

# Novel Urinary Protein Biomarkers Predicting the Development of Microalbuminuria and Renal Function Decline in Type 1 Diabetes

DANIELA SCHLATZER, BS<sup>1</sup>  
 DAVID M. MAAHS, MD, PHD<sup>2</sup>  
 MARK R. CHANCE, PHD<sup>1</sup>  
 JEAN-EUDES DAZARD, PHD<sup>1</sup>

XIAOLIN LI, MS<sup>1</sup>  
 FRED HAZLETT, MS<sup>1</sup>  
 MARIAN REWERS, MD, PHD<sup>2</sup>  
 JANET K. SNELL-BERGEON, PHD<sup>2</sup>

**OBJECTIVE**—To define a panel of novel protein biomarkers of renal disease.

**RESEARCH DESIGN AND METHODS**—Adults with type 1 diabetes in the Coronary Artery Calcification in Type 1 Diabetes study who were initially free of renal complications ( $n = 465$ ) were followed for development of micro- or macroalbuminuria (MA) and early renal function decline (ERFD, annual decline in estimated glomerular filtration rate of  $\geq 3.3\%$ ). The label-free proteomic discovery phase was conducted in 13 patients who progressed to MA by the 6-year visit and 11 control subjects, and four proteins (Tamm-Horsfall glycoprotein,  $\alpha$ -1 acid glycoprotein, clusterin, and progranulin) identified in the discovery phase were measured by enzyme-linked immunosorbent assay in 74 subjects: group A, normal renal function ( $n = 35$ ); group B, ERFD without MA ( $n = 15$ ); group C, MA without ERFD ( $n = 16$ ); and group D, both ERFD and MA ( $n = 8$ ).

**RESULTS**—In the label-free analysis, a model of progression to MA was built using 252 peptides, yielding an area under the curve (AUC) of  $84.7 \pm 5.3\%$ . In the validation study, ordinal logistic regression was used to predict development of ERFD, MA, or both. A panel including Tamm-Horsfall glycoprotein (odds ratio 2.9, 95% CI 1.3–6.2,  $P = 0.008$ ), progranulin (1.9, 0.8–4.5,  $P = 0.16$ ), clusterin (0.6, 0.3–1.1,  $P = 0.09$ ), and  $\alpha$ -1 acid glycoprotein (1.6, 0.7–3.7,  $P = 0.27$ ) improved the AUC from 0.841 to 0.889.

**CONCLUSIONS**—A panel of four novel protein biomarkers predicted early renal damage in type 1 diabetes. These findings require further validation in other populations for prediction of renal complications and treatment monitoring.

*Diabetes Care* 35:549–555, 2012

Despite tremendous progress in treatment, patients with type 1 diabetes still live 15 years shorter, experience excess morbidity, and have medical costs over 10 times higher than the general population (1). Modification of traditional risk factors has made only limited impact on development of renal complications in these patients (2). Improved glycemic control reduces progression to early stages of microvascular complications in type 1 diabetes even years later

(2), but the risk of diabetic nephropathy is not completely abolished by improved glycemic control (3) or lowering blood pressure (4). Whereas rates of end-stage renal disease have declined over recent years, after 30 years of diabetes duration, the cumulative incidence of end-stage renal disease is 7.8% among patients with type 1 diabetes (5). Further, the presence of microvascular complications is associated with increased risk of developing cardiovascular disease (6,7), which remains

the leading cause of death in people with type 1 diabetes.

The development of renal complications and eventual end-stage renal disease in people with diabetes has been thought to follow a linear path involving the development of microalbuminuria, progressing to frank proteinuria. It was thought that renal function only declines once proteinuria is present and that regression of microalbuminuria indicates reversal of the disease progress. However, clinical trials have recently demonstrated that this dogma may be incorrect (4,8,9). First, renal function as measured by glomerular filtration rate (GFR) declines before the development of proteinuria, demonstrating that there is an earlier phase of kidney damage that could be detected and targeted with interventions. Second, kidney damage can progress even when microalbuminuria has regressed (10). The identification of novel biomarkers for early renal disease is therefore a high priority.

Urine is an easily accessible biological matrix for which to identify new biomarkers. The urinary proteome is very complex, containing 1,000 of proteins and peptides that are derived from filtration/secretion/reabsorption processes within the kidney (11,12). Many of these proteins are sensitive to alterations in the kidney, and urinary proteomics provides a platform from which to simultaneously quantify hundreds of urinary proteins.

In this analysis, we have performed a longitudinal label-free protein expression study to discover early biomarkers of diabetic kidney damage before the onset of disease. Here we describe a label-free protein expression analysis in adults with type 1 diabetes who developed microalbuminuria or a significant decline in estimated glomerular filtration rate (eGFR) over 6 years of follow-up. The purpose of this study was to define a panel of proteins that can serve as novel biomarkers of early development of renal disease. In addition, we verified a selected panel of proteins identified from the discovery analysis via enzyme-linked immunosorbent assay

From the <sup>1</sup>Center for Proteomics, Case Western Reserve University, Cleveland, Ohio; and the <sup>2</sup>Barbara Davis Center for Childhood Diabetes, University of Colorado Denver, Aurora, Colorado.

Corresponding author: Janet K. Snell-Bergeon, janet.snell-bergeon@ucdenver.edu.

Received 9 August 2011 and accepted 22 November 2011.

DOI: 10.2337/dc11-1491

The article contains Supplementary Data online at <http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc11-1491/-DC1>.

© 2012 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. See <http://creativecommons.org/licenses/by-nc-nd/3.0/> for details.

(ELISA) in a larger cohort of patients with type 1 diabetes who developed MA and/or early renal decline.

## RESEARCH DESIGN AND METHODS

### Study participants

Between April 2000 and April 2002, the Coronary Artery Calcification in Type 1 Diabetes (CACTI) study enrolled 652 type 1 diabetic patients meeting the following criteria: age 19–56 years; no history of myocardial infarction, angioplasty, coronary artery bypass graft, or angina; currently on insulin therapy; diagnosed before age 30 year or a clinical course consistent with type 1 diabetes; on insulin therapy within the first year after diagnosis; and longstanding type 1 diabetes (range of duration 4–52 years). This cohort represents nearly 40% of eligible patients in the Denver metro area. Study participants were examined at baseline, 3 years, and 6 years.

**Label-free analysis.** There were 465 study participants with type 1 diabetes who were normoalbuminuric at baseline. Over 6 years of follow-up, 25 study participants with type 1 diabetes who were initially normoalbuminuric developed micro- or macroalbuminuria (MA), defined by albumin excretion rate (AER)  $\geq 20$   $\mu\text{g}/\text{min}$  or albumin-to-creatinine ratio  $\geq 30$   $\text{mg}/\text{g}$  in two overnight urine collections at the 6-year visit. We randomly selected half ( $n = 13$ ) of these study participants to conduct a proteomic discovery phase. Participants with type 1 diabetes who remained normoalbuminuric (AER  $< 7.5$   $\mu\text{g}/\text{min}$ ) at all three study visits in both overnight urine samples were frequency matched on age and diabetes duration ( $n = 11$ ). A total of 72 samples for the 24 study participants were included in this study, where each visit was considered an individual sample.

**ELISA validation study.** This validation study analyzed urine from 74 patients with type 1 diabetes at the baseline visit (visit 1). All subjects with the development of new MA from baseline to either visit 2 or visit 3 who were not previously included in the label-free analysis ( $n = 24$ ) and all subjects with early renal function decline (ERFD) from baseline to visit 2 or visit 3 were included in the validation study. ERFD was defined using the method outlined by Perkins and Krolewski (13), defined as a decline in cystatin C–based eGFR of  $\geq 3.3\%$  per year. There were 23 subjects in the CACTI

study who experienced ERFD, and all were selected for the ELISA validation study. There were 35 subjects with type 1 diabetes who completed all three study visits, had baseline urine samples available, and remained normoalbuminuric (AER  $< 7.5$   $\mu\text{g}/\text{min}$ ) and did not experience ERFD over the course of the study. All of these subjects were included as a control group with normal renal function. Study participants were categorized into groups as follows: group A, normal renal function ( $n = 35$ ); group B, ERFD but no MA ( $n = 15$ ); group C, MA but no ERFD ( $n = 16$ ); and group D, both ERFD and MA ( $n = 8$ ).

### Urine samples

**Overnight urine collections.** Study subjects were asked to collect two overnight timed urine collections at each visit. Subjects were asked to record the time at night that they last voided their bladder and to collect all urine produced overnight and first thing in the morning. The total time of the collection and the volume were recorded, and AER was calculated. In the event that subjects were unable to complete both timed samples, a spot urine sample was collected and albumin-to-creatinine ratio was calculated as an estimate of AER.

### Label-free expression

**Sample preparation.** Four milliliters of urine from the CACTI study subjects was desalted using a Microcon concentrator. Each sample was buffer exchanged three times using 3 mL of 50 mmol/L Tris, pH 8.8, to a final volume of  $\sim 100$   $\mu\text{L}$ , and protein concentrations were determined by a 2D Quant Kit (GE Healthcare, Piscataway, NJ). Ten micrograms of each sample was loaded onto a one-dimensional SDS-PAGE gel (4–20% Tris-HCl) as a quality-control measure for the desalting step. Subsequent to digestion, each sample was adjusted to 10  $\mu\text{g}$  in 50  $\mu\text{L}$ . Ten microliters of 0.2% Rapigest (Waters, Milford, MA) and dithiothreitol was added to a final concentration of 5 mmol/L. The samples were reduced at 80°C for 15 min and cooled to room temperature before alkylation with iodacetamide at a final concentration of 10 mmol/L for 30 min. Proteolytic digestion was performed with endopeptidase Lys C (Wako Chemicals, Richmond, VA) with a final enzyme-to-protein ratio of 1:10 (w/w) for 18 h at 37°C.

**Liquid chromatography and mass spectrometry.** A total of 100 nanograms of each sample was analyzed by liquid

chromatography/mass spectrometry/mass spectrometry (LC/MS/MS), and the order of sample injections was randomized over all samples. Separation and detection of peptides was performed as previously published (14). Raw LC/MS/MS data were processed via Proteomarker software (Infochromics, Toronto, Canada).

**Data processing: qualitative and quantitative.** The raw data for each run were first extracted to provide MS/MS peak lists for identification- and intensity-based profile peak lists for quantification. The MS/MS peak lists were subsequently searched by Mascot version 2.2.0 (Matrix Science, London, U.K.). The database used was the human International Protein Index (68,020 sequences). Search settings were as follows: no enzyme specificity, mass accuracy window for precursor ion, 10 ppm; mass accuracy window for fragment ions, 0.8 Da; and variable modification, including only carbamidomethylation of cysteines and oxidation of methionine. The criteria for peptide identification were a mass accuracy of  $\leq 10$  ppm and an expectation value of  $P \leq 0.05$ . Proteins that had two or more peptides matching the above criteria were considered confirmed assignments, whereas proteins identified with one peptide regardless of the Mascot score were highlighted as tentative assignments. Automated differential quantification of peptides in a set of samples was accomplished with Proteomarker as previously described (14).

**Data quality control: prefiltering, imputation, and normalization.** Subsequent to raw data acquisition and processing, data quality control and prefiltering were done for this study as previously described (14). In brief, three-step prefiltering was performed to resolve some of the peak misalignment issues and remove those peptides in the abundance matrix that received poor quality identifications or no qualitative identification at all. A peptide was rejected if 1) its consensus sequence was not assigned at all or 2) the consensus peptide sequence score was below the 74th percentile of all scores (Mascot score of  $\leq 21.125$ ). Next, intensity summaries of identical sequence peptides were integrated with annotations retained from the diffset with the least number of missing values (including retention time, charge/mass ratio, sequence, score, and protein annotation). A final prefiltering procedure was carried out to retain those peptides only for which the observed missing count per peptide was strictly  $< 50\%$  per experimental unit while

maximizing the total number of peptides remaining after selection. The final number of peptides retained was  $P = 2,584$  (1,360 proteins).

**Missing value imputation.** Missing values in LC/MS data arise because of imperfect detection and alignment of peak intensities or by true absence of expression. To account for the nonrandom nature of the “missingness mechanism” at play (nonignorable left-censoring) and its extent in this type of data (nonignorable left-censoring), we used a probability model adapted from Wang et al. (15) that describes “artifactual missing events.” This model makes inferences on the missing values of one sample on the basis of the information from other similar samples (technical replicates or nearest neighbors). It substitutes a missing measurement of intensity with its expected value of the true intensity, given that it is unobservable. Estimation of the imputation parameters was done to minimize the percentage of remaining missing values. The initial number of missing values after the above prefiltering was 60.8%. Remaining missing values after imputation (42.2%) represent truly absent peptides in the samples and were typically imputed by taking an estimate of the background noise.

**Variance stabilization of the data features.** To remove sources of systematic variation due to experimental artifacts in the measured intensities and to ensure that the usual assumptions for statistical inferences are met (normality, homoscedasticity), we applied a variance stabilization and normalization transformation on the variables (peptides). We used the joint adaptive mean variance regularization procedure recently introduced by Dazard and Rao (16). This method overcomes the lack of degrees of freedom and issues with variance-mean dependency common in high-dimensional proteomics datasets where the number of variables dominates the number of samples.

### Statistical analysis

**Unsupervised analyses.** Potential groups and outliers among the samples were checked by a principal component analysis (Supplementary Fig. 1) (17). Clustering analysis was performed using complete linkage hierarchical clustering and the gap statistics to estimate the real number of clusters in the data.

**Predictive proteomics model.** Patients were relabeled at visit 1 with their MA status at visit 3. This new response variable was then regressed onto all peptide

expression levels at visit 1 by fitting a generalized (logistic) linear model via penalized maximum likelihood (elastic net regularization) (18). Fitting was carried out by fourfold cross-validation with the help of the R implementation in the “glmnet” CRAN package (<http://cran.r-project.org>). This step allowed the selection of peptide predictors as early as visit 1 with best predictive value of progression toward MA at visit 3, as well as the determination of individual probability of MA progression by the patient.

**ELISA verification of target urine biomarkers.** Four putative urine biomarkers identified in the label-free proteomic analysis as having significant abundance at visit 1 were selected for verification via ELISA, and an additional protein progranulin was examined, which was identified in the analysis but not found significant at visit 1. Urine samples collected at visit 1 from a total of 74 study subjects were analyzed. The five biomarkers were measured by commercially available ELISA kits according to the manufacturer’s instructions: Tamms-Horsfall glycoprotein (THP) (MD Biosciences, St. Paul, MN); human progranulin (R&D Systems, Minneapolis, MN); clusterin (BioVendor, Chandler, NC); human  $\alpha$ -1 acid glycoprotein (AGP) (Assaypro, St. Charles, MO), and prostaglandin D synthase (ProstD) (BioVendor). All kits have inter- and intra-assay coefficients of variation of <15%.

## RESULTS

### Label-free expression analysis

Study participants in the label-free expression analysis who developed MA did not differ from those who remained normoalbuminuric over the 6 years of the study in terms of age, diabetes duration, sex, HbA<sub>1c</sub>, or baseline AER.

The sample preparation protocol was reproducible across individual samples and yielded sufficient protein concentrations with ranges of 0.206 to 39  $\mu\text{g}/\mu\text{L}$ . Reproducible protein patterns via one-dimensional SDS-PAGE were observed across all samples (data not shown). These samples were subsequently digested and analyzed by LC/MS/MS as described in RESEARCH DESIGN AND METHODS. Distinct chromatographic differences were observed in normoalbuminuric and MA samples (Supplementary Fig. 2). Good proteome coverage was observed with 1,115 tentative protein assignments (at least one peptide sequenced with reproducible chromatographic entities) and

246 confirmed protein assignments (at least two or more peptide sequenced). However, two samples were excluded from subsequent analyses because of poor LC/MS/MS data acquisition, leaving 22 samples for predictive model building.

To build a predictive model of the progression to MA among normoalbuminuric patients, we used data from visit 1 (252 peptides corresponding to 183 proteins) as predictors and the MA status at visit 3 as the outcome. This step formed the basis of a predictive proteomics model, which determines the individual probability of MA progression by the patient, whether the patient has already been observed or is incoming (new). The model yielded an area under the receiver operating curve (AUC) of  $84.7 \pm 5.3\%$  with a true positive rate of  $84.7 \pm 12.7\%$ , which corresponded to 19 of 22 correctly classified at visit 1. An equal distribution of increasing and decreasing peptide abundances was observed in patients who progressed to MA compared with patients who did not develop MA. Overall, 148 peptides decreased in abundance in the MA group, while 104 peptides increased.

Good correlation was observed between albumin peptide abundance and label-free albumin measures at visit 1 with values of AER measured previously in the urine samples. The median AER value for the control and MA groups was 5.64 and 9.42, respectively, whereas the label-free median peptide intensities (normalized and transformed scale) were 0.67 for the control subjects and 0.85 for the MA group. Additional example proteins that were found to be significant in the MA group at visit 1 were AGP, THP, clusterin, and ProstD (Supplementary Fig. 3). In addition, progranulin that was detected in the LC/MS/MS analysis was included in the verification analysis, since it has a similar expression profile to THP in the kidney tubule, and we detected changes in abundances of this protein in type 1 diabetic patients with MA (data not shown) (19).

**ELISA protein analysis.** Characteristics of study participants at the baseline exam were examined in patients with type 1 diabetes who remained normoalbuminuric and had normal renal function (group A,  $n = 35$ ), patients with type 1 diabetes who developed ERFD without MA (group B,  $n = 15$ ), patients with type 1 diabetes who went on to develop MA but not ERFD (group C,  $n = 16$ ), and patients with type 1 diabetes who went on to develop

both MA and ERFD (group D,  $n = 8$ ) (Table 1). There were no differences in age, diabetes duration, sex, systolic or diastolic blood pressure, total or HDL cholesterol, or waist-to-hip ratio (WHR) across groups. HbA<sub>1c</sub> was significantly lower in group A than in group C, and the use of antihypertensive medication was significantly lower in group A than in group D, whereas lipid medication use was significantly lower in group A than in group C. Baseline AER was significantly lower in group A than in groups C and D and in group B than in group D. There were no differences in baseline cystatin C or eGFR by group, but as per the study design, follow-up eGFR was significantly lower in the group with ERFD (group B) than in patients with normal renal function and no albuminuria (group A), and in the group with both ERFD and albuminuria (group D) than in the group with only albuminuria (group C).

Levels of THP, progranulin, clusterin, AGP, and ProSD were compared in groups A, B, C, and D, adjusted for age and urine creatinine (Table 2). THP was significantly higher in group D than in all other groups. AGP and ProSD were significantly higher in group B than in group A.

Standardized  $z$  scores were calculated for all proteins and were examined by group, adjusting for baseline age, diabetes duration, baseline AER, HbA<sub>1c</sub>, cystatin C, and uric acid. Significant differences were observed between groups A and D for THP and progranulin and groups A and B for AGP (Fig. 1). Both THP and progranulin followed a stepwise pattern, with the lowest levels in patients who maintained normal renal function and normoalbuminuria (group A), increasing nonsignificantly in patients with either ERFD (group B) or MA (group C) and significantly increased in patients with both ERFD and MA (group D). AGP was also lowest among patients who maintained normal renal function (group A), but was significantly increased only in patients who developed ERFD with normoalbuminuria (group B). No group differences were observed for clusterin or ProSD (data not shown).

In multivariable ordinal logistic regression modeling with study group as the outcome, THP and progranulin were significantly predictive of ERFD and MA in patients with type 1 diabetes, adjusting for baseline age, diabetes duration, uric acid and cystatin C, total and HDL cholesterol, and systolic blood pressure

as well as factors [baseline AER, BMI, ever smoking, WHR  $\times 10$ , and  $\log(\text{HbA}_{1c}) \times 10$ ] previously found to predict microalbuminuria in multiple cohorts with type 1 diabetes (20) (Table 3). AGP was marginally ( $P = 0.07$ ) predictive of ERFD and MA. In a model using stepwise selection to determine the most parsimonious model and including model 1, THP, progranulin, and clusterin were selected. Finally, a model was considered that included THP, progranulin, clusterin, and AGP. In this model, THP was significantly associated with renal outcomes, but clusterin, progranulin, and AGP were not statistically significantly predictive of the development of ERFD and MA. The C-statistic, a measure of model fit, increased from 0.841 in model 1 (without any urinary proteins except for AER) to 0.857 with progranulin, 0.871 with THP, and 0.888 in the stepwise selection model, which included model 1 + THP, progranulin, and clusterin, and to 0.889 in the final model, which included also AGP.

**CONCLUSIONS**—In this validation study using ELISA, we were able to confirm the value of several proteins previously identified through the label-free

Table 1—Characteristics of ELISA study participants at baseline examination

	Patients with normal renal function and no albuminuria (group A)	Patients with only ERFD (group B)	Patients who progressed to only albuminuria (group C)	Patients with both ERFD and albuminuria (group D)
<i>n</i>	35	15	16	8
Age (years)	34 ± 9	36 ± 10	37 ± 9	36 ± 9
Diabetes duration (years)	19 ± 7	24 ± 11	19 ± 5	23 ± 8
Sex (% male)	49	47	38	25
HbA <sub>1c</sub> (%)	7.7 ± 0.9 <sup>b</sup>	8.8 ± 2.0	8.9 ± 1.3	8.3 ± 1.5
Systolic blood pressure (mmHg)	117 ± 17	118 ± 18	118 ± 11	115 ± 12
Diastolic blood pressure (mmHg)	76 ± 9	76 ± 11	79 ± 8	73 ± 7
Antihypertensive medication use (% yes)	17 <sup>c</sup>	20	31	63
Total cholesterol (mg/dL)	167 ± 33	173 ± 44	180 ± 43	170 ± 33
HDL cholesterol (mg/dL)	57 ± 22	56 ± 19	56 ± 18	54 ± 13
LDL cholesterol (mg/dL)	94 ± 28	98 ± 38	100 ± 31	98 ± 25
Lipid medication use (% yes)	11 <sup>b</sup>	0	38	13
WHR	0.82 ± 0.09	0.80 ± 0.09	0.83 ± 0.12	0.81 ± 0.07
BMI (kg/m <sup>2</sup> )	25.0 ± 3.2 <sup>b,c</sup>	25.4 ± 3.8	27.6 ± 6.0	28.4 ± 5.2
Ever smoker (%)	14	7	13	0
Baseline AER (μg/min)	5.5 ± 3.8 <sup>b,c</sup>	7.0 ± 4.2 <sup>d</sup>	10.6 ± 4.9	9.8 ± 4.9
Final AER (μg/min)	5.0 ± 2.8 <sup>b,c</sup>	5.7 ± 4.3 <sup>d,e</sup>	61 ± 71	483 ± 809
Baseline cystatin C (mg/L)	0.66 ± 0.08	0.64 ± 0.16	0.65 ± 0.10	0.33 ± 0.14
Baseline eGFR (ml/min per 1.73 m <sup>2</sup> )	127 ± 16	136 ± 29	130 ± 20	136 ± 29
Final eGFR (ml/min per 1.73 m <sup>2</sup> )	133 ± 16 <sup>a</sup>	113 ± 24 <sup>d</sup>	135 ± 20 <sup>f</sup>	105 ± 33
Uric acid (mg/dL)	4.7 ± 0.9	5.0 ± 1.4	5.2 ± 1.2	5.1 ± 1.5

<sup>a</sup> $P < 0.05$  group A vs. group B. <sup>b</sup> $P < 0.05$  group A vs. group C. <sup>c</sup> $P < 0.05$  group A vs. group D. <sup>d</sup> $P < 0.05$  group B vs. group C. <sup>e</sup> $P < 0.05$  group B vs. group D. <sup>f</sup> $P < 0.05$  group C vs. group D.

Table 2—Urinary protein levels by ELISA, adjusted for baseline age, urine creatinine, and group

	Patients with normal renal function and no albuminuria (group A)	Patients with only ERFD (group B)	Patients who progressed to only albuminuria (group C)	Patients with both ERFD and albuminuria (group D)
<i>n</i>	35	15	16	8
THP (ng/mL)	5,490 ± 1,455 <sup>b</sup>	6,504 ± 2,214 <sup>c</sup>	5,737 ± 2,147 <sup>d</sup>	14,054 ± 3,031
Progranulin (ng/mL)	4.23 ± 0.81	4.79 ± 1.23	4.67 ± 1.19	7.60 ± 1.69
Clusterin (ng/mL)	1.76 ± 0.48	0.61 ± 0.73	0.65 ± 0.70	1.71 ± 1.0
AGP (ng/mL)	416 ± 144 <sup>a</sup>	1071 ± 219	830 ± 830	827 ± 299
ProstD (ng/mL)	296 ± 49 <sup>a</sup>	484 ± 75	318 ± 73	435 ± 10

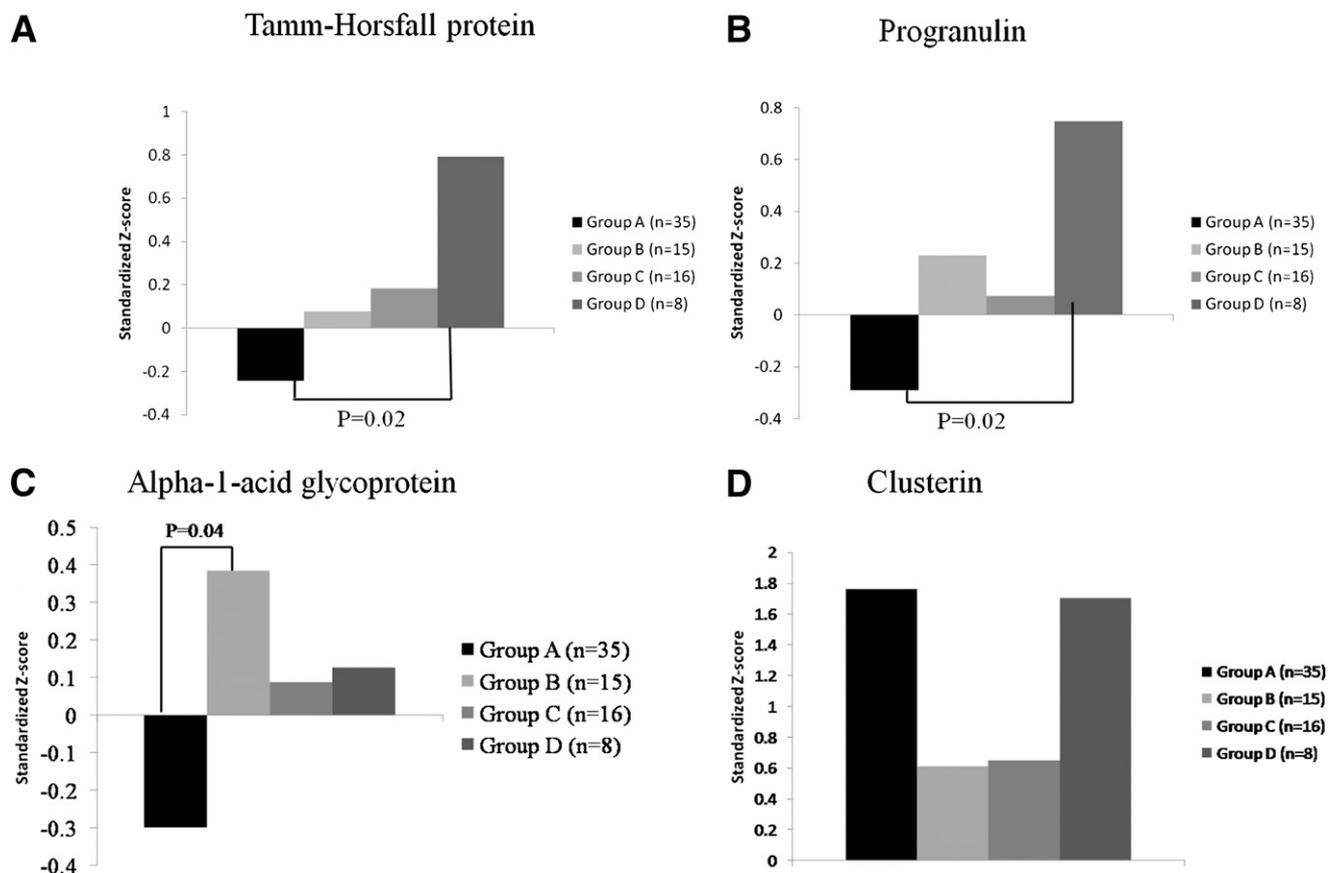
<sup>a</sup>*P* < 0.05 group A vs. group B. <sup>b</sup>*P* < 0.05 group A vs. group D. <sup>c</sup>*P* < 0.05 group B vs. group D. <sup>d</sup>*P* < 0.05 group C vs. group D.

protein expression as potential predictors of early diabetic nephropathy and that added to the model fit above and beyond previously identified risk markers. Discovery-based quantitative proteomics provide a powerful technique for identification and quantification in large-scale protein profiling for biomarker discovery, and multiple proteomic techniques are available (21–23). To date, most proteomic approaches have used two-dimensional

gel electrophoresis or capillary electrophoresis–mass spectrometry to classify type 1 diabetes and its complications (24–26).

Label-free protein expression is a peptide-based proteomic technique that capitalizes on the highly reproducible chromatography and accurate mass accuracy available in current LC/MS systems. This platform quantifies a peptide by its intensity and groups each peptide across

individual samples on the basis of its accurate mass and retention time (27,28). These intensities associated with specific mass and retention time values are organized into peptide array tables that may be further processed using statistical techniques. The label-free protein expression analysis in the current study provided a comprehensive view of proteomic changes during the development of microalbuminuria from which predictive models could be



**Figure 1**—Levels of proteins by ELISA by study group. A: Tamm-Horsfall protein levels by group. B: Progranulin protein levels by group. C: Alpha-1-acid glycoprotein levels by group. D: Clusterin protein levels by group.

**Table 3—Multivariable ordinal logistic regression analysis of urinary protein markers predicting renal outcome groups**

	OR* (95% CI)	P	C-statistic
Model 1 (baseline age, diabetes duration, HbA <sub>1c</sub> †, WHR†, BMI, ever smoker, cholesterol, HDL cholesterol, systolic blood pressure, antihypertensive treatment, anti-lipid treatment, AER, uric acid, and cystatin C)			0.841
Model 1 + THP	2.2 (1.1–4.1)	0.02	0.871
Model 1 + progranulin	1.9 (1.1–3.3)	0.03	0.857
Model 1 + AGP	1.6 (0.93–2.9)	0.09	0.851
Model 1 + ProsdD	1.3 (0.8–2.2)	0.38	0.846
Model 1 + clusterin	1.1 (0.6–1.8)	0.84	0.842
Model 1 +			0.888
THP	2.5 (1.2–5.4)	0.01	
Progranulin	2.5 (1.2–5.1)	0.01	
Clusterin	0.6 (0.3–1.1)	0.11	
Model 1 +			0.889
THP	2.9 (1.3–6.2)	0.008	
Progranulin	1.9 (0.8–4.5)	0.164	
Clusterin	0.6 (0.3–1.1)	0.090	
AGP	1.6 (0.7–3.7)	0.270	

All models were adjusted for model 1 (age, diabetes duration, HbA<sub>1c</sub>, AER, uric acid, and cystatin C). \*ORs are per 1-SD increase. †Transformed:  $\log(\text{HbA}_{1c}) \times 10$ ,  $\text{WHR} \times 10$ .

derived. The models highlight a number of clinically relevant proteins, as well as novel indicators of disease.

THP, which is produced in the thick ascending limb and the early distal convoluted tubule of the kidney, is the most abundant protein in normal urine. THP is the protein product encoded by the uromodulin gene (*UMOD*), which has been identified in genome-wide association scans with chronic kidney disease and GFR estimated from serum creatinine (29). Urinary THP has been suggested as a useful marker of renal damage and has been reported to be decreased in patients with type 1 diabetes (30–32) and in patients with kidney damage with and without diabetes (33). In the current study, we found that levels of THP were significantly increased in patients who developed both ERFD and albuminuria. Kottgen et al. (34) also found that higher levels of THP were associated with chronic kidney disease in a case-control study, with an odds ratio for chronic kidney disease of 1.72 per 1 SD increase in THP.

AGP has previously been reported as increasing in patients with diabetic nephropathy (35,36) and may serve as both an early marker of diabetic nephropathy as well as a marker of diabetic nephropathy progression. Prostaglandin D synthase, which was not predictive of renal outcomes

in the current study, was previously reported to be increased in patients with type 2 diabetes who had increased permeability of glomerular capillary walls, and higher levels of prostaglandin D synthase predicted albuminuria (37).

Other proteins included in the model (clusterin and progranulin), while not directly associated with diabetic nephropathy, were observed to be associated with renal toxicity and/or renal damage. Clusterin is a glycoprotein that may have a role in repairing kidney damage, since low levels of clusterin have been found to predict worse recovery from renal ischemia-reperfusion injury in mice (38). Clusterin may have a role in protecting organisms from apoptosis because of oxidative stress (39) and may prevent glomerulopathy associated with aging (40). Progranulin is a growth factor involved in wound healing and is known to have an anti-inflammatory effect (41). On the other hand, when progranulin is degraded into peptides by proteases, it has been shown to have a proinflammatory effect (42).

The current study provided important preliminary data on a panel of proteins that could be used to predict the early signs of diabetic nephropathy, including the development of micro- and macroalbuminuria as well as significant renal function decline. Further validation

of this protein panel is needed in other populations to verify their predictive ability for the development of both renal function decline and urinary albumin and to determine whether these could be used as biomarkers of disease progression and response to treatment.

**Acknowledgments**—This study was performed at the Barbara Davis Center for Childhood Diabetes in Denver, CO, and at the Case Center for Proteomics and Bioinformatics at Case Western Reserve University, Cleveland, OH. Support was provided by the National Institutes of Health (NIH) National Heart, Lung and Blood Institute grants R01 HL61753 and R01 HL079611, American Diabetes Association postdoctoral fellowship 7-09-CVD-06 (mass spectrometry), American Diabetes Association Junior Faculty Award 1-10-JF-50 (J.K.S.-B.), and Diabetes Endocrinology Research Center Clinical Investigation Core P30 DK57516. The study was performed at the Clinical Translational Research Center at the University of Colorado Denver, supported by NIH Grant M01 RR000051. Support was also provided in part from the Cleveland Center for Translational Science Collaborative of the NIH (UL1-RR024989).

No potential conflicts of interest relevant to this article were reported.

D.S. researched data and wrote the manuscript. D.M.M. and M.R.C. contributed to the discussion and reviewed and edited the manuscript. J.-E.D., X.L., and F.H. researched data and reviewed and edited the manuscript. M.R. reviewed and edited the manuscript. J.K.S.-B. researched data, wrote the manuscript, and contributed to the discussion. J.K.S.-B. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

## References

- Laing SP, Swerdlow AJ, Slater SD, et al. Mortality from heart disease in a cohort of 23,000 patients with insulin-treated diabetes. *Diabetologia* 2003;46:760–765
- White NH, Sun W, Cleary PA, et al. Prolonged effect of intensive therapy on the risk of retinopathy complications in patients with type 1 diabetes mellitus: 10 years after the Diabetes Control and Complications Trial. *Arch Ophthalmol* 2008;126:1707–1715
- Pambianco G, Costacou T, Orchard TJ. The prediction of major outcomes of type 1 diabetes: a 12-year prospective evaluation of three separate definitions of the metabolic syndrome and their components and estimated glucose disposal rate: the Pittsburgh Epidemiology of Diabetes Complications Study experience. *Diabetes Care* 2007;30:1248–1254

4. Mauer M, Zinman B, Gardiner R, et al. Renal and retinal effects of enalapril and losartan in type 1 diabetes. *N Engl J Med* 2009;361:40–51
5. Finne P, Reunanen A, Stenman S, Groop PH, Grönhagen-Riska C. Incidence of end-stage renal disease in patients with type 1 diabetes. *JAMA* 2005;294:1782–1787
6. Edwards NL. The role of hyperuricemia and gout in kidney and cardiovascular disease. *Cleve Clin J Med* 2008;75(Suppl. 5):S13–S16
7. van Hecke MV, Dekker JM, Stehouwer CD, et al. EURODIAB Prospective Complications Study. Diabetic retinopathy is associated with mortality and cardiovascular disease incidence: the EURODIAB Prospective Complications Study. *Diabetes Care* 2005;28:1383–1389
8. Jerums G, Panagiotopoulos S, Premaratne E, Power DA, MacIsaac RJ. Lowering of proteinuria in response to antihypertensive therapy predicts improved renal function in late but not in early diabetic nephropathy: a pooled analysis. *Am J Nephrol* 2008;28:614–627
9. Mann JF, Schmieder RE, McQueen M, et al.; ONTARGET Investigators. Renal outcomes with telmisartan, ramipril, or both, in people at high vascular risk (the ONTARGET study): a multicentre, randomised, double-blind, controlled trial. *Lancet* 2008;372:547–553
10. Steinke JM, Mauer M; International Diabetic Nephropathy Study Group. Lessons learned from studies of the natural history of diabetic nephropathy in young type 1 diabetic patients. *Pediatr Endocrinol Rev* 2008;5(Suppl. 4):958–963
11. Adachi J, Kumar C, Zhang Y, Olsen JV, Mann M. The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins. *Genome Biol* 2006;7:R80
12. Barratt J, Topham P. Urine proteomics: the present and future of measuring urinary protein components in disease. *CMAJ* 2007;177:361–368
13. Perkins BA, Krolewski AS. Early nephropathy in type 1 diabetes: the importance of early renal function decline. *Curr Opin Nephrol Hypertens* 2009;18:233–240
14. Schlatzer DM, Dazard JE, Dharsee M, et al. Urinary protein profiles in a rat model for diabetic complications. *Mol Cell Proteomics* 2009;8:2145–2158
15. Wang X, Li A, Jiang Z, Feng H. Missing value estimation for DNA microarray gene expression data by Support Vector Regression imputation and orthogonal coding scheme. *BMC Bioinformatics* 2006;7:32
16. Dazard J-E, Hua X, Rao JS. R package MVR for joint adaptive mean-variance regularization and variance stabilization. In *JSM Proceedings, Section for Statistical Programmers and Analysts*. Miami Beach, FL, American Statistical Association IMS–JSM. In press
17. Hotelling H. Relations between two sets of variables. *Biometrika* 1936;28:321–377
18. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Statist Soc* 2005;67(Series B):301–320
19. Daniel R, He Z, Carmichael KP, Halper J, Bateman A. Cellular localization of gene expression for progranulin. *J Histochem Cytochem* 2000;48:999–1009
20. Vergouwe Y, Soedamah-Muthu SS, Zgibor J, et al. Progression to microalbuminuria in type 1 diabetes: development and validation of a prediction rule. *Diabetologia* 2010;53:254–262
21. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol* 2006;24:971–983
22. Veenstra TD, Conrads TP, Hood BL, Avellino AM, Ellenbogen RG, Morrison RS. Biomarkers: mining the biofluid proteome. *Mol Cell Proteomics* 2005;4:409–418
23. Zhao Y, Lee WN, Xiao GG. Quantitative proteomics and biomarker discovery in human cancer. *Expert Rev Proteomics* 2009;6:115–118
24. Coon JJ, Zürbig P, Dakna M, et al. CE-MS analysis of the human urinary proteome for biomarker discovery and disease diagnostics. *Proteomics Clin Appl* 2008;2:964
25. Snell-Bergeon JK, Maahs DM, Ogden LG, et al. Evaluation of urinary biomarkers for coronary artery disease, diabetes, and diabetic kidney disease. *Diabetes Technol Ther* 2009;11:1–9
26. Mischak H, Delles C, Klein J, Schanstra JP. Urinary proteomics based on capillary electrophoresis-coupled mass spectrometry in kidney disease: discovery and validation of biomarkers, and clinical application. *Adv Chronic Kidney Dis* 2010;17:493–506
27. Chelius D, Bondarenko PV. Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *J Proteome Res* 2002;1:317–323
28. Wang W, Zhou H, Lin H, et al. Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal Chem* 2003;75:4818–4826
29. Köttgen A, Glazer NL, Dehghan A, et al. Multiple loci associated with indices of renal function and chronic kidney disease. *Nat Genet* 2009;41:712–717
30. Below AA, Chakraborty J, Khuder SH, Haselhuhn GD. Evaluation of urinary Tamm-Horsfall protein in post-menopausal diabetic women. *J Diabetes Complications* 1999;13:204–210
31. Torffvit O, Agardh CD, Thulin T. A study of Tamm-Horsfall protein excretion in hypertensive patients and type 1 diabetic patients. *Scand J Urol Nephrol* 1999;33:187–191
32. Torffvit O, Agardh CD. Tubular secretion of Tamm-Horsfall protein is decreased in type 1 (insulin-dependent) diabetic patients with diabetic nephropathy. *Nephron* 1993;65:227–231
33. Chakraborty J, Below AA, Solaiman D. Tamm-Horsfall protein in patients with kidney damage and diabetes. *Urol Res* 2004;32:79–83
34. Köttgen A, Hwang SJ, Larson MG, et al. Uromodulin levels associate with a common UMOD variant and risk for incident CKD. *J Am Soc Nephrol* 2010;21:337–344
35. Jiang H, Guan G, Zhang R, et al. Increased urinary excretion of orosomucoid is a risk predictor of diabetic nephropathy. *Nephrology (Carlton)* 2009;14:332–337
36. Matheson A, Willcox MD, Flanagan J, Walsh BJ. Urinary biomarkers involved in type 2 diabetes: a review. *Diabetes Metab Res Rev* 2010;26:150–171
37. Uehara Y, Makino H, Seiki K, Urade Y; L-PGDS Clinical Research Group of Kidney. Urinary excretions of lipocalin-type prostaglandin D synthase predict renal injury in type-2 diabetes: a cross-sectional and prospective multicentre study. *Nephrol Dial Transplant* 2009;24:475–482
38. Zhou W, Guan Q, Kwan CC, et al. Loss of clusterin expression worsens renal ischemia-reperfusion injury. *Am J Physiol Renal Physiol* 2010;298:F568–F578
39. Trougakos IP, Gonos ES. Regulation of clusterin/apolipoprotein J, a functional homologue to the small heat shock proteins, by oxidative stress in ageing and age-related diseases. *Free Radic Res* 2006;40:1324–1334
40. Rosenberg ME, Girton R, Finkel D, et al. Apolipoprotein J/clusterin prevents a progressive glomerulopathy of aging. *Mol Cell Biol* 2002;22:1893–1902
41. He Z, Bateman A. Progranulin (granulin-epithelin precursor, PC-cell-derived growth factor, acrogranin) mediates tissue repair and tumorigenesis. *J Mol Med (Berl)* 2003;81:600–612
42. Kojima Y, Ono K, Inoue K, et al. Progranulin expression in advanced human atherosclerotic plaque. *Atherosclerosis* 2009;206:102–108