UNIVERSITY OF CALIFORNIA SAN DIEGO

The Role of FOXA1 in Regulating Estrogen-dependent Transcription

A dissertation is submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Biological Sciences

by

Alexzander Richard Jose Sharp

Committee in charge:

Professor Michael G. Rosenfeld, Chair
Professor Steven P. Briggs
Professor James T. Kadonaga
Professor Cornelis Murre
Professor Jin Zhang

2023

The dissertation of Alexzander Richard Jose Sharp is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

DEDICATION

This dissertation is dedicated to my parents, John and Francisca Sharp, as well as my brother, Cristian Sharp.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to acknowledge Dr. Michael G. Rosenfeld, the chair of my committee and my principal source of guidance throughout this work. I would also like to acknowledge Kenneth Ohgi for his assistance with cloning FOXA1 deletion mutants into pCDH plasmids and ChIP-seq library preparation (Chapters 1-2), Dr. Lu Yang for the creation of MCF7 FOXA1 binding-site deletion mutants for the TFF1, NRIP1 and FOXC1 enhancers (Chapters 1-2), and Dr. Wei Yuan for the purification of PAG-Tn5 for AP2-γ CUT&TAG assays (Chapter 1).

VITA

2013	Bachelor of Arts and Science, University of California Davis
2014-2015	Research Assistant, University of California Davis
2016-2023	Graduate Research Assistant, University of California, San Diego
2018-2020	Teaching Assistant, University of California San Diego
2023	Doctor of Philosophy, University of California San Diego

PUBLICATIONS

Gao F, Elliott NJ, Ho J, Sharp A, Shokhirev MN, Hargreaves DC. Heterozygous Mutations in SMARCA2 Reprogram the Enhancer Landscape by Global Retargeting of SMARCA4. Mol Cell. 2019 Sep 5;75(5):891-904.e7

FIELDS OF STUDY

Major Field: Biological Sciences

Professor Michael G. Rosenfeld

ABSTRACT OF THE DISSERTATION

The Role of FOXA1 in Regulating Estrogen-dependent Transcription

by

Alexzander Richard Jose Sharp

Doctor of Philosophy in the Biological Sciences

University of California San Diego, 2023

Professor Michael G. Rosenfeld

FOXA1 is a DNA-binding protein that can regulate the estrogen transcriptional program. Although recent work has elucidated specific activities of the FOXA1 C-terminal domain, there is a dearth of information on the contribution of the N-terminal domain to FOXA1 function. By coupling FOXA1 shRNA knockdown with the expression of transgenic FOXA1 mutants, we show that the N-terminal domain is important for FOXA1's ability to assist ER-α in activating estrogen-dependent transcription at promoters and enhancers. FOXA1 binding-site mutants showed that FOXA1-binding was critical for the recruitment of the transcription factor AP2-γ. The binding of AP2-γ was also affected by deletion of the FOXA1 N-terminal intrinsically-disordered domain. Knockdown and rescue experiments also showed that the FOXA1 N-

terminus regulated H3K27 acetylation at estrogen-activated enhancers prior to ER-α recruitment without affecting the recruitment of FOXA1 or ER-α. Subsequently, proximity ligation mass spectrometry experiments suggested that FOXA1 was not only associated with a particular suite transcription factors, but also the Nuclear Remodeling and Deacetylase complex (NuRD). Estrogen-regulated enhancer FOXA1 binding-site deletion mutants were also shown to have decreased recruitment of the NuRD components CHD4 and HDAC2 compared to wild-type cells. Taken together, this data suggests that FOXA1 can regulate estrogen-dependent transcription through its N-terminus by recruiting AP2-γ and the NuRD complex.

INTRODUCTION

FOXA1 has been referred to as a pioneer factor based on its strong, sequence-specific association with DNA and its ability to displace the linker histone H1 as well as bind nucleosomes and affect their positioning around the FOXA1 binding site, which is thought to promote the recruitment of other transcription factors to enhancers (Cirillo et al., 2002; Meers et al., 2019). Furthermore, the forkhead box family of transcription factors, of which FOXA1 is a member, plays an integral role in the functions of nuclear hormone receptors that bind steroids such as estrogens, androgens, progesterone and glucocorticoids. For this reason, forkhead box transcription factors are of particular relevance to cancers of the breast, prostate, and ovaries, the majority of which feature steroid hormone receptor-dependent tumor growth patterns.

FOXA1 is composed of three structural domains: The N-terminal and C-terminal domains, which lack tertiary structure, and the globular winged-helix DNA binding domain. The winged-helix domain (WHD) is composed of a central helix-turn-helix domain which makes contact with the DNA major groove, and two flanking B-helices which are termed ‘wings’. Upon-binding the FOXA1’s forkhead domain DNA binding site, most frequently TGTTTAC, TGTTGAC, TATTTAC, or TATTGAC, the DNA binding domain B-helices wrap around the DNA and stabilize the interaction through contacts with the phosphate backbone (Cirillo et al., 2007). Mutation of the C-terminal wing of the WHD showed increased mobility in photobleaching experiments and greatly increased the expression of AR and ER-α in prostate cancer (Parolia et al., 2019). Alternative DNA motifs including AAATATTT appear to induce dimerization of FOXA1 and SNPs at these sites feature in several human pathologies (Wang et al., 2018), although the functional significance of FOXA1 dimerization is unknown.

Early work on FOXA1 showed that the C-terminal domain could bind histones H3 and H4 in vitro and help FOXA1 displace linker histone H1 at FOXA1 binding sites, contributing to its role as a pioneer factor (Cirillo et al., 2002). This C-terminal domain, but not the N-terminal domain, was also found to be necessary for the formation of a FOXA1-DNA condensate in the

absence of nucleosomes, which exerted a mechanical force in an optical tweezers experiment and was suggested to be important for maintaining promoter-enhancer contacts (Quail et al., 2021). The C-terminal domain also contains an alpha helix with a canonical TLE-binding peptide motif, which can interact directly with TLE3 and negatively regulates FOXA1's activity on Wnt-related genes (Wang et al., 2000; Parolia et al., 2019). Removal of the C-terminal domain due to frameshifts in one FOXA1 allele in prostate cancer led to increased transcription of FOXA1-regulated Wnt-related genes such as TCF-4 and LEF-1, where the mutant FOXA1 could displace full-length FOXA1. Further experiments showed that TLE3 presence at FOXA1-bound sites was drastically reduced in these cells, and that knockdown of TLE3 increased the expression of TCF-4 and LEF1 (Parolia et al., 2019). TLE3 was found to play a mixed regulatory role on estrogen-dependent transcription, and on estrogen-regulated TFF1 enhancers, shRNA knockdown of TLE3 increased H3K27ac and ER-α binding and decreased HDAC2 and CBP binding in the absence of estrogen (Jangal et al., 2014). These findings underscore the C-terminal domain's abilities to interact directly with both DNA and nucleosomes, but also regulate transcription through direct binding to one or more TLE isoforms.

At ER-α and FOXA1-bound sites, FOXA1 apparently acts as a quantitatively-important factor in determining expression levels. In one study, FOXA1 siRNA knockdown showed significantly altered expression in over 95% of ER-α-regulated genes, though the precise mechanism of this activity is not currently known. For instance, FOXA1 was found to be an important determinant of ER-α recruitment to estrogen-regulated enhancers, where 90% of ER-α binding sites exhibited a 50% or more decrease in ER-α recruitment upon treatment of MCF7 cells with siRNA targeting FOXA1, despite there being no change in ER-α expression (Hurtado et al., 2011). Additionally, upon estrogen treatment, ER-α binding can induce increased FOXA1 to nearby loci in trans, although there were very few new FOXA1 binding sites at forkhead motifs (Glont 2019). In our lab, we have identified a number of strongly activated estrogen enhancers

in MCF7 cells that exhibit not only increases in FOXA1 binding upon estrogen treatment, but also increases in several other transcription factors as a result of recruitment in trans, such as RAR-α, AP2-γ, and GATA3. We labeled them ‘Megatrans enhancers’ due to their estrogen-dependent assembly of a very large protein complex, characterized as 1-2 MDa by gel-filtration analysis, which unexpectedly also contained a number of DNA-binding transcription factors that were recruited in the absence of cognate DNA elements (Liu et al., 2014). FOXA1 was actually present in the 2 MDa Megatrans complex, as identified by gel filtration.

Although FOXA1’s role in ER-α-regulated Megatrans complex remains elusive, possible modes of regulation could include contribution to the function and formation of enhancer condensates. Proteins that contain extensive intrinsically disordered domains are thought to cooperatively mediate the stabilization of protein complexes at enhancers during transcription through phase-separated structures which are characterized by high frequency, low affinity interactions (Hnisz et al., 2017; Sabari et al., 2018). Work in our lab showed that active estrogen-dependent transcription at Megatrans enhancers was decreased by brief treatment with 1,6-hexanediol, which is commonly used to disrupt the hydrophobic interactions within these putative phase-separated structures in cells, while exerting no effect on transcription on other ER-α-regulated enhancers and non-ER-α-regulated enhancers (Nair et al., 2019). Furthermore, several components of the Megatrans complex exhibited properties of liquid-liquid phase separation *in vitro,* including fluorescence recovery after photobleaching in droplets of purified ER-α and GATA3, as well as precipitation of FOXA1, ER-α, and GATA3 by biotin-isoxazole, a compound used to isolate components of RNA granules and condensate proteomes (Nair et al., 2019; Han et al., 2012; Kato et al., 2012; Terlecki-Zaniewicz et al., 2021). Given the largely unstudied nature of FOXA1’s intrinsically-disordered N-terminus, it is not unreasonable to investigate whether it might play a role in the stabilization of protein complexes at ER-α-bound enhancers through cooperative activities within enhancer condensates.

FOXA1 plays a role in both the repression and activation of estrogen-dependent transcription (Hurtado et al., 2011). While C-terminal domain has been found to mediate particular FOXA1 pioneering activities as well as repression of Wnt-related genes, FOXA1's role in transcriptional activation is not fully understood. Recent work suggests that intrinsically disordered regions of transcription factors can stabilize protein complexes within enhancer condensates (Hnisz et al., 2017; Sabari et al., 2018; Nair et al., 2019), and FOXA1's N-terminus is a candidate for this type of interaction within the ER-α-regulated Megatrans complex. In this dissertation work, I sought to increase our understanding of the role of FOXA1's N-terminus in both activation and repression of ER-α and FOXA1-dependent transcription in the MCF7 breast cancer cell line. Using FOXA1 binding site knockouts as well as shRNA knockdown and rescue with wild-type and mutant FOXA1 as an experimental approach, I attempted to determine how FOXA1 interacts with coregulators to specify the MCF7 transcriptional program. This research is important for understanding how nuclear hormones promote organogenesis as well as cancer cell growth and survival, thereby contributing to a corpus of knowledge that will allow for improved cancer treatments.

Chapter 1: FOXA1's role in ER-α-dependent transcriptional activation

1.1. Features of strong enhancers

The evolution of multicellularity in metazoans was necessarily preceded by that of cis-acting regulatory mechanisms for complex signal and tissue specific gene expression patterns. One such category of regulatory elements referred to as transcriptional enhancers is a dominant mechanism that facilitates increased transcript diversity in metazoans. Intense research on enhancers have revealed important regulatory principles underlying various biological processes including development, tissue homeostasis and genetic basis of various disease conditions. A large fraction of enhancer-like elements can be identified using various epigenetics markers such as H3K27Ac and H3K4me1/2 (Creyghton et al., 2010; Heintzman et al., 2009; Rada-Iglesias et al., 2011) as well as higher-level chromatin binding proteins such as P300 (Visel et al., 2009) and mediator complexes (Soutourina et al., 2018). Active enhancer elements were shown to be generally nucleosome free and exist in open chromatin conformation (Boyle et al., 2008; Klemm et al., 2019). Diverse studies revealed that enhancers strength varies greatly, and one cohort of strong enhancers, referred to as super enhancers or stretch enhancers, were noted to be characterized by clusters of multiple enhancer elements in genomic window of 10-20 kb (Parker et al., 2013; Whyte et al., 2013), leading to the discovery that a proportion of very strong regulatory enhancers are themselves transcription units (Li et al., 2016; Liu et al., 2014b; Nair et al., 2019). Recent studies revealed that this class of non-coding RNA, also known as enhancer RNAs (eRNAs), are functionally important for enhancer activity and are likely present on all active enhancers (Hah et al., 2011; Kim et al., 2010; Li et al., 2013; Wang et al., 2011).

Coactivator ChIP-seq studies have revealed the recruitment of a large protein complex at most active enhancer loci, including several components of gene-regulatory assemblies such as the RNA Pol II machinery (12 subunits), mediator complex (~35 subunits), and DNA looping complexes that feature cohesin and condensin. Also included are pioneering and non-

pioneering transcription factors that are either bound directly to DNA or recruited indirectly, several transcription cofactors, components of DNA repair machinery, and nucleosome and histone modifying enzymes (Calo and Wysocka, 2013; Harlen and Churchman, 2017; Puc et al., 2017; Zaret and Carroll, 2011). Many of these proteins are decorated by post-transcriptional and translational modifications that further increase the complexity of interactions and promote local molecular condensation (Harlen and Churchman, 2017; Stillman, 2018). In addition, particular eRNAs important for the assembly of enhancer complex and enhancer function are also modified by methylation that plays crucial regulatory roles (Bi et al., 2020). Such spatial congregation of transcription factors on chromatin has been proposed to be mediated by cooperative binding of pioneer factors on DNA followed by protein-protein interactions resulting in the formation of large protein complexes (Soufi et al., 2015; Verger and Duterque-Coquillaud, 2002). Most of these clustered binding sites overlap with enhancers and super enhancers and thereby act as transcriptional hotspots (Chen et al., 2008; Siersbaek et al., 2014a; Siersbaek et al., 2014b).

There have been three prevalent models to explain the mechanisms of enhancer-promoter interactions (Blackwood and Kadonaga, 1998; Wang and Giaever, 1988): The looping model contends that tethering of enhancer to promoter regions by protein-DNA complexes allows for a decrease in distance between the two DNA elements and increased enhancer-promoter communication. Alternatively, the scanning model predicts that protein complexes bound to enhancers track along DNA until they arrive at promoters with complementary complexes resulting in protein-protein interactions that either activate or repress transcription. Although this model would explain the regulatory blocking property of insulator elements, it cannot explain observed interchromosomal enhancer-promoter associations (Lomvardas et al., 2006) and other phenomena such as chromosomal transvection (Duncan, 2002). Finally, the facilitated tracking model integrates features of both looping and scanning in which the protein complexes on enhancers scan along DNA till they engage in a looping interaction with a

cognate promoter. In recent years, the gene regulation field has made use of the advancements in super resolution microscopy and live cell imaging tools to study transcription factor dynamics and visualization of enhancer-promoter communication in live cells (Liu and Tjian, 2018).

Although the vast majority of enhancers and the cognate genes are located in the same genomic domains within transcriptionally-associated domains (TADs) or more heterogeneous insulated neighborhoods (Hnisz et al., 2016), there are many instances of long-distance genomic interactions beyond the confines of the TADs (Fanucchi et al., 2013; Mourad et al., 2014; Papantonis et al., 2012; Spilianakis et al., 2005), including inter-TAD interactions between enhancers and promoters (Proudhon et al., 2016) and multiple super enhancers activated in a cell type-specific manner (Allahyar et al., 2018; Beagrie et al., 2017; Markenscoff-Papadimitriou et al., 2014). Transcriptional activation by specific signaling events have also been shown to induce spatial proximity between co-regulatory genomic elements that are separated by millions of base pairs (Li et al., 2013; Mourad et al., 2014; Nair et al., 2019; Quintin et al., 2014). The molecular forces driving such inter-TAD or inter-chromosomal interactions are poorly understood.

Examining the details of super enhancer complex formation and function has thus far proven elusive, however other systems featuring strong, inducible transcription, such as that of estrogen-dependent enhancer activation by ER-α, continue to be powerful in understanding some of the basic mechanisms of eRNP establishment and function as well as intra-TAD and inter-TAD interactions.

1.2. Estrogen-dependent enhancer activation

Within MCF7 breast cancer cells and in response to estrogen, ER-α is recruited to 6000-8000 enhancer elements that have neighboring estrogen-induced target genes. Among these, a subset of approximately 1100 enhancers are particularly active, and are characterized by strong ChIP-seq signal for ER-α and transcriptional co-regulators as well as eRNA transcription (Li et al., 2013). Of these same genes, less than 30% overlap with the super enhancers, and a vast

majority of these enhancers had one binding site for ER-α (Liu et al., 2014b). In addition to the typical transcription machinery recruited on the enhancers, they also feature a large 2 MDa protein complex that is mostly composed of transcription factors including GATA3, AP2γ, RARα, RARγ, and FOXA1. Interestingly, these transcription factors are often recruited indirectly and not through their canonical DNA-binding motif, forming a large protein complex termed the Megatrans complex (Liu et al., 2014b). Architectural protein complexes such as condensin I and II, known to be involved in metaphase chromatin condensation, are also recruited to these sites in an estrogen-dependent manner (Li et al., 2015). The condensin complex was also shown to be critical for the formation of a network of interacting enhancers on the same chromosome (Nair et al., 2019). Furthermore, recent in vivo proximity labeling studies also showed that the nuclear effectors of the Hippo pathway, YAP/TEAD, were also specifically recruited to these enhancers through direct interaction with ER-α in an estrogen-dependent manner (Bi et al., 2020). Intriguingly, eRNAs were found to be a crucial scaffolding molecule in bringing together this multicomponent system, such that the depletion of one eRNA affected the transcription of neighboring eRNAs, suggesting the enhancer-bound complex is a ribonucleoprotein complex (Nair et al., 2019). Thus, a significant fraction of Megatrans enhancers may represent the most robust estrogen-regulated enhancers based on the acute ligand-dependent assembly of a large ribonucleoprotein complex (Liu et al., 2014b; Yang et al., 2017).

The vast majority of robust ER-α enhancers were prebound by the well-studied pioneering transcription factor FOXA1, which is also a key component of the Megatrans complex. Because of the importance of IDRs in the formation of putative RNP condensates at highly active enhancers, I have elected to investigate the critical properties underlying the actions of FOXA1 in estrogen-regulated transcriptional programs to better understand the roles of pioneer factors in the regulation of developmental and signaling programs.

1.3. FOXA1's role in estrogen-dependent enhancer activation

FOXA1 is a pioneer factor that is required for many steroid-receptor activated transcriptional events as well as other cell-type specific transcriptional programs. FOXA1 is highly expressed in male and female-specific organs, as well as in the bone marrow, liver, the gastrointestinal tract, lung, kidney and urinary tract. While this suggests that FOXA1 plays an important biological role in endodermal tissue-specific transcription, it is also expressed in mesodermally-derived sex-specific organs in the genitourinary system and in the neuroectoderm, where FOXA1 is found in the developing midbrain and has been suggested to be important for the differentiation of dopaminergic neurons (Uhlén et al., 2015, Lin et al., 2009).

There is evidence to show that FOXA1 could prime particular regulatory elements prior to the activation of transcription. In vivo footprinting revealed that FOXA1/2 are preloaded onto the Alb gene in undifferentiated hepatocytes, acting as pioneers with GATA4/6 in foregut endoderm and liver development before actual albumin expression (Bossard et al., 1998). DNA demethylation occurs following replication at FOXA1-bound sites, suggesting a passive mechanism that prevents methylation (Donaghey et al., 2018). TET1 and TET2 have also been shown to interact with FOXA1 and FOXA2 in differentiating embryonic stem cells, and may play a role in the demethylation of FOXA-bound sites (Yang et al., 2016; Ancey et al., 2017). Additionally, the replication-independent H2A.Z and FOXA1/2 co-occupy sites in developing embryonic stem cells and cancer cells, which together associate with H3K4 methylation and decreased nucleosome occupancy (Gevry et al., 2009; Aurelien et al., 2011; Li et al., 2012).

The FOXA1 DNA-binding domain is of the winged-helix family, which consists of 3 α-helices flanked by 2 specialized loop domains called 'wings' and includes other transcription-related proteins such as HSF, H1 and TFIIE/F. FOXA1's helix 3 (214-222) binds the major groove of DNA, while the wings make contact with the minor groove and DNA backbone, and it is thought that FOXA1 is allowed to bind DNA even when it is engaged by nucleosomes and can displace linker histone H1, thereby disrupting closed chromatin. Indeed, exo-ChIP-seq

revealed that FOXA1 preferentially binds to the center of nucleosomes, otherwise known as the dyad-axis, without actually removing nucleosomes from DNA (Cirillo et al., 2002; Iwafuchi-Doi et al., 2016). This is in concert with the finding that pioneer factors usually have short α-helical segments with high affinity major groove contacts with nucleosomal DNA (Fernandez Garcia et al., 2019). Furthermore, DNA features such as the repetition of AA or TT every 10 nucleotides and high GC content favor nucleosomal binding, and this high intrinsic nucleosome occupancy paired with active or neutral chromatin marks associate with FOXA1 binding through mitosis (Ioshikhes et al., 2006; Segal et al., 2006; Tilly et al., 2009; Caravaca et al., 2013). FOXA1's C-terminus (295-468) was shown to be required for it to engage and open DNAse and MNase-insensitive chromatin in vitro and for its affinity for nucleosomes at histones H3 and H4 (Cirillo et al., 2002). On the other hand, it was also found that the C-terminus may mediate sustained association with nucleosomes, while FOXA1 may not actually be impaired in the initial engagement of nucleosomes in vivo in the absence of its C-terminus from amino acids 270 to 468 (Parolia et al., 2019). While FOXA1 is often cited as binding stably and strongly to condensed sites on the genome, single molecule tracking indicates that FOXA1 association with chromatin is highly dynamic and that its residence time is only moderately longer than ER-α (Swinestead et al., 2016). Mechanisms by which initial FOXA1 binding has been proposed to regulate transcription include nucleosome phasing, such that FOXA1 binding at the dyad axis organizes nucleosome positioning (Ye et al., 2016; Cirillo et al., 2002). Recently however, particular motifs have been shown to favor direct binding by FOXA1 the dyad axis, where stronger binding at the canonical motif TGTTAC was more likely to displace nucleosomes altogether, creating a region with reduced nucleosome occupancy as determined by ATAC-seq (Meers et al., 2019).

Several FOXA1 mutations have been identified which modify the function of the FOXA1 DNA-binding domain, including the DNA-binding domain R219S mutation, show altered binding profiles from the canonical Forkhead motif and activation of non-FOXA1 target genes in prostate

cancer cells (Adams et al., 2019). Other mutations appear to increase FOXA1's ability to engage chromatin and activate transcription. These mutations were most commonly at the C-terminal end of the winged-helix domain in primary prostate tumors, with an increasing frequency of C-terminal IDR truncations and frameshifts in metastatic lines. Missense and indel mutations to the C-terminal wing of the winged-helix domain do not affect DNA-binding but increase FOXA1 diffusion to allow for faster recovery from photobleaching and increased AR transcriptional response, while mutations that lead to truncation of the FOXA1 C-terminus increased FOXA1's ability to bind DNA to the point where 13-times overexpression of the WT FOXA1 retained binding dominance of the truncation mutant at AR-bound sites in heterozygous LAPC4 cells, suggesting that the C-terminus may not only sustain but also restrict engagement with DNA. However, the FOXA1 C-terminal truncation mutants did not lead to a change to the WT cistrome, suggesting that removing the C-terminus did not increase FOXA1's pioneering activities (Parolia et al., 2019). It is unclear whether certain conditions that need to be met by chromatin to engage FOXA1 for cell-type specific binding patterns, such as H3K4 methylation, are dependent on the C-terminal domain (Lupien et al., 2008).

Recent evidence of intrinsically disordered protein domains mediating cooperativity between transcription factors points to a possible role for FOXA1's N and C-terminal domains in regulating its function on enhancers through an array of protein-protein interactions (Hnisz et al., 2018). For instance, siRNA knockdown of FOXA1 diminishes ER-α binding in 90% of sites by at least 50%, such that FOXA1 appears to be required for robust ER-α-dependent transcription at promoters (Lupien et al., 2008; Hurtado et al., 2011; Glont et al., 2019). Though the nature of the interaction between ER-α and FOXA1 is unclear, its importance in ER-α-dependent transcription underscores the possibility that its intrinsically disordered domains could play a role in cell-type specific gene-regulatory networks and estrogen-regulated transcription in MCF7 cells in particular.

1.4. FOXA1's unstructured domains

In order to perform reversible functions like catalysis and binding, proteins need structural flexibility both within and between structured domains. However, intrinsically disordered regions (IDRs) such as the FOXA1 N-terminal and C-terminal domains have incredible flexibility as well as no stable structure, and the functions of these domains are underscored by these characteristics. In fact, there is evidence to show that IDRs are important components of enhancer condensates which may undergo liquid-liquid phase-separation in an IDR-dependent manner, although it is an open question as to what determines the regulation and specificity of IDR interactions in transcription, and whether these interactions lead to liquid demixing in vivo (Boija et al., 2018, Nair et al., 2019). The existence of two 150-200 amino acid intrinsically disordered regions (IDRs) on either side of FOXA1's DNA binding domain presents an opportunity to explore the physical properties and biochemical functions of these types of protein domains as they relate to transcriptional regulation. For this dissertation work, I explore the roles of FOXA1's IDRs as well as small structured regions contained within these IDRs.

IDR amino acid sequences are commonly described by their composition, and often contain large numbers of small amino acids such as alanine, proline, glycine, and serine as well as uncompensated charge groups that disrupt the hydropathy required for the formation of stable secondary structure, usually contributing to a net negative charge (Uversky et al., 2000; Williams et al., 2001). Although FOXA1 has a net positive charge of 7 at neutral pH, its IDRs have net negative charges and the bulk of the basic residues are in FOXA1's DNA-binding domain. The PONDR VSL2 algorithm predicts two large IDRs that flank the FOXA1 winged-helix domain at positions 19-171 and 253-472. The N-terminal IDR has a net charge of -2 at pH of 7, is enriched in threonine, methionine and asparagine, and there are nine instances in which methionine and asparagine are adjacent to one another, indicating a functional relationship between these two amino acids. The C-terminal IDR has a net charge of -1 at pH of 7, and the region between positions 380 and 430 is enriched in residues such as leucine, isoleucine,

tyrosine, and phenylalanine which might facilitate interaction with hydrophobic pockets at interacting proteins.

Regions within IDRs that are thought to facilitate less promiscuous protein-protein interactions include short linear peptide motifs (SLIMs) or pre-structured motifs that contain intrinsic secondary structure (PreSMos) and regions that are predicted to undergo disorder-to-order transitions (ANCHOR) (Kim et al., 2018; Puntervoll et al., 2003; Dosztányi et al., 2009). Given that FOXA1's 380-430 is a predicted ANCHOR domain, it is potentially important for mediating more specific protein-protein interactions (Xue et al., 2010; Dosztányi et al., 2009). An IDR may have several interaction partners and stabilize protein complexes, as FOXA1's C-terminal IDR not only interacts with nucleosomes, but also contains an EH1 motif that directly binds the TLE family of WD40 domain-containing proteins (Puntervoll et al., 2003; Cirillo et al., 2002). In prostate cancer, FOXA1's C-terminus was found to bind TLE3, a corepressor in the WNT pathway, and the loss of this interaction allows for the expression of Wnt-related genes in prostate cancer and contributes to metastasis (Parolia et al., 2019). Notably, RIME ChIP-MS has demonstrated that ER-α precipitates both TLE3 and FOXA1 in MCF7 cells, indicating that the three may act together on estrogen-regulated enhancers (Papachristou et al., 2018). RIME ChIP-MS has also shown that both FOXA1 and ER-α can precipitate GRHL2 and GATA3, however the precise set of direct interactions that bring these proteins together is not known. Several other WD40 proteins, such as WDR5, WDR75, TBL3, and WDR46 were precipitated via FOXA1-RIME, indicating that involvement of these adapters in FOXA1-dependent transcription may extend beyond TLE3 (Jozwik et al., 2016; Papachristou et al., 2018).

IDR conformational energy landscapes vary greatly across the length of an IDR and are likely dependent on the cellular microenvironment. Particular IDR conformations are also dependent on modifications to specific amino acids, such as methylation, acetylation, and phosphorylation (Xie et al., 2007). FOXA1 is frequently phosphorylated in the unstructured C-terminal domain at S307, which overlaps with CDK as well as highly conserved Pin1, CK1, and

proline-directed kinase recognition motifs. CDK1 has already been confirmed to directly interact with FOXA1, although its role is unclear (Wang et al., 2018). Phosphorylation of S331 may be mediated by less conserved CDK CK1, GSK3, and proline-directed kinase motifs and proline isomerization may also be mediated by nearby Pin1 motifs (Puntervoll et al., 2003; Hornbeck et al., 2015). The placement of modified amino acids often represent key positions in the transition between collapsed and extended conformations, and the close proximity of Pin1 recognition motifs to phosphorylated residues could have functional consequences on FOXA1's ability to regulate transcription (Firman et al., 2018).

Flexibility of IDR conformation and amino acid properties in the context of different microenvironments and modifications likely allows for increased complexity of regulation and function in their interactions with other proteins, possibly playing a role in an IDR's ability to facilitate molecular crowding via high-frequency low-specificity interactions with diverse proteins and RNAs, which is thought to be important for promoting liquid-liquid or liquid-solid phase separation into membraneless organelles to perform various biochemical functions within the cell. For example, Xenopus egg nucleoli are capable of dripping and relaxing into spherical structures, having a protein concentration that is 2-fold higher than that of nucleoplasm (Brangwynne 2011, Handwerger 2005). Additionally, RNA-dependent transitions to hydrogel-like structures by FUS/EWS/TAF15, RBM3, hnRNPA1/2, and CIRBP were also dependent on low-complexity IDRs (Kwon et al., 2015; Kato et al., 2013; Molliex et al., 2015). It remains to be seen whether FOXA proteins play a role in mediating the formation of phase-separated structures in transcription, or whether their IDRs can affect the composition and/or enzymatic activities of these or similar structures.

With regards to FOXA1's N-terminal unstructured domain, methionine overrepresentation could be implicated in its activities within enhancer condensates. In yeast, oxidation of methionines in ataxin-2's low complexity domain causes liquid droplets to dissipate, which allows ataxin-2 to inhibit autophagy by releasing TORC1. Within this study, substitution of

methionine with tyrosine caused droplets to form in oxidative conditions, while substitution of serines abrogated phase separation altogether (Kato et al., 2019). Tyrosine and phenylalanine are frequently implicated in low complexity domains that promote phase separation, but methionine is less studied (Lin et al., 2017, Wang et al., 2018, Xiang et al., 2015). Methionines are typically buried within structured domains, serving important structural roles by mediating hydrophobic interactions as well as interactions with aromatic residues (Weber et al., 2019). Within the FUS protein, tyrosines flanked by glycines or serines could promote phase separation through interactions with unmethylated arginines adjacent to glycines (Kato et al., 2012; Qamar et al., 2018; Hofweber et al., 2018; Murthy et al., 2021). In this case, nearby glycines and serines were able to promote interactions between tyrosine and arginine within proteins in a condensed phase through increased hydrogen bonding (Murthy et al., 2021). It is possible that adjacent asparagines, serines, and threonines could also be important for promoting interactions between FOXA1's N-terminal methionines and other residues.

1.5. Methods for assessing FOXA1 function

The initial series of experiments I performed are based on systems that involve knocking down endogenous FOXA1 and overexpressing a series of wildtype and mutant constructs lacking portions of the N-terminal and C-terminal IDRs in MCF7 cells to investigate their effectiveness in mediating transcriptional activation as well as in recruiting FOXA1 and coregulators to sites on the genome, which will be pursued in a series of ChIP-seq and PRO-seq experiments. The PRO-seq experiments in particular have been made possible by establishing and validating permanent cell lines expressing comparable levels of each mutant under the control of doxycycline-responsive promoters. In parallel approaches, I attached the modified BirA-Turbo biotin ligase or APEX2 proteins to transduced FOXA1 and mutants, which allowed us to investigate potential near interactants with FOXA1 in the presence and absence of estrogen stimulation, and which will be explored in the second chapter of this dissertation. The goal of these experiments was to provide an understanding of how the intrinsically disordered

domains of FOXA1 may mediate cooperativity between transcription factors during estrogen-dependent activation as well as revealing more information about FOXA1 interactions than the previously-utilized RIME ChIP-MS method (Branon et al., 2018; Jozwik et al., 2016; Mohammed et al., 2013). The roles of potential short linear peptide motifs (SLIMs) within FOXA1 were also explored using these strategies.

1.5.1. Cloning FOXA1 constructs

The human FOXA1 coding sequence was cloned from PLX302_FOXA1-V5 (Addgene, cat. No. 70090) into pCDH-CMV (Addgene, cat. No. 72265), within which we cloned the 2xHA sequence for N-terminal tagging of FOXA1. Deletion mutants were created by site-directed mutagenesis with Agilent's QuickChange II kit (cat. No. 200523). APEX2 and Turbo-ID were added to the N-terminal end of FOXA1 constructs, which were cloned from pcDNA3 APEX2-NES (Addgene, cat. No. 49386) and V5-TurboID-NES_pCDNA3 (Addgene, cat. No. 107169), respectively. NEB-HiFi Assembly was used to clone methionine to alanine and tyrosine to alanine mutant N-terminal unstructured domains from synthetic fragments into pCDH-CMV which only contained the winged-helix and C-terminal domains of FOXA1 (IDT, gBlocks™; NEB, cat. No. E2621). For knockdown and rescue experiments and APEX2-mediated proximity labeling, we also used NEB HiFi Assembly to clone WT FOXA1 and mutants from pCDH-CMV plasmids into pRetroX (cat. No. 632104).

1.5.2. Preparing cells for knockdown and rescue

Trans-expressing FOXA1 cell lines are prepared using the retroviral pRetroX system. FOXA1 and mutant transgenes are cloned to pRetroX plasmids from Takara Biosciences, which express proteins in a doxycycline-dependent manner. Transgenic pRetroX plasmids are used to make lentivirus for infection of MCF7 that stably express Tet-On 3G as a result of transfection of pCMV-Tet3G from Takara Biosciences and selection with G418 (cat. No. 631335). After the Tet-On MCF7 cell line was established, 293T cells in 10 cm plates at 80% confluence were each transfected with 6 ug transgenic pRetroX, 6 ug PSPAX2, and 4 ug PMD2 for 16 hours. The

media was replaced and retrovirus was produced for 24 hours. TetOn expressing MCF7 cells were then treated with the transgene-containing pRetroX retrovirus for 24 hours, and the cells were then selected via puromycin for 2-weeks following successful transduction to generate cell lines. With the exception of the N-terminal and C-terminal IDR deletion mutants, overexpression of transgenes to levels comparable to wildtype required 1 ug/mL doxycycline. N-terminal and C-terminal IDR deletions overexpressed at this concentration, so doxycycline treatments were reduced to 0.5 ug/mL for these mutants.

For each experiment, shFOXA1 lentivirus is prepared and used to infect four 10 cm plates of transgenic MCF7 cell lines at 50% confluence for 24 hours. Sufficient 293T cells at 80% confluence are transfected with PLKO.1 plasmid containing the shRNA sequence 5'-GAGAGAAAAAATCAACAGC-3', which targets the FOXA1 mRNA 3'UTR and allows for the expression of transgenic FOXA1. Media is changed after 16 hours, and MCF7 cells are infected at a confluence of 50%. Cells are then stripped of estrogen by media replacement and treated with 50 nM 17β-estradiol for 1 hour before harvesting nuclei and commencing the Pro-seq or ChIP-seq protocol (Kwak et al., 2014). For experiments that include the degradation of the estrogen receptor to assess estrogen-independent transcription and DNA binding, the selective estrogen receptor degrader ICI-182780 was used to rapidly remove ER-α from MCF7 cells. In this case, cells were treated with 100 nM ICI for 4 hours prior to harvesting.

1.5.3. Chromatin immunoprecipitation and CUT&TAG assays

In all ChIP experiments, cells were incubated in 1% formaldehyde in PBS for 10 minutes and the reaction was quenched using 0.125 M glycine. Following nuclear isolation, sonication was performed with the diagenode BioRuptor Pico until chromatin exhibited sufficient fragmentation, and aliquoted into tubes such that each contained 1 million cells equivalent. Each aliquot of sonicated chromatin was used to perform one ChIP assay. FOXA1 ChIPs were performed using a mouse anti-HA antibody from Santa Cruz (cat. No. sc-7392) and an N-terminal antibody rabbit polyclonal antibody from Fortislife (cat. No. A305-249A). ER-α ChIPs on

FOXA1 shRNA knockdown and rescue experiments were performed with a rabbit polyclonal antibody from Millipore-Sigma (cat. No. 06-935), and ChIPs for FOXA1 binding-site deletion assays were performed with a mouse monoclonal antibody from Santa Cruz (cat. No. sc-8002). H3K27ac ChIPs were performed with a rabbit polyclonal antibody from Abcam (cat. No. ab4729). For HDAC2 ChIPs, a rabbit polyclonal from Santa Cruz was used (cat. No. sc-7899). CHD4 ChIPs were performed with a rabbit polyclonal antibody from Cell Signaling Technology (cat. No. 12011). GATA3 and AP2-γ ChIPs were performed with rabbit monoclonal antibodies from Abcam (cat. Nos. ab199428, ab218107). For the AP2-γ CUT&TAGs, purified PAG-Tn5 was prepared and the assay was performed as described previously (Kaya-Okur et al., 2019).

1.5.4. Precision Run-On sequencing

PRO-seq proceeded as described (Mahat et al., 2016), with a few modifications to the protocol. Decapping enzyme (NEB, cat. No. M0608) at 0.5 uL/sample was used to digest latent 5' caps during PNK treatment (NEB, cat. No. M0201). Libraries were prepared using the NEB small RNA library preparation kit (NEB, cat. No. E7330), where instead of performing gel size-selection and purification, magnetic size selection beads were used to remove adapters (Omega Bio-Tek, cat. No. M1378). Biotin-tagged RNA was purified using Dynabeads Streptavidin (ThermoFisher, cat. No. D65606).

1.5.8. Preparation of cell lines with FOXA1 binding site deletions

The CRISPR-Cas9 system was used to generate genomic deletions for the FOXA1 binding site for several estrogen-regulated Megatrans enhancers, including those containing the strongest ER-α peaks and enhancer RNA expression that regulate the TFF1, FOXC1, and NRIP1 genes. 2-10 20 nt sgRNAs were designed toward a 250 base pair region flanking the FOXA1 binding site, and were transfected with a donor plasmid containing the binding site deletions with 1 kb homology arms (Mali et al., 2013). After 2 weeks, colonies were assessed via PCR from the FOXA1 binding site to flanking regions. The TFF1 enhancer FOXA1 binding site is deleted within a 41 bp region, which was easily assessed via PCR (Figures 1A and 1B).

For the FOXC1 and NRIP1 enhancer FOXA1 binding site deletions, a more minimal deletion was performed consisting of the 6 base pair forkhead box, whereby mutants were assessed by using primers that contained this sequence. Successful heterozygous mutants were subjected to additional rounds of CRISPR-Cas9 mediated deletion using the same protocol. Homozygous mutants were then assessed for FOXA1 recruitment via ChIP-seq (Figure 1C).

1.6. FOXA1 modulates ER-α-dependent transcriptional activity

Previous studies have shown that ER-α may increase FOXA1 binding. Recently, this increase in ChIP signal was attributed to chromatin looping upon ER-α binding and was supported by ChIA-PET assays showing associations between estrogen-induced FOXA1-binding sites and non-induced FOXA1 binding sites (Glont et al., 2019). In our lab, we used DNA-FISH to show that ER-α could induce increases in proximity between estrogen-activated Megatrans enhancers, and this effect was reversed by the addition of 1,6-hexanediol (Nair et al., 2019). We postulated that these effects are the result of the formation of enhancer ribonucleoprotein complexes which allow for the association between enhancers via liquid-liquid phase separation and which is dependent on the intrinsically disordered regions of key transcription factors within the Megatrans complex, namely GATA3 and ER-α. Furthermore, it has been previously shown that knockdown of FOXA1 regulates ER-α binding by at least 50% on a majority of ER binding sites, and that several estrogen-regulated genes showed altered patterns expression (Hurtado et al., 2011). The mechanism by which this occurs had yet to be uncovered, however there were several hypotheses that we wanted to explore. For instance, FOXA1 could be involved in regulating nucleosome positioning or post-translational modification of histones, which then affects ER-α binding either directly or through the activities of other transcription factors. FOXA1 could also play a role in the formation of enhancer condensates, which we believe to be important for estrogen-activated transcription at Megatrans enhancers.

First, we wanted to see how Megatrans enhancers were affected by the removal of the FOXA1 binding site. We chose three genes for this analysis: FOXC1, NRIP1, and TFF1. Each

of these genes are positively regulated in MCF7 cells by strongly activated estrogen-dependent enhancers. These enhancers exhibit a number of intriguing properties, including a strongly induced transcription of enhancer RNAs as well as a strongly increased recruitment of FOXA1 upon estrogen treatment. Removing the FOXA1 binding site allows us to gauge whether knockdown of FOXA1 has an indirect effect on transcription on these enhancers, and whether the FOXA1 binding site is important for the increases in FOXA1 binding that we see in the presence of estrogen. In previous studies, FOXA1 knockdown was incomplete (Hurtado et al. 2011). Thus, removing the FOXA1 binding site could allow us to see whether low levels of FOXA1 binding exert some effect on the function of these enhancers.

As expected, we saw 50% or more increases in FOXA1 binding at these enhancers in WT cells, and FOXA1 binding was dramatically decreased. However, there still appeared to be an increase in FOXA1 binding upon estrogen treatment, which might be attributed to increases in ER-α binding. To assess the contribution of FOXA1 to estrogen-dependent transcription, we performed nuclear RT-qPCR. TFF1, NRIP1, and FOXC1 all showed decreases in estrogen-dependent transcription at both promoters and enhancers by 50% or more (Figure 1D). This shows that FOXA1 binding to TFF1, NRIP1, and FOXC1 enhancers contributes greatly to the transcription of these genes.

We wondered how this compared to the shFOXA1 knockdown, and identified a shRNA target toward the FOXA1 3'UTR that yielded an 80% knockdown (Figure S2A). We found that while TFF1 showed a decrease in both eRNA and mRNA transcription upon FOXA1 knockdown, NRIP1 showed an increase in eRNA transcription and FOXC1 showed an increase in mRNA transcription (Figure S2B). In order to assess whether FOXA1 and ER-α play independent roles in the transcriptional regulation of these genes, we also treated cells with the drug ICI-182780, which is a selective estrogen degrader commonly used to treat ER-positive breast cancer. Removal of ER-α led to increases in both NRIP1 eRNA and mRNA transcription in the ICI-treated condition compared to estrogen. Furthermore, ICI treatment coupled with

shRNA knockdown of FOXA1 led to increases in transcription in TFF1 and NRIP1 eRNA as well as FOXC1 and NRIP1 mRNA transcription. This suggests that FOXA1 is playing a repressive role in transcription at these genes, and that ER-α cooperates with FOXA1 to promote aspects of this repression (Figure S2B). The question remains as to why the deletion of the FOXA1 binding site shows different expression profiles in the estrogen condition compared to the shFOXA1 estrogen-treated condition, where NRIP1 eRNA is increased upon FOXA1 knockdown and there is no significant change in FOXC1 mRNA and eRNA transcription. It is possible in this case that latent FOXA1 binding is sufficient for rescue aspects of FOXA1 function at these enhancers in the shFOXA1 condition. Additionally, ICI treatment appears to increase transcription of TFF1 eRNA as well as NRIP1 eRNA and mRNA compared to cells that are not treated with estrogen (Figures 1D, S2B). This could be due to ER-α-mediated repression in the absence of estrogen.

Next, we wanted to see whether decreases in estrogen-dependent transcription seen in the FOXA1 binding site mutants might be attributed to decreases in ER-α binding or in the recruitment of other transcription factors to NRIP1, FOXC1, and TFF1 enhancers. We found that ER-α binding profiles change in a similar manner as those in the previous study where FOXA1 is knocked down (Hurtado et al., 2011). At TFF1 and NRIP1 enhancers, ER-α binding decreased by 50% in the FOXA1 knockdown, but the FOXA1 enhancer showed no significant decrease in ER-α binding. This suggests that FOXA1 does not play a definitive role in determining ER-α binding.

In previous work, we have established that GATA3 and AP2-γ play important roles in estrogen-dependent transcription at Megatrans enhancers, where binding sites for these transcription factors are also frequently found and recruitment is increased upon estrogen treatment (Liu et al. 2014). To ascertain FOXA1's role in recruiting these transcription factors to enhancers, we performed ChIP-qPCR. In cells treated with FOXA1 shRNA, there was no significant difference in GATA3 binding, both in the presence or absence of estrogen, on NRIP1,

TFF1, and FOXC1 enhancers when compared to cells treated with non-targeting shRNA. On the other hand, AP2-γ showed significant decreases in binding to these enhancers in the estrogen-treated condition (Figure 2B). AP2 proteins have been shown to interact directly with P300 to activate transcription (Bragança et al., 2003), and are AP2 motifs co-occur within 300 bp of FOXA1 peaks at Megatrans enhancers (Figure 2D).

H3K27ac has been previously reported to be a key determinant of enhancer function, and we wanted to know how FOXA1 and ER-α might cooperate to drive histone acetylation at TFF1, NRIP1, and FOXC1 enhancers. H3K27ac ChIP-qPCR revealed that estrogen treatment resulted in high levels of H3K27 acetylation at these enhancers, and that H3K27 acetylation was decreased both in the presence and absence of estrogen at TFF1 and NRIP1 enhancers. Surprisingly, H3K27ac increased at the FOXC1 enhancer in the estrogen-treated condition, which correlated with ER-α binding (Figure 2C). This suggests that H3K27ac is a stronger predictor of ER-α binding than FOXA1, but also that increases in ER-α or H3K27ac may not be enough to drive enhancer transcription independently of FOXA1.

1.7. FOXA1's N-terminal domains regulate Megatrans enhancer transcription

In order to understand how FOXA1 might be regulating transcription at Megatrans enhancer, we decided to perform a series of knockdown and rescue experiments with different FOXA1 mutant transgenes. We used Alphafold2 and Phyre2 to identify FOXA1 N-terminal regions of interest (Jumper et al., 2023; Kelley et al., 2015), including a small alpha-helical structured region (a.a. 23-26), and a large unstructured domain (a.a. 27-161). We reasoned that the N-terminal alpha-helix, which we called the NT-SLIM, could potentially form specific protein-protein interactions that might be involved in the regulation of transcription by FOXA1, and also that the IDR might be involved in important low-specificity protein-protein interactions that may nonetheless be an important component of FOXA1's function at enhancers (Figure 3A). To gauge the effect of the FOXA1 knockdown and rescue on estrogen-induced transcription, we used PRO-seq to measure actively transcribed RNAs at Megatrans enhancers (Figure 3B). To

define our cohort of Megatrans enhancers, we performed a FOXA1 ChIP-seq (Figure 4A) and selected only those previously identified estrogen-activated enhancers with FOXA1 peaks, yielding a total of 1158 enhancers. We found that deletion of the N-terminal intrinsically disordered region (NT-IDR) as well as the NT-SLIM showed a greatly decreased estrogen-induced transcriptional response compared to WT cells, where across these Megatrans enhancers there was approximately a 50% decrease in the average transcription, which was similar to the FOXA1 shRNA knockdown (Figure 3B). Further analysis showed that there was a significant overall decrease in Megatrans eRNA transcription both with and without estrogen in the NT-IDR deletion mutant (Figure 3D). Additionally, across 321 identified genes where there was a significant 2-fold or more induction of mRNA transcription upon estrogen treatment in WT cells, both the FOXA1 knockdown and the NT-IDR deletion mutant showed similarly significantly decreased expression profiles, although this was not the case for the NT-SLIM deletion mutant (Figure 3C).

When examining the distribution of expression values for Megatrans enhancers and the expression of the nearest mRNA genes, knocking down FOXA1 caused approximately one-third of enhancers and nearby genes to decrease their transcription in the estrogen-treated condition, while another third showed increases (Figure 3E). This is in-line with microarray analysis of FOXA1 siRNA knockdown showing mixed effects on gene expression compared to control cells (Hurtado et al., 2011). However, while NT-IDR deletion did not rescue robust estrogen-dependent transcription, it also showed reduced transcription on enhancers that increased in transcription in the FOXA1 shRNA knockdown, though not to the level seen in the WT FOXA1 rescue. On the other hand, while the NT-SLIM deletion showed a profile similar to the NT-IDR deletion for FOXA1 repressed and activated enhancers, there was a cohort of Megatrans enhancers that were robustly activated in the NT-SLIM deletion compared to WT cells and the NT-IDR deletion, many of which also showed increases in transcription in the FOXA1 knockdown (Figure 3E). Together, this data indicates that the FOXA1 N-terminus plays a role in

activating more robustly transcribed Megatrans enhancers while also repressing less actively transcribed ones. Indeed, while transcription of the highly activated TFF1 enhancer decreased significantly in the NT-IDR deletion, NRIP1 and FOXC1 are much less actively transcribed and showed less overall differences in their expression profiles (Figure 5F).

In order to better understand the biochemical mechanism by which the FOXA1 NT-IDR regulates the activation and repression of Megatrans enhancer transcription, we created NT-IDR methionine to alanine and tyrosine to alanine mutants (Figure S5A). We found that on average, there was less transcription in both the presence and absence of estrogen compared to the WT FOXA1 rescue cells (Figure S5B). The expression profiles of these mutants were similar to that of the NT-IDR deletion in that there was less robust transcription of genes that were repressed upon FOXA1 shRNA knockdown, and reduced repression of genes that were inactive in the WT FOXA1 rescue. However, like the NT-SLIM mutant, there were groups of Megatrans enhancers and adjacent genes that were more active in these mutants. Curiously, several genes and enhancers that lost expression in the FOXA1 shRNA knockdown had decreased transcription in the methionine mutant compared to the NT-IDR deletion (Figure S5C). This suggests that methionine and tyrosine might be important for specifying which genes and enhancers become activated or repressed in a FOXA1-dependent manner, perhaps through interactions with specific transcription factors or coregulators.

Next, we wanted to see whether the decreases in transcription as a result of deletion of N-terminal domains could be attributed to changes in the binding of FOXA1, ER-α, or AP2-γ to Megatrans enhancers. FOXA1 ChIP-seq shows no significant decreases to the binding of FOXA1 as a result of NT-IDR deletion, but there was a decrease in the recruitment of the FOXA1 NT-SLIM mutant. Furthermore, across the Megatrans enhancers there was a slight increase in binding as a result of estrogen treatment in the WT cells and the NT-IDR mutant, but this was not the case for the NT-SLIM mutant (Figure 4A). It is possible that the NT-SLIM plays a role in bringing FOXA1 to Megatrans enhancers through protein-protein interactions in both

the presence and absence of estrogen or is somehow involved in priming chromatin for increased FOXA1 association.

Since the NT-IDR mutant and FOXA1 shRNA knockdown show similarly decreased levels of estrogen-dependent transcription, we wanted to see whether ER-α was differentially recruited to Megatrans enhancers between WT and the mutants. Across Megatrans enhancers, however, there was no significant difference in ER-α binding. This indicates that the FOXA1 N-terminus is able to regulate transcription independent of FOXA1's ability to promote ER-α binding (Figure 5B). To see whether this difference in transcription could be attributed to another Megatrans transcription factor, we performed an AP2-γ CUT&TAG. Compared to the FOXA1 WT rescue cells, AP2-γ binding is greatly reduced in the shRNA FOXA1 knockdown (Figure 4C). Furthermore, rescue with the FOXA1 NT-IDR and NT-SLIM mutants did not increase AP2-γ binding to levels seen in the WT rescue (Figure 4D). Since AP2-γ has been shown to be an important component of Megatrans enhancer activation (Liu et al., 2014), it is possible that the FOXA1 N-terminal domain regulates activation in part through the recruitment of AP2-γ.

Previous studies have shown that knockdown of the FOXA1-associated factor TLE3 reduced HDAC2 and increased H3K27ac and RNA pol II on TFF1 enhancers in the absence of estrogen, yet we found that deletion of the FOXA1 binding site at TFF1 and NRIP1 reduces H3K27ac in both the presence and absence of estrogen (Figure 2C). We therefore wanted to investigate the role of specific regions of FOXA1 in the regulation of H3K27ac across all Megatrans enhancers. Endogenous FOXA1 knockdown and rescue with WT FOXA1 showed reduced H3K27ac in ICI-treated cells whereas estrogen-treatment increased H3K27ac (Figure 6A). While deletion of FOXA1's NT-SLIM also showed reduced H3K27ac in the ICI condition, deletion of the NT-IDR showed H3K27ac levels close to that of the estrogen-treated condition (Figures 5B, 5C). Additionally, on FOXA1 enhancers that had over 2-fold increased levels of H3K27ac, which made up 45% of all FOXA1 enhancers, in the absence of estrogen in the NT-IDR deletion when compared to WT cells there was also significant increase in transcription in

the absence of estrogen (Figure 5D). This change can be attributed to enhancers that were silenced in the WT condition but gain transcription in the FOXA1 NT-IDR deletion (Figure 5E). However, not all Megatrans genes showed high levels of H3K27ac in the absence of estrogen. TFF1, FOXC1, and NRIP1 enhancers all showed strongly induced H3K27ac in the presence of estrogen, though FOXC1 and NRIP1 had increased H3K27ac in the presence of estrogen (Figure 5F). Together, these data suggest that FOXA1 NT-IDR plays a role in the maintenance of low H3K27 acetylation on a cohort of Megatrans enhancers in the absence of estrogen.

1.8. Discussion

FOXA1 plays a complex and important role in the regulation of estrogen-dependent transcription. Much like ER-α, FOXA1 appears to behave like a switch that alternates between activation and repression in an estrogen-dependent manner. In the presence of estrogen, FOXA1 cooperates with ER-α to recruit AP2-γ to Megatrans enhancers to activate transcription, but not GATA3, which also binds nearby (Figure 2D). Furthermore, although removal of the FOXA1 binding site reduces H3K27 acetylation, the deletion of the FOXA1 NT-IDR reduces H3K27 acetylation when ER-α is also removed and increases Megatrans enhancer transcription in the absence of estrogen. This difference may be attributed to other regions of FOXA1 which could retain their functions in the N-terminal mutants, such as the C-terminal domain, which interacts with histones H3 and H4 and displaces histone H1 *in vitro* (Cirillo et al., 2002), and FOXA1's DNA-binding domain, which recruits TET proteins to promote DNA hydroxymethylation and also influences nucleosome positioning (Yeqing et al., 2016; Pierce-Benoit et al., 2017; Wang et al., 2018; Meers et al., 2019). FOXA1's ability to organize nucleosomes, which potentially requires the activities of the C-terminal domain, could be necessary for the binding of important coregulators that promote H3K27 acetylation on Megatrans enhancers such as those that regulate TFF1 and NRIP1. Additionally, FOXA1's C-terminal domain interacts with TLE3 through the structured EH1 motif. TLE3 has been shown to be important for FOXA1's ability to repress transcription on Wnt-related genes, and knockdown of TLE3 increases H3K27ac on the

TFF1 enhancer (Parolia et al., 2019; Jangal et al., 2014). It is possible that the FOXA1 N-terminus plays a role in these TLE3-dependent activities.

1.9. Acknowledgements

I would like to acknowledge Dr. Lu Yang, who constructed the MCF7 FOXA1 binding-site deletion mutant cell lines for the TFF1, NRIP1 and FOXC1 estrogen-dependent enhancers discussed in this chapter. I would also like to acknowledge Dr. Wei Yuan, who purified the PAG-Tn5 used for the AP2-γ CUT&TAGs, and Kenneth Ohgi, who constructed initial WT FOXA1 and mutant expressing pCDH plasmids and who also prepared the ChIP-seq libraries contained in this chapter.

Chapter 2: FOXA1 regulates transcriptional repression by coordination of NuRD

2.1. FOXA1-associated chromatin modification and nucleosome remodeling

In vivo, FOXA1 is required for robust ER-α binding to condensed chromatin in MCF7 cells, and the silencing of FOXA1 decreases global chromatin accessibility as assessed via FAIRE-seq (Hurtado et al., 2011). In MCF7 cells the BAF complex ATP-dependent chromatin remodeler ARID1A colocalized with FOXA1 at 78% of ChIP-seq peaks and with ER-α at 66% of peaks. Strikingly, the distribution of ARID1A peaks did not change significantly in the absence of estrogen. In this study, 12% of ARID1A peaks shared by FOXA1, SMARCA4 and ER-α were lost following treatment of MCF7 cells with FOXA1 siRNA. Interestingly, although ARID1A knockout showed no effect on H3K27ac distribution, there was increased H4 acetylation and BRD4 recruitment as well as reduced HDAC1 binding at enhancers that became active in the ARID1A knockout (Nagarajan et al., 2020). Furthermore, knockdown of FOXA1 or knockout of ARID1A stimulates expression of genes that define the basal tumor phenotype in ER-positive luminal breast cancer cell lines (Bernardo et al., 2013; Xu et al., 2020). ARID1A also played a minimal role in recruiting FOXA1 and ER-α, where ARID1A knockout reduced FOXA1 and ER-α binding at sites shared by BRG1 and BAF155 (Xu et al., 2020).

FOXA1 also appears to play a limited role in regulating activation of histone methylation. For instance, although siRNA knockdown of FOXA1 in MCF7 cells reduced genome-wide binding of the H3K4 methyltransferase KMT2C, H3K4 monomethylation only decreased at a minority of sites (Jozwik et al., 2013). While FOXA1 is recruited preferentially to H3K4 methylated chromatin, FOXA1 siRNA knockdown had a minimal effect on global H3K4 methylation and even contributed to a slight increase of the chromatin mark on a few loci in LnCAP cells (Lupien et al., 2008). In T47D cells, KMT2D was required for ER-α and FOXA1 binding to estrogen-regulated enhancers, suggesting that the presence of H3K4 methylation might actually precede FOXA1 binding (Mo et al., 2006; Toska et al., 2017).

In accord with the direct protein-protein interactions with FOXA1, TLE3 is recruited to FOXA1-regulated enhancers via its C-terminal domain to repress Wnt-related gene transcription as well as TFF1 expression in the absence of estrogen in MCF7 cells (Parolia et al., 2019; Jangal et al., 2014). It is unknown, however, whether this interaction is important for FOXA1's ability to regulate H3K27 acetylation through its N-terminal domain across all Megatrans enhancers (Figure 6C). In this chapter, we seek to increase our understanding of the regulatory modes of the FOXA1 N-terminal domain through the initial unbiased approach of proximity labeling followed by exploration of candidate proteins.

2.2. FOXA1 proximity ligation experiments

While FOXA1 is a major determinant of estrogen-dependent transcriptional events, it is unknown how FOXA1 interacts with specific partners to perform its functions. As a so-called pioneer factor, it might facilitate the binding of a suite of coregulators to enhancers to elicit particular transcriptional responses. The identification and classification of putative coregulators in this section was aided by proximity labeling assays, where proteins adjacent to Turbo-BirA or APEX2-tagged FOXA1 and our mutants of interest are biotinylated *in vivo.* This allowed for their efficient isolation using streptavidin beads and subsequent identification using mass spectrometry (Hung et al., 2016; Branon et al., 2018).

2.2.1. Preparation of cells for TURBO-ID proximity labeling

To prepare cells for proximity labeling with Turbo-ID, 6 ug FOXA1 and mutant expressing PCDH lentiviral vector with 6 ug PSPAX2 and 4 ug PMD2 were transfected into 10 cm plates that were 80% confluent with HEK293T cells for 16 hours with the Lipofectamine 2000 transfection reagent. Following production of lentivirus for 24 hours in 20 mL of fresh media per 10 cm plate, MCF7 cells are infected at 20% confluence immediately following passage. Media is replenished after 24 hours of infection, and cells are grown to 80% confluence. Each plate of cells is selected with puromycin after the next passage into two 10 cm plates such that there are two replicates for each estrogen and vehicle condition. After 24 hours

of shFOXA1 lentiviral infection starting 24 hours post-passage, cells were given stripped media which lacks steroid hormones and indicator dye, and 72 hours later estrogen is added to cells 1 hour before harvesting nuclei as described previously (Mahat et al., 2016).

2.2.2. Turbo-ID proximity labeling and mass spectrometry

Following transduction and selection of N-terminally Turbo-BIR-tagged FOXA1 and mutants, FOXA1 is then knocked down, and cells are incubated with 100 nM estrogen for 30 minutes. Cells are then incubated with 100 μM biotin for 15 minutes and harvested in ice-cold PBS. 10% of cells are saved for evaluation for transgene expression and labeling efficiency, nuclei are then extracted using the method outlined in the PRO-seq protocol, where RNase inhibitors are replaced by protease inhibitors, and biotin pulldown was performed as described previously (Kwak et al., 2014; Liu et al., 2019). Proximity labeling and pulldown efficiency was assessed via Western blot analysis with mouse monoclonal anti-biotin antibody from Santa Cruz (cat. No. sc-53179) and Pierce silver stain (cat. No. 24612) before submission for mass spectrometry.

2.2.3. APEX2-mediated proximity labeling and mass spectrometry

We used the previously described method in this dissertation to generate MCF7 lines that expressed WT FOXA1 and mutants that were N-terminally tagged with the APEX2 protein in a doxycycline-inducible manner via the RetroX tet-On system from Takara Biosciences. Each cell line was treated with shFOXA1 lentivirus for 24 hours and supplied with estrogen-deficient media for 72 hours as previously described in this dissertation, where 1 ug/mL doxycycline was added for the expression of APEX2-tagged transgenes. Prior to the biotinylation reaction, the cells were then treated with 100 nM estrogen and 500 μM biotin-phenol for 30 minutes. Hydrogen peroxide was then added for 1 minute at a final concentration of 1 mM and the cells were harvested in ice-cold PBS (Hung et al., 2016). Preparation of nuclear lysate and pulldown of biotinylated proteins was achieved using the protocol outlined for the Turbo-ID assay in the previous section.

2.2.4. FOXA1 proximity ligation reveals association with NuRD

In order to further understand the role of the FOXA1 N-terminal domains in the regulation of estrogen-dependent transcription, we elected to perform a series of proximity ligation assays coupled with mass spectrometry to identify proteins that might be acting as coregulators. To start, proximity labeling with Turbo-BirA tagged FOXA1 revealed a few complexes that likely act with FOXA1 to regulate transcription. Upon treatment with estrogen, WT FOXA1 gains several interactions with DNA binding transcription factors including C/EBPβ and the AP-1 subunits c-Jun and JunD. Several transcription-related cofactors also gain association in the estrogen treatment, including the NuRD complex factors NCOR2, HDAC2, and MBD2 as well as the BAF complex factors SMARCE1, SMARCC2 and ARID2 (Figure S4B). These complexes also feature among the 50 highest-confidence transcription-related interactors in the estrogen-treated condition, where CHD4 was strongest associated NuRD factor, and ARID1A was the BAF factor that had the most associated peptides. Within those top interactors, we also identified the NuRD complex factors CHD3, CHD5, GATAD2A, and MTA2, as well as the BAF complex factors SMARCA1, SMARCA2, SMARCA4, SMARCA5, and ARID1B. Within this list, there were also several DNA-binding transcription factors that have been identified to share ChIP-seq peaks with FOXA proteins, including CUX1, HCFC1, and C/EBPβ (Figure S4A; Tables 1,2; Partridge et al., 2020).

We also performed Turbo-ID on FOXA1 N-terminal mutants concurrently to assess differences in associated proteins. Proximity labeling with the NT-IDR deletion mutant in the presence of estrogen yielded increased peptides from several factors of the NuRD complex, especially GATAD2B, compared to WT FOXA1. Conversely, deletion of the NT-IDR decreased the number of peptides associated with ARID1A and ARID1B of the BAF complex. Proximity labeling with Turbo-BirA-tagged NT-SLIM in the presence of estrogen yielded decreased numbers of peptides from all members of the NuRD complex, with the exception of GATAD2B, compared to WT FOXA1. Furthermore, although GATAD2A peptides were abound in the WT

and NT-IDR deletion, there were none identified in the NT-SLIM deletion. The NT-SLIM deletion also yielded less peptides from members of the BAF complex, showing a similar profile to that of the NT-IDR deletion (Tables 1, 2).

We subsequently performed APEX2-mediated proximity labeling, which has a number of key advantages such as reduced labeling time and lower background biotinylation. APEX2-tagged WT FOXA1 yielded peptides from many of the same proteins as the Turbo-BirA-tagged WT FOXA1. For instance, both BAF and NuRD complexes feature prominently in peptide profiles from WT FOXA1, however the distribution of peptides among the NT-IDR and NT-SLIM differs from that of the Turbo-ID assay. While proximity labeling with APEX2-tagged FOXA1 shows reduced biotin ligation of BAF complex factors by the NT-IDR deletion mutant in the presence of estrogen, there was also reduced biotinylation of NuRD complex factors compared to APEX2-tagged WT FOXA1. Furthermore, while the NT-SLIM deletion showed similar levels of biotinylation of BAF complex factors to WT FOXA1, its association with CHD4, CHD5, and CHD6 was reduced (Table 3).

The most striking difference between Turbo-ID and APEX2-mediated proximity labeling was APEX2's capacity to biotinylate DNA-binding transcription factors that have been previously been associated with estrogen-dependent transcription. For instance, APEX2-tagged FOXA1 yielded peptides from the ER-α coregulators AP2-y, GATA3, GRHL2, TLE3, NFIX, and NFIB, although ER-α itself was not identified in either proximity labeling assay (Tables 1, 2, 3). While AP2-y and GATA3 peptides were not reduced in proximity labeling with the FOXA1 NT-IDR deletion compared to WT FOXA1, less NFIX and NFIB peptides were identified in both the NT-IDR and NT-SLIM deletion mutants. Additionally, the FOXA-associated factors HCFC1 and C/EBPβ showed decreased labeling in both APEX2-tagged N-terminal mutants, while CUX1 and TLE3 peptides were underrepresented in the NT-IDR deletion compared to WT FOXA1 (Table 3).

2.3. FOXA1 recruits NuRD to Megatrans enhancers

We previously found that the FOXA1 NT-IDR was required for reduced H3K27 acetylation across Megatrans enhancers in the absence of estrogen, which was also associated with derepression of transcription (Figures 6C, 6D). Given that the NuRD complex was overrepresented in our proximity labeling mass spectrometry data, we wondered whether this reduction in H3K27ac was the result of interactions between NuRD and FOXA1. We therefore knocked down GATAD2A, CHD4, and HDAC2 to assessed their participation at the Megatrans enhancers regulating TFF1, NRIP1, and FOXC1 expression using the listed shRNAs (Table 4). While GATAD2A was previously identified to share binding-sites with FOXA transcription factors (Meers et al., 2019), CHD4 and HDAC2 were the highest-confidence FOXA1-associated NuRD complex enzymatic effectors as they exhibited the most biotinylation in both Turbo-ID and APEX2 proximity ligation assays. While knockdown of GATAD2A caused a significant but small increase in transcription of TFF1 mRNA compared to the non-targeting shRNA in both the ICI and estrogen-treated conditions, there were decreases in transcription of NRIP1 mRNA and FOXC1 enhancer RNA in the ICI treated condition. Knockdown of CHD4 and HDAC2, however, showed clearer trends in their apparent ability to repress transcription (Figures 6B, 6C).

The knockdown of CHD4 caused several fold increases in transcription of TFF1 and FOXC1 mRNA in the ICI and estrogen-treated conditions as well as significant but less pronounced increases in NRIP1 mRNA transcription in the estrogen-treated condition when compared to the shRNA control. Similarly, CHD4 knockdown also caused increases in the transcription of TFF1, NRIP1, and FOXC1 enhancer RNA in the estrogen-treated condition, though there was little effect on transcription in the ICI-treated condition. HDAC2 knockdown showed similar trends to CHD4 knockdown for enhancer transcription, where there was little change in the ICI-treated condition, but significant increases in transcription in the estrogen-treated condition. The effect of HDAC2 knockdown differed from that of the CHD4 knockdown on mRNA transcription, where there was no change in the ICI-treated condition compared to the

control shRNA, but increases in FOXC1 mRNA and decreases in TFF1 mRNA transcription in the estrogen-treated condition (Figures 6B, 6C). Since knockdown of CHD4 and HDAC2 had the most pronounced effects on transcription at TFF1, NRIP1, and FOXC1 enhancers, we wanted to see whether FOXA1 could recruit these factors to these enhancers in the FOXA1-binding sites deletion mutants.

Using the FOXA1 binding-site knock out cell lines, we were able to see whether CHD4 and HDAC2 could bind in the absence of FOXA1 via ChIP-qPCR. HDAC2 ChIP revealed that in wild-type cells, HDAC2 is dismissed from enhancers in the presence of estrogen by a factor of about 50% when compared to ICI-treated cells. In the TFF1 and NRIP1 FOXA1 binding-site deletion lines, this trend was reversed such that HDAC2 was recruited upon estrogen treatment. Furthermore, overall HDAC2 binding levels were severely reduced at NRIP1 and FOXC1 enhancers in the FOXA1 binding site deletion mutant compared to WT cells, indicating that FOXA1 plays a key role in recruiting HDAC2 in an ER-α-independent manner. Conversely, CHD4 ChIP revealed that it is highly recruited to TFF1, NRIP1, and FOXC1 enhancers in the presence of estrogen in WT cells. Removal of the FOXA1 binding-site, however, greatly decreased this estrogen-dependent recruitment of CHD4.

2.4. Discussion

The recruitment of HDAC2 and CHD4 by FOXA1 appears to play a role in restricting the estrogen-dependent transcriptional response at TFF1, NRIP1, and FOXC1 enhancers despite the requirement of FOXA1 for robust estrogen-dependent transcription at Megatrans enhancers. This is likely a result of dual roles by FOXA1 as both an activator and repressor of transcription, which appears to be dependent on the function of estrogen-bound ER-α. Additional experiments involving the knockdown of endogenous FOXA1 and rescue with transgenic WT FOXA1 and N-terminal mutants will show whether the FOXA1 N-terminus is required for the recruitment or functions of HDAC2 and CHD4.

Regarding to the mass spectrometry results, it is unclear how differences in proximity labeling methodology reflect the distribution of peptides identified. While Turbo-BirA biotinylates lysines via conversion of biotin and ATP to biotinoyl-AMP, APEX2 converts biotin-phenol to a radical intermediate that reacts with several different amino acids within the radius of a few nanometers (Hung et al., 2016). Turbo-BirA also biotinylates proteins in the absence of excess biotin, which leads to greater background as TurboBirA-tagged FOXA1 is enzymatically active after it is translated and while it is shuttled to the nucleus. In the APEX2-mediated proximity labeling, cells expressing APEX2-tagged FOXA1 are treated with estrogen during the labeling step, so it is possible that identified peptides more closely reflect FOXA1's function within the Megatrans complex and as it engages estrogen-regulated enhancers.

In both proximity ligation assays, ARID1A is less associated with the NT-IDR, which could contribute to the decrease in gene transcription in the NT-IDR deletion compared to WT and the NT-SLIM deletion (Figure 3C, Tables 2, 3). On the other hand, if FOXA1 is associating with NuRD complex factors such as CHD4 in the Megatrans complex, it is possible that both the FOXA1 NT-SLIM and NT-IDR are important for this interaction as indicated by the APEX2-mediated labeling (Table 2). Another possibility is that the FOXA1 NT-IDR is important for the function, but not the recruitment, of NuRD complex factors, which might explain why we see increased ligation of NuRD complex factors by the TurboBirA-tagged NT-IDR deletion mutant in the presence of estrogen while there is increased H3K27 acetylation in ICI-treated cells expressing the NT-IDR mutant rescue when compared to the WT FOXA1 rescue (Figure 5C). Further work is required to elucidate the function of the FOXA1 N-terminus in its association with the NuRD complex.

2.5. Acknowledgements

I would like to acknowledge Dr. Lu Yang for his work constructing the TFF1, NRIP1 and FOXC1 enhancer mutant cell lines described in this chapter.

Chapter 3: FOXA1 plays both activating and repressive roles in transcription

3.1. A model for FOXA1-dependent regulation of Megatrans enhancers

The research findings outlined in this dissertation uncover a complex role for FOXA1 in estrogen-regulated transcription. On one hand, FOXA1 is a determinant of AP2-γ binding to enhancers, which is an important factor that regulates transcription on Megatrans enhancers (Liu et al., 2014; Figures 2B, 4C, 4D). And yet, FOXA1 can also recruit the transcriptional repressors HDAC2 and CHD4 to restrain transcription on enhancers and promoters (Figures 6B, 6C, 6D). Furthermore, in the absence of the FOXA1 N-terminal IDR, there is increased H3K27ac on a subset of enhancers in cells that are not treated with estrogen and without the estrogen receptor compared to wild-type cells (Figure 5C).

Our model for FOXA1's function at FOXA1-activated Megatrans enhancers is as follows: In the complete absence of FOXA1, GATA3 and ER-α are still able to bind Megatrans enhancers, though their ability to activate transcription is reduced (Figure 7A). In our experiments when endogenous FOXA1 is knocked down but the NT-IDR deletion mutant is expressed to wild-type levels, there is an inability to recruit HDAC2 and AP2-γ, and across FOXA1-activated Megatrans enhancers there is a reduction of estrogen-dependent transcription (Figures 3E, Figure 7B). Finally, in wild-type MCF7 cells, FOXA1 binds HDAC2 to reduce H3K27ac in the absence of estrogen, which adds an extra layer of repression. In the presence of estrogen, ER-α binds enhancers and recruits the Megatrans complex, which in conjunction with FOXA1's N-terminus recruits AP2-γ to aid in estrogen-dependent transcription. Both HDAC2 and CHD4 appear to restrain transcription to prevent overexpression of RNAs at enhancers and promoters (Figure 7C). All together, this model describes FOXA1's ability to act as a switch from transcriptional repression to activation in response to interactions with estrogen-bound ER-α.

3.2. Concluding remarks

Although this dissertation primarily focuses on the recruitment of AP2-γ and the NuRD complex as determinants of FOXA1-dependent transcriptional activity, there are likely other coactivators and corepressors that cooperate with FOXA1 to regulate transcription. For instance, our proximity ligation assays consistently reveal potential coregulators from the BAF complex, such as ARID1A, as well as other transcription factors including CUX1, HCFC1, NFIX, C/EBPβ, and AP-1 family members. Clusters of genes that are activated or repressed upon FOXA1 knockdown could be dependent on the activities of these coregulators as well as the well-established C-terminal interactor TLE3.

The mechanism by which FOXA1 cooperates with ER-α to recruit CHD4 and AP2-γ as well as dismiss HDAC2 remains unknown, however clues may be identified by examining which proteins become recruited to enhancers upon engagement by estrogen-bound ER-α. For instance, the NuRD component MTA1 was found to interact directly with ER-α via its LXXLL motif to repress transcription (Mazumdar et al., 2001; Kumar et al., 2002), although it was found that acetylation of MTA1's lysine 626 caused it to switch to an activator in a context-dependent manner (Gururaj et al., 2006).

Given FOXA1's importance in regulating expression in breast cancer, learning more about how it regulates genes that are involved in cancer progression could allow for more opportunities for treatment. More work is therefore required to fully understand its role in transcriptional regulation.

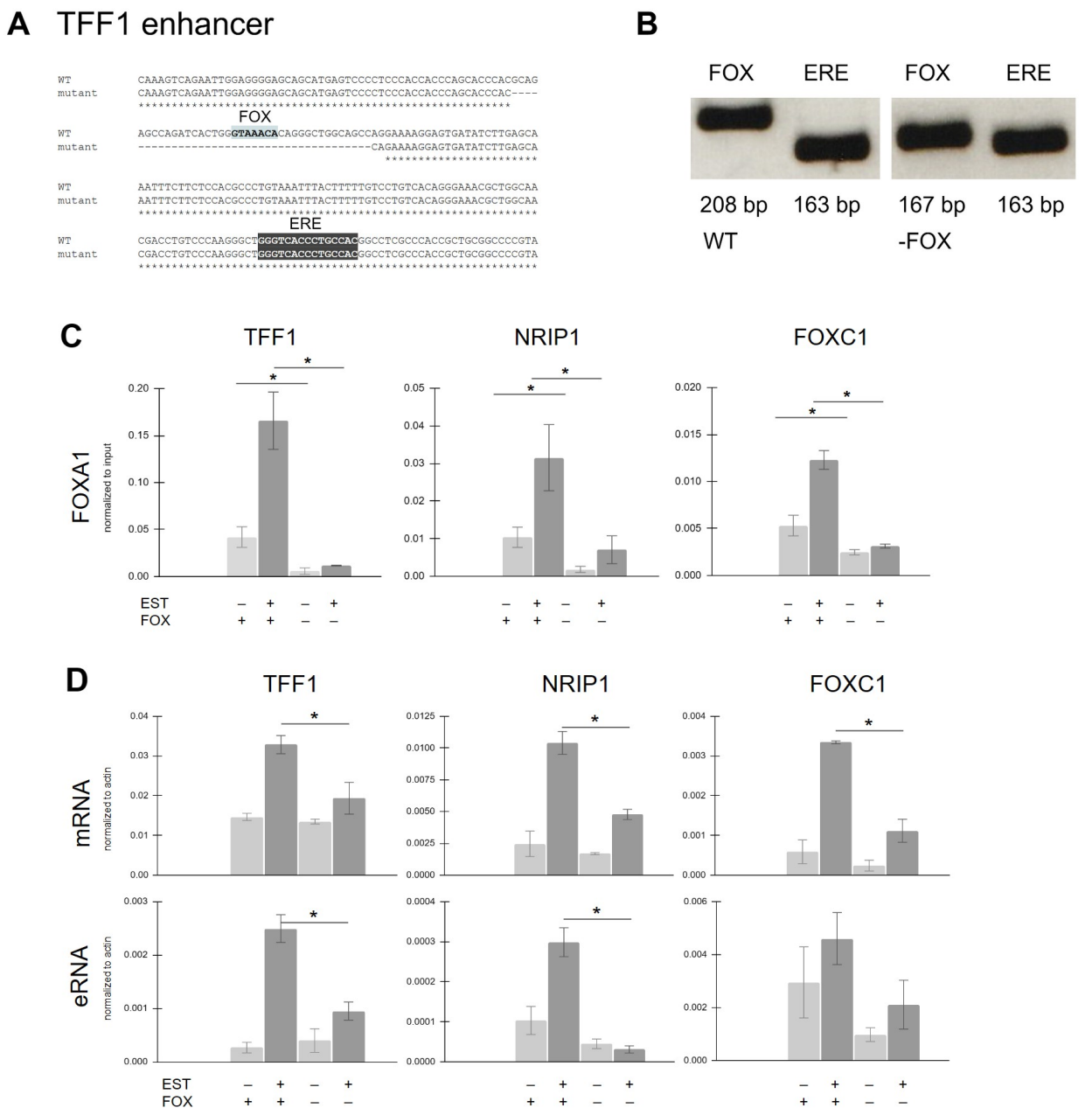
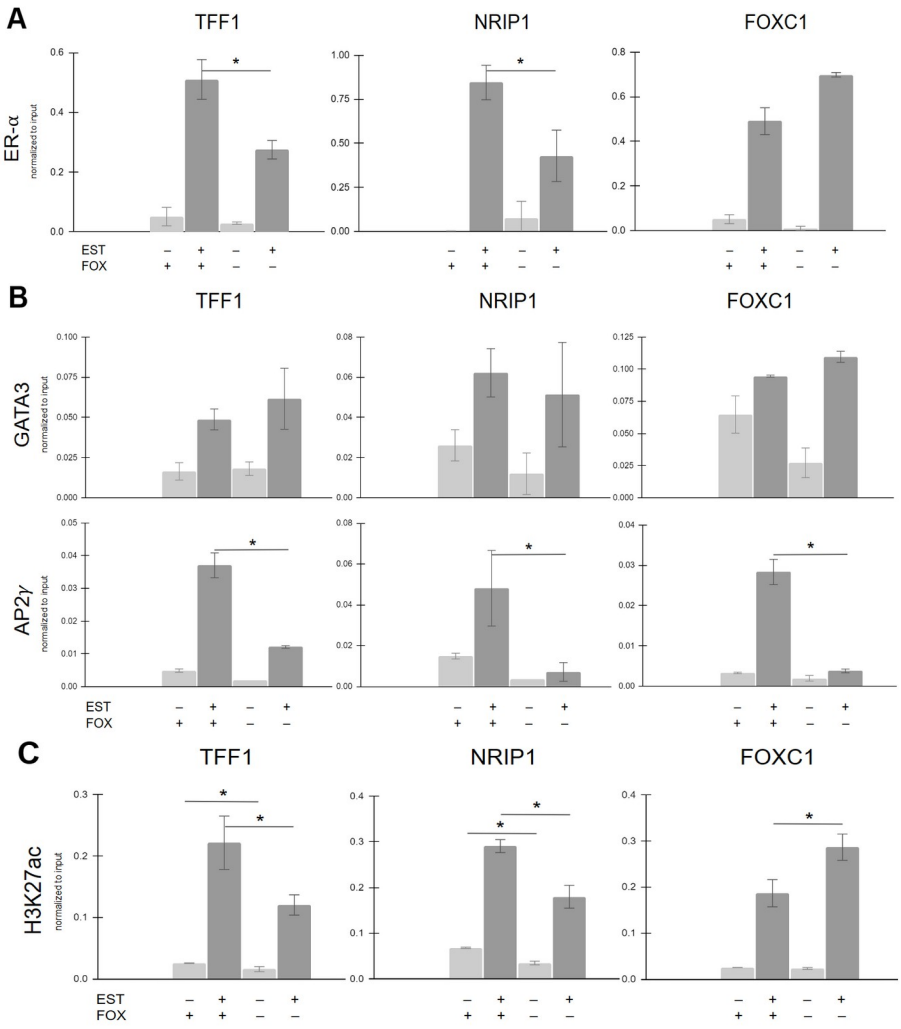


Figure 1: CRISPR-Cas9 FOX deletion mutants affect transcription. A) Genomic DNA sequence of TFF1 enhancer encompassing FOXA1 and ER-α binding sites, showing deletion of FOXA1 binding motif. B) PCR products containing WT TFF1 FOXA1 and ER-α binding site compared to binding-site deletion mutant. C) ChIP-qPCR for FOXA1 binding to TFF1, NRIP1 and FOXC1 enhancers. D) TFF1, NRIP1, and FOXC1 mRNA and eRNA expression in WT cells (FOX +) and FOXA1 binding-site mutants (FOX -), +/- 100 nM estradiol, measured by RT-qPCR. Error bars are standard deviation; one-tailed t-test for unequal variance used to calculate significance; *p<0.05 (n_{WT}=3, n_{mutant}=2).



D

Motif	Frequency	Log(p-value)	Graphic
FOX	71.9%	-1.10e^{3}	
ERE	23.4%	-5.23e^{2}	
AP-2	27.9%	-2.00e^{2}	
GRHL2	18.3%	-2.02e^{2}	
ERE (half-site)	72.0%	-1.16e^{2}	
AP-1	22.2%	-1.55e^{2}	
GATA	40.0%	-1.06e^{2}	

Figure 2: FOXA1 binding-site deletion affects recruitment of Megatrans and H3K27ac. ChIP-qPCR experiments for WT (FOX +) and FOXA1 binding-site deletion mutants (FOX-), +/- estradiol, for: A) ER-α; B) GATA3 and AP2-γ; C) H3K27ac. For all bar graphs: Error bars are standard deviation; one-tailed t-test for unequal variance used to calculate significance; *p<0.05 (n_{WT}=3, n_{mutant}=2). D) Motif frequency (HOMER, +/- 150 bp) at FOXA1 peaks (SRX3923513, peak calling by ChIP-Atlas (q<0.05)) at Megatrans enhancers (Glont et al., 2019; Zou et al., 2021; Liu et al., 2014).

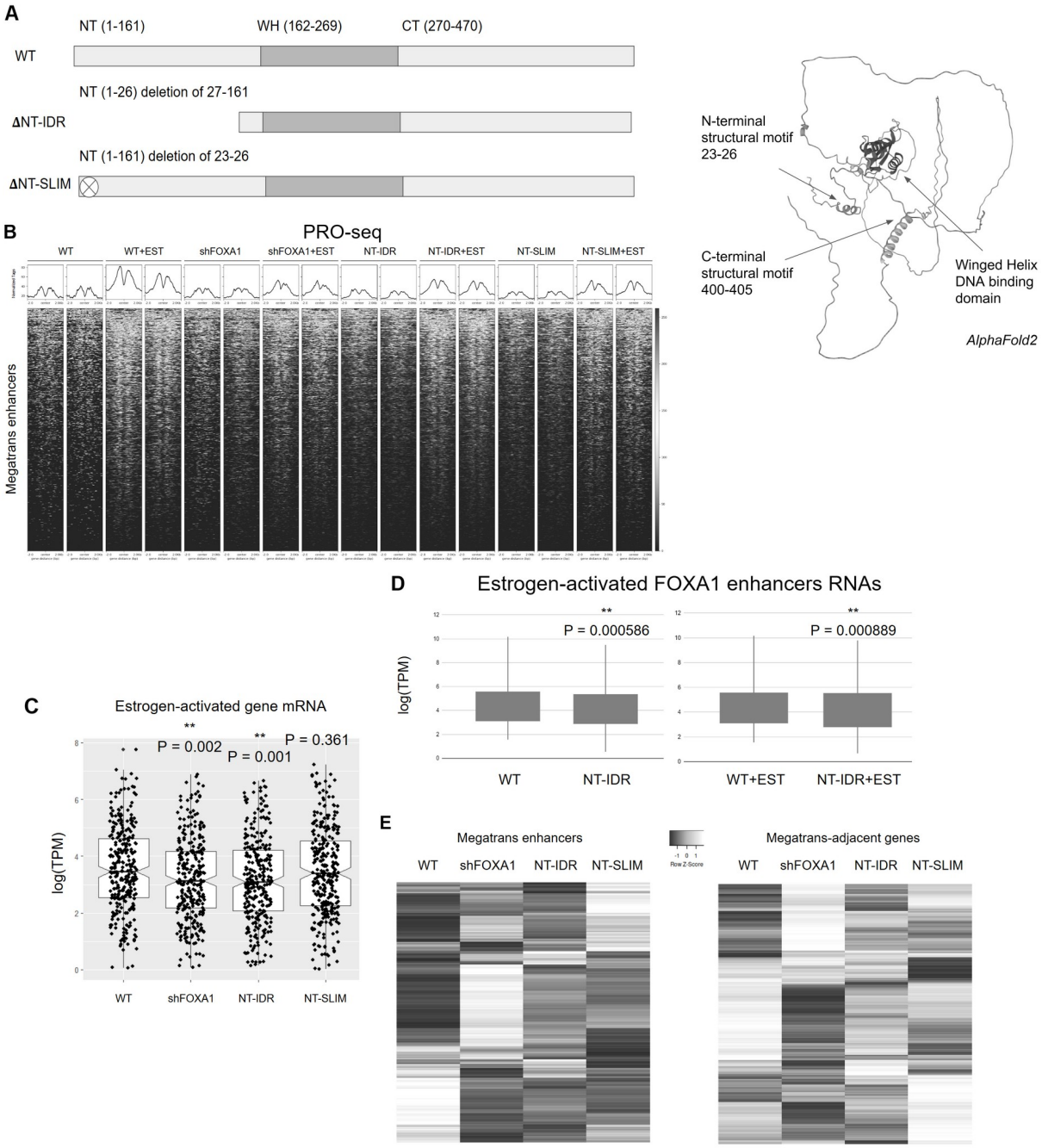


Figure 3: The FOXA1 N-terminus regulates estrogen-dependent transcription. A) left: FOXA1 N-terminal deletions; right: Alphafold2 predicted structure of FOXA1. B) Tag-density plot (deepTools) for WT and N-terminal mutation PRO-seqs, showing average tag-density line graph (top), +/- estradiol. C) Comparison of expression values for >2-fold significantly upregulated genes ($p<0.05$), n=321, + estradiol. D) eRNA expression within +/- 2 kb window of FOXA1 peak at Megatrans enhancers, +/- estradiol. E) Heatmaps comparing Megatrans enhancer eRNA (left) and adjacent mRNA (right) expression, + estradiol. Pearson's correlation is used to calculate distance for hierarchical clustering. All expression values are TMM normalized (EdgeR) for between-sample comparisons. Megatrans enhancer subset used for these analyses contains one FOXA1 peaks (n=1021). One-tailed t-test for unequal variance used for gene selection (C) and comparison between samples (C, D).

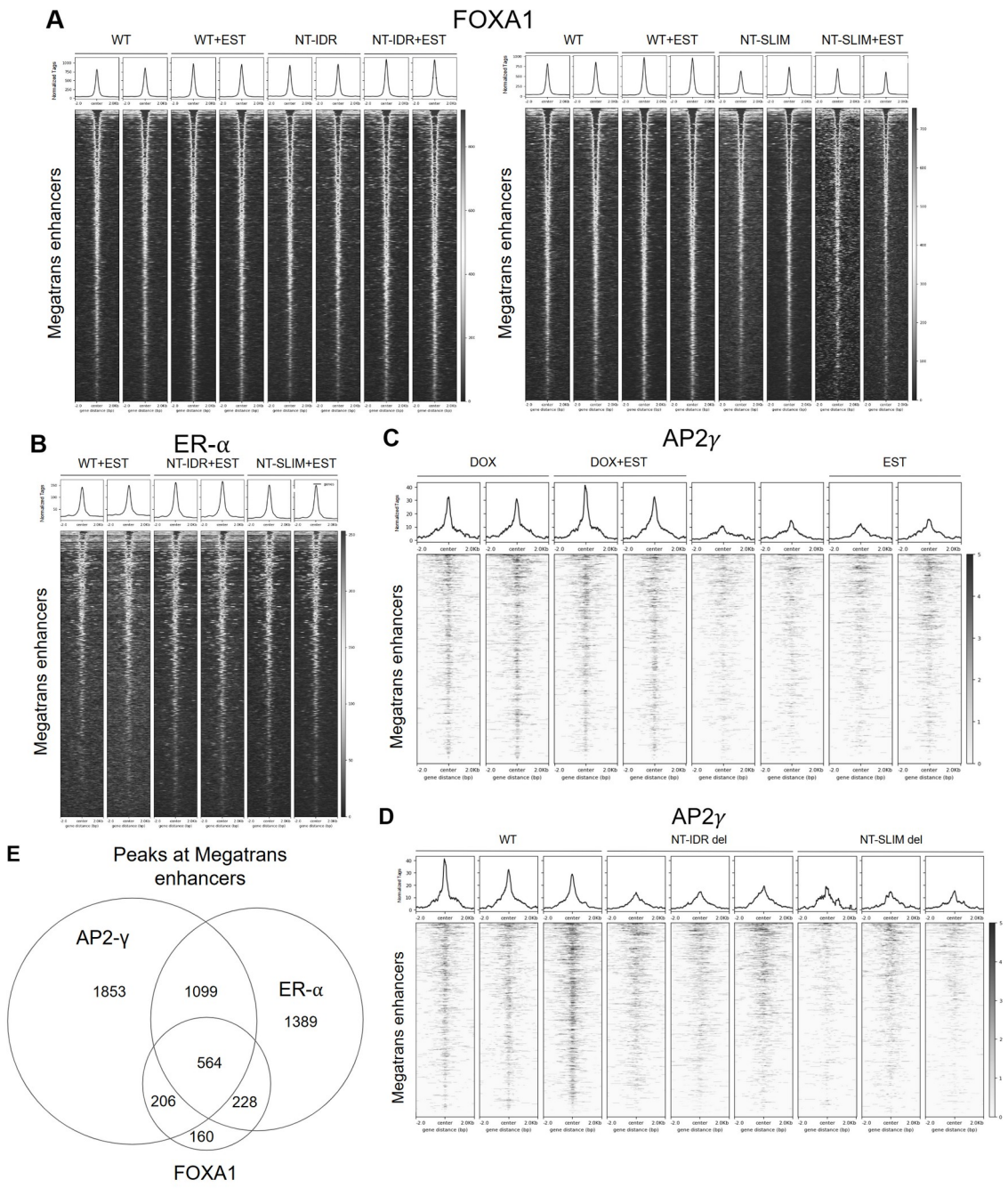


Figure 4: The FOXA1 N-terminus recruits AP2-γ independent of estrogen receptor. Tag density plots shown for WT and N-terminal mutant rescues for: A) FOXA1 (anti-HA) ChIP-seq, +/- estradiol, centered on FOXA1 peaks; B) ER-α ChIP-seq, + estradiol; C) AP2-γ CUT&TAG for WT rescue (DOX) and shFOXA1, +/- estradiol, centered on ER-α peaks; D) AP2-γ CUT&TAG for WT and N-terminal mutants, + estradiol. centered on FOXA1 peaks. ChIP-seq values are TMM normalized (EdgeR) for between sample comparisons. E) Overlap between peaks within Megatrans enhancers (+/- 2 kb) in WT cells as identified using HOMER for FOXA1 and ER-α ChIP-seqs and SEACR for AP2-γ CUT&TAG; overlapping peaks are determined to be within a 300 bp window (Analysis performed using mergePeaks from Homer).

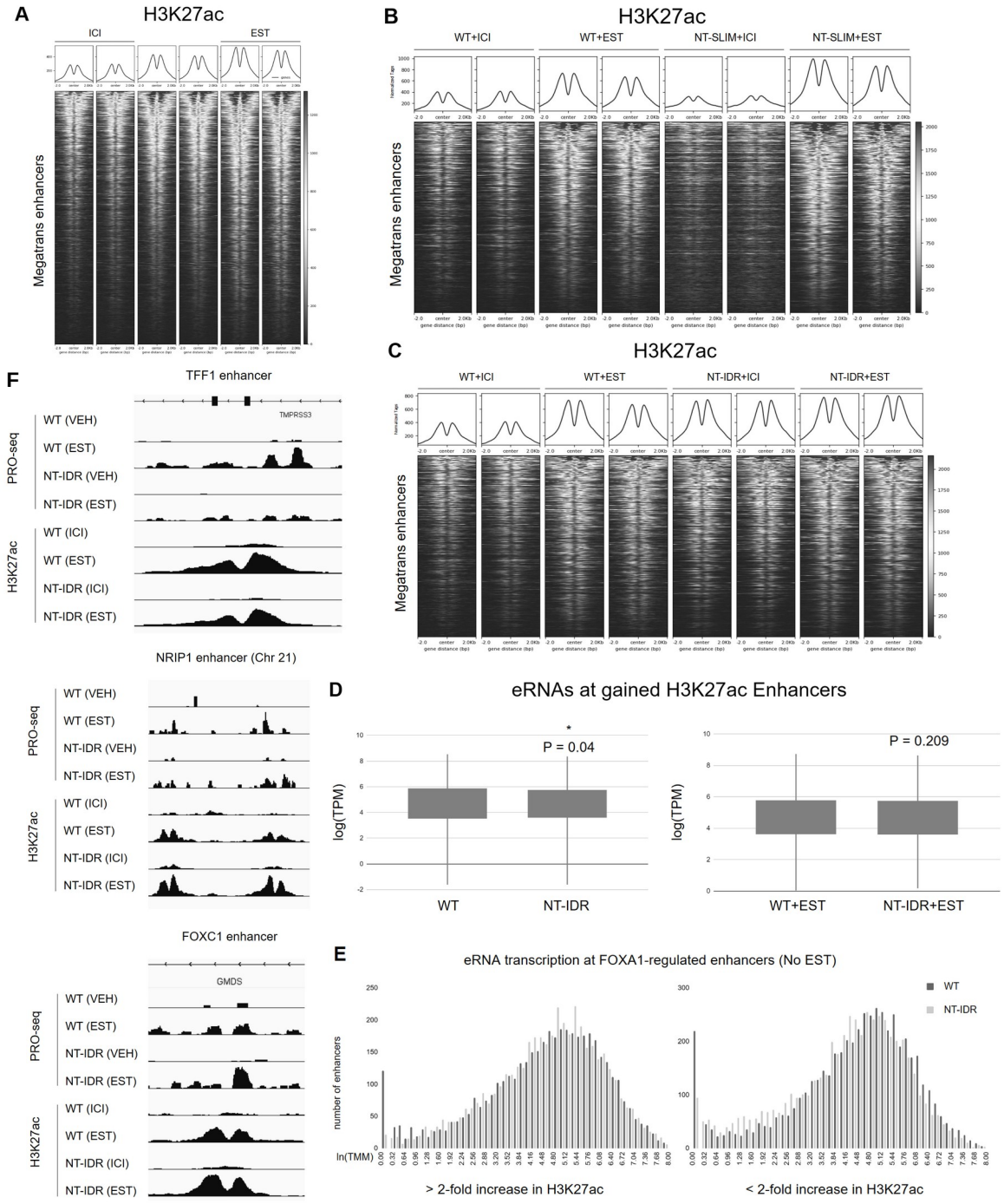


Figure 5: Effect of FOXA1 N-terminal deletions on H3K27 acetylation. H3K27ac ChIP-seq tag density plots at Megatrans enhancers, centered at FOXA1 peaks, for: A) WT FOXA1 rescue, ICI and +/- estradiol; ICI/estradiol treated WT FOXA1 rescue vs B) NT-SLIM deletion rescue and C) NT-IDR deletion rescue. D) eRNA transcription at FOXA1 enhancers with >2-fold increases in H3K27ac in the NT-IDR deletion, +/- 2 kb. Enhancers defined as containing FOXA1 peaks, H3K27ac peaks, and TSSs identified by findPeaks from HOMER in WT and NT-IDR datasets, +/- 1 kb, +/- estradiol. E) Histogram showing distribution of enhancers by transcription at >2-fold increases in H3K27ac (n=4497) and <2-fold increases (n=5397), - estradiol.

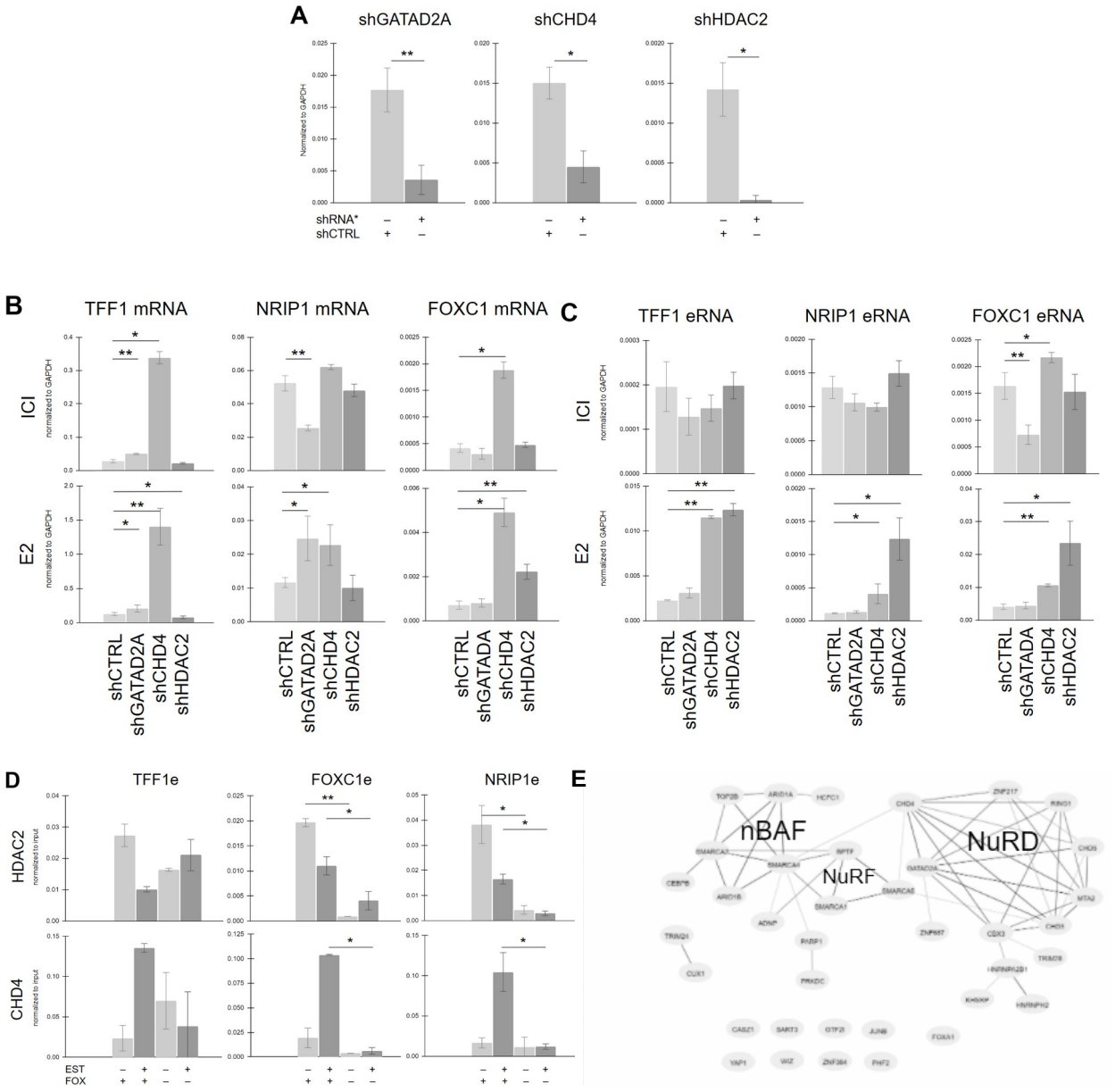


Figure 6: Contribution of the NuRD to estrogen-dependent transcription. RT-qPCR comparing negative control shRNA to shRNA targeting GATAD2A, CHD4, and HDAC2: A) Knockdown efficiency for respective mRNAs (n=4); B) TFF1, NRIP1, and FOXC1 mRNA expression for each knockdown, ICI/estradiol treatment (n=3); C) TFF1, NRIP1, and FOXC1 enhancer RNA expression profiles for each knockdown, ICI/estradiol treatment (n=3). D) ChIP-qPCR profiles at the TFF1, FOXC1, and NRIP1 enhancers comparing WT cells to FOXA1 binding-site deletion mutants (n_{WT}=3, n_{mutant}=2). E) Interactions between the top 50 biotin-labeled transcription-related proteins in the Turbo-ID assay on WT FOXA1 (Figure S4A); visualization via Cytoscape; interactions from StringDB. For all bar graphs: Error bars are standard deviation; one-tailed t-test for unequal variance used to calculate significance; *p<0.05, **p<0.005.

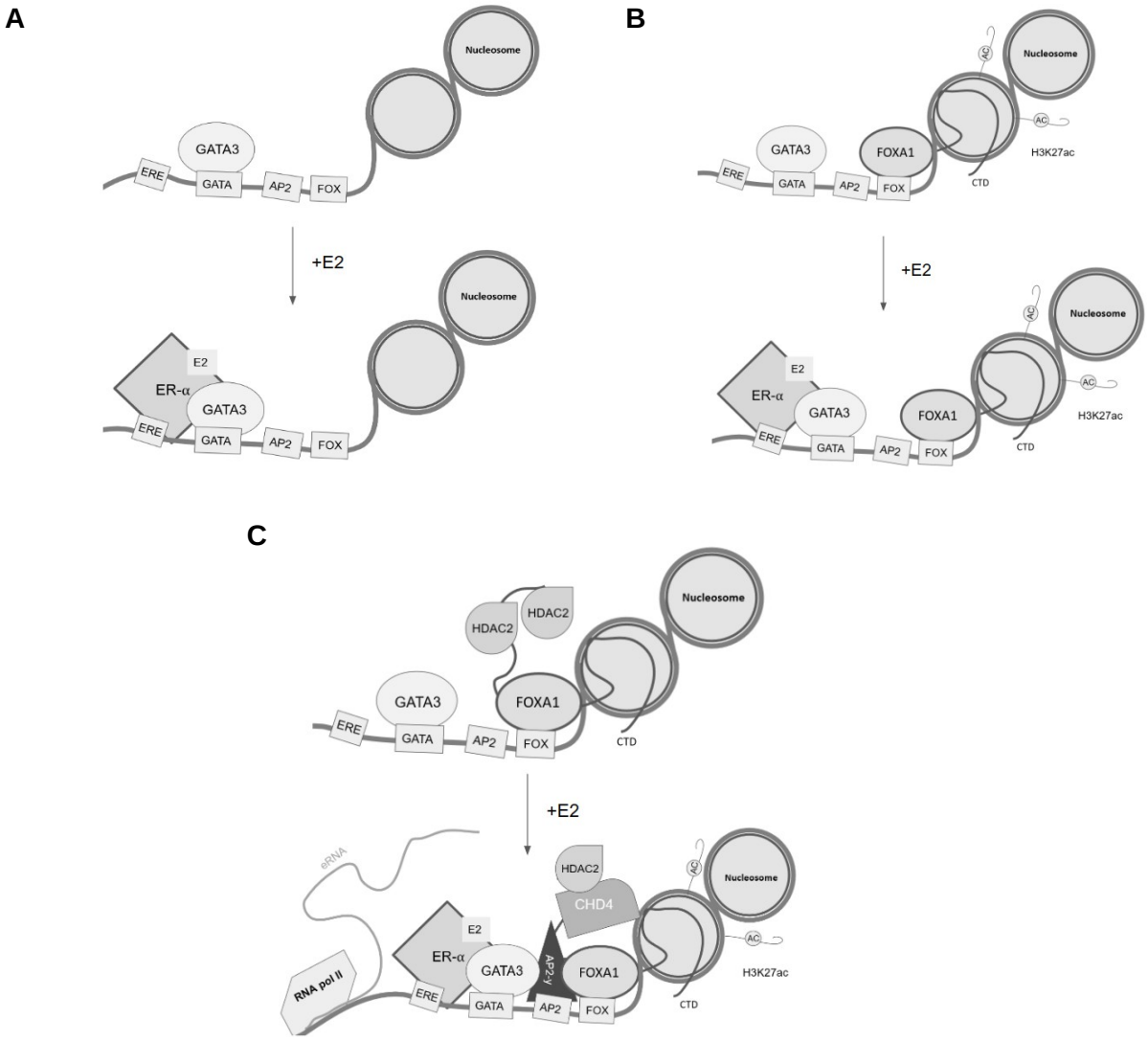


Figure 7: Model for robust FOXA1 and estrogen-dependent transcription. Estrogen-activated transcription in: A) the absence of FOXA1; B) the FOXA1 NT-IDR deletion; C) WT cells.

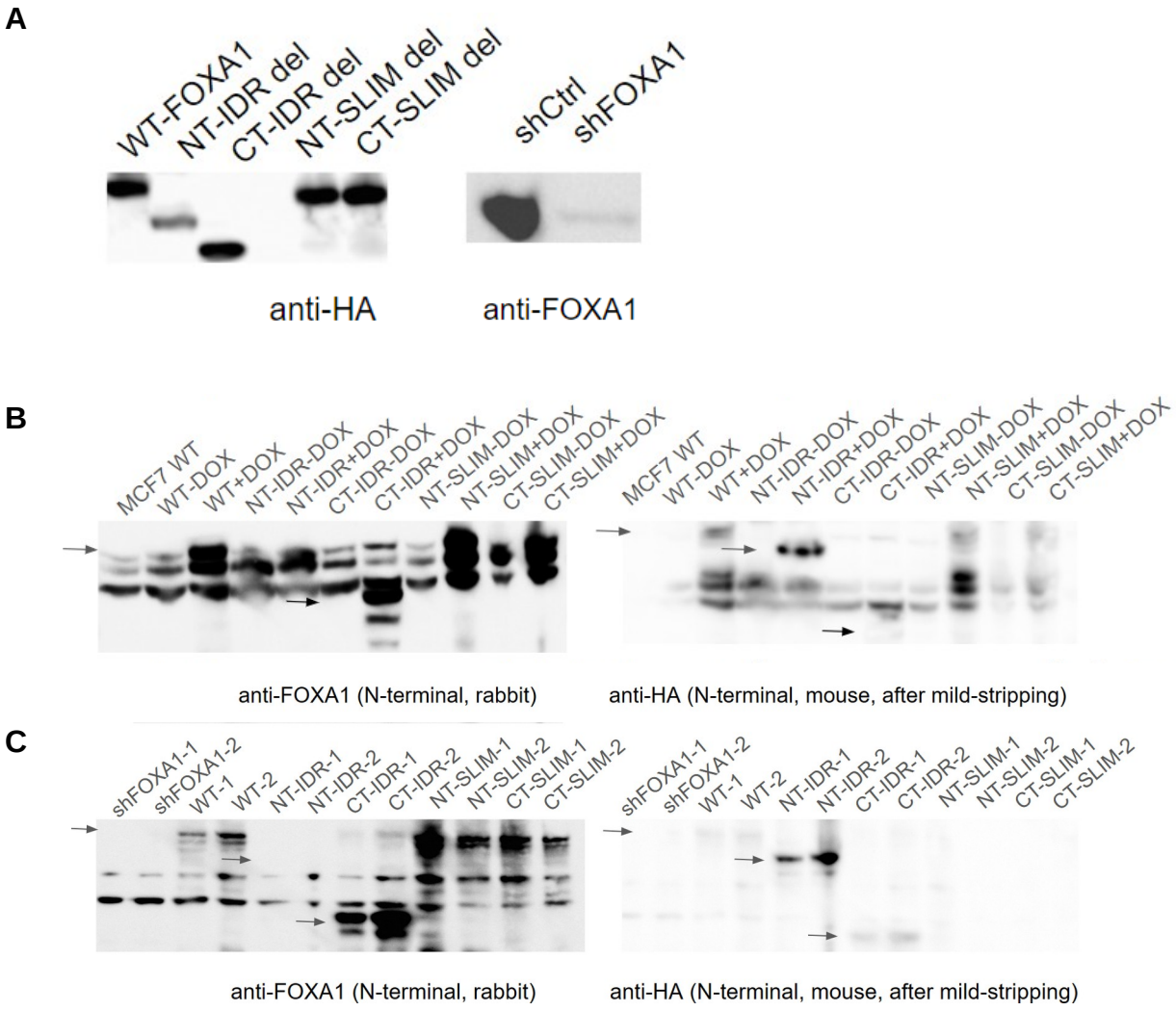


Figure S1: FOXA1 knockdown and rescue. Western blots showing: A) Left: Expression of FOXA1 as well as N-terminal and C-terminal mutants following transfection of 293T cells with pCDH plasmid, anti-HA antibody. Right: shFOXA1 knockdown, anti-FOXA1 antibody. B) Doxycycline-induced overexpression of WT FOXA1 and mutants (1 ug/mL). Arrows indicate expected bands. Left: N-terminal anti-FOXA1 antibody; Right: anti-HA antibody. C) Knockdown of endogenous FOXA1 and doxycycline-induced rescue with indicated mutants (0.5 ug/mL used NT-IDR and CT-IDR deletions and 1 ug/mL for remaining mutants). Aliquots are taken from nuclei isolated for PRO-seq experiments. Antibodies used as indicated. Western blots show mutants not included in the body of this dissertation: C-terminal SLIM deletion corresponding to TLE3-interacting EH1 motif (400-404), and C-terminal IDR deletion (271-470).

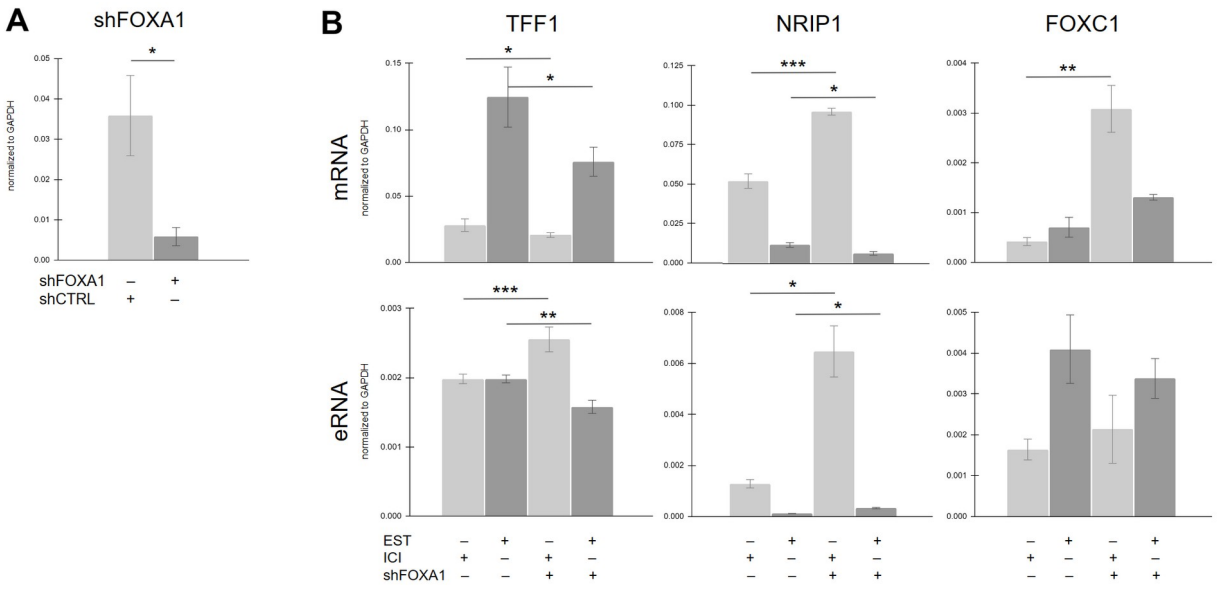


Figure S2: Effect of FOXA1 shRNA knockdown on estrogen-dependent transcription. RT-qPCR comparing negative control shRNA to shRNA targeting FOXA1: A) Knockdown efficiency for shFOXA1 (n=4); B) TFF1, NRIP1, and FOXC1 mRNA and enhancer RNA expression; ICI/estradiol treatment. Error bars are standard deviation; one-tailed t-test for unequal variance used to calculate significance; *p<0.05, **p<0.005, ***p<0.0005 (n_{shCTRL}=3, $n_{shFOXA1}$=2).

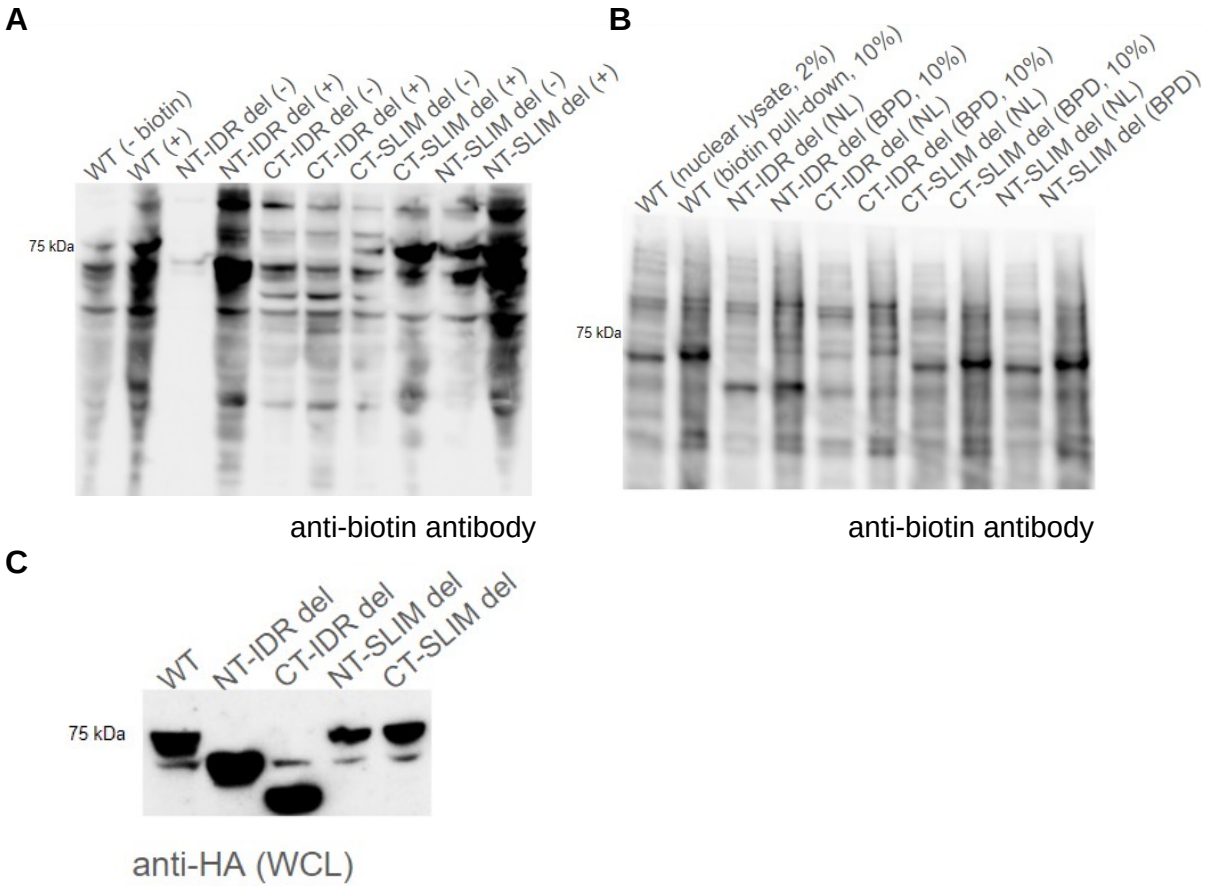


Figure S3: FOXA1 and mutant proximity labeling efficiency via Turbo-ID. Western blots showing: A) Induction of biotinylation for each mutant following addition of excess biotin; anti-biotin antibody. B) Input (nuclear lysate) and biotin pull-down (BPD) for each mutant; anti-biotin antibody. C) Doxycycline-inducible overexpression of TurboBirA-tagged WT FOXA1 and mutants in whole-cell lysate (1 ug/mL); anti-HA antibody. Western blots show mutants not included in the body of this dissertation: C-terminal SLIM deletion corresponding to TLE3-interacting EH1 motif (400-404), and C-terminal IDR deletion (271-470).

A

Protein name	-log(P)	n (P<0.05)
FOXA1	245.62	14
ARID1A	222.62	21
CHD4	217.98	17
JUNB	216.23	11
AHNAK	186.96	12
ADNP	184.65	16
TRIM24	174.69	8
SMARCA4	174.24	10
TRIM33	170.76	8
CUX1	170.65	9
CASZ1	165.59	6
HCFC1	163.5	8
SMARCA5	158.86	10
ZNF384	152.81	6
WIZ	151.17	8
PARP1	151.12	7
GTF2I	150.07	7
CBX3	149.18	5
TOP2B	148.74	5
ZNF217	141.31	5

Protein name	-log(P)	n (P<0.05)
BPTF	138.95	6
HNRNPH2	135.73	5
MTA2	127.16	9
CHD5	125.49	4
CEBPB	124.55	3
HNRNPA2B1	117.01	7
RING1	109.45	3
SART3	108.37	4
PSPC1	108.16	2
GATAD2A	106.24	7
ARID1B	104.85	4
SMARCA2	101.51	3
ZNF687	101.15	1
SMARCA1	100.43	3
KHSRP	98.66	4
PRKDC	98.4	5
PHF2	98.28	4
YAP1	98	4
CHD3	95.57	3
TRIM28	93.3	4

B

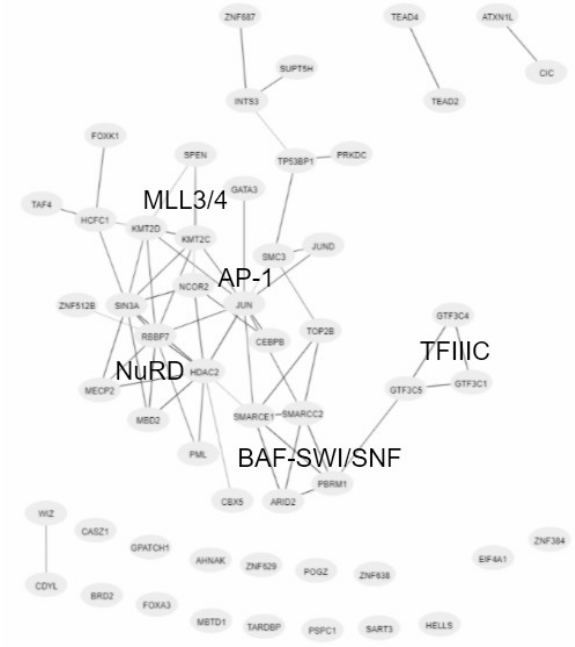


Figure S4: Turbo-ID proximity labeling results. A) Left: List contains top 50 transcription-related proteins by peptide-scoring confidence (PEAKS software suite, bioinfor.org). B) Left: Transcription-related proteins identified in estradiol-treated WT cells but not in untreated cells. Visualization via Cytoscape; interactions from StringDB.

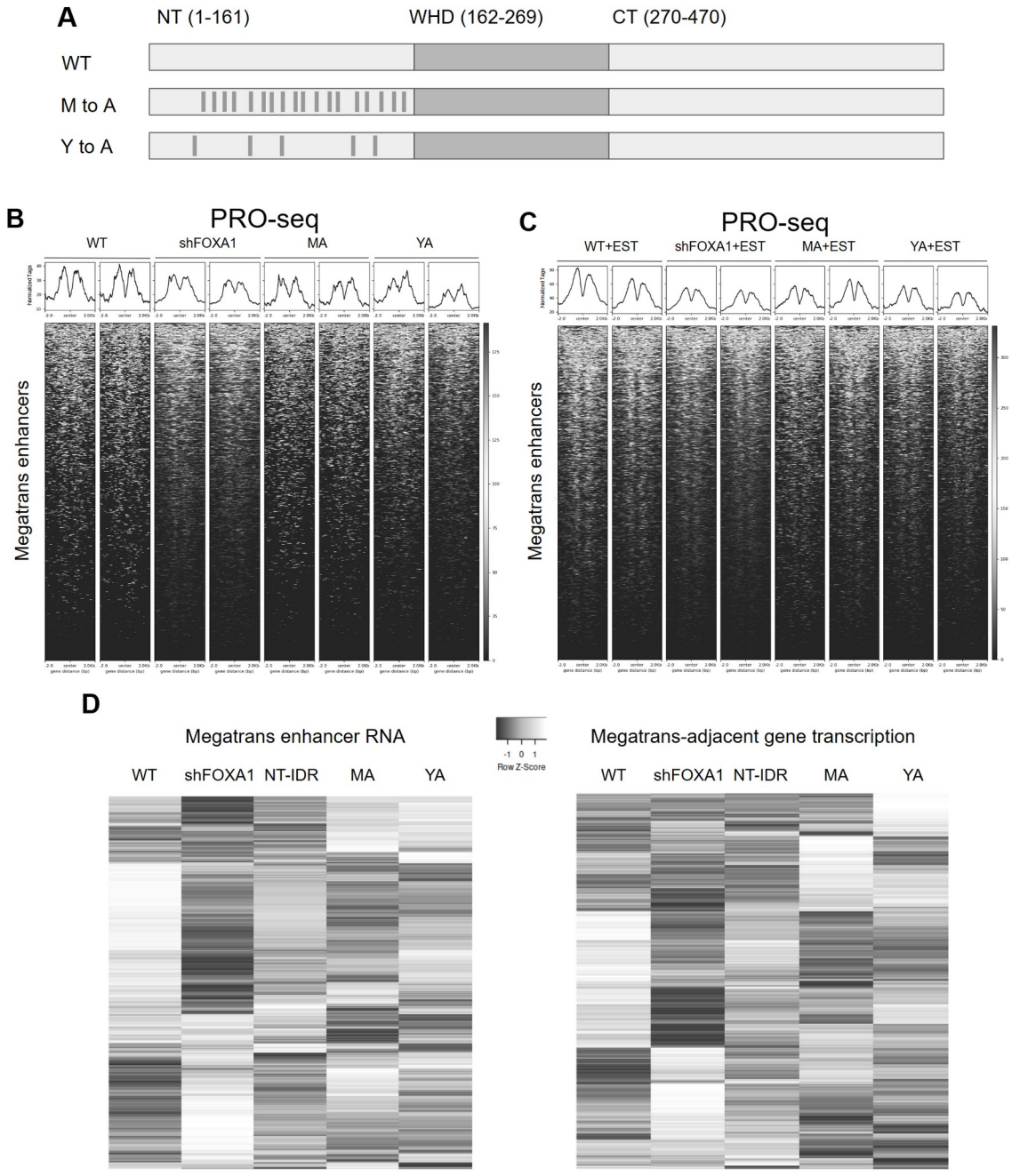


Figure S5: FOXA1 N-terminal IDR amino acid mutations. A) Schematic showing the locations of mutated amino acids in the methionine to alanine and tyrosine to alanine mutants. (B-C) Tag-density plots (deepTools) for WT and N-terminal amino acid mutation PRO-seqs, showing average tag-density line graph (top) in the absence (B) and presence of estrogen (C). D) Heatmaps comparing Megatrans enhancer eRNA (left) and adjacent mRNA (right) expression, + estradiol. Pearson's correlation is used to calculate distance for hierarchical clustering. eRNA expression measured within +/- 2 kb window of FOXA1 peak at Megatrans enhancers. All expression values are TMM normalized (EdgeR) for between-sample comparisons. Megatrans enhancer subset used for these analyses contains one FOXA1 peaks (n=1021).

Table 1: TURBO-ID shows differential association with NuRD factors between mutants

Gene Symbol	WT	WT+EST	NT-IDR+EST	NT-SLIM+EST
FOXA1	11	14	9	11
HDAC2	0	2	1	0
MTA1	2	4	7	1
MTA2	4	9	11	6
RBBP7	3	3	3	0
GATAD2A	0	7	11	0
GATAD2B	1	0	5	2
MBD2	0	2	3	0
CHD3	0	3	7	2
CHD4	10	17	26	8
CHD5	2	4	6	2

Table 2: TURBO-ID shows differential association with SWI/SNF factors between mutants

Gene Symbol	WT	WT+EST	NT-IDR+EST	NT-SLIM+EST
FOXA1	11	14	9	11
ARID1A	6	21	14	11
ARID1B	0	4	0	0
SMARCA1	1	3	5	2
SMARCA2	2	3	2	2
SMARCA4	7	10	10	7
SMARCA5	3	10	14	6
SMARCC2	0	2	2	2

Table 3: APEX2-mediated proximity labeling mass spectrometry results

Gene Symbol	WT+EST	NT-IDR+EST	NT-SLIM+EST
FOXA1	38	38	31
SPT5H	24	23	11
ARID1A	23	7	16
TRIM33	13	5	10
SMARCA4	11	2	6
NCOR2	10	3	7
SUPT16H	9	4	3
CHD4	7	1	2
CHD6	5	2	0
CHD5	5	0	0
AP2-y	7	6	5
AP2-α	8	6	5
AP2-ε	6	3	0
TLE3	7	1	10
GATA3	5	9	3
GRHL2	8	5	1
C/EBPβ	3	0	0
HCFC1	8	2	2
CUX1	7	0	4
NFIX	4	0	1
NFIB	3	0	1
SMARCC2	3	0	0
SMCHD1	3	0	6
SMCA1	3	3	0
SMARCE1	4	0	1
HDAC2	3	2	0
DNMT1	4	0	0
KDM2A	2	0	0
MBD1	4	2	2
GATAD2B	5	3	6
GATAD2A	4	1	5
MTA2	5	2	5

Table 4: List of shRNAs used

Gene Symbol	shRNA sequence
FOXA1	GAGAGAAAAAATCAACAGC
GATAD2A	TGCGGCAGAGTCAAATACAAAa
CHD4	GCGGGAGTTCAGTACCAATAA
HDAC2	GCAAATACTATGCTGTCAATT

References

Adams EJ, Karthaus WR, Hoover E, Liu D, Gruet A, Zhang Z, Cho H, DiLoreto R, Chhangawala S, Liu Y, Watson PA, Davicioni E, Sboner A, Barbieri CE, Bose R, Leslie CS, Sawyers CL. FOXA1 mutations alter pioneering activity, differentiation and prostate cancer phenotypes. Nature. 2019 Jul;571(7765):408-412.

Allahyar A, Vermeulen C, Bouwman BAM, Krijger PHL, Verstegen MJAM, Geeven G, van Kranenburg M, Pieterse M, Straver R, Haarhuis JHI, Jalink K, Teunissen H, Renkens IJ, Kloosterman WP, Rowland BD, de Wit E, de Ridder J, de Laat W. Enhancer hubs and loop collisions identified from single-allele topologies. Nat Genet. 2018 Aug;50(8):1151-1160.

Ancey PB, Ecsedi S, Lambert MP, Talukdar FR, Cros MP, Glaise D, Narvaez DM, Chauvet V, Herceg Z, Corlu A, Hernandez-Vargas H. TET-Catalyzed 5-Hydroxymethylation Precedes HNF4A Promoter Choice during Differentiation of Bipotent Liver Progenitors. Stem Cell Reports. 2017 Jul 11;9(1):264-278.

Aurélien A. et al., Epigenetic switch involved in activation of pioneer factor FOXA1-dependent enhancers. Genome Res. 2011 Apr; 21(4): 555–565.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. High-resolution profiling of histone methylations in the human genome. Cell. 2007 May 18;129(4):823-37.

Beagrie RA, Pombo A. Cell cycle: Continuous chromatin changes. Nature. 2017 Jul 5;547(7661):34-35.

Bernardo, G., Bebek, G., Ginther, C. et al., FOXA1 represses the molecular phenotype of basal breast cancer cells. Oncogene. 2013 Mar 5. 32, 554–563

Bi, M., Zhang, Z., Jiang, YZ. et al., Enhancer reprogramming driven by high-order assemblies of transcription factors promotes phenotypic plasticity and breast cancer endocrine resistance. Nat Cell Biol. 2020 May 18. 22, 701–715.

Blackwood EM, Kadonaga JT. Going the distance: a current view of enhancer action. Science. 1998 Jul 3. 281(5373):60-3.

Boija A, Klein IA, Sabari BR, Dall'Agnese A, Coffey EL, Zamudio AV, Li CH, Shrinivas K, Manteiga JC, Hannett NM, Abraham BJ, Afeyan LK, Guo YE, Rimel JK, Fant CB, Schuijers J, Lee TI, Taatjes DJ, Young RA. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell. 2018 Dec 13;175(7):1842-1855.

Bossard P, Zaret K.S. GATA transcription factors as potentiators of gut endoderm differentiation. Development. 1998 Dec 1. 125: 4909-4917

Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, Furey TS, Crawford GE. High-resolution mapping and characterization of open chromatin across the genome. Cell. 2008 Jan 25;132(2):311-22.

Bragança J, Eloranta JJ, Bamforth SD, Ibbitt JC, Hurst HC, Bhattacharya S. Physical and functional interactions among AP-2 transcription factors, p300/CREB-binding protein, and CITED2. J Biol Chem. 2003 May 2;278(18):16021-9.

Branon TC, Bosch JA, Sanchez AD, Udeshi ND, Svinkina T, Carr SA, Feldman JL, Perrimon N, Ting AY. Efficient proximity labeling in living cells and organisms with TurboID. Nat Biotechnol. 2018 Oct;36(9):880-887.

Brangwynne CP, Mitchison TJ, Hyman AA. Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. Proc Natl Acad Sci U S A. 2011 Mar 15;108(11):4334-9.

Cahan P, Li H, Morris SA, Lummertz da Rocha E, Daley GQ, Collins JJ. 2014. CellNet: network biology applied to stem cell engineering. Cell 158: 903–915.

Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why? Mol Cell. 2013 Mar 7;49(5):825-37.

Caravaca JM, Donahue G, Becker JS, He X, Vinson C, Zaret KS. Bookmarking by specific and nonspecific binding of FoxA1 pioneer factor to mitotic chromosomes. Genes Dev. 2013 Feb 1;27(3):251-60.

Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G, Sung WK, Clarke ND, Wei CL, Ng HH. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell. 2008 Jun 13;133(6):1106-17.

Cirillo LA, Zaret KS. Specific interactions of the wing domains of FOXA1 transcription factor with DNA. Journal of Molecular Biology. 2007 Feb;366(3):720-724.

Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. Mol Cell. 2002 Feb;9(2):279-89.

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A. 2010 Dec 14;107(50):21931-6.

Donaghey J, Thakurela S, Charlton J, Chen JS, Smith ZD, Gu H, Pop R, Clement K, Stamenova EK, Karnik R, Kelley DR, Gifford CA, Cacchiarelli D, Rinn JL, Gnirke A, Ziller MJ, Meissner A. Genetic determinants and epigenetic effects of pioneer-factor occupancy. Nat Genet. 2018 Feb;50(2):250-258.

Dosztányi Z, Mészáros B, Simon I. ANCHOR: web server for predicting protein binding regions in disordered proteins. Bioinformatics. 2009 Oct 15;25(20):2745-6.

Duncan IW. Transvection effects in Drosophila. Annu Rev Genet. 2002;36:521-56.

Fanucchi S, Shibayama Y, Burd S, Weinberg MS, Mhlanga MM. Chromosomal contact permits transcription between coregulated genes. Cell. 2013 Oct 24;155(3):606-20.

Fernandez Garcia M, Moore CD, Schulz KN, Alberto O, Donague G, Harrison MM, Zhu H, Zaret KS. Structural Features of Transcription Factors Associating with Nucleosome Binding. Mol Cell. 2019 Sep 5;75(5):921-932.e6.

Gévry N, Hardy S, Jacques PE, Laflamme L, Svotelis A, Robert F, Gaudreau L. Histone H2A.Z is essential for ER-α signaling. Genes Dev. 2009 Jul 1; 23(13):1522-33.

Giaever GN, Wang JC. Supercoiling of intracellular DNA can occur in eukaryotic cells. Cell. 1988 Dec 2;55(5):849-56.

Glont SE, Chernukhin I, Carroll JS. Comprehensive Genomic Analysis Reveals that the Pioneering Function of FOXA1 Is Independent of Hormonal Signaling. Cell Rep. 2019 Mar 5;26(10):2558-2565.e3.

Gururaj AE, Singh RR, Rayala SK, Holm C, den Hollander P, Zhang H, Balasenthil S, Talukder AH, Landberg G, Kumar R. MTA1, a transcriptional activator of breast cancer amplified sequence 3. Proc Natl Acad Sci U S A. 2006 Apr 25;103(17):6670-5.

Hah N, Murakami S, Nagari A, Danko CG, Kraus WL. Enhancer transcripts mark active estrogen receptor binding sites. Genome Res. 2013 Aug;23(8):1210-23.

Handwerger KE, Cordero JA, Gall JG. Cajal bodies, nucleoli, and speckles in the Xenopus oocyte nucleus have a low-density, sponge-like structure. Mol Biol Cell. 2005 Jan;16(1):202-11.

Harlen KM, Churchman LS. The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. Nat Rev Mol Cell Biol. 2017 Apr;18(4):263-273.

Harrison MM, Li XY, Kaplan T, Botchan MR, Eisen MB. Zelda binding in the early Drosophila melanogaster embryo marks regions subsequently activated at the maternal-to-zygotic transition. PLoS Genet. 2011 Oct;7(10):e1002266.

Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, Ching KA, Antosiewicz-Bourget JE, Liu H, Zhang X, Green RD, Lobanenkov VV, Stewart R, Thomson JA, Crawford GE, Kellis M, Ren B. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature. 2009 May 7;459(7243):108-12.

Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA. A Phase Separation Model for Transcriptional Control. Cell. 2017 Mar 23;169(1):13-23.

Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res. 2015 Jan;43(Database issue):D512-20.

Huang P, He Z, Ji S, Sun H, Xiang D, Liu C, Hu Y, Wang X, Hui L. Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. Nature. 2011 May 11;475(7356):386-9.

Hung V, Udeshi ND, Lam SS, Loh KH, Cox KJ, Pedram K, Carr SA, Ting AY. Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. Nat Protoc. 2016 Mar;11(3):456-75.

Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, Carroll JS. FOXA1 is a key determinant of estrogen receptor function and endocrine response. Nat Genet. 2011 Jan;43(1):27-33.

Ioshikhes IP, Albert I, Zanton SJ, Pugh BF. Nucleosome positions predicted through comparative genomics. Nat Genet. 2006 Oct;38(10):1210-5.

Jangal M, Couture JP, Bianco S, Magnani L, Mohammed H, Gévry N. The transcriptional co-repressor TLE3 suppresses basal signaling on a subset of estrogen receptor α target genes. Nucleic Acids Res. 2014 Oct;42(18):11339-48.

Jozwik K. et al., FOXA1 Directs H3K4 Monomethylation at Enhancers via Recruitment of the Methyltransferase MLL3. Cell Rep. 2016 Dec 6; 17(10): 2715–2723.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. Highly accurate protein structure prediction with AlphaFold. Nature. 2021 Aug;596(7873):583-589.

Kato M, Han TW, Xie S, Shi K, Du X, Wu LC, Mirzaei H, Goldsmith EJ, Longgood J, Pei J, Grishin NV, Frantz DE, Schneider JW, Chen S, Li L, Sawaya MR, Eisenberg D, Tycko R, McKnight SL. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell. 2012 May 11;149(4):753-67.

Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun. 2019 Apr 29;10(1):1930.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015 Jun;10(6):845-58.

Kim DH, Han KH. PreSMo Target-Binding Signatures in Intrinsically Disordered Proteins. Mol Cells. 2018 Oct 31;41(10):889-899.

Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G, Greenberg ME. Widespread transcription at neuronal activity-regulated enhancers. Nature. 2010 May 13;465(7295):182-7.

Kim HS, Tan Y, Ma W, Merkurjev D, Destici E, Ma Q, Suter T, Ohgi K, Friedman M, Skowronska-Krawczyk D, Rosenfeld MG. Pluripotency factors functionally premark cell-type-restricted enhancers in ES cells. Nature. 2018 Apr;556(7702):510-514.

Klemm SL, Shipony Z, Greenleaf WJ. Chromatin accessibility and the regulatory epigenome. Nat Rev Genet. 2019 Apr;20(4):207-220.

Kumar R, Wang RA, Mazumdar A, Talukder AH, Mandal M, Yang Z, Bagheri-Yarmand R, Sahin A, Hortobagyi G, Adam L, Barnes CJ, Vadlamudi RK. A naturally occurring MTA1 variant sequesters oestrogen receptor-alpha in the cytoplasm. Nature. 2002 Aug 8;418(6898):654-7.

Kwak H, Fuda NJ, Core LJ, Lis JT. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. Science. 2013 Feb 22;339(6122):950-3.

Kwon I, Kato M, Xiang S, Wu L, Theodoropoulos P, Mirzaei H, Han T, Xie S, Corden JL, McKnight SL. Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. Cell. 2013 Nov 21;155(5):1049-1060.

Li W, Notani D, Rosenfeld MG. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. Nat Rev Genet. 2016 Apr;17(4):207-23.

Lin W, Metzakopian E, Mavromatakis YE, Gao N, Balaskas N, Sasaki H, Briscoe J, Whitsett JA, Goulding M, Kaestner KH, Ang SL. Foxa1 and Foxa2 function both upstream of and

cooperatively with Lmx1a and Lmx1b in a feedforward loop promoting mesodiencephalic dopaminergic neuron development. Dev Biol. 2009 Sep 15;333(2):386-96.

Liu Z, Merkurjev D, Yang F, Li W, Oh S, Friedman MJ, Song X, Zhang F, Ma Q, Ohgi KA, Krones A, Rosenfeld MG. Enhancer activation requires trans-recruitment of a mega transcription factor complex. Cell. 2014 Oct 9;159(2):358-73.

Liu Z, Tjian R. Visualizing transcription factor dynamics in living cells. J Cell Biol. 2018 Apr 2;217(4):1181-1191.

Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, Axel R. Interchromosomal interactions and olfactory receptor choice. Cell. 2006 Jul 28;126(2):403-13.

Lupien M, Eeckhoute J, Meyer CA, Wang Q, Zhang Y, Li W, Carroll JS, Liu XS, Brown M. FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. Cell. 2008 Mar 21;132(6):958-70.

Magnani L, Ballantyne EB, Zhang X, Lupien M. PBX1 genomic pioneer function drives ERα signaling underlying progression in breast cancer. PLoS Genet. 2011 Nov;7(11):e1002368.

Markenscoff-Papadimitriou E, Allen WE, Colquitt BM, Goh T, Murphy KK, Monahan K, Mosley CP, Ahituv N, Lomvardas S. Enhancer interaction networks as a means for singular olfactory receptor expression. Cell. 2014 Oct 23;159(3):543-57.

Mazumdar A, Wang RA, Mishra SK, Adam L, Bagheri-Yarmand R, Mandal M, Vadlamudi RK, Kumar R. Transcriptional repression of oestrogen receptor by metastasis-associated protein 1 corepressor. Nat Cell Biol. 2001 Jan;3(1):30-7.

Meers MP, Janssens DH, Henikoff S. Pioneer Factor-Nucleosome Binding Events during Differentiation Are Motif Encoded. Mol Cell. 2019 Aug 8;75(3):562-575.e5.

Mo R, Rao SM, Zhu YJ. Identification of the MLL2 complex as a coactivator for ER-α α. J Biol Chem. 2006 Jun 9;281(23):15714-20.

Mohammed H, D'Santos C, Serandour AA, Ali HR, Brown GD, Atkins A, Rueda OM, Holmes KA, Theodorou V, Robinson JL, Zwart W, Saadi A, Ross-Innes CS, Chin SF, Menon S, Stingl J, Palmieri C, Caldas C, Carroll JS. Endogenous purification reveals GREB1 as a key estrogen receptor regulatory factor. Cell Rep. 2013 Feb 21;3(2):342-9.

Molliex A, Temirov J, Lee J, Coughlin M, Kanagaraj AP, Kim HJ, Mittag T, Taylor JP. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. Cell. 2015 Sep 24;163(1):123-33.

Mourad R, Hsu PY, Juan L, Shen C, Koneru P, Lin H, Liu Y, Nephew K, Huang TH, Li L. Estrogen induces global reorganization of chromatin structure in human breast cancer cells. PLoS One. 2014 Dec 3;9(12):e113354.

Nagarajan S, Rao SV, Sutton J, Cheeseman D, Dunn S, Papachristou EK, Prada JG, Couturier DL, Kumar S, Kishore K, Chilamakuri CSR, Glont SE, Archer Goode E, Brodie C, Guppy N, Natrajan R, Bruna A, Caldas C, Russell A, Siersbæk R, Yusa K, Chernukhin I, Carroll JS. ARID1A influences HDAC1/BRD4 activity, intrinsic proliferative capacity and breast cancer treatment response. Nat Genet. 2020 Feb;52(2):187-197.

Nair SJ, Yang L, Meluzzi D, Oh S, Yang F, Friedman MJ, Wang S, Suter T, Alshareedah I, Gamliel A, Ma Q, Zhang J, Hu Y, Tan Y, Ohgi KA, Jayani RS, Banerjee PR, Aggarwal AK, Rosenfeld MG. Phase separation of ligand-activated enhancers licenses cooperative chromosomal enhancer assembly. Nat Struct Mol Biol. 2019 Mar;26(3):193-203.

Papachristou EK, Kishore K, Holding AN, Harvey K, Roumeliotis TI, Chilamakuri CSR, Omarjee S, Chia KM, Swarbrick A, Lim E, Markowetz F, Eldridge M, Siersbaek R, D'Santos CS, Carroll JS. A quantitative mass spectrometry-based approach to monitor the dynamics of endogenous chromatin-associated protein complexes. Nat Commun. 2018 Jun 13;9(1):2311.

Papantonis A, Kohro T, Baboo S, Larkin JD, Deng B, Short P, Tsutsumi S, Taylor S, Kanki Y, Kobayashi M, Li G, Poh HM, Ruan X, Aburatani H, Ruan Y, Kodama T, Wada Y, Cook PR. TNFα signals through specialized factories where responsive coding and miRNA genes are transcribed. EMBO J. 2012 Nov 28;31(23):4404-14.

Parker SC, Stitzel ML, Taylor DL, Orozco JM, Erdos MR, Akiyama JA, van Bueren KL, Chines PS, Narisu N; NISC Comparative Sequencing Program, Black BL, Visel A, Pennacchio LA, Collins FS; National Institutes of Health Intramural Sequencing Center Comparative Sequencing Program Authors; NISC Comparative Sequencing Program Authors. Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. Proc Natl Acad Sci U S A. 2013 Oct 29;110(44):17921-6.

Parolia A, Cieslik M, Chu SC, Xiao L, Ouchi T, Zhang Y, Wang X, Vats P, Cao X, Pitchiaya S, Su F, Wang R, Feng FY, Wu YM, Lonigro RJ, Robinson DR, Chinnaiyan AM. Distinct structural classes of activating FOXA1 alterations in advanced prostate cancer. Nature. 2019 Jul;571(7765):413-418.

Partridge EC, Chhetri SB, Prokop JW, Ramaker RC, Jansen CS, Goh ST, Mackiewicz M, Newberry KM, Brandsmeier LA, Meadows SK, Messer CL, Hardigan AA, Coppola CJ, Dean EC, Jiang S, Savic D, Mortazavi A, Wold BJ, Myers RM, Mendenhall EM. Occupancy maps of 208 chromatin-associated proteins in one human cell type. Nature. 2020 Jul;583(7818):720-728.

Proudhon C, Snetkova V, Raviram R, Lobry C, Badri S, Jiang T, Hao B, Trimarchi T, Kluger Y, Aifantis I, Bonneau R, Skok JA. Active and Inactive Enhancers Cooperate to Exert Localized and Long-Range Control of Gene Regulation. Cell Rep. 2016 Jun 7;15(10):2159-2169.

Puc J, Aggarwal AK, Rosenfeld MG. Physiological functions of programmed DNA breaks in signal-induced transcription. Nat Rev Mol Cell Biol. 2017;18(8):471-476.

Puntervoll P, Linding R, Gemünd C, Chabanis-Davidson S, Mattingsdal M, Cameron S, Martin DM, Ausiello G, Brannetti B, Costantini A, Ferrè F, Maselli V, Via A, Cesareni G, Diella F, Superti-Furga G, Wyrwicz L, Ramu C, McGuigan C, Gudavalli R, Letunic I, Bork P, Rychlewski L, Küster B, Helmer-Citterich M, Hunter WN, Aasland R, Gibson TJ. ELM server: A new resource for investigating short functional sites in modular eukaryotic proteins. Nucleic Acids Res. 2003 Jul 1;31(13):3625-30.

Quintin J, Le Péron C, Palierne G, Bizot M, Cunha S, Sérandour AA, Avner S, Henry C, Percevault F, Belaud-Rotureau MA, Huet S, Watrin E, Eeckhoute J, Legagneux V, Salbert G, Métivier R. Dynamic estrogen receptor interactomes control estrogen-responsive trefoil Factor (TFF) locus cell-specific activities. Mol Cell Biol. 2014 Jul;34(13):2418-36.

Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. Nature. 2011 Feb 10;470(7333):279-83.

Segal E, Fondufe-Mittendorf Y, Chen L, Thåström A, Field Y, Moore IK, Wang JP, Widom J. A genomic code for nucleosome positioning. Nature. 2006 Aug 17;442(7104):772-8.

Sekiya S, Suzuki A. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. Nature. 2011 Jun 29;475(7356):390-3.

Siersbæk R, Baek S, Rabiee A, Nielsen R, Traynor S, Clark N, Sandelin A, Jensen ON, Sung MH, Hager GL, Mandrup S. Molecular architecture of transcription factor hotspots in early adipogenesis. Cell Rep. 2014 Jun 12;7(5):1434-1442.

Siersbæk R, Rabiee A, Nielsen R, Sidoli S, Traynor S, Loft A, Poulsen LC, Rogowska-Wrzesinska A, Jensen ON, Mandrup S. Transcription factor cooperativity in early adipogenic hotspots and super-enhancers. Cell Rep. 2014 Jun 12;7(5):1443-1455.

Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M, Zaret KS. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. Cell. 2015 Apr 23;161(3):555-568.

Soutourina J. Transcription regulation by the Mediator complex. Nat Rev Mol Cell Biol. 2018 Apr;19(4):262-274.

Spilianakis CG, Lee GR, Flavell RA. Twisting the Th1/Th2 immune response via the retinoid X receptor: lessons from a genetic approach. Eur J Immunol. 2005 Dec;35(12):3400-4.

Sun Y. et al., Zelda overcomes the high intrinsic nucleosome barrier at enhancers during Drosophila zygotic genome activation. Genome Res. 2015 Nov;25(11):1703-14.

Swinstead EE, Miranda TB, Paakinaho V, Baek S, Goldstein I, Hawkins M, Karpova TS, Ball D, Mazza D, Lavis LD, Grimm JB, Morisaki T, Grøntved L, Presman DM, Hager GL. Steroid Receptors Reprogram FoxA1 Occupancy through Dynamic Chromatin Transitions. Cell. 2016 Apr 21;165(3):593-605

Takeda DY, Spisák S, Seo JH, Bell C, O'Connor E, Korthauer K, Ribli D, Csabai I, Solymosi N, Szállási Z, Stillman DR, Cejas P, Qiu X, Long HW, Tisza V, Nuzzo PV, Rohanizadegan M, Pomerantz MM, Hahn WC, Freedman ML. A Somatically Acquired Enhancer of the Androgen Receptor Is a Noncoding Driver in Advanced Prostate Cancer. Cell. 2018 Jul 12;174(2):422-432.e13.

Tillo D, Hughes TR. 2009. G+C content dominates intrinsic nucleosome occupancy. BMC Bioinformatics 10: 442.

Toska E, Osmanbeyoglu HU, Castel P, Chan C, Hendrickson RC, Elkabets M, Dickler MN, Scaltriti M, Leslie CS, Armstrong SA, Baselga J. PI3K pathway regulates ER-dependent transcription in breast cancer through the epigenetic regulator KMT2D. Science. 2017 Mar 24;355(6331):1324-1330.

Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder

J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J, Pontén F. Proteomics. Tissue-based map of the human proteome. Science. 2015 Jan 23;347(6220):1260419.

Uversky VN, Gillespie JR, Fink AL. Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins. 2000 Nov 15;41(3):415-27.

Verger A, Duterque-Coquillaud M. When Ets transcription factors meet their partners. Bioessays. 2002 Apr;24(4):362-70.

Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, Afzal V, Ren B, Rubin EM, Pennacchio LA. ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature. 2009 Feb 12;457(7231):854-8.

Vladimir N. Uversky. Intrinsically Disordered Proteins and Their “Mysterious” (Meta)Physics. Frontiers in Physics. 2019 Feb 07. Sec. Biophysics. Vol. 7

Wang D, Garcia-Bassets I, Benner C, Li W, Su X, Zhou Y, Qiu J, Liu W, Kaikkonen MU, Ohgi KA, Glass CK, Rosenfeld MG, Fu XD. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature. 2011 May 15;474(7351):390-4.

Wang L, Ozark PA, Smith ER, Zhao Z, Marshall SA, Rendleman EJ, Piunti A, Ryan C, Whelan AL, Helmin KA, Morgan MA, Zou L, Singer BD, Shilatifard A. TET2 coactivates gene expression through demethylation of enhancers. Sci Adv. 2018 Nov 7;4(11):eaau6986.

Wang S, Singh SK, Katika MR, Lopez-Aviles S, Hurtado A. High Throughput Chemical Screening Reveals Multiple Regulatory Proteins on FOXA1 in Breast Cancer Cell Lines. Int J Mol Sci. 2018 Dec 19;19(12):4123.

Weber DS, Warren JJ. The interaction between methionine and two aromatic amino acids is an abundant and multifunctional motif in proteins. Arch Biochem Biophys. 2019 Sep 15;672:108053.

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell. 2013 Apr 11;153(2):307-19.

Williams RM, Obradovi Z, Mathura V, et al., The protein non-folding problem: amino acid determinants of intrinsic order and disorder. Pac Symp Biocomput. 2001;89-100.

Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Obradovic Z, Uversky VN. Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. J Proteome Res. 2007 May;6(5):1917-32.

Xue B, Dunbrack RL, Williams RW, Dunker AK, Uversky VN. PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. Biochim Biophys Acta. 2010 Apr;1804(4):996-1010.

Yang Y, Zhang R, Singh S, Ma J. Exploiting sequence-based features for predicting enhancer-promoter interactions. Bioinformatics. 2017 Jul 15;33(14):i252-i260.

Ye Z, Chen Z, Sunkel B, Frietze S, Huang TH, Wang Q, Jin VX. Genome-wide analysis reveals positional-nucleosome-oriented binding pattern of pioneer factor FOXA1. Nucleic Acids Res. 2016 Sep 19;44(16):7540-54.

Yang YA, Zhao JC, Fong KW, Kim J, Li S, Song C, Song B, Zheng B, He C, Yu J. FOXA1 potentiates lineage-specific enhancer activation through modulating TET1 expression and function. Nucleic Acids Res. 2016 Sep 30;44(17):8153-64.

Zaret KS, Carroll JS. Pioneer transcription factors: establishing competence for gene expression. Genes Dev. 2011 Nov 1;25(21):2227-41.

Zhaoyu Li, Paul Gadue, Kaifu Chen, Yang Jiao, Geetu Tuteja, Jonathan Schug, Wei Li, Klaus H Kaestner. Foxa2 and H2A.Z mediate nucleosome depletion during embryonic stem cell differentiation. Cell. 2012 Dec 21;151(7):1608-16.

Zou Z, Ohta T, Miura F, Oki S. ChIP-Atlas 2021 update: a data-mining suite for exploring epigenomic landscapes by fully integrating ChIP-seq, ATAC-seq and Bisulfite-seq data. Nucleic Acids Res. 2022 Mar 24;50(W1):W175–82.