Evaluation of Tubulointerstitial Lesions' Severity in Patients with Glomerulonephritides: An NMR-Based Metabonomic Study

Nikolaos G. Psihogios,¹ Rigas G. Kalaitzidis,¹ Sofia Dimou,⁹ Konstantin I. Seferiadis,†
Kostas C. Siamopoulos,‡ and Eleni T. Bairaktari*¹

Laboratory of Clinical Chemistry, Department of Nephrology, and Department of Histopathology,
Medical School, University of Ioannina, GR-451 10, Ioannina, Greece

Received March 27, 2007

An ¹H NMR-based metabonomic approach was used to investigate the correlation of histopathologically assessed tubulointerstitial lesions with the urinary metabolite profile in 77 patients with glomerulonephritides submitted to renal biopsy. The presence of renal damage was predicted with a sensitivity of 96% and a specificity of 99%. Patients with mild, moderate, and severe tubulointerstitial lesions were progressively differentiated from the healthy individuals in the Orthogonal Signal Correction Partial Least-Squares-Discriminant Analysis (OSC/PLS-DA) models with a statistically significant separation between those with mild and with severe lesions. The onset of the tubulointerstitial lesions is characterized by decreased excretion of citrate, hippurate, glycine, and creatinine, whereas further deterioration is followed by glycosuria, selective aminoaciduria, total depletion of citrate and hippurate, and gradual increase in the excretion of lactate, acetate, and trimethylamine-N-oxide. NMR-based metabonomic urinalysis could contribute to the early evaluation of the severity of the renal damage and possibly to the monitoring of kidney function.

Keywords: glomerulonephritis • tubulointerstitial • urine • ¹H NMR spectroscopy • metabonomics • kidney

Introduction

Glomerulonephritis (GN) is a group of disorders characterized by inflammation in the filtering unit of the kidney, the glomerulus, which along with a long tubule comprises the anatomical and functional unit of the kidney, the nephron. The unit is surrounded by a functionally important tissue called interstitial tissue. GN may be primary; secondary to drugs, infections, or tumors; or the presenting feature of systemic disease.¹ GN causes significant morbidity and mortality, and is a potentially preventable cause of renal failure and cardiovascular risk. Although the glomerulus is the primary site of damage, subsequent injury to the tubulointerstitium plays a major role in the overall outcome of glomerular disease.¹,²

The diagnosis of GN can be suspected by clinical and laboratory findings, such as proteinuria and abnormal urine microscopy.¹ However, renal biopsy is considered as the main tool for the evaluation of the type and degree of renal injury in almost all cases of glomerulonephritides.³ In addition, careful pathological analysis reveals the extent of tubulointerstitial damage and degree of renal tubular fibrosis, findings that in a number of cases are correlated better with the deterioration of renal function than the degree of glomerular damage itself.⁴

¹H NMR spectroscopy of urine provides overall profiles of low molecular weight (LMW) metabolites that alter characteristicly in response to changes in physiological status, toxic insult, or disease processes.⁵⁻⁷ In situations where renal damage is present in humans or experimental animals, the LMW metabolite profile of urine is significantly altered, and this is closely reflected in the ¹H NMR spectral fingerprint. Furthermore, in studies with experimental animals exposed to region-specific nephrotoxins, the NMR-generated metabolite profiles were characteristically changed according to the exact site and mechanism of the lesion (glomeruli, lower or upper regions of the proximal tubules, renal medulla).⁸,⁹ In the clinical field, NMR urinalysis has contributed to the assessment of renal transplant dysfunction,¹⁰,¹¹ to the early detection of latent tubulointerstitial distortions in glomerulonephritis,¹²,¹³ and to the detection of renal dysfunction in several pathological states.¹⁴⁻¹⁷

The exploitation of the NMR-generated metabolic data sets can be increased by the application of multivariate statistical analysis including pattern recognition (PR) methods that allow sample classification and effective interpretation.¹⁸⁻²⁰ This relatively new approach, known as metabonomics,²¹ has had major applications in clinical and biomedical topics such as drug toxicity assessment, identification of biomarkers of toxicity and disease, and the understanding of the mechanisms of metabolic responses.⁸,²²⁻²⁴

This prospective study investigates the correlation of tubulointerstitial lesions found in renal biopsies with the metabolite...
profile of urine analyzed by NMR-based metabonomics in patients with glomerulonephritides.

Materials and Methods

Subjects. The study initially included 80 consequently admitted patients to the Department of Nephrology of the University Hospital of Ioannina to be submitted to renal biopsy due to renal function abnormalities such as increased serum creatinine and/or proteinuria and chronic renal disease stages 1–3.35 Three patients were excluded during the analysis of the spectroscopic data, as described in the Results. The inclusion criteria were moderate proteinuria (<2 g/24 h) and serum creatinine <3 mg/dL. Eighty-five sex- and age-matched healthy individuals who did not require regular medication other than oral contraception or over-the-counter drugs constituted the control group. All study participants gave informed consent for the investigation, which was approved by the Ethical Committee of the University Hospital of Ioannina.

Histopathology. Patients were submitted to renal biopsy, and the renal tissues were examined by light microscopy and in certain cases with immunofluorescence and/or electron microscopy. Immunostainings were performed on formalin-fixed, paraffin-embedded tissue sections by the labeled streptavidin biotin (LSAB) method. All biopsies were reviewed by one pathologist who was blind to the NMR data. The renal biopsy diagnoses included focal segmental glomerulosclerosis in 23 patients, membranous nephropathy in 13 patients, IgA nephropathy in 8 patients, mesangiproliferative glomerulopathy in 5 patients, systemic lupus erythematosus in 8 patients, vasculitis in 5 patients, diabetic nephropathy in 9 patients, minimal change disease in 4 patients, and other causes in 5 patients. Tubulointerstitial lesions included tubular atrophy, interstitial fibrosis, and mononuclear cell infiltration. The extent of these lesions was graded as follows: mild (n = 25), moderate (n = 27), and severe (n = 25).

Samples. All subjects were requested to fast overnight and abstain from any medication (including over-the-counter drugs), alcohol, and food consumption, known to significantly affect the urinary metabolite profile.26 24 h before sampling, Blood and urine samples were obtained before patients were submitted to renal biopsy. Serum was separated by centrifugation at 1500 g for 10 min. First void urine samples were centrifuged at 1000 g for 10 min, and an aliquot was taken for clinical chemistry tests. Sodium azide (1 g/L, 100 µL) was added to the remaining urine sample to prevent bacterial contamination and stored at −80 °C until NMR analysis.

Chemical Medicine. Analysis of clinical chemistry parameters of serum and urine was carried out on an Olympus AU600 Clinical Chemistry analyzer (Olympus Diagnostica, Hamburg, Germany) by standard procedures. GFR was calculated by the MDRD equation.27

1H NMR Spectroscopy. Four hundred microliters of urine was mixed with 200 µL of phosphate buffer (0.2 M Na2HPO4/0.2 M NaH2PO4, pH 7.4) in order to minimize pH variations, and then a solution of 0.075% sodium 3-trimethylsilyl-(2,2,3,3,4,4,5,5,5)-1-propionate (TSP) in D2O was added.

1H NMR spectra were measured at 300 K on a 500 MHz Bruker DRX NMR instrument operating at 500.13 MHz and running on XWINNMR V.2.6 software. For the suppression of the water signal, the standard 1D pulse sequence NOESY-PRE-SAT (RD-90°-t1-90°-f1-90°-FID acquisition) was used.28 RD was a 3 s relaxation delay to ensure T1-relaxation between successive scans and during which the water peak was selectively irradiated; t1 represented the first increment in the NOESY experiment and was set to 3 µs; t0 was the mixing time of 150 ms, during which the water resonance was again selectively irradiated. For each spectrum, 128 scans were collected into 64K computer data points with a spectral width of 6009.6 Hz. The FIDs were multiplied with an exponential line broadening function of 0.3 Hz prior to Fourier transformation. The acquired NMR spectra were manually corrected for phase and baseline distortions by applying a simple polynomial curve fit with TopSpin 1.2 (Bruker Biospin Ltd.) and referenced to TSP (δH 0.0). The metabolites were assigned according to published literature and 2D experiments (Supplementary Figure 1 in Supporting Information).

Two-dimensional (2D) NMR spectra were carried out on selected samples for identification of urine metabolites. 1H-1H TOCSY spectra were acquired using the MLEV17 spin-lock scheme (mlevsgpphf). Fifty-six transients per increment for 800 increments were collected into 2048 data points in the F2 dimension using a spectral width of 12.02 ppm in both frequency axes and a relaxation delay of 1.2 s. A sine-bell squared function was applied to the data prior to Fourier transformation.

Statistical Analysis. Statistical analysis was performed with Statistica Ver. 6.0 (StatSoft, Inc., Tulsa, OK). Values were expressed as mean value ± standard deviation (SD) and compared by using t test. Significance levels were set at 0.05.

NMR Data Reduction and Pattern Recognition (PR). The 1H NMR spectra were automatically reduced using AMIX (Analysis of MIXtures) software package (version 3.2.4, Bruker Analytik, Rheinstetten, Germany) to 244 continuous integral segments (variables or bins) of equal width of 0.04 ppm corresponding to the chemical shift range δH, 0.2–10.0. The area between 4.38 and 6.30 ppm was excluded to remove any effect of variation from the suppression of the water resonance and from any cross-relaxation effect on the urea signal via solvent exchanging protons. The integral regions of the citrate (2.50–2.58 and 2.66–2.74) and the creatinine (3.02–3.06 and 4.02–4.06) resonances were merged to take into account the pH-dependent peak shifts and formed the “superbins” 2.54 and 2.7 for citrate and 3.04 and 4.04 for creatinine, respectively. All data was normalized by dividing each integral segment by the total area of the spectrum in order to compensate for the differences in overall concentration between individual urine samples. The resulting data matrix, consisting of 194 NMR integral segments, was exported to the SIMCA-P software package (version 10.5, UMETRICS AB, Box 7960, SE 90719, Umeå, Sweden) for the PR analysis. Prior to the analysis, the NMR data were centered and Pareto scaled (scaling factor 1/√SD).

PCA was used for the overview of the metabonomic data set and the spotting of outliers, and then for the detection of any grouping or separation trend.29 The PCA scores plot was used to reveal observations lying outside the 0.95 Hotelling’s T2 ellipse (strong outliers) and the loadings plot to interpret the patterns seen in the scores plot. The model residuals plot, DModX, was used to detect observations that exceeded the critical distance of significance <0.05 (moderate outliers).29

With Partial Least-Squares Discriminant Analysis (PLS-DA) a relationship was sought between the matrix of variables X (NMR spectral bins) and a matrix of dependent variables Y (dummy variables encoding the class membership, i.e., patient or control). The method was used to find the best possible discriminant function (model) that separates renal patients...
from controls as well as the three defined histopathology groups on the basis of their X variables. For the interpretation of the scores plot, the regression coefficients plot was used, which shows all spectral regions that contribute to the separation between the studied groups.

The technique of Orthogonal Signal Correction (OSC) was applied to remove linear combinations of variables X that were orthogonal to the Y vector of the dependent variables, to eliminate the intersubject variability and to describe maximum separation based on class.

The default method of 7-fold internal cross-validation (CV) of SIMCA-P software was applied, and the extracted parameter $Q^2$ was used to provide an estimation of the predictive capability of the PLS-DA models with $Q^2 > 0.5$ considered ‘good’ and $Q^2 > 0.9$ ‘excellent’. The parameter $R^2$ describes the explained variation and how well the data can be mathematically reproduced by the training model.

In addition, validation was performed using both held-back and external data procedures. In held-back data validation, as test set serves a portion of the training set used for the construction of the model, whereas in external data validation, as test set (prediction set) serves a new set of data not used when the model was built.

Held-back data validation for the patients—controls model was performed using 81% of the data as the training set and the remaining 19% as the test set, whereas for the model between patient groups, 68% of the data defined the training set and the remaining 32% the test set. External validation for the patients—controls model was performed using 70% of the data as the training set and the remaining 30% as the prediction set, whereas for the model between patient groups, 68% of the data defined the training set and the remaining 32% the prediction set. All observations were assigned with a class-specific numerical value to form a response Y matrix. Correct classification was based on a predicted Y cutoff of 0.5 with a 95% confidence level.

Correct and incorrect assignments were used to define True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) classification rates and then to estimate as percent sensitivity [TP/(TP + FN) × 100] and specificity [TN/(TN + FP) × 100].

**Results**

**Clinical Chemistry.** In Table 1, the main demographic and clinical chemistry parameters of the populations studied are shown. Patients with GN presented statistically significant higher levels of serum creatinine and urine total proteins and lower levels of serum albumin and GFR than the control population. There were also statistically significant differences among the three groups of patients defined by the severity of the tubulointerstitial lesions in renal biopsy estimation. Patients with moderate and severe lesions had significantly different levels of creatinine, GFR, and urine protein from those with mild lesions. However, no significant differences were observed between those with moderate and severe lesions.

### Table 1. Demographic and Clinical Chemistry Characteristics of the Populations Studied

<table>
<thead>
<tr>
<th>Subgrouping of GN patients</th>
<th>controls</th>
<th>patients</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>mild</td>
<td>25</td>
<td>27</td>
<td>0.25</td>
</tr>
<tr>
<td>moderate</td>
<td>26</td>
<td>27</td>
<td>0.17</td>
</tr>
<tr>
<td>severe</td>
<td>25</td>
<td>27</td>
<td>0.45</td>
</tr>
</tbody>
</table>

**H NMR Spectroscopy.** Four typical $^1$H NMR 500 MHz spectra of urine from a healthy individual and patients with mild, moderate, and severe renal damage indicating the different excretive profile of the LMW metabolites in each case are shown in Figure 1. The main constituents of the urine spectrum from the healthy individual are creatinine, hippurate, citrate, glycine, trimethylamine-N-oxide (TMAO), dimethylamine (DMA), and small amounts of lactate, 3-hydroxybutyrate (3-HB), N-acetyl groups from glycoproteins (N-Acs), and amino acids such as alanine, phenylalanine, and valine. The spectrum from the patient with mild renal damage indicates partial inhibition in the excretion of hippurate, citrate, and glycine, whereas the spectrum from the patient with moderate renal damage reflects further decrease in the excretion of hippurate, citrate, and glycine followed by an increase in the levels of lactate, alanine, and phenylalanine. The spectrum from the patient with severe renal damage indicates significant to complete inhibition in the excretion of hippurate, citrate, and glycine; increased levels of glucose, lactate, alanine, phenylalanine, and histidine; and a slight elevation of the spectrum’s baseline in the region (1.8–0.5 ppm) due to the resonance of the aliphatic moieties of proteins excreted in urine. A significant number of patients, apart from elevated TMAO levels, excreted one or more choline headgroup containing metabolites (between 3.20 and 3.30 ppm), but not in accordance to the severity of the renal damage.

The metabolomic data set initially consisted of 165 urine NMR spectra: 80 from patients that underwent renal biopsy and 85 from healthy individuals. The urine spectra were visually inspected, and 3 from the patient group were excluded: the first one showed 2 intense unidentified peaks at 2.16 and 2.18 ppm (probably metabolites from paracetamol ingestion) and the other two showed intense peaks within the region 3.5–3.8 ppm probably due to metabolites from drugs that the patients had received before sampling.

**Pattern Recognition Analysis.** In this data set (from 77 patients and 85 healthy individuals), PCA was applied, and the scores plot (Figure 2a) showed a separation trend between the two groups with healthy individuals clustering to the right side of the plot and patients spreading mainly to the left side of the plot. The PCA plot also revealed 2 spectra with significant alterations from patients mainly characterized by high peaks of glucose and decreased excretion of creatinine.
loadings plot (Figure 2b), variables (i.e., bins) contributing similar information are grouped together, and hence, metabolites located at the left part (glucose) are positively correlated with patients, whereas those located at the right part (hippurate, citrate, and creatinine) are positively correlated with healthy subjects.

With PLS-DA, an improved separation was achieved still containing, however, a degree of overlapping between the two classes (Figure 2c). The model parameters for the explained variation $R^2$ and the predictive capability $Q^2$ were significantly high, 0.67 and 0.61, respectively (Table 2). The relative regression coefficients plot (Figure 2d) showed that, in addition to the

---

**Figure 1.** $^1$H NMR 500 MHz spectra of urine (δ 0.3−4.6 and 6.8−8.7 ppm) from one healthy subject and patients with mild, moderate, and severe renal damage. Abbreviations: 3-HB, 3-hydroxybutyrate; Ac, acetate; Ala, alanine; Chl, choline headgroup containing metabolites; Cit, citrate; Crn, creatinine; DMA, dimethylamine; Fm, formate; Gly, glycine; Glc, glucose; Hip, hippurate; His, histidine; Lac, lactate; N-Acs, N-acetyl groups from glycoproteins; Phe, phenylalanine; TMAO, trimethylamine-N-oxide; Val, valine.
above-mentioned metabolites, glycine was found in relatively higher levels in controls (positive coefficients), whereas lactate, acetate, TMAO, and aliphatic moieties of proteins were found in relatively higher levels in patients (negative coefficients).
To determine the ability of the 1H NMR-based metabonomic analysis to distinguish the severity of the tubulointerstitial injury, PLS-DA was also applied to compare the healthy subjects with each one of the 3 patient groups characterized by mild, moderate, and severe damage (Table 2). From the corresponding R² and Q² parameters, it can be seen that the more severe the renal damage is assessed, the more the values and therefore the predictive capability of the model increase.

To minimize the possible intrinsic contribution of intersubject variability, the method of OSC filtering was applied from which two orthogonal components were removed and PLS-DA was repeated. Inspection of the two excluded orthogonal components revealed the metabolites possessing the greatest intersubject variability, that is, hippurate, creatinine, citrate, and TMAO, that are known to exhibit a large biological variation. The scores plot of the first two PCs from the resulting OSC/PLS-DA model revealed a clear separation between patients and controls, which was more distinct as the renal damage deteriorated from mild to severe (Figure 2e). The regression coefficients plot for the first component of the OSC/PLS-DA model indicated that the spectroscopic regions contributing to the clustering were almost the same as in the unfiltered data set, but with enhanced importance of the coefficients, mainly for creatinine, hippurate, citrate, and glucose (Figure 2f). The values of the corresponding R² and Q² parameters were improved to 0.80 and 0.73, respectively, as well as those between each patient group and the controls (Table 2). In Supplementary Figure 2 (Supporting Information), the scores plots between each patient group and the controls show clearly that the spectra from patients are placed away from the control region following the severity of the disease. On the basis of the values of the OSC/PLS-DA regression coefficients, controls mainly excreted higher levels of citrate, hippurate, and creatinine, whereas patients mainly excreted higher levels of glucose and a group of unidentified metabolites (3.70–3.74 ppm), choline headgroup containing metabolites, proteins, and acetate (Table 3).

Additional models were developed to compare the three patient groups together as well as pairwise (Table 4, Figure 3). As it is indicated by the R² and Q² parameters of the PLS-DA models in Table 4, the separation between the three patient groups was of low significance. In the pairwise comparison, moderate damage group was partially separated from both mild and severe groups, whereas a more distinct separation was observed between those of mild and severe damage (R² = 0.73 and Q² = 0.46). The OSC/PLS-DA models were of similar discriminating power (Table 4), except for the mild–severe model, in which a higher and significant predictive capability (0.46 vs. 0.55) was noted, seen also in the corresponding scores plot (Figure 3d). On the basis of the values of the regression coefficients, the metabolites that predominantly contributed to the separation of the moderate from the mild damage group were citrate, creatine, phenylalanine, glucose, and a group of unidentified metabolites (3.70–3.74 ppm) and from the severe damage group were creatinine, acetate, hippurate, glucose, and a group of unidentified metabolites (3.70–3.74 ppm) probably from overlapping resonances from glucose, other sugars, and α-protons of amino acids.

To minimize the possible intrinsic contribution of intersubject variability, the method of OSC filtering was applied from which two orthogonal components were removed and PLS-DA was repeated. Inspection of the two excluded orthogonal components revealed the metabolites possessing the greatest intersubject variability, that is, hippurate, creatinine, citrate, and TMAO, that are known to exhibit a large biological variation. The scores plot of the first two PCs from the resulting OSC/PLS-DA model revealed a clear separation between patients and controls, which was more distinct as the renal damage deteriorated from mild to severe (Figure 2e). The regression coefficients plot for the first component of the OSC/PLS-DA model indicated that the spectroscopic regions contributing to the clustering were almost the same as in the unfiltered data set, but with enhanced importance of the coefficients, mainly for creatinine, hippurate, citrate, and glucose (Figure 2f). The values of the corresponding R² and Q² parameters were improved to 0.80 and 0.73, respectively, as well as those between each patient group and the controls (Table 2). In Supplementary Figure 2 (Supporting Information), the scores plots between each patient group and the controls show clearly that the spectra from patients are placed away from the control region following the severity of the disease. On the basis of the values of the OSC/PLS-DA regression coefficients, controls mainly excreted higher levels of citrate, hippurate, and creatinine, whereas patients mainly excreted higher levels of glucose and a group of unidentified metabolites (3.70–3.74 ppm), choline headgroup containing metabolites, proteins, and acetate (Table 3).

Additional models were developed to compare the three patient groups together as well as pairwise (Table 4, Figure 3). As it is indicated by the R² and Q² parameters of the PLS-DA models in Table 4, the separation between the three patient groups was of low significance. In the pairwise comparison, moderate damage group was partially separated from both mild and severe groups, whereas a more distinct separation was observed between those of mild and severe damage (R² = 0.73 and Q² = 0.46). The OSC/PLS-DA models were of similar discriminating power (Table 4), except for the mild–severe model, in which a higher and significant predictive capability (0.46 vs. 0.55) was noted, seen also in the corresponding scores plot (Figure 3d). On the basis of the values of the regression coefficients, the metabolites that predominantly contributed to the separation of the moderate from the mild damage group were citrate, creatine, phenylalanine, glucose, and a group of unidentified metabolites (3.70–3.74 ppm) and from the severe damage group were creatinine, acetate, hippurate, glucose, and a group of unidentified metabolites (3.70–3.74 ppm) probably from overlapping resonances from glucose, other sugars, and α-protons of amino acids.

Table 2. The PLS-DA and OSC/PLS-DA Parameters for the Patients–Controls Models

<table>
<thead>
<tr>
<th>Groups</th>
<th>Total (Patients–Controls)</th>
<th>Mild–Control</th>
<th>Moderate–Control</th>
<th>Severe–Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R²</td>
<td>Q²</td>
<td>R²</td>
<td>Q²</td>
</tr>
<tr>
<td>PLS-DA</td>
<td>0.67</td>
<td>0.61</td>
<td>0.63</td>
<td>0.73</td>
</tr>
<tr>
<td>OSC/PLS-DA</td>
<td>0.80</td>
<td>0.73</td>
<td>0.75</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Table 3. The OSC/PLS-DA Regression Coefficients for the Patients–Controls Model

<table>
<thead>
<tr>
<th>metabolites</th>
<th>Co</th>
<th>P</th>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteins</td>
<td></td>
<td></td>
<td>0.22</td>
</tr>
<tr>
<td>Lactate</td>
<td></td>
<td></td>
<td>0.16</td>
</tr>
<tr>
<td>Acetate</td>
<td></td>
<td></td>
<td>0.20</td>
</tr>
<tr>
<td>Citrate</td>
<td></td>
<td></td>
<td>0.67</td>
</tr>
<tr>
<td>Creatinine</td>
<td></td>
<td></td>
<td>0.44</td>
</tr>
<tr>
<td>Choline metabolites</td>
<td></td>
<td></td>
<td>0.30</td>
</tr>
<tr>
<td>Glycine</td>
<td></td>
<td></td>
<td>0.16</td>
</tr>
<tr>
<td>Glucose</td>
<td></td>
<td></td>
<td>0.31</td>
</tr>
<tr>
<td>Unidentified</td>
<td></td>
<td></td>
<td>0.06</td>
</tr>
<tr>
<td>Creatine</td>
<td></td>
<td></td>
<td>0.55</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td></td>
<td></td>
<td>0.1</td>
</tr>
</tbody>
</table>

*Abbreviations: C, controls; P, patients.  A group of unidentified metabolites (3.70–3.74 ppm) probably from overlapping resonances from glucose, other sugars, and α-protons of amino acids.
The training set for the mild–severe model consisted of 34 patients (17 with mild and 17 with severe renal damage) and the test set of 16 patients (8 with mild and 8 with severe renal damage) (Supplementary Figure 6 in Supporting Information). As seen in Table 6, $R^2$ ranged from 0.81 to 0.84 and $Q^2$ from 0.41 to 0.43, whereas the classification rate was 100% for the mild renal damage sets (8 out of 8) and 88% for the severe renal damage sets (7 out of 8) in all cases.

2. **External Validation.** For each model, the corresponding training and prediction sets were randomly selected, and validation was repeated 3 times with equally numbered sets of different samples each time, which were not used when the models were built.

For the patients–controls model, 114 samples (53 patients/61 controls) from the data set defined the training set, and 48 samples (24 patients/24 controls) served as the prediction set. In Figure 4a, the $Y$-predicted scatter plot of the second repeat is shown with the $Y$ cutoff of 0.5 for classification, whereas the other two plots are shown in Supplementary Figure 7 (Supporting Information). As seen in Table 6, $R^2$ ranged from 0.81 to 0.84 and $Q^2$ from 0.41 to 0.43.
to 0.90 and $Q^2$ from 0.73 to 0.88. The average classification rate was 100% for controls (24 out of 24) and 85% for patients (20.33 out of 24). The calculated sensitivity and specificity were 96% and 99%, respectively.

Similarly for the mild–severe patient group, the training set comprised 34 patients (17 with mild and 17 with severe renal damage), and the prediction set comprised 16 patients (8 with mild and 8 with severe renal damage). In Figure 4b, the $Y$-predicted scatter plot of the second repeat is shown with the $Y$ cutoff of 0.5 for classification, whereas the other two plots are shown in Supplementary Figure 8 (Supporting Information). As seen in Table 6, $R^2$ ranged from 0.84 to 0.87 and $Q^2$ from 0.29 to 0.44. The average classification rate was 83% for the mild renal damage sets (6.67 out of 8) and 79% for the severe renal damage set (6.33 out of 8).

**Discussion**

The application of pattern recognition techniques in NMR-based urinalysis for the evaluation of renal damage has been predominantly focused on experimental studies. In the current study, an $^1$H NMR-based metabolomic approach was used for the first time to investigate the correlation of histopathologically assessed tubulointerstitial lesions with the urinary metabolite profile in patients with glomerulonephritides. The urinary metabolite profiles of the patients with glomerulonephritis at any disease stage presented distinct alterations from those recorded from healthy individuals. These alterations were more obvious in patients with severe renal disease and reflected the extent of the damage in the proximal tubules and/or the tubulointerstitial tissue as it was assessed by the histopathological analysis. Similar alterations have been seen in previously published experimental and human studies. The onset of the tubulointerstitial lesions is characterized by decreased excretion of citrate, hippurate, glycine, and creatinine, whereas further deterioration is followed by glycosuria, selective aminoaciduria, total depletion of citrate and hippurate, and gradual increase in the excretion of lactate, acetate, and TMAO.
Table 6. OSC/PLS-DA Parameters and Classification Scores of the Patients—Controls and the Mild–Severe Models Validated with Held-Back and External Data

<table>
<thead>
<tr>
<th>OSC/PLS-DA</th>
<th>Parameters</th>
<th>Classification</th>
<th>Osc/PLS-DA</th>
<th>Parameters</th>
<th>Classification</th>
<th>Mild–Severe renal damage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training set</td>
<td>Test set</td>
<td>Controls</td>
<td>Patients</td>
<td></td>
<td>Training set</td>
</tr>
<tr>
<td>First model</td>
<td>132 (62P/70C)</td>
<td>30 (15P/15C)</td>
<td>0.80 0.72 15/15</td>
<td>15/15</td>
<td>0.81 0.41 8/8</td>
<td>7/8</td>
</tr>
<tr>
<td>Second model</td>
<td>0.80 0.71 15/15</td>
<td>15/15</td>
<td>0.82 0.43 8/8</td>
<td>7/8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third model</td>
<td>0.80 0.71 15/15</td>
<td>15/15</td>
<td>0.84 0.42 8/8</td>
<td>7/8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>External data</th>
<th>Parameters</th>
<th>Classification</th>
<th>Osc/PLS-DA</th>
<th>Parameters</th>
<th>Classification</th>
<th>Mild–Severe renal damage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training set</td>
<td>Prediction set</td>
<td>Controls</td>
<td>Patients</td>
<td>Training set</td>
<td>Prediction set</td>
</tr>
<tr>
<td>First model</td>
<td>0.81 0.73 24/24</td>
<td>19/24</td>
<td>0.84 0.29 6/8</td>
<td>6/8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second model</td>
<td>0.90 0.88 24/24</td>
<td>20/24</td>
<td>0.87 0.44 7/8</td>
<td>7/8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Third model</td>
<td>0.90 0.87 24/24</td>
<td>22/24</td>
<td>0.84 0.42 7/8</td>
<td>6/8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: C, controls; P, patients; Mi, mild; Se, severe. *Fifteen out of 15 correctly classified.

Depletion of urinary citrate has been attributed to either an impairment of the tricarboxylic acid cycle or to renal tubular acidosis, which typically appears as a generalized proximal tubule dysfunction. A significant decrease of hippurate in urine may be indicative of a metabolic alteration and, even more importantly, of the efficacy of tubular secretion, whereas increased renal hippurate synthesis would require over-utilization of glycine that could account for the low levels of glycine detected in patients with moderate and severe renal damage. The urinary levels of acetate can be affected by the metabolic status of the organism, and increased excretion has been reported in proximal tubular necrosis after exposure to HgCl₂.

Lactic aciduria has been related to increased activity of anaerobic metabolic pathways, to decreased proximal tubular reabsorption, and also appears to be a general marker of renal cortical necrosis. The pattern of selective aminoaciduria, lactic aciduria, and glycosuria that were detected in the present study indicate impairment of the reabsorption mechanisms in the proximal tubular epithelial cells.

Leakage of methylamines in urine, mainly TMAO and DMA, has been reported in medullary damage and in acute graft rejection following renal transplantation. In the present study, elevated excretion of TMAO was detected mainly in the moderate and severe damage groups, but it was not followed by increased excretion of DMA indicating tubulointerstitial distortions rather than papillary necrosis.

The application of PR methods allowed the extraction of the most discriminate information from the multivariate NMR data and an effective sample classification. PCA along with visual inspection of the raw data was important for the detection of outliers in order to assess a consistent metabolic approach. Through PLS-DA, a strong separation trend between patient and control groups was detected, whereas the application of OSC-filtering led to the elimination of the intersubject variation and enabled a clear separation in the resulting models. These models were able to predict the presence of renal damage with a sensitivity of 96% and a specificity of 99% based on a 95% confidence limit for class membership. Patients with mild, moderate, and severe tubulointerstitial lesions were progressively differentiated from the healthy individuals. The comparison between groups showed a statistically significant separation between patients with mild and severe lesions and a high predictive ability of the corresponding OSC/PLS-DA models. Concerning the comparisons between moderate and severe renal damage groups, the separation was not statistically significant, but quite evident in the OSC/PLS-DA models.

It is of interest that similar findings were also observed in conventional clinical chemistry analysis. However, it is well-known that the degree of proteinuria is not always correlated to the severity of the interstitial damage. Proteinuria usually reflects an increase in glomerular permeability that allows the filtration of normally nonfiltered macromolecules such as albumin, and thus, tubular proteinuria may be masked by an overt glomerular hyperfiltration. On the other hand, serum creatinine and GFR estimation are affected by factors such as blood pressure levels and hydration of the patients. Therefore, NMR findings further support the existence of the tubulointerstitial lesions and could be a useful tool to the global assessment of kidney damage and contribute to the attenuation of the confounding factors mentioned above.

In conclusion, since the coexistence of tubular and interstitial lesions in glomerulonephritides is of crucial prognostic importance for the progress of renal glomerular function, NMR-based metabolicomic approach, as a rapid and noninvasive technique, could contribute to the early evaluation of the severity of the renal damage and possibly to the monitoring of the kidney function. This last indication is currently being evaluated in our laboratory.

Abbreviations: LMW, low molecular weight; ¹H NMR, proton nuclear magnetic resonance; PR, pattern recognition; MDRD, modification of diet in renal disease; TSP, 3-trimethylsilyl-(2,2,3,3-d³)H₁)-1-propionate; FID, free induction decay; PCA, principal component analysis; PC, principal component; PLS-DA, partial least-squares discriminant analysis; OSC, orthogonal signal correction; 3-HB, 3-hydroxybutyrate; TMAO, trimethylamine-N-oxide; DMA, dimethylamine.

Acknowledgment. The research Project is co-funded by the European Union - European Social Fund (ESF) & National Sources, in the framework of the program “HRAKLEITOS” of the “Operational Program for Education and Initial Vocational Training” of the 3rd Community Support Framework of the Hellenic Ministry of Education. The NMR spectra were recorded on a 500 MHz Bruker DRX NMR instrument, at the NMR Laboratory of NCSR "Demokritos", Agia Paraskevi, Greece.

Supporting Information Available: Figures showing the ¹H NMR 2D TOCSY spectra of urine from a patient with...
severe renal damage: OSC/PLS-DA scores plots of the urinary spectroscopic data from controls and patients with mild, moderate, and severe renal damage; OSC/PLS-DA scores plots of the urinary spectroscopic data, after the removal of the aliphatic moieties of proteins from patients with mild, moderate, and severe renal damage; OSC/PLS-DA scores plots of the urinary spectroscopic data of the aliphatic moieties of proteins from patients with mild, moderate, and severe renal damage; OSC/PLS-DA scores plots of the other 2 OSC/PLS-DA patients—controls and mild—severe models validated with held-back data; and Y-predicted scatter plots of the other 2 OSC/PLS-DA patients—controls and mild—severe models validated with external data. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(2) Nath, K. A. Tubulointerstitial changes as a major determinant in the progression of renal damage. Am. J. Kidney Dis. 1992, 20, 117.
research articles


PR070172W