Building the Interphase Nucleus: A study on the kinetics of 3D chromosome formation, temporal relation to active transcription, and the role of nuclear RNAs

Kristin N. Abramo

University of Massachusetts Medical School

Let us know how access to this document benefits you.
Follow this and additional works at: https://escholarship.umassmed.edu/gsbs_diss

Part of the Bioinformatics Commons, Cell Biology Commons, Computational Biology Commons, Genomics Commons, Laboratory and Basic Science Research Commons, Molecular Biology Commons, Molecular Genetics Commons, and the Systems Biology Commons

Repository Citation

Creative Commons License
This work is licensed under a Creative Commons Attribution-Noncommercial 4.0 License
This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in GSBS Dissertations and Theses by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.
BUILDING THE INTERPHASE NUCLEUS:
A STUDY ON THE KINETICS OF 3D CHROMOSOME FORMATION,
TEMPORAL RELATION TO ACTIVE TRANSCRIPTION,
AND THE ROLE OF NUCLEAR RNAs

A Dissertation Presented

By

KRISTIN N. ABRAMO

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

July 28, 2020

Program in Systems Biology, Interdisciplinary Graduate Program
BUILDING THE INTERPHASE NUCLEUS:

A STUDY ON THE KINETICS OF 3D CHROMOSOME FORMATION, TEMPORAL RELATION TO ACTIVE TRANSCRIPTION, AND THE ROLE OF NUCLEAR RNAs

A Dissertation Presented
By

KRISTIN N. ABRAMO

This work was undertaken in the Graduate School of Biomedical Sciences

Interdisciplinary Graduate Program

Under the mentorship of

______________________________________
Job Dekker, Ph.D., Thesis Advisor

______________________________________
Paul Kaufman, Ph.D., Member of Committee

______________________________________
Jeanne Lawrence, Ph.D., Member of Committee

______________________________________
Scot Wolfe, Ph.D., Member of Committee

______________________________________
Kenneth Zaret, Ph.D., External Member of Committee

______________________________________
Thomas Fazzio, Ph.D., Chair of Committee

______________________________________
Mary Ellen Lane, Ph.D.,
Dean of the Graduate School of Biomedical Sciences

July 28, 2020
Dedication

This thesis is dedicated to my grandparents. Omi and Opi, thank you for always encouraging me to reach for the stars and praising me along the way. Nana, thank you for being there at each step of my path and for always being ready to raise a glass to me. And lastly, to Papa. Though I’ve extremely missed your support and enthusiasm these past few years, I know you’re looking down on me, cracking jokes, singing songs, and making funny faces. I am thankful for the support from all four of my wonderful grandparents. I know you all care about me very much and I hope that I have made you proud over the years. I love you all so much. And I am happy to finally report that I’ll be done with school and you’ll finally have a doctor in the family! (But not that kind of doctor)
Acknowledgements

The work and ideas presented in this thesis were only possible due to the vast support of my advisor, colleagues, friends, and family throughout my graduate career.

I am honored to have had the privilege of being mentored by Job Dekker, PhD these past 5 years. Job’s passion for science and sharing knowledge is inspiring. I always felt more like a trusted colleague than a student due to the immense respect and scientific freedom Job provided me with. Thank you, Job, for teaching me, believing in me, challenging me, encouraging me, and ultimately helping me grow as a confident scientist.

To my Dissertation Examination Committee, thank you for agreeing to be a part of my final chapter with UMass Medical School. Thomas Fazzio, PhD, Paul Kaufman, PhD, Jeannie Lawrence, PhD, and Scot Wolfe, PhD, I am grateful that you have continued to work with and guide me since my qualifying exam in 2017.

To the Dekker lab, I am befuddled by how one lab is able to have so many smart, talented, supportive and fun scientists. Whether it was sharing the early quiet mornings in lab with Ye Zhan, Liyan Yang, and Johan Gibcus, PhD, taking Pilates/Barre classes in the afternoons with Nicki Fox, PhD, Erica Hildebrand, PhD, Betul Akgol Oksuz, and Houda Belaghzal, PhD, synchronizing over 120 dishes of cells with Anne-Laure Valton, PhD, welcoming Denis Lafontaine to the BB (best bay), learning ggplot tricks from Filipe Tavares-Cadete, PhD, drinking coffee with Yu (Sunny) Liu, PhD, forever losing my coffee mug to Sergey Venev, PhD, watching football (American and the “real” kind) with Bastiaan Dekker, skiing the Whistler bowls with Marlies Oomen, Gloucester swims with Allana Schooley, PhD, or break room chats with Hakan Ozadam, PhD, Ankita Nand, George Spracklin, PhD, Snehal Sambare, and Davood Norouzi, PhD, thank you for helping me
on my PhD journey. Special thanks to Nicki, Allana, and Marlies for proof/content reading parts of this thesis and helping in the preparation of my defense presentation.

To my other colleagues, particularly within the Program in Systems Biology, thank you for creating a fun and safe space to share, explore, criticize, and build ideas. Special shout out to Peter Cruz-Gordillo, Brent Horowitz, and Cedric (Rick) Diot, PhD mostly for providing distractions when warranted, but also for general support in science and life.

As a member of the Graduate Student Body Committee (GSBC) at UMass Med for 4 years, I took on many different hats and had the opportunity to interact with the many great students and faculty on campus. I learned so much from these great scientists. Most importantly, I was surrounded by an amazing support system of like-minded people going through similar graduate career experiences and I made lifelong friendships. Thank you to all of the students and faculty that I interacted with during my time serving on the GSBC.

Completing this graduate program would not have been possible without the constant support of my close friends inside and outside of the lab. Since day 1 of classes, I was fortunate to be surrounded by fun, supportive, and extremely smart people. Through our annual MLK weekend ski trips and the smaller events/conversations in between, this group has always been there when I’m stuck in science or need a break. I’ve enjoyed our weekly Zoom happy hours during the pandemic and nothing can stop these friendships into the future. Thank you all. Special thanks to Anne Carlisle and Meghan Spears for always being amazing and providing proper mental support and encouragement.

And finally, I would like to thank my family. First, to my partner in life and science, David Stelter, PhD, thank you for your constant support. Though we’re in very different scientific fields, we are able to discuss our science together and push each other to be
our best. David lifts me up when I get home frustrated from confusing experimental results and he celebrates with me when I have new and exciting results. He even proofread sections of this thesis! Most important, David keeps me grounded. I would not trade our weekend adventures into the White Mountains for anything in the world. I’m thankful that I have you each and every day to make life ever better.

To David’s parents, Eric and Nancy Stelter, thank you for accepting me into your family and pushing me just as much as you push David. Though mostly from far away, I appreciate all of the support you provide when we talk or are able to visit.

To my brother and sister-in-law, Anthony and Jenna Abramo, thank you for reminding me to have fun and just enjoy life. Having a big brother has certainly had its challenges, but you’re really not (that) bad and I value you and Jenna’s support.

To Asbury, the best dog ever. Asbury was the poster dog that I had been waiting for all through my childhood. Her beautiful, stereotypic Saint Bernard features made her a hit everywhere we went and she loved to take the attention. She always knew when I needed her though and would lay with me, well on me, while I relaxed and forgot about everything else. I was lucky to have her and will never forget her. Love ya berry.

Last and certainly not least, to my parents, Paul and Lorraine Abramo. Thank you for being a constant presence of support and encouragement throughout my graduate career and beyond. I appreciate that although you didn’t quite know what I was getting myself into, you took the time to understand and always showed up as my number one fans. I am extremely lucky to have you both in my life and even more fortunate that you always believe in me, support me, and love me. I could not have gotten to this stage in life and in my career without either of you.
Abstract

Following the discovery of the one-dimensional sequence of human DNA, much focus has been directed on microscopy and molecular techniques to learn about the spatial organization of chromatin in a 3D cell. The development of these powerful tools has enabled high-resolution, genome-wide analysis of chromosome structure under many different conditions. In this thesis, I focus on how the organization of interphase chromatin is established and maintained following mitosis. Mitotic chromosomes are folded into helical loop arrays creating short and condensed chromosomes, while interphase chromosomes are decondensed and folded into a number of structures at different length scales ranging from loops between CTCF sites, enhancers and promoters to topologically associating domains (TADs), and larger compartments. While the chromatin organization at these two very different states is well defined, the transition from a mitotic to interphase chromatin state is not well understood.

The aim of this thesis is to determine how interphase chromatin is organized following mitotic chromosome decondensation and to interrogate factors potentially responsible for driving the transition. First, I determine the temporal order with which CTCF-loops, TADs, and compartments reform as cells exit mitosis, revealing a unique structure at the anaphase-telophase transition never observed before. Second, I test the role of transcription in reformation of 3D chromosome structure and show that active transcription is not required for the formation of most interphase chromatin features; instead, I propose that transcription relies on the proper formation of these structures. Finally, I show that RNA in the interphase nucleus can be degraded with only slight consequences on the overall chromatin organization, suggesting that once interphase
chromatin structures are achieved, the structures are stable and RNA is only required to reduce the mixing of active and inactive compartments. Together, these studies further our understanding of how interphase structures form, how these structures relate to functional activities of the interphase cell, and the stability of chromatin structures over time.
# Table of Contents

Title Page ................................................................................................. i
Reviewer Page .......................................................................................... ii
Dedication .................................................................................................. iii
Acknowledgements ................................................................................... iv
Abstract .................................................................................................... vii
Table of Contents ...................................................................................... ix
List of Figures ........................................................................................... xiv
List of Tables ............................................................................................. xvii
Copyright Material ................................................................................... xviii
List of Abbreviations ................................................................................. xix

CHAPTER I: Introduction ........................................................................... 1

Introduction to the 3D genome .................................................................. 1

Methods to study the 3D genome .............................................................. 2
Microscopy approaches to genome organization .................................... 2
Chromosome conformation capture (3C)-based technologies ...................... 3

3D genome throughout the cell cycle ....................................................... 8
Chromatin organization in mitosis ........................................................... 8
Mechanisms of mitotic chromosome folding ............................................ 9
Chromatin organization in G1 ................................................................. 11
Compartments .......................................................................................... 11
Topologically associated domains ......................................................... 14
Dots/loops ............................................................................................... 15
Chromatin organization in S phase and G2 ............................................ 17
Mechanisms of interphase chromosome folding .................................... 18
Formation of TADs and dots by loop extrusion ..................................... 18
Formation of compartments by microphase separation .......................... 20

Roles of chromosome folding ................................................................. 22
Compartments and disease ..................................................................... 22
Topologically associated domains and disease ....................................... 23
Overview of this thesis .................................................................................................................. 24

CHAPTER II: Transition from Mitosis to Interphase 3D Genome Organization .... 26

Preface ............................................................................................................................................ 26

Abstract ........................................................................................................................................ 26

Introduction .................................................................................................................................... 27

Materials and Methods .................................................................................................................. 29

Cell Culture .................................................................................................................................... 29
Creation of Stable HeLaS3-NCAPH-dTomato Cell Line ................................................................. 29
Mitotic Synchronization .................................................................................................................. 31
Mitotic Release Timecourse ............................................................................................................ 31
Cell Cycle Analysis ....................................................................................................................... 32
Hi-C Protocol ................................................................................................................................. 32
Hi-C Data Processing ...................................................................................................................... 33
Contact probability (P(s)) plots & derivatives ................................................................................ 35
Compartment analysis .................................................................................................................... 35
TAD analysis .................................................................................................................................... 36
Loop analysis .................................................................................................................................... 37
Simulated Hi-C mixture datasets .................................................................................................... 38
Microscopy ...................................................................................................................................... 39

Immunofluorescence staining ......................................................................................................... 39

Cell Cycle Classification ................................................................................................................. 39
Protein Localization ....................................................................................................................... 40
Chromatin association .................................................................................................................... 41

Fractionation protocol .................................................................................................................... 41
Western Blots ................................................................................................................................... 42
Quantification ................................................................................................................................... 43
Statistics and Reproducibility ......................................................................................................... 43
Code Availability ............................................................................................................................ 43
Data Availability ............................................................................................................................. 43

Results ............................................................................................................................................... 44

Synchronous entry into G1 .............................................................................................................. 44
Chromosome conformational changes as cells enter G1 ................................................................. 44
Compartmentalization occurs slower than formation of TADs and loops ......................................... 47
TADs and loops form prior to G1 entry ............................................................................................ 59
An intermediate folding state during mitotic exit ............................................................................ 62
The transient intermediate folding state occurs during telophase ................................................ 71
Condensin unloading occurs during telophase while cohesin loading occurs during cytokinesis ...................................................................................................................................... 74
Inhibiting transcription as cells enter G1 
Measuring transcription during G1 entry 
Loops are maintained in nonsynchronous cells treated with transcription inhibitors 
TADs are weakened upon inhibiting transcription in nonsynchronous cells 
Compartments are slightly reduced upon inhibiting transcription in nonsynchronous cells 
Chromosome conformational changes in cells with no active transcription 
Inhibition in nonsynchronous cells 
Inhibition in cells entering G1 
Mitotic Synchronization 
Nascent RNA analysis 
Purification of nascent RNA 
Fragment Analyzer 
RT-qPCR 
Cell Cycle Analysis 
Hi-C Analysis 
Hi-C Data Processing 
Contact probability (P(s)) plots & derivatives 
Compartment analysis 
TAD analysis 
TSS pileups 
Loop analysis 
Code Availability 

Results 
Inhibiting transcription in nonsynchronous cells 
Chromosome conformational changes in cells with no active transcription 
Compartments are slightly reduced upon inhibiting transcription in nonsynchronous cells 
TADs are weakened upon inhibiting transcription in nonsynchronous cells 
Insulation around TSSs is reduced upon inhibiting transcription in nonsynchronous cells 
Loops are maintained in nonsynchronous cells treated with transcription inhibitors 
Measuring transcription during G1 entry 
Inhibiting transcription as cells enter G1
Chromosome conformational changes in cells entering G1 with no active transcription ................................................................. 124
Compartmentalization is reduced for cells entering G1 in the absence of active transcription ......................................................... 127
TADs are weaker in cells entering G1 without active transcription .......... 137
Insulation around TSSs is weaker in cells entering G1 in the absence of transcription ........................................................................ 141
Loops are established in cells entering G1 without active transcription .......... 144

Conclusion ........................................................................................................ 146

CHAPTER IV: Role of RNA in chromatin organization ........................................ 150

Abstract ........................................................................................................ 150

Introduction ..................................................................................................... 151

Materials and Methods ...................................................................................... 153
Cell Culture ...................................................................................................... 153
RNase Treatment ............................................................................................. 153
RNA isolation for fragment analyzer analysis .................................................... 154
Microscopy ...................................................................................................... 154
Hi-C Analysis .................................................................................................. 155
Hi-C Data Processing ...................................................................................... 156
Inter-chromosomal interaction analysis ............................................................ 157
Contact probability (P(s)) plots & derivatives .................................................... 157
Compartment analysis ...................................................................................... 157
iMARGI chromatin associated RNA (caRNA) analysis .................................... 158
TAD analysis .................................................................................................... 159
Loop analysis .................................................................................................... 160
Code Availability ............................................................................................. 160

Results .............................................................................................................. 161
Degrading RNA in nonsynchronous cells ........................................................ 161
Inter-chromosomal interactions change in cells with degraded RNA ............. 163
Chromosome conformational changes in cis in cells with degraded RNA ........ 167
Compartment changes occur in genomic regions thought to interact with caRNAs ............................................................................ 179
TADs are maintained in cells with degraded RNA ............................................ 186
Loops are weaker in cells with degraded RNA ................................................ 190

Conclusion ........................................................................................................ 194

CHAPTER V: Discussion .................................................................................... 197
Summary ......................................................................................................................................................... 197
Condensation ↔ Decondensation: Same processes in reverse? .............................................................. 198
Enhancer-independent transcription ............................................................................................................... 203
Replication Domains and Compartmentalization ......................................................................................... 206
The interplay of TADs, compartments, and transcription ........................................................................... 209
RNA and interphase structures ....................................................................................................................... 211
Concluding remarks ...................................................................................................................................... 213

Bibliography .................................................................................................................................................. 214
List of Figures

FIGURE 2.1: Hi-C analysis during mitotic exit and G1 entry ........................................ 45
FIGURE 2.2: Cell cycle analysis of mitotic exit time courses ........................................ 46
FIGURE 2.3: Compartment analysis for time course replicate 1 ................................. 48
FIGURE 2.4: Kinetics of loop, TAD, and compartment formation................................. 51
FIGURE 2.5: TAD and compartment domain boundaries form with similar kinetics..... 53
FIGURE 2.6: Kinetics of A and B compartment formation at various genomic distances ................................................................. 55
FIGURE 2.7: Compartment analysis by distance for chromosome 14 ....................... 57
FIGURE 2.8: TADs and loops form quicker than expected, while compartmentalization occurs slower than expected................................................................. 60
FIGURE 2.9: Kinetics of loop formation for loops of different size............................. 63
FIGURE 2.10: Analysis of time course replicate 2 ....................................................... 64
FIGURE 2.11: Analysis of time course replicate 3 ....................................................... 66
FIGURE 2.12: Formation of a transient folding intermediate ..................................... 68
FIGURE 2.13: Chromatin colocalization dynamics of condensins and cohesin during mitotic exit ............................................................................................................. 72
FIGURE 2.14: Chromatin association dynamics of CTCF, condensin, and cohesin .... 75
FIGURE 2.15: Chromatin association dynamics of condensins and cohesin during mitotic exit ............................................................................................................. 77
FIGURE 3.1: Transcription inhibition in nonsynchronous cell cultures .................... 98
FIGURE 3.2: Hi-C analysis of nonsynchronous cell cultures treated with transcription inhibitors ............................................................................................................ 100
FIGURE 3.3: Compartment boundaries and type are maintained in nonsynchronous cells treated with transcription inhibitors .............................................................................. 103
FIGURE 3.4: Compartmentalization is reduced in nonsynchronous cells treated with transcription inhibitors ............................................................................................................ 104
FIGURE 3.5: Compartmentalization using PC1 tracks generated from each Hi-C dataset also shows reduced strength in nonsynchronous cells treated with transcription inhibitors ............................................................................................................ 106
FIGURE 3.6: Compartmentalization is reduced mostly at long range in nonsynchronous cell cultures treated with transcription inhibitors ............................................................................................................ 107
FIGURE 3.7: Trans compartmentalization is maintained for nonsynchronous cells treated with transcription inhibitors ................................................................. 109

FIGURE 3.8: Insulation at TAD boundaries is reduced in nonsynchronous cells treated with transcription inhibitors .................................................................................. 111

FIGURE 3.9: Insulation at active TSSs is reduced in nonsynchronous cells treated with transcription inhibitors .................................................................................. 114

FIGURE 3.10: Loops are maintained in nonsynchronous cells treated with transcription inhibitors .................................................................................. 116

FIGURE 3.11: Kinetics of loop, TAD, and compartment formation in relation to kinetics of transcription activity upon G1 entry .................................................................. 119

FIGURE 3.12: Transcription inhibition in cells released from prometaphase arrest and entering G1 .................................................................................. 122

FIGURE 3.13: Hi-C analysis of cell cultures entering G1 in the absence of active transcription .................................................................................. 125

FIGURE 3.14: Compartment boundaries and type are properly established in cells entering G1 in the absence of active transcription .................................................................. 129

FIGURE 3.15: Compartmentalization is weaker in cells entering G1 in the absence of active transcription .................................................................................. 131

FIGURE 3.16: Compartmentalization is weaker in cells entering G1 in the absence of active transcription (replicate 2) .................................................................................. 132

FIGURE 3.17: Compartmentalization using PC1 tracks generated from each Hi-C dataset also shows weaker strength in cells entering G1 in the absence of active transcription .................................................................. 134

FIGURE 3.18: Compartmentalization is weaker at all distances in cells entering G1 in the absence of active transcription .................................................................................. 134

FIGURE 3.19: Trans compartmentalization is quicker in cells entering G1 in the absence of active transcription, however, weakens by 9 hours release from prometaphase arrest when compartmentalization should be fully established .................................................................. 136

FIGURE 3.20: Insulation at TAD boundaries is weaker in cells entering G1 in the absence of active transcription .................................................................................. 138

FIGURE 3.21: Insulation at active TSSs is weaker in cells entering G1 in the absence of active transcription .................................................................................. 142

FIGURE 3.22: Loops are established in cells entering G1 in the absence of active transcription .................................................................................. 145

FIGURE 4.1: RNA degradation in nonsynchronous cells .................................................................................. 162
FIGURE 4.2: Inter-chromosomal 3D chromatin contacts are altered in cells with degraded RNA ............................................................... 164
FIGURE 4.3: Hi-C analysis of cells with degraded RNA ........................................... 168
FIGURE 4.4: Compartment boundaries and type are maintained in cells with degraded RNA ............................................................................................................. 171
FIGURE 4.5: Compartmentalization is weakened in cells with degraded RNA ........ 174
FIGURE 4.6: Trans compartmentalization is weakened in cells with degraded RNA ... 178
FIGURE 4.7: Compartment types of chromatin regions associated with RNAs change in cells with degraded RNA ............................................................................................................. 181
FIGURE 4.8: Insulation at TAD boundaries increases for cells treated with RNase A, then decreases in cells incubated for 4 hours regardless of RNA presence .......... 187
FIGURE 4.9: Loops are weaker in cells treated with RNase A for 5 minutes, decrease overall for cells incubated for 4 hours, and are stronger for cells without RNA for 4 hours versus with RNA present .......................................................................................................................... 191
# List of Tables

Table 2.1: gRNA sequences for genome editing .......................................................... 30  
Table 2.2: PCR primers for NCAPH homology arms .................................................. 30  
Table 2.3: gBlock Gene Fragment ................................................................................ 30  
Table 2.4: Hi-C library statistics ................................................................................... 34  
Table 3.1: Firefly2 sequence ......................................................................................... 89  
Table 3.2: qPCR primers for measuring transcription inhibition ................................ 90  
Table 3.3: Hi-C library statistics ................................................................................... 93  
Table 3.4: Overview of chromatin organization in transcription inhibited cells ....... 147  
Table 4.1: Hi-C library statistics ................................................................................... 157
Copyright Material

The content and figures used in Chapter II of this dissertation have been previously published at the below citation. According to the publisher's copyright policies, ownership of copyright remains with me, the Author, and I retain the right to reproduce the work in my thesis dissertation. All other work presented in this thesis is currently unpublished.

### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3C</td>
<td>chromosome conformation capture</td>
</tr>
<tr>
<td>4C</td>
<td>chromosome conformation capture-on-chip</td>
</tr>
<tr>
<td>5C</td>
<td>chromosome conformation capture carbon copy</td>
</tr>
<tr>
<td>caRNA</td>
<td>chromatin-associated RNA</td>
</tr>
<tr>
<td>ChIA-PET</td>
<td>chromatin interaction analysis by paired-end tag sequencing</td>
</tr>
<tr>
<td>CT</td>
<td>chromosome territory</td>
</tr>
<tr>
<td>CTD</td>
<td>C-terminal domain (referred to in regards to ‘tail’ of RNAPII protein)</td>
</tr>
<tr>
<td>DRB</td>
<td>5,6-Dichlorobenzimidazole 1-ß-D-ribofuranoside</td>
</tr>
<tr>
<td>eRNA</td>
<td>enhancer RNA</td>
</tr>
<tr>
<td>FACS</td>
<td>fluorescence-activated cell sorting</td>
</tr>
<tr>
<td>FISH</td>
<td>fluorescence in situ hybridization</td>
</tr>
<tr>
<td>HP1</td>
<td>heterochromatin protein 1</td>
</tr>
<tr>
<td>Ig</td>
<td>immunoglobulin</td>
</tr>
<tr>
<td>iMARGI</td>
<td>in situ mapping of RNA-genome interactome</td>
</tr>
<tr>
<td>LAD</td>
<td>lamin-associated domain</td>
</tr>
<tr>
<td>LCR</td>
<td>locus control region</td>
</tr>
<tr>
<td>LLPS</td>
<td>liquid-liquid phase separation</td>
</tr>
<tr>
<td>IncRNA</td>
<td>long non-coding RNA</td>
</tr>
<tr>
<td>MNase</td>
<td>micrococcal nuclease</td>
</tr>
<tr>
<td>NAD</td>
<td>nucleolus-associated domain</td>
</tr>
<tr>
<td>NEB</td>
<td>nuclear envelope breakdown</td>
</tr>
<tr>
<td>ncRNA</td>
<td>non-coding RNA</td>
</tr>
<tr>
<td>PC1 / PC2</td>
<td>principal component 1 / 2 (from eigen vector decomposition)</td>
</tr>
<tr>
<td>PRC1 / PRC2</td>
<td>polycomb repressive complex 1 / 2</td>
</tr>
<tr>
<td>PRO-seq</td>
<td>precision nuclear run-on sequencing</td>
</tr>
<tr>
<td>RNAPII</td>
<td>RNA polymerase II</td>
</tr>
<tr>
<td>RNasin</td>
<td>RNase Inhibitor</td>
</tr>
<tr>
<td>SINE</td>
<td>short interspersed nuclear elements</td>
</tr>
<tr>
<td>SMC</td>
<td>structure maintenance of chromosomes</td>
</tr>
<tr>
<td>TAD</td>
<td>topologically associated domain</td>
</tr>
<tr>
<td>T-ALL</td>
<td>T cell acute lymphoblastic leukemia</td>
</tr>
<tr>
<td>T + D</td>
<td>Triptolide + DRB treatment condition (Chapter III)</td>
</tr>
<tr>
<td>TSS</td>
<td>transcription start site</td>
</tr>
<tr>
<td>ZGA</td>
<td>zygotic genome activation</td>
</tr>
</tbody>
</table>
CHAPTER I: Introduction

Introduction to the 3D genome

While sequencing the human genome was groundbreaking, it quickly became clear that the one-dimensional sequence was not a tell-all towards understanding gene regulation and disease. The human genome contains over 3 billion nucleotides split across 23 pairs of chromosomes. This turns out to be roughly 2 meters of DNA that must be contained, and able to function, within a spherical cell nucleus of roughly only 10 microns in diameter. Therefore, it was discovered that the human genome, as well as any other genome, could not simply exist as a one-dimensional polymer, but must actually be intricately folded into a three-dimensional structure capable of maintaining its ability to function. This implies that the genome must be folded in such a way to ensure proper gene expression via enhancer-promoter interactions and access to transcriptional machinery, as well as forming an organized, unentangled structure to ensure proper replication and chromosome segregation during cell division. The development of various imaging and genomic approaches has revealed the 3D organization of chromatin and how it relates to cellular functions.
Methods to study the 3D genome

Microscopy approaches to genome organization

The idea that chromosomes are organized into spatially distinct domains throughout the cell cycle has been around for a long time and was confirmed through microscopy techniques visualizing UV irradiation, pulse-labeling of nucleosides, and Giemsa-banding (Frenster et al., 1963; Zorn et al., 1979; Stack et al., 1977). In fact, these early experiments established that the interphase nucleus organizes such that repressed heterochromatin is localized to the periphery segregated from euchromatic regions at the nuclear center. The introduction of DNA fluorescence in situ hybridization (DNA-FISH) and chromosome painting soon after revolutionized research on genome organization and allowed for visualization of individual chromosomes (Gall and Pardue, 1969; Speicher et al., 1996). These experiments confirmed the previous theories and further showed that chromosomes occupy discrete domains, termed ‘chromosome territories’ (CT), with specific radial positioning (Cremer et al., 1982; Cremer and Cremer, 2001; Boyle et al., 2001; Tanabe et al., 2002; Bolzer et al., 2005). For example, small gene-rich chromosomes, such as chromosome 19, tend to be positioned at the center of the nucleus, while gene-poor nuclei, such as chromosome 18, are positioned closer to the nuclear periphery (Croft et al., 1999; Boyle et al., 2001; Tanabe et al., 2002). Interestingly, upon activation, genes can loop out of their chromosome territory creating some intermingling at the periphery of CTs and causing actively transcribed genes to colocalize (Branco and Pombo, 2006; Chambeyron and Bickmore, 2004; Osborne et al., 2004). Further, the nuclear position of loci was observed to be related to transcriptional activity. This was best observed for developmentally regulated genes, such as immunoglobulin
(Ig) loci and the β-globin locus, which are localized to the nuclear periphery in progenitor cells but with maturation these loci move towards the nuclear interior accompanied by an increase in transcription (Kosak et al., 2002; Ragoczy et al., 2006). Additionally, loci at the nuclear periphery were found to interact with lamina proteins and targeting regions to the nuclear periphery could induce relocalization and transcriptional repression of the specific locus as well as the neighboring genes (Reddy et al., 2008). While DNA-FISH allowed for three-dimensional measurements between chromatin regions and nuclear bodies, or between multiple loci, only a limited number of probes could be used at one time. More recent approaches have improved on the capability to map the spatial position of many genomic regions simultaneously, including within live cells (Wang et al., 2016; Ma et al., 2013; Ma et al., 2015; Ma et al., 2016). These improved imaging techniques reveal valuable insight on chromosome organization and have re-created interaction maps similar to what is observed using proximity ligation techniques, described in the following section. In brief, imaging approaches have confirmed the looping interactions between enhancers and promoters observed by chromosome conformation capture (3C) (Dekker et al., 2002; Barbieri, Xie et al., 2017; Beagrie, Scialdone et al., 2017), the presence of globular topologically associated domains (TADs) (Dixon et al., 2012; Nora et al., 2012; Szabo et al., 2018; Mateo et al., 2019), and the formation of spatially segregated compartment domains observed by Hi-C (Lieberman-Aiden, Berkum et al., 2009; Boettiger et al., 2016; Smeets et al., 2014; Wang et al., 2016).

**Chromosome conformation capture (3C)-based technologies**

While DNA-FISH and other imaging based methods were able to reveal fundamental aspects of chromosome organization, early techniques only allowed for very
low-throughput analyses. Therefore, biochemical approaches utilizing proximity ligation followed by DNA sequencing became a quicker and more parallel method to investigating chromosome structures. Most notably, the development of chromosome conformation capture (3C) and its refined derivatives of 4C, 5C, and Hi-C, enabled systematic analysis of contact frequencies of genomic loci in cell populations (Dekker et al., 2002; Simonis et al., 2006; Zhao, Tavoosidana et al., 2006; Dosti et al., 2006; Lieberman-Aiden, Berkum et al., 2009). These techniques were designed to probe the three-dimensional organization of chromosomes through a similar method from fixation of living cells. First, chromatin interactions are covalently crosslinked using a fixative such as formaldehyde to capture 3D contacts mediated by proteins, DNA, and RNA. Next, crosslinked chromatin is fragmented by a restriction enzyme to leave sticky DNA ends that can then re-ligate with nearby proximal chromatin regions in 3D space. This results in a library of chromatin fragments close together in 3D space and downstream analyses are then used to probe proximity in the linear genome. In the original 3C technique, this is done using PCR to search for specific ligation products and determine how frequently two known genomic regions interact (Dekker et al., 2002). We therefore term 3C as a method to study ‘one vs. one’ genomic loci and note that this approach requires prior knowledge of target sequences. Studies using 3C revealed looping interactions between transcriptional elements separated by great distances in the linear DNA sequence (Tolhuis, Palstra et al., 2002), which agrees with colocalization of distal genes in imaging studies (Osborne et al., 2004).

Several techniques have followed to build on the original ‘one vs. one’ 3C method. For example, 4C, 5C, and Hi-C variants allowed for study of ‘one vs. all’, ‘many vs. many’,
and 'all vs. all' genomic contacts, respectively, and each contributed new information on 3D chromosome structure (Simonis et al., 2006; Zhao, Tavoosidana et al., 2006; Dosti et al., 2006, Lieberman-Aiden, Berkum et al., 2009). 4C analysis built on the previous 3C technique by using a viewpoint at a region of interest ('one') to look at all other interacting regions ('many') (Simonis et al., 2006; Zhao, Tavoosidana et al., 2006). Studies using 4C revealed that both active and inactive loci have multiple long-range contacts with chromatin regions mostly on the same chromosome, i.e. intra-chromosomal cis interactions, but also in trans with other (inter-) chromosomes, and further, that loci tended to cluster together depending on their transcriptional activity, especially co-transcriptionally regulated genes (Simonis et al., 2006; Tolhuis et al., 2011). Around the same time, Dostie et al. presented 3C-Carbon Copy (5C) which utilizes deep-sequencing technology for a 'many vs. many' approach to obtain high-resolution information on all contacts with a large genomic region of interest (Dostie et al., 2006). 5C revealed old and new chromatin looping interactions, especially between long-range enhancer-promoter interactions, and the discovery of topologically associated domains (Sanyal, Lajoie et al., 2012; Nora et al., 2012). For more details on the 3D chromosome organization of TADs, see “Chromatin organization in G1” below. With slight variations to 3C and 5C methods, Capture-C was introduced and allowed for the enrichment of specific genomic loci or entire regions prior to the PCR amplification steps to get a better resolution at the specific ‘viewpoint’ or interest (Dryden et al., 2014; Hughes et al., 2014).

Finally, Hi-C utilized massively parallel sequencing for an ‘all vs. all’ approach to identify the frequency of any genomic region interacting with any other enriched-3C chromatin fragment (Lieberman-Aiden, Berkum et al., 2009). Comprehensive mapping by
Hi-C identified the spatially distinct compartments of euchromatic, active, gene-dense regions and heterochromatic, inactive, gene-poor chromatin regions. More details on chromosome organization into compartments is in the “Chromatin organization in G1” section below. Further, within the megabase spanning compartments, Hi-C (along with the previously mentioned 5C studies) revealed that chromatin organizes non-randomly into TADs (Dixon et al., 2012; Sexton, Yaffe et al., 2012). Applying this method, ChIA-PET (chromatin interaction analysis by paired-end tag sequencing) was established which allowed for the enrichment of Hi-C ligation products that were bound by proteins of interest, such as estrogen receptor, CTCF, and RNA polymerase II (Fullwood et al., 2009; Handoko et al., 2011; Li, Ruan, Auerbach, Sandhu et al., 2012). Further variants have been established from the original Hi-C protocol including updates for mammalian in situ methods (Belton et al., 2012; Rao, Huntley et al., 2014; Belaghzal et al., 2017), higher resolution chromatin maps using ‘Micro-C’ (Hsieh et al., 2015; Hsieh et al., 2016; Krietenstein et al., 2020; Hsieh et al., 2020), analysis on interaction dynamics by ‘liquid chromatin Hi-C’ (Belaghzal, Borrman et al., 2019), analysis on chromatin interactions between and within sister chromatids by ‘SisterC’ (Oomen et al., 2020; Espinosa et al., 2019; Mitter et al., 2020), and observations of multiple contacts instead of just pairwise interactions (‘C-walks’: Olivares-Chauvet et al., 2016; ‘SPRITE’: Quinodoz et al., 2018; ‘MC-3C’: Tavares-Cadete, Norouzi et al., 2020). These Hi-C based methods currently remain the most frequently used 3C-based technologies due to (1) the requirements of little to no prior knowledge of chromatin interactions, (2) use of a low number of cells, and (3) the creation of high-resolution genome wide chromatin contact-frequency maps.
Notably, the choice of restriction enzyme for fragmentation of crosslinked chromatin can greatly impact the resulting interaction library. These protocols have been carried out with enzymes that recognize a variable number of base pairs (bp), most often 4 or 6 bp, such as HindIII, BglII, BamHI, EcoRI, Acil, and DpnII (Tolhuis, Palstra et al., 2002; Tan-Wong et al., 2008; Palstra, Tolhuis et al., 2003; Miele et al., 2009; Comet et al., 2011). Enzymes with a larger recognition site have a lower frequency of cutting throughout the genome, and therefore, produce fewer chromatin fragments, have fewer possible pairwise interactions captured by re-ligation, and generate lower resolution information of chromatin contacts. In fact, a recent advancement in these protocols uses micrococcal nuclease (MNase) enabling nucleosome-resolution information on chromatin folding (Hsieh et al., 2015; Hsieh et al., 2016; Krietenstein et al., 2020; Hsieh et al., 2020).

While Hi-C and related proximity ligation techniques discussed above reveal interesting insights into topological features of the nucleus, these methods represents a genome-wide population average of contact frequencies. Methods for single cell Hi-C variations have been established (Nagano, Lubling, Stevens et al., 2013; Nagano et al., 2015; Tan, Xing et al., 2018) but still fall inferior to the observations visualized in a single cell by imaging techniques. This is especially important as vast heterogeneity has been observed for chromosome organization inside single cells (Finn et al., 2019). The combination of 3C-based and microscopy approaches, therefore, is necessary as we continue to uncover more details of the elaborate chromosome folding and organization within the nucleus.
3D genome throughout the cell cycle

Both morphology and function of chromosomes change drastically over the cell cycle. Through a combination of imaging and biochemical approaches, we now know a lot about the organization of chromosomes at each step along the way. For each cell cycle stage in the sections below, chromatin organization and its relation to functional activities of the cell is discussed, as well as possible mechanisms for the establishment of organizational features.

Chromatin organization in mitosis

The proper segregation of the genome from mother to daughter cells relies on the correct formation of mitotic chromosomes for cell division. To achieve this, chromosomes undergo a vast compaction in length (more than 100x) at mitotic entry and erase the interphase organizational features of compartments and TADs (Naumova, Imakaev, Fudenberg et al., 2013; Gibcus, Samejima, Goloborodko et al., 2018). The disappearance of TADs is not surprising since the key regulators of TADs, CTCF and cohesin, are displaced from chromatin during mitosis (Oomen et al., 2019). In contrast, it is surprising that in the absence of compartments and TADs, which correlate with and potentially drive gene regulation in interphase, some transcription start sites remain accessible and a low level of transcription is maintained during mitosis (Oomen et al., 2019; Palozola et al., 2017; Hsiung, Morrissey et al., 2015).

In addition to the disappearance of interphase chromatin organization, highly synchronized Hi-C analysis on DT-40 chicken cells revealed the appearance of a second diagonal on the chromatin interaction map (Gibcus, Samejima, Goloborodko et al., 2018). Further, the distance of this additional diagonal moved further away as cells progress
through mitosis (Gibcus, Samejima, Goloborodko et al., 2018). Together with computational modeling, these data suggest that mitotic chromosomes are a helical loop array of chromatin folded similar to a spiral staircase. This is the result of the dual activity of condensin I and condensin II. Depletion of each condensin separately revealed that prometaphase chromosome compaction is achieved through the formation of small chromatin loops (~80 kb) formed by condensin I which are encompassed within larger condensin II formed loop (~400 kb) (Gibcus, Samejima, Goloborodko et al., 2018). The second diagonal moves to larger distances with further compaction, and therefore, additional shortening of the mitotic chromosomes.

These studies confirmed earlier observations of a mitotic scaffold from which chromatin loops rotate about. Condensins and topoisomerase II were thought to be major components of the axial scaffold (Earnshaw et al., 1985; Gasser et al., 1986; Hirano and Mitchison, 1994; Saitoh et al., 1994; Hirano et al., 1997). More recent studies have revealed the precise localization and quantity of axial proteins on mitotic chromosomes (Walther et al., 2018). While both condensin I and condensin II were localized to the chromosome axis, condensin II was found in closer proximity to the axial core with ~4x less quantity of protein than condensin I. Further studies showed that condensin I is sufficient to form a compacted mitotic chromosome in an ATP-dependent manner (Strick et al., 2004; Shintomi et al., 2015; Eeftens et al., 2017).

**Mechanisms of mitotic chromosome folding**

The presence of chromatin loops in mitotic chromosomes has been observed several times, leading to the idea that chromosomes are compacted for mitosis by a loop extrusion mechanism (Paulson and Laemmli, 1977; Marsden and Laemmli, 1979; Kimura
et al., 1999; Nasmyth, 2001; Maeshima and Laemmli, 2003; Maeshima et al., 2005; Alipour and Marko, 2012, Naumova, Imakaev, Fudenberg et al., 2013; Liang et al., 2015). The loop extrusion mechanism proposes that through the binding of a ‘loop extrusion’ machine to the DNA polymer, this protein complex can then reel chromatin from each side producing a progressively larger loop and connecting more distant elements of the linear DNA strand (Goloborodko et al., 2016a; Goloborodko et al., 2016b). Computational simulations demonstrate that this would form an array of loops from a central axis, which is consistent with the organization of mitotic chromosomes observed by Hi-C (Naumova, Imakaev, Fudenberg et al., 2013; Gibcus, Samejima, Goloborodko et al., 2018). Further the presence and movement of a second diagonal on the Hi-C interaction maps of chromosomes entering mitosis is consistent with the formation of progressively larger extruded loops, nesting of loops, and the helical rotation of loops around a central axis (Gibcus, Samejima, Goloborodko et al., 2018). This is supported by the individual depletion of condensin I and condensin II. These experiments show that condensin I extrudes the small chromatin loops which are nested inside of larger condensin II extruded loops. The ability of condensin to act as the loop extruder in this model suggests that condensin is a DNA motor, which is supported by the ability of condensin to compact mitotic chromosomes in an ATP-dependent manner (Strick et al., 2004; Shintomi et al., 2015; Eeftens et al., 2017; Terakawa, Bisht, Eeftens et al., 2017; Ganji et al., 2018). Interestingly, the most elaborate of these studies revealed that condensin formed loops are asymmetric, indicating that loop extrusion in mitosis is only one sided (Ganji et al., 2018).
Chromatin organization in G1

Chromatin is often thought of as an unentangled polymer, possibly similar to a ‘fractal globule’ from polymer physics. This unknotted state is consistent with the decay seen when plotting contact frequency and linear DNA separation from Hi-C data \( P(s) \) (Lieberman-Aiden, Berkum et al., 2009; Naumova, Imakaev, Fudenberg et al., 2013; Mirny, 2011). Being in this state would allow for chromatin regions to unfold and refold as necessary for transcription activation and DNA replication. Further, this supports the idea of specific chromatin organizational units, such as chromosome territories described previously. Three features of G1 chromatin organization are described in detail in this section: (1) compartments, (2) topologically associating domains, and (3) dots or loops between specific loci.

Compartments

Within each chromosome territory (CT), chromatin is further compartmentalized by the spatial segregation of euchromatic regions from heterochromatic regions (Lieberman, Berkum et al., 2009). We refer to these as A- and B-type compartments, respectively. A compartments are categorized as regions with high GC content, gene-dense, and relatively “open” chromatin regions supported by transcriptional activity and DNase I hypersensitivity (Simonis et al., 2006; Lieberman-Aiden, Berkum et al., 2009; Hou, Li, Qin et al., 2012; Sexton, Yaffe et al., 2012). B compartments, in contrast, are regions with low GC content, gene-poor, and relatively “closed” chromatin regions with low transcriptional activity. Compartment formation is an obvious feature of a Hi-C chromatin interaction map due to the ‘plaid’ or ‘checker-board’ like pattern. This is representative of an alternating presence of these two chromatin types in the linear DNA sequence and the preference
for interactions to occur between similar compartment types. Compartment domains are on the megabase scale, often ranging from 1-10 Mb in the linear DNA sequence with an average 3 Mb domain in human cells. Notably, the plaid pattern extends into the less frequent inter-chromosomal (trans) interactions of G1 Hi-C chromatin interaction maps supporting the idea that intermingling between CTs does occur, but it is often between similar chromatin types (Branco and Pombo 2006; Chambeyron and Bickmore, 2004; Osborne et al., 2004).

A and B compartment domains can be identified by using eigen vector decomposition and aligning to gene density such that regions of high gene density are classified as positive principal component 1 (PC1) values corresponding to A compartments (Lieberman-Aiden, Berkum et al., 2009). Interestingly, these compartment tracks have been shown to correlate very well with replication timing profiles where early replicating regions overlap with A compartment domains and B compartments replicate later (Ryba et al., 2010). Compartments can also be categorized by the epigenetic marks associated with each chromatin type. For example, euchromatin (A compartments) tends to be enriched for histone marks such as H3K4me3, H3K27ac, and H3K36me3, while heterochromatic B compartments are often characterized by H3K9me3 and H3K27me3 marks (Lawrence et al., 2016). High-resolution Hi-C studies have suggested that more spatial segregation can occur within each compartment type to create ‘subcompartments’ based on the histone modifications present (Rao, Huntley et al., 2014; Belaghzal, Borrman et al., 2019).

The function and establishment of compartments is still largely unknown. Compartments are not conserved across different cell types; however, this may not be
surprising as compartments are highly related to gene expression which is cell type specific (Lieberman-Aiden, Berkum et al., 2009; Dixon, Jung, Selvaraj et al., 2015). Additionally, only recently it was shown that compartments are a feature of individual cells and not simply a result of population-based Hi-C analysis (Wang et al., 2016; Boettiger et al., 2016). Therefore, while it may not be the exact function of compartments, it seems clear that at least in some way compartmentalization is driven by transcriptional activity. For example, actively transcribed genes of A compartments are shown to colocalize to regions enriched in RNA polymerase II termed ‘transcription factories’ (Iborra et al., 1996; Sutherland and Bickmore, 2009), as well as regions enriched for splicing machinery termed ‘nuclear speckles’ (Brown et al., 2008; Rao, Huntley et al., 2014). Inactive chromatin regions, found in B compartments, also colocalize at different nuclear regions namely at the nuclear lamina and the nucleolus. B compartments localized to the nuclear lamina and nucleolus have been termed lamin-associated domains (LADs) and nucleolus-associated domains (NADs), respectively. Both LADs and NADs have been found to closely interact with lamin proteins (Vogel, Peric-Hupkes, and van Steensel, 2007; Guelen et al., 2008; Nemeth et al., 2010; van Koningsbruggen et al., 2010). While the exact position of B-type chromatin to a LAD or NAD can change between mother and daughter cells after mitosis, both of these regions maintain a transcriptionally repressed state (Thomson et al., 2004; van Koningsbruggen et al., 2010; Kind et al., 2013; Ragoczy et al., 2014). In fact, tethering a locus to the nuclear lamina can causes transcriptional repression, though this is not always the case (Finlan et al., 2008; Reddy et al., 2008). Other regions of B compartments have been observed to cluster with polycomb proteins (PRC1/PRC2) or heterochromatin proteins (HP1a) which are known to bind to and spread
the repressive histone marks common to B-type compartments (Tolhuis et al., 2006; Schwarz, Kahn et al., 2010; Zeng et al., 2010). For exploration on how compartments are established and maintained see “Mechanisms of interphase chromosome folding” below.

**Topologically associated domains**

While compartments are organizational features of up to 10 Mb of chromatin, TADs are much smaller, submegabase (<1 Mb) chromatin regions (Nora et al., 2012; Dixon et al., 2012; Sexton, Yaffe et al., 2012). TADs are often found within compartments and are defined as chromatin regions in which the interaction frequency between genomic loci contained within the domain is much higher than the interactions of those loci with regions outside of the domain. On a Hi-C chromatin interaction map, this is represented as interactions directly along the diagonal which form square blocks. TADs have also be observed in single cells through imaging and single-cell Hi-C studies (Bintu, Mateo et al., 2018; Nagano, Lubling, Stevens et al., 2013). In contrast to compartments, TADs appear to be invariant across cell types and cell differentiation (Nora et al., 2012; Zhan et al., 2017). This seems to be dependent on gene expression as loci within TADs are often coexpressed and decorated with similar histone marks (Nora et al., 2012). Further, gene regulation of loci within a TAD are constrained to interactions with enhancers of the same TAD (Anderson et al., 2014; Lupianez et al., 2015; Symmons et al., 2014). TADs are not completely conserved, however, and TAD contacts that do change upon cell differentiation, for example, often correlate with transcriptional activity and movement between A and B compartment regions (Dixon, Jung, Selvaraj et al., 2015).

TADs and TAD borders are defined by using directionality index or insulation analysis (Dixon et al., 2012; Crane, Bian, McCord, Lajoie et al., 2015; Lajoie, Dekker, and
Kaplan 2015). This determines where sharp transitions occur in the Hi-C chromatin interaction map, and we therefore define TAD borders as the local minima in interaction frequency compared to the surrounding regions. TAD borders are enriched in transcription start sites, especially housekeeping genes, SINE repetitive elements, and the binding sites for CTCF and cohesin (Nora et al., 2012; Dixon et al., 2012, Sexton, Yaffe et al., 2012; Hou, Li, Qin et al., 2012; Phillips-Cremins et al., 2013). In fact, multiple studies have established that TAD formation is disrupted in the absence of CTCF or cohesin, suggesting that the formation of TADs relies on the presence of these proteins (Schwarzer, Abdennur, Goloborodko et al., 2017; Rao et al., 2017; Gassler, Brandao et al., 2017; Wutz, Varnai, Nagasaka, Cisneros et al., 2017; Nora et al., 2017). The regulatory function of TADs is supported by their conservation as well as experiments showing that changes to TAD boundaries leads to misregulation of genes (Nora et al., 2012). The favored mechanism of TAD establishment and maintenance is described in “Mechanisms of interphase chromosome folding”.

Dots/loops

An interaction that occurs between two specific loci in most cells of a population in Hi-C analysis will appear as a ‘dot’ on the Hi-C chromatin interaction map. This is indicative of a chromatin interaction for two loci that have a preference to interact together over interacting with any other region, suggestive of forming a loop of chromatin between them. Most often this is observed between TAD boundaries and mediated by CTCF-CTCF interactions (Sanyal, Lajoie et al., 2012; Rao, Huntley et al., 2014). Interestingly, this might also require interactions with RNA (Saldana-Meyer et al., 2019; Hansen, Hsieh, Cattoglio et al., 2019). Loop interactions are also observed for long-range gene regulation
of enhancer-promoter contacts by both Hi-C and imaging studies (Dernburg et al., 1996; Tolhuis, Palstra et al., 2002; Bau, Sanyal et al., 2011; Sanyal, Lajoie et al., 2012; Hughes et al., 2014). While many of these looping interactions remain within the confinements of TADs, enhancers were only shown to interact with the nearest gene 7% of the time (Dowen, Fan, Hnisz, Ren et al., 2014, Sanyal, Lajoie et al., 2012). Therefore, chromatin looping is functionally important for gene regulation.

Possibly the best characterized chromatin looping interaction is for the β-globin locus (Tolhuis, Palstra et al., 2002; Palstra, Tolhuis et al., 2003; Palstra et al., 2008; Deng et al., 2012). Early 3C studies showed that in order to achieve proper β-globin expression during development specific long-range chromatin loops would form with the upstream ‘locus control region’ (LCR) which was enriched with enhancer elements (Tolhuis, Palstra et al., 2002). Further, these chromatin interactions are specific to erythroid cells and are not observed in cells where the β-globin gene is inactive (Palstra, Tolhuis et al., 2003). Interestingly, looping interactions of β-globin genes to the LCR were not dependent on RNA polymerase II since they were able to form during transcription inhibition (Palstra et al., 2008). Instead, forced chromatin looping of β-globin-LCR was sufficient for recruiting RNA polymerase II and activating transcription (Deng et al., 2012). Forced looping with the LCR was even sufficient to activate transcription of fetal γ-globin in adult erythrocytes (Deng, Rupon et al., 2014).

Dots have also been observed in trans between different chromosomes. These dots are likely not formed by ‘looping’ interactions per se, but actually the colocalization of active genes at transcription foci (Iborra et al., 1996; Sutherland and Bickmore, 2009; Schoenfelder, Sexton, Chakalova et al., 2010; Branco and Pombo, 2006; Chambeyron
and Bickmore, 2004; Osborne et al., 2004). For example, activated globin genes are often found colocalized, but not in a direct looping contact, with other, unrelated active genes (Bau, Sanyal et al., 2011). It has also been shown that some long-range looping interactions are mediated by polycomb complexes PRC1 and PRC2. For example, long-range looping interactions of the Hox gene clusters are dependent on PRC1 and PRC2 in embryonic stem cells to maintain developmentally regulated genes as silent but poised for activation of lineage-specific genes upon differentiation (Joshi, Wang et al., 2015; Schoenfelder, Sugar, Dimon, Javierre, Armstrong et al., 2015).

**Chromatin organization in S phase and G2**

S phase creates an additional challenge for cells with the extra function of DNA replication. As mentioned previously, the nucleus is already somewhat setup for DNA replication by the arrangement of replication timing domains into A and B compartments during G1. As cells exit G1 and enter S phase, euchromatic, transcriptionally active A compartments are replicated first, while B compartments are replicated later. Local decompaction of replicating regions occurs along the chromatin slightly weakening the strength of TADs and CTCF-CTCF loops observed by population Hi-C and in individual cells, but leaving the TAD boundaries relatively unchanged (Naumova, Imakaev, Fudenberg et al., 2013; Nagano, Lubling et al., 2017). In contrast, the ‘plaid’ pattern representative of compartmentalization gets sharper in S phase cells (Nagano, Lubling et al., 2017; Gibcus, Samejima, Goloborodko et al., 2018). After replication is complete and cells enter G2 phase, the characteristic interphase structures observed in G1 and S phase are still maintained, with the stronger compartmentalization observed in S phase also
Mechanisms of interphase chromosome folding

While the strength of interphase chromosome organizational features changes between cell cycle stages of interphase, they remain present and are thought to be organized through the same mechanisms. The following sections will explain the mechanisms currently favored by the field for interphase chromosome folding into (1) TADs, and (2) compartments. While seemingly related, perturbations in key proteins involved in the formation of these structures revealed two distinct mechanisms for their formation (Schwarzer, Abdennur, Goloborodko et al., 2017; Nuebler et al., 2018).

Formation of TADs by loop extrusion

TADs are thought to be formed via active loop extrusion, similar to the condensin-driven loop extrusion in the creation of compact mitotic chromosomes (Fudenberg, Imakaev et al., 2016; Goloborodko et al., 2016b). The history and general mechanism of loop extrusion is explained above when discussing the “Mechanisms of mitotic chromosome folding”. Unlike in mitosis, however, condensin is bound very little to interphase chromatin (Hirano and Mitchison, 1994). Therefore, loop extrusion must be driven by a different loop extrusion machine in interphase. Interestingly, cohesin, a similar SMC family protein to condensin has been implicated in the formation of TADs. In fact, depletion of cohesin in interphase causes disruption to TAD formation and recent studies have shown that the small amount of condensin II remaining on the chromatin in interphase is not sufficient to produce TADs (Schwarzer, Abdennur, Goloborodko et al., 2017; Rao et al., 2017; Gassler, Brandao et al., 2017; Abdennur, Schwarzer et al., 2018).
Further, in contrast to mitosis, rather than extrusion continuing until a barrier of another extruding factor, in interphase it is suggested that CTCF acts as an extrusion barrier to cohesin (Fudenberg, Imakaev et al., 2016; Sanborn, Rao et al., 2015; Nora et al., 2017). Experimental observations have shown that TAD boundaries contain CTCF motifs that are directional in orientation (Rao, Huntley et al., 2014; Vietri Rudan et al., 2015). Hi-C and computational simulations support the formation of chromatin loops via cohesin extrusion stopped at convergent CTCF sites (Fudenberg, Abdennur et al., 2017). Simulations are able to recapitulate the results of population Hi-C, such as ‘dots’ between CTCF-CTCF boundaries, increased interactions within TADs while insulated from other TADs, and ‘flares’ of processive extrusion (Fudenberg, Imakaev et al., 2016; Fudenberg, Abdennur et al., 2017; Sanborn, Rao et al., 2015). While recent experiments demonstrate that cohesin is able to extrude chromatin in vitro (Davidson et al., 2016; Kim, Shi et al., 2019; Davidson et al., 2019), there is still very little evidence on the ability of cohesin as an extrusion machine in cells. Therefore, the exact molecular mechanism of cohesin extrusion is still being actively explored. One potential mechanism of interest is the involvement of transcription and moving RNA polymerase II along the chromatin. Different organisms have revealed different dependencies on transcription and the formation of TADs, suggesting that domains investigated could be formed by distinct mechanisms (Davidson et al., 2016; Ulianov et al., 2016; Busslinger et al., 2017; Brandao et al., 2019; Rowley et al., 2019). While some of these studies have suggested RNA polymerase II could push cohesin along the chromatin (Busslinger et al., 2017; Heinz, Texari et al., 2018), computational simulations show that the speed of movement just by RNA
polymerase II alone would not be sufficient for loop extrusion, leaving this as an open question in the field (Fudenberg, Abdennur et al., 2017).

Interestingly, depleting CTCF and disrupting TAD formation showed no change to compartments (Nora et al., 2017). However, disrupting TADs by cohesin removal led to an enhanced compartmentalization pattern observed as the ‘checker-board’ on a Hi-C matrix (Schwarzer, Abdennur, Goloborodko et al., 2017; Rao et al., 2017; Gassler, Brandao et al., 2017). This suggests that TADs and compartments are formed through two distinct mechanisms. This is further emphasized in mouse maternal pronuclei that contain features of TADs and loops but not compartments (Flyamer, Gassler, Imakaev et al., 2017).

**Formation of compartments by microphase separation**

The most likely mechanism currently being explored for the formation of spatially segregated euchromatic and heterochromatic compartments is microphase separation. This biophysical process suggests that small alternating block monomers of A and B compartments spatially cluster together with some, but very few interactions between blocks (Leibler 1980; Matsen and Schick, 1994; Jost et al., 2014; Haddad et al., 2017; reviewed in Hildebrand and Dekker, 2020). This implies an attractive force between monomers of the same type, such as A-A interactions, or attraction mediated by chromatin binding proteins specific for one or both of the monomer types. Recent experiments have shown that many chromatin-binding proteins can form condensates in vitro, suggestive of liquid-liquid phase separation (LLPS) (Lin et al., 2015; Larson et al., 2017; Strom et al., 2017; Shin, Chang et al., 2018). This phenomenon has been proposed for nuclear structures in the past, most notably for the attraction of repetitive rDNA
sequences in the formation of the nucleolus (Brangwynne et al., 2011; Feric, Vaidya et al., 2016).

Examples of LLPS have been shown for both A and B compartments. In A compartments, RNA polymerase II (RNAPII) transcription factories may contribute to phase separation (Hilbert et al., 2018). The phosphorylation state of the CTD tail of RNAPII has also been implicated in microphase separation (Boehning, Dugast-Darzacq, Rankovic et al., 2018; Guo, Manteiga et al., 2019; Zamudio et al., 2019). While this shows attractive forces within A compartments, more evidence points towards B compartments driving microphase separation of chromatin. For example, heterochromatin 1 (HP1α/HP1β) forms liquid condensates, increases density of the chromatin, thereby condensing it, and further, excludes active chromatin regions (Verschure et al., 2005; Larson et al., 2017; Strom et al., 2017; Sanulli et al., 2019). Polycomb complex can also form condensate puncta in cells and cluster regions of chromatin decorated with H3K29me3 (Plys, Davis et al., 2019; Tatavosian et al., 2019).

The key studies suggesting that compartmentalization is driven by attractive forces between B compartments comes from lamin-associated domains (LADs) (Falk, Feodorova, Naumova et al., 2019). The localization of much heterochromatin to the nuclear lamina suggests an attraction between lamins and heterochromatin. However, studies on the interphase 3D chromatin organization in cells lacking lamin proteins show that compartmentalization is still sufficient, but localization of B compartments is changed. For example, in the ‘inverted nuclei’ of mouse rod cells heterochromatin is localized to the center of the nucleus, but still efficiently spatially compartmentalized from the euchromatic A compartments at the nuclear periphery (Falk, Feodorova, Naumova et al., 2019).
Further progeria cells or cells undergoing senescence show a lamin-dependent loss of B-B compartment interactions (Chandra, Ewels et al., 2015; Chiang et al, 2019). These data combined with computational simulations suggest that compartmentalization is due to microphase separation driven mostly by strong interactions between heterochromatic regions (Falk, Feodorova, Naumova et al., 2019; Belaghzal, Borrman et al., 2019).

Roles of chromosome folding

As introduced above, organization into compartments and TADs are fundamental mechanisms of 3D interphase chromatin organization. Assembly into these structures may facilitate proper transcription, replication, and ultimately prepare the chromosomes for condensation and segregation during cell division. Therefore, disruptions to the formation of each of these chromatin features can have pathogenic consequences. With increasing knowledge on the 3D organization of chromatin, it is more evident that the linear DNA sequence is not a tell-all for gene misregulation and disease progression. This section lays out disruptions to compartments and TADs in disease to emphasize the important role of 3D genome organization.

Compartments and disease

The spatial segregation of chromatin regions varies between cell types and with the differentiation of cells, representative of changes in gene expression and compartment switching (Dixon, Jung, Selvaraj et al., 2015). Compartment switching has also been observed in the progression of breast cancer (Barutcu et al., 2015). In this study, small gene-rich chromosomes became more “open” A compartments and genes
on these chromosomes, such as those involved in WNT signaling, were upregulated. Interestingly, heterochromatic B compartments were shown to have a higher mutation rate in cancer cells, implicating compartments as a key organizational feature to suppress mutations (Schuster-Bockler and Lehner, 2012; Fortin and Hansen, 2015). Finally, increasingly more evidence is pointing towards the role of long non-coding RNAs (IncRNA) in chromosome organization and disease. FIRRE, for example, regulates some interchromosomal compartment interactions and is required for adipogenesis (Hacisuleyman, Goff et al., 2014). Other IncRNAs, such as HOTAIR and MALAT1, also have been implicated in cancer progression due to changes in chromatin compartmentalization (Gupta et al., 2010; Ji et al., 2003). Lastly, while the precise mechanisms that drive aging are unknown, cellular senescence and progeria suggest that position of heterochromatin to the nuclear periphery is important for maintaining a repressed state (Chandra, Ewels et al., 2015; Chiang et al, 2019).

**Topologically associated domains and disease**

TADs restrict gene looping and activation to only within convergent CTCF sites. Therefore, TADs are important for regulating gene expression and maintaining correct interactions between enhancers and promoters. The most notable example of TADs in human disease is in limb formation (Lupianez et al., 2015). This studied showed that deletion, inversion, or duplication of the region surrounding the *Epha4* gene results in limb malformations. Deletion of *Epha4* gene and the downstream CTCF boundary induced chromatin looping of the *Pax3* gene in the neighboring TAD with *Epha4* associated enhancers, driving the misexpression of *Pax3* and leading to Brachydactyly. Inversion and duplication of this gene region also created inappropriate enhancer-promoter loops
and led to limb malformations of F-syndrome and Polydactyly, respectively (Lupianez et al., 2015). Similarly, deletion of TAD boundaries within the Hox gene clusters induced polycomb repressive complex 2 (PRC2) removal and transcriptional upregulation during the differentiation of embryonic stem cells to motor neurons (Narendra et al., 2015). TAD regulation, therefore, seems especially important during development to make sure proper enhancer-promoter interactions form. Disruptions to TADs have also been studied in the context of cancer. For example, T cell acute lymphoblastic leukemia (T-ALL) cells show microdeletions around normal TAD boundaries which removes the insulating effect of TADs and promotes the misregulation of proto-oncogenes (Hnisz, Weintraub et al., 2016).

**Overview of this thesis**

While powerful microscopy and biochemical approaches have unveiled vast information on the 3D spatial organization of chromosomes, many open questions remain to be addressed. This thesis will discuss the transition and maintenance of chromatin between the two well-defined chromosome states during mitosis and G1. Chapter II explores the temporal order with which CTCF-loops, TADs, and compartments reform as cells exit mitosis and reveals a unique structure at the anaphase-telophase transition. Chapter III investigates the role of transcription in reformation of G1 3D chromosome structure and shows that active transcription is not required for the establishment or maintenance of interphase chromatin features. Chapter IV probes the role of nuclear RNAs on interphase chromatin organization and demonstrates few consequences on the
overall 3D spatial organization of chromosomes. Last, Chapter V discusses how these studies further our understanding of chromosome biology, relate results of these studies to other research, and presents suggestions for key follow up experiments.
CHAPTER II: Transition from Mitosis to Interphase 3D Genome Organization

Preface

The content of this chapter encompasses work published in Nature Cell Biology, by Kristin Abramo, Anne-Laure Valton, Sergey V. Venev, Hakan Ozadam, A. Nicki Fox, and Job Dekker. The publication is entitled “A chromosome folding intermediate at the condensin-to-cohesin transition during telophase” and is under the following citation:


Abstract

Chromosome folding is extensively modulated as cells progress through the cell cycle. During mitosis, condensin complexes fold chromosomes in helically arranged nested loop arrays. In interphase, the cohesin complex generates loops that can be stalled at CTCF sites leading to positioned loops and topologically associating domains (TADs), while a separate process of compartmentalization drives spatial segregation of active and inactive chromatin domains. We used synchronized cell cultures to determine how the mitotic chromosome conformation transforms into the interphase state. Using Hi-C, chromatin binding assays, and immunofluorescence we show that by telophase condensin-mediated loops are lost and a transient folding intermediate devoid of most loops forms. By cytokinesis, cohesin-mediated CTCF-CTCF loops and positions of TADs
start to emerge rapidly. Compartment boundaries are also established early, but long-range compartmentalization is a slow process and proceeds for several hours after cells enter G1. Our results reveal the kinetics and order of events by which the interphase chromosome state is formed and identify telophase as a critical transition between condensin and cohesin driven chromosome folding.

**Introduction**

During interphase cohesin organizes chromosomes in loops, thought to be the result of a dynamic loop extrusion process (Fudenberg, Abdennur et al., 2017). Loop extrusion can occur all along chromosomes but is blocked at CTCF sites leading to detectable loops between convergent CTCF sites (Rao, Huntley et al., 2014; de Wit, Vos, Holwerda, Valdes-Quezada et al., 2015; Guo, Xu et al., 2015; Vietri Rudan et al., 2015; Sanborn, Rao et al., 2015; Fudenberg, Imakaev et al., 2016) and the formation of topologically associating domains (TADs) (Fudenberg, Imakaev et al., 2016; Nora et al., 2012; Dixon et al., 2012). At the same time long-range association of chromatin domains of similar state, within and between chromosomes, leads to a compartmentalized nuclear arrangement where heterochromatic and euchromatic segments of the genome are spatially segregated (Lieberman-Aiden, Berkum et al., 2009). Compartmentalization is likely driven by a process akin to microphase segregation and is mechanistically distinct from loop and TAD formation (Lieberman-Aiden, Berkum et al., 2009; Di Pierro, Zhang et al., 2016; Erdel and Rippe, 2018; Michieletto et al., 2016; Nora et al., 2017; Schwarzer,
Abdennur, Goloborodko et al., 2017; Rao et al., 2017; Nuebler et al., 2018; Falk, Feodorova, Naumova et al., 2019).

During mitosis cohesin mostly dissociates from chromosome arms (Sumara et al., 2000; Losada et al., 2002) and condensin complexes re-fold chromosomes into helically arranged arrays of nested loops (Paulson and Laemmli, 1977; Marsden and Laemmli, 1979; Hirano and Mitchison, 1994; Strunnikov et al., 1995; Hirano et al., 1997; Ono et al., 2003; Naumova, Imakaev, Fudenberg et al., 2013; Gibcus, Samejima, Goloborodko et al., 2018). Recently we described intermediate folding states through which cells interconvert the interphase organization into fully compacted mitotic chromosomes (Gibcus, Samejima, Goloborodko et al., 2018). The kinetics and pathway of disassembly of the mitotic conformation and re-establishment of the interphase state as cells enter G1 are not known in detail. Previous studies point to dynamic reorganization of chromosomes during mitotic exit and early G1 (Kind et al., 2013; Dileep et al., 2015). Condensin I loading, already high in metaphase, further increases during anaphase and then rapidly decreases, while condensin II colocalizes with chromatin throughout the cell cycle (Walther et al., 2018). Cohesin, mostly dissociated from chromatin during prophase and prometaphase (Sumara et al., 2000; Losada et al., 2002), re-associates with chromosomes during telophase and cytokinesis, as does CTCF (Sumara et al., 2000; Darwiche, Freeman and Strunnikov, 1999; Cai, Hossain et al., 2018). However, it is not known how these events relate to modulation of chromosome conformation.
Materials and Methods

Cell Culture

HeLa S3 CCL-2.2 cells (ATCC CCL-2.2) and HeLaS3-NCAPH-dTomato cells (see below) were cultured in DMEM, high glucose, GlutaMAX™ Supplement with pyruvate (Gibco 10569010) with 10% fetal bovine serum (Gibco 16000044) and 1% PenStrep (Gibco 15140) at 37°C in 5% CO₂.

Creation of Stable HeLaS3-NCAPH-dTomato Cell Line

We used pSpCas9(BB)-2A-Puro (PX459) V2.0 [a gift from Feng Zhang (Addgene plasmid # 62988; http://n2t.net/addgene:62988; RRID:Addgene_62988)] to construct CRISPR/Cas vectors according to the protocol of Ran, Hsu et al. (Ran, Hsu et al., 2013). gRNAs are listed in Table 2.1.

To construct donor plasmids for C-terminal integration of dTomato, plasmids were based on pUC19 and constructed using synthesized DNA and homology arms generated by PCR (primers listed in Table 2.2). Template DNA (genomic DNA from HeLa S3 cells) was amplified using Q5 High-Fidelity DNA Polymerase (New England Biolabs) to generate NCAPH homology arms. gBlock containing dTomato and Blasticidin resistance was synthesized by Integrated DNA Technologies (IDT) (sequence in Table 2.3). Homology arms and gBlocks were cloned into pUC19 by Gibson assembly, using NEBuilder® HiFi DNA Assembly Master Mix (NEB).

To generate stable cell lines, 5 x 10⁶ cells were electroporated with gRNAs and donor plasmid. 24 hours after electroporation, 1 μg/ml puromycin was added. Two days later, 1 ug/mL blasticidin was added for NCAPH-dTomato selection. After 5 days, colonies were picked for further selection in a 96-well plate.
HeLaS3-NCAPH-dTomato clone A6 cell line is available upon request, with an MTA from ATCC. Alternatively, the constructs are available to re-create the cell line in original HeLa S3 cells.

Table 2.1: gRNA sequences for genome editing

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Sequence of sense gRNA (5'-3')</th>
<th>Sequence of antisense gRNA (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>HeLa NCAPH-dTomato</td>
<td>TCTGATGTTTCTTGTGAGGCA</td>
<td>TGCCTCACAAGAACATCAGA</td>
</tr>
<tr>
<td></td>
<td>ACCTCTCTGATGTTTCTTGTG</td>
<td>CACAAAGACATCAGAGGAT</td>
</tr>
<tr>
<td></td>
<td>CAAGGAGATNGAGGTCACTA</td>
<td>TAGTGAACACTCATCATTCTTCAGAT</td>
</tr>
<tr>
<td></td>
<td>ACTATGGAGAAGTCAGCAGC</td>
<td>GCTGCTGACTTCTCCATAGT</td>
</tr>
<tr>
<td></td>
<td>TGCATGTTTGTGCTTCCCT</td>
<td>AGGGGAAGACCAAAACATCAGA</td>
</tr>
</tbody>
</table>

Table 2.2: PCR primers for NCAPH homology arms

<table>
<thead>
<tr>
<th></th>
<th>5' PCR Primer</th>
<th>3' PCR Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCAPH Homology Arm 1</td>
<td>GTAGTCCCTAGTTTCCATG</td>
<td>TTCTTGTGAGGCAAGGAGAT</td>
</tr>
<tr>
<td>NCAPH Homology Arm 2</td>
<td>TGAGTTCACATTAGGAGCATG</td>
<td>CATCTCCACAGAATGCAGCA</td>
</tr>
</tbody>
</table>

Table 2.3: gBlock Gene Fragment

<table>
<thead>
<tr>
<th>gBlock name</th>
<th>Cell line used to make</th>
<th>Sequence</th>
<th>Contains Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>dTomatoT2ABlastR</td>
<td>HeLa NCAPH-dTomato</td>
<td>gcaaggagatATGGTGAGGCAAGGGCGAGGAGGAGGGTCATAAGAAGATGCAGGAGGTTTCTTGTGAGGCA</td>
<td>dTomato, T2A, Blastocidin resistance</td>
</tr>
</tbody>
</table>
**Mitotic Synchronization**

All prometaphase synchronization of cells were done by (1) single thymidine treatment to arrest cells in S phase, (2) release into standard media to allow cell recovery and entry into late S, and (3) nocodazole treatment to arrest cells in prometaphase. On Day 1, cells were plated at $4 \times 10^6$ cells / 15 cm plate in media containing 2mM thymidine (Sigma T1895). After 24 hours, cells were washed with 1X PBS (Gibco 14190144) and standard media was added back to plates for 3 hours. Cells were then treated with media containing 100 ng/mL nocodazole (Sigma M1404) for 12 hours. Floating mitotic cells were collected and washed in 1X PBS.

**Mitotic Release Timecourse**

For prometaphase samples, washed mitotic cells were immediately prepared for downstream analysis. Remaining samples were re-cultured in standard media for synchronous release into G1 and collected at indicated times. For early time points, both floating and adherent re-cultured cells were collected for analysis. After 5 hours release from nocodazole, only adherent cells were collected.

Approximately $5 \times 10^6$ cells at each time point were fixed in 1% Formaldehyde (Fisher BP531-25) diluted in serum-free DMEM for Hi-C analysis as described in Belaghzal et al. (Belaghzal et al., 2017). For cell cycle analysis, approximately $1 \times 10^6$ cells at each time point were fixed in 86% cold ethanol (Fisher 04-355-222) and stored at -20°C. For chromatin association protein analysis, approximately $5 \times 10^6$ cells at each time point were pelleted, flash frozen, and stored at -80°C. Additional samples were collected for fluorescent microscopy. Floating mitotic cells were resuspended in 1.5 mL 4% PFA (EMS 15710) (diluted in 1X PBS), transferred onto a Poly-L-lysine-coated
coverslip (Sigma P8920) in a 6 well plate, and spun at 1500xg for 15 min. Cells adherent to coverslips at later time points were fixed in 4% PFA for 15 minutes at 20°C. All coverslips were washed 3X in 1X PBS and stored in 1X PBS at 4°C.

**Cell Cycle Analysis**

Fixed cells were washed in 1X PBS then resuspended in PBS containing 0.1% NP-40 (MP Biomedicals 0219859680), 0.5 mg/mL RNase A (Roche 10109169001) and 50 ug/mL propidium iodide (Thermo P1304MP). Samples were incubated at 20°C for 30 minutes then analyzed via LSR II or MACSQuant VYB flow cytometry. Data was analyzed using FlowJo v3. Viability gates using forward and side scatter were set on the nonsynchronous sample and applied to all samples within the set. DNA content was plotted as a histogram of the red channel. G1, S, and G2/M gates were set on nonsynchronous sample and applied to all samples within the set to get percentage of cells in each state throughout the time course release from prometaphase arrest. Values plotted for kinetics of G1 entry were normalized such that the maximum number of G1 cells = 1.

**Hi-C Protocol**

Hi-C was performed as described in Belaghzal et al. (Belaghzal et al., 2017). Briefly, flash-frozen cross-linked cell culture samples were lysed then digested with DpnII at 37°C overnight. Next, the DNA overhanging ends were filled with biotin-14-dATP at 23°C for 4 hours and ligated with T4 DNA ligase at 16°C for 4 hours. DNA was then treated with proteinase K at 65°C overnight to remove crosslinked proteins. Ligation products were purified, fragmented by sonication to an average size of 200 bp, and size selected to fragments 100 - 350 bp. We then performed end repair and dA-tailing and
selectively purified biotin tagged DNA using streptavidin beads. Illumina TruSeq adaptors were added to form the final Hi-C ligation products, samples were amplified and PCR primers were removed. Hi-C libraries were then sequenced by PE50 bases on an Illumina HiSeq4000.

**Hi-C Data Processing**

Hi-C PE50 fastq sequencing files were mapped to hg19 human reference genome using *distiller-nf* mapping pipeline ([https://github.com/mirnylab/distiller-nf](https://github.com/mirnylab/distiller-nf)). In brief, bwa mem was used to map fastq pairs in a single-side regime (-SP). Aligned reads were classified and deduplicated using *pairtools* ([https://github.com/mirnylab/pairtools](https://github.com/mirnylab/pairtools)), such that uniquely mapped and rescued pairs were retained and duplicate pairs (identical positions and strand orientations) were removed. We refer to such filtered reads as valid pairs. Valid pairs were binned into contact matrices at 10 kb, 20 kb, 40 kb, and 200 kb resolutions using *cooler* ([Abdennur and Mirny, 2019](https://doi.org/10.1093/molbev/msz012)). Iterative balancing procedure ([Imakaev, Fudenberg et al., 2012](https://doi.org/10.1126/science.2001927)) was applied to all matrices, ignoring the first 2 diagonals to avoid short-range ligation artifacts at a given resolution, and bins with low coverage were removed using MADmax filter with default parameters. Resultant ".cool" contact matrices were used in downstream analyses using *cooltools* ([https://github.com/mirnylab/cooltools](https://github.com/mirnylab/cooltools)). For downstream analyses using *cworld* ([https://github.com/dekkerlab/cworld-dekker](https://github.com/dekkerlab/cworld-dekker)), contact matrices were converted to ".matrix" using *cooltools dump_cworld*. For visualization of contact matrices (as in Fig. 2.1), .matrix files were scaled to $100 \times 10^6$ reads using *cworld scaleMatrix*. Hi-C statistics for each sample are in Table 2.4.
34
Table 2.4: Hi-C library statistics
sample name

replicate

TB-HiC-Dpn-R2-Tasyn
TB-HiC-Dpn-R2-T0
TB-HiC-Dpn-R2-T05
TB-HiC-Dpn-R2-T1
TB-HiC-Dpn-R2-T15
TB-HiC-Dpn-R2-T175
TB-HiC-Dpn-R2-T2
TB-HiC-Dpn-R2-T225
TB-HiC-Dpn-R2-T25
TB-HiC-Dpn-R2-T275
TB-HiC-Dpn-R2-T3
TB-HiC-Dpn-R2-T325
TB-HiC-Dpn-R2-T35
TB-HiC-Dpn-R2-T4
TB-HiC-Dpn-R2-T45
TB-HiC-Dpn-R2-T5
TB-HiC-Dpn-R2-T6
TB-HiC-Dpn-R2-T7
TB-HiC-Dpn-R2-T8
TB-HiC-Dpn-R2-T9
TB-HiC-Dpn-R2-T10
TB-HiC-Dpn-R2-T11
TB-HiC-Dpn-R2-T12
TB-HiC-Dpn-Tasyn
TB-HiC-Dpn-T0
TB-HiC-Dpn-T05
TB-HiC-Dpn-T1
TB-HiC-Dpn-T15
TB-HiC-Dpn-T2
TB-HiC-Dpn-T25
TB-HiC-Dpn-T3
TB-HiC-Dpn-T35
TB-HiC-Dpn-T4
TB-HiC-Dpn-T45
TB-HiC-Dpn-T5
TB-HiC-Dpn-T6
TB-HiC-Dpn-T7
TB-HiC-Dpn-T8
TB-HiC-Dpn-T9
TB-HiC-Dpn-T10
TB-HiC-Dpn-T11
TB-HiC-Dpn-T12
TB-HiC-Dpn-R3-Tasyn-R3-T1
TB-HiC-Dpn-R3-T0-R3-T1
TB-HiC-Dpn-R3-T05-R3-T1
TB-HiC-Dpn-R3-T1-R3-T1
TB-HiC-Dpn-R3-T15-R3-T1
TB-HiC-Dpn-R3-T175-R2-T1
TB-HiC-Dpn-R3-T2-R3-T1
TB-HiC-Dpn-R3-T225-R2-T1
TB-HiC-Dpn-R3-T25-R3-T1
TB-HiC-Dpn-R3-T275-R2-T1
TB-HiC-Dpn-R3-T3-R3-T1
TB-HiC-Dpn-R3-T325-R2-T1
TB-HiC-Dpn-R3-T35-R3-T1
TB-HiC-Dpn-R3-T4-R3-T1
TB-HiC-Dpn-R3-T45-R3-T1
TB-HiC-Dpn-R3-T5-R3-T1
TB-HiC-Dpn-R3-T6-R3-T1

R1
R1
R1
R1
R1
R1
R1
R1
R1
R1
R1
R1
R1
R1
R1
R1
R1
R1
R1
R1
R1
R1
R1
R2
R2
R2
R2
R2
R2
R2
R2
R2
R2
R2
R2
R2
R2
R2
R2
R2
R2
R2
R3
R3
R3
R3
R3
R3
R3
R3
R3
R3
R3
R3
R3
R3
R3
R3
R3

hours
released
nonsync
0
0.5
1
1.5
1.75
2
2.25
2.5
2.75
3
3.25
3.5
4
4.5
5
6
7
8
9
10
11
12
nonsync
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
6
7
8
9
10
11
12
nonsync
0
0.5
1
1.5
1.75
2
2.25
2.5
2.75
3
3.25
3.5
4
4.5
5
6

total reads

valid pairs

cis:trans

104,669,952
76,816,879
78,148,788
112,339,448
107,606,744
154,328,053
89,382,411
90,935,105
84,148,100
115,617,036
99,980,436
151,684,674
81,940,038
89,447,203
100,481,059
98,381,256
94,216,067
121,545,051
112,024,628
65,597,287
98,046,899
69,049,880
66,994,596
81,972,204
104,007,745
97,218,931
94,040,307
105,887,885
120,460,249
108,468,430
107,062,541
131,249,381
122,608,013
113,179,000
134,184,013
131,676,693
133,370,227
103,403,267
134,346,609
123,551,808
121,189,313
116,935,592
155,699,189
145,116,204
142,137,269
106,160,757
117,675,461
140,488,893
126,674,919
190,919,728
197,645,950
106,759,094
142,949,656
137,920,948
95,901,261
105,726,054
107,644,228
106,634,369
127,016,732

46,792,328
33,773,563
34,016,255
50,284,394
48,577,006
69,606,151
40,999,465
44,284,005
40,363,526
58,662,962
47,183,839
69,351,771
37,528,598
45,296,234
48,669,489
48,432,404
44,774,978
56,550,967
54,337,182
31,716,095
46,003,081
32,655,399
31,645,951
41,802,885
52,152,196
49,213,234
48,052,408
55,901,671
63,330,246
59,712,475
53,864,490
64,978,028
62,783,363
60,839,561
73,294,825
73,703,601
75,738,925
57,516,987
75,748,755
63,597,721
61,005,853
60,645,953
85,670,219
74,455,576
69,887,661
54,676,594
56,740,043
67,115,351
62,140,575
99,421,588
101,489,606
55,357,446
70,837,064
71,731,772
43,454,909
46,872,744
50,197,662
57,778,673
68,180,226

1.46
2.66
2.26
2.29
2.19
2.16
2.01
2.24
1.95
1.82
2.47
1.82
1.76
1.88
1.54
1.52
1.71
1.49
1.51
1.28
1.52
1.54
1.42
1.65
4.14
3.06
2.99
2.67
2.51
2.24
2.40
1.87
2.29
1.73
1.91
1.66
2.11
1.87
2.06
1.83
2.06
1.77
1.20
1.96
2.10
1.99
1.88
1.93
1.72
1.66
1.60
1.84
1.64
1.56
1.58
1.67
1.56
1.43
1.54


**Contact probability ($P(s)$) plots & derivatives**

Cis reads from the valid pairs files were used to calculate the contact frequency ($P$) as a function of genomic separation ($s$) (adapted from *cooltools*). All $P(s)$ curves were normalized for the total number of valid interactions in each data set. Corresponding derivative plots were made from each $P(s)$ plot.

**Compartment analysis**

Compartment boundaries were identified in cis using eigen vector decomposition on 200 kb binned data with *cooltools call-compartments* function. A and B compartment identities were assigned by gene density tracks such that the more gene-dense regions were labeled A compartments, and the PC1 sign was positive. Change in compartment type, therefore, occurs at locations where the value of PC1 changes sign. Compartment boundaries were defined at these locations, except for when the sign change occurred within 400 kb of another sign change.

To measure compartmentalization strength, we calculated observed/expected Hi-C matrices for 200 kb binned data, correcting for average distance decay as observed in the $P(s)$ plots (*cooltools compute-expected*). We then arranged observed/expected matrix bins according to the PC1 values of the replicate 1 Hi-C dataset from cells released from prometaphase for 8 hours. We aggregated the ordered matrices for each chromosome within a dataset then divided the aggregate matrix into 50 bins and plotted, yielding a “saddle plot” (*cooltools compute-saddle*). Strength of compartmentalization was defined as the ratio of $(A-A + B-B) / (A-B + B-A)$ interactions. Values used for this ratio were determined by calculating the mean value of the 10 bins in each corner of the saddle plot.
Values plotted for kinetics of compartment formation were normalized such that strength = 0 in prometaphase cells and the maximum value = 1.

In order to observe compartmentalization at different genomic ranges, we extracted observed/expected Hi-C data at specific distances (0-4 Mb, 4-8 Mb, 8-18 Mb, 18-38 Mb, 38-80 Mb) and made saddle plots. Since less data was used as input for each saddle plot, data was split into 20 bins instead of 50. Overall compartmentalization strength was calculated similar to above except using the mean value of the 9 bins in each corner of the saddle plot. Compartmentalization of individual compartment types was defined as the ratio of (A-A / A-B) or (B-B / A-B), where these values were determined by calculating the mean value of the 9 bins in the specified corner of the saddle plot. All values were normalized and plotted for kinetics the same as above.

**TAD analysis**

Domain boundaries were identified using insulation analysis on 40 kb binned data with `cworld matrix2insulation` and locating the minima in each profile (--is 520 kb --ids 320 kb). Domain boundaries were classified as compartment boundaries if they overlapped with the compartment boundaries defined above. All other domain boundaries were assumed to be TAD boundaries.

To measure TAD boundary formation, we aggregated 40 kb binned Hi-C data at domain boundaries identified from the replicate 1 Hi-C dataset from cells released from prometaphase for 8 hours (`cworld elementPileUp`). Insulation score was calculated by dividing the sum of interactions (with loci up to 40-500 kb away) for each bin within 500 kb of a boundary by the average of all interactions (with loci up to 40-500 kb away) for all binds located within 500 kb of a boundary.
Strength of TAD boundary formation was defined as the depletion of interactions across the boundary pileup, i.e. insulation as above. Boundary strength was calculated by measuring the average interaction of domain boundaries with regions 40-500 kb away (center vertical bin of boundary pileup) and subtracting that value from the average signal in regions immediately flanking the domain boundary (all bins left and right of domain boundary). All calculations were made after removing the bin closest to the diagonal. Values plotted for kinetics of TAD formation were normalized such that strength = 0 in prometaphase cells and the maximum value = 1.

Loop analysis

We used a previously identified set of HeLa S3 looping interactions for this analysis (Rao, Huntley et al., 2014). This set contains 3,094 total loops and 507 looping interactions are on the structurally intact chromosomes of HeLa S3 cells (Naumova, Imakaev, Fudenberg et al., 2013). To visualize looping interactions observed, we aggregated 10 or 20 kb binned data at loops larger than 200 kb to avoid the strong signal at the diagonal of the interaction matrix (cworld interactionPileUp).

Strength of loop formation was defined as the enrichment of signal at the looping interactions (center 3x3 pixels at loop position 20 kb binned data) compared to the flanking regions. Strength was calculated by averaging the signal at the looping interaction and subtracting the average signal outside. Values plotted for kinetics of loop formation were normalized such that strength = 0 in prometaphase cells and the maximum value = 1.

In order to observe formation of looping interactions at all loops sizes, we aggregated observed/expected Hi-C matrices for 20 kb binned Hi-C data at sites of
looping interactions. Using the observed/expected matrices corrects for distance decay and removes the overwhelming signal close to the diagonal, allowing us to observe smaller loops than in the observed Hi-C matrices.

**Simulated Hi-C mixture datasets**

We generated simulated Hi-C datasets for each replicate time course experiment. For each replicate the following protocol was used to randomly mix reads from prometaphase Hi-C datasets (t = 0 hours) with random Hi-C data reads from the sample having the highest percentage of G1 cells in the respective time course (t = 8 hours for replicates 1 and 2, t = 6 hours for replicate 3). Mixing ratios were determined based on cell cycle analysis of the same time course replicate, such that x% prometaphase reads + 1-x% G1 reads was representative of the experimental FACs profile observed at each time point.

First, in order to properly compare samples, all valid pair files within a single Hi-C timecourse dataset were randomly down-sampled to the lowest number of uniquely mapped reads within that timecourse dataset. Next, the down-sampled valid pairs for experimental prometaphase (t = 0 hours) and experimental G1 (t = 6 or 8 hours) were randomly sampled to yield the correct ratio of experimental cells at each time point and the same number of total reads as the down-sampled valid pairs files. This step was repeated 25 times, resulting in 25 simulated valid pairs files with the same number of reads for each time point in each replicate. $P(s)$ plots for simulated Hi-C data represent the average $P(s)$ for 25 replicate valid pair simulations. For all other analyses, valid pairs files were binned and balanced (as above) into “.cool” contact matrices and the 25 replicates from the same simulated ratios were combined using cooler merge.
**Microscopy**

**Immunofluorescence staining**

Immunofluorescence staining was performed at room temperature. Fixed cells were permeabilized with 0.1% triton (Sigma T8787) in 1X PBS for 10 minutes. Cells were blocked with 3% BSA (Sigma A7906) in 0.1% triton/PBS for 1 hour. Cells were incubated with primary antibody diluted in the blocking buffer for 2 hours [Lamin A/C (636) mouse mAb (1:800, SantaCruz sc-7292 lot C2219), Rad21 rabbit pAb (1:1000, abcam ab154769, lot GR3224138-10), CTCF rabbit pAb (1:800, Cell Signaling 2899, lot 2)]. Cells were washed with 0.1% triton/PBS 3 x 5 minutes. Cells were incubated with secondary antibodies [goat anti-rabbit IgG H&L Alexa Fluor 488 (1:1000, abcam ab15007, lot GR3225678-1), goat anti-mouse IgG H&L Alexa Fluor 700 (1:1000, invitrogen A-21036, lot 2084419)] diluted in the blocking buffer and conjugated tubulin antibody [anti-tubulin (YOL1/34)-AlexaFluor647, rat mAb, 1:100, abcam ab195884, lot GR281429-4] for 1 hour in the dark.

Cells were washed with 0.1% triton/PBS 1 x 5 minutes and then washed with 1X PBS 3 x 5 minutes. Coverslips were mounted to slides using ProLong Diamond Antifade Mountant with DAPI (Invitrogen P36962). For image acquisition, we used a Leica TCS SP5-II confocal microscope with 405 nm, 488 nm, 561 nm, and 633 nm lasers. Imaging we performed using a Leica HPX PL APO 63X/1.40-0.6 oil immersion objective with standard PMTs. Images were acquired using Leica LAS AF.

**Cell Cycle Classification**

Images were split into individual tiffs by channel and analyzed using Cell Profiler 3.1.8 and Cell Profiler Analyst 2.2.1 (Carpenter et al., 2006; Jones et al., 2008). For each
image, we identified nuclei as primary objects in the DAPI channel (‘DNA’). We then used propagation from each ‘DNA’ object to look for secondary objects in the tubulin channel (‘tubulin’). At this point, blinded tiffs of each individual cell with DAPI and tubulin staining could be isolated. Cells were manually classified into either prometaphase, metaphase, anaphase, telophase, cytokinesis, or G1 based on the morphology of DNA and tubulin. Prometaphase cells were defined as cells with condensed chromosomes and disrupted tubulin structure due to the microtubule inhibitor used for prometaphase arrest. Cells classified as metaphase had a single axis of DAPI staining with tubulin aligned on each side. Anaphase cells had tubulin on each side of the DAPI axis but must have had two distinct DAPI clusters representing the separation of two genomic copies. Telophase classification was characterized by the presence of tubulin only between the two DAPI populations and no longer on the ends. When the tubulin signal was compressed between the two DAPI clusters, we classified those as cells undergoing cytokinesis. Finally, all cells with decondensed chromatin and no nuclear tubulin were classified as G1 cells. Post-classification, cells were un-blinded and matched back to the corresponding image to allow for the measurement of cumulative counts for each cell cycle phase and the percentage of cells entering G1 over the time course. A total of 13,470 cells were classified in this study.

Protein Localization

Cell Profiler was also used to measure the localization of NCAPH-dTomato, Rad21, CTCF, and Lamin A/C. In addition to the primary nuclei objects (‘DNA’) and the secondary objects (‘tubulin’) defined above (see Cell Cycle Classification), we created a tertiary object as the region between the primary and secondary objects (‘cytoplasm’).
We calculated enrichment of NCAPH, Rad21, and CTCF co-localizing with the chromatin by measuring the mean fluorescence intensity (MFI) of each protein overlapping with the ‘DNA’ object and subtracting the MFI of each protein overlapping with the ‘cytoplasm’ object.

To measure the formation of a lamin ring, we shrunk the ‘DNA’ object and subtracted this region from the ‘DNA’ original object to create a new object (‘lamin’) at the inside edge of the ‘DNA’ where we observed lamin ring presence in nonsynchronous cells. Next, we expanded the ‘DNA’ object and subtracted the original ‘DNA’ object to create a new object (‘LamCyto’) just outside of the ‘lamin’ object. We were able to quantify the presence of a lamin ring by subtracting the MFI of lamin fluorescence in ‘LamCyto’ region from the MFI of lamin in the ‘lamin’ region. This enriched for the signal of a lamin ring, therefore, higher values correlated with the presence of a lamin ring structure at the edge of the chromatin.

**Chromatin association**

*Fractionation protocol*

Flash-frozen cell pellets from each time point of the mitotic release time course were thawed and resuspended with lysis buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1% NP-40, 1 mM DTT, 1X Halt protease inhibitor cocktail (Thermo 78430)). Samples in lysis buffer were incubated on ice for 20 minutes and then spun at 13,000 x g for 10 minutes at 4°C. The supernatant (cytoplasmic fraction) was collected and the pellet was resuspended in nuclei buffer (10 mM PIPES, pH 7.4, 10 mM KCl, 2 mM MgCl₂, 0.1% NP-40, 1 mM DTT, 1X protease inhibitor) with 0.25% triton. Samples were incubated on ice for 10 minutes and then spun at 10,000 x g for 5 minutes at 4°C. The supernatant
(nucleoplasmic fraction) was collected and the pellet (chromatin fraction) was resuspended in nuclei buffer with 20% glycerol. The chromatin fraction was then sonicated to shear the DNA using a Covaris instrument with the following parameters: 10% duty cycle, intensity 5, 200 cycles/burst, frequency sweeping, continuous degassing, 240 second process time, 4 cycles. Final chromatin-bound protein samples were stored at -20°C.

**Western Blots**

The volume for approximately the same number of cells for each sample across the mitotic release time course was loaded in each lane of a 4-12% bis-tris protein gel (Biorad 3450125) and separated in 1X MES running buffer (Biorad 1610789). Proteins were transferred to nitrocellulose membranes (Bio-Rad 1620112) at 30 V for 1.5 hours in 1X transfer buffer (Thermo 35040). Membranes were blocked with 4% milk in PBS-T (1X PBS + 0.1% tween) for 1 hour at room temperature. Membranes were then incubated with specified primary antibody diluted 1:1000 in 4% milk/PBS-T overnight at 4°C [Histone H3 (ab1791), Rad21 (ab154769), RFP (cross-reacts with dTomato for NCAPH-dTomato, Rockland 600-401-379), SMC2 (ab10412), SMC4 (ab17958), NCAPD3 (ab70349), NCAPG2 (ab70350), Lamin A (ab26300), CTCF (Cell Signaling 2899), RNA polymerase II CTD repeat phospho S2 (ab5095)]. Membranes were washed with PBS-T 3 x 10 minutes at room temperature, then incubated with secondary antibody (anti-rabbit IgG HRP-linked, Cell Signaling 7074) diluted 1:4000 in 4% milk/PBS-T for 2 hours at room temperature. Membranes were washed with PBS-T 3 x 10 minutes. Membranes were developed and imaged using SuperSignal West Dura Extended Duration Substrate (Thermo 34076) and Bio-Rad ChemiDoc.
Quantification

Band intensity for each protein was quantified using Image Lab 5.2.1. Intensities for each lane were normalized by background intensity of an equal sized area in the same lane. All protein quantifications were normalized to the Histone H3 levels for the same time course samples.

Statistics and Reproducibility

No statistical methods were used to predetermine sample size. Three replicate Hi-C time courses were performed independently with similar results. For imaging experiments, the number of cells analyzed was the maximum experimentally feasible.

Code Availability


Data Availability

Sequencing data generated in this study have been deposited in the Gene Expression Omnibus (GEO) repository under accession GSE133462. Dataset titled “R1” refers to replicate 1 which is used in Figures 2.1-2.9 and Fig. 2.12. Dataset titled “R2” refers to replicate 2 which is used in Figures 2.2, 2.4, and 2.10. Dataset titled “R3” refers to replicate 3 which is used in Figures 2.2, 2.4, and 2.11. Data from this publication can also be accessed in the 4DN Data Portal with link
Results

Synchronous entry into G1

HeLa S3 cells were arrested in prometaphase (Naumova, Imakaev, Fudenberg et al., 2013). In order to determine how chromosome conformation changes as cells exit mitosis and enter G1, prometaphase arrested cells were released in fresh media (t = 0 hours) and aliquots were harvested at subsequent time points up to 12 hours after release from prometaphase. The fraction of cells that had entered G1 was determined by FACS. We observed that about 50% of the cells had re-entered G1 between t = 3 and 4 hours and that cells began to enter S phase after about 10 hours (Fig. 2.1a, Fig. 2.2 a-b). The highest proportion of G1 cells was observed at 8 hours after release and data obtained at this time point is used as a G1 reference in this work. Replicate time courses yielded similar results (Fig 2.2 c-d).

Chromosome conformational changes as cells enter G1

We performed Hi-C on aliquots of cells taken at various time points after cells were released from prometaphase arrest (Fig. 2.1b). Hi-C chromatin interaction maps for cells in prometaphase reproduced previously identified features. First, the contact map is dominated by frequent interactions along the diagonal and the absence of locus specific features (Naumova, Imakaev, Fudenberg et al., 2013). When interaction frequencies ($P$)
FIGURE 2.1: Hi-C analysis during mitotic exit and G1 entry

(a) FACS analysis of nonsynchronous and prometaphase-arrested cultures and of cultures at different time points after release from prometaphase-arrest. Percentages in the upper right corner represent the percent of cells with a G1 DNA content. Replicate time courses yielded similar results (Fig. 2.2c-d). (b) Hi-C interaction maps for nonsynchronous and prometaphase-arrested cultures and of cultures at different time points after release from prometaphase-arrest. The order of panels is the same as in a. Data for chromosome 14 are shown for two resolutions: 200 kb (top row, for entire right arm) and 40 kb (bottom row, for 36.5 Mb – 42 Mb region). Hi-C heatmaps are all on the same color scale. (c) Left: \( P(s) \) plots for Hi-C data from nonsynchronous, mitotic (t = 0 hours), or G1 (t = 8 hours) cultures. Right: \( P(s) \) plots for Hi-C data from cells at indicated time points after release from prometaphase. Three independent experiments were performed with similar results.
FIGURE 2.2: Cell cycle analysis of mitotic exit time courses
(a) FACS analysis of nonsynchronous and prometaphase-arrested cultures and of cultures at different time points after release from prometaphase-arrest in time course replicate 1. Percentages in the upper right corner represent the number of cells with a G1 DNA content. (b-d) Quantification of the fraction of cells in G1 at each time point from time course replicate 1 normalized to t = 8 hours (b), time course replicate 2 normalized to t = 8 hours (c), and time course replicate 3 normalized to a G1 maximum assumed to be 80% (d). Three independent experiments were performed with similar results.
were plotted as a function of genomic distance \((s)\) between loci, we observed the typical decay pattern observed for mitotic cells arrested with nocodazole (Fig. 2.1c). \(P(s)\) initially decays slowly up to 10 Mb with an exponent close to -0.5, followed by a more rapid decay at larger distances.

After release from prometaphase arrest, we observed a progressive gain in features of chromatin interaction maps normally seen in interphase. First, inspection of the Hi-C interaction maps revealed the emergence of short-range interphase chromatin features, such as TADs, as quickly as 2.5 hours after release from prometaphase arrest and these become more obvious over time (Fig. 2.1b, bottom row). Second, we observed the first appearance of a checker-board pattern of longer range interactions, reflecting the formation of A and B compartments, between 3 and 4 hours (Fig. 2.1b, top row). By 8 hours, the chromatin interaction maps and the shape of \(P(s)\) strongly resembled those obtained for nonsynchronous cell cultures (Fig. 2.1c) (Naumova, Imakaev, Fudenberg et al., 2013).

**Compartmentalization occurs slower than formation of TADs and loops**

We quantified the presence and strength of specific features as they reform during mitotic exit and G1 re-entry. For these quantifications, we only used the set of structurally intact chromosomes in HeLa S3 cells as we did previously (Naumova, Imakaev, Fudenberg et al., 2013).

We used eigenvector decomposition to determine the positions of A and B compartments (Lieberman-Aiden, Berkum et al., 2009). In prometaphase-arrested cells, A and B compartments are absent (Fig 2.3). By \(t = 3\) hours, PC1 detects the presence of A and B compartments, despite the fact that in Hi-C interaction maps, the checker-board
FIGURE 2.3: Compartment analysis for time course replicate 1
(a) Principal component 1 (PC1) along Chromosome 14 for Hi-C data obtained from cells at different time points after release from prometaphase. Principal component analysis
was performed on Hi-C data binned at 200 kb resolution. PC1 detects A and B compartments starting at \( t = 3 \) hours. Lower left corner represents Pearson correlation value of each track compared to nonsynchronous PC1. (b) Principal component 3 (PC3) along Chromosome 14 for Hi-C data obtained from cells at different time points after release from prometaphase. Principal component analysis was performed on Hi-C data binned at 200 kb resolution. PC3 detects some A and B compartments starting at \( t = 2.75 \) hours, but at later time points, PC1 captures compartments. Lower left corner represents Pearson correlation value of each track compared to nonsynchronous PC1. Three independent experiments were performed with similar results.
pattern is weak (Fig. 2.1b, Fig 2.3a). For some chromosomes, PC3 corresponds to compartments at even earlier times (t = 2.75 hours) (Fig 2.3b). To quantify compartment strength, we plotted interactions between loci arranged by their PC1 values (derived from the t = 8 hours Hi-C data) and obtained “saddle plots” (Nora et al., 2017) (Fig. 2.4a, top row). In these plots, interactions in the upper left corner represent interactions between B compartments and interactions in the lower right corner represent A-A interactions. The compartment strength is calculated as the ratio of homotypic (A-A and B-B) to heterotypic (A-B) interactions. The first appearance of preferred homotypic interactions is observed as early as 2.5 hours after release (Fig. 2.4a). These preferred interactions are initially weak, but gain strength during later time points. By ~5 hours after release, compartment strength is about 50% of the maximum strength we detect at 8 hours after prometaphase release.

Next, we quantified the appearance of domain boundaries, many of which define TADs. First, we determined the positions of boundaries from the insulation profiles along chromosomes using the t = 8 hours Hi-C data (Crane, Bian, McCord, Lajoie et al., 2015; Lajoie, Dekker, and Kaplan 2015). We aggregated Hi-C data at domain boundaries (Fig. 2.4a, middle row). In nonsynchronous cells, we observe a depletion of interactions across domain boundaries (Fig. 2.4b, left). In prometaphase, insulation at boundaries is absent. As cells exit mitosis, we observe insulation at boundaries as soon as t = 2.5 hours. Insulation strength increases as time progresses and reaches 50% of maximum strength at ~3.5 hours after release. Some domain boundaries identified by insulation analysis represent compartment boundaries. When analyzed separately, we find that compartment boundaries appear with similar kinetics as TAD boundaries (Fig 2.5). We
FIGURE 2.4: Kinetics of loop, TAD, and compartment formation
(a) Top row: Saddle plots of Hi-C data binned at 200 kb resolution for nonsynchronous and prometaphase-arrested cultures and of cultures at different time points after release from prometaphase-arrest. Saddle plots were calculated using the PC1 obtained from the Hi-C data of the 8 hour time point. Numbers at the center of the heatmaps indicate compartment strength calculated as the ratio of \((AA+BB)/(AB+AB)\) using the mean values from dashed corner boxes. Middle row: Aggregate Hi-C data binned at 40 kb resolution at TAD boundaries identified from the Hi-C data of the 8 hour time point \((n = 724 \text{ boundaries})\). The order of panels is the same as the top row. Dashed lines indicate the edges of the averaged domains. Bottom row: Aggregate Hi-C data binned at 10 kb
resolution at chromatin loops on intact HeLa S3 chromosomes (n = 507 loops) identified in Rao, Huntley et al. (Rao, Huntley et al., 2014). The order of panels is the same as the top row. (b) Left: Average insulation profile across TAD boundaries shown in panel a for different time points. Right: Average Hi-C signals at and around looping interactions. Each line represents the signal from the lower left corner to the upper right corner of the loop aggregate heatmaps shown in panel a (dashed line). (c) Normalized feature strength for TADs, loops, and compartments as a function of time after release from prometaphase. For replicate time course 1 (left) and replicate time course 2 (middle) the strength of each of these features was set at 1 for the 8 hour time point. Dotted line indicates the fraction of cells in G1 at each time point, normalized to t = 8 hours. For replicate time course 3 (right) the strength for each of these features was normalized to the strength expected based on data from replicate 1. Dotted line indicates the fraction of cells in G1 at each time point, normalized to G1 maximum assumed to be 80%. Three independent experiments were performed with similar results.
FIGURE 2.5: TAD and compartment domain boundaries form with similar kinetics

(a) Aggregate Hi-C data binned at 40 kb resolution at domain boundaries (top = all 724 boundaries, middle = 657 TAD boundaries, bottom = 67 compartment boundaries) at different time points after release from prometaphase. (b) Average insulation profile across averaged domain boundaries shown in panel a (left = TAD boundaries, right = compartment boundaries) for different time points. (c) Normalized strength for domain boundaries as a function of time after release from prometaphase. The strength for each of these features was set at 1 for the 8 hour time point. TAD boundaries and compartment boundaries form with similar kinetics. Three independent replicate experiments yielded similar results.
conclude that both TAD and compartment domain boundaries are established around \( t = 2.5-3 \) hours.

Finally, we quantified the appearance of looping interactions. Rao, Huntley et al. identified looping interactions in HeLa S3 cells, the large majority of which are between CTCF sites (Rao, Huntley et al., 2014). We aggregated Hi-C data at the 507 looping interactions on structurally intact chromosomes in HeLa S3 cells (Fig. 2.4a, bottom row) (Rao, Huntley et al., 2014). While such loops are readily detected in nonsynchronous cells, they are absent in prometaphase, as observed before (Oomen et al., 2019). Loops reappear as soon as 2.5 hours after release and gain strength in the following hours. Loop strength reaches 50% of the maximum obtained over the time course after \( \sim 3.5 \) hours release from prometaphase.

To directly compare the kinetics with which TADs, loops, and compartments form, we plotted the strength of each feature at each time point as the percentage of its maximum (Fig. 2.4c). TADs and loops form with kinetics that are similar or slightly faster than the kinetics of G1 entry. In contrast, even though compartment identity is established relatively quickly (\( t = 2.5-3 \) hours), strengthening of long-range interactions between compartment domains continues for several hours with kinetics that are slower than that of cells entering G1.

To determine whether the formation of A and B compartments form with similar kinetics, we quantified A-A and B-B interaction frequencies separately as a function of time (Fig. 2.6, Fig 2.7). We find that both compartment types form with similar kinetics. Interestingly, when analyzed as a function of genomic distance between domains, B-B
FIGURE 2.6: Kinetics of A and B compartment formation at various genomic distances
(a) Saddle plots of Hi-C data for the right arm of chromosome 4 binned at 200 kb resolution for different time points and split into genomic distance bands, as shown in

(b) Normalized strength over time for different genomic distance bands.

(c) Summary of normalized strength for different compartments and genomic distance bands from 0 to 8 hours after prometaphase.
gray in the first row. (b) Normalized compartmentalization strength of different genomic distances as a function of time and split by interaction type (A-A, B-B, A-B). (c) Normalized compartmentalization strength of interaction types as a function of time and split by genomic distance. Three independent experiments were performed with similar results.
FIGURE 2.7: Compartment analysis by distance for chromosome 14
(a) Saddle plots of Hi-C data for chromosome 14 binned at 200 kb resolution for different time points and split into genomic distance bands, as shown in gray in the first row. (b)
Normalized compartmentalization strength of different genomic distances as a function of time and split by interaction type (A-A, B-B, A-B). (c) Normalized compartmentalization strength of interaction types as a function of time and split by genomic distance. Three independent experiments were performed with similar results.
interactions are most prominent between loci separated up to 38 Mb, while A-A
interactions are more prominent for loci separated by >38 Mb. For compartment
interactions up to 38 Mb, the kinetics of development of B-B interactions is faster than
that of A-A interactions. For distances larger than 38 Mb, A-A interactions develop faster.
These analyses reveal unanticipated complexities of compartmentalization.

**TADs and loops form prior to G1 entry**

TADs and loops appear somewhat earlier than cells starting to enter G1, but at
later time points TAD and loop strength follows the accumulation of G1 cells closely. We
reasoned that if the kinetics of TAD and loop formation is simply attributable to the kinetics
of cells entering G1, then the observed Hi-C data at a given time point should be very
similar to an appropriate mixture of a purely mitotic and purely G1 Hi-C dataset. Note that
this approach assumes that there is a single G1 conformation and a single mitotic
conformation. Previous analyses indicate that Hi-C captures these states with equal
efficiency so comparison of the observed Hi-C data to mixtures of Hi-C data will then test
this assumption (Naumova, Imakaev, Fudenberg et al., 2013). To generate such
mixtures, we randomly sampled reads from the prometaphase-arrested (t = 0 hours) and
8 hour released samples and mixed them according to the cell cycle distribution
(percentage of cells in G1) of each sample to obtain a simulated time course of release
from prometaphase (Fig. 2.8a). We then used the simulated time course datasets to
perform the same analyses as described above to determine TAD, loop, and
compartment strength (Fig. 2.8b-d).

In the experimental time course we observed loops at 2.5 hours after release from
prometaphase (Fig. 2.8b). However, in the simulated time course loops appear later, at
FIGURE 2.8: TADs and loops form quicker than expected, while compartmentalization occurs slower than expected

(a) Schematic diagram of simulating Hi-C data based on the percentage of G1 cells at each time point. (b) Aggregate Hi-C data binned at 20 kb resolution at chromatin loops at different time points. Top row: Experimental Hi-C data. Middle row: Simulated Hi-C data. Bottom row: The difference between experimental and simulated Hi-C data. Loops are more prominent in experimental Hi-C data than in the simulated data at $t = 2.5$ and $t = 2.75$ hours. This analysis included loops larger than 200 kb to avoid the strong signal at the diagonal of the interaction matrix. (c) Aggregate Hi-C data binned at 40 kb resolution at TAD boundaries for different time points. Top row: Experimental Hi-C data. Middle row: Simulated Hi-C data. Bottom row: The difference between experimental and simulated Hi-C data. Insulation strength is stronger in experimental Hi-C data than in simulated Hi-C data.
C data at $t = 2.5$ and $t = 2.75$ hours. (d) Saddle plots of Hi-C data binned at 200 kb resolution for different time points. Top row: Experimental Hi-C data. Middle row: Simulated Hi-C data. Bottom row: The difference between experimental and simulated Hi-C data. Saddle plots were calculated using the PC1 obtained from the experimental Hi-C data of the 8 hour time point. Compartmentalization is weaker in experimental Hi-C than in simulated Hi-C data as illustrated by the fact that A-B interactions are less depleted in the experimental data (upper right and lower left corner of saddle plots). Similar results were obtained with independent experimental and corresponding simulated time courses (Fig. 2.10a-c, 2.11a-c). Three independent experiments were performed with similar results.
about 3 hours. We find that at 2.5-2.75 hours after release, loop strength in the experimental data is greater than in the simulated data, indicating that the percentage of G1 cells is not predictive of loop strength at these early time points (Fig. 2.8b, bottom row). We did not see a difference in the kinetics of loop formation for loops of different sizes (Fig 2.9). Similarly, we quantified the appearance of insulation at boundaries as a function of time in the experimental and simulated time course datasets (Fig. 2.8c). At $t = 2.5$ and 2.75 hours, boundaries are more prominent in the experimental Hi-C data. Combined, this indicates that TADs and loops appear prior to cells entering G1. Finally, we quantified compartment strength and find that from 3 to 6 hours release, compartmentalization is weaker in the experimental Hi-C data as compared to the simulated Hi-C datasets: the simulated Hi-C data show less inter-compartment interactions (A-B) than the actual samples (Fig. 2.8d). This again illustrates that compartmentalization is a relatively slow process that continues for several hours after cells have entered G1. Similar results were obtained with independent experimental and corresponding simulated time courses (Fig 2.10a-c, Fig 2.11a-c).

**An intermediate folding state during mitotic exit**

Properties of chromosome folding can be derived from $P(s)$ plots. For example, $P(s)$ plots for interphase and mitosis are distinct (Fig. 2.12a) and have been used to test models of chromosome folding (Fudenberg, Abdennur et al., 2017; Lieberman-Aiden, Berkum et al., 2009; Naumova, Imakaev, Fudenberg et al., 2013; Gibcus, Samejima, Goloborodko et al., 2018; Dekker et al., 2002). We calculated $P(s)$ for Hi-C data obtained from cells at different times after release from prometaphase arrest. We observe a gradual transition over time from a mitotic $P(s)$ shape to that of an interphase $P(s)$ curve.
FIGURE 2.9: Kinetics of loop formation for loops of different size
Loops were grouped according to size: (a) loops less than or equal to 125 kb, (b) loops greater than 125 kb and less than or equal to 200 kb, (c) loops greater than 200 kb and less than or equal to 325 kb, (d) loops greater than 325 kb. For each panel, top row: log2(observed/expected) Hi-C data for experimental time course, middle row: log2(observed/expected) Hi-C data for simulated time course, bottom row: the difference between experimental and simulated Hi-C data. Kinetics of loop formation is similar for all loop sizes. Three independent experiments were performed with similar results.
FIGURE 2.10: Analysis of time course replicate 2
(a) Aggregated Hi-C data binned at 20 kb resolution at chromatin loops at different time points. Top row: Experimental Hi-C data. Middle row: Simulated Hi-C data. Bottom row: The difference between experimental and simulated Hi-C data. Loops are more prominent
in experimental Hi-C data than in the simulated data between 3 and 4.5 hours. This analysis included loops larger than 200 kb to avoid the strong signal at the diagonal of the interaction matrix. Simulations were performed with experimental data from this time course (mixing Hi-C data for \( t = 0 \) and \( t = 8 \) hours). 

(b) Aggregate Hi-C data binned at 40 kb resolution at TAD boundaries for different time points. Top row: Experimental Hi-C data. Middle row: Simulated Hi-C data. Bottom row: The difference between experimental and simulated Hi-C data. Insulation strength is stronger in experimental Hi-C data than in simulated Hi-C data at \( t = 3.5 \) and \( t = 4.5 \) hours. 

(c) Saddle plots of Hi-C data binned at 200 kb resolution for different time points. Top row: Experimental Hi-C data. Middle row: Simulated Hi-C data. Bottom row: The difference between experimental and simulated Hi-C data. Compartmentalization is weaker in experimental Hi-C than in simulated Hi-C data as illustrated by the fact that A-B interactions are less depleted in the experimental data (upper right and lower left corner of saddle plots). 

(d) Derivative from \( P(s) \) plots. Black lines represent the derivative of \( P(s) \) for experimental Hi-C data and the dashed green lines represent the derivative of \( P(s) \) for the simulated Hi-C datasets for corresponding time points. Three independent experiments were performed with similar results.
FIGURE 2.11: Analysis of time course replicate 3
(a) Aggregated Hi-C data binned at 20 kb resolution at chromatin loops at different time points. Top row: Experimental Hi-C data. Middle row: Simulated Hi-C data. Bottom row: The difference between experimental and simulated Hi-C data. Loops are more prominent in experimental Hi-C data than in the simulated data between 3.25 and 4 hours. This
analysis included loops larger than 200 kb to avoid the strong signal at the diagonal of the interaction matrix. Simulations were performed with experimental data from this time course (mixing Hi-C data for t = 0 and t = 6 hours). (b) Aggregate Hi-C data binned at 40 kb resolution at TAD boundaries for different time points. Top row: Experimental Hi-C data. Middle row: Simulated Hi-C data. Bottom row: The difference between experimental and simulated Hi-C data. Insulation strength is stronger in experimental Hi-C data than in simulated Hi-C data at t = 3.25 and t = 4 hours. (c) Saddle plots of Hi-C data binned at 200 kb resolution for different time points. Top row: Experimental Hi-C data. Middle row: Simulated Hi-C data. Bottom row: The difference between experimental and simulated Hi-C data. Compartmentalization is weaker in experimental Hi-C than in simulated Hi-C data as illustrated by the fact that A-B interactions are less depleted in the experimental data (upper right and lower left corner of saddle plots). (d) Derivative from $P(s)$ plots. Black lines represent the derivative of $P(s)$ for experimental Hi-C data and the dashed green lines represent the derivative of $P(s)$ for the simulated Hi-C datasets for corresponding time points. Three independent experiments were performed with similar results.
FIGURE 2.12: Formation of a transient folding intermediate

(a) Contact frequency (P) versus genomic distance (s) for read normalized Hi-C datasets for experimental mitotic and G1 data (upper left, blue and orange lines, respectively) and experimental Hi-C data obtained from cells at different time points after release from prometaphase arrest (black lines). Dashed green lines are P(s) plots for simulated Hi-C datasets for corresponding time points. At the bottom of each P(s) plot, the difference between experimental and simulated P(s) is plotted for the different time points, except for the upper left plot which shows the difference P(s) for experimental G1 and mitotic cells. Note that the difference plot for the upper left graph is on a different scale than all of the other difference plots. (b) Derivative from P(s) plots shown in panel a. In the upper left graph, we indicate features that represent the condensin mitotic loop array and the cohesin loop size and density. The blue arrow indicates loss of the condensin-dependent mitotic loop array. The orange arrow indicates the initiation of the cohesin-dependent G1 loops. Similar results were obtained with independent experimental and corresponding
simulated time courses (Fig. 2.10d, 2.11d). Three independent experiments were performed with similar results.
The transitional shapes could be the result of a mixture of mitotic $P(s)$ and interphase $P(s)$ or could represent intermediate folding states. To distinguish these possibilities, we returned to our simulated mixtures of Hi-C data described above. We calculated $P(s)$ for the simulated datasets and compared to experimental $P(s)$ at each time point (Fig. 2.12a). For most of the time points, the simulated $P(s)$ closely aligns with the experimental $P(s)$ (Fig. 2.12a, bottom graphs). Interestingly, we observed relatively large differences when we compare simulated and experimental $P(s)$ at 2.5 and 2.75 hours after release from prometaphase. This means that at those time points, the percentage of G1 cells (9% and 17%, respectively) does not explain the change in $P(s)$.

To further explore this transition and the properties of this putative folding intermediate, we calculated the derivatives of $P(s)$. Previous work has shown that the derivate of $P(s)$ can reveal the average chromatin loop size and the density of loops along the chromosome (Gibcus, Samejima, Goloborodko et al., 2018; Gassler, Brandao et al., 2017; Patel, Kang et al., 2019). The derivative of $P(s)$ for G1 cells shows a local maximum around 100 kb, indicating the average cohesin mediated loop size, followed by a relative deep minimum, indicating the linear density of chromatin loops (Fig. 2.12b). The derivative of $P(s)$ for prometaphase cells shows a local maximum around several hundred kilobases representing the condensin mediated loop array (Gibcus, Samejima, Goloborodko et al., 2018). We compared derivatives of $P(s)$ for simulated and experimental data across the time course (Fig. 2.12b, Fig 2.10d, Fig. 2.11d). We observe that experimental and simulated data are very similar for most time points. At 2.5 and 2.75 hours after release from prometaphase, however, the derivative of the experimental $P(s)$ has a unique shape. While for the simulated data evidence for a condensin loop
array is still observable, the derivative of the experimental \( P(s) \) shows a relatively constant value of -1 for genomic distances ranging from 100 kb to 1 Mb. At subsequent time points, the local maximum around 100 kb becomes more prominent and the subsequent minimum becomes deeper indicating progressive cohesin loading and loop formation. We interpret this to mean that at \( t = 2.5 \) and \( t = 2.75 \) hours, there is a transient intermediate folding state in which the condensin loop array is largely disassembled and only some cohesin loops start to form.

**The transient intermediate folding state occurs during telophase**

In order to better define the cell cycle state during which we observe the intermediate folding state we analyzed cells at different time points by microscopy using a HeLa S3 cell line expressing the condensin I subunit NCAPH fused to dTomato. The kinetics of mitotic exit for this cell line are comparable, though about 30 minutes slower, to that of HeLa S3 cells. We stained cells with DAPI to assess chromosome morphology and with antibodies against tubulin to detect spindle organization (Fig. 2.13a). Based on chromosome morphology and spindle organization, we classified cells (\( n = 13,470 \) cells) as prometaphase, metaphase, anaphase, telophase, cytokinesis, or G1. We observe that after 2.2 hours, 50% of the cells have entered metaphase and rapidly progress to anaphase (\( t = 2.55 \) hours) (Fig. 2.13b). By 2.95 hours, 50% of the cells are at the anaphase to early telophase transition. Cells spend the next \( \sim1.5 \) hours in telophase and cytokinesis and 50% of the cells have entered G1 after about 4 hours in this time course. From the timing of these events, we infer that the transient intermediate folding state occurs during telophase (\( t = 2.5-3.5 \) hours in HeLaS3-NCAPH-dTomato, \( t = 2-3 \) hours in HeLa S3).
FIGURE 2.13: Chromatin colocalization dynamics of condensins and cohesin during mitotic exit
(a) Representative images of classification of cell cycle stages based on DAPI staining and tubulin organization (see Methods). Scale bar = 5 µm. (b) Cumulative plots of HeLaS3-NCAPH-dTomato cells at different cell cycle stages (left axis) and the percent of cells in G1 (right axis) defined by imaging. At least 400 individual cells were classified for each time point: 0 minutes (n = 405 cells), 30m (n = 520), 60m (n = 780), 90m (n = 669), 105m (n = 638), 120m (n = 613), 135m (n = 601), 150m (n = 812), 165m (n = 533), 180m (n = 507), 195m (n = 650), 210m (n = 607), 240m (n = 1057), 270m (n = 760), 300m (n = 855), 360m (n = 1186), 480m (n = 959). (c) Localization of Lamin A/C, NCAPH, and Rad21 during different cell cycle stages shown in panel a. Scale bar = 5 µm. For images showing CTCF localization see Fig. 2.14a-b. (d) Quantification of NCAPH, Rad21 and CTCF colocalization with chromatin and lamin ring formation at different cell cycle stages (see Methods). P = prometaphase, M = metaphase, A = anaphase, T = telophase, C = cytokinesis, G1 = G1. Box plots represent quartiles of the dataset with a line at the median value, whiskers represent range of the dataset, and diamonds outside of whiskers are
outliers. Cell numbers for CTCF plot were P (n = 2099 cells), M (n = 1020), A (n = 199), T (n = 39), C (n = 853), and G1 (n = 2142) (see Fig. 2.14a-b). For the other three plots, the corresponding numbers were 1601, 1052, 155, 74, 927, and 2100. (e) Quantification of Rad21 and NCAPH colocalization with chromatin in single cells at different cell cycle stages. Left plot represents data from all cells with color indicating cell cycle stage. Right plots represent the data separated into each individual cell cycle stage. Four independent experiments were performed with similar results.
Condensin unloading occurs during telophase while cohesin loading occurs during cytokinesis

The derivative of $P(s)$ plots (Fig. 2.12b) combined with the cell cycle classification described above (Fig. 2.13b) indicate that the mitotic loop array is disassembled during telophase. The mitotic loop array is generated by condensins I and II, while interphase loops and TADs are mediated by cohesin (Schwarzer, Abdennur, Goloborodko et al., 2017; Rao et al., 2017; Ono et al., 2003; Gibcus, Samejima, Goloborodko et al., 2018; Hirano 2016). We determined the kinetics with which condensins dissociate and cohesin associates with chromatin as cells exit mitosis. First, we analyzed condensin binding to chromosomes by microscopy in HeLaS3-NCAPH-dTomato cells classified at different cell cycle stages (Fig. 2.13c-d). Condensin I is associated with chromosomes until late anaphase. By telophase, most of the condensin I has dissociated. In contrast, very little cohesin is observed on chromosomes up until telophase, but is increasingly colocalized with chromatin during cytokinesis when we also observe the formation of a lamin ring. CTCF is not on chromosomes during early mitosis, but becomes colocalized with chromatin during telophase and the bulk of CTCF binds during cytokinesis (Fig. 2.14a-c). These observations confirm that during telophase both condensin and cohesin are depleted from the chromatin. This is illustrated at the single cell level in Figure 2.13e.

Finally, we determined chromatin association of these complexes directly by purifying chromatin-bound proteins followed by semi-quantitative western blot analysis (Fig. 2.15a). We quantified the level of chromatin binding for proteins of interest from the western blot and normalized each to the Histone H3 level in the corresponding sample (Fig. 2.15b). We find that SMC4, a subunit of both condensin I and II, dissociates from chromatin rapidly during telophase. Condensin II (NCAPG2, NCAPD3) showed very
FIGURE 2.14: Chromatin association dynamics of CTCF, condensin, and cohesin

(a) Classification of cell cycle stages based on DAPI staining and tubulin organization. Scale bar = 5µm.

(b) Localization of Lamin A/C, NCAPH, and CTCF during different cell
cycle stages shown in panel a. Scale bar = 5µm. (c) Quantification of CTCF and NCAPH colocalization with chromatin in single cells at different cell cycle stages. Left plot represents data from all cells with color indicating cell cycle stage. Right plots represent the data separated into each individual cell cycle stage. (d) Top: Western blot analysis of chromatin-associated proteins purified from HeLaS3-NCAPH-dTomato cells at different time points after release from prometaphase. Bottom: Quantification of the western blot shown above. NCAPH and Rad21 were analyzed on the same gel. The samples for Histone H3 analysis were run on another gel. Four independent experiments were performed with similar results.
FIGURE 2.15: Chromatin association dynamics of condensins and cohesin during mitotic exit

(a) Western blot analysis of chromatin-associated proteins purified from HeLa S3 cells at different time points after release from prometaphase. (b) Quantification of the western blot shown in panel a. Protein levels were normalized to Histone H3 levels from the same samples. (c) Summary of cellular and chromosomal events as cells exit mitosis and enter G1. Top: Schematic diagrams indicate the cellular events from prometaphase into late G1. Compartment type is indicated by color: blue = A, orange = B. Red lines represent tubulin and dashed gray lines represent lamina. Bottom: Models of chromosome conformation during early mitosis, telophase, cytokinesis, and interphase. Green bar indicates abundance of condensins I and II on the chromatin at the corresponding cell cycle stages. Yellow bar indicates cohesin abundance on the chromatin at the corresponding cell cycle stages. Four independent experiments were performed with similar results.
similar dissociation kinetics, as did condensin I (NCAPH-dTomato, Fig. 2.14d). Cohesin (Rad21) started to associate with chromatin after 3 hours release from prometaphase and continued to load as cells entered and progressed through G1. Chromatin association of CTCF, Lamin A, and elongating RNAPII showed very similar binding kinetics as cohesin. The timing of chromatin association of cohesin and CTCF is consistent with earlier studies (Sumara et al., 2000; Darwiche, Freeman and Strunnikov, 1999; Cai, Hossain et al., 2018) and with more recent chromatin immunoprecipitation experiments (Cai, Hossain et al., 2018; Zhang et al., 2019).

We conclude that during telophase, most condensin has dissociated from the chromosomes and cohesin association with chromosomes is low. This is consistent with the interpretation of the Hi-C data based on the derivate of $P(s)$ that at this time point there is a transient chromatin folding intermediate with no condensin-mediated loops and only a very low density of cohesin loops. As cells progress through cytokinesis, CTCF and cohesin increasingly load on chromosomes and this continues into G1.

**Conclusion**

We identify telophase as a critical intermediate state between the mitotic and interphase chromosome conformation (Fig. 2.15c). Hi-C, immunolocalization and chromatin binding assays show loss of condensin binding prior to telophase while CTCF and cohesin start loading during cytokinesis. This intermediate conformation is characterized by the absence of most SMC-driven loops and no or very weak long-range inter-compartment interactions. Given that this intermediate occurs during telophase
which lasts approximately 20-25 minutes, the lifetime of this intermediate must be similarly short. Subsequently during cytokinesis, CTCF and cohesin re-load, CTCF-CTCF loops and TAD boundaries are re-established as are compartment domains. While TADs and loops become more prominent rapidly with kinetics faster or equal to G1 entry, long-range compartmentalization occurs slower and continues to increase for several hours after cells have entered G1.

Our data show that key features that define the interphase state, loop anchors and domain boundaries are defined prior to cells entering G1. The fact that TADs and loops form rapidly indicates that the process of loop extrusion is relatively fast, extruding loops of up to several hundreds of kb within 15-30 minutes, consistent with previous studies (1-2 Kb per second on naked DNA (Ganji et al., 2018), several Kb per minute during prophase (Gibcus, Samejima, Goloborodko et al., 2018)). In contrast, long-range compartmentalization occurs more slowly during several hours in G1, even though their boundaries and identities are detectable much earlier. This is consistent with cytological observations, which show that LADs are not yet peripherally localized during cytokinesis (Kind et al., 2013). This supports the notion that compartmentalization is mechanistically distinct from TAD and loop formation, and has been proposed to be due to phase segregation (Di Pierro, Zhang et al., 2016; Erdel and Rippe, 2018; Michieletto et al., 2016; Nuebler et al., 2018; Falk, Feodorova, Naumova et al., 2019; Jost et al., 2014). A previous study also showed that compartmentalization occurs during early G1 (Dileep et al., 2015). Our data are in line with very recent studies that independently found that domain and loop anchors are established prior to G1 entry while inter-compartment interactions develop slower (Zhang et al., 2019).
The formation of an intermediate folding state during telophase coincides with this condensin-to-cohesin transition. Hi-C data for this state shows that chromosomes are mostly devoid of loops and long-range compartmentalization is minimal. The exponent of \( P(s) \) for this intermediate fluctuates around -1 for loci separated by 100 kb up to several Mb. Interpretation of this feature is not straightforward. It could represent the fact that chromosomes are transitioning between two states, with the -1 exponent being the average of the two. Alternatively, and more interestingly, an exponent of \(~-1\) has been proposed to correspond to a largely unentangled fiber (Lieberman-Aiden, Berkum et al., 2009; Grosberg et al., 1988; Grosberg et al., 1993; Mirny 2011). How could this state be formed? One intriguing possibility is that this is a remnant of the condensin-mediated mitotic loop array that is also not entangled. Continuous loop extrusion by condensin complexes, combined with topoisomerase II activity would lead to decatenation of adjacent loops (Goloborodko et al., 2016a). Dissociation of condensin during anaphase would then leave a largely unentangled though still linearly arranged conformation. Subsequent cohesin loading would then initiate the formation of loops again. Although at this time the exact topological state of telophase chromosomes is speculative, our results demonstrate that this transient state represents a key intermediate between the mitotic and interphase genome conformations. Future examination of the molecular and physical properties of this intermediate can not only reveal mechanisms by which cells build the interphase nucleus, but may also lead to better insights into the mitotic state from which it is derived.
Acknowledgements

We thank C. Baer from the UMass SCOPE Imaging Core for advice on imaging and help with the classification pipeline on CellProfiler. We thank members of the Dekker and Mirny laboratories for discussions. We acknowledge support from the National Institutes of Health Common Fund 4D Nucleome Program (DK107980) and the National Human Genome Research Institute (HG003143). J.D. is an investigator of the Howard Hughes Medical Institute.
CHAPTER III: Role of transcription in chromatin organization

Preface

The content of this chapter encompasses unpublished work performed by Kristin Abramo, Anne-Laure Valton, and Job Dekker.

Abstract

The interphase 3D chromatin organizational structures are highly correlated with gene expression. Compartments, for example, directly correlate with transcriptional activity and form spatially distinct regions of active and inactive regions of chromatin. Further, genes separated greatly in linear DNA colocalize at RNA polymerase II clusters, termed transcription factories. We used two RNA polymerase II inhibitors to determine if cells are able to maintain, as well as establish, proper chromatin organization in the absence of transcription initiation and/or elongation. Using Hi-C, we show that overall genome organization and interphase structures of compartments, TADs, and loops remain intact in cells lacking RNA polymerase II activity; albeit, the strength of these structures is weakened. While compartment boundaries remained intact in both nonsynchronous and synchronized cell cultures entering G1 in the presence of inhibitors, compartmentalization, the strength of like compartments interacting with each other over the opposite compartment type, was weaker, especially for A-A interactions, which we typically observe as strongest for long range interactions. Further, while the location of
TAD boundaries also remained the same between conditions, insulation at domain boundaries was significantly reduced, especially for TAD boundaries located in A compartments. In contrast, we find that the interaction frequency of CTCF-CTCF loops and other looping interactions showed no effect from the lack of RNA polymerase II activity, or in some cases presented as even stronger interactions. Our results reveal that global interphase chromatin state does not rely on active transcription. Instead, we propose that the 3D genome is organized in such a way to facilitate proper transcription and upon transcriptional activation loci become more locally accessible and interact more with other active compartments.

Introduction

The 3D interphase genome is largely organized into two distinct chromatin features. First, spatial segregation of euchromatin and heterochromatin forms A and B compartments, respectively (Lieberman-Aiden, Berkum et al., 2009). A compartments are characterized by high GC content, high gene-density, and relatively “open” chromatin regions decorated with H3K4me3, H3K27ac, and H3K36me3 histone marks (Simonis et al., 2006; Lieberman-Aiden, Berkum et al., 2009; Hou, Li, Qin et al., 2012; Sexton, Yaffe et al., 2012; Lawrence et al., 2016). In contrast, chromatin regions classified as B compartments have relative low GC content, are gene-poor, and are marked by H3K9me3 and H3K27me3 resulting in relatively “closed” chromatin regions. Compartments, therefore, directly correlate with transcriptional activity of the cell and show variability between cell types and differentiation states (Lieberman-Aiden, Berkum...
et al., 2009; Dixon, Jung, Selvaraj et al., 2015). Second, submegabase chromatin regions
form topologically associated domains (TADs) in which interactions between loci within
the domain are enriched and insulated from interacting with loci outside of the domain
boundary (Nora et al., 2012; Dixon et al., 2012; Sexton, Yaffe et al., 2012). While TADs
are more conserved than compartments (Nora et al., 2012; Zhan et al., 2017), these
domains are also highly correlated with gene expression. In fact, TAD contacts that
change in cell differentiation, for example, correlate with transcriptional changes in cell-
type specific genes (Dixon, Jung, Selvaraj et al., 2015).

Our previous work, in agreement with other studies, shows that the formation of
compartments and TADs must be by two distinct mechanisms (Chapter II; Schwarzer,
Abdennur, Goloborodko et al., 2017; Nuebler et al., 2018; Abramo et al., 2019). TADs are
thought to be formed via an active cohesin-driven loop extrusion mechanism (Fudenberg,
Imakaev et al., 2016; Fudenberg, Abdennur et al., 2017). In this model, cohesin slides
along the chromatin bringing together distal loci until it is blocked by CTCF, yielding
chromatin loops between convergent CTCF sites (Rao, Huntley et al., 2014; de Wit, Vos,
Holwerda, Valdes-Quezada et al., 2015; Guo, Xu et al., 2015; Vietri Rudan et al., 2015;
Sanborn, Rao et al., 2015; Fudenberg, Imakaev et al., 2016). Recent experiments have
demonstrated the ability of similar SMC proteins enriched on mitotic chromosomes to
extrude chromatin in an ATP-dependent manner during mitosis (Strick et al., 2004;
Shintomi et al., 2015; Eeftens et al., 2017; Terakawa, Bisht, Eeftens et al., 2017; Ganji et
al., 2018). While similar experiments have demonstrated the ability of cohesin to extrude
chromatin in vitro, there is little evidence on the ability of cohesin as an extrusion machine
in cells (Davidson et al., 2016; Kim, Shi et al., 2019; Davidson et al., 2019). One potential
mechanism of interest is the involvement of transcription, in which moving RNA polymerase II (RNAPII) helps cohesin translocate along the chromatin (Ocampo-Hafalla et al., 2016; Busslinger et al., 2017; Borrie, Campor et al., 2017; Heinz, Texari et al., 2018). Computational simulations suggest that the speed of RNAPII alone would not be sufficient for loop extrusion and would require some other motor or ATP-drive cohesin movement, but this remains an open question (Fudenberg, Abdennur et al., 2017). Further, TAD boundaries are enriched in transcription start sites (Nora et al., 2012; Dixon et al., 2012, Sexton, Yaffe et al., 2012; Hou, Li, Qin et al., 2012; Phillips-Cremins et al., 2013) and depletion of mediator has been shown to weaken chromatin looping, indicating that at least some loops rely on transcriptional activity (Kagey, Newman, Bilodeau et al., Phillips-Cremins et al., 2013, Lai et al., 2013).

Compartments, in contrast, are likely driven by microphase separation of genomic regions of similar epigenetic states (Lieberman-Aiden, Berkum et al., 2009; Di Pierro, Zhang et al., 2016; Erdel and Rippe, 2018; Michieletto et al., 2016; Nora et al., 2017; Schwarzer, Abdennur, Goloborodko et al., 2017; Rao et al., 2017; Nuebler et al., 2018; Falk, Feodorova, Naumova et al., 2019). Liquid-liquid phase separation (LLPS) may indicate some of the attractive forces of both A and B compartments (Lin et al., 2015; Larson et al., 2017; Strom et al., 2017; Shin, Chang et al., 2018). For example, actively transcribed loci of A compartments colocalize to regions enriched in RNAPII termed ‘transcription factories’ (Iborra et al., 1996; Sutherland and Bickmore, 2009), which may contribute to phase separation of euchromatin (Hilbert et al., 2018). Further, the phosphorylation state of the CTD tail of RNAPII, indicative of its transcriptional status, has
also been implicated in microphase separation (Boehning, Dugast-Darzacq, Rankovic et al., 2018; Guo, Manteiga et al., 2019; Zamudio et al., 2019).

Together, these data suggest a role for transcriptional activity in genome organization. In this study, we use two RNA polymerase II inhibitors to determine if cells are able to maintain, as well as establish, proper chromatin organization in the absence of transcription initiation and/or elongation.

**Materials and Methods**

**Cell Culture**

HeLa S3 CCL-2.2 cells (ATCC CCL-2.2) were cultured in DMEM, high glucose, GlutaMAX™ Supplement with pyruvate (Gibco 10569010) with 10% fetal bovine serum (Gibco 16000044) and 1% PenStrep (Gibco 15140) at 37°C in 5% CO₂.

**Transcription Inhibition**

5,6-Dichlorobenzimidazole 1-β-D-ribofuranoside (DRB) (Sigma Aldrich, D1916-50MG) and Triptolide (Fisher Scientific, 64-590-05MG) were dissolved in DMSO (Sigma Aldrich, D2650).

*Inhibition in nonsynchronous cells*

Approximately 8x10⁶ cells were plated in fresh media on a 15 cm dish the day before the inhibitors were added. The next day, media was replaced with media containing either 200 µM DRB, 200 µM DRB + 25 µM Triptolide, or an equal volume of DMSO. Cells were incubated for 4 hours at 37°C. For nascent RNA analysis, 0.5 mM 5-
ethynyl Uridine (EU) (ThermoFisher Scientific, C10365) or an equal volume of DMSO was added to cell media for the last 1 hour of incubation. We confirmed that this concentration did not affect cell cycle status (Fig. 3.1d). After 4 hours, adherent cells were harvested for nascent RNA analysis, cell cycle analysis, and Hi-C analysis. See Fig. 3.1a for experimental setup schematic.

**Inhibition in cells entering G1**

Cells were synchronized to prometaphase using a single thymidine treatment followed by a nocodazole block (see “Mitotic Synchronization” below). Washed, mitotic cells were re-plated on to 15 cm dishes in media containing either 200 µM DRB, 200 µM DRB + 25 µM Triptolide, or an equal volume of DMSO. Cells were incubated at 37°C for the specified time lengths and then harvested for nascent RNA analysis, cell cycle analysis, and Hi-C analysis. For nascent RNA analysis, 0.5 mM 5-ethynyl Uridine (EU) (ThermoFisher Scientific, C10365) or an equal volume of DMSO was added to cell media for the last 1 hour of incubation. After 3.5 hours, half of the adherent cells plated in DMSO were harvested. The matched samples for transcription inhibition (DRB alone or Triptolide + DRB) were harvested after 4 hours release. Note, this extra 30 minutes allows cells treated with DRB to catch up and results in a similar percentage of cells in G1 as observed in the DMSO treated samples after 3.5 hours release from prometaphase arrest. Likewise, for the 9 hour samples, the remaining DMSO treated cells were harvested after 8.5 hour release from prometaphase arrest, while the transcription inhibited samples were harvested after 9 hours release. Fig. 3.12a for experimental setup schematic.
**Mitotic Synchronization**

Prometaphase synchronization of cells were done by (1) single thymidine treatment to arrest cells in S phase, (2) release into standard media to allow cell recovery and entry into late S, and (3) nocodazole treatment to arrest cells in prometaphase. On Day 1, cells were plated at 4 x 10^6 cells / 15 cm plate in media containing 2mM thymidine (Sigma T1895). After 24 hours, cells were washed with 1X PBS (Gibco 14190144) and standard media was added back to plates for 3 hours. Cells were then treated with media containing 100 ng/mL nocodazole (Sigma M1404) for 12 hours. Floating mitotic cells were collected and washed in 1X PBS containing either 200 µM DRB, 200 µM DRB + 25 µM Triptolide, or an equal volume of DMSO.

**Nascent RNA analysis**

*Purification of nascent RNA*

For all samples except prometaphase-arrested cells, plates were washed quickly with cold 1X PBS and lysed with 1 mL of TRIzol™ Reagent (ThermoFisher Scientific 15596018) for 5 minutes at room temperature. Samples were then collected in 15 mL tubes and kept at -20°C until further processing. Washed prometaphase arrested cells in suspension were lysed in a 15 mL tube for 5 minutes, then stored at -20°C.

Thawed TRIzol samples were treated with 200 µL RNAse free chloroform, mixed, and incubated for 2 minutes at room temperature. Samples were then centrifuged and the aqueous phase was transferred to a new tube. 500 µL of isopropanol was added, along with 2 µL RNaseOUT™ Recombinant Ribonuclease Inhibitor (ThermoFisher Scientific 10777019) and 1 µL glycogen. Samples were mixed and incubated for 10 minutes at room temperature to precipitate the RNAs. Samples were then spun, washed
with 75% EtOH, dried and resuspended in 20 µL RNase-free water plus 1 µL RNaseOUT™. To enhance resuspension, samples were incubated at 60°C for 15 minutes. 9 µg of RNA was used in the RiboMinus Human/Mouse transcriptome Isolation kit (ThermoFisher Scientific K155001) and ribosomal RNA was removed according to the manufacturer instructions. Spike-in RNA (Firefly2, sequence in Table 3.1) was also added at this step at 0.02 ng/µL. Nascent RNAs were then purified from the ribosomal-depleted RNA using the Click-It Nascent RNA Capture Kit (ThermoFisher Scientific C10365) following the manufacturer instructions.

### Table 3.1: Firefly2 sequence

| Firefly2 | TTCCTTCTAATACGACTCCTATAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACTGCCCTGGTGCTTGCCACCTCTTTCATCTAGGGGAGACT GCCCTGCTGCAGCACCCTAACATTTTCGACGCCGGAGTGGCCGGCCTGCCCGACGACGATGCCGGCGAGCTGCCTGCCGCCGTCGTCGTGCTGGAACACGGCAAGACCATGACCGAGAAGGAGATCGTGGACTATGTGGCCAGCCAGGTGACAACCGCCAAGAAGCTGCGCGGCGGAGGTGGTGTTCGTGGACGAGGTGCCCAAGGGCCTGACCGGCAAGCAGACGCCCGCA

**Fragment Analyzer**

For analysis of nascent RNAs by fragment analyzer, magnetic Click-It beads bound with nascent RNA were resuspended in 100 µL of 95% formamide plus 10 mM EDTA and incubated for 5 minutes at 65°C. The released RNA was then cleaned up using RNeasy MinElute Cleanup Kit (Qiagen 74204) to isolate the final RNA used on the Fragment Analyzer.
**RT-qPCR**

For analysis of nascent RNAs by RT-qPCR, magnetic Click-It beads bound with nascent RNAs were resuspended in 50 µL of wash buffer 2 and incubated for 5 minutes at 70°C. For cDNA synthesis, 20 µL of 5X VILO from the SuperScript™ VILO™ cDNA Synthesis Kit (ThermoFisher Scientific 11754050) was added and samples were incubated for 10 minutes at room temperature to anneal random primers. Super Script enzyme was then added and samples were incubated for about 90 minutes at 42°C with agitation to synthesize the cDNAs. The samples were then incubated for 5 minutes at 85°C to inactivate the enzyme and liberate the cDNA from the beads. The cDNA in the supernatant was then collected in a new tube and stored at -20°C.

qPCR was performed using Fast SYBR™ Green Master Mix (ThermoFisher Scientific 4385612) and samples were analyzed on a StepOnePlus™ Real-Time PCR System (ThermoFisher Scientific). qPCR primers (listed in Table 3.2) were designed in the exons close to the transcription start site with forward and reverse primers being in the neighboring exons.

**Table 3.2: qPCR primers for measuring transcription inhibition**

<table>
<thead>
<tr>
<th>Target</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTB</td>
<td>CCAACCGCGAGAAGATGA</td>
<td>TCCATCACGATGCCAGTG</td>
</tr>
<tr>
<td>GAPDH</td>
<td>GCAATGCATCTCTGCACCACCA</td>
<td>TTCCAGAGGGGCCATCCACA</td>
</tr>
<tr>
<td>POLR2A_1</td>
<td>TTGTGCAGGACACACTCACA</td>
<td>CAGGAGTTCACTCACTCCACC</td>
</tr>
<tr>
<td>POLR2A_2</td>
<td>GCCAGAGTTGATGATTGTCA</td>
<td>AGCCAGTTTGTGAGTCAGGTGTC</td>
</tr>
<tr>
<td>Firefly2</td>
<td>TGATCAAGTCAAGGGCTACCA</td>
<td>GCTGCAGCAGGATAGACTCC</td>
</tr>
<tr>
<td>TBP</td>
<td>CAGCAGCAAGGTTCTGTTT</td>
<td>TCATGGCACCCTAGGGAG</td>
</tr>
<tr>
<td>ERCC3</td>
<td>TCCGGATTGAGCCGGAAGT</td>
<td>TGGATTTCTTCTTGTGCGGTTGCC</td>
</tr>
<tr>
<td>CDK9</td>
<td>GTCGAAACCAAGGCTTCCCCCT</td>
<td>TCAACAGCCGACAGGTGTA</td>
</tr>
<tr>
<td>MED1</td>
<td>GGGGTCTCATCTCAGTGCAG</td>
<td>CAAGCTCCGGACAGGTCACA</td>
</tr>
</tbody>
</table>
Cell Cycle Analysis

For all samples except prometaphase-arrested cells, plates were washed with 1X PBS, adherent cells were dissociated using accutase (ThermoFisher Scientific, A11105-01), resuspended in 1X PBS, spun, and approximately 1 x 10^6 cells were resuspended in 100 µL of cold PBS. 900 µL of 95% EtOH was added slowly to fix cells in ~86% EtOH (Fisher 04-355-222). Prometaphase arrested cells in suspension did not require use of accutase, but were otherwise fixed in the same manner described above.

Fixed cells were washed in 1X PBS then resuspended in PBS containing 0.1% NP-40 (MP Biomedicals 0219859680), 0.5 mg/mL RNase A (Roche 10109169001) and 50 µg/mL propidium iodide (Thermo P1304MP). Samples were incubated at 20°C for 30 minutes then analyzed via LSR II or MACSQuant VYB flow cytometry. Data was analyzed using FlowJo v3. Viability gates using forward and side scatter were set on a nonsynchronous sample and applied to all samples within the set. DNA content was plotted as a histogram of the red channel. G1, S, and G2/M gates were set on the nonsynchronous control sample and applied to all samples within the set to get percentage of cells in each treatment and time point.

Hi-C Analysis

Approximately 5 x 10^6 cells at each time point were fixed in 1% Formaldehyde (Fisher BP531-25) diluted in serum-free DMEM for Hi-C analysis. Hi-C was performed as described in Belaghzal et al. (Belaghzal et al., 2017). Briefly, flash-frozen cross-linked cell culture samples were lysed then digested with DpnII at 37°C overnight. Next, the DNA overhanging ends were filled with biotin-14-dATP at 23°C for 4 hours and ligated with T4 DNA ligase at 16°C for 4 hours. DNA was then treated with proteinase K at 65°C overnight.
to remove crosslinked proteins. Ligation products were purified, fragmented by sonication to an average size of 200 bp, and size selected to fragments 100 - 350 bp. We then performed end repair and dA-tailing and selectively purified biotin tagged DNA using streptavidin beads. Illumina TruSeq adaptors were added to form the final Hi-C ligation products, samples were amplified and PCR primers were removed. Hi-C libraries were then sequenced by PE50 bases on an Illumina HiSeq4000.

**Hi-C Data Processing**

Hi-C PE50 fastq sequencing files were mapped to hg19 and hg38 human reference genome using `distiller-nf` mapping pipeline (https://github.com/mirnylab/distiller-nf). In brief, bwa mem was used to map fastq pairs in a single-side regime (-SP). Aligned reads were classified and deduplicated using `pairtools` (https://github.com/mirnylab/pairtools), such that uniquely mapped and rescued pairs were retained and duplicate pairs (identical positions and strand orientations) were removed. We refer to such filtered reads as valid pairs. Valid pairs were binned into contact matrices at 10 kb, 20 kb, 40 kb, and 200 kb resolutions using `cooler` (Abdennur and Mirny 2019). Iterative balancing procedure (Imakaev, Fudenberg et al., 2012) was applied to all matrices, ignoring the first 2 diagonals to avoid short-range ligation artifacts at a given resolution, and bins with low coverage were removed using MADmax filter with default parameters. Resultant “.cool” contact matrices were used in downstream analyses using `cooltools` (https://github.com/mirnylab/cooltools). Hi-C statistics for each sample are in Table 3.3.
### Table 3.3: Hi-C library statistics

<table>
<thead>
<tr>
<th>Library Name</th>
<th>Condition</th>
<th>Release Time</th>
<th>Replicate</th>
<th>Total Reads</th>
<th>Valid Pairs</th>
<th>% Cis</th>
</tr>
</thead>
<tbody>
<tr>
<td>TB-HiC-Dpn-DMSOasyn</td>
<td>DMSO</td>
<td>N/A</td>
<td>1</td>
<td>110,201,866</td>
<td>53,775,691</td>
<td>58.07%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-DRBasyn</td>
<td>DRB</td>
<td>N/A</td>
<td>1</td>
<td>140,047,670</td>
<td>66,387,638</td>
<td>57.53%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-TpDasyn</td>
<td>T+D</td>
<td>N/A</td>
<td>1</td>
<td>126,790,766</td>
<td>62,112,156</td>
<td>57.85%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-DMSOasy2</td>
<td>DMSO</td>
<td>N/A</td>
<td>2</td>
<td>106,278,046</td>
<td>58,510,389</td>
<td>59.33%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-DRBasy2</td>
<td>DRB</td>
<td>N/A</td>
<td>2</td>
<td>119,338,033</td>
<td>61,569,299</td>
<td>58.08%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-TDasy2</td>
<td>T+D</td>
<td>N/A</td>
<td>2</td>
<td>130,157,488</td>
<td>69,655,692</td>
<td>57.31%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-T0-G2M-2</td>
<td>Prometaphase</td>
<td>0 h</td>
<td>1</td>
<td>193,374,690</td>
<td>98,890,936</td>
<td>72.56%</td>
</tr>
<tr>
<td>TB-HiC-D-4h-DMSOsyn2</td>
<td>DMSO</td>
<td>3.5 h</td>
<td>1</td>
<td>191,019,863</td>
<td>101,416,500</td>
<td>58.68%</td>
</tr>
<tr>
<td>TB-HiC-D-4h-DRBsyn2</td>
<td>DRB</td>
<td>4 h</td>
<td>1</td>
<td>187,946,831</td>
<td>97,854,026</td>
<td>62.57%</td>
</tr>
<tr>
<td>TB-HiC-D-4h-TpDsyn</td>
<td>T+D</td>
<td>4 h</td>
<td>1</td>
<td>176,766,067</td>
<td>91,922,512</td>
<td>60.68%</td>
</tr>
<tr>
<td>TB-HiC-D-9h-DMSOsyn2</td>
<td>DMSO</td>
<td>8.5 h</td>
<td>1</td>
<td>136,200,410</td>
<td>74,910,320</td>
<td>64.41%</td>
</tr>
<tr>
<td>TB-HiC-D-9h-DRBsyn2</td>
<td>DRB</td>
<td>9 h</td>
<td>1</td>
<td>129,600,655</td>
<td>70,200,975</td>
<td>59.90%</td>
</tr>
<tr>
<td>TB-HiC-D-9h-TpDsyn</td>
<td>T+D</td>
<td>9 h</td>
<td>1</td>
<td>135,191,283</td>
<td>74,910,320</td>
<td>64.41%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-T0-G2M-2</td>
<td>Prometaphase</td>
<td>0 h</td>
<td>2</td>
<td>107,788,709</td>
<td>64,362,320</td>
<td>83.44%</td>
</tr>
<tr>
<td>TB-HiC-D-4h-DMSOsyn2</td>
<td>DMSO</td>
<td>3.5 h</td>
<td>2</td>
<td>116,445,843</td>
<td>66,944,410</td>
<td>59.41%</td>
</tr>
<tr>
<td>TB-HiC-D-4h-DRBsyn2</td>
<td>DRB</td>
<td>4 h</td>
<td>2</td>
<td>127,386,555</td>
<td>71,523,891</td>
<td>58.49%</td>
</tr>
<tr>
<td>TB-HiC-D-4h-TDsyn2</td>
<td>T+D</td>
<td>4 h</td>
<td>2</td>
<td>112,587,504</td>
<td>65,159,756</td>
<td>59.48%</td>
</tr>
<tr>
<td>TB-HiC-D-9h-DMSOsyn2</td>
<td>DMSO</td>
<td>8.5 h</td>
<td>2</td>
<td>123,570,952</td>
<td>70,752,437</td>
<td>58.16%</td>
</tr>
<tr>
<td>TB-HiC-D-9h-DRBsyn2</td>
<td>DRB</td>
<td>9 h</td>
<td>2</td>
<td>109,663,495</td>
<td>59,886,150</td>
<td>56.10%</td>
</tr>
<tr>
<td>TB-HiC-D-9h-TDsyn2</td>
<td>T+D</td>
<td>9 h</td>
<td>2</td>
<td>116,119,993</td>
<td>67,927,607</td>
<td>60.38%</td>
</tr>
</tbody>
</table>

### Contact probability ($P(s)$) plots & derivatives

Cis reads from the valid pairs files were used to calculate the contact frequency ($P$) as a function of genomic separation ($s$) (adapted from cooltools). Corresponding derivative plots were made from each $P(s)$ plot.

### Compartment analysis

Compartment boundaries were identified in cis using eigen vector decomposition on 200 kb binned data with cooltools call-compartments function. A and B compartment identities were assigned by gene density tracks such that the more gene-dense regions were labeled A compartments, and the PC1 sign was positive. Change in compartment
type, therefore, occurs at locations where the value of PC1 changes sign. Compartment boundaries were defined at these locations.

To measure compartmentalization strength, we calculated observed/expected Hi-C matrices for 200 kb binned data, correcting for average distance decay as observed in the $P(s)$ plots (cooltools compute-expected). We then arranged observed/expected matrix bins according to the PC1 values of either the DMSO control sample or each individual track. We aggregated the ordered matrices for each chromosome within a dataset then divided the aggregate matrix into 50 bins and plotted, yielding a “saddle plot” (cooltools compute-saddle). Overall strength of compartmentalization was defined as the ratio of $(A-A + B-B) / (A-B + B-A)$ interactions. Values used for this ratio were determined by calculating the mean value of the 10 bins in each corner of the saddle plot. Strength of A versus B compartments was defined as the ratio of $(A-A / A-B)$ or $(B-B / A-B)$, respectively.

In order to observe compartmentalization at different genomic ranges, we extracted observed/expected Hi-C data at specific distances (0-4 Mb, 4-8 Mb, 8-18 Mb, 18-38 Mb, and 38-80 Mb) and made saddle plots.

Note that some compartment analyses only use the six structurally intact HeLa S3 chromosomes (chromosome 4, 14, 17, 18, 20, and 21) determined in Naumova, Imakaev, Fudenberg et al., 2013.

**TAD analysis**

Domain boundaries were identified using insulation analysis on 40 kb binned data with cooltools diamond-insulation with a 480 kb window and locating all minima in each profile (--ignore-diags 2 --min-dist-bad-bin 2) with a threshold of log2 boundary strength $> 0.15$. Domain boundaries were classified as compartment boundaries if they overlapped
with the compartment boundaries defined above. All other domain boundaries were assumed to be TAD only boundaries.

To measure insulation strength at TAD boundaries, we aggregated 40 kb binned insulation values at domain boundaries on the six structurally intact HeLa S3 chromosomes (chromosome 4, 14, 17, 18, 20, and 21) (Naumova, Imakaev, Fudenberg et al., 2013). Average insulation across TAD domains was calculated by the averaging the aggregated domains. Insulation strength at each individual boundary was defined as the difference between the local maxima surrounding each boundary and insulation value directly at the boundary (average of 2 bins surrounding each boundary). Wilcoxon rank-sum tests were done using scipy-stats package (scipy.stats.ranksums) for each control-treatment pair.

**TSS pileups**

We determined the 500 most active TSSs on the six structurally intact HeLa S3 chromosomes (chromosomes 4, 14, 17, 18, 20, and 21) (Naumova, Imakaev, Fudenberg et al., 2013) by the nonsynchronous expression level of each transcript measured by PRO-seq analysis in HeLa S3 cells. Data from Leighton Core and Luke Wojenski at the University of Connecticut. We then oriented the TSS regions such that transcription was always left to right, flipping the negative strand. We aggregated 40 kb binned insulation values at these 500 TSSs. Average insulation across the TSS and insulation strength of each individual TSS were calculated as above. Wilcoxon rank-sum tests were done using scipy-stats package (scipy.stats.ranksums) for each control-treatment pair.
Loop analysis

We used a previously identified set of HeLa S3 looping interactions for this analysis (Rao, Huntley et al., 2014). This set contains 3,094 total loops and 507 looping interactions are on the structurally six intact chromosomes of HeLa S3 cells (chromosomes 4, 14, 17, 18, 20, and 21) (Naumova, Imakaev, Fudenberg et al., 2013). In order to observe formation of looping interactions, we aggregated observed/expected Hi-C matrices for 10 kb or 20 kb binned Hi-C data at sites of looping interactions.

Strength of loop formation was defined as the enrichment of signal at the looping interactions (center 5x5 pixels at loop position in 10 kb binned data) compared to the flanking regions. Strength was calculated by averaging the signal at the looping interaction and subtracting the average signal outside. We did this same analysis for different groupings of loops (i.e. grouping by size, grouping by presence of CTCF motif, and grouping by POLR2A ChIP-seq signal). Size and CTCF motif were determined from the original dataset. To determine which loop interactions had a least one side with a POLR2A ChIP-seq signal we used ENCODE dataset ENCFF001VJA. Wilcoxon rank-sum tests were done using scipy-stats package (scipy.stats.ranksums) for each pair.

Code Availability

**Results**

**Inhibiting transcription in nonsynchronous cells**

To test if transcription is required to maintain the structures of interphase chromosome organization, we inhibited transcription elongation and/or initiation for 4 hours in HeLa S3 cells (Fig. 3.1a). To target transcription elongation, we treated nonsynchronous cells with 5,6-Dichlorobenzimidazole 1-β-D-ribofuranoside (DRB). We measured the level of nascent RNA transcripts by EU incorporation for 1 hour. We observed a negligible amount of nascent transcription by fragment analyzer remaining after treatment as compared to a DMSO control (Fig. 3.1b) and ~90% transcription inhibition efficiency for specific transcripts measured by RT-qPCR, such as ACTB (Fig 3.1c).

Targeting transcription initiation specifically, we were never able to see a similar inhibition efficiency; therefore, to test if transcription initiation is required for maintenance of 3D interphase chromatin organization we treated cells with a combination of Triptolide and DRB (Fig. 3.1a). After 4 hours of treatment, the level of nascent RNA in these cells was undetectable by fragment analyzer compared to the DMSO control (Fig. 3.1b) and the transcription inhibition efficiency was > 90% for all RT-qPCR targets (Fig. 3.1c). In all cases, the cell cycle status of these cultures varied very little with the use of transcription inhibitors, resulting in 63-66% of cells with G1 DNA content measured by FACS analysis (Fig. 3.1d).

**Chromosome conformational changes in cells with no active transcription**

To assess chromosome conformational changes in cells with no active transcription, we performed Hi-C on nonsynchronous cells treated for 4 hours with
FIGURE 3.1: Transcription inhibition in nonsynchronous cell cultures

(a) Experiment schematic for inhibiting transcription in nonsynchronous cells to measure 3D chromosome organization. (b) Top: Fragment analyzer analysis of nascent RNA isolated from nonsynchronous cells treated with DMSO, DRB, or Triptolide + DRB. Control sample with no EU added for 1 hour prior to harvesting shows that nascent RNA isolation was successful. Bottom: Quantification of fragment analyzer results normalized to the maximum amount of RNA pulled down in the DMSO control sample. (c) RT-qPCR measurements for ACTB, GAPDH, and POLR2A transcripts in nascent RNA isolated from nonsynchronous cells treated with DMSO, DRB, or Triptolide + DRB. (d) Top: FACS analysis of nonsynchronous cells, nonsynchronous cells treated with DMSO for 4 hours, and nonsynchronous cells treated with DMSO for 4 hours with the addition of EU for the final hour before harvesting. Percentages in the upper right corner represent the percentage of cells with a G1 DNA content. Bottom: FACS analysis of nonsynchronous cells treated with DMSO, DRB, or Triptolide + DRB for 4 hours with the addition of EU for the final hour before harvesting. Percentages in the upper right corner represent the percentage of cells with a G1 DNA content. Similar results were seen in a replicate experiment.
transcription inhibitors (Fig. 3.2). Hi-C chromatin interaction maps for cells treated with DRB or Triptolide + DRB appear very similar to cells treated with DMSO by eye (Fig. 3.2a). All chromatin interaction maps appear to have typical features expected for interphase cells. First, we observe a checker-board pattern off of the diagonal, representing spatially distinct interactions between active and inactive regions separated by large linear distances, termed compartments (Fig. 3.2a, top). These features remain seen in Hi-C libraries of cells that do not have activate transcription. Next, at shorter distances, we observe the presence of topologically associated domains (TADs) in which interactions are mostly contained within square regions along the diagonal of the interaction map and limited outside of these boundaries and between square regions (Fig. 3.2a, bottom). TADs visually appear very similar between Hi-C libraries treated with DMSO and transcription inhibitors.

In addition to visual features of chromatin interaction maps, we plotted interaction frequencies ($P$) as a function of genomic distance ($s$) between loci to derive properties of chromosome folding (Fig. 3.2b-c, left). We observe the typical decay pattern of chromatin interaction maps of interphase cells in which $P(s)$ shows two distinct organizations by distance (Lieberman-Aiden, Berkum et al., 2009). First, for short distances <1 Mb, the $P(s)$ curves show a shallow decay representing the local interactions contained within TADs. Second, for interactions between regions 1-10 Mb apart, the $P(s)$ decays quicker before hitting another plateau corresponding to the limit of long-range compartment interactions due to spatial compaction. To further explore the role of transcription on active processes in interphase chromatin organization, we calculated the derivatives of $P(s)$ (Fig. 3.2b-c, right). Prior work has demonstrated that the derivative of $P(s)$ for 10 kb
FIGURE 3.2: Hi-C analysis of nonsynchronous cell cultures treated with transcription inhibitors

(a) Hi-C interaction maps for nonsynchronous cell cultures treated with DMSO, DRB, Triptolide + DRB for 4 hours. Data for chromosome 14 are shown for two resolutions: 200 kb (top row, for entire right arm) and 40 kb (bottom row, for 36.5 Mb – 42 Mb region). (b) Left: Contact frequency $P$ versus genomic distance (s) for cis Hi-C data from nonsynchronous cell cultures treated with transcription inhibitors for 4 hours (top) and percent trans reads in each Hi-C library (bottom). Right: Derivative from $P(s)$ plots shown on left. (c) $P(s)$ and derivative of $P(s)$ plots for replicate 2, as in panel b.
– 10 Mb interactions can reveal the average size of extruded loops as the local maximum and the density of extrusion proteins, such as cohesin, as the depth of the local minimum (Gibcus, Samejima, Goloborodko et al., 2018; Gassler, Brandao et al., 2017; Patel, Kang et al., 2019). Using this method, we determined that cells treated with DMSO have an average cohesin extruded loop size of about 100 kb, which is consistent with our previous findings for this cell type (Chapter II; Abramo et al., 2019). Cell treated with DRB or Triptolide + DRB showed a local maximum in the derivative of $P(s)$ that was slightly shifted to the right of that of DMSO treated cells, suggesting cohesin extruded loops that are slightly larger than the average 100 kb loop formed in DMSO treated cells. Further, while very slight, the depth of the minimum of the derivative of $P(s)$ increases for cells treated with transcription inhibitors, indicating that the cohesin linear density on the chromatin may be slightly higher in cells lacking active transcription.

**Compartments are slightly reduced upon inhibiting transcription in nonsynchronous cells**

Visual inspection of the checker-board pattern in the chromatin interaction maps suggests very little changes in compartment maintenance during transcription inhibition (Fig. 3.2a, top). To confirm these observations, we determined the location of compartment types and the strengths of intra-compartment interactions for nonsynchronous cells treated with transcription inhibitors. First, we determined the positions of active (A) and inactive (B) compartments using eigenvector decomposition aligned with gene density such that A compartments yielded positive PC1 values (Lieberman-Aiden, Berkum et al., 2009). We find that the PC1 values called for the Hi-C libraries of transcription inhibited cells correlate very well with that of the DMSO treated
cells (Fig. 3.3a) with Pearson correlation r-values of 0.959 for DRB and 0.957 Triptolide + DRB. Since we detect similar correlation between replicate samples (DMSO rep 1 v. rep 2: 0.961, DRB rep 1 v. rep 2: 0.96, T+D rep 1 v. rep 2: 0.962), the assignment of A and B compartments appears to be maintained for cells treated with transcription inhibitors. Replicate treatments yielded similar results (DMSO x DRB rep 2: 0.959, DMSO x T+D rep 2: 0.961). We confirmed that the compartment tracks are similar between the three treatment conditions by plotting PC1 for a small chromosome, chromosome 14 (Fig. 3.3b) and a larger chromosome, chromosome 4 (Fig. 3.3c), and observe very little differences between the PC1 tracks.

To quantify the strengths of compartment interactions for nonsynchronous cells treated with transcription inhibitors, we created “saddle plots” (Fig. 3.4) (Nora et al., 2017) using the set of six structurally intact chromosomes in HeLa S3 cells as we did previously (Naumova, Imakaev, Fudenberg et al., 2013; Abramo et al., 2019). Saddle plots allow us to visualize and quantify the amount of interactions between similar compartment domains, A-A and B-B interactions, as well as interactions between opposite compartment types, A-B. We plotted interactions between loci arranged by the DMSO PC1 values which results in a saddle shape where interactions are enriched in the top left and bottom right corners representing strong B-B and A-A interactions, respectively, compared to A-B interactions (Fig. 3.4a). We quantified compartment strength as the ratio of the mean homotypic interaction (A-A and B-B) to the mean heterotypic interaction (A-B) for increasing bin sizes (Fig. 3.4b) and determined that taking the mean of the 10x10 square in each corner of the saddle plot (bin size = 10) yielded the cleanest results. We find that compartment strength is decreased > 20% in cells treated with DRB or Triptolide
FIGURE 3.3: Compartment boundaries and type are maintained in nonsynchronous cells treated with transcription inhibitors

(a) Control, DMSO treated principal component 1 (PC1) values from genome wide eigen vector decomposition on 200 kb binned Hi-C datasets versus PC1 values generated from the Hi-C libraries of transcription inhibited cell cultures (left: DRB, right: Triptolide + DRB). Lower right corner represents Pearson correlation value for each plot. Similar r-values can be seen between replicate Hi-C libraries (DMSO R1 v. R2: 0.961, DRB R1 v. R2: 0.96, T+D R1 v. R2: 0.962.). Similar results between conditions were also seen in a replicate experiment (DMSO R2 v. DRB R2: 0.959, DMSO R2 v. T+D R2: 0.961). (b) PC1 along Chromosome 14 for Hi-C data obtained from nonsynchronous cells treated with DMSO, DRB, or Triptolide + DRB for 4 hours. (c) PC1 along Chromosome 4 for Hi-C data obtained from nonsynchronous cells treated with DMSO, DRB, or Triptolide + DRB for 4 hours. Legend is the same as in panel b. Similar results were seen in a replicate experiment.
FIGURE 3.4: Compartmentalization is reduced in nonsynchronous cells treated with transcription inhibitors

(a) Saddle plots of Hi-C data binned at 200 kb resolution for nonsynchronous cell cultures treated with DMSO, DRB, or Triptolide + DRB for 4 hours. Saddle plots were calculated using the PC1 obtained from the Hi-C data of the DMSO control in each experiment set. Dashed corner boxes represent the regions used to calculate compartment strength from the saddle plots in panels c, d, f, and g. (b) Plots quantifying compartmentalization strength (AA+BB)/(AB+BA) using the mean values of an increasingly larger square (bin size) in each corner of the saddle plot. Dashed line represents bin size = 10 which is also shown as the dashed 10x10 squares in each corner of the DMSO saddle plot in panel a. (c) Compartment strength (bin size = 10) values normalized to compartment strength of the DMSO control for replicate 1 and (d) replicate 2. (e) Plots quantifying compartmentalization strength of A compartments (AA/AB) and B compartments (BB/AB) separately using the mean values of an increasingly larger square, as in panel b. Dashed line represents bin size = 10. (f) Compartment strength of A (solid) and B (striped) regions normalized to the strength of A and B in the DMSO control for replicate 1 and (g) replicate 2.
+ DRB (Fig. 3.4c) and that this is due to a decrease in both types of homotypic interactions, but strength is especially reduced in A-A interactions (Fig. 3.4e-f). Replicate treatments yielded similar results (Fig. 3.4d, g).

For consistency between samples, the analysis in Fig. 3.4 used the PC1 track derived from the DMSO control of each experimental set. However, since the correlation of PC1 values is high (Fig. 3.3a), we reasoned that saddles could also be created in which loci were arranged by the PC1 values derived from each individual Hi-C library (i.e. using their own PC1 track) (Fig. 3.5). Interestingly, when arranged according to their own PC1 tracks, we find compartment strength of transcription inhibited cells is reduced but not to the same extent as using the DMSO PC1 track for all samples. This further confirms that the maintenance of compartment interactions is disrupted in cells lacking active transcription.

We next quantified compartmentalization as a function of genomic distance (Fig. 3.6). As we have shown previously, B-B interactions are stronger than A-A interactions for loci separated up to 18 Mb, while A-A interactions are more prominent for loci separated by > 18 Mb in the DMSO control (Chapter II; Abramo et al., 2019). When cells are treated with transcription inhibitors, either DRB alone or Triptolide + DRB, we find that overall compartment strength (A-A + B-B / A-B + B-A) is decreased at all distances, but especially decreased for loci separated by > 18 Mb where we expect A-A interactions to be most prominent (Fig. 3.6b, left). The changes specifically occurring in A-A strength are emphasized when quantifying A-A and B-B interaction frequencies separately as a function of genomic distance (Fig. 3.6b, middle/right). While the strength of B-B compartment interactions remains relatively constant across conditions (Fig. 3.6b, right),
FIGURE 3.5: Compartmentalization using PC1 tracks generated from each Hi-C dataset also shows reduced strength in nonsynchronous cells treated with transcription inhibitors

Compartment strength (bin size = 10) for (a) replicate 1 and (b) replicate 2 calculated from saddle plots of Hi-C data binned at 200 kb resolution for nonsynchronous cell cultures treated with DMSO, DRB, or Triptolide + DRB for 4 hours and normalized to compartment strength of the DMSO control. Saddle plots were generated using the PC1 values obtained from the Hi-C data of each individual dataset.
FIGURE 3.6: Compartmentalization is reduced mostly at long range in nonsynchronous cell cultures treated with transcription inhibitors

(a) Saddle plots of Hi-C data at 200 kb resolution for nonsynchronous cells treated with DMSO, DRB, or Triptolide + DRB for 4 hours and split into genomic distance bands, as shown in Chapter II of this thesis and in Abramo et al., 2019. Saddle plots were calculated using the PC1 obtained from the Hi-C data of the DMSO control in each experiment set.

(b) Compartment score (bin size = 10) at different genomic distances in each condition, split by interaction type. Overall compartmentalization represents the mean values for (AA+BB)/(AB+BA), A-A compartment scores use the mean values for AA/AB, and B-B compartment scores are calculated using the mean values for BB/AB. Similar results were seen in a replicate experiment.
A-A interactions are reduced for loci separated by > 8 Mb (Fig. 3.6b, middle). For the longest range interactions > 38 Mb, we find that a decrease in both A-A and B-B compartment interactions contributes to the overall decrease in compartment strength. This may suggest that previous A-A interactions mediated by active transcription are now able to interact with B compartments and causing B-B interactions to dissociate.

Lastly, we quantified the strength of interchromosomal compartment interactions using “saddle plots” with only trans interaction frequencies (Fig. 3.7a). In contrast to cis compartmentalization, we find that compartment strength in trans is overall maintained or even slightly stronger in cells without active transcription (Fig. 3.7b-c). The quantification of A-A and B-B compartment strengths separately suggests that B-B interactions are slightly increasing in trans after transcription is inhibited, while A-A interactions slightly decrease (Fig. 3.7a). This suggests that the maintenance of compartment interactions in trans relies minimally on active transcription and is likely regulated by a separate mechanism.

**TADs are weakened upon inhibiting transcription in nonsynchronous cells**

To test if TADs are maintained in cells lacking active transcription, which is suggested to be true by the similarities of the chromatin interaction maps (Fig 3.2a, bottom), we quantified the strength of domain boundaries. First, we measured the relative frequency of chromatin interactions across each locus along chromosomes, generating insulation profiles for each Hi-C sample (Crane, Bian, McCord, Lajoie et al., 2015; Lajoie, Dekker, and Kaplan 2015). We find that the insulation values derived from the Hi-C libraries of nonsynchronous cells treated with transcription inhibitors correlate very well with that of cells treated with DMSO (Fig. 3.8a) with Pearson correlation r-values of 0.964.
FIGURE 3.7: Trans compartmentalization is maintained for nonsynchronous cells treated with transcription inhibitors

(a) Saddle plots of trans Hi-C data binned at 200 kb resolution for nonsynchronous cell cultures treated with DMSO, DRB, or Triptolide + DRB for 4 hours. Saddle plots were calculated using the PC1 obtained from the Hi-C data of the DMSO control in each experiment set. Values in the top left corner represent compartment strength of B-B interactions calculated as the ratio of the mean of the 10x10 square in the top left (BB) to the mean of the 10x10 square for AB interactions. Values in the bottom right corner represent compartment strength of A-A interactions calculated as the ratio of the mean of the 10x10 square in the bottom right (AA) to the mean of the 10x10 square for AB interactions. (b) Trans compartment strength values normalized to trans compartment strength of the DMSO control for replicate 1 and (c) replicate 2.
FIGURE 3.8: Insulation at TAD boundaries is reduced in nonsynchronous cells treated with transcription inhibitors

(a) Control, DMSO treated insulation values generated from genome wide insulation analysis on 40 kb binned Hi-C datasets versus insulation values generated from the Hi-C libraries of transcription inhibited cell cultures (left: DRB, right: Triptolide + DRB). Lower right corner represents Pearson correlation value for each plot. Similar results between conditions were also seen in a replicate experiment (DMSO R2 v. DRB R2: 0.967, DMSO R2 v. T+D R2: 0.966). (b) Top: Insulation profile along Chromosome 14 for Hi-C data obtained from nonsynchronous cells treated with DMSO, DRB, or Triptolide + DRB for 4 hours. Bottom: Insulation boundaries defined as local minima in the insulation profiles for DMSO replicate 1 (top blue bar), DMSO replicate 2 (lower blue bar), DRB (orange bar), and Triptolide + DRB (green bar). Black bars represent the regions defined as A compartments in a G1 HeLa S3 Hi-C dataset. (c) Pileup of the average insulation profile across TAD boundaries on six structurally intact HeLa S3 chromosomes defined in the Hi-C data of DMSO treated cells in each set. Left: replicate 1, Right: replicate 2. The number of boundaries used in each pileup is noted at the bottom of each plot. (d) Violin plots of the insulation strength at each individual boundary, separated by location in either A (left) or B (right) compartment type. P-value from Wilcoxon rank-sum test measured between each control-treatment pair. The horizontal white bars represent the median.
Boxplots within violin plots (gray) represent the first and third quartiles of the dataset. This panel represents data from the pooled Hi-C datasets of 2 replicates.
for DRB treatment and 0.966 for Triptolide + DRB. We detect similar results in replicates (DRB: 0.967, T+D: 0.966). We plotted the insulation values along Chromosome 14 to directly visualize the correlation of these tracks and find that the tracks look very similar in general between the three conditions with very slight variations in the amplitude of the tracks (Fig. 3.8b). We defined domain boundaries as local minima in the insulation track (Lajoie, Dekker, and Kaplan 2015). The domain boundaries defined in Hi-C libraries from transcription inhibited cell cultures do not differ very much from those defined Hi-C libraries from the control cultures, especially considering some of these differences can also be seen between DMSO replicates (Fig. 3.8b). Examples of both missing boundaries and new boundaries are observed across all samples, but most of these are due to a 1-2 bin shift in regions of a lot of noise in the insulation tracks or regions of very small domains. Differences in the location of domain boundaries defined in each sample do not appear to be due to location in a specific compartment type (Fig. 3.8b).

To quantify the strength of TAD boundaries, we aggregated insulation profiles at domain boundaries that did not overlap with compartment boundaries (Fig. 3.8c-d). The average insulation profile across these boundaries was slightly weaker in both replicates of cells treated with either DRB or Triptolide + DRB (Fig. 3.8c). We measured the insulation strength of each individual boundary and separated these values based on the compartment type for each boundary (Fig. 3.8d). We find that the distribution of the insulation strengths around boundaries in A and B compartments is significantly reduced in cells treated with Triptolide + DRB, compared to DMSO control treated cells. For cells treated with DRB alone, we only observe a significant reduction in insulation around TAD boundaries that are located in B compartments, compared to DMSO control treated cells,
but do not see a significant reduction in insulation around TAD boundaries located in A compartments (though close, p=0.05), where active transcription would be occurring. This suggests that transcription elongation has some role in the maintenance of TAD boundaries, but this oddly seems more relevant in regions where active polymerases are less likely to be. Transcription initiation or RNA polymerase II itself, since Triptolide induces proteasomal degradation of RNA polymerase II (Bensaude, 2011), appear to be involved in the maintenance of TADs and restricting chromatin interactions across TAD boundaries, even in inactive B compartments.

**Insulation around TSSs is reduced upon inhibiting transcription in nonsynchronous cells**

To quantify the strength of insulation around TSSs, we plotted the insulation values 500 kb upstream and downstream of the highest expressed genes on six structurally intact chromosomes in HeLa S3 cells from Hi-C libraries of nonsynchronous cells treated with DMSO, DRB, or Triptolide + DRB for 4 hours (Fig. 3.9a, top). In general, these heatmaps show a depletion of signal at the TSS, suggesting that these regions are insulated from other regions. The average insulation profile across these TSSs was slightly weaker in both replicates of cells treated with either DRB or Triptolide + DRB (Fig. 3.9a, bottom, Fig. 3.9c-d). To compare the insulation at each of the 500 TSSs individually, we subtracted the insulation values 500 kb upstream and downstream of these regions in the DMSO control Hi-C library from the insulation values in the Hi-C libraries of either DRB or Triptolide + DRB treated cells (Fig. 3.9b, top). We observe that insulation values surrounding many TSSs are higher in the Hi-C libraries generated from transcription inhibited cells, meaning that insulation is reduced. The average difference in insulation
FIGURE 3.9: Insulation at active TSSs is reduced in nonsynchronous cells treated with transcription inhibitors

(a) Top: Heatmap showing insulation values 500 kb upstream and downstream of the 500 most active TSSs on structurally intact chromosomes in HeLa S3 cells for 40 kb binned Hi-C data generated from nonsynchronous cell cultures treated with DMSO, DRB, or Triptolide + DRB for 4 hours. TSSs used in this analysis were determined using PRO-seq data generated by Leighton Core and Luke Wojenski at the University of Connecticut. Bottom: Average insulation profile across the 500 most active TSSs for each condition.

(b) Top: Heatmap showing the difference in insulation values 500 kb upstream and downstream of the 500 most active TSSs for Hi-C datasets of transcription inhibited cells versus DMSO control Hi-C data. Bottom: Average insulation profile of the difference in insulation values across the 500 most active TSSs for Hi-C datasets of transcription inhibited cells versus DMSO control Hi-C data.

(c) Pileup of the average insulation profile across the 500 most active TSSs for HeLa S3 cells, as in the bottom panels of a, for replicate 1 and (d) replicate 2.

(e) Violin plots of the insulation strength at each of the 500 most active TSSs in replicate 1 and (f) replicate 2 of Hi-C on nonsynchronous cells treated with DMSO, DRB, or Triptolide + DRB for 4 hours. P-value from Wilcoxon rank-sum test measured between each control-treatment pair. The horizontal white bars represent the median. Boxplots within violin plots (gray) represent the first and third quartiles of the dataset.
profiles across these TSSs is slightly higher for the cells treated with DRB or Triptolide + DRB (Fig. 3.9b, bottom), suggested less insulation at these sites and more interactions across these TSSs. We quantified the insulation strength of each individual TSS used in this analysis (Fig. 3.9e-f). We find that the distribution of insulation strengths for nonsynchronous cells treated with DRB or Triptolide + DRB for 4 hours is lower than the distribution of insulation strengths in control DMSO treated cells. Therefore, in the absence of transcription, insulation for the top 500 most active TSSs is reduced and these TSSs are potentially less accessible. This suggests that active transcription is responsible for maintaining some of these regions as accessible, restricting their interactions with other chromatin regions, and restricting interactions of chromatin across TSSs. However, for some of these regions, the remaining insulation may be due to CTCF proximity to the TSSs.

**Loops are maintained in nonsynchronous cells treated with transcription inhibitors**

Last, we tested if looping interactions are maintained in cells lacking active transcription (Fig. 3.10). We aggregated Hi-C data for nonsynchronous cells treated with DMSO, DRB, or Triptolide + DRB for 4 hours at the 507 looping interactions identified for the six structurally intact HeLa S3 chromosomes from Rao et al. (Fig. 3.10a) (Rao, Huntley et al., 2014). We observed little difference from visual inspection of the aggregate heatmaps. To better compare these looping interactions, we quantified the strength of looping interactions as the enrichment of Hi-C signal at the pairwise interaction over background signal from the flanking regions for each individual looping interaction (Fig. 3.10b). We find, in two independent replicates, that loops are maintained for nonsynchronous cells treated with transcription inhibitors for 4 hours compared to control
FIGURE 3.10: Loops are maintained in nonsynchronous cells treated with transcription inhibitors

(a) Aggregate Hi-C data binned at 10 kb resolution at chromatin loops on intact HeLa S3 chromosomes (n = 507 loops) identified in Rao, Huntley et al., 2014. (b) Loop strength
values for replicate 1 (left) and replicate 2 (right). Bar plots represent mean loop strengths with error bars at +/- 95% confidence interval. Wilcoxon rank-sum test between each control-treatment pair showed no significant differences in loop strength. (c) Aggregate Hi-C data binned at 10 kb resolution at chromatin loops from panel a separated by loop size: loops less than or equal to 125 kb, loops greater than 125 kb and less than or equal to 200 kb, loops greater than 200 kb and less than or equal to 325 kb, and loops greater than 325 kb. (d) Loop strength values for each treatment and loop size category for replicate 1 (left) and replicate 2 (right). Bar plots represent mean loop strengths with error bars at +/- 95% confidence interval. Wilcoxon rank-sum test between each control-treatment pair showed no significant differences in loop strength. (e) Aggregate Hi-C data binned at 10 kb resolution at chromatin loops from panel a separated by the presence of a CTCF motif in the loop: “CTCF-CTCF” corresponds to loops having a CTCF motif at each side of the loop, “no CTCF-CTCF” loops are categorized as those loops having a CTCF motif on only one side of the loop, and “no CTCF-no CTCF” loops have no CTCF motif on either side of the loop. (f) Loop strength values for each treatment and loop motif category for replicate 1 (left) and replicate 2 (right). Bar plots represent mean loop strengths with error bars at +/- 95% confidence interval. Wilcoxon rank-sum test between each control-treatment pair showed no significant differences in loop strength. (g) Aggregate Hi-C data binned at 10 kb resolution at chromatin loops from panel a separated by the presence of a POLR2A ChIP-seq signal: “POLR2A-ChIP” category corresponds to those loops that have a ChIP signal for at least one side of the loop bases, and “no POLR2A-ChIP” category represents the loops that have no ChIP signal for POLR2A on either side of the loop. (h) Loop strength values for each treatment and loop POLR2A category for replicate 1 (left) and replicate 2 (right). Bar plots represent mean loop strengths with error bars at +/- 95% confidence interval. Wilcoxon rank-sum test between each control-treatment pair showed no significant differences in loop strength.
cells treated with DMSO. In some cases, we even see an increase in loop strength in the absence of transcription, though we did not find these values to be significantly higher than strength values of DMSO treated cells. We next aggregated these Hi-C data at the 507 looping interactions separated by size (Fig. 3.10c). We find, in two independent replicates, that looping interactions at all distances are maintained in the absence of transcription (Fig. 3.10d). Next, we aggregated these Hi-C data at the 507 looping interactions separated by the presence of CTCF motifs at the loop base(s) (Fig. 3.10e). Again, we find that in two independent replicates the looping interactions are maintained in the absence of transcription independent of having a CTCF motif at one or both of the loop bases (Fig. 3.10f). Last, we aggregated the Hi-C data for nonsynchronous cells treated with DMSO, DRB, or Triptolide + DRB for 4 hours at looping interactions separated by their overlap with a POLR2A ChIP-seq peak on at least one loop base (Fig. 3.10g). We find that looping interactions are maintained in the absence of transcription regardless of the presence of POLR2A at the base of the loop (Fig. 3.10h). These analyses suggest that active transcription is not necessary for the maintenance of looping interactions.

**Measuring transcription during G1 entry**

To assess the extent that transcription is involved in the establishment of interphase chromatin organization, we first measured transcription in the cell cultures used in Chapter II of this thesis (Abramo et al., 2019). This allowed us to directly compare transcription in relation to the appearance of interphase structures of loops, TADs, and compartments as cells enter G1 (Fig. 3.11a). We find that global transcription is low for cells in mitosis (up to 2 hours release from prometaphase arrest), as expected and seen
FIGURE 3.11: Kinetics of loop, TAD, and compartment formation in relation to kinetics of transcription activity upon G1 entry

(a) Normalized feature strength for TADs, loops, and compartments as a function of time after release from prometaphase, from Chapter II Fig. 2.4c, middle. The strength of each of these features was set at 1 for \( t = 8 \) hours release from prometaphase arrest. Black line indicates the fraction of cells in G1 at each time point, normalized to \( t = 8 \) hours. (b) Normalized nascent transcription measurements as a function of time for cell cultures of the same experimental time course release from prometaphase shown in panel a. The black line represents global nascent RNA measurements by fragment analyzer analysis. The blue, red, and dashed green lines represent relative measurements of nascent transcripts for housekeeping genes (ACTB and GAPDH), transcription related genes (TBP, ERRC3, CDK9, and MED1), and POLR2A, respectively, by RT-qPCR in relation to the \( t = 8 \) hour sample.
previously (Palozola et al., 2017). Transcription increases gradually as cells enter G1 from 2 hours on after release from prometaphase arrest. Global nascent transcription measured by fragment analyzer reaches about 50% of the nonsynchronous transcript levels between t = 4 hours and t = 6 hours and reaches the full nonsynchronous transcript levels by about t = 8 hours. This coincides with the kinetics of compartmentalization, albeit transcription kinetics are slower, possibly requiring compartment formation for transcription. In multiple replicates, we saw a spike in global nascent transcript levels around t = 4 hours after prometaphase release.

We confirmed that the fragment analyzer measurements were a good representation of global nascent transcription by also doing RT-qPCR for some specific transcripts (Fig. 3.11b). We find that the kinetics of ACTB and GAPDH transcription follows the same trend of global transcription measured by fragment analyzer (Fig. 3.11b, ‘housekeeping genes’). In contrast, we found that transcription of the POLR2A locus spikes very high as soon as cells enter G1 after 2 hours release from prometaphase arrest and remains high up to about t = 6 hours release from prometaphase arrest (Fig. 3.11b, ‘POLR2A’). This is consistent with the early spike in cell maintenance genes prior to bulk transcription seen previously (Palozola et al., 2017). We wondered if this was true for all transcription-related genes and could explain why global transcription is a much slower and more gradual process, therefore, we measured the nascent transcript levels for TBP, ERCC3, CDK9, and MED1 by RT-qPCR for all samples in the time course (Fig. 3.11b, ‘transcription related genes’). We find that these genes do not follow the same transcription kinetics as POLR2A and instead these loci are transcribed slowly and gradually as cells enter G1 just as the housekeeping genes and global transcript
measurements. Therefore, while it seems transcription of the POLR2A locus itself likely does not require global 3D interphase chromosome organization, global transcription may rely on the proper organization of chromatin into TADs and loops, and even possibly compartments in order to regulate gene expression.

**Inhibiting transcription as cells enter G1**

To test if transcription is required to establish 3D interphase chromosome organization, we inhibited transcription elongation and/or initiation during prometaphase release in HeLa S3 cells (Fig. 3.12a). We harvested cells after 4 hours release from prometaphase arrest in transcription inhibitors to test if TADs and CTCF-CTCF loops are able to form properly and after 9 hours release from prometaphase arrest in transcription inhibitors to test if compartmentalization of active and inactive domains is achieved. Cells treated with either DRB alone or Triptolide + DRB showed negligible amounts of total nascent transcripts measured by fragment analyzer after being released from prometaphase arrest and treated with inhibitors for 4 hours or 9 hours (Fig. 3.12b). We also measured nascent transcripts of housekeeping genes ACTB and GAPDH by RT-qPCR (Fig. 3.12c). We measured 65-75% transcription inhibition efficiency in cells treated with DRB for 4 hours after release from prometaphase arrest, and 82-88% in cells treated with Triptolide + DRB for 4 hours after release (Fig. 3.12c, left). We note that transcription as a whole is low at these early time points contributing to an apparent lower inhibition efficiency than expected (Fig. 3.12d, 3.12b, lower). We confirm this with measurements of POLR2A, which is high already in control cells at 4 hours release from prometaphase arrest compared to other transcripts, such as ACTB and GAPDH, and we quantify 96-98% inhibition efficiency of POLR2A for cells treated with DRB or Triptolide + DRB. For
FIGURE 3.12: Transcription inhibition in cells released from prometaphase arrest and entering G1

(a) Experimental schematic for inhibiting transcription in cells arrested in prometaphase and synchronously released into G1 for 4 hours and 9 hours. (b) Top: Fragment analyzer
analysis of nascent RNA isolated from cells released from prometaphase arrest in the presence of DMSO, DRB, or Triptolide + DRB for 4 hours (left) or 9 hours (right). Middle: Quantification of fragment analyzer results normalized to the maximum amount of RNA pulled down in the DMSO control sample after 4 hours (left) and after 9 hours (right) release from prometaphase arrest. Bottom: Quantification of the fragment analyzer results normalized to the maximum amount of RNA pulled down in the DMSO control sample after 9 hours release from prometaphase arrest. (c) RT-qPCR measurements for ACTB, GAPDH, and POLR2A transcripts on nascent RNA isolated from cells treated with DMSO, DRB, or Triptolide + DRB during release from prometaphase arrest for 4 hours (left) or 9 hours (right). Transcript levels are relative to nascent RNA pulled down in the DMSO control for each transcript at each time point. (d) RT-qPCR measurements from panel c, relative to nascent RNA pulled down in the DMSO control sample after 9 hours release from prometaphase arrest. (e) FACS analysis of cells released from prometaphase arrest in DMSO, DRB, or Triptolide + DRB for 4 hours (left) and 9 hours (right). Percentages in the upper right corner represent the percent of cells with a G1 DNA content. Similar results were seen in a replicate experiment.
cells treated with transcription inhibitors for 9 hours during release from prometaphase arrest, we measured $>78\%$ efficiency by RT-qPCR in cells treated with DRB and $>93\%$ efficiency in cells treated with Triptolide + DRB (Fig. 3.12c, right). The cell cycle status of these cultures varied very little with the use of transcription inhibitors, resulting in 77-81% of cells with G1 DNA content measured by FACS analysis after 4 hours release from prometaphase arrest and 81-84% G1 in cultures released for 9 hours (Fig. 3.12e). We note that these experiments already corrected for any changes in the cell cycle as described in the Methods.

**Chromosome conformational changes in cells entering G1 with no active transcription**

To assess chromosome conformational changes in cells released from prometaphase arrest and entering G1 with no active transcription, we performed Hi-C on cells released in the presence of DMSO, DRB, or Triptolide + DRB for 4 hours (Fig. 3.13a) or 9 hours after prometaphase arrest (Fig. 3.13d). Hi-C chromatin interaction maps for cells in these different treatment conditions appear very similar by eye, all having the typical features expected for cells in G1. We note, however, that the checker-board pattern off of the diagonal representing spatially distinct compartments is weaker in all samples released for 4 hours compared to those released for 9 hours (Fig. 3.13a, top, Fig. 3.13d, top), as seen previously (Chapter II; Abramo et al., 2019).

We next plotted interaction frequencies ($P$) as a function of genomic distances ($s$) between loci to derive properties of chromosome folding for prometaphase arrested cells released in the presence of DMSO, DRB, or Triptolide + DRB for 4 hours (Fig. 3.13b-c, left) and 9 hours (Fig. 3.13e-f, left). We observe the typical decay pattern of chromatin
FIGURE 3.13: Hi-C analysis of cell cultures entering G1 in the absence of active transcription

(a) Hi-C interaction maps for prometaphase arrested cells released in the presence of DMSO, DRB, or Triptolide + DRB for 4 hours. Data for chromosome 14 are shown for two resolutions: 200 kb (top row, for entire right arm) and 40 kb (bottom row, for 36.5 Mb – 42 Mb regions). (b) Left: Contact frequency \( P \) versus genomic distance \( s \) for cis Hi-C data from synchronous cell cultures released from prometaphase arrest in the presence of transcription inhibitors for 4 hours (top) and percent trans reads in each Hi-C library (bottom). Right: Derivative from \( P(s) \) plots shown on left. (c) \( P(s) \) and derivative \( P(s) \) plots for replicate 2, as in panel b. (d) Hi-C interaction maps, as in panel a, for prometaphase arrested cells released in the presence of DMSO, DRB, or Triptolide + DRB for 9 hours. (e) Left: \( P(s) \) plots for cis Hi-C data from synchronous cell cultures released from prometaphase arrest in the presence of transcription inhibitors for 4 hours (top) and
percent trans reads in each Hi-C library (bottom), as in panel b. Right: Derivative from $P(s)$ plots shown on left. (f) $P(s)$ and derivative $P'(s)$ plots for replicate 2, as in panel e.
interaction maps of G1 cells for all conditions. To further explore the role of transcription on active processes of G1 chromatin formation, we calculated the derivatives of $P(s)$ for prometaphase cells released in transcription inhibitors for 4 hours (Fig. 3.13b-c, right) and 9 hours (Fig. 3.13e-f, right). We determined that cells entering G1 have an average cohesin extruded loop size of about 100 kb at both 4 hours and 9 hours release from prometaphase arrest (DMSO treated), which is consistent with our findings for nonsynchronous cells above, as well as in Chapter II of this thesis (Abramo et al., 2019).

Cells treated with DRB or Triptolide + DRB during release from prometaphase arrest for 4 hours showed a local maximum in the derivative of $P(s)$ that was similar to that of the DMSO treated cells, suggesting the formation of similarly sized loops (Fig. 3.13b-c, right). In contrast, cells treated with transcription inhibitors during release from prometaphase arrest for 9 hours showed a local maximum in the derivative $P(s)$ that was slightly shifted to the right of that of the DMSO treated cells (Fig. 3.13e-f, right), suggesting that cohesin extruded loops formed in the absence of transcription are slightly larger than the average 100 kb loop formed in the DMSO treated cells. Further, the depth of the minimum of the derivative of $P(s)$ decreases for cells treated with transcription inhibitors, indicating that the cohesin linear density on the chromatin might be reduced for cells released from prometaphase arrest for 9 hours in the absence of active transcription.

**Compartmentalization is reduced for cells entering G1 in the absence of active transcription**

Visual inspection of the checker-board pattern in the chromatin interaction maps suggests very little changes in compartment establishment during transcription inhibition for 4 hours (Fig. 3.13a, top) or 9 hours (Fig. 3.13d, top) after release from prometaphase
arrest. To confirm these observations, we used eigenvector decomposition, as done previously for the nonsynchronous experiments, to determine the locations of A and B compartments and the strength of compartmentalization in cells entering G1 in the absence of transcription. We find that the PC1 values defined for the Hi-C libraries of prometaphase arrested cells released in the presence of DRB or Triptolide + DRB for 4 hours (Fig. 3.14a) or 9 hours (Fig. 3.14c) correlate very well with that of the DMSO treated cells during prometaphase release for 4 hours or 9 hours, respectively. This suggests that the assignment of A and B compartments is established properly in the absence of transcription. We plotted the PC1 values along chromosome 14 for Hi-C libraries are cells released from prometaphase arrested in the presence of DMSO, DRB, or Triptolide + DRB for 4 hours (Fig. 3.14b) and 9 hours (Fig. 3.14d). As expected from the correlation values in panels a and c, we observe very little differences between the PC1 tracks at each release time point.

Next, we quantified the strength of compartmentalization for cells entering G1 in the absence of active transcription by creating "saddle plots" (Nora et al., 2017) of interactions between loci arranged by the DMSO PC1 values for either prometaphase arrested cells released for 4 hours (Fig. 3.15a) or 9 hours (Fig. 3.15d). We find that after 4 hours release from prometaphase arrest, compartmentalization in the DMSO control is not yet fully achieved and increases much more in t = 9 hours, as expected and seen previously (Chapter II; Abramo et al., 2019). Even so, we observe that the compartment scores are lower in cells treated with transcription inhibitors for 4 hours as they enter G1, suggesting that the kinetics of compartmentalization are delayed in cells entering G1 in the absence of transcription (Fig. 3.15a-b, Fig. 3.16a). By 9 hours release from
FIGURE 3.14: Compartment boundaries and type are properly established in cells entering G1 in the absence of active transcription
(a) Control, DMSO treated principal component 1 (PC1) values from genome wide eigen vector decomposition on 200 kb binned Hi-C datasets versus PC1 values generated from the Hi-C libraries of cell cultures released from prometaphase arrest for 4 hours in the presence of transcription inhibitors (left: DRB, right: Triptolide + DRB). Lower right corner
represents Pearson correlation value for each plot. Similar r-values can be seen between replicate Hi-C libraries (DMSO R1 v. R2: 0.946, DRB R1 v. R2: 0.96, T+D R1 v. R2: 0.967.). Similar results between conditions were also seen in a replicate experiment (DMSO R2 v. DRB R2: 0.951, DMSO R2 v. T+D R2: 0.938). (b) PC1 along Chromosome 14 for Hi-C data obtained from prometaphase arrested cells released and treated with DMSO, DRB, or Triptolide + DRB for 4 hours. (c) Control, DMSO treated principal component 1 (PC1) values from genome wide eigen vector decomposition on 200 kb binned Hi-C datasets versus PC1 values generated from the Hi-C libraries of cell cultures released from prometaphase arrest for 9 hours in the presence of transcription inhibitors (left: DRB, right: Triptolide + DRB), as in panel a. Lower right corner represents Pearson correlation value for each plot. Similar r-values can be seen between replicate Hi-C libraries (DMSO R1 v. R2: 0.977, DRB R1 v. R2: 0.971, T+D R1 v. R2: 0.979.). Similar results between conditions were also seen in a replicate experiment (DMSO R2 v. DRB R2: 0.958, DMSO R2 v. T+D R2: 0.962). (d) PC1 along Chromosome 14 for Hi-C data obtained from prometaphase arrested cells released and treated with DMSO, DRB, or Triptolide + DRB for 9 hours, as in panel b. Similar results were seen in a replicate experiment.
FIGURE 3.15: Compartmentalization is weaker in cells entering G1 in the absence of active transcription

(a) Saddle plots of Hi-C data binned at 200 kb resolution for prometaphase arrested cells released in the presence of DMSO, DRB, or Triptolide + DRB for 4 hours. Saddle plots were calculated using the PC1 obtained from the Hi-C data of the prometaphase arrested cells released in the presence of DMSO for 4 hours. Numbers at the center of the heatmaps indicate compartment strength calculated as the ratio of \((AA+BB)/(AB+BA)\) using the mean values from dashed corner boxes (bin size = 10). (b) Compartment strength values, as shown on heatmaps in panel a, normalized to the compartment strength of the DMSO control. (c) Compartment strength of A (solid) and B (striped) regions normalized to the strength of A and B in the DMSO control. (d) Saddle plots of Hi-C data binned at 200 kb resolution, as in panel a, for prometaphase arrested cells released in the presence of DMSO, DRB, or Triptolide + DRB for 9 hours. Saddle plots were calculated using the PC1 obtained from the Hi-C data of the prometaphase arrested cells released in the presence of DMSO for 9 hours. Numbers at the center of the heatmaps indicate compartment strength calculated as the ratio of \((AA+BB)/(AB+BA)\) using the mean values from dashed corner boxes (bin size = 10). (e) Compartment strength values, as shown on heatmaps in panel d, normalized to the compartment strength of the DMSO control. (f) Compartment strength of A (solid) and B (striped) regions normalized to the strength of A and B in the DMSO control.
FIGURE 3.16: Compartmentalization is weaker in cells entering G1 in the absence of active transcription (replicate 2)
Compartment strength values calculated from saddle plots of Hi-C data binned at 200 kb resolution, as in Fig. 3.15, for replicate 2 of prometaphase arrested cells released in the presence of DMSO, DRB, or Triptolide for (a) 4 hours and (b) 9 hours and normalized to compartment strength of the DMSO control of each set.
prometaphase arrest, we expect that compartmentalization is fully achieved, as discussed in Chapter II of this thesis (Abramo et al., 2019). We find that prometaphase cells released in the presence of DRB or Triptolide + DRB for 9 hours do not reach the same compartmentalization level as their DMSO treated counterparts (Fig. 3.15d-e, Fig. 3.16b), suggesting that compartmentalization is severely delayed or can never be achieved to the same level as the control when cells are entering G1 in the absence of transcription. We further analyzed the independent contribution of A-A interactions in compartmentalization and B-B interactions to see if there is a delay or change specifically in the interactions of one compartment in comparison to the other. We find that the lower compartment strength values for prometaphase cells released in the presence of transcription inhibitors for 4 hours (Fig. 3.15c) and 9 hours (Fig. 3.15f) are largely due to changes in A-A interactions, however, interactions for both A-A and B-B are reduced in all conditions.

Since the PC1 values correlated so well, we also created “saddle plots” of interactions between loci arranged by the PC1 values determined from each individual sample (Fig. 3.17). Similar to the previous results, we find that cells released from prometaphase arrest in the presence of transcription inhibitors for 4 hours (Fig. 3.17, left) and 9 hours (Fig. 3.17, right) establish compartments that are less strong in their intra-compartment interaction compared to inter-compartment interactions, especially for the later t = 9 hours by which time compartmentalization is achieved in control cells.

We next quantified compartmentalization as a function of genomic distance to test if the delay/reduction in compartment strength is distance dependent (Fig. 3.18). We find that for cells released from prometaphase arrest and entering G1 for 4 hours in the
FIGURE 3.17: Compartmentalization using PC1 tracks generated from each Hi-C dataset also shows weaker strength in cells entering G1 in the absence of active transcription

Heatmaps represent the compartment strength values calculated from saddle plots of Hi-C data binned at 200 kb resolution for prometaphase arrested cells released in the presence of DMSO, DRB, or Triptolide + DRB for 4 hours (left) and 9 hours (right). Saddle plots were generated either using the PC1 values obtained from the Hi-C data of the DMSO control in each set (“DMSO PC1”) or the Hi-C data of each individual sample (“Own PC1 track”).

FIGURE 3.18: Compartmentalization is weaker at all distances in cells entering G1 in the absence of active transcription

Heatmaps of compartment scores (bin size = 10) at different genomic distances for prometaphase arrested cells released in the presence of DMSO, DRB, or Triptolide + DRB for (a) 4 hours and (b) 9 hours, split by interaction type. Overall compartmentalization represents the mean values for (AA+BB)/(AB+BA), A-A compartment scores use the mean values for AA/AB, and B-B compartment scores are calculated using the mean values for BB/AB. Saddle plots were calculated using the PC1 obtained from the Hi-C data of the DMSO control in each experiment set. Similar results were seen in a replicate experiment.
absence of active transcription overall compartment strength is reduced at all distances (Fig. 3.18a, left). While this seems to be due to a reduction in compartment strength of both A compartments and B compartments at all distances (Fig. 3.18a, middle-right), we observe a larger effect in B-B compartment interactions for shorter distances (< 8 Mb), where we have shown previously that B-B compartmentalization is strongest in untreated cells (Chapter II; Abramo et al., 2019). In contrast, we find that transcription inhibition for 4 hours as cells are released from prometaphase arrest primarily reduces compartmentalization at longer distances (>8 Mb) in A-A interactions, where these compartment interactions are known to be more prominent. For prometaphase cells released in the presence of DMSO, DRB, or Triptolide + DRB for 9 hours, we find that compartment strength is also reduced for all distances (Fig. 3.18b, left). When analyzed separately, we find that A-A compartmentalization is reduced much more than B-B compartment interactions for transcription inhibited cells compared to the DMSO control treated cells (Fig. 3.18b, middle-right). This suggests that inhibiting transcription as cells enter G1 affects the ability of the cell to fully compartmentalize genomic regions into active and inactive sub-nuclear neighborhoods, preferentially influenced by the inability of A-A compartment interactions to form fully.

Lastly, we quantified the strength of compartmentalization for interchromosomal contacts by generating "saddle plots" of trans chromatin interaction frequencies arranged according to the PC1 value of the DMSO control sample for each set (Fig. 3.19). We find that the kinetics of compartmentalization in trans is quicker in the absence of transcription in cells released from prometaphase arrest for 4 hours (Fig. 3.19a-c). This faster compartmentalization occurs in both A-A and B-B chromatin interactions in trans (Fig.
FIGURE 3.19: Trans compartmentalization is quicker in cells entering G1 in the absence of active transcription, however, weakens by 9 hours release from prometaphase arrest when compartmentalization should be fully established. 

(a) Saddle plots of trans Hi-C data binned at 200 kb resolution for prometaphase arrested cell cultures released in the presence of DMSO, DRB, or Triptolide + DRB for 4 hours. Saddle plots were calculated using the PC1 obtained from the Hi-C data of the DMSO control in each experiment set. (b) Trans compartment strength (bin size = 10) values normalized to trans compartment strength of the DMSO control for replicate 1 and (c) replicate 2. (d) Saddle plots of trans Hi-C data binned at 200 kb resolution, as in panel a, for prometaphase arrested cell cultures released in the presence of DMSO, DRB, or Triptolide + DRB for 9 hours. (e) Trans compartment strength (bin size = 10) values normalized to trans compartment strength of the DMSO control for replicate 1 and (f) replicate 2.
3.19a). We find a slightly different result for prometaphase arrested cells released in transcription inhibitors for 9 hours (Fig. 3.19d-f). For cells released for 9 hours and entering G1 in the presence of DRB, we find B-B compartment interactions are unchanged, but A-A compartmentalization is reduced yielding a reduced trans compartment strength compared to DMSO treated cells (Fig. 3.19d-f, “DRB”). In contrast, we find that cells released from prometaphase arrested in the presence of Triptolide + DRB for 9 hours are able to establish A-A interactions in trans and have increased B-B interactions, yielding an overall increased compartment score in trans in the absence of transcription. This suggests that the establishment of compartment interactions in trans may rely slightly on active transcription, though these values are still very similar. Further compartmentalization in trans may be directly related to compartmentalization in cis such that a decrease in cis strength of compartmentalization in the absence of transcription leads to an increase in trans.

**TADs are weaker in cells entering G1 without active transcription**

To test if TADs are properly established in cells entering G1 in the absence of transcription, we derived insulation values and quantified strength of insulation at domain boundaries. We find that the insulation values derived from the Hi-C libraries of prometaphase arrested cells released in the presence of transcription inhibitors for 4 hours (Fig. 3.20a) and 9 hours (Fig. 3.20e) correlate very well with those of cells released from prometaphase arrested in the presence of DMSO for the same amount of time. We plotted the insulation values along Chromosome 14 to directly visualize the correlation of these tracks and find very few differences between the three conditions for t = 4 hours (Fig. 3.20b) and t = 9 hours (Fig. 3.20f). The domain boundaries, defined as the local
FIGURE 3.20: Insulation at TAD boundaries is weaker in cells entering G1 in the absence of active transcription

(a) Control, DMSO treated insulation values generated from genome wide insulation analysis on 40 kb binned Hi-C datasets versus insulation values generated from the Hi-C libraries of prometaphase arrested cell cultures released for 4 hours in the presence of transcription inhibitors (left: DRB, right: Triptolide + DRB). Lower right corner represents Pearson correlation value for each plot. Similar results between conditions were also seen in a replicate experiment (DMSO R2 v. DRB R2: 0.969, DMSO R2 v. T+D R2: 0.968). (b) Top: Insulation profile along Chromosome 14 for Hi-C data obtained from prometaphase arrested cells released in the presence of DMSO, DRB, or Triptolide + DRB for 4 hours. Bottom: Insulation boundaries defined as local minima in the insulation profiles for DMSO replicate 1 (top blue bar), DMSO replicate 2 (lower blue bar), DRB (orange bar), and Triptolide + DRB (green bar). Black bars represent the regions defined as A compartments in a G1 HeLa S3 Hi-C dataset. (c) Pileup of the average insulation profile across TAD boundaries on intact HeLa S3 chromosomes defined in the Hi-C data of DMSO treated cells in each set. Left: replicate 1, Right: replicate 2. The number of
boundaries used in each pileup is noted at the bottom of each plot. (d) Violin plots of the insulation strength at each individual boundary, separated by location in either A (left) or B (right) compartment type. P-value from Wilcoxon rank-sum test measured between each control-treatment pair. The horizontal white bars represent the median. Boxplots within violin plots (gray) represent the first and third quartiles of the dataset. This panel represents data from the pooled Hi-C datasets of 2 replicates. (e) Control, DMSO treated insulation values generated from genome wide insulation analysis on 40 kb binned Hi-C datasets versus insulation values generated from the Hi-C libraries of prometaphase arrested cell cultures released for 9 hours in the presence of transcription inhibitors (left: DRB, right: Triptolide + DRB). Lower right corner represents Pearson correlation value for each plot. Similar results between conditions were also seen in a replicate experiment (DMSO R2 v. DRB R2: 0.960, DMSO R2 v. T+D R2: 0.947). (f) Top: Insulation profile along Chromosome 14 for Hi-C data obtained from prometaphase arrested cells released in the presence of DMSO, DRB, or Triptolide + DRB for 9 hours. Bottom: Insulation boundaries defined as local minima in the insulation profiles for DMSO replicate 1 (top red bar), DMSO replicate 2 (lower red bar), DRB (purple bar), and Triptolide + DRB (brown bar). Black bars represent the regions defined as A compartments in a G1 HeLa S3 Hi-C dataset. (g) Pileup of the average insulation profile across TAD boundaries on intact HeLa S3 chromosomes defined in the Hi-C data of DMSO treated cells in each set. Left: replicate 1, Right: replicate 2. (h) Violin plots of the insulation strength at each individual boundary, separated by location in either A (left) or B (right) compartment type. P-value from Wilcoxon rank-sum test measured between each control-treatment pair. The horizontal white bars represent the median. Boxplots within violin plots (gray) represent the first and third quartiles of the dataset. This panel represents data from the pooled Hi-C datasets of 2 replicates.
minima in the insulation track (Lajoie, Dekker, and Kaplan 2015) for Hi-C libraries of prometaphase arrested cells released in DMSO, DRB, or Triptolide + DRB for 4 hours were very similar between samples (Fig. 3.20b). Though regions of both missing boundaries and new boundaries are observed in the cells entering G1 in the absence of transcription, we also see changes in these areas between replicates of DMSO treatment and it is likely due to a slight shift in regions called as boundaries and not due to boundary changes in a specific compartment type. We observe similar results for the cells released from prometaphase arrest for 9 hours in the absence of transcription (Fig. 3.20f).

To quantify the strength of TAD boundaries, defined as the lack of interactions across boundaries, we aggregated insulation profiles at domain boundaries that did not overlap with compartment boundaries for prometaphase arrested cells released in DMSO, DRB, or Triptolide + DRB for 4 hours (Fig. 3.20c) and 9 hours (Fig. 3.20g). We find the average insulation profile across domain boundaries is reduced in two replicates for all conditions. Next, we quantified the insulation strength of each individual boundary separated by their location in either an A or B compartment for cells released from prometaphase arrest in DMSO, DRB, or Triptolide + DRB for 4 hours (Fig. 3.20d) and 9 hours (Fig. 3.20h). While we find that the distribution of the insulation strengths around boundaries in both A and B compartments is significantly reduced in cells entering G1 in the absence of transcription, this is much more significant for boundaries located in A compartments (Fig. 3.20d, left, Fig. 3.20h, left). While this may suggest that active transcription is involved in establishing insulation at TAD boundaries, especially in A compartments, this effect is still very small. We also note that a slight variability in the
time of harvesting these cells could lead to slight changes in the quantification of these structures, especially at early time points (see ‘Conclusion’ below).

**Insulation around TSSs is weaker in cells entering G1 in the absence of transcription**

To quantify the strength of insulation around TSSs, we aggregated the insulation values 500 kb upstream and downstream of the highest expressed genes on the six structurally intact chromosomes in HeLa S3 cells from Hi-C libraries of prometaphase arrested cells released in the presence of DMSO, DRB, or Triptolide + DRB for 4 hours (Fig. 3.21a-b) and 9 hours (Fig. 3.21f-g). We find that insulation surrounding TSSs is slightly weaker in cells that entered G1 in the absence of transcription. To compare insulation at each of the 500 TSSs individually, we plotted the difference of insulation values 500 kb upstream and downstream of TSSs from Hi-C libraries of cells released from prometaphase arrest in DMSO from the insulation values from Hi-C libraries of prometaphase arrested cells released in transcription inhibitors for 4 hours (Fig. 3.21c) and 9 hours (Fig. 3.21h). We observe that insulation values surrounding many TSSs are higher in the Hi-C libraries generated from transcription inhibited cells, especially those cells that were released from prometaphase arrest for 9 hours when transcription is known to be high in untreated cells (Fig. 3.11). The average difference in insulation profiles across these TSSs is slightly higher for transcription inhibited cells, with a maxima around the location of the TSSs, suggesting that there is less insulation at these sites in the absence of active transcription (Fig. 3.21c, lower, Fig. 3.21h, lower). We quantified the insulation strength at each of these individual TSSs and find that the distribution of insulation strengths for prometaphase arrested cells released in transcription inhibitors
FIGURE 3.21: Insulation at active TSSs is weaker in cells entering G1 in the absence of active transcription

(a) Pileup of the average insulation profile across the 500 most active TSSs for HeLa S3 Hi-C data of prometaphase arrested cells released in the DMSO, DRB, or Triptolide + DRB for 4 hours replicate 1 and (b) replicate 2. (c) Top: Heatmap showing the difference in insulation values 500 kb upstream and downstream of the 500 most active TSSs. Bottom: Average insulation profile of the difference in insulation values across the 500 most active TSSs for Hi-C datasets of transcription inhibited cells versus DMSO control Hi-C data. (d) Violin plots of the insulation strength at each of the 500 most active TSSs in replicate 1 and (e) replicate 2 of Hi-C on prometaphase arrested cells released in the presence of DMSO, DRB, or Triptolide + DRB for 4 hours. P-value from Wilcoxon rank-sum test measured between each control-treatment pair. The horizontal white bars represent the median. Boxplots within violin plots (gray) represent the first and third quartiles of the dataset. (f) Pileup of the average insulation profile across the 500 most active TSSs for HeLa S3 Hi-C data of prometaphase arrested cells released in the DMSO,
DRB, or Triptolide + DRB for 9 hours replicate 1 and (g) replicate 2. (h) Top: Heatmap showing the difference in insulation values 500 kb upstream and downstream of the 500 most active TSSs, as in panel c, for Hi-C datasets of transcription inhibited cells versus DMSO control Hi-C data upon prometaphase release for 9 hours. Bottom: Average insulation profile of the difference in insulation values across the 500 most active TSSs for Hi-C datasets of transcription inhibited cells versus DMSO control Hi-C data. (i) Violin plots of the insulation strength at each of the 500 most active TSSs in replicate 1 and (j) replicate 2 of Hi-C on prometaphase arrested cells released in the presence of DMSO, DRB, or Triptolide + DRB for 9 hours. P-value from Wilcoxon rank-sum test measured between each control-treatment pair. The horizontal white bars represent the median. Boxplots within violin plots (gray) represent the first and third quartiles of the dataset.
for 4 hours (Fig. 3.21d-e) and 9 hours (Fig. 3.21i-j) is much lower than the distribution of insulation strengths for control DMSO treated cells. Therefore, cells entering G1 depend on active transcription to establish insulation around active TSSs. It is also possible that active transcription establishes insulation at these sites. This is also supported by the results showing a much larger difference in insulation after 9 hours release which is when transcription is much more active.

**Loops are established in cells entering G1 without active transcription**

Last, we tested if looping interactions are established in cells entering G1 in the absence of active transcription. We aggregated Hi-C data at 507 looping interactions (Rao, Huntley et al., 2014) for prometaphase arrested cells released in the presence of DMSO, DRB, or Triptolide + DRB for 4 hours (Fig. 3.22a, left) and 9 hours (Fig. 3.22a, right). We observe very little difference from visual inspection of the aggregate heatmaps, therefore, we quantified the strength of interactions between each pairwise interaction (Fig. 3.22b). We find, in two independent replicates, that loops from Hi-C libraries of cells released for only 4 hours from prometaphase arrest are established as normal, with some very slight non-significant variation. In contrast, for loops in Hi-C libraries of cells released for 9 hours from prometaphase arrest, we find that loops are established similarly between DMSO control treated cells and Triptolide + DRB treated cells, however, cells treated with DRB alone had increased strength for these looping interactions. This could be an effect of using DRB which should freeze RNA polymerase II in place at the pausing site, whereas using Triptolide will induce proteasomal degradation of RNA polymerase II. We tried to further parse this out by separating loops by size, presence of CTCF motif, and ChIP-seq signal for POLR2A, as we did in Fig. 3.10 for nonsynchronous cells,
FIGURE 3.22: Loops are established in cells entering G1 in the absence of active transcription

(a) Aggregate Hi-C data binned at 10 kb resolution at chromatin loops on intact HeLa S3 chromosomes (n = 507 loops) identified in Rao, Huntley et al., 2014 for prometaphase arrested cells (left) released in the presence of DMSO, DRB, or Triptolide + DRB for 4 hours (middle) and 9 hours (right). (b) Loop strength values for replicate 1 (top) and replicate 2 (bottom). Bar plots represent mean loop strengths with error bars at +/- 95% confidence interval. Wilcoxon rank-sum test between each control-treatment pair showed no significant differences in loop strength.
however we saw no significant differences in the distribution of loop strengths separated in this way for prometaphase arrested cells released in DMSO, DRB, or Triptolide + DRB for 4 hours and 9 hours.

**Conclusion**

In this chapter, we describe the typical chromatin features observed in Hi-C analysis of interphase cells and how these 3D structures change interactions upon transcription inhibition. While we find some quantitative effects in the strength of each of these structures (Table 3.4), overall we conclude that transcription is not required for either the maintenance or establishment of compartments, TADs, and loops. These features can be observed on Hi-C interaction maps regardless of transcription activity. The data presented here supports the findings of many related studies (Li, Lyu, Hou et al., 2015; Hug et al., 2017, Du et al., 2017; Ke, Xu, Chen, Feng et al., 2017; Barutcu et al., 2019). For example, chemical inhibition of transcription in Drosophila shows compartment structures, however compartmentalization is decreased (Li, Lyu, Hou et al., 2015; Hug et al., 2017). More recent experiments in mammalian cells also reveal the presence of compartments and TADs after transcription inhibition, but observe less of a difference to compartmentalization of genomic loci and a decrease in the strength of TAD boundaries (Barutcu et al., 2019). This might be the result of displaced cohesin from CTCF sites, therefore disrupting TAD boundaries. This is in agreement with our samples treated with transcription inhibitors for 9 hours that show $P(s)$ curves which could
Table 3.4: Overview of chromatin organization in transcription inhibited cells

<table>
<thead>
<tr>
<th>Synchronization State</th>
<th>Nonsynchronous</th>
<th>Released from prometaphase arrest 4 hours</th>
<th>Released from prometaphase arrest 9 hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DRB</td>
<td>Triptolide + DRB</td>
<td>DRB</td>
</tr>
<tr>
<td>Chromatin interaction scaling, ( P(s) )</td>
<td>maintained</td>
<td>maintained</td>
<td>established similarly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>dramatically.cohesin extruded loops; less cohesin linear density</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>slightly larger cohesin extruded loops; less cohesin linear density</td>
</tr>
<tr>
<td>Cis compartmentalization</td>
<td>weaker, mostly A-A</td>
<td>weaker, mostly A-A</td>
<td>weaker, B-B more weak long distances; A-A more weak long distances</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>slightly larger cohesin extruded loops; less cohesin linear density</td>
</tr>
<tr>
<td>Trans compartmentalization</td>
<td>maintained/ stronger</td>
<td>maintained/ stronger</td>
<td>established similarly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>weaker</td>
</tr>
<tr>
<td>Insulation strength at TAD boundaries</td>
<td>weaker, mostly for boundaries in B comps</td>
<td>weaker, mostly for boundaries in B comps</td>
<td>weaker, mostly for boundaries in A comps</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>weaker, mostly for boundaries in A comps</td>
</tr>
<tr>
<td>Insulation strength at highly active TSSs</td>
<td>weaker</td>
<td>weaker</td>
<td>weaker</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>weaker</td>
</tr>
<tr>
<td>Looping interactions</td>
<td>maintained/ stronger</td>
<td>maintained/ stronger</td>
<td>established similarly</td>
</tr>
</tbody>
</table>

represent slightly bigger cohesin extruded loops and slightly reduced cohesin linear density (Fig. 3.13, Table 3.4).

Interestingly, we observe only very few differences between Hi-C data of cells inhibited for just transcription elongation (using DRB) or for both transcription initiation and elongation (using Triptolide + DRB). A major difference in the use of these chemicals is the presence of RNAPII. Triptolide has been shown to induce proteasomal degradation of RNAPII (Bensaude, 2011). While in most analyses we observe similar effects to the chromatin organization, a 9 hours treatment with Triptolide + DRB shows increased trans compartmentalization compared to 9 hour treatment with DRB alone (Table 3.4). Since it has been shown that many trans interactions are regulated by clustering of loci at RNAPII transcription factories (Iborra et al., 1996; Sutherland and Bickmore, 2009; Schoenfelder, Sexton, Chakalova et al., 2010; Branco and Pombo, 2006; Chambeyron and Bickmore,
2004; Osborne et al., 2004), this is a very surprising result. This result implies that the presence of RNAPII is inhibiting some of the trans contacts that could be occurring. Interestingly, we reveal that the increased trans compartmentalization is specifically do to an increase in B-B contacts while A-A contacts in trans are maintained (Fig. 3.19). This, therefore, suggests that A-A contacts in trans are able to form after treatment with Triptolide + DRB (and with the assumption that RNAPII transcription factories are dissolved). Further, this may suggest that some B-B interactions are normally drawn into the transcription factories due to close association with active loci. Similar conclusions were found when RNAPII was depleted by an auxin-inducible degron system (Jiang, Huang, Lun, Li et al., 2020). However, since exact measurements of RNAPII are not shown in the experiments we present here, we cannot assume complete degradation of RNAPII.

Based on these results, global interphase chromatin state does not seem to rely on active transcription. Instead, we propose the 3D genome is organized in to compartments and TADs to facilitate proper transcription. In a normal cell, A compartments will re-localize to transcription factories. This will bring together active loci in both cis and trans. However, these regions still become spatially separated in transcription inhibited cells, likely by phase separation due to the epigenetic marks or other RNAPII-independent transcription-related proteins localized to the chromatin. Further, in the absence of transcription, TSSs are less insulated than upon transcriptional activation in which these regions bound by RNAPII and other transcription machinery become locally more accessible but do not change the global chromatin organization. We note however, that a large caveat to these experiments is that the transcription block is
not complete (Fig. 3.1 and Fig. 3.12). Therefore, it is possible that a very low level of transcription is sufficient to establish and maintain compartments, TADs, and loops. In addition, early measurements suggest that a 30 minute delay in harvested cells released from mitotic arrest and treated with DRB would result in cells at similar stages in G1. If entry into G1 was delayed by even the slightest bit, then measurements of these structures at early time points could also vary. However, by 8.5 hours release from prometaphase compartments, TADs, and loops should all be established fully, and we therefore, do not expect that these results would be impacted by any delay in G1 entry.
CHAPTER IV: Role of RNA in chromatin organization

Abstract

RNA transcripts can interact with chromatin and influence changes in chromatin organization. The most common example of this is XIST RNA that is transcribed from and coats the inactive X chromosome inducing formation of a heterochromatic Barr body for dosage compensation. Several long non-coding RNA molecules, such as NEAT1 and FIRRE, have also been proposed as mediators of phase separation in the formation of active and inactive compartments, respectively. Further, the RNA binding regions of chromatin proteins, such as CTCF, have suggested a role for RNA in TAD establishment and maintenance. We used RNase A to degrade RNA in permeabilized cells and determine which features of chromatin organization rely on the presence of RNA molecules. While microscopy of permeabilized cells treated with RNase A for only 5 minutes shows dramatic morphological changes of chromatin, using Hi-C we find that RNA degradation causes much less dramatic changes in 3D chromatin contacts. While RNA does not seem to be required for the maintenance of chromosome structure, we describe quantitative effects on the strength of compartments, topologically associated domains, and looping interactions. We propose that RNA essentially coats each chromosome both with specific and non-specific interactions while filling the volume of the nucleus. In this model, upon RNA degradation chromosomes initially collapse onto themselves where RNA was previously coating each chromosome and forcing them into
a specific shape, then chromosomes slowly expand after 4 hours of dissociation into the newly empty nuclear space.

**Introduction**

The nucleus is a highly complex environment consisting of DNA, RNA, and many proteins. These components must coordinate within the nucleus to create a highly organized chromatin state that allows for proper DNA replication, chromosome segregation, and gene expression. As early as 30 years ago, RNA was shown to be a critical presence in the nucleus, without which chromatin collapsed and aggregated around the nucleoli, suggesting that RNA is not only a key structural element of the nuclear matrix but may also be involved in the 3D structure of chromatin (Nickerson et al., 1989). Since then, RNA has been implicated with many features of interphase chromatin organization. At a local chromatin level, RNA has been shown to bind directly to mediator proteins and coordinate enhancer-promoter interactions. Mediator proteins were observed to specifically interact with transcripts of enhancer elements (enhancer RNA, eRNA), thereby bridging enhancers with RNA polymerase II and their target promoter to form a loop (Lai et al., 2013). Depleting eRNAs or mediator complex caused a reduction in the loop formation between enhancers and their target promoter (Lai et al., 2013; Phillips-Cremens et al., 2013). Similarly, CTCF has also been shown to directly interact with non-coding RNA (ncRNA) in the formation of CTCF-CTCF loops and topologically associated domains (TADs). TADs are thought to be formed through an active loop extrusion process which is blocked at convergent sites of CTCF leading to a
loop between CTCF sites and an enrichment of chromatin interactions within the loop compared to outside of the loop (Rao, Huntley et al., 2014; de Wit, Vos, Holwerda, Valdes-Quezada et al., 2015; Guo, Xu et al., 2015; Vietri Rudan et al., 2015; Sanborn, Rao et al., 2015; Fudenberg, Imakaev et al., 2016). A distinct class of these CTCF-CTCF loops was discovered to require ncRNA for their establishment and maintenance (Saldana-Meyer et al., 2014; Kung et al., 2014; Hansen, Hsieh, Cattoglio et al., 2019; Saldana-Meyer et al., 2019; Thakur et al., 2019). Further, several RNA species have been shown to directly interact with DNA or DNA-binding proteins and form spatially segregated nuclear domains of heterochromatic and euchromatic compartments. *XIST*, for example, directly interacts in cis to coat the inactive X chromosome, recruit repressive proteins, such as PRC2, and form a heterochromatic Barr body (Clemson and Lawrence 1996; Engreitz et al., 2013; Simon, Pinter, Fang et al., 2013). Other long non-coding RNAs (lncRNA), such as *FIRRE* (functional intergenic repeating RNA element) contain repetitive RNA domains which can enable multivalent interactions and form a network of colocalized compartment regions from different chromosomes (Hacisuleyman et al., 2014). Finally, some lncRNAs can nucleate and maintain nuclear bodies. *MALAT1* and *NEAT1*, for example, concentrate proteins for transcription and RNA processing forming nuclear speckles and paraspeckles, respectively (Hutchinson, Ensminger et al., 2007; Clemson et al., 2009; West, David et al., 2014). Chromatin localization to each of the above-described nuclear domains allows for the proper regulation of cell-type specific gene expression. While much has been determined about the specific role for these individual and highly expressed lncRNAs, most lncRNAs are expressed at very low levels (Hangauer, Vaughn, and McManus, 2013). Therefore, it is unlikely for all lncRNAs to have a unique function in
chromatin organization, and much more likely for lncRNA functions to be redundant and act cooperatively. In this study, we aim to determine how RNA as a whole shapes nuclear organization.

Materials and Methods

Cell Culture

HeLa S3 CCL-2.2 cells (ATCC CCL-2.2) were cultured in DMEM, high glucose, GlutaMAX™ Supplement with pyruvate (Gibco 10569010) with 10% fetal bovine serum (Gibco 16000044) and 1% PenStrep (Gibco 15140) at 37°C in 5% CO₂.

RNase Treatment

Adherent nonsynchronous cells were plated and cultured overnight. The next day cells were washed in ice-cold CSK buffer (100 mM NaCl, 300 mM sucrose, 3 mM MgCl₂, 10 mM PIPES, pH 6.8). For control 'RNasin' treated samples, RNasin Plus RNase Inhibitor (Promega N2615) was added to the ice-cold CSK buffer at 1:25 dilution. After this quick wash, cells were permeabilized in ice-cold CSK buffer with 0.5% Triton-X-100 (Sigma T8787) for 3 minutes on ice. Again, for control 'RNasin' treated samples, RNasin Plus RNase Inhibitor was added to the permeabilization buffer at 1:25 dilution. Permeabilized cells were then quickly washed again in CSK buffer +/- RNasin and then incubated at 37°C for 5 minutes or 4 hours in pre-warmed treatments. Pre-warmed treatment for the 'Mock' (permeabilize only) condition was CSK buffer warmed at 37°C.
‘RNasin’ control samples were treated with pre-warmed CSK buffer + RNasin Plus RNase Inhibitor at 1:25. And lastly, ‘RNase A’ samples were treated with pre-warmed CSK buffer + 100 ug/mL RNase A (Roche 10109169001). All cells were fixed, according to the protocol for individual analyses, only after they were treated as described above.

**RNA isolation for fragment analyzer analysis**

Mock, RNasin, or RNase A treated cells were washed quickly with cold 1X PBS and lysed with 1 mL of TRIzol™ Reagent (ThermoFisher Scientific 15596018) for 5 minutes at room temperature. Samples were then collected in 15 mL tubes and kept at -20°C until further processing. Thawed TRIzol samples were treated with 200 µL RNase free chloroform, mixed, and incubated for 2 minutes at room temperature. Samples were then centrifuged and the aqueous phase was transferred to a new tube. 500 µL of isopropanol was added, along with 2 µL RNaseOUT™ Recombinant Ribonuclease Inhibitor (ThermoFisher Scientific 10777019) and 1 µL glycogen. Samples were mixed and incubated for 10 minutes at room temperature to precipitate the RNAs. Samples were then spun, washed with 75% EtOH, dried and resuspended in 20 µL RNase-free water plus 1 µL RNaseOUT™. To enhance resuspension, samples were incubated at 60°C for 15 minutes. 9 µg of RNA was used in the RiboMinus Human/Mouse transcriptome Isolation kit (ThermoFisher Scientific K155001) and ribosomal RNA was removed according to the manufacturer instructions. RNA was analyzed by fragment analyzer.

**Microscopy**

Cells used in microscopy analysis were grown on poly-L-lysine (PLL) coated coverslips prior to Mock, RNasin, or RNase A treatment. Coverslips were prepared by
coating coverslips in 1 mg/mL poly-L-lysine (PLL) in ddH₂O (Sigma P4832) overnight at 4°C. Coverslips were then washed in ddH₂O 10X, rinsed in 100% EtOH, and allowed to dry in a sterile hood. Dried PLL-coated coverslips were then carefully transferred into new tissue culture dishes to grow cells. Cells were plated in 35 mm dishes with PLL-coated coverslips at 0.5 x 10⁶ cells and allowed to adhere to the coverslip at 37°C overnight.

After treatment, cells were fixed in 4% paraformaldehyde in PBS for 15 minutes at room temperature. Coverslips were washed 3X in 1X PBS and cells were stained in 2 μg/mL DAPI (in 1X PBS) (Thermo D1306) for 1 minute at room temperature. Coverslips were quickly rinsed in PBS, then ddH₂O, and mounted to slides using ProLong™ Gold Antifade Mountant (Thermo P36934). For image acquisition, we used a Nikon Eclipse Ti microscope. Imaging was performed using an Apo TIRF, N.A. 1.49, 60x oil immersion objective (Nikon) and a Zyla sCMOS camera (Andor). Images were acquired using NIS-Elements 4.4.

Hi-C Analysis

Approximately 5 x 10⁶ cells at each time point were fixed in 1% Formaldehyde (Fisher BP531-25) diluted in serum-free DMEM for Hi-C analysis. Hi-C was performed as described in Belaghzal et al. (Belaghzal et al., 2017). Briefly, flash-frozen cross-linked cell culture samples were lysed then digested with DpnII at 37°C overnight. Next, the DNA overhanging ends were filled with biotin-14-dATP at 23°C for 4 hours and ligated with T4 DNA ligase at 16°C for 4 hours. DNA was then treated with proteinase K at 65°C overnight to remove crosslinked proteins. Ligation products were purified, fragmented by sonication to an average size of 200 bp, and size selected to fragments 100 - 350 bp. We then performed end repair and dA-tailing and selectively purified biotin tagged DNA using
streptavidin beads. Illumina TruSeq adaptors were added to form the final Hi-C ligation products, samples were amplified and PCR primers were removed. Hi-C libraries were then sequenced by PE50 bases on an Illumina HiSeq4000.

**Hi-C Data Processing**

Hi-C PE50 fastq sequencing files were mapped to hg19 and hg38 human reference genome using distiller-nf mapping pipeline (https://github.com/mirnylab/distiller-nf). In brief, bwa mem was used to map fastq pairs in a single-side regime (-SP). Aligned reads were classified and deduplicated using pairtools (https://github.com/mirnylab/pairtools), such that uniquely mapped and rescued pairs were retained and duplicate pairs (identical positions and strand orientations) were removed. We refer to such filtered reads as valid pairs. Valid pairs were binned into contact matrices at 10 kb, 20 kb, 40 kb, 200 kb, and 1 Mb resolutions using cooler (Abdennur and Mirny, 2019). Iterative balancing procedure (Imakaev, Fudenberg et al., 2012) was applied to all matrices, ignoring the first 2 diagonals to avoid short-range ligation artifacts at a given resolution, and bins with low coverage were removed using MADmax filter with default parameters. Resultant “.cool” contact matrices were used in downstream analyses using cooltools (https://github.com/mirnylab/cooltools). Hi-C statistics for each sample are in Table 4.1.

Note that libraries using condition “Perm Only” are not shown in any of the figures for this Chapter. This control shows that without the addition of RNasin, cells will naturally collapse likely due to the activation of endogenous endonucleases, and therefore, the ‘RNasin’ Hi-C libraries serve as a better control.
### Table 4.1: Hi-C library statistics

<table>
<thead>
<tr>
<th>Library Name</th>
<th>Condition</th>
<th>Treatment Time</th>
<th>Replicate</th>
<th>Total Reads</th>
<th>Valid Pairs</th>
<th>% Cis</th>
</tr>
</thead>
<tbody>
<tr>
<td>TB-HiC-Dpn-Perm</td>
<td>Perm Only</td>
<td>5 min</td>
<td>1</td>
<td>115,486,479</td>
<td>64,641,960</td>
<td>55%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-RNasin</td>
<td>RNasin</td>
<td>5 min</td>
<td>1</td>
<td>132,930,781</td>
<td>73,954,019</td>
<td>56%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-RNase</td>
<td>RNase A</td>
<td>5 min</td>
<td>1</td>
<td>109,422,428</td>
<td>60,626,793</td>
<td>56%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-Perm-R2-T1</td>
<td>Perm Only</td>
<td>5 min</td>
<td>2</td>
<td>103,465,047</td>
<td>56,645,401</td>
<td>55%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-Rnasin-R2-T1</td>
<td>RNasin</td>
<td>5 min</td>
<td>2</td>
<td>133,312,880</td>
<td>70,976,130</td>
<td>57%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-RNase-R2-T1</td>
<td>RNase A</td>
<td>5 min</td>
<td>2</td>
<td>118,137,012</td>
<td>57,110,382</td>
<td>56%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-Perm4h-R1-T1</td>
<td>Perm Only</td>
<td>4 hours</td>
<td>1</td>
<td>115,485,533</td>
<td>63,493,291</td>
<td>51%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-RNasin4h-R1-T1</td>
<td>RNasin</td>
<td>4 hours</td>
<td>1</td>
<td>127,002,590</td>
<td>68,305,028</td>
<td>59%</td>
</tr>
<tr>
<td>TB-HiC-Dpn-RNase4h-R1-T1</td>
<td>RNase A</td>
<td>4 hours</td>
<td>1</td>
<td>115,630,362</td>
<td>62,636,554</td>
<td>45%</td>
</tr>
</tbody>
</table>

### Inter-chromosomical interaction analysis

To measure inter-chromosomial interactions we calculated the expected Hi-C matrices in trans for 1 Mb binned data, correcting for the length of each chromosome (*cooltools compute-expected*). We then plotted the “balanced.avg” values for each pair of chromosomes from the expected files.

### Contact probability (P(s)) plots & derivatives

Cis reads from the valid pairs files were used to calculate the contact frequency (\(P\)) as a function of genomic separation (s) (adapted from *cooltools*). Corresponding derivative plots were made from each \(P(s)\) plot.

### Compartment analysis

Compartment boundaries were identified in cis using eigen vector decomposition on 200 kb binned data with *cooltools call-compartments* function. A and B compartment identities were assigned by gene density tracks such that the more gene-dense regions were labeled A compartments, and the PC1 sign was positive. Change in compartment
type, therefore, occurs at locations where the value of PC1 changes sign. Compartment boundaries were defined at these locations.

To measure compartmentalization strength, we calculated observed/expected Hi-C matrices for 200 kb binned data, correcting for average distance decay as observed in the $P(s)$ plots \textit{(cooltools compute-expected)}. We then arranged observed/expected matrix bins according to the PC1 values of either the DMSO control sample or each individual track. We aggregated the ordered matrices for each chromosome within a dataset then divided the aggregate matrix into 50 bins and plotted, yielding a “saddle plot” \textit{(cooltools compute-saddle)}. Overall strength of compartmentalization was defined as the ratio of $(A-A + B-B) / (A-B + B-A)$ interactions. Values used for this ratio were determined by calculating the mean value of the $10 \times 10$ square in each corner of the saddle plot. Strength of A versus B compartments was defined as the ratio of $(A-A / A-B)$ or $(B-B / A-B)$, respectively.

In order to observe compartmentalization at different genomic ranges, we extracted observed/expected Hi-C data at specific distances (0-4 Mb, 4-8 Mb, 8-18 Mb, 18-38 Mb, and 38-80 Mb) and made saddle plots.

Note that some compartment analyses only use the six structurally intact HeLa S3 chromosomes (chromosome 4, 14, 17, 18, 20, and 21) determined in Naumova, Imakaev, Fudenberg et al., 2013.

\textbf{iMARGI chromatin associated RNA (caRNA) analysis}

We used a previously identified set of HFF RNA-chromatin interactions determined by iMARGI since no datasets specifically for HeLa S3 are publicly available (GEO accession number GSM3478206) (Yan, Huang, Wu et al., 2019; Wu, Yan et al., 2019).
This dataset contained 16,688 RNAs that each had anywhere from 1 to 193,561 chromatin associating regions. For each RNA, we measured that percentage of chromatin regions it associates with that are defined as A compartment type for the Hi-C libraries of cells treated with RNASin or RNase A for 5 minutes and 4 hours. We then filtered out any RNAs that associated with chromatin regions of missing value in the Hi-C datasets which yielded 16,525 caRNAs in the final analysis. To cluster the percentage of chromatin interactions caRNAs have that are in A compartments in RNASin versus RNase A treatment conditions, we used kmeans clustering algorithm from scikit-learn (sklearn.cluster.KMeans) with n_clusters = 3 (Pedregosa et al., JMLR 12, pp. 2825-2830, 2011).

**TAD analysis**

Domain boundaries were identified using insulation analysis on 40 kb binned data with cooltools diamond-insulation with a 480 kb window and locating all minima in each profile (--ignore-diags 2 --min-dist-bad-bin 2) with a threshold of log2 boundary strength > 0.15. Domain boundaries were classified as compartment boundaries if they overlapped with the compartment boundaries defined above. All other domain boundaries were assumed to be TAD only boundaries.

To measure insulation strength at TAD boundaries, we aggregated 40 kb binned insulation values at domain boundaries on the six structurally intact HeLa S3 chromosomes (chromosome 4, 14, 17, 18, 20, and 21) (Naumova, Imakaev, Fudenberg et al., 2013). Average insulation across TAD domains was calculated by the averaging the aggregated domains. Insulation strength at each individual boundary was defined as the difference between the local maxima surrounding each boundary and insulation value.
directly at the boundary (average of 2 bins surrounding each boundary). Wilcoxon rank-sum tests were done using scipy-stats package (scipy.stats.ranksums) for each control-treatment pair.

**Loop analysis**

We used a previously identified set of HeLa S3 looping interactions for this analysis (Rao, Huntley et al., 2014). This set contains 3,094 total loops and 507 looping interactions are on the structurally six intact chromosomes of HeLa S3 cells (chromosomes 4, 14, 17, 18, 20, and 21) (Naumova, Imakaev, Fudenberg et al., 2013). In order to observe formation of looping interactions, we aggregated observed/expected Hi-C matrices for 10 kb or 20 kb binned Hi-C data at sites of looping interactions.

Strength of loop formation was defined as the enrichment of signal at the looping interactions (center 5x5 pixels at loop position in 10 kb binned data) compared to the flanking regions. Strength was calculated by averaging the signal at the looping interaction and subtracting the average signal outside. We did this same analysis for different groupings of loops (i.e. grouping by size and grouping by presence of CTCF motif). Size and CTCF motif were determined from the original dataset. Wilcoxon rank-sum tests were done using scipy-stats package (scipy.stats.ranksums) for each pair.

**Code Availability**

Results

Degrading RNA in nonsynchronous cells

To test if RNA is required to maintain the structures of interphase chromosome organization, we degraded RNA molecules in permeabilized nonsynchronous HeLa S3 cells by treatment with RNase A and compared these cells to control ‘Mock’ treated cells and cells treated with an RNase inhibitor, ‘RNasin’ (Fig. 4.1a). Cells treated with RNase A, for as little as 5 minutes, show condensed chromatin aggregates and collapse of chromatin around the nucleoli (Fig. 4.1a, right), as observed previously (Nickerson et al., 1989; Hall, Carone et al., 2014). This collapse is enhanced with a longer 4 hour incubation in RNase A. We note some morphological changes are also observed in the perm only ‘Mock’ treated cells when compared to cells treated with RNasin, suggesting that RNA might also be degraded in these cells. We quantified the amount of RNA degraded in each of these conditions on a fragment analyzer and find that RNA molecules are efficiently degraded after 5 minutes of RNase A treatment, and further degraded after 4 hours (Fig. 4.1b). Cells treated with RNasin maintain the integrity of the RNA molecules present, however, we were not able to recover RNA as well after 4 hours of treatment in any of the conditions. Nevertheless, we find that RNasin is necessary to use as a control due to the RNA degradation that occurs in the ‘Mock’ permeabilize only cells after both 5 minutes and 4 hours at 37°C. We believe that this is due to the activation of endogenous endonucleases which can be inhibited by RNasin.
FIGURE 4.1: RNA degradation in nonsynchronous cells

(a) Experiment schematic for degrading RNA in nonsynchronous HeLa S3 cells to measure 3D chromosome organization. Imaging shows DAPI staining of cells for each treatment after 5 minutes or 4 hours, revealing morphology changes due to RNA degradation. Scale bar = 5um. (b) Fragment analyzer analysis of ribo-depleted RNA isolated from cells after treatment with mock, RNasin, or RNase A for 5 minutes and 4 hours. Bottom panel shows quantification of bands in the top panel.
Inter-chromosomal interactions change in cells with degraded RNA

To assess chromosome conformational changes in cells with degraded RNA molecules, we performed Hi-C on permeabilized nonsynchronous cells treated with RNasin or RNase A for 5 minutes and 4 hours. While we also performed Hi-C on the ‘Mock’ treated cells, these data are not included in the remaining analyses due to the RNA degradation shown in Fig. 4.1b. The statistics for reads from all Hi-C libraries are included in Table 4.1.

Since nuclei of permeabilized cells treated with RNase A undergo a noticeable morphology change, we first wondered if the genome-wide chromatin interactions between different chromosomes changed. We created Hi-C chromatin interaction maps of the average interaction frequency of each chromosome pair in the Hi-C libraries of cells treated with RNasin or RNase A for 5 minutes (Fig. 4.2a) and 4 hours (Fig. 4.2c). All four conditions have inter-chromosomal chromatin interaction maps with the typical features expected for interphase cells. Most notably, we observe that small chromosomes interact more with each other than any other chromosome pairs (Boyle et al., 2001; Tanabe et al., 2002; Lieberman-Aiden, Berkum et al., 2009). To determine changes in inter-chromosomal interactions, we subtracted the Hi-C chromatin interaction map of RNasin treated cells from the interaction map of RNase A treated cells for 5 minutes of treatment (Fig. 4.2b) and 4 hours of treatment (Fig. 4.2d). We find that after 5 minutes of treatment with RNase A, overall inter-chromosomal contacts are reduced compared to cells treated with RNasin, with a few exceptions like chromosome 19 that slightly increases interactions with some chromosomes (Fig. 4.2b). This implies that the morphological changes of cells treated with RNase A for 5 minutes could represent a collapse of
FIGURE 4.2: Inter-chromosomal 3D chromatin contacts are altered in cells with degraded RNA

(a) Average interactions are plotted for each pair of chromosomes for cells treated with RNasin (left) or RNase A (right) for 5 minutes. Values were calculated using the average expected interactions of 1 Mb binned Hi-C data. (b) The difference in the average inter-chromosomal interactions of RNase A treated cells and cells treated with RNasin for 5 minutes. (c) Average interactions for each pair of chromosomes for cells treated with RNasin (left) or RNase A (right) for 4 hours, calculated as in panel a. (d) The difference in the average inter-chromosomal interactions of cells treated with RNase A for 4 hours and cells treated with RNasin for 4 hours. (e) The difference in the average inter-chromosomal interactions of cells treated with RNasin for 4 hours and cells treated with RNasin for 5 minutes to highlight changes that occur just from permeabilized cells incubated at 37°C for 4 hours. (f) The difference in the average inter-chromosomal
interactions of cells treated with RNase A for 4 hours and cells treated with RNasin for 5 minutes revealing that the difference observed between samples at 4 hours is still true despite the changes between RNasin samples for 5 minutes and 4 hours.
individual chromosomes mostly on themselves, thereby limiting interactions between chromosomes. In contrast, cells treated with RNase A for 4 hours reveal an overall large increase in inter-chromosomal interactions, with the exception of small chromosomes which interact with each other less than in cells treated with RNasin (Fig. 4.2d). This suggests that given 4 hours to diffuse and form new interactions, chromosomes are able to interact with each other more in an environment with degraded RNA than in the presence of RNA. Therefore, small chromosomes reduce interactions with each other and form new interactions with larger chromosomes, just as other larger chromosomes form new chromosomal interactions. This is also supported by an overall increase in trans interactions observed in the Hi-C libraries of RNase A treated cells versus RNasin treated cells for 4 hours (Table 4.1).

We note, however, that treatment with RNasin for 5 minutes and 4 hours do not appear completely the same (Fig. 4.2e). Cells incubated for 4 hours in RNasin show decreased overall inter-chromosomal interactions compared to cells incubated for 5 minutes, except for the small chromosomes which increase their interactions with each other in 4 hours. This suggests that after 4 hours and in the presence of RNA molecules, chromosome interactions become stronger with themselves, including the hub of small chromosomes which interact stronger together as a complex than in 5 minutes. Finally, we compared RNase A treatment for 4 hours with RNasin treatment for 5 minutes and find that overall inter-chromosomal interactions are increased in those cells with degraded RNA and with the ability for chromatin to diffuse and form new interactions for 4 hours (Fig. 4.2f). This confirms that although we were not able to recover as much RNA after 4 hours, the cells treated with RNasin and RNase A for 4 hours have very different global
chromatin interactions. Surprisingly, we did not find an enrichment for inter-chromosomal interactions involved in nucleoli assembly in any of the conditions, which we expected from the morphological collapse of chromatin around nucleoli when treated with RNase A (Fig. 4.1a).

**Chromosome conformational changes in cis in cells with degraded RNA**

To assess chromosome conformational changes within each individual chromosome in cells with RNA molecules degraded, we created Hi-C chromatin interaction maps for cis contacts (Fig. 4.3). We observe the typical features expected for interphase cells in all cis chromatin interaction maps. First, we observe a checker-board like pattern off the diagonal which represents spatially distinct long-range interactions (1-10 Mb) between active and inactive compartment regions (Fig. 4.3a-b, top). Deep analysis of compartment features will be presented in the following section. We next visually observe interactions between loci separated by relatively small genomic distances (closer to the diagonal) representative of topologically associated domains (Fig. 4.3a-b, bottom). Deep analysis of TADs will be in a subsequent section. In brief, the Hi-C chromatin interaction maps for cis contacts show the expected features of chromatin organization into compartments and TADs and differences between RNasin and RNase treated cells will be explored in the following sections.

We next plotted interaction frequencies ($P$) as a function of genomic distance ($s$) between loci to derive properties of chromatin folding from the Hi-C chromatin interaction maps (Fig. 4.3c). We find two distinct organizations by distance in the $P(s)$ plots for all 4 conditions, as expected for chromatin interaction maps of interphase cells (Lieberman-Aiden, Burkum et al., 2009). First, we observe a shallow decay for distances < 1Mb.
FIGURE 4.3: Hi-C analysis of cells with degraded RNA

(a) Left: Hi-C cis interaction maps for nonsynchronous HeLa S3 cells treated with RNasin or RNase A for 5 minutes. Data for chromosome 14 are shown for two resolutions: 200 kb (top row, for entire right arm) and 40 kb (bottom row, for 36.5 Mb – 42 Mb region). Right: Difference in the Hi-C interaction maps of cells treated with RNase A for 5 minutes versus RNasin for 5 minutes for the two resolutions / regions of chromosome 14 shown on the left. (b) Left: Hi-C cis interaction maps for nonsynchronous HeLa S3 cells treated with RNasin or RNase A for 4 hours. Data for chromosome 14 are shown for two resolutions: 200 kb (top row, for entire right arm) and 40 kb (bottom row, for 36.5 Mb – 42 Mb region). Right: Difference in the Hi-C interaction maps of cells treated with RNase A for 4 hours versus RNasin for 4 hours for the two resolutions / regions of chromosome 14 shown on the left. (c) Top: Contact frequency $P$ versus genomic distance ($s$) for cis Hi-C data from nonsynchronous cells treated with RNasin or RNase A for 5 minutes (solid lines) and 4 hours (dashed line). Bottom: Percent trans reads in each Hi-C library represented in the $P(s)$ curves above. (d) Derivative from $P(s)$ plots show in panel c.
representing local interactions contained within TADs. Second, we observe a quicker decay for interactions 1-10 Mb apart corresponding to long-range compartment interactions. The slight shift along the y-axis for these curves is due to differences in the percent of trans contacts in each of the Hi-C libraries (Fig. 4.3c, bottom). We find that the Hi-C libraries of cells treated with RNasin or RNase A for 5 minutes, as well as cells treated with RNasin for 4 hours have a similar percent of trans reads. In contrast, we find that the Hi-C libraries of cells treated with RNase A for 4 hours have increased interactions in trans, compared to the others, shifting the $P(s)$ curve of cis interactions down. We calculated the derivatives of the $P(s)$ and find that the shape of these curves is similar across conditions (Fig. 4.3d). These plots can be used to determine the average size of extruded loops and the density of cohesin extrusion proteins (Gibcus, Samejima, Goloborodko et al., 2018; Gassler, Brandao et al., 2017; Patel, Kang et al., 2019). We find that the average size of extruded loops, determined by the local maximum in the derivative of $P(s)$, is very similar between all 4 conditions at about 100 kb which is consistent with our previous findings for this cell type (Chapter II; Abramo et al., 2019).

Interestingly, we find that while the depth of the minimum of the derivative of $P(s)$ is maintained within treatment sets of 5 minutes and 4 hours, the depth decreases in cells incubated for 4 hours suggesting that the linear density of cohesin on the chromatin is decreased. We propose that permeabilized cells incubated for 4 hours at 37°C have less cohesin bound due to leakage of free protein out of the nucleus.

**Compartments are weaker in cells with degraded RNA**

Compartment features are observed in the Hi-C libraries of cells treated with RNasin or RNase A for 5 minutes and 4 hours (Fig. 4.3a-b, top). However, when looking
specifically at the differences in these heatmaps (Fig. 4.3a-b, top right), we find that cells treated with RNase A for 5 minutes have slightly changed cis contacts compared to cells treated with RNasin, with some specific contacts increasing, while others are decreasing. In contrast, we find that cells treated with RNase A for 4 hours have decreased contacts overall compared to cells treated with RNasin, especially contacts close to the diagonal.

To further analyze the checker-board pattern of compartments observed on the Hi-C interaction maps, we determined the location of active (A) and inactive (B) compartments and the strength of homotypic (A-A and B-B) compartment interactions in cells treated with RNasin or RNase A for 5 minutes (Fig. 4.4a-d) and 4 hours (Fig. 4.4e-i). First, we used eigenvector decomposition aligned with gene density to determine the positions of A and B compartments as positive and negative PC1 values, respectively (Lieberman-Aiden, Berkum et al., 2009). We find that the PC1 values called for the Hi-C libraries of RNase A treated cells for 5 minutes do not correlate very well with that of RNasin treated cells with Pearson correlation r-value of 0.736 (Fig. 4.4a). Previous experiments have indicated that well correlated PC1 values, such as replicate experiments, should have correlation r-values around 0.95. We took a closer look at the compartment tracks of Hi-C libraries for RNasin versus RNase A treated cells for 5 minutes and find that the compartment tracks overlap well for most chromosomes, such as chromosome 14 (Fig. 4.4b), however, the PC1 values do not overlap for chromosome 4 (Fig. 4.4c). Instead, we find that the PC1 values of the Hi-C library of cells treated with RNase A for 5 minutes does not pick up compartments, and instead, we pick up compartments in PC2 for chromosome 4. When corrected for chromosome 4, we find that the compartment tracks of the Hi-C libraries of RNasin and RNase A treated cells for 5
FIGURE 4.4: Compartment boundaries and type are maintained in cells with degraded RNA
(a) Control, 5 minute RNasin treated principal component 1 (PC1) values from eigenvector decomposition on 200 kb binned Hi-C datasets of the six structurally intact HeLa
S3 chromosomes (chr 4, 14, 17, 18, 20, 21) (Naumova 2013) versus the PC1 values generated from the Hi-C libraries of 5 minute RNase A treated cells. Value in the lower right corner shows the Pearson correlation r-value. Correlation of these two samples is lower than expected. (b) PC1 along chromosome 14 for Hi-C data obtained from cells treated with RNasin or RNase A for 5 minutes. (c) PC1 along chromosome 4 for Hi-C data obtained from cells treated with RNasin or RNase A for 5 minutes. Compartments were not picked up in PC1 for chromosome 4 of cells treated with RNase A for 5 minutes, however, we find that compartment tracks are picked up in PC2 for this sample. This was only an issue for chromosome 4, the largest of the structurally intact chromosomes for HeLa S3. (d) PC1 values of the 5 minute RNasin treated sample versus the PC1 track of 5 minute RNase A treated samples, corrected for chromosome 4 to use the value in PC2. This yields a much higher Pearson correlation r-value than in panel a. (e) Control, 4 hour RNasin treated principal component 1 (PC1) values from eigen vector decomposition on 200 kb binned Hi-C datasets of the six structurally intact HeLa S3 chromosomes versus the PC1 values generated from the Hi-C libraries of 4 hour RNase A treated cells. Value in the lower right corner shows the Pearson correlation r-value. (f) PC1 along chromosome 14 for Hi-C data obtained from cells treated with RNasin or RNase A for 4 hours. (g) PC1 along chromosome 4 for Hi-C data obtained from cells treated with RNasin or RNase A for 4 hours. Compartments were not picked up in the PC1 values of either of these samples. (h) Principal component 2 (PC2) values plotted along chromosome 4 pick up the compartment tracks for Hi-C data obtained from cells treated with RNasin or RNase A for 4 hours. This was only a problem for chromosome 4 of the structurally intact chromosomes for HeLa S3 cells. (i) PC1 values of the 4 hour RNasin treated samples, corrected for chromosome 4 to use the PC2 values, versus the PC1 track of the 4 hour RNase A treated samples, corrected for chromosome 4 to use the PC2 values. This actually decreases the Pearson correlation r-value because the PC1 values of chromosome 4 for RNasin and RNase treated cells for 4 hour correlated better than the PC2 tracks for these samples. We also note that the corrected tracks for RNasin and RNase A treated cells for 5 minutes and 4 hours correlate well to each other, having very similar compartment tracks. Pearson correlation r-value for RNasin 5 minutes x RNasin 4 hours = 0.951, RNasin 5m x RNase A 4h = 0.91, RNasin 4h x RNase A 5m = 0.926, and RNase A 5m x RNase A 4h = 0.906.
minutes correlate much more with a Pearson correlation r-value of 0.948 (Fig. 4.4d). For cells treated with either RNasin or RNase A for 4 hours, we find that the compartment tracks determined for the respective Hi-C libraries correlate highly (Fig. 4.4e). However, once again the compartments are picked up in PC2 instead of PC1 for chromosome 4 (Fig. 4.4g-h), with the PC1 values for the remaining chromosomes successfully picking up compartments (ex. chromosome 14, Fig. 4.4f). When corrected for chromosome 4, we find that the compartment tracks for the Hi-C libraries of RNasin and RNase A treated cells for 4 hours actually correlate less well, but still have a Pearson correlation r-value of 0.9 (Fig. 4.4i). This suggests that although the tracks seem to call compartment types similarly by positive versus negative PC1 values, the amplitude of the tracks differ, suggesting compartment strength might change.

To quantify the strength of compartmentalization, or the likelihood of homotypic (A-A and B-B) interactions over that of heterotypic (A-B) compartment interactions, we created "saddle plots" (Fig. 4.5) (Nora et al., 2017). When we plotted interactions between loci arranged by the PC1 values determined from the Hi-C interaction map of cells treated with RNasin for 5 minutes, we observed an enrichment in interactions in the top left and bottom right corners and a depletion in interactions in the opposite corners, resembling a saddle (Fig. 4.5a). This was true for Hi-C data of both cells treated with RNasin and cells treated with RNase A for 5 minutes. We quantified compartment strength as the ratio of homotypic to heterotypic interactions (Fig. 4.5b). We find that cells treated with RNase A for 5 minutes have weaker compartment strength than those treated with RNasin and that while this is due to a decrease in both A-A and B-B compartment interactions, A-A contacts seem slightly more effected by RNA degradation (Fig. 4.5c). For the Hi-C
FIGURE 4.5: Compartmentalization is weakened in cells with degraded RNA
(a) Saddle plots of Hi-C data binned at 200 kb resolution for cells treated with RNasin or RNase A for 5 minutes. Saddle plots were calculated using the PC1 obtained from the Hi-
C data of the RNasin 5m control. Dashed corner boxes represent the regions used to calculate compartment strength from the saddle plots. (b) Plots quantifying compartment strength \((AA+BB)/(AB+BA)\) using the mean values of 10x10 squares in each corner of the saddle plot for Hi-C libraries of cells treated with RNasin or RNase A for 5 minutes. (c) Plots quantifying compartment strength of A compartments \((AA/AB)\) (solid bars) and B compartment \((BB/AB)\) (striped bars) separately using the mean values of 10x10 squares for each corner of the saddle plot for Hi-C libraries of cells treated with RNasin or RNase A for 5 minutes. (d) Saddle plots of Hi-C data binned at 200 kb resolution for cells treated with RNasin or RNase A for 4 hours. Saddle plots were calculated using the PC1 obtained from the Hi-C data of the RNasin 4h control, corrected with the PC2 of chromosome 4. Note that using the compartment track generated from the 4 hour RNasin sample does not look as “saddle”-like as using the 5 minute RNasin treated sample (see panels g-i). (e) Plots quantifying compartment strength \((AA+BB)/(AB+BA)\) using the mean values of 10x10 squares in each corner of the saddle plot for Hi-C libraries of cells treated with RNasin or RNase A for 4 hours. (f) Plots quantifying compartment strength of A compartments \((AA/AB)\) (solid bars) and B compartment \((BB/AB)\) (striped bars) separately using the mean values of 10x10 squares for each corner of the saddle plot for Hi-C libraries of cells treated with RNasin or RNase A for 4 hours. (g) Saddle plots of Hi-C data binned at 200 kb resolution for cells treated with RNasin or RNase A for 4 hours. Saddle plots were calculated using the PC1 obtained from the Hi-C data of the RNasin 5 minute control since this yields a better saddle-like shape and the compartment tracks correlate, as shown in Fig. 4.4. (h) Plots quantifying compartment strength \((AA+BB)/(AB+BA)\) using the mean values of 10x10 squares in each corner of the saddle plot for Hi-C libraries of cells treated with RNasin or RNase A for 4 hours and aligned by the PC1 values of the RNasin 5 minute sample. (i) Plots quantifying compartment strength of A compartments \((AA/AB)\) (solid bars) and B compartment \((BB/AB)\) (striped bars) separately using the mean values of 10x10 squares for each corner of the saddle plot for Hi-C libraries of cells treated with RNasin or RNase A for 4 hours and loci aligned using the PC1 values of the RNasin 5 minute sample. (j) Compartment score at different genomic distances in each condition calculated from saddle plots using the RNasin 5 minute PC1 values and split by interaction type. Overall compartmentalization represents the mean values for \((AA+BB)/(AB+BA)\), A-A compartment scores using the mean values of the 10x10 bins for AA/BB, and B-B compartment scores are calculated using the mean values for BB/AB.
libraries of cells treated with RNasin or RNase A for 4 hours, we created saddle plots by plotting interactions between loci arranged by the corrected PC1 values determined from the 4 hour RNasin sample (Fig. 4.5d). These plots did not resemble saddles as well as the samples treated for 5 minutes which is surprising because the PC1 values of the RNasin 5 min and RNasin 4 hour samples correlate (Pearson correlation r-value = 0.926). This implies that the PC1 values determined from the Hi-C library of cells treated with RNasin for 4 hours are different enough that the arrangement of loci in the saddle plots is not as good as for the RNasin 5 minute sample. Nevertheless, we quantified the saddle plots created in this way and find that the compartment strength of cells treated with RNase A for 4 hours is weaker than that of cells treated with RNasin for 4 hours (Fig. 4.5e). We observe this decrease specifically in the compartment strength of A-A interactions more than B-B (Fig. 4.5f). Since the PC1 values for the RNasin 5 minute and corrected PC1 values of the RNasin 4 hours samples correlated so well, we reasoned that we could use the RNasin 5 minute PC1 track to better arrange loci and create saddles for the 4 hour Hi-C data (Fig. 4.5g). When the Hi-C data of cells treated with RNasin or RNase A for 4 hours is arranged according the PC1 track of the RNasin 5 minute Hi-C library, we observe a much more prominent saddle plot, similar to what is seen for the 5 minute samples. We quantified compartment strength from these saddle plots and confirm our findings that compartmentalization is weaker in cells treated with RNase A than those treated with RNasin (Fig. 4.5h), and can attribute this to a higher decrease in A-A compartment type interactions, though B-B interactions are also greatly decreased (Fig. 4.5i). We also note that compartment strength is slightly decreased between the Hi-C data from cells treated with RNasin for 5 minutes and 4 hours, presumably due to the
incubation time. Regardless the difference is much more prominent with treatment of RNase A.

Since interactions observed visually in the Hi-C chromatin interactions maps (Fig. 4.3) changed depending on distance from the diagonal, we quantified compartmentalization as a function of genomic distance (Fig. 4.5j). As we have shown previously (Chapter II; Abramo et al., 2019), compartmentalization is strongest for loci separated by 8-38 Mb, B-B interactions are stronger than A-A interactions for loci separated up to 18 Mb, and A-A interactions are more prominent for loci separated by > 18 Mb. When RNA molecules are degraded in cells by RNase A treatment for 5 minutes, we find that overall compartment strength \((A-A + B-B / A-B + B-A)\) is decreased at all distances (Fig. 4.5j, left). We next quantified compartmentalization of A and B compartments separately and find that compartment strength for cells treated with RNase A versus RNasin for 5 minutes is weakened more for A-A interactions compared to B-B interactions for loci separated by up to 18 Mb (Fig. 4.5j, middle/right). In contrast, we find that for loci separated by > 18 Mb, B-B interactions are weakened more than A-A interactions. For the Hi-C libraries of cells treated for 4 hours with either RNasin or RNase A, we find that overall compartment strength is decreased at all distances in cells with degraded RNA (Fig. 4.5j, left) and A-A interactions weaken more than B-B interactions at all distances (Fig. 4.5j, middle/right). This suggests that RNA molecules might have a larger role in the chromatin organization of A compartments than B compartments.

Lastly, we quantified the strength of compartmentalization for inter-chromosomal trans contacts (Fig. 4.6). We find that compartment strength in trans for the Hi-C libraries of cells treated with RNasin or RNase A for 5 minutes is similar (Fig. 4.6a-c). This
FIGURE 4.6: Trans compartmentalization is weakened in cells with degraded RNA

(a) Saddle plots of trans Hi-C data binned at 200 kb resolution for cells treated with RNasin or RNase A for 5 minutes. Saddle plots were calculated using the PC1 obtained from the Hi-C data of the RNasin 5 minute control. (b) Trans compartment strength (AA+BB)/(AB+BA) calculated by the mean of the 10x10 square of each corner of the saddle plot. (c) Trans compartment strength separated by compartment type. A compartment strength (solid bars) calculated as the ratio of the mean of the 10x10 square in the bottom right of the saddle plot to the mean of the 10x10 square for AB interactions. B compartment strength (striped bars) calculated as the ratio of the mean of the 10x10 square in the top left of the saddle plot to the mean of the 10x10 square for AB interactions. (d) Saddle plots of trans Hi-C data binned at 200 kb resolution for cells treated with RNasin or RNase A for 4 hours. Saddle plots were calculated using the PC1 obtained from the Hi-C data of the RNasin 5 minute control. (e) Trans compartment strength (AA+BB)/(AB+BA) calculated by the mean of the 10x10 square of each corner of the saddle plot for cells treated with RNasin or RNase A for 4 hours. (f) Trans compartment strength separated by compartment type for cells treated with RNasin or RNase A for 4 hours. A compartment strength (solid bars) calculated as the ratio of the mean of the 10x10 square in the bottom right of the saddle plot to the mean of the 10x10 square for AB interactions. B compartment strength (striped bars) calculated as the ratio of the mean of the 10x10 square in the top left of the saddle plot to the mean of the 10x10 square for AB interactions.
suggests that trans compartment interactions do not change after a 5 minute treatment with RNase A. In contrast, we find that trans compartment strength for the Hi-C libraries of cells treated with RNase A for 4 hours is much lower than that of cells treated with RNasin for 4 hours (Fig. 4.6d-e). Interestingly, the trans compartment strength for the Hi-C data of cells treated with RNase A for 4 hours is similar to the trans compartment strength for the Hi-C data of cells treated with RNasin or RNase A for 5 minutes. Therefore, the changes in compartment strength observed here are due to an increase in trans compartmentalization in the Hi-C data for cells treated with RNasin for 4 hours. This suggests that during the 4 hour incubation, chromatin is better segregated into A and B compartments between chromosomes. We find that the increased compartmentalization is occurring in both A-A and B-B contacts (Fig. 4.6f). Further, when compared to the 5 minute treatments, we find that the Hi-C data from cells treated with RNase A for 4 hours have decreased A-A compartmentalization in trans, but relatively maintained B-B contacts (Fig. 4.6c, 4.6f). This strengthens the findings above and implies that RNA molecules might have a larger role in the chromatin organization of A compartments than B compartments.

Compartment switches occur in genomic regions thought to interact with caRNAs

We next checked if the chromatin regions known to interact with chromatin-associated RNAs (caRNAs) changed upon RNA degradation. We used a previously published dataset of RNA-DNA interactions mapped by iMARGI for HFF cells since no datasets are currently available for HeLa S3 cells (Yan, Huang, Wu et al., 2019). For each caRNA in the datasets, we determined the percentage of chromatin loci it has been shown to interact with that are classified as A compartments in our control Hi-C dataset from
cells treated with RNasin for 5 minutes (Fig. 4.7a). We find that overall the loci defined to interact with caRNAs in this dataset are classified more as A compartments than B compartments (Fig. 4.7a, top). Further, we find that the majority of the caRNAs fall into the bins at either end of the histogram, implying that 90-100% of the chromatin regions these caRNAs bind to are classified as the same compartment type. To be sure that this was not due to caRNAs with only a single chromatin binding site, we compared the percentage of A compartments each caRNA is thought to interact with by the number of chromatin binding sites found previously for that caRNA (Fig. 4.7a, bottom). While we find that some caRNAs of this dataset only had 1 chromatin associating region, leading to a classification of 100% A compartment interacting or 100% B compartment (0% A compartment) interacting, most of the caRNAs were determined to have many more than one chromatin associating region.

To observe how loci normally thought to associate with caRNA change with RNase A treatment, we measured the percentage of chromatin loci defined as interacting with caRNA that were classified as A compartments in the Hi-C datasets of cells treated with RNase A for 5 minutes and compared this to the measurements of cells treated with RNasin for 5 minutes (Fig. 4.7b). We find three distinct groups by k-means clustering of these datasets. In cluster 1, we observe no change in the percentage of A compartments defined in the Hi-C data of RNasin versus RNase A treated cells, implying no compartment type switching of the genomic loci thought to interact with these caRNAs (Fig. 4.7b, top). This represents 52.7% of the caRNAs of the iMARGI dataset (Fig. 4.7c), therefore over half of the RNAs thought to bind DNA seem to have no effect on the overall 3D structural organization of chromatin. Cluster 2 contained genomic loci thought to be
FIGURE 4.7: Compartment types of chromatin regions associated with RNAs change in cells with degraded RNA

(a) Top: Distribution of the percentage of chromatin interactions for a previously published iMARGI dataset of chromatin-associated RNAs (caRNAs) that are classified as A compartments in Hi-C libraries of cells treated with RNasin for 5 minutes. Color scale represents 0% interactions with regions in A compartments (dark purple) to 100% of the caRNA interactions with chromatin in A compartments (yellow), as shown in panel b. Histogram represents a total of 16,525 caRNAs with each RNA ranging from 1 to 193,561 chromatin associating regions. Bottom: Scatter plot comparing the percentage of chromatin interactions for caRNAs that are classified as A compartments, as in the top, with the number of chromatin associated regions for each of the 16,525 caRNAs. We note that there is no bias for only few chromatin associated regions for caRNAs that strictly interact with regions of A (100%) or B (0%) compartment type. We also note the discrete behavior of this plot at low values of chromatin associating regions in y which is due to only certain percentages in x being allowed based on the values in y. (b) Kmeans cluster (k=3) on heatmaps of the percentage of caRNA interactions with A compartments for Hi-C libraries of cells treated with RNasin for 5 minutes versus RNase A for 5 minutes. We find the clusters with 3 distinct features. Cluster 1 (n=8757) contains caRNAs that have the exact same percentage of interactions with A type chromatin regions in RNasin and
RNase A treated conditions. Cluster 2 (n=4664) contains caRNAs with chromatin regions that become more B-like upon RNase A treatment for 5 minutes. Lastly, cluster 3 (n=3104) contains caRNAs with chromatin regions that are classified more as A compartments in the RNase A treated cells than they were in the RNasin treated cells. Arrows on the right highlight regions where the chromatin regions that a particular caRNA interacted with completely change compartment type when treated with RNase A for 5 minutes compared to the RNasin control. (c) Pie chart showing the percentage of caRNAs that fall into each category from the k-means clustering of Hi-C data of cells treated with RNasin or RNase A for 5 minutes. Clusters 2 and 3 are further separated to highlight those caRNAs that interact with chromatin regions that completely switch compartment type when treated with RNase A for 5 minutes. (d) Distribution of the difference between percentages each caRNA interacts with regions classified as compartment type A in cells treated with RNase A versus RNasin for 5 minutes. This creates a plot such that cluster 1 is on the left with difference values of 0 to -100 and cluster 2 is on the right with difference values of 0 to +100. Any negative value represents caRNAs in which the chromatin regions they associate with become more B type compartments when treated with RNase A, while positive values represent more A type compartments for the chromatin regions of caRNAs. (e) Scatter plot comparing the difference values from panel d with the number of chromatin associated regions for each caRNAs. We note the same behaviors observed in the bottom of panel a. (f) Distribution of the number of chromatin associating regions for each of the caRNAs in which the chromatin region(s) with which it associates completely change compartment type after RNase A treatment for 5 minutes. (g) Pie chart showing the percentage of caRNAs that fall into each category from the k-means clustering (k=3) of Hi-C data of cells treated with RNasin or RNase A for 4 hours. Clusters 2 and 3 are further separated to highlight those caRNAs that interact with chromatin regions that completely switch compartment type when treated with RNase A for 4 hours. (h) Pie chart showing the percentage of caRNAs that fall into each category from the k-means clustering (k=3) of Hi-C data of cells treated with RNasin for 5 minutes or 4 hours. Clusters 2 and 3 are further separated to highlight those caRNAs that interact with chromatin regions that completely switch compartment type when treated with RNasin for 4 hours. (i) Pie chart showing the percentage of caRNAs that fall into each category from the k-means clustering (k=3) of Hi-C data of cells treated with RNasin or RNase A for 4 hours. This differs from panel g because we first removed the 46.5% of caRNAs in which the chromatin regions they associate with change between RNasin treatment for 5 minutes and 4 hours. Therefore, this pie chart should only show changes that are specific to treatment with RNase A for 4 hours and not any general effects of incubation for 4 hours. Clusters 2 and 3 are further separated to highlight those caRNAs that interact with chromatin regions that completely switch compartment type when treated with RNase A for 4 hours compared to RNasin for 4 hours.
bound by caRNA which became more enriched in B-type compartments upon RNase A treatment for 5 minutes compared to RNASin treated cells (Fig. 4.7b, middle). This cluster represents 28% of the caRNAs defined in the iMARGI dataset, suggesting that 28% of the RNAs thought to bind DNA might be responsible for keeping chromatin regions open in an A-compartment like conformation. Thus, upon RNase A treatment for 5 minutes, many of these chromatin regions get more condensed and switch to a B compartment status. In fact, we find that 1.4% of the caRNAs of the total iMARGI dataset may be fully responsible for keeping chromatin loci in an A compartment conformation because these regions show a complete compartment switch from A to B after RNase A treatment for 5 minutes (Fig. 4.7c). Lastly, cluster 3 is classified as genomic regions bound by caRNA that are classified more as A compartments following RNA degradation by RNase A for 5 minutes (Fig. 4.7b, bottom). This cluster represents the smallest portion of the caRNAs (19.3%) defined in the previously published iMARGI dataset and suggests that these RNAs might be responsible for maintaining chromatin regions as B compartments (Fig. 4.7c). While this cluster overall is smaller than the others, it represents almost twice as many RNAs than cluster 2 which may be fully responsible for compartment maintenance. This cluster shows that genomic regions bound by 2.2% of the total caRNAs defined in the iMARGI dataset completely switch compartments from B to A after RNA degradation with RNase A for 5 minutes (Fig. 4.7c).

We examined clusters 2 and 3 further to determine how much the chromatin regions thought to be bound by these caRNA change upon RNase A treatment (Fig. 4.7d-f). First, we measured the difference in the percentage of A compartment regions thought to be bound by caRNAs in the Hi-C datasets of cells treated with RNase A for 5 minutes
versus RNasin (Fig. 4.7d). Negative values in this analysis represent genomic loci of our Hi-C datasets that are thought to be bound by caRNA and are classified more often as B compartments when treated with RNase A (cluster 2), while positive values represent loci classified more as A compartments (cluster 3) compared to the Hi-C data from cells treated with RNasin for 5 minutes. We find an enrichment in the values close to 0, meaning very few compartment changes, and at the edges of the plot, meaning complete or close to a complete compartment switch, but do also find many genomic loci that interact with RNAs falling in between. We next plotted these differences against the number of chromatin associating regions each caRNA is thought to have from the iMARGI dataset, similar to the lower panel of Fig. 4.7a, to see if there was any bias in this analysis (Fig. 4.7e). We observe a similar pattern to Fig. 4.7a and conclude that there is no bias based on the number of chromatin associated regions and changes in the compartment type of genomic loci defined in the Hi-C datasets. Lastly, we analyzed specifically those genomic loci thought to interact with caRNAs that cause complete compartment switching from A to B (cluster 2) or from B to A (cluster 3) upon RNA degradation (Fig. 4.7f). We find that the majority of caRNAs responsible for complete compartment switches in the genomic loci of Hi-C data were shown to interact with 0-25 chromatin regions by iMARGI, with many of these RNAs shown to have more than 1 chromatin binding region (not shown). This again confirms that there is no bias for the number of chromatin associating regions in this analysis.

We performed the same analysis for the Hi-C data obtained from cells treated with RNasin or RNase A for 4 hours. We found that the genomic regions thought to be bound by more than half of the RNAs defined in the previously published iMARGI dataset were
not affected by RNA degradation (Fig. 4.7g). These chromatin regions were classified as the same compartment type in the Hi-C data from cells treated with RNasin and RNase A for 4 hours. We found that 17.9% of the RNAs thought to bind DNA might be responsible for maintaining chromatin regions as A compartments (cluster 2) and 24.9% responsible for maintaining B compartments (cluster 3) (Fig. 4.7g). These genomic regions were classified as switching from A to B or B to A, respectively, upon RNase A treatment for 4 hours compared to the Hi-C data from cells treated with RNasin for 4 hours.

Since some morphology changes seemed to occur in these cells simply because they were permeabilized and unfixed for 4 hours, as described previously, we next determined how genomic loci thought to be bound by RNA changed specifically in the Hi-C datasets from cells treated with RNasin for 5 minutes versus 4 hours (Fig. 4.7h). We find that the majority of genomic regions thought to interact with RNA (53.5% of the caRNAs from iMARGI) do not change compartment type between the Hi-C dataset of cells treated for 5 minutes and 4 hours with RNasin. Interestingly, cluster 2 is much higher (about 35% of the caRNAs) in this analysis compared to others, representing a higher amount of genomic regions thought to be bound by RNA that switch to a more B-like compartment status after 4 hours in RNasin treatment, compared to the Hi-C data from cells treated with RNasin for only 5 minutes. We removed these caRNAs, as well as the about 10% of caRNAs that were in cluster 3 of this kmeans clustering analysis since changes in these genomic regions in the Hi-C datasets were due to the incubation time. After removal, we re-analyzed the Hi-C datasets from cells treated with RNasin or RNase A for 4 hours to compare changes in the genomic loci thought to interact with RNA specifically from treatment with RNase A (Fig. 4.7i). In contrast to the observations in Fig.
4.7g, we find that an overwhelming larger percentage (80.9%) of genomic regions thought
to interact with caRNAs do not switch compartment type in the Hi-C data from cells treated
with RNasin versus RNase A for 4 hours. Further, we find a roughly similar percentage of
caRNAs in clusters 2 and 3, implying that the compartment type changes that occur in
the Hi-C data from treatment with RNase A are not specific to a particular compartment.
In other words, the genomic regions that seem to rely on the presence of RNA for their
3D organization are equally found in A and B compartments.

**TADs are maintained in cells with degraded RNA**

TADs are shorter range interactions (up to hundreds of kb) that appear on the Hi-
C chromatin interaction maps as squares along the diagonal in which interactions are
enriched within each square and limited outside of the boundaries of each square. TADs
visually appear to be present in all 4 conditions (Fig. 4.3a-b, bottom). When looking
specifically at the differences in these TADs (Fig. 4.3a-b, bottom right), we find that
interactions within TADs are increased in cells with RNA degraded for 5 minutes
compared to RNasin treated cells and that TAD interactions visually appear to be
decreased in cells treated with RNase A for 4 hours compared to RNasin.

To quantify how TAD interactions change in cells with degraded RNA, we
measured the strength of domain boundaries. First, we generated insulation profiles for
each Hi-C sample (Crane, Bian, McCord, Lajoie et al., 2015; Lajoie, Dekker, and Kaplan
2015). We find that the insulation values derived from the Hi-C data from cells treated
with RNase A for 5 minutes correlates very well with that of cells treated with RNasin for
5 minutes (Fig. 4.8a, left). The correlation of the insulation values from Hi-C data from
cells treated for 4 hours correlates slightly less well (Fig. 4.8a, middle). This is likely due
FIGURE 4.8: Insulation at TAD boundaries increases for cells treated with RNase A, then decreases in cells incubated for 4 hours regardless of RNA presence.

(a) Insulation values generated from genome wide analysis on 40 kb binned Hi-C data of cells treated with RNasin for 5 minutes versus cells treated with RNase A for 5 minutes (left), cells treated with RNasin versus RNase A for 4 hours (middle), and cells treated with RNasin for 5 minutes versus cells treated with RNase A for 4 hours (right). Value in
the lower right corner represents Pearson correlation r-values for each plot. (b) Top: Insulation profile along Chromosome 14 for Hi-C data obtained from cells treated with RNasin or RNase A for 5 minutes or 4 hours. Bottom: Insulation boundaries defined as local minima in the insulation profiles for cells treated with RNasin or RNase A for 5 minutes or 4 hours. Black bars represent the regions defined as A compartments in the Hi-C dataset of cells treated with RNasin for 5 minutes. (c) Top: Heatmap showing insulation values 1 Mb upstream and downstream of the TAD boundaries on the six structurally intact HeLa S3 chromosomes defined in the Hi-C data of cells treated with RNasin for 5 minutes. Bottom: Average insulation profile across the TAD boundaries for each condition. (d) Pileup of the average insulation profile across TAD boundaries on the six structurally intact HeLa S3 chromosomes defined in the Hi-C data of cells treated with RNasin for 5 minutes. The number of boundaries used in the pileup is noted at the bottom of the plot. (e) Violin plots of the insulation strength at each individual boundary, separated by location in either A (top) or B (bottom) compartment type. The horizontal white bars represent the median. Boxplots within violin plots (gray) represent the first and third quartiles of the dataset. Asterisks represent significant p-values from Wilcoxon rank-sum test measured between each pair of datasets. If no line is drawn between pairs, no significance was found. (*** = p<2x10^{-12}, ** = p<2x10^{-7}, * = p < 2x10^{-4})
specifically to the RNase A treatment since we find that the insulation values derived from the Hi-C data from cells treated with RNasin for 4 hours correlate just as well with those of cells treated with RNasin for 5 minutes (Fig. 4.8a, right) as the 5 minute RNase A/RNasin insulation values. Nevertheless, when we plotted the insulation values for chromosome 14 in each of the conditions we found remarkably similar values all along the chromosome across the four treatment conditions (Fig. 4.8b, top), suggesting similar location of TADs and domain boundaries. To compare the location of domain boundaries, we defined boundary sites as the local minima in the insulation tracks (Lajoie, Dekker, and Kaplan 2015). We find that the domain boundaries defined in each Hi-C library look very similar to each other, yet, we observe boundaries that are both missing or new across samples (Fig. 4.8b, bottom). These domain boundaries that change in different conditions appeared to mostly be located within regions defined as being A compartments in the Hi-C library of cells treated with RNasin for 5 minutes.

To quantify the strength of the domain boundaries, we aggregated insulation profiles at the TAD boundaries defined by the Hi-C library of cells treated with RNasin for 5 minutes (Fig. 4.8c). We find that insulation at many of these sites changes for cells treated with RNase A for 5 minutes (Fig. 4.8c, left), however, the average insulation profile across these boundaries changes very little (Fig. 4.8c, left bottom and Fig. 4.8d). Insulation at TAD boundaries was greatly reduced in all Hi-C data generated from cells treated for 4 hours (Fig. 4.8c-d), but the average insulation profile across the boundaries between RNasin and RNase A treated cells for 4 hours was very similar. We next measured the insulation strength of each individual boundary (Fig. 4.8e) as the depletion of Hi-C chromatin interactions at the boundary site compared to its surrounding region.
We find that the strength of boundaries in both A and B compartments weaken significantly between the Hi-C libraries of cells treated for 5 minutes versus 4 hours. Within each set of treatment times (5 minutes or 4 hours), however, we observe very few changes in the insulation strength at boundaries between Hi-C data from cells treated with RNasin and cells treated with RNase. For 5 minutes of treatment with either RNasin or RNase A, we find a significant increase in the strength of boundaries in B compartments when RNA is degraded (Fig. 4.8e, bottom), but observe no other significant changes. This increase in boundary strength suggests that fewer chromatin interactions are occurring across TAD boundaries, especially in B compartments, for cells with degraded RNA. With a 4 hour treatment, we observe no significant changes when RNA is degraded versus treatment with RNasin. Overall, we find that the biggest changes in insulation occur due to the incubation time and not necessarily RNA degradation. We also note that all effects on insulation are very small.

**Loops are weaker in cells with degraded RNA**

Last, we tested if looping interactions are maintained in cells with degraded RNA (Fig. 4.9). We used the 507 looping interactions identified by Rao et al. (Rao 2014) for the six structurally intact HeLa S3 chromosomes and aggregated Hi-C data at these sites cells treated treated with RNasin or RNase A for 5 minutes (Fig. 4.9a) and 4 hours (Fig. 4.9c). Upon visual inspection, we observed slight changes in the aggregate interaction maps. To better compare these changes, we quantified the strength of the looping interactions as the prominence of each individual loop over its background signal. We find that loop strength is significantly decreased in cells treated with RNase A for 5 minutes compared to RNasin (Fig. 4.9b). In contrast, we find that cells treated with RNase A for 4
FIGURE 4.9: Loops are weaker in cells treated with RNase A for 5 minutes, decrease overall for cells incubated for 4 hours, and are stronger for cells without RNA for 4 hours versus with RNA present
(a) Aggregate Hi-C data for cells treated with RNasin or RNase A for 5 minutes binned at 10 kb resolution at chromatin loops on intact HeLa S3 chromosomes (n = 507 loops) identified in Rao, Huntley et al., 2014. (b) Loop strength values for cells treated with RNasin or RNase A for 5 minutes. Bar plots represent mean loop strengths with error
bars at +/- 95% confidence interval. Wilcoxon rank-sum test between RNasin and RNase A treated cells shows a significant decrease in loop strength for cells treated with RNase A for 5 minutes. (c) Aggregate Hi-C data for cells treated with RNasin or RNase A for 4 hours binned at 10 kb resolution at the same 507 chromatin loops as in panel a. (d) Loop strength values for cells treated with RNasin or RNase A for 4 hours. Bar plots represent mean loop strengths with error bars at +/- 95% confidence interval. Wilcoxon rank-sum test between RNasin and RNase A treated cells shows a significant increase in loop strength for cells treated with RNase A for 4 hours. (e) Loop strength values for aggregate Hi-C data binned at 10 kb resolution at chromatin loops from panels a and c separated by loop size for cells treated with RNasin or RNase A for 5 minutes (left) and cells treated with RNasin or RNase A for 4 hours (right). Loop size groups are: loops less than or equal to 125 kb (n=148 loops), loops greater than 125 kb and less than or equal to 200 kb (n=122 loops), loops greater than 200 kb and less than or equal to 325 kb (n=112 loops), and loops greater than 325 kb (n=125 loops). Bar plots represent mean loop strengths with error bars at +/- 95% confidence interval. Wilcoxon rank-sum test between each pair shows only a significant differences in loop strength for very small and very large loops in 5 minute treatment and very large loops in the 4 hour treatment. (f) Loop strength values for aggregate Hi-C data binned at 10 kb resolution at chromatin loops from panels a and c separated by the presence of a CTCF motif for cells treated with RNasin or RNase A for 5 minutes (left) and cells treated with RNasin or RNase A for 4 hours (right). Loop motif groups are: “CTCF-CTCF” which corresponds to loops having a CTCF motif at each side of the loop (n=139 loops), “CTCF-noCTCF” which is categorized as those loops having a CTCF motif on only one side of the loop (n=255 loops), and “noCTCF-noCTCF” loops which have no CTCF motif on either side of the loop (n=113 loops). Bar plots represent mean loop strengths with error bars at +/- 95% confidence interval. Wilcoxon rank-sum test between each pair shows a significant decrease in loop strength for all loop motif categories in 5 minute treatment and a significant increase in loop strength for “CTCF-CTCF” loops in the 4 hour treatment.
hours had significantly increased loop strength compared to cells treated with RNasin for 4 hours (Fig. 4.9d). Further, when comparing the raw values of loop strength across all 4 samples, we observe that the initial decrease in loop strength after treatment with RNase A for 5 minutes is maintained after a 4 hour treatment and that the difference in 4 hour samples is due to a drastic decrease in loop strength for RNasin treated cells (Fig. 4.9a-d). This suggests that RNA is not necessary to maintain most looping interactions, however, looping interactions are drastically decreased when in the presence of RNA, using RNasin, and given time for chromatin to dissociate.

We next separated these 507 looping interactions by loop size and again measured the loop strength (Fig. 4.9e). We find that while the strength of looping interactions was decreased at all loop sizes, significant decreases occur at the smallest and largest loops for cells treated with RNase A for 5 minutes compared to RNasin (Fig. 4.9e, left). For the Hi-C libraries of cells treated with RNasin or RNase A for 4 hours, we find that the loop strength for loops of all sizes is stronger for cells treated with RNase A, but only find a significant difference for the largest loops (Fig. 4.9e, right). Last, we aggregated the Hi-C data cells treated with RNasin or RNasin at looping interactions separated by the presence of CTCF motifs at the base(s) (Fig. 4.9f). We find that with a 5 minute treatment of RNase A, looping interactions in all the groups have significantly reduced loop strength compared to loops in the Hi-C libraries of RNasin treated cells (Fig. 4.9f, left). This suggests that RNA is important for the maintenance of these loops regardless of the presence of a CTCF motif at the loop base. In contrast, we show that degradation of RNA in cells for 4 hours significantly increases the loop strength for looping interactions with a CTCF motif on each side compared to cells treated with RNasin, but
find no other significant changes (Fig. 4.9f, right). These analyses suggest that RNA is necessary for the maintenance of some looping interactions, but not all, and that the long incubation in the presence of intact RNA can have some effect on decreasing loop stability.

**Conclusion**

We determine that RNA is not required for the global maintenance of 3D chromosome structure. While morphology changes dramatically in cells treated with RNase A compared to control (Fig. 4.1a), as seen previously (Nickerson et al., 1989; Hall, Carone et al., 2014), few changes are observed in the chromatin interactions of cells with intact versus degraded RNA. We find that features of compartment and TAD organization are preserved in the Hi-C chromatin interaction maps of cells treated with Rnasin or RNase A for 5 minutes and 4 hours (Fig. 4.3). Further, we find that compartment and TAD boundaries are defined at similar locations across these four different conditions (Fig. 4.4, Fig. 4.8). While these features are maintained, we measure quantitative differences in the strengths of chromatin contacts dependent on intact RNA. For example, compartmentalization interactions are weaker in cells with degraded RNA (Fig. 4.5) and insulation around TAD boundaries is stronger in cells with degraded RNA (Fig. 4.8). These results differ somewhat to previous reports of Hi-C analysis on RNase A treated cell populations (Barutcu et al., 2019) which is likely due to the length of RNase A treatment. Additionally, since we have two time points after RNase A treatment, we are
able to observe both initial changes due to shock to the chromatin after RNA degradation, and the long-term changes to chromosome organization.

Our data indicate a model in which DNA and RNA often interact due to the limited volume inside of the cell nucleus. We propose that RNA thereby coats each chromosome through both specific and non-specific interactions. When RNA is degraded in this study, chromosomes initially collapse on themselves. This is similar to a balloon shrinking after being popped simply because there are no longer as many forces to maintain the previous volume. This is supported by a slight decrease in inter-chromosomal interactions (Fig. 4.2b), an increase in heterotypic (A-B) compartment interactions (Fig. 4.5a), and an increase in interactions within TADs and thereby the strengthening of insulation between TADs (Fig. 4.3a, Fig. 4.8). After 4 hours of RNase A treatment, we find that chromosomes are able to diffuse into the space previously filled by RNA and form new contacts with other chromosomes. This is supported by a large increase in trans interactions (Table 4.1, Fig. 4.2d) and changes in specific cis contacts such as decreased compartmentalization, suggestive of additional intermingling of different compartment types (Fig. 4.5d-i). While we also note dissociation of chromatin in cells with intact RNA after 4 hours, we attribute this to the slower natural decay of RNA and proteins within the nucleus allowing for some chromatin movement and the formation of a few non-specific contacts, but not as much as in the absence of intact RNA.

Surprisingly, this study did not reveal a defined role for RNA in the maintenance of CTCF-CTCF loops as suggested previously (Hansen, Hsieh, Cattoglio et al., 2019; Saldana-Meyer et al., 2019; Thakur et al., 2019). While we find some loops are weaker when RNA is degraded, we show that after 4 hours CTCF-CTCF loops become even
weaker in the presence of RNA (Fig. 4.9f). In contrast, we reveal a role for RNA in the maintenance of compartmentalization. We find that RNA has a larger role in the chromatin organization of A compartments than B compartments (Fig. 4.5j). We suggest that the presence of RNA near regions of active transcription (A compartments) is much higher than in inactive regions (B compartments), which is supported by the iMARGI dataset used in which genomic regions thought to interact with caRNA were classified more often as A compartments in our Hi-C datasets (Fig. 4.7a). Therefore, the removal of RNA impacts A compartment regions more. We acknowledge that this result differs from previous reports which suggest that RNA has a larger role in heterochromatin (Barutcu et al., 2019; Thakur et al., 2019). Contrary to these studies, we observe changes due to RNA in both heterochromatic B compartments, as well as euchromatic A compartments, but observe a larger effect on the maintenance of A interactions. This is certainly not the singular driving force of compartmentalization due to (1) the preservation of compartments in the absence of RNA, although reduced in strength, and (2) stronger evidence that compartmentalization is driven by interactions of B compartments (Nuebler et al., 2018). Further, this idea would imply that RNA specifically from active transcription plays into compartmentalization and in this study we cannot specify if interacting RNA is generated during this interphase stage or if it is RNA that is maintained through mitosis.
CHAPTER V: Discussion

Summary

This thesis aimed to address a large gap in the field of chromosome biology: the transition from a mitotic to interphase chromatin state. Though we know a lot about the structures in each individual state, mitosis and interphase (presented in “Chapter I: Introduction”), how chromosomes are able to decondense from a mitotic helical loop array and establish interphase compartments, TADs, and loops is not well understood. Chapter II of this thesis defined the temporal order with which CTCF-loops, TADs, and compartments reform as cells exit mitosis. This revealed a novel structure of chromatin unique to the anaphase-telophase transition. Further, we found that TADs are established much quicker than compartments, providing further support to the field that TADs and compartments form by two distinct mechanisms. Chapter III probed active transcription as a driving force for the establishment and maintenance of interphase structures after mitotic exit. While RNA polymerase II has been proposed to facilitate the movement of cohesin in TAD formation and the clustering of euchromatin at transcription factories, we find that active transcription is not required for the establishment or maintenance of TAD and compartment chromatin structures. Interestingly, we do find a quantitative difference in cells with versus without active transcription suggesting that moving RNA polymerase II or the RNA transcripts produced impact the strength of TAD and compartment interactions. The presence of RNA transcripts and their role on interphase chromatin
organization is then interrogated in Chapter IV of this thesis. Cells were treated with RNase A and only slight changes to the strength of TAD and compartment interactions were observed, similar to inhibiting transcription. This, therefore, suggests that newly synthesized RNA or transcripts from a previous cell cycle are not required for maintenance of TAD or compartment domains, though they do contribute to the strength of contacts within domains. Together, the results shown in Chapters II, III, and IV further our understanding of how interphase structures form, how these structures relate to functional activities of interphase cells, and the stability of chromatin structures over time. This last chapter will highlight some of the major results of the work presented in Chapters II, III, and IV in the context of other research and provide future directions to apply these findings and move the field forward.

Condensation ↔ Decondensation: Same processes in reverse?

Chapter II of this thesis presents the study of 3D chromosome organization as cells exit mitosis and enter G1, analyzing decondensation of chromosomes. We wondered if the process of decondensation is simply the reverse process of chromosome condensation. Recently our lab revealed the transition of spatial chromosome organization as cells exit G2 and enter mitosis, demonstrating the condensation of chromosomes (Gibcus, Samejima, Goloborodko et al., 2018). Therefore, in this section, we compare the process of chromosome condensation and decondensation.

Cells in the G2 phase of the cell cycle have very similar chromosome organization to G1 as determined by Hi-C (Naumova, Imakaev, Fudenberg et al., 2013; Nagano,
Lubling et al., 2017; Gibcus, Samejima, Goloborodko et al., 2018). These cells show features of chromosome territories, spatial segregation into A and B compartments, and interactions enriched into TADs. The most notable difference of 3D chromosome organization in G2 versus G1 is that compartmentalization is stronger, yielding a more defined plaid pattern in the Hi-C interaction maps and representing more of a complete separation of A and B compartments with limited heterotypic (A-B) interactions (Nagano, Lubling et al., 2017; Gibcus, Samejima, Goloborodko et al., 2018). This could represent the microphase separation of replicated regions of closely aligned sister chromatids into similar compartments (Leibler 1980; Matsen and Schick, 1994; Jost et al., 2014; Liang et al., 2015; Nagaska et al., 2016; Haddad et al., 2017). As cells progress into prophase of mitosis, these well-known interphase chromatin structures are quickly lost and the Hi-C interaction maps show only interactions very close to the diagonal with no off-diagonal features (Gibcus, Samejima, Goloborodko et al., 2018). Interestingly, TADs and compartments disappear with very similar kinetics as chromosomes condense. In contrast, in Chapter II of this thesis, we show the quick appearance of TADs and the much slower progression to fully compartmentalized chromatin as cells proceed further into G1. The breakdown of compartments and TADs as cell enter mitosis, therefore, seems to be very different from the establishment of these structures entering G1.

To study why these processes seem to occur differently, we need to understand what is actually happening at each state and what proteins are binding and regulating chromosome conformation. During prophase, sister chromatids appear as long linear structures which are closely aligned and mixed (Liang et al., 2015; Nagaska et al., 2016; Gibcus, Samejima, Goloborodko et al., 2018). The linear axis of each chromatid is bound
by 3 key components: cohesin, topoisomerase II, and condensins (Tanaka et al., 2000; Ono et al., 2004; Hirota, Gerlich et al., 2004; Liang et al., 2015). In fact, many of these proteins are bound prior to this state. For example, during interphase, topoisomerase II is mainly responsible for releasing supercoil induced by DNA replication and transcription (Wang 2002). Further, cohesin is thought to be the loop extruding machine during interphase establishment and maintenance of TADs (Fudenberg, Abdennur et al., 2017). With these interphase proteins bound, why then do we lose the chromosome organization of compartments and TADs?

It turns out that the functions of these proteins shift dramatically during mitosis and this is an area of active study. For example, in prophase, topoisomerase II is mainly responsible for the decatenation of sister chromatids before progressing into metaphase (Liang et al., 2015; Nagaska et al., 2016). Cohesin is an interesting story itself. While the bulk of cohesin begins to bind in telophase, as shown in Chapter II here and also by others (Sumara et al., 2000; Darwiche, Freeman and Strunnikov, 1999), a replication-dependent, Esco1/Esco2-dependent cohesin binds during S phase to establish sister chromatid cohesion (Uhlmann and Nasmyth, 1998; Skibbens et al., 1999; Hou and Zou 2005). During prophase, the majority of cohesin is removed by Wapl, coinciding with the disappearance of TADs (Peters and Nishiyama, 2012; Gibcus, Samejima, Goloborodko et al., 2018) and the remaining cohesin molecules are thought to be cohesive cohesins which are stabilized on the chromatin via acetylation by Esco1/Esco2 (Hou and Zou 2005). Therefore, cohesin is only bound during prophase to maintain sister chromatid cohesion and ensure proper segregation of genetic information into daughter cells (Tanaka et al., 2000; Liang et al., 2015; Nagaska et al., 2016). In fact, cohesive cohesin
is removed from the chromosome arms already by late prophase, only localizing at the centromeres until anaphase when sisters separate, and the remaining cohesive cohesin is removed (Waizenegger et al., 2000). Interestingly, the absence of SMC2 (condensin I/II subunit) delays the disappearance of compartments and TADs in prophase even though cells are able to progress through nuclear envelope breakdown (NEB) and chromosomes condense (Gibcus, Samejima, Goloborodko et al., 2018). This implies that condensin I and II are necessary for the disappearance of compartments and TADs, and therefore, may play a role in the removal of cohesin.

Another interesting feature observed during chromosome condensation occurs further into mitosis. As cells progress into prometaphase, after nuclear envelope breakdown, a second diagonal appears parallel to the main diagonal for all loci and chromosomes (Gibcus, Samejima, Goloborodko et al., 2018). This feature of the Hi-C chromatin interaction map is thought to represent the interactions of every loci that are separated by ~3 Mb in the linear DNA sequence via helical loop chromosome organization. As cells progress even further into mitosis, this second diagonal moves to longer distances representing larger chromatin loops as chromosomes condense into shorter and wider chromosomes, similar to observations by microscopy (Gibcus, Samejima, Goloborodko et al., 2018; Liang et al., 2015; Nagaska et al., 2016). This leads to an interesting question in terms of condensation versus decondensation: could a similar mechanism happen in the reverse as cells enter G1? With better synchronization techniques, one could take a finer look at cells exiting mitosis and see if chromosomes slowly get longer and thinner. This result would be opposite to the results seen via condensation and instead we would hope to observe a 2nd diagonal on the Hi-C chromatin
interaction map moving towards the main diagonal, representing progressively smaller loops in each helical loop turn for each chromosome.

In Chapter II, we present a unique, transient chromosome structure during telophase which has never been observed before (Abramo et al., 2019). This chromosome conformation appears to be free of SMC proteins, and therefore, SMC-driven loops. If the progressive mechanism described for the condensation of chromosomes going into mitosis is true in the reverse, decondensation would have to be achieved by telophase when a loop-less chromatin structure is observed. This therefore raises the question of how condensin is removed from chromatin and what the function of a loop-less state could be. The factor(s) required for condensin I and condensin II removal from mitotic chromosomes are still unknown. One could imagine a scenario where cohesin binds to chromosomes and immediately starts extruding chromatin, knocking off condensin. This cannot be the case, however, since it seems in our cells that there is mutual exclusion of condensins and cohesin on the chromosomes. Condensins are bound to the chromosomes until late anaphase, telophase presents an SMC-free state, then immediately following exit from telophase cohesin is able to bind to chromatin and begin extruding. Therefore, it seems that a ‘naked’ state of chromatin in telophase gives cohesin the ability to bind anywhere and quickly extrude without the interference of other SMC proteins.

Further, the unique structure we observed in telophase provides information on a clear difference between chromosomes in prophase versus chromosomes in telophase. During prophase, chromosomes are bound by both cohesin and condensins for a significant amount of time. In contrast, as cells exit mitosis, chromosomes can be bound
by either condensins or cohesin, but never both. As stated above, condensins may remove extruding cohesin during prophase, however, and only leave cohesive cohesin bound for sister chromatid cohesion during mitosis. Therefore, condensins and extruding cohesin may not be compatible, and thus mitotic exit requires a ‘naked’ chromosome state free of SMC proteins for cohesin to bind following telophase and begin extrusion.

An obvious caveat of these experiments is that the phenomena of a chromatin state free of SMC proteins was only observed in one cell type. Therefore, the experiments carried out in Chapter II should be repeated in other cell types to confirm this structure and determine if this process of decondensation can be applied as a general principle of chromosome organization. We attempted to look for this unique chromatin structure in single cell Hi-C data, but were unable to acquire enough knowledge on the exact mitotic state of the cells (Nagano, Lubling et al., 2017).

Enhancer-independent transcription

Mitotic chromosomes in their maximally condensed state during metaphase are short and wide helical loop arrays with no observed short range TAD or loop interactions, and a clearly visible, but more diffuse than during prometaphase, second diagonal positioned at ~12 Mb (Gibcus, Samejima, Goloborodko et al., 2018). This second diagonal represents interactions between loop layers which are ~12 Mb apart in the linear DNA sequence (Gibcus, Samejima, Goloborodko et al., 2018). In order to achieve this fully condensed state, the compartments and TADs normally observed in interphase chromatin organization disappear (Naumova, Imakaev, Fudenberg et al., 2013; Gibcus,
Samejima, Goloborodko et al., 2018; Abramo et al., 2019). In addition, early studies suggested a global arrest in transcription (Prescott and Bender, 1962) and this is believed to be facilitated by the removal or dissociation of most chromatin regulators from the condensed chromosomes (Martinez-Balbas et al., 1995). Recent evidence, however, shows that a low level of transcription is maintained during mitosis (Palozola et al., 2017). This is supported by the idea of ‘mitotic bookmarking’ (Michelotti et al., 1997), in which the more active genes of interphase remain transcriptionally active during mitosis, albeit at a much lower level. The complete mechanism of mitotic bookmarking is still unknown, however, some transcription factors and histone modifications remain bound to mitotic chromosomes and some TSSs remain accessible during mitosis (Martinez-Balbas et al., 1995; Segil et al., 1996; Michelotti et al., 1997; Kadauke and Blobel, 2013; Oomen et al., 2019; Palozola et al., 2019). Interestingly, a transient burst of transcription was shown to come from these ‘bookmarked’ promoters upon G1 entry. Therefore, this suggests that mitotic bookmarking allows genes to quickly restore gene expression during mitotic exit. Further analysis revealed that these first spikes in transcription as cells progress to G1 are primarily involved in general cell function, while cell-type specific gene expression largely does not occur until later G1 (Palozola et al., 2017). Interestingly, genes involved in general cell function, more commonly referred to as housekeeping genes, are known to require very few enhancers for their transcription, especially as compared to cell-type specific genes such as those involved in limb development (Osterwalder et al., 2018). Therefore, in alignment with the lack of close range chromatin interactions, mitotic transcription likely does not require the formation of enhancer-promoter loops. This is also supported by the lack of eRNA transcription during mitosis which is thought to facilitate
chromatin looping between promoters and enhancers during interphase, and therefore, drive transcription (Palozola et al., 2017). Further, earlier studies also showing a transient spikes in transcription following mitotic exit did not observe a spike in enhancer-promoter contacts (Hsiung et al., 2016).

If transcription of these housekeeping genes is able to occur in the absence of enhancer-promoter loops, we wondered then what the relationship between chromatin loops in G1 and transcription is. Analysis of the mitotic exit time course Hi-C data, presented in Chapter II, resulted in no bias in loops that form first during G1 re-entry and global transcription reactivation (Abramo et al., 2019). Deeper analysis of these Hi-C data should be done, however. Possibly with better synchronization techniques and also more read depth via sequencing, we would be able to call ‘dots’ in each sample to determine if we (1) observe progressively more loops, and (2) observe enhancer-promoter loops at early transcribing genes. Experiments presented in Chapter III of this thesis tried to address the same question by measuring chromatin loops (and other interphase structures) as cells enter G1 with active transcription chemically inhibited. Analysis of these data revealed that chromatin loops were established in the absence of active transcription, however, only a specific set of previously determined loops were used (Rao, Huntley et al., 2014) and these loops appear to be mostly CTCF-CTCF loops. Therefore, enhancing the resolution of these Hi-C datasets via deeper sequencing may be a quick way to determine chromatin loops specific to these treatment conditions. We note, however, that if no differences in chromatin loops are observed between cells with or without active transcription, this supports the idea of enhancer-independent transcription of early housekeeping genes. Since cell-type specific genes take longer before they are
fully expressed, transcription of these loci is likely to follow the kinetics of bulk transcription presented in Chapter III. Interestingly, we found that the kinetics of bulk transcription correlates better with compartmentalization of cells than with establishment of TADs and chromatin loops as cells enter G1. Since compartments are not conserved across cell types and differentiation states (Lieberman-Aiden, Berkum et al., 2009; Dixon, Jung, Selvaraj et al., 2015), it makes sense for compartmentalization to correlate with the more cell-type specific gene expression. Overall, this further supports the idea that transcription is not required for the establishments of TADs and chromatin loops (at least those that likely involve CTCF), though it does affect the strength of intra-TAD interactions and compartmentalization due to cell-type specific gene expression.

Replication Domains and Compartmentalization

There is a defined spatiotemporal program to the order with which chromosomal regions are replicated during S phase (Hatton et al., 1988; Manuelidis, 1990; Dimitrova and Gilbert, 1999). Early replicating loci tend to be transcriptionally active and are localized to the nuclear interior, while later replicating domains are localized to the nuclear periphery and tend to be transcriptionally inactive (Hiratani and Gilbert, 2009). Replication profiles, therefore, have been found to correlate highly with the spatial segregation of chromatin into A and B compartments (Hatton et al., 1988; Lieberman-Aiden, Berkum et al., 2009; Ryba et al., 2010). Further, the boundaries of replication domains were shown to correlate with TAD boundaries (Pope, Ryba et al., 2014) and interestingly, inhibiting DNA replication can inhibit the establishment of TADs (Ke, Xu, Chen, Feng et al., 2017).
With compartments and TADs highly correlated with transcription, this suggests a model in which DNA replication and transcription are closely coordinated. The first loci to be replicated in S phase are those that are located within TADs that are permissive for transcription, and therefore, localized to the nuclear interior with other loci of A type compartments. Interestingly, though G2 chromatin is spatially organized similar to G1 and S and localization within the nucleus does not change, G2 phase chromatin does not have the ability to replicate according to the correct replication timing, and therefore has been deemed to lack the determinants of replication timing (Lu et al., 2010). We and others have observed increased compartmentalization of chromatin during S phase and G2 (Chapter II; Nagano, Lubling et al., 2017; Gibcus, Samejima, Goloborodko et al., 2018; Abramo et al., 2019), which may be due to changes in chromatin state after replication. One might imagine a simple model in which once a TAD, and potentially its neighboring TADs of the same compartment type, are replicated, those regions, which are now in duplicate, can all interact in the local spatial environment and create stronger homotypic compartment interactions. This would therefore increase our definition of compartmentalization and yield a more defined plaid pattern on the Hi-C chromatin interaction maps.

Determinants of the replication timing program are lost in G2 phase and thus have to be re-established prior to the next S phase (Lu et al., 2010). Consistent with their correlation to interphase 3D chromatin features, the replication timing program is established coincident with TADs and compartments in early G1 (Dileep et al., 2015). Therefore, this leads to the question of why compartments take so long to establish in our time course Hi-C data on cells entering G1 in Chapter II. First, similar to Dileep et al., we
observe the establishment of compartments at the same time as TAD formation (Dileep et al., 2015; Abramo et al., 2019). This is done using eigen vector decomposition and plotting the gene-density corrected PC1 track (Abramo et al., 2019), which was also observed in a similar study (Zhang et al., 2019). However, this just defines the compartment boundaries. While some interactions will begin to bring compartment domains together that are otherwise separated linearly by the opposite compartment type, these will generally be short-range, intra-TAD interactions of homotypic compartments. These interactions are confirmed in Dileep et al. by using FISH (Dileep et al., 2015). In contrast, in Chapter II of this thesis, we determine that the kinetics of compartment establishment is slower than the kinetics of TAD formations because we are measuring compartmentalization which is the measurement of overall nuclear separation between A and B compartments (Abramo et al. 2019). Our results are further supported by other studies which suggest that TADs and compartments are formed by two distinct mechanisms (Schwarzer, Abdennur, Goloborodko et al., 2017; Nuebler et al., 2018; Abramo et al., 2019; Zhang et al., 2019). Regardless, this still questions why such a delay exists between defining replication timing and compartmentalization. It also suggests that licensing of replication origins in late mitosis could influence 3D chromosome architecture.

Further, since replication timing is also correlated with gene regulation, this presents the question of the relationship between replication domains, transcription, and 3D chromosome structure. Interestingly, cell differentiation causes replication timing changes concordant with transcriptional activity (Hiratani et al., 2008; Desprat et al., 2009). In Chapter III of this thesis, we show that compartmentalization is reduced upon transcription inhibition. Therefore, as a follow up to these experiments it would be
interesting to compare the replication timing as cells enter S phase in the presence or absence of active transcription. While the inhibition of transcription in early G1 may prevent the cells from having the appropriate proteins for DNA replication, those proteins could be provided as was done to force DNA replication in G2 phase (Lu et al., 2010). Further, it would be interesting to explore if chromosomes are then able to properly condense for mitosis. Overall, since TAD and compartment boundaries are properly formed, we hypothesize that replication timing will not be greatly impacted by transcription and cells will be able to progress into S, G2, and M. Therefore, an open question would still remain on why certain chromatin regions are replicated before others.

The interplay of TADs, compartments, and transcription

The growing evidence that TADs and compartments form by two distinct mechanisms (Schwarzer, Abdennur, Goloborodko et al., 2017; Nuebler et al., 2018; Abramo et al., 2019; Zhang et al., 2019), which is also shown in Chapter II of this thesis, has emphasized an antagonistic relationship between the two chromatin features. For example, while depleting cohesin (RAD21) or the cohesin loader (NIPBL) results in the disappearance of TADs, it also results in an enhanced compartmentalization (Schwarzer, Abdennur, Goloborodko et al., 2017; Rao et al., 2017). Since depleting these factors did not result in changes to histone modifications, enhanced compartmentalization, especially into smaller defined compartments, is likely the effect of a more complete phase separation of active and inactive chromatin domains. Depletion of cohesin influenced the interchromosomal colocalization of many chromatin contacts at
superenhancers, further supporting the idea of more efficient phase separation of chromatin regions with similar epigenetic marks (Rao et al., 2017). Interestingly, cohesin depletion resulted in relatively unchanged gene expression overall, but the loci that relocated to the superenhancers were greatly down-regulated (Rao et al., 2017). This implies cohesin-dependent TAD formation can influence and promote gene expression, however, transcription does not rely on the formation of TADs. This is consistent with previous findings in early mouse embryos which show that zygotic genome activation (ZGA) occurs prior to the establishment of TADs (Du et al., 2017; Ke, Xu, Chen, Feng et al., 2017). Activation of these genes could be similar to the enhancer-independent gene regulation discussed at the beginning of this chapter. In addition, these studies showed that ZGA is not required for the establishment of TADs in embryos (Du et al., 2017; Ke, Xu, Chen, Feng et al., 2017; Hug et al., 2017).

The results presented in Chapter II of this thesis show that TADs are established in early G1 prior to maximum compartmentalization of chromatin reached much later (Abramo et al., 2019). This, along with the data presented above, suggest that the establishment of TADs is antagonistic to phase separation—without TADs chromatin fully segregates into clusters with different epigenetic marks, but with active loop extrusion forming TADs, the chromatin is unable to fully spatially segregate. An interesting follow up experiment, therefore, would be to perform the same Hi-C time course as cells exit mitosis, however, this time in the absence of cohesin or the cohesin loader. As was shown previously, TADs should not form and compartmentalization should be enhanced with more smaller compartments. The interesting analysis of this experiment is in measuring the kinetics of compartmentalization. TAD depletion could have two impacts on the
kinetics of compartmentalization: (1) compartmentalization may be quicker than previously observed, or (2) compartmentalization could still have slow kinetics but the end result is stronger, smaller compartments. These experiments would reveal more of the complexities of phase separation. Inhibiting transcription, as in Chapter III, could also be done in these experiments to see how the combination of cohesin and transcription depletion affect phase separation and compartmentalization. While TADs and compartments were able to become established in cells entering G1 in the absence of active transcription, these interactions were weaker. The follow up experiment presented here could determine if the weaker compartmentalization was truly due to the lack of active transcription or if the absence of transcription just delays phase separation and compartmentalization.

RNA and interphase structures

Chapter IV of this thesis determined that RNA is not required to maintain the interphase chromatin organizational structures of loops, TADs, and compartments. Although these structures were stably present regardless of the RNA state (degraded or not), we did observe a quantitative difference in the interactions within each 3D structure. Interestingly, contrary to previous studies (Barutcu et al., 2019; Thakur et al., 2019), these experiments revealed a larger dependency on RNA for A-A interactions than interactions between loci in B compartments. We attribute this to (1) a difference in the timing of RNase A treatment and (2) a difference in the techniques used. For example, Thakur et al. use CUT&RUN and measure the accessibility of chromatin at specific protein binding
sites or modified histones (Thakur et al., 2019). Similar to our studies, they find that chromatin regions with heterochromatic histone marks (B compartments) are decondensed upon RNase A treatment. Our study differs however, because we observe larger changes in A-A interactions, which are already accessible. Therefore, how might one pick up changes to already accessible genomic regions that become more accessible using such a technique? Combining these data, a more likely model is that RNA in A compartments functions to counteract the natural tendency of chromatin to condense (Hall and Lawrence, 2016). Therefore, in an experiment such as ours with different time points after RNA degradation, we can measure (1) the initial collapse of chromosomes into a more condensed state with itself, and (2) the later dissociation of chromatin into newly open space. The later dissociation is able to re-establish the accessibility of A compartments, introduces more accessibility in B compartments, and also causes mixing between A and B compartments.

Recently, our lab developed a new technique called liquid chromatin Hi-C that can be used to determine the stability of chromatin interactions in different chromatin regions (Belaghzal, Borrman et al., 2019). Therefore, a future experiment could be to apply this method to cells after RNA degradation. This should allow for the identification of precise chromatin regions that display decreased stability in the chromatin contacts observed by chromosome conformation capture. Further, while a previously published dataset of chromatin associated RNAs was used in the analysis presented in Chapter IV, iMARGI could be performed on these exact cells. In this case, the transcripts that normally bind to the chromatin regions which displayed decreased stability upon RNase A treatment will reveal the most important RNA transcripts for 3D chromosome organization.
Concluding remarks

This thesis has addressed changes in 3D organization as chromosomes transition from condensed mitotic helical loop arrays in to decondensed chromatin organizational features of compartments, TADs, and loops. While gene regulation is highly correlated with interphase chromatin organization, we show that active transcription does not drive the establishment of these features. Further, we find that the RNA products of active transcription have minimal function in the maintenance of compartments, TADs, and loops. While these results further our understanding of how 3D interphase chromatin structures are established and maintained, many questions (addressed in the above discussion) remain unanswered and present a great opportunity for further study.


extrusion organizes zygotic genome architecture. Embo j 36:3600-3618. doi: 10.15252/embj.201798083


Hou F and Zou H (2005) Two human orthologues of Eco1/Ctf7 acetyltransferases are both required for proper sister-chromatid cohesion. Mol Biol Cell 16:3908-18. doi: 10.1091/mbc.e04-12-1063


Mammalian Chromosome Architecture. Molecular Cell 78:554-565.e7. doi: https://doi.org/10.1016/j.molcel.2020.03.003


Miele A, Bystricky K and Dekker J (2009) Yeast Silent Mating Type Loci Form Heterochromatic Clusters through Silencer Protein-Dependent Long-Range Interactions. PLOS Genetics 5:e1000478. doi: 10.1371/journal.pgen.1000478


Sanulli S, Trnka MJ, Dharmarajan V, Tibble RW, Pascal BD, Burlingame AL, Griffin PR, Gross JD and Narlikar GJ (2019) HP1 reshapes nucleosome core to promote phase


Tavares-Cadete F, Norouzi D, Dekker B, Liu Y and Dekker J (2020) Multi-contact 3C data reveal that the human genome is largely unentangled. bioRxiv:2020.03.03.975425. doi: 10.1101/2020.03.03.975425

Terakawa T, Bisht S, Eeftens JM, Dekker C, Haering CH and Greene EC (2017) The condensin complex is a mechanochemical motor that translocates along DNA. Science 358:672-676. doi: 10.1126/science.aan6516


insights into mitotic chromosome architecture. J Cell Biol 217:2309-2328. doi: 10.1083/jcb.201801048


