Biomarker amplification by serum carrier protein binding

Arpita I. Mehta[a,b,*], Sally Ross[b,c], Mark S. Lowenthal[b], Vincent Fusaro[b,c], David A. Fishman[d], Emanuel F. Petricoin III[b] and Lance A. Liotta[b]

[a]*NIH-Howard Hughes Research Scholar, Howard Hughes Medical Institute, Bethesda, MD 20814, USA*
[b]*FDA-NCI Clinical Proteomics Program, Office of the Director, Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, MD 20892, USA*
[c]*FDA-NCI Clinical Proteomics Program, Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892, USA*
[d]*National Ovarian Cancer Early Detection Program, Northwestern University Medical School, Chicago, IL 60611, USA*

Abstract. Mass spectroscopic analysis of the low molecular mass (LMM) range of the serum/plasma proteome is a rapidly emerging frontier for biomarker discovery. This study examined the proportion of LMM biomarkers, which are bound to circulating carrier proteins. Mass spectroscopic analysis of human serum following molecular mass fractionation, demonstrated that the majority of LMM biomarkers exist bound to carrier proteins. Moreover, the pattern of LMM biomarkers bound specifically to albumin is distinct from those bound to non-albumin carriers. Prominent SELDI-TOF ionic species (m/z 6631.7043) identified to correlate with the presence of ovarian cancer were amplified by albumin capture. Several insights emerged: a) Accumulation of LMM biomarkers on circulating carrier proteins greatly amplifies the total serum/plasma concentration of the measurable biomarker, b) The total serum/plasma biomarker concentration is largely determined by the carrier protein clearance rate, not the unbound biomarker clearance rate itself, and c) Examination of the LMM species bound to a specific carrier protein may contain important diagnostic information. These findings shift the focus of biomarker detection to the carrier protein and its biomarker content.

1. Introduction

Despite the urgent clinical need to discover serum biomarkers for the early detection of disease, the number of new biomarkers reaching routine clinical use remains unacceptably low [1–3]. The low molecular mass range ($<$ 10 kDa) of the plasma proteome, while largely uncharacterized, promises to contain a rich source of previously undiscovered biomarkers [4–16]. Low molecular mass molecules exist below 10 kDa, the lower limit of effective resolution achieved by conventional 2-D gel electrophoresis [4,5]. Conse-

quently, investigators have turned to mass spectrometry, which exhibits its optimal performance in the low molecular mass range. Initial attempts to employ mass spectrometry for the identification of biomarkers have been very promising [6–13].

Traditionally, mass spectrometry analysis of complex protein mixtures involves an upfront chromatographic separation step [4–6], followed by an enzymatic fragmentation of the separated proteins for direct MS-MS identification [14–16]. In contrast to such traditional approaches, recent applications of mass spectrometry biomarker analysis may have been successful because no enzymatic treatment was conducted and have utilized the native undigested serum proteome as a launch point for biomarker discovery [9–13]. Small molecular mass biomarkers may in fact be created by specific disease related enzymatic cleav-

*Corresponding author: Arpita I. Mehta, 10 Center Drive, Bldg 10, Room 2A33, Laboratory of Pathology, NCI, NIH, Bethesda, MD 20892-1500, USA. Tel.: +1 847 530 8230; Fax: +1 301 480 0853; E-mail: arpita.mehta@tufts.edu.

age or posttranslational modification of larger proteins. Enzymatic treatment prior to analysis may destroy or mask this information content by cleavage of disease biomarkers and by creating large quantities of enzymatic fragments from high abundance proteins.

Several groups of investigators have reported the discovery of low molecular mass diagnostic biomarkers using direct mass spectrometry analysis of non-enzymatically treated serum or plasma [8–13]. Since the vast majority of the proteins in the test sample are above the range of accurate detection (> 10 kDa) by the mass spectrometer, the low molecular mass biomarkers that emerge from the analysis must be derived from two possible sources: a) free in solution or b) bound to larger carrier molecules. SELDI-TOF and MALDI-TOF mass spectrometry analysis involves the laser-induced ionization of dried mixtures of molecules [17,18] adherent to a surface. The generated ions can represent biomarkers originally existing in the free unbound phase or existing in a bound state with larger proteins. Since the low molecular mass serum/plasma proteome is largely uncharacterized, there has been no previous analytical or experimental estimate of the relative proportion of small molecular mass biomarkers that exist in the free versus the bound phase.

Under the assumption that the low molecular mass biomarkers contain important diagnostic information and that protein biomarkers useful in disease detection are of very low abundance, the search for biomarkers usually begins with a separation step to remove the abundant high molecular mass "contaminating" proteins such as albumin, thyroglobulin, and immunoglobulins so that the analysis can focus on the lower molecular mass region [14–16]. The purpose of the present study is to examine the proportion of the low molecular mass species bound to the high molecular mass fraction of the serum/plasma proteome. From a physiologic perspective, free phase low molecular mass molecules (< 30 kDa) should be rapidly cleared through the kidney [19–20]. Consequently such rapid physiologic excretion may significantly reduce the concentration of free phase low molecular mass species to a level below detection. In the face of the vast excess of high molecular mass serum proteins, it may be likely that low abundance and low mass species will tend to bind large carrier proteins. The abundant high molecular mass carrier proteins exist above the cut-off for kidney clearance [19,20], and hence possess a half-life that is many orders of magnitude larger than small molecules. Circulating carrier proteins may thus become the reser-

voir for the accumulation and amplification of bound low mass biomarkers. Non-covalent association with albumin has been shown to extend the half-life of short-lived proteins introduced into the circulation [21–23]. The fact that many investigators now are employing and/or developing methods by which the higher abundance proteins above 30 kDa are specifically subtracted from native serum/plasma prior to analysis [4–6] may dramatically diminish the chances of finding the important low abundance and low molecular mass disease biomarkers.

We examined the proportion of low molecular mass species detectable by SELDI (surface enhanced laser desorption and ionization) that are associated with the higher molecular mass serum proteome. Human serum was fractionated into high molecular mass and low molecular mass native fractions. Each fraction was assayed by SELDI to assess whether the preponderance of low molecular mass ions is found in the low or the high molecular mass fraction. We further examined the subpopulation of molecular species bound to albumin compared to the total carrier protein fraction. We explored whether SELDI-TOF identified biomarkers correlating with presence of ovarian cancer were associated with high molecular mass carrier proteins. Finally, we explored the theoretical implications of biomarker amplification due to carrier protein binding. The overall aim was to evaluate the use of carrier proteins as an affinity capture means for disease relevant biomarkers.

2. Materials and Methods

2.1. Materials

Montage Albumin Deplete Kit, Centriplus centrifugal filters with average MWCO (molecular weight cutoff) of 30 kDa, and ZipTip C_{18} reversed-phase desalting columns (6 μm bed) were purchased from Millipore Corporation (Bedford, MA); human serum acquired and processed in house; sequencing grade modified trypsin purchased from Promega (Madison, WI); dithiothreitol (DTT) and iodoacetamide, (97%) were purchased from Aldrich Chemical Co (Milwaukee, WI); glacial acetic acid and formic acid (88%) were purchased from Maillinckrodt Baker Inc (Phillipsburg, NJ); trifluoroacetic acid (TFA, sequanal grade) purchased from Pierce (Rockford, IL); acetonitrile (ACN) spectrometric grade purchased from Merck (Darmstadt, GE); ammonium bicarbonate (NH4HCO3) certified purchased from Fisher Scientific (Fair Lawn, NJ); urea, enzyme grade, purchased from BRL Life Tech (Gaithersburg, MD).

2.2. Serum samples

Serum samples were derived from the ovarian cancer clinical study set of the National Ovarian Cancer Early Detection Program, Northwestern University. The full characteristics of this study set have been described previously [9], and are posted in detail on http://clinicalproteomics.steem.com. Bioinformatic serum proteomic pattern analysis was conducted as described previously [9]. Further details are provided on http://clinicalproteomics.steem.com.

2.3. Mass spectrometry

Surface Enhanced Laser Desorption and Ionization (SELDI) was conducted using a PBSII (Ciphergen Systems) as described previously [17,18]. Human serum was collected and anonymized as previously reported [17,18]. Analysis was conducted on a WCX2 (weak cationic exchange) chip. The serum was fractionated into molecular mass classes under native conditions.

2.4. Mass fractionation

Thirty microliters of unfractionated human native serum was introduced into a Sephadex G-25® or a Sephadex G-50 molecular sieve spin column according to the manufacturer's instructions. The column was centrifuged at 3000 x g for three minutes, and approximately 30 microliters of eluate containing the high molecular mass fraction was collected. The eluate was treated with 50% acetonitrile (w/w in water) to dissociate bound molecules for 30 minutes and was transferred to the inlet of a molecular filtration microcolumn. (Microcon YM-30 Millipore Centrifugal Filter Device) The column was centrifuged at 1000 x g. The eluate containing the low molecular weight fraction was collected. All fractions at each stage were sampled and one microliter was analyzed by SELDI.

2.5. Albumin separation

Segregation of albumin and its low molecular mass binding constituents was conducted using the Montage Albumin Deplete Kit. 100 μL of human serum was diluted one to one with Equilibration Buffer provided with the kit for a final volume of 200 μL, and vortexed. The column was rehydrated twice with 400 μL of Equilibration Buffer and centrifuged through the column insert for 2 minutes at 2,000 rpm. 200 μL of diluted serum was introduced into the rehydrated albumin column and centrifuged for 2 minutes at 2,000 rpm. The eluate from the column contained the serum without albumin. The bound fraction contained the albumin and the low molecular weight species bound to the albumin. We then added 400 μL EAM solution composed of 50% acetonitrile and 0.1% TFA to the column to strip the column and dissociate albumin from its bound species. After 30 minutes, EAM solution was centrifuged through the column at 2,000 rpm for 3 minutes. The eluate contained the dissociated albumin and low molecular weight species that bind to albumin.

Analysis of the proteins bound to the column using ion trap mass spectrometry was performed in line with an LCQ Classic MS (ThermoFinnigan, San Jose, CA) with a modified nanospray source. Dynamic exclusion of the three most abundant peptide hits from a full MS scan were selected for MS/MS analysis by collision induced dissociation with normalized collision energy of 35% and an activation time of 30 ms. Ion spray voltage was 2.00 kV with a capillary voltage of 26.20 V and a capillary temperature of 160°C. Results for MS/MS scans were searched and compared with theoretical spectra in the Sequest Browser database specified for human proteins.

2.6. SELDI/TOF

WCX2 protein arrays were processed in a bioprocessor (Ciphergen Biosystems, Inc). 100 μl of 10 mM HCl was applied to the protein arrays in the bioprocessor and allowed to incubate for 5 minutes. The HCl wash was aspirated and discarded and 100 μl of H_2O was applied and allowed to incubate for one minute. The H_2O was aspirated and discarded, then reapplied for another minute. 100 μl of 10 mM ammonium acetate with 0.1% TritonX was applied to the surface and allowed to incubate for 5 minutes. The ammonium acetate was aspirated and discarded. A second application of ammonium acetate was applied and allowed to incubate for 5 minutes. The chip surfaces were then dried using a vacuum to remove any excess amount of liquid. Five microliters of raw sera, or molecular mass fraction, or eluate, was then applied to each chip surface and allowed to incubate for 55 minutes. Each protein chip was washed six times with 150 μl of PBS and H_2O and then vacuum dried. Cross contamination was eliminated between spots by using a bioprocessor gasket. The gasket was removed and 1.0 μl of a saturated solution of the Energy Absorbing Molecule cinnamic acid (25% saturation) in 50% (v/v) acetonitrile, 0.5%
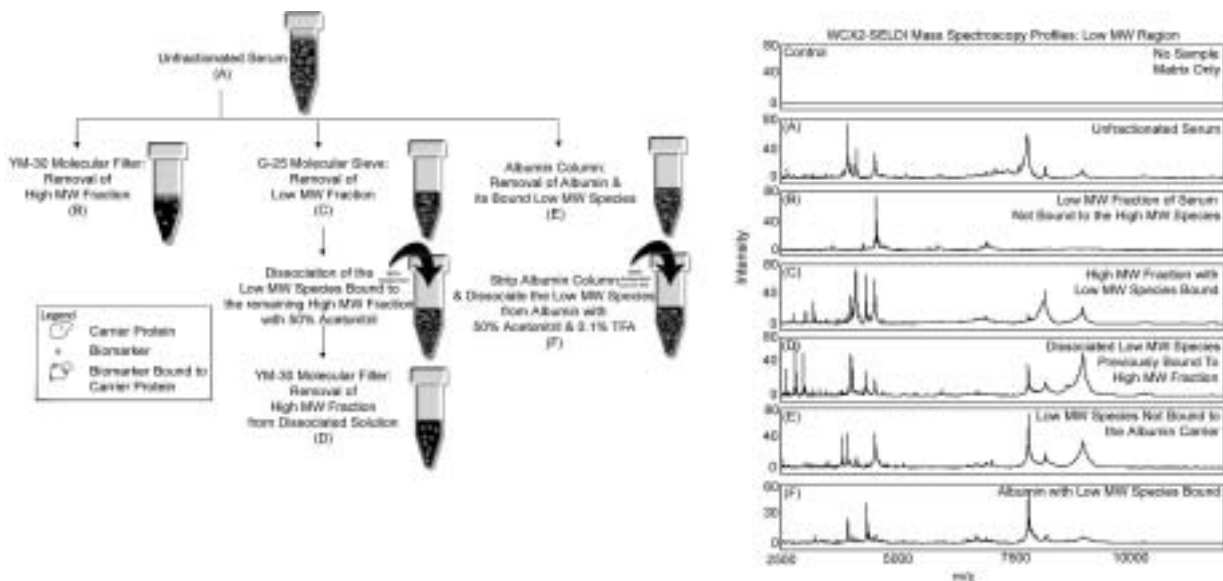
Fig. 1. Mass spectrometry experimental analysis of high and low molecular mass serum protein fractions. Left: separation protocol. Right: SELDI-TOF profiles of fractions. The preponderance of mass spectrometry detected low molecular weight ions are retained within the high molecular weight protein carrier fraction (C-F), and do not exist in the free phase (B). Horizontal axis: Mass/Charge (m/z). Vertical axis: Intensity. SELDI - MS Analysis (weak cationic exchange, WCX2): Each peak in the MS profiles represents the ionization of the low molecular weight species within the following serum fractions. Control: Matrix alone, background spectra (A) Spectra generated from whole, native, unfractionated serum applied directly to the WCX2 chip surface. (B) Spectra after removal of all species greater than 30,000 MW by Microcon YM-30 Molecular Filter from the unfractionated serum. (C) Direct analysis of high molecular weight native species by MS after fractionation through the Sephadex G-25 Molecular Sieve. The MS profile displays the low molecular weight species ionized and desorbed away from the high molecular weight carriers. (D) The high molecular weight fraction eluted from the G-25 sieve was treated with a 50% acetonitrile solution to dissociate potential low molecular weight species bound to their high molecular weight carriers. This dissociated eluate was then passed through the YM-30 Filter. Only the low molecular weight species previously bound to the G-25 sieved high molecular weight fraction should pass through the filter. (E) The ion species associated with HMW carrier species depleted of albumin and albumin binding partners. (F) Stripped albumin and ion species associated specifically with albumin only.

trifluoroacetic acid was applied to each spot on the protein array twice allowing the solution to dry between applications.

2.7. Mathematical modeling

The kinetics of biomarker production, carrier protein(s) binding, and clearance, was modeled as a deterministic compartmental model with first order kinetics.

3. Results

3.1. Association of LMW species with HMW carrier proteins

Analysis of native human serum fractionated into high and low molecular mass fractions revealed that the vast majority of low molecular mass serum / plasma biomarkers detectable as MS ions, exist bound to large carrier proteins. SELDI analysis of native serum frac-

tions of high and low molecular mass, shown in Fig. 1, demonstrate that virtually all of the detectable ions are derived from molecular species bound to large carrier proteins. In fact, removal of the high molecular mass proteins under native conditions (Fig. 1(B)), a common method used for biomarker discovery [14–16], removes a significant proportion of the ions generated by SELDI-TOF. Comparing the spectra of Figs 1(A) to 1(C) indicates that the majority of ions generated from unfractionated serum are derived from species associated with larger carrier proteins. Figure 1(D), displays the ion spectra of species previously bound, and then dissociated and separated from the higher molecular mass fraction. The intensity and number of many ion species is augmented comparing Figs 1(A), and 1(C).

3.2. Populations of LMW species associated specifically with albumin

Figure 1(E) displays the ions associated with the non-albumin carrier proteins, and Figure 1F displays the

ions generated from species bound only to albumin. A significant proportion of the ions in the spectra appear to be derived from species associated with albumin, compared to non-albumin carrier proteins.

We verified through microcapillary LC MS/MS that our albumin bound fraction acquired through stripping the Montage Albumin Deplete Column was entirely albumin and its bound low molecular mass species. Since there was no indication of a significant proportion of other high molecular mass proteins bound to the albumin specific column, we can assume that the low molecular mass species detected were derived from a specific association with albumin, or at least aggregated with and co-separated with albumin. Furthermore, we have positively identified hundreds of low molecular mass species after dissociation from their albumin carrier. Additional studies which will be submitted for publication elsewhere involve characterizing the entire repertoire of low molecular mass species bound to individual serum carrier proteins by LC MS/MS.

We next addressed the question as to whether the ions generated from the species bound to albumin contained disease biomarker information. SELDI-TOF ion patterns, generated on the WCX2 chips, correlating with ovarian cancer were identified by methods described previously [9] (http://clinicalproteomics.steem.com). Two clinical sera data sets were employed and all spectra are provided on the website for downloading as follows: Dataset 8-7-02: Ovarian Dataset 8-7-02.zip, Ovarian Sample Info 8-7-02.xls. The sample set included 91 unaffected controls and 162 ovarian cancers. The following selection of ion mass/charge (m/z) values generated a pattern that was 100% predictive in the training and blinded testing – 2760.6685, 19643.409, 465.56916, 6631.7043, 14051.976, 435.46452, and 3497.5508. http://clinicalproteomics.steem.com.

Randomly selected representative serum samples from this study set were analyzed by MALDI-TOF comparing the spectra generated by unfractionated sera to the spectra generated only from the species bound to albumin. As demonstrated in Fig. 2 the spectra generated from the species bound to albumin is complex and exhibits a number of differences between the cancer and the unaffected ("normal") cases shown in the example. Comparing the peak intensities between the unfractionated serum (containing all the carrier proteins with their associated or bound species), and the albumin-bound fraction (Fig. 2) indicates that a significant proportion of putative disease biomarkers may be associated with albumin.

Figure 3 is an example of ion 6631.7043, a member of the ion pattern 100% correlated with ovarian cancer

in this clinical study set. Matched for dilution and amplitude, the predicate ion is highly associated with albumin, and the ionization intensity is augmented in the albumin bound fraction. This demonstrates the albumin binding selectivity of a specific SELDI-TOF ion associated with ovarian cancer.

Figure 4 displays the ion spectra for a pooled ovarian cancer serum sample in which the ion species bound only to albumin are compared for different amounts of albumin captured on the column. The captured albumin with its associated species was denatured and its binding partners were dissociated. When a higher number of albumin molecules were stripped of their associated species, the amplitude and complexity of the LMW species, including those in the region of putative ovarian cancer biomarkers, were augmented. These data indicate that albumin capture is a feasible method for biomarker enrichment.

4. Theoretical analysis

4.1. Carrier protein concentration, $C_r(t)$: Dependency on clearance rate

At any point in time, the total concentration of the biomarker is dependent upon the biomarker production rate, the biomarker clearance / excretion rate, the binding of the biomarker to a circulating carrier protein, and the clearance / excretion rate of the carrier protein. We can view the blood intravascular space as a single compartment with volume V. We define the concentration of the carrier protein r as Cr, where the rate of carrier protein production is $k_{in,r}$, and the rate of its elimination or removal is $k_{out,r}$.

Then the change in the carrier concentration can be expressed as:

$$\frac{dC_r(t)}{dt} = \frac{k_{in,r}}{V} - k_{out,r}C_r(t) \qquad (1)$$

Using the LaPlace Transform, the concentration of the carrier protein, at time t is:

$$C_r(t) = \frac{k_{in,r}}{Vk_{out,r}} - \frac{k_{in,r}}{Vk_{out,r}}\left(e^{-k_{out,r}t}\right) \qquad (2)$$

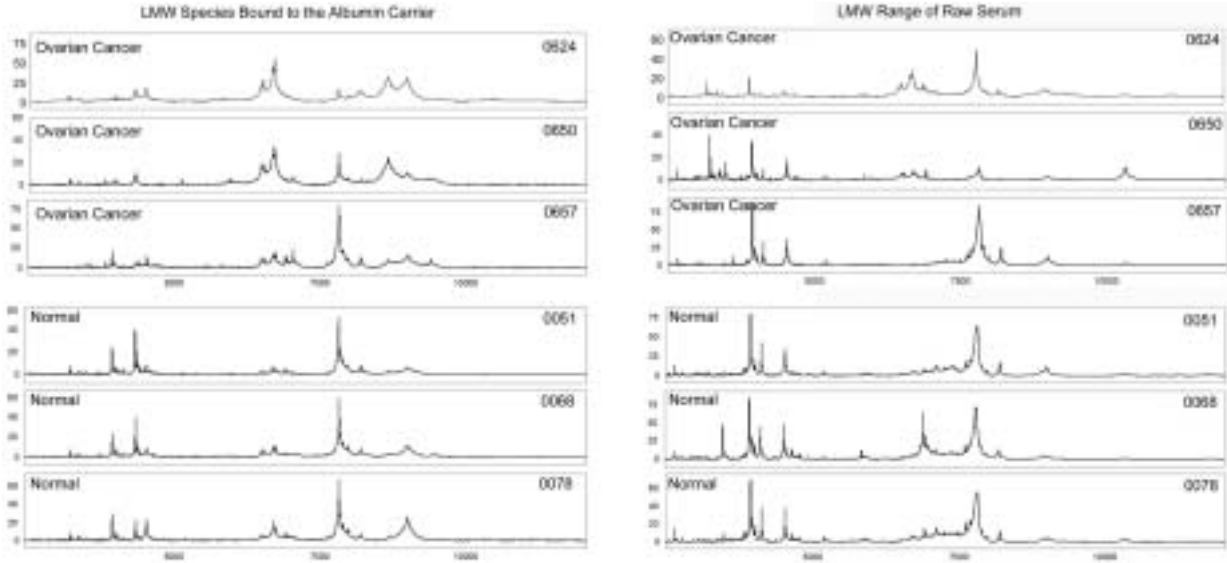The initial conditions are $C_r(t) = 0$ at $t = 0$.

Fig. 2. SELDI-TOF spectra of serum derived from the ovarian cancer sera study set, using the methods described for Figs 1(A) and 1(F).

4.2. Amplification of biomarker concentration in the presence of the carrier

We assume that a biomarker is continuously produced or shed from the tissue source over time. As shown by the experimental data, biomarker molecules can accumulate over time in a carrier-bound form. At steady state, the total concentration of a biomarker measured in a blood sample can therefore become elevated due to its association with the carrier protein. The level of amplification (A) of the biomarker concentration at steady state, due to the presence of the carrier protein can be defined as the following ratio where C_B is the concentration of the biomarker.

$$A = \qquad\qquad\qquad\qquad (3)$$
$$\frac{C_B \text{ in the presence of carrier protein}(t)}{C_B \text{ in the absence of carrier protein}(t)}.$$

Plasma biomarkers reflecting a physiologic or disease state of perfused tissues are expected to exist at concentrations many orders of magnitude below the concentration of large carrier proteins such as albumin and immunoglobulins. The experimental findings are logical because, even with a low affinity for the carrier protein, the majority of the biomarker molecules will tend to be associated with the vast excess of circulating carrier protein. Consequently, it is also logical that the biomarker will take on the clearance rate of the carrier protein it is associated with.

The concentration of the biomarker $C_B(t)$ as a function of time can therefore be described as the balance between the biomarker input production rate $k_{in,b}$ distributed in the blood volume V, and the loss or clearance of the biomarker bound to the carrier protein $(k_{out,br}C_{br}(t))$. If $k_{in,b}$ is assumed to be a simple constant production rate, a linear function of time, similar to the assumption for the carrier protein (Eq. (1)) assuming first order kinetics, where the clearance rate is a constant proportion of the carrier bound biomarker, $k_{out,br}C_{br}(t)$, where "br" refers to bound biomarker, then

$$\frac{dC_B(t)}{dt} = \frac{k_{in,b}}{V} - k_{out,br}C_{br}(t)$$
$$C_B(t) = \frac{k_{in,b}}{V k_{out,br}} - \frac{k_{in,b}}{V k_{out,br}} \left(e^{-k_{out,br}t}\right) \qquad (4)$$

Initial conditions $C_B(t) = 0$ at $t = 0$.
At $t \to \infty$ or steady state, we have

$$C_B(t) = \frac{k_{in,b}}{V k_{out,br}} \qquad\qquad\qquad (5)$$

Because the biomarker bound to the carrier protein acquires the clearance rate of the carrier protein, $k_{out,br} = k_{out,r}$, the total steady state concentration of biomarker in the plasma becomes a simple function of the biomarker production rate and the clearance, excretion rate of the carrier protein.

$$C_B(t) \approx \frac{k_{in,b}}{k_{out,r}V}. \qquad\qquad\qquad (6)$$

The results of this analysis reveal that the final total concentration of the biomarker measurable in a blood
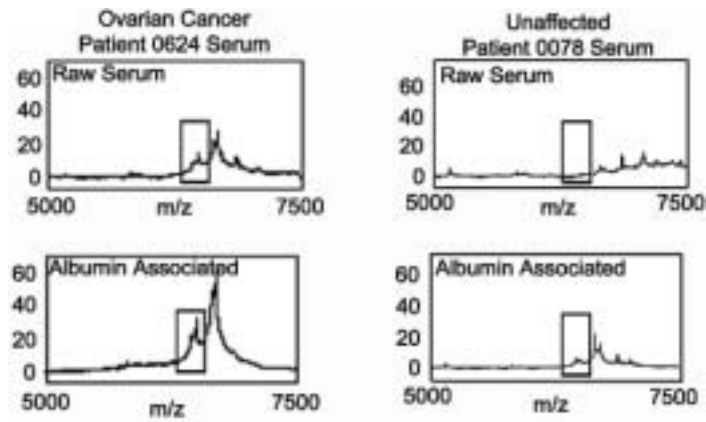
Fig. 3. Magnification of spectra region 5000–7500 m/z comparing the albumin associated ions matched for sample concentration and amplitude.
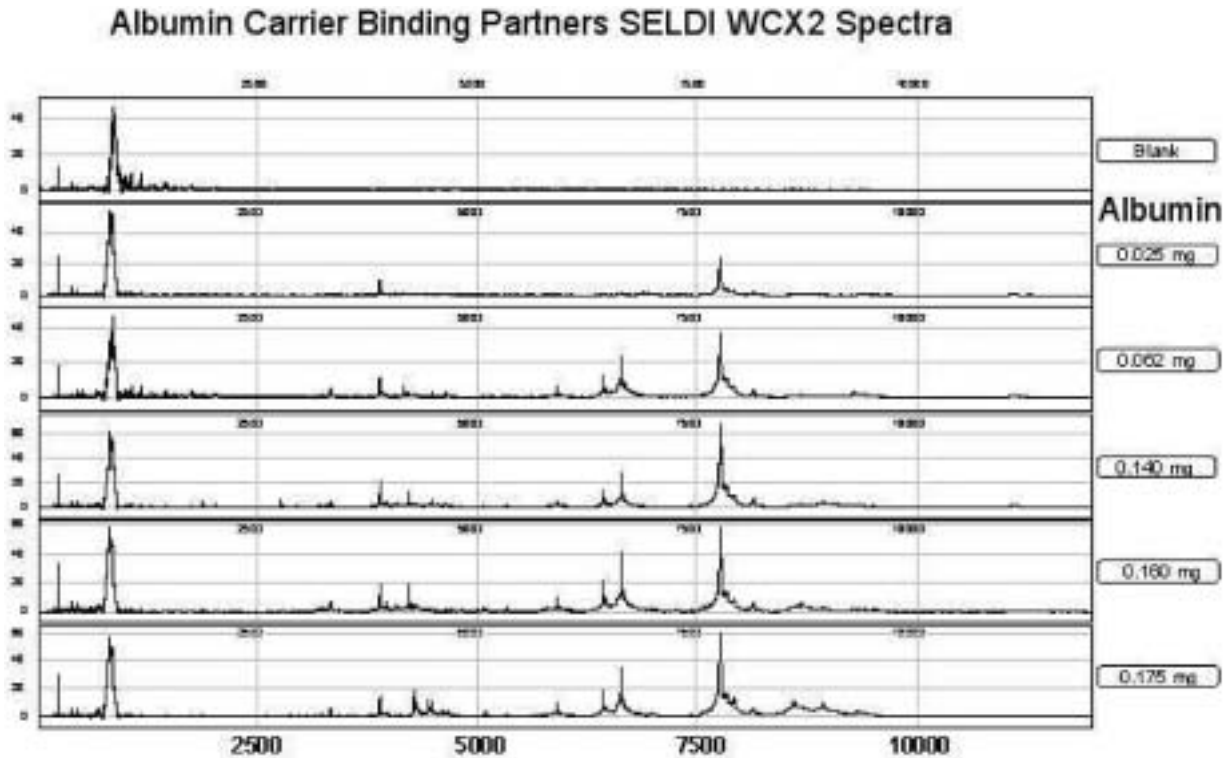

Fig. 4. SELDI-TOF spectra: Influence of ion amplitude on quantity of albumin captured (0.025–0.175 mg) from the serum. Ovarian cancer pooled serum sample. The ion species increase in amplitude and complexity when they are derived from larger quantities of bound albumin. This supports the concept that the ion species are quantitatively associated with the albumin.

sample is inversely proportional to the clearance rate of the carrier protein to which the biomarker is bound. Table 1 is a series of computed solutions to Eq. (6) for a range of hypothetical biomarker production rates, and for a series of different named carrier proteins. The clearance rates for serum carrier proteins listed in Table 1 are known [21]. The clearance and excretion rate for free biomarkers was chosen to span the known range for small molecules [19–21,27]. For a carrier protein such as albumin, with a long half-life, the resulting amplification (Table 1) can be several orders of magnitude. Carrier protein amplification thus becomes a major factor determining whether a low abundance biomarker can reach a threshold of concentration that

Table 1
Theoretical prediction based on Eq. (6) of measured total biomarker concentration for selected high abundance serum carrier proteins, as a function of biomarker production rate, and carrier protein half-life. The experimental results indicated that the majority of the biomarker species exists in a state of association with carrier proteins, whereby the clearance rate of the biomarker takes on the clearance rate of the carrier protein

Biomarker	Biomarker concentration (fmol/L)						
Production rate (fmol/day)	Absence of any carrier protein $t/2$: 1/8 days	Albumin $t/2$: 18 days	Ig $t/2$: 21 days	IgG $t/2$: 23 days	IgA/M $t/2$: 5 days	IgD $t/2$: 2.8 days	IgE $t/2$: 2 days
0.01	0.0005	0.07	0.08	0.09	0.02	0.01	0.01
0.10	0.005	0.69	0.80	0.88	0.19	0.11	0.08
1	0.05	6.88	8.02	8.79	1.91	1.07	0.76
10	0.5	68.8	80.2	87.9	19.1	10.7	7.6
100	5	688	802	879	191	107	76
1000	48	6877	8023	8788	1910	1070	764
10000	478	68773	80235	87876	19103	10698	7641
100000	4776	687725	802346	878760	191035	106980	76414
1000000	47759	6877254	8023463	8787602	1910348	1069795	764139

is above the lower limits of detection.

5. Discussion

A growing body of scientific studies supports the importance of the low molecular mass region of the serum proteome as an uncharted resource for biomarker discovery [9,28]. The experimental data of the present study supports the concept that the vast majority of small mass ions detected by mass spectrometry of native human serum exist in association with circulating carrier proteins of higher molecular weight. This conclusion has several important implications for biomarker physiology and biomarker measurement technology.

Experimental data presented in Figs 1–4 reveal that the majority of ions generated by SELDI-TOF analysis are found to be associated with carrier proteins, rather than free in solution phase (1B versus 1C-F). Moreover, as shown in Figs 2 and 3, ion species altered in disease study sets may be those specifically captured on a single carrier protein. In the example, the carrier protein is albumin. In the past, extensive effort has been placed on separating and discarding the high abundance large carrier proteins in the native plasma so that the remaining low abundance, disease-related markers could be discovered. The present results demonstrate that the search for biomarkers must be directed to those molecules bound to the carrier proteins. Removal of high abundance serum/plasma proteins prior to proteomic analysis should be conducted after, not before, dissociation from binding partners. This separation approach has been attempted for 2-D gel analysis [4,5]. As shown in Fig. 4, albumin capture can be used as a means to enrich for disease relevant biomarkers specifically associated with albumin. This now provides a novel method to harvest the necessary quantities of biomarker species required for sequencing and identification.

The first implication from this study is that the concentration of a biomarker measured in serum or plasma is directly related to the clearance rate or half life of the carrier protein, not the biomarker clearance rate itself. As shown in Eq. (6), the concentration of the biomarker is a function of the ratio between the biomarker production rate from the tissue and the clearance rate of the carrier protein. This means that carrier protein binding amplifies the total biomarker concentration levels measured in serum or plasma. Amplification occurs because the carrier protein acts as a reservoir to accumulate the biomarker over time, as the tissue is continuously producing the biomarker. Thus a biomarker produced by a small volume of tissue such as the ovary [9,10], prostate [8,12,13], or breast [11], at a low concentration (e.g. one femtomole per day) can accumulate to a concentration of one picomole in the serum because it binds with a carrier protein with a much longer half-life. In this example the existence of the carrier protein can raise the concentration of the biomarker to a range detectable by conventional assay technology [29]. Without the carrier protein, the free biomarker would be rapidly cleared by the kidney and would therefore reside at a steady state concentration many fold below the detection limits of assay technology.

The impact of this conclusion extends beyond current mass spectrometry detection technology. Small biomarkers are commonly not the province of two-site sandwich immunoassays [30,31]. This is because it is

difficult to develop two antibody-binding sites on the same small molecule. In contrast, if the first half of the immunoassay sandwich was the carrier protein and the second half was the small biomarker, a sandwich immunoassay could be achieved.

It is logical that the biomarker clearance rate becomes the carrier protein clearance rate because the carrier protein, even if it has low affinity for the biomarker, is in vast excess. This means that if we know the clearance rate of a given carrier protein, and we know the serum/plasma concentration of the biomarker bound to that carrier protein; we can estimate a lower limit for the continuous production rate by the tissue (Eq. (6)). Thus, if the concentration of a generic, bound biomarker α is 3.8 ng/mL (42.22 fmol/mL, where the biomarker α with its carrier protein have a combined molecular weight of 90 kDa) and the carrier protein half-life is 2.43 days, then the production rate of α from the tissue is at least 45,000 femtomoles per day [21,25,27,32,33]. If α is produced by a one cubic centimeter tumor composed of 10^9 cells, then each cell would produce approximately 16,000 molecules per day. This approximation is consistent with previous experimental findings [32].

These data indicate that the low molecular mass proteome, existing within the range detectable by MALDI-TOF, exists predominately in the bound phase. We propose that technologies that focus on efficient capture of the carrier proteins and specific elution of the low molecular weight biomarkers will yield the greatest amount of diagnostic information. The bound biomarkers may exist in concentrations ten to 500 times greater compared to their free counterparts. Since the carrier proteins exist in vast excess compared to the biomarkers, it is unlikely that the carrier proteins will become saturated with bound biomarkers. Moreover, based on its unique affinity topology [34], each carrier protein may have its own constellation of bound biomarkers. Indeed, the distribution of biomarkers among specific plasma/ serum carrier proteins may have important diagnostic information. Finally, these findings lead to the concept of artificial carrier molecules designed to harvest specific populations of biomarkers associated with target organs or diseases.

References

[1] N.L. Anderson and N.G. Anderson, The human plasma proteome: history, character, and diagnostic prospects, *Mol Cell Proteomics* **1** (2002), 845–867.

[2] J.B. Ward Jr and R.E. Henderson, Identification of needs in biomarker research, [Review], *Environ Health Perspect* **104** (1996), 895–900.

[3] R. Etzioni, N. Urban, S. Ramsey, M. McIntosh, S. Schwartz, B. Reid, J. Radich, G. Anderson and L. Hartwell, Early detection: The case for early detection, *Nat Rev Cancer* **3**(4) (April 2003), 243–252.

[4] Y.Y. Wang, P. Cheng and D.W. Chan, A simple affinity spin tube filter method for removing high-abundant common proteins or enriching low-abundant biomarkers for serum proteomic analysis, *Proteomics* **3**(3) (March 2003), 243–248.

[5] H.J. Issaq, T.P. Conrads, G.M. Janini and T.D. Veenstra, Methods for fractionation, separation and profiling of proteins and peptides, *Electrophoresis* **23** (2002), 3048–3061.

[6] D.L. Rothemund, V.L. Locke, A. Liew, T.M. Thomas, V. Wasinger and D.B. Rylatt, Depletion of the highly abundant protein albumin from human plasma using the Gradiflow, *Proteomics* **3**(3) (March 2003), 279–287.

[7] E.F. Petricoin III and L.A. Liotta, Mass Spectrometry – Based Diagnostics: The Upcoming Revolution in Disease Detection [Editorial], *Clin Chem* **49** (2003).

[8] B.L. Adam, A. Vlahou, O.J. Semmes and G.L. Wright Jr, Proteomic approaches to biomarker discovery in prostate and bladder cancers, [Review], *Proteomics* **1** (2001), 1264–1270.

[9] E.F. Petricoin, A.M. Ardekani, B.A. Hitt, P.J. Levine, V.A. Fusaro and S.M. Steinberg et al., Use of proteomic patterns in serum to identify ovarian cancer, *Lancet* **359** (2002), 572–577.

[10] E.F. Petricoin III, G.B. Mills, E.S. Kohn and L.A. Liotta, Proteomic patterns in serum and identification of ovarian cancer [Reply], *Lancet* **360** (2002), 170–171.

[11] J. Li, Z. Zhang, J. Rosenzweig, Y.Y. Wang and D.W. Chan, Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer, *Clin Chem* **48** (2002), 1296–1304.

[12] E.F. Petricoin, D.K. Ornstein, C.P. Paweletz, A. Ardekani, P.S. Hackett and B.A. Hitt et al., Serum proteomic patterns for detection of prostate cancer, *J Natl Cancer Inst* **94** (2002), 1576–1578.

[13] B.L. Adam, Y. Qu, J.W. Davis, M.D. Ward, M.A. Clements and L.H. Cazares et al. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men, *Cancer Res* **62** (2002), 3609–3614.

[14] B.A. Lollo, S. Harvey, J. Liao, A.C. Stevens, R. Wagenknecht and R. Sayen et al., Improved two-dimensional gel electrophoresis representation of serum proteins by using Proto-Clear, *Electrophoresis* **20** (1999), 854–859.

[15] A.B. Kantor, Comprehensive phenotyping and biological marker discovery, [Review], *Dis Markers* **18** (2002), 91–97.

[16] W.H. McDonald and J.R. Yates 3rd, Shotgun proteomics and biomarker discovery, [Review], *Dis Markers* **18** (2002), 99–105.

[17] T. Bonk and A. Humeny, MALDI-TOF-MS analysis of protein and DNA, [Review], *Neuroscientist* **7** (2001), 6–12.

[18] S.R. Weinberger, E. Boschetti, P. Santambien and V. Brenac, Surface-enhanced laser desorption-ionization retentate chromatography mass spectrometry (SELDI-RC-MS): A new method for rapid development of process chromatography conditions, *J Chromatogr B Analyt Technol Biomed Life Sci* **782** (2002), 307–316.

[19] C. Cojocel, K. Maita, K. Baumann and J.B. Hook, Renal processing of low molecular weight proteins, *Pflugers Arch* **401** (1984), 333–339.

[20] T. Maack, Renal handling of low molecular weight proteins, [Review], *Am J Med* **58** (1975), 57–64.

[21] M.S. Dennis, M. Zhang, Y.G. Meng, M. Kadkhodayan, D. Kirchhofer and D. Combs et al., Albumin binding as a general strategy for improving the pharmacokinetics of proteins, *J Biol Chem* **277** (2002), 35035–35043.

[22] P. Yeh, D. Landais, M. Lemaitre, I. Maury, J.Y. Crenne and J. Becquart et al., Design of yeast-secreted albumin derivatives for human therapy: Biological and antiviral properties of a serum albumin-CD4 genetic conjugate, *Proc Natl Acad Sci USA* **89** (1992), 1904–1908.

[23] B.L. Osborn, H.S. Olsen, B. Nardelli, J.H. Murray, J.X. Zhou and A. Garcia et al., Pharmacokinetic and pharmacodynamic studies of a human serum albumin-interferon-alpha fusion protein in cynomolgus monkeys, *J Pharmacol Exp Ther* **303** (2002), 540–548.

[24] W.M. Saltzman, *Drug delivery: Engineering principles for drug therapy*, Oxford: University Press, 2001, pp. 175–196.

[25] J.D. Veldhuis, M.L. Johnson, L.M. Faunt, M. Mercado and G. Baumann, Influence of the high-affinity growth hormone (GH)-binding protein on plasma profiles of free and bound GH and on the apparent half-life of GH. Modeling analysis and clinical applications, *J Clin Invest* **91** (1993), 629–641.

[26] Y. Ding, B. Lin and C.W. Huie, Binding studies of porphyrins to human serum albumin using affinity capillary electrophoresis, *Electrophoresis* **22** (2001), 2210–2216.

[27] V. Ravery, A. Meulemans and L. Boccon-Gibod, Clearance of free and total serum PSA after prostatic surgery, *Eur Urol* **33** (1998), 251–254.

[28] M.B. Daly and R.F. Ozols, The search for predictive patterns in ovarian cancer: Proteomics meets bioinformatics, *Cancer Cell* **1** (2002), 111–112.

[29] B.N. Swanson, Delivery of high-quality biomarker assays, [Review], *Dis Markers* **18** (2002), 47–56.

[30] N.B. Kiviat and C.W. Critchlow, Novel approaches to identification of biomarkers for detection of early stage cancer, [Review], *Dis Markers* **18** (2002), 73–81.

[31] M.S. Tockman, P.K. Gupta, N.J. Pressman and J.L. Mulshine, Considerations in bringing a cancer biomarker to clinical application, [Review], *Cancer Res* **52** (1992), 2711s–2718s.

[32] N.L. Simone, A.T. Remaley, L. Charboneau, E.F. Petricoin, J.W. Glickman and M.R. Emmert-Buck et al., Sensitive immunoassay of tissue cell proteins procured by laser capture microdissection, *Am J Pathol* **156** (2000), 445–452.

[33] T. Okegawa, H. Noda, K. Nutahara and E. Higashihara, Comparisons of the various combinations of free, complexed, and total prostate-specific antigen for the detection of prostate cancer, *Eur Urol* **38** (2000), 380–387.

[34] R.G. Reed, R.C. Feldhoff, O.L. Clute and T. Peters Jr, Fragments of bovine serum albumin produced by limited proteolysis. Conformation and ligand binding, *Biochemistry* **14** (1975), 4578–4583.