Biomarkers discovery by peptide and protein profiling in biological fluids based on functionalized magnetic beads purification and mass spectrometry

Fulvio Magni[1]*, Yuri E.M. Van Der Burgt[2]*, Clizia Chinello[1], Veronica Mainini[1], Erica Gianazza[1,2], Valeria Squeo[1], André M. Deelder[2], Marzia Galli Kienle[1]

[1] Department of Experimental Medicine, University of Milano-Bicocca, Monza, Italy; [2]Department of Parasitology, Biomolecular Mass Spectrometry Unit, Leiden University Medical Center, Leiden, The Netherlands.

*equally contributing authors

Introduction

Proteomics aims for the full identification and quantification of all expressed proteins in any organism. This is however an extremely tedious task since one gene often accounts for multiple proteins due to gene splicing and processing of proteins, such as the addition of post-translational modifications. Moreover, the concentration range of occurring proteins varies more than a factor of one million. For these reasons, protein profiling was considered a promising technique in the early days of proteomics. Ideally a protein profile can be observed in one single measurement. In various clinical studies profiling methods have been successful in the detection of proteome variations as a consequence of an altered homeostasis. Proteins that are differentially expressed as a consequence of a disease are very useful in medical science as they can be used as new biomarkers for the diagnosis, prognosis and as possible therapeutic targets. In order to find such proteins or biomarkers two different kinds of biological material have been used: tissue samples and body fluids. Tissues are obtained from biopsies, from stable cell lines or cell cultures, or from subcellular fractions. Despite their large usage tissues suffer from several disadvantages. Tissue samples are difficult to obtain and are comprised of several different type of cells. Standardization of the methods to obtain subcellular fraction that affects its preparation and purity is a challenge not yet solved. The difference between a cell culture and its corresponding wild type present in the body limits the translation of information derived from the first to the latter.

On the contrary body fluids do not suffer from these limitations inherent to tissue samples. Fluids are very easily accessible with non- or very low-invasive methods at relatively low cost. They perfuse all the organs in the body carrying secreted protein from tissues. Therefore the protein profile of the biological fluids can reflect the status of the body. Among biological fluids serum, plasma and urine are the most analyzed samples but also cerebrospinal fluid (CSF), saliva, amniotic fluids have been used. Moreover classical methods to investigate the tissue proteome, aiming at biomarker discovery, are generally based on two-dimensional electrophoresis (2DE) and are not suitable for clinical chemistry lab requirements in which large sample cohorts have to be analyzed in a short time. This addresses another great potential of body fluids profiling : the analysis can be carried out high-throughput without sacrificing robustness and quality of the method. In fact 2DE is a laborious process that is difficult to automate. It still suffers from several technical limitations in terms of repeatability and reproducibility even though progress has been made using three different fluorescent labels that enables simultaneous migration of three samples on the same gel (e.g proteins extracted from control and disease, and the internal standard).Since the beginning of the 1990ties, when this new term (proteomics) was coined, a lot of progress has been made. Among them, several strategies to search these biomarkers in biological fluids have been developed in order to try to tackle some of the limitations of the current methods.

Nowadays, mass spectrometry (MS) is the method of choice for the analysis of proteins, and as a consequence the field is now often referred to as MS-based proteomics. Direct analysis of the biological fluids with mass spectrometry is a challenging

approach due to the sample complexity. To carry out a repeatable and robust mass spectrometric analysis of proteins in body fluids a suitable clean-up procedure is required in which salts and detergents are removed. The presence of salts can suppress the ionization in the mass spectrometer and chromatographic profiles may be influenced by from tailing due to co-elution of contaminants[1]. Therefore a pre-fractionation of the fluids is essential in order to increase the number of proteins that can be detected within a single MS-experiment, thus facilitating the discovery of new markers. Moreover, the fractionation of the biological fluids will also enrich low abundant proteins in fractions. These approaches lead to build the protein profile of the different biological fluids. Variations observed in patient profiles of body fluids compared to those of controls can be used to find the best pattern of signals that allows to discriminate two populations or to stratify the patients according to tumour stage or to the response to the therapy. One the major advantages of this strategy is that no pre-knowledge of the identity of signals selected for the cluster is needed to allow their use as biomarkers[2].

A specific agent to capture proteins enriches the sample and thus contributes to sensitivity enhancement. In general, protein separation techniques are based on different protein physical properties, such as size, isoelectric point, solubility and affinity. Materials known from different chromatographic platforms are coupled to the surface of a carrier in order to obtain peptides and proteins. One of the first approaches to pre-fractionate the body fluid proteome using an activated surface was the Surface-Enhanced Laser Desorption/Ionization (SELDI) technique. The SELDI technique for protein profiling is probably the most known and widely used approach in which biological fluids are applied directly to a target plate that is later introduced into a mass spectrometer. After removing unbound material to the modified surface of the SELDI chip, the molecular weight of the captured proteins on the target plate is determined using a time-of-flight (TOF) mass analyzer[3]. In this way the body fluid protein profile for the studied population is obtained. This technology is not free of criticism. In particular not very good reproducibility of the results due to drift, noise or the use of different lots of chips are reported. Moreover the direct identification of these markers cannot be carried out using the SELDI-TOF system. Their identity has to be determined with different analytical approaches. Promising alternatives to this technology are based on magnetic beads with a functionalized or activated surface or on miniaturized chromatographic systems that allow off-line fractionation of the proteome present in the fluids before MS analysis. The combination of magnetic bead purification and matrix-assisted laser desorption ionization (MALDI) TOF-MS has been shown a powerful alternative to the SELDI-platform: the active surface of magnetic beads is much larger, resulting in a higher binding capacity, and identification of captured peptides and protein is possible through the use of a more advanced TOF mass analyzer. Moreover, only a small part of the eluted peptide/proteins fractions are used for the protein profile and the remaining sample can be to use to identify markers with other MS-approaches (e.g. MALDI-TOF/TOF or LC-ESI-MS/MS) without the need of additional purification. This review is mainly focussed on the pre-fractionation based on magnetic beads and their applications.

Magnetic beads

In the middle of the 1990ties one of the first report about this approach is that by Girault S. et al. describing the use of magnetic beads with the external surface modified with streptavidin to capture biotinylated peptides[4]. Peptides were then eluted, mixed with matrix and analyzed by MALDI-TOF. Authors described this technology as a promising strategy to study peptide/proteins interactions. Later, an example of application of such approach based on magnetic beads but applied to biomarker discovery was reported[5]. Anti-PSA-IgG-biotin was bound to streptavidin coated beads and used to enrich and to purify PSA from serum. The authors reported that purification was satisfactory being the co-elution of non-specifically bound proteins highly reduced. Then activities were rather quiescent until the beginning of XX century when, driven from the SELDI technology success, they started to be used in clinical studies more frequently as indicated by the increment of the number of publications[3]. A review describing the clinical applications of mass spectrometry to biomarker discovery using body fluids was published in the 2003[6]. The strategy to search, find and validate new biomarkers is suggested to be generally divided in

three phases: in the first step, discovery of a cluster of signal able to differentiate patients from controls is carried out, in the second phase, the diagnostic capability of the cluster build in the first phase, is evaluated with a new cohort of patients and controls and, finally, in the third step unknown samples are used to test the cluster in blind-experiments.

Villanueva and co-workers pioneered in using functionalized magnetic beads, describing the serum peptide profiling strategy based on magnetic particles with the external surface modified with ligands of increasing carbon chain length (C1, C2, C3, C8 and C18) and with porous or non-porous surface[7]. Binding capacity of the different beads, expressed as number of peaks observed in MALDI-TOF spectrum, showed that C8 was the best performance extraction phase. The amounts of beads and serum and their volume ratio, sample pre-treatment before extraction, best eluent, and others aspects of the procedure were investigated in detail. Initially they obtained few peaks despite the extensive tuning of the mass spectrometer. However, after applying few shots (20-30) at rather high laser power, the acquired spectrum with the following 100 shots, at the normal operating laser power in the same position, resulted in many additional peaks. The addition of the n-octyl-glucoside in the samples before extraction has been shown to give the highest number of peaks either in the 0.8-4kDa or in the 4-15kDa mass range. Preliminary evaluation of the data reproducibility with C8 magnetic beads in serum showed a CV for inter-assay measurements evaluated in four different days of about 20%[8]. To correct for variability spiking body fluids with synthetic peptides was suggested. Recently, a study was reported aiming at finding the best strategy to perform a comparative study of serum from breast patients and healthy control persons with magnetic beads with various functionalizations: -hydrophobic interactions (C8), -weak cation exchange (WCX) and -immobilized metal-ion affinity (IMAC-Cu)[9]. The authors suggested to use WCX beads as a starting approach since these resulted in the highest number of signals. Later, Callesen et al[10] concluded differently after evaluating not only commercially available magnetic beads (C8 and IMAC-Cu) but also custom made microcolumns (C8, C18, IMAC-Cu and anion-exchange) used according to the manufacturer instructions or after manual modifications of the protocols. After a careful adjustment of the analytical procedures the highest number of reproducible peaks (more than 400) was obtained with IMAC-Cu purified fractions analysed by MALDI-TOF with 2,6-DHAP as matrix. A detailed study on the optimal conditions for serum pre-fractionation was described for new C18 magnetic beads[11]. A similar value of the reproducibility was observed analyzing either crude serum or n-octyl-glucoside treated with an average CV of about 10-12% and 24-30% for the intra-assay and inter-assay experiments, respectively.

Very soon it was clear to scientists involved in biomarker discovery using proteomics approaches that several factors affect the mass spectrum such as the origin and collection of biological material, the clean-up and MS-technique used. In this respect, in 2005 Baumann et al showed that anticoagulant, temperature and freeze-thaw cycles greatly influence the serum proteome profile obtained using magnetic beads (C3, C8, C18) following the manufacture instructions[12]. However, with a highly standardized protocol for blood collection, serum preparation, time of thawing, the intra and inter-assay variability, measured with ions in the mass range 1,000-10,000 and with low, medium and high intensity, were in the range of 3%-23% and of 6%-33% (CV%), respectively. Several studies have been carried out to investigate the effects of serum manipulation on the spectrum profile. In particular, two of these studies were based on fluids pre-fractionation with activated beads[13,14]. The first showed that storage temperature is the main factor, among those been tested responsible for mass spectra changes. It was concluded that samples kept at -80 °C are more stable than those at -20 °C[13]. The second study was carried out not only with C8 but also with IMAC-Cu beads. Different conditions were evaluated (tubes, clotting time and temperature) for serum and plasma sample preparation[14]. It was shown that procedures for serum preparation had the greatest impact on the profile compared to meals, handling procedures and storage conditions. Moreover data about the effects of different anti-coagulant on the protein profile of plasma are also described. Finally, according to their results both serum and plasma can be kept at 4 °C up to 24 hours without any significant modification of the protein profile within 1 to 25 kDa mass range. In another study factors such as temperature and time

for serum preparation as well as analytical humidity and temperature for the crystallization process were investigated with respect to their effect on the profiles[15]. The results showed that a large variations in protein profiles were due to leaving serum at room temperature for more than 30 min. In fact, mass spectra variation evaluated using WCX beads under very well controlled temperature (22 °C + 0.5 °C) and humidity (35 °C + 1 °C) for 5 consecutive days gave an overall CV of 0.27%. Statistical analysis of the results showed that 73% of the mass spectra variability derives from the humidity. Indeed when humidity is below that value (e.g. 15%) the crystals morphology was heavily affected thus resulting in unsuitable mass spectra. Recently, the influence of using different matrices (alpha-cyano-4-hydroxycinnamic acid, sinapinic acid, 2,5-dihydroxybenzoic acid and 2,5-dihydroxyacetophenone) on the mass spectra variability was evaluated for a standard peptides/proteins mixture and plasma samples[16]. The authors showed that matrix preparations had the greatest impact on the spectra compared to sample process before MS analysis using C8 magnetics beads. Earlier, Villanueva et al. reported an optimized protocol for biomarker discovery based on magnetic beads[17]. Serum preparation, storage and shipping conditions, MS tuning, mass spectra processing, and statistical analysis are described in detail showing possible sources of error. Later, from the same group, a study of the protein profile in serum of bladder, prostate and breast cancer patients using C8 coated magnetic beads was carried out[18]. The authors showed that in the serum protein there is a mixture of peptides originating both from ex-vivo coagulation and complement-degradation processes and from exoproteases released by cancer cells. Therefore serum can be a source not only of cancer-specific but also of cancer-type serum peptides that can be used as biomarkers after a validation study.

Alternative to beads

Besides functionalized magnetic beads other approaches can be used to fractionate the proteome before MS analysis[19]. Among them we present here two alternatives: ProteoMiner and micro-packed column.

The first alternative is the recently developed ProteoMiner pre-fractionation system based on hexapeptide combinatorial ligand libraries bound to the external surface of beads[20,21]. The rationale was to improve the protein separation capacity from fluids using not a single amino acid as early reported in the 1970 by Porath et al[22,23] but a small peptide. It was expected that the number of proteins bound to the beads was proportional to the length of the peptides. Surprisingly, it was found that this was true only up to a peptide length of six residues. Longer peptide chains did not result in an increment of captured proteins/peptides. Large numbers of copies of the same peptide are fixed on the each bead at a concentration of about 10-15 pmol[21]. A binding capacity of about 1-3 ng proteins per bead was estimated, obviously varying for different molecular masses[21].

The second alternative to magnetic beads consists of custom-made microcolumns with different stationary phases (C8, Cu-IMAC)[10]. Highly reproducible protein profiles were obtained for human serum samples using 2,6-dihydroxyacetophenone (2,6-DHP) as a matrix instead of the commonly applied alpha-cyano-4-hydroxycinnamic acid. From this study it was suggested to use magnetic beads with immobilized metal affinity chromatography (Cu) in combination with 2,6-DHP because these provide the most comprehensive list of candidate biomarkers. In another study, carried out by Tiss and co-workers, serum proteome fractionation procedures based on various stationary phases were compared[24]. Two different strategies based on magnetic beads (C8-C18-WCX) or on pipette tips pre-packed with a stationary solid phase (ZipTip C4-C18 and OMIX C4-C18) were evaluated. Their results using magnetic beads were different from those above described in terms of reproducibility of MS profile with obtained CV's of about 62%, 50% and 42 % for C8, C18 and WCX, respectively. It was suggested to use ZipTip because of better results with respect to the lower background noise and better signal-to-noise. The intra-assay and inter-assay CV for ZipTip C18 evaluated with a standard serum sample was of about 9% and 10%, respectively. An intra-run and inter-run CV of 9% and 7.5% was obtained analyzing 30 serum samples.

Examples of clinical applications

Among the body fluids such as serum, plasma, urine, amniotic fluids, saliva, tear, and cerebrospinal

fluid, the most widely used for biomarker discovery studies are those derived from blood. Serum is successfully used to search differences in the protein profile between healthy control individuals and patients with bladder[18,1], metastatic thyroid[25], hepatocellular[26], ovarian[27], gastric[28], breast [18], prostate[18], and renal cell carcinoma[29].

Although cancer is the most studied pathology other frontiers are open in this respect. Serum levels of two peptides were found altered in dystrophin-deficient mice[30]. One of these signals was identified as the N-terminal portion of Factor XIIIa and as such suggested as a candidate biomarker for muscular dystrophy[30]. Three peaks in the serum protein profile of 16 patients affected by autistic spectrum disorder (ASD) were identified after application of C8 magnetic beads[31]. This discovery opens a new technological approach in the diagnosis and possibly also in the treatment of this devastating psychiatric disorder. Unfortunately the identity of these peptides is still unknown.

Similarly, plasma can be a source of potential markers. Differences in peptide levels have been reported using plasma from oral cancer[32], nasopharyngeal tumour[33], and asthmatic patients[34]. An example of markers that have been found in other body fluids are three peptides in urine obtained from renal cell carcinoma patients[35]. One of these peptides was identified as a fragment of the Uromodulin precursor. Furthermore, protein profiles were obtained from amniotic fluids of 60 normal karyotypic and of 20 aneuploid women[36]. Eight peaks were chosen using a genetic algorithm based on their discriminating capability. The cluster display in the training phase, used to build the diagnostic cluster with 30 normal subjects, 100% and 72 % of sensitivity and specificity, respectively while in the cluster validation 100% of sensitivity and 96.6% of specificity. Finally, protein profiles of the tear fluid have been obtained using a pre-fractionation based on activated magnetic beads[37].

Acknowledgments

The present work has been supported by grants PRIN 2004, FIRB 2001 n. RBNE01HCKF, FIRB n. RBRN07BMCT and FAR 2004-2008 (ex-60% MURST) of the Italian Ministry of Research.

Keywords: ClinProt, biological fluids, proteomics, magnetic beads, mass spectrometry.

References

1) Tiss A, Smith C, Camuzeaux S et al. Serum peptide profiling using MALDI mass spectrometry. Practical proteomics 2007; **1**: 77-89.
2) Villar-Garea A, Griese M, Imhof A. Biomarker discovery from body fluids using mass spectrometry. Journal of Chromatography B 2007; **849**: 105-14.
3) Chen L, Fatima S, Peng J et al. SELDI protein chip technology for the detection of serum biomarkers for liver disease. Protein Pept Lett 2009; **16**(5): 467-72.
4) Girault S, Chassaing G, Blais JC et al. Coupling of MALDI-TOF mass analysis to the separation of biotinylated peptides by magnetic streptavidin beads. Anal Chem 1996; **68**(13): 2122-6.
5) Jochen P, Unverzagt C, Lenz H et al. Purification of prostate-specific antigen from human serum by indirect immunosorption and elution with a hapten. Analytical Biochemistry 1999; **273**: 98-104.
6) Pusch W, Flocco MT, Leung SM et al. Mass spectrometry-based clinical proteomics. Pharmacogenomics 2003; **4**(4): 463-76.
7) Villanueva J, Philip J, Entenberg D et al. Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. Anal Chem 2004; **76**: 1560-70.
8) De Noo ME, Tollenaar RAEM, Özalp A et al. Reliability of human serum protein profiles generated with C8 magnetic beads assisted Maldi-TOF mass spectrometry. Anal Chem 2005; **77**: 7232-41.
9) Qiu F, Liu HY, Zhang XJ et al. Optimization of magnetic beads for MALDI-TOF-MS analysis. Frontiers in Bioscience 2009; **14**: 3712-23.
10) Callesen AK, DePont Christensen R, Madsen JS et al. Reproducibility of serum protein profiling by systematic assessment using solid-phase extraction and matrix-assisted laser desorption/ionization mass spectrometry. Rapid communications in mass spectrometry 2008; **22**: 291-300.
11) Jimenez CR, El Filali Z, Knol JC et al. Automated serum peptide profiling using novel magnetic C18 beads off-line coupled to MALDI-TOF-MS. Proteomics Clin Appl 2007; **1**: 1-7.
12) Baumann S, Ceglarek U, Fiedler GM et al. Standardized approach to proteome profiling of human serum based on magnetic bead separation and Matrix-Assisted Laser Desorption/Ionization Time-of-Flight mass spectrometry. Clinical Chemistry 2005; **51**(6): 973-80.
13) West-Nielsen M, Hogdall EV, Marchiori E at al. Sample handling for mass spectrometric proteomic investigation of human Sera. Anal Chem 2005; **77**: 5114-23.
14) Hsieh SY, Chen RK, Pan YH et al. Systematical evaluation of the effects of sample collection procedures on low-molecular-weight serum/plasma proteome profiling. Proteomics 2006; **6**: 3189-3198.
15) West-Norager M, Kelstrup CD, Schou C et al. Unravelling in vitro variables of major importance fort he outcome of mass spectrometry-based serum proteomics. Journal of chromatography B 2007; **847**: 30-7.

16) Penno MAS, Ernst M, Hoffmann P. Optimal preparation methods for automated matrix-assisted laser desorption/ ionization time-of-flight mass spectrometry profiling of low molecular weight proteins and peptides. Rapid Commun. Mass Spectrom 2009; **23**: 2656-2662.
17) Villanueva J, Philip J, Chaparro CA et al. Correcting common errors in identifying cancer-specific serum peptide signatures. Journal of Proteome Research 2005; **4**: 1060-72.
18) Villanueva J, Shaffer DR, Philip J et al. Differential exoprotease activities confer tumor-specific serum peptidome patterns. The journal of clinical investigation 2006; **116**: 271-84.
19) Righetti PG, Castagna A, Antonioli P et al. Prefractionation techniques in proteome analysis: the mining tools of the third millenium. Electrophoresis 2005; **26**: 297-319.
20) Boschetti E, Righetti PG. The proteoMiner in the proteomics arena: a non-depleting tool for discovering low-abundance species. Journal of Proteomics 2008; **71**: 255-64.
21) Boschetti E, Righetti PG. Hexapeptide combinatorial ligand libraries: the march of the detection of the low-abundance proteome continues. BioTechniques 2008; **44:** 663-65.
22) Porath J, Fryklund L. Cromatography of proteins in dipolar ion adsorbants. Nature 1970; **226**: 1169-70.
23) Porath J, Fornstedt N. Group fractionation of plasma proteins on dipolar ion-exchangers. J Chromatogr 1970; **51**: 479-89.
24) Schwanborn K, Krieg RC, Grosse J, et al. Serum proteomic profiling in patient with bladder cancer. Eur Urol 2009; **56** (6): 989-997.
25) Villanueva J, Martorella AJ, Lawlor K, et al. Serum peptidome patterns that distinguish metastatic tyroid carcinoma from cancer-free controls are unbiased by gender and age. Mol Cell Proteomics 2006; **5**:1840-52.
26) Orvisky E, Drake SK, Martin BM et al. Enrichment of low molecular weight fraction of serum for MS analysis of peptides associated with hepatocellular carcinoma. Proteomics 2006; **6**: 2895-2902.
27) Ziganshin RH, Alexeev DG, Arapidi GP et al. Serum proteome profiling for diagnostics of ovarian cancer using ClinProt magnetic technique and MALDI-TOF mass spectrometry. Biomedical Chemistry 2008; **2** (4): 335-342.
28) Ebert MPA, Niemeyer D, Deininger SO et al. Identification and confirmation of increased fibrinopeptide a serum protein levels in gastric cancer sera by magnetic bead assisted MALDI-TOF mass-spectrometry. Journal of Proteome Research 2006; **5**: 2152-8.
29) Chinello C, Gianazza E, Zoppis I, et al. Serum Biomarkers of Renal Cell Carcinoma Assessed Using a Protein Profiling Approach Based on ClinProt Technique. Urology 2009; DOI:10.1016/j.urology.2009.09.050.
30) Alagaratnam S, Mertens BJA, Dalebout JC, et al. Serum protein profiling in mice: identification of factor XIIa as a potential biomarker for muscular dystrophy. Proteomics 2008; **8**: 1552-1563.
31) Taurines R, Dudley E, Conner AC, et al. Serum protein profiling and proteomics in autistic spectrum disorder using magnetic bead-assisted mass spectrometry. Eur Arch Psychiatry Clin Neurosci; 2009; DOI:10.1007/ s00406-009-0066-5.
32) Cheng AJ, Chen LC, Chien KY, et al. Oral cancer plasma tumor marker identified with bad-based affinity-fractionated proteomic technology. Clinical Chemistry 2005; **51**: 2236-44.
33) Chang JT, Chen LC, Wei SY, et al. Increase diagnostic efficacy by combined use of fingerprint markers in mass spectrometry-plasma peptidomes from nasopharyngeal cancer patients for example. Clin Biochem 2006; **12**: 1144-51.
34) Zhang X, Leung SM, Morris CR, et al. Evaluation of a novel, integrated approach using functionalized magnetic beads, bench-top MALDI-TOF-MS with prestructurated sample supports, and pattern recognition software for profiling potential biomarkers in human plasma. J Biomolecular Techniques 2004; **15:** 176-175.
35) Bosso N, Chinello C, Picozzi SCM. Human urine biomarkers of Renal Cell Carcinoma evaluated by ClinProt Proteomics Clinical Application 2008; **2**: 1036-1046.
36) Wang TH, Chang YL, Peng HH, et al. Rapid detection of fetal aneuploidy using proteomics approaches on amniotic fluid supernatant. Prenat Diagn 2005; **25**: 559-566.
37) Sekiyama E, Matsuyama Y, Higo D et al. Applying magnetic bead separation/MALDI-TOF mass spectrometry to human tear fluid proteome analysis. J Protomics Bioinform 2008; **1**(7): 368-373.

Correspondence : Fulvio Magni
Department of Experimental Medicine
University of Milano-Bicocca
Via Cadore, 48
20052 Monza, Italy
E-mail: Fulvio.Magni@unimib.it