

Protein Sample Treatment with Peptide Ligand Library: Coverage and Consistency

Lei Li¹, ChengJun Sun¹, Steve Freeby¹, Dennis Yee¹, Sylvie Kieffer-Jaquinod²,
Luc Guerrier³, Egisto Boschetti^{3*}, Lee Lomas¹

¹Bio-Rad Laboratories, Hercules, CA 94547, USA

²DSV/IRTSV Laboratoire EdyP, CEA, 38054 Grenoble, France

³Bio-Rad Laboratories, 92430 Marnes-la-Coquette, France

Abstract

Low-abundance protein detection in biological samples is one of the main challenges in proteomics investigations. One approach that makes the detection of rare species possible is the treatment of biological samples with solid-phase combinatorial peptide ligand libraries. However, the use of combinations of ligands opens an uncertainty in that, since the diversity of the library is very large, aliquots of beads sampled from the library might not have fully comparable bead species each time. Reproducibility of experimental data with highly diverse libraries is therefore a main concern to address.

This paper reports reproducibility data when aliquots of similar and different volumes of libraries are used at a certain sample to library ratio. Eluates from ligand libraries and other fractions are analyzed using various complementary methods such as two-dimensional gel electrophoresis, immunoassay and mass spectrometry.

The collected data show a high level of consistency from sample to sample when processed with similar and variable bead volumes. Analytical determinations are all convergent with each other in considering the similarity of results. It is anticipated that this demonstration reinforces the possibility that differential proteomics studies, in particular for the discovery of protein targets of interest, can effectively be accomplished with combinatorial peptide libraries.

Keywords: Proteomics; Peptide library; Cytokine; Human plasma; Reproducibility; Quantitation

Introduction

Turk et al., (2003) pioneered the idea that peptide libraries would provide a valuable set of tools able to enhance the understanding of protein interaction domains and consequently contribute to the elucidation of protein interaction networks. In the past few years, this vision has become realized in at least one important aspect of proteomics investigations and several published papers now describe the use of combinatorial peptide ligand libraries for the treatment of protein extracts. This technology is based on libraries of very large ligand diversity and, when used in overloading conditions, allows for reducing the concentration difference between the most concentrated proteins and the rarest proteins. This is accomplished by a concomitant dilution of high-abundance proteins, due to specific ligand saturation, and a concentration of low-abundance species

(Thulasiraman et al., 2005; Guerrier et al., 2008; Righetti et al., 2009). By that way it becomes possible to detect species that are normally undetectable either due to detection interference of the most abundant proteins (e.g. the case of albumin in serum) or because their concentration is below the sensitivity of the analytical methods used.

An increasing number of examples now illustrate the utility of this approach, not only to enlarge the number of protein species that can be detected within a complex protein sample, but also in the area of allergen detection (Bachi et al., 2009) and impurity trace analysis (Fortis et al., 2007). Most recently the use of peptide ligand libraries has for the first time provided a link between genomics and proteomics (Bianchi et al., 2009) where it was possible to contribute to the understanding of congenital dyserythropoietic anemia type II (CDAII) symptom.

Many low-abundance new gene products have been found after sample treatment with peptide libraries. Some of them were known and/or expected (see for example Castagna et al., 2005, Guerrier et al., 2007; Roux-Dalvai et al., 2008), while some others found in biological fluid samples from various species, were unexpected as reported in human serum (Sennels et al., 2007), chicken egg yolk (D'Ambrosio et al., 2008) and snake venom (Calvete et al., 2009). Moreover after peptide library treatment, the cell extracts revealed traces of gene products that are normally repressed in mature cells (Roux-Dalvai et al., 2008).

Generally most peptide libraries are composed of tens of millions of diversomers (each unique peptide attached to a single bead); therefore to be certain that the entire library used for the treatment of a sample is statistically represented, a large volume of beads should be taken. This is in contrast to what is possible in practice and also with observed data. To elucidate the situation a progressive explanation was given with the support from different experimental results (Boschetti et al., 2007). Basically it was observed that one single bead (hence one peptide diversomer) is capable to interact with more than one protein and also that one single protein can be captured by several beads,

***Corresponding author:** Egisto Boschetti, Bio-Rad Laboratories, 92430 Marnes-la-Coquette, France, E-mail: Egisto.boschetti@bio-rad.com

Received November 08, 2009; **Accepted** December 17, 2009; **Published** December 18, 2009

Citation: Li L, Sun C, Freeby S, Yee D, Kieffer-Jaquinod S, et al. (2009) Protein Sample Treatment with Peptide Ligand Library: Coverage and Consistency. *J Proteomics Bioinform* 2: 485-494. doi:10.4172/jpb.1000110

Copyright: © 2009 Li L, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

each of them carrying a different peptide. Additional studies on peptide libraries of different length supported the principle of a large competition between proteins present in the biological samples and hence generating competition effects during the protein capturing process on the beads (Simo et al., 2008).

In spite of several published reproducibility demonstrations, no single report has so far provided an exhaustive study of reproducibility and quantitation of captured proteins, particularly when the necessity to process small sample volumes dictates the use of small bead volumes that provide low total peptide library representation. It is thus the aim of the present work to make links between reproducibility and sample volume using the most known analytical methods for proteomics investigations.

Reported data herein contribute to the interpretation of peptide ligands-protein interaction phenomena, answering thus to questions of coherence between the extremely large size of peptide libraries and their member representation when bead draws are smaller than the size of the ligand collection.

Materials and Methods

Materials

Chemicals and biologicals such as ProteoMiner (combinatorial hexapeptide ligand library beads), molecular mass standards, SDS-PAGE precasted gels, 11-cm long IPG strips, Criterion 8-16% gels, CM-10 ProteinChip arrays, mass spectrometry calibration kit "All-in-one" and Human Cytokine 27-plex assay kit, were all from Bio-Rad Laboratories, Hercules, CA. All other chemicals used in the experimental work were pure analytical grade products and purchased from Sigma-Aldrich, St Louis, MO. Sequencing grade trypsin was from Promega, Madison, WI.

Human serum (without anticoagulants) and plasma (with K-2 EDTA anticoagulant) were from Bioreclamation, Hicksville, NY.

Apparatus such as electrophoresis systems, Bio-Plex and PCS 4000 mass spectrometer along with their dedicated accessories, software and chemicals were from Bio-Rad, Hercules, CA.

NanoLC-MS/MS system was composed of an HPLC Ultimate 3000, Dionex (Amsterdam, The Netherlands) and LTQ-Orbitrap system, Thermo Fischer Scientific (Bremen, Germany).

Human serum or plasma treatment with peptide ligand library (ProteoMiner)

One mL of human serum or plasma was centrifuged to perfect clarity and then used without preliminary treatments. ProteoMiner treatment was performed according to manufacturer's recommendation. Briefly, 100 μ L aliquots of ProteoMiner beads were pre-equilibrated with phosphate buffered saline (PBS, pH 7.4), loaded with 1 ml plasma samples and incubated for two hours at room temperature under constant rotation. After incubation, the beads were washed four times with 500 μ L of PBS to eliminate the unbound proteins. Captured proteins were eluted using a solution composed of 8M urea, 2% CHAPS and 5% acetic acid. Collected proteins were then analyzed according to methods described below.

A number of experiments with the same procedure were performed using 20 μ L and 50 μ L of peptide library beads instead

of 100 μ L; plasma loaded was adjusted to 200 μ L and 500 μ L respectively to maintain constant the sample volume to bead volume ratio.

2D-PAGE analysis

The desired amount of protein sample was solubilized in the "2D sample rehydration buffer" (7 M urea, 2 M thiourea, 2% CHAPS, 50 mM DTT, 2 mM TBP (tri-butylphosphine), 0.2% biolyte 5-8, and 0.002% bromophenol blue). 11-cm long IPG strips (Bio-Rad Laboratories), pH 5-8, were passively rehydrated with 100 μ g of total protein in 185 μ L of rehydration buffer for 12 hrs. Isoelectric focusing (IEF) was carried out at 250 Volts for 30 minutes followed by a linear rapid voltage gradient to 8000 Volts until reaching 35000 Vhr (current limit was set at 50 μ A) on IEF Cell (Bio-Rad Laboratories, CA). Focused strips were held under 500 V until ready for equilibration. For the second dimension, the focused IPG strips were first blotted against damp filter paper to remove excess mineral oil and then equilibrated for 15 min in Equilibration Buffer I containing 6 M urea, 375 mM Tris-HCl (pH 8.8), 2% SDS, 20% glycerol, and 2% (w/v) DTT followed by 15 min in Equilibration Buffer II containing 6 M urea, 375 mM Tris-HCl (pH 8.8), 2% SDS, 20% glycerol, and 2.5% (w/v) iodoacetamide under gentle shaking. The IPG strips were then briefly dipped into 1X TGS buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3), laid on top of 8-16% Criterion Tris-HCl gels (IPG + 1), and sealed with overlay agarose (0.5% low melting agarose in TGS buffer with 0.003% bromophenol blue). Electrophoretic run was performed at 200 V constant voltage until the dye front reached the bottom of the gel. Gels were fixed in 40% ethanol and 10% acetic acid for 3 hr and then stained overnight with Flamingo fluorescent stain. Stained 2-DE gels were scanned with a PharosFx imaging system (Bio-Rad) and gel images were captured via QuantityOne software (Bio-Rad) at 100 μ m resolution. Images were cropped and processed in QuantityOne and analyzed by SameSpot software (Non-linear). Signal intensity of each spot was compared across all the gels and statistical analyses used were ANOVA and PCA (principal component analysis) from the SameSpot software.

Immunoassay of cytokines by Bio-Plex

Each vial of cytokine standards from the Bio-Plex human cytokine 27-plex assay kit was reconstituted in 500 μ L human serum on ice for 30min. 800 μ L of the reconstituted cytokine standards was added into 7.2 mL of human serum. One mL of the cytokine spiked serum was incubated with 100 μ L of the equilibrated ProteoMiner beads in spin column at room temperature for 2 hrs. 200 μ L of above spiked serum was used for 20 μ L ProteoMiner beads. Triplicate spin columns were processed in the experiment. The flowthrough was collected and the columns were washed 3 times with PBS according to the standard ProteoMiner protocol. The bound proteins were then eluted with 100 μ L (20 μ L for the 20 μ L beads) elution buffer containing 8 M urea, 2% CHAPS and 5% acetic acid for 3 times. Cytokine signals from the spiked serum, flowthrough, combined washes and elution fractions were diluted in the sample diluent and analyzed using the standard Bio-Plex cytokine protocol. Briefly, 27-plex human cytokine panel beads were diluted 25 fold in the Bio-Plex assay buffer and 50 μ L was loaded into 96 well plate. 100 μ L of Bio-Plex wash buffer was used to equilibrate the beads. The cytokine standards were serially diluted and loaded

into each well and used as the reading standards. The spiked serum, flowthrough and combined washes were diluted 30 fold while the elution fractions 90 fold in the Bio-Plex sample diluent. 100 μ L of each diluted sample was added in triplicate into each well containing the 27-plex cytokine beads. The plate was incubated at room temperature on a plate shaker for 30min. Then the wells were washed 3 times with 100 μ L of Bio-Plex wash buffer. The detection antibody solution was diluted 10 fold in the detection antibody diluent and 25 μ L was applied into each well. After incubation on a plate shaker for 30min with aluminium foil cover, the plate was washed 3 times with 100 μ L of Bio-Plex wash buffer. The streptavidin-PE solution was diluted 100 fold in the Bio-Plex assay buffer and 50 μ L was added into each well. After incubation on a plate shaker for 10min with aluminium foil cover, the plate was washed 3 times with 100 μ L of Bio-Plex wash buffer. The beads in each well were then resuspended with 125 μ L of Bio-Plex assay buffer. The plate was read on the Bio-Plex system and the signals of each cytokine from different samples were calculated using the standard curve from the diluted cytokine standards.

NanoLC-MS/MS protocol and analysis

200 μ L of plasma sample was treated with 20 μ L of ProteoMiner library following the same protocol as mentioned above. For each set of nanoLC-MS/MS experiment, an estimated quantity of 30 μ g of human plasma protein obtained from peptide library were separated by SDS-PAGE using a 4-12 % gradient polyacrylamide precast gel plate.

Protein bands were manually excised (two trials: 10 bands and 20 bands) from the gels and transferred into 96-well microtitration plates. The following sample preparations were performed automatically using Freedom EVO150 robot, (Tecan Tracing AG, Switzerland). Excised gel samples were washed several times by incubation in 25 mM NH_4HCO_3 for 15 min and then in 50% (v/v) acetonitrile containing 25 mM NH_4HCO_3 for 15 min. Gel pieces were then dehydrated with 100% acetonitrile and then incubated with 7% hydrogen peroxide for 15 min before being washed again with the solution described above. 0.15 μ g of modified sequencing grade trypsin in 25 mM NH_4HCO_3 was added to the dehydrated gel bands. After 30 min incubation at room temperature, 20 μ L of 25 mM NH_4HCO_3 were added on gel pieces before overnight incubation at 37°C. Peptides were then extracted from gel pieces in three 15 min sequential extraction steps in 30 μ L of 50% acetonitrile, 30 μ L of 5% formic acid and finally 30 μ L of 100% acetonitrile. The pooled supernatants were then transferred into microcentrifuge tubes and dried under vacuum.

Dried extracted peptides were resuspended in water containing 2.5% acetonitrile and 0.1% trifluoroacetic acid before being transferred in vials compatible with nanoLC-MS/MS analysis system. The method consisted of a 60 minute gradient at a flow rate of 300 nL/min using a gradient from two solvents: A (2 % acetonitrile and 0.1% formic acid in water) and B (80% acetonitrile and 0.08% formic acid in water). The system included: a 300 μ m X 5 mm PepMap C18 precolumn in order to pre-concentrate peptides and a 75 μ m X 150 mm C18 column (Gemini C18 phase for in-house built columns) used for peptide elution. The instrument was calibrated each week with a mixture of caffeine, MRFA and Ultramark and was stable during one week with

a mass shift precision below 5 ppm. MS and MS/MS data were acquired using Xcalibur (Thermo Fischer Scientific) and processed using Mascot Daemon software (Matrix Science) Intranet version 2.0. We used a SwissProt/Trembl_decoy database with *Homo sapiens* as the taxonomy. Peptide variable modifications allowed during the searches were: N-acetyl (protein), dioxidation (M), oxidation (M) and cysteic acid (C). The other parameters were: peptide tolerance = 10 ppm, MS/MS tolerance = 1 Da, 2 missed cleavage site by trypsin allowed.

Each .dat file resulting from the Mascot processing of one gel band was filtered through IRMa (a homemade parsing solution) (Dupierris et al., 2009) with a p value of 0.05 and using a filter on the rank (only rank 1 peptides were retained). The filtered results were downloaded into a MSI database. The average FP rate of the results was less than 2 %. A homemade tool (Heidi) was used for the compilation, grouping and comparison of the proteins. At final, all proteins with less than 2 peptides were filtered out leaving no more FP protein identifications.

SELDI-TOF-MS protein pattern analysis

Human serum eluates from various experiments (peptide beads from three different lots and six sample treatment replicates from a single peptide beads lot) were analyzed by mass spectrometry using CM-10 ProteinChip arrays associated with SELDI-TOF-MS reader.

15 μ g of each protein sample from the elutions was diluted in 200 μ L CM-10 low stringency buffer and deposited on spots of the array surface in the Bioprocessor device. After sample application and incubation at room temperature for 60 min on a microplate shaker, the chip surface was washed three times with the CM-10 low stringency buffer and one time with de-ionized water to remove non-associated protein and dried. Then the arrays were prepared for the analysis by application of two times 1 μ L of energy adsorbing matrix solution composed of a half-saturated solution of sinapinic acid in 50% acetonitrile and 0.5% trifluoroacetic acid. Arrays were then analyzed with a PCS 4000 ProteinChip Reader. The instrument was used in a positive ion mode, with an ion acceleration potential of 25 kV. The laser intensity was set at 1,500 nJ. The mass range investigated was from 2 to 15 kDa. The instrument was mass calibrated using "All-in-one" protein standard kit.

Results and Discussion

The demonstration of the capability of peptide ligand libraries to decrease high-abundance species while enriching low- or very low-abundance species had quite frequently generated questions about the reproducibility of experimental data. Reviewers of submitted papers for publications often requested explanations and data unambiguously indicating that the treatment of biological samples with peptide libraries gave consistent results.

Indications of reproducibility were brought in various circumstances, as for instance in the discovery of novel proteins from human serum (Sennels et al., 2007) and proteins from red blood cell lysate (Roux-Dalvai et al., 2008) where either electrophoretic analysis or mass spectrometry were performed as analytical methods. Nevertheless the reproducibility and the quantitation aspects of the technology remained one of the most frequently asked questions.

The dilemma is generated by the fact that, the library being an assembly of beads each of them carrying a distinct ligand, when a certain volume of beads is taken out of a bulk, the proportions of peptide ligands represented on the beads are not statistically identical. The dilemma becomes even more concerning when one considers that working with very small volumes of biological samples also restricts the volume of beads that can be used and in these situations, the total number of library diversomers largely exceeds the number of beads drawn practically for a given experiment. In these conditions reproducibility of data and rational explanations are legitimized requests.

Logically the number of diverse peptides depends on the number of building blocks (selected amino acids in this case) and on the length of the peptide. As described by Boschetti and Righetti, 2008, the number of diversomers used in the current combinatorial hexapeptide ligand library (see Figure 1A solid line) is about 16.8 million of units. Since each peptide ligand is attached to a different bead, the entire collection of beads covering all diversity represents at least between 4.5 and 5 mL of settled material. This calculated volume results from the bead diameter that is here considered of being 65 μm in average. In practice this situation is less realistic because when randomly sampling the solid-phase library out of a bulk, there is no certainty to get one representative for each diversomers. Statistically the

phenomenon is well represented by Maillard et al., 2009, where it appears reasonable to admit that the representation of about 90% of the library is reached when the number of beads exceeds by about a factor of 2.43 the theoretical number of diversomers (see Figure 1B). Thus the bead volume necessary to comprise about 90% of the library would be 10.9-12.1 mL instead of 4.5-5 mL. Such a situation is not workable in practice because the amount of proteins required to treat such a large volume of beads would be unpractical. Most reported experiments involving various proteomes were performed using either 1 mL or 100 μL of beads. One mL comprises about 3.6 million beads and therefore only about 21% of the entire library. In spite of this “small” volume of beads the experimental data have always been observed to be reproducible. To counterbalance this assessment, it has been several times demonstrated that a given protein can easily be captured by various peptide structures (Huang et al., 1996; Kaufman et al., 2002) and also that a single peptide structure (single bead) can easily capture various species (Boschetti et al., 2007).

It is within this context that we started an experimental comparative work focusing exclusively on reproducibility of sample treatment with the described peptide ligand library. The same volume of beads was repeatedly taken, on one hand, and comparisons of data when decreasing the volume of beads were made, on the other hand.

The first set of experiments (Figure 2) was intended to demonstrate both the intrinsic reproducibility (same lot beads in replicates) and extrinsic reproducibility (bead library from different lots) of ProteoMiner peptide library. In each case 100 μL of beads were taken for the treatment of human serum. Protein eluates obtained with 8 M urea containing 2% CHAPS and 5% acetic acid were analyzed by two-dimensional gel electrophoresis as described in the Material and Method section. As a first comment the difference of patterns from the initial non-treated sample (Control) and library-treated sample was very significant in terms of spot number and spot positioning throughout the isoelectric points and the mass ranges. Manual and scanner examination of the maps after the library treatment did not result in significant difference in spot count, positioning and intensity. The reproducibility of data appeared thus very satisfactory and confirmed what was already reported in the literature for various biological samples and under different conditions for both mass ranges and peptides identification (Sennels et al., 2007; Roux-Dalvai et al., 2008). These data add to reproducibility of masses of intact species proteins by mass spectrometry before trypsination (Boschetti et al., 2007). Nonetheless it is to be noticed that the literature reported reproducibility data only for a single bead lot. Although it is not easy to argue about the spot intensity, it appeared that only few spots from different experiments showed a little different staining intensity.

From arguments related to peptide diversity and number of representative beads according to their volume, it was interesting to understand that a decrease of the volume of bead library and hence the increased risk of diversity differences between samples of combinatorial beads, would contribute to differences in protein patterns from one experiment to another.

Figure 3 represents two-dimensional gel electrophoresis of eluates obtained out of 20, 50 and 100 μL of ProteoMiner beads

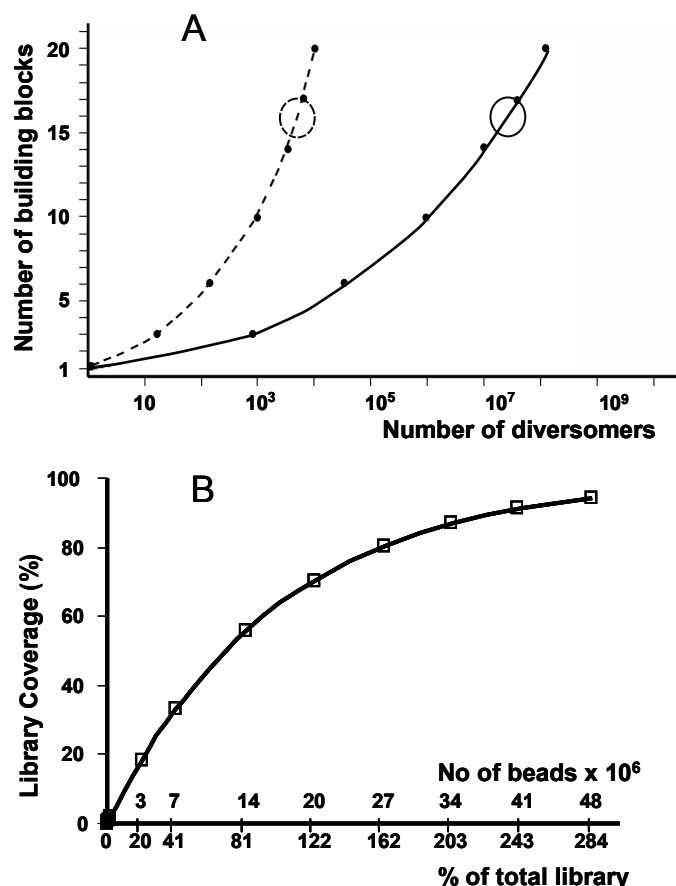


Figure 1: Combinatorial peptide library considerations. **A:** representation of the possible number of diversomers obtained using various amino acids. Broken line refers to the preparation of tripeptides and continuous line refers to the preparation of hexapeptide. When using 16 amino acids the number of hexapeptidomers is 16.8 million, while the number of tri-peptidomers is 4096 (see circles). **B:** representation of statistical library coverage as a function of the volume of beads taken (or the percent of volume of a hexapeptide library made using 16 amino acids) (see Maillard et al., 2009). For details see comments in the text.

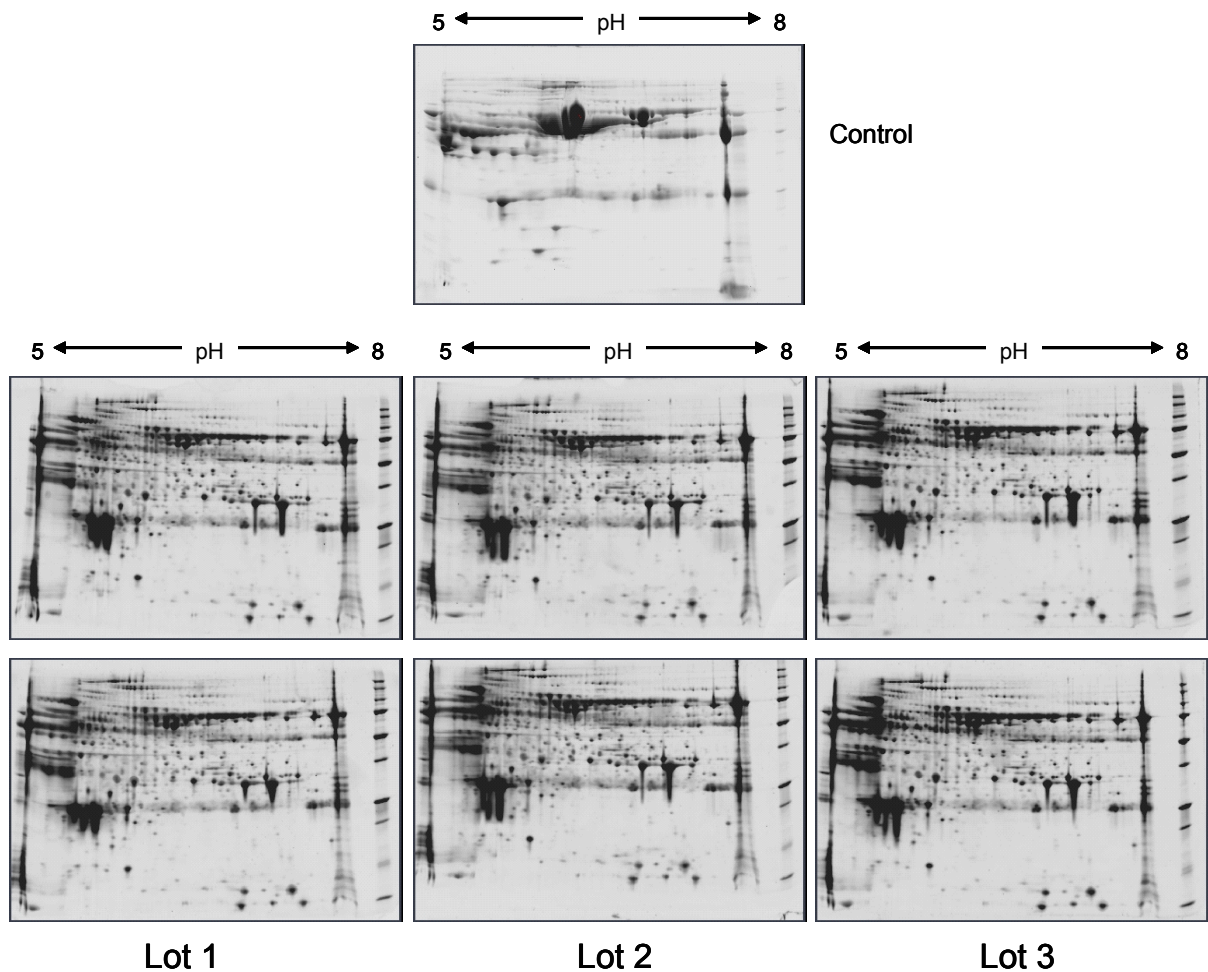


Figure 2: Two-dimensional gel electrophoresis of eluted proteins from peptide libraries in duplicate (middle and bottom images) by using all the time 100 μ L of beads from different batches ("Lot 1", "Lot 2" and "Lot 3"). Biological sample treated with peptide libraries was human serum. The volume of serum used each time was 1 ml for 100 μ L of beads. Upper panel (Control) represents the initial non treated biological sample.

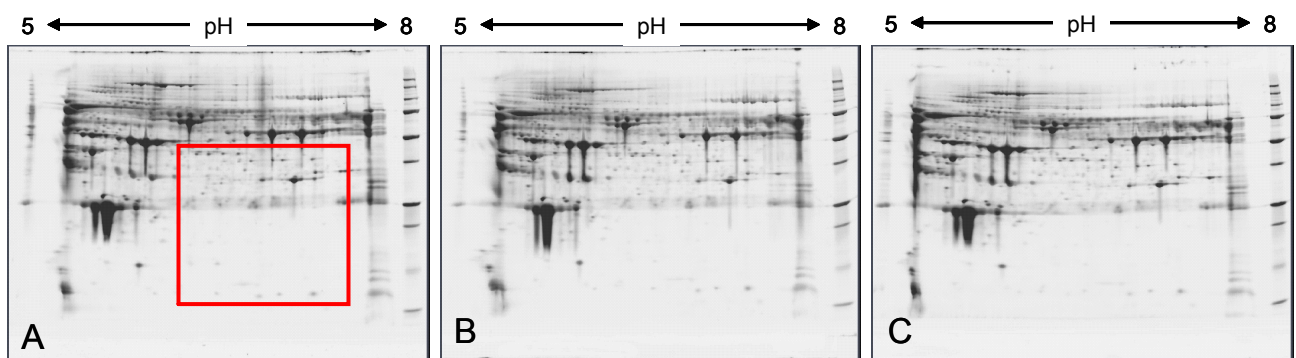


Figure 3: Two-dimensional gel electrophoresis of eluted proteins from peptide libraries when using various volumes of beads. The plasma volume/ beads volume ratio was always the same (1 mL per 100 μ L of beads). The volume of beads taken was of 20, 50 and 100 μ L for respectively "A", "B" and "C". Protein patterns did not significantly differ from each other. All trials were performed in duplicate (not shown). Red frame represents the zoom area where a more in-depth analysis was performed (see following figure).

treated with respectively 0.2 mL, 0.5 mL and 1 mL of human plasma. Three gels were run for each of the three columns per library volume. Overall spot patterns and count obtained did not show significant differences. When comparing the 2D gel runs for each bead library volume within a restricted zone as illustrated on Figure 4 (in duplicates), out of the 155 spots detected, there were three spots (2%) with intensity change of over two fold between 20 μ L vs. 50 μ L beads and 20 μ L vs. 100 μ L beads. Principal component analyses using SameSpot software showed no significant difference between the three volumes of combina-

torial beads. The same experiment was performed on 100 μ L and 20 μ L library using a human serum sample. Results showed that out of the 382 spots detected, only three spots (<1%) had intensity change of over two fold between the two volumes (data not shown). Therefore, there was no significant difference found among the various library volume tested.

Even with limited representation of ligands from the library, their performance as represented by the number of protein species captured from the plasma or serum was very comparable. Nonetheless two-dimensional gel electrophoresis gives only a qualitative image

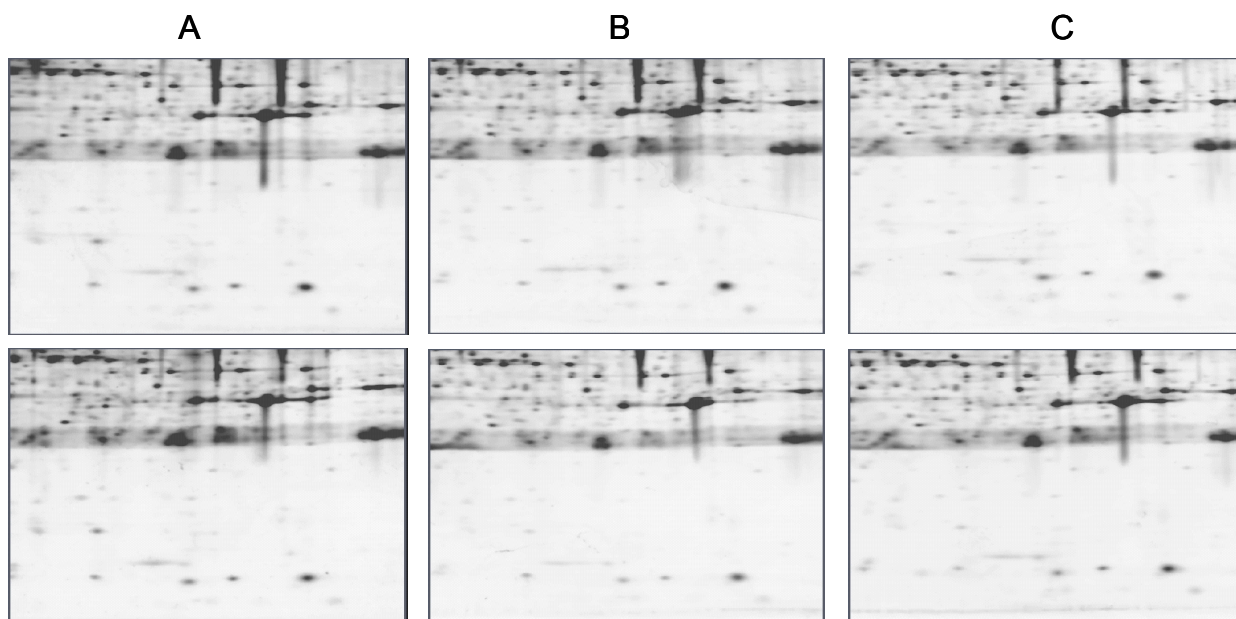


Figure 4: Regional analysis of two-dimensional electrophoresis of eluted plasma proteins from peptide libraries: blows-up of specific areas for an in-depth analysis of the spot patterns from Figure 3. Upper and lower series are duplicates; “A”, “B” and “C” are respectively images from 20, 50 and 100 μ L of beads.

of the reality even if spot intensity could be interpreted as a possible variation in the relative concentration of a given protein. To bring an answer to the quantitative reproducibility question, other experiments were performed using a multiplexed quantitative immunoassay of well-known proteins such as a panel of human cytokines using Bio-Plex system.

First, the serum samples containing spiked cytokine standards were treated with peptide library under different conditions and then the spiked serum, flowthrough, combined washes and the eluates assayed for the quantitative determination of each cytokine component in each fraction (for details see Material and Method section). The goal was not to know how much each cytokine was concentrated or not, but rather to see if the amount of given cytokines was reproducible from trial to trail and using various volumes of bead library. For exhaustive information for the reader it should be underlined that elution of proteins from bead libraries involving strong dissociating reagents (in this case

acidic urea containing a detergent) may have engendered denaturation of the eluted proteins with possible enhancement or reduction of molecular recognition towards their specific antibodies. This is the reason why in the present work we calculated the binding of cytokines to the beads by subtraction of the signal in the flow-through and washes from the original spiked serum to avoid possible misinterpretations. Nonetheless, the signals of each considered cytokines in the eluates showed very tight standard variation from different treatments as mentioned below.

Figure 5 assembles experimental data obtained in triplicate in the collected fractions from 20 μ L and 100 μ L bead libraries. What has been observed is that (i) the intrinsic reproducibility of each cytokine considered was quite satisfactory with a pooled standard variation of less than 7.6% in the eluates for both series of trials using 20 μ L and 100 μ L of beads, (ii) very similar proportionality of the cytokines was maintained whatever the volume of beads taken and (iii) the majority of the input up to more than 90%, even in ng range, of the most cytokines shown here, bound to the ProteoMiner beads, which indicated that the ProteoMiner beads can indeed enrich really low-abundance protein targets. Some of the human cytokines didn't bind significantly to the beads, which could be due to either competition with other high-abundance proteins, association with other proteins and thus depleted together with those proteins, or simply lack of the specific binding bead partners.

These quantitative data confirm what was also observed with two-dimensional gel electrophoresis of bead eluates (see above). Results from quantitative determinations are important when assessing the reproducibility of the library treatments especially when comparative analysis of biological samples are considered as for instance with the aim of biomarker discovery from drug effect or from targeted pathologies. Some preliminary attempts were made in a recent published paper by spiking biological extracts with exogenous proteins (Roux-Dalvai et al., 2008) prior to peptide library treatment. Excellent reproducible results were reported using HPLC and mass spectrometry analytical methods.

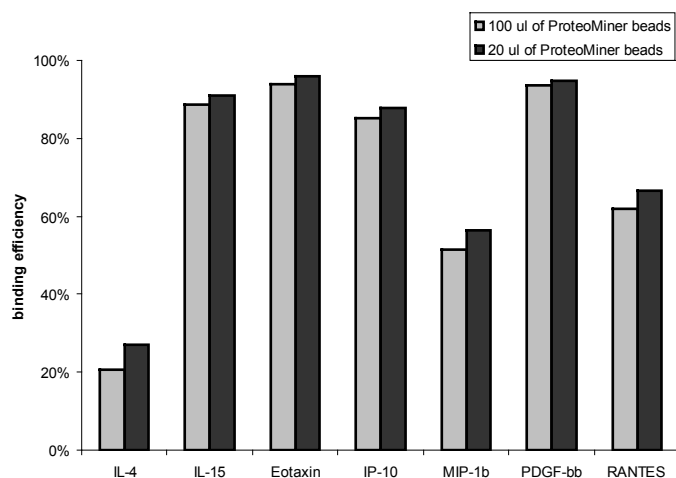


Figure 5: Multiplexed immunoassays of several bound human cytokines after treatment of cytokine spiked samples with peptide ligand libraries. The volume of ProteoMiner beads used was either 100 μ L (light grey bars) or 20 μ L (dark grey bars). Experiments were made in triplicates and shown values averaged.

To complete the present reproducibility work another set of experiments was performed involving nanoLC-MS/MS as the analytical method. Two parallel trials performed using 20 μ L of ProteoMiner beads and 200 μ L of human plasma produced eluates that were both analyzed by nanoLC-MS/MS (see Material and Method section). For each eluate three separate analyses were performed so that comparisons could be made between replicates (reproducibility of the analytical method) and between two eluates obtained using two different samples of peptide library beads.

As shown in Figure 6A and 6B, replicate data are quite comparable with an extremely large overlap of gene products found for both plasma eluates. Replicates from the first sample show an overlap of species of about 88%; replicates from the second sample show an overlap of about 85%. The comparison of merged data from each sample illustrated in Figure 6C also shows a very large overlap of common gene products with only 10 to 17% species outside the common pool. Considering that in this series of experiments four different types of reproducibility are summed up (sample treatment, SDS-PAGE separation, HPLC fractionation of peptides and MS/MS analysis), it can be assessed the good consistency of sample treatment with ProteoMiner confirming all other data discussed above using alternative analytical approaches. As an illustration of reproducibility of HPLC, Figure 7 reports HPLC patterns of two distinct eluates from two separate experiments after SDS-PAGE and trypsin digestion. After elution of captured proteins followed by SDS-PAGE analysis, gel slicing, trypsin breakdown of separated proteins and fractionation of resulting peptides by HPLC, it is remarkable to observe that the patterns are very consistent in terms of elution of species positioning. Proportionality of signal intensity of fractionated peptides is also well conserved. Data of HPLC peptides from trypsinized serum before and after treatment with ProteoMiner were recently reported in an earlier published

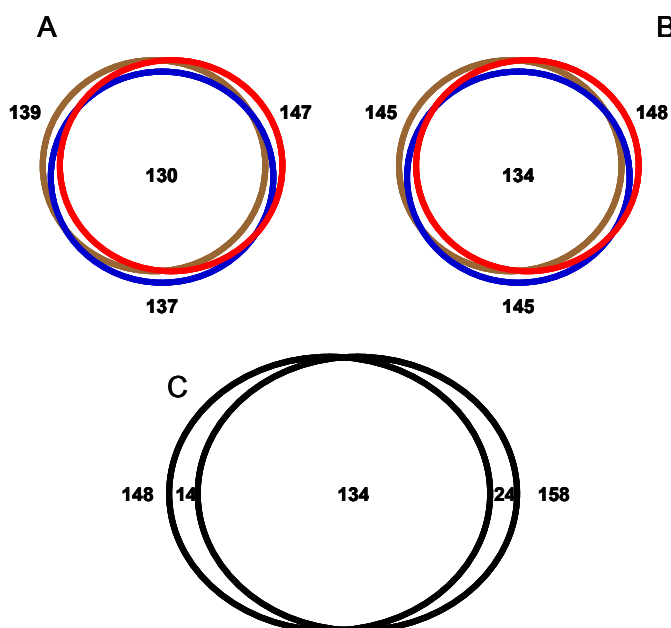


Figure 6: Overlap diagrams of gene products from nanoLC-MS/MS results. **A:** replicates (three) from a first peptide ligand beads eluate (overall 148 unique proteins). **B:** replicates (three) from a second peptide ligand beads eluate performed under the same conditions (overall 158 unique proteins). **C:** Comparison of global gene products data from the first (left circle) and the second eluate (right circle).

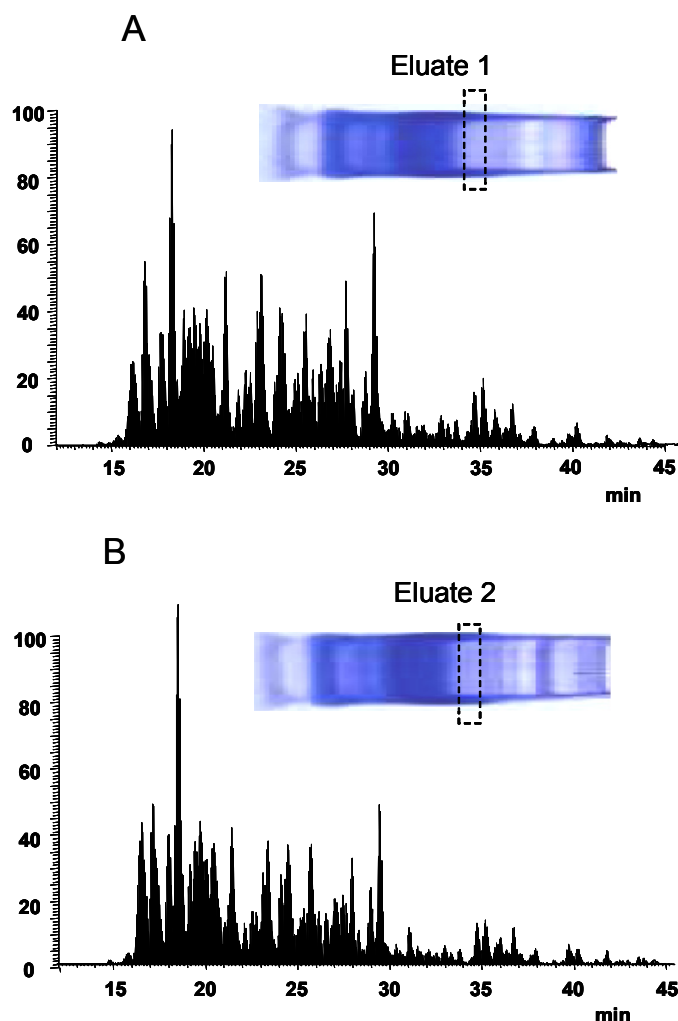


Figure 7: Nano-HPLC comparative patterns of two distinct ProteoMiner experiments with human plasma. Eluates (A: eluate 1 and B: eluate 2) were first separated by SDS-PAGE, then a slice of gel was treated with trypsin (see box) and then resulting peptides fractionated by HPLC. For more details see "Materials and Methods" section.

paper (Righetti et al., 2009) where the authors showed a strong modification of patterns demonstrating the huge effect of peptide library in evidencing many more peptide species. This fact brings even more emphasis on the reproducibility of HPLC data after sample treatment.

NanoLC-MS/MS data reported above were somewhat expected since the data resulting from SELDI-TOF analysis of several dozens of individual peak signals (whole non-trypsinized proteins) were also very consistent. Here also the analysis was performed in triplicates and intensities of same peaks (exactly same mass), corresponding to most likely the same polypeptides, were plotted. For both experiments, comparison of different bead library lots (Figure 8A) and six replicates from a same bead library (20 μ L beads with 200 μ L serum, Figure 8B), dispersion of data was very narrow with pooled CV not larger than 20.4% of the signal intensity from 34 distinct species. These very similar data were confirmed by nanoLC-MS/MS after sample treatment (Figure 6). Naturally the automatic approach to treat each sample with well-defined protocols played a role in the reduction of error margins. What nanoLC-MS/MS brought over SELDI-TOF data was the certainty that patterns were well compared by using reconstituted identity of gene products from the sequence of peptides.

Overall the present study with human serum or plasma treated with peptide library beads analyzed by very different approaches unambiguously demonstrate the high reproducibility of the technology as it was already perceived here and there in published application reports. For instance plasma protein recovery upon peptide treatment was very consistent with CVs lower than 10% (Sihlbom et al., 2007) when using a single elution process. Authors reported experimental data concluding not only on improved signal intensity of low-abundance species, but also “in a reproducible fashion”. Lastly a published paper reported the demonstration that when spiking a sample of human serum with various amounts of *E. coli* extract the relative proportionality between proteins was very well maintained (Hartwing et al., 2009). Authors assessed the reproducibility of the sample treatment with ProteoMiner by using 100 μ L of beaded peptide library with respect to proteins of low abundance (spiked *E. coli* proteins). Pre-labelled *E. coli* proteins allowed demonstrating the preservation of quantitative ratios between species and between samples. Quantitative curves were built with the

impression that the quantitation was met as long as considered proteins are not present at saturating amount.

Although the reported work was performed with the aim of demonstrating the possibility to use a library for quantitative purposes, repeated sample treatments with bead volumes under representing the entire library (library coverage was estimated close to 2-3%), results confirmed what is reported in our experiments.

Unlike already published papers where an initial native biological sample was compared to the same treated sample and differences discussed and used for novel gene product discoveries, the present work focuses on data consistency. Since reported data cross over various conditions and obtained samples are analyzed by different methods, this report contributes to a better understanding of the complex interaction mechanism between ligand library and proteins from crude biological samples. Reproducible results are actually of utmost importance when considering proteomics differential studies and quantitative determinations.

All the reported analytical data may need explanations, some of them rational and some others more speculative. If one considers that each single peptide ligand (basically a single bead) can capture one protein, clearly volumes of beads used do not represent the whole library structural coverage and hence the reproducible data obtained cannot be rationally explained. Therefore the concept based on one-ligand-one-protein is here to be modulated a bit.

As already reported, a single bead is capable of capturing more than a single species (Huang et al., 1996; Miyamoto et al., 2008) with presumably different association constants. The interaction phenomenon being governed by the mass action law, displacement effects are not only dependent on affinity constants but also on the relative concentration of species having affinity for the same peptide ligand. This phenomenon was described when elucidating the influence of peptide length on protein capture (Simo et al., 2008), on one hand, and more recently when reporting protein patterns when the interaction was performed at different pHs (Fasoli et al., 2009), on the other hand. Similar peptide libraries were described also as a source of affinity ligands for protein purification with impressive results for a number of proteins (Bastek et al., 2000; Kaufman et al., 2002; Yang et al., 2005). For a given protein more than one single peptide was systematically identified with different adsorption-desorption properties. Structurally identified peptides although not identical were quite similar and they generally shared three amino acids. As reported by Huang et al., (1996), about two dozens of peptides were identified for the capture of von Willebrand factor from human plasma. Some peptides released the captured proteins upon sodium chloride wash, others after contact with a solution of acetic acid. Partition effect of the target protein between the solid and the liquid phase was also reported at the capturing phase as well at the elution stage. These phenomena illustrate the influence of the affinity strength with the putative peptide ligand. In the present work it should be reminded that the interaction between the peptide library and the complex sample is made under very large overloading conditions with a binding capacity saturation phenomenon (Thulasiraman et al., 2005). Massive amounts of proteins influence the effect of the

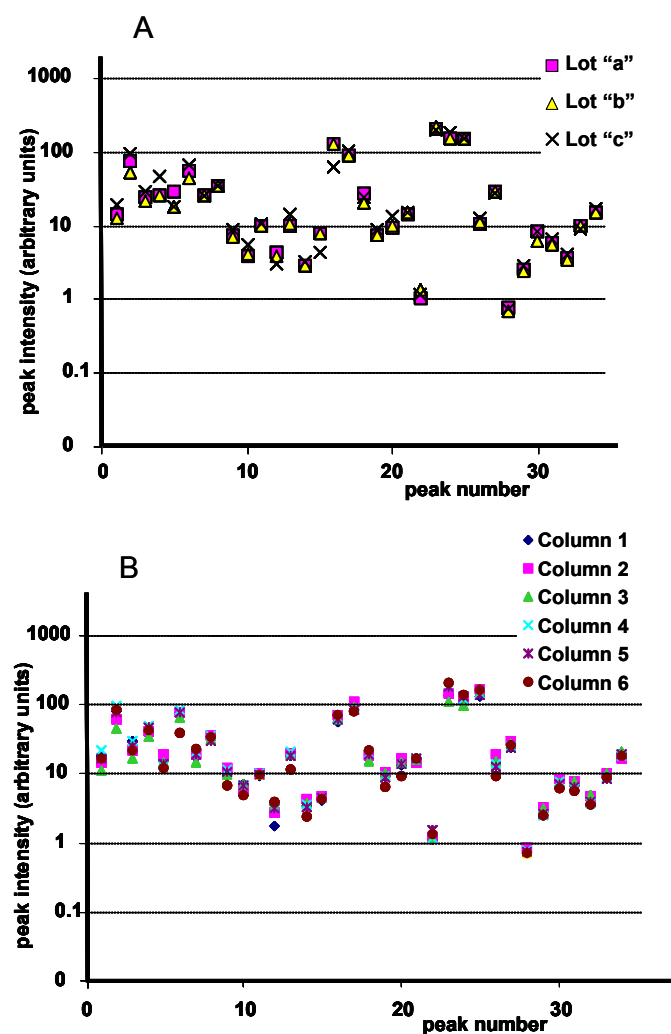


Figure 8: Variation of mass spectrometry signal intensity as a function of various trials. Experiments were performed using SELDI-TOF-MS with CM-10 ProteinChip array.

A: Comparison between three different lots of peptide ligand beads (same lots “1”, “2” and “3” from Figure 2). For details see Material & Method section. Sample treatments were performed using each time 1000 μ L human serum loaded on 100 μ L peptide beads.

B: Comparison between six different serum treatment replicates with peptide ligand beads. Sample treatments were performed using each time 200 μ L human serum loaded on 20 μ L peptide beads.

dissociation constant especially for species that are very concentrated such as albumin, and others that are present only in trace amount. The latter being submitted to a more effective competition, must have strong affinity for the peptide ligand to be effectively captured. It is the association of affinity constants and relative concentrations of species, both thermodynamic parameters, that justify the co-capture of more than a species by a single bead, hence a single peptide structure (Boschetti et al., 2007; Fasoli et al., 2009).

To fully understand the reproducibility question a couple of other considerations are important. One of them is the real diversity of peptides in terms of functionality for affinity dockings and the length of the peptide. Peptides that differ from each other because one glycine is replaced by an alanine or because isoleucine is replaced by valine have probably very similar capturing properties for a given protein. Since sequences of three amino acids seemed enough for the capture of a given protein (see above), similar properties are expected with hexapeptides having the same three distal amino acids even if the three other proximal amino acids are different. As recently discussed (Simo et al., 2008), the length of the peptide ligand beyond the trimer does not impact significantly the capturing protein pattern. The exclusive contribution of elongated chains to tetra, penta- and hexa-peptides has in fact been described as being relatively marginal. This phenomenon is interesting because, while with hexapeptides obtained with for example 16 amino acids the number of diversomers is 16.8 millions, with tripeptides the number of diversomers is reduced to 4096. Translated into the number of beads and their representative volume it means that the minimum volume of 65 μm beads comprising 16.8 millions of diversomers is 4.6 mL while the minimum volume of beads comprising 4096 diversomers is only as little as 1.12 μL of beads. Considering the probability to have the largest representation of the diversomers it would need about two times the minimum volume (see Figure 1B), a good reproducibility would be reached using 2.5-3 μL of beads.

Integrating the phenomenon of displacement described above and the question of the peptide length involvement, both converge towards the use of a number of beads much smaller than the size of the library without major impact on the detectable pattern of the treated sample.

Conclusion

It is believed that the presented work unambiguously demonstrated the reproducibility of sample treatment with peptide ligand libraries. There is a clear convergence of data from various types of analysis of eluates from ProteoMiner substantiating high reproducibility whatever the analytical method used (SDS-PAGE, two-dimensional gel electrophoresis, LC-MS/MS, SELDI-TOF and immunoassays). Replicates from a defined experiment, parallel experiments with various lots of bead libraries as well as changes in volumes of peptide library beads with proportionally the same amount of sample, always resulted in highly similar results. These experimental facts engendered complementary considerations around the mechanism of interaction of the technology. Discussions around the library coverage highlighted the fact that it is not necessary to deal with the entire library due to the multiple interaction possibilities involved and, at the same time, to the concentration of species and

their relative affinity constant. Peptide structures, although different from each other and in very large number, show situations of functional similarity especially when similar compositions of amino acids are present. In these situations the number of diverse peptides representing the library does not result from statistical calculations; rather samples of hexapeptide library as small as 1-10% of coverage is enough to obtain consistent results from sample to sample. Nonetheless reproducibility of sample treatment is dependent on a strict control of physicochemical environmental parameters such as pH and ionic strength as it is recently extensively discussed (Fasoli et al., 2009).

With this possibility in mind it becomes feasible to use the library with very small biological samples especially for comparative investigations as demonstrated for instance in differential profiling studies for the detection of markers of diagnostic interest (Sihlbom et al., 2007; Au et al., 2007; Petri et al., 2009).

Reproducible methods are important when considering quantitative determinations; given the data and discussion presented in this work, the conservation of relative concentration of species after library treatment occurs even with very small bead volumes.

As a final argument it could also be indicated that, since proteins are captured by various interaction forces, sequential elutions would not only simplify the analysis of treated samples, but also would allow focusing on classes of proteins with common structural or functional properties.

Acknowledgements

The author team wishes to thank Dr Qian-Shu Wang from Bio-Rad Laboratories, Hercules, for his appreciated help for BioPlex technical recommendations and data interpretation.

References

1. Au JS, Cho WC, Yip TT, Yip C, Zhu H, et al. (2007) Deep proteome profiling of sera from never-smoked lung cancer patients. *Biomed Pharmacother* 61: 570-577. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
2. Bachi A, Restuccia U, Fasoli E, Boschetti E, Peltre G, et al. (2009) In-depth exploration of cow's whey proteome via combinatorial peptide ligand libraries. *J Prot Res* 8: 3925-3936. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
3. Bastek PD, Land JM, Baumbach GA, Hammond DH, Carbonell RG (2000) Discovery of alpha-1-proteinase inhibitor binding peptide from the screening of a solid phase combinatorial library. *Sep Sci Technol* 35: 1681-1706. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
4. Bianchi P, Fermo Z, Vercellati C, Barcellini W, Iurlo A, et al. (2009) Congenital Dyserythropoietic Anemia Type II (CDAI) is Caused by Mutations in the SEC23B Gene. *Human Mutat* 30: 1292-1298. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
5. Boschetti E, Lomas L, Righetti PG (2007) Romancing the "hidden proteome", Anno Domini two zero zero six. *J Chromatogr A* 1153: 277-290. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
6. Boschetti E, Righetti PG (2008) Hexapeptide combinatorial ligand libraries: the march for the detection of the low-abundance proteome continues. *BioTechniques* 44: 663-665. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
7. Calvete J, Fasoli E, Sanz L, Boschetti E, Righetti PG (2009) Exploring the venom proteome of the western diamondback rattlesnake, *Crotalus atrox*, via snake venomomics and combinatorial peptide ligand library approaches. *J Prot Res* 8: 3055-3067. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
8. Candiano G, Bruschi M, Musante L, Santucci L (2004) Blue Silver: a very

- sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis* 25: 1327-1333. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
9. Castagna A, Cecconi D, Sennels L, Rappsilber J, Guerrier L, et al. (2005) Exploring the hidden human urinary proteome via ligand library beads. *J Prot Res* 4: 1917-1930. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
10. D'Ambrosio C, Arena A, Scaloni A, Guerrier L, Boschetti E, et al. (2008) Exploring the chicken egg white proteome with combinatorial peptide ligand library. *J Prot Res* 7: 3461-3474. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
11. Dupierris V, Masselon C, Court M, Kieffer-Jaquinod S, Bruley C (2009) A toolbox for validation of mass spectrometry peptide identification of database: IRMa. *Bioinformatics* 25: 1980-1981. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
12. Fasoli E, Farinazzo F, Sun CJ, Kravchuck AV, Guerrier L, et al. (2009) Interaction between proteins and peptide libraries in proteome analysis: pH involvement for a larger capture of species. *J Proteomics* [Epub ahead of print]. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
13. Fortis F, Guerrier L, Rinalducci S, Zolla L, Antonioli P, et al. (2007) Capturing and amplifying impurities from purified recombinant monoclonal antibodies via peptide libraries: a proteomics study. *Proteomics* 7: 1624-1633. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
14. Guerrier L, Claverol S, Fortis F, Rinalducci S, Timperio AM, et al. (2007) Exploring the platelet proteome via combinatorial hexapeptide ligand library. *J Prot Res* 6: 4290-4303. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
15. Guerrier L, Righetti PG, Boschetti E (2008) Reduction of dynamic protein concentration range of biological extracts for proteomics studies for the discovery of low-abundance species. *Nature Protocols* 3: 883-890. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
16. Hartwig S, Czibere A, Kotzka J, Paßlack W, Haas R, et al. (2009) Combinatorial hexapeptide ligand libraries (ProteoMiner™): An innovative fractionation tool for differential quantitative clinical proteomics. *Arch Physiol Biochem* 115: 1-6. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
17. Huang PY, Baumbach GA, Dadd CA, Buettner JA, Masecar BL, et al. (1996) Affinity purification of von Willebrand factor using ligands derived from peptide libraries. *Bioorg Med Chem* 4: 699-708. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
18. Kaufman DB, Hentsch ME, Baumbach GA, Buettner JA, Dadd CA, et al. (2002) Affinity purification of fibrinogen using a ligand from a peptide library. *Biotechnol Bioengin* 77: 278-289. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
19. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680-685. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
20. Maillard N, Clouet A, Darbre T, Reymond JL (2009) Combinatorial libraries of peptide dendrimers: design, synthesis, on-bead high-throughput screening, bead decoding and characterization. *Nature Protocols* 4: 132-142. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
21. Myamoto S, Liu R, Hung S, Wang X, Lam KS (2008) Screening of a one bead-one compound combinatorial library for beta-actin identifies molecules active toward Ramos B-lymphoma cells. *Anal Biochem* 374: 112-120. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
22. Petri AL, Simonsen AH, Yip TT, Hogdall E, Fung ET, et al. (2009) Three new potential ovarian cancer biomarkers detected in human urine with equalizer bead technology. *Acta Obstet Gynecol Scand* 88: 18-26. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
23. Righetti PG, Boschetti E (2009) The art of observing rare protein species with peptide libraries. *Proteomics* 9: 1492-1510. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
24. Roux-Dalvai F, Gonzalez de Peredo A, Simo C, Guerrier L, Bouyssié D, et al. (2008) Extensive analysis of the cytoplasmic proteome of human erythrocytes using the peptide ligand library technology and advanced spectrometry. *Mol Cell Proteomics* 7: 2254-2269. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
25. Sennels L, Salek M, Lomas L, Boschetti E, Righetti PG, et al. (2007) Proteomic analysis of human blood serum using peptide library beads. *J Prot Res* 6: 4055-4062. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
26. Sihlbom C, Kanmert I, von Bahr H, Davidsson P (2007) Evaluation of the Combination of Bead Technology with SELDI-TOF-MS and 2-D DIGE for Detection of Plasma Proteins. *J Prot Res* 7: 4191-4198. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
27. Simo C, Bachi A, Cattaneo A, Guerrier L, Fortis F, et al. (2008) Performance of combinatorial peptide libraries in capturing the low-abundance proteome of red blood cells. I. Behavior of mono- to hexa-peptides. *Anal Chem* 80: 3547-3556. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
28. Thulasiraman V, Lin S, Gheorghiu L, Lathrop J, Lomas L, et al. (2005) Reduction of concentration difference of proteins from biological liquids using combinatorial ligands. *Electrophoresis* 26: 3561-3571. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
29. Turk BE, Cantley LC (2003) Peptide libraries: at the crossroads of proteomics and bioinformatics. *Curr Opin Chem Biol* 7: 84-90. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)
30. Yang H, Gurgel PV, Carbonell RG (2005) Hexamer peptide affinity resins that bind the Fc region of human immunoglobulin G. *J Pept Res* 66: 120-137. » [CrossRef](#) » [PubMed](#) » [Google Scholar](#)