Phosphoproteins with Stability against All Urinary Phosphatases as Potential Biomarkers

Mindi Zhao, Kehui Liu, Youhe Gao

Urine, by accumulating all kinds of changes, was proposed to be a better source for biomarker discovery. As one of the most common post-translational modifications, phosphorylation plays a vital role in many biological activities. However, the urine phosphoproteome has been largely neglected due to the low abundance of phosphoproteins and the presence of various phosphatases in urine. The low level of background phosphorylation in urine is actually advantageous, as urinary phosphopeptides/proteins that are stable to the phosphatases present in urine have the potential to serve as valuable disease biomarkers. Using a TiO2 enrichment strategy, this study aimed to create a comprehensive proteomic profile of human urinary phosphoproteins and to characterize the changes in the urine phosphoproteome after incubation of urine with cell lysates. In total, 106 urine phosphorylation sites corresponding to 64 proteins, including 80 previously unidentified human urine protein phosphorylation sites, were identified by mass spectrometry. Fifteen phosphopeptides, together averaging 47% of the total phosphopeptides, were found in samples from three individuals. Addition of cellular proteins to urine did not significantly change the phosphorylation level of urine proteins. But there were still a few phosphopeptides from cell lysates survived urinary phosphatases; such phosphopeptides represent potential biomarkers in urine.
Phosphoproteins with Stability against All Urinary Phosphatases as Potential Biomarkers in Urine

Mindi Zhao¹, Kehui Liu²* and Youhe Gao³.¹*

1. National Key Laboratory of Medical Molecular Biology, Department of Pathophysiology, Institute of Basic Medical Sciences Chinese Academy of Medical Sciences, School of Basic Medicine Peking Union Medical College, Beijing, China.
2. State Key Laboratory of Molecular and Developmental Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China.
3. Department of Biochemistry and Molecular Biology, Beijing Normal University, Gene Engineering and Biotechnology Beijing Key Laboratory, Beijing, China

* Corresponding author

Youhe Gao Email: gaoyouhe@bnu.edu.cn
Kehui Liu Email: khliu@genetics.ac.cn
Introduction

Urine, by accumulating all kinds of changes, was proposed to be a better source for biomarker discovery (Gao 2013; Gao 2014a; Gao 2014b; Li et al. 2014). However, due to the masking effects of highly abundant proteins in urine and to ion suppression, mass spectrometry (MS) is substantially limited in its ability to detect low-abundance urine proteins and post-translational modifications (PTMs) without enrichment. Thus, many urine proteomic studies have focused on comparing protein abundance between healthy subjects and patients with specific medical conditions but have not considered PTMs. Only urinary glycoproteins were used in some biomarker studies, as high-abundance albumin is not N-glycosylated and can be subsequently removed (Wang et al. 2006; Wang et al. 2008).

Phosphorylation, one of the most common PTMs, is a major determinant of protein-protein interactions, signaling pathways, and physiological and cellular functions (Cohen 2002). However, the urine phosphoproteome has largely been neglected due to the low abundance of urine phosphoproteins probably due to the presence of various phosphatases in urine and the poor stability of post-translational modifications (Lopez et al. 2012). However, the low level of background phosphorylation in urine is actually advantageous, as urinary phosphopeptides and proteins that are able to survive the phosphatases present in urine can potentially serve as valuable disease biomarkers.

The phosphorylation of proteins present in many body fluids, including CSF, serum and saliva, has been studied. Although more than 2,000 proteins have been identified in urine, only 59 phosphopeptides were identified in urine from healthy human subjects in a proteome study by using phosphorylation as a criterion in database searching (Li et al. 2010). Urine contains more than 20 different phosphatases, and previous studies have shown that the levels of several urinary...
phosphoproteins increase in certain diseases and conditions. As one example, 23 unique phosphopeptides were found to be differentially expressed in pregnant and non-pregnant women (Zheng et al. 2013). The phosphorylated sodium chloride co-transporter in urinary exosomes was considered as a potential marker for aldosteronism (van der Lubbe et al. 2012), whereas profilin 1, a phosphorylated urinary protein, is a candidate biomarker for aggressive bladder cancer (Zoidakis et al. 2012).

In this study, we aimed to profile human urinary phosphoproteins. To determine which phosphoprotein species potentially present in urine can survive urinary phosphatases, we also set to illustrate changes in the urine phosphoproteome after incubation of urine with cell lysate (Figure 1).

**Materials & Methods**

**Ethics statement**

Prior to enrollment in the study, all volunteers were given a verbal explanation of the study and signed an informed consent document. The consent procedure and the research protocol in this study were approved by the Medical Ethics Committee of Peking Union Medical College (#018-2013).

**Urine sample preparation**

Three healthy volunteers were enrolled in the study. All subjects were asked to maintain a comfortable supine position at a room temperature between 20°C and 24°C, and each subject was instructed to collect a 24-h urine sample. Seven hundred milliliters of urine was collected from each study volunteer. Four hundred milliliters of urine from each volunteer was pooled, and the pooled sample was divided into three equal parts. To two of these portions, 1 mg or 100 µg of protein from cell lysate was added, and each sample was then incubated at 37°C for 4 hours. The
remaining 300 mL urine from each volunteer was centrifuged at 3,500 g for 30 min. After
discarding the debris, the supernatant was centrifuged at 12,000 g for 30 min. After the removal
of precipitates, urinary proteins were extracted by acetone precipitation and re-dissolved in 25
mM ammonium bicarbonate.

Cell lysates

Human A498 cells (Cell Resource Center, Chinese Academy of Medical Sciences) were cultured
in DMEM media with 10% fetal bovine serum and 1% penicillin and streptomycin. The cell
pellets (1×10^8 cells) were lysed in 3 mL of lysis buffer containing 25 mM Tris (pH 8) and 8 M
urea. The mixture was subjected to ultrasonication at 60 W using 30 cycles at 50% duty. Cell
debris was removed by centrifugation at 12,000 g for 15 min at 4°C. Protein concentration was
measured with a Bradford assay. Four milligrams of total protein was resuspended in 25 mM
ammonium bicarbonate.

In-solution digestion and phosphopeptide enrichment

The proteins in the sample were reduced in 20 mM DTT at 56°C for 1 hour and were
subsequently alkylated in the dark with 50 mM iodoacetamide at room temperature for 45 min.
Trypsin was added (1:50, w/w), and digestion was conducted overnight at 37°C before quenching
by the addition of formic acid to a final concentration of 0.1%. The digestions were centrifuged
to remove insoluble materials, and the supernatants were desalted by passage over an Oasis HLB
cartridge (Waters, Milford, MA) and then dried by vacuum evaporation.

TiO$_2$ enrichment of the peptide mixtures was performed using the TiO$_2$ Phosphopeptide
Enrichment Kit (Thermo Fisher Scientific, Bremen, Germany). The peptide samples were re-
dissolved in 150 µL of 28% lactic acid and applied to the TiO$_2$ spin column, and phosphopeptide
isolation was performed according to the manufacturer’s instructions. The phosphopeptides were
eluted with 50 µL of 5% pyrrolidine.
LC-MS/MS analysis

Samples were analyzed on an LTQ Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). The peptide mixture was loaded onto a C18 trap column (5 cm, 5 µm, 200 Å, Proxeon, Co.). Captured peptides were directed to a reversed-phase analytical column (25 cm, 5 µm, 120 Å, homemade) with an integral fritted nanospray emitter and eluted using a linear gradient of acetonitrile (3-30% in 80 min) in 0.1% formic acid at a flow rate of 300 nL/min. Spectra were recorded in data-dependent acquisition mode and the dynamic exclusion option of 30 seconds was enabled. MS scan in the Orbitrap system with high resolution was set to r=120,000 (all Orbitrap system resolution values are given at m/z 400), followed by Collision Induced Dissociation fragmentation of the top 15 most abundant precursor ions or MSA (neutral losses of m/z 98, 49, and 32.6) both with target values of 1,000 and 35% normalized collision energy. Each sample was analyzed three times.

Data analysis

The data was analyzed using a pre-release version of Thermo Scientific Proteome Discoverer™ software version 1.3. SEQUEST and Mascot search engines were used for identification against a Uniprot_2013_06 database. Validation based on separate target and decoy searches and subsequent calculation of classical score-based false discovery rates or Percolator (Kall et al. 2007) determined q-values were used for assessing the statistical significance of the identifications. The Percolator algorithm uses a set of features related to the quality of the peptide-spectrum matches (PSMs) to classify between correct and incorrect matches. The mass tolerance was set at 10 ppm for precursors and 1 Da for product ions. Modifications selected were set as fixed carbamidomethylation of cysteines, variable oxidation of methionine, phosphorylation of serine, threonine, and tyrosine residues (S, T and Y) were chosen. Maximum missed cleavage sites were set to be 2. False-positive assignments were removed by the
thresholds of two parameters, Xcorr and ΔCn (obtained using the target-decoy database search method) with the false-discovery rate (FDR) of 1%. Peptides were assigned again using the PhosphoRS algorithm (Shiromizu et al. 2013), which calculated the possibility of the phosphorylation site from the spectra matched to the identified peptides. The score threshold for peptide identification was set at a 1% false discovery rate (FDR) and a 70% PhosphoRS site probability. Peptides <7 amino acids in length were discarded due to their reduced specificity for a single protein. Further filtering criteria for peptides included detection in at least 2 of 3 technical replicates per biological sample.

Results and Discussion

Urine phosphorylation sites identified in this study

The three individual urine samples were analyzed by LC-MS/MS in triplicate; the average technical variation of the analysis was 64%. In total, 106 phosphorylation sites (97 pSer, 6 pThr and 3 pTyr) corresponding to 64 phosphoproteins were identified in the urine proteome. The ratio of pSer:pThr:pTyr was close to the reported ratio in cells (Olsen et al. 2006). Approximately 50% of the high-confidence phosphorylation sites showed an acidophilic SXE motif (Motif-x algorithm (Amanchy et al. 2007)). Of the 64 phosphorylated proteins, 54 (84%) contained one or two unique phosphorylated peptides. Ten proteins yielding more than 3 different phosphopeptides were detected, and one protein (osteopontin) was detected that yielded more than 40 phosphorylated peptides. Some of the high-confidence phosphoproteins (> 3 phosphopeptides) are of great importance for biomarker discovery. For example, expression of human fibulin-2 was reported to be associated with acute kidney injury (Jin et al. 2012), whereas upregulation of human aquaporin in urine has been associated with nephrogenic diabetes insipidus (Kanno et al. 1995).
A total of 15 phosphopeptides derived from 8 phosphoproteins were found to be common to the three individual urine samples; these peptides contributed, on average, 47% of the phosphopeptide abundance in each sample. The spectral counts of each peptide from each individual are listed in Table 1. Among the 8 phosphoproteins identified as common to the three individuals in this study, alpha-2-HS-glycoprotein has been considered a reliable indicator of many glomerular diseases, such as diabetic nephropathy, in previous studies not considering PTMs (Rao et al. 2007). A relationship between osteopontin and renal calculi has also been confirmed (Nene et al. 2013).

Of the 64 phosphoproteins identified in urine, 19 are present in the Urinary Protein Biomarker Database (Shao et al. 2011). Twelve of these are associated with renal diseases such as obstructive nephropathy, IgA nephropathy and renal calculi (Table S1).

The data obtained in this study were compared with two previous phosphoproteome data sets. In one previous urine proteome study, 59 phosphorylation sites were identified without enrichment (Li et al. 2010). Although different enrichment strategies and identification methods were adopted and different samples were included in our study, 40.7% of the sites reported in previous studies (24 phosphorylation sites) were also found in our study. In another urinary exosome study, 19 phosphorylation sites were identified after IMAC enrichment (Gonzales et al. 2009). Although our study analyzed urine rather than exosome phosphopeptides and used a TiO$_2$ enrichment method, 4 phosphorylation sites common to previous exosome studies were identified.

Because urine is considered a filtrate of blood, we compared the phosphorylation sites we identified in this study with phosphorylation sites reported in previous studies of serum phosphoproteins (Carrascal et al. 2010; Jaros et al. 2012; Zhou et al. 2009). Some phosphorylation sites, including those in kininogen-1 (ETTCSKEsNEELTESCETK), AMBP protein (FSVVYAKCDSSPDsAEDVR), SPARC-like protein 1 (ISHELDSAsSEVN), protein Z-
dependent protease inhibitor (VVQAPKEEEEDEQE-AsEEKASEEEK) and albumin (TCVADEsAENC'DK), were identical to sites identified in those studies. However, the overlap between serum and urine phosphorylation sites is low; this may be due to the low abundance of many phosphoproteins in serum and urine, a condition that makes phosphoproteins in blood and urine difficult to identify. The low overlap between the sites found in urine in this study and previously reported serum phosphoproteins may also be because not all phosphoproteins in the blood are passed into the urine. The phosphorylation levels of urinary proteins may reflect pathological renal states and/or abnormal physiological states. Because urine reflects many of the changes that occur in the body, the phosphoproteins identified in this study may have been phosphorylated in their original tissues prior to release into blood and urine and thus may be related to both blood condition and renal function.

Changes in urine phosphopeptides after the addition of cell lysates

Phosphatase inhibitors were not used in the urine or in cell lysates. A total of 792 phosphopeptides were identified in crude cell samples, whereas 74 phosphopeptides were identified in the pooled urine. After incubating urine with 100 µg or 1 mg of cell lysate at 37°C for 4 h, 74 or 76 phosphopeptides were identified, respectively. In addition, one peptide (KGAGDGsDEEVDGKADGAEAKPAE) corresponding to myosin-9 was present in the cell lysate and in the cell lysate-supplemented urine samples but not in the phosphorylation-enriched urine samples (Figure 2). This finding indicates that some phosphopeptides from the cell lysate are likely stable upon exposure to phosphatases in urine and that they can be identified when added to urine. When 1 mg of cell lysate was added to the urine samples, 10 phosphopeptides corresponding to 8 proteins could be identified both in the cell lysate and in the urine-cell lysate mixture but not in the untreated urine sample. The eight proteins (AHNAK, CTTN, FLNC, G3BP1, MAP1B, NDRG1, PEA15, and PGRMC1) identified after adding the larger amount (1
mg) of cell lysate to the urine samples are closely related to cell growth, cellular movement and cancer. Figure 3 shows the peptides and proteins identified in urine alone, urine with added cell lysate and cell lysate alone.

Although many phosphoproteins were present in the cell lysate, the addition of cellular proteins did not significantly change the level of protein phosphorylation in urine, most likely because urine contains more than 20 different phosphatases (Marimuthu et al. 2011). It is worth noting that several phosphorylated peptides were stable to incubation with urine phosphatases. Phosphorylation varies among different cells. In different cells from which proteins are released into urine, some phosphorylation sites will survive the urine phosphatases; these may represent potential biomarkers. The overall phosphatase activity of urine changes due to alterations in the combination of phosphatases present or to the activity of individual phosphatases. Such changes are also reflected in the phosphorylation status of proteins in urine. Increasing the detection sensitivity of phosphorylation may aid in the discovery of additional urine phospho-biomarkers. Phosphoproteins that can survive all urine phosphatases, some of which are identified here, potentially represent valuable urinary biomarkers.

References


Figure Legends

Figure 1. Overview of the workflow. Urine samples from three individuals, urine samples incubated with cell lysates, and cell lysates were separately digested, enriched and analyzed by LC-MS/MS.

Figure 2. Phosphopeptides identified in urine, urine with added cell lysate and cell lysates. A. Urine. B. Urine incubated with 100 µg of cell lysate. C. Urine incubated with 1 mg of cell lysate. D. Cell lysates.

Figure 3. Representative MS/MS spectra of phosphopeptides from selected proteins that were identified only in cell lysate and in urine with added cell lysate. A. Myosin-9. B. Ras GTPase-activating protein-binding protein 1. C. Membrane-associated progesterone receptor component 1. D. Neuroblast differentiation-associated protein AHNAK.

Supporting Information Legends

Table S1. Phosphorylation sites identified in this study and their clinical uses.

Figure S1. Representative MS/MS spectra of peptides from selected proteins that were identified based on single-peptide evidence.
Overview of the workflow.

1

Urine S1 → Urine S2 → Urine S3 → Pooled Urine S1+S2+S3 → Cell lysate

Urine → Urine+100ug cell lysates → Urine+1mg cell lysates → Incubation for 4 hour at 37°C → Urine proteins extraction, tryptic digestion, TiO2 enrichment → LC-MS/MS identification
Phosphopeptides identified in urine, urine with added cell lysate and cell lysates.

Representative MS/MS spectra of phosphopeptides from selected proteins that were identified only in cell lysate and in urine with added cell lysate.

Table 1 (on next page)

Fifteen phosphopeptides corresponding to 8 phosphoproteins were common to urine samples from three individuals.
<table>
<thead>
<tr>
<th>Accession</th>
<th>Protein name</th>
<th>Peptides</th>
<th>Modifications</th>
<th>S1 spectra</th>
<th>S2 spectra</th>
<th>S3 spectra</th>
</tr>
</thead>
<tbody>
<tr>
<td>P04792</td>
<td>Heat shock protein beta-1</td>
<td>QLSGVSEIR</td>
<td>S3(Phospho)</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>P01042</td>
<td>Kininogen-1</td>
<td>ETTCsKEsNEELTESCETK</td>
<td>S8(Phospho)</td>
<td>12</td>
<td>22</td>
<td>17</td>
</tr>
<tr>
<td>P02765</td>
<td>Alpha-2-HS-glycoprotein</td>
<td>CDSSPDsAEDVR</td>
<td>S7(Phospho)</td>
<td>31</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>P10451</td>
<td>Osteopontin</td>
<td>ANDEsNEHSDVIDsQELSK</td>
<td>S5(Phospho); S14(Phospho)</td>
<td>6</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ISHELDsAsSEVN</td>
<td>S7(Phospho); S9(Phospho)</td>
<td>3</td>
<td>29</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KANDESNEHSDVIDsQELSKVSR</td>
<td>S15(Phospho)</td>
<td>37</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AIPVAQDLNAPSDWDsR</td>
<td>S16(Phospho)</td>
<td>23</td>
<td>24</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QNLLAPQNAVsSEETNDFKQETLPSK</td>
<td>S11(Phospho)</td>
<td>18</td>
<td>5</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>QNLLAPQNAVsSEETNDFK</td>
<td>S12(Phospho)</td>
<td>48</td>
<td>17</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GKDSYETSQLDQsAETHSHK</td>
<td>S14(Phospho)</td>
<td>10</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KANDESNEHSDVIDsQELSK</td>
<td>S6(Phospho); S15(Phospho)</td>
<td>15</td>
<td>30</td>
<td>13</td>
</tr>
<tr>
<td>Q8NBJ4</td>
<td>Golgi membrane protein 1</td>
<td>GFGGAGELGQTPQVQAALSVsQENPE</td>
<td>S21(Phospho)</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Q969E1</td>
<td>Liver-expressed antimicrobial peptide 2</td>
<td>RCsLSVAQE</td>
<td>S3(Phospho)</td>
<td>5</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>Q9NZH0</td>
<td>G-protein coupled receptor family C group 5</td>
<td>TAGFPNGsLGK</td>
<td>S8(Phospho)</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Q9ULI3</td>
<td>Protein HEG homolog 1</td>
<td>SYSESSSTSsSESLNSSAPR</td>
<td>S10(Phospho)</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>