Published in final edited form as:

Expert Rev Proteomics. 2018 April ; 15(4): 353–366. doi:10.1080/14789450.2018.1450631.

Affinity Enrichment for MS: Improving the yield of low abundance biomarkers

Brianna Kim¹, Robyn Araujo², Marissa Howard³, Ruben Magni³, Lance Liotta³, and Alessandra Luchini³

¹Ceres Nanosciences, Manassas, Virginia, USA

²School of Mathematical Sciences, Queensland University of Technology, Brisbane, QLD 4000, Australia

³Center for Applied Proteomics and Molecular Medicine, George Mason University, 10920 George Mason Circle, Manassas, Virginia 20110, USA

Abstract

Introduction—Mass spectrometry (MS) is the premier tool for discovering novel disease-associated protein biomarkers. Unfortunately, when applied to complex body fluid samples, MS has poor sensitivity for the detection of low abundance biomarkers ($\ll 10$ ng/mL), derived directly from the diseased tissue cells or pathogens.

Areas covered—Herein we discuss the strengths and drawbacks of technologies used to concentrate low abundance analytes in body fluids, with the aim to improve the effective sensitivity for MS discovery. Solvent removal by dry-down or dialysis, and immune-depletion of high abundance serum or plasma proteins, is shown to have disadvantages compared to positive selection of the candidate biomarkers by affinity enrichment. A theoretical analysis of affinity enrichment reveals that the yield for low abundance biomarkers is a direct function of the binding affinity (Association/Dissociation rates) used for biomarker capture. In addition, a high affinity capture pre processing step can effectively dissociate the candidate biomarker from partitioning with high abundance proteins such as albumin.

Expert Commentary—Properly designed high affinity capture materials can enrich the yield of low abundance (0.1–10 picograms/mL) candidate biomarkers for MS detection. Affinity capture and concentration, as an upfront step in sample preparation for MS, combined with MS advances in software and hardware that improve the resolution of the chromatographic separation can yield a transformative new class of low abundance biomarkers predicting disease risk or disease latency.

1. Introduction

Low abundance is the greatest roadblock to the discovery of protein body fluid biomarkers for the detection of early stage infectious diseases, cancer, and neurodegenerative disorders.

A critical need within the biochemical and biomedical research sector is the identification of low abundance biomarkers that are predictive of early stage cancer, early stage neurologic disorders, infectious disease, or correlate with therapeutic outcome or toxicity (1). While the interest in the potential and value of biomarkers has never been greater, the research

investment in biomarker discovery and clinical validation has yielded a very poor return to date (1, 2). This poor return is due in large part to the low abundance of early disease biomarkers that exist at a concentration below the detection limit of biomarker discovery platforms. Protein biomarker discovery and quantitation by mass spectrometry (MS(1)) and multiple/parallel reaction monitoring (MRM(3)) are powerful approaches (1, 2) but are severely limited in their practical application for complex clinical samples because of their poor effective sensitivity (lower limit of detection) (4) for complex body fluids. The analyte detection sensitivity for MS or MRM applied directly to a complex body fluid is typically greater than 50 ng per mL (6).

In contrast, the vast majority of diagnostic analytes measured in the clinical laboratory by immunoassay platforms fall in the range between 5 pg/mL and 10 ng/mL (7). Thus, the most important protein biomarkers, particularly those derived from early stage disease (8), are invisible to conventional MS or MRM (9). MS and MRM lack practical sensitivity because of technical and physiological constraints. Proteins and peptides are masked by a billion-fold excess quantities of resident proteins such as immunoglobulin and albumin. The MS input sample is strictly limited in the maximum total protein (< 5 ug) content, a value lower than the serum or plasma protein content in the microliter volume of the MS input. Consequently, increasing the sensitivity is not simply a matter of concentrating the sample (for example by drying the sample to remove the water), because this will overwhelm the total protein capacity introduced into the MS. A further barrier to biomarker discovery is the lability and perishability of candidate biomarkers ex vivo following clinical sample collection. Diagnostic proteins and peptides in body fluids are subject to rapid enzymatic degradation, or aggregation and precipitation, following collection (10).

A limitation of many cancer markers used in the clinics is the lack of specificity. As in the case of PSA, the marker can be shed by the healthy prostate tissue and by non-malignant disorders. A more sensitive approach to discovery biomarkers would permit the identification of markers that are exquisitely specific to the tumor tissue (5). Tissue homogenates and cell lysates, presenting a much lower dynamic range in their protein content than bodily fluids, are samples more favorable to mass spectrometry analysis (3, 5, 11). Mass spectrometry techniques have significantly contributed to the detection and quantification of biomarkers from tissue biopsy samples (when available) in order to confirm the tumor origin of the biomarkers (3, 5). MS applied to tumor biopsies, perhaps in combination with cell enrichment techniques such as microdissection, are critically important additions to the workflow for biomarker scientists aiming to validate that the biomarker they find in a body fluid is actually derived from the diseased lesion itself.

1.1 Physiologic roadblocks causing low abundance

Circulatory volume dilution and diffusion barriers are responsible for the expected low abundance of cancer biomarkers that have the most clinical importance (Figure 1). An important example of this constraint is the quest for biomarkers in the blood useful for early cancer detection. An overarching goal of the cancer biomarker field is to detect early stage pre-metastatic cancer. In theory, surgery alone can effectively cure the cancer if it is removed prior to metastasis. After metastasis has occurred the prognosis declines rapidly. Stage zero

(pre invasive) and stage 1 solid tumors are frequently less than 0.5 cm and exist in an admixture with dense host stroma and immune cells. Biomarkers elaborated by the small volume of neoplastic cells in the tumor lesion microenvironment must diffuse across multiple cellular and extracellular barriers before they enter the venous drainage, and are then immediately diluted into a large blood volume (Figure 1). Circulating biomarkers can be cleared in the liver or secreted in the kidney, further reducing their concentration in the plasma. Lutz et al considered these physiologic challenges when they developed a mathematical prediction of the expected concentration of tumor cell biomarkers in the blood as a function of tumor size (8). They constructed a linear one compartment model representing homogeneous well mixed plasma. The tumor biomarker is defined as a protein that is secreted by tumor cells into the extracellular fluid compartment. It is assumed that a certain percentage of the secreted protein will enter the intravascular space (plasma) at a continuous rate. In the plasma, the protein of interest has a distinct half-life caused by degradation, hepatic sequestration, or glomerular filtration.

Lutz considered two potential pathophysiologic conditions: (1) the tumor biomarker is secreted only by the tumor cells, or 2) both the tumor and the host cells of the lesion produce the biomarker (8). The author's calculations for the biomarker PSA assumed that the analytical detection cut-off for the clinical immunoassay was 50 pg/mL. Following this assumption, the smallest detectable tumor ranged in size from 27mm³ to 3.45x10⁵ mm³. The analytical PSA test assumed for the authors' cut-off calculation is a standard of care clinical immunoassay. If they had chosen a direct mass spectrometer assay, the concentration cut off would have been at least 100 times higher, since clinical immunoassays are much more sensitive compared to mass spectrometry for the direct measurement of proteins in complex body fluids. Brown et al conducted a similar calculation of cancer blood biomarker concentrations relative to the size of the primary tumor (12). These authors concluded: "We estimated that the tumors we would need to detect to achieve even 50% sensitivity are more than 200 times smaller than the clinically apparent serous cancers typically used to evaluate performance of candidate biomarkers; none of the biomarker assays reported to date comes close to the required level of performance."

Two general conclusions emerge from the analysis by Lutz et al (8) and Brown and Palmer (12). The first is that the direct application of mass spectrometry to body fluids alone is not a method of choice to discover cancer biomarkers when the object is to detect low abundance biomarkers derived from small pre metastatic lesions or early stage disease. The second is that new technology is required to increase the sensitivity of biomarker detection by MS by at least 200 fold, if the goal is to discover biomarkers derived from a pre-metastatic cancer lesion.

1.2 Enrichment and concentration methods enhance the effective sensitivity of mass spectrometry for biomarker discovery

Faced with a low abundance protein or metabolite analyte, the options for enhancing the sensitivity of detection are generally limited to two approaches. The first approach is to increase the intrinsic sensitivity of the detection method. A variety of approaches are available for increasing the analytical sensitivity, including changing the detection device, or

the label, or employing enzymatic amplification of the signal (13, 14). The drawbacks are that such optimization may be specific to the analyte, that the background will increase along with the signal, or that the resolution (in the case of MS) will be reduced. The second approach is to enrich or concentrate the analytes of interest in the sample so that the input to the detection system contains a higher number of analyte molecules (15, 16). The advantage of this second approach is that the sensitivity improves without any changes in the detection or measurement system. The present review focuses on this second approach, analyte enrichment, as a means to elevate the effective sensitivity of the MS.

2. Non - affinity based concentration methods

Concentrating the total protein content of the dilute starting material is often used as an early step in protein extraction, purification, and analysis. Several methods exist for concentrating the protein solute, and de-salting the sample, by removal of the fluid solvent along with small molecules (Figure 2).

2.1 Membrane filtration concentrators

The dilute protein sample is loaded into the concentrator cartridge which comprises two compartments separated by a membrane filter with a size sieving function (17). When a centrifugal force is applied, the fluid solvent passes through the membrane into the second chamber, while the protein solute is held back and remains in the first chamber. The main drawback of this method is non-specific binding of the biomarkers of interest to the membrane, so that they are not accessible for analysis. Further drawbacks include membrane clogging, and small processing volumes that are constrained by the centrifuge tube size (18).

2.2 Dialysis

For large volumes concentration can be achieved by placing the input sample in a dialysis bag that is submerged in a large volume of water. Alternatively, dialysis can be conducted in a dialysis device or chamber. The dialysis membrane retains the proteins inside the dialysis bag or device and the sample is concentrated as the water is pulled across the dialysis membrane and retained by a water absorbing gel or polymer. An ideal water absorbing polymer can be polyethylene glycol. The drawbacks of dialysis are the length of time needed for dialysis of big volumes and the large output volume resulting in a poor concentration factor (19).

2.3 Precipitation, dry-down, or Salting Out to concentrate a sample

Precipitation or “salting out” of proteins is another common method for concentrating proteins from large volumes. Ammonium sulfate is a common reagent used for this purpose. Increasing the concentrations of ammonium sulfate elevates the ionic strength of the solution. The increase in ionic strength reduces the repulsion of same-charges of protein molecules and reduces the solvation shell around proteins. Once these forces are sufficiently reduced the proteins will interact, aggregate, and precipitate. The precipitated proteins can be collected by centrifugation and be resolubilized in a smaller volume to achieve a concentration effect. A drawback of salting out is functional disruption of biologic activity when the protein is in an insoluble state. Moreover, the proteins may not return to solution

phase upon resolubilization. Commercial kits exist to conduct salting-out using proprietary reagents that protect the biological activity of the proteins, and improve resolubilization (20).

2.4 Molecular size sieving chromatography or gel electrophoresis for concentrating a sample (21)

Size sieving chromatography can be accomplished using small spin columns, or large sieving columns. The larger proteins emerge at the front, and depending on the volume of the elution fraction, can be partially concentrated. A further option for non-affinity protein concentration is to use some form of gel electrophoresis to drive the sample mixture through a porous acrylamide gel matrix. In this case the smaller proteins and molecules travel deeper into, or through, the gel. This is usually done with a two chamber apparatus separated by a gel with a defined pore size. An electrode is introduced into the top and the bottom chamber and an electrical charge is used to drive the small molecules or small proteins from the top to the bottom chamber. This can result in concentration of large species in the top chamber to achieve a size sieving effect. The drawbacks are the high voltage/current requirements, the time delay for separation and migration, and the difficulty in accessing or eluting the proteins that enter the gel plug. Proteins separated into the gel can be subject to in gel digestion for MS with a detection sensitivity limit usually in the range of 20 ng for a single protein in a gel band. For a complex mixture of proteins concentrated in a gel, the sensitivity is much poorer, typically in the μg range (21).

3. Affinity Enrichment Methods

Affinity enrichment has a greater record of success for low abundance biomarker discovery compared to simple concentration by dry-down, or salting-out. A variety of affinity capture based methods have been developed to partition the serum proteins so that the biomarker candidates are separated from the unwanted contaminating proteins (22–28).

Affinity methods fall into two major categories (Figure 3), negative and positive selection:

- a. Affinity immunodepletion removal of the unwanted species such as albumin or immunoglobulins (22, 29), or
- b. Affinity capture and enrichment of the biomarker of interest. Affinity baits range from antibodies (23, 24, 30), protein or nucleic acid ligands (26), peptide libraries (28), to chemical baits (16, 31) (e.g. dyes or metals).

Immunodepletion is a widely used method of sample preprocessing for mass spectrometry based biomarker discovery studies (29, 32).

Bandow (15) compared immunodepletion of high abundance proteins (Seppro IgY14 column followed by the Seppro IgY-SuperMix system) to plasma protein fractionization by ProteoMiner, an affinity capture system using bead bound peptide hexamers. The output of these enrichment or depletion strategies was studied by 2-D gels and MS identification. Bandow found striking differences in the 2-D and 1-D gel pattern of protein species that emerged from the output of these methods. Nevertheless, MS sequencing of the detectable

protein spots in the different plasma fractions were exclusively high-abundant proteins normally present in plasma at concentrations between 1 ug and 40 mg/mL, with a high preponderance of common plasma proteins involved in wound healing and coagulation, such as fibrinogen, apolipoprotein, vitamin D binding protein, transferrin, and prothrombin.

Bandow's results parallel the experiences of many biomarker discovery labs who find that immunodepletion or peptide library fractionization does not adequately reveal low abundance proteins expected to be derived from diseased tissues. In fact such discovery workflows for MS analysis of plasma too often do not even reveal the majority of common moderate to low abundance diagnostic plasma proteins routinely measured with high precision by clinical grade immunoassay instruments (15).

3.1 Drawbacks of Immune depletion

Although widely used, column immunodepletion methods have two major weaknesses, beyond the high cost of the columns. The first is that low abundance analytes of interest can form a complex with the high abundance proteins such as albumin and immunoglobulin and are therefore removed and depleted (Figure 3B) (25). This is likely even if the affinity of the candidate biomarker for albumin is low, since the stoichiometry favors association between a low abundance protein and albumin, for example, that is in a billion-fold excess (see Theoretical Analysis below). A large number of studies characterizing the non-specific removal of low abundance proteins via immunodepletion have been published (33–38). Tu and colleagues (39) analyzed the high abundance and low abundance fraction after immunodepletion via LC-MS/MS. They showed that immunodepletion has an off-target retention of 20 proteins and permits the discovery of 36 additional proteins in comparison to unfractionated plasma on a total 216 of proteins. They concluded that the approach is not efficient for discovery of low abundance biomarkers in plasma. Ahn and colleagues also investigated the utility of immunodepletion of the 14 most abundant proteins from human plasma for the detection of 27 cytokines, chemokines and growth factors in 15 plasma samples. They concluded that immunodepletion is counterproductive because 21 of 27 cytokines were detected in the bound fraction and 22 cytokines had higher levels in the neat plasma with respect to the flow through (40). The second weakness of immunodepletion columns applied to clinical samples is the danger, and prevalence, of carry-over from successive samples run over the immune-depletion columns (Figure 3B). Carry over from the previous column runs can be insidious because changing flow pockets in the column can dislodge carryover which appears in the output of the elution many samples after the original contamination. Too often carry over is not monitored or considered. The danger of immunodepletion column carry-over is not necessarily reduced by harsher column washing between runs, because this can damage the column antibodies and reduce the life of the column, while having no significant improvement in carry-over. The danger of carryover severely reduces the compliance for CAP/CLIA certification, and can be lead to sample bias depending on the order of the unknown sample in the column runs.

4. Classes of affinity chemistries

Two general classes of affinity molecules have been successfully applied to capture and concentrate low abundance biomarker proteins for MS. The first class is protein or nucleic acid binding molecules such as antibodies, aptamers, enzyme substrates, or natural ligands (23, 24, 26) that have a specific known binding partner target. The advantage is that these molecules, when used in a flow through column format can effectively sequester and concentrate their cognate biomarker with relatively high efficiency and yield. As noted above, the efficiency, yield, and low abundance reach of the method will be a function of the affinity of the binding molecule. The drawback of these molecules is that they cannot be used to discover new classes of low abundance biomarkers that are not their natural binding partner, they can only be used for protein analytes that are known targets in advance. A further drawback emerges if this class of affinity ligands is formatted in a reusable column. Reusable affinity columns containing molecules susceptible to denaturation, will, by definition, decline in performance after each cycle of stripping. Moreover, as noted above, carryover is a major problem that can lead to severe bias for MS biomarker detection. The carryover problem can be solved by disposable single-use columns, if they can be of low cost and high batch to batch uniformity.

The second class of affinity capture reagents are molecules that bind general classes of proteins, but do not target a single specific analyte (16, 31). The affinity material can be a dye, a metal, a drug, or molecule that recognizes an affinity tag (e.g. His-Tag, relevant for cell line and animal studies (41)) inserted on the molecules of interest. Examples include organic dyes that bind multiple classes of proteins (Figure 5), metal chelating, or anti-phosphoTyr columns that recognize general phosphorylation sites (31), drugs with unknown binding partners, or peptide libraries (28) that are comprised of random variations in sequence. A lead example of this approach is the solid phase extraction method of Ueda (41,42) which is designed label, isolate, and characterize, cell surface displayed glycosylated proteins. The method is a dynamic covalent chemistry that reacts hydrazide chemistry with cell surface glycans. The impact and usefulness of the technology is documented by published studies that show a correlation of glycan structure alterations and glycan abundance in different disease states (41,42).

Although present in any fractionation method, the risk of inadvertent loss of potential minor fractions of low abundance biomarkers of unknown sequence is reduced when using an affinity enrichment strategy that employs promiscuous affinity baits. Affinity capture enrichment methods that recognize broad classes of proteins have advantages in terms of total coverage of the biomarker repertoire for discovery of previously unknown proteins or peptides in a sample. Non ligand-specific binding mechanisms that can be exploited for use in an affinity matrix chemistry include ionic, hydrophobic, aromatic, and sterically active, binding sites. A combination of multivalent mechanisms ensures an unselective capture that is usually complemented by high binding capacity. Examples of this non selective chemistries include small molecule chemical baits such as dyes. Cibacron blue covalently bound to agarose beads is a widely used method for protein purification (44). Disperse yellow 9, disperse orange 3, acid black 48 and a high number of other dyes covalently incorporated in buoyant hydrogel nanoparticles were demonstrated to interact individually

with hundreds of different proteins in serum or plasma (45). Importantly, the combination of different chemical baits yielded an increased number of interacting proteins by combining their respective protein repertoire (45).

The defining characteristic of this class of affinity capture is that they can be used to find new classes of proteins in a mixture that are unknown ahead of time. For this class, the specificity and the identity of the molecules harvested is fully determined by the downstream analytical device. In the present case this is the MS. Once again, the low abundance reach of this method is largely determined by the affinity of the binding. One way to achieve high affinity binding is to genetically or chemically tag the population of proteins in the sample ahead of time. Examples are HIS tagging or biotinylation (31). The affinity capture ligands for these modifications are of high affinity (e.g. cobalt-based HIS-Tag binding chemistry) and are specific for the introduced tag. The drawback is the requirement for genetically modifying the cells or chemically labeling all of the unknown proteins in a body fluid sample as a first step in the discovery process.

We can see that chemical or genetic tagging for MS discovery of low abundance proteins is not ideal when the goal is to discover novel disease markers of clinical value. A better approach would be based on a chemical that generally binds all classes of NATIVE UNMODIFIED proteins and peptides with very high affinity. Fortunately, many types of low cost organic dyes have been discovered which bind proteins (and other macromolecules) with very high affinity (46–50) (Figure 5). The dyes can be packaged into porous colloidal hydrogel particles that float buoyant in solution for optimal exchange with the solvent (45, 51–53). Commercially available Nanotrap® particles employ this principle, and have been successfully used to identify MS low abundance molecules 100 fold lower in abundance compared to conventional immunoaffinity methods (Figure 6 and (45, 48–51, 53–56)).

4.1 Affinity Capture Nanoparticles

The authors have created the harvesting hydrogel Nanotrap® particles and have used them for a variety of biomarker discovery applications (45, 48–51, 53–56). The workflow is depicted in Figure 5. Hydrogel dye bait nanoparticles (Nanotrap® particles) are introduced into the body fluid sample where they harvest general classes of molecules, antigens, proteins, exosomes, or microorganisms (54, 57, 58). After incubation in the fluid sample the Nanotrap® particles are separated by centrifugation, magnetization, or filtration, and their cargo is eluted into a small volume. The output is a small volume containing a concentrated set of biomolecules harvested from the entire volume of the starting volume. De-salting and preservation of the captures molecules is also achieved. Following the theoretical principles of presented above, the Nanotrap® particles containing high affinity dye baits can effectively sequester proteins of interest away from albumin or immunoglobulins (45). The porosity of the hydrogel particles is tuned to exclude large unwanted proteins, while permitting the entry of low molecular weight molecules. A further advantage is that the Nanotrap® particles are single-use, disposable, and low cost. Thus the issue of carryover is solved at this sample preparation step. A weakness is that the Nanotrap® particles cannot capture large proteins greater than 60kDa if they are tuned to exclude albumin (Figure 5). This weakness may in reality be a strength when the goal is to discover serum or urine proteins that enter body

fluids from the tissues. Normally the blood vessel basement membrane is not permeable to tissue proteins larger than albumin. Among other biologic purposes this serves to keep albumin within the plasma and to maintain an osmotic gradient in the correct direction. Proteins emanating from tumor or diseased tissue that diffuse into the blood or lymph are likely to be smaller than albumin or may be fragments of proteins that are smaller than 60kDa. Consequently, biomarkers derived from tissue cells that penetrate basement membrane and cell junction diffusion barriers to enter the blood stream are likely to be small enough to be captured by porous hydrogel affinity particles that exclude albumin.

4.2 Theoretical analysis of affinity capture

A theoretical analysis of affinity capture can provide important insights into the parameters required to isolate low abundance biomarkers existing in a complex mixture of plasma, serum, or urine proteins. Our analysis assumes that particles composed of hydrogel material and functionalized with affinity capture binding moieties are introduced the fluid sample containing a complex mixture of high abundance proteins. The particles are porous and constituted by cross-linked polymer chains; 95% of the internal volume is constituted by solvent. As shown below this analysis reveals the feasibility of using custom designed affinity capture particles to harvest and sequester low molecular weight or low abundance biomarkers, even though the biomarker is complexed with high abundance proteins (e.g. albumin).

At $t=0$, a harvesting capture molecule is introduced to the system, at which time an equilibrium is assumed to exist between the free (uncomplexed) biomarker, b , the 'natural' capture carrier protein, C_N (e.g. albumin or IgG), and the biomarker-carrier protein complex, bC_N . In the absence of an affinity harvesting molecule, the reaction scheme is as depicted below,

where k_N and k_{-N} are the forward and reverse rate constants for the reaction, respectively. Here it is assumed that the carrier protein (albumin) exists in such vast excess over the biomarker that the free carrier protein concentration, $[C_N]$, may be considered a constant and may be absorbed into the forward rate constant. Thus, the equilibrium ratio will be given by where K_N is the affinity constant for the biomarker and its natural carrier protein. Note that $k_N \gg k_{-N}$ or $K_N \gg 1$ since the relatively large free carrier protein concentration typically amplifies the forward reaction rate by several orders of magnitude[15]. When the harvesting affinity capture particle is introduced, on the other hand, the reaction scheme is as depicted in Scheme 1, where k_H and k_{-H} are the forward and reverse rate constants, respectively, for the reaction between the biomarker and the harvesting capture particle, C_H , producing the complex bC_H . Of note, the dye bait concentration in the inner volume of the capturing particles is very high, in the high millimolar range. This value can be derived from published data (55, 59). One sphere of radius $0.25 \mu\text{m}$ occupies a volume of $0.065 \mu\text{m}^3$. One milliliter of suspension contains 230 billion particles and, therefore, $2.87\text{E-}7$ mol of dye. The total

volume occupied by the particles is therefore $15 \mu\text{m}^3$. The molar concentration of the dye in the internal volume of the particles is $2.87\text{E-}7 \text{ mol}/1.5\text{E-}5 \text{ L} = 0.019 \text{ mol/L} = 0.019\text{M}$. In comparison, the concentration of albumin in serum is $7.5\text{E-}4 \text{ M}$ two orders of magnitude lower. In Scheme 2, the reaction scheme includes a diffusion parameter depicting the transit of the biomarker b from the outside to the inside of the nanoparticles, where b^* is the biomarker that entered the harvesting particle to reach the capturing affinity bait. For this analysis, however, since the particles are buoyant, open cage, and in a colloidal suspension, we assume that the $b^* = b$ and that the diffusion effects are negligible. Note that the rate constants k_H and k_{-H} account for the rate of biomarker transit through the particle pores as well as the affinity of the biomarker for the encapsulated bait molecule.

Here again, the harvesting particles are present in such vast excess over the low abundance biomarker that their (very large) concentration may be absorbed into the forward rate constant, k_H , so that $k_H \gg k_{-H}$. The introduced harvesting affinity particles have a much greater innate affinity for the small biomarkers, in addition to existing at a higher concentration compared to the natural carrier proteins, so that $k_H \gg k_N$. For this reason, an association between the biomarker and the harvesting particle is markedly preferred over the complex with the natural carrier protein. Since the harvesting particle concentration itself contributes to the forward rate constant, k_H , this bias towards the harvesting particle will be enhanced further as the concentration of harvesting particles in the mixture is increased. The temporal variations of the free and complexed forms of the biomarker are now described by the suite of Equations (2) through (4) below, subject to the specified initial conditions and parameter set.

Parameters and Initial Conditions for Equations (2)-(4)

$$\frac{d[b]}{dt} = -k_N[b] + k_{-N}[bC_N] - k_{-H}[bC_H], \quad (2)$$

$$\frac{d[bC_N]}{dt} = k_N[b] - k_{-N}[bC_N], \quad (3)$$

$$\frac{d[bC_H]}{dt} = k_H[b] - k_{-H}[bC_H] \quad (4)$$

$$\begin{aligned} k_N &= 0.1; & [bC_N](0) &= 1; \\ k_H &= 1; & [b](0) &= \frac{[bC_N](0)}{K_N} = 0.1; \\ k_{-N} &= k_{-H} = 0.01; & [bC_H](0) &= 0; \\ (K_N &= 10). \end{aligned}$$

Note that this parameter set has been chosen so that $k_N \gg k_{-N}$ ($K_N \gg 1$), $k_H \gg k_{-H}$ ($K_H \gg 1$) and $k_H \gg k_N$. In other words, biomarkers markedly prefer complexed forms over free forms, and the formation of a complex with the harvesting affinity molecule is markedly preferred over a complex with the natural carrier protein. The solutions to these equations are depicted in Figure 4, for a harvesting particle with a biomarker affinity that is ten times greater than that of the natural carrier protein, i.e., $k_H/k_N=10$.

The mechanism underlying these model solutions may be explained as follows: When the harvesting particles are first introduced to the mixture, the bound and free-phase biomarkers are initially in the ratio determined by Equation (1), i.e., $[bC_N]=K_N[b]$. The harvesting particles immediately begin to sequester the free-phase biomarker, attempting to establish the corresponding ratio $[bC_H]=K_H[b]$ between the complexed and uncomplexed biomarker forms. This removal of free-phase biomarker from the mixture perturbs the equilibrium ratio for the natural carrier protein, so that $[bC_N]<K_N[b]$, which in turn generates a driving force for the transfer of biomarker from the natural carrier protein to the free-phase form. This transfer sponsors the continued binding of free-phase biomarker to the harvesting particles. Biomarker transfer from the natural carrier protein to the harvesting particle via the free-phase form will continue in this way until an overall equilibrium is reached, with $[bC_H]=K_H[b]$ and $[bC_N]=K_N[b]$. The proportion of biomarker in complexed association with the harvesting particle will then be $K_H/(1+K_N+K_H) \approx K_H/(K_N+K_H)$. Thus, the greater the biomarker affinity for the harvesting particle, K_H , in comparison with its affinity for the natural carrier protein, K_N , the closer this ratio will be to unity (ie. all biomarker in complexed association with the harvesting particle.)

These theoretical considerations lead to several important conclusions about the parameters and performance characteristics of affinity capture systems for low abundance biomarker capture and detection:

1. The primary determinant of the efficiency and yield of the biomarker harvesting is the affinity of the capture bait, regardless of whether it is an antibody, a chemical (e.g. a dye), or a nucleic acid (e.g. an aptamer). The lower the abundance of the biomarker the higher the affinity that is required to capture all of the low abundance molecules of interest. A sufficiently high affinity capture ligand, or capture particle, can effectively sequester and concentrate all of the low abundance biomarker molecules from the input biofluid.
2. In the absence of a high affinity capture of the low abundance biomarker analyte, the analyte of interest will tend to partition with the high abundance carrier proteins including albumin and immunoglobulin, that exist in vast excess. Even if the affinity of the low abundance biomarker for the carrier protein is low it will be driven to associate with the high abundance natural carrier protein (albumin). Thus an immune depletion method which removes the high abundance proteins will deplete the low abundance biomarkers along with the high abundance proteins. This is counterproductive for biomarker discovery by MS.

3. For a sufficiently high affinity capture harvesting particle, the low abundance biomarker will partition off of the high abundance carrier (e.g. albumin or IgG), thereby obviating the need for immune depletion.

5. Vision for the Future

5.1 Affinity Enrichment improves detection sensitivity and reduces analytical variability

The improvement in technologies for up-front enrichment promises to greatly accelerate the discovery, and routine measurement of low abundance biomarkers that were previously invisible to mass spectrometry. In support of this view, affinity capture enrichment methods have permitted the detection and measurement of low abundance hormone, infectious disease, cancer, and cardiac markers that could not be detectable by conventional methods (60–63). Nanotrap® particles have been successfully employed to capture, concentrate and preserve different classes of molecules and antigens, including lipophosphoglycans(49, 50), nucleic acids(54), cytokines(64, 65), viral and exosome particles(58, 66), greatly improving the detection capabilities of current analytical technologies.

The ideal workflow for biomarker discovery, employing affinity enrichment, should include a disposable enrichment step of sufficient high affinity for the category of target analytes. (Figure 6) A disposable single-use technology will reduce bias caused by carryover and degradation of the affinity matrix following regeneration or stripping. Moreover, routine enrichment and concentration up front for harvesting and measuring proteins less than one nanogram/mL will reduce analytical variability because the total concentration of the target protein into the MS or MRM will be elevated, falling into an analytical range that is well above the background. For conventional methods, the limits of sensitivity exist close to the concentration of the low abundance biologically relevant biomarker. In contrast, increasing the total biomarker concentration input will bring the detection system into a range of higher precision. For an immunoassay, or a MS MRM assay, increasing the total analyte concentration input can also bring the assay from the non-linear to the linear region of the dose response curve (Figure 5).

5.2 Low abundance biomarker discovery, a new frontier for disease risk assessment

The major limitation of conventional MS biomarker discovery to date is the high abundance of the candidate biomarkers that emerge (67). Studying the biologic identity of published MS candidate biomarkers reveals that markers of coagulation, immune cell function, stress, and extracellular matrix predominate. While these markers may predict an indirect response of the host to the disease, it is unlikely that they will be highly specific, since they are not directly derived from the diseased cells (67). Routine discovery of novel low abundance biomarkers in the picogram/mL range has the potential to revolutionize early disease diagnosis because this can reach down into the low concentration region, comparable to the sensitivity of PCR, required for detection of premalignant lesions or latent infections (8, 9). This will herald of a new class of low abundance biomarkers for predicting RISK of disease. Instead of restricting our biomarker candidates to those associated with growing tumors, we can envision biomarkers that are derived from the cancer premalignant field which precedes the frank emergence of neoplasms. Thus, protein biomarkers of the future might predict who

are at higher risk to get cancer in the next 5 years, rather than who has advanced cancer today (Fig 7). The same philosophy applies to infectious disease. Latent or early stage infections are currently diagnosed by serology. Serology, by definition, does not reveal active infectious disease. Employing affinity enrichment, MS of the future will monitor protein antigens shed by the pathogen into blood or urine. Using MS in a discovery mode, or MS in a quantitative MRM mode, it will be possible to detect latent or active infections, co-infections, or infections that persist after therapy. This will realize the promise of MS for routine clinical diagnostics.

Sample fractionation methods (e.g. PRISM-SRM(32, 68)) can greatly improve the intrinsic sensitivity of mass spectrometry. Shi and colleagues demonstrated that high resolution fractionation based on high-pH reverse phase liquid chromatography, in absence of immunoaffinity depletion, achieved a limit of detection for single reaction monitoring in the low nanogram per milliliter range using a TSQ Quantum Ultra triple quadrupole mass spectrometer. The fractionation time used in this paper was 103 minutes. A workflow that combines enhanced chromatographic separation with sample concentration will reach the sensitivity that approaches PCR and that is required to detect markers of risk (Figure 6).

Conclusion

We recommend a fully automated MS biomarker discovery workflow that incorporates a disposable, affinity based enrichment step. The enrichment step for positive selection of the candidate biomarkers, as discussed above, is advantageous compared to immunodepletion methods which have the drawback of high cost, carryover contamination, and removal of the low abundance biomarkers that are bound to high abundance proteins captured by the depletion columns. Following affinity enrichment, the next step in the workflow should include enhanced chromatographic separation methods such as illustrated in the Figure 7 that contribute to the maximization of sensitivity by improving the resolution of the chromatographic separation. Finally, we recommend that biomarker candidates are stringently filtered for biological relevance and that high abundance, common host proteins (including clotting factors, stress related protein, etc.) are deleted from consideration. This will maximize the chance that the candidate biomarkers derive from the diseased tissue itself, thereby maximizing clinical relevance.

Expert Commentary

The low abundance of diseased derived protein biomarkers shed into biologic fluids is the greatest roadblock to detecting early stage infectious disease. In theory, non-affinity based concentration, and affinity enrichment methods, can improve the effective sensitivity of mass spectrometry for low abundance biomarker discovery. Nevertheless, each has significant advantages and drawbacks.

Non-affinity based

- Membrane filtration has low yield due to nonspecific binding of biomarkers to the filter matrix. Dialysis of large volumes takes an extended amount of time and may not adequately enrich the desired low-abundant biomarker. Salting out

disrupts biological activity and may hinder resolubilization of precipitated proteins. Lastly, size-sieving chromatography presents a range of logistical issues, from high voltage/current requirements, time delay for separation and migration, and difficulty in accessing or eluting proteins that enter the gel plug.

Affinity enrichment

- Negative affinity enrichment using immunodepletion. Biomarkers of interest often bind to high abundance carrier proteins such as immunoglobulins and albumin. The low abundance biomarkers remain bound to the carrier protein and are removed during immunodepletion, severely reducing the yield for MS discovery. Additionally, immunodepletion can be associated with column carryover, resulting in contamination from previous samples passed through the column. Carryover issues can violate CAP/CLIA compliance guidelines.
- **Positive affinity enrichment:** Affinity capture columns and nanoparticles avoid the aforementioned issues with immunodepletion, and can improve detection sensitivity while reducing analytical variability. Chemical, heavy metal, or genetic tagging of proteins can provide labeling options for affinity enrichment, but require prior treatment or genetic manipulation of cultured cells. When the goal is to discover novel, previously unknown protein disease markers of clinical value genetic or chemical tagging of the patient's diseased cells may not be feasible. Thus affinity capture methods which use a chemical agent to capture broad classes of native, unmodified proteins and peptides with high affinity, is preferred for discovery of novel low abundance disease associated proteins in body fluids. Many types of low-cost, organic dyes perform this function, and thus can be packaged into hydrogel particles. Commercially available Nanotrap® particles employ this principle and have already successfully been used to identify low abundant biomarkers in MS/MRM.

Five-year view

Affinity enrichment followed by mass spectrometry promises to realize the dream of non-antibody MS diagnostics. Future discoveries of very low abundance specific proteins derived from disease precursors, and early stage disease, will usher in a new class of RISK associated biomarkers applicable to disease prevention. To discover these biomarkers of the future, we recommend a fully automated MS biomarker discovery workflow that enriches positively selected candidate biomarkers and maximizes sensitivity through enhanced chromatographic separation methods, and advanced bioinformatics tools. All biomarker candidates must be rigorously filtered for biological function, looking for a plausible mechanistic connection to the diseased cells, while eliminating the high abundant, common host proteins. This approach emphasizes the yield of candidate biomarkers that are derived from the diseased tissues, maximizing clinical relevance.

Specifically for infectious diseases, we envision that mass spectrometry will play a larger role in culture-free pathogen diagnosis and functional characterization in vivo. The MS specificity criteria for the future will move from number of hits per protein to number of hits

per organism. Under this scenario, a true positive sample will be defined as showing more than one peptide greater than 7 amino acids in length, displaying 100% identity in manually validated databases of the pathogenic organism but not of every other annotated biological entity. As analytical sensitivity increases, we expect that the secretome of pathogens will be monitored in complex body fluids such as urine or blood; it should be possible to routinely document co-infections in the same patient, define the organism clade or subtype and monitor therapy success.

References

1. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*. 2003; 422:198–207. [PubMed: 12634793]
2. Poste G. Bring on the biomarkers. *Nature*. 2011; 469:156–157. [PubMed: 21228852]
3. Kennedy JJ, et al. Optimized Protocol for Quantitative Multiple Reaction Monitoring-Based Proteomic Analysis of Formalin-Fixed, Paraffin-Embedded Tissues. *J Proteome Res*. 2016; 15:2717–2728. [PubMed: 27462933]
4. Gerszten RE, et al. Challenges in translating plasma proteomics from bench to bedside: update from the NHLBI Clinical Proteomics Programs. *Am J Physiol Lung Cell Mol Physiol*. 2008; 295:L16–22. [PubMed: 18456800]
5. Johann DJ, et al. Combined blood/tissue analysis for cancer biomarker discovery: application to renal cell carcinoma. *Anal Chem*. 2010; 82:1584–1588. [PubMed: 20121140]
6. Kuhn E, et al. Developing multiplexed assays for troponin I and interleukin-33 in plasma by peptide immunoaffinity enrichment and targeted mass spectrometry. *Clin Chem*. 2009; 55:1108–1117. [PubMed: 19372185]
7. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics*. 2002; 1:845–867. [PubMed: 12488461]
8. Lutz AM, Willmann JK, Cochran FV, Ray P, Gambhir SS. Cancer screening: a mathematical model relating secreted blood biomarker levels to tumor sizes. *PLoS Med*. 2008; 5:e170. [PubMed: 18715113]
9. Pan S, et al. Mass spectrometry based targeted protein quantification: methods and applications. *J Proteome Res*. 2009; 8:787–797. [PubMed: 19105742]
10. Marshall J, et al. Processing of serum proteins underlies the mass spectral fingerprinting of myocardial infarction. *J Proteome Res*. 2003; 2:361–372. [PubMed: 12938926]
11. Cifani P, Kentsis A. High Sensitivity Quantitative Proteomics Using Automated Multidimensional Nano-flow Chromatography and Accumulated Ion Monitoring on Quadrupole-Orbitrap-Linear Ion Trap Mass Spectrometer. *Mol Cell Proteomics*. 2017; 16:2006–2016. [PubMed: 28821601]
12. Brown PO, Palmer C. The preclinical natural history of serous ovarian cancer: defining the target for early detection. *PLoS Med*. 2009; 6:e1000114. [PubMed: 19636370]
13. Wang YC, Han J. Pre-binding dynamic range and sensitivity enhancement for immuno-sensors using nanofluidic preconcentrator. *Lab Chip*. 2008; 8:392–394. [PubMed: 18305855]
14. Holstein CA, Griffin M, Hong J, Sampson PD. Statistical method for determining and comparing limits of detection of bioassays. *Anal Chem*. 2015; 87:9795–9801. [PubMed: 26376354]
15. Bandow JE. Comparison of protein enrichment strategies for proteome analysis of plasma. *Proteomics*. 2010; 10:1416–1425. [PubMed: 20127685]
16. Schumacher JA, Crockett DK, Elenitoba-Johnson KS, Lim MS. Evaluation of enrichment techniques for mass spectrometry: identification of tyrosine phosphoproteins in cancer cells. *J Mol Diagn*. 2007; 9:169–177. [PubMed: 17384208]
17. Onnerfjord P, Eremin SA, Emnéus J, Marko-Varga G. High sample throughput flow immunoassay utilising restricted access columns for the separation of bound and free label. *J Chromatogr A*. 1998; 800:219–230. [PubMed: 9561764]
18. Tudorache M, Co M, Lifgren H, Emnéus J. Ultrasensitive magnetic particle-based immunosupported liquid membrane assay. *Anal Chem*. 2005; 77:7156–7162. [PubMed: 16285661]

19. Lin ZJ, Desai-Krieger D, Shum L. Simultaneous determination of glipizide and rosiglitazone unbound drug concentrations in plasma by equilibrium dialysis and liquid chromatography-tandem mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2004; 801:265–272.
20. de Godoy LM, et al. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*. 2008; 455:1251–1254. [PubMed: 18820680]
21. Meleady P. 2D gel electrophoresis and mass spectrometry identification and analysis of proteins. *Methods Mol Biol*. 2011; 784:123–137. [PubMed: 21898217]
22. Apori AA, Herr AE. Homogeneous immunosubtraction integrated with sample preparation enabled by a microfluidic format. *Anal Chem*. 2011; 83:2691–2698. [PubMed: 21375345]
23. Guo A, et al. Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. *Mol Cell Proteomics*. 2014; 13:372–387. [PubMed: 24129315]
24. Kaboord B, Smith S, Patel B, Meier S. Enrichment of low-abundant protein targets by immunoprecipitation upstream of mass spectrometry. *Methods Mol Biol*. 2015; 1295:135–151. [PubMed: 25820720]
25. Keilhauer EC, Hein MY, Mann M. Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). *Mol Cell Proteomics*. 2015; 14:120–135. [PubMed: 25363814]
26. Richards SL, et al. Aptamer based peptide enrichment for quantitative analysis of gonadotropin-releasing hormone by LC-MS/MS. *Talanta*. 2016; 150:671–680. [PubMed: 26838458]
27. Wang KY, et al. Multiplexed immunoassay: quantitation and profiling of serum biomarkers using magnetic nanoprobe and MALDI-TOF MS. *Anal Chem*. 2008; 80:6159–6167. [PubMed: 18642877]
28. Whiteaker JR, et al. Sequential multiplexed analyte quantification using peptide immunoaffinity enrichment coupled to mass spectrometry. *Mol Cell Proteomics*. 2012; 11 M111.015347.
29. Prieto DA, et al. Mass spectrometry in cancer biomarker research: a case for immunodepletion of abundant blood-derived proteins from clinical tissue specimens. *Biomark Med*. 2014; 8:269–286. [PubMed: 24521024]
30. Kulasingam V, Smith CR, Batruch I, Diamandis EP. Immuno-mass spectrometry: quantification of low-abundance proteins in biological fluids. *Methods Mol Biol*. 2011; 728:207–218. [PubMed: 21468950]
31. Zarei M, Sprenger A, Rackiewicz M, Dengjel J. Fast and easy phosphopeptide fractionation by combinatorial ERLIC-SCX solid-phase extraction for in-depth phosphoproteome analysis. *Nat Protoc*. 2016; 11:37–45. [PubMed: 26633130]
32. Shi T, et al. Targeted quantification of low ng/mL level proteins in human serum without immunoaffinity depletion. *J Proteome Res*. 2013; 12:3353–3361. [PubMed: 23763644]
33. Polaskova V, Kapur A, Khan A, Molloy MP, Baker MS. High-abundance protein depletion: comparison of methods for human plasma biomarker discovery. *Electrophoresis*. 2010; 31:471–482. [PubMed: 20119956]
34. Roche S, et al. Depletion of one, six, twelve or twenty major blood proteins before proteomic analysis: the more the better? *J Proteomics*. 2009; 72:945–951. [PubMed: 19341827]
35. Shen Y, et al. Characterization of the human blood plasma proteome. *Proteomics*. 2005; 5:4034–4045. [PubMed: 16152657]
36. Granger J, Siddiqui J, Copeland S, Remick D. Albumin depletion of human plasma also removes low abundance proteins including the cytokines. *Proteomics*. 2005; 5:4713–4718. [PubMed: 16281180]
37. Bellei E, et al. High-abundance proteins depletion for serum proteomic analysis: concomitant removal of non-targeted proteins. *Amino Acids*. 2011; 40:145–156. [PubMed: 20495836]
38. Prieto DA, et al. Preparation and Immunoaffinity Depletion of Fresh Frozen Tissue Homogenates for Mass Spectrometry-Based Proteomics in the Context of Drug Target/Biomarker Discovery. *Methods Mol Biol*. 2017; 1647:71–90. [PubMed: 28808996]
39. Tu C, et al. Depletion of abundant plasma proteins and limitations of plasma proteomics. *J Proteome Res*. 2010; 9:4982–4991. [PubMed: 20677825]

40. Ahn SB, Khan A. Detection and quantitation of twenty-seven cytokines, chemokines and growth factors pre- and post-high abundance protein depletion in human plasma. *EuPA Open Proteomics*. 2014; 3:78–84.
41. Mirzaei H, Schieler JL, Rochet JC, Regnier F. Identification of rotenone-induced modifications in alpha-synuclein using affinity pull-down and tandem mass spectrometry. *Anal Chem*. 2006; 78:2422–2431. [PubMed: 16579629]
42. Ueda K. Glycoproteomic strategies: From discovery to clinical application of cancer carbohydrate biomarkers. *Proteomics Clin Appl*. 2013; 7:607–617. [PubMed: 23640819]
43. Yang S, Zhang H. Solid-phase glycan isolation for glycomics analysis. *Proteomics Clin Appl*. 2012; 6:596–608. [PubMed: 23090885]
44. Thompson ST, Cass KH, Stellwagen E. Blue dextran-sepharose: an affinity column for the dinucleotide fold in proteins. *Proceedings of the National Academy of Sciences of the United States of America*. 1975; 72:669–672. [PubMed: 164664]
45. Tamburro D, et al. Multifunctional Core-Shell Nanoparticles: Discovery of Previously Invisible Biomarkers. *J Am Chem Soc*. 2011; 133:19178–19188. [PubMed: 21999289]
46. Luchini A, Espina V, Liotta LA. Protein painting reveals solvent-excluded drug targets hidden within native protein-protein interfaces. *Nat Commun*. 2014; 5:4413. [PubMed: 25048602]
47. Luchini A, et al. Official Gazette of the United States Patent and Trademark Office Patents. 2013
48. Douglas TA, et al. The use of hydrogel microparticles to sequester and concentrate bacterial antigens in a urine test for Lyme disease. *Biomaterials*. 2011; 32:1157–1166. [PubMed: 21035184]
49. Castro-Sesquen YE, et al. Use of a Chagas Urine Nanoparticle Test (Chunap) to Correlate with Parasitemia Levels in T. cruzi/HIV Co-infected Patients. *PLoS Negl Trop Dis*. 2016; 10:e0004407. [PubMed: 26919324]
50. Castro-Sesquen YE, et al. Use of a novel chagas urine nanoparticle test (chunap) for diagnosis of congenital chagas disease. *PLoS Negl Trop Dis*. 2014; 8:e3211. [PubMed: 25275534]
51. Shafagati N, et al. The use of Nanotrap particles for biodefense and emerging infectious disease diagnostics. *Pathogens and disease*. 2014; 71:164–176. [PubMed: 24449537]
52. Patanarut A, et al. Synthesis and characterization of hydrogel particles containing Cibacron Blue F3G-A. *Colloids Surf A Physicochem Eng Asp*. 2010; 362:8–19. [PubMed: 20871782]
53. Magni R, et al. Application of Nanotrap technology for high sensitivity measurement of urinary outer surface protein A carboxyl-terminus domain in early stage Lyme borreliosis. *J Transl Med*. 2015; 13:346. [PubMed: 26537892]
54. Narayanan A, et al. Reactive oxygen species activate NFκB (p65) and p53 and induce apoptosis in RVFV infected liver cells. *Virology*. 2014; 449:270–286. [PubMed: 24418562]
55. Luchini A, et al. Application of Analyte Harvesting Nanoparticle Technology to the Measurement of Urinary HGH in Healthy Individuals. *J Sports Med Doping Stud*. 2012; 2
56. Fredolini C, et al. A novel technological approach for the discovery of low molecular weight serum biomarkers based on “smart” hydrogel nanoparticles applied to prostate cancer. *Proceedings of the American Association for Cancer Research Annual Meeting*. 2009; 50:1081.
57. Shafagati N, et al. The use of NanoTrap particles as a sample enrichment method to enhance the detection of Rift Valley Fever Virus. *PLoS Negl Trop Dis*. 2013; 7:e2296. [PubMed: 23861988]
58. Jaworski E, et al. The use of Nanotrap particles technology in capturing HIV-1 virions and viral proteins from infected cells. *PLoS One*. 2014; 9:e96778. [PubMed: 24820173]
59. Longo C, et al. Core-shell hydrogel particles harvest, concentrate and preserve labile low abundance biomarkers. *PLoS One*. 2009; 4:e4763. [PubMed: 19274087]
60. Liu C, et al. Quantification of circulating Mycobacterium tuberculosis antigen peptides allows rapid diagnosis of active disease and treatment monitoring. *Proc Natl Acad Sci U S A*. 2017; 114:3969–3974. [PubMed: 28348223]
61. Miller WL, et al. Comparison of mass spectrometry and clinical assay measurements of circulating fragments of B-type natriuretic peptide in patients with chronic heart failure. *Circ Heart Fail*. 2011; 4:355–360. [PubMed: 21292992]

62. Waybright TJ, Xu X, Faupel-Badger JM, Xiao Z. Preparation of human serum for prolactin measurement by multiple reaction monitoring mass spectrometry. *Methods Mol Biol.* 2013; 1002:195–203. [PubMed: 23625405]
63. Kushnir MM, et al. Measurement of thyroglobulin by liquid chromatography-tandem mass spectrometry in serum and plasma in the presence of antithyroglobulin autoantibodies. *Clin Chem.* 2013; 59:982–990. [PubMed: 23396140]
64. Popova TG, Teunis A, Espina V, Liotta LA, Popov SG. Chemokine-Releasing Microparticles Improve Bacterial Clearance and Survival of Anthrax Spore-Challenged Mice. *PLoS One.* 2016; 11:e0163163. [PubMed: 27632537]
65. Popova TG, et al. Chemokine-Releasing Nanoparticles for Manipulation of the Lymph Node Microenvironment. *Nanomaterials.* 2015; 5:298–320. [PubMed: 25878893]
66. Narayanan A, et al. Exosomes derived from HIV-1-infected cells contain trans-activation response element RNA. *J Biol Chem.* 2013; 288:20014–20033. [PubMed: 23661700]
67. Crutchfield CA, Thomas SN, Sokoll LJ, Chan DW. Advances in mass spectrometry-based clinical biomarker discovery. *Clin Proteomics.* 2016; 13:1. [PubMed: 26751220]
68. Shi T, et al. Antibody-free, targeted mass-spectrometric approach for quantification of proteins at low picogram per milliliter levels in human plasma/serum. *Proc Natl Acad Sci U S A.* 2012; 109:15395–15400. [PubMed: 22949669]

Key issues

- Mass spectrometry is the leading approach for discovering novel, disease associated protein biomarkers, but has poor sensitivity when detecting low abundant biomarkers in body fluids.
- The majority of diagnostic protein analytes measured in clinical laboratories exist in low concentrations less than one nanogram/mL, and most are not currently detectable by MS. There is a critical need to improve this effective sensitivity if MS is going to become a routine clinical diagnostic tool.
- Affinity enrichment and concentration methods enhance the effective sensitivity of mass spectrometry for biomarker discovery, but also come with significant drawbacks that must be taken into consideration.
- The most promising approach for the future is positive selection through high affinity binding of low abundance candidate protein biomarkers. Affinity capture nanoparticles have been documented to improve detection sensitivity and reduce analytical variability, enabling low-abundance biomarker discovery

Figure 1. Disease specific biomarkers exist at exceedingly low concentration in blood. Biomarkers that emanate from a small lesion enter the blood circulation, are diluted in a large blood volume, and can be cleared by the liver or kidneys.

Figure 2.

Non-affinity based concentration methods. Protein precipitation (A) and volume reduction by dialysis (B) are example methods that do not improve signal to noise ratio (C). High abundance resident proteins (HAP, e.g., albumin, complement cascade) are concentrated to the same extent as the low abundance proteins (LAP), which constitute the biomarkers of interest. The net result is that LAP are still invisible to mass spectrometry because their signal is below the dynamic range of the machine.

Figure 3.

Working principle of immuno-enrichment and immuno-depletion is depicted in panel (A). In immuno-enrichment separations, antibodies or affinity molecules specific for a single protein or a protein tag are immobilized onto a solid phase. All the unbound analytes are washed away and the target proteins are subsequently eluted. In immuno-depletion separations, antibodies or affinity molecules specific for high abundance, unwanted proteins are immobilized onto a solid phase. The flow-through is therefore deprived of the high abundance proteins. Common challenges and pitfalls of antibody-based concentration

methods are depicted in panel (B). In blood and urine, low abundance biomarkers naturally form complexes with high abundance resident proteins that function as carrier proteins. This poses a challenge for immuno-enrichment and immuno-depletion strategies because it causes the target biomarker to follow the faith of high abundance proteins and to be washed away or retained in solid phase, respectively. Moreover, antibody columns and resins can not be disposable because of their high cost. Therefore, instances of carry over contamination from previous samples can arise.

Figure 4. Transfer of a low molecular weight biomarker from a natural carrier protein to a harvesting particle with only a ten times greater affinity for the biomarker.

Figure 5.

Nanotrap® particle biomarker harvesting, concentration, and preservation followed by MS analysis for detection of protein biomarkers. (A) Open meshwork core Nanotrap® particles rapidly sequester analytes and exclude high abundance proteins. Chemical baits are not specific for a single protein but can capture a variety of proteins simultaneously. Specificity of analysis is provided by MS. (B) The concentration factor (up to 1000 fold) is determined by the ratio of the sample input fluid volume to the input volume of the assay. (C) Signal to

noise ratio is greatly enhanced because the Nanotrap® particles reduce the background signal and effectively increase the concentration of target biomarkers.

Figure 6.

Nanotrap® particles increase the effective sensitivity of multiple reaction monitoring (MRM) and western blot analysis. (A) Standard curve for the OspA peptide GYVLEGLTAEK obtained with a triple quadrupole mass spectrometer and the stable isotope dilution (SID) method. *Borrelia burgdorferi* protein lysate was spiked into human urine and the sample was subjected to Nanotrap® particle pre-processing. The linearity range spans the concentration interval from 5 to 100 pg/mL. (B) Standard curve for the OspB peptide ATIDQVELK obtained with a triple quadrupole mass spectrometer and the SID method. *Borrelia burgdorferi* protein lysate was spiked into human urine and the sample was subjected to Nanotrap® particle pre-processing. The linearity range spans the concentration interval from 10 to 100 pg/mL. (C) Western blot analysis shows that OspA and OspB proteins yield a very high signal after Nanotrap® particle pre-processing (sample E) but they are invisible in the initial solution (sample IS). (D) Nanotrap® particles increase the effective detection limit of western blot analysis. (E) Image density analysis of the western blot bands was performed with ImageJ. Panels (C, D and E) were adapted from (48).

Figure 7.

The ideal workflow for MS based biomarker discovery should include an affinity capture concentration method, an enhanced chromatography system to increase sensitivity (e.g. PRISM-SRM(32, 68)), and a statistical analysis that emphasizes biological relevance (A). Increased sensitivity of MS biomarker discovery platforms will open the way to discovery of biomarkers of disease risk (B).

$$\frac{[bC_N]}{[b]} = \frac{k_N}{k_{-N}} = K_N,$$

Scheme 1.

Scheme 2.