Published in final edited form as:

J Proteome Res. 2009 February 6; 8(2): 483–492. doi:10.1021/pr8007013.

Pancreatic Cancer Serum Detection Using A Lectin/Glyco-Antibody Array Method

Chen Li¹, Diane M Simeone², Dean E. Brenner³, Michelle A. Anderson³, Kerby A Shedden⁴, Mack T. Ruffin⁵, and David M. Lubman^{1,2,6}

¹ Department of Chemistry, The University of Michigan, Ann Arbor, MI 48109

² Department of Surgery, The University of Michigan Medical Center, Ann Arbor, MI

³ Department of Internal Medicine, The University of Michigan Medical Center, Ann Arbor, MI

⁴ Department of Statistics, The University of Michigan, Ann Arbor, MI, 48109

⁵ Department of Family Medicine, The University of Michigan Medical Center, Ann Arbor, MI

⁶ Comprehensive Cancer Center, The University of Michigan, Ann Arbor, MI

Abstract

Pancreatic cancer is a formidable disease and early detection biomarkers are needed to make inroads into improving the outcomes in these patients. In this work lectin antibody microarrays were utilized to detect unique glycosylation patterns of proteins from serum. Antibodies to four potential glycoprotein markers that were found in previous studies were printed on nitrocellulose coated glass slides and these microarrays were hybridized against patient serum to extract the target glycoproteins. Lectins were then used to detect different glycan structural units on the captured glycoproteins in a sandwich assay format. The biotinylated lectins used to assess differential glycosylation patterns were Aleuria aurentia lectin (AAL), Sambucus nigra bark lectin (SNA), Maackia amurensis lectin II (MAL), Lens culinaris agglutinin (LCA), and Concanavalin A (ConA). Captured glycoproteins were evaluated on the microarray in situ by on-plate digestion and direct analysis using MALDI QIT-TOF mass spectroscopy. Analysis was performed using serum from 89 normal controls, 35 chronic pancreatitis samples, 37 diabetic samples and 22 pancreatic cancer samples. We found that this method had excellent reproducibility as measured by the signal deviation of control blocks as on-slide standard and 41 pairs of pure technical replicates. It was possible to discriminate cancer from the other disease groups and normal samples with high sensitivity and specificity where the response of Alpha-1- β glycoprotein to lectin SNA increased by 69% in the cancer sample compared to the other non-cancer groups (95% confidence interval 53% to 86%). These data suggest that differential glycosylation patterns detected on high throughput lectin microarrays are a promising biomarker approach for the early detection of pancreatic cancer.

Keywords

Glycoproteins; Pancreatic cancer; Lectins; Antibody Array; Cancer Markers

Introduction

Pancreatic cancer continues to have a high mortality rate due to detection at a late stage of the disease[1]. In fact, 85% of patients initially present with advanced, non-resectable disease, highlighting the importance of identifying early detection biomarkers. In addition, in a subset of patients, it may be quite difficult to distinguish chronic pancreatitis and pancreatic cancer, necessitating unnecessary surgery in some patients that otherwise might not require it if an adequate biomarker to distinguish these two diseases was available. A serum biomarker test is expected to improve the efficiency of the diagnosis, where the blood contains the unique secretome of the tumor cells. Several serum markers have been investigated for pancreatic cancer. Elevated CA19-9 level has been cited as a potential marker of disease although it generally does not have the specificity or sensitivity for general screening[2-8]. It has been frequently utilized as a marker to monitor a patient's progress after surgery[9]. Other existing biomarkers relate to the inflammation that associates with the tumor and other pancreatic diseases that may be present[10-12]. It should be noted that no individual biomarker has been found to be conclusive at diagnosis to distinguish chronic pancreatitis and pancreatic cancer. [13,14] To our knowledge, there is no study comparing the serum of pancreatic cancer and diabetes which is a widely existing disease in patients at risk of pancreatic cancer. Discovery of new early detection biomarkers that are specific for pancreatic cancer remains a major challenge.

Post translational modification of the proteome in serum analysis has become an important area in biomarker research[15]. Of particular interest is the study of glycoproteins where unique protein glycosylation patterns are associated with cancer[16-25]. Glycans are involved in many biological processes including protein-protein interactions, protein folding, immune recognition, cell adhesion and inter-cellular signaling[26]. Alteration of glycan structure and coverage on several major glycoproteins in serum has been shown to contribute to the progression of cancer. In previous work, fucosylated haptoglobin was suggested as a biomarker for early detection of pancreatic cancer[27]. Also the glycoforms of alpha-1-acid glycoprotein have been found to vary in cancer patients compared to the healthy controls[28]. These biomarkers can be used to improve the confidence of the diagnosis through identification of disease-related glycan structures by various separation and mass spectrometry techniques [29-32]. In one such study using lectin extraction and mass spec analysis the glycosylated isoforms of alpha-antitrypsin were shown to change in cancer compared to normal samples or pancreatitis[33]. Other studies have removed the glycan groups from the glycoprotein content of the cell and used glycan profiling to show distinct differences between cancer and normal samples based on changes in carbohydrate structures in serum, although association with a particular protein is lost[34]. In other studies hydrazide columns have been used to extract glycoproteins from serum which were digested and analyzed by LCMS/MS. In this report glycoproteins associated with cancer were found although the actual glycan structural information was not delineated[35].

Recently, various microarray formats have been utilized for studying glycosylation patterns. In one study examining sera samples from patients with colon and pancreatic cancers, glycoproteins extracted from serum were printed on glass slides and hybridized against various lectins to study changes in the glycan patterns during cancer progression[36,37]. This method provides a means of studying subtle changes in glycan structure and is an excellent discovery platform but does not provide a high throughput mode for further validation. Other methods have included the use of glycan arrays where glycans are directly printed on glass slides[38] or alternatively lectin arrays where lectins are printed on a slide and glycoproteins or whole cells hybridized against them. The lectin array approach has been used to identify differences in glycoprotein surface markers for cancer cells compared to normal cells and between different types and stages of cancer in several studies[39,40]. Alternatively an antibody array approach

has been used to capture proteins from serum and a lectin hybridized against the glycoprotein to study changes in glycan structure[41]. This method can screen large numbers of samples from serum for such changes but requires a discovery platform to choose the antibodies on the array for screening.

The antibody microarray is a favorable format for high throughput analysis, with a high level of specificity and reproducibility[42-44]. In the present study, we selected four glycoproteins as our target proteins for the antibody microarray based on our previous work[36,45]. Antibodies to these glycoprotein biomarkers were printed on nitrocellulose coated glass slides. The glycans on the printed antibodies were first blocked to eliminate their interference in the hybridization with lectins. The target proteins in the serum were then captured on the antibody array and probed with several biotinylated lectins where streptavidinylated fluorescent dyes were used for detection. Eighty nine samples from normal controls, 35 chronic pancreatitis samples, 37 diabetics samples and 22 pancreatic cancer samples were processed using this method where non-cancer samples were randomly selected and all cancer samples available were used. Antibody specificity was verified by on-target digestion of the captured glycoproteins with subsequent on-slide MALDI-MS identification. The data was subjected to statistical analysis to decompose the observed variation into effects attributable to disease groups, experimental blocks, patient heterogeneity, and technical noise.

2. Experimental

Sera

Inclusion criteria for the study included patients with a confirmed diagnosis of pancreatic cancer, chronic pancreatitis, long-term (for 10 or more years) Type II diabetes mellitus, or healthy adults with the ability to provide written, informed consent, and provide 40 ml of blood. Exclusion criteria included inability to provide informed consent, patients' actively undergoing chemotherapy or radiation therapy for pancreatic cancer, and patients with other malignancies diagnosed or treated within the last 5 years. The sera samples were obtained from patients with a confirmed diagnosis of pancreatic adenocarcinoma who were seen in the Multidisciplinary Pancreatic Tumor Clinic at the University of Michigan Comprehensive Cancer Center. All cancer sera samples used in this study were obtained from patients with stages III/IV pancreatic cancer. The mean age of the tumor group was 65.4 years (range 54-74 years). The sera from the normal, pancreatitis, and diabetes groups was age and sex-matched to the tumor group. The chronic pancreatitis group was sampled when there were no symptoms of acute flare of their disease. All sera were processed using identical procedures. The samples were permitted to sit at room temperature for a minimum of 30 minutes (and a maximum of 60 minutes) to allow the clot to form in the red top tubes, and then centrifuged at $1,300 \times g$ at 4°C for 20 minutes. The serum was removed, transferred to polypropylene, capped tubes in 1 ml aliquots, and frozen. The frozen samples were stored at -70°C until assayed. All serum samples were labeled with a unique identifier to protect the confidentiality of the patient. None of the samples were thawed more than twice before analysis. This study was approved by the Institutional Review Board for the University of Michigan Medical School.

Microarray preparation and serum hybridization

Alpha-1- β glycoprotein antibody was purchased from Novus, while Amyloid p component antibody and Antithrombin antibody were from Abcam. All the antibodies are monoclonal, raised in mice, targeting human proteins. Antibodies were diluted to 50 $\mu\text{g}/\text{mL}$ in PBS and spotted on ultra-thin nitrocellulose coated slides (PATH slides, GenTel Bioscience) with a piezoelectric non-contact printer (Nano plotter; GESIM). Each spotting event that resulted in 500 pL of sample being deposited was programmed to occur 5 times/spot to ensure that 2.5 nL was being spotted per sample. The spots used by the MALDI-MS experiment were printed 50

times. Each antibody was printed in triplicate. The spot diameters were 280 μm and 700 μm for the spots that were printed 5 times and 50 times respectively. The spacing between the spots was 0.7mm. 14 blocks were printed on each slide in a 2 $\times$ 7 format and the block distance was 9.4mm.

Figure 1 presents an experimental flow chart of the microarray processing and on-target digestion for MALDI-MS. The antibody arrays on the slides were first chemically derivatized with a method similar to previous work [22] but modified for this work. The printed slides were dried in an oven at 30°C for 1h before gently being washed with PBST 0.1 (100% PBS with 0.1% tween 20) and coupling buffer (0.02M sodium acetate, pH 5.5), and then oxidized by 200 mM NaIO₄(Sigma) solution at 4°C in the dark. After 3 hours the slides were removed from the oxidizing solution and rinsed with coupling buffer. The slides were immersed in 1 mM 4-(4-N-maleimidophenyl)butyric acid hydrazide hydrochloride (MPBH; Pierce Biotechnology) at room temperature for 2 hours to derivatize the carbonyl groups. 1 mM Cys-Gly dipeptide(Sigma) was incubated with the antibodies on the slides at 4°C overnight. The slides were blocked with 1% BSA for 1 hour and dried by centrifugation.

The slides were inserted into the SIMplex (Gentel) 16 Multi-Array device which separates the blocks and prevents cross contamination when different samples are applied on neighboring wells. Serum samples were diluted 10 times with PBST 0.1 containing 0.1% Brij. 100 μL of each sample was applied to the antibody array manually and left in a humidified chamber for 1 hour to prevent evaporation. Slides were rinsed with PBST 0.1 for 3 times to remove unbound proteins. The arrays were then treated with different detection biotinylated lectins (Vector Laboratory) to determine lectin response and streptavidinylated fluorescent dye (Alexa555; Invitrogen Biotechnology) was used for detection. After a final wash, the slides were dried and scanned with a microarray scanner (Genepix 4000A; Axon). The program Genepix Pro 6.0 was used to extract the numerical data. A threshold of signal to background ratio was set at 10 and less than 1% of the spots were under this threshold and excluded. The mean of the intensity in each spot was taken as a single data point into analysis.

On-target digestion and MALDI-QIT-TOF

The microarray slides were incubated with 0.5 M lactose for 10 min and washed with PBST 0.1 to remove the captured lectin from the glycoprotein. After an additional wash with water the slides were dried with centrifugation. Trypsin was diluted to 50 ng/ μL with 50 mM ammonium bicarbonate and printed on the microarray spots. The printed slides were moved into a humidity chamber and incubated at 37° C for 5min. 35 mg/mL 2,5-dihydroxybenzoic acid (DHB) (LaserBio Labs, France) in 50% acetonitrile was printed on the microarray by the microarray printer and allowed to dry.

Mass spectrometric analysis of the microarray slides was performed using the Axima quadrupole ion trap-TOF (MALDI-QIT) (Shimadzu Biotech, Manchester, UK). The microarray slide was analyzed directly by taping the slide onto the stainless steel MALDI plate and inserting it into the instrument for analysis. Acquisition and data processing were controlled by Launchpad software (Kratos, Manchester, UK). A pulsed N₂ laser light (337 nm) with a pulse rate of 5 Hz was used for ionization. Each profile resulted from 2 laser shots. Argon was used as the collision gas for CID and helium was used for cooling the trapped ions. TOF was externally calibrated using 500 fmol/ μL of bradykinin fragment 1-7 (757.40 m/z), angiotensin II (1046.54 m/z), P14R (1533.86 m/z), and ACTH (2465.20 m/z) (Sigma-Aldrich). The mass accuracy of the measurement under these conditions was 50 ppm.

Results and Discussion

Microarray printing and processing

The antibodies were printed on ultrathin nitrocellulose slides and hybridized with serum in a 14 multi-array device, then visualized with biotinylated lectin and Alexafluor-555. In a reproducibility test, a common sample selected at random was applied to all 14 arrays. Lectin SNA was used for detection. The fluorescent image of the slide is shown in Figure 2a to illustrate the quality of the printed spots. There is a very small variation of the signal between different arrays over the slides. The intensity of the signal in each block was calculated as shown in Figure 2b. The standard deviations of the signal of all three antibodies within the slides were about 5% of the average. In order to estimate the variation of the signal on different slides, 2 blocks on each slide were hybridized with the same two samples. The signals of these two blocks were compared across slides to calculate the variation. In multiple-slide experiments, the value of the slide to slide variation was about 10% of the average signal.

Different dilutions of serum were tested to determine the optimum concentration of the target glycoproteins. The same lectin SNA was used for detection in this test. There were seven dilutions of serum sample from 2 to 600 times dilution that were applied to the arrays. Figure 3 depicts how the intensity of the signal changes for the three antibodies with a decreasing dilution fold. A rising trend was noted from the 600X dilution to the 50X dilution for the three glycoproteins shown. In the 50X dilution to the 20X dilution the signal was relatively unchanged except for Antithrombin-III, where the signal increased 20% from the 50X dilution to the 20X. The signal remained the same from the 20X dilution until it reached the 5X dilution, where a saturation of the signal has occurred. A decrease of signal for all three glycoproteins from the 5X dilution to the 2X dilution of serum sample can be seen in the figure 3, likely due to competing non-specific binding on the antibodies.

The result of the dilution test demonstrated that the antibodies were saturated by their target protein at 20X dilution or above in the process of hybridization (1hour, room temperature and gentle shaking). Below 50X dilution the antibodies were not completely occupied, so the signal decreased with additional dilution. The nonlinear relation between the concentration of the serum and the intensity of the signal could be attributed to various factors that may affect the antibody-antigen reaction, including accessibility of the antibodies, diffusion rate and solubility of the antigen in the hybridization buffer. Nonspecific binding on the antibodies was also considered as a possibility, but was further investigated and excluded by on-target digestion and MALDI-MS analysis.

To analyze the difference of the glycosylation on potential biomarker proteins, protein expression levels must be normalized. The protein level was estimated by antibody assay. In our experiment the three potential biomarkers were all relatively high abundance proteins in human serum (concentration >20mg/L) which could easily saturate the antibodies printed on the microarray. Under saturation conditions, the amount of target biomarkers captured on the antibody spots was equal to the capacity of the printed antibody which should be the same in all the replicate blocks. As a result, the need for protein assay was avoided and the intensity of the signal on the microarray directly represented the level of glycosylation.

Antibody specificity test with MALDI-QIT-TOF

In order to validate the specificity of the antibodies we performed on-target digestion and MALDI-QIT-TOF of the spots after elution of biotinylated lectins captured on the glycoproteins with a concentrated sugar solution. A trypsin solution with 50mM ammonium bicarbonate was printed with the microarray printer using the same spot layout as in the antibody printing. The volume of the trypsin solution was 4nL which in a humidity chamber

lasts about 5 minutes before drying out. Ammonium bicarbonate usually decomposes at the same time. 2,5-dihydroxybenzoic acid was then dissolved in 50% acetonitrile and printed on the digested spots. The matrix solution itself is very acidic and stops the digestion to prevent further digestion of antibodies and trypsin autolysis. Acetonitrile also partially dissolved the nitrocellulose film and the digested peptides on the film were extracted and mixed with matrix. Nitrocellulose film has been reported as an excellent substrate for MALDI-MS[46]. The presence of nitrocellulose in the mixture did not affect the crystallization of DHB.

The specificity (specific binding vs. non-specific binding) of the antibody as a function of the dilution times of the serum can be determined by comparing the spectrum from the arrays processed with different conditions. In the experiment one control array (incubated with blocking buffer) and two sample arrays which were hybridized with 2X and 10X dilution of the same serum were tested. The presented figures are the spectra of Amyloid p component antibody spot. Figure 4a shows the spectrum of the Amyloid p component spot in the control array which only contained the antibody(anti human Amyloid p component). All the peaks in the spectrum are the peptides digested from the antibody and the enzyme itself. The top 3 peaks are attributed to the antibody digest. The intensity of the other peaks was too low to be identified. The spectra in figures 4b and 4c are generated from the Amyloid p component spots in the sample arrays. In the mass spectrum of 10X dilution, 3 new peaks appeared which were all identified by MS/MS to be tryptic peptides of Amyloid p component. This result indicated that no other protein was captured on the antibody or the amount was too low to be detected. In the case of the 2X dilution, 2 additional peaks emerged in the spectrum where one of them was identified as human serum albumin while the other one was not identified. The extra peaks are a sign of nonspecific binding on the antibody. Thus, only when the concentration of the sample was increased to 2X the dilution of the serum does non-specific binding begin to affect the specificity of the antibody.

Detecting glycosylation on captured protein by blocked antibody arrays

The chemical derivatization method was employed to block the glycans on the antibodies to eliminate their binding with the lectins used for detection of glycoproteins[41]. The cis-diol groups on the glycans were gently oxidized and converted to aldehyde groups which then reacted with hydrazide-maleimide bifunctional cross-linking reagent and capped with a Cys-Gly dipeptide. After the derivatization reaction the lectins could not recognize the modified oligosaccharide group.

All the antibodies were tested against several samples and lectins to evaluate the effectiveness of the protocol. The underivatized antibodies responded to some of the lectins, but after derivatization the binding greatly decreased or disappeared. The serum solution was incubated against the derivatized antibody array where the spots showed lectin binding on proteins captured by the antibodies, indicating that the antibodies maintained their function after derivatization.

A previous study described ten potential biomarkers in the sera of normal and cancer patients that significantly changed their response to several lectins[18]. We chose four of these target proteins (Antithrombin-III, Amyloid p component, alpha-1- β glycoprotein and kininogen) based on their response in previous work as a proof of concept to determine the proteins which provided the best discrimination of samples from patients in different groups based on lectin response. The biotinylated lectins used were Aleuria aurentia lectin (AAL), Sambucus nigra bark lectin (SNA), Maackia amurensis lectin II (MAL), Lens culinaris agglutinin (LCA), and Concanavalin A (ConA). AAL and LCA bind fucose linked (α -1,6) to N-acetylglucosamine or (α -1,3) to N-acetyllactosamine related structures. Both MAL and SNA recognize sialic acid on the terminal branches. MAL detects glycans containing NeuAc-Gal-GlcNAc with sialic acid at the 3 position of galactose while SNA binds preferentially to sialic acid attached to

terminal galactose in an (α -2,6) and an (α -2,3) linkage to a lesser degree. ConA recognizes α -linked mannose including high-mannose-type and hybrid-type structures. These lectins were selected since fucosylation and sialylation have been shown to be related to cancer development [27,33] and ConA binds to almost all the N-linked glycoproteins where its signal translates into a general level of glycosylation.

An initial test of the four antibodies against five lectins was performed (see fig1 supplementary information). The lectins used were AAL, LCA, SNA, MAL and ConA. For LCA, AAL, SNA and MAL the three cancer samples all showed a stronger response than the pancreatitis and normal samples, whereas for ConA there was equal signal in the three groups. A binding pattern was shared between LCA and AAL, which agreed with their same specificity on fucosylated N-linked glycans, though the signal of LCA was lower in intensity. These lectins were found to preferentially distinguish normal and chronic pancreatitis samples from cancer samples (data not shown). MAL was not used for subsequent analysis due to its low sensitivity with these antibodies. Of the 4 antibodies, 3 of them (A1BG, Amyloid p component and Antithrombin-III) displayed a signal-to-background ratio of higher than 20, and were chosen for large set analysis.

High throughput analysis and data quality test

183 samples from patients with various genders, fasting status and disease classes were processed in 4 batches on 16 slides. Since the signal to background ratio for all the valid spots were higher than 10, the signals were directly used for analysis without taking into account the background. 41 of the patients in the groups of normal, chronic pancreatitis and diabetics contributed three samples with two samples collected twice under fasting conditions and the other sample was collected under non-fasting conditions. The lectin used in this study is SNA. Two patients provided only double fasting samples which are used for the data quality test. For the other samples including some of the normal, pancreatitis and all the cancer patients, the information of the gender and fasting status is not available. After adjusting for fasting status, gender, and disease category, the data points were compared to a normal reference distribution (figure not shown). Based on this comparison, two outlying data points from the antibody of Antithrombin-III were excluded from all subsequent analysis.

The accuracy of the antibody microarray analysis is heavily dependent on the reproducibility of the technique which is also used as a means to filter out unreliable antibodies in distinguishing cancer from other disease classes. Reproducibility is assessed by fitting a linear mixed effects model to log2 scale expression data, separately for each antibody. This is a type of ANOVA model in which fixed effects for fasting status, gender, and disease category are included along with random effects for patients, and batches within patients. Thus the expression variation for every antibody around the mean for its fasting/gender/disease group is described in terms of three variance components (residual, patient and batch within patient), which characterize the reproducibility at different levels of the experiment. Residual variance represents variation for technical replicates (same person, batch, and fasting status). Batch variance represents technical variation for the same person and fasting status across batches. Patient variance represents stable biological variation across people. Table 1 shows the three variance components on the standard deviation scale, for the three antibodies. For example, the residual SD for A1BG is 0.21, which means that two thirds of the replicates will lie within $(2^{0.21}-1) \times 100\% = 16\%$ of the true values and 95% of the replicates will lie within $(2^{0.42}-1) \times 100\% = 34\%$ of the true value. Alternatively, the reproducibility could be exhibited by the correlation of the replicate spots in log2 scale which is presented in Figure 5. The scatterplots suggest that the technical error is not limited to a handful of outliers, consistent with our finding of an approximately normal distribution of residual variance, as discussed above. Figure 5 shows data for all non-cancer patients and antibodies pooled.

Examination of potential bias

Sex, fasting status and other related diseases could all possibly become the source of bias in biomarker validation.[47] As discussed above, linear mixed effects models were built separately for each of the three antibodies, with these potentially biasing factors modeled as fixed effects. As always, one level of each factor variable is omitted, so the implicit fixed effect for a normal, non-fasting female is zero, and all other factor settings are interpreted as deviations from this arbitrary baseline setting. The results are listed in Table 2. Fixed effect estimates are shown along with the standard error. The T-value is the ratio of the effect estimate to the standard error. Since log2 scale data are analyzed, we also converted the effects to the raw scale, given in terms of percent change relative to the baseline category. For A1BG the factors have small and non-significant effects. For Amyloid there is a significant effect for fasting, and for Antithrombin-III there is a significant effect for sex and disease (mainly due to pancreatitis). These effects are statistically significant but are small in magnitude relative to the residual and patient variation, and to the response in cancer.

Antibody performance in distinguishing cancer and non-cancers

Table 3 provides information concerning the discrimination between cancer and non-cancer samples. These estimates are obtained through the same linear mixed effects model discussed above. A fixed effect estimate b on the log2 scale implies a percent change of $(2^b - 1) \times 100\%$ on the raw scale, where negative and positive percent changes correspond to decreased, and increased expression, respectively. The A1BG signal increases by 69% $\% = (2^{0.76} - 1) \times 100\%$ in cancer samples compared to normal, chronic pancreatitis, and diabetic samples. The Amyloid signal increases 33%, and Antithrombin-III is essentially unchanged. The total standard deviation from technical and biological variation (within disease classes) is around $0.32 = \sqrt{0.216^2 + 0.207^2 + 0.127^2}$ for A1BG, summing the variance components from table 1. Thus approximately 95% of the time, the experimental result for a particular non-cancer sample is expected to fall between 35% below the mean level or 56% above the mean level for all non-cancer samples $((2^{-0.64} - 1) \times 100 = -35$ and $(2^{0.64} - 1) \times 100 = 56$, where $0.64 = 2 \times 0.32$ is two standard deviations. Thus the 69% average increase in A1BG associated with pancreatic cancer falls well outside the range of normal variation. ROC curves in Figure 6 were also constructed for each of the three markers, based on their ability to distinguish pancreatic cancer from non-cancer samples (a pool of normals, pancreatitis, and diabetes). All three markers show some discrimination where only A1BG is potentially useful on its own. A1BG distinguished cancer and non-cancer samples with a 100% sensitivity and a 98% specificity. The AUC value measuring the area under the ROC curve for A1BG is 0.998. For Amyloid p component the cancer samples were distinguished from non-cancer samples with an 88% sensitivity and a 68% specificity and its AUC value is 0.875. The discrimination for Antithrombin-III is due to the differences between cancer and pancreatitis and it would be unable to distinguish cancer from diabetes based on these data. According to the scatter plot in Figure 7 where the signals of A1BG and Amyloid were used as X and Y axes, the overlap of the cancer samples with the non-cancer groups is around 20%. The extent of the difference in 4 patient groups is also shown in Figure 8 which depicts the distribution of the measurement for the antibody A1BG. The boxplot provides the upper and lower quartiles of the measurements with respect to the median value (red line in the middle of each box). The lines provide the ranges of the measurements, excluding outliers (+).

The use of the antibody microarray to capture potential biomarkers available in cancer serum provides a means for high throughput and analysis of glycosylation patterns. Because of the specific goal of quantifying the glycans in this study, antibodies were saturated with the analytes by optimizing the dilution times of the sera according to the saturation curve. Thus the response of the lectin from the microarray directly represented the level of the particular glycosylation without concern about the various concentrations of the proteins in different

samples. This strategy also defined the sensitive steps in the experiment where the serum was aliquoted, diluted and hybridized with the microarray, while in other applications of antibody microarrays, factors such as precipitation, heterogeneity of the serum and conditions in hybridization may vary and lead to bias in the method.

In this study antibody specificity was confirmed by direct MALDI-MS of the microarray spots. Traditional immunoblotting is based on the same interaction as in the antibody microarray and does not exclude undesirable binding. MALDI-MS can identify the tryptic peptides of any captured abundant protein on the target. The microarray printer was essential in precisely depositing the extremely small amount of enzyme and matrix on top of the antibody spots [48]. In this experiment, the nitrocellulose surface generated high quality mass spectra. In spite of peaks from the antibody that dominated the mass spectra, target proteins were readily identified and non-specific binding was also found when the serum was not sufficiently diluted.

To access the technical error of the assay, a comprehensive reproducibility test was applied by using two fasting samples from the same patients (drawn at two times) as pure technical replicates. The samples were disordered before being incubated on the antibody arrays. In most other duplicate studies[42-45], variations from an entire slide or batch was more likely to be detected while the individual variability of the single blocks within the slide and batch were ignored. In this work, by statistically comparing the pairs of technical replicates that were distributed across the slides, we were able to evaluate the divergence of the signal from the ideal value that resulted from the technical error.

The pancreatic cancer samples could be clearly distinguished from other disease states and normal samples. The ROC curves showed that Alpha-1- β glycoprotein response to SNA resulted in specific detection of pancreatic cancer with high sensitivity and specificity. A combined ROC curve of Alpha-1- β glycoprotein and amyloid did not provide any improvement in discrimination. However, a limitation of the analysis was that the cancer samples studies were all from late stage pancreatic cancer, thus the study served as a proof of principle experiment. To have an impact on the morbidity and mortality associated with pancreatic cancer, we will need to examine an adequately sized test and training set of early stage pancreatic cancer samples and compare results with appropriate age and sex matched normals and disease controls (i.e chronic pancreatitis and diabetics). Such a serum set is currently being collected as part of a multi-institution collaborative effort by the Early Detection Research Network (EDRN).

Conclusion

In this work, an antibody/glycoprotein/lectin sandwich assay was developed for screening potential markers of pancreatic cancer. These markers were chosen for study based upon our previous work using a lectin glycoarray approach. Three potential markers were chosen and their corresponding antibodies were printed on coated glass slides. They were exposed to sera from 89 normal samples, 35 chronic pancreatitis samples, 37 diabetic samples and 22 pancreatic cancer samples. The captured glycoproteins were analyzed against four different lectins where SNA was found to provide the best results.

Further, MALDI QIT-TOF MS was used for direct analysis of the captured glycoproteins to optimize dilution conditions of the serum and for minimizing nonspecific binding. It was shown that the pancreatic cancer samples could be clearly distinguished from other disease states and normal samples. The ROC curves showed that Alpha-1- β glycoprotein response to SNA resulted in specific detection of pancreatic cancer with high sensitivity and specificity. The resulting scatterplots also showed the ability to clearly distinguish pancreatic cancer from chronic pancreatitis, diabetics or normal samples. The protein Amyloid also showed an acceptable ability to discriminate pancreatic cancer according to the ROC curve whereas

Antithrombin-III could not provide such discrimination. A combined ROC curve of Alpha-1- β glycoprotein and Amyloid did not provide any improvement in discrimination due to correlation between the two markers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

We would like to thank Brian Haab and Dr. Chen Songming of the Van Andel Institute for helpful suggestions on preparation of the antibody arrays and procedures for blocking the antibody carbohydrate groups. We would also like to thank Stephanie Laurinec, Jes Pedroza and Missy Tuck for collection of the samples used in this work.

Credit: This work was supported in part by the National Cancer Institute under grants 1R21CA124441 (DML), R01 CA106402(DML), National Institutes of Health under grant RO1GM49500(DML), the Great Lakes New England Clinical Epidemiology and Validation Center of the EDNRN under CA86400(DEB) and the CCOP under CA74648 (DEB).

References

1. Jemal A, Siegel R, Ward E, Murray T, et al. CA Cancer J. Clin 2006;56:106–130. [PubMed: 16514137]
2. Mann DV, Edwards R, Ho S, Lau WY, Glazer G. Eur. J. Surg. Oncol 2000;26:474–479. [PubMed: 11016469]
3. Ferrone CR, Finkelstein DM, Thayer SP, Muzikansky A, Fernandez-del Castillo C, Warshaw AL. J. Clin. Oncol 2006;24:2897–2902. [PubMed: 16782929]
4. Duffy MJ. Ann. Clin. Biochem 1998;35(Pt 3):364–370. [PubMed: 9635101]
5. Boeck S, Stieber P, Holdenrieder S, Wilkowski R, Heinemann V. Oncology 2006;70:255–264. [PubMed: 16899980]
6. Dalglish AG. British Medical Journal 2000;321:380.
7. Chang CY, Huang SP, Chiu HM, Lee YC, Chen MF, Lin JT. Hepatogastroenterology 2006;53:1–4. [PubMed: 16506366]
8. Kim JE, Lee KT, Lee JK, Paik SW, Rhee JC, Choi. KWJ. Gastroenterol. Hepatol 2004;19:182–186.
9. Riker A, Libutti SK, Bartlett DL. Surg. Oncol 1998;6
10. Wigmore SJ, Fearon KC, Sangster K, Maingay JP, Garden OJ, Ross JA. Int. J. Oncol 2002;21:881–886. [PubMed: 12239630]
11. Fearon KC, Barber MD, Falconer JS, McMillan DC, Ross JA, Preston T. World J. Surg 1999;23:584–588. [PubMed: 10227928]
12. Dube DH, Bertozzi CR. Nat. Rev. Drug Discov 2005;4:477–488. [PubMed: 15931257]
13. Garcea G, Neal CP, Pattenden CJ, Steward WP, Berry DP. Eur. J. Cancer 2005;41:2213–2236. [PubMed: 16146690]
14. Rustgi AK. Gastroenterology 2005;129:1344–1347. [PubMed: 16230087]
15. VanMeter A, Signore M, Pierobon M, Espina V, Liotta LA, Petricoin EF. Expert Review of Molecular Diagnostics 2007;5:625–633. [PubMed: 17892368]
16. An HJ, Peavy TR, Hedrick JL, Lebrilla CB. Anal. Chem 2003;75:5628–5637. [PubMed: 14710847]
17. Block TM, Comunale MA, Lowman M, Steel LF, Romano PR, Fimmel C, Tennant BC, London WT, Evans AA, Blumberg BS, Dwek RA, Mattu TS, Mehta AS. Proc. Natl. Acad. Sci. U. S. A 2005;102:779–784. [PubMed: 15642945]
18. Ransom HW, Varghese RS, Goldman L, An Y, Loffredo CA, Kyselova, Mechref Y, Novotny M, Drake SK, Goldman R. Journal of Proteome Research 2008;7:603–610. [PubMed: 18189345]
19. Gessner P, Riedl S, Quentmaier A, Kemmner W. Cancer Lett 1993;75:143–149. [PubMed: 8313349]
20. Gorelik E, Galili U, Raz A. Cancer Metastasis Rev 2001;20:245–277. [PubMed: 12085965]
21. Morelle W, Michalski JC. Curr. Pharm. Des 2005;11:2615–2645. [PubMed: 16101462]

22. Peracaula R, Tabares G, Royle L, Harvey DJ, Dwek RA, Rudd PM, de Llorens R. *Glycobiology* 2003;13:457–470. [PubMed: 12626390]
23. Poland DC, Kratz E, Vermeiden JP, De Groot SM, Bruyneel B, De Vries T, Van Dijk W. *Prostate* 2002;52:34–42. [PubMed: 11992618]
24. Marrero JA, Romano PR, Nikolaeva O, Steel L, Mehta A, Fimmel CJ, Comunale MA, D'Amelio A, Lok AS, Block TM. *J. Hepatol* 2005;43:1007–1012. [PubMed: 16137783]
25. Ahmed N, Oliva KT, Barker G, Hoffmann P, Reeve S, Smith IA, Quinn MA, Rice GE. *Proteomics* 2005;5:4625–4636. [PubMed: 16220531]
26. Bertozzi CR, Kiessling LL. *Science* 2001;291:2357–2364. [PubMed: 11269316]
27. Okuyama N, Ide Y, Nakano M, Nakagawa T, Yamanaka K, Moriwaki K, Murata K, Ohigashi H, Yokoyama S, Eguchi H, Ishikawa O, Ito T, Kato M, Kasahara A, Kawano S, Gu JG, Miyoshi E. *Int. J. Cancer* 2006;118:2803–2808. [PubMed: 16385567]
28. Lacunza I, Kremmer T, Diez-Masa JC, Sanz J, de Frutos M. *Electrophoresis* 2007;23:4447–4451.
29. Yang ZP, Hancock WS. *Journal of Chromatography A* 2004;1–2:79–88
30. Drake RR, Schwegler EE, Malik G, Diaz J, Block T, Mehta A, Semmes OJ. *Molecular & Cellular Proteomics* 2006;10:1957–1967. [PubMed: 16760258]
31. Cho W, Jung K, Regnier FE. *Anal. Chem* 2008;14:5286–5292. [PubMed: 18558770]
32. Kyselova Z, Mechref Y, Kang P, Goetz JA, Dobrolecki LE, Sledge GW, Schnaper L, Hickey RJ, Malkas LH, Novotny MV. *Clinical Chemistry* 2008;7:1166–1175. [PubMed: 18487288]
33. Zhao J, Simeone DM, Heidt D, Anderson MA, Lubman DM. *Journal of Proteome Research* 2006;7:1792–1802. [PubMed: 16823988]
34. Zhao J, Qiu WL, Simeone DM, Lubman DM. *Journal of Proteome Research* 2007;3:1126–1138. [PubMed: 17249709]
35. Zhang H, Li XJ, Martin DB, Aebersold R. *Nature Biotechnology* 2003;6:660–666.
36. Zhao J, Patwa TH, Qiu WL, Shedden K, Hinderer R, Misek DE, Anderson MA, Simeone DM, Lubman DM. *Journal of Proteome Research* 2007;5:1864–1874. [PubMed: 17428079]
37. Qiu Y, Patwa TH;L, Xu L, Shedden K, Misek DE, Tuck M, Jin G, Ruffin MT, Turgeon DK, Synal S, Bresalier R, Marcon N, Brenner DE, Lubman DM. *Journal of Proteome Research* 2008;7(4):1693–1703. [PubMed: 18311904]
38. Alvarez RA, Blixt O. *Glycobiology* 2006;415:292–310.
39. Kuno A, Uchiyama N, Koseki-Kuno S, Ebe Y, Takashima S, Yamada M, Hirabayashi J. *Nature Methods* 2005;11:851–856. [PubMed: 16278656]
40. Chen P, Liu Y,K, Kang XN, Sun L, Yang PY, Tang ZY. *Journal of Cancer Research and Clinical Oncology* 2008;8:851–860. [PubMed: 18264723]
41. Chen SM, LaRoche T, Hamelinck D, Bergsma D, Brenner D, Simeone D, Brand RE, Haab BB. *Nature Methods* 2007;5:437–444. [PubMed: 17417647]
42. Ingvarsson J, Wingren C, Carlsson A, Ellmark P, Wahren B, Engstrom G, Harmenberg U, Krogh M, Peterson C, Borrebaeck CAK. *Proteomics* 2008;11:2211–2219. [PubMed: 18528842]
43. Haab BB. *Current Opinion in Biotechnology* 2006;4:415–421. [PubMed: 16837184]
44. Orzechowski R, Hamelinck D, Li L, Gliwa E, VanBrocklin M, Marrero JA, Woude GFV, Feng ZD, Brand R, Haab BB. *Cancer Research* 2005;23:11193–11202. [PubMed: 16322270]
45. Kreunin P, Zhao J, Rosser C, Urquidi V, Lubman DM, Goodison S. *Journal of Proteome Research* 2007;7:2631–2639. [PubMed: 17518487]
46. Liang XL, Lubman DM, Rossi DT, Nordblom GD, Barksdale CM. *Anal. Chem* 1998;3:498–503. [PubMed: 9470484]
47. Ransohoff DF. *Nat. Rev. Cancer* 2005;5:142, 149. [PubMed: 15685197]
48. Evans-Nguyen KM, Tao SC, Zhu H, Cotter RJ. *Anal. Chem* 2008;5:1448–1458. [PubMed: 18254611]

Figure 1.
An outline of the experimental flow of microarray processing and on-target digestion.

Figure 2.

a The fluorescent image of 14 identical antibody arrays on a slide where in each box, the first column is A1BG, the second column is Amyloid p component, the third column is Antithrombin III. There were thus 42 identical spots per antibody on the array. They were hybridized with the same sample to test the array-to-array reproducibility. SNA lectin was used for detection; **b** The intensity of the fluorescent signals in the slide shown in **a** was computed and presented in the three charts by spot number for A1BG, Amyloid p component and Antithrombin-III.

Figure 3.

This is a saturation curve which shows how the three antibodies respond to different dilution of serum with SNA lectin detection. The numbers on X-axis are the folds that the serum has been diluted before being hybridized with the antibody array. The y-axis is the intensity of the signal.

Figure 4.

The MALDI-MS spectra generated on the microarray spots of Amyloid p component antibody after on-target digestion. The peaks identified as Amyloid p component were marked with green arrows where the extra peaks appearing in c were marked with black arrows. **a** control spot, without incubation of serum; **b** incubated with 10X diluted serum; **c** incubated with 2X diluted serum.

Figure 5.

Scatter plot in log2 scale between every pair of technical replicates (a replicate is two distinct points same patient, same antibody, same fasting status and same batch) i.e. for any spot in the figure, its values on the axes are the log2 intensities of any pair of replicate spots. The lectin used in this study is SNA.

Figure 6.
ROC curves for the three antibodies alone and A1BG and Amyloid combined.

Figure 7. Scatter plot of sialylation level detected by lectin SNA on A1BG and Amyloid p component. Samples from different disease classes are differently colored. Normal spots are orange, pancreatitis spots are blue, diabetes spots are pink and cancer spots are dark blue.

Figure 8. The boxplot depicts the distribution of the measurements for antibody A1BG. The boxes provide the upper and lower quartiles of the measurements with respect to the median value (red line in each box). The lines provide the ranges of the measurements, excluding outliers (+).

Table 1

Residual, batch and patient standard deviations of individual antibodies. These are the estimated standard deviations for variance components in the linear mixed effects model, as described in the text. Only samples with fasting replicates were analyzed.

Std. Dev. (log2)	A1BG	Amyloid	Antithrombin-III
Residual	0.21559	0.19667	0.22877
Batch	0.20701	0.16376	0.15382
Patient	0.12681	0.05973	0.19776

Table 2

Fixed effects for sex, disease classes (excluding cancer) and fasting status. Point estimates, standard error and T value (ratio of point estimate to standard error) are given. The effect estimate (column 2) is on the log2 scale, and the percent effect on the raw scale is given in column 3. T-values greater than 2 in magnitude correspond to two-sided p-values smaller than 0.05. Only samples with fasting replicates were analyzed. The reference category is a normal (free of pancreatitis and diabetes), non-fasting female. Fixed effects are relative to this category.

A1BG	Estimate	Percent	Std. Error	T value
Male	0.2948	2%	0.05324	0.55
Female	0	0%	---	---
Pancreatitis	-0.03688	-3%	0.65551	-0.56
Diabetic	-0.3033	-2%	0.06616	-0.46
Normal	0	0%	---	---
Fasting	0.06761	5%	0.03157	2.14
Non-fasting	0	0%	---	---
Amyloid	Estimate	Percent	Std. Error	T value
Male	0.04110	3%	0.03808	1.1
Female	0	0%		
Pancreatitis	-0.03545	-2%	0.04829	-0.7
Diabetic	-0.00562	0%	0.04719	-0.1
Normal	0	0%	---	---
Fasting	0.06383	4%	0.02758	2.3
Non-fasting	0	0%	---	---
Antithrombin-III	Estimate	Percent	Std. Error	T value
Male	0.1134	8%	0.04489	2.5
Female	0	0%	---	---
Pancreatitis	-0.16112	-12%	0.0564	-2.9
Diabetic	0.03839	3%	0.05640	0.7
Normal	0	0%	---	---
Fasting	0.01663	1%	0.03088	0.5
Non-fasting	0	0%	---	---

Table 3

Fixed effects of disease classes including cancer. Point estimates, standard error and T value (ratio of point estimate to standard error) are given. The effect estimate (column 2) is on the log2 scale, and the percent effect on the raw scale is given in column 3. T-values greater than 2 in magnitude correspond to two-sided p-values smaller than 0.05. The reference category is an individual free of diabetes and pancreatitis. Fixed effects are relative to this category.

A1BG	Estimate	Percent	Std. Error	T value
Cancer	0.75515	69%	0.06972	10.8
Pancreatitis	0.02404	−2%	0.06323	−0.4
Diabetic	−0.01072	−1%	0.07169	−0.1
Normal	0	0%	---	---
Amyloid	Estimate	Percent	Std. Error	T value
Cancer	0.41395	33%	0.06065	6.8
Pancreatitis	−0.04436	−3%	0.05510	−0.8
Diabetic	0.00274	0%	0.06065	0.0439
Normal	0	0%	---	---
Antithrombin-III	Estimate	Percent	Std. Error	T value
Cancer	0.03193	2%	0.07523	0.4
Pancreatitis	−0.1629	−11%	0.06808	−2.4
Diabetic	0.02831	2%	0.07523	0.4
Normal	0	0%	---	---