Discovery and Validation of Serum Protein Changes in Type 1 Diabetes Patients Using High Throughput Two Dimensional Liquid Chromatography-Mass Spectrometry and Immunoassays

Wenbo Zhi†§§, Ashok Sharma†§§, Sharad Purohit‡, Eric Miller‡, Bruce Bode¶, Stephen W. Anderson∥, John Chip Reed**, R. Dennis Steed**, Leigh Steed‡, Diane Hopkins‡, and Jin-Xiong She‡§‡‡

Type 1 diabetes (T1D) is expected to cause significant changes in the serum proteome; however, few studies have systematically assessed the proteomic profile change associated with the disease. In this study, a semi-quantitative spectral counting-based two-dimensional liquid chromatography mass spectrometry platform was used to analyze serum samples from T1D patients and controls. In this discovery phase, significant differences were found for 21 serum proteins implicated in inflammation, oxidation, metabolic regulation, and autoimmunity. To assess the validity of these findings, six candidate proteins including adiponectin, insulin-like growth factor binding protein 2, serum amyloid protein A, C-reactive protein, myeloperoxidase, and transforming growth factor beta induced were selected for subsequent immune assays for 1139 T1D patients and 848 controls. A series of statistical analyses using cases and controls matched for age, sex, and genetic risk confirmed that T1D patients have significantly higher serum levels for four of the six proteins: adiponectin (odds ratio (OR) = 1.95, \( p = 10^{-27} \)), insulin-like growth factor binding protein 2 (OR = 2.02, \( p < 10^{-28} \)), C-reactive protein (OR = 1.13, \( p = 0.007 \)), serum amyloid protein A (OR = 1.51, \( p < 10^{-19} \); whereas the serum levels were significantly lower in patients than controls for the two other proteins: transforming growth factor beta induced (OR = 0.74, \( p < 10^{-5} \) and myeloperoxidase (OR = 0.51, \( p < 10^{-41} \)). Compared with subjects in the bottom quartile, subjects in the top quartile for adiponectin (OR = 6.29, \( p < 10^{-37} \)), insulin-like growth factor binding protein 2 (OR = 7.95, \( p < 10^{-49} \)), C-reactive protein (OR = 1.38, \( p = 0.025 \)), serum amyloid protein A (OR = 3.36, \( p < 10^{-16} \)) had the highest risk of T1D, whereas subjects in the top quartile of transforming growth factor beta induced (OR = 0.41, \( p < 10^{-11} \)) and myeloperoxidase (OR = 0.10, \( p < 10^{-43} \)) had the lowest risk of T1D. These findings provided valuable information on the proteomic changes in the sera of T1D patients. Molecular & Cellular Proteomics 10.1074/mcp.M111.012203, 1–10, 2011.

Type 1 diabetes (T1D) is an autoimmune disease that occurs when insulin-producing \( \beta \) cells in the pancreas are destroyed by one’s own immune system. Significant efforts have been devoted to developing biomarkers that could predict the risk for T1D or diabetic complications (1). Because pathological tissues from T1D patients are very difficult to obtain because of ethical and practical concerns (2), serum is an excellent alternative for biomarkers relevant to diseases because serum proteins may be secreted by cells at the pathological sites. However, comprehensive analysis of the serum proteome is a challenging task using current proteomic technologies because of its extraordinary complexity and high dynamic range in concentration (3, 4). Proteomic analysis is further complicated by the potentially

From the †Center for Biotechnology and Genomic Medicine; ‡Department of Pathology, Georgia Health Sciences University, Augusta, GA; †Atlanta Diabetes Associates, Atlanta, GA; §§Pediatric Endocrine Associates, Atlanta, GA; ¶Southeastern Endocrine and Diabetes, Atlanta, GA

Received June 27, 2011, and in revised form, September 2, 2011
Published, MCP Papers in Press, September 6, 2011, DOI 10.1074/mcp.M111.012203

Molecular & Cellular Proteomics 10.11
10.1074/mcp.M111.012203-1

Want to cite this article? Please look on the last page for the proper citation format.
large variation between individuals and relatively small differences between study groups. Furthermore, serum protein levels may be influenced by a large number of parameters such as subject age, sex, geographic or ethnic origin, environmental factors, and disease status. Therefore, quantification of the serum proteome requires well-designed strategies that incorporate different technologies and experimental designs at different phases.

The two dimensional liquid chromatography (2DLC) mass spectrometry platform has been widely used in the discovery of protein candidates that may serve as biomarkers (5). Prior to the MS analysis, serum samples have to be processed to reduce the dynamic range and enrich for low-abundance proteins (e.g. normalization of protein levels). Multiple normalization approaches have been used in previous biomarker discovery studies. We and others have successfully used a combinatorial hexapeptide library-based technique for sample preprocessing (5–7). There are also several different approaches that can be used to derive quantitative information from MS analysis. Spectral counting is a simple, low-cost and semiquantitative approach to assess MS-based protein differences.(8) This approach was used in this study to compare serum protein differences between T1D patients and autoantibody-negative (AbN) control subjects.

Irrespective of the technical platforms used in the discovery phase, the selected biomarker candidates have to be validated in subsequent studies. Ideally, the validation studies should use a different set of measurement techniques and independent sample set(s) with reasonably large sample size. Our validation phase included six candidate proteins that showed significant differences in the discovery phase: adiponectin (ADIPOQ), insulin-like growth factor binding protein 2 (IGFBP2), serum amyloid protein A (SAA), C-reactive protein (CRP), myeloperoxidase (MPO), and transforming growth factor beta induced (TGFBI). ELISA or Luminex assays were used to analyze 1139 T1D patients and 848 AbN controls. This study provided convincing evidence that the six proteins were significantly altered in T1D patients.

**EXPERIMENTAL PROCEDURES**

**Human Subjects and Serum Samples**—This study was approved by the institutional review board of the Georgia Health Sciences University and informed consent was obtained from every subject or his/her legally authorized representative. Peripheral blood was collected in serum separator tubes (BD Biosciences, San Jose, CA). After clotting for 30–120 min, the tubes were centrifuged and serum was collected in serum separator tubes (BD Biosciences, San Jose, CA). After clotting for 30–120 min, the tubes were centrifuged and serum was immediately aliquoted and stored in −80 °C freezers. From blood collection to assay, every sample had less than three freeze/thaw cycles. Under these storage and processing conditions, serum proteins are usually stable and do not show significant signs of degradation.

A total of 30 T1D patients and 30 AbN controls were used in the discovery phase of this study. These subjects were matched for age and sex. Because of the long analysis time (around 16 h per sample) for each sample, serum samples in the same phenotypic group were pooled for proteomic analysis. We created three serum pools for each phenotypic group: AbN (AbN-P1, AbN-P2, and AbN-P3), T1D (T1D-

**FIG. 1. Distribution of age and gender for subjects analyzed in this study.** Female (top panels) and male (bottom panels) were presented separately. The left panels are for controls whereas the right panels are for T1D subjects.

P1, T1D-P2, and T1D-P3). Each pool contained equal amounts of serum from all 10 subjects.

The subjects used in the confirmation phase included 1139 T1D patients and 848 AbN controls. The controls consisted of 529 subjects who were first-degree relatives (FDR) of T1D patients and 319 general population (GP) subjects without a diabetic FDR. All subjects (patients, FDR, and GP controls) were recruited in the state of Georgia, USA, mainly in the Atlanta and Augusta areas. They were recruited from the same period of time and using identical procedures for sample processing and storage. The controls are subjects who participated in our longitudinal study, the Prospective Assessment in Newborns of Diabetes Autoimmunity (PANDA). The sex and age distributions for control and T1D subjects are presented in Fig. 1. The stored samples were aliquoted into 96-well v-bottom plates, known as “master” plates. Each master plate contained similar numbers of T1D patients and AbN controls. From the master plates, multiple working plates were created and frozen to avoid repeated freeze/thaw cycles. All sample plates were stored at −80 °C until use.

**2DLC MS**—All six sample pools were first processed to normalize the concentrations of serum proteins (decrease abundant proteins and increase low-abundance proteins) using the ProteoMiner kit (Bio-Rad Laboratories, Hercules, CA), which is a library of random hexamer peptide conjugated to beads (9). Normalized samples were digested with trypsin and then analyzed with a ProteomeX™ Work station (Thermo, San Jose, CA), which consists of an on-line 2DLC system and a linear ion-trap mass spectrometry (LTQ). Peptides were resolved online by a strong cation exchange (SCX) capillary column and then a reverse phase (RP) capillary column. Eluted peptide ions were detected in one survey scan (from 400 to 1600 m/z) followed by ten data-dependent MS/MS scans (ion isolation width 2 m/z, normalized collision energy 35%) with dynamic exclusion settings (repeat count 2, repeat duration 30s and exclusion duration 60s). Because the MS instrument is not capable of analyzing all peptide ions passing through the machine (10), each of the six sample pools was analyzed three times.
Serum Protein Profiles in Type 1 Diabetes Patients

and the average data from the three replicates were used for statistical analyses. All 18 MS/MS runs were completed in 3 weeks.

**Protein Identification and Quantification**—The tandem MS data sets were searched against the NCBI nonredundant human database (Date: 07/24/2007, number of proteins: 133123) using the embedded TurboSEQUEST algorithm of the Bioworks software (version 3.2, Thermo, San Jose, CA). MS and MS/MS tolerances were set at ±1 Da and ±0.5 Da, respectively. Missed cleavage sites were set to be 2. Carbamidomethylation of cysteine was set as fixed modification and oxidation of methionine was set to be variable modification. Only peptides as a result of complete trypsin digest and with a minimum of four amino acids in length were considered for protein identification. False positive rates (FPR) were calculated using a previously reported method (11) (See also supplemental Methods). The confident protein lists were generated using a FPR threshold of 1%.

The peptide counts (spectral numbers) were used as a semiquantitative measure for protein levels. The quantitative data for all 18 LC-MS runs (six pools × three replicates) were assembled together using in-house software. The data were first normalized using the “total peptide count” method, which assumes that the total peptide count for each LC-MS run should be identical. After data normalization across all 18 LC-MS runs, the mean peptide count for the three replicates is calculated for each protein and each sample pool. The average peptide counts for each sample pool were used for statistical analysis. The differences between the T1D and AbN groups were assessed by student t test. To calculate the T1D/AbN ratios for proteins that were not detected in the AbN group (spectral number = 0), the expression levels were arbitrarily set to be 0.05. The raw peptide counts, normalized peptide counts, T1D/AbN ratios, and values for all detected proteins are shown in supplemental Table S1.

**Luminex Assays**—SAA and MPO were measured using Luminex kits from Millipore Company (Billerica, MA) according to manufacturer’s recommendation. Properly diluted serum samples (1:1000 and 1:100 dilutions for SAA and MPO, respectively) were incubated with the antibody-coupled microspheres and then with biotinylated detection antibody before the addition of streptavidin-phycoerythrin. The captured bead-complexes were measured with the Bioplex 100 array reader (Bio-Rad Laboratories, Hercules, CA) with the following instrument settings (events/bead: 50, minimum events: 0, Flow rate: 60 μl/min, sample size: 50 μl, discriminator gate: 8,000 to 15,000). Median fluorescence intensity (MFI) was collected and used for calculating protein concentration.

**ELISA Assays**—Four serum proteins (TGFBI, CRP, ADIPOQ, and IGFBP2) were measured using ELISA kits from R&D Systems Inc. (Minneapolis, MN). All kits are two-step “sandwich” assays. Briefly, serum samples were diluted with dilution factors of 1:19,200, 1:20,000, 1:640 and 1:2000 for ADIPOQ, CRP, IGFBP2, and TGFBI, respectively. Diluted samples were incubated in microtiter plates coated with the capture antibody for each molecule. After incubation and washing, the detection antibody labeled with horse radish peroxidase (HRP) was added to each well. After a second incubation and then washing, substrate tetramethylbenzidine (TMB) was added to the wells and then an acidic stopping solution was added to terminate the reactions. The plates were read with a plate reader to determine the absorbance at 450 nm (target) and 630 nm (reference) wavelength.

**Statistical Analyses**—All statistical analyses were performed using the R language and environment for statistical computing (R version 2.12.1; R Foundation for Statistical Computing; www.r-project.org). For Luminex/ELISA data, protein concentrations were estimated using a four-parameter logistic regression fit to the standard curve with the known concentration included on each plate using a serial dilution series. The MFI and concentration data were log-transformed to achieve more normal distributions. Wells with individual bead counts below 30, or bead CV above 200 were flagged for exclusion. The coefficient of variation of replicate wells was also checked and wells with CV > 25% were not included in further analyses. To automatically detect outliers in the standard curve, we used the “rdstuend” function, which calculates the studentized residuals. Because the standard deviations of residuals vary greatly, these residuals are renormalized by “rdstuend” to have unit variance, using leave-one-out measure of the error variance. This is an important technique in the detection of outliers. Finally, all standard curves were manually inspected for any abnormalities.

The pairwise correlation between individual protein levels was computed using Pearson correlation coefficient. The effect of age on the serum levels of each candidate molecule was determined using a linear regression of protein concentration with age as covariate on data stratified by sex and disease status. For further analyses, case-control matching was performed with respect to age, sex, and genetic risk (based on HLA genotypes) using the “matching” R package (12). The differences in the levels of the proteins between different study groups were assessed using a paired t test on the case-control paired data. To estimate the relative risk of diabetes at different protein concentrations, we performed conditional logistic regression on matched paired data. The odds ratios and 95% confidence intervals (CI) were computed for different models using one or a combination of proteins in each model. For analyses using protein concentration as continuous variable, the data were log transformed and scaled to unit standard deviation. For analyses using protein concentration as categorical variables, values 1, 2, 3, 4 were assigned using the quartile values in controls as cutoff points. A p value for trend across quartiles was assessed in separate models using lowest quartile as referent.

**RESULTS**

To discover and validate serum proteomic differences between T1D and AbN control groups, we designed our experiments comprising of a discovery phase and a validation phase (Fig. 2). In the discovery phase, a small set of samples was analyzed using spectral counting-based 2DLC MS technique, whereas in the validation phase six candidate proteins were analyzed using ELISA or Luminex assays.

**Discovery of Serum Biomarker Candidates Using 2DLC MS**—To discover serum protein alterations associated with T1D, six pools of serum samples (three AbN and three T1D) were generated by pooling sera from 10 subjects for each pool. At first, a combinatorial hexapeptide library conjugated to beads was used to reduce the dynamic range of serum proteins and increase the chance of discovering medium- to low-abundance proteins (5). Normalized samples were digested with trypsin and the resulting peptides were separated on a 2D-HPLC and analyzed by MS. A total of 2477 proteins were identified in the entire data set (supplemental Table S1) and each MS run identified on average 655 proteins. As expected, the number of proteins that can be identified in each sample increases as the number of replicate runs increases (supplemental Fig. S1). After three replicates, further increase in numbers of detected proteins slows significantly.

Thirty one proteins showed significant differences between the T1D and AbN groups using our predefined threshold (p < 0.05 and ≥ 1.5-fold change). At least 21 of the 31 proteins may be functionally relevant to T1D as they may be implicated in innate immunity, inflammation and...
Validation of Six Candidate Proteins—To assess the validity of the discovery data, six proteins were selected for the validation studies based on the degree of differences in the discovery data set, assay availability and functional relevance to T1D. These six candidate proteins were analyzed using either ELISA or Luminex assays for 848 AbN controls and 1139 T1D patients. The age and sex distributions of these subjects are shown in Fig. 1. Before examining the differences between AbN and T1D groups, we determined whether sex and sampling age of the subjects were potential confounding factors. We performed regression of protein concentration with age as covariate (Fig. 3). Data for each molecule were divided into four subgroups based on sex and disease phenotype. There were significant correlations with age in both phenotypic groups (AbN and T1D) and for both sexes for several serum proteins (supplemental Table S2). SAA and CRP showed the greatest overall correlations with age for different sexes and different phenotypes (supplemental Table S2). These two molecules are positively correlated with age in both phenotypic groups and for both sexes. Serum levels of IGFBP2 and ADIPOQ are negatively correlated with age in AbN subjects in both sexes; however, the levels of these two molecules increase with age in T1D subjects. The patterns of correlations are quite similar for the four proteins with related functions (SAA and CRP; ADIPOQ and IGFBP2). MPO and TGFBI levels did not differ significantly by age.

Because age and sex are confounding factors for four proteins, age and sex matching was then performed using a multivariate and propensity score matching software (12). Each T1D patient was paired with a closest AbN control with respect to age, sex, and genetic risk based on HLA genotypes (Table II). Paired T-tests were used to compare the differences in the mean levels of the six proteins between T1D and AbN subjects (Fig. 4). The serum levels of ADIPOQ (T1D/AbN = 1.43, p = 6.2e-46), IGFBP2 (T1D/AbN = 1.98, p = 2.1e-62), SAA (T1D/AbN = 1.77, p = 3.4e-25), and CRP (T1D/AbN = 1.25, p = 9.4e-05) were significantly higher in T1D cases as compared with matched AbN controls. The mean TGFBI (T1D/AbN = 0.86, p = 2.6e-21) and MPO (T1D/AbN = 0.53, p = 3.4e-55) levels were significantly lower in T1D cases as compared with matched AbN controls.

Next, we examined the pairwise correlations for all six proteins (Fig. 5). There were significant positive correlations between functionally related proteins (r = 0.66 for SAA-CRP and r = 0.43 for ADIPOQ-IGFBP2). SAA and CRP are known to be directly related to the inflammatory process. On the other hand, ADIPOQ and IGFBP2 are known to be involved in glucose intolerance, insulin resistance, and inhibition of insulin-like growth factor-dependent signaling.

Serum Protein Changes in T1D Patients—To assess serum protein changes associated with T1D, we performed conditional (matched pairs) logistic regression analysis adjusting for age, sex, and genetic risk. The odds ratios of T1D per S.D. increment of protein levels were computed for different models (Table III). We found that all six molecules were significantly associated with increased or decreased risk of having T1D: ADIPOQ (OR = 1.95), IGFBP2 (OR = 2.02), CRP (OR = 1.13), SAA (OR = 1.51), TGFBI (OR = 0.74), and MPO (OR = 0.51). To assess the relative risk of T1D across different quartiles of protein levels, we categorized protein levels using...
the quartile values in controls as cutoff points. Subjects in the top quartile had the highest risk of diabetes compared with subjects in the bottom quartile for four proteins: ADIPOQ (OR/H11005 6.29), IGFBP2 (OR/H11005 7.95), CRP (OR/H11005 1.38), and SAA (OR/H11005 3.36) (Table III). In contrast, subjects in the top quartile had the lowest risk of diabetes compared with subjects in the bottom quartile for the other two proteins: TGFBI (OR/H11005 0.41) and MPO (OR/H11005 0.10) (Table III).

Protein Combinations and T1D—We also evaluated models using different combinations of proteins utilizing at least one of the proteins with the most positive association (IGFBP2) or the most negative association (MPO). We performed conditional logistic regression analyses using all possible combinations and evaluated models for improvement in the ORs. Addition of SAA, IGFBP2, or CRP improved the MPO model fit (OR/H11005 0.45, 0.48, and 0.49 respectively) (Table III). The best improvement was observed when MPO was combined with SAA and IGFBP2 (OR = 0.42). No further improvement was observed when all six proteins were included (OR = 0.42). Combination of IGFBP2 with TGFBI (OR = 2.18), MPO (OR = 2.10), or CRP (OR = 2.07) also improved the models significantly. After examining all possible combinations, the highest odds ratio (OR = 2.44) was obtained with four proteins (IGFBP2, TGFBI, MPO, and CRP). We also evaluated different models after transforming the quantitative data (protein levels) into categorical data based on quartiles (Table III). IGFBP2 had the largest OR when comparing the fourth quartile with the first quartile (OR/H11005 7.95). When IGFBP2 is combined with other proteins (TGFBI, MPO or CRP), the OR increased modestly (Table III) and the highest OR was obtained for a four protein model that includes IGFBP2, TGFBI, MPO, and CRP (OR/H11005 11.32). On the other hand, lowest OR was obtained with the quartile data for the model that includes MPO, SAA, and IGFBP2 (OR = 0.05) (Table III).

<table>
<thead>
<tr>
<th>Accession ID</th>
<th>Protein Name</th>
<th>Symbol</th>
<th>ABN mean</th>
<th>T1D mean</th>
<th>T1D/ABN</th>
<th>p value</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innate immunity</td>
<td>kininogen 1 isoform 1</td>
<td>KNG1</td>
<td>1.39</td>
<td>3.52</td>
<td>2.48</td>
<td>0.0014</td>
<td>DC activation</td>
</tr>
<tr>
<td>complement C3b</td>
<td>C3</td>
<td>3.22</td>
<td>1.56</td>
<td>0.49</td>
<td>0.0061</td>
<td>Complement</td>
<td></td>
</tr>
<tr>
<td>complement component C4</td>
<td>C4</td>
<td>2.33</td>
<td>1.11</td>
<td>0.49</td>
<td>0.0125</td>
<td>Complement</td>
<td></td>
</tr>
<tr>
<td>Inflammation/oxidation</td>
<td>serum amyloid A protein*</td>
<td>SAA</td>
<td>0.28</td>
<td>1.38</td>
<td>4.34</td>
<td>0.0002</td>
<td>Proinflammatory</td>
</tr>
<tr>
<td>peroxiredoxin 2 isoform a</td>
<td>PRDX1</td>
<td>0.51</td>
<td>0.00</td>
<td>0.09</td>
<td>0.0480</td>
<td>Anti-oxidant</td>
<td></td>
</tr>
<tr>
<td>annexin I</td>
<td>ANXA1</td>
<td>0.00</td>
<td>0.46</td>
<td>10.23</td>
<td>0.0163</td>
<td>Inflammatory response</td>
<td></td>
</tr>
<tr>
<td>C-Reactive Protein*</td>
<td>CRP</td>
<td>0.54</td>
<td>2.97</td>
<td>5.13</td>
<td>0.0227</td>
<td>Proinflammatory</td>
<td></td>
</tr>
<tr>
<td>Lymphocyte activation/proliferation</td>
<td>erbB3 binding protein</td>
<td>PA2G4</td>
<td>0.50</td>
<td>0.00</td>
<td>0.09</td>
<td>0.0485</td>
<td>Cell proliferation</td>
</tr>
<tr>
<td>Lymphocyte trafficking/infiltration</td>
<td>apolipoprotein A-IV precursor</td>
<td>APOA4</td>
<td>0.45</td>
<td>1.41</td>
<td>2.89</td>
<td>0.0311</td>
<td>Leukocyte adhesion</td>
</tr>
<tr>
<td>WNT1 inducible signaling pathway protein 2 precursor</td>
<td>WISP2</td>
<td>0.09</td>
<td>1.05</td>
<td>8.04</td>
<td>0.0071</td>
<td>Cell adhesion</td>
<td></td>
</tr>
<tr>
<td>heparin binding protein</td>
<td>HBP</td>
<td>0.30</td>
<td>0.93</td>
<td>2.79</td>
<td>0.0304</td>
<td>Vascular permeability</td>
<td></td>
</tr>
<tr>
<td>tropomyosin 4</td>
<td>TPM4</td>
<td>0.93</td>
<td>0.12</td>
<td>0.18</td>
<td>0.0352</td>
<td>Endothelial function</td>
<td></td>
</tr>
<tr>
<td>Glucose regulation/homeostasis</td>
<td>insulin-like growth factor binding protein 2*</td>
<td>IGFBP2</td>
<td>0.19</td>
<td>1.13</td>
<td>4.87</td>
<td>0.0180</td>
<td>Regulation of cell growth</td>
</tr>
<tr>
<td>adiponectin precursor</td>
<td>ADIPOQ</td>
<td>5.61</td>
<td>8.61</td>
<td>1.53</td>
<td>0.0112</td>
<td>Glucose homeostasis</td>
<td></td>
</tr>
<tr>
<td>solute carrier family 2</td>
<td>SLC2A2</td>
<td>1.91</td>
<td>0.65</td>
<td>0.36</td>
<td>0.0144</td>
<td>Glucose transport</td>
<td></td>
</tr>
<tr>
<td>Autoimmunity/immune-regulation</td>
<td>myeloperoxidase*</td>
<td>MPO</td>
<td>0.38</td>
<td>1.43</td>
<td>3.47</td>
<td>0.0489</td>
<td>Immuno-regulation</td>
</tr>
<tr>
<td>transforming growth factor, beta-induced*</td>
<td>TGFB1</td>
<td>0.53</td>
<td>0.09</td>
<td>0.25</td>
<td>0.0412</td>
<td>Autoimmunity</td>
<td></td>
</tr>
</tbody>
</table>

Serum Protein Profiles in Type 1 Diabetes Patients

Molecular & Cellular Proteomics 10.11 10.1074/mcp.M111.012203–5
DISCUSSION

This is the first comprehensive study that integrates multiple state-of-the-art proteomic technologies to discover and validate serum proteomic changes in T1D patients. Using a 2DLC MS platform, 21 serum proteins with functional relevance to T1D were found to be significantly different between T1D and controls. Using two different immunoassays, six selected proteins (SAA, CRP, MPO, TGFBI, ADIPOQ, and IGFBP2) were further validated in a confirmation data set consisting of 848 control subjects and 1139 T1D patients. Several technologies and designs were critical to the success of the study. Normalization of serum protein concentrations with a combinatorial hexapeptide library and efficient separation of peptides with 2DLC allowed the analyses of proteins with moderate levels in serum. The large numbers of samples analyzed in the validation phase with the high throughput ELISA/Luminex assays was essential to the successful confirmation of the candidate proteins.

From a functional point of view, all six serum proteins are potentially relevant to T1D as they are implicated in inflammation, oxidation, metabolic regulation, and autoimmunity. Inflammation is a hallmark of many diseases including diabetes and autoimmune diseases. Two well-known inflammatory mediators, SAA and CRP, were discovered and confirmed to be increased in T1D patients compared with controls in our study. The elevation of CRP serum levels in T1D may be because of the increased inflammation and microcirculatory abnormalities in T1D patients (13, 14). However, SAA has not been properly investigated in T1D. SAA is a classic acute-phase protein predominantly produced by the liver in response to injury, infection, and inflammation (15). In vitro, SAA induces the expression and release of proinflammatory cytokines such as tumor necrosis factor-alpha (TNF-α), interleukin 1 beta (IL-1β), and interleukin-8 (IL-8) in neutrophils and monocytes. The basal SAA levels are persistently elevated in type 2 diabetes (T2D) and other chronic conditions (16). It is also found that neutrophils and monocytes from T2D patients are more responsive to SAA for the induction of the proinflammatory cytokine IL-1β and the proangiogenic and chemotactic protein IL-8 (17). The increased production of cytokines and increased migration of leukocytes in response to

![Fig. 3. Changes of mean serum levels according to sex and age.](image)

**TABLE II**

<table>
<thead>
<tr>
<th></th>
<th>Before Matching</th>
<th>After Matching</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AbN</td>
<td>T1D</td>
</tr>
<tr>
<td>Age (years)</td>
<td>22.188</td>
<td>33.377</td>
</tr>
<tr>
<td>Sex (Female %)</td>
<td>54.36%</td>
<td>52.15%</td>
</tr>
<tr>
<td>Genetic Risk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>60.14%</td>
<td>36.35%</td>
</tr>
<tr>
<td>Medium</td>
<td>20.99%</td>
<td>19.05%</td>
</tr>
<tr>
<td>High</td>
<td>16.39%</td>
<td>41.53%</td>
</tr>
</tbody>
</table>
SAA may contribute to a sustained accumulation and activation of inflammatory cells. Accordingly, the hyper-responsiveness of leukocytes to SAA may be relevant to the proinflammatory conditions associated with vascular and other complications in diabetic patients (18, 19). Therefore, SAA may also be a very important molecule in T1D and its complications. Although the mean SAA level in the plasma of 38 T1D patients was not different from the mean level in 41 control subjects measured in a previous study (20), the current study with 1139 T1D patients and 848 controls revealed significantly higher SAA levels in T1D patients. Although SAA and CRP levels in serum are correlated (Fig. 4), SAA seems to be a better indicator of the inflammatory state of diabetic patients than CRP because the difference between AbN and T1D is higher for SAA than CRP.

IGFBP2 is a member of the insulin-like growth factor binding protein family. These proteins bind to insulin-like growth factor-I (IGF-I) and -II (IGF-II) with very high affinity and control the distribution of the IGFs (21). The IGFBPs play important roles in the biological functions of IGFs including prolonging the half-time and regulating the bioavailability and bioactivity of IGFs. IGFBP2 is the major form of IGFBPs in cerebrospinal fluid but also exists in human serum in the form of a 50-kDa complex with IGF-I or IGF-II (22). Several studies reported circulating IGFBP2 levels under different pathological conditions. One study found a nonsignificant increase of serum IGFBP2 in a small group of 58 young T1D patients (23). In the same study, the insulin treated group has a significantly lower serum IGFBP2 than the untreated group, suggesting that insulin may influence IGFBP2 levels. Several studies also found that growth hormones and insulin can independently inhibit IGFBP2 expression (24–26).

Adiponectin (ADIPOQ) belongs to the family of adipocytkines, adipocyte-derived biologically active molecules, which
may influence the function and the structural integrity of other tissues (27). Serum adiponectin levels are significantly reduced in obese subjects (28) and patients with diseases frequently associated with obesity such as T2D (29) and coronary artery disease (30). Consistent with our findings, several studies reported higher serum/plasma adiponectin levels in patients with T1D and certain diabetic complications including nephropathy and cardiovascular disease (31–35). Adiponectin accumulates in damaged vascular walls and beneficially modulates the endothelial inflammatory response to vascular injury and can serve as a predictor of microvascular complications in T1D patients (36). The mechanisms that control the synthesis of adiponectin have not yet been determined. The differences in fat distribution and function between diabetic and nondiabetic subjects, alleviation of suppression effect by insulin and decreased elimination of adiponectin due to impaired renal function may potentially account for the difference in adiponectin levels (36, 37). The correlation between ADIPOQ and IGFBP2 levels (Fig. 4) may be because of the coregulation of these molecules by insulin.

TGFB1 is an extracellular matrix protein initially found in the human adenocarcinoma cells after treatment with Transforming growth factor beta (TGF-β). Its precise function remains obscure, although it appears to be an intriguingly versatile molecule with a role in a wide range of physiological and pathological conditions including corneal dystrophy (38), tumorigenesis (39), and nephropathy (40). Its serum concentration and possible role in T1D or its complications have not been reported. This is the first study to show that the serum level of TGFB1 is significantly lower in T1D patients.

Myeloperoxidase (MPO) is a hemoprotein released by activated neutrophils, monocytes, and tissue-associated macrophages after inflammatory stimuli. MPO catalyzes a reaction between chloride and hydrogen peroxide to produce hypohalous acid, which reacts with tyrosine residues to form 3-chlorotyrosine. Both hypochlorous acid and 3-chlorotyrosine are potent antimicrobial agents (41). MPO may play a protective role against the development of autoimmune diseases as shown in MPO knockout mice that are more susceptible to experimental autoimmune encephalomyelitis (EAE), an animal model for human multiple sclerosis (42). Oxidant hypochlorous acid produced by MPO reacts with taurine to form taurine chloramines that is a long lived molecule with immunomodulatory properties including suppression of lymphocyte proliferation, inhibition of the generation of proinflammatory cytokines (TNF-α, IFN-γ, IL-6, and IL-12), and promotion of anti-inflammatory interleukin production (43). Consistent with these findings, MPO activity was found to be significantly reduced in multiple sclerosis patients (44). This study provides unambiguous evidence that T1D patients have significantly lower serum MPO levels compared with controls. In support of our results, MPO activity was found to be lower in the neutrophils from T1D patients, an indication of impaired neutrophil function (45, 46) and lower intracellular MPO of neutrophils was also observed in T2D patients (47). However, two previous studies with very small sample sizes (30 and 37 patients, respectively) found marginally higher MPO activity in T1D patients compared with controls. In support of our results, MPO activity was found to be lower in the neutrophils from T1D patients, an indication of impaired neutrophil function (45, 46) and lower intracellular MPO of neutrophils was also observed in T2D patients (47). Because there is a wide range of MPO levels in T1D patients and controls, results from studies with small number of subjects can be very unreliable. In this study, we found convincing evidence that the mean serum MPO levels are lower in T1D patients than in controls. Subjects in the top quartile had the lowest risk of having diabetes compared with subjects in the bottom quartile (OR = 0.10).

**Acknowledgments**—We thank Dr. Richard McIndoe for developing the software to extract spectral information. We have no conflicts of interest.

* This work was supported by grants from the National Institutes of Health (4R33HD050196, 4R33-DK069878, and 2RO1HD37800) and Health (4R33HD050196, 4R33-DK069878, and 2RO1HD37800) and
Juvenile Diabetes Research Foundation (JDRF 1-2004-661) to Dr. Jin-Xiong She. Wenbo Zhi is supported by a postdoctoral fellowship from the Juvenile Diabetes Research Foundation (JDRF 3-2009-275). Sharad Purohit is a recipient of a JDRF Advanced Postdoctoral Fellowship (JDRF 10-2006-792) and Career Development Award (JDRF 2-2011-153).

This article contains supplemental Fig.S1 and Tables S1 and S2.

To whom correspondence should be addressed: Center for Biotechnology and Genomic Medicine, Georgia Health Sciences University, 1120 15th Street, Augusta, GA 30912. Tel: 706-721-3410; Fax: 706-721-3688; E-mail: jshe@georgiahealth.edu.

REFERENCES


In order to cite this article properly, please include all of the following information: Zhi, W., Sharma, A., Purohit, S., Miller, E., Bode, B., Anderson, S. W., Reed, J. C., Steed, R. D., Steed, L., Hopkins, D., and She, J.-X. (2011) Discovery and Validation of Serum Protein Changes in Type 1 Diabetes Patients Using High Throughput Two Dimensional Liquid Chromatography-Mass Spectrometry and Immunoassays. *Mol. Cell. Proteomics* 10(11):M111.012203. DOI: 10.1074/mcp.M111.012203.