August 3, 2012
Revised: August 6, 2012

Quantitative Phosphoproteomics in Nuclei of Vasopressin-Sensitive
Renal Collecting Duct Cells

Steven J. Bolger*, Patricia A. Gonzales Hurtado*, Jason D. Hoffert, Fahad Saeed, Trairak Pisitkun, and Mark A. Knepper

*Co-First Authors

Epithelial Systems Biology Laboratory,
NHLBI, National Institutes of Health,
Bethesda, MD 20892-1603, USA

Running title: Quantitative Phosphoproteomics in Collecting Duct

*Correspondence to Mark A. Knepper, MD PhD, National Institutes of Health, Bldg. 10, Room 6N260,
10 CENTER DR, MSC-1603, Bethesda, MD 20892-1603,
Phone: (301) 496-3064, FAX (301) 402-1443, e-mail: knep@helix.nih.gov

Copyright © 2012 by the American Physiological Society.
Abstract

Vasopressin regulates transport across the collecting duct epithelium in part via effects on gene transcription. Transcriptional regulation occurs partially via changes in phosphorylation of transcription factors, transcriptional coactivators, and protein kinases in the nucleus. To test whether vasopressin alters the nuclear phosphoproteome of vasopressin-sensitive cultured mouse mpkCCD cells, we used stable isotope labeling and mass spectrometry to quantify thousands of phosphorylation sites in nuclear extracts and nuclear pellet fractions. Measurements were made in the presence and absence of the vasopressin analog dDAVP. Of the 1,251 sites quantified, 39 changed significantly in response to dDAVP. Network analysis of the regulated proteins revealed two major clusters (“cell-cell adhesion” and “transcriptional regulation”) that were connected to known elements of the vasopressin signaling pathway. The hub proteins for these two clusters were the transcriptional coactivator β-catenin and the transcription factor c-Jun. Phosphorylation of β-catenin at Ser552 was increased by dDAVP (log₂[dDAVP/vehicle] = 1.79), and phosphorylation of c-Jun at Ser73 was decreased (log₂[dDAVP/vehicle] = -0.53). The β-catenin site is known to be targeted by either protein kinase A or Akt, both of which are activated in response to vasopressin. The c-Jun site is a canonical target for the MAP kinase Jnk2, which is downregulated in response to vasopressin in the collecting duct. The data support the idea that vasopressin-mediated control of transcription in collecting duct cells involves selective changes in the nuclear phosphoproteome. All data are available to users at http://helixweb.nih.gov/ESBL/Database/mNPPD/.
Introduction

Vasopressin is a nine-amino acid peptide hormone that mediates the regulation of water and solute transport in the kidney. In the renal collecting duct, vasopressin regulates key transporter proteins including the water channel aquaporin-2 (AQP2) (14). Vasopressin has been shown to regulate water permeability in the collecting duct in part by triggering the redistribution of AQP2 channels from intracellular locations to the apical plasma membrane through short-term trafficking effects (34). Additionally, long-term vasopressin stimulation has been shown to increase the transcription of the Aqp2 \( (\text{footnote 1}) \) gene (29, 31, 64) resulting in increases in Aqp2 mRNA (17, 21, 63) and protein (35) in the kidney. Most reviews indicate a role for the transcription factor Creb1 in this process, presumably via protein kinase A-mediated phosphorylation (24) (8) (36). However, this general model has recently been called into question (30). Vasopressin also increases the mRNA (33) and protein (16) abundances of the \( \beta \) and \( \gamma \) subunits of the epithelial sodium channel (ENaC). In general, regulation of transcription for a particular gene occurs via transcription factors that bind to enhancer or repressor cis-elements, usually upstream from the coding region of the gene (2). Transcriptional regulation can also involve coregulators and other regulatory proteins as well as chromatin modification. These regulatory events are typically mediated by differential posttranslational modifications and/or by nuclear translocation of regulatory proteins.

The complex process by which the translocation and/or modification of transcriptional regulators influence transcription in response to vasopressin is poorly understood. The low abundance of many transcriptional regulators relative to other categories of proteins has made them difficult to detect by protein mass spectrometry carried out at a whole-cell level (40). This problem can be addressed by enriching the transcriptional regulators through isolation of nuclei.
from cells prior to proteomic analysis. We were previously able to identify 154 proteins
classified as “transcription factor” proteins in native rat inner medullary collecting duct cells
using tandem mass spectrometry coupled to high performance liquid chromatography (LC-
MS/MS) by enriching the nuclear proteins prior to analysis (56). Similarly, 379 proteins with the
Gene Ontology Biological Process term “Regulation of transcription” were identified in the
nuclear extract fraction of cultured mpkCCD clone 11 cells (49). (This clonal line was developed
to maximize AQP2 response to the vasopressin analog 1-desamino-8-D-arginine vasopressin
dDAVP) both in terms of membrane trafficking and transcription (66.) In the same study, we
identified 65 proteins that exhibit significant changes in nuclear abundance in response to
dDAVP in mpkCCD clone 11 cells including transcription factors (JunB, Elf3, Gatad2b,
Hmbox1), transcriptional coregulators (Ctnnb1 and Crebbp), subunits of the mediator complex,
an E3 ubiquitin ligase (Nedd4), a nuclear transport regulator (RanGap1), and several proteins
associated with tight junctions and adherens junctions (49).

Vasopressin has also been shown to regulate the phosphoproteome of collecting duct
cells at the whole-cell level (5), (26), (27), (45). However, the effect of vasopressin on
phosphorylation specifically in the nucleus is unknown. The activities of many transcriptional
regulators are controlled by phosphorylation. A well-known example is the bZIP transcription
factor Creb (gene symbol: Creb1), which is regulated through phosphorylation at Ser133 by
ribosomal S6 kinase (20), calmodulin-dependent kinase 2 (53), or protein kinase A (23). To
identify proteins whose phosphorylation is altered by vasopressin in the nucleus, we employed
Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC) and LC-MS/MS to quantify
changes in protein phosphorylation in biochemically-isolated nuclei of mpkCCD clone 11 cells.
Methods

Cell Culture and SILAC Labeling

The mouse renal cortical collecting duct clonal cell line (mpkCCD, clone 11) (66) was grown in SILAC culture medium containing the amino acids arginine and lysine labeled with either “heavy” ($^{13}$C $^{15}$N arginine, $^{13}$C lysine) or “light” ($^{12}$C $^{14}$N arginine, $^{12}$C lysine) stable isotopes for 16 days (37). In previous studies, it was confirmed that 16 days is a sufficient period to result in uniform stable isotope labeling in these cells (45). These immortalized cells were originally grown from transgenic mice expressing the transforming large tumor (T) and small tumor (t) antigens (7). All SILAC reagents and media were obtained from Invitrogen (Carlsbad, CA). The cells were grown on membrane supports (100 mm, 0.4 µm pore size, Transwell, Corning), and transepithelial resistance was measured daily to identify confluence. Cells were pretreated and re-treated with dDAVP using a protocol similar to that of Schenk at al. (49) with the following changes. Once confluent, all cells were cultured for three additional days (with 1 nM dDAVP) to ensure the same starting proteomic profile. After this initial equilibration, the dDAVP was withdrawn for all cells for 6 hours. Following withdrawal, the cells were exposed to either 0.1 nM dDAVP or vehicle for 30 minutes prior to harvesting. Our previous studies have demonstrated that most phosphorylation changes in IMCD cells are seen within 15 minutes of dDAVP exposure (26). In different replicates, vasopressin was added alternately to the heavy-isotope labeled cells or the light-isotope labeled cells.

Nuclear Isolation

The cells were rapidly harvested in ice-cold PBS with protease and phosphatase inhibitors (HaltTM Protease Inhibitor Cocktail, Pierce) and were processed on ice to block to
block post-harvest phosphatase and kinase activities. The cells were centrifuged at 4°C at 500 g for 5 minutes. The supernatant was removed, and the pellet was weighed in a tared tube. Heavy and light samples (dDAVP and vehicle) were combined 1:1 into a single sample. Thus, beyond this point both dDAVP samples and vehicle-treated samples were, by definition, exposed to the identical processing steps, accounting for the high degree of precision of the SILAC method (37). The samples were fractionated into the nuclear extract (NE), nuclear pellet (NP), and cytoplasmic extract (CE) using the commercially available the NE-PER detergent-based Nuclear Protein Extraction Kit (Pierce). The fractions were separated following the instructions provided with the kit with minor modifications. After adding the Cytoplasmic Extraction Reagent (CER)-I reagent and vortexing, the sample was pipetted up and down 200 times to further lyse the cells. After adding the CER II reagent and vortexing, the sample was pipetted up and down 50 times. Before adding the Nuclear Extraction Reagent, the sample was washed four times with ice-cold PBS, vortexed, and centrifuged. The nuclear extract (NE) was concentrated and the buffer was exchanged to 6M urea using Microcon tubes (YM-3, 3kDa nominal molecular weight limit; Millipore). The nuclear pellet (NP) was sonicated several times for 3 seconds with 0.5 second pulses to break down DNA. Protein amount in nuclear extract and nuclear pellet was measured using the BCA protein assay (Pierce). Typically, 700 µg of protein was isolated in the NE and 1000 µg in the NP from 6 transwell plates. Mass spectrometry analysis was carried out for both of the nuclear fractions.

Sample preparation and LC-MS/MS analysis

Protein samples from NE and NP fractions were reduced, alkylated, and digested with trypsin as previously described (49). The resulting peptides were separated into 24 fractions
using strong cation exchange chromatography as previously described (26). The fractionated samples were dried in vacuo and resuspended in 0.1% formic acid. Phosphopeptides were enriched using Fe-NTA phosphopeptide-enrichment IMAC columns (Pierce) following the manufacturer’s protocol. Dried samples were resuspended in 0.1% formic acid and desalted using Graphite Spin Columns (Pierce) before analysis by LC-MS/MS. Tryptic peptides were analyzed on an LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific) equipped with a nano-electrospray ion source. The peptides were fractionated with a reverse-phase PicoFrit column (New Objective, Woburn, MA) using a linear gradient of 5%–35% ACN in 0.1% FA in 45 min at a flow rate of 0.25 μl/min. Precursor mass spectra (MS1) were acquired in the Orbitrap at 60,000 resolution, and product mass spectra (MS2) were acquired with the ion trap.

Bioinformatic Analysis

Peptide sequences were assigned to the spectra using two different search algorithms: InsPecT (55) and SEQUEST (65). For MS2 spectra, the search parameters were set for a single fixed modification (carbamidomethylation of cysteine, (+57.02146 Da) as well as several variable modifications, namely phosphorylation (+79.96633) of Ser, Thr, and Tyr, isotope labeling of lysine (+6.02013 Da) and arginine (+10.0083 Da), and oxidation of methionine (+15.99491 Da) with a maximum of four modifications and three missed cleavages allowed per peptide. SEQUEST searches were performed using Proteome Discoverer ver 1.3 (Thermo) running the most recent mouse RefSeq Database. False positives were controlled using target-decoy analysis (6) to set search filters for < 1% FDR. The InsPecT search was performed on the National Institutes of Health Biowulf cluster (http://biowulf.nih.gov). Phosphorylation site
assignment was performed using the following scoring algorithms: PhosSA for SEQUEST data (48) and Phosphate Localization Score for InsPecT data (38). Quantification of phosphopeptides based on the area under the curve of reconstructed chromatograms for corresponding heavy and light precursors was performed using QUIL software (59).

The software *Venn Diagram Plotter* (version 1.4.3740, downloaded from http://omics.pnl.gov/software/VennDiagramPlotter.php) was employed to construct Venn diagrams depicting the shared and unique proteins among lists of proteins. The software *NHLBI-ABE* (*National Heart, Lung and Blood Institute’s Automated Bioinformatics Extractor*) (available for download: http://helixweb.nih.gov/ESBL/ABE/) was used to extract *Gene Ontology Biological Process* terms from lists of proteins in this study. We employed an online tool called *PhosphoLogo* (http://helixweb.nih.gov/PhosphoLogo/) to generate logos to illustrate amino acids in positions upstream and downstream of a phosphorylation site that are overrepresented in a list of phosphorylated peptides (15).

Network analysis of potential protein interactions among the proteins whose phosphorylation changed significantly with dDAVP was carried out using the online computational tool and database called *STRING 9.0* (54), which builds functional protein-association networks based on compiled available experimental evidence. The online resource *Database for Annotation, Visualization and Integrated Discovery* (DAVID, ver 6.7, http://david.abcc.ncifcrf.gov/) was employed to classify the nuclear proteins regulated by dDAVP into functional categories. Proteins were grouped according to Biological Process (GOTERM_BP_ALL), and a list of peptides in the flow-through fractions was used as background. The in-house program *cPhos* (67) was employed to determine the number of orthologs for input phosphopeptides and the amino acid present at the corresponding site.
**Immunoblotting**

Proteins were resolved by SDS-PAGE on Criterion polyacrylamide gels (Bio-Rad, Hercules, CA) at 200 V for 45 min and transferred electrophoretically onto nitrocellulose membranes at 80 V for 40 min. Blots were blocked for 1 h with a proprietary blocking buffer (Li-Cor, Lincoln, NE), rinsed, and probed with the respective affinity-purified antibodies at proper dilution (in Licor blocking buffer containing 0.1% Tween 20) overnight at 4°C. After 1 h incubation with secondary antibody (Alexa fluor 680 anti-rabbit immunoglobulin G) at 1:5,000 dilution, sites of antibody-antigen reaction were detected using the Odyssey infrared imager (Li-Cor).
Results

Figure 1 is a flow diagram showing the experimental procedures. We performed subcellular fractionation of mpkCCD cells (clone 11) using a commercially available kit yielding the nuclear extract (NE) and the nuclear pellet (NP). We have previously characterized these fractions via immunoblotting (49). While the NE contains soluble nuclear proteins, the NP is expected to consist of predominantly DNA-bound proteins, proteins of the nuclear envelope, and parts of the endoplasmic reticulum (51). In both fractions, we carried out immobilized metal affinity chromatography (IMAC) to enrich the phosphopeptides present. In addition, we analyzed the peptides that did not bind to the IMAC column (“flow through”) to assess whether a change in the abundance of a given phosphopeptide was due to a change in the abundance of the protein from which the phosphopeptide was derived.

We employed SILAC coupled to LC-MS/MS to quantify changes in the nuclear phosphoproteome of mpkCCD cells in response to 30 minutes of treatment with the vasopressin analog dDAVP at a physiological concentration (0.1 nM) in four separate experiments. Statistical filters were employed to restrict the false discovery rate to 1% at the peptide level. The proteins that were identified in 3 or more experiments were placed in an online database to provide a public resource (http://helixweb.nih.gov/ESBL/Database/mNPPD/). Figure 2 shows a summary of the data from all experiments. Of the 1,251 proteins identified from the phosphorylated peptides eluted from the IMAC columns, 344 proteins were shared between the nuclear fractions, 401 were unique to the NE, and 506 were unique to the NP. A smaller percentage of proteins from the flow-through samples were common to both the NE and NP.

To confirm the success of the subcellular fractionation, we extracted the Biological Process Gene Ontology terms for the proteins in the flow-through (non-phosphorylated) and eluate (phosphorylated) lists of peptides. The ten most frequent Gene Ontology Biological
Process terms for the NE and NP fractions are displayed in Figure 2B. The most frequent processes were “Regulation of transcription, DNA-dependent” and “Transcription, DNA-dependent,” and the lists included primarily processes confined to the nucleus, as expected. In all, 15 transcription factors were identified in this study (Jun, Gata3, Gatad2a, Foxk2, Klf3, Rlf, Tcf20, Tcfap4, Zbtb44, Zfp148, Zfp185, Zfp219, Zfp608, Zc3h13, Zc3h18). Interestingly, in 14 out of these 15 transcription factors, the identified phosphorylation site had a proline in position +1 downstream from the phosphorylated Ser or Thr, suggesting that these are phosphorylated by protein kinases in either the MAP kinase family or cyclin-dependent kinase family (‘proline-directed’ kinase families) (45). In addition, several phosphopeptides corresponding to protein kinases were identified in nuclear fractions including Gsk3b, Mylk, Mark2, Dyrk1b, Cdk2, Bckdk and Baz1b.

To provide a stringent read-out of regulated proteins, we employed two simultaneous statistical criteria to identify the phosphorylated peptides that changed significantly in nuclear abundance (Figure 3): p < 0.05 by a t test (horizontal dashed line) and log2(dDAVP/vehicle) outside of a 95% confidence interval for vehicle-versus-vehicle experiments (vertical dashed lines) (49). Using these criteria, 11 of the 116 phosphorylation sites identified in the NE were classified as significantly changed in abundance in response to dDAVP (gray background). Table 1 summarizes the sites that were increased: S1021 of Srcin1, S1191 of Tcof1, S5522 and S5525 of Ahnak, S572 of Srrm1, and S67 of Krt19. Table 2 shows the sites that were decreased, namely S364 of myosin light chain kinase (Mylk), S73 of transcription factor c-Jun, S5504 and S5536 of Ahnak, and S30 of Sept9. Thirty five of the 356 phosphorylation sites of the NP fraction were identified as significantly changed. The 24 phosphorylation sites that were increased included 4 sites in 3 members of the catenin family: S552 of β-catenin, S862 and S864 of δ-catenin, and
S665 of Jup (γ-catenin) (Table 1). The 11 phosphorylation sites that were decreased in the NP fraction included S364 of Mylk and S1172/S1182 of Mllt4 (Afadin) (Table 2). (The presence of a phosphopeptide from the water channel Aqp2 in Table 2 likely owes to the fact that the outer membrane of the nucleus is continuous with the endoplasmic reticulum, where Aqp2 is initially formed and processed. The presence of both Ser256 phosphorylation and Ser261 phosphorylation suggests therefore that these phosphorylation events can occur in the endoplasmic reticulum. Ser261 phosphorylation is strongly decreased by vasopressin accounting for the decrease in the doubly phosphorylated peptide seen in Table 2.) As indicated in Tables 1 and 2, most of the observed phosphorylated amino acids were conserved over a range of species pointing to likely functional importance.

Similar criteria were employed to identify proteins detected in the flow-through list of peptides that exhibited significant abundance changes in response to 30-min exposure to dDAVP. Only 3 proteins of the NE (Gtf3c1, Smarcb1, and Ctnna2) and 6 proteins of the NP (Nckap1, Trpv4, Baiap2l1, Naa40, Epn3, and Synpo) changed significantly in abundance. Importantly, none of the proteins with phosphorylation sites that changed significantly exhibited corresponding changes in total abundance at the 30-minute time point used in this study. Furthermore, phosphopeptides exhibited more extreme abundance changes than non-phosphopeptides. A comparison of changes in phosphopeptides and the corresponding holo-proteins are included in the online database (http://helixweb.nih.gov/ESBL/Database/mNPPD/).

Finally, more of detected phosphorylation sites (9.96%) exhibited significant changes in abundance compared to holo-proteins (0.32%). Together, these results suggest that the changes in abundance of the phosphorylated peptides are, in general, due to changes in phosphorylation rather than changes in the underlying abundances of the corresponding proteins.
Of the 39 phosphorylation sites that were significantly changed in abundance in response to dDAVP (based on the two statistical criteria listed above), 7 were common between the two nuclear fractions, 4 were unique to the NE, and 28 were unique to the NP (Figure 4A). The abundance changes for the 7 common phosphorylation sites were approximately equal between the two fractions (Figure 4B). For these proteins, the redundancy of the quantification provides an even higher level of confidence in the observed abundance changes.

To examine the properties of the phosphorylation sites that increased or decreased in abundance, we constructed sequence logos that show overrepresented amino acids in positions upstream and downstream of the phosphorylation site using the online tool PhosphoLogo (15) (Figure 4C). The logo for the sites that were decreased showed a preference for proline at the +1 position (‘proline-directed motif’). A different logo was seen for the sites that were increased, exhibiting a preference for basic amino acids namely arginine (R) and lysine (K) at the -2 and -3 positions (‘basophilic motif’). The observed motifs overall indicate properties of the kinases responsible for phosphorylating these proteins and suggest decreased activation of MAP kinases and/or cyclin-dependent kinases (proline-directed kinase family) and increased activation of AGC family kinases and/or calmodulin-dependent kinases (basophilic kinase family). A similar conclusion was drawn in our previous study of phosphoproteomics in unfractionated mpkCCD cells (45).

Figure 5 summarizes the proteins that undergo changes in phosphorylation in the two nuclear fractions, organized according to functional categories. This figure was manually constructed based on Swiss-Prot annotations. Notably, Figure 5 includes three members of the catenin family (Ctnnb1, Ctnnd1, and Jup), the transcription factor c-Jun, and a protein kinase (Mylk) (see Discussion). Figure 5 includes proteins involved in cell-cell adhesion, intracellular...
movement, chromatin modification, cell spreading and migration, and vesicular transport.

Several proteins are involved in signal transduction including inositol 1,4,5-trisphosphate receptor type 3 (Itpr3), SRC kinase signaling inhibitor 1 (Srcin1), septin 9 (Sept9), the G protein-coupled receptor Gprc5c, tight junction protein 1 (ZO-1 or Tjp1), and the nucleoprotein Ahnak.

Ser364 of myosin light chain kinase (Mylk) exhibited the largest decrease in abundance in both the NE and NP fractions. Analysis using DAVID software (Methods) showed that the category of regulated proteins classified by the Gene Ontology Biological Process term “Cell adhesion” is enriched 14.3 fold (relative to all proteins found in the flow-through fractions from NE and NP samples). Figure 6 shows a protein-protein interaction network generated by STRING 9.0 online software (see Methods). The input was the list of all proteins in Tables 1 and 2 (red nodes) in addition to two protein lists identified in previous studies: 1) the list of proteins found in the core vasopressin signaling network described by Pisitkun et al. (41) (gray nodes) and 2) the list of proteins that were found to translocate into the nucleus in response to vasopressin exposure (49) (blue nodes). The edges indicate predicted functional links based on experimental evidence, databases, or text-mining of literature as explained on the STRING website. (The user may regenerate Figure 7 at the STRING website using the dataset available in Supplemental Dataset 1.) The network generated has three main clusters: a cell-cell adhesion cluster shown on the left, the core vasopressin-signaling cluster shown in the center, and a transcriptional regulation cluster shown on the right. There are two major hubs (highest connectivity) namely β-catenin and c-Jun. Because of the presumed central role of these proteins, we carried out immunoblotting to confirm the vasopressin-induced changes in phosphorylation of c-Jun and β-catenin (Figure 7).
Discussion

Vasopressin regulates water transport in the renal collecting duct in part by inducing changes in the rate of transcription of the \( Aqp2 \) gene. Concomitantly, it regulates transcription of a number of other genes including the \( \beta \) and \( \gamma \) subunits of the ENaC (16, 33) and the calcium-binding protein calcyclin (12). In addition, large-scale studies using the SAGE technique in mpkCCD cells have identified 59 vasopressin-regulated transcripts that are believed to change as a result of transcriptional regulation (12). One process involved in transcriptional regulation is the phosphorylation of regulatory proteins such as transcription factors and transcriptional co-regulators. These proteins act at \(\text{cis}\)-elements (enhancers or repressors) to indirectly regulate RNA polymerase II-mediated transcription. In this study, we employed a SILAC-based LC-MS/MS method to quantify changes in the nuclear phosphoproteome of mpkCCD cells (recloned to maximize vasopressin responses (66)) following 30-min exposure to the V2-receptor selective vasopressin analog dDAVP. In total, we identified 3,674 phosphorylation sites in 1,251 proteins from biochemically isolated nuclei. The proteins that were identified in at least 3 (out of 4) experiments were placed in an online database to provide a public resource (http://helixweb.nih.gov/ESBL/Database/mNPPD/). Only 39 phosphorylation sites on 29 proteins satisfied the two statistical criteria (Methods) used to select likely true-positive results. It is probable that this statistical approach excluded some regulated sites, but it emphasizes a small subset of nuclear proteins that appear most likely to be involved in regulation of transcription in the collecting duct. In the remainder of this discussion, we relate the findings of this study to the existing literature on the effects of vasopressin on water transport.
An important intermediate goal in seeking the possible physiological significance of a particular proteomic data set is integration of the data with other relevant data sets. One approach to this is to classify the proteins that were found to be regulated with regard to their annotated functions. When this was done (Figure 6) several protein groups were evident. First, there were a number of classical nuclear proteins with diverse functions in the nucleus, including chromatin modification, RNA processing, ribosome biogenesis and transcriptional regulation. Second, there are a number of proteins with predominantly cytoplasmic functions, but which may nevertheless translocate into the nucleus, either as a constitutive process or in a regulated manner. The prototypical example is β-catenin (Ctnnb1), which is a component of adherens junctions but also is a signaling molecule (Wnt pathway) that can translocate to the nucleus where it acts as a transcriptional coactivator (9). This list includes several adhesion proteins including the tight junction protein ZO-1 (Tjp1), which was also previously demonstrated in nuclei (42, 49). It is perhaps surprising that there were so many regulated nuclear phosphoproteins that are known as cell-adhesion proteins. Indeed, Gene Ontology term enrichment analysis revealed that these categories of proteins are present substantially out of proportion relative to other categories of proteins that normally function in the cytoplasm. In general, only cell adhesion proteins that do not span the plasma membrane were found to be present in nuclear fractions and regulated including δ-catenin, β-catenin, γ-catenin (Jup), ZO-1 (Tjp1) and afadin (Mllt4). Interestingly, afadin contains a so-called FHA domain (Forkhead associated domain), which is known to mediate nuclear translocation and to bind phosphorylated amino acids to regulate a number of transcription factors and protein kinases (28). Thus, the
results raise the possibility that nuclear translocation and phosphorylation of non-integral-
membrane cell-adhesion proteins may be a more general phenomenon than currently appreciated.

Another useful tool in data integration is network analysis. We used STRING 9.0 (54) to
identify and map relationships between our list of vasopressin-regulated nuclear phosphoproteins
(Figure 7). Though not all proteins had connections, the network manifested two major clusters
of proteins. One is composed of proteins involved in the regulation of gene expression. A second
is composed of proteins involved in cell-cell junctions, mainly adherens junctions, tight junctions
and desmosomes. The hub proteins for these two protein clusters were c-Jun and β-catenin,
respectively.

Although all of the proteins found to have altered phosphorylation in response to
vasopressin are potentially important in vasopressin-mediated transcriptional regulation, for the
reasons summarized in the preceding paragraphs, β-catenin and c-Jun appear to be the best
candidates for roles in transcriptional regulation of Aqp2 by vasopressin. Hence, we discuss these
proteins in greater detail in what follows. In addition, we consider a third phosphoprotein,
myosin light chain kinase (Mylk), because its phosphorylation exhibited the greatest decrease
among all of the phosphorylation sites quantified in this study.

β-Catenin. β-catenin acts in the Wnt signaling pathway to activate the transcription of
crucial target genes responsible for cellular proliferation and differentiation. It additionally
controls E-cadherin-mediated cell adhesion at the plasma membrane and mediates the attachment
of adherens junctions to the actin cytoskeleton (9). In the nucleus, β-catenin has been shown to
act as a transcriptional coactivator in conjunction with the T-cell-specific transcription factor
(TCF)/lymphoid enhancer binding factor family of proteins (LEF) (32). Phosphorylation of
Ser552 of β-catenin by Akt and protein kinase A (PKA) has been shown to promote interactions
with additional transcriptional factors such as cAMP-response element-binding protein (CREB) and to enhance \( \beta \)-catenin/TCF reporter activation through association with histone deacetylases (13). Phosphorylation by Akt at Ser552 has also been shown to increase its transcriptional activity (19). Notably, in mpkCCD cells, both PKA catalytic subunit genes (\textit{Prkaca} and \textit{Prkacb}) and all three Akt genes (\textit{Akt1}, \textit{Akt2} and \textit{Akt3}) are expressed at moderate levels. An Akt-related kinase, Sgk3, is also strongly expressed in mpkCCD cells (66), whereas a different Sgk gene, Sgk1, is expressed in native rat IMCD cells (10) (58). Studies in amphibian A6 cells (3, 18), and mammalian cell lines (52) have supported the view that Sgk orthologs are involved in the response to antidiuretic hormone. Although direct evidence for activation of Sgk isoforms by vasopressin (e.g. by demonstrating increased enzyme activity or activation loop phosphorylation) is to our knowledge lacking, studies in mammalian COS7 cells have demonstrated that cyclic AMP increases Sgk activity (39).

Figure 8A shows the general structure of mouse \( \beta \)-catenin indicating the position of the Ser552 phosphorylation site relative to the remainder of the protein. The site is at the COOH-terminal end of a series of so-called “Armadillo repeats”. This region is involved with the binding of \( \beta \)-catenin to E-cadherin in the adherens junction complex and to a nuclear protein called ICAT (\textit{Inhibitor of \( \beta \)-Catenin And TCF4}) (61). ICAT inhibits Wnt signaling by interfering with the formation of appropriate \( \beta \)-catenin-containing transactivation complexes (61). The COOH-terminal end of \( \beta \)-catenin possesses a putative PDZ-ligand sequence and the NH2-terminal end has a “Vinculin Binding” motif. Perhaps, the most well characterized part of \( \beta \)-catenin is a series of serines near the NH2-terminus that, as part of the Wnt pathway, becomes sequentially phosphorylated by GSK3\( \beta \), leading to its degradation. GSK3\( \beta \) has previously been implicated in regulation of \textit{Aqp2} gene expression in mouse collecting duct (44).
We have previously shown an increase in phosphorylation of Ser552 of β-catenin at the whole-cell level in mpkCCD cells (45). The present work, coupled with that of Schenk et al (49), suggests that this phosphorylated form is also present in the nucleus of mpkCCD cells. Indeed, phosphorylation at Ser552 has been shown to promote β-catenin nuclear localization (19).

Phosphorylation of β-catenin at Ser552 by Akt has also been found to enhance β-catenin protein levels and nuclear signaling by standard reporter assays (19, 57). Both protein kinases that have been shown to phosphorylate this site, Akt and PKA, are known to be activated in response to vasopressin in inner medullary collecting duct (41). Overall, the data support the view, previously presented, that the Wnt pathway is active in mature collecting duct cells and that the Wnt signaling cascade is upregulated by vasopressin (46).

The Ser552 phosphorylation site of β-catenin (QRRTS*MGGT) displays a “basophilic” motif (R in positions P-2 and P-3) consistent with the conclusion that a so-called basophilic protein kinase (e.g. from the AGC or calmodulin-dependent kinase families) mediates the phosphorylation in response to vasopressin. Figure 8B shows the high level of conservation of this sequence among species ranging from humans to mosquitos. Besides PKA and Akt, other kinases can phosphorylate serines and threonines within a similar motif, including Sgk1, protein kinase Cδ, calmodulin-dependent kinase 2δ, and another Ca^{2+}/calmodulin dependent kinase, namely death-associated protein kinase 1 (Dapk1) (15). Thus, we cannot conclude which protein kinase is the physiologically important mediator of Ser552 phosphorylation of β-catenin in the nuclei of collecting duct cells.

We also observed that the phosphorylation of two other members of the catenin family of proteins, δ-catenin and γ-catenin (gene symbol: Jup) was significantly increased in nuclei in
response to dDAVP. Members of the catenin family contribute to the indirect association of
cadherins with the underlying cytoskeleton at adherens junctions, are involved with linkages to
intermediate filaments at desmosomes, and modulate cadherin endocytosis (32). The abundance
of the catenin family of proteins among the list of proteins with phosphorylation sites that change
in response to dDAVP likely reflects their central role in the vasopressin signaling network. In
addition, it seems possible that δ-catenin and γ-catenin, like β-catenin, may play roles as
transcriptional coactivators in the nucleus.

c-Jun. The transcription factor c-Jun (gene symbol: Jun) commonly dimerizes with c-Fos
or one of its homologs to form the AP-1 transcription factor. AP-1 regulates many cellular
processes including differentiation, proliferation and apoptosis (4). We have observed a similar
decrease in phosphorylation of Ser73 of c-Jun at both the nuclear (this study) and whole-cell (45)
levels. This site is the canonical site for phosphorylation by Jnk and other related MAP kinases
(43), and its phosphorylation is required for activity of the enzyme (1) Phosphorylation of c-Jun
at Ser73 has been shown to enhance transcriptional activation (43) (50). Figure 9A shows the
general structure of c-Jun from mouse pinpointing the Ser73 phosphorylation site relative to the
bZIP domain structure of the protein and to other phosphorylation sites including an annotated
GSK3β site. Transcriptomic analysis has demonstrated that c-Jun, JunB, and c-Fos are
expressed in both native rat IMCD (58) and cultured mpkCCD (clone 11) cells (66) at levels well
above the median for all transcripts (see databases accessible at
https://intramural.nhlbi.nih.gov/labs/LKEM_G/LKEM/Pages/
AP-1 has already been implicated in regulation of Aqp2 transcription. Yasui et al. demonstrated that vasopressin induces both adenosine 3’,5’-cyclic monophosphate (cAMP) responsive element binding protein (CREB) phosphorylation and c-Fos expression and that the c-Fos/c-Jun heterodimer binds to the AP-1 element leading to APQ2 promoter activation in cultured LLC-PK1 cells (64). The results of the study, therefore, implicate c-Jun as a transcriptional regulator of Aqp2. However, the observed decrease in Ser73 phosphorylation would then suggest the opposite conclusion, raising the possibility that some other AP-1 site plays a repressive role. We have previously used bioinformatics to identify one putative binding site for members of the Jun family of proteins in the 5’-flanking region of Aqp2 815 bp upstream from the transcription site (49). However, data from the ENCODE project (47) reporting ChIP-seq data for a number of human cell lines shows c-Fos, Fosl2, JunD and JunB binding approximately 4000-5000 bp upstream from the Aqp2 transcription start site pointing to a likely AP-1 binding site (See URL: http://genome.ucsc.edu/cgi-bin/hgTracks?position=chr12:50339048-50354957&hgsid=287643327&wgEncode=full). The functional role of this site has not been evaluated in collecting duct cells.

The Ser73 phosphorylation site of c-Jun identified (LAS*PELER) furthermore exhibits a “proline-directed” motif consistent with a role for either MAP kinases or cyclin-dependent kinases in the response to vasopressin. Figure 9B shows a very high level of conservation of this sequence among species ranging from humans to mosquitoes. We have previously shown that phosphopeptides decreasing in abundance in response to dDAVP had a predominance of “proline-directed” motifs (45). The family of c-Jun NH2-terminal kinases (JNKs), examples of proline-directed protein kinases, have been shown to bind to and phosphorylate the NH2-terminal activation domain of c-Jun (1) (25). We have previously shown that vasopressin treatment in
native collecting duct cells decreases the phosphorylation of Jnk2 (Mapk9) and p38α (Mapk14) at their canonical activation sites (26). The Jnk signal transduction pathway contributes to the regulation of cell proliferation and apoptosis, is required for embryonic morphogenesis, and contributes to the function of differentiated cells.

Myosin Light Chain Kinase. Among all phosphorylation sites quantified in nuclear fractions, we observed the largest decrease at Ser364 of myosin light chain kinase (gene symbol: Mylk) in response to dDAVP. The myosin light chain kinases are a family of serine/threonine, calcium/calmodulin-dependent kinases that phosphorylate myosin regulatory light chain (22). We have previously shown that Mylk phosphorylates two myosin light chain isoforms in native inner medullary collecting duct cells and is an important downstream component of vasopressin signaling in inner medullary collecting duct cells via its effect on AQP2 trafficking (11).

Mylk is the only kinase that underwent a significant change in phosphorylation in response to dDAVP in the present study of nuclear proteins. The finding of regulated phosphorylation of Mylk in the nucleus raises the possibility that Mylk may play a role in the elements of vasopressin signaling that govern transcriptional regulation of Aqp2. In a previous study (49), we identified Mylk’s target myosin regulatory light chain B as well as the corresponding conventional non-muscle myosin heavy chains coded by the Myh9 and Myh10 genes in nuclear fractions of mpkCCD (clone 11) cells (Web search: mpkCCD Nucleus Proteome Database), although their abundances did not change with vasopressin. It remains a possibility that Mylk phosphorylates other target proteins in the nucleus such as cell-adhesion kinase beta (Pyk2β) (60).

The Ser364 phosphorylation site of Mylk has not been previously characterized to our knowledge. The sequence surrounding this site (AIGSFS*PGEDR) is compatible with a “proline
directed” kinase (that is, a MAP kinase or a cyclin-dependent kinase). This site is shown in Figure 10 relative to the overall domain structure of Mylk. It sits amid a series of so-called “immunoglobulin-like” domains far upstream of the actin-binding and calmodulin-binding domains as well as the catalytic domain of the kinase. This region has been associated with binding of Mylk to actin microfilaments (stress fibers) (62).
Footnote 1. Throughout the text, we use italics to indicate gene symbols. Non-italicized strings are common abbreviations.

Acknowledgements

Mass spectrometry was performed in the National Heart, Lung, and Blood Institute Proteomics Core Facility (Director: Marjan Gucek). We would like to thank Guanghui Wang of the Proteomics Core Facility for assistance performing mass spectrometry. This work was funded by the Division of Intramural Research of the National Heart, Lung, and Blood Institute (Project Z01-HL001285 to M.A.K.).

Disclosures

The authors declare no conflicts of interest.
References


Table 1. Phosphopeptides increased in abundance in response to dDAVP. The abundance changes are reported as log₂(dDAVP/vehicle) for both the phosphopeptide and the corresponding total protein. Conservation of phosphorylation sites was determined as the proportion of orthologs with a corresponding serine, threonine, or tyrosine at the conserved site. Phosphopeptides identified in the NE fraction are shaded gray. Ambiguous phosphorylation site assignments are labeled with an asterisk.

<table>
<thead>
<tr>
<th>Nuclear Fraction</th>
<th>Gene Symbol</th>
<th>RefSeq</th>
<th>Peptide Sequence</th>
<th>Phosphorylation Site</th>
<th>log₂(dDAVP/vehicle)</th>
<th>Number of Orthologs</th>
<th>Proportion of Conserved Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>Clmn</td>
<td>NP_001035772</td>
<td>TGS*IAEPTPESSILSTR</td>
<td>S397*</td>
<td>5.08</td>
<td>0.07</td>
<td>8</td>
</tr>
<tr>
<td>NP</td>
<td>Sec22b</td>
<td>NP_003472</td>
<td>NILGS*INTELQVQOR</td>
<td>S137</td>
<td>3.92</td>
<td>-0.12</td>
<td>18</td>
</tr>
<tr>
<td>NP</td>
<td>Itp3</td>
<td>NP_542120</td>
<td>VSS*FSMPSSSR</td>
<td>S1932</td>
<td>2.49</td>
<td>0.1</td>
<td>8</td>
</tr>
<tr>
<td>NP</td>
<td>Hst1h1d</td>
<td>NP_663759</td>
<td>KAS*GPPVSELITK</td>
<td>S37</td>
<td>2.33</td>
<td>-0.01</td>
<td>6</td>
</tr>
<tr>
<td>NP</td>
<td>Crnd1</td>
<td>NP_031641</td>
<td>SQSS*HXYDDSTPLDR</td>
<td>S862</td>
<td>2.23</td>
<td>0.26</td>
<td>8</td>
</tr>
<tr>
<td>NP</td>
<td>Hst1h1d</td>
<td>NP_663759</td>
<td>KASGPPVS*ELITK</td>
<td>S42*</td>
<td>2.12</td>
<td>-0.01</td>
<td>6</td>
</tr>
<tr>
<td>NE</td>
<td>Krt19</td>
<td>NP_032407</td>
<td>GGS*FSGLAVSGLALLGNEK</td>
<td>867</td>
<td>2.00</td>
<td>-0.01</td>
<td>7</td>
</tr>
<tr>
<td>NP</td>
<td>Clmn</td>
<td>NP_001035772</td>
<td>RTS*MGGTQQQFVEGR</td>
<td>S552*</td>
<td>1.79</td>
<td>0.12</td>
<td>12</td>
</tr>
<tr>
<td>NP</td>
<td>Sec22b</td>
<td>NP_003472</td>
<td>SQSS*HXYDDSTPLDR</td>
<td>S864</td>
<td>1.78</td>
<td>0.26</td>
<td>8</td>
</tr>
<tr>
<td>NP</td>
<td>Itp3</td>
<td>NP_542120</td>
<td>VSS*FSMPSSSR</td>
<td>S1932</td>
<td>2.49</td>
<td>0.1</td>
<td>8</td>
</tr>
<tr>
<td>NP</td>
<td>Crnd1</td>
<td>NP_031641</td>
<td>SQSS*HXYDDSTPLDR</td>
<td>S862</td>
<td>2.23</td>
<td>0.26</td>
<td>8</td>
</tr>
<tr>
<td>NP</td>
<td>Cbx5</td>
<td>NP_001035772</td>
<td>RKS*FSNSADDIK</td>
<td>S93</td>
<td>1.62</td>
<td>-0.01</td>
<td>10</td>
</tr>
<tr>
<td>NP</td>
<td>Krt19</td>
<td>NP_032407</td>
<td>GGS*FSGLAVSGLALLGNEK</td>
<td>867</td>
<td>1.50</td>
<td>-0.01</td>
<td>7</td>
</tr>
<tr>
<td>NP</td>
<td>Ags2</td>
<td>NP_033829</td>
<td>ROS*VELHSOSQGLPR</td>
<td>S256</td>
<td>1.50</td>
<td>N/A</td>
<td>10</td>
</tr>
<tr>
<td>NP</td>
<td>Nop56</td>
<td>NP_0077155</td>
<td>KFS*EEPEVAANFTK</td>
<td>S554</td>
<td>1.50</td>
<td>-0.05</td>
<td>18</td>
</tr>
<tr>
<td>NP</td>
<td>Cbx5</td>
<td>NP_001035772</td>
<td>S*SFNSADDIK</td>
<td>S92</td>
<td>1.44</td>
<td>-0.01</td>
<td>10</td>
</tr>
<tr>
<td>NP</td>
<td>Hst1h1e</td>
<td>NP_056602</td>
<td>KTS*GPPVSELITK</td>
<td>S36*</td>
<td>1.09</td>
<td>-0.05</td>
<td>6</td>
</tr>
<tr>
<td>NE</td>
<td>Snc1n</td>
<td>NP_001035772</td>
<td>RGS*DELTVPR</td>
<td>S1021</td>
<td>1.08</td>
<td>0.14</td>
<td>8</td>
</tr>
<tr>
<td>NE</td>
<td>Snc1n</td>
<td>NP_001035772</td>
<td>RGS*DELTVPR</td>
<td>S1021</td>
<td>1.08</td>
<td>0.14</td>
<td>8</td>
</tr>
<tr>
<td>NE</td>
<td>Ahnak</td>
<td>NP_033773</td>
<td>AS*LGSLEGEVEAASSPK</td>
<td>5522</td>
<td>0.86</td>
<td>0.15</td>
<td>7</td>
</tr>
<tr>
<td>NE</td>
<td>Lmna</td>
<td>NP_001035772</td>
<td>GSHCSGSGDPAEYNLR</td>
<td>S570</td>
<td>0.81</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>NE</td>
<td>Tcof1</td>
<td>NP_035692</td>
<td>KLS*KDEAGAPK</td>
<td>S1191</td>
<td>0.81</td>
<td>-0.22</td>
<td>7</td>
</tr>
<tr>
<td>NE</td>
<td>Ahnak</td>
<td>NP_033773</td>
<td>AS*LGSLEGEVEAASSPK</td>
<td>5522</td>
<td>0.77</td>
<td>0.12</td>
<td>7</td>
</tr>
<tr>
<td>NE</td>
<td>Snn1</td>
<td>NP_058079</td>
<td>S*PSAPPPPPPPPPPPPPPR</td>
<td>S572</td>
<td>0.72</td>
<td>-0.04</td>
<td>13</td>
</tr>
<tr>
<td>NE</td>
<td>Ahnak</td>
<td>NP_033773</td>
<td>AS*LGSLEGEVEAASSPK</td>
<td>5522</td>
<td>0.70</td>
<td>0.12</td>
<td>7</td>
</tr>
<tr>
<td>NP</td>
<td>Hmpk</td>
<td>NP_079555</td>
<td>GSYGDLGGPIPTQVIPK</td>
<td>S379*</td>
<td>0.63</td>
<td>-0.01</td>
<td>9</td>
</tr>
<tr>
<td>NP</td>
<td>Jup</td>
<td>NP_034723</td>
<td>RVS*VELNSLFK</td>
<td>S665</td>
<td>0.60</td>
<td>0.16</td>
<td>8</td>
</tr>
<tr>
<td>NP</td>
<td>Tip3</td>
<td>NP_038797</td>
<td>S*GGGSEANGLDLVSGYK</td>
<td>3156</td>
<td>0.59</td>
<td>0.14</td>
<td>8</td>
</tr>
<tr>
<td>NP</td>
<td>Hst1h1e</td>
<td>NP_056602</td>
<td>TSGPPVSELITK</td>
<td>T36*</td>
<td>0.58</td>
<td>-0.05</td>
<td>6</td>
</tr>
<tr>
<td>NP</td>
<td>Tip3</td>
<td>NP_038797</td>
<td>RSS*GGGSEANGLDLVSGYK</td>
<td>3157*</td>
<td>0.57</td>
<td>0.14</td>
<td>8</td>
</tr>
</tbody>
</table>
Table 2. Phosphopeptides decreased in abundance in response to dDAVP. The abundance changes are reported as log2(dDAVP/vehicle) for both the phosphopeptide and the corresponding total protein. Conservation of phosphorylation sites was determined as the proportion of orthologs with a corresponding serine, threonine, or tyrosine at the conserved site. Phosphopeptides identified in the NE fraction are shaded gray. Ambiguous phosphorylation site assignments are labeled with an asterisk.

<table>
<thead>
<tr>
<th>Nuclear Fraction</th>
<th>Gene Symbol</th>
<th>RefSeq</th>
<th>Peptide Sequence</th>
<th>Phosphorylation Site</th>
<th>log2(dDAVP/vehicle)</th>
<th>Number of Orthologs</th>
<th>Proportion of Conserved Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>NE</td>
<td>Mylk</td>
<td>NP_647461</td>
<td>VPAIGSFS*PGEDR</td>
<td>S364</td>
<td>-1.36</td>
<td>-0.03</td>
<td>7</td>
</tr>
<tr>
<td>NP</td>
<td>Mylk</td>
<td>NP_647461</td>
<td>VPAIGSFS*PGEDR</td>
<td>S364</td>
<td>-1.27</td>
<td>0.04</td>
<td>7</td>
</tr>
<tr>
<td>NP</td>
<td>Aqp2</td>
<td>NP_033829</td>
<td>RQS<em>VELHS</em>PQSLPR</td>
<td>S256, S261</td>
<td>-1.13</td>
<td>N/A</td>
<td>10, 10</td>
</tr>
<tr>
<td>NP</td>
<td>Fam207a</td>
<td>NP_598759</td>
<td>AVEPALPQS*PAGGAGAK</td>
<td>S38</td>
<td>-1.09</td>
<td>N/A</td>
<td>9</td>
</tr>
<tr>
<td>NP</td>
<td>Ahnak</td>
<td>NP_033773</td>
<td>ASLGSLEGVEAEASS*PK</td>
<td>S5356</td>
<td>-1.04</td>
<td>0.15</td>
<td>7</td>
</tr>
<tr>
<td>NP</td>
<td>Tjp1</td>
<td>NP_033412</td>
<td>IDS*PQLKASQQK</td>
<td>S912</td>
<td>-0.95</td>
<td>0.09</td>
<td>9</td>
</tr>
<tr>
<td>NE</td>
<td>Ahnak</td>
<td>NP_033773</td>
<td>ASLGSLEGVEAEASS*PK</td>
<td>S5356</td>
<td>-0.93</td>
<td>0.12</td>
<td>7</td>
</tr>
<tr>
<td>NP</td>
<td>Mllt4</td>
<td>NP_034936</td>
<td>S<em>SPNAVANOPPS</em>PGGK</td>
<td>S1172*, S1182*</td>
<td>-0.83</td>
<td>0.17</td>
<td>11, 11</td>
</tr>
<tr>
<td>NP</td>
<td>Cobi</td>
<td>NP_766084</td>
<td>TPGNCEPHS*PK</td>
<td>S1128</td>
<td>-0.83</td>
<td>N/A</td>
<td>8</td>
</tr>
<tr>
<td>NP</td>
<td>Tjp1</td>
<td>NP_001157046</td>
<td>IDS*PQLKASQQVYR</td>
<td>S912</td>
<td>-0.77</td>
<td>N/A</td>
<td>9</td>
</tr>
<tr>
<td>NP</td>
<td>Gprc5c</td>
<td>NP_671750</td>
<td>ATANSQVMGS*ANSTLR</td>
<td>S402</td>
<td>-0.73</td>
<td>0.19</td>
<td>9</td>
</tr>
<tr>
<td>NE</td>
<td>Ahnak</td>
<td>NP_033773</td>
<td>GGTGVS*PEASISGSK</td>
<td>S5504</td>
<td>-0.57</td>
<td>0.12</td>
<td>7</td>
</tr>
<tr>
<td>NP</td>
<td>Ahnak</td>
<td>NP_033773</td>
<td>GGTGVS*PEASISGSK</td>
<td>S5504</td>
<td>-0.57</td>
<td>0.15</td>
<td>7</td>
</tr>
<tr>
<td>NP</td>
<td>Ahnak</td>
<td>NP_033773</td>
<td>LPSGS<em>GPAS</em>PTTGSAVDIR</td>
<td>S213, S217</td>
<td>-0.54</td>
<td>0.15</td>
<td>7, 7</td>
</tr>
<tr>
<td>NE</td>
<td>Jun</td>
<td>NP_034721</td>
<td>LAS*PHELER</td>
<td>S73</td>
<td>-0.53</td>
<td>N/A</td>
<td>11</td>
</tr>
<tr>
<td>NE</td>
<td>Sept9</td>
<td>NP_001100958</td>
<td>S*FEVEEIEPNNSTPR</td>
<td>S30</td>
<td>-0.52</td>
<td>-0.08</td>
<td>6</td>
</tr>
</tbody>
</table>
Figure 1. Flow diagram of methodology. (A) SILAC labeling and isolation of nuclear fractions. (B) Sample preparation for LC/MS/MS. (C) Analysis of output from LC-MS/MS.

Figure 2. Comparison of proteins identified by MS from eluate and flow-through fractions. (A) Venn diagrams show proteins that are common to nuclear extract and nuclear pellet as well as proteins that are unique to one of the two samples. (B) Ten most common Gene Ontology Biological Process terms assigned to the proteins in the eluate and flow-through sub-fractions of nuclear fractions.

Figure 3. Plots showing -log(p) vs. log$_2$(dDAVP/vehicle) for the (A) eluate and (B) flow-through peptides. Peptides satisfying two statistical criteria ($p < 0.05$ by $t$-test and log$_2$(dDAVP/vehicle) outside of a 95% confidence interval for vehicle-versus-vehicle experiments) are found within the shaded areas.

Figure 4. Distribution of phosphorylation sites that changed significantly in response to dDAVP. (A) Venn diagram of regulated sites identified in the nuclear fractions. (B) Plot of values for the 7 peptides quantified in both the nuclear extract and the nuclear pellet. (C) Consensus logos for the peptides that were decreased (left) or increased (right) in abundance in response to dDAVP.

Figure 5. Classification of proteins with peptides that were significantly changed in abundance in response to dDAVP. Proteins are identified by their official gene symbols and are classified according to the “[FUNCTION]” annotations in the respective Swiss-Prot records. Abundance
changes for each phosphorylation site shown in Tables 1 and 2 were reported as log2(dDAVP/vehicle). Positive abundance changes were labeled green, and negative abundance changes were labeled red. Values in bold refer to changes in the nuclear extract.

**Figure 6.** Vasopressin signaling network constructed using STRING 9.0. The core vasopressin signaling network (see text) is indicated by the white nodes. The red nodes correspond to proteins identified in this study, while the blue nodes indicate proteins from Schenk et al. (49). The pink nodes indicate proteins common to both studies. Many of the proteins fell into one of two functional categories (“Cell-Cell Adhesion” and “Transcriptional Regulation”).

**Figure 7.** Immunoblotting of phospho-β-catenin and phospho-c-Jun (n=3). (A) Immunoblotting of phospho-β-catenin (Ser552) showed a significant increase in abundance in response to dDAVP. (B) Immunoblotting of phospho-c-Jun (Ser73) revealed a significant decrease in abundance in response to dDAVP.

**Figure 8.** β-catenin structure and alignment. (A) NHLBI-AbDesigner (http://helixweb.nih.gov/AbDesigner/) was employed to construct a visual representation of the domains, phosphorylation sites, and other structural information. (B) Alignment across 11 species of β-catenin sequences spanning the Ser552 phosphorylation site.

**Figure 9.** c-Jun structure and alignment. (A) NHLBI-AbDesigner (http://helixweb.nih.gov/AbDesigner/) was used to construct a visual representation of the
domains, phosphorylation sites, and other structural information. (B) Alignment across 11 species of c-Jun sequences spanning the Ser73 phosphorylation site.

**Figure 10. Mylk structure.** NHLBI-AbDesigner (http://helixweb.nih.gov/AbDesigner/) was employed to construct a visual representation of the domains, phosphorylation sites, and other structural information of *Mylk.*
**Phosphopeptide Enrichment**

**Light Medium**
- $^{12}$C $^{14}$N arginine
- $^{12}$C lysine

**Heavy Medium**
- $^{13}$C $^{15}$N arginine
- $^{13}$C lysine

**SILAC labeling**
- (16 days, >6 doubling times)

**Cell Culture**
- (transmembrane supports, 3 days, 1 nM dDAVP)

**dDAVP Withdrawal**
- (6 hrs)

**dDAVP or Vehicle Treatment**
- (0.1 nM, 30 min)

**Heavy and Light Samples Pooled 1:1**

**Nuclear-Cytoplasmic Separation (NE-PER)**

**Nuclear Pellet**

**Nuclear Extract**

**In-Solution Trypsin Digestion, Fractionation**

**Phosphopeptide Enrichment (Fe-NTA IMAC column)**

**Flow Through**
- (Nonphosphorylated Peptides)

**Eluate**
- (Phosphorylated Peptides)

**LC-MS/MS (LTQ Orbitrap Velos)**

**Peptide Identification**
- (SEQUEST, InsPecT)

**Phosphorylation Site Assignment (PhosSA, PLS)**

**Peptide Quantification (QUIL)**
A

Eluate (1251 proteins)

Nuclear extract

401 344 506

Nuclear pellet

Flow Through (4121 proteins)

Nuclear extract

686 2065 1370

Nuclear pellet

B

<table>
<thead>
<tr>
<th>Eluate Count</th>
<th>Nuclear Pellet Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regulation of transcription, DNA-dependent</td>
<td>127</td>
</tr>
<tr>
<td>Transcription, DNA-dependent</td>
<td>122</td>
</tr>
<tr>
<td>Transport</td>
<td>75</td>
</tr>
<tr>
<td>mRNA processing</td>
<td>72</td>
</tr>
<tr>
<td>RNA splicing</td>
<td>59</td>
</tr>
<tr>
<td>Positive regulation of transcription from RNA polymerase II promoter</td>
<td>44</td>
</tr>
<tr>
<td>Negative regulation of transcription, DNA-dependent</td>
<td>43</td>
</tr>
<tr>
<td>Cell cycle</td>
<td>40</td>
</tr>
<tr>
<td>Metabolic process</td>
<td>38</td>
</tr>
<tr>
<td>Negative regulation of transcription from RNA polymerase II promoter</td>
<td>38</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Flow Through Count</th>
<th>Nuclear Pellet Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regulation of transcription, DNA-dependent</td>
<td>421</td>
</tr>
<tr>
<td>Transcription, DNA-dependent</td>
<td>420</td>
</tr>
<tr>
<td>Metabolic process</td>
<td>244</td>
</tr>
<tr>
<td>Transport</td>
<td>226</td>
</tr>
<tr>
<td>mRNA processing</td>
<td>178</td>
</tr>
<tr>
<td>RNA splicing</td>
<td>146</td>
</tr>
<tr>
<td>Cell cycle</td>
<td>139</td>
</tr>
<tr>
<td>Protein transport</td>
<td>155</td>
</tr>
<tr>
<td>Oxidation-reduction process</td>
<td>125</td>
</tr>
<tr>
<td>RNA splicing</td>
<td>154</td>
</tr>
<tr>
<td>Apoptotic process</td>
<td>121</td>
</tr>
<tr>
<td>Oxidation-reduction process</td>
<td></td>
</tr>
<tr>
<td>RNA splicing</td>
<td></td>
</tr>
<tr>
<td>Apoptotic process</td>
<td></td>
</tr>
</tbody>
</table>
A) Venn diagram showing the overlap of nuclear extract and nuclear pellet.

B) Graph showing the log2(dDAV/vehicle) for various proteins.

C) Graphs showing decreased and increased residue positions with corresponding bits.
A  
**c-Jun**

<table>
<thead>
<tr>
<th>Normalized Area</th>
<th>- dDAVP</th>
<th>+ dDAVP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>0.6</td>
</tr>
<tr>
<td>0.4</td>
<td>1</td>
<td>0.6</td>
</tr>
<tr>
<td>0.6</td>
<td>1</td>
<td>0.6</td>
</tr>
<tr>
<td>0.8</td>
<td>1</td>
<td>0.6</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.6</td>
</tr>
</tbody>
</table>

B  
**β-catenin**

<table>
<thead>
<tr>
<th>Normalized Area</th>
<th>- dDAVP</th>
<th>+ dDAVP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>2.5</td>
</tr>
<tr>
<td>0.4</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>0.6</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>0.8</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

* Significant difference.
Mouse *Mylk* (Myosin Light Chain Kinase, Q6PDN3)

**Kyte-Doolittle Hydropathy**

- Ig Domain 1
- Ig Domain 2
- Ig Domain 3
- Ig Domain 4
- Ig Domain 5
- Ig Domain 6
- Actin Binding
- Phosphorylation Sites
- Ser364

**Kyte-Doolittle Hydropathy**

- Ig Domain 7
- Ig Domain 8
- Fibronectin
- Kinase Catalytic Domain
- Ig Domain 9
- Actin Binding
- Calmodulin Binding
- Phosphorylation Sites
- GSK3β
- ERK1/2