The application of protein microarrays to serum diagnostics: Prostate cancer as a test case

Jeremy C. Miller\textsuperscript{a}, E. Brian Butler\textsuperscript{b}, Bin Sing Teh\textsuperscript{b} and Brian B. Haab\textsuperscript{a,}\textsuperscript{*}
\textsuperscript{a}The Van Andel Research Institute, 333 Bostwick NE, Grand Rapids, MI 49503, USA
\textsuperscript{b}Baylor College of Medicine, Houston, TX 77030, USA

1. Introduction

Reliable and specific serum disease markers have great value as non-invasive, rapid and inexpensive assays. The discovery of new disease markers is particularly necessary for diseases that are difficult to detect or diagnose at an early and curable stage. For example, the early detection of pancreatic cancer and the differentiation of malignant from benign disease are extremely difficult using current imaging and cytological methods. Improved screening tools would permit the avoidance of unnecessary pancreaticoduodenectomies and allow the opportunity to perform the procedure at a curative stage [1]. The challenge to the discovery of new serum markers lies in the difficulties of high-throughput detection and quantitation of proteins. A new tool that is potentially well suited to meet this challenge is the protein microarray. The feasibility of accurate, sensitive and specific protein microarray detection of multiple proteins in a serum background was recently demonstrated [2], and efforts are now underway to apply this technology to marker discovery. The technology as described by Haab et al. [2] was built upon the existing DNA microarray platforms that are present in many labs (see http://cmgm.stanford.edu/pbrown/), making the method practical and easy to implement.

In addition, further availability of protein microarray technology is coming through the many commercial ventures that are actively working to get various types of protein chips to market. This article addresses the use of protein microarrays for serum marker detection and discovery, using prostate cancer as a model disease. We initially describe protein microarray technology and its suitability for serum analysis, then discuss the existing serum markers for prostate cancer and the potential advantages of using multiple markers, and finally describe serum protein studies using protein microarrays.

2. Protein microarray technology for highly-parallel serum protein detection

The microarray format has many beneficial features for protein analysis, such as highly parallel detection, low sample consumption, and the potential for highly accurate and sensitive detection in multiple wavelength regions using scanning fluorescence microscopy, as recently demonstrated [2]. Certain aspects of the technology make it particularly well suited to the analysis and discovery of serum markers. For example, the ability to run many microarray experiments rapidly enables studies on the large populations of samples that are needed for good statistics on new markers. Additionally, the sophisticated software tools that are continually under development to analyze DNA microarray data also may be used to analyze protein microarray data. Many of these tools are specifically designed for the identification of genes or sets of genes that have diagnostic utility. It has been noted that new markers may be comprised of combinations of genes rather than individual genes [3]. Microarrays provide a highly effective tool to analyze the relationship between many genes to evaluate their combined value.

Multiplexed protein detection using spotted antibodies and antigens has been demonstrated for a variety
of applications with diverse technological implementations. Protein arrays on poly(vinylidene fluoride) (PVDF) and nitrocellulose membranes have been used to screen binding specificities of a protein expression library [4–6] and to detect DNA, RNA, and protein binding targets [7]. Phage displayed antibodies were arrayed onto filters for high-throughput screening of their specificities [8]. Derivatized glass slides have been used to attach microarrays of antibodies and antigens for high-throughput ELISA [9], to detect autoantibodies [10], and to detect protein-protein [2] and proteinsmall molecule interactions [11]. Since the technology is relatively new, most of the published reports focus on feasibility studies and technological characterization rather than biological studies. Efforts are underway to apply the technology to biological studies and to address the issues necessary to make the method more robust and practical.

The primary experimental challenge in obtaining useful protein microarray data is the acquisition of high specificity and high affinity protein capture reagents. The specificity and affinity of the capture reagents define the sensitivity and accuracy of the assay. Having many high quality capture reagents adds to usefulness of protein microarray data, but several aspects of protein chemistry make the collection of a such a set difficult. Unlike nucleic acids, for which binding interactions are well characterized and predictable, protein binding interactions must be identified empirically. Since protein-protein interactions have a wide variety in binding strengths, stabilities and specificities, finding a suitable binding partner to a particular protein may be difficult in some cases. Additionally, proteins are expensive and time-consuming to produce and purify.

Several approaches have been put forth to address the generation of protein capture reagents for arrays. High-throughput protein expression and purification methods have been developed, based on recombinant baculoviruses [12] or the Gateway™ recombinant cloning system [13]. The proteins are produced in 96-well microtiter plates and efficiently purified through the amino- or carboxyl-terminal attachment of an epitope tag, such as poly-histidine or Glu-Glu. An efficient method to test for proper protein expression and folding is based on the arraying of individual bacterial colonies of a cDNA library onto membranes [4]. The arrayed colonies were induced for protein expression, the cells were lysed on the membrane, and the proteins were tested for proper expression, folding, and antibody specificity by antibody staining. High-throughput, rapid and less expensive antibody production for microarrays may be possible using phage display libraries [14]. Antibodies to specific antigens can be selected from a diverse library of antibodies displayed on the surface of phage clones, and after selection the selected clones can be amplified. The feasibility of multiplexed antigen detection using arrayed scFv phage display clones on membranes was recently shown [8].

There are other technological challenges in the development of practical protein microarrays. Because proteins have an almost unlimited variety in charges, polarities and structures, the efficient attachment of specific spotted proteins while repelling the adsorption of non-specific background proteins can be difficult. Various protein attachment methods, surface blocking methods and new surfaces that are resistant to non-specific protein binding have been evaluated. A particularly effective method appears to be the attachment of biotinylated proteins through a streptavidin-biotin bridge on the end of poly(ethylene glycol) (PEG) polymer strands [15]. The PEG, which is attached to a poly-l-lysine coating on glass, efficiently repels non-specific background proteins, yet specific attachment of the capture proteins is achieved through the biotin-streptavidin junction. A simpler strategy for protein attachment is the adsorption of spotted proteins to poly-l-lysine coated glass, followed by the blocking of the surface with milk or BSA proteins [2]. No modification of the spotted protein is required with this approach, but non-specific binding to the poly-l-lysine may be higher than to PEG. The most widely used method to attach proteins to glass is the covalent reaction of protein amine groups with silane cross-linkers [9,11]. For reviews on various implementations of protein microarray technology and advantages and disadvantages for particular applications, see references [16–19].

3. Serum markers for the diagnosis of prostate cancer

The best demonstration of the utility of serum markers for cancer diagnosis is the prostate cancer marker prostate specific antigen (PSA). PSA tests are used to screen men over age 50 in the general population and at younger ages in patients who have a familial history of prostate cancer or other risk factors. 67–80% of men with developing prostate cancers are identified [20], depending on patient age and mode of the PSA test, making the PSA test the most sensitive serum test available. PSA is a useful marker of recurrence in post-operative
The main shortcoming of the PSA test is low specificity, leading to many otherwise unnecessary biopsies. Since PSA is an organ-specific rather than a cancer-specific marker, conditions such as simple hypertrophy, prostatitis, or other benign conditions produce positive PSA tests. In concentrations between 4 and 10 ng/ml, considered abnormal levels, PSA has only a 25% specificity for prostate cancer [21]. Above 10 ng/ml, PSA is more specific for prostate cancer, giving accurate diagnoses in about 67% of cases [21]. Another shortcoming of the PSA test is a lack of information about the stage and aggressiveness of the cancer, regardless of the PSA concentration. At present, doctors often have little indication whether a radical prostatectomy is necessary to prevent aggressive growth of the cancer and to prolong the patient’s life, or whether surgery is unnecessary and would give no benefit to the patient.

Due to the limitations of the PSA test, much research is being devoted to the discovery of an improved prostate cancer serum test. The proteins Keratinocyte growth factor (KGF) [22], Human glandular kallikrein 2 (hK2) [23,24], PSA complexed with alpha (2)-macroglobulin (PSA-A2M) [25] and Interleukin-8 (IL-8) [26] have been investigated as markers to improve the differentiation between benign prostatic hyperplasia (BPH) and prostate cancer. A recent study found an increase in specificity from 9% to 28% (at a 95% sensitivity) to differentiate benign prostatic hyperplasia from prostate cancer using hK2 combined with free and total PSA measurements [27]. Another study similarly found that measurements of hK2 along with free and total PSA improved the identification of prostate cancer in patients with low total PSA (2.5–4.5 ng/mL) [24].

Others serum proteins have been investigated for information on prognosis or stage of prostate cancer. The carboxy-terminal propeptide of type I procollagen (PICP), a biochemical marker of bone formation, was shown to be a significant marker for bone metastasis and poor prognosis [28]. In addition, serum levels of urokinase-type plasminogen activator [29], interleukin 6 [30] and neuron-specific enolase [31] were shown to have prognostic value for prostate cancer. The serum levels of testosterone [32] and hK2 [33] have been evaluated to assess the stage of prostate cancer. hK2, which seems to have a higher serum concentration in men with prostate cancer as compared to men with BPH (see above), also appears to have a higher serum concentration in patients with non-prostate confined cancer as compared to prostate confined cancer [33]. These findings, taken together, show that many changes in addition to high PSA levels occur in the serum of prostate cancer patients. Although no single test is sensitive or specific for all situations, the combination of the markers could yield a test with greatly enhanced diagnostic utility. A summary of serum markers and their potential utility in diagnosing prostate cancer is provided in Table 1.

### 4. The use of combined markers in diagnosis

As noted above, the use of combinations of markers has great potential to improve diagnostic specificity and sensitivity over individual markers. Table 2 summarizes a number of examples of the uses of multiple markers to enhance the diagnosis or prognosis, discussed in more detail below.

A study of prognostic markers for small cell lung cancer found that patients could be classified into four groups of 5-year survival rates based on the combined expression of cyclin E, Ki-67, and ras p21 [34]. Patients with no expression alterations in the three markers had a 96% survival rate, and patients with all three altered had a 41% survival rate. Although each marker individually provided some prognostic information, the three together significantly enhanced the accuracy of risk stratification.

CA-125 has been used for many years as a serum marker for malignant pelvic masses. When taken alone, CA-125 at an abnormal level (> 35 U/ml) gives a fairly high sensitivity and specificity of 78.1% and 76.8%, respectively. However, a significant enhancement in diagnostic performance was shown using a panel of five markers (CA-125, OVX1, LASA, CA15-3, and CA72-4) [35]. When two of these markers were elevated, the sensitivity and specificity increased to 83.3% and 84.0% respectively. These markers were further enhanced using a regression analysis of the values of all five of the markers, giving a sensitivity of 90.6% and a specificity of 93.2%.

The specificity of prostate cancer diagnosis seems to be improved by the use of multiple markers, using the measurements of both the free and bound forms of PSA and the measurement of hK2 (described above). When biopsies were not performed on patients with free/total PSA ratio of > 0.25, there was a 20% reduction in the number of unnecessary biopsies performed [36]. The sensitivity of the test for prostate cancer remained at levels similar to that of PSA alone.
Table 1

Examples of reported prostate cancer serum markers and their potential uses

<table>
<thead>
<tr>
<th>Comments</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detection of prostate cancer or risk</td>
<td></td>
</tr>
<tr>
<td>Prostate specific antigen (PSA)</td>
<td>High sensitivity for prostate cancer, but not specific or prognostic</td>
</tr>
<tr>
<td>Insulin-like growth factor 1 (IGF-1)</td>
<td>Low serum levels correlated with prostate cancer</td>
</tr>
<tr>
<td>25-hydroxyvitamin D (25-VD)</td>
<td>Low serum 25-VD correlated with increased risk for prostate cancer</td>
</tr>
<tr>
<td>Prostatic acid phosphatase (PAP)</td>
<td>The first marker for prostate cancer, supplanted by PSA</td>
</tr>
<tr>
<td>To improve specificity</td>
<td></td>
</tr>
<tr>
<td>Human glandular kallikrein 2 (hK2)</td>
<td>Higher in prostate cancer than BPH, used with free and total PSA</td>
</tr>
<tr>
<td>Keratinocyte growth factor (KGF)</td>
<td>Lower in prostate cancer than BPH</td>
</tr>
<tr>
<td>PSA complexed with alpha(2)-macroglobulin (PSA-A2M)</td>
<td>The percentage of PSA complexed with A2M was lower in prostate cancer than BPH</td>
</tr>
<tr>
<td>Interleukin-8 (IL-8)</td>
<td>Serum IL-8 was higher in prostate cancer than BPH; also may correlate with stage</td>
</tr>
<tr>
<td>Information on grade, stage and prognosis</td>
<td></td>
</tr>
<tr>
<td>Carboxy-terminal propeptide of type I procollagen (PICP)</td>
<td>High PICP correlated with bone metastasis and poor outcome</td>
</tr>
<tr>
<td>Urokinase-type plasminogen activator (uPA)</td>
<td>High uPA correlated with disease progression and poor prognosis</td>
</tr>
<tr>
<td>Interleukin 6 (IL-6)</td>
<td>High IL-6 correlated with stage and poor prognosis</td>
</tr>
<tr>
<td>Neuron-specific enolase (NSE)</td>
<td>Pretreatment-elevated NSE correlated with poor prognosis</td>
</tr>
<tr>
<td>Testosterone</td>
<td>Low free serum testosterone correlated with extent and grade of disease</td>
</tr>
<tr>
<td>Human glandular kallikrein 2 (hK2)</td>
<td>Predicted non-organ confined versus organ confined cancer</td>
</tr>
<tr>
<td>Serum luteinizing hormone (LH)</td>
<td>Low levels correlated with advanced disease</td>
</tr>
<tr>
<td>Bone alkaline phosphatase (BAP)</td>
<td>Correlated with the extent of bone disease</td>
</tr>
<tr>
<td>Matrix metalloproteinase-2 (MMP-2)</td>
<td>Higher in patients with metastasis</td>
</tr>
</tbody>
</table>

Table 2

Examples of investigations of the use of combinations of markers to enhance the diagnostic/prognostic characteristics of single markers

<table>
<thead>
<tr>
<th>Comments</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improve sensitivity/specificity of diagnosis</td>
<td></td>
</tr>
<tr>
<td>Colorectal Carcinoma:</td>
<td></td>
</tr>
<tr>
<td>Creatine kinase-BB, Homoaarginine-sensitive alkaline phosphatase, Salivary-type amylase, Macro-creatine kinase type 2, Ferritin, Alpha 1-acid glycoprotein, C-reactive protein, Alpha 1-antitrypsin, Ceruloplasmin, CEA, and Beta human chorionicadotropin</td>
<td>Improved sensitivity of diagnosis of colorectal carcinoma by 17% and 64% for early- and late-stage disease respectively.</td>
</tr>
<tr>
<td>Ovarian Cancer:</td>
<td></td>
</tr>
<tr>
<td>CA-125, OVX1, LASA, CA15-3, and CA72-4</td>
<td>Improved sensitivity and specificity over CA-125 for the identification of malignant pelvic masses from 78.1% and 76.8%, respectively, to 90.6% and 93.2%, respectively.</td>
</tr>
<tr>
<td>Pancreatic Cancer:</td>
<td></td>
</tr>
<tr>
<td>CA 19-9, CEA</td>
<td>Combined serum levels of CA 19-9 and CEA provided 85.8% sensitivity and 91.2% specificity. This was compared with 30% and 83% sensitivity for the individual markers respectively.</td>
</tr>
<tr>
<td>Improve prognosis:</td>
<td></td>
</tr>
<tr>
<td>Non-Small Cell Lung Cancer:</td>
<td></td>
</tr>
<tr>
<td>Cyclin E, Ki-67, and Ras p21</td>
<td>Segregated patients based on five-year survival patients with 0 altered – 96% survival, all three altered 41% survival.</td>
</tr>
</tbody>
</table>

5. Protein microarrays for highly multiplexed protein detection

A recent study showed the feasibility of sensitive and accurate protein microarray detection of multiple specific antibodies and antigens in a serum background [2]. A robotic device (identical to that used to spot cDNA arrays [37]) was used to print hundreds of specific antibody or antigen solutions in an array on the surface of derivatized microscope slides. Two complex protein samples, one serving as a standard for comparative quantitation, and the other representing an experimental sample in which the protein quantities were to be measured, were labeled by covalent attachment of spectrally-resolvable fluorescent dyes. Specific antibody-antigen interactions localized specific com-
ponents of the complex mixtures to defined cognate spots in the array, where the relative intensity of the fluorescent signal representing the experimental sample and the reference standard provided a measure of each protein’s abundance in the experimental sample.

The specificity, sensitivity and accuracy of the assay were evaluated using 115 antibody/antigen pairs. Six different mixtures of the 115 antibodies and six different mixtures of 115 antigens were prepared so that the concentration of each species varied in a unique pattern across the protein mixtures over a range of three orders of magnitude. Each of the six protein mixtures was labeled with the dye Cy5 (red fluorescence) and mixed with a Cy3-labeled (green fluorescence) “reference” mixture containing each of the same 115 proteins at a constant concentration. The variation across the six microarrays in the red-to-green (R/G) ratio measured for each antibody or antigen spot should reflect the variation in the concentration of the corresponding binding partner in the set of mixes. By comparing the

Fig. 1. Relationship between the Cy5/Cy3 fluorescence ratios measured using antigen microarrays and the concentration ratio of the cognate antibodies. The solid line represents the median of the log10 transformed Cy5/Cy3 fluorescence ratios from 6–12 replicate antigen spots, with the error bars representing the standard deviation between the replicate spots. The dashed line represents the ideal linear relationship between R/G ratio and concentration ratio. (From BB Haab, MJ Dunham, PO Brown. Genome Biology 2:1–13, 2001.)
observed variation in the concentration ratios with the known variation in the concentration ratios, the performance of each antibody/antigen pair could be assayed. Figure 1 presents this relationship for 12 different arrayed antigens detecting their respective cognate antibodies in complex solutions. The dashed line represents the ideal linear response in R/G ratio with respect to analyte concentration, and the solid line is the median log10(R/G) ratio of 6–9 replicate spots, with the error bars representing the standard deviation in the log10(R/G) ratio.

For many of the antigens, the experimental data very closely followed the ideal response (represented by the dashed line). For antigens such as P38 delta, Numb, and AIM-1, the measurements were reproducible and accurate over the entire three orders of magnitude concentration range. These antigens have detection limits of less than 1 ng/mL for their respective antibodies. The ratios measured at replicate spots were highly consistent and exhibited low standard deviations, except in some cases at low concentrations where the dispersion appeared more random (e.g. G3BP and ARNT1).

These data demonstrated accurate and specific quantitation of protein ligands in a complex, physiologically relevant background. The detection limit of the assay depends on the level of background protein binding and on the affinity and specificity of the antibody/antigen interaction. Some of the antibody/antigen pairs allowed detection of the cognate ligands at absolute concentrations below 1 ng/ml, sensitivities sufficient for measurement of many clinically important proteins in patient blood samples. Since many potentially inter-
Fig. 3. Investigation of experimental reproducibility and the accuracy in reverse labeling of protein microarray experiments. The R/G ratio of each antibody spot in experiment (a) of Fig. 2 was plotted with respect to the R/G ratio of its counterpart spot in experiment (b) of Fig. 2. The diagonal line (not fitted to the data) indicates the ideal inverse relationship between the “color swapped” experiments.

The data in Fig. 3 indicate that, with the exception of a few antibody spots, reverse labeling of protein microarray experiments is accurate and reproducible. The diagonal line represents the ideal inverse relationship between the R/G ratios of the two experiments. The data points close to the diagonal line indicate that reverse labeling is accurate, while data points far from the diagonal line indicate inaccuracies in labeling.

Detecting proteins that have serum concentrations below that level will be important to further improve the sensitivity of the protein microarray assay. Approaches to increase detection sensitivity include using surfaces that are more resistant to background protein binding, selecting antibodies that have optimized affinities, and amplifying the fluorescent signal of the bound antibody. Rolling circle amplification (RCA) is a technique that has been shown to lower detection limits in solid phase immunoassays by over 100-fold [38]. Application of RCA to protein microarrays, along with the use of optimized antibodies and surfaces, should allow the quantification of many low abundance and potentially important serum proteins.

6. Highly parallel serum protein detection using protein microarrays

Preliminary studies are under way in our laboratory to apply protein microarray technology to prostate cancer serum marker discovery. Prostate cancer serum samples with characterized PSA concentrations provide a positive control, and other serum proteins that have been studied in connection with prostate cancer give many good leads for testing the utility of highly multiplexed marker detection. Over 200 antibodies to putative serum markers, known serum proteins and known cancer genes were collected through collaborations and printed in arrays on poly-l-lysine derivatized microscope slides. As a reference protein solution, 34 prostate cancer serum samples and 20 serum samples from healthy patients were pooled together. The individual serum samples were then fluorescently labeled and co-incubated on the microarrays with the differentially labeled reference.

Initial experiments have confirmed multiplexed specific detection of proteins in the serum samples. Figure 2 presents two microarrays in which the fluorescent labeling of a prostate cancer sample and the reference pool were swapped. In panel (a), the prostate cancer serum and the reference pool were labeled with Cy5 (red fluorescence) and Cy3 (green fluorescence), respectively, and in panel (b) the labeling is reversed. Many antibody spots have fluorescence clearly above background in red, yellow and green colors. The expanded regions of the images in the lower panels of Fig. 2 show that many of the antibody spots that are primarily red in one array are primarily green in the
other, consistent with reproducible and specific labeling. Significantly, the PSA antibody appears more red when the prostate cancer serum was labeled red, and more green when the same sample was labeled green. The analysis of additional prostate cancer serum samples is needed to confirm quantitative detection of this protein.

To quantitatively compare the color swapped experiments, the R/G ratios of matched spots on the two arrays were plotted with respect to each other (Fig. 3). The solid diagonal line represents the ideal inverse relationship. The general trend of the R/G ratios follows the solid line, showing that for most of these spots, the labeling and detection was consistent and reproducible. The scatter around the solid line reflects the level of noise inherent in the measurement, which can be evaluated to determine thresholds for accepting or rejecting spots for further analysis. Some of the spots fell well outside the scatter on the inverse diagonal, such as those in the lower left of the figure, and would be rejected from further analysis.

These data demonstrate high signal-to-noise and reproducible detection of multiple proteins in human cancer serum. After a large set of prostate cancer sera samples have been analyzed, the aim of the analysis will be to identify patterns of proteins that significantly correlate with particular clinical parameters. Methods developed for the analysis of RNA expression profiles from cDNA microarrays will be applied to the data. For example, individual proteins that distinguish two sample sets (such as BPH versus prostate cancer) could be identified using a permutation t-test [39], and patterns of proteins that make the same distinction could be identified using methods such as ‘tree harvesting’ [40] or a ‘cluster identification tool’ [41]. With the right antibodies on the arrays, previous demonstrations of the value of multiple markers and power of microarrays indicate that significant advances in serum marker discovery and validation should be achievable with this new tool.

Acknowledgements

J.C.M. is funded through a Van Andel Institute fellowship. We thank the Van Andel Institute for funding and the staff of the Van Andel Institute for helpful interactions.

References


Submit your manuscripts at http://www.hindawi.com