The Glucocorticoid-Mediated Dynamics of Genome Architecture

by

Anthony M. D'Ippolito

University Program in Genetics and Genomics
Duke University

Date:_______________________

Approved:

___________________________
Timothy Reddy, Supervisor

___________________________
Gregory Crawford

___________________________
Charles Gersbach

___________________________
Donald McDonnell

___________________________
Jenny Tung

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy
in the University Program in Genetics and Genomics
in the Graduate School
of Duke University

2018
ABSTRACT

The Glucocorticoid-Mediated Dynamics of Genome Architecture

by

Anthony M. D'Ippolito

University Program in Genetics and Genomics
Duke University

Date:_______________________

Approved:

___________________________

Timothy Reddy, Supervisor

___________________________

Gregory Crawford

___________________________

Charles Gersbach

___________________________

Donald McDonnell

___________________________

Jenny Tung

An abstract of a Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the University Program in Genetics and Genomics in the Graduate School of Duke University

2018
Abstract

Human cells are perpetually receiving and responding to a variety of intrinsic and extrinsic signals. A primary mechanism by which cells carry out these responses is via changes in the regulation of gene expression. Many studies have examined gene regulation in steady state systems, but few have investigated the genomic response to stimuli. Therefore, it is less well understood how cellular stimuli elicit dynamic gene expression responses. Here, we investigate how extracellular stimuli mediate gene expression responses via: 1) Changes in transcription factor configurations at enhancer elements; and 2) Changes in chromatin looping between putative enhancers and their target gene promoters. To study these phenomena, we used glucocorticoid (GC) treatment as a model transcriptional stimulus. This hormone steroid is known to bind to and activate the GC receptor (GR), a ligand-induced transcription factor (TF), and is therefore a highly tractable system for studying stimulus responsive gene regulation. Using this model system, we first used high-resolution TF-binding site mapping approaches to elucidate the genomic binding locations of GR and its associated cofactors. Using these approaches, we found evidence that: 1) The GR binds to the genome as both a monomer and dimer; and 2) The GR binds to the genome with AP-1 in a more relaxed configuration, while it binds FOXA1 in a more constrained configuration. We next interrogated the role of chromatin looping in mediating dynamic
transcriptional responses. For this work we used high-throughput genomics methods to assay chromatin conformation across a time course of GC treatment. These studies resulted in several main findings: 1) Chromatin loops do not form in response to stimulus, but are instead pre-formed before GC treatment; 2) Chromatin looping interactions increase between distal GR binding sites and GC-responsive genes; 3) The insulator protein CTCF is depleted at stimulus responsive looping interactions; and 4) GC treatment mediates changes in higher-order chromosome compartmentalization that are concordant with gene expression responses. Together these results provide evidence for a genome topology that is pre-wired to respond to stimulus, and that subsequent transcriptional responses are mediated through GR binding to putative enhancer elements with other TFs, in a variety of genomic binding configurations.
Dedication

I would like to dedicate this work to my wife, and best friend, Tessa. I would not have been able to complete these work without your kindness, strength, and unconditional support. Your companionship truly makes me better. I am beyond thrilled to start our journey together as parents. Our son will be very lucky to have such an exceptional woman in his life.
Contents

Abstract ........................................................................................................................................ iv

List of Tables ................................................................................................................................ xi

List of Figures ............................................................................................................................ xii

Acknowledgements .................................................................................................................... xiv

1. Introduction ............................................................................................................................. 1

   1.1 Gene expression dynamics are a core component of cellular responses ...................... 1

   1.2 The non-coding genome ............................................................................................... 2

   1.3 The fundamental DNA elements of eukaryotic gene regulation ................................ 2

   1.4 Transcription factors are DNA-binding proteins that regulate gene expression ... 5

   1.5 Chromatin modifications in transcriptional regulation ............................................... 6

   1.6 Enhancers interact with promoters via chromatin loops .............................................. 8

   1.7 Enhancer activity is often constrained to topologically associating domains ........... 12

   1.8 Active and inactive chromatin spatially segregate into different chromosomal compartments ........................................................................................................... 15

   1.9 Glucocorticoid treatment as a model transcriptional stimulus .................................. 17

2. Binding configurations of the glucocorticoid receptor revealed by ChIP-exo ................ 19

   2.1 Introduction ..................................................................................................................... 19

   2.2 Results ............................................................................................................................ 21

      2.2.1 ChIP-exo identifies both dimer and monomer binding configurations of the GR ................................................................. 21

      2.2.2 Motif analyses reveal putative GR-cofactor binding configurations ................. 27
2.2.3 Putative GR-FOXA1 composite sites revealed by 2D footprinting ..................29

2.3 Discussion ..............................................................................................................31

2.4 Methods ..................................................................................................................33

2.4.1 Experimental .......................................................................................................33

2.4.1.1 Cell culture ....................................................................................................33

2.4.1.2 ChIP-exo .......................................................................................................34

2.4.2 Computational .....................................................................................................36

2.4.2.1 ChIP-exo data processing .............................................................................36

2.4.2.2 Binding site discovery via GEM .....................................................................36

2.4.2.3 2D Footprinting ............................................................................................36

3. Glucocorticoid treatment mediates chromatin loop dynamics concordant with gene expression responses ........................................................................................................38

3.1 Introduction ............................................................................................................38

3.2 Results ....................................................................................................................41

3.2.1 Thousands of loops change their interaction frequency in response to GC treatment ........................................................................................................................41

3.2.2 Dex-responsive loops are pre-existing ..............................................................44

3.2.3 Networks of chromatin loops have coordinated responses to dex treatment ..45

3.2.4 GC-induced genes are enriched for increased looping interactions ..................47

3.2.5 Chromatin looping dynamics are associated with low CTCF occupancy .......51

3.2.6 Increased and decreased loops are associated with chromatin state and TF binding ......................................................................................................................54
3.2.7 GR binding at distal enhancers is associated with gene activation and repression .................................................................55
3.2.8 GC-regulated genes change chromosome compartmentalization ..................59
3.3 Discussion ......................................................................................................................................................................................61
3.4 Methods ......................................................................................................................................................................................65
3.4.1 Experimental .............................................................................................................................................................................65
  3.4.1.1 Cell culture ........................................................................................................................................................................65
  3.4.1.2 In situ Hi-C ........................................................................................................................................................................65
  3.4.1.3 Targeted repression of chromatin loop anchors ..................................................70
3.4.2 Computational ..........................................................................................................................................................................72
  3.4.2.1 Differential loop analyses ..................................................................................................................................................72
  3.4.2.2 Visualization of dynamic interaction matrices ..................................................73
  3.4.2.3 Aggregate dynamic interaction matrices .........................................................74
  3.4.2.4 Pre-existing chromatin interaction analyses ..................................................74
  3.4.2.5 Analysis of loop network sizes ......................................................................75
  3.4.2.6 Coordination of dex-responsive networks ....................................................75
  3.4.2.7 Dynamic loop enrichment at dex-regulated genes .........................................76
  3.4.2.8 Enrichment of increased loop anchors at gene bodies of induced genes ...76
  3.4.2.9 Targeted repression of chromatin loop anchors ............................................77
  3.4.2.10 Analysis of TF occupancy vs. CTCF occupancy at dynamic loops ..........78
  3.4.2.11 Predicting differential expression with chromatin interactions .................78
  3.4.2.12 Predicting differential interactions via epigenomic marks .........................81
3.4.2.13 Differential compartment analyses ................................................. 82
3.4.2.14 Interaction dynamics between compartments .............................. 83
3.4.2.15 Enrichment of dex-regulated genes among compartment types .......... 83

4. Conclusions and Future Directions ................................................................. 85
4.1 The topologic pre-programming of transcriptional responses ................. 85
4.2 Stimulus-induced chromatin looping dynamics ........................................ 87
4.3 Temporal and spatial dynamics of chromatin loops and looping networks ... 89
4.3.1 How is the 3D distance between loop anchors altered upon GC treatment? .. 89
4.3.2 How do looping kinetics relate to gene expression kinetics? .................... 91
4.3.3 How do multiple enhancers regulate a single gene? .............................. 92
4.4 Mechanistic function of chromatin looping ................................................ 93
4.5 GR-mediated transcriptional repression ...................................................... 94
4.5.1 Why do putative GR-repressive sites lie so far from the genes they interact with? ............................................................................................................. 95
4.5.2 Are these GR sites looped to GC-repressed genes functional? ................. 96
4.5.3 What is the role of nGREs in global gene expression responses to GCs? .... 97

Appendix A ............................................................................................................. 98
References ............................................................................................................. 113
Biography ............................................................................................................. 137
List of Tables

Table 1: Reagents and Resources ........................................................................................................ 107

Table 2: gRNA Sequences for dCas9-KRAB experiments ................................................................. 111
List of Figures

Figure 1: Chromosome conformation capture ................................................................. 10

Figure 2: The loop extrusion model of chromatin loop formation................................. 11

Figure 3: Topological associating domains are a primary unit of genome organization.. 13

Figure 4: The genome is organized into at least two distinct compartments ............... 16

Figure 5: Overview of ChIP-exo biochemistry ............................................................... 22

Figure 6: ChIP-exo is comparable to ChIP-seq in assaying the repertoire of GR binding sites .......................................................... 23

Figure 7: ChIP-exo reveals putative dimer and monomer binding of the GR.............. 26

Figure 8: 2D footprinting reveals GR binding footprints .............................................. 27

Figure 9: GEM reveals putative GR-cofactor binding configurations ......................... 29

Figure 10: ChIP-exo reveals putative GR-FOXA1 binding configuration..................... 31

Figure 11: Thousands of loops change their interaction frequency in response to GC treatment ......................................................................... 42

Figure 12: Networks of chromatin loops have coordinated responses to dex treatment .46

Figure 13: Dex-induced genes are enriched for increased looping ............................. 50

Figure 14: Chromatin looping dynamics are associated with low CTCF occupancy ....53

Figure 15: GR binding at distal enhancers is associated with gene activation and repression ......................................................................... 56

Figure 16: Dex treatment mediates changes in chromatin compartmentalization ....60

Figure 17: Thousands of loops respond to dex treatment ........................................... 98

Figure 18: Networks of dynamic loops change concordantly across multiple time points ........................................................................ 99
Figure 19: Anchors of dex-increased loops function in dex-mediated gene induction... 100

Figure 20: Dex-mediated looping dynamics are associated with low CTCF occupancy 102

Figure 21: GR binding at distal enhancers is associated with gene activation and repression... 103

Figure 22: Compartmental dynamics are reproducible across resolutions... 105
Acknowledgements

This work would have not been possible without the exceptional mentorship of Dr. Timothy Reddy. I am extremely grateful for Dr. Reddy’s patience and generosity as a mentor. The opportunities he has granted me while working in his lab have been invaluable for my technical and professional growth as a scientist. I would also like to thank the members of my dissertation committee for their guidance throughout the development of this work: Dr. Gregory Crawford, Dr. Charles Gersbach, Dr. Jenny Tung, and Dr. Donald McDonnell.

I also want to acknowledge my many collaborators, including: Chris Vockley, Ian McDowell, Isaac Hilton, and Pratiksha Thakore. I’m very grateful for both the opportunities they have granted me to be involved in their work, as well as the contributions they have made to my own. The scientific discussion and collaboration I’ve had with these individuals has been a highlight of my graduate career.

I would also like to acknowledge the many exceptional mentors I’ve had early in my career. I specifically want to thank Dr. Christopher Phiel for his training in molecular biology, and for giving me the opportunity to work in his lab as an undergraduate. Finally, I would like to thank my high school biology teacher, Eric Joreski, for his challenging but inspirational instruction.
1. Introduction

1.1 Gene expression dynamics are a core component of cellular responses

Human cells are perpetually receiving and responding to a variety of extracellular stimuli. The mechanisms governing these responses are complex and diverse, depending on the specific perturbation. However, a common core component of these responses is a change in the expression of the cell’s genes. This process is thought to be mediated by a complex interplay of DNA sequence, DNA-binding proteins, epigenetic modifications, and the 3D conformation of the genome. Though enormous efforts have been taken to better understand the genomics of gene regulation in the context of static cellular systems, how the dynamics of such components are involved in cellular responses is less well understood (Consortium, 2012, Roadmap Epigenomics et al., 2015). A deeper understanding of these processes will not only provide insight into the fundamental biological mechanisms of gene regulation, but has direct implications for human health and disease (Degner et al., 2012, Trynka et al., 2013). Here we use glucocorticoids as a model cellular stimulus to investigate how DNA-binding proteins and dynamics in the 3D topology of the genome mediate transcriptional responses. In this chapter, we will review how the above-described features of the non-coding genome are thought to come together to regulate gene expression.
1.2 The non-coding genome

The haploid human genome is comprised of ~3 billion base pairs of DNA sequence. The sequencing of the human genome revealed that ~1% of sequence is protein coding, ~24% is contained within introns, and the remainder, ~75% is intergenic (Venter et al., 2001, Lander et al., 2001). Since then, large consortiums have dispensed massive efforts to characterize this non-coding genomic sequence (Consortium, 2012, Roadmap Epigenomics et al., 2015). These studies have revealed a highly complex genomic landscape, riddled with DNA-binding proteins, covalent modifications of DNA and histone proteins, and long-distance interactions between linearly distal genomic loci. Although we have learned much about the composition of the non-coding genome, how all of the above-described features of the genome are integrated to produce cellular heterogeneity, and to allow for the rapid response to intrinsic and extrinsic stimuli is less well understood.

1.3 The fundamental DNA elements of eukaryotic gene regulation

There are four broad classes of DNA regulatory sequences: promoters, enhancers, insulators, and silencers. Beginning with the former, promoters are located at the transcriptional start site (TSS) of genes, and are composed of a variety of DNA sequences that recruit RNA polymerase II and other general transcription factors (GTFs)
(Ippen et al., 1968, Stevens, 1960, Conaway et al., 1991a). One notable sequence is the TATA box, an AT-rich sequence that binds the TATA-binding protein (TBP). TBP binding leads to the recruitment of a variety of TATA-associated factors (TAFs) and RNA polymerase II, culminating in the formation of the transcriptional pre-initiation complex (PIC) (Conaway et al., 1991b). Phosphorylation of the C-terminal domain (CTD) of RNA polymerase II triggers transcriptional elongation (Cadena and Dahmus, 1987). Phosphorylation of the CTD is in part regulated by Mediator, a multi-subunit complex that helps to stabilize the PIC, and serves as a bridge to integrate signals between the PIC and transcription factors bound in cis (Kelleher et al., 1990).

Cis-bound TF binding sites that positively modulate the transcriptional output of a given promoter are called enhancers (Banerji et al., 1981b). These sequences typically contain DNA motifs for several TFs (Panne et al., 2007), and have been observed several kb away from their target promoter (Kleinjan et al., 2006, Lettice et al., 2002, Consortium, 2012, Banerji et al., 1981a, Banerji et al., 1981b). The presence of different TF binding sites allows enhancer elements to integrate both intrinsic and extrinsic cellular signals. Therefore, the repertoire of active enhancer elements in a given cell is highly cell-type and condition specific (Consortium, 2012, Roadmap Epigenomics et al., 2015). Two popular models exist for explaining combinatorial TF binding: 1) A tightly constrained regulatory “grammar”; and 2) A more flexible “billboard” model of TF binding. In the former model, the activity of an enhancer is highly dependent on the number,
orientation, spacing, and order of TF binding sites. One representative example of such enhancer architecture is at the interferon-β (IFN-β) gene. Here, enhancer activity depends on the binding of eight different TFs (Panne et al., 2007). Conversely, in the billboard model of enhancer activity, TFs are thought to bind cooperatively, but with few orientation and spacing constraints (Kulkarni and Arnosti, 2003). Moreover, it has been proposed that the billboard model of enhancer architecture should be reexamined to be even more flexible, accompanying multiple, distal enhancer sequences that come together via protein-protein interactions (Vockley et al., 2017)

Next, whereas enhancer elements function to increase transcription, silencer elements decrease the transcription of their target gene (Brand et al., 1985). Silencers have proved difficult to detect with current genomic methods (Vockley et al., 2016), and as such, are less well studied than their enhancer counterparts. However, one example of a transcriptional silencer is the negative glucocorticoid (GC) response element (nGRE). Some studies have shown that the GC receptor (GR), a TF activated by GCs, can bind to an nGRE, recruit corepressor proteins, and subsequently repress gene transcription (Surjit et al., 2011, Hudson et al., 2013, Ramamoorthy and Cidlowski, 2013)

Finally, insulator elements are sequences commonly bound by CCCTC-binding factor (CTCF), and prevent upstream enhancers from acting on downstream promoters

---

1 This is a manuscript I co-authored as part of my graduate studies.
2 This is also a manuscript I co-authored as part of my graduate studies.
(and vice-versa) (Bell et al., 1999). A prototypic example is the imprinting control region (ICR) at the H19-Igf2 locus (Bell and Felsenfeld, 2000). Here, the ICR is positioned downstream of the Igf2 gene, and upstream of the H19 gene and an enhancer element. Importantly, the ICR is differentially methylated, depending on the parental origin of the allele. When unmethylated, the ICR binds CTCF and acts as an insulator, preventing the downstream enhancer from acting on Igf2, resulting in H19 expression. Conversely, when the ICR is methylated, it can no longer bind CTCF, thus allowing the downstream enhancer to act on Igf2, resulting in its expression. Notably, CTCF-bound insulator elements are a hallmark of topologically associating domains (Dixon et al., 2012), which are described in further detail in the sections to follow.

1.4 Transcription factors are DNA-binding proteins that regulate gene expression

Transcription factors (TFs) are DNA-binding proteins that are found at transcriptional regulatory elements (Payvar et al., 1981). These proteins are generally comprised of two domains: a DNA-binding domain, and an effector domain (Brent and Ptashne, 1985). DNA-binding domains recognize and bind to short DNA motifs, which typically are 6-12 bp in length, and often contain degenerate bases (Stormo, 2000). Notably, DNA-binding domains are thought to not only to recognize the sequence of DNA, but also shape of DNA (Slattery et al., 2014). The effector domains of TFs are involved in the recruitment of other TFs and transcriptional co-activators/repressors
(Fryer and Archer, 1998, Chen and Evans, 1995). The activities of DNA-binding domains and effector domains can sometimes be coupled. For example, the glucocorticoid receptor (GR) is a TF observed to bind a ~15 bp degenerate motif, known as a glucocorticoid response element (GRE) (Strahle et al., 1987). GR binding to the GRE results in the recruitment of various co-regulator proteins and chromatin modifying enzymes (Fryer and Archer, 1998, Kamei et al., 1996, Ronacher et al., 2009, Guenther et al., 2001). Slight variations in this motif have been found to induce conformational changes in the protein structure, propagating through to its effector domains, and resulting in the recruitment of different co-regulator proteins (Meijsing et al., 2009).

1.5 Chromatin modifications in transcriptional regulation

As alluded to above, another core component of eukaryotic gene regulation is the compaction and covalent modification of chromatin (Lorch et al., 1987). The fundamental packaging unit of the genome is the nucleosome, comprised of 8 histone octamers (Kornberg and Thomas, 1974, Finch et al., 1977). These octamers are made up of 2 sets of H2A, H2B, H3, and H4 proteins, and are wrapped with approximately 147 bp of DNA. Each histone octamer has an N-terminal “tail” that frequently undergoes a variety of different covalent modifications (Durrin et al., 1991). Notable modifications implicated in gene regulation are methylation of lysine 4 of H3, acetylation of lysine 27 of H3, methylation of lysine 27 of H3, and methylation of lysine 9 of H3. H3K4 mono-
methylation is associated with enhancer elements, while tri-methylation of this residue is frequently found at promoters (Liang et al., 2004). H3K27 acetylation is found at active enhancer elements and TSSs (Creyghton et al., 2010). This mark is deposited by p300/CBP, which in itself is highly predictive of enhancer elements (Visel et al., 2009, Tie et al., 2009). Moreover, both H3K27ac and H3K4me1 are commonly used together to designate putative active enhancers (Creyghton et al., 2010). H3K27 tri-methylation is associated with transcriptional repression via the Polycomb proteins (Aranda et al., 2015, Kuzmichev et al., 2002). Lastly, H3K9me tri-methylation is typically associated with transcriptionally silent heterochromatin (Nakayama et al., 2001).

The abovementioned histone modifications are deposited by a variety of enzymes, including histone acetyltransferases (HATs) and histone methyltransferases (HMTs) (Kleff et al., 1995, Tie et al., 2009, Wang et al., 2009, Rea et al., 2000). Conversely, these marks can be removed by histone deacetylases (HDACs) and histone demethylases (HDMs) (Taunton et al., 1996, Shi et al., 2004). HDACs and HDMs can be recruited to chromatin by TFs, transcriptional co-regulators, and in some cases, long non-coding RNAs (lncRNAs) (Arias et al., 1994, Kamei et al., 1996, Kaneko et al., 2010). One mechanism by which these histone modifications regulate transcription is via the recruitment of co-regulator proteins. For example, several co-regulators (including p300/CBP) possess bromodomains (BRDs) that recognize acetylated lysine residues (Sanchez and Zhou, 2009). Another mechanism by which histone modifications are
thought to affect gene regulation is by directly affecting chromatin compaction. For example, some evidence suggests that acetylation of histone tails reduces the positive charge of the affected lysine residues, thus decreasing the association of histones with negatively charged DNA (Fenley et al., 2010). This dissociation is thought to allow for increased accessibility of regulatory elements to bind their cognate TFs.

1.6 Enhancers interact with promoters via chromatin loops

There are various mechanisms that could explain how distal enhancer elements communicate with their target promoters (Wang and Giaever, 1988, Ptashne, 1986). These mechanisms include: 1) Protein tracking, where transcription factors bound at a distal enhancer element translocate across the DNA until encountering a target promoter; 2) Oozing, where proteins at distal enhancers perpetually recruit additional proteins to their adjacent sequences, until reaching a target promoter; and 3) DNA looping, where proteins bound to enhancer elements physically interact with basal transcriptional machinery at the promoter, looping out the intervening DNA. Mounting experimental evidence points to the latter model as the predominant form of communication between enhancers and promoters. For example, the first observation of such a DNA loop was by Griffith and colleagues, studying the λ repressor, a transcription factor found in bacteriophage (Griffith et al., 1986). Using electron microscopy, they observed a loop between two λ repressor binding sites located ~50 bp
apart. Since then, observations of eukaryotic chromatin loops have become common in studies of gene regulation. A notable example is the β-globin gene cluster, where 5 globin genes are regulated by a series of DHSs over 50 kb upstream, known as the locus control region (LCR) (Levings and Bungert, 2002, Maniatis et al., 1978, Tuan et al., 1985). Here it was observed that these hypersensitive sites loop to the β-globin genes in a cell-type specific manner (Carter et al., 2002, Tolhuis et al., 2002). One method used to observe such looping is chromosome conformation capture (3C) (Dekker et al., 2002). This method entails cross-link interacting pieces of DNA, and a subsequent ligation step, where those fragments locked in a looping interaction will ligate together more frequently than to other fragments (Figure 1). These chimeric fragments can then be assayed via PCR, qPCR, or DNA sequencing. This assay has seen widespread use in genomics studies of chromosome architecture (Sanyal et al., 2012, Fullwood et al., 2009, Lieberman-Aiden et al., 2009, Dixon et al., 2012, Jin et al., 2013, Schmitt et al., 2016), and is a primary source of data in the experiments outlined below.
Figure 1: Chromosome conformation capture

(A) Transcription factors often bind distal to their target promoters. (B) In chromosome conformation capture, looping interactions between distal enhancers and promoters are captured by: 1) cross-linking protein-DNA complexes, and treating with a restriction enzyme; 2) treating with a DNA ligase such that intra-complex ligations are favored over inter-complex ligations; and 3) releasing these chimeric DNA molecules. (C) The resultant chimeric molecules can then be assayed by PCR or DNA-sequencing, and interactions can be inferred by the enrichment of chimeras containing both sequences thought to participate in the chromatin loop.

Recent studies have converged on a model of chromatin loop formation termed the “loop-extrusion” model (Sanborn et al., 2015, Fudenberg et al., 2016). Under this model, cohesin is loaded onto the genome and proceeds to extrude DNA through its ring structure. This process continues until cohesin is sterically blocked by a boundary element, such as a CTCF binding site (Figure 2). Notably, this process appears to depend on the orientation of the pair of CTCF motifs. Convergent motifs appear to be more amenable to loop formation, where divergent motifs, or motifs positioned in the same direction do not appear to form chromatin loops (Rao et al., 2014). Experiments altering
the orientation of convergent CTCF motifs at chromatin loops via CRISPR-genome editing validate this model (Sanborn et al., 2015). Furthermore, computational simulation studies using this model of chromatin loop formation are able to accurately reproduce the structures of topologically-associating domains (see below) (Fudenberg et al., 2016).

![Diagram of the loop extrusion model of chromatin loop formation](image)

**Figure 2: The loop extrusion model of chromatin loop formation**

Under the loop-extrusion model of chromatin loop formation, cohesin is first loaded onto the genome, and begins to extrude DNA. This process is thought to continue until cohesin is sterically blocked, in this case, by CTCF binding sites. CTCF binding sites at chromatin loops are largely observed in the convergent orientation.
1.7 Enhancer activity is often constrained to topologically associating domains

Soon after the advent of high-throughput sequencing technologies, studies began to move beyond single-locus interrogation of chromatin looping and investigate genome architecture on a global scale. Some of the first global studies of genome architecture utilized a derivative of 3C, called Hi-C. These studies revealed previously unknown topological features of the genome (Lieberman-Aiden et al., 2009, Dixon et al., 2012). One such feature is that chromosomes are demarcated into megabase-scale genomic neighborhoods called topologically associating domains (TADs). These regions have strict boundaries, where loci within the boundaries physically interact more with each other than those regions outside of the boundaries (Figure 3). Moreover, these domains are frequently demarcated by a chromatin loop, where the boundaries of a TAD are enriched for physical interactions with one-another, compared to the already increased interaction levels within the TAD itself. For this reason, TADs are sometimes referred to as “loop domains” in the literature (Rao et al., 2014). The boundaries of such regions are commonly bound by CTCF. It was recently demonstrated that the inducible degradation of CTCF results in the loss of TAD formation, highlighting this protein’s role in the maintenance of genome architecture (Nora et al., 2017).
**Figure 3: Topological associating domains are a primary unit of genome organization**

Hi-C data from A549 cells at a region on chromosome 17, visualized by Juicebox (Durand et al., 2016a). Each pixel represents an interaction between one locus in this region (x-axis) with another region in this locus (y-axis). The color intensity of the pixels indicates the number of observed interactions between the two respective loci. This number has been capped at 97 for visualization purposes. TADs have been outlined with black boxes.
Mounting evidence suggests that TADs are a core unit of genome organization that serve important roles in gene regulation, DNA replication, and VDJ recombination (Narendra et al., 2015, Pope et al., 2014, Hu et al., 2015). TADs have been observed to be highly conserved between species, and highly invariant across different cell types of the same organism (Vietri Rudan et al., 2015, Dixon et al., 2015, Schmitt et al., 2016). In studies of transcriptional regulation, TADs appear to compartmentalize the gene-regulatory activity of enhancer elements, and prevent the activity of enhancers in adjacent TADs from exerting regulatory effects on genes within the TAD. For example, the HoxA gene cluster is divided into two TADs, each flanked by CTCF binding sites. When an intervening CTCF site is deleted via CRISPR genome editing, active chromatin marks in one TAD are observed to spread into the adjacent TAD. Furthermore, new enhancer-promoter interactions are formed between the adjacent TADs, and these interactions are associated with increased gene expression (Narendra et al., 2015).

Additional evidence supporting the gene-regulatory role of TADs is the discovery of so-called TADopathies (Matharu and Ahituv, 2015). For example, the disruption of TAD boundaries at the WNT6/IHH/EPHA4/PAX3 locus leads to ectopic expression of said genes, and can subsequently result in a variety of limb malformation syndromes (Lupianez et al., 2015).
1.8 Active and inactive chromatin spatially segregate into different chromosomal compartments

With the advent of Hi-C, Lieberman-Aiden and colleagues observed two distinct chromosomal compartments, termed compartment “A” and compartment “B” (Lieberman-Aiden et al., 2009). Such compartmentalization can be observed via the plaid patterns in a Hi-C interaction map (Figure 4A and 4B). Compartment “A” was observed to contain gene-dense regions enriched for epigenetic marks indicative of active transcription. Conversely, compartment “B” was observed to be gene-poor and have epigenetic marks indicative of heterochromatin. Furthermore, regions in compartment “B” are concordant with lamina-associated domains (LADs) (Kind et al., 2015). Such compartmentalization has also been observed in imaging studies, where domains belonging to each compartment spatially segregate for each chromosome (Wang et al., 2016). Whereas TADs are highly cell-type invariant, chromosome compartments appear to be less so. Studies of genome topology during cellular differentiation show that TADs can switch their compartment identity during differentiation, and that these changes are associated with concordant changes in gene expression (Dixon et al., 2015). Similar results have been observed in the meta-analyses of several human cell lines and primary tissues (Schmitt et al., 2016).
Figure 4: The genome is organized into at least two distinct compartments

(A) In situ Hi-C data from A549 cells at chromosome 20. The plaid patterns of alternating blocks of increased interactions are indicative of chromosome compartments, where large regions appear to skip over some regions to engage in physical interactions.

(B) The same region as in A, except the data has been transformed for better visualization of compartments. First the observed interaction counts were normalized by the interaction frequency expected by chance, and then the Pearson correlation matrix was taken.

The functional role of chromosome compartments in the regulation of gene expression is unclear. In studies where either CTCF or cohesin are degraded, chromosome compartments remain intact, while most TADs and loop structures disappear (Nora et al., 2017, Rao et al., 2017). These findings provide evidence that compartmentalization is independent from mechanisms of chromatin loop formation.

Some have proposed that compartmentalization is driven by the aggregation of like-histone modifications and chromatin-associated proteins, termed phase-separation (Jost et al., 2014, Strom et al., 2017, Hnisz et al., 2017). Indeed, Strom and colleagues recently
demonstrated that heterochromatin protein 1 (HP1) displays liquid-liquid demixing properties, self-nucleating into distinct foci (Strom et al., 2017). Therefore, it is unclear whether chromatin compartmentalization is an active process, driving gene-regulatory programs, or is a structure that is merely reflective of the aggregation of like-chromatin features in the nucleus.

1.9 Glucocorticoid treatment as a model transcriptional stimulus

Glucocorticoids (GCs) are steroid hormones that are naturally occurring in the body (Mason, 1936, Reichstein, 1936). The predominant form of GCs in the body is cortisol, and is synthesized in the adrenal glands (Kendall, 1951). GCs function in a variety of biological processes, including metabolism, circadian rhythms, and inflammation (Kadmiel and Cidlowski, 2013). Notably, GCs are potent repressors of inflammatory responses (Hench et al., 1949). As such, synthetic GCs such as dexamethasone (dex) have been utilized in the clinic to treat a variety of inflammatory ailments (Coutinho and Chapman, 2011, Bunim et al., 1958).

The molecular mechanism by which GCs mediate cellular responses is via activation of the GC receptor (GR) (Hollenberg et al., 1985). Prior to hormone binding, the GR resides in the cytoplasm bound to chaperone proteins (Sanchez et al., 1985). Once bound by GCs, the GR is released from this protein complex, translocates to the nucleus, and binds to the genome to both activate and repress transcription (Ringold et al., 1977,
Pfahl, 1982, Drouin et al., 1989). This ligand-inducible activity of the GR makes it a highly tractable model TF to study the gene-regulatory responses of extracellular stimuli. Also, it primarily binds distal to GC-regulated genes, at putative enhancer elements (Reddy et al., 2009), and cooperates with a variety of other TFs (Biddie et al., 2011, Herzig et al., 2001, Nitsch et al., 1993). Together these characteristics make the GC response a well-suited model stimulus to study: 1) TF binding site architecture at enhancer elements; and 2) Long-distance interactions between enhancers and promoters. These two features of gene regulation will be the focus of the following two chapters. In Chapter 2 we will examine the various binding configurations in which the GR and its associated cofactor TFs bind to the genome. In Chapter 3 we will examine how GC treatment affects global chromatin looping interactions, and how these interactions inform mechanisms of GC-responsive gene regulation.
2. Binding configurations of the glucocorticoid receptor revealed by ChIP-exo

2.1 Introduction

Gene regulatory loci are frequently occupied by several transcription factors and cofactor proteins. Some studies suggest that the spacing and orientation of different transcription factor binding sites must be highly constrained for proper regulatory element function (Panne et al., 2007). This model of TF organization is said to adhere to a binding-site “grammar”. Other studies suggest that spacing and orientation are not important for function, and merely the presence/absence of TFs is important for proper transcriptional activity. This type of organization has been described as the “billboard model” of transcription factor binding (Kulkarni and Arnosti, 2003). Although some high-throughput studies have examined how TF binding configurations mediate transcriptional responses using reporter constructs (Smith et al., 2013), few have examined the precise binding configurations in the endogenous genome. It is therefore unclear how often TFs adhere to each of these models in their natural genomic context.

A representative transcription factor that is known to bind to the genome with several other TFs is the glucocorticoid receptor (GR). The GR has been observed to bind with TFs such as AP-1, FOXA1, CREB1, and NF-κB (Yang-Yen et al., 1990, Nitsch et al., 1993, Herzig et al., 2001, Ray and Prefontaine, 1994). Furthermore, these different binding configurations are thought to be important for eliciting specific transcriptional responses (Newton and Holden, 2007). Though these single-locus studies of the
glucocorticoid receptor have been insightful in understanding the mechanistic basis of GR-mediated transcriptional regulation, the distribution of these various binding site configurations across the genome is unknown. Therefore, it is unclear how different GR binding modes elicit transcriptional responses to glucocorticoids. Although current high-throughput technologies such as ChIP-seq reveal the genomic loci bound by a specific transcription factor, its resolution is limited to a few hundred base pairs (Johnson et al., 2007, Robertson et al., 2007).

To better understand the genomic distribution of the different GR binding configurations, we used ChIP-exo, a high-resolution variant of ChIP-seq (Rhee and Pugh, 2012). This method uses an exonuclease digestion to trim back the DNA fragments to precisely where the protein of interest binds to the DNA. Using this approach, we found evidence for both monomeric and dimeric genomic binding of the GR. We also found evidence for the GR participating in co-binding with other TFs, reminiscent of either a strict regulatory grammar, or billboard model, depending on the cofactor protein. We also were able to observe a robust DNA footprint at GR-FOXA1 binding sites, distinct from monomeric/dimeric GR binding.
2.2 Results

2.2.1 ChIP-exo identifies both dimer and monomer binding configurations of the GR

To better understand the different binding configurations of the GR, we chose to use ChIP-exo, a derivative of ChIP-seq that has been shown to have higher precision in identify the genomic coordinates of DNA-binding proteins (Rhee and Pugh, 2012). This method utilizes λ exonuclease to digest away DNA until the enzyme reaches a protein-DNA cross-link (Figure 5). This step allows for the identification of the boundaries of DNA-bound proteins via high-throughput sequencing. In this study, we treated A549 cells with 100 nM of dexamethasone, a synthetic glucocorticoid, and performed ChIP-exo on the resulting chromatin. We attempted to construct libraries using chromatin from a vehicle control treatment, however, this did not yield a detectable library. This is in line with the GR’s requirement for glucocorticoid for genomic localization (Picard and Yamamoto, 1987), and the low background signal observed from previous ChIP-exo experiments (Rhee and Pugh, 2012).
Figure 5: Overview of ChIP-exo biochemistry.

With the extensive number of biochemical reactions, specifically exonuclease digestions, we were concerned that low affinity GR binding sites may be more difficult to detect via ChIP-exo. To address this concern, we compared the coverage of GR binding sites from ChIP-exo versus traditional ChIP-seq (Figure 6) (Reddy et al., 2009). We found the coverage from ChIP-exo to be positively correlated with the coverage from ChIP-seq ($R^2 = 0.66$).

Next, we sought to use the precision of ChIP-exo to interrogate how the GR binds to its cognate motif, the glucocorticoid response element (GRE). To do this, we aggregated the coverage of the 5’ position of read 1 at the best match to the GRE at each GR binding site (Figure 7A). The 5’ position of read 1 should be indicative of where λ exonuclease is stopped, presumably by DNA-protein formaldehyde cross-links. We also performed the same analysis for the 5’ position of read 2, whose position should only be
Figure 6: ChIP-exo is comparable to ChIP-seq in assaying the repertoire of GR binding sites

Each data point represents a GR binding site called with ChIP-seq data. Values represent the number of sequencing reads that fall within a GR binding site, normalized to the total number of mapped sequencing reads. The result of sonication, and is not specific to the location of DNA-bound proteins (Figure 7B). Read 2 positions should be similar to what is observed with traditional ChIP-seq, and was therefore included as a control. From here forward, we will refer to read 1 as the ChIP-exo read and read 2 as the ChIP-seq read.

Using the ChIP-exo read we observed an aggregate signal that could be indicative of both the classical dimer binding configuration of the GR, as well as the
binding of individual monomers (Figure 7C). However, it is unclear if this signal is indicative of bona fide monomer binding events, or just incomplete cross-linking of the GR dimer. Using the ChIP-seq read, we observed a much broader aggregate signal, consistent with the fact that this read should not be specific to the site of protein-DNA cross-links. We did observe some signal with the ChIP-seq read that is consistent with GR monomer/dimer binding, presumably due to a slightly higher affinity of dsDNA breaks at the TF binding site via sonication. Indeed, sonication has been used to map accessible chromatin regions without the need for an exonuclease or endonuclease treatment (Auerbach et al., 2009).

The aggregate ChIP-exo signal discerns the boundaries of both dimeric and monomeric GR binding. However, it is unclear whether the observed boundaries reflect each side of the same DNA-protein complexes, or result from heterogeneous GR binding across the cell population. In other words, the information from each strand has been kept separate, and pairing such information would help discern the full footprints at each GR binding site. To test the hypothesis that the GR binds to the genome as a full dimer and not just as two adjacent monomers, we sought to develop an approach to combine information across strands to pair the upstream and downstream protein boundaries into a full DNA footprint. Briefly, we first locally smoothed read pile-ups for each strand using kernel density estimation, and then multiplied these vectors to produce a matrix, or 2D DNA footprint (Figures 8A and 8B). We next aggregated these
footprints across all GR binding sites, and then used information about the local maxima
to extrapolate the binding configurations of the GR (Figures 8C and 8D). Using this
approach, we were able to discern the full dimeric, as well as monomeric, binding
configurations of the GR. These observations support the hypothesis that the GR can
indeed bind to the genome as a dimer.
Figure 7: ChIP-exo reveals putative dimer and monomer binding of the GR

(A, B) The average pile-up of ChIP-exo reads at GR binding sites, in respect to the center of the strongest GR motif. The 5’ position of read 1 and read 2, respectively, were used, and pile-ups were separated by strand. (C) Schematic of putative binding configurations of the GR. ChIP-exo signal could represent the classical dimer configuration, either monomer binding, or dimer binding with incomplete cross-linking of one monomer.
Figure 8: 2D footprinting reveals GR binding footprints

(A) Example of read pile-ups at the binding site of a GR dimer. Prior to footprinting analysis, the pile-ups of the 5’ end of read 1 are determined for each strand at a genomic window centered on the best GR motif. (B) For each strand, pile-ups are smoothed via kernel density estimation, and the product of these distributions are used to create a matrix of probabilities. (C) Footprinting analysis of GR ChIP-exo data reveals 3 primary binding configurations. (D) These 3 regions of enrichment can be interpreted as dimer, and either monomer binding configurations of the GR.

2.2.2 Motif analyses reveal putative GR-cofactor binding configurations

Next, we wanted to test the hypothesis that the GR and its cofactor TFs bind to the genome in distinct configurations. To test this hypothesis, we decided to use Genome Wide Binding Event Finding and Motif Discovery (GEM) (Guo et al., 2012).
This method uses the spatial distribution of ChIP-exo reads to simultaneously elucidate the location of binding events and putative explanatory motifs at those binding events. Using GEM, we recovered the GR motif, motifs for known GR cofactors, and one motif with an unknown binding protein (Figure 9). We found that these motifs varied in their positioning in respect to the predicted binding coordinates. For example, while the GR motif was found to be highly enriched at the precise location of binding, AP-1, a transcription factor shown to frequently co-occupy GR binding sites (Biddie et al., 2011), was found to be broadly enriched at GR binding sites. In contrast to AP-1, motifs for TEAD1 and FOXA1 were found to be relatively constrained in respect to the binding location. The latter observation has since been reproduced in a separate study (Starick et al., 2015).
Putative binding site motifs as determined by GEM. Each row represents a different motif found to be enriched at GR ChIP-exo binding events. The first column are the motif PWMs as sequence logos. The second column are lollipop diagrams of the distribution of motif occurrences in respect to the binding event. The highest number of occurrences on either strand are labeled with both the distance from the binding event, and the number of occurrences at that position.

2.2.3 Putative GR-FOXA1 composite sites revealed by 2D footprinting

We noticed that the cofactor motifs discovered by GEM were similar to GR’s motif. Therefore, we hypothesized that these motif occurrences may not represent cofactor binding, but are merely slight variations of the GRE to which the GR binds. To determine whether these motifs represent DNA footprints distinct from the GR monomer/dimer footprints, we decided to apply our 2D footprinting method to one of
these cofactor motifs, FOXA1. Prior to footprinting, we first subset GR binding sites shown to be co-bound by FOXA1 (Reddy et al., 2012). We observed a footprint at these GREs that is smaller than the dimeric GR configuration, but larger than monomeric GR binding (Figure 10A). We hypothesized this footprint may represent either a GR-FOXA1 composite site, or a tethered binding site at which the GR indirectly binds to the genome via FOXA1, or vice versa (Figure 10C). To validate that this specific footprint was occupied by FOXA1, and to help elucidate the binding configuration at these GR-FOXA1 binding sites, we performed ChIP-exo for FOXA1. When performing footprinting with this data, on the same set of binding sites, we found a strong footprint at the same position as observed with GR ChIP-exo data, providing further evidence for FOXA1 occupancy at this specific footprint (Figure 10B). Notably, we observed two other prominent FOXA1 footprints at these sites, spanning ~40 and ~80 bps, centered on the GRE. It is unclear what specific binding configurations these footprints represent, but they may include the additional binding of cofactors and co-activator proteins to these GR-FOXA1 co-occupied sites.
Figure 10: ChIP-exo reveals putative GR-FOXA1 binding configuration

(A) Footprinting analysis at GR binding sites known to be co-occupied by FOXA1 display a unique binding configuration. (B) Using FOXA1 ChIP-exo data for footprinting provides evidence for FOXA1 occupancy at these binding sites. (C) These footprinting enrichments could be interpreted as a FOXA1-GR monomer composite site, or tethered binding, where one TF binds indirectly to the DNA via protein-protein interactions. Here the orange and teal ovals represent the GR and FOXA1, respectively.

2.3 Discussion

Results from previous studies of the GR have demonstrated that the configurations with which it binds the genome with its protein cofactors is important for gene regulation (Newton and Holden, 2007). Here we used a high-resolution derivative of ChIP-seq, ChIP-exo, to elucidate the different binding configurations of the GR across the human genome. Although site-specific TF configurations were difficult to discern
with this method, we arrived at the following conclusions: 1) The GR binds as both a dimer and a monomer across the genome. It is unclear, however, if our evidence for monomeric binding is merely the result of incomplete cross-linking of the GR dimer; 2) We observed evidence for billboard-like GR binding with AP-1, as well as stricter, grammar-like binding with proteins such as FOXA1 and TEAD1; and 3) We observed evidence for a GR-FOXA1 composite or tethered binding configuration.

Although we arrived at the abovementioned findings, we found it difficult to elucidate GR binding configurations at individual GR binding sites. The ChIP-exo signal at specific GR binding sites was often complex, without single, clear upstream and downstream boundaries. As a result, we relied heavily on the aggregate analysis of many GR binding sites. However, theses analyses were also difficult to design without a prior knowledge of: 1) The specific binding site motif at which to center the footprinting analysis; and 2) Which specific set of binding sites were candidates for said binding configuration.

Altogether we concluded that this approach may not be entirely suitable for the de novo elucidation of novel or unknown binding configurations of proteins at enhancer elements. However, with improved methods, and a more suitable study design, this type of approach may be appropriate for such a task. For example, a recently developed variant of the ChIP-exo method, ChIP-NEXUS (He et al., 2015), may yield better results. This method utilizes a more efficient adapter ligation step, as well as molecular
barcodes, increasing the overall complexity and signal compared to ChIP-exo. Also, as we found with FOXA1, ChIP’ing multiple transcription factors may further aid the elucidation of the multi-protein binding configurations. Lastly, we would anticipate that using ChIP-exo in the context of a perturbation to the protein structure itself may also be a viable approach. For example, in a recent study the authors performed ChIP-exo on a GR mutant thought to have a reduced ability to dimerize, GR\textsuperscript{dim} (Lim et al., 2015, Reichardt et al., 1998). With this approach, they found GR\textsuperscript{dim} did indeed have reduced binding to dimeric sites, and favored monomeric binding at putative composite sites with other cofactor proteins.

2.4 Methods

2.4.1 Experimental

2.4.1.1 Cell culture

A single seed stock of A549s was first plated into a 15 cm dish and grown under standard culture conditions using Ham's F-12K (Kaighn's) Medium, 10% FBS, 1% penicillin-streptomycin. Once confluence was reached (approximately 25 M cells), 500 \(\mu\)M dex in 100% EtOH was added to a final concentration of 100 nM for a 1 hour. As a negative control, another plate was treated with an equivalent volume of 100% EtOH for 1 hour. However, these vehicle control samples did not yield a detectable GR ChIP-exo sequencing library.
At the end of the 1 hour incubation, cells were treated with 1% formaldehyde for 10 minutes, and subsequently treated with 0.125 M glycine for 5 minutes. Media was aspirated, and cells were washed with an equivalent volume of cold 1X PBS. Next, 5 mL of Farnham Lysis Buffer (5 mM PIPES pH 8.0, 85 mM KCl, 0.5% NP-40, with a Roche Protease Inhibitor Cocktail Tablet) was added to the plate, and cells were scraped into 15 mL conical tubes. After centrifugation at 2,000 rpm for 5 minutes at 4°C, the buffer was aspirated, and cells were either snap frozen with in a dry-ice bath and stored at -80°C, or continued through to the sonication step.

2.4.1.2 ChIP-exo

Pellets were resuspended in 300 µL of RIPA buffer (1X PBS, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, filtered, with a Roche Protease Inhibitor Cocktail Tablet). Samples were sonicated with a Diagenode Bioruptor Twin, at high power, for a total time of 45 minutes (cycles of 30 seconds on, 30 seconds off). Samples were then centrifuged at 13,000 rpm for 15 minutes at 4°C. The supernatant was transferred to a new tube, and the total volume was brought to 1 mL with additional RIPA buffer.

Prior to immunoprecipitation, 200 µL of M-280 Sheep Anti-Rabbit or Sheep Anti-Mouse IgG Dynabeads (Thermo Fisher Scientific, 11203D or 11201D) were washed 3 times with 1 mL of 1X PBS (5 mg/mL BSA). Beads were resuspended in 1 mL of 1X PBS (5 mg/mL BSA). Next, either 5 µg of rabbit polyclonal GR antibody (Santa Cruz, sc-
1003x) or 5 µg FOXA1 mouse monoclonal antibody (Santa Cruz, sc-101058) was added to the beads, and incubated with rotation, overnight at 4°C. Next, beads were again washed 3 times with 1 mL of 1X PBS (5 mg/mL BSA) and resuspended in 100 µL.

Chromatin (in 1 mL of RIPA buffer) was then added to the antibody-beads, and incubated at 4°C overnight, with rotation. Beads were next washed 5 times at 4°C with 1 mL of LiCl wash buffer (100 mM Tris pH7.5, 500 mM LiCl, 1% NP-40, 1% sodium deoxycholate, filtered). Beads were then washed with 1 mL of 1X TE buffer, and resuspended in 200 µL of IP Elution buffer (1% SDS, 0.1 M NaHCO₃, filtered). Next, beads were incubated at 65°C for 1 hour, with vortexing every 15 minutes. They were then centrifuged at 13,000 rpm for 3 minutes, and the supernatant was removed and the ChIP’d DNA was purified with a Qiagen MinElute Kit.

ChIP-exo library construction was performed as described in the literature with a few modifications (Rhee and Pugh, 2012): 1) At the first adapter ligation step, an adapter with a 5’-phosphate on both strands was used to increased λ exonuclease activity; 2) At the second strand synthesis step (via Φ29 polymerase) a 5’-biotinylated primer was used, such that the resulting dsDNA could be captured with streptavidin beads, on which all subsequent reactions would be performed; and 3) An exonuclease T treatment was performed after the second strand synthesis to remove any free primer. Libraries were sequenced using pair-end 25 bp reads on an Illumina MiSeq instrument.
2.4.2 Computational

2.4.2.1 ChIP-exo data processing

After read quality assessment via FastQC (Andrews, 2010), read pairs were aligned with Bowtie 2 (Langmead and Salzberg, 2012) with the options: --no-discordant - -no-mixed --no-contain --no-dovetail -X 700. Next, potential PCR duplicates were removed with SAMtools’ “rmdup” (Li et al., 2009), and ENCODE’s blacklisted regions were removed (Consortium, 2012). Finally, alignments were split into separate files according to strand, and only the pile-ups of the 5’ end of each read were used for aggregate plots and footprinting analyses.

2.4.2.2 Binding site discovery via GEM

To elucidate GR binding sites and putative cofactor motifs, we used GEM (Guo et al., 2012), with the following options: --k_min 6 --k_max 21 -nrf.

2.4.2.3 2D Footprinting

Interrogating the 5’-read pile-ups on either the positive or negative strand at GR binding sites was informative for discerning the boundaries of DNA-protein complexes. However, when using this approach, it is unclear whether these boundaries exist in the same DNA-protein complexes, or are the result of heterogeneity of TF binding. To
address this problem, we sought to pair boundary information from both the positive and negative strands at each GR binding site to perform a type of TF footprinting.

To achieve this goal, we performed the following analysis steps: 1) Find the best match to the GR motif at each GR ChIP-seq peak using the MEME Suite (Bailey et al., 2006); 2) Find the 5′-read pile-ups of both the positive and negative strand at these regions, extending +/- 50 bps from the best GR motif; 3) Perform kernel density estimation (using a Gaussian kernel) for these 5′-read pile-ups, on both the positive and negative strand; 4) Multiply these vectors of probabilities to construct a matrix, or 2D footprint; and 5) Average these 2D footprints across all GR binding sites to determine the aggregate binding configurations of the GR.
3. Glucocorticoid treatment mediates chromatin loop dynamics concordant with gene expression responses

3.1 Introduction

Cells employ a variety of mechanisms to rapidly respond to environmental stimuli. Oftentimes, a core component of such responses involves modulating gene-regulatory mechanisms. A common element of human gene regulation involves transcription factors (TFs) binding to distal regulatory elements that physically interact with their target gene’s promoter to form chromatin loops. At steady-state, there are as many as millions of chromatin loops in the human genome (Jin et al., 2013). Higher-order chromatin topology is reorganized during cellular differentiation, and chromatin loops often connect distal regulatory elements to target gene promoters in a cell-type specific manner (Hakim et al., 2009, Rao et al., 2014, Dixon et al., 2015, Phanstiel et al., 2017). Furthermore, studies creating, inhibiting and deleting the anchors of loops demonstrate that those loops may have a mechanistic role in steady-state gene regulation (Deng et al., 2012, Li et al., 2013, Narendra et al., 2015).

The role of chromatin looping in response to short-term environmental signals is still unclear. To investigate how such responses affect enhancer-promoter interactions, we focused on a paradigmatic human gene regulatory system, the transcriptional response to glucocorticoid (GC) steroid hormones. Absent GCs, the glucocorticoid

---

1 This chapter has been adapted from a primary-author manuscript that is currently in review (D'Ippolito et al., In review)
receptor (GR) resides in the cytoplasm bound to chaperone proteins (Sanchez et al., 1990). Once stimulated by GCs, the GR translocates to the nucleus, binds the genome at thousands of locations, and regulates the expression of hundreds of genes (Reddy et al., 2009). Therefore, using GCs as a model environmental stimulus allows for precise temporal control over TF binding and the subsequent effects on chromatin looping.

Previous studies have shown that where the GR binds in response to GCs is largely predetermined by the binding of other TFs and chromatin accessibility prior to exposure (Biddie et al., 2011, John et al., 2011). It is currently unclear whether this same preprogramming is observed at the level of chromatin looping. Several studies have investigated how genome topology responds to GC treatment. However, global responses to GCs have not yet been assayed at the resolution of individual chromatin loops. Chromatin looping between enhancers and promoters at a few genes regulated by GCs were found to be pre-established (Hakim et al., 2011, Stavreva et al., 2015). In another study focused specifically on EP300 binding sites, the authors also observed pre-established loops, which increased looping upon GC treatment (Kuznetsova et al., 2015). Furthermore, genome-wide studies describing the effects of progesterone and estrogen treatment showed broad changes in interactions within topologically associating domains (TADs) and chromosome compartments, respectively (Le Dily et al., 2014, Mourad et al., 2014). These studies, however, did not resolve changes in individual loops.
To resolve the question as to whether enhancer-promoter interactions at latent GR binding sites exist prior to hormone treatment, or are instead formed in response to hormone treatment, we performed high-resolution in situ Hi-C across a 12-hour time course of GC treatment. Unlike previous studies, this approach allowed us to assay genome topology at the scale of chromatin loops, in a genome-wide, TF-agnostic manner. Overall, we found that GCs cause changes in the interaction frequencies of loops that exist prior to treatment rather than formation of entirely new loops. We observed that induced genes increase interactions with distal GR binding sites in response to GC treatment. Targeting those loops with the programmable heterochromatin-inducing factor dCas9KRAB reduced both baseline and dex-mediated induction of expression, suggesting that the pre-existing loops contribute functionally to both processes. At repressed genes, we also observed looping to distal GR binding sites. While loops at GC-repressed genes were generally unresponsive, we observed modest increases in looping at those genes with distal GR binding sites. We also found that both induced and repressed genes were involved in changes in chromosome compartmentalization concordant with their expression response. Together, these findings support a model in which the cellular response to GCs is pre-wired in chromatin interactions, and in which GR binding at putative enhancers is indicative of both gene activation and repression.
3.2 Results

3.2.1 Thousands of loops change their interaction frequency in response to GC treatment

To measure changes in chromatin interactions at high-resolution, in a TF-agnostic fashion, we performed in situ Hi-C in human A549 cells after 0, 1, 4, 8, and 12 hours of treatment with 100 nM dexamethasone (dex), a GC highly specific to the GR (Reul et al., 2000) (Figure 1A). We then called loops using Hi-C Computational Unbiased Peak Search (HiCCUPS) (Rao et al., 2014, Durand et al., 2016b) for each time point, binning the genome into both 5 kb and 10 kb windows. We found
Figure 11: Thousands of loops change their interaction frequency in response to GC treatment

(A) Overview of experimental design. (B, C) Dynamics of chromatin loops across the time course at loci near ZFP36, and near DIO3OS, respectively. Contact maps represent the log₂ fold-change in interaction frequency between each time point and time point 0. Dynamic loops have been annotated with circles. Dex-induced genes and dex-repressed genes are colored in blue and red, respectively. (D) Dynamics of chromatin loops across the time course at the ZFP36 locus. (E) Dynamics of chromatin loops across the time course at the DIO3OS locus. (F) The total number of chromatin loops that either increase, or decrease their interactions at each time point. Solid lines indicate the total number of dynamic loops, and dashed lines represent the number of
dynamic loops not previously observed. (G) The mean log$_2$ fold-change in interactions at
dex-increased, dex-decreased, and static loop coordinates, +/- 25 kb. Intensities represent
the mean log$_2$ fold-change across all regions of that loop class. (H) Interaction
frequencies of chromatin loops at time point 0.

48,422 unique chromatin loops across the entire time course. These interactions were
reproducible across replicates ($r \geq 0.7$; Appx. A; Figure 17A). After filtering for low read
counts across replicates, we tested 41,467 loops for differential interactions in response
to dex treatment. Across the time course, dex significantly increased the interaction
frequencies of 1,491 (3.60%) loops (Figures 11B and 11D, and Appx. A; Figure 17B) and
significantly decreased the interaction frequencies of 1,744 (4.21%) loops (Figures 11C
and 11E, and Appx. A; Figure 17C) at a false discovery rate (FDR) $\leq 5\%$.

Compared to the pre-dex condition, most differential interactions were observed
after 1 hour of dex treatment, with 857 increased and 960 decreased loops. The fewest
differential interactions were observed after 12 hours of dex treatment, with 170
increased and 137 decreased loops, suggesting that cells are returning to steady state
near the end of the time course (Figure 11F). This trend in looping dynamics is
independent of the FDR cutoff used to call differential loops (Appx. A; Figure 17D). The
distribution of fold changes for both increased and decreased loops were similar in
magnitude (Appx. A; Figure 17E).

To determine whether these dynamic interactions were constrained to specific
foci, or were instead representative of broader changes in chromatin interactions, we
examined the dynamics of interactions +/- 25 kb from the loop coordinates (Figure 11G). Overall, we found that the observed looping dynamics were highly specific to the loop coordinates compared to nearby interaction bins (Appx. A; Figure 17F).

3.2.2 Dex-responsive loops are pre-existing

Previous studies of a few GC-regulated genes suggest that chromatin loops exist prior to hormone treatment, and are not formed de novo (Hakim et al., 2009, Stavreva et al., 2015). To test that model genome-wide, we compared pre-dex loop interaction frequencies between static, and latent increased and decreased loops using linear regression (Appx. A; Figure 17G). Increased loops had slightly lower interaction frequencies relative to static loops before treatment ($\beta = -9.8 \times 10^{-2}$, FDR = $4.4 \times 10^{-4}$; Figure 11H and Appx. A; Figure 17G). This was particularly true for loops that increased interactions with longer dex treatment ($\beta = -2.5 \times 10^{-2}$, FDR = $7.4 \times 10^{-8}$). The pre-dex interaction frequencies of increased loops, however, were still much higher than those of randomly chosen non-loop interactions ($\beta_{\text{Increased}} = -9.8 \times 10^{-2}$ vs. $\beta_{\text{Non-loop}} = -3.3$). In contrast, decreased loops had interaction frequencies higher than static loops before treatment ($\beta = 7.5 \times 10^{-1}$, FDR < $1 \times 10^{-100}$). These results suggest that dex treatment does not result in de novo loop formation, but instead results in quantitative changes in the interaction frequencies of pre-established loops.
3.2.3 Networks of chromatin loops have coordinated responses to dex treatment

Previous studies have found that chromatin loops frequently share anchor points to form networks of interactions (Sanyal et al., 2012). We noticed that dex-responsive chromatin loops were often found within the same networks of interactions (Figures 12A-C, and Appx. A; Figure 18A). When defining networks of interactions using only dex-responsive loops, we found that the resulting network sizes were larger than would be expected by chance \( P = 3.1 \times 10^{-10} \), Kolmogorov-Smirnov (KS) test; Figure 12D]. Furthermore, the networks of dex-responsive loops generally change interactions concordantly, either all increasing or all decreasing their interaction frequencies (Figure 12E, and Figures 18B-D). Together, these results suggest that dex-mediated changes in chromatin looping are highly coordinated among clusters of genomic loci.
Figure 12: Networks of chromatin loops have coordinated responses to dex treatment

(A) Dex-responsive chromatin interaction dynamics at a locus near DUSP1. Contact maps have been annotated as in Figures 11B-C. (B) Interaction dynamics of chromatin loops at the DUSP1 locus annotated in Figure 12A. (C) Gene expression dynamics of dex-responsive genes at the DUSP1 locus. (D) Distribution of network sizes for dynamic loops (plus a pseudo-count of 1 for each column). (E) Mean vs. variance of log₂ fold-changes in interaction frequencies of loops in dex-responsive networks after 1 h dex treatment.
3.2.4 GC-induced genes are enriched for increased looping interactions

Previous studies of chromatin interactions in response to hormones have shown that hormone-induced genes increase interactions with distal hormone-receptor binding sites (Fullwood et al., 2009, Hakim et al., 2011, Kuznetsova et al., 2015). Based on those studies, we hypothesized that dex-induced genes would increase interactions with distal regulatory elements, while dex-repressed genes would decrease interactions with distal regulatory elements. To test that hypothesis, we used RNA-seq to measure gene expression after 0, 1, 4, 8, and 12 hours of dex exposure (McDowell, In review), and then compared changes in gene expression to changes in loop interaction frequencies. Both induced and repressed genes were positioned at more chromatin loop anchors than genes not regulated by dex (Appx. A; Figure 19A). Induced genes were significantly enriched for anchors of increased loops at all time points (Figure 13A) and were significantly depleted for anchors of decreased loops at several time points (Appx. A; Figure 19B). Induced genes were also enriched for anchors of increased loops at their transcription start and end sites (TSSs and TESs, respectively; Appx. A; Figure 19C). Moreover, we found that dex-induced genes were enriched for larger networks of increased interactions than expected by chance ($P = 0.01$, KS test; Figure 13B). These results are in line with our initial hypothesis that increased interactions are associated

2 This is a manuscript I co-authored as part of my graduate studies.
with increased gene expression genome wide, and further support recent findings that multiple TF binding sites may interact together to synergistically activate gene expression (Vockley et al., 2016).

Contrary to our hypothesis that repressed genes would involve decreases in chromatin looping, repressed genes were not significantly enriched for the anchors of decreased loops, nor those of increased loops (Appx. A; Figures 19D-E). We hypothesized that repressed genes may not be positioned at the anchors of decreased chromatin loops, but may localize within the intervening DNA of decreased loops. However, we found no significant enrichment of dex-repressed genes within decreased loops (Appx. A; Figure 19F). Furthermore, dex-repressed genes were depleted within increased loops (Appx. A; Figure 19G). These findings indicate that decreased loops may not directly associate with dex-dependent gene repression.

To empirically test if the anchors of dynamic loops contribute to gene regulation and GC-induction, we targeted a synthetic heterochromatin-inducing factor and transcriptional repressor, dCas9\textsuperscript{KRAB} (Gilbert et al., 2013), to the anchors of increased loops at dex-induced genes. The KRAB domain recruits a histone methyltransferase complex that methylates H3K9 and blocks enhancer activity. Therefore, targeting the KRAB domain to loop anchors via fusion to the dCas9 protein allows for site-specific
blocking of potential enhancer activity (Thakore et al., 2015). We targeted loop anchors near the dex-induced genes MT1A, ARL4A, and TCEAL1 (Appx. A; Figures 19H-J), which had increased interaction frequencies after 1 hour of dex-treatment (Appx. A; Figures 19C-E). For each target anchor, we assayed gene expression after 0 or 1 hour of 100 nM dex treatment. In all of loci that we tested, recruiting dCas9KRAB to these anchor regions decreased baseline expression of one or more adjacent genes (Figures 13F-K and Appx. A; Figures 19K-M; left panels). To test if dCas9KRAB decreased the fold-change in expression mediated by dex treatment, and not just absolute expression levels, we used linear regression with interaction terms modeling the effects of dCas9KRAB on dex-mediated gene induction (Appx. A; Figures 19N-P). Using this approach, we found that targeting dCas9KRAB to these loop anchors repressed dex-mediated induction of many genes at these loci (Figures 13F-K and Appx. A; Figures 19K-M; right panels). These results suggest that these interacting loci are important for maintaining gene expression levels both prior to, and in response to dex treatment.

---

3 This is a manuscript I co-authored as part of my graduate studies.
Figure 13: Dex-induced genes are enriched for increased looping

(A) Number of dex-induced genes at the anchors of dex-increased looping (*: $P \leq 5\%$; n.s.: $P > 5\%$). (B) The distributions of network sizes for all increased loops, and for increased loops at dex-induced genes. (C-E) dCas9KRAB was targeted to the anchors of increased loops near MT2A, ARL4A and TCEAL1, respectively. gRNA target sites for either "A" or "B" anchor of increased loops have been labeled under the contact map.
Genes whose expression was assayed via RT-qPCR have been labeled below the gene annotation track. (F-K) Gene expression of genes near targets of dCas9KRAB from Figures 13C-E, as measured by RT-qPCR. Gene expression at 0 h was compared to the negative control (no gRNA; gray) and gene expression after 1 h dex treatment was compared to 0 h dex for each target (blue). Bar plots designate mean log2 fold-change and error-bars designate standard error of the mean. Significance of gene expression changes was determined via linear regression (*: FDR ≤ 10%, n.s.: FDR > 10%).

3.2.5 Chromatin looping dynamics are associated with low CTCF occupancy

The DNA-binding protein CTCF has been previously shown to mediate chromatin loop formation (Ong and Corces, 2014, Nora et al., 2017). As expected, CTCF binds most loop anchors – 82% with an FDR < 1% – in A549 cells (Figure 14A) (McDowell, In review). Compared to the GR and other TFs, CTCF binding is stable in response to dex (McDowell, In review). One possibility is that CTCF maintains loops that do not change in response to dex, whereas other TFs regulate loops that are dynamic in response to dex. Based on that model, we hypothesized that loops that are more dynamic in response to dex have less CTCF and more binding of other TFs. We found that larger increases or decreases in looping were associated with lower CTCF occupancy at loop anchors (Figure 14B and Appx. A; Figures 20A-C). We also found that loops with less CTCF bound to their anchors were more often bound by the GR and other TFs (Figures 14C-D). For example, a dex-increased loop at the dex-induced gene EDN3 has low CTCF ChIP-seq signal, but high GR signal at its anchors (Figures 14E-G and Figure 20D). Furthermore, loops with evidence of CTCF at only one loop anchor,
with the exception of JUN, had higher occupancy of TFs at the other non-CTCF-bound anchor (Figures 14H-I). An example of such asymmetric CTCF occupancy can be seen at the dex-induced gene CEBPD (Figures 14J-L and Figure 20E). Together, these results show that although loops are highly occupied by CTCF, loops that change interaction frequencies in response to dex are associated with binding of other TFs.
Figure 14: Chromatin looping dynamics are associated with low CTCF occupancy

(A) Proportion of loops with CTCF binding sites within their anchor regions. (B) Change in interaction frequency of loops vs. mean CTCF occupancy at loop anchors after 1 h dex treatment. (C) Mean normalized ChIP-seq signal of loop anchors for the GR vs. CTCF. (D) Coefficients from regression models testing the association of TF occupancy at loop anchors with the level of CTCF binding at loop anchors (n.s.: FDR > 5%). (E) A loop at the EDN3 locus with low CTCF occupancy and high GR occupancy is increased in response to dex treatment. (F) The dex-mediated interaction dynamics of the loop annotated in Figure 14E. (G) The dex-mediated gene-expression dynamics of EDN3. (H) GR ChIP-seq signal at loop anchors for loops with evidence of only one anchor bound by CTCF. (I) Coefficients from regression models testing the association of TF occupancy at loop anchors bound vs. not bound by CTCF, for loops with only one anchor bound by CTCF (n.s.: FDR > 5%). (J) A loop at the CEBPD locus with high CTCF occupancy at one anchor and high GR occupancy at the opposite anchor is increased in response to dex treatment. (K) The dex-mediated interaction dynamics of the loops annotated in Figure 14J. (L) The dex-mediated gene-expression dynamics of CEBPD.
3.2.6 Increased and decreased loops are associated with chromatin state and TF binding

We next evaluated whether chromatin state or TF occupancy at loop anchors before and after dex treatment corresponded to changes in loop interactions. We did so by integrating our Hi-C data with numerous complementary genomic datasets, TF binding motifs, and sequence features such as GC content, CpG density, and G-quadruplexes (G4) (Pfeifer et al., 1996, Lam et al., 2013, White et al., 2013, McDowell, In review). The presence of the activation-associated histone marks H3K27ac, H3K4me1, and H3K4me2 prior to dex treatment positively correlated with increased looping interactions (Appx. A; Figure 21A). Similarly, occupancy of GR cofactors prior to dex treatment and motifs for GR-associated TFs were also associated with increased looping. In contrast, pre-bound CTCF and GC content inversely correlated with changes in looping interactions. Increases in activation-associated histone marks and TF occupancy after dex treatment were also associated with increased looping interactions. These associations were much stronger than those found for the same marks, pre-dex treatment. Loops that decreased interactions in response to dex had largely the opposite trends, but weaker associations overall. These interactions were depleted for activation-associated histone marks and TF binding prior to dex treatment and were enriched for high GC content sequence features (Appx. A; Figure 21B). Together, these findings indicate that latent, dex-increased loops are enriched for activation-associated chromatin features that further increase with dex treatment. Meanwhile, loops that decrease
interactions are less strongly associated with chromatin states and DNA sequence features that we studied.

3.2.7 GR binding at distal enhancers is associated with gene activation and repression

Determining which chromatin features correspond to dex-mediated gene expression responses is typically challenging without a priori knowledge of gene-enhancer relationships. Using our high-resolution interaction maps, we were able to assign distal regulatory elements to dex-regulated genes, enabling us to evaluate which distal genomic features associate with gene activation and repression (Figure 15A). Prior to dex treatment, we observed stronger enrichment for activation-associated histone marks, TF binding, and TF motifs at the distal enhancers of induced genes than at the genes themselves (Figure 15B). After dex treatment, TF occupancy and activation-associated histone marks increased at both the gene-distal and gene-proximal anchors of those loops. For example, sites looped to the dex-induced gene FOSL2 gain GR, EP300, JUN and H3K27ac (Figure 15D and Appx. A; Figure 21C). Overall, dynamics in the binding of AP-1 components, such as JUN, had the highest association with gene induction. Although increases in activation-associated histone marks such as H3K27ac and H3K4 methylation also associated with gene induction, the effect was most prominent at later time points. Dynamic CTCF occupancy was the least predictive of gene expression responses, consistent with our observation that dex does not influence
CTCF occupancy over the time course we studied (McDowell, In review). These results support a model where TFs act across chromatin loops to mediate gene induction.

Figure 15: GR binding at distal enhancers is associated with gene activation and repression

(A) Genomic features at loop anchors proximal (overlapping the gene body) or distal to dex-regulated genes were used to predict dex-mediated gene expression responses. (B, C) Coefficient values from elastic net logistic regressions using genomic features of loop anchors to predict whether a gene is induced or repressed by dex, respectively. (D) The dex-induced genes, FOSL2 and RP11-373D23.3 interact with distal loci that exhibit increases in GR, EP300, JUN, and CEBPB occupancy. (E) The dex-
repressed gene, \textit{ARRDC3}, interacts with distal loci that exhibit increases in GR and HES2, and decreases in JUN and H3K27ac. (F) The average change in ChIP-seq signal at anchors looped to dexamethasone-induced genes. Anchors have been subset based on evidence of GR occupancy. (G) Motif enrichments at GR binding sites looped to repressed genes. Motif occurrences were compared to those found using shuffled sequences (Fisher’s exact test; * : FDR ≤ 5%). (H) The change in the interaction frequency of loops at dexamethasone-repressed genes after 1 hour of dexamethasone treatment. Loops have been subset by GR occupancy, and distributions were compared by a Wilcoxon Rank-sum test (*: \( P \leq 0.05 \)).

For dexamethasone-repressed genes, gene-proximal regions and distal enhancers were also enriched for TF binding and activation-associated histone marks prior to dexamethasone treatment (Figure 15C). After dexamethasone treatment, distal enhancers gained GR and HES2 occupancy, and both the proximal and distal regions lost AP-1 and H3K27ac. However, when surveying specific examples of dexamethasone-repressed genes, we noticed that these dynamics in chromatin marks do not occur at the same loop anchors. For example, at the \textit{ARRDC3} locus, the distal loci that gain GR binding exhibit modest increases in JUN, while those with weak GR binding lose JUN occupancy, as well as H3K27ac (Figure 15E and Appx. A; Figure 21D). These trends are representative of all GR binding sites looped to repressed genes (Figure 15F). Those anchors with evidence of GR binding generally gained AP-1, while those without GR binding lost AP-1 binding. Both types of loop anchors lost activation-associated histone marks in response to dexamethasone. This increase in AP-1 binding does not appear to be a result of unintentionally assaying anchors that are also looped to nearby dexamethasone-induced genes. Such anchors generally gain both TF occupancy and active histone marks (Appx. A; Figure 21F). Previous studies have proposed a model where the GR
interacts with genome-bound AP-1 via protein-protein interactions to repress transcription (Yang-Yen et al., 1990). To help discern whether these GR binding sites are the result of such tethering or are instead motif-driven binding, we performed motif searching on GR binding sites loops to repressed genes (Figure 15G). We found that the AP-1 motif, as well as the FOXA1 motif, were significantly enriched compared to shuffled controls. Furthermore, these motifs were more highly enriched at GR sites looped to repressed genes than those sites without evidence of GR binding (Appx. A; Figure 21G). We did not find evidence for enrichment of the GR motif in either of these analyses. Lastly, although we were unable to find an enrichment of dynamic looping at repressed genes (Appx. A; Figures 19D-G), we did observe that those loops bound by the GR at repressed genes had modest increases in looping after dex treatment (Figure 15H and Appx. A; Figure 21H). Loops at repressed genes not bound by the GR were largely unaffected by dex treatment. Overall, these results support previous evidence of GR’s function as a transcriptional repressor, and provide new evidence that such repressive GR binding sites may interact via chromatin looping. In addition, these findings demonstrate that chromatin dynamics at the enhancers of dex-repressed genes are diverse, and can be discriminated by their level of GR occupancy.
3.2.8 GC-regulated genes change chromosome compartmentalization

Previous studies have shown that chromosomes are organized into two primary compartments, termed A and B, that are enriched or depleted for evidence of active transcription, respectively (Lieberman-Aiden et al., 2009). We hypothesized that dex treatment may result in dex-induced genes interacting more with transcriptionally-active compartments, while dex-repressed genes would interact more with transcriptionally-inactive compartments. To test this hypothesis, we first measured compartmentalization by performing principal component (PC) analysis on the interaction matrices, and then examining the values of the first PC (PC1). Overall, we did not find evidence of complete switches in compartment classification (Figure 16A), as is observed during cellular differentiation (Dixon et al., 2015). We did, however, observe quantitative changes in compartment associations for 4% of all 50 kb windows genome-wide (Figures 16B-C). These dynamic regions were more likely than static regions to reside in compartment A prior to dex treatment (Figure 16D and Appx. A; Figure 22A). Overall, 87% became more B-like (decreased PC1 value) after 4 hours of dex treatment (Figure 16C). These regions also increased interactions with compartment B according to a linear regression model ($\beta = 3.8 \times 10^{-2}$, $P < 1 \times 10^{-100}$; Figure 22B). The regions that dex caused to become more B-like were enriched for dex-repressed genes compared to static regions (Figure 16E). Conversely, the 13% of dynamic regions that became more A-like (increased PC1 value) increased interactions with compartment A.
(β = 3.3 x 10^{-2}, P < 1 x 10^{-100}), decreased interactions with compartment B (β = -3.3 x 10^{-2}, P < 1 x 10^{-100}), and were enriched for dex-induced genes (Figure 16F, and Appx. A; Figure 22C). These findings are reproducible using different resolutions of chromatin interactions (Appx. A; Figures 22D-I). Together, these results show that dex mediates dynamics in higher-order chromatin topology, and that those changes are associated with both gene repression and gene activation.

Figure 16: Dex treatment mediates changes in chromatin compartmentalization

(A) A/B compartment classification of chromosome 1. The first PC was oriented such that positive values positively correlate with gene expression. Values were averaged across biological replicates for each time point. (B) Variation in PC1 values across the dex time course at a locus on chromosome 1. Variation with the duration of dex treatment was evaluated via an ANOVA, and regions with an FDR ≤ 5% are shaded red. (C) Regions with dynamic PC1 values. The PC1 values of each dynamic region in the genome were averaged across biological replicates, standardized across time points, and then hierarchically clustered. (D) Non-standardized PC1 values of the dynamic regions shown in Figure 5C. (E) Number of dex-repressed genes within a given
compartment type (*: \( P \leq 0.05 \); n.s.: \( P > 0.05 \)). \( \textbf{F} \) Number of dex-induced genes within a given compartment type (*: \( P \leq 0.05 \); n.s.: \( P > 0.05 \)).

### 3.3 Discussion

Previous studies have shown that GR binding sites are largely pre-determined by other TF binding sites or by DNase accessibility (Biddie et al., 2011, John et al., 2011). In addition, many focused studies have observed examples of pre-existing chromatin loops between latent GR binding sites and promoters. Our work generalizes that model to chromatin looping across the genome. As a consequence of pre-existing loop structure, TF binding at latent GR binding sites likely contributes to baseline expression of GC-target genes. That finding could explain previous observations that GC-induced genes are typically already expressed prior to hormone treatment (Reddy et al., 2009).

Our findings agree with previous studies of the hormone-mediated dynamics of individual chromatin loops, suggesting that our findings may extend to other nuclear receptors (Fullwood et al., 2009, Hakim et al., 2011, Kuznetsova et al., 2015). However, our results contradict previously published results on the dynamics of chromatin looping in response to TNF-\( \alpha \) treatment (Jin et al., 2013). That study observes pre-existing chromatin topology, but they do not detect significant changes in looping after TNF-\( \alpha \) treatment. It is unclear if this discrepancy is due to technical differences between \textit{in situ} Hi-C and in-solution Hi-C assays, or biological differences between glucocorticoid and TNF-\( \alpha \) responses.
Why loops with decreased interactions are not associated with changes in gene expression also remains to be understood. One intriguing explanation is that decreased loops result from GC-dependent inhibition of DNA replication. GCs are well known to decrease cellular proliferation in certain cellular contexts, and previous studies have shown that DNA replication timing is correlated with genome topology (Pope et al., 2014, Hofmann et al., 1995). Indeed, many of the canonical genes involved in DNA replication are transcriptionally repressed over the dex exposure time course studied here (Table S9).

Our observations suggest that stimulus-responsive TFs such as the GR may mediate dynamic chromatin looping, while CTCF is required for the maintenance of basal chromatin interactions. In a recently developed model of chromatin loop formation, the molecular motor cohesin extrudes loops through its ring structure until it is blocked by genome-bound CTCF (Sanborn et al., 2015, Fudenberg et al., 2016). If this model is borne out as a primary mechanism of loop formation, our findings would suggest that binding of stimulus responsive TFs such as the GR may also function to block cohesin and thus stabilize chromatin looping. Two recent studies show that the inducible degradation of either CTCF or cohesin renders most chromatin looping undetectable (Nora et al., 2017, Rao et al., 2017). Notably, neither of these perturbations have large effects on gene expression. In light of these findings, it is interesting to speculate whether the same effects would be observed in the context of an
environmental stimulus, such as GC treatment. For example, one function of chromatin loop formation may be to dynamically sample the regulatory landscape so that new signals are properly integrated into transcriptional outputs. Therefore, loss of cohesin may ablate a GC transcriptional response due to an inadequate exposure of new GR binding sites to their target genes. Conversely, loss of CTCF may have less dramatic effects, given that dynamic loops are generally depleted for CTCF occupancy.

Additional experiments perturbing putative loop-forming machinery will help elucidate the mechanisms involved in chromatin looping dynamics in response to environmental stimuli.

The mechanisms of GC-mediated repression have remained elusive. Several studies have proposed a transrepressive model in which GR binding to DNA-bound factors such as AP-1 repress transcriptional activity (Yang-Yen et al., 1990). Efforts to identify transrepressive sites via ChIP-seq have been obfuscated by the observation that GR binding is depleted around repressed genes (Reddy et al., 2009, Langlais et al., 2012). Our study helps to resolve that apparent conflict by showing that there are indeed long-range loops between the promoters of GC repressed genes and distal GR binding sites.

Early efforts to describe the specific mechanism of transrepression provided evidence that GR binding reduces occupancy of AP-1 (Yang-Yen et al., 1990). However, later studies demonstrated that AP-1 occupancy is unaffected by GR binding (Rogatsky et al., 2001). Our results show that both of these events may occur at transrepressed regions,
and that AP-1 occupancy actually increases at those sites with high GR occupancy. Furthermore, our findings demonstrate that GR binding sites at repressed genes interact more frequently in response to dex treatment. It is unclear how these different genomic events may be coordinated to produce a repressive response. Some studies of repression by steroid hormones provide evidence for a squelching mechanism, whereby nuclear receptors at induced genes compete away coactivator proteins that are also used at repressed genes (Onate et al., 1995). One interpretation of our results could involve a hybrid of both the transrepressive and squelching models, whereby GR tethering stabilizes AP-1 binding, preventing its recruitment to other nearby enhancers. However, we did not observe a loss of looping between non-GR binding sites at repressed genes, indicating that such a mechanism would leave chromatin looping intact. Ultimately, additional studies are required to better understand how the chromatin dynamics observed at repressed genes relate to decreased transcription. Notably, we previously were unable to detect direct repression in reporter assays of isolated GR binding sites, suggesting that context beyond the DNA sequence of the GR binding site contributes to the loss in activity of enhancers of repressed genes (Vockley et al., 2016). Therefore, future investigation of putative repressive GR binding sites may benefit from perturbations at the endogenous loci.
3.4 Methods

3.4.1 Experimental

3.4.1.1 Cell culture

A single seed stock of A549s was first plated into a 500 cm\(^2\) dish and grown under standard culture conditions using Ham's F-12K (Kaighn's) Medium, 10% FBS, 1% penicillin-streptomycin. After expanding to 5 new 500 cm\(^2\) plates, cells were grown to confluence. Once confluence was reached, 500 µM dex in 100% EtOH was added to a final concentration of 100 nM to start the 12-hour time point. Dex was subsequently added to the 8, 4, and 1 hour time point plates such that all incubations would end at the same time.

3.4.1.2 In situ Hi-C

Library construction was performed as previously described with minor modifications (Rao et al., 2014).

Cells were fixed on the plate at a final concentration of 1% formaldehyde for 10 minutes at room temperature. Fixation was quenched with the addition of 2.5 M glycine to a final concentration 0.125 M for 5 minutes, at room temperature. After media aspiration, 15 mL of 1X PBS was added to the plate, and cells were scraped into a 50 mL conical tube. After centrifugation at 300 x g at 4°C for 5 minutes, the supernatant was discarded, and cell were resuspended in 10 mL of cold 1X PBS. Cells were aliquoted into
5 x 10⁶ cell equivalents (estimated from a count plate) into 1.5 mL microcentrifuge tubes, and then centrifuged at 300 x g at 4ºC for 5 minutes. The supernatant was discarded, and cells were snap-frozen in a dry-ice/ethanol bath. Aliquots were stored at -80ºC until the time of library construction. This procedure was repeated 4 different times using different A549 seed stocks.

Frozen cell pellets were thawed on ice and resuspended in a mixture of 250 µl ice-cold Hi-C lysis buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.2% Igepal CA630) and 50 µl of protease inhibitors (Sigma, P8340). After a 15-minute incubation on ice, cells were pelleted at 2,500 x g for 5 minutes. The supernatant was removed, the pellet was resuspended in 500 µl ice-cold Hi-C lysis buffer and pelleted again at 2,500 x g for 5 minutes. After removing the supernatant, the pellet was resuspended in 50 µl of 0.5% sodium dodecyl sulfate (SDS) and incubated at 62ºC for 10 minutes. To quench the SDS, 145 µl water and 25 µl 10% Triton X-100 was gently mixed with the suspension and incubated at 37ºC for 15 minutes. Next, 25 µl of 10X NEBuffer2 and 100 U of MboI restriction enzyme (NEB, R0147) were added to the sample and incubated overnight at 37ºC with rotation.

Samples were incubated at 62ºC for 20 minutes to inactivate MboI, then cooled to room temperature for 15 minutes. To fill in the restriction fragment overhangs and mark the DNA ends with biotin, we added 37.5 µl of 0.4 mM biotin-14-dATP (Life Technologies, 19524-016), 1.5 µl each of 10 mM dCTP, dGTP, and dTTP, and 8 µl of DNA
Polymerase I, Large (Klenow) Fragment (40 U; NEB M0210), and incubated at 37ºC for 45 minutes with rotation. To ligate free DNA ends, we then added 669 µl water, 120 µl 10X NEB T4 Ligase Buffer (NEB, B0202), 100 µl 10% Triton X-100, 6 µl 20 mg/mL BSA (NEB, XXX), 5 µl T4 DNA Ligase (NEB, B000S), and incubated at room temperature for 4 hours with slow rotation. After ligation was complete, we centrifuged samples at 2,500 x g for 5 minutes, discarded the supernatant, resuspended in 300 µl 1% SDS and 12.5 µl of 20 mg/mL Proteinase K (Thermo Fisher Scientific, EO0492), and incubated at 55ºC for 30 minutes. Next, we added 32.5 µl of 5 M sodium chloride and incubated at 68ºC overnight.

Samples were cooled to room temperature and mixed with 552 µl of 100% ethanol and 34.5 µl 3 M sodium acetate (pH 5.2). After incubation at -80ºC for 15 minutes, samples were centrifuged at 20,000 x g at 4ºC for 15 minutes. The supernatant was discarded and pellets were washed twice with 800 µl ice-cold 70% ethanol, centrifuging at 20,000 x g at 4ºC for 5 minutes. The DNA pellets were then resuspended in 140 µl 10 mM Tris-HCl (pH 8) and incubated at 37ºC for 15 minutes. DNA was then transferred to 130 µl Covaris microTUBE®es and sheared on a Covaris S2 machine (cycle: 10%, intensity: 4, cycles/burst: 200, time: 55 seconds). Sheared DNA was transferred to a new 1.5 mL tube. The microTUBE® was washed with 70 µl water, and then combined with the sheared DNA. Samples were then size-selected using Agencourt AMPure XP.
beads (Beckman Coulter, A63882) using a 2-sided selection (bead:sample ratios: 0.55/0.7). DNA was eluted off the beads with 300 µl of 10 mM Tris-HCl (pH 8).

First, 150 µl of 10 mg/mL Dynabeads MyOne Streptavidin T1 beads (Life technologies, 65602) were washed with 400 µl of 1X Tween Washing Buffer (1X TWB; 5 mM Tris-HCl (pH 7.5); 0.5 mM EDTA; 1 M NaCl; 0.05% Tween 20). Beads were then separated on a magnet, the supernatant was discarded, and the beads were resuspended in 300 µl of 2X binding buffer (10 mM Tris-HCl, pH 7.5; 1 mM EDTA; 2 M NaCl). Beads were then added to DNA samples, and incubated at room temperature for 15 minutes with rotation. After separation on a magnet, and removing the supernatant, beads were washed with 600 µl 1X TWB, transferred to a new tube, and heated to 55ºC for 2 minutes (with one inversion). After one additional wash, beads were resuspended in 100 µl 1X T4 DNA ligase buffer (NEB, B0202), and transferred to a new tube. Next, the supernatant was discarded, and beads were resuspended in 8.8 µl 10X T4 DNA ligase buffer, 2 µl 25 mM dNTPs, 5 µl T4 PNK (NEB, M0201), 4 µl T4 DNA polymerase I (NEB, M0203), 1 µl DNA polymerase I, Large (Klenow) Fragment (NEB, M02010), 79.2 µl water, and incubated at room temperature for 30 minutes.

The beads were then washed twice with 1X TWB as previously described, and resuspended in 100 µl 1X NEBuffer 2 (NEB, B7002S). Next, the supernatant was discarded, and beads were resuspended in 9 µl 10X NEBuffer 2, 5 µl 10 mM dATP, 5 µl Klenow exo⁻ (NEB, M0212), 81 µl water, and incubated at 37ºC for 30 minutes.
The beads were then washed twice with 1X TWB as previously described, and resuspended in 100 µl 1X Quick Ligation reaction buffer (NEB, B6058). Next, the supernatant was discarded, and beads were resuspended in 50 µl 1X Quick Ligation reaction buffer, 2 µl DNA Quick ligase (NEB, M2200), 3 µl of 15 uM Illumina TruSeq adapter (hybridized in-house), and incubated at room temperature for 15 minutes.

The beads were then washed twice with 1X TWB as previously described, and resuspended in 100 µl of 10 mM Trix-HCl (pH 8). Next, the supernatant was discarded, and beads were resuspended in 10 µl 5X Q5 buffer, 10 µl GC Enhancer, 1 µl 10 mM dNTPs, 1 µl 10 uM i5 primer, 1 µl 10 uM i7 primer, 26.5 µl water, and 0.5 µl Q5 polymerase (NEB, M0491). Libraries were amplified using the following conditions: 98ºC 30 seconds, 12 cycles of (98ºC 10 seconds, 65ºC 30 seconds, 72ºC 30 seconds), and 72ºC 2 minutes. Beads were separated on a magnet and the supernatant was saved in a new 1.5 mL tube. The beads were washed with 200 µl water, which was combined with PCR product. The PCR product was then size selected using Agencourt AMPure XP beads with a bead:sample volume ratio of 0.7. DNA was eluted in 35 µl of 10 mM Tris-HCl (pH 8.0), and quantified via Qubit dsDNA HS Assay (Thermo Fisher Scientific, Q32851) and High Sensitivity D1000 ScreenTape (Aglient, 5067-5585). The quality of libraries was further assessed via sequencing on an Illumina MiSeq, using 150 bp paired-end reads, before sequencing on a HiSeq4000 using 150 bp paired-end reads.
3.4.1.3 Targeted repression of chromatin loop anchors

Singleton chromatin loops whose interactions increased after 1 hour of dexamethasone (dex) exposure, and had anchors overlapping dex-induced genes (1 h dex), were first chosen for targeted repression (loops near MT2A, ARL4A, and TCEAL1). Four guide-RNAs (gRNAs) for each anchor of each loop were designed using the CRISPR Design Tool (http://crispr.mit.edu/) (Table 2). gRNAs were prioritized by specificity-score and proximity to GR binding sites or DNase hypersensitive sites (DHSs) within the loop anchor region. gRNAs were cloned into either phU6-gRNA, pmU6-gRNA, phH1-gRNA, or ph7SK-gRNA (Kabadi et al., 2014), then sequence confirmed by Sanger sequencing. For a negative control, transcriptional terminator sequences were cloned into each gRNA position. A destination vector for each 4-gRNA set was constructed using pLV hUbC-dCas9 VP64-T2A-GFP as a starting point (Kabadi et al., 2014). The dCas9-VP64-T2A-GFP sequence was replaced with a hygromycin phosphotransferase gene using XbaI and AgeI. Next, an Esp3I site within the hygromycin phosphotransferase gene was disrupted using site-directed mutagenesis. Each 4-gRNA set was then cloned into this destination vector via Golden Gate assembly, as described previously (Kabadi et al., 2014). Clones were validated via colony PCR.

Lentivirus containing each 4-gRNA construct was produced using the 2nd generation lentivirus plasmids psPAX2 and pMD2.G. Briefly, 293T cells were grown in 6-well plates at $2.5 \times 10^5$ cells per well, under standard culturing conditions in DMEM
(high glucose, pyruvate; Thermo Fisher Scientific), 10% FBS and 1% penicillin-streptomycin. Cells were transfected with 500 ng psPAX2, 50 ng pMD2.G, and 500 ng of gRNA plasmid using FuGene HD (Promega). One day following transfection, media was replaced with high-FBS media (20%). After 1 day of culturing, media was collected, replaced, and stored at 4°C. After another day of culturing, media was collected again, combined with the previous day’s aliquot, and lentivirus was concentrated using Lenti-X Concentrator (Clontech). Lentivirus was stored in 1/10th volume 1X PBS at -80°C until transduction.

A549 cells were transduced with lentivirus (produced as described above) containing the pLV hU6-sgRNA hUbC-dCas9-KRAB-T2a-Puro construct (Thakore et al., 2015). Cells were selected with 2 ng/µL puromycin for 7 days. Surviving cells were then transduced with each of the above mentioned 4-gRNA constructs, and treated with 2 ng/µL puromycin and 500 µg/mL hygromycin B for 7 days. Surviving cells where then re-plated into 96 well plates, and treated with 100 nM dex for either 0 or 1 hours. Each drug treatment for each cell line was performed in sextuplet.

Expression of genes proximal to chromatin loop anchors was assessed via RT-qPCR, using the Cells-to-CT Kit 1-Step TaqMan Kit (Thermo Fisher Scientific). Each sample was assayed in duplicate, and each FAM-labeled gene probe was multiplexed with a VIC-labeled GAPDH probe as an endogenous control. qPCR was performed on a StepOnePlus Real-Time PCR System (Thermo Fisher Scientific).
3.4.2 Computational

3.4.2.1 Differential loop analyses

The Juicer analysis tool was used to perform read alignment (hg38), read pairing, and deduplication for each in situ Hi-C library (Durand et al., 2016b). Interaction pairs for technical and biological replicates for each time point were combined, and matrix normalization was performed using the Juicebox command-line tool (Durand et al., 2016a). Chromatin loops were called for each time point using the HiCCUPs function in Juicer. Specifically, chromatin loops were called on contact matrices normalized using the VC_SQRT method, at both 5 kb and 10 kb resolution, with an FDR of ≤ 0.1%. Other options used for HiCCUPs loop calling were as follows: -p 4,2 -i 7,5 -t 0.02,1.5,1.75,2 -d 20000,20000. Loop calls for each time point were then merged to create a union set of chromatin loops across the time course. To reduce the presence of spurious loop calls due to artifacts in matrix normalization, loops within 25 kb of 5 kb bins with low mappability (< 100 interaction counts per 1 billion mapped reads) were removed. Raw interaction counts within these loop calls were then tallied for each biological replicate and time point. The MboI restriction enzyme used to fragment DNA in situ is inhibited by CpG methylation overlapping its restriction site. To minimize bias in interaction counts due to differential DNA methylation across the time course, any interaction pair with a ligation junction containing a CpG dinucleotide was removed from the data. In addition, we also sought to minimize bias from differential chromatin accessibility in response to dex treatment. Thus, interaction pairs with a MboI site found within a
DNase hypersensitive site found to be differentially accessible across the dex time course were removed (FDR ≤ 10%; (McDowell, In review)). The resulting interaction counts were then used to find differentially interacting chromatin loops between time points, using a generalized linear model (GLM) likelihood ratio test from the R package edgeR (Robinson et al., 2010). Biological replicate was also included in the GLM formula to account for replicate-specific effects across time points.

### 3.4.2.2 Visualization of dynamic interaction matrices

To generate interaction matrices representing the change in chromatin interactions in response to dex treatment, we first used the `predFC` function in the edgeR package to estimate the log₂(fold-change) in interaction counts at each bin in each sub-matrix we chose to interrogate. To reduce variability in these estimates at sparse interaction bins, we chose a “prior.count” of 10 and 20, for matrices at 5 kb and 10 kb resolution, respectively. We also included the biological replicate in the GLM formula to account for replicate-specific effects across time points. The resulting moderated log₂(fold-change) values were further smoothed using a 3x3 weighted-sliding window, where the smoothed value of each bin can be represented as follows:

\[
\hat{b}_{i,j} = \frac{7}{16} b_{i,j} + \frac{1}{16} \sum_{l=0}^{2} \sum_{m=0}^{2} b_{i+l-1,j+m-1}
\]

The smoothed value of bin \(b\) at coordinates \(i,j\), is effectively the summation of \(1/2\) its original value and \(1/16\) the value of each of its adjacent bins.
3.4.2.3 Aggregate dynamic interaction matrices

In previous work by Rao and Huntley et al. (Rao et al., 2014), the authors developed an approach to visualize the average interaction signal across several regions of the genome. This approach was termed aggregate peak analysis (APA), and was shown to be advantageous in assessing general trends in focal enrichments of interactions, particularly with low-depth Hi-C datasets. As an extension of this method, we sought to visualize the average dynamic interaction signal across several regions of the genome, specifically at dynamic or static loops (Figure 11G). To generate these plots, we first determined interaction dynamics at each individual loop, +/- 25 kb, using the procedures described above, save the inclusion of a “prior.count”, as well as submatrix smoothing. Next, we simply found the mean log2(fold-change) interaction frequency at each bin. To determine the extent to which dynamics at the loop coordinates differ from the dynamics of the local background, we used the value of the center pixel and the values of the 3 x 3 region in the lower-left corner of the sub-matrix to generate a z-score.

3.4.2.4 Pre-existing chromatin interaction analyses

To assess the interaction frequencies of latent dynamic loops we first sought to remove biases in interaction counts due to variability in loop anchor width and interaction distance. Matrix normalized interaction counts at time point 0 were regressed onto loop anchor width and log10(interaction distance), as well as the time point of
interaction dynamics, and interaction class (either increased, decreased, static, or non-loop). Time point-specific effects were assessed by including a coefficient modeling the interaction between time point and interaction class. Effects from anchor width and interaction distance were then subtracted from the observed interaction counts to visualize the pre-dex interaction frequencies of dex-mediated dynamic chromatin loops (Figure 11H).

3.4.2.5 Analysis of loop network sizes

To determine if the size distribution for networks of dynamic loops differs from what is expected by chance, we used a random sampling approach. We first randomly chose an equivalent number of loops from the list of all loops tested for differential interactions across the dex time course. Next, we determined a distribution of network sizes by grouping loops together based on whether or not their loop anchors overlapped. This process was repeated for 100 random sets of loops, and the resultant pooled distribution of network sizes was compared to the observed distribution of network sizes via a KS test.

3.4.2.6 Coordination of dex-responsive networks

To assess the coordination of looping dynamics within a network of dex-responsive loops, we first found the mean and the variance of the log2 fold-changes in
interaction frequency within networks. To determine the mean and variance expected by chance, we shuffled the log$_2$ fold-changes of interaction frequencies across all dynamic loops and recalculated mean and variance for each time point. We then compared these shuffled distributions to the observed distributions via a KS test.

3.4.2.7 Dynamic loop enrichment at dex-regulated genes

To determine the enrichment of dex-responsive chromatin loops at dex-responsive genes, we first determined the number of induced/repressed genes that were within 2 kb of an anchor of an increased/decreased loop (FDR ≤ 5%), at each time point. We then determined how many genes you would expect by chance to fall within 2 kb of a differentially loop anchor. To generate this expected distribution, we first randomly chose an equivalent number of loops from the list of all loops tested for differential interactions across the dex time course, and then determined their association with differentially expressed genes as described above. We repeated this process 1000 times to determine an empirical $P$-value for each gene/loop comparison, at each time point. An equivalent analysis was performed for determining gene enrichment in the intervening regions of loops.

3.4.2.8 Enrichment of increased loop anchors at gene bodies of induced genes

To determine if specific loci at dex-induced genes are enriched for increased looping, we first determined the distribution of loop anchors of increased loops at dex-
induced genes, +/- 10 kb, using anchor midpoints as their coordinates. We then scaled all
gene bodies to a width of 10 kb to make anchor localization comparable across different
gene sizes. These regions were then divided into 200 bp bins, and we used a sliding
window to compare anchor localization of every bin to the 5 upstream and 5
downstream bins. More specifically, we used counts from these flanking bins to
approximate a Poisson distribution and then estimated a $P$-value for anchor localization
at a given bin.

3.4.2.9 Targeted repression of chromatin loop anchors

To determine how dex treatment and/or dCas9KRAB affected gene expression, we
normalized by endogenous control (GAPDH) and compared to the reference sample (0
hour of the negative control, or the 0-hour condition of each dCas9KRAB condition) using
the ΔΔCT method. These values were used for visualization of gene expression effects in
Figure 13. To determine the extent to which dCas9KRAB affected either basal or dex-
mediated gene expression, we used multiple linear regression:

$$-\Delta CT = \beta_0 + \beta_1 \text{Duration} + \beta_2 \text{Target}_A + \beta_3 \text{Target}_B + \beta_4 \text{Target}_A \times \text{Duration} +$$

$$\beta_5 \text{Target}_B \times \text{Duration} + \epsilon$$
Where ΔCT values were regressed onto the duration of 100 nM dex treatment (0 or 1 hours) and the anchor targeted by dCas9^{KRAB} (none, A, or B). To test whether dCas9^{KRAB} specifically altered dex-mediated expression changes, we included coefficients modeling the interaction between the dCas9^{KRAB} target and the duration of treatment.

### 3.4.2.10 Analysis of TF occupancy vs. CTCF occupancy at dynamic loops

To determine if higher TF occupancy is associated with lower CTCF occupancy at loop anchors, we again used linear regression. For each ChIP-seq factor, time point, loop class (increased, decreased, or static), we first calculated the mean of the normalized ChIP-seq signal at the anchors of each loop. We then used these values to regress TF occupancy onto CTCF occupancy. To test the relationship between CTCF occupancy and TF occupancy at loop anchors with evidence of only one anchor occupied by CTCF, we used an analogous approach. Here, instead of using the mean normalized CTCF ChIP-seq signal, we regressed TF signal onto a categorical covariate describing CTCF as either present (1) or absent (0).

### 3.4.2.11 Predicting differential expression with chromatin interactions

To determine which genomic features at loop anchors are associated with dex-mediated gene expression dynamics, we used elastic-net logistic regression. We
included 4 general classes of predictors in our models: 1) basal TF occupancy, histone modification, and DNase accessibility; 2) TF occupancy, histone modification, and DNase accessibility in response to dex; 3) TF motifs; and 4) GC-associated sequence features.

ChIP-seq signal for TF occupancy and histone modifications, and DNase-seq signal at loop anchors at time point 0 was first normalized by the total number of mapped reads per experiment, by anchor width, and by input ChIP-control signal (for ChIP-seq data). Normalized values were then averaged across biological replicates. Dex-responsive ChIP-seq signal for TF occupancy and histone modifications, and DNase-seq signal for loop anchors was calculated by computing the log2 fold-change in signal at each loop anchor, while accounting for the total number of mapped reads per experiment. Motif scores for each loop anchor were found using MAST (Bailey and Gribskov, 1998) with PWMs for GR and associated co-factor proteins from JASPAR (Mathelier et al., 2014), as well as the negative glucocorticoid response element (nGRE) (Surjit et al., 2011). Search space was limited to DHSs overlapping anchor regions at any time point. For anchors without an annotated DHS, the entire anchor region was used. Repetitive elements where removed prior to motif scanning with Repeat Masker (Smit, 2013-2015) and the highest motif score at each loop anchor was used as a predictor. The G4 density (number of G4s per bp) for each loop was calculated by using previously published data (Chambers et al., 2015).
The unique set of loop anchors from the union set of loops were classified as either interacting with a dex-induced gene, a dex-repressed gene, or only static genes, as determined by the overlap of the anchor coordinates with the gene body. For each class, anchors were further subsetted into those “proximal” to a gene (those overlapping a gene body) and those “distal” to gene (those looped to a gene body). DNA anchor classification was regressed onto the above-described genomic predictors using elastic net logistic regression via the R package glmnet (Friedman et al., 2010). Anchors interacting with induced genes and repressed genes were modeled separately, using anchors interacting with static genes as a negative set. Each time point and anchor position (proximal/distal) was also modeled separately. Predictors were mean-centered and scaled to a standard deviation of 1, and optimal λ values for a range of α values were found using the cv.glmnet function. The lowest λ value (and corresponding α value) within the standard error of the λ value with the lowest mean squared-error across all cross-validations was chosen for each model. We used the mean optimal α and λ values across all models as common parameter values so that the scale of the coefficients across models are comparable.
3.4.2.12 Predicting differential interactions via epigenomic marks

To determine which genomic features at loop anchors are associated with dexamethasone-mediated looping dynamics, we used elastic-net logistic regression. We included the same 4 general classes of predictors as described above.

ChIP-seq signal for TF occupancy and histone modifications, and DNase-seq signal at loop anchors at time point 0 was calculated as described above. The maximum value at time point 0 among loop anchors was used as a predictor for each loop. Dexamethasone-responsive ChIP-seq signal for TF occupancy and histone modifications, and DNase-seq signal for loop anchors was calculated as described above. The value for the loop anchor with the greatest absolute change in signal, compared to time point 0, was used as the predictor for each loop. Motif scores for each loop were determined as described above, and the maximum score among loop anchors was used for each chromatin loop. The G4 density for each anchor was calculated as described above, and the average G4 density across both loop anchors was used as the predictor for each loop.

Chromatin loop log₂ fold-changes were regressed onto these predictors using elastic net regression via the R package glmnet (Friedman et al., 2010). Positive and negative log₂ interaction fold changes, as well as each time point, were modeled separately. Predictors were mean-centered and scaled to a standard deviation of 1, and optimal λ values for a range of α values were found using the cv.glmnet function. The lowest λ value (and corresponding α value) within the standard error of the λ value
with the lowest mean squared-error across all cross-validations was used for each model. We used the mean optimal \( \alpha \) and \( \lambda \) values across all models as common parameter values so that the scale of the coefficients across models are comparable.

3.4.2.13 Differential compartment analyses

A/B compartments were calculated for each biological replicate and time point using previously described methods with some modifications (Dixon et al., 2015): 1) Observed/Expected (O/E) contact matrices were generated for each chromosome at 50 kb, 100 kb, 250 kb, and 500 kb resolution, using Juicer (Durand et al., 2016b); 2) The Pearson correlation matrix of each O/E was calculated; 3) The covariance matrix of each correlation matrix was calculated; 4) PCA was carried out on each covariance matrix; 5) The PC corresponding to A/B compartments (typically the first PC), for each chromosome was oriented such that positive values positively correlated with regions of active gene expression (RNA-seq data from time point 0; (McDowell, In review)). To find regions whose PC1 value varied significantly with duration of dex treatment, we performed an analysis of variance (ANOVA) for time point and biological replicate, for each genomic bin. Dex-responsive genomic bins were called using an FDR cutoff of \( \leq 5\% \).
3.4.2.14 Interaction dynamics between compartments

To determine the extent to which genomic regions with variable A/B values altered their interactions with A/B compartments across the time course, we used multiple linear regression. We first subset regions with dynamic A/B values into those that either decreased or increased their A/B value across the time course. For each of these sets, we regressed the log(O/E interaction count) onto log(interaction distance), chromosome, duration of dex treatment, and class of interacting compartment (either A or B). To assess compartment-specific interactions across time, we included a coefficient to model the interaction between compartment class and duration of dex treatment. Separate models were fit using data at 50, 100, 250 and 500 kb resolutions. Interactions were visualized by first removing effects of chromosome and interaction distance (Appx. A; Figure 22B-C).

3.4.2.15 Enrichment of dex-regulated genes among compartment types

To evaluate whether dex-induced or dex-repressed genes were enriched within dynamic compartments, we used a sub-sampling approach. For each gene class, we randomly sampled 200 genomic bins from each bin class: increased PC1, decreased PC1, or static PC1. We then determined the number of induced/repressed found within each set of genomic bins. This was repeated 100 times to generate distributions of gene counts.
for each bin class. The distributions using dynamic regions were compared to that of static regions via Wilcoxon Rank-Sum test (Figure 16E-F and Appx. A; Figure 22G-I).
4. Conclusions and Future Directions

In this work, we have revealed new insights into how the GR mediates genomic responses to GC treatment. Using high-resolution TF-mapping approaches, we provide evidence that the GR binds to the genome in both dimeric and monomeric configurations. We also revealed evidence that the GR binds to the genome with other TFs in both highly constrained, and more flexible configurations. Next, using high-throughput approaches to measure chromosome conformation, we revealed several insights into GC-mediated changes in genome topology. We observed that GC-responsive chromatin looping interactions do not form de novo, but are formed prior to GC-treatment. We also observed a concordance between gene expression responses and GR-associated chromatin looping responses, suggesting a functional link between the two. Furthermore, we observed higher-order GC-mediated changes in chromosome compartmentalization, suggesting a role for chromosome territories in the genomic response to GCs. Together, these findings support a model where pre-formed chromatin loops pre-program the transcriptomic response to GC treatment, and that these interactions are associated with a repertoire of GR-cofactor binding configurations.

4.1 The topologic pre-programming of transcriptional responses

Here we have demonstrated that the chromatin loops connecting distal GR binding sites to GC-responsive genes are largely formed prior to GC treatment. This
observation suggests that the transcriptomic response to GCs has been pre-programmed into the 3D architecture of the genome. Similar observations have been made in respect to TF occupancy and chromatin state. The majority of GR binding sites are pre-determined by AP-1 binding, as well as DNase accessibility (Biddie et al., 2011, John et al., 2011). DNase accessibility and GR binding have been demonstrated to vary across different cell types, along with GC-responsive transcription (Gertz et al., 2013, Polman et al., 2012, Consortium, 2012). If GC-responsive loops are also cell-type specific, this could serve as an additional layer of regulation to propagate cell-type specific transcriptomic responses to GCs. However, it is currently unclear to what extent latent GC-responsive chromatin loops vary across cell types, and how such variability may contribute cell-type specific gene expression responses.

Recent high-resolution studies of basal chromatin looping across cell types have demonstrated that 40-75% of chromatin loops may be shared across cell types (Rao et al., 2014, Sanyal et al., 2012). A current gap in our fundamental knowledge of stimulus-responsive chromatin looping is what proportion of latent-dynamic loops are cell-type specific. Furthermore, what is the concordance between cell-type specific looping and cell-type specific gene expression responses? An investigation of chromatin looping dynamics across multiple cell types would help resolve such a question. Not only would this provide insight into the fundamental mechanisms of cell-type specific gene regulation, but these studies could help elucidate the mechanistic function chromatin
looping itself. For example, what is the prevalence of cell-type shared loops at cell-type specific stimulus-responsive genes? These data could address questions of the sufficiency for chromatin looping in gene expression responses. Moreover, integrating the heterogeneity of cell-type specific loops with the heterogeneity of cell-type specific epigenetic marks may allow you to infer mechanistic relationships between these different features of gene regulation. Such inferential hypotheses could then be tested via synthetic epigenetic modifiers such as dCas9KRAB and dCas9P300 (Thakore et al., 2015, Hilton et al., 2015).

4.2 Stimulus-induced chromatin looping dynamics

In this work, we find that ~7% of looping interactions change their interaction frequency in response to glucocorticoid treatment. Although several studies have examined stimulus responsive looping at specific loci (Hakim et al., 2009, Li et al., 2013, Ramamoorthy and Cidlowski, 2013), few studies have thus far examined stimulus responsive chromatin looping on a global scale. However, in the context of these genomic studies of chromatin looping, the dynamics we observe with GC treatment appear to particularly robust. In one study, Jin and colleagues assayed chromatin interactions before and after TNF-α treatment, observing minimal changes in interactions upon treatment (Jin et al., 2013). In another study, Phanstiel and colleagues

---

4 These are both manuscripts I co-authored as part of my graduate studies.
assay chromatin looping in THP-1 cells, a monocytic leukemia cell line, before and after differentiation to macrophages via phorbol myristate acetate (PMA) treatment (Phanstiel et al., 2017). Here they only observed ~1% of loops change interaction frequency after PMA treatment. In the context of our results, it is somewhat surprising that PMA treatment incurred so few dynamic loops, especially since treatment induces cellular differentiation and is not just a transient stimulus as with GCs. Nonetheless, comparing our results with the abovementioned studies is made difficult by inherent differences in methodology and biological system. More specifically, Jin and colleagues used in-dilution Hi-C, where chromatin interactions are assayed after cell lysis. Compared to in situ Hi-C, the method used by us and Phanstiel and colleagues, in-dilution Hi-C has been documented as less sensitive in detecting distinct looping interactions (Rao et al., 2014). Furthermore, it is unclear if TF activation via GCs or TNF-α, and the subsequent genomic consequences of such activation, are comparable in these cell types (A549 and IMR90 cells, respectively). Though the biochemistry is comparable between our study and that of Phanstiel and colleagues, the details of the computational analysis of dynamic looping are quite different. As the cost of sequencing continues to decline, and chromatin conformation assays advance, it will be interesting to see the extent to which different stimuli mediate chromatin looping dynamics.
4.3 Temporal and spatial dynamics of chromatin loops and looping networks

In this work, we use proximity-based 3C assays to determine genomic conformational responses to GC treatment. When interpreting dynamic chromatin looping data, it is important to consider the limitations of 3C technology (Fudenberg and Imakaev, 2017): 1) 3C is performed on a population of cells, therefore observed interaction frequencies should be considered an aggregation of this population; 2) 3C does not directly provide information about 3D distances between loci. Observed contacts are dependent on the cross-linking radius of formaldehyde, as well as the chromatin composition at the respective loci; and 3) 3C generally only assays 2-way interactions (a very small percentage of chimeric DNA molecules will contain fragments from more than 2 locations) (Rao et al., 2017). Most higher-order interaction networks can only be inferred. Therefore, many questions about the spatiotemporal aspects of stimulus-induced chromatin looping remain unanswered. Below, I discuss some of these outstanding challenges and how recent technological advances could be used to address them.

4.3.1 How is the 3D distance between loop anchors altered upon GC treatment?

Our data shows that thousands of loops change interaction frequency in response to GC treatment. As mentioned above, these results indicate that the respective
loci involved in a looping interaction cross-link together at different frequencies in response to GC treatment. It is unclear if this increased cross-linking is due to actual decreases in the 3D distance between two loci, or are merely the result of changes in chromatin composition that favors or disfavors cross-linking conformations. Moreover, it is unclear how dynamics in looping are distributed across the population of cells. Addressing these issues would provide insight into the precise mechanisms of dynamic looping. Specifically, does GC treatment increase the co-localization of two loci, or are the two loci already interacting, and GC treatment just increases the stability of that interaction? Furthermore, elucidating the cellular heterogeneity underlying looping responses will allow for more focused investigations of the relationship between chromatin looping and gene expression responses.

Although some studies have used Hi-C data to model 3D distances computationally (Lieberman-Aiden et al., 2009, Sanborn et al., 2015), the most direct means for elucidating 3D distances, while preserving information about cellular heterogeneity, is via microscopy. Such methods have revealed valuable insights into dynamic chromatin looping (Fullwood et al., 2009, Jubb et al., 2017), however, these methods are typically low-throughput. The recent development of highly-multiplexed fluorescence in situ hybridization (FISH) methods such as FISSEQ and MERFISH have shown promise for the in situ interrogation of the transcriptome (Lee et al., 2014, Chen et
Future advancements in such technology will likely prove valuable in interrogating the 3D positioning of chromatin loops en masse.

**4.3.2 How do looping kinetics relate to gene expression kinetics?**

Although our time course study of GC treatment provided useful insights into the genomic response to GCs, these approaches are unable to resolve the precise temporal relationships between chromatin looping and dynamic gene expression responses. Gene transcription has been observed to occur in short spurts, termed “transcriptional bursting” (McKnight and Miller, 1979). The frequency of these bursts can be increased by the addition of adjacent enhancer elements (Fukaya et al., 2016). Furthermore, studies of synthetically “forced” chromatin loops have demonstrated that looping interactions between enhancers and promoters can increase the rate of transcriptional bursting (Bartman et al., 2016). Such approaches applied to systems of stimulus responsive transcription could prove valuable in coupling the dynamics of chromatin looping to rates of transcriptional regulation. Understanding this relationship would likely aid in the elucidation of the causative role of chromatin looping in gene expression responses.
4.3.3 How do multiple enhancers regulate a single gene?

Finally, as mentioned above, Hi-C is generally only capable of interrogating 2-way interactions. We and others have used these data to infer networks of looping interactions, via shared anchor points (Figure 12) (Sanborn et al., 2015). However, it is unclear if these higher-order interactions occur simultaneously, or are the result of multiple 2-way interactions aggregated over a population of cells. A better understanding of these structures is important for elucidating how multiple distal enhancers regulate a single gene. Specifically, do multiple enhancers interact and cooperate in putative interaction networks to regulate gene expression, or does each individual enhancer interact with its target gene separately to contribute to the overall gene expression level? Addressing these questions would be particularly useful for interpreting the functional impacts of mutations in enhancer elements.

Genomic methods for the direct interrogation of such higher-order interactions are just beginning to enter the scientific literature (Beagrie et al., 2017). Such methods applied to dynamic systems like the GC-response would be highly informative in elucidating the dynamics within such looping networks, and how this relates to gene expression responses.
4.4 Mechanistic function of chromatin looping

Despite the concordance we observe between dynamic chromatin looping and dynamic gene expression responses, the literature is mixed as to whether chromatin looping has a causative role in regulating gene expression. For example, in one study where the authors synthetically force chromatin looping between the LCR and β-globin genes, the authors observed gene induction in some cell types, but not others, indicating that looping itself is not sufficient to drive gene expression responses (Morgan et al., 2017). Furthermore, in studies where CTCF or cohesin are depleted from the cell, authors observed an almost complete loss of chromatin looping (Nora et al., 2017, Rao et al., 2017). Notably, gene expression is only mildly affected in these experiments, suggesting looping is dispensable for gene regulation altogether. It is currently unclear how these results can be reconciled with the enormous amount of genomics data showing strong enrichments between chromatin looping and gene expression (Fullwood et al., 2009, Sanyal et al., 2012, Heidari et al., 2014, Li et al., 2012, Dowen et al., 2014, Javierre et al., 2016). As we begin to converge on the precise mechanisms of loop formation, experiments perturbing such machinery will prove useful in gaining insight into the role of chromatin looping in gene regulation. Furthermore, a better understanding of the mechanisms of loop formation could aide in the development of new CRISPR-Cas9-based tools aimed at site-specific loop disruption and formation.
4.5 GR-mediated transcriptional repression

In this work, we demonstrate that GR-bound putative enhancer elements physically interact with both GC-induced and GC-repressed genes. Furthermore, we show those GR binding sites looped to each class of GC-responsive gene have unique characteristics. Although both are bound by the GR and its associated cofactor TFs, GR binding sites looped to GC-induced genes are enriched for the GR motif. Conversely, those GR binding sites looped to GC-repressed genes are enriched for AP-1 and FOXA1 motifs, suggesting the GR may bind to the genome through protein-protein interactions. Moreover, these observations specifically provide evidence for a “transrepressive” model of GC-mediated transcriptional repression. In this model, a GR monomer binds to DNA-bound TFs, and can recruit co-repressors to ablate transcriptional activity (Ray and Prefontaine, 1994, Ronacher et al., 2009). An alternative model of GR-mediated repression, is repression via GR binding to a nGRE. At these response elements, the GR is thought to bind in an alternative dimer configuration, resulting in the recruitment of co-repressors and subsequent transcriptional repression (Surjit et al., 2011, Ramamoorthy and Cidlowski, 2013, Hudson et al., 2013). Notably, we did not find evidence for the enrichment of nGREs at GR binding sites looped to repressed genes, nor did we find evidence for nGREs in our analysis of GR ChIP-exo data using GEM.

Thus far, evidence from genomic studies has failed to generalize either transrepression or nGREs as canonical mechanisms of direct, GR-mediated
transcriptional repression. The first GR ChIP-seq revealed that GR binding sites bind further away from GC-repressed genes that what would be expected by chance (Reddy et al., 2009). In addition, these GR binding sites failed to show enrichment for the nGRE motif. Moreover, high-throughput reporter assays measuring the transcriptional activity of GR binding sites have failed to provide evidence for GR-mediated transcriptional repression (Vockley et al., 2016). However, our chromatin looping data demonstrates that distal GR binding sites are enriched for interactions with GC-repressed genes. These results are, to our knowledge, the first to demonstrate a global enrichment of the GR at GC-repressed genes. While this advancement is important for reconciling single-gene studies of the GR with more recent genomic studies, many unanswered questions remain in respect to GR’s repressive activity. In the sections below, I will discuss some of these outstanding questions and experimental approaches that could address them.

4.5.1 Why do putative GR-repressive sites lie so far from the genes they interact with?

Although we observed looping between GR binding sites and GC-repressed genes, these loops spanned longer distances than what is expected by chance. It is unclear why this discrepancy exists. Addressing this discrepancy could provide insights into the mechanisms of GR-mediated transcriptional repression. For example, one possible explanation could be differences in the ability to detect each class of GR binding site. Current models of GR-repression involve the recruitment of corepressors and the
removal of active chromatin marks (Chen and Evans, 1995, Horlein et al., 1995, Guenther et al., 2001). This activity may reduce the accessibility of such loci, and reduce the occupancy of nearby TFs. This combination of chromatin compaction and reduced TF binding could result in the decreased detection of these GR binding sites via ChIP-seq. This hypothesis could be tested via a short-term GC time course, in an attempt to “catch” the GR binding sites that may be closer to GC-repressed genes, before such a negative-feedback mechanism could take hold.

### 4.5.2 Are these GR sites looped to GC-repressed genes functional?

Although we observe an enrichment of GR binding at sites looped to repressed genes, it is still unclear if these sites actually function in gene repression, or are just transient inconsequential binding events. In support of the latter argument is the lack of repressive activity observed in high-throughput reporter assays of ChIP’d GR binding sites (Vockley et al., 2016). However, since these experiments assay binding site activity outside their genomic context, it is possible that GR-mediated repression may require elements in the endogenous genome for proper functioning (Inoue et al., 2017). In this case, genome-editing experiments may help validate these putative repressive binding sites. Even so, since these sites appear to be tethered, and not direct GR binding, genome-editing approaches may still prove difficult in mechanistically separating the basal activity of GR-tethered TFs, and the putative GR-mediated repression of such TFs.
4.5.3 What is the role of nGREs in global gene expression responses to GCs?

Though anecdotal and structural evidence exists for the repressive function of GR-bound nGREs (Ramamoorthy and Cidlowski, 2013, Surjit et al., 2011, Hudson et al., 2013), we do not observe an enrichment for this motif at our putative repressive GR binding sites, nor do we find evidence for nGREs in our ChIP-exo data. One explanation for this lack of enrichment could be a decreased capability to detect these GR binding sites, for reasons similar to those described above. Another hypothesis could be that nGREs are not a general mechanism of GR-mediated repression, but a more specialized, gene-specific means to repress transcription. Such a hypothesis could be tested via a gene-ontology analysis of those genes interacting with GR-bound nGREs. Data across different cell types may also help discern any specialized function for nGREs.
Figure 17: Thousands of loops respond to dex treatment

(A) Correlation of interaction counts at chromatin loops across biological replicates and time points. (B, C) Chromatin interactions at time point 0, at loci near ZFP36 and DIO3OS, respectively. Matrix-normalized (VC_SQRT method) interaction maps were visualized with Juicebox (Durand et al., 2016a). Latent dynamic loops have been annotated with circles. Dex-induced genes and dex-repressed genes are colored in blue and red, respectively. (D) The number of dynamic loops across the time course at different FDR cutoffs. Values have been scaled to a mean of 1 to make trends comparable. (E) Distribution of log2 fold-changes in loop interaction frequencies across the dex time course. (F) Overall enrichment of loop dynamics compared to local background, represented by z-scores. Each z-score was calculated using the value of the center pixel and the values of the 3 x 3 region in the lower left corner of each submatrix. (G) Regression statistics for determining the association of interaction frequencies of chromatin loops at time point 0 to loop dynamics and time.
Figure 18: Networks of dynamic loops change concordantly across multiple time points

(A) Chromatin loops at time point 0 near DUSP1 frequently share anchor regions. Matrix-normalized (VC_SQRT method) interaction maps were visualized with Juicebox (Durand et al., 2016a). Latent dynamic loops have been annotated with circles. Dex-induced genes have been labeled below the gene annotation track. (B-D) Mean vs. variance of log2 fold-changes of interaction frequencies of loops within networks of dynamic loops after 4, 8, and 12 hours of dex treatment, respectively.
Figure 19: Anchors of dex-increased loops function in dex-mediated gene induction

(A) The number of loops at a given dex-regulated gene. Loop-gene pairings were assigned based on whether the coordinates of loop anchors overlapped gene bodies, and differences in distributions were assessed via Wilcoxon Rank-sum test (*: \( P \leq 0.05 \)). (B) Number of dex-induced genes at the anchors of dex-decreased loops (*: \( P \leq 0.05 \); n.s.: \( P > 0.05 \)). (C) Top: Distribution of anchors of dex-increased loops at dex-induced genes. The gene bodies of dex-induced genes were all scaled to 10 kb. Bottom: Statistical significance of loop anchor enrichment at induced genes. The red horizontal line designates a \( P \)-value of 0.05. (D, E) Number of dex-repressed genes at the anchors of either decreased or increased loops (*: \( P \leq 0.05 \); n.s.: \( P > 0.05 \)). (F, G) Number of dex-repressed genes encompassed by decreased or increased loops, respectively (*: \( P \leq 0.05 \); n.s.: \( P > 0.05 \)).
n.s.: $P > 0.05$). (H-J) Chromatin interactions at time point 0 near MT2A, ARL4A, and TCEAL1, respectively. Matrix-normalized (VC_SQRT method) interaction maps were visualized with Juicebox (Durand et al., 2016a). Latent dynamic loops have been annotated with circles. (K-M) Gene expression of genes near targets of dCas9KRAB from Figure 13, as measured by RT-qPCR. Gene expression at 0 h was compared to the negative control (no gRNA; gray) and gene expression after 1 h dex treatment was compared to 0 h dex for each target (blue). Bar plots designate mean log2 fold-change and error-bars designate standard error of the mean. Significance of gene expression changes was determined via linear regression (* : FDR ≤ 10%, n.s. : FDR > 10%). (N-P) The effects of dCas9KRAB on basal and dex-mediated gene expression of genes near MT1A, ARL4A and TCEAL1, respectively. ΔCT values were regressed onto treatment duration and dCas9KRAB target. dCas9KRAB effects on dex-mediated gene induction were modeled using an interaction term between target site and treatment duration (*: FDR > 10%).
Figure 20: Dex-mediated looping dynamics are associated with low CTCF occupancy

(A-C) Change in interaction frequency of loops vs. mean CTCF occupancy at loop anchors after 4, 8, and 12 h dex treatment, respectively. (D-E) Chromatin interactions at time point 0 near EDN3 and CEPBD, respectively. Matrix-normalized (VC_SQRT method) interaction maps were visualized with Juicebox (Durand et al., 2016a). Latent dynamic loops have been annotated with circles, as done in Figures 14E and 14J.
Figure 21: GR binding at distal enhancers is associated with gene activation and repression

(A, B) Coefficient values from elastic net regressions using genomic features at loop anchors to predict whether a loop increases or decreases its interaction frequency after dex treatment, respectively. (C, D) Dynamics of chromatin loops at loci near FOSL2, and near ARRDC3, respectively. Contact maps represent the log₂ fold-change in interaction frequency between time point 4 and time point 0. (E) The cumulative distribution of the distance between distal GR binding sites and the genes they are looped to. (F) The average change in ChIP-seq signal at anchors looped to dex-induced
genes. Anchors have been subset based on evidence of GR occupancy. (G) Motif enrichments at GR binding sites looped to repressed genes. Motif occurrences were compared to those found at binding sites looped to repressed genes with no GR (Fisher’s exact test; * : FDR ≤ 5%). (H) The change in the interaction frequency of loops at dexamethasone-repressed genes. Loops have been subset by GR occupancy.
Figure 22: Compartmental dynamics are reproducible across resolutions

(A) Distributions of PC1 values for regions that either increase, decrease, or are static in response to dex. All classes of dynamic regions were enriched over static regions (Wilcoxon Rank-Sum test, $P \leq 0.05$). (B) Interaction counts of regions that decrease their PC1 value in response to dex with regions in either compartment A or B. Interaction counts were normalized by interaction distance and chromosome via multiple linear regression. Outliers have been excluded from this plot. (C) Interaction counts of regions that increase their PC1 value in response to dex with regions in either
compartment A or B. Interaction counts were normalized by interaction distance and chromosome via multiple linear regression. Outliers have been excluded from this plot. (D-F) Regions with dynamic PC1 values. The PC1 values of each dynamic region in the genome were averaged across biological replicates, standardized across time points, and then hierarchically clustered. (G-I) Number of dex-induced or dex-repressed genes within a given compartment type (*: $P \leq 0.05$; n.s.: $P > 0.05$).
<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chemicals, Peptides, and Recombinant Proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ham’s F-12K Medium</td>
<td>Thermo Fisher Scientific</td>
<td>21127022</td>
</tr>
<tr>
<td>Fetal Bovine Serum</td>
<td>Thermo Fisher Scientific</td>
<td>16000044</td>
</tr>
<tr>
<td>Penicillin-Streptomycin (10,000 U/mL)</td>
<td>Thermo Fisher Scientific</td>
<td>15140122</td>
</tr>
<tr>
<td>Dexamethasone</td>
<td>Sigma-Aldrich</td>
<td>D4902-25MG</td>
</tr>
<tr>
<td>Formaldehyde</td>
<td>Sigma</td>
<td>F8775-500ML</td>
</tr>
<tr>
<td>Protease inhibitors</td>
<td>Sigma</td>
<td>P8340</td>
</tr>
<tr>
<td>NEBuffer 2</td>
<td>New England Biolabs</td>
<td>B7002S</td>
</tr>
<tr>
<td>MboI</td>
<td>New England Biolabs</td>
<td>R0147M</td>
</tr>
<tr>
<td>biotin-14-dATP</td>
<td>Life Technologies</td>
<td>19524-016</td>
</tr>
<tr>
<td>DNA Polymerase I, Large (Klenow) Fragment</td>
<td>New England Biolabs</td>
<td>M0210</td>
</tr>
<tr>
<td>10X NEB T4 Ligase Buffer</td>
<td>New England Biolabs</td>
<td>B0202</td>
</tr>
<tr>
<td>BSA (20 mg/mL)</td>
<td>New England Biolabs</td>
<td>B000S</td>
</tr>
<tr>
<td>Proteinase K (20 mg/mL)</td>
<td>Thermo Fisher Scientific</td>
<td>EO0492</td>
</tr>
<tr>
<td>T4 PNK</td>
<td>New England Biolabs</td>
<td>M0201</td>
</tr>
<tr>
<td>T4 DNA polymerase I</td>
<td>New England Biolabs</td>
<td>M0203</td>
</tr>
<tr>
<td>Klenow exo-</td>
<td>New England Biolabs</td>
<td>M0212</td>
</tr>
<tr>
<td>1X Quick Ligation reaction buffer</td>
<td>New England Biolabs</td>
<td>B6058</td>
</tr>
<tr>
<td>Quick Ligase</td>
<td>New England Biolabs</td>
<td>M2200</td>
</tr>
<tr>
<td>Q5 High-Fidelity DNA Polymerase</td>
<td>New England Biolabs</td>
<td>M0491L</td>
</tr>
<tr>
<td>XbaI</td>
<td>New England Biolabs</td>
<td>R0145</td>
</tr>
<tr>
<td>AgeI</td>
<td>New England Biolabs</td>
<td>R0552</td>
</tr>
<tr>
<td>DMEM, high glucose, pyruvate</td>
<td>Thermo Fisher Scientific</td>
<td>11995065</td>
</tr>
<tr>
<td>FuGENE HD Transfection Reagent</td>
<td>Promega</td>
<td>E2311</td>
</tr>
<tr>
<td>Item</td>
<td>Supplier</td>
<td>Catalog No.</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>---------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Lenti-X Concentrator</td>
<td>Clontech</td>
<td>631232</td>
</tr>
<tr>
<td>Puromycin dihydrochloride</td>
<td>Sigma</td>
<td>P8833-10MG</td>
</tr>
<tr>
<td>Hygromycin B</td>
<td>Thermo Fisher Scientific</td>
<td>10687010</td>
</tr>
<tr>
<td><strong>Critical Commercial Assays</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agencourt AMPure XP beads</td>
<td>Beckman Coulter</td>
<td>A63882</td>
</tr>
<tr>
<td>Dynabeads MyOne Streptavidin T1 beads</td>
<td>Life Technologies</td>
<td>65602</td>
</tr>
<tr>
<td>Qubit dsDNA HS Assay</td>
<td>Thermo Fisher Scientific</td>
<td>Q32851</td>
</tr>
<tr>
<td>High Sensitivity D1000 ScreenTape</td>
<td>Aglient</td>
<td>5067-5585</td>
</tr>
<tr>
<td>MiSeq Reagent Kit v2 (300-cycles)</td>
<td>Illumina</td>
<td>MS-102-2002</td>
</tr>
<tr>
<td>Cells-to-CT Kit 1-Step TaqMan Kit</td>
<td>Thermo Fisher Scientific</td>
<td>A25603</td>
</tr>
<tr>
<td>TaqMan Assay - MT2A (FAM-MGB)</td>
<td>Thermo Fisher Scientific</td>
<td>Hs02379661_g1</td>
</tr>
<tr>
<td>TaqMan Assay - MT1A (FAM-MGB)</td>
<td>Thermo Fisher Scientific</td>
<td>Hs00831826_s1</td>
</tr>
<tr>
<td>TaqMan Assay - MT1X (FAM-MGB)</td>
<td>Thermo Fisher Scientific</td>
<td>Hs00745167_sH</td>
</tr>
<tr>
<td>TaqMan Assay - TCEAL1 (FAM-MGB)</td>
<td>Thermo Fisher Scientific</td>
<td>Hs02378857_s1</td>
</tr>
<tr>
<td>TaqMan Assay - TCEAL3 (FAM-MGB)</td>
<td>Thermo Fisher Scientific</td>
<td>Hs03645841_gH</td>
</tr>
<tr>
<td>TaqMan Assay - TCEAL4 (FAM-MGB)</td>
<td>Thermo Fisher Scientific</td>
<td>Hs02520318_gH</td>
</tr>
<tr>
<td>TaqMan Assay - MORF4L2 (FAM-MGB)</td>
<td>Thermo Fisher Scientific</td>
<td>Hs01856765_s1</td>
</tr>
<tr>
<td>TaqMan Assay - ARL4A (FAM-MGB)</td>
<td>Thermo Fisher Scientific</td>
<td>Hs01932504_s1</td>
</tr>
<tr>
<td>TaqMan Assay - SCIN (FAM-MGB)</td>
<td>Thermo Fisher Scientific</td>
<td>Hs00263961_m1</td>
</tr>
<tr>
<td>TaqMan Assay - GAPDH (VIC-PL)</td>
<td>Thermo Fisher Scientific</td>
<td>Hs02758991_g1</td>
</tr>
<tr>
<td><strong>Deposited Data</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hi-C on A549 cell line treated with 100 nM dexamethasone for 0 hours.</td>
<td>This paper</td>
<td>GEO: GSE92819</td>
</tr>
<tr>
<td>Hi-C on A549 cell line treated with 100 nM dexamethasone for 1 hours.</td>
<td>This paper</td>
<td>GEO: GSE92793</td>
</tr>
<tr>
<td>Hi-C on A549 cell line treated with 100 nM dexamethasone for 4 hours.</td>
<td>This paper</td>
<td>GEO: GSE92804</td>
</tr>
<tr>
<td>Hi-C on A549 cell line treated with 100 nM dexamethasone for 8 hours.</td>
<td>This paper</td>
<td>GEO: GSE92825</td>
</tr>
<tr>
<td>Hi-C on A549 cell line treated with 100 nM dexamethasone for 12 hours.</td>
<td>This paper</td>
<td>GEO: GSE92811</td>
</tr>
</tbody>
</table>

**Experimental Models: Cell Lines**

| Human: A549 cell line | ATCC | CCL-185 |
| Human: 293T cell line | ATCC | CRL-3216 |

| Recombinant DNA |
|-----------------|------|--------|
| phU6-gRNA       | Addgene | 53188 |
| pmU6-gRNA       | Addgene | 53187 |
| phH1-gRNA       | Addgene | 53186 |
| ph7SK-gRNA      | Addgene | 53189 |
| pLV hUbC-dCas9 VP64-T2A-GFP | Addgene | 53192 |
| psPAX2          | Addgene | 12260 |
| pMD2.G          | Addgene | 12259 |
| pLV hU6-sgRNA hUbC-dCas9-KRAB-T2a-Puro | Addgene | 71236 |

<table>
<thead>
<tr>
<th>Software and Algorithms</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRISPR Design Tool</td>
</tr>
<tr>
<td>Juicer</td>
</tr>
<tr>
<td>Juicebox</td>
</tr>
<tr>
<td>BWA (v.0.7.13)</td>
</tr>
<tr>
<td>edgeR (v.3.8.6)</td>
</tr>
<tr>
<td>Tool</td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>MAST (v.4.10.0)</td>
</tr>
<tr>
<td>glmnet (v.2.0-10)</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>JASPAR</td>
</tr>
<tr>
<td>G-quadruplex positions</td>
</tr>
<tr>
<td>RNA-seq, ChIP-seq, and DNase-seq across a 100 nM dex time course in A549 cells.</td>
</tr>
<tr>
<td>Locus</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>MT2A</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>TCEAL1</td>
</tr>
<tr>
<td>ARL4A</td>
</tr>
<tr>
<td>ARL4A</td>
</tr>
<tr>
<td>ARL4A</td>
</tr>
<tr>
<td>ARL4A</td>
</tr>
</tbody>
</table>

111
<table>
<thead>
<tr>
<th>ARL4A</th>
<th>A</th>
<th>3</th>
<th>F</th>
<th>TCCCAACGTCGTCTAGGGCGTCCGC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARL4A</td>
<td>A</td>
<td>3</td>
<td>R</td>
<td>AAACGCGGACGCCCTAGACGACGTT</td>
</tr>
<tr>
<td>ARL4A</td>
<td>A</td>
<td>4</td>
<td>F</td>
<td>CCTCGAGTTCGAATCGGCCTAGGC</td>
</tr>
<tr>
<td>ARL4A</td>
<td>A</td>
<td>4</td>
<td>R</td>
<td>AAACGCGCTAGGCCGATTCGAACTC</td>
</tr>
<tr>
<td>ARL4A</td>
<td>B</td>
<td>1</td>
<td>F</td>
<td>CACCGAGCAAGACAAATCTCGGCG</td>
</tr>
<tr>
<td>ARL4A</td>
<td>B</td>
<td>1</td>
<td>R</td>
<td>AAACGCGAGATTTGTCTTTGCTC</td>
</tr>
<tr>
<td>ARL4A</td>
<td>B</td>
<td>2</td>
<td>F</td>
<td>TTGTTTGCTAATCGCTACTGAT</td>
</tr>
<tr>
<td>ARL4A</td>
<td>B</td>
<td>2</td>
<td>R</td>
<td>AAACATCAGTAAGCAATTAGGCAAA</td>
</tr>
<tr>
<td>ARL4A</td>
<td>B</td>
<td>3</td>
<td>F</td>
<td>TCCCCACCGATTAACAGAGCGTTG</td>
</tr>
<tr>
<td>ARL4A</td>
<td>B</td>
<td>3</td>
<td>R</td>
<td>AAACGCGAAGCTTTGCTCGGT</td>
</tr>
<tr>
<td>ARL4A</td>
<td>B</td>
<td>4</td>
<td>F</td>
<td>CCTCGCATGTGCGGTCCTGAAACATT</td>
</tr>
<tr>
<td>ARL4A</td>
<td>B</td>
<td>4</td>
<td>R</td>
<td>AAACAAATGGCTAGACGCCCACATGC</td>
</tr>
</tbody>
</table>
References


Gluocorticoids modulate the pre-existing chromatin structure to regulate gene expression.


Mcdowell, I. In review. Glucocorticoid Receptor Recruits to Enhancers and Drives Activation by Motif-directed Binding.


mammalian regulatory sequences supports a flexible organizational model. *Nat Genet*, 45, 1021-1028.


STEVENS, A. 1960. Incorporation of the Adenine Ribonucleotide into Rna by Cell Fractions from, E-Coli B. *Biochemical and Biophysical Research Communications*, 3, 92-96.


Biography

Anthony Michael D’Ippolito was born on May 4, 1990 in Columbus, Ohio. He earned a Bachelor’s of Science in Molecular Genetics at The Ohio State University in 2012. As an undergraduate, Anthony performed research at the Center for Injury Research and Policy at Nationwide Children’s Hospital. His work there earned him a Young Epidemiology Regional Semifinalist Scholarship, and a primary author manuscript: “Epidemiology of pediatric holiday-related injuries presenting to US emergency departments”. Anthony also performed research in a molecular biology lab as an undergraduate. He worked under the mentorship of Dr. Christopher Phiel, at Nationwide Children’s Hospital, where he co-authored several papers. In addition, Anthony was awarded an internship at the Advanced Sequencing Applications group at Life Technologies in Carlsbad, California.

Anthony then joined the University Program in Genetics and Genomics at Duke University, and performed research under the mentorship of Dr. Tim Reddy. There he co-authored several papers, including a primary author manuscript titled “Glucocorticoid treatment mediates chromatin loop dynamics concordant with gene expression responses”. Anthony was also awarded the Ruth L. Kirschstein National Research Service Award from the NIAID, as well as an ASHG Charles J. Epstein Trainee Semifinalist Award for Excellence in Human Genetics Research.