Integrated Fourier Transform Infrared Imaging and Proteomics for Identification of a Candidate Histochemical Biomarker in Bladder Cancer

Kathrin E. Witzke,* Frederik Großerueschkamp, Hendrik Jütte, Melanie Horn, Florian Roghmann, Nicolas von Landenberg, Thilo Bracht, Angela Kallenbach-Thieltges, Heiko Käfferlein, Thomas Brüning, Karin Schork, Martin Eisenacher, Katrin Marcus,* Joachim Noldus, Andrea Tannapfel, Barbara Sitek,* and Klaus Gerwert

From the Medizinisches Proteom-Center,* the Department of Biophysics, the Institute of Pathology, the Department of Urology, Marien Hospital Herne, and the Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Ruhr University Bochum, Bochum, Germany

Accepted for publication November 26, 2018.
Address correspondence to Klaus Gerwert, Dr.rer.nat., Department of Biophysics; or Barbara Sitek, Dr.rer.nat., Medizinisches Proteom-Center, Ruhr University Bochum, Universitätstrasse 150, 44801 Bochum, Germany. E-mail: gerwert@bph.ruhr-uni-bochum.de or barbara.sitek@rub.de.

Bladder cancer is the second most common urogenital malignancy, with approximately 430,000 new cases diagnosed worldwide in 2012.1,2 Approximately 70% to 75% of patients are newly diagnosed with nonmuscle invasive, mostly low-grade bladder cancer, whereas 25% to 30% of de novo diagnoses are high-grade bladder cancer at the stage of infiltration of the muscularis, with metastases observed in 10% of cases.3−5 A significant percentage (50% to 70%) of patients diagnosed with nonmuscle invasive cancer will have multiple recurrences, and 10% to 15% will even progress to invasion.3,6,7 A special case of bladder cancer is carcinoma in situ (CIS), which is a flat, noninvasive tumor with a rate of progression to an invasive state of approximately 54%.8

The prognosis of bladder cancer (BC) is stage and grade dependent; survival rate is higher for lower stages (T2 or less).6,9 Because there are currently no established biomarkers for bladder cancer,10 a diagnosis is established...
solely on the basis of microscopic examination of biopsied lesions. However, visual scoring of tumor grade and stage shows interobserver variability.\textsuperscript{11} The differentiation of reactive urothelial atypia as part of an acute or chronic urocytis and CIS can be particularly difficult, especially after previous treatment with attenuated mycobacteria (\textit{Bacillus Calmette-Guérin}) for deliberate induction of an inflammatory reaction, which is a primary treatment option for CIS aside from early cystectomy.\textsuperscript{2} Correct discrimination of nonmuscle invasive low-grade bladder cancer and CIS is not trivial; in some cases, even differentiating between low- and invasive high-grade tumors can be challenging.\textsuperscript{12} There is therefore a need for a sensitive and specific biomarker that can support pathologic findings in the diagnosis of different subtypes of bladder cancer.

Extensive proteomic studies of urine, serum, and tissue samples have been performed to identify bladder cancer biomarkers,\textsuperscript{13–17} and candidate proteins have been proposed as immunohistochemical markers for diagnosis and prognosis, including CD44, p53, and cytokeratin CK20. p53 Accumulation is correlated with tumor stage and grade and is observed in invasive high-grade tumors,\textsuperscript{18} whereas cytokeratin CK20 is up-regulated early in abnormal urothelial differentiation and is therefore also a biomarker for low-grade tumors.\textsuperscript{19} The use of an immunohistochemistry panel consisting of CD44, cytokeratin CK20, and p53 has been proposed to differentiate between benign and neoplastic entities.\textsuperscript{20} However, in practice, categorizing an immunohistochemical stain as strictly positive or negative can be difficult, a fact that is further aggravated by trying to rely on a three-marker panel. A reliable single marker to differentiate benign from neoplastic flat urothelial lesions is not found yet. Visual inspection of a hematoxylin and eosin (H&E)—stained tissue section remains the gold standard for tumor grading and staging.

Given the limitations of earlier attempts to identify appropriate biomarkers for bladder cancer, a novel approach was used in the present study. To limit interfering factors and use only homogeneous cases of severe urocytis or invasive high-grade carcinoma tissue samples, laser capture microdissection (LCM) was used to obtain a sample from a region of interest (ROI). Mass spectrometry (MS) combined with LCM has been successfully applied to the identification of candidate biomarkers in various cancers,\textsuperscript{21–25} including bladder cancer.\textsuperscript{25} However, to date, it has not been possible to use label-free annotated tissue samples with minimal interobserver and intraobserver variation for liquid chromatography—tandem MS (LC-MS/MS) analyses. Typically, histologic or pathologic examinations are necessary to delineate ROIs in tissue samples, which are then transferred to unstained adjacent sections for LCM. However, this invariably leads to deviations in annotated ROIs because these will only partly overlap with diseased tissue annotated by the pathologist, resulting in more or less heterogeneous samples.

To circumvent this problem, label-free Fourier transform infrared (FTIR)—guided LCM was used to obtain specimens for subsequent proteomic analysis. This approach was recently implemented in a validation study of biomarkers in immunohistochemical panels to differentiate between pleural mesothelioma subtypes.\textsuperscript{26} Tissue classification by FTIR imaging allows label-free annotation of thin sections with high accuracy (>90% sensitivity and specificity) and without interobserver variability as opposed to pathologic annotation.\textsuperscript{27–31} Differential analyses can also be performed by label-free FTIR imaging, as demonstrated for lung, colon, and breast cancer.\textsuperscript{32–34}

As a first step, an IR classifier for bladder tissue was developed to identify ROIs. Ten invasive high-grade bladder cancer samples with no accompanying CIS and seven cystitis tissue samples were classified label free by FTIR imaging. This classification was subsequently transferred to an LCM microscope to collect label-free samples of the same thin tissue section for subsequent LC-MS/MS analysis. This yielded more homogeneous samples for proteomic analysis than would be obtained from adjacent sections. The proteomics analysis revealed three candidate biomarkers that were then verified in the discovery phase by immunohistochemistry (IHC) in three sets of samples, including 20 invasive high-grade and 20 cystitis formalin-fixed, paraffin-embedded (FFPE) tissue specimens as well as 20 low-grade samples. In a third step, the best-performing candidate biomarker from the discovery IHC study, AHNAK2, was further verified in a second independent and much larger cohort of four groups, including FFPE tissue samples of 51 high-grade muscle invasive BC, 67 CIS, 84 low-grade BC, and 108 patients with severe cystitis containing reactive urothelial atypia, low grade. AHNAK2 was found to achieve 97% sensitivity and 69% specificity in differentiating between severe cystitis with reactive urothelial atypia (RUA) and CIS; 80% sensitivity and 86% specificity in differentiating between low and invasive high grades; and 97% sensitivity and 55% specificity in differentiating between low grade and CIS.

\section*{Materials and Methods}

\subsection*{Sample Sets}

Five sample sets were used in this study. Sample set 1 ($N = 11$) was the training set for the FTIR imaging random forest (RF) classifier and included invasive high-grade ($n = 6$), low-grade ($n = 3$), and severe cystitis with RUA ($n = 2$) specimens. Sample set 2 ($N = 103$) was used for verification of the FTIR imaging tissue classification and comprised freshly frozen samples of severe cystitis with RUA ($n = 41$), low-grade carcinoma ($n = 19$), and invasive high-grade carcinoma ($n = 43$). For the discovery phase, sample set 3 was used. Sample set 3 ($N = 17$) was used for combined FTIR imaging and label-free LC-MS/MS and included freshly frozen tissue from invasive high-grade ($n = 10$) and severe cystitis with RUA ($n = 7$) patients who were between 48 and 85 years of age, with a mean age
of 67 years. Of the 17 patients, 4 were female and 13 were male (Supplemental Table S1). Sample set 4 (N = 60) for the first candidate biomarker verification step included FFPE tissues from severe cystitis with RUA and invasive high-grade and also low-grade patients (n = 20 each), all male and between 60 and 80 years of age (mean, 70 years) (Supplemental Table S2). Sample set 5 (N = 310) for the second candidate biomarker verification step consisted of FFPE tissues from severe cystitis with RUA (n = 108), invasive high-grade without CIS (n = 51), low-grade (n = 84), and CIS (n = 67) patients between 23 and 90 years of age, with a mean age of 69 years (Supplemental Table S3).

Because urothelial dysplasia is not fully understood and its significance is unclear, it was not included in this analysis.

Clinical Data

Sample sets 1 to 4 as well as some high-grade bladder cancer patients (n = 20) from sample set 5 were collected at the Department of Urology, Ruhr University Bochum, Marien Hospital Herne (Bochum, Germany), and were managed and distributed by the Institute for Prevention and Occupational Medicine of the German Social Accident Insurance (Ruhr University Bochum). The local ethics committee approved the study (approval number 3674-10). The remaining specimens in sample set 5 were directly obtained at the Department of Urology, Ruhr University Bochum, Marien Hospital Herne, with approval from the local ethics committee (approval number 4047-11). Written, informed consent was obtained from each patient, and the study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki.

Sample Preparation for Discovery

Sample sets 1 to 3 for the discovery phase were collected during surgery, according to a standard operation procedure. The samples were washed with isotonic saline solution, slowly frozen on the surface of liquid nitrogen within 8 minutes, and stored at −80°C until further processing. Frozen sections were obtained using an HM550 cryostat (Thermo Fisher Scientific, Waltham, MA) at −20°C. Sections cut at a thickness of 10 μm were collected on polyethylene terephthalate frame slides. The cryostat was cleaned between samples with 70% ethanol to avoid cross-contamination. All preparatory steps were performed as quickly as possible, according to standard operation procedures, to guarantee high-quality samples with minimal tissue and protein degradation.

FTIR Imaging

FTIR imaging spectroscopy was performed in reflection mode on low-e microscope slides (Kevely Technologies, Chesterland, OH) or in transmission mode on polyethylene terephthalate frame slides using a Cary 620 IR microscope (Agilent Technologies, Santa Clara, CA) equipped with a 128 × 128-pixel liquid nitrogen-cooled mercury cadmium telluride focal plane array detector. The collected wave number region was 3700 to 950 cm⁻¹ at a spectral resolution of 4 cm⁻¹. For sample and background spectra, 128 scans were coadded. The mapped pixel resolution was approximately 5.5 μm for a tissue sampling area of approximately 715 μm² for each focal plane array field. This resulted in oversampling due to a lower optical resolution in the IR region. The results of this and previous studies have shown that oversampling does not affect cancer detection.35 The Cary 670 spectrometer (Agilent Technologies), microscope, and sample chamber were continuously flushed with dry air to avoid spectral contributions from atmospheric water. In addition, a 24/7 liquid nitrogen cooling supply (Norhof, Maarssen, the Netherlands) was included in the system, permitting continuous measurement. FT was performed with Mertz phase correction and Blackman-Harris four-term apodization. Measurements were made in the mosaic mode of Agilent Resolutions Pro software version 5.3.0.1694. Individual mosaic tiles, each measuring 128 × 128 pixels, were stitched together automatically after the measurement. Each raw spectral vector consisted of 1428 data points (resolution, 4 cm⁻¹; zero filling, 4; upper folding limit, 5266 cm⁻¹). The stitching was performed in Matlab 2017b (MathWorks, Natick, MA). The RF was established on low-e slides and extended to polyethylene terephthalate frame slides for FTIR-guided LCM.

Spectral Preprocessing and Analysis

Preprocessing of the raw data sets was performed in Matlab as follows; it is possible to work with an oversampled data set collected at 5.5-μm spatial resolution.35 All images presented herein are from the original data set of unbinned data. The first step after stitching was a quality test based on the integral of the amide I band and the signal/noise ratio (noise, 2100 to 2000 cm⁻¹; signal, 1600 to 1500 cm⁻¹). Pixel spectra from the vicinity of voids or cracks in the tissue and regions without tissue were removed in this step.36 All spectra were then subjected to extended multiplicative scattering correction—based Mie and resonance-Mie scattering correction from 2300 to 950 cm⁻¹ in one iteration.37 Higher iteration numbers were tested, but because of low scattering effects, the classification results were not further improved. Thus, processing time was more important for this study than perfect spectral correction. The spectral range from 1800 to 950 cm⁻¹ was used for unsupervised hierarchical and k-means clustering and the supervised RF algorithm. For unsupervised methods, the second derivative of the smoothed spectra was examined. The smoothing was performed using a nine-point Savitzky-Golay filter.38 The supervised RF was trained on
unsmeared absorption spectra. Tests with the second derivative did not result in higher accuracy for this classification.

Coordinate Transfer

The coordinates of ROIs selected from the FTIR imaging results were transferred to LCM by two-dimensional Helmert transformation on the basis of three reference points. Disparities in pixel size and rotation between the FTIR imaging microscope and LCM were adjusted. For this purpose, the FTIR imaging microscope was calibrated using a 1951 United States Airforce Target glass slide resolution target (USAF 1951 1 ×; Edmund Optics, Barrington, NJ). The resultant pixel size of the focal plane array detector was determined to be 5.65 × 5.75 μm, and the rotation between the visible and IR image microscope was approximately 0.06 degrees.26

Selection of ROIs and LCM

The results of the RF are presented in an index color image representing different tissue types, and the color representing the tissue of interest was defined as the ROI. The thin tissue section was classified according to its spectral fingerprint without further processing. Thus, FTIR-guided LCM enabled collection of label-free sample pools without interobserver and intraobserver variability. Samples of 1 mm² were collected in 20 experimental runs, interobserver and intraobserver variability. Samples of 1 mm² were collected in 20 experimental runs. For each ROI, a 1951 United States Airforce Target glass slide resolution target was imaged in the FTIR imaging microscope and LCM were adjusted. For this purpose, the FTIR imaging microscope was calibrated using a 1951 United States Airforce Target glass slide resolution target (USAF 1951 1 ×; Edmund Optics, Barrington, NJ). The resultant pixel size of the focal plane array detector was determined to be 5.65 × 5.75 μm, and the rotation between the visible and IR image microscope was approximately 0.06 degrees.26

Histologic Analysis

H&E staining was performed on the same slides that were previously used for FTIR spectroscopy analysis. Tissue samples were washed with Milli-Q ultrapure water (Merck Millipore, Darmstadt, Germany), stained with Harris hematoxylin (VWR, Bruchsal, Germany) for 50 seconds, washed with water, counterstained with eosin (Merck Millipore), dehydrated in increasing gradients of alcohol, and mounted with Euparal (Carl Roth GmbH, Karlsruhe, Germany). Stained sections were imaged automatically with a BX41 microscope (Olympus, Tokyo, Japan).

Sample Preparation for LC-MS/MS Analysis

LCM vials were sonicated upside down on ice for 1 minute and then centrifuged for 1 minute in the upright position to transfer the sample from the lid into the vial itself. A 3.7-μL volume of dithiothreitol (20 mmol/L) was added to the sample for reduction (30 minutes, 60°C), followed by 2.2 μL iodoacetic acid (100 mmol/L) for alkylation (30 minutes, room temperature in the dark). Lysed proteins were tryptically digested overnight at 37°C (0.6 μL trypsin; 0.033 μg/μL). For acidification, 1.3 μL of trifluoroacetic acid (10%) was added (30 minutes, 37°C), and samples were transferred to glass vials, dried in a vacuum centrifuge, and dissolved in 0.1% trifluoroacetic acid. A sample amount corresponding to an approximately 0.5-mm² area of laser microdissected tissue was used for one MS measurement.

LC-MS/MS Analysis

LC-MS/MS analysis was performed on an LTQ Orbitrap Elite instrument (Thermo Fisher Scientific) coupled online to an upstream-connected Ultimate 3000 RSLCnano high-performance liquid chromatography system (Dionex, Idstein, Germany), as previously described.39 Briefly, pre-concentration of peptides was performed on a C18 trap column (Acclaim PepMap 100; 100 μm × 2 cm, 5 μm, 100 Å; Thermo Fisher Scientific) within 7 minutes at a flow rate of 30 μL/minute with 0.1% trifluoroacetic acid. The peptides were then transferred to a Nano Viper C18 analytical column (Acclaim PepMap RSLC; 75 μm × 50 cm, 2 μm, 100 Å; Thermo Fisher Scientific). Peptide separation was achieved with a gradient from 5% to 40% solvent B over 98 minutes at 400 nL/minute and 60°C (solvent A, 0.1% formic acid; solvent B, 0.1% formic acid and 84% acetonitrile). MS/MS spectra were acquired in data-dependent mode. Full-scan mass spectra in the Orbitrap analyzer were acquired in profile mode at a resolution of 60,000 at 400 m/z and within a mass range of 350 to 2000 m/z. For MS/MS measurements, the 20 most abundant peptide ions were fragmented by collision-induced dissociation and measured for tandem mass spectra in the linear ion trap.

Protein Identification and Quantification

Proteins were identified using Proteome Discoverer version 1.4 (Thermo Fisher Scientific). The spectra were searched against the UniProtKB/Swiss-Prot database (release 2015_10; 70,075 entries) using Mascot version 2.5 (Matrix Science, London, UK). The taxonomy setting was Homo sapiens, and mass tolerance was 5 ppm and 0.4 Da for precursor and fragment ions, respectively. Dynamic and static modifications were considered for methionine (oxidation) and cysteine (carbamidomethyl), respectively. The false discovery rate (FDR) was calculated with the Proteome Discoverer Percolator function, and identifications with an FDR > 1% were rejected. The software Progenesis QI version 2.0.5387.52102 (Nonlinear Dynamics, Durham, NC) was used for label-free quantification. The obtained .raw files were aligned to a reference run, and a master map of common features was applied to all experimental runs to adjust for differences in retention time. Peak picking in Progenesis was performed in default mode, where a feature is only generated when the corresponding peak is present in most experimental runs. If no peak is detected in a minority of experimental runs, the value for the feature in the
corresponding runs is set to 0. Ion charge states of 2+, 3+, and 4+ with a minimum of three isotope peaks were considered, and raw ion abundances were normalized for automatic correction of technical or experimental variations between runs. Quantified features were identified using the obtained Proteome Discoverer identifications. All nonconflicting peptides were considered for protein quantification. An in-house R script (R version 3.4.0) was applied to the normalized protein abundances obtained from the software for t-test calculations. Ratios of means between groups were determined on the basis of normalized abundances, whereas t-test calculations used arcsinh-transformed data to maintain consistency with Progenesis QI software. t-Test P values were adjusted for FDR control (pFDR) with the method of Benjamini and Hochberg. Proteins were considered as differentially abundant between groups when the absolute ratio of means was ≥1.5 and the pFDR was ≤0.05. Proteomics data have been deposited as a complete submission in the ProteomeXchange Consortium via the PRIDE partner repository (http://www.proteomexchange.org, last accessed August 24, 2018; data set identifier: PXD009203). The .msf files obtained in Proteome Discoverer were converted into the mzIdentML standard format using ProCon PROteomics Conversion tool version 0.9.718.

Verification by IHC

IHC was performed as previously described. Samples were fixed in 4% formaldehyde solution and embedded in paraffin. Sections were prepared at a thickness of 4 μm, deparaffinized in xylene, and then rehydrated. Endogenous peroxidase activity was blocked for 30 minutes with 3% hydrogen peroxide in methanol. After a short rinse in phosphate-buffered saline, sections were preincubated with avidin-biotin (SP-2001; Vector Laboratories, Burlingame, CA) for 15 minutes to reduce non-specific background staining. The sections were incubated with normal goat serum for 20 minutes, followed by primary antibodies against cytokeratin CK5/6 (Dako, Glostrup, Denmark; M7237) at 1:500; asporin (ASPN; Abcam, Cambridge, MA; ab201208) at 1:100; and AHNAK2 (Atlas Antibodies, Bromma, Sweden; HPA004145) at 1:1000. After washing with phosphate-buffered saline, the sections were incubated with biotinylated goat anti-mouse IgG (BioGenex, Hamburg, Germany) for 30 minutes and then with peroxidase-conjugated streptavidin (Dako). The peroxidase reaction was allowed to proceed for 8 minutes, with 0.05% 3,3-diaminobenzidine tetrahydrochloride solution added as substrate. Sections were counterstained with hematoxylin. Negative controls were prepared by replacing the primary antibody with mouse or goat ascites fluid (Sigma-Aldrich, St. Louis, MO). The slides were examined and scored. At least eight randomly selected ×20 high-power fields with a minimum of 4000 cells were scored for degree of expression. The staining in tumor cells was scored on the basis of the immunoreactivity score (IRS). Intensity was measured from 0 (no reaction) to 3 (strong reaction). The relative proportion of positive tumor cells was scored as follows: 0 (no reaction), 1 (<10% expression), 2 (10% to 50% expression), 3 (51% to 80% expression), and 4 (>80% expression). The two individual results were then multiplied to obtain the IRS.

Statistical Analysis

Statistical analysis for verification was performed with R scripts written in house. For IRSs, nonparametric statistics were applied because of the discrete nature of their measurement. The Kruskal-Wallis test was used to identify overall differences in marker expression between groups. To account for multiple testing of different markers, P-value adjustment was applied using the Bonferroni method. When the result was significant, the Wilcoxon rank-sum test was used for pairwise comparisons of groups and the corresponding P values were again adjusted by applying the Bonferroni correction. The power of the markers to distinguish between cystitis and different tumor grades was evaluated by receiver operating characteristic (ROC) curve analyses using the R package pROC, which revealed the balance between sensitivity (true positive rate) and specificity (true negative rate) of a marker in a defined comparison. Values for area under the curve as well as corresponding 95% CIs were determined. Youden’s

<table>
<thead>
<tr>
<th>Histologic diagnosis</th>
<th>FTIR classification, n</th>
<th>Samples/patients, n</th>
<th>Statistical measures, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Negative</td>
<td>Cancerous samples: 62</td>
</tr>
<tr>
<td></td>
<td>Correct positive: 59</td>
<td>False negative: 3</td>
<td>Noncancerous samples: 41</td>
</tr>
<tr>
<td></td>
<td>False positive: 2</td>
<td>Correct negative: 39</td>
<td>Total: 103</td>
</tr>
<tr>
<td></td>
<td>Positive predictive value: 96.7%</td>
<td>Negative predictive value: 92.9%</td>
<td></td>
</tr>
</tbody>
</table>

Data are expressed as n, unless otherwise indicated.

FTIR, Fourier transform infrared.
criterion—which maximized the sum of sensitivity and specificity—was applied to determine the optimal threshold point from an ROC curve.

**Results**

**FTIR-Guided LCM**

FTIR-guided LCM was used for label-free tissue annotation and harvesting. The spatially resolved FTIR spectra of label-free, freshly frozen thin tissue sections were used as a molecular fingerprint of the proteome, genome, transcriptome, lipidome, and metabolome of the underlying tissue.\(^{26}\) Index colors were assigned to the spectra to classify the tissue by digital labeling on the basis of an RF algorithm.\(^{32,35}\) This classifier was trained using 11 samples (sample set 1) and verified using 103 freshly frozen samples (sample set 2) that included 41 cystitis, 19 low-grade carcinoma, and 43 invasive high-grade carcinoma. The samples were compared to H&E-stained images—the clinical gold standard—annotated by a trained pathologist, and showed a specificity of 95%, a sensitivity of 95%, and an accuracy of 95% (Table 1).

The identified cancer ROIs of the first RF were classified as low grade (\(n = 18\)) and invasive high grade (\(n = 41\)) by a second RF with a specificity of 94%, a sensitivity of 93%, and an accuracy of 93% (Supplemental Table S4) compared with the annotation by the pathologist. The label-free classification was then used to microdissect 10 invasive high-grade BC and 7 cystitis homogeneous tissue samples (sample set 3). The thin tissue sections were classified label free by FTIR imaging. Homogeneous ROIs in both specimens (cystitis and invasive high grade) were selected, transferred to the LCM, and then microdissected for subsequent proteome analysis, according to a previously described workflow.\(^{26}\) The label-free tissue classification is exemplarily shown in Figure 1. The first RF classification differentiated between healthy and cancerous tissue. The cancer tissue is shown in red, connective tissue in cyan, inflammation in brown, and urothelial cells in purple.

![Figure 1](image1.png)

**Figure 1** Fourier transform infrared (FTIR) imaging index color images obtained label free before hematoxylin and eosin (H&E) staining of bladder cancer tissue. **A** and **B**: Low-grade and invasive high-grade carcinoma (red, bladder cancer; cyan, connective tissue; brown, inflammation; purple, urothelial cells). **C** and **D**: Corresponding H&E staining after FTIR imaging. The label-free classification matches the cancerous regions. **E** and **F**: Results of the second random forest, which classified the cancer as invasive high-grade (purple) or low-grade (olive) carcinoma. **E**: No invasive high-grade carcinoma is observed. Scale bars = 200 \(\mu\)m.

![Figure 2](image2.png)

**Figure 2** Differential analysis of mass spectrometry data. Thresholds for differentially expressed proteins for \(p_{\text{FDR}}\) were 0.05 and 1.5 for absolute ratios of means (RoM), as indicated by gray dashed lines. Positive and negative ratios of means indicate higher protein abundance in invasive high grade and severe cystitis with reactive urothelial atypia, respectively. All quantified proteins are shown. Proteins selected for further verification are annotated and shown in red; differentially expressed proteins are shown in dark purple; proteins with more than one unique peptide are depicted in light purple; and residual quantified proteins with only one unique peptide are depicted in gray. IHC, immunohistochemistry.
These match the annotation in the corresponding H&E-stained section (Figure 1, C and D). A second RF to classify the grading further analyzed the cancerous class: low- and invasive high-grade carcinoma (Figure 1, E and F). No invasive high-grade carcinoma was observed in the low-grade carcinoma section, whereas a mixture of low- and invasive high-grade carcinoma was detected in high-grade carcinoma. This is in agreement with histopathology and proves the usability of label-free classification by FTIR imaging.

Quantitative Proteomics Analysis of LCM Tissue Samples

The next step was an unbiased quantitative proteomic analysis using samples harvested by FTIR-guided LCM. A total of 3515 proteins were identified using Proteome Discoverer software, and 2481 were quantified by Progenesis QI, of which 1739 had more than one unique peptide. There were 74 proteins that showed differential abundance (adjusted $P \leq 0.05$ and absolute ratio of means $\geq 1.5$) (Figure 2); of these, 62 were higher abundant in cystitis samples and 12 were more abundant in the invasive high-grade group (Figure 2 and Supplemental Table S5).

Selection of Candidate Biomarkers for Verification

The 12 proteins that were significantly higher abundant in invasive high-grade tumor samples were analyzed in greater detail as potential biomarkers (Table 2). Several factors were considered for this selection, including the ratio of means or their known/unknown association in the context of bladder cancer.

The proteins with the highest ratios of means [ie, keratin, type II cytoskeletal (KRT)6A and AHNAK2] that were selected for verification were 58.3- and 29.4-fold higher abundant in high-grade cancer compared with cystitis. In contrast, ASPN had no known association with bladder cancer and was therefore included in the verification. Regulation profiles of the candidates are shown in Figure 3.

Evaluation of Three Candidate Biomarkers by IHC

AHNAK2, KRT6A, and ASPN expression in severe cystitis and low-grade and invasive high-grade bladder cancer specimens were examined by IHC (sample set 4) to assess their potential as biomarkers (Figure 4A). KRT6A was highly expressed in most cystitis and some low-grade cancer cells but was undetectable in invasive high-grade cancer cells. ASPN showed moderate immunoreactivity in many cystitis cells, moderate-to-high expression in all low-grade cancer cells, and high expression in all invasive high-grade cancer cells. AHNAK2 expression was absent in cystitis, with only some background staining in connective tissue, faintly.

![Figure 3](image_url) Regulation profiles of selected candidate proteins. Data were transformed to obtain a normal distribution. Boxes represent the 25th and 75th percentiles; whiskers extend to the most extreme data points; the median is shown as a horizontal line; and the mean value is shown as a square within the box. Green boxes indicate cystitis; red boxes, high-grade bladder cancer. $^*P \leq 0.05$.

<table>
<thead>
<tr>
<th>UniProt ID*</th>
<th>Gene name</th>
<th>Protein name</th>
<th>$P_{FDR}$</th>
<th>Ratio of means</th>
</tr>
</thead>
<tbody>
<tr>
<td>P02538</td>
<td>KRT6A</td>
<td>Keratin, type II cytoskeletal 6A</td>
<td>0.0236</td>
<td>58.3</td>
</tr>
<tr>
<td>Q81VF2</td>
<td>AHNAK2</td>
<td>Protein AHNAK2</td>
<td>0.0236</td>
<td>29.4</td>
</tr>
<tr>
<td>Q02743</td>
<td>HTRA1</td>
<td>Serine protease HTRA1</td>
<td>0.0203</td>
<td>28.1</td>
</tr>
<tr>
<td>P04264</td>
<td>KRT1</td>
<td>Keratin, type II cytoskeletal 1</td>
<td>0.0408</td>
<td>8.9</td>
</tr>
<tr>
<td>Q08X1N</td>
<td>ASPN</td>
<td>Asporin</td>
<td>0.0243</td>
<td>8.4</td>
</tr>
<tr>
<td>Q9NR30</td>
<td>DDX21</td>
<td>Nucleolar RNA helicase 2</td>
<td>0.0203</td>
<td>7.4</td>
</tr>
<tr>
<td>P35527</td>
<td>KRT9</td>
<td>Keratin, type I cytoskeletal 9</td>
<td>0.0216</td>
<td>6.3</td>
</tr>
<tr>
<td>P21810</td>
<td>BGN</td>
<td>Biglycan</td>
<td>0.0473</td>
<td>3.5</td>
</tr>
<tr>
<td>Q96RP9</td>
<td>GFM1</td>
<td>Elongation factor G, mitochondrial</td>
<td>0.0438</td>
<td>2.2</td>
</tr>
<tr>
<td>P41250</td>
<td>GARS</td>
<td>Glycine-tRNA ligase</td>
<td>0.0446</td>
<td>2.0</td>
</tr>
<tr>
<td>095782</td>
<td>AP2A1</td>
<td>AP-2 complex subunit $\alpha$-1</td>
<td>0.0468</td>
<td>1.7</td>
</tr>
<tr>
<td>00341</td>
<td>HDLBP</td>
<td>Vigilin</td>
<td>0.0035</td>
<td>1.7</td>
</tr>
</tbody>
</table>

*Searchable at [https://www.uniprot.org](https://www.uniprot.org).

$p_{FDR}$, FDR-corrected $P$ value.
Figure 4  Immunohistochemical verification of AHNAK2, KRT6A, and ASPN. A: Representative staining for three candidate biomarkers in each group (severe cystitis with reactive urothelial atypia (RUA) and low- and invasive high-grade bladder cancer). B: Boxplots of immune reactive score (IRS) for severe cystitis with RUA and low-grade and invasive high-grade bladder cancer. Boxes represent the 25th and 75th percentiles; whiskers extend to the most extreme data points; and the median is shown as a horizontal line. Green indicates cystitis; orange, low-grade bladder cancer; red, high-grade bladder cancer. C: Receiver operating characteristic curves showing best specificity and sensitivity for each marker in comparisons with significant differences. Area under the curve (AUC) is shown, and ranges in parentheses show the 95% CI. *n = 20 (B, severe cystitis with RUA and low-grade and invasive high-grade bladder cancer). **P ≤ 0.01, ***P ≤ 0.001. Scale bars: 200 μm (A, top row); 50 μm (A, middle and bottom rows).
detected in some low-grade bladder cancer cells, and high in many invasive high-grade bladder cancer cells.

An IRS—which encompasses both the percentage of stained cells and staining intensity—was assigned to each sample for each of the candidate biomarkers (Figure 4B). Only AHNAK2 expression showed significant differences between all three groups (Supplemental Table S6). ROC curves were generated to evaluate the relationship between sensitivity (true positive rate) and specificity (true negative rate) of the markers (Supplemental Table S7). The largest area under the curve for the comparison between cystitis and invasive high-grade bladder cancer was obtained for AHNAK2 (0.87), with a sensitivity of 75% and a specificity of 95% (Figure 4C). By comparison, ASPN has a smaller area under the curve (0.80), with a sensitivity and a specificity of 90% and 75%, respectively; and KRT6A had the smallest area under the curve (0.76), with a sensitivity and a specificity of 75% and 85%, respectively. All other comparisons were only significant for AHNAK2, and this protein was therefore selected for an additional verification step in a larger cohort that included a CIS group.

### Second Verification Step for AHNAK2

AHNAK2 was further verified using FFPE tissue samples from a larger, independent patient cohort comprising four groups, including severe cystitis with RUA (n = 108), invasive high grade with no concomitant CIS (n = 51), CIS (n = 67), and low grade (n = 84). AHNAK2 expression (Figure 5A) was, again, absent in cystitis, with only some background staining in connective tissue, not detectable in most low-grade bladder cancer samples, well detectable in carcinoma in situ, and high in many high-grade bladder cancer cells. AHNAK2 levels differed most significantly between RUA and CIS, and between RUA and invasive high grade (Figure 5B and Supplemental Table S8). The ROC curves between RUA and CIS (Figure 6A), low and invasive high grades (Figure 6B), and low grade and CIS (Figure 6C) revealed that for the differentiation between RUA and CIS, AHNAK2 had a sensitivity of 97% and a specificity of 69% at a cutoff value of 0.5; for differentiating between low and high grades, the sensitivity and specificity were 80% and 86%, respectively, at a cutoff value of 2.5; and for differentiating between low-grade carcinoma and CIS, the sensitivity and specificity were 97% and 55%, respectively, at a cutoff value of 0.5. All other comparisons are shown in Supplemental Table S9 and Supplemental Figure S1.

### Discussion

There is a need for a sensitive and specific biomarker for bladder cancer because of high interobserver variability in established diagnostic approaches, which can lead to a failure to distinguish between histologically similar subtypes. In particular, differentiating between reactive urothelial atypia in the setting of acute or chronic urocystitis or low-grade nonmuscle invasive-type bladder cancer and the noninvasive but high-grade CIS is important for therapeutic decision making. Because these entities may be histologically similar to each other but show immensely different prognoses, they can mean a diagnostic dilemma. Tumor grade is an important prognostic factor for risk stratification, and distinguishing between low- and invasive high-grade bladder cancer is
clinically important. At present, there are no reliable IHC biomarkers in regular clinical use that can aid pathologists in difficult cases, whose only option in cases of uncertainty is to seek a second opinion from their peers.

To identify novel candidate biomarkers, an unbiased approach integrating label-free automated tissue annotation and global label-free LC-MS/MS proteomic analysis was used. The material of choice for the discovery phase was fresh frozen tissue; as such, the analyzed disease stages in this phase were limited to severe urocystitis and invasive high-grade bladder cancer without accompanying CIS. Tissue biopsies obtained by cystectomy in the case of invasive high-grade bladder cancer are suitable for LCM with subsequent proteome analysis, unlike low-grade bladder cancer that is usually only obtained in small amounts by transurethral resection. However, cystectomy samples tend to contain large regions of invasive high-grade carcinoma and cystitis with RUA, which were analyzed in the discovery phase of the present study.

Label-free FTIR imaging identifies not only the cancer region as would an IHC marker panel, but also all other tissue types, including a clear depiction of regions of inflammation. Multiplex FTIR imaging thus replaces not only a single biomarker but also a multi-IHC panel. FTIR-guided LCM allows label-free, observer-independent, and precise harvesting of homogeneous tissue for subsequent LC-MS/MS analysis.

The proteomic analysis identified three potential biomarkers for bladder cancer. Despite the fact of promising biomarker candidates, also for bladder cancer, in which lower abundance correlates to tumor progression, only proteins higher abundant in bladder cancer were considered for further analysis. This is attributed to the fact that the discovery cohort consisted of severe urocystitis and invasive high-grade bladder cancer samples, in which LCM provided homogeneous samples not confounded by tissue other than stated. However, high-grade bladder cancer is often accompanied by areas of urocystitis, which could confound the biomarker results in clinical routine if the marker is more highly abundant in urocystitis samples. In addition, although negative biomarkers also provide valuable biological information, evaluation of positive biomarkers, which stain the diseased state, is much easier in clinical routine. Although KRT6A and AHNAK2 were selected on the basis of their high ratios of means, ASPN was chosen because it has not yet been linked to bladder cancer. The candidate proteins were verified by the well-established method of IHC using FFPE tissue. The advantage herein was the inclusion of additional bladder cancer grades because FFPE tissue is suitable for IHC analysis irrespective of the method of tissue collection. Therefore, low-grade samples and cystitis samples containing reactive urothelial atypia were included in the first verification step in a small cohort of 20 samples per group. Of the three candidate proteins, only AHNAK2 showed significant differences in all comparisons and, more important, in the comparison of low- versus invasive high-grade bladder cancer. Therefore, only AHNAK2 was retained for the second verification step with a larger patient cohort that included CIS samples. Discrepancies between the results of the verification step and proteomics data, as observed for KRT6A, can be attributed to random biological variance given that two different patient cohorts were used for identification and subsequent verification. An alternative explanation is that in bladder cancer, keratin 6A is overexpressed only in malignant cells with activation of growth factor signaling. This could be true of invasive high-grade cases in the discovery study and for cases with high IRSs in the verification of KRT6A by IHC.

In this study, AHNAK2 was identified and verified in two steps as a candidate biomarker for bladder cancer. AHNAK2 is a 600-kDa protein with a high number of

Figure 6  Statistical analysis of immune reactive score (IRS) determined from receiver operating characteristic (ROC) curves. ROC curves show specificity and sensitivity for each evaluated comparison. A: Severe cystitis with reactive urothelial atypia versus carcinoma in situ (CIS). B: Differentiation between low grade and invasive high grades. C: Comparison between low grade and CIS. IRS cutoff values are indicated at the turning point of the curve, with specificity and sensitivity values shown in parentheses. Area under the curve (AUC) is shown, and ranges in parentheses indicate the 95% CI.

Witzke et al

628  ajp.amjpathol.org  The American Journal of Pathology
conserved repeats; although its function is largely unknown, it is thought to associate with calcium channel proteins. AHNAK2 has already been proposed as a potential prognostic biomarker for clear renal cell and pancreatic cancers and is part of a urinary mRNA panel for the diagnosis of bladder cancer and prediction of tumor aggressiveness. In the present study, AHNAK2 showed promise for differentiating between low- and invasive high-grade muscle invasive and noninvasive cancer (CIS), and especially between the oftentimes hard-to-distinguish RUA and CIS. A sensitivity of 97% was achieved for the discrimination between cystitis and CIS and between low grade and CIS, with specificities of 69% and 55%, respectively. Misdiagnosis of CIS can delay treatment of an aggressive malignancy or could lead to unnecessary treatment by intravesical instillation therapy with Bacillus Calmette-Guérin or even early radical cystectomy (the recommended treatment option for Bacillus Calmette-Guérin—refractory CIS), with all of the associated adverse effects. AHNAK2 could be a helpful tool for detecting CIS recurrence or persistence after Bacillus Calmette-Guérin therapy for CIS, in which an inflammatory response is deliberately induced in the bladder.

Although the cutoff value in both cases was low at 0.5, differentiation was usually possible because most cystitis and low-grade cases had an IRS of 0. There were a few outliers in the RUA cohort with IRSs of up to 9, but in at least one of these cases, we had follow-up information that the patient was diagnosed with low-grade bladder cancer 6 months later. This suggests that AHNAK2 could serve as a prognostic biomarker, which should be investigated in future studies. AHNAK2 had a sensitivity of 80% and a specificity of 86% for the differentiation between low- and invasive high-grade bladder cancer, demonstrating that it can be useful for therapeutic decision making by pathologists in borderline cases.

Besides its detection by IHC in tissue sections, AHNAK2 should be considered as a urinary marker because there are no others that are accepted for diagnosis or follow-up in routine practice or in clinical guidelines. Some urinary biomarker tests have been proposed for bladder cancer detection, such as the protein-based BTA STAT or BTA TRAK assays (both from POLYMEDCO, New York, NY) or NMP22BladderChek/NMP22 (Alere, Scarborough, ME), and the cell-based UroVysion (Abbott Molecular, Inc., Des Plaines, IL) or ImmunoCyt/uCyt+ (Scimedx Corp., Denville, NJ) test that allow noninvasive diagnosis of bladder cancers in contrast to invasive cytologic methods. However, all these tests yield results with high variance and are therefore not recommended by clinical guidelines. A study is currently underway to evaluate the diagnostic value of AHNAK2 in cytologic analyses. However, the results demonstrate that AHNAK2 could be a useful candidate biomarker for IHC that can be used as an adjunct to H&E staining in cases in which it is difficult to differentiate between bladder cancer subtypes.

Overall, the FTIR imaging approach described herein achieves high sensitivity and specificity for tissue classification in bladder cancer and allows precise selection of ROIs for subsequent proteomics-based analyses. This novel approach can be extended to other cancer types to provide new biomarker candidates.

Acknowledgments

The authors would like to thank Lidia Janota, Kristin Rosowski, Stephanie Tautges, and Birgit Zülch for their excellent technical assistance.

Supplemental Data

Supplemental material for this article can be found at https://doi.org/10.1016/j.ajpath.2018.11.018.

References


69. Kinders R, Jones T, Root R, Bruce C, Murchison H, Corey M, Ellis WJ, Blumenstein BA, Ishak LM, En...