Toward Molecularly Targeted Therapeutics for Renal Disease

by

Edward Charles Grimley

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Molecular and Cellular Pathology)
In the University of Michigan
2017

Doctoral Committee:

Professor Gregory R. Dressler, Chair
Associate Professor Scott E. Barolo
Professor Eric R. Fearon
Associate Professor Maria (Ken) E. Figueroa, University of Miami
Associate Professor Zaneta Nikolovska-Coleska

Edward Charles Grimley

egrimley@umich.edu

ORCID iD: 0000-0003-0438-506X

Dedication

To my parents, my wife, and my children who made me feel like this was possible and for whom I strive to be my very best.

Acknowledgements

I would like to especially thank my esteemed mentor, Greg Dressler, for his guidance, support, and wisdom. I am both grateful and honored to have had the opportunity to be a member of his lab. He has inspired me to think creatively, fostered my intellectual growth, and has always challenged me to live up to my potential as a developing scientist. While much of my scientific success is a direct result of Greg's skillful supervision, he has also played a significant role in my personal development. He has consistently been there to lend an ear and provide encouragement. I appreciate all that he has done, and I aspire to grow into a scientist of his ilk.

I would like to thank former and current members of the Dressler lab. I am particularly grateful for Egon Ranghini and Abdul Soofi who played instrumental roles in my technical training and who provided me with their friendship. These two made every day enjoyable and were always there to talk things through.

I have been fortunate to have had Scott Barolo, Eric Fearon, Ken Figueroa, and Zaneta Nikolovska-Coleska serving on my thesis committee. I tremendously appreciate their time and insight, which has contributed considerably to the quality of my thesis.

I would like to extend my thanks to the many others who have contributed to this work and my training. This includes, but is not limited to, Adam Stein, Sanj Patel, Puneet Garg, and the members of their labs for providing insightful conversation

regarding this work and ways to improve it. I am thankful for Chenzhong Liao, a postdoctoral fellow in Zaneta Nikolovska-Coleska's lab, for his expertise in setting up and running the virtual screening campaign. Martha Larson from the University of Michigan Center for Chemical Genomics has been invaluable for her guidance and support with developing the HTS campaign. Scott Larson and Andy White provided medicinal chemistry support for which I am thankful. Also, thank you to Laura Labut from the MCP program for her assistance.

I owe a great deal of my success to my grandma for cultivating my curiosity and desire to learn. When I was growing up she seemed to have new projects for me to do every time that I would visit. When she ran out of projects we would take trips to a hobby shop, visit a nature center, or go to a museum. Our time together helped to form my interest in science. She has always believed in me and assured me that I could do anything. I am grateful to have had such a wonderful person in my life.

My parents, Edward and Teresa, have always done their best to set a great example for me. They have worked tirelessly to see myself and my siblings do well in life. They have shown me that with dedication and a little hard work anything is possible. They have been an endless supply of love and encouragement, and I truly appreciate everything that they have done to get me to this point.

To my children, Preston and Ethan, I want you to know how proud you make me. You guys inspire me and fill me with hope. Your patience and encouragement

throughout my educational career has been tremendous and without it none of this would have been possible. You guys are awesome and I love you both!

Finally, my wife, Lynn, has been by my side for each and every step of this journey. Having her with me has provided comfort and relief in times of stress. She has always believed in me, even when I was not so sure of myself. It was often Lynn who kept me on track, usually with gentle reminders. I am grateful for her outstanding editing skills, but most of all I am thankful for her undying love. There is no better partner to have been on this journey with. I love you, Lynn!

Table of Contents

List of Tables

Table

List of Figures

Figure

Abstract

Pax2 is a developmental control gene that is essential for organogenesis of the kidney and urogenital tract. Genes of this nature are generally suppressed in healthy adult tissues. However, developmental genes and pathways are frequently reactivated in renal diseases like cancer or polycystic kidney disease. Furthermore, following acute renal injury, regenerating renal epithelial cells require reactivation of embryonic genes to drive proliferation and differentiation of the new tubular cells. Throughout this thesis, knowledge of Pax genes in renal development and disease is leveraged to generate innovative therapeutic interventions for an assortment of kidney diseases. Using a novel reporter cell line, that responds to Pax2 to drive transcription of luciferase, I developed an assay compatible with high-throughput screening (HTS). An HTS campaign of 69,125 unique small-molecules yielded 48 molecules that negatively regulate Pax2 mediated transcription activation and two molecules that positively regulate this process. In collaboration with Dr. Liao from Dr. Nikolovska-Coleska's lab, a virtual screening campaign based on homology modeling was developed to identify small-molecules with the potential for binding to the Pax2 DNA-binding domain. This campaign yielded 227 candidate molecules that were subsequently examined in the Pax2 reporter assay. Following confirmation, counterscreen, and titration assays, five molecules remained. Analogs of these molecules yielded a more potent compound, EG1. This molecule was further characterized and shown to inhibit Pax2 DNA-binding through a direct interaction

with the paired domain. Furthermore, EG1 treatment recapitulates aspects of Pax2 loss of function in an embryonic organ culture model system. Embryonic kidneys grown in the presence of EG1 exhibit defects in branching morphogenesis and have gene expression changes consistent with loss of Pax2 activity. Embryonic lungs, which undergo a Pax independent branching morphogenesis, were unaffected by EG1. Similarly, Pax2 negative cancer cells were unaffected by EG1 treatment while Pax2 positive cancer cells exhibited a significant decrease in viability. These findings suggest that I have identified a *bona fide* small-molecule regulator of Pax2 transcription activation. Furthermore, my findings represent an important proof of concept, demonstrating that developmental regulatory transcription factors, like Pax2, can be targeted by small-molecules and that doing so may provide an avenue for developing novel therapeutic approaches.

CHAPTER 1

Introduction

Abstract

Cellular phenotypes are largely determined by the set of genes that a cell expresses. This makes regulating transcriptional programs a vital part of cellular fate determination. Transcriptional regulation is an intricate and tightly controlled process. It is multifaceted and includes epigenetic processes such as DNA methylation and post-translational modification of histone proteins. Transcription factors often provide the link between signaling pathways and transcriptional regulation. They are DNA-binding proteins that can provide locus specificity to complexes of proteins with the ability to reorganize chromatin such that particular genes are expressed or silenced accordingly. The Pax family of transcription factors represents a critical class of developmental regulatory proteins. While Pax proteins are indispensable for a host of developmental processes, aberrant expression of these factors often drive pathogenesis. For example, Pax2 is essential for nephrogenesis but the persistent expression of Pax2 is implicated in several cystic disorders of the kidney and a variety of renal cancers. Despite being implicated in countless disease states, the Pax proteins have not been investigated as therapeutic targets although several lines of evidence suggest that they may, in fact, be excellent targets. Targeting tissue-specific developmental control genes represents a

unique therapeutic approach with tremendous potential for improving clinical outcomes while reducing harmful off-target effects.

Introduction

Pax genes are critical developmental regulators and are frequently dysregulated in cancer

The mammalian Pax gene family is a nine-member (*Pax1-9*) collection of developmental regulatory gene that are expressed in a variety of tissues during embryogenesis. The precise timing and positioning of Pax gene expression is tightly regulated, with the majority occurring during embryonic development. However, in the stem cell compartments of various tissues Pax gene expression continues into adulthood where these genes contribute to stem cell maintenance and tissue regeneration [1-4]. Pax genes are intimately involved with the patterning of the developing organism and the progressive regionalization of body tissues. These genes are frequently essential for proper organogenesis of tissues in which they are expressed. For example, members of the group II subfamily (Pax2/5/8) are necessary for organogenesis of the eye, ear, kidney, thyroid, CNS, and B lymphocytes [5-9]. Except for *Pax1* and *Pax9*, all of the Pax genes are involved in development and specification of the central nervous system (CNS) [2, 10]. The role of Pax genes in the CNS is beyond the scope of this thesis but is presented elsewhere [2, 4, 11].

Development of complex organisms requires intricately coordinated methods for defining body segmentation patterns and constructing functional tissues. These methods utilize fundamental biological processes including proliferation, migration, apoptosis, lineage specification, and differentiation. When executed properly these processes synchronize to form a functional organism; however, dysregulation of these processes is often the driving force behind diseases like cancer [12, 13]. Notably, the Pax family of

transcription factors have been implicated in regulating many of these processes and their misregulation has been linked to an assortment of neoplasias.

The aberrant expression of Pax2 is seen in rapidly proliferating renal epithelial cells, such as in polycystic kidney disease (PKD) [14, 15], juvenile cystic and dysplastic diseases [16], Wilms' tumor [17, 18], and renal cell carcinoma (RCC) [19-21]. Pax2 has been described as a pro-survival factor that fuels proliferation and provides tumor cells with resistance to programmed cell death. Thus, decreased Pax2 gene dosage proved beneficial for reducing cyst formation and slowing PKD progression in mouse models [14, 15]. Moreover, following suppression of Pax2 gene expression renal cancer cell lines are sensitized to chemotherapeutic agents, exhibit an increase in apoptosis, and have a reduced proliferative capacity [19, 22, 23].

Pax5 is abnormally expressed in leukemias such as Hodgkin's lymphoma, non-Hodgkin's lymphoma, follicular lymphoma, B-cell acute lymphoblastic leukemia (B-ALL), and hairy cell leukemia [24]. Pax5 is the most commonly mutated gene in B-ALL [25]. The majority of these mutations result in reduced levels of functional Pax5 protein. These findings suggest that Pax5 plays a tumor suppressor role in B-ALL. In contrast, non-Hodgkin's lymphomas almost universally show high levels of Pax5 expression [26, 27]. This can partially be explained by a recurrent translocation, t(9;14)(p13;q32), that places the *Pax5* gene under the control of the immunoglobulin heavy-chain (*IGH*) enhancers and/or promoter [28-30]. Additionally, several solid tumors exhibit Pax5 expression. A malignant subset of neuroblastoma cells (N-type) is one such example. While Pax5 expression is not seen in the benign subset of neuroblastoma cells (S-type), inducing its expression in these cells causes them to adopt an N-type phenotype [31]. S-type cells

expressing Pax5 show increased cell division and gain colony forming abilities whereas knockdown of Pax5 in N-type cells significantly slows proliferation [31], suggesting that Pax5 acts as an oncogene in tumors of this type.

Like Pax2, Pax8 is frequently expressed at high levels in cancers of the kidney [32, 33]. Pax8 is also expressed in most thyroid cancers and is associated with increased risk of recurrence [34-36]. Additionally, an oncogenic fusion protein that results from a chromosomal translocation, t(2;3)(q13;p35), which combines the *Pax8* and peroxisome proliferator-activated receptor gamma (*PPARγ*) genes, is observed in approximately 36% of follicular thyroid carcinomas and roughly 11% of follicular thyroid adenomas [37-40]. Expression of this fusion protein is associated with amplified cell cycling, enhanced cell growth, increased invasive potential, and reduced cell death [41, 42].

For the purposes of this thesis, I have only highlighted the subgroup II family members, but other Pax proteins are also implicated in oncogenesis. While the precise role of Pax proteins in cancer continues to be investigated, the data suggests that dysregulation of these proteins is a major component of pathogenesis. I was interested in exploring the potential of Pax proteins as therapeutic targets. I was particularly interested in Pax2 since the Dressler lab has acquired vast knowledge of this protein over the past 20 plus years and has developed numerous tools for investigating Pax2 activity in the kidney and in renal diseases that require better treatment options.

<u>Molecularly targeted therapeutics for renal diseases are desperately needed</u>

With up to 12.5 million individuals affected worldwide, autosomal dominant polycystic kidney disease (ADPKD) is the most commonly inherited disorder of the kidney

[43] and one of the leading causes of end-stage renal disease [44]. ADPDK is characterized by relentless proliferation of tubular epithelial cells. This uncontrolled proliferation leads to the formation of cysts, progressive kidney enlargement, and ultimately decline in kidney function. In addition to excessive proliferation, renal epithelial cells exhibit abnormal fluid secretion, have altered cell morphology, and make pathological interactions with the extracellular matrix [45]. Most human cases of PKD are caused by mutations in *Pkd1* or *Pkd2* genes; however, there are many other juvenile and adult onset diseases that result in renal cysts [46]. Characterization of *Pkd1/2* and their encoded Polycystin proteins identified defects in primary cilium as a major determinant for pathology. The Hedgehog, Wnt, cAMP, and planar cell polarity pathways are just a few of the pathways that are known to be regulated by cilia. Yet it remains unclear how these pathways contribute to PKD pathogenesis. One common feature of cystic epithelial cells in the kidney is persistent overexpression of Pax2. Interestingly, decreased Pax2 gene dosage proved beneficial for reducing cyst formation and slowing PKD progression in mouse models [14, 15]. These findings suggest that Pax2 could be a useful target for treatment of PKD for which treatment options remain quite limited. In fact, most PKD treatment regimens are aimed at reducing disease-associated complications such as hypertension, cyst infection, and chronic pain in order to increase the quality of life and extend lifespan [43]. Unfortunately, the net result of this approach is that roughly 70% of patients with ADPKD end up requiring renal transplant therapy sometime between 40 and 70 years of age [47]. Only recently has mechanistic knowledge led to the identification of novel targets for treatment of ADPKD. These targets include HMG-CoA reductase, mTOR, and somatostatin receptors. Perhaps the most promising data comes from a

large randomized clinical trial in which vasopressin V2 receptor antagonism by tolvaptan was found to slow the increase in total kidney volume and the decline in glomerular filtration rate [48]. Tolvaptan has since been approved for ADPKD treatment in Japan. However, the United States Food and Drug Administration raised concerns regarding 1) total kidney volume as an established surrogate, 2) the uncertainty introduced by missing data and a post-treatment baseline for the key secondary endpoint, and 3) the small improvement in renal function decline [49]. While the current group of mechanistically identified targets is encouraging, there remains a desperate need for better treatment options.

Likewise, current treatment regimens for RCC are insufficient. Worldwide, RCC accounts for 90% of kidney cancer cases with greater that 270,000 new patients diagnosed and 116,00 patients succumbing to their disease each year [50]. Close to 95% of RCCs have high levels of Pax2 expression and this expression is associated with metastasis [19, 20, 51-53]. Until very recently, the treatment of choice for RCC was immunotherapy using a combination of interferon alpha and interleukin-2 with high dose interleukin-2 as the only accepted option for patients presenting with metastasis [54]. Since then, inactivation of Von Hippel-Lindau (VHL), a tumor suppressor, was found to be present in the majority of clear cell RCCs [55]. Inactivation of VHL produces an increase in angiogenic factors including Vascular Endothelial Growth Factor (VEGF). The discovery of the VHL-VEGF link in these cancers revolutionized the field. Now, frontline therapies commonly include the VEGF inhibitors sunitinib, pazopanib, or bevacizumab [56]. While this change in treatment preference has proven benefits, rates of response are limited, patients frequently develop resistance, and the overall long-term survival rates

have not dramatically improved. Newly emerging options for second-line therapy include mTOR inhibitors (temsirolimus or everolimus), multi-tyrosine kinase inhibitors (sorafenib, axitinib, or cabozantinib), and programmed cell death protein 1 (PD-1) inhibition (nivolumab). Although each has shown benefits in phase III trials, the median overall survival was well below 3 years in any case [54]. The field has made tremendous progress recently; however, additional efforts are needed to greatly improve patient outcomes.

It is evident that molecularly targeted therapies are in desperate need for combating proliferative disorders of the kidney. Pax2 has been linked to a number of these disorders and has been shown to be intimately involved with survival, proliferation, and resistance to chemotherapy. Since Pax2 is a novel but well-validated target that is expressed in a cell-type specific manner, I hypothesize that it will be an outstanding candidate for developing targeted therapeutics against. To begin to address this need it is imperative to know as much as possible about the proposed target so to determine how best to attack it.

Pax genes characterized by an evolutionarily conserved DNA-binding domain known as the paired domain

Pax proteins are sequence-specific DNA-binding proteins that predominately bind to regulatory regions of the genome [57]. Such proteins were thought to be transcription factors and are classified by the structure of their DNA-binding domains. This classification has been valuable for determining how certain classes of transcription factors bind to DNA in a sequence-specific manner and has provided insight into how these factors have developed specialized functions throughout evolution [58]. A number

of DNA-binding domain superfamilies have been described. These include the helix-turn-helix (HTH), zinc-finger, β-scaffold, and basic domain classes of transcription factors [59]. The HTH superfamily represents an important class of transcription factors that includes critical developmental regulators like homeo, forkhead, E-twenty-six, and paired domain containing proteins.

The mammalian Pax gene family was originally identified based on an evolutionarily conserved DNA-binding domain termed the paired-box or paired domain (PD) [60]. All nine of the mammalian Pax genes encode an N-terminal PD, but downstream structural elements, including an octapeptide (OP) sequence and homeodomain (HD), which can be observed in various combinations amongst the family members, divide the family into four subgroups [61, 62]. Group I genes (*Pax1* and *Pax9*) have the OP sequence but lack the HD, Group II genes (*Pax2*, *Pax5*, and *Pax8*) have the OP sequence and a partial HD, Group III genes (*Pax3* and *Pax7*) have the OP sequence and a full HD, and Group IV genes (Pax4 and Pax6) lack the OP sequence but have a full homeodomain (Figure 1.1).

The structure of the paired domain and its nature of DNA-binding has been defined

The highly-conserved 128-amino acid PD, which confers DNA-binding ability, defines the Pax proteins. The structure of the *Drosophila* Prd [63], human Pax5 [64], and human Pax6 [65] PDs were elucidated by co-crystallization with their cognate DNA sequences (Figure 1.2). These studies revealed that the PD is composed of two globular subdomains that are connected by a disordered linker. Interestingly, both of the globular subdomains adopt a structural motif, helix-turn-helix (HTH), that has classically been

known as a DNA-binding motif. Additionally, the co-crystal structures indicate that the N-terminal HTH subdomain (PAI subdomain) and the C-terminal HTH subdomain (RED subdomain) make contacts with nucleotide bases in adjacent major grooves of the DNA double helix while the extended linker makes substantial contact with bases in the intervening minor groove. Nuclear magnetic resonance analysis of the Pax8 PD in the absence of DNA largely confirmed the structure observed in the previously mentioned x-ray crystallography studies [66]. Given the high degree of concordance within these structural determinations, it is likely that this structure is representative of the entire Pax family. However, it is important to note that the native states may vary among family members, as evidenced by the Pax8 PD being largely folded while the Pax6 PD was found to be primarily unfolded in the absence of DNA [66, 67]. While each of the HTH subdomains are capable of independently binding DNA [68], both are required for comprehensive regulation of target genes [69].

DNA-binding specificity of several Pax family members is confounded by the presence of a second DNA-binding domain known as the homeodomain

The importance of the PD subdomains in defining DNA-binding specificity is further complicated by the existence of an HD in some members of the Pax family. A number of Pax family members contain a full (Pax3, Pax4, Pax6, and Pax7) or partial HD (Pax2, Pax5, and Pax8). Crystal structures of the paired-type HD reveal that it is composed of 3 α-helices which adopt an HTH structural motif [70, 71]. Additionally, these structures show that the paired-type HD is quite similar to HDs of other members of the homeobox superfamily. One key difference appears to be the existence of dimer inducing

intramolecular interactions that are only present in the paired-type HD [70, 71]. Another key difference is the presence of a serine residue at amino acid position 50 which is a critical position for defining DNA-binding specificity within the homeobox superfamily [72, 73]. The HD is a sequence-specific DNA-binding domain which recognizes the TAAT motif. Due to the dimer forming capability of the paired-type HD, it recognizes palindromic sites containing two TAAT motifs that are separated by two to three base pairs [74]. While the functional significance of the HD has not been fully elucidated, it does appear to be tissue and context dependent. For example, the HD of Pax6b is required for lens formation but is expendable for endocrine pancreatic cell differentiation in zebrafish [75].

The two distinct binding modalities potentially provide versatility in Pax target gene regulation and may help to explain the diversity of developmental programs which are regulated by these transcription factors. DNA-binding by Pax proteins could be primarily mediated by either the PD or HD or could involve a combination of the two domains. In fact, targeted mutagenesis experiments looking at one domain or the other have identified examples of all three cases. Yet cases of one domain or the other being solely responsible for DNA-binding specificity appear to be in the minority. Indeed, both PD and HD binding motifs are frequently seen together in target gene regulatory regions indicating that DNA-binding is highly-dependent on cooperative interactions between the two domains [76-79]. Furthermore, *in vivo* analysis indicates both DNA-binding domains are required for complete functionality [80, 81]. Moreover, biochemical analysis along with numerous mutational studies have described the Pax PD/HD interdependence and have shown that alterations within the PD lead to structural changes in the HD and ultimately result in diminished binding efficiency [82-87]. For example, deletion of amino acids 29-34 within

the PD totally eliminates DNA-binding by both the PD and HD of Pax3 [87]. Taken together, these studies underscore the critical importance of the PD in shaping DNA-binding specificity in Pax family members with or without an HD.

Additional conserved regions contribute to Pax mediated transcriptional regulation through interactions with adapter proteins PTIP and GRG4

Transcription factors play a key role in regulating transcriptional programs. They bind to DNA in a sequence-specific manner and recruit coactivators, chromatin modifying complexes, and the transcriptional machinery to initiate transcription [88]. Alternatively, transcription factors may repress gene expression by directly interfering with the transcriptional machinery or through the recruitment of corepressor complexes. Such is the case with Pax2, which can regulate transcription positively or negatively depending on cellular context and the availability of cofactors. These activities appear to be regulated by two distinct regions of the Pax proteins.

The carboxy-terminal portion of the Pax proteins has been found to confer transcription activation potential to most of the Pax family members [17, 89-94]. This so-called transactivation domain contains regions rich in proline, serine, threonine, and tyrosine residues but is otherwise not well conserved [95]. Interestingly, this domain is critical for achieving maximum transcription activation, yet point mutations within this region of the Pax proteins generally have little effect on transactivation potential. This suggests that multiple regions within this domain contribute to transcription activation either through cooperative binding or through conformational effects. This domain tends to be less defined and has not been shown to contain any conserved structural motifs or

consensus sequences that would identify binding partners. However, a yeast two-hybrid assay with the Pax2 transactivation domain as the bait identified Pax Transactivation-Domain Interacting Protein (PTIP) as a binding partner [95].

Another conserved structural element is found in all but two members, Pax4 and Pax6, of the Pax family. This element, known as the octapeptide sequence, is an eight amino acid element that is located within the region between the two DNA-binding domains. The OP is evolutionarily related to the *Drosophila* engrailed family Eh1 repression domain, found in Engrailed proteins, and the Gsc-En homology element, found in Goosecoid proteins [96]. The Eh1 motif is a binding site for transcriptional corepressors and was found to interact with members of the Groucho-related protein (Grg)/Transducin-like enhancer of split (Tle) family [97]. Consistent with this finding, Pax2, Pax3, and Pax5 have been found to convert from transcriptional activators to transcriptional repressors by virtue of a direct interaction between Grg4 and the OP [98-101].

The interaction of Pax transcription factors with the adapter proteins PTIP and GRG4 link them to epigenetic complexes that regulate chromatin structure. The regulation of chromatin structure is a key mechanism of controlling gene expression and is discussed in more detail later.

Storage of DNA as chromatin within the nucleus represents a mechanism for controlling how it is utilized

Transcriptional regulation is a complex process that occurs at numerous levels, starting at the level of DNA structure. (Figure 1.3). In eukaryotes, DNA within the nucleus is maintained in a highly-structured state, known as chromatin. Nucleosomes, a 147 base-

pair stretch of DNA wrapped 1.65 turns around a core histone octamer, form the basic unit that chromatin is comprised of [102, 103]. The core histone octamer is made up of two units of each of the highly-conserved core histone proteins (H2A, H2B, H3, and H4) [102, 103]. The DNA, through electrostatic forces, interacts with and wraps around the histone octamer leading to compaction of the DNA [102, 104]. A linker histone, H1, binds to the nucleosome and intervening DNA sequence to promote even further condensation [105]. This spatial compaction of DNA promotes stability, reliable repair, fidelity in replication, and allows for the information encoded within the genome to be utilized in a controlled and organized manner [106, 107].

While the chromatin structure is critical for packing DNA into the nucleus, the presence of nucleosomes can restrict the transcriptional machinery from accessing the DNA [108, 109]. This suggests that in order for efficient transcription of genomic information to take place, chromatin structure must be altered such that the transcriptional machinery can gain access to the DNA [110]. Chromatin within a non-dividing cell can be classified as either euchromatin or heterochromatin based on how tightly it is packaged. Euchromatin is a region where the DNA is loosely packaged, easily accessible, and poised for transcriptional activity [111]. Conversely, heterochromatin is a region where DNA is highly-condensed, inaccessible, and transcriptionally inactive [111]. The presence of multiple chromatin states within a cell suggests the existence of a mechanism for controlling compaction/decompaction of genomic material. Moreover, it suggests that genomic output could be controlled by regulating access to certain genomic regions.

Epigenetic modifications alter chromatin structure to regulate gene expression and consequently, cellular phenotype

Structural alteration of the chromatin, or chromatin remodeling as it is otherwise known, can be regulated by post-translational modifications of histone tails or through methylation of DNA [112, 113]. DNA methylation of cytosine residues was the first and best characterized mechanism that correlated with gene expression and chromatin structure. However, post-translational modification of histone tails has been an area of intense interest for the past two decades. This interest has expanded our understanding of the many types of modifications and their associated functional consequences.

Functionally, histone and DNA modifications control gene expression and phenotype without impacting the underlying information that is encoded within the sequence of the DNA [112, 114]. The phenomenon of controlling genetic output and heritable traits in the absence of genetic alteration is known as epigenetics and can help to explain how such a diversity of cell and tissue types within an organism can arise even though they all carry the same genetic information [115, 116]. Interestingly, epigenetic modifications can be passed from one generation to another or from mother cell to daughter cell. [117-119]. Despite being heritable, epigenetic modifications are not permanent. In fact, they are the result of an ever-evolving process in which sets of genes can be targeted for activation or repression in response to environmental or developmental cues. While highly-dynamic, epigenetic modifications are not placed randomly. Instead, various protein complexes regulate them.

Histone post-translational modification as a mechanism for regulating chromatin structure to effect gene expression

The core histones are small and positively charged proteins consisting of a globular domain and a flexible tail. While the core of the histone octamer adopts a central helical fold, known as the histone fold, the disordered N-terminal tails project out from the body of the nucleosome and into solution where they are accessible to nuclear proteins [120]. The histone proteins are evolutionarily conserved, suggesting that they have an important physiological role. In fact, histone tails serve as a signaling hub for information regarding chromatin architecture and gene expression. Histone tails are a major site for post-translational modification with well over 100 different modifications being observed across all four core histone tails [113]. The types of modifications to histone tails include acetylation, methylation, phosphorylation, ubiquitination, and sumoylation [111, 121]. The covalent addition of these functional groups to the side chains of amino acids within the histone tails is linked to significant biological consequences. Ultimately, these consequences can result in altered gene expression patterns.

Post-translational modifications are believed to impact chromatin structure, and consequently, gene expression in one of several ways. One way is through altering the charge status of the nucleosome. This can be accomplished through the addition of a charged functional group, as is the case with histone acetylation. In the case of acetylation, the addition of the negatively charged acetyl group to the positively charged histone neutralizes the charge on the histones and subsequently alters the degree of association between DNA and the histones thereby making the DNA more accessible [122]. Thus, histone acetylation is generally associated with transcriptional activity. A

second way the post-translational modification of histone tails affects gene expression is by creating or destroying binding sites for effector molecules [122]. A number of conserved histone binding domains have been identified, many of which are capable of distinguishing between post-translationally modified and unmodified histones. This type of discretion allows for effector molecules to be recruited to target genes in a manner dependent on the modification state and provides a mechanism for regulating expression of specific genes [123].

It is imperative that gene expression is regulated spatially and temporally as an organism develops and ages so that they can meet the ever-changing demands that a cell encounters. The highly-dynamic nature of histone post-translational modification provides an excellent system for responding to developmental or environmental cues in a quick and efficient manner. Thus, gene expression programs can rapidly adapt and cellular phenotypes can be reprogrammed to fit current needs.

Histone methyltransferases and demethylases specifically regulate histone methylation to facilitate changes in gene expression patterns

Histone methylation takes on several forms. The overwhelming majority of methylation events occur on the side chains of lysine or arginine residues. Lysine residues are subject to the covalent addition of one, two, or three methyl groups while arginine residues can be methylated once or twice [124]. Thus, lysine residues can exist as un-, mono-, di-, or trimethylated whereas arginine residues may be un-, mono-, symmetric di-, or asymmetric dimethylated. Histone methylation is a dynamic process that

is moderated by the histone methyltransferases (HMTs), which place the methyl marks, and the histone demethylases, which remove them.

Histone methylation is carried out in a specific manner with each lysine and arginine being methylated by a unique family of enzymes [124]. These enzymes modify residues at a particular location but are unable to modify residues at other positions [125]. Methylation of each residue carries with it a different biological consequence. These marks, working in combinatorial fashion, dictate overall chromatin structure and transcriptional activity [125].

Histone modifications may signal for additional modifications to establish an environment that is conducive or detrimental to gene expression

Specific histone modifications correlate with transcriptional activation while others correlate with transcriptional repression [126]. A classic example of this phenomenon is trimethylation of lysine 4 of histone H3 (H3K4me3) which is prevalent near transcription start sites of actively expressed genes [127]. In contrast to H3K4me3, which promotes activation of gene transcription, trimethylation of lysine 27 of histone H3 (H3K27me3) is associated with transcriptional repression [126]. In fact, H3K27me3 is most frequently observed at promoter regions of silenced genes [128].

Often, the presence of one histone modification will influence the way adjacent marks are placed or removed [121, 125]. In this case, a certain histone modification may promote recruitment of enzymes that place modifications with similar effects while obstructing enzymes that catalyze modifications with conflicting effects. For example, H3K4me3 is enhanced following H2B monoubiquitination [129]. Additionally, H3K4me3

promotes recruitment of a histone demethylase with activity against H3K9 [130]. Since H3K4me3, an activating mark, is being added and methylation of H3K9, a repressive mark, is being removed the net result is enhanced transcriptional activity.

Homologs of epigenetic modifiers discovered in *Drosophila* are identified as histone methyltransferases with opposing functions in gene regulation

A critical breakthrough in the understanding of epigenetic processes came from the discovery and characterization of histone methyltransferases. Genetics and biochemistry collided to expose two fundamental groups of epigenetic modifiers. These two groups, the Polycomb family of repressors and the Trithorax family of activators, were discovered in genetic screens performed in *Drosophila* [119, 131-133]. While it was known that these groups play a role in regulating transcription, it was not immediately evident how they did so. This changed when the methyltransferases were cloned, sequenced, and found to be members of the Polycomb and Trithorax groups which had been previously identified in the genetic screens. Remarkably, the histone methyltransferases associated with euchromatin and gene expression were found among the Trithorax family of activators [119, 131]. On the contrary, the histone methyltransferases linked to heterochromatin and gene silencing were identified as members of the Polycomb family of repressors [119, 132, 133].

Proteins from the Polycomb and Trithorax groups are well conserved. Indeed, homologs of these two groups have been described in mammals. Included in the list of Trithorax homologs are the mixed-lineage leukemia (MLL) proteins which catalyze the methylation of H3K4 [131]. Within the list of Polycomb homologs are enhancer of zeste

homolog 2 (EZH2) and suppressor of variegation 3-9 homolog 1 (SUV39H1) which catalyze methylation of H3K27 and H3K9, respectively [132].

While EZH2 catalyzes the methylation reaction that results in H3K27 methylation, it has no activity on its own. For robust methylation activity, EZH2 requires interaction with at least two additional cofactors [134-136]. These two cofactors, suppressor of zeste 12 (SUZ12) and embryonic ectoderm development (EED), along with retinoblastoma-associated protein 46 and 48 (RbAp46/48) and EZH2 make up the core components of a larger complex of proteins termed the polycomb repressive complex 2 (PRC2) [137]. The four core components are sufficient for methylation activity; however, numerous other cofactors have been found to play a role in fine-tuning PRC2 activity [137].

The MLL family of histone methyltransferases has six members including MLL1-4, which catalyze methylation of H3K4. Like EZH2, the MLL methyltransferases are the catalytic subunits of large protein complexes [138]. Also, like EZH2, MLL proteins have little activity on their own and require several key cofactors for full functionality [139-141]. Each of these cofactors, Ash2-like (ASH2L), WD repeat containing protein 5 (WDR5), and RB binding protein 5 (RBBP5), are essential for assembly of the core complex and induction of methyltransferase activity.

The PRC2 and MLL complexes are essential for life and their activities have been linked to many aspects of cell physiology including fate determination, stem cell self-renewal, cell cycle regulation, and tumorigenesis [131, 132]. Even though these complexes are vitally important it is not clear how they identify gene targets. Interestingly, the core complexes do not contain sequence-specific DNA-binding proteins [142]. Therefore, targeting of these complexes must be carried out by additional factors.

Transcription factors like Pax2 provide locus specificity for epigenetic modifiers

As mentioned previously, the interaction of Pax2 with the adapter proteins PTIP and GRG4 link them to epigenetic complexes that regulate chromatin structure (Figure 1.4). Biochemical analysis has shown that the Pax2 transactivation domain is phosphorylated by c-Jun N-terminal Kinase (JNK) in response to external signals such as WNT signals [99, 143]. Phosphorylation of the transactivation domain primes it for binding by PTIP, a multi-BRCT domain containing protein [95]. The BRCT domain was first discovered in the C-terminal region of BRCA1 [144] and has since been shown to be a phosphoserine binding domain [145, 146]. The BRCT domain mediated binding of PTIP to Pax2 links Pax2 to the MLL histone methyltransferase complex [147]. This interaction is associated with an increase in H3K4me3 and results in amplified transcription.

In addition to facilitating gene activation, Pax2 can also mediate gene silencing through an interaction with GRG4. The Grg4 mediated conversion of Pax2 from positive to negative transcriptional regulator has been looked at mechanistically [101, 148]. These studies propose a model in which Grg4-dependent recruitment of a phosphatase, PPM1B, leads to dephosphorylation of the Pax2 transactivation domain with subsequent displacement of the PTIP/MLL complex. The displacement of the PTIP/MLL complex is concurrent with Grg4-dependent recruitment of PRMT5, an arginine methyltransferase, and the polycomb repressive complex 2 (PRC2). Ultimately, this leads to alterations in the chromatin landscape as the activating H3K4me3 marks placed by the PTIP/MLL complex are exchanged for the silencing H4R3me2 and H3K27me3 marks placed by PRMT5 and PRC2 respectively.

Together these findings show that Pax proteins exhibit versatility in regulating gene expression and that they can provide the locus specificity for a variety of epigenetic complexes. Furthermore, these data point to additional aspects of Pax biology that could potentially be targeted to regulate their activity.

Potential ways of targeting Pax2 with small-molecules to regulate Pax2 mediated gene regulation

Pax proteins are quite complex. They possess the ability to alternate between positive and negative transcriptional regulation depending on cellular context and cofactor availability. This provides an opportunity to manipulate their activity in one direction or the other by interfering with particular protein-protein interactions. For example, blocking the Pax2-PTIP interaction would limit gene activation while disruption of the Pax2-GRG4 interaction would inhibit Pax2 target gene repression. The same effect could be observed by disrupting a myriad of interactions that are required to assemble the respective epigenetic complexes or through inhibiting their enzymatic activity.

The Pax factors contain multiple DNA-binding domains that combine to define target specificity. While both DNA-binding domains contribute to the overall functionality of Pax proteins, the paired domain is what defines these proteins and is perhaps the most critical. Indeed, the vast majority of disease-related mutations described in Pax genes have been found to be localized to the PD [87, 89, 149-156]. Interestingly, several of these mutations were found to be clustered in one region of the Pax2 paired domain and have been shown to abolish DNA-binding ability of Pax2, which suggests that small-

molecules targeting this region may inhibit Pax2 activity by interfering with Pax2 localization to target genes.

Transcription factors like Pax2 act, in response to a complex network of signaling pathways which are activated by intrinsic or extrinsic stimuli, to regulate the expression of a particular set of genes so that the appropriate biological response can be mounted [157, 158]. The Pax2 transactivation domain is phosphorylated by c-Jun N-terminal Kinase (JNK) in response to external signals such as WNT signals [99, 143]. Phosphorylation of the transactivation domain in associated with Pax2 mediated activation of target genes. Targeting the kinases involved in the upstream signaling cascades that eventually lead to phosphorylation of the transactivation domain could block the activating ability of Pax2 while not affecting its gene silencing ability.

While the above examples represent rational approaches based on existing knowledge of Pax2, there are also unbiased approaches that could be employed. Unbiased approaches have the potential to identify small-molecule regulators of Pax2 activity that do not function through any of the previously described methods. If found, these molecules could useful probes for shedding light on novel aspects of Pax2 regulation.

Throughout the remainder of this thesis I will discuss the advantages and disadvantages of these approaches and describe in detail the methods that I chose to employ. Additionally, I will address the positives and negatives of several screening approaches and talk about an assay the I developed to look for small-molecule with activity against Pax2. I will discuss the results of the screening approaches and highlight a novel small-molecule inhibitor of Pax2 mediated transcription activation that I identified.

Finally, I will talk about my plans for future studies and ways to improve upon the hits that I have found.

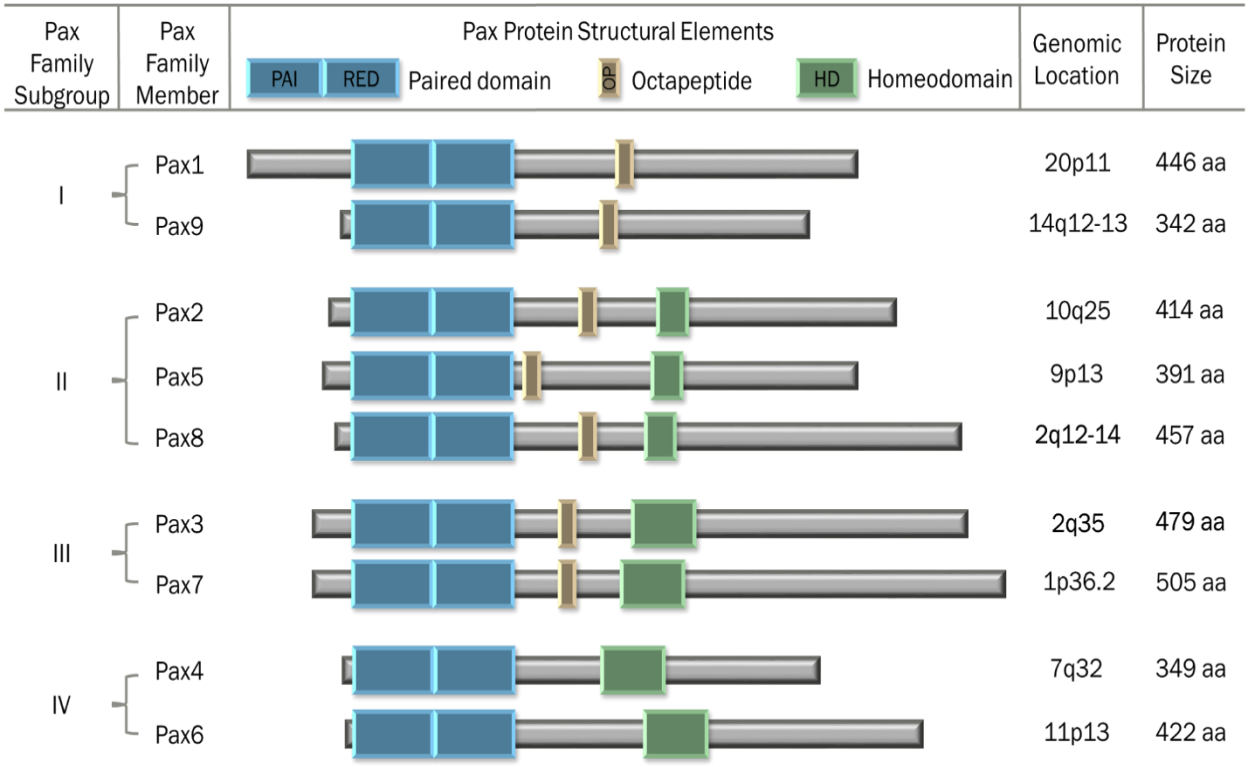


Figure 1.1. The mammalian Pax family of transcription factors. Schematic drawings of the nine Pax family members showing the conserved structural domains. These structural domains include the highly-conserved paired domain (blue) which endows these proteins with DNA-binding capability and is the defining domain for this family of transcription factors. The Pax family of transcription factors are further divided into four subgroups based on the presence of additional conserved structural elements. These elements include the octapeptide sequence (gold) which is found in all members except Pax4 and Pax6. Pax3, Pax4, Pax6, and Pax7 also contain a full paired type homeodomain (green) while Pax2, Pax5, and Pax8 contain only a truncated version of the paired type homeodomain.

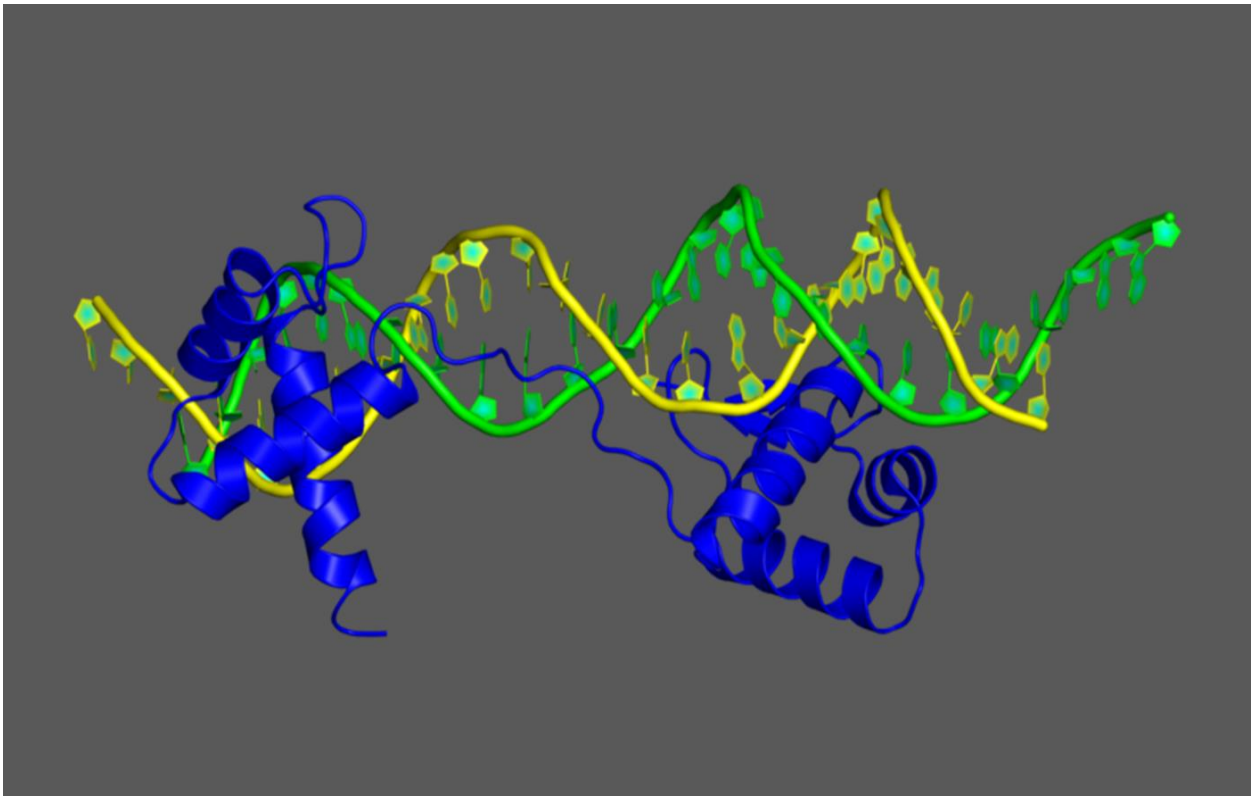


Figure 1.2. Structure of Pax5 paired domain bound to DNA (PDB ID: 1K78). Pax5 paired domain (Blue) was elucidated by co-crystallization with its cognate DNA sequence (green and yellow) and Ets1 (not shown) [64]. This structure shows that the paired domain is composed of two globular HTH subdomains that are connected by a disordered linker. Additionally, the co-crystal structures indicate that the N-terminal HTH subdomain and the C-terminal HTH subdomain make contacts with nucleotide bases in adjacent major grooves of the DNA double helix while the extended linker makes substantial contact with bases in the intervening minor groove.

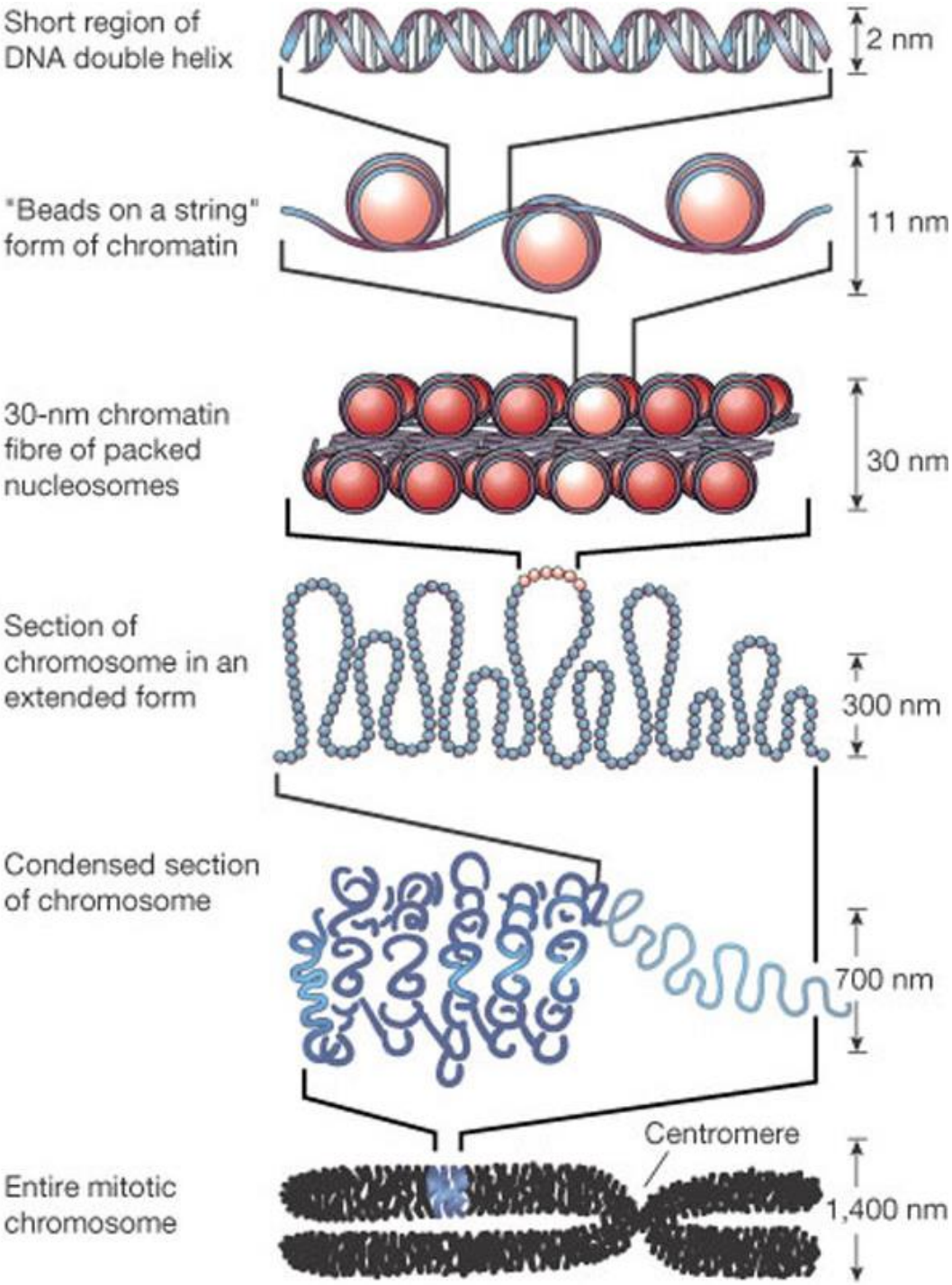


Figure 1.3. Organization of DNA within a cell. DNA is wrapped around an octamer of core histones to form nucleosomes, the basic unit of chromatin. Loosely connected nucleosomes form the structure known as beads on a string. These nucleosomes can be folded into a fiber that is roughly 30 nm in diameter. The 30 nm fiber can be further compacted into higher-order structures that eventually produce the mitotic chromosome. This figure is taken from Felsenfeld, G and Groudine, M *Nature,* 2003 [159].

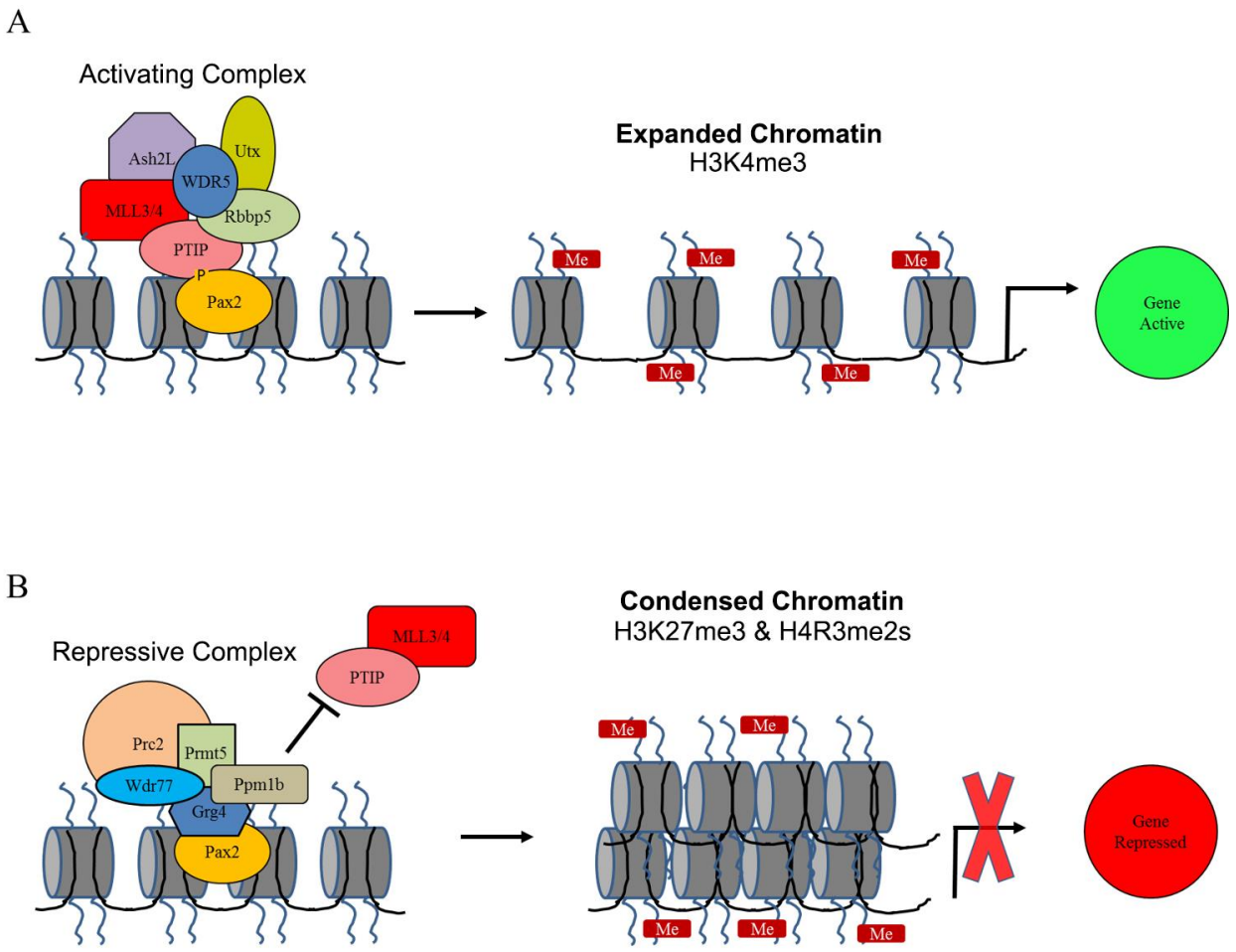


Figure 1.4. Model of Pax mediated gene regulation. Cell and tissue-type specific DNA-binding proteins such as Pax2 can provide the locus specificity for chromatin modifying complexes and ultimately control which genes are expressed or repressed. **A)** Pax2 recruits the adapter protein PTIP and the MLL complex which catalyzes methylation of H3K4. These methylation marks trigger chromatin relaxation and allow for gene transcription. **B)** However, in cells with high levels of Grg4, the PTIP/MLL complex can be replaced by another set of epigenetic modifiers that include PRMT5 and PRC2. These complexes initiate methylation of H4R3 and H3K27, respectively. The placement of these marks triggers chromatin condensation and prohibits gene expression.

Bibliography

1. Seale, P., et al., *Pax7 is required for the specification of myogenic satellite cells.* Cell, 2000. **102**(6): p. 777-86.

2. Blake, J.A., et al., *Perplexing Pax: from puzzle to paradigm.* Dev Dyn, 2008. **237**(10): p. 2791-803.

3. Lang, D., et al., *PAX genes: roles in development, pathophysiology, and cancer.* Biochem Pharmacol, 2007. **73**(1): p. 1-14.

4. Blake, J.A. and M.R. Ziman, *Pax genes: regulators of lineage specification and progenitor cell maintenance.* Development, 2014. **141**(4): p. 737-51.

5. Dressler, G.R., et al., *Pax2, a new murine paired-box-containing gene and its expression in the developing excretory system.* Development, 1990. **109**(4): p. 787-95.

6. Plachov, D., et al., *Pax8, a murine paired box gene expressed in the developing excretory system and thyroid gland.* Development, 1990. **110**(2): p. 643-51.

7. Sanyanusin, P., et al., *Mutation of the PAX2 gene in a family with optic nerve colobomas, renal anomalies and vesicoureteral reflux.* Nat Genet, 1995. **9**(4): p. 358-64.

8. Mansouri, A., K. Chowdhury, and P. Gruss, *Follicular cells of the thyroid gland require Pax8 gene function.* Nat Genet, 1998. **19**(1): p. 87-90.

9. Adams, B., et al., *Pax-5 encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS, and adult testis.* Genes Dev, 1992. **6**(9): p. 1589-607.

10. Wang, Q., et al., *Pax genes in embryogenesis and oncogenesis.* J Cell Mol Med, 2008. **12**(6A): p. 2281-94.

11. Stoykova, A. and P. Gruss, *Roles of Pax-genes in developing and adult brain as suggested by expression patterns.* J Neurosci, 1994. **14**(3 Pt 2): p. 1395-412.

12. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57-70.

13. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.

14. Ostrom, L., et al., *Reduced Pax2 gene dosage increases apoptosis and slows the progression of renal cystic disease.* Dev Biol, 2000. **219**(2): p. 250-8.

15. Stayner, C., et al., *Pax2 gene dosage influences cystogenesis in autosomal dominant polycystic kidney disease.* Hum Mol Genet, 2006. **15**(24): p. 3520-8.

16. Winyard, P.J., et al., *The PAX2 tanscription factor is expressed in cystic and hyperproliferative dysplastic epithelia in human kidney malformations.* J Clin Invest, 1996. **98**(2): p. 451-9.

17. Dressler, G.R. and E.C. Douglass, *Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor.* Proc Natl Acad Sci U S A, 1992. **89**(4): p. 1179-83.

18. Eccles, M.R., et al., *Expression of the PAX2 gene in human fetal kidney and Wilms' tumor.* Cell Growth Differ, 1992. **3**(5): p. 279-89.

19. Gnarra, J.R. and G.R. Dressler, *Expression of Pax-2 in human renal cell carcinoma and growth inhibition by antisense oligonucleotides.* Cancer Res, 1995. **55**(18): p. 4092-8.

20. Tong, G.X., et al., *PAX2: a reliable marker for nephrogenic adenoma.* Mod Pathol, 2006. **19**(3): p. 356-63.

21. Tong, G.X., et al., *PAX8 and PAX2 immunostaining facilitates the diagnosis of primary epithelial neoplasms of the male genital tract.* Am J Surg Pathol, 2011. **35**(10): p. 1473-83.

22. Hueber, P.A., et al., *PAX2 inactivation enhances cisplatin-induced apoptosis in renal carcinoma cells.* Kidney Int, 2006. **69**(7): p. 1139-45.

23. Hueber, P.A., et al., *In vivo validation of PAX2 as a target for renal cancer therapy.* Cancer Lett, 2008. **265**(1): p. 148-55.

24. Morgan, E.A., et al., *PAX8 and PAX5 are differentially expressed in B-cell and T-cell lymphomas.* Histopathology, 2013. **62**(3): p. 406-13.

25. Mullighan, C.G., et al., *Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia.* Nature, 2007. **446**(7137): p. 758-64.

26. Krenacs, L., et al., *Transcription factor B-cell-specific activator protein (BSAP) is differentially expressed in B cells and in subsets of B-cell lymphomas.* Blood, 1998. **92**(4): p. 1308-16.

27. Zhang, X., Z. Lin, and I. Kim, *Pax5 expression in non-Hodgkin's lymphomas and acute leukemias.* J Korean Med Sci, 2003. **18**(6): p. 804-8.

28. Morrison, A.M., et al., *Deregulated PAX-5 transcription from a translocated IgH promoter in marginal zone lymphoma.* Blood, 1998. **92**(10): p. 3865-78.

29. Busslinger, M., et al., *Deregulation of PAX-5 by translocation of the Emu enhancer of the IgH locus adjacent to two alternative PAX-5 promoters in a diffuse large-cell lymphoma.* Proc Natl Acad Sci U S A, 1996. **93**(12): p. 6129-34.

30. Iida, S., et al., *The t(9;14)(p13;q32) chromosomal translocation associated with lymphoplasmacytoid lymphoma involves the PAX-5 gene.* Blood, 1996. **88**(11): p. 4110-7.

31. Baumann Kubetzko, F.B., et al., *The PAX5 oncogene is expressed in N-type neuroblastoma cells and increases tumorigenicity of a S-type cell line.* Carcinogenesis, 2004. **25**(10): p. 1839-46.

32. Davies, J.A., A.D. Perera, and C.L. Walker, *Mechanisms of epithelial development and neoplasia in the metanephric kidney.* Int J Dev Biol, 1999. **43**(5): p. 473-8.

33. Igarashi, T., et al., *Aberrant expression of Pax-2 mRNA in renal cell carcinoma tissue and parenchyma of the affected kidney.* Int J Urol, 2001. **8**(2): p. 60-4.

34. Scouten, W.T., et al., *Cytoplasmic localization of the paired box gene, Pax-8, is found in pediatric thyroid cancer and may be associated with a greater risk of recurrence.* Thyroid, 2004. **14**(12): p. 1037-46.

35. Puglisi, F., et al., *Expression of Pax-8, p53 and bcl-2 in human benign and malignant thyroid diseases.* Anticancer Res, 2000. **20**(1A): p. 311-6.

36. Castro, P., et al., *PAX8-PPARgamma rearrangement is frequently detected in the follicular variant of papillary thyroid carcinoma.* J Clin Endocrinol Metab, 2006. **91**(1): p. 213-20.

37. Kroll, T.G., et al., *PAX8-PPARgamma1 fusion oncogene in human thyroid carcinoma [corrected].* Science, 2000. **289**(5483): p. 1357-60.

38. Cheung, L., et al., *Detection of the PAX8-PPAR gamma fusion oncogene in both follicular thyroid carcinomas and adenomas.* J Clin Endocrinol Metab, 2003. **88**(1): p. 354-7.

39. Marques, A.R., et al., *Expression of PAX8-PPAR gamma 1 rearrangements in both follicular thyroid carcinomas and adenomas.* J Clin Endocrinol Metab, 2002. **87**(8): p. 3947-52.

40. Placzkowski, K.A., et al., *The Role of the PAX8/PPARgamma Fusion Oncogene in Thyroid Cancer.* PPAR Res, 2008. **2008**: p. 672829.

41. Gregory Powell, J., et al., *The PAX8/PPARgamma fusion oncoprotein transforms immortalized human thyrocytes through a mechanism probably involving wild-type PPARgamma inhibition.* Oncogene, 2004. **23**(20): p. 3634-41.

42. Nikiforova, M.N., et al., *PAX8-PPARgamma rearrangement in thyroid tumors: RT-PCR and immunohistochemical analyses.* Am J Surg Pathol, 2002. **26**(8): p. 1016-23.

43. Chapman, A.B., et al., *Autosomal-dominant polycystic kidney disease (ADPKD): executive summary from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference.* Kidney International, 2015. **88**(1): p. 17-27.

44. Torres, V.E., P.C. Harris, and Y. Pirson, *Autosomal dominant polycystic kidney disease.* Lancet, 2007. **369**(9569): p. 1287-301.

45. Torres, V.E. and P.C. Harris, *Strategies targeting cAMP signaling in the treatment of polycystic kidney disease.* J Am Soc Nephrol, 2014. **25**(1): p. 18-32.

46. Hildebrandt, F., T. Benzing, and N. Katsanis, *Ciliopathies.* N Engl J Med, 2011. **364**(16): p. 1533-43.

47. Spithoven, E.M., et al., *Renal replacement therapy for autosomal dominant polycystic kidney disease (ADPKD) in Europe: prevalence and survival--an analysis of data from the ERA-EDTA Registry.* Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association, 2014. **29 Suppl 4**: p. iv15-25.

48. Torres, V.E., et al., *Tolvaptan in patients with autosomal dominant polycystic kidney disease.* The New England journal of medicine, 2012. **367**(25): p. 2407-18.

49. Torres, V.E., *Vasopressin receptor antagonists, heart failure, and polycystic kidney disease.* Annual Review of Medicine, 2015. **66**: p. 195-210.

50. Ljungberg, B., et al., *The epidemiology of renal cell carcinoma.* European Urology, 2011. **60**(4): p. 615-21.

51. Daniel, L., et al., *Pax-2 expression in adult renal tumors.* Hum Pathol, 2001. **32**(3): p. 282-7.

52. Mazal, P.R., et al., *Expression of aquaporins and PAX-2 compared to CD10 and cytokeratin 7 in renal neoplasms: a tissue microarray study.* Mod Pathol, 2005. **18**(4): p. 535-40.

53. Memeo, L., et al., *Immunohistochemical analysis for cytokeratin 7, KIT, and PAX2: value in the differential diagnosis of chromophobe cell carcinoma.* Am J Clin Pathol, 2007. **127**(2): p. 225-9.

54. Molina, A.M. and D.M. Nanus, *Recent advances in the management of renal cell carcinoma.* F1000Research, 2016. **5**.

55. Gnarra, J.R., et al., *Mutations of the VHL tumour suppressor gene in renal carcinoma.* Nature Genetics, 1994. **7**: p. 85-90.

56. Ljungberg, B., et al., *EAU guidelines on renal cell carcinoma: 2014 update.* European Urology, 2015. **67**(5): p. 913-24.

57. Latchman, D.S., *Transcription factors: an overview.* Int J Biochem Cell Biol, 1997. **29**(12): p. 1305-12.

58. Vaquerizas, J.M., et al., *A census of human transcription factors: function, expression and evolution.* Nat Rev Genet, 2009. **10**(4): p. 252-63.

59. Luscombe, N.M., et al., *An overview of the structures of protein-DNA complexes.* Genome Biol, 2000. **1**(1): p. REVIEWS001.

60. Bopp, D., et al., *Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila.* Cell, 1986. **47**(6): p. 1033-40.

61. Noll, M., *Evolution and role of Pax genes.* Curr Opin Genet Dev, 1993. **3**(4): p. 595-605.

62. Walther, C., et al., *Pax: a murine multigene family of paired box-containing genes.* Genomics, 1991. **11**(2): p. 424-34.

63. Xu, W., et al., *Crystal structure of a paired domain-DNA complex at 2.5 A resolution reveals structural basis for Pax developmental mutations.* Cell, 1995. **80**(4): p. 639-50.

64. Garvie, C.W., J. Hagman, and C. Wolberger, *Structural studies of Ets-1/Pax5 complex formation on DNA.* Mol Cell, 2001. **8**(6): p. 1267-76.

65. Xu, H.E., et al., *Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding.* Genes Dev, 1999. **13**(10): p. 1263-75.

66. Codutti, L., et al., *The solution structure of DNA-free Pax-8 paired box domain accounts for redox regulation of transcriptional activity in the pax protein family.* J Biol Chem, 2008. **283**(48): p. 33321-8.

67. Epstein, J., et al., *Identification of a Pax paired domain recognition sequence and evidence for DNA-dependent conformational changes.* J Biol Chem, 1994. **269**(11): p. 8355-61.

68. Epstein, J.A., et al., *Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing.* Genes Dev, 1994. **8**(17): p. 2022-34.

69. Pellizzari, L., G. Tell, and G. Damante, *Co-operation between the PAI and RED subdomains of Pax-8 in the interaction with the thyroglobulin promoter.* The Biochemical journal, 1999. **337 (Pt 2)**: p. 253-62.

70. Birrane, G., A. Soni, and J.A. Ladias, *Structural basis for DNA recognition by the human PAX3 homeodomain.* Biochemistry, 2009. **48**(6): p. 1148-55.

71. Wilson, D.S., et al., *High resolution crystal structure of a paired (Pax) class cooperative homeodomain dimer on DNA.* Cell, 1995. **82**(5): p. 709-19.

72. Treisman, J., et al., *A single amino acid can determine the DNA binding specificity of homeodomain proteins.* Cell, 1989. **59**(3): p. 553-62.

73. Hanes, S.D. and R. Brent, *DNA specificity of the bicoid activator protein is determined by homeodomain recognition helix residue 9.* Cell, 1989. **57**(7): p. 1275-83.

74. Wilson, D., et al., *Cooperative dimerization of paired class homeo domains on DNA.* Genes Dev, 1993. **7**(11): p. 2120-34.

75. Verbruggen, V., et al., *The Pax6b homeodomain is dispensable for pancreatic endocrine cell differentiation in zebrafish.* J Biol Chem, 2010. **285**(18): p. 13863-73.

76. Soleimani, V.D., et al., *Transcriptional dominance of Pax7 in adult myogenesis is due to high-affinity recognition of homeodomain motifs.* Dev Cell, 2012. **22**(6): p. 1208-20.

77. Corry, G.N. and D.A. Underhill, *Pax3 target gene recognition occurs through distinct modes that are differentially affected by disease-associated mutations.* Pigment Cell Res, 2005. **18**(6): p. 427-38.

78. Jun, S. and C. Desplan, *Cooperative interactions between paired domain and homeodomain.* Development, 1996. **122**(9): p. 2639-50.

79. Budry, L., et al., *The selector gene Pax7 dictates alternate pituitary cell fates through its pioneer action on chromatin remodeling.* Genes Dev, 2012. **26**(20): p. 2299-310.

80. Clements, J., et al., *Mutational analysis of the eyeless gene and phenotypic rescue reveal that an intact Eyeless protein is necessary for normal eye and brain development in Drosophila.* Dev Biol, 2009. **334**(2): p. 503-12.

81. Bertuccioli, C., et al., *In vivo requirement for the paired domain and homeodomain of the paired segmentation gene product.* Development, 1996. **122**(9): p. 2673-85.

82. Fortin, A.S., D.A. Underhill, and P. Gros, *Reciprocal effect of Waardenburg syndrome mutations on DNA binding by the Pax-3 paired domain and homeodomain.* Hum Mol Genet, 1997. **6**(11): p. 1781-90.

83. Underhill, D.A. and P. Gros, *The paired-domain regulates DNA binding by the homeodomain within the intact Pax-3 protein.* J Biol Chem, 1997. **272**(22): p. 14175-82.

84. Underhill, D.A., K.J. Vogan, and P. Gros, *Analysis of the mouse Splotch-delayed mutation indicates that the Pax-3 paired domain can influence homeodomain DNA-binding activity.* Proc Natl Acad Sci U S A, 1995. **92**(9): p. 3692-6.

85. Apuzzo, S., et al., *Cross-talk between the paired domain and the homeodomain of Pax3: DNA binding by each domain causes a structural change in the other domain, supporting interdependence for DNA Binding.* J Biol Chem, 2004. **279**(32): p. 33601-12.

86. Apuzzo, S. and P. Gros, *Site-specific modification of single cysteine Pax3 mutants reveals reciprocal regulation of DNA binding activity of the paired and homeo domain.* Biochemistry, 2002. **41**(40): p. 12076-85.

87. Chalepakis, G., et al., *Molecular basis of splotch and Waardenburg Pax-3 mutations.* Proc Natl Acad Sci U S A, 1994. **91**(9): p. 3685-9.

88. Jones, K.A. and J.T. Kadonaga, *Exploring the transcription-chromatin interface.* Genes Dev, 2000. **14**(16): p. 1992-6.

89. Kozmik, Z., et al., *Alternative splicing of Pax-8 gene transcripts is developmentally regulated and generates isoforms with different transactivation properties.* Mol Cell Biol, 1993. **13**(10): p. 6024-35.

90. Chalepakis, G., et al., *Pax-3 contains domains for transcription activation and transcription inhibition.* Proc Natl Acad Sci U S A, 1994. **91**(26): p. 12745-9.

91. Czerny, T. and M. Busslinger, *DNA-binding and transactivation properties of Pax-6: three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5).* Mol Cell Biol, 1995. **15**(5): p. 2858-71.

92. Fujitani, Y., et al., *Identification of a portable repression domain and an E1A-responsive activation domain in Pax4: a possible role of Pax4 as a transcriptional repressor in the pancreas.* Mol Cell Biol, 1999. **19**(12): p. 8281-91.

93. Kalousova, A., et al., *DNA binding and transactivating properties of the paired and homeobox protein Pax4.* Biochem Biophys Res Commun, 1999. **259**(3): p. 510-8.

94. Lechner, M.S. and G.R. Dressler, *Mapping of Pax-2 transcription activation domains.* J Biol Chem, 1996. **271**(35): p. 21088-93.

95. Lechner, M.S., I. Levitan, and G.R. Dressler, *PTIP, a novel BRCT domain-containing protein interacts with Pax2 and is associated with active chromatin.* Nucleic Acids Res, 2000. **28**(14): p. 2741-51.

96. Smith, S.T. and J.B. Jaynes, *A conserved region of engrailed, shared among all en-, gsc-, Nk1-, Nk2- and msh-class homeoproteins, mediates active transcriptional repression in vivo.* Development, 1996. **122**(10): p. 3141-50.

97. Tolkunova, E.N., et al., *Two distinct types of repression domain in engrailed: one interacts with the groucho corepressor and is preferentially active on integrated target genes.* Mol Cell Biol, 1998. **18**(5): p. 2804-14.

98. Eberhard, D., et al., *Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family.* EMBO J, 2000. **19**(10): p. 2292-303.

99. Cai, Y., et al., *Groucho suppresses Pax2 transactivation by inhibition of JNK-mediated phosphorylation.* EMBO J, 2003. **22**(20): p. 5522-9.

100. Lang, D., et al., *Pax3 functions at a nodal point in melanocyte stem cell differentiation.* Nature, 2005. **433**(7028): p. 884-7.

101. Patel, S.R., et al., *Epigenetic mechanisms of Groucho/Grg/TLE mediated transcriptional repression.* Mol Cell, 2012. **45**(2): p. 185-95.

102. Kornberg, R.D., *Chromatin structure: a repeating unit of histones and DNA.* Science, 1974. **184**(4139): p. 868-71.

103. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 A resolution.* Nature, 1997. **389**(6648): p. 251-60.

104. Kornberg, R.D. and J.O. Thomas, *Chromatin structure; oligomers of the histones.* Science, 1974. **184**(4139): p. 865-8.

105. Ramakrishnan, V., *Histone H1 and chromatin higher-order structure.* Crit Rev Eukaryot Gene Expr, 1997. **7**(3): p. 215-30.

106. Groth, A., et al., *Chromatin challenges during DNA replication and repair.* Cell, 2007. **128**(4): p. 721-33.

107. Li, B., M. Carey, and J.L. Workman, *The role of chromatin during transcription.* Cell, 2007. **128**(4): p. 707-19.

108. Hayes, J.J. and A.P. Wolffe, *The interaction of transcription factors with nucleosomal DNA.* Bioessays, 1992. **14**(9): p. 597-603.

109. Workman, J.L. and R.E. Kingston, *Nucleosome core displacement in vitro via a metastable transcription factor-nucleosome complex.* Science, 1992. **258**(5089): p. 1780-4.

110. Sims, R.J., 3rd, S.S. Mandal, and D. Reinberg, *Recent highlights of RNA-polymerase-II-mediated transcription.* Curr Opin Cell Biol, 2004. **16**(3): p. 263-71.

111. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.

112. Goldberg, A.D., C.D. Allis, and E. Bernstein, *Epigenetics: a landscape takes shape.* Cell, 2007. **128**(4): p. 635-8.

113. Bernstein, B.E., A. Meissner, and E.S. Lander, *The mammalian epigenome.* Cell, 2007. **128**(4): p. 669-81.

114. Bird, A., *Perceptions of epigenetics.* Nature, 2007. **447**(7143): p. 396-398.

115. Jaenisch, R. and R. Young, *Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming.* Cell, 2008. **132**(4): p. 567-82.

116. *Moving AHEAD with an international human epigenome project.* Nature, 2008. **454**(7205): p. 711-715.

117. Bird, A., *DNA methylation patterns and epigenetic memory.* Genes Dev, 2002. **16**(1): p. 6-21.

118. Grewal, S.I. and D. Moazed, *Heterochromatin and epigenetic control of gene expression.* Science, 2003. **301**(5634): p. 798-802.

119. Ringrose, L. and R. Paro, *Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins.* Annu Rev Genet, 2004. **38**: p. 413-43.

120. Allis, C.D., T. Jenuwein, and D. Reinberg, *Epigenetics*. 2007: Cold Spring Harbor Laboratory Press.

121. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications.* Nature, 2000. **403**(6765): p. 41-5.

122. Bannister, A.J. and T. Kouzarides, *Regulation of chromatin by histone modifications.* Cell Res, 2011. **21**(3): p. 381-95.

123. Jenuwein, T. and C.D. Allis, *Translating the histone code.* Science, 2001. **293**(5532): p. 1074-80.

124. Martin, C. and Y. Zhang, *The diverse functions of histone lysine methylation.* Nat Rev Mol Cell Biol, 2005. **6**(11): p. 838-49.

125. Zhang, Y. and D. Reinberg, *Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails.* Genes Dev, 2001. **15**(18): p. 2343-60.

126. Ernst, J., et al., *Mapping and analysis of chromatin state dynamics in nine human cell types.* Nature, 2011. **473**(7345): p. 43-9.

127. Santos-Rosa, H., et al., *Active genes are tri-methylated at K4 of histone H3.* Nature, 2002. **419**(6905): p. 407-11.

128. Cao, R., et al., *Role of histone H3 lysine 27 methylation in Polycomb-group silencing.* Science, 2002. **298**(5595): p. 1039-43.

129. Lee, J.S., E. Smith, and A. Shilatifard, *The language of histone crosstalk.* Cell, 2010. **142**(5): p. 682-5.

130. Kleine-Kohlbrecher, D., et al., *A functional link between the histone demethylase PHF8 and the transcription factor ZNF711 in X-linked mental retardation.* Mol Cell, 2010. **38**(2): p. 165-78.

131. Schuettengruber, B., et al., *Trithorax group proteins: switching genes on and keeping them active.* Nat Rev Mol Cell Biol, 2011. **12**(12): p. 799-814.

132. Lanzuolo, C. and V. Orlando, *Memories from the polycomb group proteins.* Annu Rev Genet, 2012. **46**: p. 561-89.

133. Schotta, G., et al., *Position-effect variegation and the genetic dissection of chromatin regulation in Drosophila.* Semin Cell Dev Biol, 2003. **14**(1): p. 67-75.

134. Cao, R. and Y. Zhang, *SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex.* Mol Cell, 2004. **15**(1): p. 57-67.

135. Pasini, D., et al., *Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity.* EMBO J, 2004. **23**(20): p. 4061-71.

136. Montgomery, N.D., et al., *The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation.* Curr Biol, 2005. **15**(10): p. 942-7.

137. Margueron, R. and D. Reinberg, *The Polycomb complex PRC2 and its mark in life.* Nature, 2011. **469**(7330): p. 343-9.

138. Avdic, V., et al., *Structural and biochemical insights into MLL1 core complex assembly.* Structure, 2011. **19**(1): p. 101-8.

139. Wysocka, J., et al., *WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development.* Cell, 2005. **121**(6): p. 859-72.

140. Steward, M.M., et al., *Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes.* Nat Struct Mol Biol, 2006. **13**(9): p. 852-4.

141. Dou, Y., et al., *Regulation of MLL1 H3K4 methyltransferase activity by its core components.* Nat Struct Mol Biol, 2006. **13**(8): p. 713-9.

142. Schuettengruber, B., et al., *Genome regulation by polycomb and trithorax proteins.* Cell, 2007. **128**(4): p. 735-45.

143. Cai, Y., et al., *Phosphorylation of Pax2 by the c-Jun N-terminal kinase and enhanced Pax2-dependent transcription activation.* J Biol Chem, 2002. **277**(2): p. 1217-22.

144. Couch, F.J. and B.L. Weber, *Mutations and Polymorphisms in the familial early-onset breast cancer (BRCA1) gene.* Human Mutation, 1996. **8**(1): p. 8-18.

145. Manke, I.A., et al., *BRCT Repeats As Phosphopeptide-Binding Modules Involved in Protein Targeting.* Science, 2003. **302**(5645): p. 636-639.

146. Yu, X., et al., *The BRCT Domain Is a Phospho-Protein Binding Domain.* Science, 2003. **302**(5645): p. 639-642.

147. Patel, S.R., et al., *The BRCT-domain containing protein PTIP links PAX2 to a histone H3, lysine 4 methyltransferase complex.* Dev Cell, 2007. **13**(4): p. 580-92.

148. Abraham, S., et al., *The Groucho-associated phosphatase PPM1B displaces Pax transactivation domain interacting protein (PTIP) to switch the transcription factor Pax2 from a transcriptional activator to a repressor.* J Biol Chem, 2015. **290**(11): p. 7185-94.

149. Balling, R., U. Deutsch, and P. Gruss, *undulated, a mutation affecting the development of the mouse skeleton, has a point mutation in the paired box of Pax 1.* Cell, 1988. **55**(3): p. 531-5.

150. Epstein, D.J., M. Vekemans, and P. Gros, *Splotch (Sp2H), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3.* Cell, 1991. **67**(4): p. 767-74.

151. Hill, R.E., et al., *Mouse Small eye results from mutations in a paired-like homeobox-containing gene.* Nature, 1992. **355**(6362): p. 750.

152. Glaser, T., D.S. Walton, and R.L. Maas, *Genomic structure, evolutionary conservation and aniridia mutations in the human PAX6 gene.* Nat Genet, 1992. **2**(3): p. 232-9.

153. Shimajiri, Y., et al., *A missense mutation of Pax4 gene (R121W) is associated with type 2 diabetes in Japanese.* Diabetes, 2001. **50**(12): p. 2864-9.

154. Jumlongras, D., et al., *A novel missense mutation in the paired domain of PAX9 causes non-syndromic oligodontia.* Hum Genet, 2004. **114**(3): p. 242-9.

155. Mostowska, A., et al., *Novel mutation in the paired box sequence of PAX9 gene in a sporadic form of oligodontia.* Eur J Oral Sci, 2003. **111**(3): p. 272-6.

156. Zhao, J., et al., *A novel missense mutation in the paired domain of human PAX9 causes oligodontia.* Am J Med Genet A, 2007. **143A**(21): p. 2592-7.

157. Brivanlou, A.H. and J.E. Darnell, Jr., *Signal transduction and the control of gene expression.* Science, 2002. **295**(5556): p. 813-8.

158. Carrera, I. and J.E. Treisman, *Message in a nucleus: signaling to the transcriptional machinery.* Curr Opin Genet Dev, 2008. **18**(5): p. 397-403.

159. Felsenfeld, G. and M. Groudine, *Controlling the double helix.* Nature, 2003. **421**(6921): p. 448-53.

CHAPTER 2

Developing a cell-based assay for identifying small-molecule regulators of Pax2 transactivation activity

Abstract

Transcription factors, like Pax2, represent a novel and exciting class of therapeutic targets. However, the discovery and development of small-molecules capable of regulating transcription factor activity is a daunting task. Therefore, these factors continue to be under-investigated. Despite the lack of attention, recent screening campaigns have yielded several small-molecule inhibitors of transcription factors. Some of these molecules target the signaling pathways upstream of the transcription factor or the epigenetic complexes that are downstream. While these methods have proven useful, approaches directly targeting the transcription factor are preferred. Since Pax2 is a diagnostic marker and a validated target for cancers of the kidney, I wanted to find small-molecule modulators of Pax2 activity. To address this goal, I established a cell-based reporter assay that is dependent on Pax2 mediated transcription activation. PRS4-Luc reporter cells transiently transfected with a Pax2 expression vector show strong induction of luciferase. A clonal PRS4-Luc cell line was isolated and used to develop the reporter assay. Once developed, the assay was optimized and miniaturized so that it was suitable

for HTS. Mock runs of this assay in 384-well plates determined that I had developed a robust and reproducible assay, as demonstrated by a Z' of 0.71. This assay provides an attractive method to screen small-molecules for those with antiPax2 activity.

Introduction

Transcription factors are critical regulators of disease processes and represent an appealing class of therapeutic targets

Given the critical role of transcription factors in influencing cellular phenotypes, it makes sense that mutations in or dysregulation of these proteins would have dire consequences. In fact, aberrant gene expression is a hallmark of countless human conditions including cancers [1]. For example, Pax2 is overexpressed in Wilms' tumor [2, 3] and renal cell carcinoma [4-6] where it promotes proliferation and restricts apoptosis. Decreasing Pax2 levels in renal cancer cell lines increases apoptosis, reduces proliferative capacity, and enhances the effects of chemotherapeutic agents [4, 7, 8]. Together these findings make Pax2 an attractive target and suggest that developing small-molecules with the ability to effect transcription factor function may lead to novel therapies. Moreover, these molecules would provide useful research tools for understanding how gene expression is regulated at a molecular level.

Transcription factors represent an under-investigated class of therapeutic targets. In fact, transcription factors have classically been considered "undruggable" [9-11] because of their nuclear localization, the charged nature of their interaction with DNA, and their lack of classical binding pockets. With the exception of the nuclear receptor class of transcription factors (retinoic acid receptor, androgen receptor, estrogen receptor, etc.,), most others do not have endogenous ligands and therefore lack a well-defined pocket for small-molecule binding. This makes them much less amenable to traditional approaches of inhibition with small-molecules. For this reason, indirect approaches to targeting transcription factors have been the most utilized methods. These indirect

approaches often work through targeting kinases involved in the upstream signaling cascades that would otherwise lead to activation of the transcription factor. A major drawback to this approach is that there is a high degree of redundancy and cross-talk amongst signaling cascades which means that targeting one pathway could have far-reaching effects on other signaling pathways. Since these non-specific effects are highly undesirable and because signaling pathways converge at the level of the ultimate effector molecule, the transcription factor, there is a strong motivation for directly targeting transcription factors.

Strategies for targeting transcription factors

Targeting transcription factors is a challenging but exciting strategy for cancer therapeutic development. Perhaps one of best ways to inhibit transcription factors would be to block the interaction between the transcription factor and DNA. This could be accomplished by either targeting the DNA-binding domain of the transcription factor or targeting the binding site within the DNA. While the DNA-binding domain may provide a pocket to target, the charged nature of the protein/DNA interaction could make this difficult. Yet, several DNA-binding domain inhibitors have been found. These include inS3-54, a small-molecule inhibitor of Signal Transducer and Activator of Transcription 3 (STAT3) which was identified recently and was used as a backbone to develop a more potent and selective molecule, inS3-54A18 [12, 13]. These molecules were shown to directly interact with the STAT3 DNA-binding domain and block STAT3 from binding to DNA. Importantly, disruption of the STAT3-DNA interaction induced apoptosis, hindered migration, and limited invasion in cancer cells treated with these molecules [12, 13].

A natural product, echinomycin, which intercalates preferentially in 5'-A/TCGT-3' sites [14] inhibits Hypoxia Inducible Factor 1 (HIF-1) binding and transcription activation [15]. The hypoxia response element (HRE) is found in the promoter of many HIF-1 target genes and is considered the preferential DNA motif for HIF-1 [16]. The HRE site contains 5'-ACGTG-3' which is consistent with the echinomycin recognition sequence. In fact, it was shown that the antiHIF-1 activity of echinomycin was due to intercalation within the HRE and inhibition of HIF-1 binding [15]. Interestingly, it was later shown that echinomycin exhibits similar activity against c-Myc which recognizes 5'-CACGTG-3' sites [17]. Unfortunately, molecules such as echinomycin are unlikely to be able to reach the level of discrimination required to be selective. Thus, they are likely to have effects well beyond those of the preferred targets and therefore are not expected to provide much help in the clinic.

While echinomycin may not provide a viable approach for inhibiting c-Myc in patients, c-Myc inhibitors are of great interest, because c-Myc regulates expression of genes involved in many fundamental cellular processes that are dysregulated in cancer including cell cycle progression, growth, and cell death [18, 19]. c-Myc does not regulate gene expression alone; instead, it does so through heterodimerization with Myc Associated Factor X (Max). Since c-Myc/Max dimerization is required for gene regulation, this interaction provides a potential mechanism for inhibiting c-Myc activity. Indeed, several small-molecule inhibitors of this interaction have been discovered [20-26]. This demonstrates the potential effectiveness of targeting protein-protein interactions that are required for transcription factor activity. Inhibitors of protein-protein interactions could be used to target any number of cofactors or even inhibit recruitment of the epigenetic

complexes to target genes. Targeting protein-protein interactions provides an opportunity for developing highly-specific inhibitors, but this method does not come without its own challenges. Protein-protein interaction interfaces tend to be expansive with surface areas in the 1,500-3000 $Å^2$ range [27, 28]. In addition to the large surface area, these interfaces are generally shallow and typically lack grooves or pockets [10, 29, 30]. The combination of these features pose real obstacles for small-molecule inhibitors of protein-protein interactions.

Transcription factor activity can be indirectly controlled by targeting epigenetic regulators. In terms of DNA methylation, perhaps the best example comes from the azanucleoside class of compounds that include 5-azacytidine and decitabine. These molecules are cytidine analogs which get incorporated into DNA and covalently bind to the DNA methyltransferases [31]. This leads to exhaustion of the DNMTs and eventual reversal of the abnormal DNA methylation signature [32-34]. Despite early problems with gastrointestinal toxicity [33], lower dose 5-azacytidine was recently approved for use in patients with myelodysplastic syndromes (MDS) [35]. The benefits of 5-azacytidine treatment of patients with MDS includes increased time to conversion to acute myeloid leukemia (AML), better quality of life, and improved survival when compared to conventional therapy [36, 37].

In addition to the DNA modifiers, numerous histone modifying enzymes have been targeted. In fact, A pair of HDAC inhibitors, vorinostat and romidepsin, have proven beneficial for treatment of cutaneous T-cell lymphoma and have been approved for use in humans [38, 39]. Furthermore, small-molecule inhibitors of HMTs are of tremendous interest. This interest has led to the discovery of EPZ-5676 and EPZ-6438 (tazemetostat)

which inhibit DOT1L and EZH2 respectively and are currently in clinical trials [40-42]. Of note, these two enzymes place marks with diametrically opposed functions. DOT1L catalyzes methylation of H3K79, which is considered an activating mark, while EZH2 is responsible for H3K27 methylation, which is a well-known repressive mark. Beyond the methyltransferases, demethylases are also under investigation. In particular, Lysine Specific Demethylase 1 (LSD1; also known as KDM1A), which has been implicated in leukemia and cancers of the prostate, colon, and breast, has been the subject of numerous studies [43-46]. While modulators of histone modifying enzymes have generated great interest and have well-documented benefits, there remain questions regarding specificity when targeting enzymes that are not unique to disease.

Screening Approaches

While there are many strategies available for targeting transcription factors, there are even more screening methods that can be utilized to find active small-molecules. Depending on the biological context, availability of information regarding molecular mechanism, availability of high-resolution structural data about the target, and overall goals of the project, the approach to drug discovery can be based on either forward or reverse chemical genetics [47]. Forward chemical genetic approaches are largely phenotype or function driven and frequently make use of cell-based assays for screening. While forward chemical genetic approaches are good at identifying molecules that produce the desired responses, they require follow-up experiments to determine the target and mechanism of action. Reverse chemical genetics approaches are ordinarily centered on engagement with the target and routinely rely on biochemical or biophysical

binding studies. While reverse chemical genetic approaches identify molecules with a high affinity for the target of interest, they require follow-up experiments to confirm intracellular target engagement, assess the functional consequences related to this interaction, and to evaluate cytotoxicity. In any case, whether forward or reverse chemical genomics approaches are desired, there are a multitude of screening platforms that have been successfully employed for the identification of efficacious small-molecules.

Target-based screening methods are terrific for small-molecule drug discovery when there is a well-defined activity or interaction that could be manipulated in order to get the desired response. Often, target-based screening approaches utilize biochemical or virtual screening. Biochemical screening typically involves enzymatic activity assays or direct binding assays. These assays come in many forms, but some of the most commonly used ones are fluorescence polarization (FP), enzyme-linked immunosorbent assay (ELISA), Förester/fluorescence resonance energy transfer (FRET), and AlphaLISA [48]. It is important to note that biochemical screening is performed using purified proteins in the absence of cellular context. Thus, these assays face serious limitations that must be considered before they are used. For example, the intracellular target engagement and consequences thereof are impossible to determine using biochemical screening approaches. These are not fatal flaws as they can be investigated after hits have been identified, and some chemical problems may even be overcome with medicinal chemistry help.

Phenotype-based screening overcomes many of the challenges of biochemical screening. These screening approaches generally rely on cellular or organism-based assays. These assays are not cheap or trivial to set up and run, but they provide lead

compounds which have greater translational potential. This is because these assays assess compounds in the context of a fully functional cell and, therefore, more closely resemble the *in vivo* setting [49]. One benefit of phenotype-based assays is that they allow many targets to be gauged at once. The downside to this is that there are many possible targets for any particular hit compound, which makes determining a mechanism of action much more challenging and opens the door to increased potential for off target effects. This task is becoming less and less daunting as new methods for identifying the target(s) of candidate molecules are being developed and utilized [50]. Another challenge facing cell-based assays is cell permeability, which could mask important information from compounds that cannot penetrate cells but would otherwise be active [51].

Understanding the benefits and challenges of each of these screening methods is vital for selecting the most appropriate method for the given circumstance and is essential for completing a successful screening campaign. As such, careful consideration must be taken when deciding on a screening platform. In an attempt to maximize our chances of finding novel small-molecules that regulate Pax2 transactivation I decided to utilize both a target-centered approach and a phenotype-based approach. By taking these two approaches I expect to identify small-molecules that modulate Pax2 activity both directly and indirectly. I chose to employ a molecular docking based virtual screen for the target-centered approach and a cell-based reporter assay for the phenotypic screen. Some of the main reasons that these methods were selected is because there were high-resolution crystal structures of several Pax family paired domains [52-54], information from mutation studies suggesting a specific region of the paired domain may be amenable to small-

molecule inhibition [55, 56], and a very powerful reporter cell line, which had provided in-depth mechanistic knowledge of Pax2 transcription activation [56-61].

A DNA motif recognized by the paired domain provided critical insight into Pax biology and

In the early 1990s, studies investigating the DNA-binding specificity of Pax1 and Pax2 identified a nucleotide sequence which was termed the paired domain recognition site 4 (PRS4) [62, 63]. Identification of this binding site proved invaluable for furthering our understanding of Pax2 biochemistry. Integration of this element into reporter cells was instrumental for deciphering Pax2 mediated regulation of transcription [56-61]. These studies found that Pax2 can regulate gene expression positively or negatively depending on the cellular context. Pax2 can positively regulate gene expression through its connection with the MLL complex, which places activating methylation marks and enhances transcription (Figure 2.1). The connection to the MLL complex is facilitated by the adapter protein PTIP. The Pax2-PTIP interaction is potentiated by Wnt signaling and JNK mediated phosphorylation of Pax2. On the other hand, in the presence of high levels of the corepressor protein Grg4, PTIP and the MLL complex can be displaced (Figure 2.2). This displacement is a consequence of Grg4 recruitment of the phosphatase PPM1B and subsequent dephosphorylation of Pax2. Concomitantly, Grg4 recruits the PRC2 and the arginine methyltransferase, PRMT5, which place repressive methylation marks and inhibits transcription.

It is evident that this reporter has been a useful tool in dissecting how Pax2 regulates target gene expression at the molecular level. Due to the extensive

characterization and my familiarity with this reporter, I sought to co-opt it for use in my small-molecule screening campaign. To accomplish this, the PRS-EGFP reporter that was previously used to describe Pax2 mediated transcription activation was cloned, the EGFP cassette was exchanged for a luciferase cassette, and the new construct was put back into HEK293 cells. As previously described, under normal circumstances these cells do not express Pax proteins, and as a result, the reporter gene exhibits only a low level of basal expression. However, transient transfection of these cells with a Pax2 expression vector leads to a marked increase in reporter gene expression. These Pax responsive cells provide an excellent way for us to monitor Pax2 transcription activation.

Results

Generation and screening of Pax2 reporter cell lines identifies a responsive line, PRS4-Luc clone #33

Oligonucleotides containing five copies of the PRS4 element were synthesized and cloned upstream of a Thymidine Kinase (TK) promoter and a luciferase gene in an expression vector containing a neomycin resistance cassette. This vector was incorporated into HEK293 cells through transfection. Following transfection these cells were placed under antibiotic selection to generate stably transfected cells. These cells were subjected to limiting dilution to generate clonal lines. The clonal lines were expanded and screened for their potential to respond to transient Pax2 transfection (Figure 2.3).

I sought to identify a PRS-Luc cell line that exhibited a large change in luciferase expression following transient Pax2 transfection. Numerous PRS-Luc clones were screened and several showed an increase in luciferase expression following transient

Pax2 transfection. One cell line, PRS-Luc clone #33, stood out. This cell line had a relatively low basal level of luciferase expression and exhibited just over a fivefold change in luminescence when comparing cells transfected with wildtype or mutant Pax2 (Figure 2.3 A-B). While this was more than twice the level of activation that was seen in any of the other lines tested, it is still less than I was aiming for to ensure a quality HTS assay. Given this finding, I decided to proceed with assay development using this cell line but was interested in finding ways to improve the level of activation.

Comparison of transfection reagents leads to selection of Lipofectamine 2000 as the reagent of choice for our reporter assay

Having identified a PRS-Luc reporter cell that responded well to transient Pax2 transfection, I next sought to identify ways to improve the level of activation. As a first step, I began comparing the level of activation following transfection with an array of transfection reagents. I was fortunate to receive samples of several reagents including Effectene and PolyFect from Qiagen, TransIT-LT1 and TransIT-293 from Mirus Bio, GeneJuice from EMD Millipore, PureFection from System Biosciences, and Lipofectamine 2000 from Thermo Fisher Scientific. These reagents along with FuGENE 6, which the Dressler lab had been previously using, were used to transiently transfect PRS-EGFP cells with Pax2. At this point, all of the reagents were being used according to the manufacturer's recommendations. Following transfection, the EGFP levels were investigated using either flow cytometry or Western blotting (Figure 2.4).

The flow cytometry results show that, depending on the reagent used, between 7.5% and 36.6% of the cells analyzed exhibited an increase in green fluorescence as

indicated by their shift to the right of the plot (Figure 2.4 A). This number is representative of the transfection efficiency. The top 3 reagents were FuGENE 6, Lipofectamine 2000, and PolyFect which showed 29.1%, 31.4%, and 36.6% efficiency respectively. To confirm these results, cells were once again transfected using the various reagents, and whole cell lysates were generated from these cells. The lysates were used for Western blot analysis which confirmed that FuGENE 6, Lipofectamine 2000, and PolyFect led to the greatest levels of EGFP production (Figure 2.4 B).

Considering that these three reagents performed at a comparable level, I calculated the cost per transfection for each of the reagents to help determine which reagent I would ultimately use (Figure 2.4 C). Based on the volume of reagent needed per plate and the cost per milliliter of reagent I determined that it would cost $5.18 per transfection if Lipofectamine 2000 was used, $6.46 per transfection if FuGENE 6 was used, and $11.20 per transfection if PolyFect was used. While the PolyFect reagent led to the greatest transfection efficiency, the small difference in efficiency did not justify the large increase in cost per transfection. Thus, Lipofectamine 2000 was selected as the transfection reagent of choice for the HTS assay.

Determination of optimal conditions for PRS4-Luc reporter assay

Having settled on a transfection reagent, I next set out to optimize the parameters of the HTS assay. The goal for the following experiments was to determine the conditions that would generate the most robust and reproducible assay possible. Using the PRS-Luc cell line, I optimized multiple parameters to improve the luminescence assay (Figure 2.5).

To enhance transfection efficiencies, I started by varying the transfection reagent to plasmid ratio. I found that increasing the ratio from 3:1 to 3.5:1 led to an increase in the level of luciferase activation. Furthermore, I found that decreasing the ratio from 3:1 to 2.5:1 led to a similar decrease in the level of luciferase activation (Figure 2.5 A). Based on these findings, I determined that having a slight increase in the amount of transfection reagent would be best going forward.

Next, I began to look at pre-transfection seeding densities. While previous transfection experiments were performed following visual inspection of confluence, I set out to standardize the number of cells that were present at the time of transfection. I thought that this would limit some of the variability by ensuring there were always the same number of cells being subjected to the transfection. On the afternoon of the day before transfection, I plated 3, 6, or 9 million cells in a 100 mm cell culture dish. The next day I proceeded with transfecting the cells as described previously. By visually inspecting the plates the following morning I determined that these numbers corresponded to roughly 60%, 80%, and 100% confluence respectively. I found an inverse relationship between the pre-transfection seeding density and the level of luciferase activation. As the pre-transfection seeding density increased the level of luciferase activation decreased (Figure 2.5 B). Thus, I determined that I would perform transfections on plates that were seeded with 3 million cells on the day before transfection.

The next parameter that I optimized was post-transfection seeding density. I suspected that I would need to balance the increase in luciferase activation associated with an increased cell number with the decrease in luciferase activation linked to reduced cellular activity caused by contact inhibition. Not surprisingly, that was what I found. While

increasing the post-transfection seeding density from 20,000 cells/well to 30,000 cells/well did produce an increase in luciferase activation, a further increase in seeding density to 40,000 cells/well actually led to a level of luciferase activation that was less than what was observed when 20,000 cells were plated per well (Figure 2.5 C). This led me to selecting 30,000 cells/well as the standard for the assay.

Another parameter that I thought might affect the level of luciferase activation that was able to be achieved was post-transfection incubation time. Since the assay depended on transiently transfected cells taking up the plasmid DNA, transcribing the DNA into mRNA that could subsequently be translated into protein, binding of this newly synthesized Pax2 protein to the PRS element, recruitment of cofactors required for transcription of the luciferase reporter gene, and translation of this product into the active enzyme, I was not sure how long it would take to obtain optimal activation. I investigated the level of luciferase activation at 24 h, 36 h, and 48 h after transfection and found an inverse relationship between the post-transfection incubation time and the level of luciferase activation. As the post-transfection incubation time increased, the level of luciferase activation decreased (Figure 2.5 D). These findings suggested that it was best to keep the incubation time to 24 h.

Since the previous optimization experiments only produced minor changes in the level of luciferase activation, I sought additional ways of improving our assay. The fact that reducing the number of cells prior to transfection led to an increased level of luciferase activation suggested that perhaps I was only achieving low level transfection efficiency because I was not using enough plasmid. To address this potential issue, I prepared a series of transfections in which the amount of plasmid was titrated from one

to ten micrograms. Indeed, a substantial increase in the level of luciferase activation was obtained by increasing the amount of plasmid. The increase appeared to level off after 4 μg and actually began to dip as I reached 10 μg (Figure 2.5 E). These results suggested that too much foreign DNA or too much of the transfection reagent were cytotoxic. Furthermore, these results suggest that I was achieving near complete transfection efficiency when 4 μg of plasmid was used per transfection. Because there were only minimal gains in luciferase activation when the amount of plasmid was increased beyond 4 μg and higher doses appeared to be cytotoxic, I thought it best to proceed using only 4 μg of plasmid per plate of cells transfected.

While the parameters were optimized to produce the greatest level of luciferase activation, I noticed that leaving the transfection reagent on the cells overnight, as suggested in the manufacturers recommended protocol, led to some of the cells detaching from the plate and others looking generally unhealthy. To avoid this problem, I thought that I might be able to perform the transfections over a shorter time period and still obtain a large enough luciferase induction to develop a quality HTS assay. For convenience sake, I reduced the amount of time that the transfection reagent was left on the cells to only six hours. This change allowed us to transfect cells in the morning and harvest them in the afternoon for use in our HTS assay. Importantly, this change only led to a small difference in the level of luciferase induction.

Adaptation of the optimized PRS4-Luc reporter assay to HTS format demonstrates development of a robust and reproducible assay

Following the assay optimization steps outlined previously, I was satisfied that I had developed a robust and reproducible assay that was primed for adapting to HTS. To adapt the assay to an HTS format, I simply moved from 96-well cell culture dishes to 384-well cell culture dishes by scaling down the number of cells from 30,000 cells/well to 6,000 cells/well. To determine if the assay was ready for HTS, I performed mock runs with test plates (Figure 2.6). For these plates, I followed all of the steps that I would normally do for the HTS, but instead of adding compounds to the wells only DMSO (vehicle) was added. These runs allowed me to investigate well-to-well variability, determine the level of luciferase induction, and calculate Z'. Z' is a dimensionless statistical parameter used for evaluating HTS assays [64]. It is a parameter that is dependent on the variation between measurements (well-to-well variability) and the dynamic range of the assay (the separation between the positive and negative controls). Calculating Z' provides a number between 0 (meaningless assay) and 1 (perfect assay). For an assay to be considered ready for HTS, a Z' of between 0.5 and 0.8 should be obtained. Assays scoring below 0.5 would make it difficult to discriminate between true hits and random variability while assays scoring over 0.8 could be overdeveloped. Importantly, I found that the assay generated a Z' of 0.71, a score indicative of a high-quality HTS assay and a score seldom observed with cell-based HTS assays (Figure 2.6 B). Furthermore, these test plates showed me that the well-to-well variability was relatively low. In fact, no wells exhibited greater than 19% difference. Notably, no false positives were identified in the test plate. This suggests that I should expect to have a low level of false positive hits in our screen. I was highly satisfied with the HTS assay that had been developed and was ready to put it to use.

Discussion

Pax2 is overexpressed or ectopically expressed in multiple proliferative disorders of the kidney including polycystic kidney disease [65, 66], juvenile cystic and dysplastic diseases [67], Wilms' tumor [2, 3], and renal cell carcinoma [4-6]. In these disorders Pax2 encourages proliferation and endows tumor cells with resistance to apoptosis and chemotherapy. Decreasing Pax2 levels in mouse models of PKD cut down on cyst formation and slowed disease progression [65, 66]. In addition, renal cancer cell lines are sensitized to chemotherapeutic agents, exhibit an increase in apoptosis, and have a reduced proliferative capacity following suppression of Pax2 gene expression [4, 7, 8]. Together these findings make Pax2 an attractive therapeutic target.

While small-molecule drug discovery for transcription factors is not an easy task, the benefit that such molecules could provide makes this a worthwhile venture. Transcription factors are activated by a complex network of signaling pathways. Following activation these factors localize to DNA and recruit cofactors, often epigenetic complexes, to modify gene expression. This provides numerous opportunities for intervention. Whether targeting the upstream signaling, the transcription factor, recruitment of cofactors, or the downstream epigenetic modifiers is desired, there are many screening platforms that could be employed. I considered both direct and indirect methods for targeting Pax2 and weighed the options for screening methods. Ultimately, I decided to use two approaches: 1) unbiased HTS and 2) molecular docking based virtual screening. This decision was predicated on the availability of and my familiarity with the PRS4 reporter construct. The unbiased HTS allows us to screen for molecules with the potential

to alter Pax2 activity at any step in the process which greatly increases our chances of identifying a promising lead compound. Any hit identified by the HTS would have to be further validated, and a mechanism of action would have to be determined. To offset some of this work and to increase the chances of identifying direct inhibitors of Pax2 I decided to develop a virtual screening approach. Since I was already developing an assay to assess *in vitro* activity of small-molecules for the HTS campaign and because high-resolution crystal structures of several Pax paired domains were available, the virtual screen added little to the upfront workload but significantly narrowed the scope of mechanistic studies for hits identified through this approach.

In this chapter, I established a Pax2 responsive reporter cell line. This cell line contains a PRS4 binding site upstream of a gene coding for luciferase. Upon transient transfection with a Pax2 expression vector, the reporter is activated and an increase in luminescent signal is observed. Several parameters were investigated and optimized to maximize luciferase induction. Following optimization, this assay was miniaturized so that it was in an HTS compatible format. This assay was tested and found to be robust and reproducible, as indicated by a Z' of 0.71. A Z' of this magnitude is not typically achieved for cell-based assays, and it suggests that our assay is ready for use.

Despite the inherent difficulty in discovering small-molecule inhibitors of Protein-DNA interactions, the demand for such molecules is very likely to continue to rise based on the sheer volume of disease mechanisms driven by these types of interactions. Just as protein-protein interactions were avoided until several successful campaigns were performed, I suspect that Protein-DNA interactions will become more popular targets as success in this area grows.

Materials and Methods

Cell culture

HEK293 cells containing a PRS-EGFP or PRS-Luc reporter, described previously [60], were cultured in Dulbecco's Modified Eagle Medium (DMEM, Gibco) supplemented with 10% fetal bovine serum (Sigma), 100 U/mL penicillin, 100 ng/mL streptomycin (Gibco), and 400 ng/mL geneticin (Gibco) and were maintained under humidified 5% CO2/95% air at 37 ^{o}C.

Transfection protocols

Initial transfection experiments were performed according to individual manufacturers protocols. HEK293 cells were transiently transfected with CMV-Pax2b or CMV-Pax2b^{C63Y} expression vector. Briefly, 3x10^{6} cells were plated in a 100 mm tissue culture dish (Corning) and incubated overnight. The next morning the media was replaced with 5 mL of antibiotic free DMEM, and transfection reactions were prepared. In a 2 mL Eppendorf tube containing 896 µL of Opti-MEM I (Gibco), 4 µL of 1µg/µL plasmid was added. In a second 2 mL Eppendorf tube containing 886 µL Opti-MEM I, 14 µL of Lipofectamine2000 (Life Technologies) was added and allowed to incubate at room temperature for 5 min. The contents of the two tubes were mixed by pipetting and allowed to incubate at room temperature for 20 min. The transfection mix was added dropwise to a 100 mm dish of cells. The plate was shaken gently and placed in the incubator for 6 h. The plate of transfected cells was trypsinized with 0.05% trypsin (Gibco) for 3 min. The cells were harvested, counted, plated in 6-well culture dishes (500,000 cells/well for

Western blot analysis) or 96-well dishes (30,000 cells/well for luminescence assays), and were placed in the incubator overnight.

Flow Cytometry

Transiently transfected PRS4-EGFP cells were washed with PBS (Gibco) and trypsinized with 0.05% trypsin (Gibco). Cells were resuspended in DMEM with 10% FBS and centrifuged for 5 min. at 400g. The media was removed and the cell pellet was resuspended in ice cold 0.3% BSA (Proliant) in PBS. The centrifugation and resuspension steps were repeated and cells were filter through a 70 µm mesh to remove clumps. Propidium iodide (PI, Thermo) was added to a final concentration of 10 µg/mL before sorting cells on FACS Aria II (BD). The cells were sorted based on viability dye (PI) and EGFP. Cells were collected into 1.5 ml tubes containing PBS and viewed to ensure EGFP expression.

Western blot analysis

For Western blot analysis cell culture media was removed from test plates, and cells were lysed with in 2x SDS buffer (20% glycerol, 4% sodium dodecyl sulfate, 0.2 M dithiothreitol, 125 mM Tris, pH 6.8), as described [68]. Total protein was separated on a 4-12 % Bis-Tris gels (Life technologies), transferred to PVDF membranes, and immunoblotted with anti-Pax2 [2] or anti-EGFP antibody (Santa Cruz).

Luminescence assay

For luciferase reporter assays, test plates were removed from the incubator and equilibrated to room temperature for ten minutes before media was reduced and an equal volume of SteadyGlo (Promega) was added. The plates were covered to protect them from light and incubated at room temperature for an additional ten minutes. Plates were subsequently read on a PHERAstar multimode plate reader (BMG).

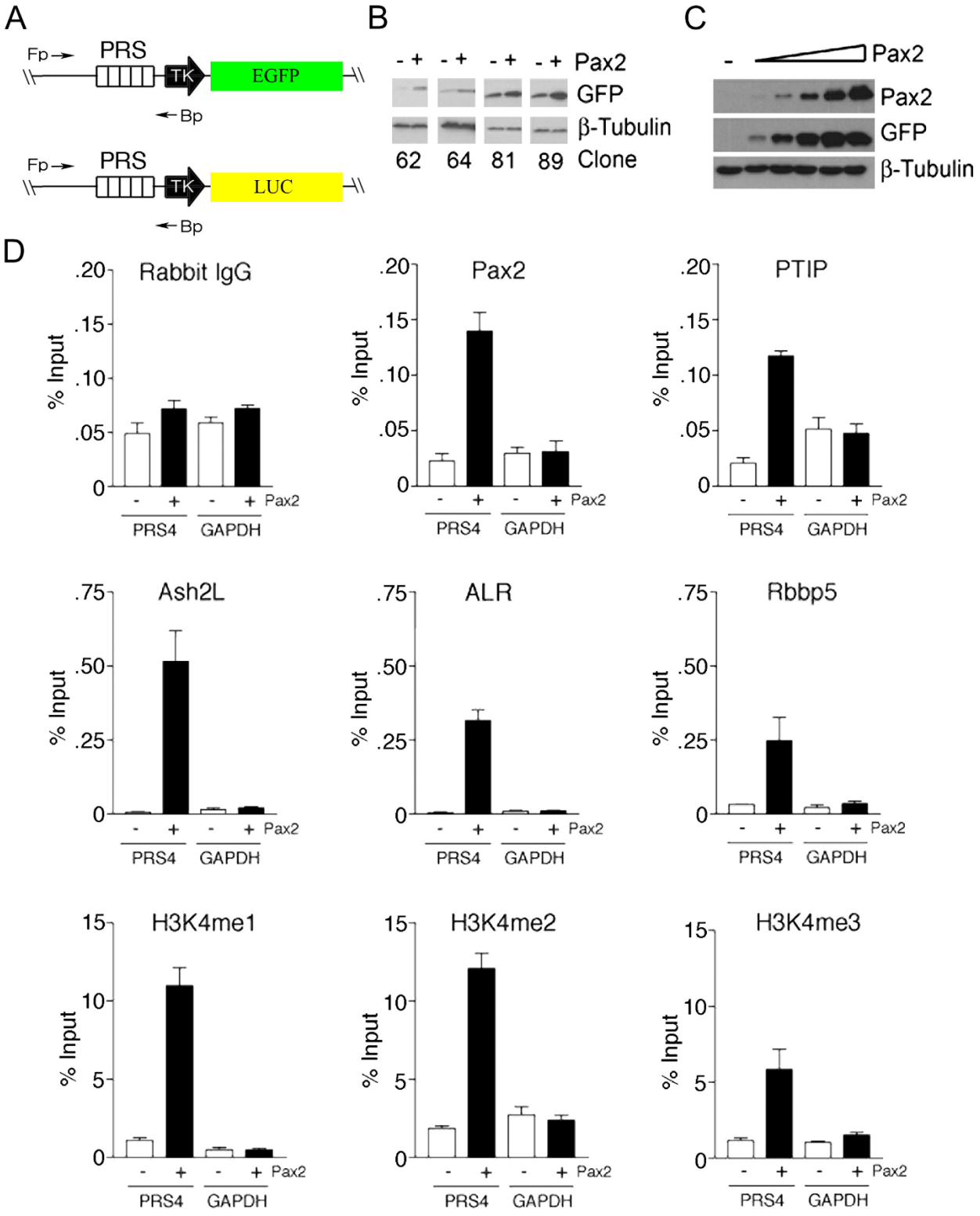


Figure 2.1. Characterization of PRS reporter activation. A) Schematic of Pax responsive reporter genes: PRS, Paired domain recognition site; TK, minimal HSV-TK promoter; EGFP, enhanced green fluoresencent protein coding; Luc, luciferase coding; Fp, forward primer used in ChIP analysis; Bp, reverse primer used in ChIP analysis. **B)**

Clones carrying the PRS-EGFP integrated reporter are responsive to transient Pax2 transfection. Whole cell lysates were blotted for the proteins indicated. **C)** HEK293 cells containing the integrated PRS-EGFP reporter that were transiently transfected with increasing amounts of Pax2 show dose dependent levels of EGFP. Whole cell lysates were blotted for the proteins indicated. **D)** ChIP from cells with or without Pax2 transfection. Primer pairs against the PRS and a control promoter (GAPDH) were used for RT-qPCR. Relative amounts of PCR product are expressed as a percent of input chromatin. Notably, Pax2 binds to the PRS element. Furthermore, Pax2 dependent recruitment of PTIP, ASH2L, ALR, and RBBP5 leads to an increase in H3K4 mono-, di-, and trimethylation. Averages from triplicates are shown with error bars indicating 1 SEM. This figure is modified from Patel, S.R. et al. *Dev Cell*, 2007 [59].

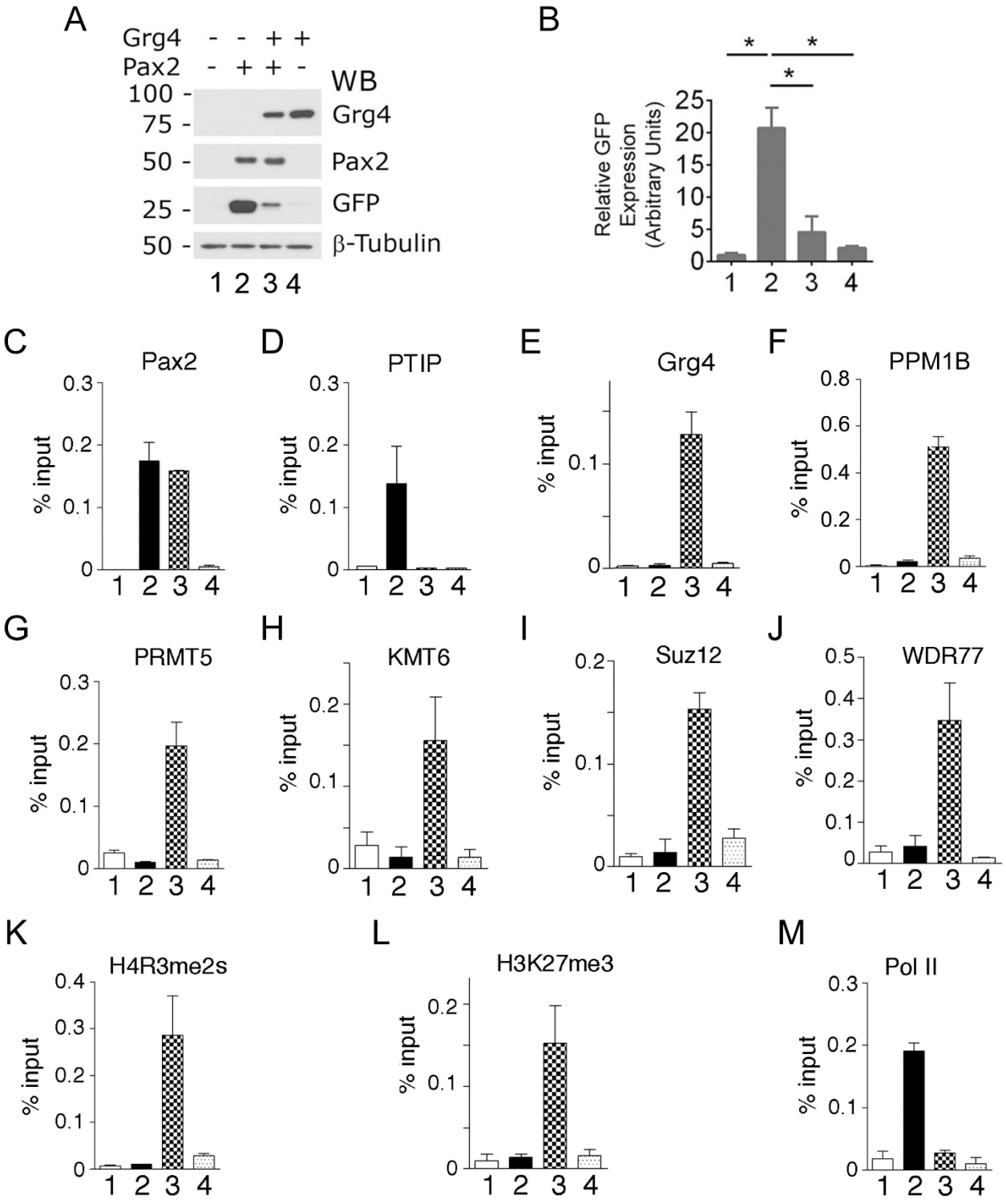
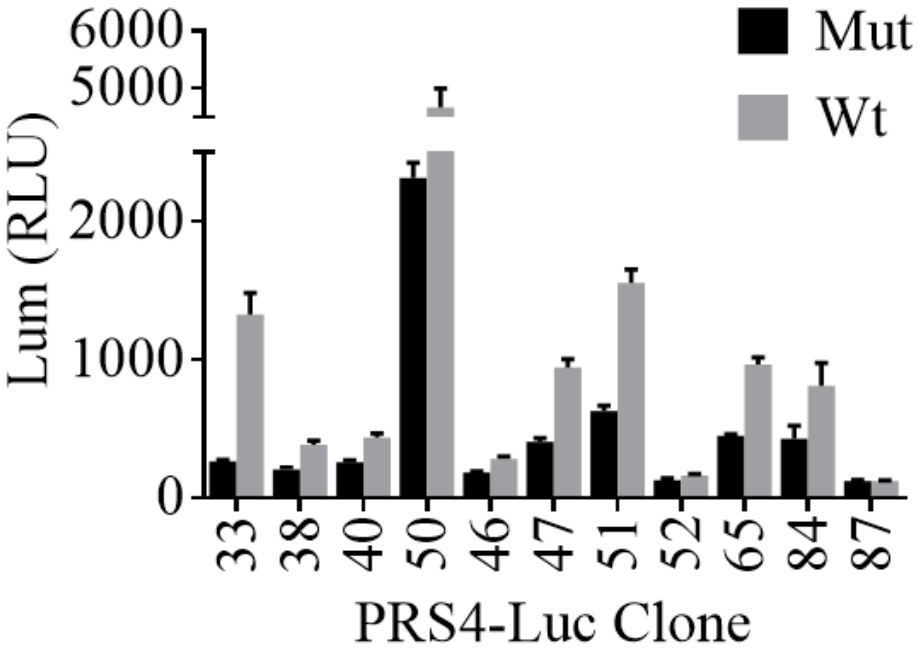


Figure 2.2. Characterization of PRS reporter repression. A) HEK293 cells containing the integrated PRS-EGFP reporter were transiently transfected with control vector, Pax2 expression vector, Pax2 and Grg4 expression vectors, or just Grg4 expression vector. Whole cell lysates were collected and blotted for the proteins indicated. **B)** Densitometry

based quantification of GFP expression in cells transfected as described in A. Cells transfected with the control vector were used to normalize results. Note the significant increase in GFP expression when Pax2 is transfected alone and the suppression of this increase in Pax2-Grg4 co-transfected cells. *, $p<0.01$ **C-M)** ChIP assays from PRS-EGFP cells transfected as described in A. Antibodies used for each ChIP experiment are shown above their respective graphs. Averages from triplicate experiments are shown with error bars indicating 1 SEM. **C)** Note that Pax2 is localized to the PRS4 element in cells transfected with a Pax2 expression vector. **D)** PTIP is found at the PRS4 element in a Pax2-dependent manner. Notably, PTIP localization is disrupted in cells co-transfected with Pax2 and Grg4 expression vectors. **E)** Note that Grg4 localization to the PRS4 element occurs in a Pax2-dependent manner. **F-J)** The localization of Grg4 to the PRS4 element is accompanied by PPM1B (phosphatase), PRMT5 (arginine methyltransferase), KMT6 (EZH2, PRC2 catalytic subunit), Suz12 (PRC2 subunit), and WDR77 (PRC2 subunit). **K-L)** The recruitment of the factors shown in F-J leads to symmetric dimethylation of H3R4 and trimethylation of H3K27. **M)** RNA Pol2 recruitment is increased in cells transfected with a Pax2 expression vector, but this increase is abolished when Grg4, PPM1B, PRMT5, PRC2, and their associated methylation marks are present. **K)** This figure is modified from Patel, S.R. et al. *Mol Cell*, 2012 and Abraham, S. et al. *J Biol Chem*, 2015 [60, 61].

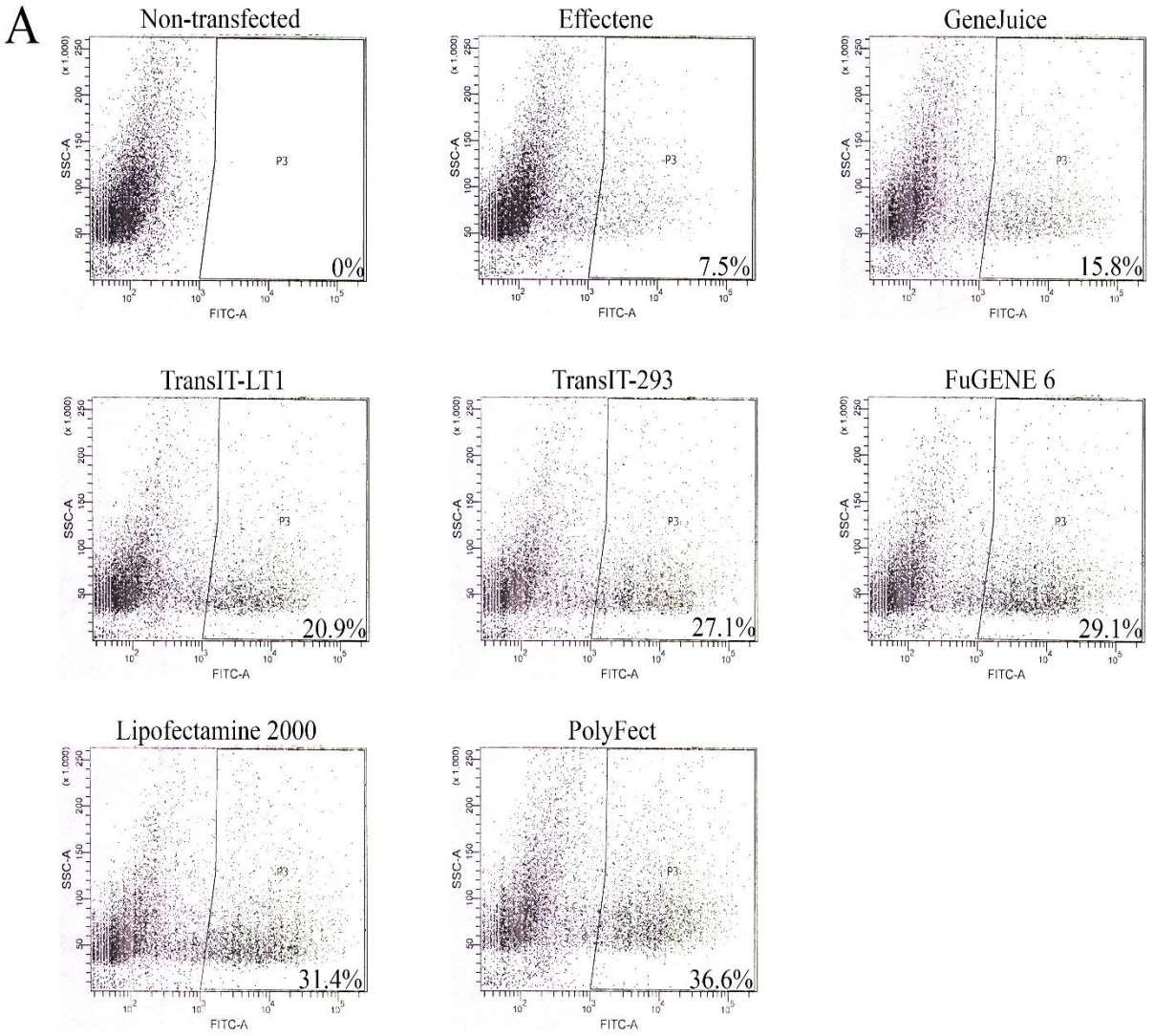
A

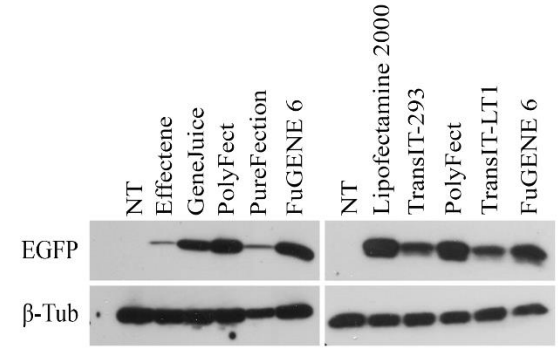


B

PRS4-Luc Clone	Luciferase Activation (Fold change)
33	5.04 ± 0.11
38	1.85 ± 0.10
40	1.70 ± 0.08
46	1.53 ± 0.06
47	2.33 ± 0.08
50	2.02 ± 0.08
51	2.47 ± 0.08
52	1.27 ± 0.13
65	2.15 ± 0.05
84	1.89 ± 0.27
87	1.00 ± 0.05

Figure 2.3. Response of PRS-Luc cell lines to transient Pax2 transfection. PRS-Luc cell lines were transiently transfected with either CMV-Pax2b or CMV-Pax2b^{C63Y} and luciferase induction was monitored using the SteadyGlo luminescence assay. **A)** Graph of raw luciferase activity assay results **B)** Table indicating luciferase induction based on the fold change observed between cells transfected with Wt and Mut Pax2b. I would like to thank Egon Ranghini for generating the stably transfected PRS4-Luc clones that I screened for Pax2 responsiveness.





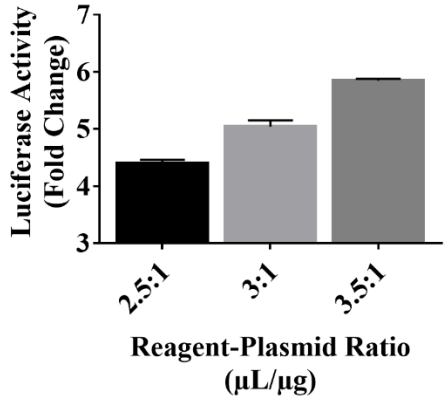
C

Reagent	Unit Cost ($/mL)	Volume of Reagent per Transfection (µL)	Number of Transfections per Unit	Cost per Transfection ($)
Effectene	278	60	16.67	16.68
GeneJuice	323	18	55.56	5.81
TransIT-LT1	210	18	55.56	3.78
TransIT-293	249	18	55.56	4.48
FuGENE6	359	18	55.56	6.46
Lipofectamine 2000	288	18	55.56	5.18
PolyFect	140	80	12.50	11.20

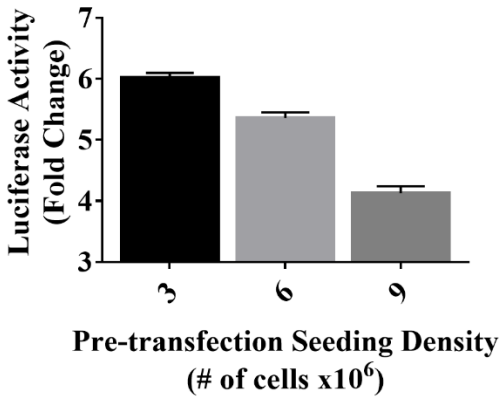
Figure 2.4. Comparison of transfection efficiency for several transfection reagents. Numerous transfection reagents used to transiently transfect PRS-Luc cell lines with CMV-Pax2b to investigate transfection efficiency. **A)** Flow cytometric analysis

of EGFP induction **B)** Western blot analysis of EGFP induction. **C)** Cost analysis for the various transfection reagents used.

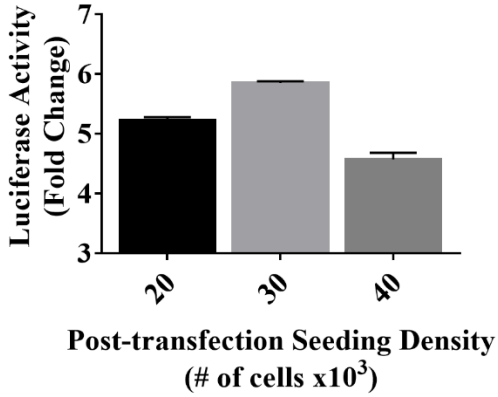
A


Luciferase Activity
(Fold Change)
7
6
5
4
3
2.5:1
3:1
3.5:1
Reagent-Plasmid Ratio
(µL/µg)

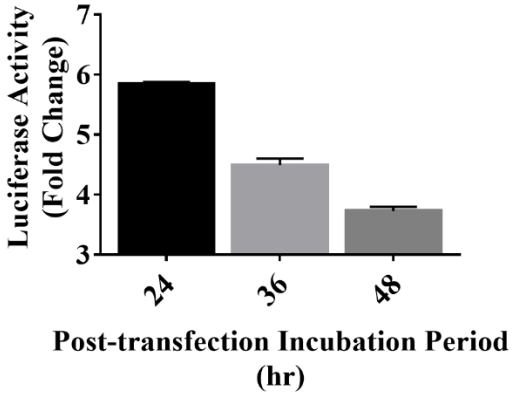
B


Luciferase Activity
(Fold Change)
7
6
5
4
3
3
6
9
Pre-transfection Seeding Density
(# of cells x10^6)

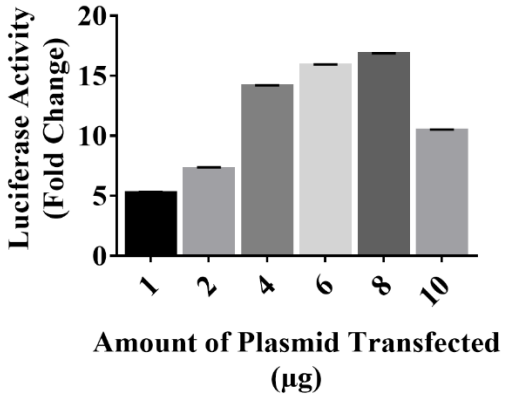
C


Luciferase Activity
(Fold Change)
7
6
5
4
3
20
30
40
Post-transfection Seeding Density
(# of cells x10^3)

D


Luciferase Activity
(Fold Change)
7
6
5
4
3
24
36
48
Post-transfection Incubation Period
(hr)

E


Luciferase Activity
(Fold Change)
20
15
10
5
0
1
2
4
6
8
10
Amount of Plasmid Transfected
(µg)

F

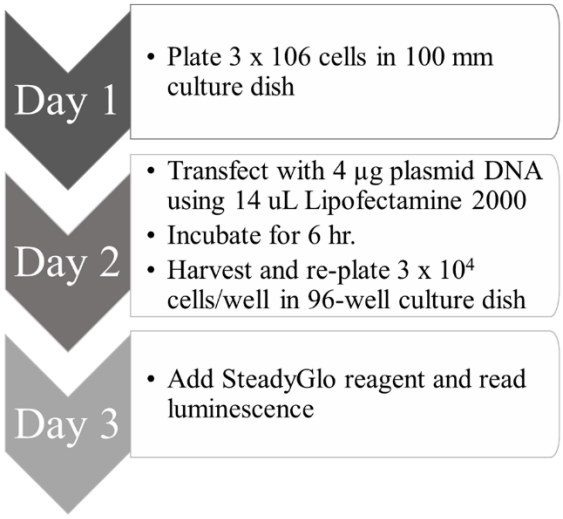
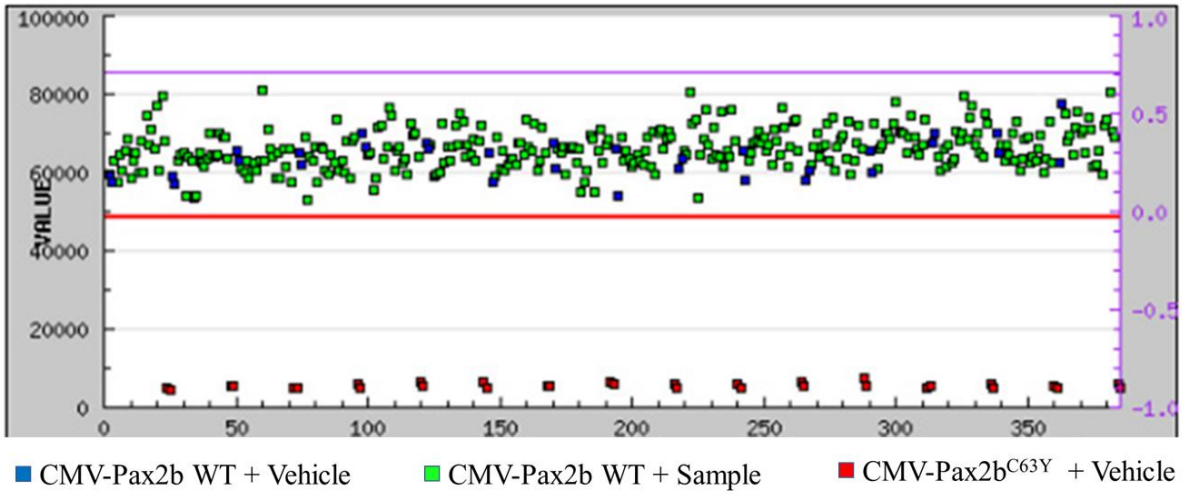

Day 1
• Plate 3 x 106 cells in 100 mm culture dish
Day 2
• Transfect with 4 µg plasmid DNA using 14 uL Lipofectamine 2000
• Incubate for 6 hr.
• Harvest and re-plate 3 x 10^4 cells/well in 96-well culture dish
Day 3
• Add SteadyGlo reagent and read luminescence

Figure 2.5. Optimization of PRS-Luc assay. PRS-Luc clone #33 cells were used for comparison of luciferase induction **A)** following transient transfection of CMV-Pax2b using different ratios of transfection reagent to DNA, **B)** at assorted pre-transfection seeding densities, **C)** at various post-transfection seeding densities, **D)** at several timepoints following transient CMV-Pax2b transfection, **E)** and following transient transfection with varying amounts of CMV-Pax2b expression plasmid. **F)** Diagram describing optimized assay workflow.

A



B

Z'	Hits	Total	< 0%	0 - 19%	20-39%	40-59%	60-79%	80-100%	>100%
0.71	0	320	219	101	0	0	0	0	0

Figure 2.6. Adaptation of PRS-Luc assay to HTS format. A mock HTS run was performed. In this mock run, all of the wells that would normally have had test compound added received only DMSO. The 384-well test plate was generated and PRS-Luc clone #33 cells which were transiently transfected with CMV-Pax2b or CMV-Pax2b^{C63Y} were added. **A)** Screenshot depicting raw luminescence values for the test plate. **B)** Table of test plate statistics.

Bibliography

1. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.

2. Dressler, G.R. and E.C. Douglass, *Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor.* Proc Natl Acad Sci U S A, 1992. **89**(4): p. 1179-83.

3. Eccles, M.R., et al., *Expression of the PAX2 gene in human fetal kidney and Wilms' tumor.* Cell Growth Differ, 1992. **3**(5): p. 279-89.

4. Gnarra, J.R. and G.R. Dressler, *Expression of Pax-2 in human renal cell carcinoma and growth inhibition by antisense oligonucleotides.* Cancer Res, 1995. **55**(18): p. 4092-8.

5. Tong, G.X., et al., *PAX2: a reliable marker for nephrogenic adenoma.* Mod Pathol, 2006. **19**(3): p. 356-63.

6. Tong, G.X., et al., *PAX8 and PAX2 immunostaining facilitates the diagnosis of primary epithelial neoplasms of the male genital tract.* Am J Surg Pathol, 2011. **35**(10): p. 1473-83.

7. Hueber, P.A., et al., *PAX2 inactivation enhances cisplatin-induced apoptosis in renal carcinoma cells.* Kidney Int, 2006. **69**(7): p. 1139-45.

8. Hueber, P.A., et al., *In vivo validation of PAX2 as a target for renal cancer therapy.* Cancer Lett, 2008. **265**(1): p. 148-55.

9. Darnell, J.E., Jr., *Transcription factors as targets for cancer therapy.* Nat Rev Cancer, 2002. **2**(10): p. 740-9.

10. Hopkins, A.L. and C.R. Groom, *The druggable genome.* Nat Rev Drug Discov, 2002. **1**(9): p. 727-30.

11. Imming, P., C. Sinning, and A. Meyer, *Drugs, their targets and the nature and number of drug targets.* Nat Rev Drug Discov, 2006. **5**(10): p. 821-834.

12. Huang, W., et al., *A small molecule compound targeting STAT3 DNA-binding domain inhibits cancer cell proliferation, migration, and invasion.* ACS Chem Biol, 2014. **9**(5): p. 1188-96.

13. Huang, W., et al., *Small-molecule inhibitors targeting the DNA-binding domain of STAT3 suppress tumor growth, metastasis and STAT3 target gene expression in vivo.* Oncogene, 2016. **35**(6): p. 783-92.

14. Berg, T., *Inhibition of transcription factors with small organic molecules.* Curr Opin Chem Biol, 2008. **12**(4): p. 464-71.

15. Kong, D., et al., *Echinomycin, a small-molecule inhibitor of hypoxia-inducible factor-1 DNA-binding activity.* Cancer Res, 2005. **65**(19): p. 9047-55.

16. Chapman-Smith, A. and M.L. Whitelaw, *Novel DNA binding by a basic helix-loop-helix protein. The role of the dioxin receptor PAS domain.* J Biol Chem, 2006. **281**(18): p. 12535-45.

17. Vlaminck, B., et al., *Dual effect of echinomycin on hypoxia-inducible factor-1 activity under normoxic and hypoxic conditions.* FEBS J, 2007. **274**(21): p. 5533-42.

18. Adhikary, S. and M. Eilers, *Transcriptional regulation and transformation by Myc proteins.* Nat Rev Mol Cell Biol, 2005. **6**(8): p. 635-45.

19. Cowling, V.H. and M.D. Cole, *Mechanism of transcriptional activation by the Myc oncoproteins.* Semin Cancer Biol, 2006. **16**(4): p. 242-52.

20. Xu, Y., et al., *A credit-card library approach for disrupting protein-protein interactions.* Bioorg Med Chem, 2006. **14**(8): p. 2660-73.

21. Lu, X., et al., *Disruption of the MYC transcriptional function by a small-molecule antagonist of MYC/MAX dimerization.* Oncol Rep, 2008. **19**(3): p. 825-30.

22. Wang, H., et al., *Improved low molecular weight Myc-Max inhibitors.* Mol Cancer Ther, 2007. **6**(9): p. 2399-408.

23. Yin, X., et al., *Low molecular weight inhibitors of Myc-Max interaction and function.* Oncogene, 2003. **22**(40): p. 6151-9.

24. Kiessling, A., et al., *Selective inhibition of c-Myc/Max dimerization and DNA binding by small molecules.* Chem Biol, 2006. **13**(7): p. 745-51.

25. Kiessling, A., et al., *Selective inhibition of c-Myc/Max dimerization by a pyrazolo[1,5-a]pyrimidine.* ChemMedChem, 2007. **2**(5): p. 627-30.

26. Berg, T., et al., *Small-molecule antagonists of Myc/Max dimerization inhibit Myc-induced transformation of chicken embryo fibroblasts.* Proc Natl Acad Sci U S A, 2002. **99**(6): p. 3830-5.

27. Jones, S. and J.M. Thornton, *Principles of protein-protein interactions.* Proc Natl Acad Sci U S A, 1996. **93**(1): p. 13-20.

28. Lo Conte, L., C. Chothia, and J. Janin, *The atomic structure of protein-protein recognition sites.* J Mol Biol, 1999. **285**(5): p. 2177-98.

29. Russ, A.P. and S. Lampel, *The druggable genome: an update.* Drug Discov Today, 2005. **10**(23-24): p. 1607-10.

30. Thompson, A.D., et al., *Fine-tuning multiprotein complexes using small molecules.* ACS Chem Biol, 2012. **7**(8): p. 1311-20.

31. Quintas-Cardama, A., F.P. Santos, and G. Garcia-Manero, *Therapy with azanucleosides for myelodysplastic syndromes.* Nat Rev Clin Oncol, 2010. **7**(8): p. 433-44.

32. Momparler, R.L., et al., *5-aza-2'-deoxycytidine therapy in patients with acute leukemia inhibits DNA methylation.* Leuk Res, 1984. **8**(2): p. 181-5.

33. von Hoff, D.D. and M. Slavik, *5-azacytidine--a new anticancer drug with significant activity in acute myeloblastic leukemia.* Adv Pharmacol Chemother, 1977. **14**: p. 285-326.

34. Wilson, V.L., P.A. Jones, and R.L. Momparler, *Inhibition of DNA methylation in L1210 leukemic cells by 5-aza-2'-deoxycytidine as a possible mechanism of chemotherapeutic action.* Cancer Res, 1983. **43**(8): p. 3493-6.

35. Silverman, L.R., et al., *Randomized controlled trial of azacitidine in patients with the myelodysplastic syndrome: a study of the cancer and leukemia group B.* J Clin Oncol, 2002. **20**(10): p. 2429-40.

36. Kornblith, A.B., et al., *Impact of azacytidine on the quality of life of patients with myelodysplastic syndrome treated in a randomized phase III trial: a Cancer and Leukemia Group B study.* J Clin Oncol, 2002. **20**(10): p. 2441-52.

37. Fenaux, P. and L. Ades, *Review of azacitidine trials in Intermediate-2-and High-risk myelodysplastic syndromes.* Leuk Res, 2009. **33 Suppl 2**: p. S7-11.

38. Bolden, J.E., M.J. Peart, and R.W. Johnstone, *Anticancer activities of histone deacetylase inhibitors.* Nat Rev Drug Discov, 2006. **5**(9): p. 769-84.

39. O'Connor, O.A., et al., *Clinical experience with intravenous and oral formulations of the novel histone deacetylase inhibitor suberoylanilide hydroxamic acid in patients with advanced hematologic malignancies.* J Clin Oncol, 2006. **24**(1): p. 166-73.

40. Daigle, S.R., et al., *Potent inhibition of DOT1L as treatment of MLL-fusion leukemia.* Blood, 2013. **122**(6): p. 1017-25.

41. Knutson, S.K., et al., *Selective inhibition of EZH2 by EPZ-6438 leads to potent antitumor activity in EZH2-mutant non-Hodgkin lymphoma.* Mol Cancer Ther, 2014. **13**(4): p. 842-54.

42. Morera, L., M. Lubbert, and M. Jung, *Targeting histone methyltransferases and demethylases in clinical trials for cancer therapy.* Clin Epigenetics, 2016. **8**: p. 57.

43. Ueda, R., et al., *Identification of cell-active lysine specific demethylase 1-selective inhibitors.* J Am Chem Soc, 2009. **131**(48): p. 17536-7.

44. Rotili, D. and A. Mai, *Targeting Histone Demethylases: A New Avenue for the Fight against Cancer.* Genes Cancer, 2011. **2**(6): p. 663-79.

45. Cortez, V., et al., *Targeting the PELP1-KDM1 axis as a potential therapeutic strategy for breast cancer.* Breast Cancer Res, 2012. **14**(4): p. R108.

46. McGrath, J. and P. Trojer, *Targeting histone lysine methylation in cancer.* Pharmacol Ther, 2015. **150**: p. 1-22.

47. Stockwell, B.R., *Chemical genetics: ligand-based discovery of gene function.* Nat Rev Genet, 2000. **1**(2): p. 116-25.

48. Janzen, W.P., *Screening technologies for small molecule discovery: the state of the art.* Chem Biol, 2014. **21**(9): p. 1162-70.

49. Powell, D.J., R.P. Hertzberg, and R. Macarromicronn, *Design and Implementation of High-Throughput Screening Assays.* Methods Mol Biol, 2016. **1439**: p. 1-32.

50. Eggert, U.S., *The why and how of phenotypic small-molecule screens.* Nat Chem Biol, 2013. **9**(4): p. 206-209.

51. Blass, B.E., *Chapter 4 - In vitro Screening Systems*, in *Basic Principles of Drug Discovery and Development*. 2015, Academic Press: Boston. p. 143-202.

52. Xu, W., et al., *Crystal structure of a paired domain-DNA complex at 2.5 A resolution reveals structural basis for Pax developmental mutations.* Cell, 1995. **80**(4): p. 639-50.

53. Xu, H.E., et al., *Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding.* Genes Dev, 1999. **13**(10): p. 1263-75.

54. Garvie, C.W., J. Hagman, and C. Wolberger, *Structural studies of Ets-1/Pax5 complex formation on DNA.* Mol Cell, 2001. **8**(6): p. 1267-76.

55. Macchia, P.E., et al., *PAX8 mutations associated with congenital hypothyroidism caused by thyroid dysgenesis.* Nat Genet, 1998. **19**(1): p. 83-6.

56. Cai, Y., et al., *Groucho suppresses Pax2 transactivation by inhibition of JNK-mediated phosphorylation.* EMBO J, 2003. **22**(20): p. 5522-9.

57. Cai, Y., et al., *Phosphorylation of Pax2 by the c-Jun N-terminal kinase and enhanced Pax2-dependent transcription activation.* J Biol Chem, 2002. **277**(2): p. 1217-22.

58. Lechner, M.S., I. Levitan, and G.R. Dressler, *PTIP, a novel BRCT domain-containing protein interacts with Pax2 and is associated with active chromatin.* Nucleic Acids Res, 2000. **28**(14): p. 2741-51.

59. Patel, S.R., et al., *The BRCT-domain containing protein PTIP links PAX2 to a histone H3, lysine 4 methyltransferase complex.* Dev Cell, 2007. **13**(4): p. 580-92.

60. Patel, S.R., et al., *Epigenetic mechanisms of Groucho/Grg/TLE mediated transcriptional repression.* Mol Cell, 2012. **45**(2): p. 185-95.

61. Abraham, S., et al., *The Groucho-associated phosphatase PPM1B displaces Pax transactivation domain interacting protein (PTIP) to switch the transcription factor Pax2 from a transcriptional activator to a repressor.* J Biol Chem, 2015. **290**(11): p. 7185-94.

62. Chalepakis, G., et al., *The molecular basis of the undulated/Pax-1 mutation.* Cell, 1991. **66**(5): p. 873-84.

63. Fickenscher, H.R., G. Chalepakis, and P. Gruss, *Murine Pax-2 protein is a sequence-specific trans-activator with expression in the genital system.* DNA Cell Biol, 1993. **12**(5): p. 381-91.

64. Zhang, J.H., T.D. Chung, and K.R. Oldenburg, *A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays.* J Biomol Screen, 1999. **4**(2): p. 67-73.

65. Ostrom, L., et al., *Reduced Pax2 gene dosage increases apoptosis and slows the progression of renal cystic disease.* Dev Biol, 2000. **219**(2): p. 250-8.

66. Stayner, C., et al., *Pax2 gene dosage influences cystogenesis in autosomal dominant polycystic kidney disease.* Hum Mol Genet, 2006. **15**(24): p. 3520-8.

67. Winyard, P.J., et al., *The PAX2 tanscription factor is expressed in cystic and hyperproliferative dysplastic epithelia in human kidney malformations.* J Clin Invest, 1996. **98**(2): p. 451-9.

68. Zhang, P., et al., *Activation of Wnt11 by transforming growth factor-beta drives mesenchymal gene expression through non-canonical Wnt protein signaling in renal epithelial cells.* J Biol Chem, 2012. **287**(25): p. 21290-302.

CHAPTER 3

Targeting the paired domain to identify a small-molecule inhibitor of Pax2 mediated transcription activation

Abstract

The promise of rational drug design lies in leveraging fundamental knowledge of tissue and disease-specific pathways to find innovative paths of intervention while curtailing systemic side effects. Perhaps of the most efficient methods to inhibit Pax2 transcription activation would be to interfere with Pax2 localization at target genes. We developed a virtual screening approach to identify small-molecules with the potential to interact with the Pax2 DNA-binding domain and abolish DNA-binding ability. Candidate molecules identified in the virtual screen were analyzed in a cell-based reporter assay to confirm their antiPax2 activity. Next, analogs were tested and a more potent compound was identified. This compound, termed EG1, was found to bind to the Pax2 paired domain blocking DNA-binding activity. EG1 was found to phenocopy loss of Pax2 function in an *ex vivo* embryonic kidney culture model. In this model, we observed disruption in kidney development and gene expression changes that were consistent with genetic loss of function studies. Moreover, we showed that EG1 treatment leads to reduced proliferative potential in Pax2 expressing urogenital cancer cells but has little effect on urogenital

cancer cells lacking Pax2 expression. Our results suggest that we have identified a bona fide small-molecule that can, at low micromolar concentrations, inhibit Pax2 activity. EG1 can serve as a backbone for developing more potent and selective inhibitors which could help in treating a variety of kidney diseases for which specific therapies are lacking.

Introduction

Virtual screening is a powerful tool for hit discovery

In silico screening, also referred to as virtual screening or computational screening, has proven to be a useful tool for identification of small-molecules with the potential to interact with and regulate a protein of interest. In fact, an escalating number of success stories using virtual screening have been reported [1-5]. Additionally, virtual screening has become an integral part of many drug discovery pipelines. By and large, virtual screening can be divided into two categories, ligand-based virtual screening (LBVS) and structure-based virtual screening (SBVS) [6, 7]. LBVS relies on structural knowledge of known ligands and active molecules 1) for searching through chemical databases for similar molecules or 2) for designing novel molecules with similar structure. On the contrary, SBVS depends on high-quality structural information for the protein of interest to assess binding potential based on predicted free energy changes associated with the binding [4]. Perhaps the most widely used method of virtual screening is docking-based virtual screening, a form of SBVS that models the interaction between a small-molecule and the protein of interest at the atomic level. Molecular docking approaches require a high-resolution structure of the target protein. Typically, X-ray crystallographic structures, NMR structures, or structures determined by homology modeling are used in molecular docking studies [6, 8, 9]. The quality of hits obtained through molecular modeling is directly dependent on the quality of the structures used for docking studies [10].

Pax2 is a validated target and the paired domain is a great candidate for SBVS approaches to identifying small-molecule inhibitors

A fundamental first step in any successful virtual screening campaign is defining the target [11]. This step requires in-depth knowledge of disease biology to identify a target protein. Furthermore, it requires an understanding of what this protein does, how it accomplishes these tasks, and how this activity could be manipulated with small-molecules. In this study, we chose Pax2 as our target protein. Pax2 was selected because it was previously shown to be overexpressed in renal cell carcinoma [12], polycystic kidney disease [13], Wilms' tumor [14], and other proliferative disorders of the urogenital system. While overexpression may be a cause or consequence of the disease, Pax2 overexpression in renal cell carcinoma was shown to contribute to increased proliferative potential, reduced apoptosis, and resistance to chemotherapeutic intervention [15, 16]. These findings suggest that inhibition of Pax2 could provide therapeutic benefit and validate it as a target.

Having selected Pax2 as our target protein, the next step was to determine how to best target Pax2 with small-molecules. Based on our knowledge of Pax2 biology we determined that one of the productive ways to inhibit Pax2 activity would be to block Pax2 from binding to DNA thereby inhibiting Pax2 mediated regulation of gene expression. This led us to focus our attention on the Pax2 paired domain. A series of point mutations within the paired domain had been found to severely limit DNA-binding ability and lead to disease [17, 18]. Targeting this region of the paired domain with small-molecules may be effective in preventing Pax2 from interacting with DNA.

Strengths and weaknesses of virtual screening

The goal of virtual screening is to find molecules that bind tightly to the target protein and accomplish the desired effect. With virtual screening, many compounds can be screened in a short period of time. The benefit of using virtual screening is that most of the compounds showing no potential for binding to the target can be easily dismissed. This limits the load of compounds that require *in vitro* testing. Given the reduced list of potential compounds, virtual screening can save significant amounts of time and money [19]. Another benefit to virtual screening is that a probable mechanism of action is inherently provided by the modeling [20]. Moreover, this modeling provides information regarding the binding mode and orientation of the candidate molecule that could prove useful for guiding rational hit-to-lead development.

Although virtual screening has many benefits and has been successfully utilized in the drug discovery arena, it is not without limitations. First and foremost, virtual screening hits are merely predictions based on a model system. These hits are only as good as the modeling and docking that was performed. The modeling, even with a high-resolution structure and the best of intentions, is limited because the structures that the models are based on are often based on a solid state version of the target [21]. Thus, it is possible that the structure is not biologically relevant. Given these limitations, it is clear that hits need to be experimentally examined for their ability to bind the target protein *in vitro* and to accomplish the desired effect. In this study, we utilized the *in vitro* luciferase reporter assay described in Chapter 2 to initially test the hits identified by our virtual screen. Moreover, we made use of several biochemical and biophysical binding assays to investigate target engagement and mechanism of action.

Despite being implicated in numerous disease states, Pax2 has not previously been looked at as a therapeutic target. In general, targeting tissue-specific developmental control genes is rare. Regardless, we propose that this could represent an excellent approach with the potential to avoid the harmful side effects that plague many of the current treatment regimens.

Results

Virtual screening to identify small-molecules with the potential to interact with the paired domain

Our goal was to find small-molecules that can inhibit Pax2 transcription activation through blocking Pax2 from binding to its cognate DNA sequence. To limit the number of compounds that we had to physically screen we chose to address this goal using an *in silico* screening approach. Since the structure of Pax2 has yet to be solved, we turned to homology modeling. Homology modeling followed by structural refinement using molecular mechanics and molecular dynamics simulations provide a quality alternative to solving the structure of Pax2. We were fortunate that crystal structures for the paired domains of Pax5 (1K78) [22] and Pax6 (6PAX) [23] were available. Amino acid identity between Pax2 and Pax5 is 97% throughout the paired domain while 76% amino acid identity is observed between Pax2 and Pax6. Considering that Pax2 and Pax5 have a greater degree of amino acid identity than Pax2 and Pax6 and that the Pax5 structure was resolved to 2.25 Å, the structure of the Pax5 paired domain was utilized as the template for our Pax2 homology model.

On the basis of the homology model and several point mutations (C63Y and Q47P) that were previously found to hinder DNA-binding, a DNA-binding pocket was identified [17, 18] (Figure 3.1 A). Hydrophobic residues including L32, V59, and I67 form a pocket surrounding the well-defined DNA-binding pocket that contains the C63 residue (Figure 3.1 B). The Glide docking program was used to perform structural database searching of CCG compound library. The screening was centered around the C63 residue for compounds with the potential to dock within this region of the Pax2 DNA-binding domain. The screening was performed first using Glide 5.5 standard-precision (SP) mode and the top scoring molecules were rescreened using the extra-precision (XP) mode. The top scoring 227 compounds with reasonable docking poses were selected as potential small-molecule inhibitors of Pax2.

Cell-based assays confirm virtual screening discovery of a novel inhibitor of Pax2 transcription activation

The cell-based PRS4-Luc assay described in Chapter 2 was employed to test the candidate inhibitors that were identified by the *in silico* screen. Of the 227 potential inhibitors, 225 were available at the CCG and were investigated for their ability to inhibit Pax2 mediated transcription. Initially, we analyzed the compounds that were prepared and housed at the CCG in dose response assays. From the 225 compounds tested in the dose response assays, 31 showed concentration-dependent inhibition. These 31 hits had IC_{50} values that ranged from 3 μM to over than 500 μM. To confirm the activity of these 31 hits, fresh powder stocks were ordered and retested in our primary activity assay. This was an important step for confirming that the observed activity was, in fact, due to the

desired compound and not due to mislabeling, contamination, or a degradation product. After retesting the fresh powder stocks to confirm activity, we wanted to rule out compounds whose activity was not specific for Pax2 but was instead due to inhibition of the luciferase enzyme or due to interfering with basic cellular processes such as transcription, translation, proliferation, etc. To address this concern, we established a counterscreen in which HEK293 cells were transiently transfected with a luciferase expression vector whose expression was driven by the constitutive CMV promoter (CMV-Luc). Following transfection, these cells were treated with the desired compound and assayed for any loss in luciferase activity.

Of the 31 hits that were retested and subjected to the counterscreen, five were confirmed and found to inhibit luciferase in a Pax2 dependent manner (Table 3.1). The five hits were made up of two pairs of structurally similar compounds and a single compound that was not related to the other two pairs. The IC_{50} values for these compounds ranged from about 50 µM to 125 µM. Since these compounds had relatively mild inhibitory effects and relatively high IC_{50} values, we purchased a number of commercially available analogs. The analogs were tested in our primary and counterscreen assays for their ability to inhibit the luciferase signal. A total of 64 analogs were purchased and tested. The IC50 values for the analogs ranged from approximately 10 µM to greater than 150 µM. From the list of analogs, EG1 (Figure 3.2 B, PubChem CID 2193203) was selected for further study due to its increased potency and selectivity. Pax2 mediated expression of the luciferase reporter gene was inhibited by EG1 in a dose-dependent manner. While EG1 showed an IC_{50} of approximately 10 µM in the PRS4-Luc assay, there was little to no inhibitory effect observed in the CMV-Luc counterscreen

(Figure 3.2 C). Using our PRS4-EGFP cell line we were able to show that EG1 was capable of inhibiting Pax2 mediated expression of EGFP without significant effects on Pax2 protein levels in a second independent Pax2 reporter cell line (Figure 3.2 F).

Considering that the paired domains of the other group II Pax members, Pax5 and Pax8, are 97% and 92% identical to Pax2 respectively, we sought to determine if EG1 could inhibit their transactivation activity or if it was selective for Pax2. Not surprisingly, we found that EG1 inhibited all three of the group II Pax family members with similar IC_{50} values (Figure 3.2 D). To investigate potential EG1 activity against other transcription factors, we turned to a BMP-dependent luciferase reporter (BRE-Luc) cell line. This HEK293 cell line contains an integrated luciferase reporter gene that depends on BMP signaling and P-Smad mediated transactivation. Notably, this cell line was unaffected by titration of EG1 (Figure 3.2 E). Taken together, these data suggest that we have identified a specific and efficacious inhibitor of group II Pax proteins.

<u>EG1 interacts directly with the paired domain of Pax2 and inhibits Pax2 from interacting with DNA</u>

Since EG1 was identified through our *in silico* screen for small-molecules with the potential to interact with the paired domain of Pax2, we predicted that the mechanism of action for EG1 was through binding the paired domain and blocking DNA-binding. To investigate the possible interaction between EG1 and the Pax2 paired domain we turned to biolayer interferometry. Biolayer interferometry is a label-free technique for determining the binding affinity (K_d) between two molecules. For this experiment, we immobilized biotinylated Pax2 paired domain on streptavidin biosensors and measured the binding

affinity of EG1 to the immobilized Pax2 paired domain using the Octet RED96 system (Figure 3.3 A). A K_d of 1.35 μM was calculated using the K_{on} and K_{off} rates, whereas steady-state analysis of the binding curves provided a K_d of 1.5 μM for this interaction (Figure 3.3 B).

Having identified a direct interaction between EG1 and the Pax2 paired domain, we turned our attention toward determining what effect this interaction had on Pax2 DNA-binding ability. To address this, we set up an electrophoretic mobility shift assay (EMSA, Figure 3.3 C). Recombinant Pax2 paired domain was incubated with increasing concentrations of EG1. After a short incubation, an isotopically labeled double stranded PRS4 oligo was added. Following another short incubation, these samples were loaded and run on native polyacrylamide gels. A shift in molecular weight shows that the paired domain is able to bind to the isotopically labeled PRS4 probe. Notably, 10 μM EG1 reduces this binding by more than 50% while 100 μM EG1 nearly completely abolishes this binding. Together these data show EG1 specifically binds to the DNA-binding pocket of the Pax2 paired domain and prevents Pax2 from binding to DNA.

<u>EG1 disrupts embryonic kidney development in a manner that is consistent with loss of Pax2 function</u>

Pax2 activity is essential for the development of the kidney. It plays a role in morphogenetic processes such as ureteric bud branching and in converting metanephric mesenchymal cells into epithelial cells. Losing Pax2 activity hinders branching morphogenesis, prevents metanephric mesenchymal cells from aggregating around the ureteric bud tips, and completely inhibits metanephric mesenchymal cells from

undergoing a mesenchymal to epithelial transition [24-26]. If EG1 is a Pax2 inhibitor, as we suspect, then treating developing kidneys with EG1 should recapitulate these findings. We made use of a well-described *ex vivo* organ culture model system to resolve whether this was the case or not (Figure 3.4). For these experiments, embryonic day 11.5 (E11.5) kidney rudiments were isolated by microdissection. The embryonic kidney rudiments were cultured for 48 h on transwell filters in the presence of EG1 or vehicle. Embryonic lung rudiments, which undergo a branching morphogenesis but do not express and Pax family members, were isolated along with the embryonic kidney rudiments. Since the branching morphogenesis of the lungs proceeds in a Pax independent manner, lung branching should not be affected by EG1 treatment. Following the 48 h culture period, the rudiments were immunostained with antibodies against Pax2 and Cytokeratin (Figure 3.4 A). The Pax2 staining illuminates both the metanephric mesenchyme and the epithelial cells of the ureteric tree while Cytokeratin only marks the epithelial cells of the ureteric tree (Figure 3.4 A). This staining shows that the Pax2 positive metanephric mesenchymal cells failed to aggregate completely at the ureteric bud tips. Instead of aggregating at the ureteric bud tips these cells could be found dispersed throughout the kidney rudiment. This failure to aggregate was accompanied by defects in branching. Ureteric bud tips were counted and found to be significantly reduced following treatment with EG1 (Figure 3.4 B) Importantly, these changes occurred despite Pax2 positive cells being dispersed throughout the kidney rudiments which suggests that EG1 treatment disrupted aspects of renal development that are consistent with a loss of Pax2 function without affecting Pax2 expression levels. These data suggest a Pax2 specific effect of EG1 which is entirely consistent with loss of Pax2 activity.

EG1 treatment leads to gene expression changes that are consistent with reduced Pax2 function

Recently, a study investigating genes that are regulated by Pax2 in renal progenitor cells, which were isolated from embryonic kidneys, found that many key regulators of nephrogenesis were downregulated following the loss of Pax2 [27]. Furthermore, this study showed that in addition to the downregulation of nephrogenic factors a number of genes associated with paraxial mesoderm or interstitial stroma were upregulated. We decided to look at several of the genes identified in that study to see if *ex vivo* organ cultured embryonic kidneys exposed to EG1 would recapitulate gene expression changes caused by loss of Pax2. For these experiments, embryonic day 11.5 (E11.5) kidney rudiments were isolated by microdissection. The embryonic kidney rudiments were cultured for 48 h on transwell filters in the presence of EG1 or vehicle. Following the 48 h culture period, Pax2 target gene expression was analyzed by RT-qPCR or whole mount *in situ* hybridization (Figure 3.5). Expression of a direct target of Pax2, *Cited1*, was found to be severely reduced in renal progenitor cells following EG1 treatment (Figure 3.5 A). While the expression of *Cited1* was greatly reduced, staining indicates that residual expression is in the correct cell types. This suggests that EG1 is not simply killing the progenitor cells. Gene expression changes assessed by RT-qPCR reveals a clear reduction in *HNF4a* and confirms results from the *in situ* hybridization for *Cited1* (Figure 3.5 B). While these two genes are only expressed in renal progenitor cells, *Bmp7* is expressed in both epithelial progenitor cells and the stroma and shows less of a reduction following EG1 treatment (Figure 3.5 B). In addition to looking at previously

identified genes that were shown to be downregulated following the loss of Pax2, we examined several of the stromal genes that were previously found to exhibit higher expression levels following the loss of Pax2. In doing so, we found that *FoxD1*, *Msx1*, and *Twist2* were all upregulated in *ex vivo* cultured kidney rudiments following EG1 treatment (Figure 3.5 C). These data indicate that EG1 is acting in a way that is completely consistent with a loss of Pax2 activity and suggests EG1 is acting in a Pax2 specific manner.

Pax2 positive but not Pax2 negative urogenital cancer cell lines exhibit reduced viability when exposed to EG1

Human renal cancer cell lines are sensitized to chemotherapeutic agents such as cisplatin and exhibit a decrease in proliferation following suppression of Pax2 gene expression [12, 15, 16]. Based on these findings, we investigated whether EG1 treatment has an impact on the viability of urogenital cancer cells that express or do not express Pax2 (Figure 3.6). Pax2 positive cell lines included a renal cell carcinoma line (RCC111) and two ovarian carcinoma cell lines (TOV112D and E2S) while Pax2 negative lines included two prostate cancer cell lines (22Rv1 and PC3) and an ovarian cancer cell line (SK-OV-3). Pax2 expression was investigated by Western blot analysis (Figure 3.6 A). We treated the cancer cell lines with 12.5 µM or 25 µM EG1 and checked viability 48 h post-treatment (Figure 3.6 B). A significant decrease in viability was observed for all three of the Pax2 positive cell lines, whereas none of the Pax2 negative cells lines were affected. To better understand how viability was affected, we measured levels of phosphor-histone H3 (P-H3). P-H3 is a commonly used marker to measure proliferation

of cell cycle arrest. Similar to what was observed in the viability assays, the Pax2 positive cell lines showed decreased levels of P-H3, whereas P-H3 levels in the Pax2 negative cell lines was unchanged (Figure 3.6 C). Interestingly, the effects of EG1 appear to be dependent on the levels of Pax2 proteins with RCC111 having the highest level of Pax2 and being the least sensitive to EG1 treatment. These data are entirely consistent with a role for Pax2 in regulating proliferation and survival in renal and ovarian cancers. Furthermore, these results confirm that Pax2 is a viable target for anticancer therapeutics.

Discussion

In this chapter, we utilized the crystal structure of Pax5 [22] to prepare a homology model of the Pax2 paired domain. This model was used for structure-based virtual screening with the hope of identifying small-molecules capable of binding to the Pax2 paired domain. This virtual screening yielded 227 potential interactors with reasonable docking poses. Of these 227 potential interactors, 225 were readily available and were tested in our cell-based luciferase reporter assay for their ability to inhibit Pax2 mediated transactivation of the reporter gene. The primary screen yielded 31 molecules capable of inhibiting the luciferase signal. Fresh powder stocks of these 31 molecules were ordered and used for confirmation and counterscreening assays. Following these assays, five specific inhibitors remained. Due to the mild inhibitory effects and high IC_{50} values, a number of analogs were ordered and tested. These analogs led to the discovery of a more potent inhibitor. This molecule, which we called EG1, was selected for further analysis and characterization. As predicted by our homology model and confirmed by biolayer interferometry, EG1 was found to bind directly to the Pax2 paired domain. The

binding studies determined this interaction had an affinity of 1.35 µM to 1.5 µM. Gel shift assays determined that this interaction abolishes the DNA-binding ability of Pax2.

Pax2 activity is essential for the development of the kidney. It plays a role in morphogenetic processes like ureteric bud branching and in converting metanephric mesenchymal cells into epithelial cells. Losing Pax2 activity hinders branching morphogenesis, prevents metanephric mesenchymal cells from aggregating around the ureteric bud tips, and completely inhibits metanephric mesenchymal cells from undergoing a mesenchymal to epithelial transition [24-26]. An *ex vivo* organ culture model system shows that EG1 disrupts these Pax2 dependent processes but does not have a similar effect on branching of lung epithelia which proceeds without Pax activity. An early step in nephrogenesis is the formation of the ureteric bud which grows outward toward and eventually invades the metanephric mesenchyme. The Pax2 positive mesenchymal cells condense around the tips of the ureteric bud and in response to Wnt signals undergo a mesenchymal to epithelial transition which contributes to the growing ureteric bud tree [28]. Similar to Pax2 genetic mutants, EG1 blocks the Pax2 positive metanephric mesenchymal cells from condensing around the ureteric bud tips, prevents these cells from transitioning to epithelial cells, and limits branching of the ureteric bud tree. The ability of EG1 to phenocopy the findings from genetic loss of function studies provides some of the strongest evidence for it being a Pax specific inhibitor. Moreover, we saw that EG1 treatment phenocopied gene expression changes associated with genetic loss of function. These changes included loss of Cited1, a marker of renal progenitor cells, and an increase in the stromal gene FoxD1 [29, 30].

Aberrant Pax2 expression is seen in rapidly proliferating renal epithelial cells, such as in polycystic kidney disease [13, 31], juvenile cystic and dysplastic diseases [32], Wilms' tumor [14, 33], and renal cell carcinoma [12, 34, 35]. Pax2 has been described as a pro-survival factor that fuels proliferation and provides tumor cells with resistance to programmed cell death. In line with those thoughts, decreased Pax2 gene dosage proved beneficial for reducing cyst formation and slowing PKD progression in mouse models [13, 31]. Moreover, renal cancer cell lines are sensitized to chemotherapeutic agents, exhibit an increase in apoptosis, and have a reduced proliferative capacity following suppression of Pax2 gene expression [12, 15, 16]. Likewise, we found that EG1 treatment led to decreased proliferation and reduced viability in Pax2 positive cancer cells but not in Pax2 negative cancer cells.

One of the advantages of Pax2 as a therapeutic target is that the number of cell and tissue types that express Pax2 are relatively limited. Furthermore, Pax2 expression is absolutely required for kidney development but is suppressed in healthy adult renal epithelial cells. This limited expression pattern could help to minimize systemic side effects. One concern regarding this matter is the degree of homology observed amongst the Pax family members within the paired domain. The high-degree of similarity could allow EG1 to target Pax family members beyond Pax2. In fact, we found that the other group II family members, which have the highest degree of homology, are also inhibited by EG1. Pax8 expression is found in the thyroid and partially matches Pax2 expression in the urogenital tract [36], while Pax 5 is expressed in B-cell progenitors [37]. Inhibition of other Pax family members could increase the chances of developing negative side effects. However, pan-Pax activity could be useful. For instance, Pax5 overexpression

has been reported in a number of hematological malignancies [38]. Moreover, Pax8 is aberrantly expressed in T-cell lymphomas [39] and multiple urogenital cancers [40].

Since *in silico* screening requires a high-resolution structure to develop a quality model, we targeted the paired domain of Pax2. Pax2 mediated transcription activation is not limited to Pax2 binding. In fact, this process requires interaction with the adapter protein PTIP, the recruitment of mixed lineage leukemia (MLL) histone methyltransferase complex, and subsequent chromatin remodeling [41, 42]. Pax2 mediated transcription activation is enhanced by c-Jun N-terminal kinase phosphorylation of the Pax2 transactivation domain [17, 43]. In addition to its role in transcription activation, Pax2 also plays a role in transcription repression. Pax2 has been found to convert from transcriptional activators to transcriptional repressors by virtue of a direct interaction between Grg4 and the OP [17, 44-46]. The Grg4 mediated conversion from positive to negative transcriptional regulator has been looked at mechanistically for Pax2 [46, 47]. These studies identified a model in which Grg4-dependent recruitment of a phosphatase, PPM1B, leads to dephosphorylation of the Pax2 transactivation domain with subsequent displacement of the PTIP/MLL complex. The displacement of the PTIP/MLL complex is concurrent with Grg4-dependent recruitment of PRMT5, an arginine methyltransferase, and the polycomb repressive complex 2 (PRC2). Ultimately, this leads to alterations in the chromatin landscape as the activating H3K4me3 marks placed by the PTIP/MLL complex are exchanged for the silencing H4R3me2 and H3K27me3 marks placed by PRMT5 and PRC2 respectively. These changes result in condensation of chromatin and repression of gene expression. It is clear that the biochemical regulation of Pax2 is very complex. This complexity may provide a multitude of avenues for modulating Pax2 activity

that goes well beyond directly targeting Pax2. Even though there are many benefits to virtual screening, an unbiased screening approach for Pax2 inhibitors could produce a diversity of compounds which may act alone to regulate Pax2 activity or in combination to enhance the activity of EG1.

EG1 can serve as a backbone for developing more potent and selective inhibitors which could help in treating a variety of kidney diseases for which specific therapies are lacking.

Materials and Methods

Virtual Screening

The three-dimensional (3D) structure of Pax2 has not been determined yet, thus we used homology modeling to model the 3D structure of the target molecule. There are two related structures reported, the Pax5/Ets-1 in complex with DNA (PDB ID: 1K78) and Pax6 in complex with DNA (PDB ID: 6PAX). Amino acid identity throughout the paired DNA-binding domain is 97% between Pax2 and Pax5 and 76% between Pax2 and Pax6. Based on the high degree of identity and the higher resolution of the Pax5 structure bound to DNA in complex with the ETS1 protein (2.25 Å), Pax5 (PDB ID 1K78, chain A) was used as the template for the structural homology model of Pax2. The initial models were generated employing Prime 3.0 of Schrödinger. The homology modeled 3D structure of the Pax2 paired domain underwent a 10 ns molecular dynamics using Amber 11. The produced trajectories were clustered using the MMTSB toolset. In each cluster, a structure with the minimum energy was chosen for virtual screening.

Virtual screening was performed against the compound library (151,634 compounds) available at the Center for Chemical Genomics (CCG) at University of Michigan. The library was filtered in order to avoid possible problematic compounds (known as PAINS) and to use drug-like compounds for the virtual screening by applying multiple criteria including: molecular weight in the range of 300 to 670 Da; calculated logP in the range 2 to 7; Aromatic rings in the range 0 to 5; number of hydrogen bond donors from 0 to 6; number of hydrogen bond acceptors from 2 to 12; and other. A total of 97,378 molecules were selected and taken for structure-based virtual screening (molecular docking approach) using the Glide docking program. Cys63, which is involved in the binding of DNA, was defined as the center of the active site. A cubic box (the active site) with a length of side of 10 Å represented the volume of the protein for which grids were calculated using Glide 5.5. The produced grids were used for docking studies. Glide 5.5 standard-precision (SP) mode and extra-precision (XP) of Schrödinger was used for the virtual screening. The top 1000 candidate small-molecules with the best scores obtained by Glide SP were rescreened using the XP mode. After the re-ranking, the top 227 scored compounds with reasonable docking poses were considered as potential small-molecule inhibitors of Pax2. Of these, 225 were available for screening in a cell-based assay for the inhibition of Pax2 mediated transactivation.

Cell culture

HEK293 cells containing PRS4-EGFP or PRS4-Luc reporter, described previously [48], were cultured in Dulbecco′s Modified Eagle Medium (DMEM) (Gibco) supplemented

with 10% FCS (Sigma), 100 U/ml penicillin, 100 mg/ml streptomycin (Gibco), 400 ng/mL geneticin (Gibco) and were maintained under humidified 5% CO_2/95% air at 37°C.

Transfection

HEK293 cells were transiently transfected with CMV-Pax2b, CMV-Pax2b^{C63Y}, CMV-Pax5, CMV-Pax8, or CMV-Luciferase expression vector. Briefly, $3x10^6$ cells were plated in a 100 mm tissue culture dish (Corning) and incubated overnight. The next morning the media was replaced with 5 mL of antibiotic-free DMEM and transfection reactions were prepared. In a 2 mL Eppendorf tube containing 896 µL of Opti-MEM I (Gibco) 4 µL of 1µg/µL plasmid was added. In a second 2 mL Eppendorf tube containing 886 µL Opti-MEM I 14 µL of Lipofectamine2000 (Life Technologies) was added and allowed to incubate at room temperature for 5 min. The contents of the two tubes were mixed by pipetting and allowed to incubate at room temperature for 20 min. The transfection mix was added dropwise to a 100 mm dish of cells. The plate was shaken gently and placed in the incubator for 6 h. The plate of transfected cells was trypsinized with 0.05% trypsin (Gibco) for 3 min. The cells were harvested, counted, plated in 6-well culture dishes (500,000 cells/well for Western blot analysis) or 96-well dishes (30,000 cells/well for luminescence assays) containing DMSO or increasing concentrations of compound, and were placed in the incubator overnight.

Luciferase reporter assays

For luciferase reporter assays, test plates were removed from the incubator and equilibrated to room temperature for 10 min. before the media was reduced and an equal

volume of Steady-Glo (Promega) was added. The plates were incubated at room temperature for another 10 min. and were read on a PHERAstar multimode plate reader (BMG).

Western blot analysis

For Western blot analysis, the media was removed from test plates and cells were lysed in 2X SDS buffer (20% glycerol, 4% sodium dodecyl sulfate, 0.2 M dithiothreitol, 125 mM Tris, pH 6.8), as described [49]. Total protein was separated on 4-12% Bis-Tris gels (Life Technologies), transferred to PVDF membranes and immunoblotted with anti-Pax2 antibody [50], anti-EGFP antibody (Santa Cruz), anti-actin (Cell Signaling Tech.), anti-P-Histone H3 (Cell Signaling Tech.), anti-Histone H3 (Cell Signaling Tech.), or anti-Actin (Sigma).

Electrophoretic Mobility Shift Assays

PRS4 probe was obtained by generating a double stranded oligo, end labeling the double stranded oligo with [^{32}P] dATP using the polynucleotide kinase reaction, and purifying the product with a mini Quick Spin DNA column (Roche). The following oligos were used to generate the PRS4 probe: TCGAGATATCTAGAGCGGAAGGTGAGCCCAGTGA, TCACTGGGCTCACCGTTCCGCTCTAGATATCTCGA.

Binding reactions were performed in 0.5X Z-Buffer (10% glycerol, 12.5 mM HEPES pH 7.8, 6.25 mM $MgCl_2$, 0.5 mM DTT, 0.05 M KCl, 0.05% NP4) and contained 3 pg

recombinant Pax2 Paired domain, 100 ng poly (dI-dC), ^{32}P-labeled probe (10 000 c.p.m.), and DMSO or increasing concentrations of EG1. Binding reactions were carried out at room temperature for 20min. Samples were resolved at room temperature on 6% native polyacrylamide gels in 0.5X TBE at 120 V.

Affinity Binding Kinetics

The Pax2 paired domain protein containing the first DNA-binding pocket defined in our homology model (aa 1-81) was expressed as a His-tagged protein in E. coli and purified by Ni-affinity. After biotinylation, the protein was bound to saturation to the biosensor of an Octet RED Multichannel platform (ForteBio). All binding and equilibration were done in PBS with 0.1% DMSO. For binding to compound, sensors were incubated for 300 seconds in EG1 at the following concentrations: 0.5, 1, 2, 4, 8, 16, 31, 62, 125, and 250 μM. Disassociation was then measured for 350 seconds in PBS to determine affinity constants.

Organ Cultures

Mice were kept according to NIH guidelines. All procedures were approved by the University Committee on Use and Care of Animals at the University of Michigan. Wild-type FVB mice aged 8-12 weeks were used in this study (Jackson Lab). Kidney and lung rudiments were microdissected at embryonic day 11.5 (E11.5) and were cultured on 0.4 μm Transwell filter inserts (Costar) in Dulbecco′s Modified Eagle Medium (Gibco) supplemented with 10% FCS (Sigma), 100 U/ml penicillin, 100 mg/ml streptomycin

(Gibco), and DMSO (Sigma) or increasing concentrations of EG1 under humidified 5% CO_2/95% air at 37°C for 2 days.

In situ hybridization

In situ hybridization was performed on 2-day *ex vivo* cultured embryonic kidney rudiments as described previously [27]. Briefly, the samples were fixed in 4% PFA in PBS for 1 h at 4 °C, then washed three times in PBS+0.1% Tween-20 (PBT) for 10 min at room temperature and dehydrated through a PBT/methanol series. Templates for digoxigenin (DIG)-labeled riboprobes were generated by PCR amplification of E11.5-E15.5 mouse embryo cDNAs and sequenced. The following primer pair was used for riboprobe template as described [27]:

Cited1-ATGCCAACCAGGAGATGAAC,

CGATGTTAATACGACTCACTATAGGGCAACAGAATCGGTGGCTTTT.

RNA reverse transcription and real-time qPCR

Total RNA was extracted from 2-day *ex vivo* cultured kidney rudiments using TRIzol reagent (Life Technologies) and RNeasy Mini Kit (Qiagen). Following extraction, 200 ng of total RNA was reverse transcribed into cDNA using SuperScript Vilo Reverse Transcriptase (Life Technologies). cDNA templates were amplified with iTaq Universal SYBR Green Supermix (Bio-Rad) in a Mx3005P Real-Time PCR System (Stratagene). The following primer pairs which were described previously [27] were used in this study:

Bmp7–CAGCCAGAATCGCTCCAAGA, GCAATGATCCAGTCCTGCCA;

Cited1– CTCTGGGAAGGAGGATGCC, CCAGAGGAGCTAGTGGGAAC;

Foxd1 – TTCGGATTCTTGGACCAGAC, CAAGTCAGGGTTGCAGCATA

Hnf4a – TACTCCTGCAGGTTTAGCCG, CAGCCCGGAAGCACTTCTTA;

Hprt – GTTGGGCTTACCTCACTGCT, TCATCGCTAATCACGACGCT;

Msx1 – GCCCCGAGAAACTAGATCGG, GGACTCAGCCGTCTGGC;

Twist2 –GTCTCAGCTACGCCTTCTCC, CAGGTGGGTCCTGGCTTG.

Immunofluorescence

Immunofluorescence was performed on 2-day *ex vivo* cultured embryonic kidney and lung rudiments as described previously [51]. Briefly, the samples were fixed in ice-cold methanol for 20 min, then washed in PBS+0.1% Tween-20 (PBT), immunolabeled with anti-Pax2 (1:400) and anti-Cytokeratin (1:400) (Sigma) antibodies overnight at 4°C in PBT containing 5% goat serum (Sigma), washed with PBT, incubated with Alexa Fluor 488 and Alexa Fluor 594 secondary antibodies (1:500) (Life Technologies) overnight at 4°C, washed with PBT, and were mounted on microscope slides. Images were taken at 10X and 20X magnification.

Viability and Apoptosis Assays

Cancer cell lines were cultured in T75 flasks (Greiner) containing RPMI-1640 (Gibco) supplemented with 10% FCS (Sigma), 100 U/ml penicillin, 100 mg/ml and streptomycin (Gibco). Cancer cell lines were maintained under humidified 5% CO_2/95% air at 37°C. Cancer cell lines were trypsinized with 0.25% trypsin, plated in 12-well culture dishes (Corning) or 96-well culture dishes (Greiner) at 50,000 cells/well or 500 cells/well respectively, and were incubated overnight. DMSO or increasing concentrations of EG1

and vehicle were added to the wells and the plates were once again incubated overnight. Cells from the 12-well plates were lysed in 2X SDS lysis buffer and used for Western blotting. 96-well plates were removed from the incubator and allowed to equilibrate to room temperature for 10 min. before CellTiter-Glo (Promega) was added and the luminescence was read on the PHERAstar multimode plate reader.

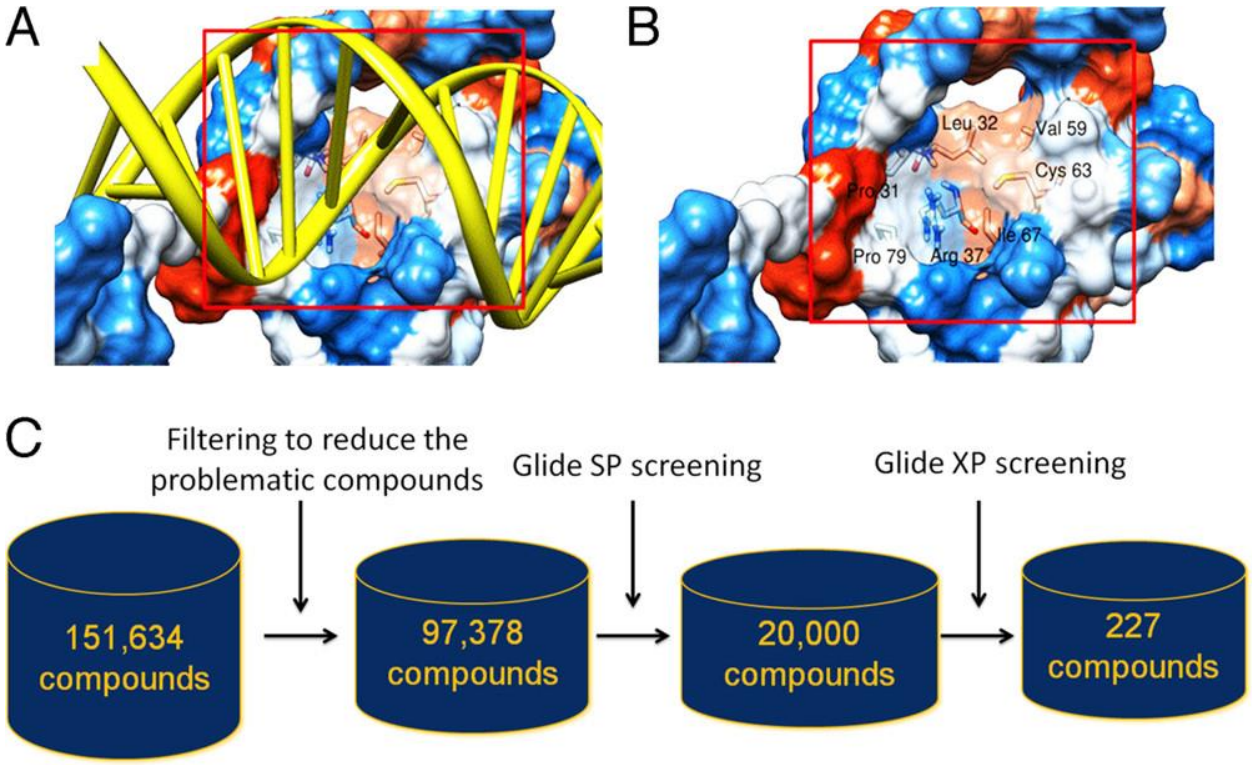


Figure 3.1. Homology model of Pax2 paired domain and virtual screen. A) A three-dimensional homology model of the Pax2 paired domain bound to DNA based on the crystal structure of the Pax5/ETS-1 complex with DNA. **B)** The red box indicated the DNA binding pocket that was targeted in the virtual screening. Critical amino acid residues within the DNA-binding pocket are indicated. For cell based screening of candidates, a Cys63 to tyrosine mutant was used as a control due to its hindered DNA-binding ability and loss of Pax2 mediated transactivation activity [17]. **C)** Flowchart describing the virtual screening process and the number of compounds identified at each step. Ultimately, 227 candidate molecules were discovered through the *in silico* screening. I would like to thank Chenzhong Liao for establishing the homology model and performing the virtual screening that led to the 227 compounds. This figure was taken from Grimley, E. et al. *ACS Chem Biol*, 2017 [52].

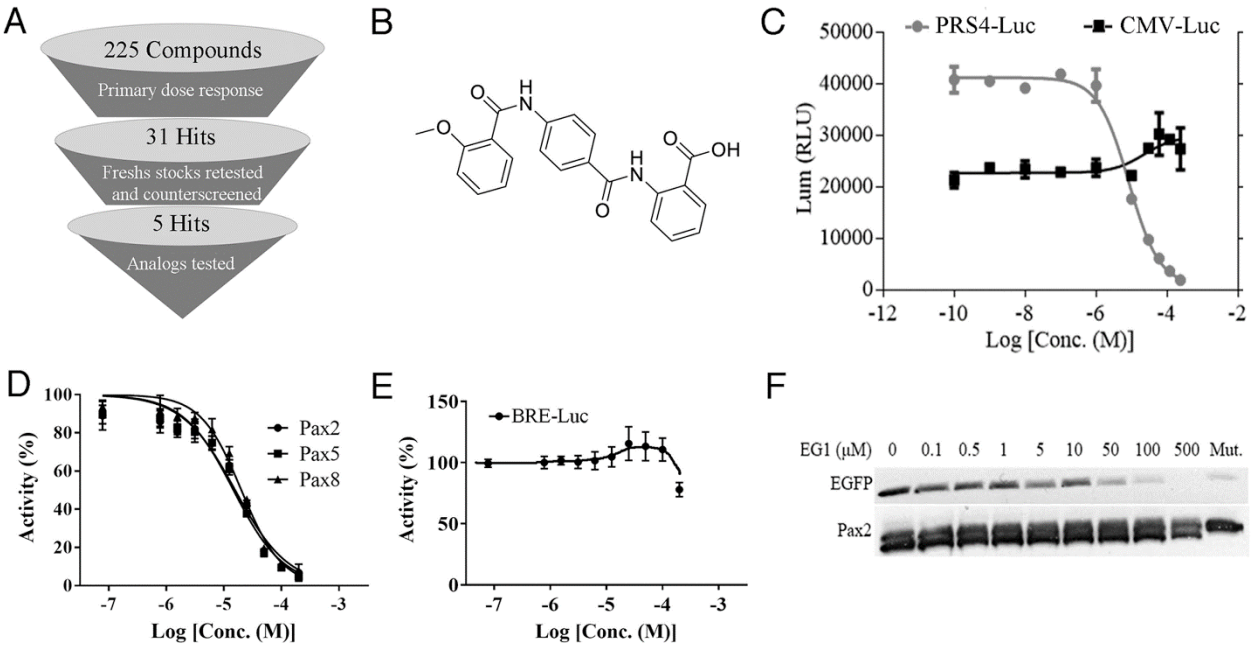


Figure 3.2. *In vitro* confirmation of *in silico* screening results and identification of a small-molecule Pax2 inhibitor. A) Diagram of *in vitro* validation workflow. **B)** Chemical structure of lead compound EG1. **C)** Titration curves for PRS4-Luc reporter cells transiently transfected with CMV-Pax2 expression vector (gray) or CMV-Luc expression vector (black) and treated with increasing concentrations of EG1. Notably, these curves show dose responsive inhibition of luminescence that is dependent on Pax2. **D)** Experiment like that shown in C but with the closely related proteins Pax5 and Pax8 activating the PRS4-Luc reporter. Note, EG1 inhibits transactivation by all 3 Pax proteins. **E)** Titration of EG1 has little effect on the ability of BMP4 and Smad1 to activate a BMP response element in BRE-Luc reporter cell line exposed to BMP4. **F)** Western blot of PRS4-EGFP cells transfected with CMV-Pax2b or CMV-Pax2b^{C63Y} (Mut) and exposed to increasing concentrations of EG1 shows that EG1 dose dependently inhibits Pax2 transactivation activity in these cells. I would like to thank Martha Larsen and Nick Santoro for generating the compound plates that were used for *in vitro* screening. Additionally, I would like to thank Scott Larsen for selecting the analogs that I purchased and tested for antiPax2 activity. This figure was taken from Grimley, E. et al. *ACS Chem Biol*, 2017 [52].

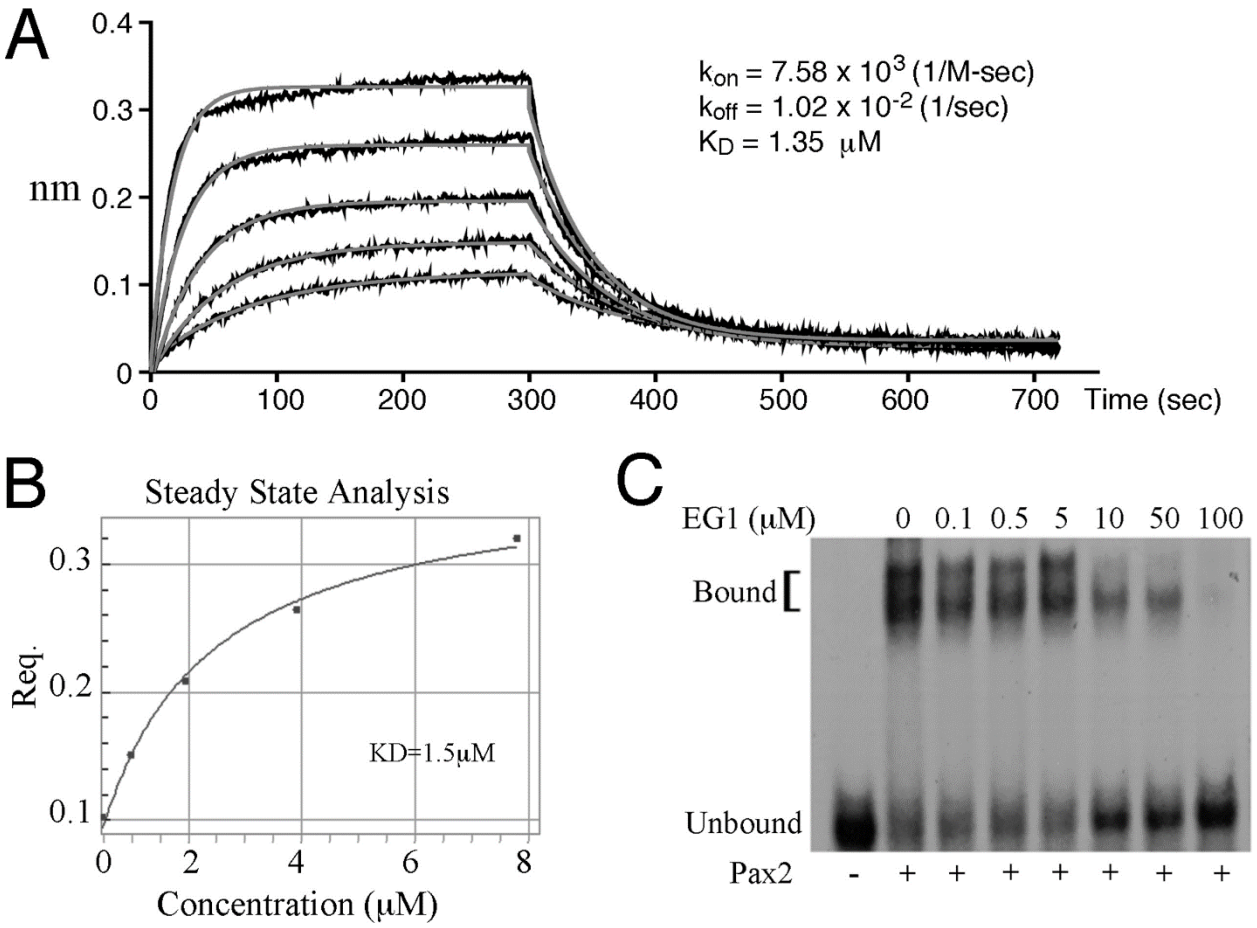


Figure 3.3. EG1 inhibits Pax2 from binding DNA through a direct interaction with the paired domain A) The EG1-Pax2 interaction was investigated by biolayer interferometry in order to determine the kinetics of binding and binding affinity. Streptavidin sensors containing immobilized Pax2 paired domain were exposed to EG1 at concentrations of 0.5, 1.0, 2.0, 4.0, and 8.0 μM. Global fitting of the binding curves was performed using the Octet RED96 instrument software. A K_d of 1.35 μM was determined based on the K_{on} and K_{off} rates. **B)** Steady state analysis calculates a K_d of 1.5 μM which is consistent with the K_{on} and K_{off} rates determined in A. **C)** Gel shift assay performed with isotope labeled PRS4 oligo that was incubated with Pax2 paired domain shows that EG1 dose dependently disrupts the Pax2-DNA interaction. I would like to thank Greg Dressler for performing the OctetRed binding experiments and Zaneta Nikolovska-Coleska for analyzing the OctetRed data. This figure was taken from Grimley, E. et al. *ACS Chem Biol*, 2017 [52].

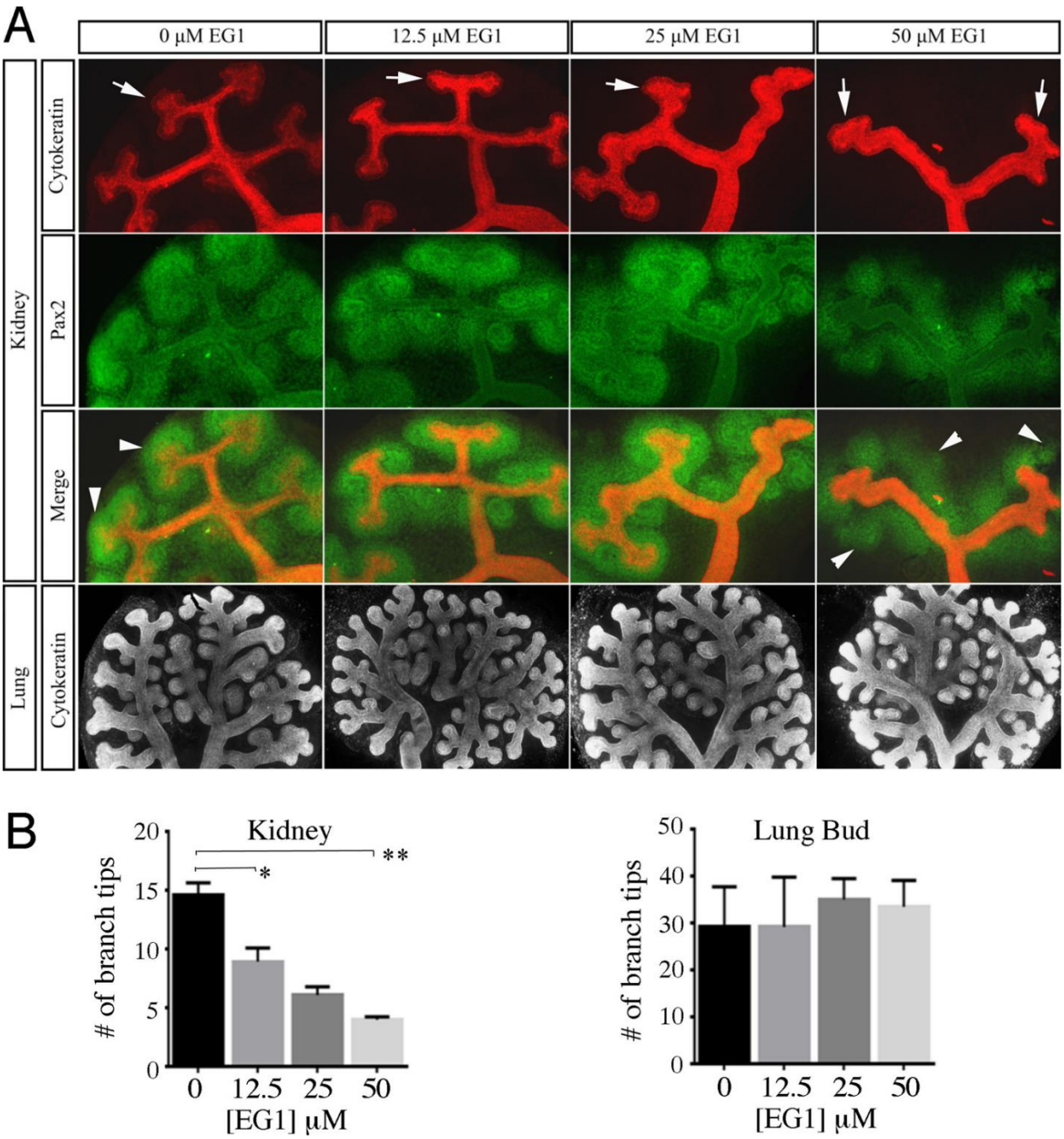

A
0 μM EG1
12.5 μM EG1
25 μM EG1
50 μM EG1
Kidney
Cytokeratin
Pax2
Merge
Lung
Cytokeratin
B
Kidney
# of branch tips
20
15
10
5
0
*
**
0
12.5
25
50
[EG1] μM
Lung Bud
# of branch tips
50
40
30
20
10
0
0
12.5
25
50
[EG1] μM

Figure 3.4 EG1 disrupts branching morphogenesis of *ex vivo* cultured embryonic kidneys but not embryonic lungs A) *Ex vivo* culture of kidney and lung rudiments from E11.5 mouse embryos were exposed to increasing concentrations of EG1. While kidney rudiments treated with EG1 showed a reduction in branching morphogenesis and a disruption in metanephric mesenchyme condensation around the ureteric bud tips, lung rudiments appeared unchanged. Importantly, these findings are consistent with the role of Pax2 in kidney development and the lack of Pax proteins in the lung. **B)** Embryonic lung and kidney rudiments were cultured in the presence of EG1 and branch tips were quantified. Notably, EG1 exposure led to significant changes in kidney (n= 37, 24, 25, and 38, respectively) but not lung (n= 10, 7, 7, and 12, respectively) branching. *$p < 0.001$, **$p < 0.0001$. This figure was taken from Grimley, E. et al. *ACS Chem Biol*, 2017 [52].

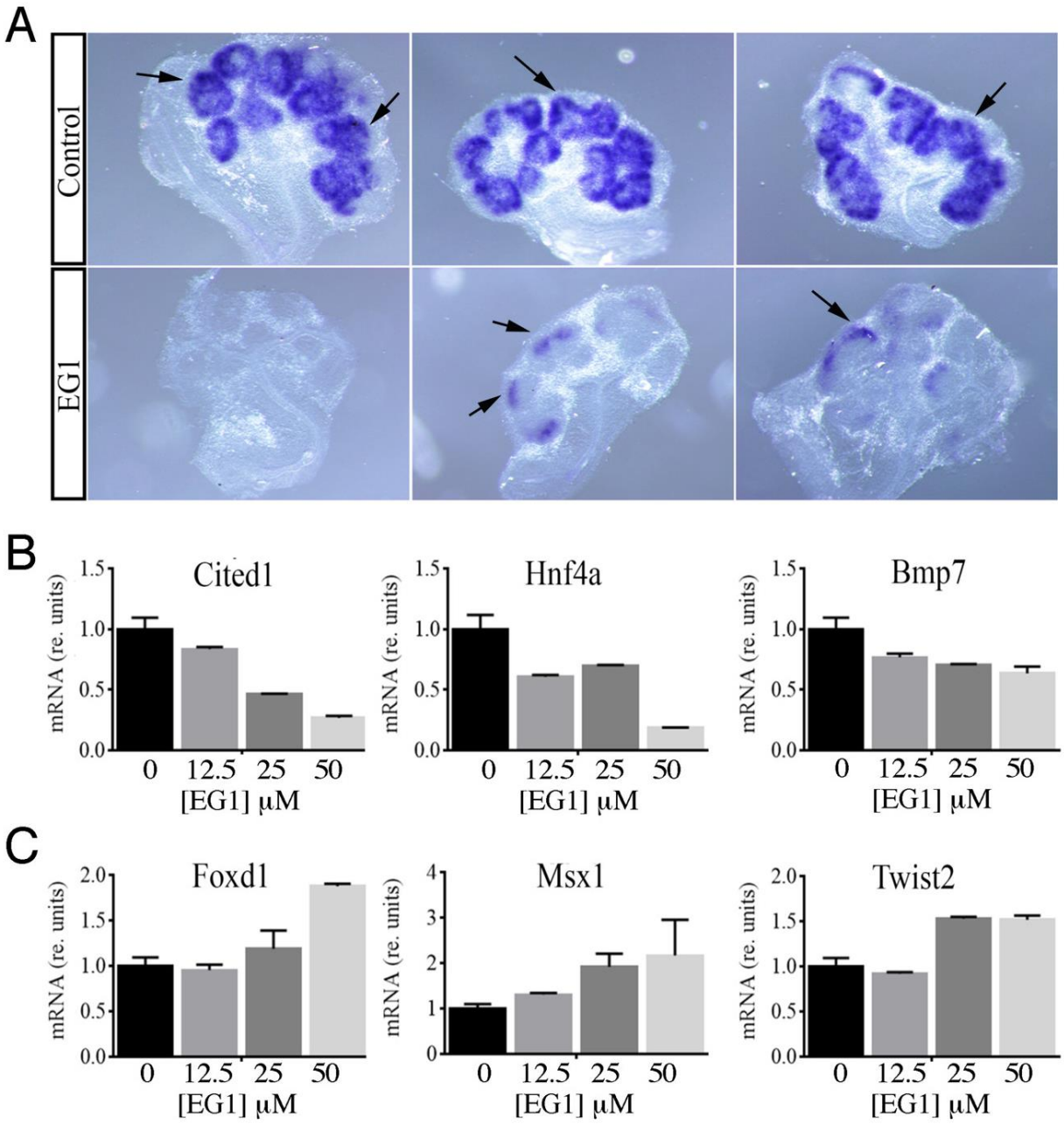


Figure 3.5 EG1 exposure produces gene expression changes consistent with reduced Pax2 transactivation activity in an *ex vivo* organ culture model A) Whole mount *in situ* hybridization for *Cited1* performed on *ex vivo* cultured embryonic kidney rudiments grown in the presence of EG1 shows dramatic reduction in expression of a Pax2 target gene. Note, EG1 exposure strongly inhibits *Cited1* expression in epithelia progenitor cells (arrows) found at the ureteric bud tips. **B)** RT-qPCR for mRNAs of critical nephrogenic regulators show decreased expression after EG1 exposure. **C)** RT-qPCR for mRNAs typically expressed in the interstitial stroma and paraxial mesoderm show increased expression after EG1 exposure. This figure was taken from Grimley, E. et al. *ACS Chem Biol*, 2017 [52].

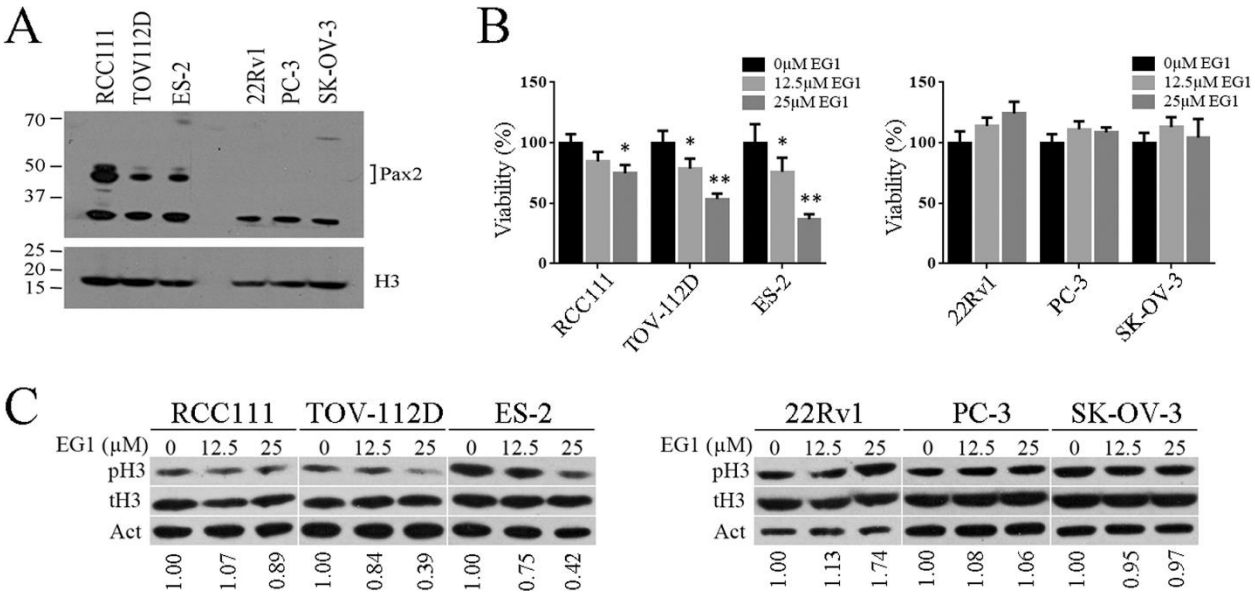


Figure 3.6 Pax2 expressing urogenital cancer cell lines exhibit reduced viability following EG1 treatment. A) Western blot using whole cell lysates from renal (RCC111), ovarian (TOV-112D, ES-2, and SK-OV-3), and prostate (22Rv1 and PC-3) cancer cell lines. Antibodies used are indicated on the right. Notably, RCC111, TOV-112D, and ES-2 cells produced the characteristic Pax2 doublet at 46-48 kDa while 22Rv1, PC-3, and SK-OV-3 cells did not show expression of Pax2. Histone H3 and the nonspecific band at 30 kDa serve as loading controls. **B)** CellTiter-glo luminescent cell viability assay indicates significant reduction in Pax2 positive but not Pax2 negative cancer cell lines following EG1 titration. *$p < 0.05$. **C)** Pax2 positive but not Pax2 negative cancer cell lines treated with increasing concentrations of EG1 show decreased levels of phosphorylated histone H3 (pH3) by Western blot. Levels of pH3 were quantified and normalized based on total H3 levels. These values are shown below the individual blots. This figure was taken from Grimley, E. et al. *ACS Chem Biol*, 2017 [52].

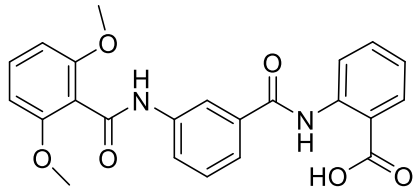
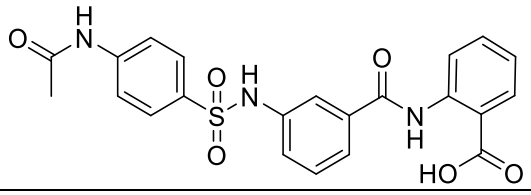
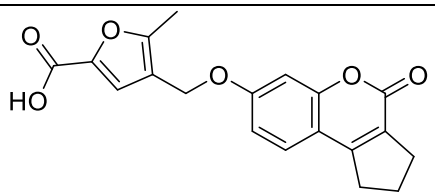
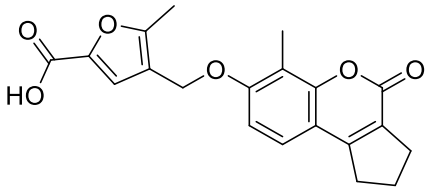
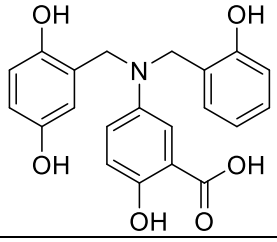
		PRS4-LUC	CMV-LUC
Structure	CCG ID	IC_{50} (μM)	IC_{50} (μM)
	3036	106.7 ± 11.3	>500
	3647	126.5 ± 26.5	>500
	21766	126.4 ± 19.8	>500
	21767	51.9 ± 4.9	>500
	100613	48.6 ± 6.5	291.3 ± 75.6

Table 3.1. Pax2 Specific Inhibitors Identified Through *In Silico* Screening and Confirmed in Cell-Based Activity Assays. This table was taken from Grimley, E. et al. *ACS Chem Biol*, 2017 [52].

Bibliography

1. Villoutreix, B.O., R. Eudes, and M.A. Miteva, *Structure-based virtual ligand screening: recent success stories.* Comb Chem High Throughput Screen, 2009. **12**(10): p. 1000-16.

2. Tuccinardi, T., *Docking-based virtual screening: recent developments.* Comb Chem High Throughput Screen, 2009. **12**(3): p. 303-14.

3. Cheng, T., et al., *Structure-based virtual screening for drug discovery: a problem-centric review.* AAPS J, 2012. **14**(1): p. 133-41.

4. Sliwoski, G., et al., *Computational methods in drug discovery.* Pharmacol Rev, 2014. **66**(1): p. 334-95.

5. Macalino, S.J., et al., *Role of computer-aided drug design in modern drug discovery.* Arch Pharm Res, 2015. **38**(9): p. 1686-701.

6. Kalyaanamoorthy, S. and Y.P. Chen, *Structure-based drug design to augment hit discovery.* Drug Discov Today, 2011. **16**(17-18): p. 831-9.

7. Lionta, E., et al., *Structure-based virtual screening for drug discovery: principles, applications and recent advances.* Curr Top Med Chem, 2014. **14**(16): p. 1923-38.

8. Anderson, A.C., *The process of structure-based drug design.* Chem Biol, 2003. **10**(9): p. 787-97.

9. Hillisch, A., L.F. Pineda, and R. Hilgenfeld, *Utility of homology models in the drug discovery process.* Drug Discov Today, 2004. **9**(15): p. 659-69.

10. Forli, S., *Charting a Path to Success in Virtual Screening.* Molecules, 2015. **20**(10): p. 18732-58.

11. Cerqueira, N.M., et al., *Virtual screening of compound libraries.* Methods Mol Biol, 2009. **572**: p. 57-70.

12. Gnarra, J.R. and G.R. Dressler, *Expression of Pax-2 in human renal cell carcinoma and growth inhibition by antisense oligonucleotides.* Cancer Res, 1995. **55**(18): p. 4092-8.

13. Ostrom, L., et al., *Reduced Pax2 gene dosage increases apoptosis and slows the progression of renal cystic disease.* Dev Biol, 2000. **219**(2): p. 250-8.

14. Dressler, G.R. and E.C. Douglass, *Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor.* Proc Natl Acad Sci U S A, 1992. **89**(4): p. 1179-83.

15. Hueber, P.A., et al., *PAX2 inactivation enhances cisplatin-induced apoptosis in renal carcinoma cells.* Kidney Int, 2006. **69**(7): p. 1139-45.

16. Hueber, P.A., et al., *In vivo validation of PAX2 as a target for renal cancer therapy.* Cancer Lett, 2008. **265**(1): p. 148-55.

17. Cai, Y., et al., *Groucho suppresses Pax2 transactivation by inhibition of JNK-mediated phosphorylation.* EMBO J, 2003. **22**(20): p. 5522-9.

18. Bower, M., et al., *Update of PAX2 mutations in renal coloboma syndrome and establishment of a locus-specific database.* Hum Mutat, 2012. **33**(3): p. 457-66.

19. Leelananda, S.P. and S. Lindert, *Computational methods in drug discovery.* Beilstein J Org Chem, 2016. **12**: p. 2694-2718.

20. Kireev, D., *Structure-Based Virtual Screening of Commercially Available Compound Libraries.* Methods Mol Biol, 2016. **1439**: p. 65-76.

21. Blass, B.E., *Chapter 1 - Drug Discovery and Development: An Overview of Modern Methods and Principles*, in *Basic Principles of Drug Discovery and Development*. 2015, Academic Press: Boston. p. 1-34.

22. Garvie, C.W., J. Hagman, and C. Wolberger, *Structural studies of Ets-1/Pax5 complex formation on DNA.* Mol Cell, 2001. **8**(6): p. 1267-76.

23. Xu, H.E., et al., *Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding.* Genes Dev, 1999. **13**(10): p. 1263-75.

24. Rothenpieler, U.W. and G.R. Dressler, *Pax-2 is required for mesenchyme-to-epithelium conversion during kidney development.* Development, 1993. **119**(3): p. 711-20.

25. Torres, M., et al., *Pax-2 controls multiple steps of urogenital development.* Development, 1995. **121**(12): p. 4057-65.

26. Soofi, A., I. Levitan, and G.R. Dressler, *Two novel EGFP insertion alleles reveal unique aspects of Pax2 function in embryonic and adult kidneys.* Dev Biol, 2012. **365**(1): p. 241-50.

27. Ranghini, E.J. and G.R. Dressler, *Evidence for intermediate mesoderm and kidney progenitor cell specification by Pax2 and PTIP dependent mechanisms.* Developmental Biology, 2015. **399**(2): p. 296-305.

28. Carroll, T.J., et al., *Wnt9b plays a central role in the regulation of mesenchymal to epithelial transitions underlying organogenesis of the mammalian urogenital system.* Dev Cell, 2005. **9**(2): p. 283-92.

29. Boyle, S., et al., *Fate mapping using Cited1-CreERT2 mice demonstrates that the cap mesenchyme contains self-renewing progenitor cells and gives rise exclusively to nephronic epithelia.* Dev Biol, 2008. **313**(1): p. 234-45.

30. Hatini, V., et al., *Essential role of stromal mesenchyme in kidney morphogenesis revealed by targeted disruption of Winged Helix transcription factor BF-2.* Genes Dev, 1996. **10**(12): p. 1467-78.

31. Stayner, C., et al., *Pax2 gene dosage influences cystogenesis in autosomal dominant polycystic kidney disease.* Hum Mol Genet, 2006. **15**(24): p. 3520-8.

32. Winyard, P.J., et al., *The PAX2 tanscription factor is expressed in cystic and hyperproliferative dysplastic epithelia in human kidney malformations.* J Clin Invest, 1996. **98**(2): p. 451-9.

33. Eccles, M.R., et al., *Expression of the PAX2 gene in human fetal kidney and Wilms' tumor.* Cell Growth Differ, 1992. **3**(5): p. 279-89.

34. Tong, G.X., et al., *PAX2: a reliable marker for nephrogenic adenoma.* Mod Pathol, 2006. **19**(3): p. 356-63.

35. Tong, G.X., et al., *PAX8 and PAX2 immunostaining facilitates the diagnosis of primary epithelial neoplasms of the male genital tract.* Am J Surg Pathol, 2011. **35**(10): p. 1473-83.

36. Plachov, D., et al., *Pax8, a murine paired box gene expressed in the developing excretory system and thyroid gland.* Development, 1990. **110**(2): p. 643-51.

37. Urbanek, P., et al., *Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP.* Cell, 1994. **79**(5): p. 901-12.

38. Krenacs, L., et al., *Transcription factor B-cell-specific activator protein (BSAP) is differentially expressed in B cells and in subsets of B-cell lymphomas.* Blood, 1998. **92**(4): p. 1308-16.

39. Morgan, E.A., et al., *PAX8 and PAX5 are differentially expressed in B-cell and T-cell lymphomas.* Histopathology, 2013. **62**(3): p. 406-13.

40. Ozcan, A., et al., *PAX2 and PAX8 expression in primary and metastatic renal tumors: a comprehensive comparison.* Arch Pathol Lab Med, 2012. **136**(12): p. 1541-51.

41. Patel, S.R., et al., *The BRCT-domain containing protein PTIP links PAX2 to a histone H3, lysine 4 methyltransferase complex.* Dev Cell, 2007. **13**(4): p. 580-92.

42. Patel, S.R., E. Ranghini, and G.R. Dressler, *Mechanisms of gene activation and repression by Pax proteins in the developing kidney.* Pediatr Nephrol, 2014. **29**(4): p. 589-95.

43. Cai, Y., et al., *Phosphorylation of Pax2 by the c-Jun N-terminal kinase and enhanced Pax2-dependent transcription activation.* J Biol Chem, 2002. **277**(2): p. 1217-22.

44. Eberhard, D., et al., *Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family.* EMBO J, 2000. **19**(10): p. 2292-303.

45. Lang, D., et al., *Pax3 functions at a nodal point in melanocyte stem cell differentiation.* Nature, 2005. **433**(7028): p. 884-7.

46. Patel, S.R., et al., *Epigenetic mechanisms of Groucho/Grg/TLE mediated transcriptional repression.* Mol Cell, 2012. **45**(2): p. 185-95.

47. Abraham, S., et al., *The Groucho-associated phosphatase PPM1B displaces Pax transactivation domain interacting protein (PTIP) to switch the transcription factor Pax2 from a transcriptional activator to a repressor.* J Biol Chem, 2015. **290**(11): p. 7185-94.

48. Patel, S.R., et al., *Epigenetic mechanisms of Groucho/Grg/TLE mediated transcriptional repression.* Molecular Cell, 2012. **45**(2): p. 185-95.

49. Zhang, P., et al., *Activation of Wnt11 by Transforming Growth Factor-beta Drives Mesenchymal Gene Expression through Non-canonical Wnt Protein Signaling in Renal Epithelial Cells.* J Biol Chem, 2012. **287**(25): p. 21290-302.

50. Dressler, G.R. and E.C. Douglass, *Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**: p. 1179-1183.

51. Cho, E.A., et al., *Differential expression and function of cadherin-6 during renal epithelium development.* Development, 1998. **125**: p. 4806-4815.

52. Grimley, E., et al., *Inhibition of Pax2 Transcription Activation with a Small Molecule that Targets the DNA Binding Domain.* ACS Chem Biol, 2017.

CHAPTER 4

Implementation of a high-throughput screen designed to identify small-molecule modulators of Pax2 mediated transcription regulation

Abstract

Pax2 is an embryonic regulatory protein that is essential for the development of the kidney and urogenital system. While Pax2 activity is high during embryonic development, it is generally repressed in healthy epithelial cells of the nephron. However, Pax2 is reexpressed in surviving proximal tubule epithelia following an acute injury where it is proposed to play a pro-survival role and be important for regeneration. Continued or ectopic Pax2 expression is associated with a number of neoplastic disorders. High levels of Pax2 appear to be critical for maintaining the proliferative state of renal cancer cells which suggest that blocking Pax2 activity may be beneficial for patients suffering from these tumors. Considering this, we developed an HTS campaign aimed at identifying small-molecule regulators of Pax2 activity. HTS is a powerful tool for discovery of active molecules that relies on physically interrogating thousands of molecules. Using the PRS4-Luc reporter assay that we previously optimized for HTS, we screened 69,125 molecules and found 48 inhibitors and 2 activators with EC_{50} values ranging from 1.76 to 121 µM.

From these hits, 3 separate clusters of inhibitors with similar structures were found. Overall, these data demonstrate the feasibility of targeting Pax2 transactivation activity.

Introduction

Pax2 is essential for renal development, is dysregulated in disease, and is a potential target for developing novel therapeutics

The kidney and reproductive tract are derived from the intermediate mesoderm [1, 2], which express *Pax2* and its homolog *Pax8* among the earliest markers for this region [3]. The Wolffian duct, also known as the nephric duct, is the first distinguishable structure derived from the *Pax2* positive intermediate mesoderm. As this duct expands caudally from about the 12^{th} somite toward to the cloaca two transient excretory organs, the pronephros and mesonephros, develop and degenerate after the metanephric, or adult, kidney begins to form at the caudal end. Metanephric kidney development begins when the ureteric bud grows outward from the nephric duct, invades the *Pax2* positive metanephric mesenchyme surrounding it, and initiates induction. Through a complex series of reciprocal interactions, the ureteric bud signals to the metanephric mesenchyme to aggregate around the tips of the ureteric bud to form the Pax2+/Six2+/Cited1+ cap mesenchyme, the progenitor cells of the nephrons [4, 5]. The cap mesenchymal cells undergo a mesenchymal to epithelial transition leading to the formation of renal tubules. Concurrently, the metanephric mesenchyme signals to the ureteric bud to undergo an intricate pattern of branching morphogenesis that ultimately produces the collecting duct system [6].

Pax2 is essential for proper development of the urogenital system and is necessary for renal cell proliferation [7, 8]. While Pax2 expression is critical in developing structures of the nephron, it is down-regulated in mature nephron epithelial cells as kidney development terminates [9, 10]. Pax2 expression is, however, reactivated in regenerating

epithelial cells following acute tubular injury resulting from ischemia or nephrotoxic substances [11]. Epithelial cells that have survived the initial insult have been identified as the source for these regenerating tubular cells [12, 13]. This suggests that dedifferentiation and induction of cell cycle progression is a prerequisite for tubular regeneration. Pax2 expression is likely to play a role in these processes as its expression appears to be required for the expansion of the epithelial cells that repopulate the injured tubules.

While Pax2 is highly expressed during development and in response to acute tubular injury, it is largely absent in the nephrons of healthy adults. However, as discussed previously, misexpression is observed in polycystic kidney disease [14, 15], juvenile cystic and dysplastic diseases [16], Wilms' tumor [17, 18], and renal cell carcinoma [19-21]. In fact, Pax2 is the standard immunodiagnostic marker for tumors originating from the kidney [20-25]. Reducing Pax2 activity by lowering gene dosage in mouse models of PKD or in renal carcinoma cell lines slows proliferation and disease progression [14, 15, 19, 26, 27].

Taken together these findings suggest that Pax2 could be an excellent candidate for developing molecularly targeted therapeutics. We anticipate that compounds which enhance Pax2 activity will be valuable for improving the regenerative capability of injured nephrons while those that restrict Pax2 activity will be useful for treating an array of kidney tumors. In Chapter 3, we discussed a virtual screening approach that was used to identify a small-molecule which prohibits Pax2 from binding to its cognate DNA sequence. While the virtual screening approach selectively targets Pax2 activity, it is far from the only way to target Pax2. In this chapter, we will discuss a second screening campaign that was performed.

High-throughput screening is a powerful tool for hit discovery

High-throughput screening (HTS) is the process by which tens of thousands of compounds are physically evaluated for their biological or biochemical activity. Since its development in the 1990's, HTS has been a major component for hit identification and lead discovery within the pharmaceutical industry [28]. Likewise, HTS has become commonplace for academic researchers [29]. Improvements in automation and the commercial availability of chemical libraries has made it so that even laboratories that do not specialize in chemistry can partake in HTS. This has greatly expanded the search space and allowed for the discovery of tool compounds for targets that are not typically pursued by the pharmaceutical industry.

Historically, pharmaceutical companies have preferentially targeted cell surface receptors, ion channels, and enzymes [30]. This is largely due to the presence of active site cavities that make these types of targets predisposed to small-molecule binding [31]. Additionally, difficult targets and rare diseases are frequently avoided by pharmaceutical companies because of their associated risks or lack of profitability. This leaves a huge opening for academic research labs which are not concerned about shareholder interest and find value in pursuing these neglected targets. One area where this is evident is in protein-protein interaction inhibitors. Once thought to be undruggable, success in targeting these types of interactions over the past decade has shown it be anything but and has helped to reshape the screening landscape [32]. This is just one example of many that demonstrate the value of academic screening programs.

Strengths and weaknesses of HTS

HTS has been the most successful method for discovery of hit compounds [29]. HTS is an unbiased brute force method for physically assessing thousands of compounds rapidly. The major benefit to HTS is that compounds are physically screened which means hits are more relevant. However, physically screening huge numbers of compounds can get expensive and, even though the hits are more relevant, they still require extensive validation. This validation is often more involved, time-consuming, and expensive than the initial screening.

Using the PRS4-Luc assay for HTS allows for discovery of small-molecule regulators of Pax2 activity that may function at any level

The goal of HTS is to identify drug-like compounds that 1) specifically bind to and modulate the activity of a biological target or 2) selectively induce a desired biological effect. The first steps in HTS include selecting and validating a target, choosing a method of targeting, and deciding on and establishing a suitable assay for measuring the desired effect. These aspects were all considered and discussed in Chapter 2 of this thesis. Depending on the goals of the project, target-based or phenotypic screening approaches may be taken. The benefits of these approaches were discussed in more detail in Chapter 2 as were some of the major assays used by each approach. In the end, we decided that a phenotypic screen utilizing the PRS4 reporter would be best for our purposes.

The PRS4 reporter assay is a cell-based assay that monitors Pax2 mediated transcription of a luciferase reporter gene. This assay provides us the opportunity to investigate many aspects of Pax2 biology simultaneously. Using this assay, we have the

potential to identify small-molecules that effect Pax2 activity at numerous levels. For instance, small-molecules that disrupt activation of the reporter gene could be inhibiting the upstream signaling that leads to Pax2 phosphorylation, interfering with Pax2 recognition of the PRS4 element, obstructing one of the many protein-protein interactions needed for transcription to occur, inhibiting the methyltransferase activity of the MLL complex, inhibiting transcription or translation of Pax2, altering Pax2 nuclear localization, or increasing Pax2 degradation. Another major benefit of the PRS4-Luc reporter assay is the ability of the assay to identify small-molecules that increase Pax2 mediated transcription activation.

Results

Primary screening

The optimized PRS4-Luc assay described in Chapter 2 was used as a means for screening small-molecules contained within the chemical library of the Center for Chemical Genomics at the University of Michigan (Figure 4.1). This assay relies on the binding of Pax2 to the PRS4 element, recruitment of PTIP and the MLL complex, histone methylation, and transcription of the luciferase reporter gene (Figure 4.1 A). It is a robust and reproducible assay capable of monitoring compounds for their ability to enhance or repress Pax2 mediated transcription (Figure 4.1 B).

In total, 69,125 unique compounds were examined in our primary screening campaign (Figure 4.1 C). These compounds were assessed across 226 separate 384-well test plates. At this stage, compounds were only tested once at a single concentration of 8 µM. Each test plate contained wells for positive and negative controls. The signal

generated within the positive and negative control wells were each averaged and then used to calculate the percent activation/inhibition of the sample wells on a per plate basis. Compounds that generated a luminescent signal that was at least 50% greater than that observed in the negative control wells or those that were at least 3 standard deviations greater than the negative control wells were called activator hits. Compounds that produced a luminescent signal that was at least 3 standard deviations below the negative control wells or those whose signal was at least 35% less than the negative control wells were called inhibitor hits. A total of 728 activators and 1,960 inhibitors were identified through this campaign (Figure 4.1 C). This correlates with a hit rate of 1.1% for activators and 2.8% for inhibitors. While the 2.8% hit rate for inhibitors is a bit higher than the 1-2% that is typical of primary screening campaigns, it is not outrageous and it may reflect the complexity of the system. We could have made the cutoff criteria more stringent to reduce the number of hits, but we wanted to conserve them at this point. One reason that we did not want to eliminate them was that we felt we had a high-quality assay that could easily distinguish between hits and non-hits. This was evidenced by the fact that we obtained an average Z' of 0.67 for the primary screen.

Confirmation and counter screening

We retested the molecules that showed activity in our primary screen to confirm that they were true hits and not false positives. The compounds were retested as they had been originally and at the same concentration, but this time they were each tested in triplicate. Since a significant number of compounds in screening libraries have activity against luciferase, a counterscreen was run in parallel to our confirmation screen [33, 34].

The same PRS4-Luc cells that were used for the primary and confirmation screen were used for the counterscreen. To accomplish this, the cells were transfected with a plasmid encoding the luciferase gene. Luciferase expression was driven by the constitutive cytomegalovirus (CMV) promoter. Thus, luciferase expression in these cells is completely independent of Pax2 activity which means that any hits that alter the luminescent signal in these cells are nonspecific and should not be considered further. The same criteria that was used to identify hits in the primary screen was used to confirm hits and identify nonspecific compounds. Hits that were confirmed but showed no activity in the counterscreen were subjected to cheminformatic review. After cheminformatic-based triaging and removal of compounds with reactive functional groups or those likely to aggregate, a total of 233 inhibitors and 110 activators remained (Figure 4.1 C). This corresponds to a 11.9% confirmation rate for inhibitors (0.34% overall) and a 15.1% confirmation rate for activators (0.16% overall).

The confirmation and counterscreen campaigns were completed across 13 test plates each. Importantly, these campaigns maintained a high average Z' of 0.67 and 0.69, respectively. Based on these results, we were confident that we had identified hits worth testing in titration experiments.

Dose response and confirmation of dose response

The 233 confirmed inhibitors and 110 confirmed activators were selected for dose response experiments. These experiments were performed with the same compound stocks used for the primary, confirmation, and counterscreen assays. Compounds were tested in duplicate across an 8-point dilution series for the dose response assays. The

dilution series started at 50 µM and ended at 1.4 µM using a dilution factor of 1.67. A total of 146 inhibitors and 8 activators showed concentration dependent effects on Pax2 mediated transactivation of the reporter gene and were commercially available (Figure 4.1 C). Fresh powder stocks of these compounds were purchased and the dose response assay was repeated using the newly purchased compounds. Concomitantly, these compounds were also tested for dose dependent effects in the CMV-Luc counterscreen. Following dose response confirmation and counterscreening, a total of 48 inhibitors and two activators remained (Figure 4.1 C). This corresponds to an overall hit rate of 0.069% for inhibitors and 0.003% for activators.

Summary of HTS results

The 50 hit compounds represent a spectrum of activity ranging from 1.76 to 121 µM (Table 4.1). Half of the hits have an EC_{50} of less than 10 µM, a third of them have EC_{50} values between 10 and 50 µM, and only 8 compounds have EC_{50} values greater than 50 µM. Analysis of the chemical structures of the 48 inhibitors showed that several of the compounds could be clustered based on their similarity. Using the DataWarrior software (openmolecules.org) and a 70% similarity cutoff, three clusters were identified. One of the clusters contains seven compounds with similar chemical scaffolds while the other two clusters each have two members. Visual inspection of the two smaller clusters shows that they share a core scaffold that could allow them to be grouped together. Both of the activators and the 37 remaining inhibitors have unique structures.

Discussion

In this chapter, we described an HTS campaign that was run. This campaign was performed in hopes of finding small-molecules capable of altering Pax2 transactivation activity. Since enhancing Pax2 activity may be beneficial for treating patients with acute kidney injury and inhibiting its activity may be useful for treating renal tumors, we selected an assay capable of identifying both positive and negative regulators of Pax2 activity. A total of 69,125 unique compounds were assayed in our primary screen. This primary screen yielded 2,688 hits, of which1,960 were inhibitors and 728 were activators of Pax2 activity. These hits were then tested in triplicate in our confirmation and counter screening assays. This was done in order to rule out false positives and non-specific hits. The compounds that survived the confirmation and counterscreening assays, 233 inhibitors and 110 activators, were subsequently tested in titration experiments. These experiments were first completed using the same stock compound used in the previous steps of the campaign and then using compound from fresh powder stocks. These compounds were tested across a concentration range starting at 50 µM and ending at 1.4 µM using a dilution factor of 1.67. A total of 48 inhibitors and two activators showed concentration dependent effects that were specific for Pax2 mediated transactivation of the reporter gene.

Generation of dose response curves and calculation of EC_{50} values shows that the hits have a range of activity. EC_{50} values ranging from 1.76 to 121 µM were calculated. An EC_{50} of less than 10 µM was observed for 25 compounds and another 17 compounds have EC_{50} values between 10 and 50 µM. Analysis of the 48 inhibitors identified 3 clusters based on their similarity. One of the clusters contains seven compounds with similar

chemical scaffolds while the other two clusters each have two members but share a common core scaffold that could allow them to be considered as a single group.

Overall these data demonstrate the feasibility of targeting Pax2 transactivation activity and shows that transcription factors can be good targets if enough is known about they function at both a molecular and biological level. Additionally, these data demonstrate that PRS4-Luc assay is suitable for HTS. We have shown that it is possible to find small-molecules that enhance or repress Pax2 activity using this assay. This is an important proof of concept and a critical first step in developing small-molecule regulators of Pax2.

Materials and Methods

Cell culture

HEK293 cells containing a PRS-Luc reporter, described previously [35], were cultured in Dulbecco's Modified Eagle Medium (DMEM, Gibco) supplemented with 10% fetal bovine serum (Sigma), 100 U/mL penicillin, 100 ng/mL streptomycin (Gibco), and 400 ng/mL geneticin (Gibco) and were maintained under humidified 5% CO2/95% air at 37 ^{o}C.

Transfection protocols

HEK293 cells were transiently transfected with CMV-Pax2b (used for negative control and sample wells), CMV-Pax2b^{C63Y} (used for positive control wells), or CMV-Luc (used for counterscreen) expression vector. Briefly, 3x10^{6} cells were plated in a 100 mm tissue culture dish (Corning) and incubated overnight. The next morning the media was

replaced with 5 mL of antibiotic free DMEM and transfection reactions were prepared. In a 2 mL Eppendorf tube containing 896 µL of Opti-MEM I (Gibco) 4 µL of 1µg/µL plasmid was added. In a second 2 mL Eppendorf tube containing 886 µL Opti-MEM I 14 µL of Lipofectamine2000 (Life Technologies) was added and allowed to incubate at room temperature for 5 min. The contents of the two tubes were mixed by pipetting and allowed to incubate at room temperature for 20 min. The transfection mix was added dropwise to a 100 mm dish of cells. The plate was shaken gently and placed in the incubator for 6 hr. The plate of transfected cells was trypsinized with 0.05% trypsin (Gibco) for 3 min. The cells were harvested, counted, diluted to 300 cells/µL. A Multidrop 384 (Thermo Scientific) was used to dispense 20 µL/well of freshly diluted cells to 384-well test plates (Greiner).

384-well test plate preparation

A Multidrop 384 was used to add 30 µL of serum and antibiotic free DMEM to 384-well plates. A Biomek FX liquid handler (Beckman Coulter) equipped with a 384-well pin tool was used to transfer 200 nL of 2 mM compound or DMSO from stock plates to the test plates. After addition of cells to the test plates, the final volume for each well was 50 µL and final concentrations of 8 µM for compounds, 0.4% for DMSO, and 4% for serum were obtained.

Luminescence assay

For luciferase reporter assays, test plates were removed from the incubator and equilibrated to room temperature for ten minutes before media was reduced using a

ELx405 plate washer (BioTek). An equal volume of SteadyGlo (Promega) was added using a Multidrop 384. The plates were covered to protect them from light and incubated at room temperature for an additional ten minutes. Luminescence from the plates was subsequently read on a PHERAstar multimode plate reader (BMG).

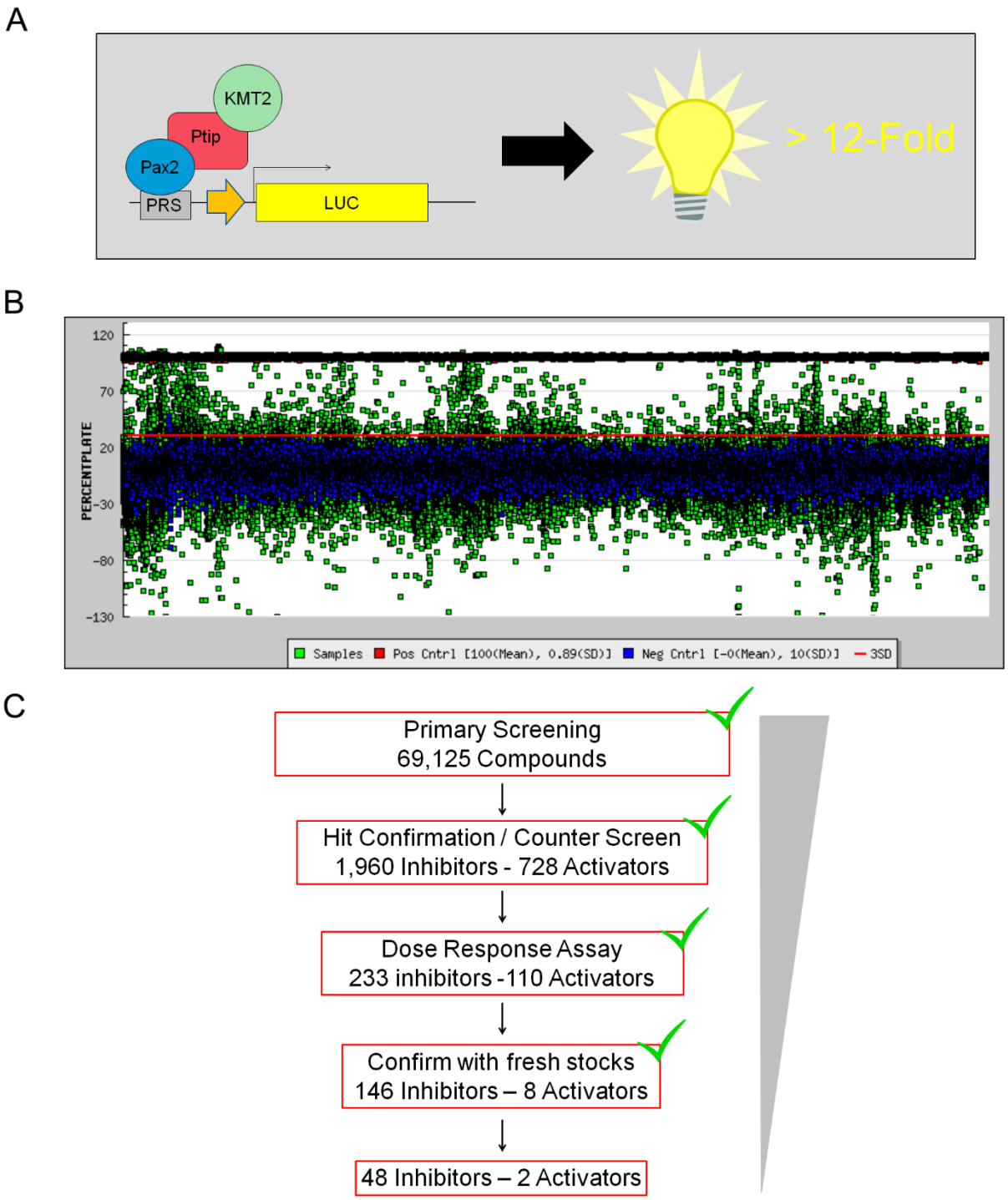


Figure 4.1. HTS campaign for small-molecule modulators of Pax2 transcription activation. HEK293 cells carrying the integrated PRS4-Luc reporter construct were transiently transfected with either a Wt or Mut Pax2 expression vector. The transfected cells were plated in 384-well plates and treated with either vehicle (DMSO) or compound for 18 hours before luminescence was read. **A)** Cartoon diagram illustrating how the PRS4-Luc reporter functions. Note, Pax2 binding to the PRS4 element and subsequent

recruitment of the PTIP/MLL complex induces expression of the reporter gene. **B)** Screenshot of HTS campaign showing negative control wells (blue; CMV-Pax2b transfected cells treated with DMSO), positive control wells (red; CMV-Pax2b^{C63Y} transfected cells treated with DMSO), and sample wells (green; CMV-Pax2b transfected cells treated with compound). The wells are graphed according to their percent activity which is determined on a per plate basis and is based on the averaged raw luminescence values for the positive and negative control wells. Since the positive control used in this screen is the DNA-binding deficient Pax2 mutant, samples with positive activity values (top of graph) are inhibitors while those with negative activity values (bottom of graph) are activators of Pax2 mediated transcription activation. **C)** Summary of HTS campaign results. Green check-marks highlight steps that have been completed. I would like to thank Martha Larsen and Nick Santoro for preparing the compound plates that were used for the screening campaign.

Compound ID	PRS4-Luc		CMV-Luc	
	EC_{50}	Hill Slope	EC_{50}	Hill Slope
EG-H1	98.0	-1.7	ND	ND
EG-H2	121	-1.2	ND	ND
EG-H3	54.0	-1.0	ND	ND
EG-H4	14.7	-1.4	ND	ND
EG-H5	17.1	-1.2	ND	ND
+ EG-H6	9.21	-2.1	ND	ND
+ EG-H7	8.15	-2.9	ND	ND
EG-H8	75.2	-0.8	ND	ND
EG-H9	29.1	-6.6	ND	ND
EG-H10	29.0	-1.0	ND	ND
EG-H11	6.29	-1.3	ND	ND
EG-H12	21.1	-0.7	ND	ND
EG-H13	2.99	-2.6	ND	ND
EG-H14	4.91	-2.7	ND	ND
EG-H15	29.1	-0.7	ND	ND
EG-H16	47.1	-0.7	ND	ND
EG-H17	15.5	-2.5	ND	ND
EG-H18	22.6	-1.6	ND	ND
EG-H19	6.50	-1.0	ND	ND
EG-H20	5.96	-1.8	ND	ND
EG-H21	86.8	-1.1	ND	ND
EG-H22	65.0	-1.2	ND	ND
EG-H23	43.4	-0.6	ND	ND
EG-H24	8.99	-1.0	ND	ND
EG-H25	1.76	-1.2	18.6	-1.9

Compound ID	PRS4-Luc		CMV-Luc	
	EC_{50}	Hill Slope	EC_{50}	Hill Slope
EG-H26	21.8	-0.6	ND	ND
† EG-H27	6.20	-1.9	ND	ND
† EG-H28	8.58	-1.9	ND	ND
EG-H29	19.0	-0.5	ND	ND
EG-H30	5.21	-3.4	53.8	-21
‡ EG-H31	5.72	-1.8	ND	ND
‡ EG-H32	3.97	-2.2	ND	ND
EG-H33	3.80	-69	ND	ND
‡ EG-H34	7.38	-3.7	ND	ND
‡ EG-H35	3.06	-2.1	ND	ND
‡ EG-H36	6.36	-1.7	ND	ND
‡ EG-H37	4.00	-2.1	ND	ND
‡ EG-H38	9.98	-36	ND	ND
EG-H39	11.5	-1.9	ND	ND
EG-H40	20.3	-0.9	ND	ND
EG-H41	5.98	-1.8	ND	ND
EG-H42	83.9	-0.7	ND	ND
EG-H43	6.57	-2.1	60.2	-4.1
EG-H44	101	-0.8	ND	ND
EG-H45	8.32	-0.7	ND	ND
EG-H46	12.3	-0.9	121	-3
EG-H47	1.83	-1.0	43.4	-1
EG-H48	16.8	-1.5	ND	ND
EG-H49	9.29	1.0	101	0.7
EG-H50	10.6	0.6	137	0.8

Table 4.1. Small-Molecules Modulators of Pax2 Activity Identified Through High-Throughput Screening. (+ indicates molecules in cluster 1, † indicates molecules in cluster 2, and ‡ indicates molecules in cluster 3)

Bibliography

1. Dressler, G.R., *The cellular basis of kidney development.* Annu Rev Cell Dev Biol, 2006. **22**: p. 509-29.

2. Dressler, G.R., *Advances in early kidney specification, development and patterning.* Development, 2009. **136**(23): p. 3863-74.

3. Bouchard, M., et al., *Nephric lineage specification by Pax2 and Pax8.* Genes Dev, 2002. **16**(22): p. 2958-70.

4. Boyle, S., et al., *Fate mapping using Cited1-CreERT2 mice demonstrates that the cap mesenchyme contains self-renewing progenitor cells and gives rise exclusively to nephronic epithelia.* Dev Biol, 2008. **313**(1): p. 234-45.

5. Kobayashi, A., et al., *Six2 defines and regulates a multipotent self-renewing nephron progenitor population throughout mammalian kidney development.* Cell Stem Cell, 2008. **3**(2): p. 169-81.

6. Costantini, F. and R. Kopan, *Patterning a complex organ: branching morphogenesis and nephron segmentation in kidney development.* Developmental cell, 2010. **18**(5): p. 698-712.

7. Favor, J., et al., *The mouse Pax2(1Neu) mutation is identical to a human PAX2 mutation in a family with renal-coloboma syndrome and results in developmental defects of the brain, ear, eye, and kidney.* Proc Natl Acad Sci U S A, 1996. **93**(24): p. 13870-5.

8. Torres, M., et al., *Pax-2 controls multiple steps of urogenital development* Development, 1995. **121**: p. 4057-4065.

9. Dressler, G.R. and E.C. Douglass, *Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**: p. 1179-1183.

10. Ryan, G., et al., *Repression of Pax-2 by WT1 during normal kidney development.* Development, 1995. **121**: p. 867-875.

11. Imgrund, M., et al., *Re-expression of the developmental gene Pax-2 during experimental acute tubular necrosis in mice 1.* Kidney Int, 1999. **56**(4): p. 1423-31.

12. Humphreys, B.D., et al., *Intrinsic epithelial cells repair the kidney after injury.* Cell Stem Cell, 2008. **2**(3): p. 284-91.

13. Humphreys, B.D., et al., *Repair of injured proximal tubule does not involve specialized progenitors.* Proc Natl Acad Sci U S A, 2011. **108**(22): p. 9226-31.

14. Ostrom, L., et al., *Reduced Pax2 gene dosage increases apoptosis and slows the progression of renal cystic disease.* Dev Biol, 2000. **219**(2): p. 250-8.

15. Stayner, C., et al., *Pax2 gene dosage influences cystogenesis in autosomal dominant polycystic kidney disease.* Hum Mol Genet, 2006. **15**(24): p. 3520-8.

16. Winyard, P.J., et al., *The PAX2 tanscription factor is expressed in cystic and hyperproliferative dysplastic epithelia in human kidney malformations.* J Clin Invest, 1996. **98**(2): p. 451-9.

17. Dressler, G.R. and E.C. Douglass, *Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor.* Proc Natl Acad Sci U S A, 1992. **89**(4): p. 1179-83.

18. Eccles, M.R., et al., *Expression of the PAX2 gene in human fetal kidney and Wilms' tumor.* Cell Growth Differ, 1992. **3**(5): p. 279-89.

19. Gnarra, J.R. and G.R. Dressler, *Expression of Pax-2 in human renal cell carcinoma and growth inhibition by antisense oligonucleotides.* Cancer Res, 1995. **55**(18): p. 4092-8.

20. Tong, G.X., et al., *PAX2: a reliable marker for nephrogenic adenoma.* Mod Pathol, 2006. **19**(3): p. 356-63.

21. Tong, G.X., et al., *PAX8 and PAX2 immunostaining facilitates the diagnosis of primary epithelial neoplasms of the male genital tract.* Am J Surg Pathol, 2011. **35**(10): p. 1473-83.

22. Tong, G.X., et al., *Expression of PAX2 in papillary serous carcinoma of the ovary: immunohistochemical evidence of fallopian tube or secondary Mullerian system origin?* Mod Pathol, 2007. **20**(8): p. 856-63.

23. Ozcan, A., et al., *PAX-2 in the diagnosis of primary renal tumors: immunohistochemical comparison with renal cell carcinoma marker antigen and kidney-specific cadherin.* Am J Clin Pathol, 2009. **131**(3): p. 393-404.

24. Ordonez, N.G., *Value of PAX2 immunostaining in tumor diagnosis: a review and update.* Adv Anat Pathol, 2012. **19**(6): p. 401-9.

25. Ozcan, A., et al., *PAX2 and PAX8 expression in primary and metastatic renal tumors: a comprehensive comparison.* Arch Pathol Lab Med, 2012. **136**(12): p. 1541-51.

26. Hueber, P.A., et al., *In vivo validation of PAX2 as a target for renal cancer therapy.* Cancer Lett, 2008. **265**(1): p. 148-55.

27. Hueber, P.A., et al., *PAX2 inactivation enhances cisplatin-induced apoptosis in renal carcinoma cells.* Kidney Int, 2006. **69**(7): p. 1139-45.

28. Mayr, L.M. and D. Bojanic, *Novel trends in high-throughput screening.* Current Opinion in Pharmacology, 2009. **9**(5): p. 580-588.

29. Macarron, R., et al., *Impact of high-throughput screening in biomedical research.* Nat Rev Drug Discov, 2011. **10**(3): p. 188-195.

30. Carnero, A., *High throughput screening in drug discovery.* Clin Transl Oncol, 2006. **8**(7): p. 482-90.

31. Bakail, M. and F. Ochsenbein, *Targeting protein–protein interactions, a wide open field for drug design.* Comptes Rendus Chimie, 2016. **19**(1–2): p. 19-27.

32. Arkin, Michelle R., Y. Tang, and James A. Wells, *Small-Molecule Inhibitors of Protein-Protein Interactions: Progressing toward the Reality.* Chemistry & Biology, 2014. **21**(9): p. 1102-1114.

33. Bittker, J.A., N.T. Ross, and T. Brown, *High Throughput Screening Methods*. 2016, [Place of publication not identified]: National Book Network International : Royal Society of Chemistry. 1 online resource.

34. Ho, P.I., et al., *Reporter enzyme inhibitor study to aid assembly of orthogonal reporter gene assays.* ACS Chem Biol, 2013. **8**(5): p. 1009-17.

35. Patel, S.R., et al., *Epigenetic mechanisms of Groucho/Grg/TLE mediated transcriptional repression.* Mol Cell, 2012. **45**(2): p. 185-95.

CHAPTER 5

Discussion and future directions

Discussion

Pax2 is essential for renal development, is dysregulated in disease, and is a potential target for developing novel therapeutics

Pax2 expression is one of the earliest markers for the intermediate mesoderm [1], the tissue from which the kidney and reproductive tract are derived [2, 3]. The first observed structure from the Pax2 positive intermediate mesoderm is the nephric duct. As this duct forms, it expands caudally toward to the cloaca. The kidney begins to form at the caudal end of the nephric duct when a portion of the nephric duct buds outward, invades the *Pax2* positive metanephric mesenchyme surrounding it, and initiates induction. Through a complex series of reciprocal interactions, the ureteric bud signals to the metanephric mesenchyme to aggregate around the tips of the ureteric bud to form the Pax2+/Six2+/Cited1+ cap mesenchyme, the progenitor cells of the nephrons [4, 5]. The cap mesenchymal cells undergo a mesenchymal to epithelial transition to form the renal tubules. At the same time, the metanephric mesenchyme signals to the ureteric bud to undergo a branching morphogenesis event that ultimately produces the collecting duct system [6].

Pax2 is essential for proper development of the urogenital system and is necessary for renal cell proliferation [7, 8]. While *Pax2* expression is critical in developing structures of the nephron, it is down-regulated in mature nephron epithelial cells as kidney development terminates [9, 10]. Pax2 expression is, however, reactivated in regenerating epithelial cells following acute tubular injury resulting from ischemia or nephrotoxic substances [11]. Epithelial cells that have survived the initial insult have been identified as the source for these regenerating tubular cells [12, 13]. This suggests that dedifferentiation and induction of cell cycle progression is a prerequisite for tubular regeneration. Pax2 expression is likely to play a role in these processes as its expression appears to be required for the expansion of the epithelial cells that repopulate the injured tubules.

Pax2 is highly-expressed during development and in response to acute tubular injury, but this expression largely subsides as development and regeneration concludes. However, aberrant Pax2 expression is seen in PKD [14, 15] and renal cancers [16-21]. Interestingly, a reduction in Pax2 gene dosage proved beneficial for inhibiting cyst formation and slowing disease progression in mouse models of PKD [14, 15]. Furthermore, reducing Pax2 levels in renal cancer cell lines sensitized them to chemotherapeutic agents, induced apoptosis, and slowed cell cycling [18, 22, 23].

Taken together, these findings validate Pax2 as a candidate for developing molecularly targeted therapeutics. I posited that compounds which enhance Pax2 activity would be valuable for improving the regenerative capability of injured nephrons while those that restrict Pax2 activity would be useful for treating an array of kidney tumors.

Determining how best to attack the target and how to screen for active molecules are critical for success

Targeting transcription factors is a challenging but exciting strategy for cancer therapeutic development. While transcription factors have classically been considered undruggable because of their nuclear localization, the nature of their interaction with DNA, and the lack of well-defined pockets for targeting [24-26], several have been successfully targeted with small-molecules recently [27-33]. The approaches used include targeting upstream signaling pathways that lead to transcription factor activation, targeting the interaction between the transcription factor and its cognate DNA sequence, targeting the protein-protein interactions that occur between the transcription factor and its cofactors, or targeting the enzymes that catalyze the post-translational modification of histone tails.

Understanding the benefits and challenges of each of these methods is vital for completing a successful screening campaign. As such, careful consideration was given to deciding on a screening platform. To maximize the chances of finding novel small-molecules that regulate Pax2 transactivation I decided to utilize both a target-centered approach and a phenotype-based approach. I selected molecular modeling based virtual screening as the target-centered approach and reporter cell-based HTS as the phenotype-centered approach. By taking these two approaches I expected to identify small-molecules that modulate Pax2 activity both directly and indirectly.

HTS has been the most successful method for discovery of hit compounds [34] and is an unbiased brute force method for physically assessing thousands of compounds rapidly. Cell-based HTS allows for interrogation of multiple aspects of target biology simultaneously. Therefore, cell-based HTS can provide hits with a variety of mechanisms.

This can be both a blessing and a curse as one of the more challenging aspects of HTS is characterizing hits and determining a mechanism of action. While physically screening huge numbers of compounds can get expensive, the follow-up validation is often more involved, time consuming, and costly than the initial screening.

The major benefit to virtual screening is that many compounds can be screened in a short time, and the majority of compounds with no potential for binding to the target can be easily removed thereby limiting the number of compounds that require *in vitro* testing and saving significant amounts of time and money [35]. Another significant benefit to virtual screening is that a probable mechanism of action is inherently provided by the modeling [36]. Moreover, this modeling provides information regarding the binding mode and orientation of the candidate molecule that could prove useful for guiding rational hit-to-lead development. On the other hand, virtual screening hits are merely predictions based on a model system. These hits are only as good as the modeling and docking that was performed. Even with a high-resolution structure, modeling is limited because the structure is generally a solid state version of the target [37]. Thus, it is possible that the structure does not represent a biologically relevant version of the target. These limitations require candidate hits to be experimentally examined for their ability to bind the target protein *in vitro* and elicit the desired response.

I predicted that VS and HTS would be complementary and the combination of the two approaches would provide the best chance of identifying attractive hits. The main reason that the HTS approach was selected was because the Dressler lab had previously developed a reporter cell line that I thought could easily be co-opted for my screening purposes. Having this cell line as the basis for the HTS campaign allows us to

simultaneously look for molecules that affect any step in the process of Pax2 mediated transcription activation. I chose the molecular docking based virtual screening approach because there were high-resolution crystal structures of several Pax family paired domains [38-40] and information from mutation studies suggesting that a specific region of the paired domain may be amenable to small-molecule inhibition [41, 42]. Moreover, directly targeting the DNA-binding domain to interfere with its ability to interact with DNA is likely one of the most effective methods for affecting Pax2 activity. While the molecular docking base virtual screening approach allows us to screen molecules for their potential to directly interact with Pax2 and regulate its activity, this method requires that the identified molecules be assayed for *in vitro* activity. Since I had to go through the process of developing the HTS assay anyway, performing the virtual screening added very little to the workload and provided a potential starting point for mechanistic studies.

Development and optimization of a cell-based reporter assay for monitoring Pax2 mediated transcription activation

In Chapter 2, I discussed the importance of the PRS4 element which was identified in the early 1990's [43, 44] and was used to enhance understanding of Pax2 biochemistry [42, 45-49]. These studies showed that Pax2 binds to the PRS4 element to regulate transcription. They showed that missense mutations within the paired domain, that were known to be disease causing in humans, led to diminished binding potential. These studies found that Pax2 can regulate gene expression positively or negatively depending on the cellular context. They showed that positive regulation involves a connection with the MLL complex, which places activating methylation marks and enhances transcription.

It was shown that the connection to the MLL complex is facilitated by the adapter protein PTIP and that the Pax2-PTIP interaction is potentiated by Wnt signaling and JNK mediated phosphorylation of Pax2. To the contrary, in the presence of high levels of the corepressor protein Grg4, PTIP and the MLL complex can be displaced. This displacement is a consequence of Grg4 recruitment of the phosphatase PPM1B and subsequent dephosphorylation of Pax2. Concomitantly, Grg4 recruits the PRC2 and the arginine methyltransferase PRMT5, which place repressive methylation marks and inhibit transcription.

I used the PRS4 element to create a Pax2 responsive luciferase reporter gene which was subsequently integrated into HEK293 cells. These PRS4-Luc cells were then clonally selected and screened for their ability to respond to Pax2. Clonal cell line #33 was found to have a low basal level of luciferase expression and a large response to transient transfection with a Pax2 expression vector. This cell line was subsequently used for developing the HTS assay. Several parameters were investigated and optimized resulting in a strong induction of luciferase expression in response to Pax2. Following optimization, the assay was miniaturized so that it was suitable for HTS. Mock runs, in which all the wells intended for sample compound were instead only given DMSO, showed that I had developed a robust and reproducible assay that was ready for HTS. This was evidenced by the absence of false positive hits in these test plates and a Z' which was calculated to be 0.71.

<u>Using the PRS4-Luc assay for HTS led to the discovery of small-molecule regulators of Pax2 activity</u>

The PRS4 reporter assay is a cell-based assay that monitors Pax2 mediated transcription of a luciferase reporter gene. This assay provides us the opportunity to investigate compounds for their effect on numerous aspects of Pax2 biology at once. For instance, small-molecules that disrupt activation of the reporter gene could be inhibiting the upstream signaling that leads to Pax2 phosphorylation, interfering with Pax2 recognition of the PRS4 element, obstructing one of the many protein-protein interactions needed for transcription to occur, inhibiting the methyltransferase activity of the MLL complex, inhibiting transcription or translation of Pax2, altering Pax2 nuclear localization, or increasing Pax2 degradation. Another major benefit of the PRS4-Luc reporter assay is the ability of the assay to identify small-molecules that increase Pax2 mediated transcription activation.

In Chapter 4, I described an HTS campaign of 69,125 unique compounds assayed. This primary screen yielded 2,688 active molecules, 1,960 inhibitors and 728 activators. These molecules were then tested in triplicate in the confirmation and counter screening assays. This was done to rule out false positives and non-specific hits. The compounds that survived the confirmation and counterscreening assays, 233 inhibitors and 110 activators, were subsequently tested in titration experiments. These experiments were first completed using the same stock compound used in the previous steps of the campaign and then using compound from fresh powder stocks. These compounds were tested across a concentration range starting at 50 µM and ending at 1.4 µM using a dilution factor of 1.67. A total of 48 inhibitors and two activators showed concentration dependent effects that were specific for Pax2 mediated transactivation of the reporter gene.

Generation of dose response curves and calculation of EC_{50} values shows that the hits discovered in the HTS campaign have a range of activity. EC_{50} values ranging from 1.76 to 121 µM were observed. An EC_{50} of less than 10 µM was observed for 25 compounds and another 17 compounds have EC_{50} values between 10 and 50 µM. Analysis of the 48 inhibitors identified 3 clusters based on their similarity. One of the clusters contains seven compounds with similar chemical scaffolds while the other two clusters each have two members but share a common core scaffold that could allow them to be considered as a single group. The compounds identified through this HTS campaign merit further characterization and could serve as starting points for developing lead molecules.

A critical analysis of the PRS4-Liuc assay and the HTS campaign

I felt that I had developed a strong assay for assessing chemical modulation of Pax2 activity; however, it is important to recognize some of the shortcomings of the assay and discuss aspects of the screening campaign that could have been improved upon. Based on the knowledge gained by going through this process, one thing that I would suggest doing differently next time would be to use a cell line that does not require transfection. There are several reasons for this: 1) transfection has the potential to increase cell permeability which may affect the way compounds work in other models that do not require transfection, 2) transfection produces expression levels that may not be representative of physiological conditions and may mask some of the weaker inhibitors that may be effective against lower protein levels, 3) transfection greatly increases the time required, workload, and cost of performing the screen which could limit the number

of compounds that are screened, and 4) transfection efficiency is one of the most challenging aspects to keep consistent.

In terms of the screening efforts, it is difficult for us to assess the impact of points one and two since I have not examined the HTS hits in models where transfection was not used and Pax2 was not overexpressed; however, points three and four did play a role. The impact of cost on the screening efforts can be observed in two ways. First, I began screening molecules in 2012 with a set of 4160 compounds, but funding concerns put these efforts on hold until the end of 2015 when the remainder of the compounds were screened. Second, while I initially had plans to screen the entire CCG chemical library. I ended up only screening roughly a third of the available compounds. Had the cost of materials been less, it is possible that the I could have screened the entire collection of compounds and completed screening in 2012. Had that been the case, I have no doubts that the efforts would have been much further along by now.

The impact of transfection efficiency variability can be observed by the campaign Z' which was 0.42. The 69,125 compounds were screened over 13 separate runs, and while the average plate and average run Z' values are excellent at 0.67 and 0.59 respectively, the campaign Z' suffered from variability between negative control values between the separate runs. These numbers show that I achieved limited variability amongst controls on a particular plate or amongst plates within a particular run but had an increased variability between separate runs. I suspect that this is a direct result of variability associated with transfection efficiencies despite my best efforts to keep the transfection procedure consistent.

There are a couple of ways that I could have avoided a transfection-based assay. One method would have been to generate a reporter cell line using cells that express endogenous Pax2 such as IMCD or MDCK cells. This could have been accomplished by generating stably transfected cells carrying the PRS4-Luc reporter gene, just as the HEK293 PRS4-Luc cells were generated. Alternatively, I could have generated stably transfected cells carrying both an inducible Pax2 expression cassette and the PRS4-Luc reporter gene.

Another limitation is related to the counterscreen assay that was employed. This assay uses the same PRS4-Luc cells that were utilized in the primary screen which is nice for keeping things consistent. The concern with this assay is that it relies on transient transfection of a plasmid encoding a constitutively expressed luciferase gene driven by the CMV promoter. While this set-up is great for removing most nonspecific hits, there is the potential that some are missed because of differences between the primary and counterscren assays in the manner by which the luciferase gene is transcribed. In the primary assay the luciferase cassette is expressed from a native chromatin environment while in the counterscreen assay it is being expressed from a plasmid. Since plasmids are not subjected to all of the same epigenetic regulators that control native chromatin, there is the potential that compounds that affect these regulators would not be removed by the counterscreen assay. This could have been avoided had I taken a little more time and generated a stably transfected cell line carrying an integrated CMV-Luc cassette.

While the limitations described above represent areas for improvement, I do not think that they invalidate the approach that I ultimately employed. This is evidenced by the ability to identify both positive and negative regulators of Pax2 activity from the HTS

assay and the ability to identify molecules with activity against Pax2 that were able to be validated by a series of assays that were independent from the PRS4-Luc assay for the virtual screening campaign.

The overall hit rate for the HTS campaign was 0.069% for inhibitors and 0.002% for activators. Typical hit rates for HTS campaigns range from 0.01% to 0.14% [50]. While the hit rate for inhibitors falls right in the middle of the anticipated range, the hit rate for activators was well below this range. This may be due to the limited number of mechanisms for increasing Pax2 activity that are specific for Pax2 and could be looked at by the PRS4-Luc assay.

One way to activate Pax2 is through activation of the JNK pathway which leads to phosphorylation of its transactivation domain and subsequent recruitment of PTIP and the MLL complex. To affect this process, small-molecules would most likely have to target enzymes within the JNK signaling pathway, which has a role in regulating apoptosis and cell growth [51]. Given these roles, small-molecules with activity in this pathway are likely to produce a similar effect in both the PRS4-Luc and CMV-Luc assay. These molecules are not specific for Pax2 and have a high likelihood of being removed during counterscreening. Another way to increase the activating potential of Pax2 would be to inhibit recruitment of GRG4, PPM1B, PRMT5, and PRC2. Unfortunately, the reporter cell line does not express GRG4 so these aspects of Pax2 activity were not able to be investigated.

My primary interest was in identifying inhibitor of Pax2 activity. Had my focus been on identifying activators, I would have taken precautions to ensure that these aspects

were represented in the screening assay. Despite these limitations, I was able to identify a couple of molecules that positively regulate Pax2 in a specific manner.

Virtual screening identifies a novel small-molecule that targets the Pax2 paired domain

In Chapter 3, I discussed a virtual screening campaign that was aimed at identifying small-molecules that target the paired domain to disrupt Pax2 activity. Since the structure of Pax2 has yet to be solved, we turned to homology modeling and structural refinement using the paired domain of Pax5 (1K78) [40], which has 97% amino acid identity throughout the paired domain with Pax2 as a template. The Glide docking program was used to perform structural database searching of CCG compound library. Screening was centered around the C63 residue for compounds with potential to dock within this region of the Pax2 DNA-binding domain. The screening was performed first using Glide 5.5 standard-precision (SP) mode, and the top scoring molecules were rescreened using the extra-precision (XP) mode. The top scoring 227 compounds with reasonable docking poses were selected as potential small-molecule inhibitors of Pax2.

The hits identified by the virtual screening campaign were subsequently examined in the PRS4-Luc assay for activity against Pax2. This yielded 31 active molecules that were ordered for confirmation and counterscreening assays. These assays yielded five molecules with mild inhibitory effects. A number of analogs of these molecules were purchased, and a more potent molecule was identified. This molecule, which we called EG1, was selected for further analysis and characterization.

As predicted by the homology model and confirmed by biolayer interferometry, EG1 was found to bind directly to the Pax2 paired domain. The binding studies

determined this interaction had an affinity of 1.35 μM to 1.5 μM. Gel shift assays determined that this interaction abolishes the DNA-binding ability of Pax2. EG1 treatment was shown to phenocopy loss of Pax2 in developing kidney rudiments. The rudiments treated with EG1 exhibited a marked decrease in branching morphogenesis and showed gene expression changes consistent with reduced Pax2 activity. Lung buds, which do not require Pax proteins for their branching, were unaffected by EG1 treatment. Similarly, Pax2 positive but not Pax2 negative cancer cell lines of urogenital origin exhibited reduced viability when treated with EG1. With further characterization and optimization, EG1 could serve as a backbone for developing more potent and selective inhibitors of Pax2.

A critical analysis of the homology-based virtual screening campaign

The virtual screening campaign yielded 227 candidate inhibitors that were subsequently examined in the Pax2 reporter assay. Following confirmation, counterscreen, and titration assays, five molecules remained. These numbers represent a hit rate of just over 2.2%. This number may not be representative of the true hit rate since my approach involved testing the candidates in *in vitro* activity assays instead of binding assays, which are traditionally used to confirm candidates identified through virtual screening. While the hit rate may not be completely accurate, it is within the range typically observed with virtual screening [50]. This validates the modeling and docking that were performed.

The paired domain is composed of two globular subdomains that are connected by a disordered linker [38-40, 52]. Both globular subdomains adopt a HTH structural motif

that has classically been known as a DNA-binding motif. Each of the HTH subdomains are capable of independently binding DNA [53]. I elected to target the N-terminal HTH subdomain based on mutational studies which indicated that missense mutations within this domain severely affect DNA-binding. While the virtual screening campaign successfully identified several small-molecule regulators of Pax2 activity, these molecules have relatively high IC_{50} values. The contribution of the C-terminal HTH subdomain to binding may help to explain the weak activity of the virtual screening hits. Given the success with targeting the N-terminal HTH subdomain, it may be worthwhile to use virtual screening to target the C-terminal HTH subdomain. If this were to be successfully completed, it is possible that these molecules could work synergistically with molecules targeting the N-terminal HTH subdomain.

Advantages and disadvantages of small-molecules targeting Pax2

One of the advantages of Pax2 as a therapeutic target is that the number of cell and tissue types that express Pax2 are relatively limited. Furthermore, Pax2 expression is observed during nephrogenesis but is suppressed in healthy adult renal epithelial cells. This limited expression pattern could help to minimize systemic side effects. One concern regarding this matter is the degree of homology observed amongst the Pax family members within the paired domain. The high-degree of similarity could allow compounds targeting this domain to affect multiple Pax family members. In fact, the paired domain inhibitor that was identified in Chapter 3 disrupted transactivation activity of Pax5 and Pax8 in addition to Pax2. Pax8 expression is found in the thyroid and partially matches Pax2 expression in the urogenital tract [54], while Pax 5 is expressed in B-cell progenitors

[55]. As discussed previously, loss of Pax5 is one of the most frequently observed alterations in B-cell acute lymphoblastic leukemia (B-ALL) [56]. Loss of Pax5 in B-ALL is associated with a defect in B-cell maturation and an accumulation of B-cell progenitors. It is possible that sustained inhibition of Pax5 by small-molecules could lead to the development of leukemia. However, the accumulation of B-cell progenitors should be a transient phenomenon that is relieved when treatment is stopped. Therefore, small-molecule inhibitors which affect Pax5 could still be useful if an optimized dosing schedule is determined. The same could be true for other Pax proteins that are expressed in adult tissues including Pax1/9 in the thymus, Pax3/7 in the stem cell compartment of muscle, and Pax8 in the thyroid [57].

While cross-reactivity represents a legitimate concern for increased toxicity, pan-Pax activity could actually be useful for treating cancers driven by other Pax family members. For instance, Pax5 over expression has been reported in a number of hematological malignancies [58], and genomic rearrangements involving Pax3 or Pax7 are associated with development of pediatric rhabdomyosarcoma [59, 60]. Furthermore, Pax8 is aberrantly expressed in T-cell lymphomas [61], urogenital cancers [62], and gliomas [63].

Concerns about cross-reactivity could be alleviated by inhibitors that target regions within the Pax proteins that are not well-conserved, like the C-terminal transactivation domain. While the amino acid sequence within this domain is not well conserved across the Pax family, the function is. The transactivation domain, as its name implies, confers transcription activation potential to most of the Pax family members [16, 64-69]. While Pax2 mediated gene activation is facilitated by an interaction between the transactivation

domain and PTIP, which links Pax2 to the MLL complex which places activating histone marks, the lack of conservation within the transactivation domain suggests that the various members of the Pax family may have unique interaction partners and utilize different methods for achieving gene activation. The structures of the transactivation domains from the Pax family members have not been elucidated, so it is not possible to specifically target this region. However, the HTS screening approach that was employed could identify small-molecules that prevent PTIP binding or interfere with recruitment of the MLL complex. These molecules have the potential to be more specific for a particular Pax family member since they target interaction interfaces that may be unique. Such molecules could be used alone or in combination with the paired domain inhibitors. The combination approach may allow for a lower dose of each compound to be used while still achieving an enhanced effectiveness.

Summary

Overall this thesis demonstrates the feasibility of targeting Pax2 transactivation activity. Furthermore, it shows that transcription factors can be targeted if enough is known about their molecular biology. Additionally, this thesis demonstrates that PRS4-Luc assay is valuable for assessing small-molecule activity against Pax2 and is suitable for HTS. I have shown that, using this assay, it is possible to find small-molecules that enhance or repress Pax2 activity. This thesis represents an important proof of concept and a critical first step in developing small-molecule regulators of Pax2.

Transcription factors are a challenging class of drug targets, but there have been clear inroads made in this area recently. This thesis represents just one example, but

coupled with the success of others, it signifies a shift in how the drug discovery community should view transcription factors as targets. No longer should they be considered undruggable. This field is still in its early stages, but the progress that has been achieved and the importance of this class of targets make it an attractive area of research that is sure to grow in the future.

Future Directions

Virtual Screen follow-up

I am already working with the University of Michigan Center for Structural Biology to obtain a crystal structure of EG1 bound to the Pax2 paired domain. If we are successful in obtaining this structure, we will have a better idea of how EG1 binds to Pax2 and how best to manipulate this compound to improve binding affinity. Working closely with medicinal chemists, we will synthesize a series of compounds for use in structure-activity relationship studies. These compounds will be tested in binding experiments and in the PRS4-Luc assay to determine their potency and efficacy. In addition, their solubility, metabolic stability, and physiochemical properties will be assessed and used to help determine the most promising lead compound for further development.

In the process of developing better lead molecules, I will need to investigate specificity within the Pax family. The high-degree of homology within the paired domains of Pax family members makes this a difficult task. Since numerous Pax family members are implicated in tumorigenesis, it may be useful to have pan Pax inhibitors. However, the ultimate goal would be to develop small-molecules that preferentially target each of the individual Pax genes. Using the PRS4-Luc reporter, I have shown that all the group II

subfamily members are inhibited by EG1 with nearly identical IC_{50} values. The effects of EG1 on transactivation activity of other Pax family members were investigated using the PRS4-Luc reporter cell line; however, the other Pax family members were unable to activate the PRS4-Luc reporter. While I was discouraged that I was not able to determine EG1 activity against the other Pax family members, I was encouraged that there were DNA-binding specificity differences, which suggests that it may be possible to at least obtain compounds with specificity for a particular subfamily. To further address specificity amongst the Pax family, I would need to generate a series of reporter cell lines that are responsive to the other members. Alternatively, I could obtain purified protein from the other members and examine specificity in the EMSA or BLI binding.

HTS follow-up

This work identified several chemical scaffolds as inhibitors of Pax2 mediated transcription activation but did not generate enough data for chemical optimization. I propose purchasing commercially available analogs with structural similarity to the three scaffolds identified by the HTS campaign. These compounds would be examined in the PRS4-Luc and CMV-Luc assays in order to generate initial structure-activity relationship data. In terms of the singletons that were found by the HTS campaign, I propose cherry-picking the CCG chemical library for molecules with structural similarity. Since I only screened 69,125 unique molecules and there are over 160,000 compounds in the CCG library, I suspect that there will be many similar structures that have not been tested previously. Testing of these molecules could help us to identify additional scaffolds that may be useful for developing better lead compounds.

The Pax proteins are attractive targets for development of small-molecules not only because of their central role in renal development and disease but also because I have detailed knowledge of DNA-binding, protein-protein interactions, and structure-function relationships that contribute to Pax mediated gene regulation. This knowledge will be critical for mechanistic studies. There are several methods at my disposal for investigating Pax2-DNA interactions. Using ChIP, EMSA, or BLI, I can probe this interaction and determine if activity of the hit compound is due to effects on Pax2-DNA binding. Some of these methods were established and utilized in Chapter 3 suggesting that I can mimic those conditions and obtain quality results. If I find that the activity of the hits is not due to effects on Pax2-DNA binding, I can begin investigating the myriad of protein-protein interactions required for Pax2 activity. Perhaps the best way to investigate these interactions would be by performing ChIP on PRS4-Luc cells that have been transiently transfected with Pax2. In the same way that previous studies characterized the events and interactions required for Pax2 mediated transcription activation, I can use antibodies specific for each of the individual proteins involved and monitor assembly of each of these proteins. In this way, I can determine if protein-protein interactions are disrupted. These studies should provide an idea about which interactions are affected by the compound. If I find that Pax2 localizes to the PRS4 element but PTIP does not, then the compound could be interfering with the Pax2-PTIP interaction or could be affecting Pax2 phosphorylation. Pax2 phosphorylation status can be interrogated by labeling with ^{32}P orthophosphate and immunoprecipitation of Pax2 protein, as described previously [45]. If Pax2 and PTIP are present, I will look for members of the epigenetic complexes and their associated marks. These experiments should tell which, if any, members of the epigenetic

complexes are missing. If none of the members are missing, it is possible that the compound is exhibiting its effects through altering the enzymatic ability of the epigenetic complexes. This should be evident by looking at the marks that are present but could be examined using enzymatic assays. The ChIP studies could be confirmed with traditional immunoprecipitation experiments or with biophysical binding experiments like BLI or SPR. Additionally, upstream signaling pathways could be investigated by Western blotting for phosphorylation of known substrates or by testing compounds in a kinase array. It is possible that there are unknown mechanisms of Pax2 regulation that the compounds could be working through that none of the methods described previously would identify. In this case, I may have to turn to quantitative proteomics approaches to identify candidate targets of the compounds.

One of the strengths of the approach is that there are several well characterized assays for testing efficacy of Pax2 regulators. One such example is the kidney culture system that was described in Chapter 3. Since branching morphogenesis and the conversion of mesenchymal cells to epithelia require Pax2 activity, I can culture embryonic kidneys in the presence of compound and monitor these processes. I can grow embryonic lungs as a control, just as I did in Chapter 3. Also in Chapter 3, I described a panel of urogenital cancer cell lines that either express Pax2 and are sensitive to Pax2 inhibition or do not express Pax2 and were unaffected by exposure to the Pax2 inhibitor. These cell lines could be used for testing of the HTS hits. Following extensive characterization and optimization, lead compounds could be tested in mouse models of PKD or acute injury to assess their affects *in vivo*.

Bibliography

1. Bouchard, M., et al., *Nephric lineage specification by Pax2 and Pax8.* Genes Dev, 2002. **16**(22): p. 2958-70.

2. Dressler, G.R., *The cellular basis of kidney development.* Annu Rev Cell Dev Biol, 2006. **22**: p. 509-29.

3. Dressler, G.R., *Advances in early kidney specification, development and patterning.* Development, 2009. **136**(23): p. 3863-74.

4. Boyle, S., et al., *Fate mapping using Cited1-CreERT2 mice demonstrates that the cap mesenchyme contains self-renewing progenitor cells and gives rise exclusively to nephronic epithelia.* Dev Biol, 2008. **313**(1): p. 234-45.

5. Kobayashi, A., et al., *Six2 defines and regulates a multipotent self-renewing nephron progenitor population throughout mammalian kidney development.* Cell Stem Cell, 2008. **3**(2): p. 169-81.

6. Costantini, F. and R. Kopan, *Patterning a complex organ: branching morphogenesis and nephron segmentation in kidney development.* Developmental cell, 2010. **18**(5): p. 698-712.

7. Favor, J., et al., *The mouse Pax2(1Neu) mutation is identical to a human PAX2 mutation in a family with renal-coloboma syndrome and results in developmental defects of the brain, ear, eye, and kidney.* Proc Natl Acad Sci U S A, 1996. **93**(24): p. 13870-5.

8. Torres, M., et al., *Pax-2 controls multiple steps of urogenital development* Development, 1995. **121**: p. 4057-4065.

9. Dressler, G.R. and E.C. Douglass, *Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**: p. 1179-1183.

10. Ryan, G., et al., *Repression of Pax-2 by WT1 during normal kidney development.* Development, 1995. **121**: p. 867-875.

11. Imgrund, M., et al., *Re-expression of the developmental gene Pax-2 during experimental acute tubular necrosis in mice 1.* Kidney Int, 1999. **56**(4): p. 1423-31.

12. Humphreys, B.D., et al., *Intrinsic epithelial cells repair the kidney after injury.* Cell Stem Cell, 2008. **2**(3): p. 284-91.

13. Humphreys, B.D., et al., *Repair of injured proximal tubule does not involve specialized progenitors.* Proc Natl Acad Sci U S A, 2011. **108**(22): p. 9226-31.

14. Ostrom, L., et al., *Reduced Pax2 gene dosage increases apoptosis and slows the progression of renal cystic disease.* Dev Biol, 2000. **219**(2): p. 250-8.

15. Stayner, C., et al., *Pax2 gene dosage influences cystogenesis in autosomal dominant polycystic kidney disease.* Hum Mol Genet, 2006. **15**(24): p. 3520-8.

16. Dressler, G.R. and E.C. Douglass, *Pax-2 is a DNA-binding protein expressed in embryonic kidney and Wilms tumor.* Proc Natl Acad Sci U S A, 1992. **89**(4): p. 1179-83.

17. Eccles, M.R., et al., *Expression of the PAX2 gene in human fetal kidney and Wilms' tumor.* Cell Growth Differ, 1992. **3**(5): p. 279-89.

18. Gnarra, J.R. and G.R. Dressler, *Expression of Pax-2 in human renal cell carcinoma and growth inhibition by antisense oligonucleotides.* Cancer Res, 1995. **55**(18): p. 4092-8.

19. Tong, G.X., et al., *Expression of PAX2 in papillary serous carcinoma of the ovary: immunohistochemical evidence of fallopian tube or secondary Mullerian system origin?* Mod Pathol, 2007. **20**(8): p. 856-63.

20. Tong, G.X., et al., *PAX2: a reliable marker for nephrogenic adenoma.* Mod Pathol, 2006. **19**(3): p. 356-63.

21. Tong, G.X., et al., *PAX8 and PAX2 immunostaining facilitates the diagnosis of primary epithelial neoplasms of the male genital tract.* Am J Surg Pathol, 2011. **35**(10): p. 1473-83.

22. Hueber, P.A., et al., *PAX2 inactivation enhances cisplatin-induced apoptosis in renal carcinoma cells.* Kidney Int, 2006. **69**(7): p. 1139-45.

23. Hueber, P.A., et al., *In vivo validation of PAX2 as a target for renal cancer therapy.* Cancer Lett, 2008. **265**(1): p. 148-55.

24. Darnell, J.E., Jr., *Transcription factors as targets for cancer therapy.* Nat Rev Cancer, 2002. **2**(10): p. 740-9.

25. Hopkins, A.L. and C.R. Groom, *The druggable genome.* Nat Rev Drug Discov, 2002. **1**(9): p. 727-30.

26. Imming, P., C. Sinning, and A. Meyer, *Drugs, their targets and the nature and number of drug targets.* Nat Rev Drug Discov, 2006. **5**(10): p. 821-834.

27. Huang, W., et al., *A small molecule compound targeting STAT3 DNA-binding domain inhibits cancer cell proliferation, migration, and invasion.* ACS Chem Biol, 2014. **9**(5): p. 1188-96.

28. Huang, W., et al., *Small-molecule inhibitors targeting the DNA-binding domain of STAT3 suppress tumor growth, metastasis and STAT3 target gene expression in vivo.* Oncogene, 2016. **35**(6): p. 783-92.

29. Kong, D., et al., *Echinomycin, a small-molecule inhibitor of hypoxia-inducible factor-1 DNA-binding activity.* Cancer Res, 2005. **65**(19): p. 9047-55.

30. Xu, Y., et al., *A credit-card library approach for disrupting protein-protein interactions.* Bioorg Med Chem, 2006. **14**(8): p. 2660-73.

31. Lu, X., et al., *Disruption of the MYC transcriptional function by a small-molecule antagonist of MYC/MAX dimerization.* Oncol Rep, 2008. **19**(3): p. 825-30.

32. Wang, H., et al., *Improved low molecular weight Myc-Max inhibitors.* Mol Cancer Ther, 2007. **6**(9): p. 2399-408.

33. Kiessling, A., et al., *Selective inhibition of c-Myc/Max dimerization and DNA binding by small molecules.* Chem Biol, 2006. **13**(7): p. 745-51.

34. Macarron, R., et al., *Impact of high-throughput screening in biomedical research.* Nat Rev Drug Discov, 2011. **10**(3): p. 188-195.

35. Leelananda, S.P. and S. Lindert, *Computational methods in drug discovery.* Beilstein J Org Chem, 2016. **12**: p. 2694-2718.

36. Kireev, D., *Structure-Based Virtual Screening of Commercially Available Compound Libraries.* Methods Mol Biol, 2016. **1439**: p. 65-76.

37. Blass, B.E., *Chapter 1 - Drug Discovery and Development: An Overview of Modern Methods and Principles*, in *Basic Principles of Drug Discovery and Development*. 2015, Academic Press: Boston. p. 1-34.

38. Xu, W., et al., *Crystal structure of a paired domain-DNA complex at 2.5 A resolution reveals structural basis for Pax developmental mutations.* Cell, 1995. **80**(4): p. 639-50.

39. Xu, H.E., et al., *Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding.* Genes Dev, 1999. **13**(10): p. 1263-75.

40. Garvie, C.W., J. Hagman, and C. Wolberger, *Structural studies of Ets-1/Pax5 complex formation on DNA.* Mol Cell, 2001. **8**(6): p. 1267-76.

41. Macchia, P.E., et al., *PAX8 mutations associated with congenital hypothyroidism caused by thyroid dysgenesis.* Nat Genet, 1998. **19**(1): p. 83-6.

42. Cai, Y., et al., *Groucho suppresses Pax2 transactivation by inhibition of JNK-mediated phosphorylation.* EMBO J, 2003. **22**(20): p. 5522-9.

43. Chalepakis, G., et al., *The molecular basis of the undulated/Pax-1 mutation.* Cell, 1991. **66**(5): p. 873-84.

44. Fickenscher, H.R., G. Chalepakis, and P. Gruss, *Murine Pax-2 protein is a sequence-specific trans-activator with expression in the genital system.* DNA Cell Biol, 1993. **12**(5): p. 381-91.

45. Cai, Y., et al., *Phosphorylation of Pax2 by the c-Jun N-terminal kinase and enhanced Pax2-dependent transcription activation.* J Biol Chem, 2002. **277**(2): p. 1217-22.

46. Lechner, M.S., I. Levitan, and G.R. Dressler, *PTIP, a novel BRCT domain-containing protein interacts with Pax2 and is associated with active chromatin.* Nucleic Acids Res, 2000. **28**(14): p. 2741-51.

47. Patel, S.R., et al., *The BRCT-domain containing protein PTIP links PAX2 to a histone H3, lysine 4 methyltransferase complex.* Dev Cell, 2007. **13**(4): p. 580-92.

48. Patel, S.R., et al., *Epigenetic mechanisms of Groucho/Grg/TLE mediated transcriptional repression.* Mol Cell, 2012. **45**(2): p. 185-95.

49. Abraham, S., et al., *The Groucho-associated phosphatase PPM1B displaces Pax transactivation domain interacting protein (PTIP) to switch the transcription factor Pax2 from a transcriptional activator to a repressor.* J Biol Chem, 2015. **290**(11): p. 7185-94.

50. Zhu, T., et al., *Hit identification and optimization in virtual screening: practical recommendations based on a critical literature analysis.* J Med Chem, 2013. **56**(17): p. 6560-72.

51. Sabapathy, K., *Role of the JNK Pathway in Human Diseases*, in *Progress in Molecular Biology and Translational Science*, S. Shirish, Editor. 2012, Academic Press. p. 145-169.

52. Codutti, L., et al., *The solution structure of DNA-free Pax-8 paired box domain accounts for redox regulation of transcriptional activity in the pax protein family.* J Biol Chem, 2008. **283**(48): p. 33321-8.

53. Epstein, J.A., et al., *Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing.* Genes Dev, 1994. **8**(17): p. 2022-34.

54. Plachov, D., et al., *Pax8, a murine paired box gene expressed in the developing excretory system and thyroid gland.* Development, 1990. **110**(2): p. 643-51.

55. Urbanek, P., et al., *Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP.* Cell, 1994. **79**(5): p. 901-12.

56. Mullighan, C.G., et al., *Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia.* Nature, 2007. **446**(7137): p. 758-64.

57. Li, C.G. and M.R. Eccles, *PAX Genes in Cancer; Friends or Foes?* Front Genet, 2012. **3**: p. 6.

58. Krenacs, L., et al., *Transcription factor B-cell-specific activator protein (BSAP) is differentially expressed in B cells and in subsets of B-cell lymphomas.* Blood, 1998. **92**(4): p. 1308-16.

59. Barr, F.G., *Gene fusions involving PAX and FOX family members in alveolar rhabdomyosarcoma.* Oncogene, 2001. **20**(40): p. 5736-46.

60. Barr, F.G., et al., *Rearrangement of the Pax3 paired box gene in the paediatric solid tumour alvelar rhabdomyosarcoma.* Nature Genetics, 1995. **3**: p. 113-117.

61. Morgan, E.A., et al., *PAX8 and PAX5 are differentially expressed in B-cell and T-cell lymphomas.* Histopathology, 2013. **62**(3): p. 406-13.

62. Ozcan, A., et al., *PAX2 and PAX8 expression in primary and metastatic renal tumors: a comprehensive comparison.* Arch Pathol Lab Med, 2012. **136**(12): p. 1541-51.

63. Hung, N., et al., *Increased paired box transcription factor 8 has a survival function in glioma.* BMC Cancer, 2014. **14**: p. 159.

64. Kozmik, Z., et al., *Alternative splicing of Pax-8 gene transcripts is developmentally regulated and generates isoforms with different transactivation properties.* Mol Cell Biol, 1993. **13**(10): p. 6024-35.

65. Chalepakis, G., et al., *Pax-3 contains domains for transcription activation and transcription inhibition.* Proc Natl Acad Sci U S A, 1994. **91**(26): p. 12745-9.

66. Czerny, T. and M. Busslinger, *DNA-binding and transactivation properties of Pax-6: three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5).* Mol Cell Biol, 1995. **15**(5): p. 2858-71.

67. Fujitani, Y., et al., *Identification of a portable repression domain and an E1A-responsive activation domain in Pax4: a possible role of Pax4 as a transcriptional repressor in the pancreas.* Mol Cell Biol, 1999. **19**(12): p. 8281-91.

68. Kalousova, A., et al., *DNA binding and transactivating properties of the paired and homeobox protein Pax4.* Biochem Biophys Res Commun, 1999. **259**(3): p. 510-8.

69. Lechner, M.S. and G.R. Dressler, *Mapping of Pax-2 transcription activation domains.* J Biol Chem, 1996. **271**(35): p. 21088-93.