycero-B-mano-heptose-6-epimerase	4.8	34 871

In an additional example, human white cells obtained from the buffy coat were stimulated and cultured in medium supplemented with human albumin and transferrin. After culture, the supernatant was recovered and shown to contain supplemental proteins, including a number of trace components at ng/mL concentrations. When this supernatant was mixed with a combinatorial library, many trace proteins were enriched while the concentration of supplemental proteins decreased as shown in Fig. 6. When 1 mL of cell culture supernatant containing 1.5 mg/mL of protein was loaded on 1 mL of beads, all material was adsorbed and no difference in relative protein concentration was observed between the starting material and the eluted fraction (lanes 2 and 6); however, as the volume of the supernatant and resin:supernatant ratio increased and exceeded the binding capacity of the beads, the concentration of supplemental proteins (high abundance) was not increased, while trace proteins continued to increase in their proportion up to the limits of the experiment, thereby decreasing the relative concentration range.

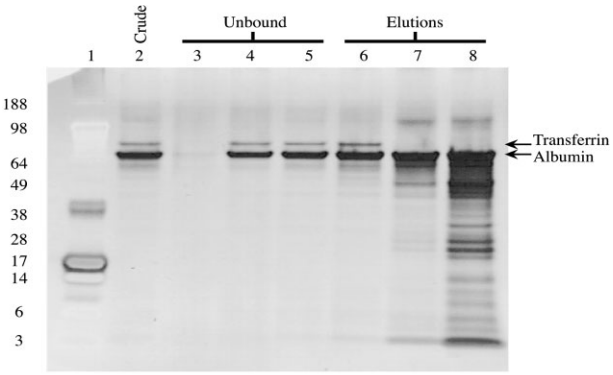

Figure 6. Cell culture supernatant analyzed by SDS-PAGE, before and after treatment with the ligand library. Lane 1: M_r standards; lane 2: crude supernatant; lanes 3 and 6 represent, respectively, the supernatant and the extract from combinatorial bead treatment using a bead-to-sample volume ratio of 1:1; lanes 4 and 7 represent the supernatant and the extract, respectively, from combinatorial bead treatment using a bead-to-sample volume ratio of 1:10; lanes 5 and 8 represent, the supernatant and the extract, respectively, from combinatorial bead treatment using a bead-to-sample volume ratio of 1:100.

3.5 Serial processing of serum samples–dynamic range reduction followed by anionic exchange fractionation

To further demonstrate the capability of the described method in increasing the detection of the number of proteins, specific comparative experiments were performed on serum proteins fractionated by ion exchange with or without pretreatment using the combinatorial ligand library. The fractionation method utilized a classical anion exchange chromatography under a stepwise pH gradient as described earlier [28, 29] and collected fractions were analyzed by SELDI-TOF-MS.

Table 3 shows relevant peak count data comparing fractions from untreated and ligand library treated serum. Nonfractionated, treated serum showed 31% additional peaks compared to nonfractionated, nontreated sample. This difference is also present when comparing the fractionated samples before or after treatment with the ligand library, where 72% additional peaks were detected. It should be stressed that even the treated, nonfractionated sample showed a larger number of peaks compared to the fractionated serum (+7%). Most relevant data were obtained using CM10 and IMAC30 biochips.

4 Discussion

A major goal of proteomics is the detection of species of low abundance. Difficulties in achieving this are related to (i) low sensitivity of detection methods, (ii) the very large number of proteins that constitute the mixture, and (iii) suppression, or masking, of signal in the presence of proteins of extremely high concentration. Current detection methods have a dynamic range of about four orders of magnitude, which is insufficient to cover the overall dynamic range of protein abundance in blood estimated to be greater than ten orders of magnitude. This constitutes a critical impediment to the detection of protein species that may represent diagnostic/prognostic indicators or therapeutic targets [30]. An approach to improve detection sensitivity of low-concentration proteins is to reduce the concentration of the most abundant proteins, which interfere with trace protein detection [11]; however, this approach suffers from the concomitant removal of associated low-concentration species [10] and from the

Table 3. Peak count before and after anion exchange fractionation of native serum and serum treated with ligand library

Analyzed protein mixture	Peak count with different types of biochips				
	CM10	IMAC30	Q10	H50	Total
Nonfractionated whole serum	134	108	83	45	370
Anionic exchange resin fractionated whole serum	176	145	76	56	453
Nonfractionated treated serum	130	214	106	36	486
Anionic exchange resin fractionated treated serum	259	231	140	114	780

Analysis made using SELDI-TOF-MS using different types of surfaces

fact that even an extremely selective extraction does not concentrate proteins present in trace amounts. Indeed, the depletion further dilutes trace proteins including important serum protease inhibitors, which may compromise the stability of the serum sample. Other affinity-based enrichment methods [31] based on heparin or hydroxyapatite, enrich only for those classes of proteins with the selected general affinity. Nevertheless, the process of abundant protein depletion has some interest because it represents a way to analyze species associated with the major proteins [10].

The method described in this report simultaneously increases the concentration of low-abundance proteins and decreases the concentration of high-abundance proteins from the same mixture in a single operation while maintaining protein diversity. This method is based on solid phase, selective adsorption of proteins for affinity ligands under controlled saturation conditions.

SDS-PAGE analysis demonstrated the enrichment of high-molecular-weight proteins, while MS demonstrated the enrichment of low-molecular-weight proteins. These low-molecular-weight species may be of diagnostic relevance when they result from proteolytic breakdown of proteins due to perturbed biological pathways. This enrichment was evidenced in a variety of very different biological samples, including low-concentration cell culture supernatant and egg white. Thus, this method largely overcomes the problem of the dynamic range to proteomic analyses.

Despite the large diversity of ligands in this library, there are some proteins that are undetectable following incubation of serum with the library, including a few proteins that were detectable before the equalization, but no longer visible after the operation. This suggests that the concentration of the analyte may not be wholly responsible for the low level of interaction, but rather the absence of an appropriate ligand in this library. Based on the results above, this phenomenon is estimated to affect about 2% of total proteins present in serum. The library used in these experiments, the backbone resin on which the ligands are synthesized as well as the design of the ligands themselves, was one of several specifically designed with complementary properties to bind plasma proteins [22–24, 26]. It appears highly efficient in the main, but it is likely that supplementation with the addition of alternate libraries will further increase the overall diversity and protein absorption. For example, libraries containing branched or circular ligands or ligands comprised additionally of unnatural amino acids, carbohydrates, fatty acids, and/or nucleic acid moieties may provide additional affinity motifs that are not represented in the current affinity library but necessary to bind a greater number of proteins within a biological sample. Interestingly, apolipoprotein A represents a relatively abundant serum protein, and under saturating ligand conditions, one would predict a decrease in apolipoprotein A concentration; however, 2-DE clearly shows an increase in the spot colorimetric density corresponding to apolipoprotein A (Fig. 3B). This may be due, in part, to apolipoprotein A1 (ApoA1) being a component of high-density lipoproteins (HDLs) and low-density lipoproteins (LDLs) in serum; ligands to individual proteins in a complex will bind associated members of the complex, thereby concentrating all proteins in the complex beyond their individual concentration on their individual ligands. An additional consideration is that more than one protein may interact with the same peptide ligand and thus reduce the ability of the library to increase the concentration of every low-concentration protein species.

When the serum proteins are bound to library and sequentially eluted by different modes of displacement, the complexity of the eluate is significantly reduced and, as a result, the number of detectable species again increased compared to a treatment followed by a single ion exchange step (data not shown); thus, the described method simultaneously decreases the concentration range and fractionates the samples. Repeated treatments of eluates with the ligand library should further decrease the concentration range differential and further improve detection of trace species. The entire method is very straightforward, rapid, and applicable to a wide variety of starting materials, potentially including whole blood immediately upon collection, thereby improving sample handling and facilitating biomarker detection.

The greatest increase in detectable species results with the largest sample/library ratio. Although this may be seen as a technological limitation, a reduction in the particle size of the combinatorial solid-phase library resolves the question quite easily while maintaining the diversity of the ligands in a very small volume. For example, reduction of bead size to < 1 μm provides sufficient diversity in 20 μL of beads.

The combinatorial library approach increases the number of species detectable and represents a process to either replace, or supplement, current fractionation methods such as ion exchange. Based on MS, the number of detected species following incubation of serum with library was higher compared to fractionation of serum with a regular anion exchange column (486 peaks against 453; representing a 7% increase). When the serum is incubated with library followed by anion exchange, the number of detectable proteins compared with starting serum increased even further by about 72%.

We expect the described process will enhance the investigation of low-abundance proteins. Used as a standalone method or in association with established fractionation methods (generic or specific) as described in a recent review [32], it will constitute an additional tool of discovery for further proteomic investigations. It has not escaped our attention that cells, viruses, organelles, and nonproteinaceous components can also be processed using this technology and, with the ability to elute proteins nondestructively under mild conditions, it is possible to evaluate them more readily for the biological activity of trace components.

It is envisioned that the described approach, with or without additional fractionation methods, will greatly improve the detection of biomarkers of pathological relevance that are currently unavailable.

Authors thank Professor P.G. Righetti and Dr. A. Castagna from University of Verona (Italy) for their assistance in 2-DE map analysis and Dr. Scot Weinberger (Ciphergen Biosystems) for his critical support and valuable suggestions in the preparation of the manuscript.

Received February 21, 2005

5 References

[1] Anderson, L. N., Anderson, N. G., *Mol. Cell. Proteomics* 2002, *1*, 845–867.

[2] Pieper, R., Gatlin, C. L., Makusky, A. J., Russo, P. S., *Proteomics* 2003, *3*, 1345–1364.

[3] Lopez, M. F., *Electrophoresis* 2000, *21*, 1082–1093.

[4] Wilkins, M. R., Williams, K. L., Appel, R. D., Hochstrasser, D. F. (Eds.), *Proteome Research: New Frontiers in Functional Genomics*, Springer, Berlin 1997, p. 243.

[5] Hamdan, H., Righetti, P. G., *Proteomics Today*, Wiley, Hoboken 2005, p. 429.

[6] James, P. (Ed.), *Proteome Research, Mass Spectrometry*, Springer, Berlin 2001, p. 274.

[7] Leatherbarrow, R. J., Dean, P. D. G. A., *Biochem. J.* 1980, *189*, 27–34.

[8] Birch, R. M., O'Byrne, C., Booth, I. R., Cash, P., *Proteomics* 2003, *3*, 764–776.

[9] Maccarrone, G., Milfay, D., Birg, I., Rosenhagen, M., Holsboer, R., Grimm, R., Bailey, J., *et al. Electrophoresis* 2004, *25*, 2402–2412.

[10] Mehta, A. I., Ross, S., Lowenthal, M. S., Fusaro, V., *Dis. Markers* 2002, *18*, 1–10.

[11] Björhall, K., Miliotis, T., Davidsson, P., *Proteomics* 2005, *5*, 307–317.

[12] Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., Aebersold, R., *Proc. Natl. Acad. Sci. USA* 2000, *97*, 9390–9395.

[13] Yamada, M., Murakami, K., Wallingford, J. C., Yuki, Y., *Electrophoresis* 2002, *23*, 1153–1160.

[14] Bjellqvist, B., Sanchez, J. C., Pasquali, C., Ravier, F., *Electrophoresis* 1993, *14*, 1375–1378.

[15] Westbrook, J. A., Yan, J. X., Wait, R., Welson, S. Y., Dunn, M. J., *Electrophoresis* 2001, *22*, 2865–2871.

[16] Cho, M. J., Jeon, B. S., Park, J. W., Jung, T. S., *Electrophoresis* 2002, *23*, 1161–1173.

[17] Wilchek, M., Miron, T., Kohn, J., *Methods Enzymol.* 1984, 3–55.

[18] Merrifield, R. B., *Science* 1965, *150*, 178–185.

[19] Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., *Nature* 1991, *354*, 82–84.

[20] Furka, A., Sebestyen, F., Asgedom, M., Dibo, G., *Int. J. Pept. Protein Res.* 1991, *37*, 487–493.

[21] Watts, A. D., Hunt, N. H., Hambly, B. D., Chaudhri, G., *Electrophoresis* 1997, *18*, 1086–1091.

[22] Buettner, J. A., Dadd, C. A., Baumbach, G. A., Masecar, B. L., Hammond, D. J., *Int. J. Pept. Protein Res.* 1966, *47*, 70–83.

[23] Huang, P. Y., Baumbach, G. A., Dadd, C. A., Buettner, J. A., *Bioorg. Med. Chem.* 1996, *4*, 699–708.

[24] Kaufman, D. B., *Biotechnol. Bioeng.* 2002, *77*, 278–289.

[25] Buettner, J. A., Dadd, C. A., Baumbach, G. A., Masecar, B. L., Hammond, D. J., *Int. J. Pept. Protein. Res.* 1996, *47*, 70–83.

[26] Bastek, P. D., Land, J. M., Baumbach, G. A., Hammond, D. H., Carbonnel, R. G., *Sep. Sci. Technol.* 2000, *35*, 1681–1706.

[27] Link, A. J., Robison, K., Church, G. M., *Electrophoresis* 1997, *18*, 1259–1313.

[28] Wiesner, A., *Curr. Pharmacol. Biotechnol.* 2004, *5*, 45–67.

[29] Guerrier, L., Lomas, L., Boschetti, E., *J. Chromatogr.* 2005 (in press) *1073*, 25–33.

[30] Corthal, G. L., Wasinger, V. C., Hochstrasser, D. F., Sanchez, J.-C., *Electrophoresis* 2000, *21*, 1104–1115.

[31] Fountoulakis, M., Takas, M. F., Berndt, P., Langen, H., Tacaks, B., *Electrophoresis* 1999, *20*, 2181–2195.

[32] Righetti, P. G., Castagna, A., Antonioli, P., Boschetti, E., *Electrophoresis* 2005, *26*, 297–319.