Genome-Wide Mapping of DNA Accessibility and Binding Sites for CREB and C/EBPβ in Vasopressin-Sensitive Collecting Duct Cells

Hyun Jun Jung,1 Viswanathan Raghuram,1 Jae Wook Lee,2 and Mark A. Knepper1

1Epithelial Systems Biology Laboratory, Systems Biology Center, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland; and 2National Cancer Center, Goyang, Gyeonggi-do, Republic of Korea

ABSTRACT

Background Renal water excretion is controlled by vasopressin, in part through regulation of the transcription of the aquaporin-2 gene (Aqp2).

Methods To identify enhancer regions likely to be involved in the regulation of Aqp2 and other principal cell–specific genes, we used several next generation DNA-sequencing techniques in a well characterized cultured cell model of collecting duct principal cells (mpkCCD). To locate enhancers, we performed the assay for transposase-accessible chromatin using sequencing (ATAC-Seq) to identify accessible regions of DNA and integrated the data with data generated by chromatin immunoprecipitation followed by next generation DNA-sequencing (ChIP-Seq) for CCCTC binding factor (CTCF) binding, histone H3 lysine-27 acetylation, and RNA polymerase II.

Results We identified two high-probability enhancers centered 81 kb upstream and 5.8 kb downstream from the Aqp2 transcriptional start site. Motif analysis of these regions and the Aqp2 promoter identified several potential transcription factor binding sites, including sites for two b-ZIP transcription factors: CCAAT/enhancer binding protein-β (C/EBPβ) and cAMP-responsive element binding protein (CREB). To identify genomic binding sites for both, we conducted ChIP-Seq using well characterized antibodies. In the presence of vasopressin, C/EBPβ, a pioneer transcription factor critical to cell-specific gene expression, bound strongly at the identified enhancer downstream from Aqp2. However, over multiple replicates, we found no detectable CREB binding sites within 390 kb of Aqp2. Thus, any role for CREB in the regulation of Aqp2 gene transcription is likely to be indirect.

Conclusions The analysis identified two enhancer regions pertinent to transcriptional regulation of the Aqp2 gene and showed C/EBPβ (but not CREB) binding.


Vasopressin regulates water excretion largely through control of water permeability in the renal collecting duct. At a molecular level, water transport is controlled through regulation of the water channel protein, aquaporin-2 (AQP2). AQP2 is regulated in at least two ways: (1) control of trafficking of AQP2 to the apical plasma membrane of collecting duct principal cells1,2 and (2) control of the absolute abundance of the AQP2 protein.3 The latter process results largely from the regulation of Aqp2 gene transcription,4–7 although regulation of AQP2 protein half life also plays a role.8–10 Defective regulation of Aqp2 gene transcription has been found to be responsible for many acquired water balance disorders,11 including polyuric disorders (e.g., lithium-induced nephrogenic diabetes

Received May 19, 2017. Accepted February 18, 2018.

Published online ahead of print. Publication date available at www.jasn.org

Correspondence: Dr. Mark A. Knepper, Epithelial Systems Biology Laboratory, Systems Biology Center, National Heart, Lung, and Blood Institute, National Institutes of Health, 10 Center Drive, MSC-1603, Bethesda, MD 20892-1603. Email: knep@helix.nih.gov

Copyright © 2018 by the American Society of Nephrology
insipidus, hypercalcemic polyuria, hypokalemic polyuria) and dilutional hyponatremia as seen with the syndrome of inappropriate antidiuresis and congestive heart failure. To understand the pathophysiologic basis of these disorders, it is important to discover the normal physiologic mechanisms by which vasopressin, working by binding and activation of the vasopressin V2 receptor, stimulates Aqp2 gene transcription.

The V2 receptor couples to the heterotrimeric G-protein α-subunit GoS, which in collecting duct principal cells, activates adenylyl cyclase 6, thereby increasing intracellular cAMP levels. Actions of cAMP are mediated chiefly through activation of one or both protein kinase A catalytic proteins (coded by Pkaca and Prkacb genes). When these two genes were deleted from mpkCCD cells using CRISPR-Cas9, AQP2 mRNA and protein were nearly undetectable even in the presence of vasopressin, suggesting that Aqp2 gene transcription depends on protein kinase A.12

The mouse mpkCCD cell line has been used extensively in the investigation of regulation of collecting duct function.6,13–15 It manifests both major elements of vasopressin-mediated regulation of AQP2, namely regulated AQP2 trafficking and regulated Aqp2 gene transcription.15 On the basis of RNA-Seq and RNA polymerase II chromatin immunoprecipitation followed by DNA sequencing (ChIP-Seq) in mpkCCD cells, it seems that vasopressin mediates highly selective regulation of Aqp2 gene transcription, requiring the joint action of at least four DNA binding transcription factors.4 An intermediate goal on the way to an understanding of the mechanism of vasopressin-mediated Aqp2 gene transcription is to identify these transcription factors. A number of hypothesis-driven studies have proposed roles for a variety of transcription factors, including cAMP-responsive element binding (CREB) protein,16,17 AP-1,16 NF-κB,18 GATA family,19–21 ETS family,15,22 NFAT5, and NFATc proteins.23–25 However, a general solution to the problem is challenging, because the mammalian genome codes for approximately 1500 transcription factors and a trial and error approach are unlikely to be definitive. In this paper, we present a systems-level analysis using both prior data (histone H3 acetylation at lysine 27 [H3K27Ac] ChIP-Seq and RNA polymerase II ChIP-Seq) and new data from assay for transposase-accessible chromatin using sequencing (ATAC-Seq) experiments to identify enhancers near the Aqp2 gene. Enhancers are segments of DNA that are typically a few hundred base pairs in length, and they are occupied by multiple transcription factors that can recruit RNA polymerase II and transcriptional coactivators to target genes.26 ATAC-Seq is a new methodology designed to precisely identify regions of DNA that are open, allowing access to transcription factors and other DNA binding proteins.27 After enhancer identification, transcription factor ChIP-Seq experiments focused on two b-ZIP transcription factors, CCAAT/enhancer binding protein-β (C/EBPβ) and CREB, that are predicted to bind to either the identified enhancers or the Aqp2 promoter.

Significance Statement
Renal water excretion is controlled by vasopressin, in part through transcriptional regulation of the Aqp2 gene in collecting duct cells. Defects in this mechanism are responsible for several water balance disorders. Using mouse collecting duct cells, the authors used several next generation sequencing (NGS) techniques to investigate Aqp2 regulation. These approaches identified two enhancer regions vicinal to the Aqp2 gene centered 81 kb upstream and 5.8 kb downstream from the Aqp2 transcriptional start site. ChIP-Seq analysis for the transcription factor C/EBPβ identified a strong binding site within the second region. ChIP-Seq analysis for another transcription factor CREB, which widely held to regulate Aqp2 gene transcription, identified many canonical CREB binding sites throughout the genome but none within 390 kb of Aqp2.

METHODS

Cell Culture
AQP2-expressing mpkCCD clone 11 (mpkCCD) cells were grown on permeable membrane supports as described previously.15 Cells were exposed to 0.1 nM [deaminos-Cys1, D-Arg8]-Vasopressin (dDAVP) or vehicle for 30 minutes.

Preparation of Nuclear and Cytosolic Fractions
Nuclear and cytosolic fractions were prepared from mpkCCD cells using the NE-PER Nuclear and Cytoplasmic Extraction Reagent Kit (Pierce) as described previously.28

Immunoblotting
Immunoblotting was performed as described.29 Primary antibodies were CREB (17–600; Millipore), C/EBPβ (sc-150; Santa Cruz), p-CREB (S133, 17–10131; Millipore), ATF1 (sc-243; Santa Cruz), β-catenin (9562; Cell Signaling), and p-β-catenin (S552, 9566; Cell Signaling).

ATAC-Seq
ATAC-Seq was performed following a method published by Buenrostro et al.27 Nuclear pellets (from 50,000 cells) were incubated in “transposition reaction mixture” using Tagment DNA Buffer (FC-121–1030; Illumina) and Tn5 transposase (FC-121–1030) for 30 minutes at 37°C. The transposed DNA fragments were purified and amplified using PCR. Purified DNA libraries were sequenced on an Illumina HiSeq 3000 platform.

ChIP-Seq
Chromatin immunoprecipitation (ChIP) was performed using the truChIP Chromatin Shearing Reagent Kit (Covaris) and the SimpleChIP Enzymatic Chromatin IP Kit (Cell Signaling) following the manufacturer’s protocols. Protein/DNA was crosslinked using 1.1% formaldehyde. Nuclei were isolated and sheared using an ultrasonicator. The sheared chromatin was incubated with 4 μg of antibody (CREB, 17–600 or C/EBPβ, sc-150) and additionally incubated with Protein G magnetic beads. After elution of ChIP products, DNA was purified and subjected to library preparation using the Ovation SP Ultralow Library System (NuGEN) following...
the manufacturer’s protocol. The DNA libraries were gel purified and sequenced on an Illumina HiSeq 2000 platform.

**Data Processing**

The sequence reads were mapped on the mouse reference genome (mm10) using Burrows–Wheeler Aligner and Samtools. All uniquely mapped reads were used in downstream peak calling (MACS2; Bedtools), motif analysis (MEME-ChIP), and bioinformatic analysis (R package ChIPpeakAnno).

For ATAC-Seq data processing, adapter sequences were trimmed for both forward and reverse reads using cutadapt and aligned on mouse reference genome mm10 using Bowtie2. Duplicates were removed from mapped reads using Picard. Peaks were identified using MACS2. Read distribution was visualized on the UCSC genome browser using UCSC bedClip and bedGraphToBigWig.

**Data Mining and Bioinformatic Analyses**

A dataset of CCCTC binding factor (CTCF) ChIP-Seq was downloaded from the mouse ENCODE database (GSE36027). mpkCCD ChIP-Seq datasets were downloaded from GSE95009 for ATAC-Seq, H3K27Ac ChIP-Seq, and RNA polymerase II ChIP-Seq were integrated by multiplying the complements of their minimum Bayes factors.

RESULTS

Transcription factor proteins that regulate transcription of Aqp2 and other genes in the renal collecting duct may bind directly to gene promoters or enhancers that are not immediately adjacent to the regulated gene. Recent studies summarized by Hnisz et al. provide information that narrows the
possible locations of transcription factor binding sites that regulate a particular gene, such as \textit{Aqp2}. This constraint is on the basis of the concept of CTCF-bounded loops that define insulated neighborhoods. CTCF is a DNA binding protein that binds to DNA as a monomer during interphase but homodimerizes with upstream or downstream DNA-bound CTCF molecules to create CTCF loops (Figure 1A). The locations of these loops are highly conserved among mammalian species and relatively invariant among cell types.\textsuperscript{36} There are approximately 13,000 such insulated loops across the genome, with an average of two to three genes per insulated loop. The enhancers that regulate a given gene are found within the same CTCF loop. Prior studies done as part of the mouse ENCODE project have used ChIP-Seq to identify CTCF binding sites.\textsuperscript{34} These data provide the information needed to identify the insulated loop that contains \textit{Aqp2} (Figure 1B). As expected, CTCF binding sites were the same among kidney, liver, lung, and small intestine. The loop that contains \textit{Aqp2} is 99 kb in length and also contains two other genes: \textit{Faim2} and \textit{Aqp5}. Thus, enhancers that control \textit{Aqp2} gene transcription are highly likely to reside within this 99-kb insulated loop.

To assure that the mpkCCD cells used in this study respond appropriately to vasopressin, we carried out Western blotting for known vasopressin-regulated phosphorylation sites. We tested the ability of dDAVP (0.1 nM for 30 minutes) to increase phosphorylation of the transcriptional coregulator \(\beta\)-catenin at Ser552, a vasopressin-responsive site\textsuperscript{12,28,37} (Figure 2). Band density corresponding to this target was more than doubled. In the same samples, we probed for protein kinase A target sites on CREB and ATF1, which are recognized by the same antibody but can be discriminated on the basis of molecular weight. Both sites showed increases in phosphorylation in response to vasopressin (Figure 2), confirming the viability and vasopressin responsiveness of the cells.

We previously found evidence that 714 distinct transcription factors are expressed in mpkCCD cells\textsuperscript{12} (Supplemental Dataset 1). All could be considered candidates to play roles in regulation of \textit{Aqp2} gene transcription. However, 714 transcription factors are too many to allow efficient one by one evaluation by ChIP-Seq. To narrow the possible candidates, it

![Figure 3](https://www.jasn.org)  
**Figure 3.** Mapping ChIP-Seq data to three collecting duct genes: (A) adenylyl cyclase 6 (\textit{Adcy6}), (B) V2 vasopressin receptor (\textit{Avpr2}), and (C) aquaporin-2 (\textit{Aqp2}). The ATAC-Seq sites indicating accessible DNA are indicated by the track MACS2 Peaks. (General feature of ATAC-Seq analysis are in Supplemental Figure 1.) Read distribution normalized by signal per million reads was visualized on the UCSC genome browser using two-reference genome information: mouse mm9 for CTCF ChIP-Seq in kidney downloaded from mouse ENCODE database and mouse mm10 for ATAc-Seq, H3K27Ac ChIP-Seq, and RNA polymerase II ChIP-Seq in mpkCCD cells. Gray shading indicates the region corresponding to gene body. dDAVP, [deamino-Cys1, D-Arg8]-Vasopressin. CTCF, CCCTC binding factor; ATAC-Seq, assay for transposase-accessible chromatin using sequencing; H3K27Ac, histone H3 acetylation at lysine 27; ChIP-Seq, chromatin immunoprecipitation followed by DNA sequencing; RNA PolII, RNA polymerase II.

www.jasn.org | BASIC RESEARCH

\textbf{J Am Soc Nephrol 29:} 1490–1500, 2018 Genomic Binding Sites for CREB and C/EBPβ
is desirable to identify subregions of the *Faim2-Aqp2-Aqp5* insulated loop that are likely to contain enhancers. Doing so will allow us to identify potential transcription factor binding sites in these regions. We present three types of data that are informative in this regard, viz. ATAC-Seq, histone H3K27Ac ChIP-Seq, and RNA polymerase II ChIP-Seq (explained below). Relevant profiles for these measures in mpkCCD cells are shown in Figure 3.

ATAC-Seq is a recently devised method for assessing DNA accessibility, providing better spatial resolution than DNase I hypersensitivity profiling. Figure 3 shows ATAC-Seq data for mpkCCD cells mapped to the regions around three collecting duct genes: *Adcy6*, *Avpr2*, and *Aqp2*. *Adcy6* and *Avpr2* are strongly expressed in collecting duct principal cells, but their transcription is not regulated by vasopressin. Mouse ENCODE data for CTCF ChIP-Seq are included to provide a reference. The 37-kb putative CTCF insulated region surrounding *Adcy6* encloses a major ATAC-Seq peak centered 1200 bases upstream from *Adcy6* plus several less prominent peaks that are indicated by the “MACS2 Peaks” track. The 97.5-kb putative CTCF insulated region surrounding *Avpr2* encloses three discernable ATAC-Seq peaks centered at 11,300, 43,800, and 67,200 bases upstream from *Avpr2*. The *Faim2-Aqp2-Aqp5* insulated loop contains a major peak overlapping the unexpressed *Faim2* gene 81 kb upstream from the *Aqp2* transcriptional start site. It also contains several less prominent peaks between *Faim2* and *Aqp2* as well as between *Aqp2* and *Aqp5*. Full ATAC-Seq data can be downloaded from GEO (GSE108786).

Figure 3 also shows data from H3K27Ac ChIP-Seq experiments for the same three regions. (A small subset of these data was previously reported in the work by Isobe et al.) H3K27Ac mapping is widely used to identify open regions of DNA overlapping both promoters and enhancers. To a

---

**Figure 4.** Identification of enhancer elements for aquaporin-2 (*Aqp2*) gene. High open probability regions in the *Faim2-Aqp2-Aqp5* insulated loop were identified using minimum Bayes factors to integrate ATAC-Seq, histone H3K27Ac ChIP-Seq, and RNA polymerase II ChIP-Seq data into single “open probability” (Po) values as described in Supplemental Methods. The calculation is on the basis of the signal-to-noise ratio for each measure at each point in the loop. Noise values for this calculation were estimated as two times the median read density value across the loop. (Supplemental Figures 2 and 3 show sensitivity analysis.) The Po calculation identified two regions of high probability as indicated. These regions were analyzed using MEME-ChIP to identify JASPAR transcription factor motifs present. CTCF, CCCTC binding factor; dDAVP, [deamino-Cys1, D-Arg8]-Vasopressin. ATAC-Seq, assay for transposase-accessible chromatin using sequencing; H3K27Ac, histone H3 acetylation at lysine 27; ChIP-Seq, chromatin immunoprecipitation followed by DNA sequencing; RNA PolII, RNA polymerase II.
large extent, the H3K27Ac peaks overlap the ATAC-Seq peaks, providing redundancy for identification of potential enhancers. In response to dDAVP, there was a demonstrable increase in H3K27 acetylation in the Aqp2 promoter, gene body, and downstream from the gene body.

Figure 3 also shows binding sites for RNA polymerase II (Polr2a subunit) mined from Sandoval et al. RNA polymerase II has been shown to localize not only to gene bodies of actively transcribed genes but also, to active enhancers. dDAVP increases RNA polymerase II occupancy over the Aqp2 gene body as previously shown and also in two other regions of the Faim2-Aqp2-Aqp5 insulated loop, namely a region just downstream from Aqp2 and a region overlapping the last exon of Faim2. (Of note, the Faim2 mRNA is not detectable in mpkCCD cells.)

In Figure 4, we integrate the three measures (ATAC-Seq, histone H3K27Ac ChIP-Seq, and RNA polymerase II ChIP-Seq) to obtain an aggregate “open probability” (Po) indicating the regions in the Faim2-Aqp2-Aqp5 insulated loop that are most likely to be enhancers. Among the open regions, two stand out as being the densest and longest as highlighted in Figure 4. We analyzed these two enhancers using the JASPAR database of transcription factor binding motifs to identify DNA sequence motifs associated with specific transcription factors. In the upstream enhancer, potential sites for several transcription factors were identified, including some known to be expressed in mpkCCD cells viz. NFYA, RREB1, E2F3, YY1, CTCF, and SMAD4 (Supplemental Dataset 1). In the identified enhancer between the Aqp2 and Aqp5 genes, potential sites for the transcription factors TFAP2A, PBX3, CEBPB, CEBPG, NRF1, ELF1, ELF4, KLF9, GLIS2, NFIX, NFIA, and HIF1A were identified. Some of these have previously been implicated in regulation of Aqp2 gene transcription: AP2 (TFAP2A), ETS family (ELF1 and ELF4), and KLF (KLF9). Among the sites found, C/EBPβ is a so-called “pioneer” transcription factor that can initiate regulated changes in chromatin structure before binding of other transcription factors.

To identify whether C/EBPβ binds anywhere in the Faim2-Aqp2-Aqp5 insulated loop or elsewhere in the genome (Supplemental Figure 4), we carried out ChIP-Seq experiments (Figure 5) using a well tested C/EBPβ antibody. The full datasets are provided permanently at https://hpcwebapps.cit.nih.gov/ESBL/Database/ChIP-Seq_CREB_CEBPB/ and the GEO (GSE98076). As shown in Figure 5, 2401 C/EBPβ binding sites were identified in dDAVP-treated cells (0.1 nM, 30 minutes), whereas 209 reproducible sites were found in vehicle-treated cells. Read densities for many of the 209 sites common to the dDAVP and vehicle samples were increased in the dDAVP-treated versus vehicle-treated cells: mean log2(dDAVP/vehicle) =1.14 (Figure 5A) as expected on the basis of the observation that dDAVP causes translocation of C/EBPβ into the nucleus of mpkCCD cells (Supplemental Figure 4) (https://hpcwebapps.cit.nih.gov/ESBL/Database/QuantNucProteomics/). In mpkCCD cells, C/EBPβ binding sites are not limited to gene promoters, but they are also frequently found in intergenic regions and introns (Figure 5B), contrasting with CREB (see below). The consensus motif identified from C/EBPβ binding sites matched canonical binding motifs for C/EBPβ.
the C/EBP family. These data support the fidelity of the ChIP-Seq methodology applied to C/EBPβ binding site localization. Vasopressin-sensitive C/EBPβ binding sites were identified in the vicinity of several known C/EBPβ-regulated genes, namely Hif1a,43 Itpr2,44 Smg7,42 and Idh145 (Figure 6, A–E), further supporting the fidelity of the method.

A C/EBPβ binding site was identified just 400 bp downstream from Aqp2, and the read density appeared to be increased by dDAVP (Figure 6F). This corresponds to a predicted C/EBPβ binding site in one of the high open probability DNA regions identified in Figure 4. The same site was identified in two additional samples treated for 24 hours with dDAVP (Supplemental Figure 5). Supplemental Tables 1 and 2 list all transcription factors with C/EBPβ binding sites at their promoters as well as protein kinase and transporter genes with C/EBPβ binding sites (Supplemental Figure 6).

The fraction of C/EBPβ binding sites associated with collecting duct–specific genes was greater than in the whole transcriptome, suggesting a role in tissue-specific gene expression (Supplemental Figure 7). Tissue-specific genes tend to have TATA boxes rather than CpG islands in their promoter proximal regions.46,47 Consistent with this, genes with TATA boxes are more frequent among genes with C/EBPβ binding sites than would be expected from a random selection of expressed genes, whereas genes with promoter proximal CpG islands are less frequent (Supplemental Figure 7). These observations lend support to the possible role of C/EBPβ in cell-specific gene expression in collecting duct principal cells.

Figure 6. C/EBPβ ChIP-Seq: examples of binding sites found. Read distribution normalized by signal per million reads was visualized on the UCSC genome browser using two-reference genomes: mouse mm9 for CTCF ChIP-Seq in kidney downloaded from the mouse ENCODE database and mouse mm10 for C/EBPβ ChIP-Seq in mpkCCD cells. Gray shading indicates the region corresponding to gene body. (A–E) Results confirmed previous findings for Hif1a, Itpr2, Smg7, Idh1, and Cebpb. (F) A C/EBPβ binding site was identified just 400 bp downstream from the aquaporin-2 (Aqp2) gene in the dDAVP–treated sample. C/EBPβ, CCAAT/enhancer binding protein-β; CTCF, CCCTC binding factor; dDAVP, [deamino-Cys1, D-Arg8]-Vasopressin; ChIP-Seq, chromatin immunoprecipitation followed by DNA sequencing.
Because vasopressin signals chiefly through increases in cAMP in collecting duct principal cells, it has been widely assumed that the transcription factor CREB is responsible for vasopressin-mediated regulation of Aqp2 gene transcription.\textsuperscript{11,48–51} A cAMP response element has been reported to be present within the promoter of Aqp2,\textsuperscript{16,17} although CREB binding to this site has not been directly shown. To identify CREB binding sites throughout the genome, we have carried out ChIP-Seq experiments in mpkCCD cells (four control-dDAVP pairs done on different days).

For ChIP-Seq, we used a CREB antibody previously used for ChIP-Seq in other tissues\textsuperscript{52} (Supplemental Figure 4). Binding sites were identified as genomic positions in which mapped reads for ChIP samples significantly exceeded the input control using a stringent cutoff ($q<0.01$ using MACS2). The full datasets are provided at https://hpccwebapps.cit.nih.gov/ESBL/Database/ChIP-Seq_CREB_CEBPB/ and at GEO (GSE98076). A total of 793 CREB binding sites were identified in mpkCCD, and most of them were seen under both conditions. In general, normalized read densities were virtually unchanged or slightly reduced in the dDAVP-treated cells (0.1 nM for 30 minutes) versus vehicle-treated cells (Figure 7A). This finding is consistent with prior studies showing that CREB is regulated chiefly by changes in phosphorylation in its “pKID domain” rather than by differential binding or nuclear translocation.\textsuperscript{53,54} Over 90% of total CREB binding sites were located at promoter regions (1 kb upstream to +100 bp downstream from the transcription start site) (Figure 7B). Sequences extracted from binding site peak centers were used to identify binding motifs (Figure 7C). A consensus half-cAMP response element sequence was identified that matches prior data for b-ZIP transcription factors CREB3, ATF1, and ATF4, confirming the fidelity of the method. The motif is the same for vehicle- and dDAVP-treated cells, although shifted by three bases. Motifs matched previously reported motifs for CREB3, ATF1, and ATF4 found in the JASPAR motif database; e values in parentheses are the motif P values (computed by a Fisher exact test for enrichment of the motif in the sequences) times the number of candidate motifs tested. CREB, cAMP-responsive element binding protein; dDAVP, [deaminom-Cys1, D-Arg8]-Vasopressin; ChIP-Seq, chromatin immunoprecipitation followed by DNA sequencing.

**DISCUSSION**

Studies in model organisms have shown a high level of complexity in transcriptional networks with multiple cross-regulated transcription factors affecting transcription.\textsuperscript{58} Sandoval et al.\textsuperscript{4} have estimated that a minimum of four transcription factors are required to explain selective regulation of Aqp2 gene transcription in collecting duct cells. mpkCCD cells have been found to express >700 transcription factors,
making a protein by protein identification of relevant transcription factors using ChIP-Seq and CRISPR-Cas9 deletions impractical. Here, we have used an unbiased, systems-level approach using several next generation sequencing modalities to identify likely enhancers within the same insulated neighborhood as \textit{Aqp2}. These \textit{Aqp2}-associated enhancers are likely to bind the transcription factors involved in cell type–specific expression and regulation by vasopressin. Motif analysis of the enhancers identified relatively few transcription factors, thus reducing the number of candidates that will require investigation in future studies (Figure 4). We have carried out transcription factor ChIP-Seq for two candidates, \textit{viz.} CREB and C/EBP\(\beta\), identifying a strong binding site for the latter in one of the identified enhancer regions. We provide comprehensive genome-wide binding data as an online resource for investigators interested in regulation of genes other than \textit{Aqp2}. Future studies can use next generation sequencing techniques, advanced reporter systems, and CRISPR-Cas9 to examine roles of all transcription factor candidates in the \textit{Aqp2} gene regulatory network.

\section*{ACKNOWLEDGMENTS}

Next generation sequencing was done in the National Heart, Lung and Blood Institute DNA Sequencing Core Facility (Yuesheng Li, Director). The work was primarily funded by Division of Intramural Research, National Heart, Lung, and Blood Institute projects ZIA-HL001285 (to M.A.K.) and ZIA-HL006129 (to M.A.K.). All authors approved the final version of the manuscript.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure8}
\caption{CREB ChIP-Seq: examples of binding sites found. Read distribution normalized by signal per million reads was visualized on the UCSC genome browser using two-reference genomes: mouse mm9 for CTCF ChIP-Seq in kidney downloaded from the mouse ENCODE database and mouse mm10 for CREB ChIP-Seq in mpkCCD cells. Gray shading indicates the region corresponding to gene body. (A–E) Results confirmed previous findings for \textit{Fos}, \textit{Mcm5}, \textit{Cbwd1}, \textit{Jund}, and \textit{Creb}. (F) No CREB binding sites were found within the Faim2-aquaporin-2 (\textit{Aqp2})-\textit{Aqp5} insulated loop. (The closest CREB binding site from the \textit{Aqp2} gene was 390 kb upstream.) CREB, cAMP-responsive element binding protein; CTCF, CCCTC binding factor; dDAVP, [deamino-Cys1, D-Arg8]-Vasopressin; ChIP-Seq, chromatin immunoprecipitation followed by DNA sequencing.}
\end{figure}
DISCLOSURES
None.

REFERENCES


www.jasn.org | BASIC RESEARCH

J Am Soc Nephrol 29: 1490–1500, 2018
Genomic Binding Sites for CREB and C/EBPβ

1499

This article contains supplemental material online at http://jasn.asnjournals.org/lookup/suppl/doi:10.1681/ASN.2017050545/-/DCSupplemental.