Identification and Validation of Urinary Biomarkers for Differential Diagnosis and Evaluation of Therapeutic Intervention in ANCA associated Vasculitis

Marion Haubitz1*, David M. Good2*, Alexander Woywodt1, Hermann Haller1, Harald Rupprecht4, Dan Theodorescu5, Mohammed Dakna6, Joshua J. Coon2,3, and Harald Mischak6

1 Department of Nephrology, Hannover Medical School, Hannover, Germany
2 Departments of Chemistry and 3Biomolecular Chemistry, University of Wisconsin – Madison, Madison, WI, USA
4 Department of Nephrology, LMU Munich, Germany
5 Department of Urology, University of Virginia, Charlottesville, VA, USA
6 Mosaiques Diagnostics GmbH, Hannover, Germany
* contributed equally

To whom correspondence should be addressed:
Marion Haubitz, M.D.
Department of Nephrology
Medical School Hannover
30623 Hannover
Germany
Tel. 0511/532 6319
Fax 0511/5328108
E-mail: Haubitz.Marion@MH-Hannover.de

Running title: Urinary biomarkers in Vasculitis
Summary

Renal activity and smoldering disease is difficult to assess in ANCA associated vasculitis (AAV) due to renal scarring. Even repeated biopsies suffer from sampling errors in this focal disease - especially in patients with chronic renal insufficiency. We have applied capillary electrophoresis coupled to mass spectrometry towards urine samples from patients with active renal AAV to identify and validate urinary biomarkers that enable differential diagnosis of disease and assessment of disease activity. The data were compared to healthy individuals, patients with other renal and non-renal diseases, and patients with AAV in remission. 113 potential biomarkers were identified that differed significantly between active renal AAV and healthy individuals and patients with other chronic renal diseases. Of these, 58 could be sequenced. Sensitivity and specificity of models based on 18 sequenced biomarkers were validated using blinded urine samples of 40 patients with different renal diseases. Discrimination of AAV from other renal diseases in blinded samples was possible with 90% sensitivity and 86.7 to 90% specificity, depending on the model. Ten patients with active AAV were followed for 6 months after initiation of treatment. Immunosuppressive therapy led to a change of the proteome towards “remission”. 47 biomarkers could be sequenced that underwent significant changes during therapy, together with regression of clinical symptoms, normalization of CRP, and improvement of renal function. Proteomic analysis with CE-MS represents a promising tool for fast identification of patients with active AAV, indication of renal relapses and monitoring for ongoing active renal disease and remission without renal biopsy.
Introduction

Systemic vasculitides are a heterogeneous group of disorders with inflammation of the blood vessel wall as their common hallmark. These disorders often pose difficulties with regard to diagnosis and monitoring of disease activity both at the initial presentation and during follow-up. In one subgroup of small-vessel vasculitides, the advent of antineutrophil cytoplasmic antibodies (ANCA) in the 1980s not only provided a new pathogenetic concept, but also a diagnostic marker (1, 2). In this group, the granulomatous ANCA associated vasculitis (GAAV) (previously named Wegener’s granulomatosis) and the microscopic polyangiitis (MPA) share several common features, including pauci-immune focal crescentic necrotizing glomerulonephritis and often a pulmonary capillaritis (3). Because of the association with ANCA, these diseases (together with the Churg-Strauss-Syndrome) are sometimes collectively referred to as ANCA-associated vasculitis (AAV). Recently, circulating endothelial cells have emerged as an important marker correlating with severity and activity of the systemic vasculitic disease and their clinical use in ANCA associated small-vessel vasculitis (AAV) has been demonstrated (4, 5). Regarding renal involvement, which is found in up to 80 - 90% of the patients with AAV, activity is defined by kidney biopsy with pauci-immune necrotizing glomerulonephritis. Renal involvement may occur or recur at every point of the disease and the follow-up, even if other organ involvement was controlled by immunosuppressive therapy. Early detection is important as renal prognosis depends on early administration of immunosuppressive treatment (6), and scarring and relapses increase the risk for terminal renal failure, which itself is a risk factor for patient survival. As kidney biopsy is invasive and the risk of bleeding increases with chronic renal damage, surrogate markers, such as rising creatinine, increasing proteinuria, and most importantly erythrocytes and erythrocyte cast in the urinary sediment, were used. However, these markers have limitations. Microhematuria might persist despite remission, proteinuria might increase
Urinary biomarkers in Vasculitis
despite improvement in renal function, and other renal diseases can also develop (7). Therefore, new markers for renal disease activity are eagerly awaited.

Recently, proteome analysis of urine has presented itself as a promising tool in the definition of chronic renal diseases (8, 9). We have developed an analytical platform for human urine analysis using capillary electrophoresis coupled on-line to an Electrospray Ionization Time of Flight Mass Spectrometer (CE-ESI-TOF-MS) (8, 10, 11). This approach permits the rapid analysis of the low molecular weight urinary proteome/peptidome in a single step and enabled identification and validation of several urinary biomarkers in patients with different renal diseases (11-14).

In this study, we aimed towards identification of biomarkers for active AAV and response to immunosuppressive treatment. The results indicate that urinary proteome analysis enables differential diagnosis and monitoring of renal disease activity of AAV.
Experimental Procedures

Patients, Procedures and Demographics

Informed consent was obtained from all patients and healthy controls after local ethics committee approval. These studies were also performed in accordance with the Helsinki Declaration.

Urine samples were collected in the morning, after voiding the first urine of the day, from 18 patients with active AAV with a pauci-immune necrotizing glomerulonephritis [8 granulomatous ANCA associated vasculitis (GAAV) and 10 microscopic polyangiitis (MPA)] and 19 patients with inactive AAV all with previous biopsy-proven pauci-immune necrotizing glomerulonephritis and comparable therapy. Diagnosis of GAAV (previously named Wegener’s granulomatosis) and MPA was made according to the American College of Rheumatology (ACR) criteria and the Chapel Hill consensus Conference definition (15). Five of the 18 patients with active AAV were treated with immunosuppressive drugs, 4 with steroids prior to the urinary analysis. All renal biopsies were analyzed by a triple diagnostic procedure including immunohistology as well as electron microscopy. Clinical characteristics of all patients with AAV are summarized in Table 1.

Sensitivity and specificity were evaluated in 40 blinded samples from individuals with different, biopsy-proven active glomerular diseases: 10 with active AAV, 29 with other glomerular diseases (membranous glomerulonephritis n=9, IgA-nephropathy n=4, focal segmental glomerular sclerosis n=6, minimal change disease n=2, lupus nephritis n=4, others n=4) and one normal control.

In 10 of the active AAV patients (median age 61.5 range 19 to 69 years, 5 female), repeated urine samples were collected prior to, and 1, 3, and 6 months after initiation of immunosuppressive treatment with prednisolone and intravenous cyclophosphamide pulses together with the documentation of creatinine, urinary sediment, proteinuria, and C-reactive...
protein (CRP). Clinical activity was expressed by the Birmingham vasculitis activity score (BVAS) (16).

In addition, CE-MS data from 225 patients with non-vasculitis kidney related diseases and 200 healthy controls were utilized. Further controls for specificity were samples from patients with CMV infection after renal transplantation (influence of a viral infection), nephrolithiasis, bladder cancer, renal cancer (patients with other causes of microhematuria), and patients with hypertension and a diabetes type II without microhematuria (influence of blood pressure). Samples and patients’ demographic data are described in recent manuscripts (12, 13, 17-22). All samples were midstream urine, following a standardized collection protocol, as also outlined in (23). In short, second urine of the day was collected and frozen immediately after collection without the addition of any preservatives.

**Sample preparation and CE-MS analysis**

All urine samples for CE-MS analyses were from spontaneously voided urine and were stored at -20°C for up to 3 years until analysis. Of note: we were unable to detect significant storage-related degradation in samples that were stored at -20°C for > 10 years, when comparing these to similar samples that were stored for only several weeks. For proteomic analysis, a 0.7 mL aliquot of urine was thawed immediately before use and diluted with 0.7 mL of 2 M urea, 10 mM NH₄OH containing 0.02% SDS. To remove higher molecular mass proteins, samples were filtered using Centrisart ultracentrifugation filter devices (20 kDa molecular weight cut-off; Sartorius, Goettingen, Germany) at 3,000 rcf until 1.1 ml of filtrate was obtained. This filtrate was applied onto a PD-10 desalting column (Amersham Bioscience, Uppsala, Sweden) equilibrated in 0.01% NH₄OH in HPLC-grade H₂O (Roth, Germany) to remove urea, electrolytes, salts, and to enrich polypeptides present. Finally, all samples were lyophilized, stored at 4°C, and suspended in HPLC-grade H₂O shortly before CE-MS analysis, as described (20).
CE-MS analysis was performed as described (20, 24) using a P/ACE MDQ capillary electrophoresis system (Beckman Coulter, Fullerton, USA) on-line coupled to a Micro-TOF MS (Bruker Daltonic, Bremen, Germany). Data acquisition and MS acquisition methods were automatically controlled by the CE via contact-close-relays. Spectra were accumulated every 3 s, over a range of m/z 350 to 3000 Th. Accuracy, precision, selectivity, sensitivity, reproducibility, and stability are described in detail elsewhere (20, 25). In short, the detection limit is in the range of 1 fmol, depending on the ionization properties of the individual peptide. In a urine sample, the detection limit (in the crude sample, before processing) is in the range of 100 - 1000 fmol/ml. In 15 consecutive analyses the standard deviation was between 2.5 and 4% and intermediate precision was 5.5%. Stability and variability over time were also assessed and are available in the Supplementary Documentation.

Data processing and cluster analysis

Mass spectral ion peaks representing identical molecules at different charge states were deconvoluted into single masses using MosaiquesVisu software (26), accessible at www.proteomiques.com. Migration time and ion signal intensity (amplitude) were normalized based on 29 collagen fragments that serve as internal standards (25). These internal polypeptide standards are the result of normal biological processes and appear to be unaffected by any disease state studied to date (greater than 5000 samples analyzed to date) (23). The resulting peak list characterizes each polypeptide by its molecular mass [kDa], normalized migration time [min], and normalized signal intensity. All detected polypeptides were deposited, matched, and annotated in a Microsoft SQL database, allowing further analysis and comparison of multiple samples (patient groups). To establish identity of polypeptides observed in different samples, a linear function was employed that allowed, depending on the mass of the polypeptide, a 50 ppm absolute mass deviation for peptides of 800 Da that increased linearly to 100 ppm absolute mass deviation for peptides with a
maximum mass of 20 kDa. A similar linear function was used when comparing CE migration times, allowing a 5% absolute deviation. CE-MS data of all individual samples can be accessed in **Supplementary Table 3**.

Disease-specific polypeptide patterns were generated using both support-vector-machine (SVM) based MosaCluster software (14), and also a method using a linear combination of log-transformed data (27). To perform the linear combination, all normalized signal intensity values for possible biomarkers below 1 were substituted with a value of 1. The average signal intensity for a specific biomarker over all cases was compared to the average intensity for the same biomarker over all controls. To avoid artificial weighting of specific biomarkers in the set due to the difference in observed signal intensities for case and control, the relative distance between the two averages (case and control) was always set to two. This relative distance of signal intensities between the disease and control samples was provided using the formula: 

\[
\frac{2}{|\bar{x}_{\text{case}} - \bar{x}_{\text{control}}|} \left( A_i - \text{mean}_{\text{averages}} \right),
\]

where \( A_i \) is the log-transformed signal intensity of the \( i^{th} \) biomarker in the \( k^{th} \) sample in either the test set or the blinded set, \( \text{mean}_{\text{averages}} \) is the average of the mean intensity of all possible markers for test set samples, \( \bar{x}_{\text{case}} \) represents the mean observed signal intensity of the possible biomarker from all vasculitis samples and \( \bar{x}_{\text{control}} \) represents the mean signal intensity of the possible biomarker from the combined control samples (apparently healthy individuals and patients with chronic renal disease other than vasculitis). This linear classification procedure is in some way a simplified version of the nearest centroids classification approaches. Variants of these approaches are those that include shrinkage as implemented in the SAM (significance analysis of microarrays) and PAM (prediction analysis of microarrays) classifiers (28), as well as those without shrinkage like the classical Fishers DLDA (diagonal linear discriminant analysis) and the ClaNC (Classification to Nearest Centroids) procedures (29).
Statistical methods, definition of biomarkers and sample classification

Estimates of sensitivity and specificity were calculated based on tabulating the number of correctly classified samples. Confidence intervals (95% CI) were based on exact binomial calculations and were carried out in MedCalc version 8.1.1.0 (MedCalc Software, Mariakerke, Belgium, http://www.medcalc.be). The Receiver Operating Characteristic (ROC) plot was obtained by plotting all sensitivity values (true positive fraction) on the y-axis against their equivalent (1-specificity) values (false positive fraction) for all available thresholds on the x-axis (MedCalc Software). The area under the ROC curve (AUC) was evaluated, as it provides a single measure of overall accuracy that is not dependent upon a particular threshold (30). The reported unadjusted p-values were calculated using the natural logarithm-transformed intensities and the Gaussian approximation to the t-distribution. Statistical adjustment for multiple testing was performed by adjusting according to the strong control of the family wise error using the Bonferroni procedure (31). In addition, the p-values according to the false discovery rate (FDR) adjustments of Benjamini-Hochberg were calculated (case vs. control; (32)).

Sequencing

Candidate biomarkers were sequenced using LC-MS/MS analysis as recently described in detail (13, 33), using instruments with electron transfer dissociation (ETD) capability (34-36). Spectral data were searched against the IPI human non-redundant database using the Open Mass Spectrometry Search Algorithm (OMSSA; free from NCBI, http://pubchem.ncbi.nlm.nih.gov/omssa/), using an e-value cut-off of 1.00E-2. All matched sequences were manually validated. All sequences obtained from human urine can be assessed at http://mosaiques-diagnostics.de/diapatpcms/mosaiquescms/front_content.php?idcat=257, as described in (23).
Results

The samples used and the flow of data are schematically shown in Figure 1.

Identification of urinary biomarkers for differential diagnosis of vasculitis

Data from CE-MS analysis of 18 patients with active ANCA associated vasculitis were compared to data obtained from healthy volunteers \([n=200]\), focal segmental glomerulosclerosis \([\text{FSGS}; n=30]\), diabetic nephropathy \([\text{DN}; n=78]\), IgA nephropathy \([\text{IgAN}; n=57]\), minimal change disease \([\text{MCD}; n=25]\), and membranous glomerulonephritis \([\text{MNGN}; n=35]\) patients. Clinical data of these patients have been described in several recent manuscripts \((12, 17, 27, 37)\). This resulted in 6 datasets in the control groups, one from healthy volunteers, and 5 from disease controls. In total, 425 samples were utilized as controls with the aim to define biomarkers that enable differentiation between vasculitis and healthy controls, as well as other renal diseases. The compiled data from the CE-MS analysis of the 7 groups are shown in Figure 2. A minimum frequency of 0.3 in cases or controls was required for any feature to be accepted as a potential biomarker.

Upon application of the student’s t-test adjusted for multiple testing according to the stringent method of Bonferroni \((31)\), 113 possible biomarkers with an adjusted p-value of 0.05 or lower could be defined. A graphic depiction of the distribution of these 113 potential biomarkers in the 7 different groups is shown in Figure 3; all data, including the AUC values for the individual biomarkers and the p-value after adjustment of the FDR according to the method of Benjamini and Hochberg, are given in Supplementary Table 1.

58 of the 113 initially defined potential biomarkers could be sequenced, as also shown in Supplementary Table 1. The most prominent observation is the appearance of C- and N-terminal hemoglobin fragments, while no fragments from the middle of the hemoglobin protein could be identified yet. Further, up-regulation of alpha1-antitrypsin and albumin fragments could be observed, while several collagen fragments appear to be underrepresented in comparison to the controls.
Limiting the available data to only the sequenced biomarkers, we attempted to combine these to a vasculitis-specific biomarker model. As the active vasculitis group consisted of only 18 patients, we reduced the number of biomarkers used in the panel without decreasing accuracy. To achieve this lower number of biomarkers included in the model, a take-one-out method was employed, whereby the model was run for all biomarkers minus a specific biomarker for each of the 58 initial possibilities. Biomarkers which did not influence the accuracy, specificity, and sensitivity of the model upon self-validation when taken out were left out of the final panel; this ensured only the top-scoring biomarkers were included in the final panel, which consisted of 18 biomarkers (Table 2).

Two algorithms were used to combine the 18 biomarkers into a disease-specific model. A model using support vector machines (SVM) resulted in correct classification of 424/443 of the samples in the training set upon self-validation, giving an accuracy of 95.7% with sensitivity and specificity of 100 % and 95.5%, respectively. 94.4 % sensitivity and 95.1% specificity were obtained upon complete cross-validation, using an ideal cutoff of 0.2. Alternatively, linear combination was used. This model resulted in 94.4 % sensitivity and 97.4% specificity at a cutoff of 4.8. Receiver operating characteristic (ROC) curves from both models are shown in Figure 4A.

The biomarker models were subsequently validated in a blinded dataset. To this end, 40 samples were analyzed using CE-MS and evaluated using both the “vasculitis specific” SVM-based model as well as the linear combinations (results see Table 3). Blinded analysis with these 2 classification algorithms correctly identified 9 out of the 10 active AAV samples as active AAV. Of 30 patients without active AAV (9 MNGN, 6 FSGS, 4 IgAN, 4 proliferative lupus nephritis (LN), 2 MCD, 2 membranoproliferative glomerulonephritis (MPGN), 2 glomerular sclerosis (GS), and 1 NC), 27 or 26 patients (SVM-model: 4 false classified, Lin-model: 3 false classified) were correctly identified as not having AAV. Of the
false positive scoring patients, the majority (two or three, depending on the algorithm used) had a progressive IgA nephropathy.

As evident from Table 3 and Figure 4, classification based on the linear combination or SVMs gave similar results; the linear combination appeared to show slightly higher accuracy, but the difference was insignificant. The misclassified patients were generally the same, irrespective of model or algorithm. ROC curves from the blinded application of both models onto the 40 samples are shown in Figure 4B.

As the data on the blinded cohort suggested that the biomarkers identified may lack specificity in differentiating IgAN from vasculitis, we have investigated the performance of the biomarker model in a set of 18 IgA patients matched on a case to case basis concerning creatinine and proteinuria (compared to the 18 active vasculitis patients). Of these samples, 14 scored negative (correct) and 4 positive (false) for vasculitis, resulting in a specificity of 78% in the linear model, 13 scored negative (correct) and 5 positive (false) for vasculitis in the SVM-model.

We also assessed the significance of the tentatively identified 113 biomarkers for AAV in the test set. As the number of cases and controls was much smaller in the test set, we did not adjust for multiple testing. Of the 113 biomarkers, 101 were found above the required frequency threshold of 0.3 in the test set. Of these, 55 were also found to be significant (p-value below 0.05) in the test set. An additional 12 biomarkers showed a trend (p-value between 0.05 and 0.1), and 34 did not show any significant correlation with vasculitis in the test set. The unadjusted p-values of the biomarkers in the test set are all given in Supplementary Table 1.

To further examine the specificity of the biomarker panel regarding the influence of virus infections on the one hand and microhematuria and hypertension on the other, we have examined additional patient groups. Urine samples obtained from 51 patients with CMV infection after renal transplant (22) scored negative in all but one patient. Further controls
were urine samples from patients with microhematuria due to nephrolithiasis (20) (30 of 33 negative) or bladder cancer (20) (95 of 110 negative) and patients with renal cancer (20) (111 of 113 negative). Urine from 11 patients with proliferative lupus nephritis was investigated and only one patient was positive. 17 Urine samples from patients with hypertension and type II diabetes without microalbuminuria all scored negative for the vasculitis pattern (21). To investigate if the slightly reduced specificity observed in the cohort of the samples from patients with bladder cancer could be attributed to specific biomarkers, we investigating only the 15 samples that scored positive in this group. We were unable to correlate a specific biomarker with the false positive scoring of these patients, but rather the combination of fragments from hemoglobin, alpha-1-antitrypsin, fibrinogen, and alpha-2-HS-glycoprotein. Overall, the data indicate a high specificity of the biomarker pattern used.

**Biomarkers for disease activity**

We subsequently sought to identify biomarkers that could potentially be used to address response to therapy/activity of disease. 18 samples from patients having active vasculitis and 19 samples from patients who had undergone treatment for the disease and were in stable clinical remission for more than 18 months were used to identify potential biomarkers. Comparison of the datasets resulted in 166 possible biomarkers when applying an adjusted p-value cut-off of 0.05. As potentially useful biomarkers should also display significant differences between active vasculitis and apparently healthy normal controls, we also compared the 18 patients with active disease to the 200 normal controls using the same stringent statistics. This resulted in the identification of 266 potential biomarkers, 122 of these found to be significant in both comparisons. The complete data on these 122 potential biomarkers are given in **Supplementary Table 2**. 47 of these potential biomarkers could be sequenced. As these biomarkers should enable assessment of disease activity, we decided not
Urinary biomarkers in Vasculitis

to reduce their number further, but instead use all 47 sequenced biomarkers to enable a maximum distance between disease and control with respect to the dataspace. Upon complete cross-validation, this panel of biomarkers gave 37 (100%) correct identifications in the training set. As SVMs do not give any level of confidence, and may not be ideally suited to address linear changes as expected in progression, we also combined the 47 potential biomarkers in a linear model, which enabled classification of the training set with 100% accuracy.

In ten patients, measurements were done before and 1, 3 and 6 months after initiation of immunosuppressive treatment. With successful treatment, clinical activity expressed by the Birmingham vasculitis activity score (BVAS), C-reactive protein (CRP), creatinine, and proteinuria declined after 6 months as compared to the value at month 0: BVAS, \( p=0.002 \) (median at 0, 1, 3 and 6 months 19.5, 6, 0 and 0, respectively), CRP, \( p=0.002 \) (median at 0, 1, 3 and 6 months 117, 15, 3 and 1.5 mg/l, respectively), creatinine, \( p=0.098 \) (median at 0, 1, 3 and 6 months 193, 128, 118 and 131 \( \mu \)mol/l, respectively) and proteinuria, \( p=0.0098 \) (median at 0, 1, 3 and 6 months 1.0, 1.2, 0.6 and 0.4 g/24h, respectively). ANCA titer declined significantly from a median of 1:32 (range 1:8 to 1:4096) at month 0, to 1:16, 1:8 and 0 at month 1, 3 and 6, respectively (\( p=0.002 \)). Prednisolone doses decreased to a median of 50mg at month 1, 15mg at month 3 and 10mg at month 6. After 6 months all but one patient (still receiving cyclophosphamide pulses) were on maintenance treatment with azathioprine or mucophenolat mofetil. The urinary proteome patterns changed gradually from active to inactive. This change was delayed in two patients, 2255 and 2259. These two patients experienced a transient rise in creatinine that was attributed to an infection, or a subsequent relapse, respectively. While these data further support the value of proteome-based monitoring, no definitive conclusion can be drawn from two patients only.

The assessment of all samples utilizing an SVM-based model or a linear combination is given in Figure 5. These results are indicative of decreasing disease activity with a decline in
Urinary biomarkers in Vasculitis

BVAS, CRP, and renal parameters with increasing time of immunosuppressive treatment. To examine if these results merely reflect changes in proteinuria, microhematuria, or renal function, we have investigated the correlation between the scores obtained from the two biomarker models and proteinuria, microhematuria, and creatinine. While the distribution of data in the comparison of proteinuria and creatinine with the proteomic patterns clearly indicated a random distribution and no correlation at all, we were able to find some correlation between erythrocytes in urine and proteomic scores. Since both proteomic pattern and microhematuria are expected to change as a response to treatment, it is not clear if these two variables are dependent. We have examined all 22 hemoglobin fragments. In general, the reduction was >10-fold after the first month of therapy. After 3 months, most fragments were generally below detection limit in the samples. Certainly, in general, erythrocyturia/hematuria only is not reflected by the proteomic pattern, as patients with other diseases who have microhematuria do not score positive with the vasculitis-specific proteomic pattern.

Discussion

In this study we show that urinary proteome analysis with CE-MS permits differentiating of patients with active AAV versus healthy individuals and patients with other chronic renal diseases. A panel of biomarkers permits distinguishing between patients with active AAV and patients in remission. Initiation of immunosuppressive treatment results in a change of the pattern from active to inactive disease, correlating with clinically decreasing vasculitis activity and the achievement of remission.

Both the cross-sectional and longitudinal studies resulted in the identification of >100 potential biomarkers based on statistical evaluation, even when adjusted for multiple testing. As a consequence, we limited ourselves to only the biomarkers that could be identified. Employing the 18 biomarkers in the SVM-based model and in the linear model produced similar results. However, in all cases the linear model insignificantly outperformed the SVM-
based model. This is likely due to the fact that SVM-based models tend to over-fit data, especially when the training sets used are not overly large and may therefore not reflect the true diversity of a real population. Consequently, the use of SVM-derived models should be limited to large datasets that, due to their diversity, are less vulnerable to overfitting. This assumption could be further confirmed in experiments involving larger sets of cases and controls, where SVM and other multiparametric machine learning algorithms consistently outperformed linear combination, while linear combination gave better results if only a smaller subset of cases and controls was used (Mischak et al., manuscript in preparation).

To enable assessment of the validity of the single biomarker independent from any classification algorithm involved, we also tested if the prediction of validity (based on statistics) was correct for each of the 113 individual biomarkers. As a rough estimate, we expected that 6 biomarkers (5%) would turn out to be not significant, while the others should again show significance. As evident, the expectations were not fully met by the dataset. Only 55% of the assessed biomarkers again showed significant differences, 66% if a trend was also valued acceptable. However, 34% revealed no significant association with AAV. These findings further support an observation that we made in several previous studies: even adjustment for FDR and for multiple testing results in an overestimation of the actual quality of the biomarkers in the training set, likely due to unknown bias introduced. These results further imply that any proteomic results must be confirmed in a blinded test set, mere statistical testing (even if performed properly) is insufficient to establish association of a biomarker for disease with good confidence.

It should be noted that the majority of the false positive results from both models were from samples which, upon unblinding, were shown to be from patients with severe IgA nephropathy, which may be interpreted as a limited form of Schoenlein-Henoch vasculitis. Patients with severe IgA nephropathy may present quite similarly to patients with ANCA associated vasculitis regarding the kidney with a rapidly declining renal function. In a
matched comparison, 4 or 5 (depending on the model used) of 18 patients with IgA nephropathy scored positive. Using samples from patients with bladder cancer, nephrolithiasis or renal cancer, we show that the reduced specificity observed in IgA nephropathy is not merely due to microhematuria, as these additional cohorts with microhematuria did not show an increase in false positive results.

No correlation between creatinine, proteinuria, and proteomic scores for vasculitis could be shown. However, erythrocyturia correlated with the proteomic scores (results are shown in Supplementary Figure 1), which is not unexpected as microhematuria increases in active renal vasculitis and is linked to vasculitis activity. A missing influence of renal function on a specific pattern is also supported by our previous study in IgA nephropathy (18) where no systematic differences in frequency of the discriminating polypeptides could be shown when patients with creatinine <120µmol and ≥120µmol were compared. These IgA nephropathy patients had a degree of renal insufficiency and proteinuria not different from the 18 AAV patients of this study (18). The results also indicate that proteome analysis delivers an additional parameter that can be used for clinical evaluation, independent from creatinine levels, and may enable a more accurate assessment of disease and disease state.

58 of the biomarkers used within the differential diagnosis pattern and 47 biomarkers used in the disease state classification panel could be sequenced. In both of the models, the most frequently observed peptides were proteolytic products of hemoglobin, including multiple fragments from both the alpha and beta subunits. The existence of these fragments is to be expected, as microhematuria is a characteristic finding in vasculitis. Of the 37 fragments of the hemoglobin alpha and beta chain that could be identified to date in human urine, 21 were found to be statistically significant biomarkers in this study. Remarkably, only C- and N-terminal fragments could be defined as biomarkers, while fragments from the core of the molecules, although present in urine, showed no significant value as biomarkers. The existence of specific fragments may be a result of the specifically released proteases in ANCA.
associated vasculitis after activation of neutrophiles by ANCA (38). Consequently, these fragments, being indicators of such specific protease activity, appear suitable in differentiating vasculitis from other pathological situations where hematuria is observed, but results in different fragments due to divergent proteolytic activity. This is also reflected by the data on patients with microhematuria-associated diseases like nephrolithias, bladder, and renal cancer, where >90% of the samples scored negative for vasculitis.

In addition to these hemoglobin fragments, we found several fragments from albumin and alpha-1-antitrypsin increased in active vasculitis. Of the 21 alpha-1-antitrypsin peptides identified, 14 serve as biomarkers for vasculitis. Alpha-1-antitrypsin has a central role in controlling tissue damage by inhibiting proteases including elastase and proteinase 3, both released after activation of primed leukocytes by ANCA in vitro (54). Without inhibition, proteinase 3 and elastase are capable of inducing endothelial cells apoptosis, cytokine and tissue factor production (39-41). Moreover, proteinase 3 and elastase have been shown in kidney biopsies of ANCA associated pauci-immune necrotizing vasculitis (42) and elevated levels of proteinase 3 and elastase could be demonstrated in patients with ANCA associated vasculitis (43, 44). Thus these results point to a role of alpha1-antitrypsin as a mediator of activity in ANCA associated vasculitis.

Of the 16 fragments from albumin identified to date in human urine, 5, all starting at the same position at the N-terminus, could be identified as biomarkers in this study. As also observed in diabetic nephropathy (27), specific collagen fragments were found reduced in comparison to the controls. Though collagen itself is very common, the ability to use collagen fragments as specific biomarkers for multiple different disease states indicates that these fragments are the result of variations in the in vivo protease activity and are therefore directly linked to the disease activity. As outlined in a recent review (45), it is likely that the urinary peptides reflect to a substantial degree turnover of extracellular matrix. (Patho)physiological changes in such turnover, e.g. in the case of fibrosis, but also in
Urinary biomarkers in Vasculitis

inflammation, will display as indicative alterations of several urinary peptides. When comparing the changes observed in the kidney in diabetic nephropathy and vasculitis, the inflammatory component appears as one hallmark of vasculitis, while fibrosis is in contrast much more commonly found in diabetic nephropathy. As a consequence, we were able to detect the significant decrease of several collagen fragments in diabetic nephropathy (indicating reduced degradation of collagen, which results in increase of extracellular matrix), which was by far less pronounced in the samples from vasculitis patients reported here. At the same time, the up-regulation of peptides from proteins involved in inflammatory processes appear to be a hallmark of urine from patients with active vasculitis. Candidates for those proteins are alpha-1-antitrypsin (see above), CD99 (a cell surface glycoprotein involved in leukocyte migration and T-cell adhesion both processes that play a role in the pathogenesis of ANCA associated vasculitis) and also hemoglobin fragments only found in this patient group pointing to the specific proteases, active in ANCA associated vasculitis.

In the light of other recent datasets on urinary protein fragments (12, 14, 20, 27), it appears that the combination of peptides that alone are not specific for a particular disease with proteolytic activity results in the generation of a specific panel of degradation products that can be utilized to identify the disease with high specificity, and also assess disease progression/therapeutic benefits. Similar findings, the presence of specific fragments of albumin and alpha-1-antitrypsin in chronic renal disease, have recently been reported by Candiano et al (46).

A major medical need in patients with ANCA associated vasculitis and renal involvement is assessment of the activity of the disease. Here, urinary proteome analysis may be of significant value. If the disease is in remission, immunosuppressive treatment may be reduced. This is important for the patients to minimize the risk of side effects. Microhematuria may be used as a parameter for activity, but its value is limited as it may persist in remission. Repeated biopsies are generally not indicated due to the associated risk
Urinary biomarkers in Vasculitis

(mostly of severe bleeding). Moreover, as the disease is focal, active lesions might be missed. ANCA has been used as a marker of active disease but the results were disappointing (47). Circulating endothelial cells will need time to decrease (48), probably as repair mechanisms lead to release of cells during the healing process. CRP might be increased as a result of an ongoing infection or severe arteriosclerosis (49). Endothelial microparticles may represent a promising tool, but there are not enough data to draw a final conclusion (50, 51). The data presented here suggest that urinary proteomic biomarkers may be an excellent tool to overcome the above mentioned shortcomings. They indicate that urinary proteome analysis does not only enable non-invasive diagnosis and differential diagnosis of AAV, but also allows non-invasive monitoring of disease activity in the kidney.
AUTHOR CONTRIBUTIONS

Dr. Haubitz and Dr. Mischak had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study design. Haubitz, Haller, Mischak, Dakna,

Acquisition of data: Haubitz, Good, Woywodt, Rupprecht, Theodorescu, Coon

Analysis and interpretation of data. Haubitz, Mischak

Manuscript preparation. Haubitz, Mischak, Good

Statistical analysis. Dakna.

Acknowledgements

MH and AW were supported in part by a grant of the DFG. HM was supported in part by EUROTRANS-BIO grant ETB-2006-016 and EU Funding through InGenious HyperCare (LSHM-C7-2006-037093) and PREDICTIONS (1272568). DMG was supported through a pre-doctoral fellowship from the National Institutes of Health; the Biotechnology Training Program, NIH 5T32GM08349.

Statement of Competing Financial Interests

HM is the founder and co-owner of Mosaiques Diagnostics, who developed the CE-MS technology. MD is an employee of Mosaiques Diagnostics.
Figure Legends

Figure 1: Usage of samples and flow of information. A Identification and validation of diagnostic biomarkers and biomarker models. 18 cases of AAV were compared to 425 controls (healthy individuals and patients with different chronic renal diseases), which resulted in the definition of 113 potential biomarkers. Of these, 58 could be sequenced, and 18 were employed in two biomarker models, based on linear combination or SVM-driven. All 113 potential biomarkers and the two biomarker models were evaluated in a test set of 40 blinded samples that consisted of 10 samples from patients with AAV and 30 controls. B Identification and validation of biomarkers and biomarker models for therapy assessment. CE-MS data from 18 urine samples from patients with active AAV were compared to data from urine samples from 19 patients with AAV in complete remission. In parallel, the data from the 18 samples from active AAV were also compared to data from urine samples from 200 healthy controls. 166 and 226 biomarkers that changed significantly with active AAV (p-value < 0.05 upon adjustment for multiple testing) could be defined in the two comparisons, 122 of these were significant in both analyses. Of these 122 potential biomarkers associated with disease activity, 47 could be sequenced. Employing linear combination or SVM, respectively, biomarker models indicative of disease activity were established based upon these 47 urinary peptides. These two biomarker models were subsequently tested in a set consisting of samples longitudinally collected (at month 0, 1, 3, and 6) from 10 patients where therapy of AAV was initiated.

Figure 2: Compiled protein patterns of the CE-MS analysis of urine samples from patients and controls examined in this study. Shown are compiled patterns consisting of all samples from patients with active vasculitis and each of the six control groups. (AAV = ANCA associated vasculitis; NC = apparently healthy normal control, MNGN = membranous
glomerulonephritis, DN = diabetic nephropathy, FSGS = focal segmental glomerular sclerosis, IgAN = IgA nephropathy, MCD = minimal change disease). The molecular mass on a logarithmic scale (0.8 – 25 kDa, indicated on the left) is plotted against normalized migration time (18-45 min, indicated on the bottom). Signal intensity is encoded by peak height and color. While the difference in peptides present between apparently healthy controls and the six different chronic renal diseases is evident, distinct biomarkers that can distinguish between AAV and the other chronic renal diseases are not easily identified.

**Figure 3:** Distribution of only the 113 peptides that reveal significant differences between AAV and controls (p-value < 0.05 after adjustment for multiple testing), shown in the patients with active vasculitis and patients with other chronic renal disease or healthy controls. The molecular mass on a logarithmic scale (0.8 – 25 kDa, indicated on the left) is plotted against normalized migration time (18-45 min, indicated on the bottom). Signal intensity is encoded by peak height and color. All statistically significant biomarkers from Supplementary Table 1 are shown. (AAV = ANCA associated vasculitis; NC = normal control, MNGN = membranous glomerulonephritis, DN = diabetic nephropathy, FSGS = focal segmental glomerular sclerosis, IgAN = IgA nephropathy, MCD = minimal change disease). When examining only these 113 potential biomarkers, differences between AAV and the other chronic renal diseases become evident.

**Figure 4:** Performance of the proteomic biomarker models based on 18 urinary peptides specific for vasculitis. A: ROC analysis of the classification results obtained upon complete cross-validation in the training set consisting of 18 cases and 425 controls. The numerical values of the classification obtained employing either linear combination of the 18 biomarkers (left panel), or the SVM-driven model based on the same 18 biomarkers (right panel) was examined. The 95% confidence interval is indicated by the dashed line. B: performance of the
biomarker models in the test set of the 40 blinded samples. Upon unblinding, 10 were found to be AAV patients, 29 harbored other chronic renal diseases, one sample was from a healthy control (see also Table 3).

**Figure 5:** Scoring AAV activity based on proteome analysis during therapy in a longitudinal assessment. Shown is the scoring of the samples obtained at the indicated period of time after initiation of treatment in the models based on SVM (A), or linear combination of the biomarkers (B). The allocation of the 10 patients (indicated by their internal ID-tag) to the different colors and symbols used in the graph is shown on the right.

**Supplementary Figure 1:** Correlation between erythrocyturia (given as erythrocytes per high power field) and the scoring (as box and whisker plots) of individual samples in the SVM model and the LIN model.
Table 1: Characteristics of patients with active ANCA associated vasculitis (AAV) and patients in long-term remission (GAAV = granulomatous AAV; MPA = microscopic polyangiitis).

Table 2: 18 biomarkers utilized for the discriminatory models between patients with vasculitis and controls (apparently healthy or patients with other chronic renal diseases). Shown are the protein/peptide identification number in the dataset (ID), mass (in Da) and normalized migration time (CE-T; in min), the unadjusted p-values, p-values [adjusted using Bonferroni (Bonferroni) and Benjamini-Hochberg (BH)], the AUC value in the ROC analysis (AUC), and the frequency and mean (median) amplitude based on datasets where the biomarker could be detected (omitting all 0 values) in the two groups of the training set, AAV, and control (Contr). Further, sequence, original protein, and the position of the first and last amino acid of the peptide in the respective protein sequence are given.

Table 3: Classification result of the biomarker models in the validation set consisting of 40 blinded samples using the cutoff values obtained from the results of the complete cross-validation in the training set (0.2 for the SVMmodel and 4.8 for the Linmodel, respectively). Given are the sample ID, the clinical diagnosis, and the classification result using the SVMmodel and the Linmodel. Samples scoring positive for vasculitis in either model are indicated in bold, correct positive scores are labelled in red. (AAV = ANCA associated vasculitis; FSGS = focal segmental glomerular sclerosis, LN = lupus nephritis, IgAN = IgA nephropathy, MPGN = membranoproliferative glomerulonephritis, MNGN = membranous glomerulonephritis, MCD = minimal change disease, GS = glomerular sclerosis, NC = normal control).
**Supplementary Table 1** Panel of statistically significant biomarkers when discriminating between patients with vasculitis and controls (apparently healthy or patients with other renal diseases). Shown are the protein/peptide identification number in the dataset (ID), mass (in Da) and normalized migration time (CE-T, in min), the unadjusted p-values (unadj p-value), p-values after adjustment according to Bonferroni (Bonferroni), p-values after adjustment according to Benjamini-Hochberg (BH) and the AUC value in the ROC analysis (AUC); the frequency in the two groups of the training set, vasculitis (Freq AAV) and control (Freq contr) and mean (median) amplitude based on datasets where the biomarker could be detected, (omitting all 0 values). Where available, sequence and original protein are given, as well as the position of the first and last amino acid of the peptide in the respective protein sequence.

**Supplementary Table 2** Panel of biomarkers that show statistically significant differential distribution in both the comparison of patients with active AAV vs. patients with AAV and no apparent sign of activity, and patients with active AAV vs. apparently healthy controls. Shown are the protein/peptide identification number in the dataset (ID), mass (in Da) and normalized migration time (CE-T, in min), the unadjusted p-values (unadj p-value), p-values after adjustment according to Bonferroni (Bonferroni), p-values after adjustment according to Benjamini-Hochberg (BH) and the AUC value in the ROC analysis (AUC); the frequency and mean (median) amplitude based on datasets where the biomarker could be detected, (omitting all 0 values) in the two groups, active AAV (act_AAV) and inactive AAV (inact_AAV). Where available, sequence and original protein are shown, as well as the position of the first and last amino acid of the peptide in the respective protein sequence.

**Supplementary Table 3:** Pivot Table consisting of 5 worksheets which includes all CE-MS data for every sample in the study. The worksheet Protein assignment shows the mass (Mass, in Da) and migration time (CE-T, in min) of peptides assigned to a certain prot_ID, which is
subsequently utilized as unique identifier in the database. Sample assignment indicates the
unique patient ID (IDpatient), respective group (Study) and assigned sample evaluation ID
(eval_ID). Worksheets T1, T2 and T3 list the amplitudes of each polypeptide in the individual
samples. The table is accessible at http://www.mosaiques-
diagnostics.com/Supplementary_Table3.exe, the password to extract the data is MosaR08.

Supplementary Document: CE-MS analytical platform validation
Reference List


Figure 4

Figure 5
Table 1

<table>
<thead>
<tr>
<th></th>
<th>Active AAV Patients (n=18)</th>
<th>AAV Patients in Remission (n=19)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age [years]</td>
<td>64 (range 19-74)</td>
<td>67 (range 43 – 83)</td>
</tr>
<tr>
<td>Gender</td>
<td>9 female, 9 male</td>
<td>13 female, 6 male</td>
</tr>
<tr>
<td>Disease</td>
<td>9 GAAV, 9 MPA</td>
<td>15 GAAV, 4 MPA</td>
</tr>
<tr>
<td>Proteinuria [g/24h]</td>
<td>0.8 (range 0.1 – 2.1)</td>
<td>0.1 (range 0.1 – 0.6)</td>
</tr>
<tr>
<td>ID</td>
<td>Mass</td>
<td>CE-T</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>------------</td>
</tr>
<tr>
<td>15800</td>
<td>1068.51</td>
<td>21.75</td>
</tr>
<tr>
<td>25219</td>
<td>1213.61</td>
<td>27.2</td>
</tr>
<tr>
<td>30174</td>
<td>1292.59</td>
<td>28.28</td>
</tr>
<tr>
<td>35339</td>
<td>1378.61</td>
<td>28.82</td>
</tr>
<tr>
<td>38879</td>
<td>1439.66</td>
<td>29.82</td>
</tr>
<tr>
<td>40075</td>
<td>1448.78</td>
<td>20.27</td>
</tr>
<tr>
<td>51120</td>
<td>1629.85</td>
<td>24.89</td>
</tr>
<tr>
<td>56033</td>
<td>1705.89</td>
<td>20.73</td>
</tr>
<tr>
<td>60320</td>
<td>1797.93</td>
<td>23.9</td>
</tr>
<tr>
<td>62557</td>
<td>1847.95</td>
<td>21.2</td>
</tr>
<tr>
<td>63772</td>
<td>1873.96</td>
<td>21.43</td>
</tr>
<tr>
<td>66483</td>
<td>1923.97</td>
<td>21.6</td>
</tr>
<tr>
<td>70633</td>
<td>2013.89</td>
<td>31.76</td>
</tr>
<tr>
<td>89325</td>
<td>2356.15</td>
<td>19.52</td>
</tr>
<tr>
<td>90840</td>
<td>2389.24</td>
<td>22.4</td>
</tr>
<tr>
<td>94543</td>
<td>2478.25</td>
<td>23.03</td>
</tr>
<tr>
<td>98089</td>
<td>2559.18</td>
<td>19.41</td>
</tr>
<tr>
<td>118597</td>
<td>3021.35</td>
<td>23.42</td>
</tr>
</tbody>
</table>
Table 3

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Diagnosis</th>
<th>SVMmodel</th>
<th>Linmodel</th>
</tr>
</thead>
<tbody>
<tr>
<td>21310</td>
<td>NC</td>
<td>-2.168</td>
<td>-20.73</td>
</tr>
<tr>
<td>21092</td>
<td>MCD</td>
<td>-2.147</td>
<td>-18.56</td>
</tr>
<tr>
<td>20915</td>
<td>FSGS</td>
<td>-2.015</td>
<td>-17.19</td>
</tr>
<tr>
<td>21144</td>
<td>GS</td>
<td>-1.938</td>
<td>-19.28</td>
</tr>
<tr>
<td>21309</td>
<td>LN</td>
<td>-1.834</td>
<td>-18.78</td>
</tr>
<tr>
<td>21139</td>
<td>LN</td>
<td>-1.798</td>
<td>-20.43</td>
</tr>
<tr>
<td>21312</td>
<td>MNGN</td>
<td>-1.734</td>
<td>-9.24</td>
</tr>
<tr>
<td>20918</td>
<td>FSGS</td>
<td>-1.682</td>
<td>-16.44</td>
</tr>
<tr>
<td>20970</td>
<td>FSGS</td>
<td>-1.676</td>
<td>-10.20</td>
</tr>
<tr>
<td>21068</td>
<td>MNGN</td>
<td>-1.599</td>
<td>-11.58</td>
</tr>
<tr>
<td>21224</td>
<td>AAV</td>
<td>-1.528</td>
<td>-12.20</td>
</tr>
<tr>
<td>21279</td>
<td>FSGS</td>
<td>-1.528</td>
<td>-10.39</td>
</tr>
<tr>
<td>20972</td>
<td>MPGN</td>
<td>-1.493</td>
<td>-8.47</td>
</tr>
<tr>
<td>20980</td>
<td>MNGN</td>
<td>-1.471</td>
<td>-12.53</td>
</tr>
<tr>
<td>21070</td>
<td>MNGN</td>
<td>-1.45</td>
<td>-10.46</td>
</tr>
<tr>
<td>21313</td>
<td>MPGN</td>
<td>-1.421</td>
<td>-5.75</td>
</tr>
<tr>
<td>20917</td>
<td>FSGS</td>
<td>-1.366</td>
<td>-15.55</td>
</tr>
<tr>
<td>21287</td>
<td>MNGN</td>
<td>-1.282</td>
<td>-16.22</td>
</tr>
<tr>
<td>21278</td>
<td>MNGN</td>
<td>-1.27</td>
<td>-10.99</td>
</tr>
<tr>
<td>21482</td>
<td>FSGS</td>
<td>-1.203</td>
<td>-9.84</td>
</tr>
<tr>
<td>21014</td>
<td>MNGN</td>
<td>-1.195</td>
<td>-7.61</td>
</tr>
<tr>
<td>20903</td>
<td>MCD</td>
<td>-0.818</td>
<td>-12.42</td>
</tr>
<tr>
<td>21157</td>
<td>MNGN</td>
<td>-0.296</td>
<td>-0.55</td>
</tr>
<tr>
<td>21102</td>
<td>IgAN</td>
<td>-0.292</td>
<td>-2.60</td>
</tr>
<tr>
<td>20971</td>
<td>MNGN</td>
<td>0.009</td>
<td>2.79</td>
</tr>
<tr>
<td>20976</td>
<td>LN</td>
<td>0.131</td>
<td>1.41</td>
</tr>
<tr>
<td>21217</td>
<td>GS</td>
<td>0.187</td>
<td>1.54</td>
</tr>
<tr>
<td>9619</td>
<td>AAV</td>
<td>0.259</td>
<td>11.14</td>
</tr>
<tr>
<td>21197</td>
<td>AAV</td>
<td>0.274</td>
<td>9.22</td>
</tr>
<tr>
<td>20933</td>
<td>IgAN</td>
<td>0.359</td>
<td>4.52</td>
</tr>
<tr>
<td>21311</td>
<td>AAV</td>
<td>0.903</td>
<td>8.70</td>
</tr>
<tr>
<td>21277</td>
<td>AAV</td>
<td>1.029</td>
<td>8.51</td>
</tr>
<tr>
<td>21065</td>
<td>IgAN</td>
<td>1.316</td>
<td>16.61</td>
</tr>
<tr>
<td>20932</td>
<td>LN</td>
<td>1.427</td>
<td>14.41</td>
</tr>
<tr>
<td>21195</td>
<td>IgAN</td>
<td>1.606</td>
<td>21.65</td>
</tr>
<tr>
<td>21192</td>
<td>AAV</td>
<td>2.153</td>
<td>22.29</td>
</tr>
<tr>
<td>21199</td>
<td>AAV</td>
<td>2.211</td>
<td>21.10</td>
</tr>
<tr>
<td>15701</td>
<td>AAV</td>
<td>2.225</td>
<td>21.81</td>
</tr>
<tr>
<td>21488</td>
<td>AAV</td>
<td>2.388</td>
<td>24.11</td>
</tr>
<tr>
<td>21143</td>
<td>AAV</td>
<td>2.58</td>
<td>29.97</td>
</tr>
</tbody>
</table>