Research Article

Urine Glycoprotein Profile Reveals Novel Markers for Chronic Kidney Disease

Anuradha Vivekanandan-Giri,1 Jessica L. Slocum,1 Carolyn L. Buller,1 Venkatesha Basrur,2 Wenjun Ju,1 Rodica Pop-Busui,3 David M. Lubman,2,4,5 Matthias Kretzler,1,4 and Subramaniam Pennathur1,4

1 Division of Nephrology, University of Michigan, Ann Arbor, MI 48105, USA
2 Department of Pathology, University of Michigan, Ann Arbor, MI 48105, USA
3 Division of Metabolism, Endocrinology and Diabetes, Department of Internal Medicine, University of Michigan, Ann Arbor, MI 48105, USA
4 Department of Computational Medicine and Biology, University of Michigan, Ann Arbor, MI 48105, USA
5 Department of Surgery, University of Michigan, Ann Arbor, MI 48105, USA

Correspondence should be addressed to Subramaniam Pennathur, spennath@umich.edu

Received 30 June 2011; Accepted 30 July 2011

Chronic kidney disease (CKD) affects approximately 11% of the US population with over 100,000 individuals progressing to end-stage renal disease (ESRD) annually [1, 2]. Despite this significant and growing public health problem, it remains difficult to predict which individuals will progress to ESRD. As ESRD carries a substantial increase in morbidity and mortality, it is critical to identify this high-risk patient population that would most benefit from early and aggressive therapy.

Current strategies for predicting CKD progression are limited. Pathologic examination of renal tissue provides valuable information on degree of interstitial fibrosis and predilection for ESRD. However, renal biopsy is invasive with a limited role for longitudinal followup. Quantitative measures of proteinuria have long been used as noninvasive markers of CKD progression [3], yet these largely albumin-based methods detect nonselective proteinuria and incompletely correlate with disease. With recent advances in high through-put technology and mass spectrometry techniques, urine proteomic investigation is an attractive tool in the pursuit for noninvasive and specific markers of CKD progression [4, 5].

Numerous investigators have successfully applied broad-scale urine proteomic strategies to kidney disease. The urine
Table 1: Patient characteristics of study subjects.

| Variable                  | Healthy control (n = 6) | CKD (n = 6) | P     |
|---------------------------|-------------------------|-------------|-------|
| Age (years)               | 46.3 (13.5)             | 47.2 (14.2) | 0.92  |
| Sex (male/female)         | 2/4                     | 2/4         | 1.00  |
| Body mass index (kg/m²)   | 24.3 (3.0)              | 30.5 (4.8)  | 0.02  |
| Serum creatinine (mg/dL)  | 0.85 (0.16)             | 1.75 (1.09) | 0.07  |
| eGFR (mL/min)*            | 83.0 (15.0)             | 52.0 (27.4) | 0.05  |
| Protein/creatinine ratio  | 0.03 (0.02)             | 2.15 (1.44) | 0.01  |

All data expressed as mean ± SD.

*eGFR estimated glomerular filtration rate.

proteome predicts nephropathy and decline in renal function in diabetic subjects [6, 7]. It also correlates with early changes of focal segmental nephrosclerosis [8], can identify IgA nephropathy and renal allograft rejection [9, 10], and predicts treatment response and disease activity in nephrotic syndrome and lupus nephritis [11, 12]. Despite these advances, analysis of the entire urine proteome is particularly difficult in CKD. With disruption of the glomerular filtration barrier and leakage of abundant plasma proteins into the urine, a nonselective, largely albumin predominant, pattern often results [13]. To overcome this, methods to increase the detection of low-abundance proteins have been developed to provide disease specificity and clinical relevance of urine profiling and to mechanistically understand factors influencing disease progression.

Glycoprotein enrichment techniques allow depletion of albumin and other abundant plasma proteins while providing a more thorough analysis of a subfraction of the urine proteome. As glycosylated proteins are critical for cellular interactions and signaling cascades, disease states are likely to cause early and specific alterations in urinary glycoprotein excretion. Indeed, glycoproteins are now important markers of autoimmunity and malignancy [14, 15]. More recently, the plasma glycoproteome has been used to predict nephropathy in diabetic subjects [16]. Despite this promising role as a noninvasive and specific biomarker of disease, little is known about the urinary glycoproteome in CKD.

We hypothesized that the urinary glycoproteome would be altered in CKD compared to healthy controls and that specific glycoprotein alterations might be useful in predicting CKD progression. The overall goal of this study was to perform an initial exploratory analysis of the urine glycoproteins in patients with CKD compared to healthy controls. We present a comprehensive profiling of the urinary glycoproteome in control and CKD subjects utilizing a hydrazide enrichment technique combined with tandem mass spectrometry identification of the glycoproteins.

2. Methods

2.1. Sample Collection and Processing. Clean catch urine samples were obtained from six CKD subjects and six age-matched healthy controls following written informed consent approved by the University of Michigan Institutional Review Board. Samples were stored at −80°C and thawed immediately prior to proteomic analysis. An initial 5000 g centrifugation was performed at 4°C for 10 minutes to remove cellular debris. Approximately, 30–50 mL healthy control samples and 1-2 mL CKD samples were concentrated using a 3 kDa filter cut-off membrane (Vivaspin 3 kDa MWCO, GE healthcare, Buckinghamshire, UK and Amicon ultra 0.5 mL, Millipore, Ireland resp.). As CKD subjects had higher urinary protein content (Table 1), the processed volumes were lower.

Urine protein concentration was determined using Coomassie Protein Assay Reagent with BSA standard (Thermo Scientific, Rockford, Illinois). 200 μg of concentrated protein were utilized for downstream processing. Protein samples were exchanged into 50 mM ammonium bicarbonate buffer (pH 7.4). Urine creatinine concentration was determined by tandem mass spectrometry (MS/MS) as described previously by our group [17]. To determine the level of creatinine, a known amount of [2H3]creatinine was spiked into each sample. A full-scan mass spectrum revealed molecular ions of m/z 114 and 117 for authentic creatinine and [2H3]creatinine, respectively. The transitions of the m/z 114 to 44 and m/z 117 to 47 were monitored in multiple-reaction monitoring mode for authentic and [2H3]creatinine, respectively, utilizing an Agilent Technologies (New Castle, DE) 6410 Triple Quadrupole mass spectrometer system, equipped with an Agilent 1200 series HPLC.
Table 2: Urinary glycoproteins unique to CKD or healthy control subjects.

| Unique proteins in healthy controls | Unique proteins in CKD |
|-------------------------------------|-------------------------|
| 70 kDa lysosomal alpha-glucosidase (GAA) | Antithrombin-III (SERPINC1) |
| Alpha-1B-glycoprotein (A1BG) | Complement factor H-related 1 (CFHR1) |
| Basigin (BSG) | Desmoglein-2 (DSG2) |
| Beta-galactosidase (GLB1) | Lumican (LUM) |
| Beta-sarcoglycan (SGCB) | Lymphatic vessel endothelial hyaluronic acid receptor 1 (LYVE1) |
| Butyrophilin (BTN2A1) | Pigment epithelium-derived factor (SERPINF1) |
| Carboxypeptidase M (CPM) | Thyroxine-binding globulin (SERPINA7) |
| CD276 antigen (CD276) | Zinc-alpha-2-glycoprotein (AZP1) |
| Complement component C4B (C4B) | |
| Cubilin (CUBN) | |
| Colony stimulating factor 1 (macrophage) (CSF1) | |
| Delta and notch-like epidermal growth factor-related receptor (DNER) | |
| Desmocollin-2 (DSC2) | |
| Desmoglein-1 (DSG1) | |
| Epidermal growth factor (EGFR) | |
| Secreted frizzled-related protein-4 (SFRP4) | |
| Fibronectin 1 (FN1) | |
| Folate receptor alpha (FOLR1) | |
| Golgi phosphoprotein 2 (GOLPH2) | |
| Glutamyl aminopeptidase (ENPEP) | |
| Hepatitis B virus receptor binding protein (Q6PYX1) | |
| Hepatic asialoglycoprotein receptor 1 transcript variant b (ASGR1) | |
| Heparan sulfate proteoglycan 2 (HSPG2) | |
| Intercellular adhesion molecule 1 (ICAM1) | |
| Kallikrein-1 (KLK1) | |
| Kallikrein 3 (APS) | |
| Lysosomal alpha-glucosidase (GAA) | |
| Lysosomal-associated membrane protein 2 (LAMP2) | |
| Maltese-glucosamylase (MGAM) | |
| Microfibril-associated glycoprotein 4 (MFP4) | |
| Mucin-6 (MUC6) | |
| Neuronal pentraxin receptor (NPTXR) | |
| Neuropilin and tolloid-like protein 1 (NETO1) | |
| Probable serine carboxypeptidase (CPVL) | |
| Sex hormone binding globulin (SHBG) | |

system. The creatinine concentration in the urine sample was determined by comparing the peak areas for authentic and $[^2]$H$_3$ creatinine for the above transitions.

2.2. Glycoprotein Separation and Enrichment. In order to assess recovery following the enrichment procedure, 5 μg of invertase from Saccharomyces cerevisiae (Sigma, St. Louis, MO) was spiked into 200 μg of protein in every sample. Glycoproteins were enriched from urinary proteins utilizing the hydrazide resin capture protocol as described previously by Zhang et al. [18]. Briefly, samples were oxidized with 10 mM sodium metaperiodate then incubated with hydrazide resin overnight at room temperature. Samples were then centrifuged at 3000 g for 2 minutes and the resin was washed successively with equal volumes 50 mM ammonium bicarbonate buffer (pH 7.4; Buffer A) supplemented with 8 M urea, followed by Buffer A alone and then water. The beads were resuspended in water, and the protein was reduced with 5 mM DTT followed by alkylation with 15 mM iodoacetamide. Trypsin (sequencing grade modified trypsin, Promega Corporation, Madison, WI) at 1:20 μg ratio was added to the samples and incubated overnight at 37°C for digestion. Following digestion, the beads were centrifuged at 3000 g for 2 minutes and the resin was then washed successively with 1.5 M NaCl, 80% acetonitrile, 100% methanol, and Buffer A. The resin was then resuspended in Buffer A and incubated with 5 units of PNGaseF (New England Biolabs, Ipswich, MA) overnight at 37°C for glycopeptide release. The glycopeptides were cleaned using a reverse phase column and
Figure 2: Mass Spectra of glycopeptides derived from Zinc alpha 2 Glycoprotein (a) and Golgi phosphoprotein (b) in CKD urine. Tryptic digests of urine glycoproteins were subjected to LC-ESI-MS/MS analysis as described in Section 2. The mass spectra of peptides DIVEYNDSNGSHVLQGR from zinc alpha 2 glycoprotein which is upregulated (a) and those of peptide AVLVNiTTGER from Golgi phosphoprotein which is significantly downregulated in CKD subjects (b) are shown. The N-linked glycosylation site of each peptide is depicted in red.

eluted with 50% acetonitrile/0.1% TFA followed by elution with 80% acetonitrile/0.1% TFA. The peptides were then dried at 60°C in a vacuum centrifuge and stored for mass spectrometric analysis.

2.3. Liquid Chromatography Electrospray Ionization (ESI/LC) MS/MS Analysis. Peptide samples were resuspended in 0.1% formic acid and loaded onto an in-house packed reverse phase separation column (0.075 × 100 mm, MAGIC C18 AQ particles, 5 μm, Michrom Bioresources). The peptides were separated on a 1% acetic acid/acetonitrile gradient system (5–50% acetonitrile for 75 min, followed by a 10 min 95% acetonitrile wash) at a flow rate of ~300 nl/min. Peptides were directly sprayed onto the MS using a nanospray source. An LTQ Orbitrap XL (Thermo Fisher Scientific, Waltham, MA) was run in automatic mode collecting a high resolution MS scan (FWHM 30,000) followed by data-dependent acquisition of MS/MS scans on the 9 most intense ions (relative collision energy ~35%). Dynamic exclusion was set to collect 2 MS/MS scans on each ion and exclude it for an additional 2 min. Charge state screening was enabled to exclude +1 and undetermined charge states.

2.4. Data Processing and Statistical Analysis. The Human UniProt database (Release 2011-5) was appended with a reverse database, a common contaminant list, and yeast invertase. Raw files were converted to mzXML format and searched against the database using X!Tandem with a k-score plug-in, an open-source search engine developed by the Global Proteome Machine (http://www.thegpm.org/). The search parameters were as follows: (1) precursor mass tolerance window of 100 ppm and fragment mass
| No. | Protein | Charge state | Peptide sequence with NxS/T motif |
|-----|---------|--------------|----------------------------------|
| 1   | 155 kDa platelet multimerin (MMRN1) | 2⁺ | LQN[115]LTLPTN[115]ASIK |
| 2   | 70 kDa lysosomal alpha-glucosidase (GAA) | 2⁺ | GVFITN[115]ETGQPLGK |
| 3   | Afamin (AFAM) | 2⁺ | DIENFN[115]STQK |
| 4   | Aminopeptidase N (AMPN) | 3⁺ | KLN[115]YTLSQGHR |
| 5   | Attractin (ATRN) | 2⁺ | IDSTG[115]VTNLK |
| 6   | Apolipoprotein D (APO D) | 2⁺ / 3⁺ | ADGTVNQIEGATPVN[115]LTEPAK |
| 7   | Apolipoprotein F (APO F) | 2⁺ | Q[111]GGVN[115]ATQVLQLHNL |
| 8   | Apolipoprotein J (APO J) | 2⁺ | LAN[115]LTQGEDQYLYR |
| 9   | Alpha-1-antichymotrypsin (AACT) | 3⁺ / 4⁺ | GLKFN[115]LLETSEAEIHQSFGHLRL |
| 10  | Alpha-2-HS-glycoprotein (FETUA) | 2⁺ / 3⁺ | AALAFAFNAQNN[115]GSNFQLEEISR |
| 11  | Alpha-1-acid glycoprotein 1 (ORM1) | 3⁺ | QDQC[160]IYN[115]TTYLNVQR |
| 12  | Alpha-1-acid glycoprotein 2 (ORM2) | 2⁺ / 3⁺ | QNQC[160]FYNN[115]SSYLNVRQ |
| 13  | Alpha-1B-glycoprotein (A1BG) | 3⁺ / 4⁺ | EGDHFELEVPQAEQEDVEATPHQPN[115]YSC[160]SYR |
| 14  | Antithrombin-III (SERPINC1) | 2⁺ / 3⁺ | LGAC[160]N[115]DTILQQLM[147]EYKFTDTESEK |
| 15  | Basigin (BSG) | 3⁺ | ITTDSEDKAL[147]N[115]GSERS |
| 16  | Beta-galactosidase (GB1) | 2⁺ | NNVITLN[115]ITGK |
| 17  | Biotinidase (BTJ) | 2⁺ / 3⁺ | NPVGLGAEN[115]ATGETDPHSHK |
| 18  | Butyrophilin, subfamily 2, member A1 (BTN2A1) | 2⁺ | GSVALVIHN[115]ITAQEN[115]GTYR |
| 19  | Cathepsin D heavy chain (CTSD) | 2⁺ | GSLSYLN[115]VTR |
| 20  | Cathepsin L (CTSL) | 3⁺ | YSVAN[115]DTGFVDPQKEK |
| 21  | Carboxypeptidase B2 (CBPB2) | 2⁺ / 3⁺ | QTVHFFYN[115]ASDVANKV |
| 22  | Carboxypeptidase M (CBPM) | 2⁺ | NFPDAFEYNN[115]VSR |
| 23  | Calcium binding protein 39 (CAB39) | 2⁺ | HN[115]FTIM[147]TK |
| 24  | CD276 antigen (CD276) | 2⁺ | VVLGANE[115]GYSC[160]LVR |
| 25  | CD163 antigen (CD163) | 2⁺ | APGWAN[115]SSAGSGR |
| 26  | CD44 protein (CD44) | 2⁺ | AFN[115]STLPTM[147]AQM[147]EK |
| 27  | CD7 antigen (CD7) | 3⁺ | GRIDFSGQDNN[115]LTTIM[147]HR |
| No. | Protein                               | Charge state | Peptide sequence with NxS/T motif |
|-----|---------------------------------------|--------------|-----------------------------------|
| 29  | Cell adhesion molecule 1 (CADM1)      | 2+           | VSLTN[115]VSISDEGR                |
|     |                                       | 2+           | FQLLN[115]FSSSELK                |
| 30  | Ceruloplasmin (CP)                    | 2+/3+        | EHEGAIVPDN[115]TTDFQR             |
|     |                                       | 3+/4+        | ELHHLQEQN[115]VSNFLDKGEMYK       |
|     |                                       | 2+           | EN[115]LTAPGSDAVFEEQGTR          |
| 31  | Complement component C4B (C4B)        | 2+           | GLN[115]VTLSSGR                  |
| 32  | Complement factor H (CFH)             | 3+           | IPC[160]SQPPQIEHTIN[115]SSR       |
| 33  | Complement factor H-related 1 (CFHR1) | 2+           | LQNNENN[115]ISC[160]VER          |
| 34  | Complement factor I (CFI)             | 2+           | FLNN[115]GTC[160]TAEGK           |
| 35  | Cubilin (CUBN)                        | 2+           | LC[160]SSV[115]VSNEIK             |
|     |                                       | 2+           | AGFN[115]ASFHK                   |
| 36  | Corticosteroid-binding globulin (SERPINA6) | 2+           | AQLIQQLQFN[115]LTTER            |
|     |                                       | 3+           | AVQLNNEEGYDTAGSTGVTIN[115]LTSKPIILR |
| 37  | Colony stimulating factor 1 (macrophage) (CSF1) | 2+          | VKNVFN[115]ETK                  |
| 38  | Delta and notch-like epidermal growth factor-related receptor (DNER) | 2+          | LVSFEVPQN[115]TSVK               |
| 39  | Desmocollin-2 (DSC2)                 | 2+           | LKAIN[115]DTAAR                  |
|     |                                       | 2+           | AN[115]YTILK                     |
| 40  | Desmoglein-1 (DSC1)                   | 2+           | DYNTKN[115]GTIK                  |
| 41  | Desmoglein-2 (DSG2)                   | 2+           | IN[115]ATDADEPNTLNSK             |
|     |                                       | 2+           | YVQN[115]GTYTVK                  |
| 42  | DNA ligase 4 (LIG4)                   | 2+           | APN[115]LTVNKK                   |
| 43  | Dual specificity protein phosphatase CDC14B (CDC14B) | 2+          | NHN[115]VTIIIR                  |
| 44  | Epidermal growth factor (EGF)         | 2+           | GN[115]NHSILLSALK                |
| 45  | Epididymis secretory sperm binding protein Li 44a (SERPINA1) | 2+/3+/4+/3+/4+/3+/4+ | YLGN[115]ATAIFLPDKEGLQHLENELTHDIK |
|     |                                       | 3+/4+        | ADTHDEIEGLNFIN[115]LTEIPEAQIHEFQQEL |
| 46  | Extracellular link domain containing 1 (XLKD1) | 2+/3+       | KANQQLN[115]FTEAK                |
| 47  | Fibrillin 1 (FBN1)                    | 2+           | TAIFAPN[115]ISHVSNK              |
| 48  | Fibrinopeptide A (FGA)                | 2+           | M[147]DGSLNFIN[115]RT            |
| 49  | Fibronectin type III domain-containing protein 5 (FNDC5) | 2+          | FIQEVV[115]TTTR                  |
| 50  | Frizzled protein 4 (FRP4)             | 2+           | ISM[147]C[160]QNLGYN[115]VTK     |
| 51  | Fibronectin 1 (FN1)                   | 3+           | DQC[160]IVDDITYNVN[115]DTFHK     |
| 52  | Folate receptor alpha (FOLR1)         | 2+           | GWN[115]WTSGFKN                  |
| 53  | Galectin-3-binding protein (LGALS3BP) | 2+           | ALGFEN[115]ATQALGR                |
|     |                                       | 2+           | AAIPSAĐTIN[115]SSK               |
|     |                                       | 2+           | GLN[115]LEDPTYKPR                |
|     |                                       | 2+           | TVIRPFPYTN[115]SGVVD             |
| 54  | Glutaminyl-peptide cyclotransferase (QPCT) | 2+/3+       | NYHQPAILN[115]SSALR              |
|     |                                       | 3+/4+        | YFQN[115]SYGGVIIQDDHIPFLR        |
| 55  | Golgi phosphoprotein 2 (GOLPH2)       | 3+           | LQQDLQLFQKN[115]QTNLER           |
|     |                                       | 1+/2+        | AVLVN[115]N[115]TTGER            |
| No. | Protein                                      | Charge state | Peptide sequence with NxS/T motif                                      |
|-----|----------------------------------------------|--------------|-----------------------------------------------------------------------|
| 56  | Glutamyl aminopeptidase (ENPEP)              | 2⁺          | HTAEYAAN[115]ITK                                                      |
| 57  | Haptoglobin beta chain (HP)                  | 2⁺/3⁺       | VVLHPN[115]YSQVDI GLIK                                               |
|     |                                              | 3⁺/4⁺       | MYSHHN[115]ETTGTATLINEQWLLTTAK                                       |
|     |                                              | 2⁺/3⁺/4⁺    | NLFN[115]HSEN[115]ATAKDIAPLTLYVGKK                                    |
|     |                                              | 3⁺          | Q[111]LVEIEKVLHPN[115]YSQVDI GLIK                                     |
| 58  | HEG homolog 1 (HEG1)                         | 2⁺          | SYSESSTSSSES[LN[115]]SSAPR                                         |
| 59  | Hemopexin (HPX)                              | 3⁺/4⁺/2⁺    | GHGHRN[115]GTGHGN[115]STHHGPEYM[147]R                                |
|     |                                              | 2⁺          | ALPQPQN[115]VTSLLGC[160]TH                                           |
| 60  | Hepatitis B virus receptor binding protein (Q6YPX1) | 2⁺          | EEQYN[115]STYR                                                       |
| 61  | Hepatic asialoglycoprotein receptor 1 transcript variant b (ASGR1) | 2⁺          | ETFSN[115]F T A S T E A Q V K                                        |
| 62  | Heparan sulfate proteoglycan 2 (HSPG2)       | 2⁺          | ALVN[115]F T R                                                       |
| 63  | Ig alpha-1 chain C region (IGHA11)           | 3⁺          | LAGKPThyN[115]VSVVM[147]AEVDGT[160]Y                                  |
| 64  | Ig gamma-1 chain C region (IGHG1)            | 2⁺          | EEQYN[115]STYR                                                       |
|     |                                              | 2⁺/3⁺       | TKPREEQYN[115]STYR                                                  |
| 65  | Ig gamma-2 chain C region (IGHG2)            | 2⁺          | EEQFN[115]STFR                                                       |
|     |                                              | 2⁺/3⁺       | TKPREEQFN[115]STFR                                                  |
| 66  | Ig gamma-4 chain C region (IGHG4)            | 2⁺          | EEQFN[115]STFR                                                       |
|     |                                              | 2⁺/3⁺       | TKPREEQFN[115]STFR                                                  |
| 67  | Ig mu chain C region (IGHM)                  | 2⁺          | YKN[115]NSDISSTR                                                    |
| 68  | Inducible T-cell co-stimulator ligand (ICOSLG) | 2⁺          | TVVTYHIPQN[115]SLENVDSR                                             |
| 69  | Insulin-like growth factor-binding protein 3 (IGFBP3) | 2⁺          | GLC[160]VN[115]ASAVSR                                              |
| 70  | Intercellular adhesion molecule 1 (ICAM1)    | 2⁺          | LNPTVTVGN[115]DSFSAK                                               |
|     |                                              | 2⁺          | AN[115]TVVLNR                                                      |
| 71  | Intercellular adhesion molecule 2 (ICAM2)    | 2⁺          | GN[115]ETLHYETF GK                                                 |
| 72  | Kallikrein-1 (KLK1)                          | 4⁺          | HNLFDDDEN[115]TAQFVHVSESEPFGN[115]M[147]SLEN[115]HTR               |
| 73  | KALLIKREIN-2 (KLK2)                          | 2⁺          | N[115]KS V I L GR                                                   |
| 74  | Kininogen 1 (KNG1)                           | 2⁺          | LNAENN[115]ATFYFK                                                  |
|     |                                              | 2⁺          | ITYSIVQT[115]C[160]SK                                              |
|     |                                              | 3⁺/4⁺       | HGIQFYFN[115]NTQHSSLTLNEVKR                                       |
|     |                                              | 2⁺          | YNSQN[115]QSNQFVLYR                                              |
| 75  | Leucine-rich alpha-2-glycoprotein (LRG1)     | 2⁺          | MFSQN[115]D TR                                                     |
|     |                                              | 2⁺/3⁺       | KLPPGGLAN[115]FTLLR                                               |
| 76  | Leukocyte-associated immunoglobulin-like receptor 1 (LAIR1) | 2⁺/3⁺       | STYN[115]D TEDV SQ ASPSESEAR                                        |
| 77  | Lumican (LUM)                                | 3⁺          | KLHINHNN[115]LTESVG PLPK                                            |
| 78  | Lysosomal acid phosphatase (ACP2)            | 2⁺          | YEQLQN[115]ETR                                                      |
| 79  | Lysosomal alpha-glucosidase (GAA)            | 2⁺          | GVFTTN[115]ETGQPLIGK                                              |
|     |                                              | 2⁺          | LEN[115]LSSSEM[147]GYTATL TR                                        |
| 80  | Lysosome-associated membrane glycoprotein 1 (LAMP1) | 2⁺          | GHTLTLN[115]F T R                                                |
| No. | Protein                                               | Charge state | Peptide sequence with NxS/T motif                                      |
|-----|-------------------------------------------------------|--------------|--------------------------------------------------------------------------------|
| 81  | Lysosomal-associated membrane protein 2, (LAMP2)      | 2⁺           | VASVININPN[115]TTHSTGSC[160]R                                           |
| 82  | Lymphatic vessel endothelial hyaluronic acid receptor 1 (XLKD1) | 2⁺           | ANQQLN[115]FTEAK                                                      |
| 83  | Lysyl oxidase (LOX)                                    | 3⁺           | AE[N[115]QTAPGEVPALSNLPPSRRDPGAAVPGAAN[115]ASAQQPR                     |
| 84  | Major prion protein (PRNP)                             | 2⁺           | Q[111]HTVTTTTTGEN[115]FTETDVK                                         |
| 85  | Membrane protein FAM174A (FAM174A)                     | 2⁺           | GSEGGN[115]GSNPVAGLETDDHGGK                                           |
| 86  | Maltase-glucoamylase (MGAM)                            | 2⁺           | ILGM[147]EEPSN[115]VTVK                                               |
|     |                                                       | 2⁺           | VILILPAGSN[115]ETQPYPAFT                                               |
| 87  | Microfibril-associated glycoprotein 4 (MFAP4)          | 2⁺           | VDLEDFN[115]NTAYAK                                                   |
| 88  | Monocyte differentiation antigen CD14                 | 2⁺           | LRN[115]VSWATGR                                                  |
| 89  | Mucin-6 (MUC6)                                         | 2⁺           | GC[160]M[147]AN[115]VTVTR                                             |
| 90  | N-acetylgalcosamine-6-sulfatase (GNS)                  | 2⁺           | YYN[115]YTLSIN[115]GK                                                 |
|     |                                                       | 2⁺           | TPMTN[115]SSIQFGLNAFR                                                |
| 91  | N-acylsphingosine amidohydrolase (ASAH1)               | 2⁺           | TVLEN[115]STSYEEAK                                                   |
| 92  | Neuronal pentraxin receptor (NPTXR)                    | 2⁺           | ALPGGADN[115]ASVAGAAASPGQ                                            |
| 93  | Neuropilin and tollloid-like protein 1 (NETO1)         | 2⁺           | HESEYN[115]TTR                                                    |
| 94  | Peptidase inhibitor 16 (PI16)                          | 2⁺           | SLPNFPN[115]TSATAN[115]ATGGR                                         |
| 95  | Pigment epithelium-derived factor (SERPINF1)           | 3⁺           | VTQ[115]LTLIEE LTSFHIHDIDR                                             |
| 96  | Plasma protease C1 inhibitor (SERPING1)                | 2⁺/3⁺        | GVTVSQIFHSPDLAIRDFTVN[115]ASR                                         |
|     |                                                       | 2⁺           | VLSN[115]NSDANLEIINTVWAK                                              |
|     |                                                       | 3⁺           | VGQLQLSHN[115]LSLIVVPQNLK                                             |
| 97  | Plasma serine protease inhibitor (SERPIN5A)             | 2⁺           | VVGVYPQQN[115]ATALFILPSEGK                                           |
| 98  | Platelet-derived growth factor subunit B (PDGFB)       | 3⁺           | LLHGDPGEEDGAELDLN[115]M[147]TR                                       |
| 99  | Polytrophin (TROPH)                                    | 2⁺           | N[115]N[115]VTEDIK                                                  |
| 100 | Probable G-protein coupled receptor 116 (GPR116)       | 2⁺           | ANEQVQQLN[115]QTYK                                                 |
|     |                                                       | 2⁺           | YEEQQLEIQN[115]SSR                                                   |
| 101 | Probable serine carboxypeptidase (CPVL)                | 2⁺           | Q[111]AIHVGN[115]QTFNDGTIVEK                                         |
| 102 | Prosaposin (PSAP)                                      | 2⁺/3⁺        | NLEKN[115]STKQEILAALEK                                              |
| 103 | Prostaglandin D2 synthase 21 kDa (PTGDS)               | 2⁺/3⁺        | SVVAPATDGGGLN[115]LTSTFLR                                             |
|     |                                                       | 2⁺           | WFSAGLASN[115]SSWLR                                                  |
| 104 | Prostatic acid phosphatase (ACP)                      | 3⁺           | FLN[115]ESYKHEQVYIR                                                 |
| 105 | Proteinase-activated receptor 1 (F2R)                  | 2⁺           | ATN[115]ATLDPDR                                                  |
| 106 | Protein shisa-7 (SHISA7)                               | 2⁺           | LTGALTGGGAAASPQAN[115]GTR                                             |
| 107 | RING finger protein 10 (RNF10)                         | 2⁺           | N[115]ESFN[115]N[115]QSR                                          |
Table 3: Continued.

| No. | Protein                                                         | Charge state | Peptide sequence with NxS/T motif                                                                 |
|-----|-----------------------------------------------------------------|--------------|--------------------------------------------------------------------------------------------------|
| 108 | Secretory component (Polymeric IG Receptor) (PIGR)             | 3+           | AN[115]LTNFPE[115]GTFVNVIAQLSQQDSGRRQ[111]GLYPVVLIDSSGYVNP[115]YTGRVPGN[115]VTAVGLGETLKYKCGLG[115]SR |
| 109 | Slit homolog 1 (SLIT1)                                         | 2+           | LELN[115]GN[115]N[115]ITR                                                                         |
| 110 | Sushi domain-containing protein 2 (SUSD2)                      | 2+           | SELVN[115]ETR                                                                                        |
| 111 | Sex hormone binding globulin (SHBG)                            | 2+           | LDVDQALN[115]RT                                                                                      |
| 112 | Transferrin (TF)                                               | 2+/3+        | Q[111]QQLFLGSN[115]VTDC[160]SGNFC[160]LFR                                                        |
| 113 | Thrombin heavy chain (F2)                                     | 4+           | YPHKPEIN[115]STTHPGADLQENFC[160]R                                                                  |
| 114 | Tripeptidyl-peptidase I variant (TPP1)                         | 3+           | FLSSPHLPPSSYFN[115]ASGR                                                                            |
| 115 | Tyrosine-protein kinase receptor UFO (AXL)                     | 2+           | SLHVPGLN[115]KT                                                                                  |
| 116 | TIMP metallopeptidase inhibitor 1 (TIMP1)                      | 2+           | FVGTPEVN[115]QTTLYQR                                                                              |
| 117 | Thyroxine-binding globulin (SERPINA7)                          | 2+           | TLYETEVFSTDFSN[115]ISAK                                                                         |
| 118 | Trypstatin (AMBP)                                              | 2+/3+        | SKWN[115]ITM[147]ESYVVHTNYDEYAIFLTK                                                              |
| 119 | Transmembrane protein 108 (TMEM108)                            | 4+           | KGAGN[115]SSRPVPPAPGGHRS                                                                           |
| 120 | Uromodulin (UMOD)                                              | 2+/3+        | Q[111]DFN[115]ITDISLLEHR                                                                       |
| 121 | Vasorin (VASN)                                                 | 2+           | LHEITN[115]ETFR                                                                                  |
| 122 | Zinc-alpha-2-glycoprotein (AZGP1)                              | 2+/3+        | DIVEYN[115]DSN[115]GSHVLQGR                                                                     |
|       |                                                                  | 3+/4+        | AREDFIM[147]ETLKDIVEYN[115]DSN[115]GSHVLQGR                                                       |
|       |                                                                  | 2+           | FGCEIENN[115]RS                                                                               |

tolerance of 0.8 Da; (2) allowing two missed cleavages; (3) variable modification: oxidation of methionine (+15.9949 Da), carbamidomethyl cysteine (57.0214 Da), and +0.9840 Da, reflecting the conversion of asparagine in the NxS/T motif to aspartate due to the release of the N-linked glycopeptides from their oligosaccharides. All proteins with a ProteinProphet probability of greater than 0.9 were considered as positive identifications [19]. Only proteins containing peptides with the NxS/T sequence motif were included for statistical analysis.

Baseline characteristics of the control and CKD subjects were compared using Fisher’s exact test for categorical variables and Student’s t-test for continuous variables. Data is presented as means (±SD). Spectral counts for individual proteins were normalized to Saccharomyces cerevisiae invertase and to urine creatinine content. Spectral counts were compared across the two subject groups using the nonparametric Mann-Whitney test, and P values were adjusted for multiple comparisons using the False Discovery Rate (FDR) with reported q-values. All statistical analyses were performed with the use of SAS software, version 9.2.

2.5. Gene Ontology Analysis. Significant proteins of interest were analyzed using the Gene Ontology Database (Gene Ontology Consortium, http://www.geneontology.org, Princeton University, New Jersey, US; [20]). For a given Gene Ontology (GO) category, the relative enrichment of genes encoding the proteins detected in CKD relative to all reference genes in that category were calculated as previously described using GO Tools made available by the Bioinformatics Group at the Lewis-Sigler Institute (Princeton University, New Jersey, US; [21]). A cutoff value of P < 0.01 was used to report a functional category as significantly overrepresented. To address the multiple comparisons problem that arises when many processes are evaluated simultaneously, the analysis included calculation of the FDR [21]. To improve statistical confidence in our results, all enriched functional categories were required to be significant using both methods (P < 0.01 and FDR < 0.05).

3. Results

3.1. Study Subject Characteristics. Urine was isolated from six subjects with CKD and six age-matched healthy controls.
Baseline subject characteristics are provided in Table 1. Two important issues were considered with patient selection. First, the etiology of CKD was chosen to be diverse. This would ensure robustness of the putative markers as a CKD marker rather than a disease-specific marker. Second, we specifically targeted early Stage 3 CKD subjects to identify early disease markers that would potentially indicate pathways dysregulated early in the course of disease. This might offer mechanistic insights into disease pathogenesis and progression and have implications in therapeutic strategies.

The six subjects had biopsy-proven diabetic nephropathy, lupus nephritis (n = 2), postacute tubular necrosis damage, NSAID nephropathy, and membranoproliferative glomerulonephritis, respectively. The mean estimated glomerular filtration rate (eGFR) was 83 mL/min in control subjects and 52 mL/min in CKD subjects.

3.2. Glycoprotein Spectral Count Normalization. Glycoproteins were extracted and enriched from the twelve urinary samples. To account for variations in the glycoprotein extraction efficiency, 5 μg of the yeast protein invertase from Saccharomyces cerevisiae was added to each sample prior to extraction. After addition to the database, invertase spectral count served as a surrogate marker for extraction efficiency in each individual sample. Invertase spectral counts ranged from 31 to 122 in the twelve samples with an average spectral count of 86 (±31). Each sample was normalized independently to the invertase spectral counts.

To account for intersubject urine concentration variability, spectral counts were then normalized to urine creatinine content. This provides standardization for urinary creatinine excretion and concentration differences which can vary with volume status, stress, diet, activity level, age, gender, and overall health status [22]. Indeed, this normalization is commonly followed in clinical practice where degree of urinary protein is normalized to creatinine to obtain protein excretion rates [23]. Final spectral counts were expressed per mmol creatinine.

3.3. Urine Glycoproteome Is Altered in CKD. Urinary glycoproteins were isolated from six subjects with CKD and six healthy controls using a hydrazide technique as described in Section 2. A total of 122 glycoproteins were identified, of which 35 proteins were unique to healthy control patients, 8 were unique to CKD subjects, and 79 were common proteins.

3.4. Gene Ontology Analysis Reveals Enrichment for Distinct Biological Functions of Differentially Expressed Urinary Glycoproteins. The 23 proteins with differential expression in CKD were subjected to a GO Database search and further analyzed with GO Tools [20, 21]. GO Term Finder (http://go.princeton.edu/cgi-bin/GOTermFinder) allowed for clustered identification of proteins annotated to specific GO biological process, location, and function classifications. A subsequent GO Term Mapper (http://go.princeton.edu/cgi-bin/GOTermMapper) analysis of significantly altered proteins was performed to bin the proteins to GO parent terms or GO Slim terms (http://www.geneontology.org/GO.slims.shtml).

GO analysis (Figure 3) for biological processes demonstrated that 16 of the 23 proteins were linked to immune/stress response and biological process regulation (P < 1 × 10−4). 9 of the 23 were acute-phase and inflammatory response proteins (P < 1 × 10−3). Six proteins were regulators of hemostasis, platelet degranulation and coagulation (P < 1 × 10−4), and 10 were involved in...
Figure 3: Global view of biological processes of differentially expressed urinary glycoproteins in CKD. Urinary glycoproteins that were differentially detected in CKD subjects were associated with biological functions using GO process annotations. This approach demonstrated significant overrepresentation of proteins involved in several categories, including regulation of response to stress, platelet activation/hemostasis/coagulation, acute-phase response, regulation of biological processes, localization, secretion, transport, and cell death. A bipartite network (generated using Cytoscape [24]) showing the relationship between GO process annotations (yellow hexagon nodes) and differentially regulated proteins in CKD subjects (white/red/green circular nodes). The size of the GO nodes is proportional to the number of edges (lines) that connect them to proteins. The 10 proteins that are altered with q-value of <0.05 are depicted in red (up-regulation) and green (down-regulation) in CKD subjects. GAA, 70 kDa lysosomal alpha-glucosidase; APOD, Apolipoprotein D; FETUA: Alpha-2-HS-glycoprotein chain B; ORM1, Alpha-1-acid glycoprotein 1; SERPINC1, Antithrombin-III; GLB1, Beta-galactosidase; CP, Ceruloplasmin; CUBN, Cubilin; EGF, Epidermal growth factor; E9KL23, Epididymis secretory sperm binding protein Li 44a; LGALS3BP, Galectin-3-binding protein; GOLPH2, Golgi phosphoprotein 2; HP, Haptoglobin beta chain; IGHG1, Ig gamma-1 chain C region; IGHG2, Ig gamma-2 chain C region; KNG1, Kininogen 1; LRG, Leucine-rich alpha-2-glycoprotein; SERPINC1, Plasma protease C1 inhibitor; PTGDS, Prostaglandin D2 synthase 21 kDa; TF, Transferrin; AMBP, Trypsatin; UMOD, Uromodulin; AZGP1, Zinc-alpha-2-glycoprotein.
Table 4: Differentially regulated proteins identified in CKD subjects.

| Protein Code | Name of the protein identified | P value  | q-value  | Direction of change in CKD subjects |
|--------------|---------------------------------|----------|----------|-------------------------------------|
| APOD         | Apolipoprotein D                | 0.0022   | 0.0224   | Up                                  |
| FETUA        | Alpha-2-HS-glycoprotein chain B | 0.0022   | 0.0224   | Up                                  |
| ORM1         | Alpha-1-acid glycoprotein 1     | 0.0022   | 0.0224   | Up                                  |
| E9KL23       | Epididymis secretory sperm binding protein Li 44a | 0.0022 | 0.0224 | Up |
| LGALS3BP     | Galectin-3-binding protein      | 0.0022   | 0.022    | Up                                  |
| GOLPH2       | Golgi phosphoprotein 2          | 0.0022   | 0.0224   | Down                                |
| HP           | Haptoglobin                     | 0.0022   | 0.0224   | Up                                  |
| KNG1         | Kininogen 1                     | 0.0022   | 0.0224   | Up                                  |
| LRG          | Leucine-rich alpha-2-glycoprotein | 0.0022 | 0.0224 | Up |
| SERPING1     | Plasma protease C1 inhibitor    | 0.0022   | 0.0224   | Up                                  |
| PTGDS        | Prostaglandin D2 synthase 21kDa | 0.0022  | 0.0224   | Up                                  |
| AZGP1        | Zinc-alpha-2-glycoprotein       | 0.0022   | 0.0224   | Up                                  |
| GAA          | 70 kDa lysosomal alpha-glucosidase | 0.0152 | 0.13     | Down                                |
| SERPINC1     | Antithrombin-III                | 0.0152   | 0.103    | Up                                  |
| GLB1         | Beta-galactosidase              | 0.0152   | 0.103    | Down                                |
| CUBN         | Cubilin                         | 0.0152   | 0.103    | Down                                |
| EGF          | Epidermal growth factor         | 0.0152   | 0.13     | Down                                |
| UMOD         | Uromodulin                      | 0.0152   | 0.103    | Up                                  |
| TF           | Transferrin                     | 0.0216   | 0.1387   | Up                                  |
| AMBP         | Trypsatin                       | 0.0411   | 0.18     | Up                                  |
| CP           | Ceruloplasmin                   | 0.0433   | 0.18     | Up                                  |
| IGHG1        | Ig gamma-1 chain C region       | 0.0433   | 0.18     | Up                                  |
| IGHG2        | Ig gamma-2 chain C region       | 0.0433   | 0.18     | Up                                  |

localization, transport, and secretion ($P < 1 \times 10^{-4}$). Other processes involved include metal ion homeostasis (4 proteins) and cell death (3 proteins).

Table 5 displays function and location for the 23 proteins which were differentially expressed in CKD. 18 out of the 23 proteins localized to the extracellular region consistent with possible extracellular matrix remodeling that typifies renal disease. The analysis also revealed 2 major clusters of molecular function: 20 out of the 23 proteins were involved in binding and protein-protein interactions ($P = 5 \times 10^{-4}$). 5 proteins were endopeptidase inhibitors ($P < 1 \times 10^{-6}$). Collectively, these observations implicate the inflammatory/acute-phase response and extracellular matrix remodeling in CKD. They also strongly support the proposal that glycoproteomic analysis of urine might reveal mechanisms underpinning CKD.

4. Discussion

CKD is a growing public health problem with dramatic increases in morbidity and mortality following progression to ESRD. Given this, there is a tremendous need for the development of biomarkers to predict CKD progression and allow for early therapeutic intervention. Urine proteomic strategies are now at the forefront of this search due to the sensitivity of MS/MS analysis and the ability to develop noninvasive biomarkers from a readily available biofluid. Significant progress has been made, particularly in diabetes, where urine proteomic analysis can predict nephropathy [6, 25, 26]. Despite these developments, the majority of proteomic studies have relied on two-dimensional (2D) differential in-gel electrophoresis for protein separation. Resulting samples, particularly in CKD subjects, contain large amounts of highly abundant plasma proteins due to nonspecific leakage through the glomerular filtration barrier. Targeted analyses of low-abundance proteins will likely lead to more disease-specific and clinically relevant protein biomarkers.

We therefore focused our attention on the urinary N-linked glycoproteome. Glycoproteins are an important protein subfraction accounting for up to 50% of the human proteome at any given time [27]. Due to their critical role in cell-cell interactions and signaling cascades, glycoproteins are promising markers for identifying kidney disease activity and progression. In this study we present an initial examination of the urinary N-linked glycoproteome in CKD subjects compared to healthy control subjects. We successfully isolated N-linked glycoproteins from twelve urine samples utilizing a hydrazide capture technique. 122 unique glycosylated proteins were detected amongst the twelve subjects (Table 3). This number is similar to other recent glycoproteome analyses. Ahn et al. recently reported isolating 164–174 unique proteins from human diabetic plasma using a multi-lectin column enrichment technique [16]. Yang et al. isolated 265 urinary glycoproteins from bladder cancer subjects and healthy controls also utilizing a multi-lectin column for enrichment, but larger sample sizes were used than in our current study [15]. These
Table 5: Component location and function associated with significantly up- or downregulated proteins identified in CKD patients.

| Gene ontology term | Location | Cluster frequency | P-value | FDR | Proteins annotated to the GO Term | Gene ontology term | Cluster frequency | P-value | FDR | Proteins annotated to the GO Term |
|--------------------|----------|-------------------|---------|-----|-----------------------------------|--------------------|-------------------|---------|-----|-----------------------------------|
| Extracellular region | 18 of 23 proteins, 78.3% | $7.36 \times 10^{-15}$ | 0.00% | | EGF, IGHG1, LRG, AMBP, AZGP1, KNG1, SERPING1, TF, LGALS3BP, FETUA, CP, HP, SERPINC1, IGHG2, ORM1, APOD, PTGDS, UMOD | Binding | 20 of 23 proteins, 87.0% | 0.00051 | 0.89% | EGF, IGHG1, AMBP, CUBN, AZGP1, KNG1, SERPING1, TF, LGALS3BP, GAA, FETUA, CP, SERPINC1, HP, IGHG2, GLB1, ORM1, APOD, PTGDS, UMOD |
| Extracellular space | 11 of 23 proteins, 47.8% | $5.47 \times 10^{-11}$ | 0.00% | | EGF, LGALS3BP, CP, FETUA, SERPINC1, APOD, ORM1, PTGDS, UMOD, KNG1, SERPING1 | Protein binding | 15 of 23 proteins, 65.2% | 0.00066 | 0.80% | |
| Extracellular region part | 11 of 23 proteins, 47.8% | $1.13 \times 10^{-9}$ | 0.00% | | EGF, LGALS3BP, CP, FETUA, SERPINC1, APOD, ORM1, PTGDS, UMOD, KNG1, SERPING1 | Enzyme regulator activity | 6 of 23 proteins, 26.1% | 0.00024 | 0.00% | EGF, SERPINC1, FETUA, KNG1, AMBP, SERPING1 |
| Cytoplasmic vesicle | 6 of 23 proteins, 26.1% | $6.81 \times 10^{-5}$ | 0.20% | | TF, EGF, CUBN, UMOD, KNG1, SERPING1 | Endopeptidase inhibitor activity | 5 of 23 proteins, 21.7% | 3.78E-07 | 0.00% | SERPINC1, FETUA, KNG1, AMBP, SERPING1 |
| Vesicle | 6 of 23 proteins, 26.1% | $8.57 \times 10^{-5}$ | 0.18% | | TF, EGF, CUBN, UMOD, KNG1, SERPING1 | Endopeptidase regulator activity | 5 of 23 proteins, 21.7% | 4.29E-07 | 0.00% | SERPINC1, FETUA, KNG1, AMBP, SERPING1 |
| Cell fraction | 5 of 23 proteins, 21.7% | 0.00405 | 2.29% | | EGF, IGHG2, IGHG1, CUBN, AMBP | Peptidase inhibitor activity | 5 of 23 proteins, 21.7% | 4.84E-07 | 0.00% | SERPINC1, FETUA, KNG1, AMBP, SERPING1 |
| Cytoplasmic membrane-bounded vesicle | 5 of 23 proteins, 21.7% | 0.00055 | 0.15% | | TF, EGF, CUBN, KNG1, SERPING1 | Peptidase regulator activity | 5 of 23 proteins, 21.7% | 1.10E-06 | 0.00% | SERPINC1, FETUA, KNG1, AMBP, SERPING1 |
| Membrane-bounded vesicle | 5 of 23 proteins, 21.7% | 0.00062 | 0.29% | | TF, EGF, CUBN, KNG1, SERPING1 | Enzyme inhibitor activity | 5 of 23 proteins, 21.7% | 8.50E-06 | 0.00% | SERPINC1, FETUA, KNG1, AMBP, SERPING1 |
| Gene ontology term                      | Cluster frequency | Location | Function                              | Gene ontology term | Cluster frequency | Function                              | Gene ontology term | Cluster frequency | Function                              |
|----------------------------------------|------------------|----------|---------------------------------------|--------------------|------------------|---------------------------------------|--------------------|------------------|---------------------------------------|
| Cytoplasmic vesicle part               | 4 of 23 proteins, 17.4% |          | Transporter activity                  | EGF, CUBN, KNG1, SERPING1 | 0.00024 | 0.17% | CUBN, APOD, AZGP1, PTGDS, AMBP          |
| Stored secretory granule               | 4 of 23 proteins, 17.4% |          | Carbohydrate binding                  | TF, EGF, KNG1, SERPING1 | 5.33E−05 | 0.22% | SERPINC1, GAA, KNG1                    |
| Membrane fraction                      | 4 of 23 proteins, 17.4% |          | Lipid binding                         | IGHG2, IGHG1, CUBN, AMBP | 0.00796 | 2.69% | APOD, AZGP1, PTGDS                     |
| Insoluble fraction                     | 4 of 23 proteins, 17.4% |          | Serine-type endopeptidase inhibitor activity | IGHG2, IGHG1, CUBN, AMBP | 0.00897 | 2.81% | SERPINC1, AMBP, SERPING1, E9KL23      |
| Apical plasma membrane                 | 3 of 23 proteins, 13.0% |          | Hemoglobin binding                    | TF, CUBN, UMOD     | 0.00069 | 0.27% | HP, CUBN                               |
| Perinuclear region of cytoplasm        | 3 of 23 proteins, 13.0% |          | Fatty acid binding                    | TF, GLB1, PTGDS    | 0.00565 | 2.58% | AZGP1, PTGDS                           |
| Apical part of cell                    | 3 of 23 proteins, 13.0% |          | Cysteine-type endopeptidase inhibitor activity | TF, CUBN, UMOD     | 0.00146 | 0.94% | FETUA, KNG1                            |
| Lytic vacuole                          | 3 of 23 proteins, 13.0% |          | Monocarboxylic acid binding           | CUBN, GLB1, GAA    | 0.0026  | 1.89% | AZGP1, PTGDS                           |
| Lysosome                               | 3 of 23 proteins, 13.0% |          | Hydrolyase activity, hydrolyzing O-glycosyl compounds | CUBN, GLB1, GAA    | 0.0026  | 1.79% | GLB1, GAA                              |
| Platelet alpha granule lumen           | 3 of 23 proteins, 13.0% |          | Protein kinase regulator activity      | EGF, KNG1, SERPING1 | 2.55E−05 | 0.00% | EGF, FETUA                             |
Table 5: Continued.

| Gene ontology term          | Cluster frequency | Location                  | P-value | FDR  | Proteins annotated to the GO Term | Gene ontology term          | Cluster frequency | Function                                         | P-value | FDR  | Proteins annotated to the GO term |
|-----------------------------|-------------------|---------------------------|---------|------|-----------------------------------|-----------------------------|-------------------|-------------------------------------------------|---------|------|-----------------------------------|
| Secretory granule lumen    | 3 of 23 proteins, | 1.37E−05                  | 0.00%   |      | EGF, KNG1, SERPING1               | Kinase regulator activity   | 2 of 23 proteins | Kinase regulator activity                        | 0.0054  | 2.71%| EGF, FETUA                        |
|                             | 13.0%             |                           |         |      |                                   | activity                    | 8.7%              |                                                  |         |      |                                   |
| Cytoplasmic membrane-bound | 3 of 23 proteins, | 1.55E−05                  | 0.00%   |      | EGF, KNG1, SERPING1               | Hydrolase activity          | 2 of 23 proteins | Hydrolase activity, acting on glycosyl bonds     | 0.0056  | 2.78%| GLB1, GAA                         |
| vesicle lumen               | 13.0%             |                           |         |      |                                   | activity                    | 8.7%              |                                                  |         |      |                                   |
| Vesicle lumen               | 3 of 23 proteins, | 1.74E−05                  | 0.00%   |      | EGF, KNG1, SERPING1               | Tetrapyrrole binding        | 2 of 23 proteins | Tetrapyrrole binding                            | 0.0081  | 2.67%| CUBN, AMBP                        |
|                             | 13.0%             |                           |         |      |                                   |                             | 8.7%              |                                                  |         |      |                                   |
| Platelet alpha granule      | 3 of 23 proteins, | 2.55E−05                  | 0.00%   |      | EGF, KNG1, SERPING1               | Heparin binding             | 2 of 23 proteins | Heparin binding                                 | 0.0060  | 2.63%| SERPIN1, KNG1                     |
|                             | 13.0%             |                           |         |      |                                   |                             | 8.7%              |                                                  |         |      |                                   |
| Vacuole                     | 3 of 23 proteins, | 0.00423                   | 2.27%   |      | CUBN, GLB1, GAA                   |                             |                   |                                                 |         |      |                                   |
|                             | 13.0%             |                           |         |      |                                   |                             |                   |                                                 |         |      |                                   |
| Extrinsic to membrane       | 2 of 23 proteins, | 0.00532                   | 2.61%   |      | CUBN, UMOD                        |                             |                   |                                                 |         |      |                                   |
| Lysosomal membrane vesicle  | 8.7%              |                           |         |      |                                   |                             |                   |                                                 |         |      |                                   |
| Endocytic vesicle           | 2 of 23 proteins, | 0.00699                   | 2.64%   |      | CUBN, GAA                        |                             |                   |                                                 |         |      |                                   |
|                             | 8.7%              |                           |         |      |                                   |                             |                   |                                                 |         |      |                                   |
| Coated pit                  | 2 of 23 proteins, | 0.00332                   | 2.10%   |      | TF, CUBN                         |                             |                   |                                                 |         |      |                                   |
|                             | 8.7%              |                           |         |      |                                   |                             |                   |                                                 |         |      |                                   |
| Cell projection membrane    | 2 of 23 proteins, | 0.00113                   | 0.88%   |      | TF, CUBN                         |                             |                   |                                                 |         |      |                                   |
|                             | 8.7%              |                           |         |      |                                   |                             |                   |                                                 |         |      |                                   |
|                             |                   |                           |         |      |                                   |                             |                   |                                                 |         |      |                                   |

International Journal of Proteomics
results support a successful hydrazide based technique for glycoprotein isolation in human urine. Further studies are required to identify optimal extraction strategies.

We detected 8 glycoproteins unique to CKD subjects and 35 unique to healthy controls (Table 2). Additionally, of the 122 total proteins identified, 23 glycoproteins were differentially expressed in CKD subjects versus healthy controls. 18 were upregulated in CKD while 5 were downregulated (Table 4). Many of the differentially expressed proteins have been previously linked to kidney disease supporting a potential role as a CKD biomarker. Two of the most significantly upregulated proteins in our CKD samples were AZGP1 and LRG, both of which are established inflammatory mediators. Alteration of AZGP1 and LRG expression is predictive of acute kidney injury in postsurgical patients [28]. AZGP1 has also been shown to be increased in diabetes and diabetic nephropathy [13, 29]. PTGDS, a known extracellular transporter for lipophilic molecules, is formed de novo in renal tubules [30]. PTGDS is upregulated in early diabetes [31] and is a marker of hypertension and latent renal injury [32]. SERPING1, an extracellular matrix regulator, is increased in acute renal allograft rejection perhaps suggesting an important role for collagen remodeling [33]. KNG1, a bradykinin precursor, has also been shown to be upregulated in acute renal allograft rejection [34], and gene variation induces altered aldosterone sensitivity in hypertensive subjects [35]. Interestingly, LUM, a proteoglycan, is a protein unique to CKD but without statistically significant upregulation. Altered regulation of LUM has been linked with abnormal collagen fibril morphology as a mediator of fibrotic disease in diabetic nephropathy [36, 37]. CUBN, an apical protein in proximal tubule cells, was unique and downregulated in CKD. Recent investigation supports a role of CUBN in albumin reabsorption with genetic variance at this locus predicting microalbuminuria [38]. The decreased urinary CUBN excretion found in our CKD population may represent a dysfunctional variant or potentially a causative factor responsible for increasing proteinuria.

We used annotations by the GO Consortium and GO Tools to connect the complex array of proteins identified in CKD subjects to biological processes, protein function, and cellular location. Many of the multiprotein pathways differentially expressed in CKD are involved in coagulation, inflammation, and acute-phase response (Table 3, Figure 3). Twenty proteins were linked to protein-protein interactions and binding. Remarkably, there were altered levels of proteins that were involved in acute-phase response and immune/stress response proteins (18 out of 23), implicating a possible mechanistic role for these pathways in CKD. Our detection of the several extracellular proteins and matrix remodeling proteases likely reflects matrix remodeling that occurs in CKD. These findings are consistent with previous literature, as CKD is known to have increased propensity for atherosclerosis, endothelial dysfunction, increased basal inflammation, and altered stress response [39, 40].

In this study, we have established normalization techniques which will be essential to future urine glycoproteome analyses. To account for variations in the glycoprotein extraction efficiency of individual samples, yeast invertase (yeast glycoprotein with several glycopeptides) was added to each sample prior to extraction. In this way, glycopeptides derived from invertase serve as an internal marker for the extraction efficiency in each sample. Our samples were also normalized for urine creatinine content. This is of particular importance as marked intersubject variability can exist in creatinine excretion in random urine specimens consistent with different concentrations due to hydration status. Indeed, such normalization would be essential to extrapolate net excretion rates of a given protein in 24 hours and is commonly employed in clinical practice to quantify albumin excretion rates [23].

In summary, we have utilized a hydrazide-based approach to enrich the urinary glycoproteome with subsequent identification of the urinary glycoproteins in a human CKD population for the first time. Our results indicate that urine carries a distinct population of glycoproteins that function in protease inhibition, protein binding, and the acute-phase/immune-stress response in subjects with CKD. It will be of interest to study a larger number of subjects to determine whether urinary levels of these proteins might be useful indicators of CKD and to investigate the proposal that these proteins could be markers of disease progression.

**Abbreviations**

| Acronym | Description |
|---------|-------------|
| CKD     | Chronic kidney disease |
| ESRD    | End-stage renal disease |
| FDR     | False discovery rate |
| GO      | Gene ontology |
| LC-ESI-MS/MS | Liquid chromatography-electrospray ionization tandem MS analysis |
| MS      | Mass spectrometry |

**Acknowledgments**

This work is supported in part by Grants from the National Institutes of Health (HL094230, DK079912, HLI02334, GM49500, CA108597, DK082841, and K12HD001438), the Doris Duke Foundation Clinical Scientist Development Award (S. Pennathur), the American Diabetes Association (1-08-CR-48; R. Pop-Busui), the George O’Brien Kidney Center (DK089503), and the Molecular Phenotyping Core of the Michigan Nutrition and Obesity Research Center (DK089503). A. Vivekanandan-Giri and J. L. Slocum contributed equally to this work.

**References**

[1] J. Coresh, E. Selvin, L. A. Stevens et al., “Prevalence of chronic kidney disease in the United States,” *Journal of the American Medical Association*, vol. 298, no. 17, pp. 2038–2047, 2007.

[2] United States Renal Data System, “Atlas of chronic kidney disease and end-stage renal disease in the United States,” Annual Data Report USRDS 2010, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Md, USA, 2010.

[3] P. Ruggenenti, A. Perna, L. Mosconi, R. Pisoni, and G. Remuzzi, “Urinary protein excretion rate is the best independent predictor of ESRF in non-diabetic proteinuric chronic kidney disease,” *Kidney International*, vol. 51, no. 1, pp. 16–22, 1997.
nephropathies,” *Kidney International*, vol. 53, no. 5, pp. 1209–1216, 1998.

[4] D. Fliser, J. Novak, V. Thongboonkerd et al., “Advances in urinary proteome analysis and biomarker discovery,” *Journal of the American Society of Nephrology*, vol. 18, no. 4, pp. 1057–1071, 2007.

[5] J. Klein, P. Kavvadas, N. Prakoura et al., “Renal fibrosis: insight from proteomics in animal models and human disease,” *Proteomics*, vol. 11, no. 4, pp. 805–815, 2011.

[6] K. Rossing, H. Mischak, M. Dakna et al., “Urinary proteomics in diabetes and CKD,” *Journal of the American Society of Nephrology*, vol. 19, no. 7, pp. 1283–1290, 2008.

[7] M. L. Merchant, B. A. Perkins, G. M. Boratyn et al., “Urinary peptidome may predict renal function decline in type 1 diabetes and microalbuminuria,” *Journal of the American Society of Nephrology*, vol. 20, no. 9, pp. 2065–2074, 2009.

[8] H. A. Shui, T. H. Huang, S. M. Ka, P. H. Chen, Y. F. Lin, and A. Chen, “Urinary proteome and potential biomarkers associated with serial pathogenesis steps of focal segmental glomerulosclerosis,” *Nephropathy Dialysis Transplantation*, vol. 23, no. 1, pp. 176–185, 2008.

[9] M. Haubitz, S. Wittke, E. M. Weissinger et al., “Urine protein patterns can serve as diagnostic tools in patients with IgA nephropathy,” *Kidney International*, vol. 67, no. 6, pp. 2313–2320, 2005.

[10] E. Banon-Maneus, F. Diekmann, M. Carrascal et al., “Two-dimensional difference gel electrophoresis urinary peptidome profile in the search of nonimmune chronic allograft dysfunction biomarkers,” *Transplantation*, vol. 89, no. 5, pp. 548–558, 2010.

[11] R. P. Wroniecki, T. N. Orlova, N. Mendelev et al., “Urinary peptidome of steroid-sensitive and steroid-resistant idiopathic nephrotic syndrome of childhood,” *American Journal of Nephrology*, vol. 26, no. 3, pp. 258–267, 2006.

[12] X. Zhang, M. Jin, H. Wu et al., “Biomarkers of lupus nephritis determined by serial urine proteomics,” *Kidney International*, vol. 74, no. 6, pp. 799–807, 2008.

[13] S. Jain, A. Rajput, Y. Kumar, N. Uppuluri, A. S. Arvind, and U. Tatu, “Proteomic analysis of urine protein markers for accurate prediction of diabetic kidney disorder,” *Journal of Association of Physicians of India*, vol. 53, pp. 513–520, 2005.

[14] J. B. Lowe, “Glycosylation, immunity, and autoimmunity,” *Cell*, vol. 104, no. 6, pp. 809–812, 2001.

[15] N. Yang, S. Feng, K. Shedden et al., “Urinary glycoprotein biomarker discovery for bladder cancer detection using LC/MS-MS and label-free quantification,” *Clinical Cancer Research*, vol. 17, no. 10, pp. 3349–3359, 2011.

[16] J. M. Ahn, B. G. Kim, M. H. Yu, I. K. Lee, and J. Y. Cho, “Identification of diabetic nephropathy-selective proteins in human plasma by multi-lectin affinity chromatography and LC-MS/MS,” *Proteomics—Clinical Applications*, vol. 4, no. 6-7, pp. 644–653, 2010.

[17] H. Zhang, J. Saha, J. Byun et al., “Rosiglitazone reduces renal and plasma markers of oxidative injury and reverses urinary metabolite abnormalities in the amelioration of diabetic nephropathy,” *American Journal of Physiology—Renal Physiology*, vol. 295, no. 4, pp. F1071–F1081, 2008.

[18] H. Zhang, X. J. Li, D. B. Martin, and R. Aebersold, “Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry,” *Nature Biotechnology*, vol. 21, no. 6, pp. 660–666, 2003.

[19] T. Vaisar, S. Pennathur, P. S. Green et al., “Shotgun proteomics implicates protease inhibition and complement activation in the anti-inflammatory properties of HDL,” *Journal of Clinical Investigation*, vol. 117, no. 3, pp. 746–756, 2007.

[20] M. Ashburner, C. A. Ball, J. A. Blake et al., “Gene ontology: tool for the unification of biology. The gene ontology consortium,” *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000.

[21] E. I. Boyle, S. Weng, J. Gollub et al., “GO::termfinder—open source software for accessing gene ontology information and finding significantly enriched gene ontology terms associated with a list of genes,” *Bioinformatics*, vol. 20, no. 18, pp. 3710–3715, 2004.

[22] R. C. Miller, E. Brindle, D. J. Holman et al., “Comparison of specific gravity and creatinine for normalizing urinary reproductive hormone concentrations,” *Clinical Chemistry*, vol. 50, no. 5, pp. 924–932, 2004.

[23] G. Eknayan, T. Hostetter, G. L. Bakris et al., “Proteinuria and other markers of chronic kidney disease: a position statement of the National Kidney Foundation (NKF) and the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK),” *American Journal of Kidney Diseases*, vol. 42, no. 4, pp. 617–622, 2003.

[24] M. E. Smoot, K. Ono, J. Ruscheinski, P. L. Wang, and T. Ideker, “Cytoscape 2.8: new features for data integration and network visualization,” *Bioinformatics*, vol. 27, no. 3, pp. 431–432, 2011.

[25] K. Sharma, S. Lee, S. Han et al., “Two-dimensional fluorescence difference gel electrophoresis analysis of the urine proteome in human diabetic nephropathy,” *Proteomics*, vol. 5, no. 10, pp. 2648–2655, 2005.

[26] P. V. Rao, X. Lu, M. Standley et al., “Proteomic identification of urinary biomarkers of diabetic nephropathy,” *Diabetes Care*, vol. 30, no. 3, pp. 629–637, 2007.

[27] D. Vanderschaeghe, N. Festjens, J. Delanghe, and N. Callewaert, “Glucose profiling using modern glycomics technology: technical aspects and applications,” *Biological Chemistry*, vol. 391, no. 2-3, pp. 149–161, 2010.

[28] F. Aregger, C. Pilop, D. E. Uehlinger et al., “Urinary proteomics before and after extracorporeal circulation in patients with and without acute kidney injury,” *Journal of Thoracic and Cardiovascular Surgery*, vol. 139, no. 3, pp. 692–700, 2010.

[29] S. Riaz, S. S. Alam, S. K. Srai, V. Skinner, A. Riaz, and M. W. Akhtar, “Protemic identification of human urinary biomarkers in diabetes mellitus type 2,” *Diabetes Technology and Therapeutics*, vol. 12, no. 12, pp. 979–988, 2010.

[30] N. Nagata, K. Fujimori, I. Okazaki et al., “De novo synthesis, uptake and proteolytic processing of lipocalin-type prostaglandin D synthase, β-trace, in the kidneys,” *FEBS Journal*, vol. 276, no. 23, pp. 7146–7158, 2009.

[31] H. Oda, Y. Shina, K. Seiki, N. Sato, N. Eguchi, and Y. Urade, “Development and evaluation of a practical ELISA for human urinary lipocalin-type prostaglandin D synthase,” *Clinical Chemistry*, vol. 48, no. 9, pp. 1445–1453, 2002.

[32] N. Hirawa, Y. Uehara, M. Yamakado et al., “Lipocalin-type prostaglandin D synthase in essential hypertension,” *Hypertension*, vol. 39, no. 2, part 2, pp. 449–454, 2002.

[33] X. B. Ling, T. K. Sigdel, K. Lau et al., “Integrative urinary peptidomics in renal transplantation identifies biomarkers for acute rejection,” *Journal of the American Society of Nephrology*, vol. 21, no. 4, pp. 646–653, 2010.

[34] G. V. Cohen Freue, M. Sasaki, A. Meredith et al., “Proteinomic signatures in plasma during early acute renal allograft rejection,” *Molecular and Cellular Proteomics*, vol. 9, no. 9, pp. 1894–1967, 2010.

[35] M. Barbalic, G. L. Schwartz, A. B. Chapman, S. T. Turner, and E. Boerwinkle, “Kininogen gene (KNG) variation has a
consistent effect on aldosterone response to antihypertensive drug therapy: the GERA study,” *Physiological Genomics*, vol. 39, no. 1, pp. 56–60, 2009.

[36] S. Chakravarti, T. Magnuson, J. H. Lass, K. J. Jepsen, C. LaMantia, and H. Carroll, “Lumican regulates collagen fibril assembly: skin fragility and corneal opacity in the absence of lumican,” *Journal of Cell Biology*, vol. 141, no. 5, pp. 1277–1286, 1998.

[37] L. Schaefer, I. Raslik, H. J. Grone et al., “Small proteoglycans in human diabetic nephropathy: discrepancy between glomerular expression and protein accumulation of decorin, biglycan, lumican, and fibromodulin,” *FASEB journal*, vol. 15, no. 3, pp. 559–561, 2001.

[38] C. A. Böger, M. H. Chen, A. Tin et al., “CUBN is a gene locus for albuminuria,” *Journal of the American Society of Nephrology*, vol. 22, no. 3, pp. 555–570, 2011.

[39] A. S. Go, G. M. Chertow, D. Fan, C. E. McCulloch, and C. Y. Hsu, “Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization,” *The New England Journal of Medicine*, vol. 351, no. 13, pp. 1296–1305, 2004.

[40] A. Recio-Mayoral, D. Banerjee, C. Streather, and J. C. Kaski, “Endothelial dysfunction, inflammation and atherosclerosis in chronic kidney disease—a cross-sectional study of predialysis, dialysis and kidney-transplantation patients,” *Atherosclerosis*, vol. 216, no. 2, pp. 446–451, 2011.